(12) STANDARD PATENT APPLICATION (11) Application No. **AU 2017200439 A1**
(19) AUSTRALIAN PATENT OFFICE

(54) Title
HOST CELLS AND METHODS FOR PRODUCTION OF ISOBUTANOL

(51) International Patent Classification(s)
C12N 15/61 (2006.01) ***C12P 7/16*** (2006.01)
C12N 15/53 (2006.01) ***C12P 7/52*** (2006.01)
C12N 15/63 (2006.01)

(21) Application No: **2017200439** (22) Date of Filing: **2017.01.20**

(43) Publication Date: **2017.02.16**
(43) Publication Journal Date: **2017.02.16**

(62) Divisional of:
2012230737

(71) Applicant(s)
Butamax (TM) Advanced Biofuels LLC

(72) Inventor(s)
ANTHONY, Larry Cameron;GIBSON, Katharine J.;HE, Hongxian;HUANG, Lisa L.;O'KEEFE, Daniel P.;KRUCKEBERG, Arthur Leo;LI, Yougen;MAGGIO-HALL, Lori Ann;McELVAIN, Jessica;NELSON, Mark J.;PATNAIK, Ranjan;ROTHMAN, Steven Cary

(74) Agent / Attorney
Houlihan² Pty Ltd, PO Box 611, BALWYN NORTH, VIC, 3104, AU

2017200439 20 Jan 2017

ABSTRACT

Provided herein are recombinant yeast host cells and methods for their use for production of isobutanol. Yeast host cells provided comprise an isobutanol biosynthetic pathway and at least one of reduced or eliminated aldehyde dehydrogenase activity, reduced or eliminated acetolactate reductase activity; or a heterologous polynucleotide encoding a polypeptide having ketol-acid reductoisomerase activity.

2017200439 20 Jan 2017

P/00/011
Regulation 3.2

AUSTRALIA

Patents Act 1990

COMPLETE SPECIFICATION

FOR A DIVISIONAL PATENT

ORIGINAL

Name of Applicant: BUTAMAX (TM) ADVANCED BIOFUELS LLC

Actual Inventors: Larry Cameron ANTHONY
Katharine J. GIBSON
Hongxian HE
Lisa L. HUANG
Daniel P. O'KEEFE
Arthur Leo KRUCKEBERG
Yougen LI
Lori Ann MAGGIO-HALL
Jessica McELVAIN
Mark J. NELSON
Ranjan PATNAIK
Steven Cary ROTHMAN

Address for Service: Houlihan^2, Level 1, 70 Doncaster Road, Balwyn North, Victoria 3104, Australia

Invention Title: HOST CELLS AND METHODS FOR PRODUCTION OF ISOBUTANOL

The following statement is a full description of this invention, including the best method of performing it known to the Applicant:-

---

2017200439 20 Jan 2017

TITLE

HOST CELLS AND METHODS FOR PRODUCTION OF ISOBUTANOL

CROSS-REFERENCE TO RELATED APPLICATIONS

The present Application is a Divisional Application from Australian Patent Application No. 2012230737. The entire disclosures of Australian Patent Application No. 2012230737 and its corresponding International Patent Application No. PCT/US2012/030479, are incorporated herein by reference.

This Application claims priority under 35 U.S.C. § 119(e) to U.S. Application Nos. 61/472,484 (filed April 6, 2011), 61/467,261 (filed March 24, 2011), 61/472,487 (filed April 6, 2011), 61/467,271 (filed March 24, 2011), 61/570,513 (filed December 14, 2011), 61/467,249 (filed March 24, 2011), 61/472,497 (filed April 6, 2011), and 61/472,474 (filed April 6, 2011), each of which is in incorporated herein by reference in its entirety.

GOVERNMENT LICENSE RIGHTS

This invention was made with Government support under Agreement DE-AR0000006 awarded by the United States Department of Energy. The Government has certain rights in this invention.

FIELD OF THE INVENTION

The invention relates to recombinant host cells and methods for fermentative production of isobutanol.

REFERENCE TO SEQUENCE LISTING SUBMITTTED ELECTRONICALLY VIA EFS-WEB

The content of the electronically submitted Sequence Listing (Name: 20120323_CL5367USNA_SEQLIST.txt; Size 2,003,893 bytes; Date of Creation March 23, 2012) is herein incorporated by reference in its entirety.

BACKGROUND OF THE INVENTION

Butanol is an important industrial chemical, useful as a fuel additive, as a feedstock chemical in the plastics industry, and as a food grade extractant in the food and flavor industry. Each year 10 to 12 billion pounds of butanol are produced by petrochemical means and the need for this commodity chemical will likely increase in the future.

Methods for the chemical synthesis of isobutanol are known, such as oxo synthesis, catalytic hydrogenation of carbon monoxide (Ullmann's Encyclopedia of Industrial Chemistry, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp. 716-719) and Guerbet condensation of methanol with n-propanol (Carlini et al., J. Molec. Catal. A. Chem. 220:215-220, 2004). These processes use starting materials derived from petrochemicals, are generally expensive, and are not environmentally friendly. The production of isobutanol from plant-derived raw materials would minimize green house gas emissions and would represent an advance in the art.

Isobutanol is produced biologically as a by-product of yeast fermentation. It is a component of "fusel oil" that forms as a result of the incomplete metabolism of amino acids by fungi. Isobutanol is specifically produced from catabolism of L-valine. After the amine group of L-valine is harvested as a nitrogen source, the resulting α-keto acid is decarboxylated and reduced to isobutanol by enzymes of the so-called Ehrlich pathway (Dickinson et al., J. Biol. Chem. 273:25752-25756, 1998).

Improvements and alternatives for the biosynthesis of butanol directly from sugars would improve economic viability and would represent an advance in the art.

SUMMARY OF THE INVENTION

Provided herein are recombinant yeast host cells and methods for the production of isobutanol.

In some embodiments, a recombinant host cell comprises an engineered isobutanol production pathway and (a) at least one of (i) a heterologous polypeptide with ketol-acid reductoisomerase (KARI) activity

selected from the group consisting of (1) a polypeptide having at least about 90% identity to a KARI enzyme derived from *Bifidobacterium angulatum, Bifidobacterium dentium, Zymomonas mobilis, Clostridium beijerinckii* or *Anaerostipes caccae*, or an active fragment thereof (2) a polypeptide having at least about 90% identity or at least about 95% identity to SEQ ID NO: 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, or 65; or (ii) a heterologous polynucleotide encoding the heterologous polypeptide with KARI activity of (a); and (b) at least one host cell modification that enhances performance of the engineered isobutanol production pathway. In some embodiments, the combination of (a) and (b) results is a synergistic increase in isobutanol production pathway performance.

In some embodiments, a recombinant host cell comprises an isobutanol biosynthetic pathway and (a) a heterologous polypeptide with ketol-acid reductoisomerase (KARI) activity selected from the group consisting of (i) a polypeptide having at least about 90% identity to a KARI enzyme derived from *Bifidobacterium angulatum, Bifidobacterium dentium, Zymomonas mobilis, Clostridium beijerinckii* or *Anaerostipes caccae*, or an active fragment thereof, (ii) a polypeptide having at least about 90% identity or at least about 95% identity to SEQ ID NO: 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, or 65, (iii) a polypeptide having at least about 90% identity or at least about 95% identity to a KARI enzyme derived from *Bifidobacterium angulatum, Bifidobacterium dentium, Enterococcus gallinarum, Streptococcus thermophiles, Zymomonas mobilis, Clostridium beijerinckii, Anaerostipes caccae*, or *Lactococcus lactis subsp. cremoris MG1363* or an active fragment thereof, wherein the polypeptide has a K_M for NADH less than about 50, (iv) a polypeptide having at least about 90% identity or at least about 95% identity to a KARI enzyme derived from *Staphylococcus capitis SK14, Staphylococcus epidermidis M23864-W1, Staphylococcus hominis SK119, Staphylococcus aureus subsp. aureus TCH130, Staphylococcus warneri L37603, Staphylococcus epidermidis W23144, Staphylococcus saprophyticus subsp. Saprophyticus ATCC15305, Staphylococcus*

carnosus subsp. Carnosus TM300, Listeria monocytogenes EGD-e, Listeria grayi DSM 20601, Enterococcus casseliflavus EC30, Enterococcus gallinarum EG2, Macrococcus caseolyticus JCSC5402, Streptococcus vestibularis, Streptococcus mutans UA159, Streptococcus gordonii str, cgakkus sybstr. CH1, Streptococcus suis 89/1591, Streptococcus infantarius subsp. infantarius ATCC BAA-102, Lactococcus lactis subsp cremoris MG1363, Lactococcus lactis, Leuconostoc mesenteroides subsp mesenteroides ATCC8293, Lactobacillus buchneri ATCC 11577, Staphylococcus haemolyticus JCSC1435, Staphylococcus epidermidis ATCC12228, Streptococcus pneumoniae CGSP14, Streptococcus pneumoniae TIGR4, Streptococcus sanguinis SK36, Streptococcus salivarius SK126, Streptococcus thermophilus LMD-9, Streptococcus pneumoniae CCRI 1974M2, Lactococcus lactis subsp. lactis II1403, Leuconostoc mesenteroides subsp cremoris ATCC19254, Leuconostoc mesenteroides subsp cremoris, Lactobacillus brevis subsp. gravesensis ATCC27305, or *Lactococcus lactis subsp lactis NCDO2118* or an active fragment thereof, wherein the heterologous polypeptide has a K_M for NADH less than about 50, (v) a heterologous polypeptide with KARI activity having at least about 90% identity or at least about 95% identity to SEQ ID NO: 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, or 135 or an active fragment thereof, wherein the heterologous polypeptide has a K_M for NADH less than about 50; (b) a heterologous polynucleotide encoding the heterologous polypeptide with KARI activity of (a); (c) reduced or eliminated aldehyde dehydrogenase activity; (d) reduced or eliminated aldehyde oxidase activity; (e) reduced or eliminated acetolactate reductase activity; or (f) a combination thereof.

In one embodiment, the recombinant host cell comprises reduced or eliminated aldehyde dehydrogenase expression activity and reduced or eliminated acetolactate reductase expression or activity.

In another embodiment, the recombinant host cell comprises (i) reduced or eliminated aldehyde dehydrogenase expression or activity or reduced or eliminated acetolactate reductase expression or activity and (ii)

a heterologous polynucleotide encoding a polypeptide having KARI activity and K_M for NADH less than 300 μM.

In another ebmobodiment, the recombinant host comprises a heterologous polypeptide with KARI activity that has at least about 90% or at least about 95% identity to SEQ ID NO: 27, 29, 141, 143, 275, or 277.

In some embodiments, the recombinant host cell comprises a heterologous polypeptide with KARI activity that comprises substitutions in amino acids corresponding to S56 and S58 of SEQ ID NO: 27. In some embodiments, the polypeptide with KARI activity further comprises a substitution of one or more of the amino acids corresponding to I86, N87, T131, or T191 of SEQ ID NO: 27. In some embodiments, the polypeptide with KARI activity having at least 90% identity or at least 95% identity to SEQ ID NO: 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, or 65.

In some embodiments, the recombinant host cell has an effective isobutanol productivity of at least about 3, at least about 4, or at least about 5 grams per gram of cells after about 48 hours, wherein at least the last about 24 of the 48 hours are under anaerobic conditions.

In some embodiments, the recombinant host cell comprises a heterologous polypeptide with KARI activity that has a K_M for NADH less than about 350, less than about 100, less than about 50, or less than about 10 μM at pH 6.8.

In some embodiments, the recombinant host cell comprises a heterologous polypeptide with KARI activity that has at least about 90% identity or at least about 95% identity to SEQ ID NO: 376, 382, 378, or 275.

In some embodiments, the recombinant host cell comprises a heterologous polypeptide with KARI activity comprises an amino acid substitution at one or more of the positions corresponding to amino acids A41, S56, S58, I87, T131, T191, R227, or Q246 of a KARI enzyme derived from *Anaerostipes caccae* (SEQ ID NO:27).

In some embodiments, the recombinant host cell comprises a heterologous polypeptide with KARI activity that comprises SEQ ID NO: 33 or SEQ ID NO:35 or an active fragment thereof.

In another embodiment, the recombinant host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity. In some embodiments, the polypeptide having aldehyde dehydrogenase activity catalyzes the conversion of isobutyraldehyde to isobutyric acid. In some embodiments, the polypeptide having aldehyde dehydrogenase activity corresponds to Enzyme Commission Number EC 1.21.3, EC 1.2.1.4, and/or EC1.2.1.5. In some embodiments, the host cell is *S. cerevisiae* and the polypeptide having aldehyde dehydrogenase activity is ALD2, ALD3, ALD4, ALD5, ALD6 or a homolog thereof. In some embodiments, the host cell *is K. lactis* and the polypeptide having aldehyde dehydrogenase activity is is KLLA0F00440, KLLA0E23057, KLLA0D10021, or KLLA0D09999G. In some embodiments, the host cell is *P. stipitis* and the polypeptide having aldehyde dehydrogenase activity is ALD2, ALD3, ALD4, ALD5, or ALD7. In some embodiments, the host cell is *Lactobacillus plantarum* and said polypeptide having aldehyde dehydrogenase activity is AldH. In some embodiments, the host cell is *E. coli* and the polypeptide having aldehyde dehydrogenase activity is aldA, aldB, or aldH.

In another embodiment, the host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde oxidase activity. In some embodiments, the polypeptide having aldehyde oxidase activity catalyzes the conversion of isobutyraldehyde to isobutyric acid. In some embodiments, the polypeptide having aldehyde oxidase activity corresponds to Enzyme Commission Number EC 1.2.3.1. In some embodiments, the polypeptide having aldehyde oxidase activity is AOX1 and/or AOX2.

In another embodiment, the host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide encoding a polypeptide having acetolactate reductase activity. In some embodiments,

the polypeptide having acetolactate reductase activity comprises a polypeptide encoded by a polynucleotide selected from the group consisting of SEQ ID NO: 676, SEQ ID NO: 678, SEQ ID NO: 680, SEQ ID NO: 682, SEQ ID NO: 684, SEQ ID NO: 686, SEQ ID NO: 688, SEQ ID NO: 690, SEQ ID NO: 692, SEQ ID NO: 694, SEQ ID NO: 696, SEQ ID NO:702, SEQ ID NO: 704, SEQ ID NO: 706, SEQ ID NO: 708, SEQ ID NO: 710, SEQ ID NO: 712, SEQ ID NO: 714, SEQ ID NO: 716, SEQ ID NO: 718, SEQ ID NO: 720, SEQ ID NO: 722, SEQ ID NO: 724, SEQ ID NO:726 , SEQ ID NO:728, and SEQ ID NO: 730. In some embodiments, the polypeptide having acetolactate reductase activity is YMR226C.

In another embodiment, the recombinant host cell is a yeast host cell. In some embodiments, the yeast is selected from the group consisting of *Saccharomyces, Schizosaccharomyces, Hansenula, Candida, Kluyveromyces, Yarrowia, Issatchenkia*, or *Pichia*. In some embodiments, the host cell is *Saccharomyces cerevisiae*.

In another embodiment, the host cell is a bacterial cell. In some embodiments, the bacterial cell is a *Clostridium, Zymomonas, Escherichia, Salmonella, Rhodococcus, Pseudomonas, Bacillus, Lactobacillus, Enterococcus, Pediococcus, Alcaligenes, Klebsiella, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Lactococcus, Leuconostoc, Oenococcus, Pediococcus*, or *Streptococcus* cell. In some embodiment, the bacterial cell is not *E. coli*.

In some embodiments, the engineered isobutanol production pathway of the recombinant host cell comprises the following substrate to product conversions: (a) pyruvate to acetolactate, (b) acetolactate to 2,3-dihydroxyisovalerate, (c) 2,3-dihydroxyisovalerate to 2-ketoisovalerate, (d) 2-ketoisovalerate to isobutyraldehyde, and (e) isobutyraldehyde to isobutanol and more than one of the substrate to product conversions is catalyzed by an enzyme that is heterologous to the host cell. In some embodiments, all of the substrate to product conversions are catalyzed by enzymes heterologous to the host cell. In some embodiments, at least one heterologous polynucleotide encoding an enzyme heterologous to the host cell is chromosomally integrated into the host cell. In some

embodiments, the substrate to product conversions are catalyzed by enzymes substantially localized to the cytosol. In some embodiments, the substrate to product conversion for isobutyraldehyde to isobutanol is catalyzed by an alcohol dehydrogenase enzyme that utilizes NADH as a cofactor. In some embodiments, the converstion of acetolactate to 2,3-dihydroxyisovalerate is catalyzed by a KARI that can use NADH as a cofactor.

In some embodiments, the host cell comprises the plasmid pLH702 or pLH701 or a plasmid having the same coding regions. In some embodiments, the host cell comprises the plasmid pBP915 or a plasmid having the same coding regions. In some embodiments, the host cell comprises the plasmid pYZ067ΔkivDΔhADH or a plasmid having the same coding regions.

In some embodiments, the host cell comprises reduced, disrupted, or eliminated ability to convert acetolactate to 2,3-dihydroxy-2-methylbutyrate.

In some embodiments, the host cell is yeast and has reduced or eliminated pyruvate decarboxylase expression or activity. In some embodiments, the host cell has reduced or eliminated PDC1, PDC5, or PDC6 activity or a combination thereof.

In some embodiments, the host cell has reduced or eliminated NAD-dependent glycerol-3-phosphate dehydrogenase expression or activity. In some embodiments, the host cell has reduced GPD2 activity.

In some embodients, the host cell has reduced or eliminated FRA2 expression or activity.

In some embodiments, the host cell produces isobutanol under anaerobic conditions and the molar ratio of isobutanol to glycerol is greater than 1.

In some embodiments, the polypeptide having ketol-acid reductoisomerase activity matches the profile HMM given provided in Table Z with a profile HMM E value of $<10^{-3}$.

In some embodiments, the host cell produces isobutanol at a yield greater than about 25%, about 50%, about 75%, or about 90% of theoretical yield.

In some embodiments, isobutanol and ethanol are produced.

Methods for producing isobutanol include methods comprising providing a recombinant host cell as described above and contacting the host cell with a carbon substrate under conditions whereby isobutanol is produced. In some embodiments, at least a portion of the contacting occurs under anaerobic conditions. In some embodiments, the contacting occurs in the presence of an extractant. In some embodiments, the contacting occurs in the presence of a sufficient quantity of organic extractant to form a two-phase system comprising an aqueous phase and an organic phase. In some embodiments, one or more of the effective rate, effective titer, or effective yield of isobutanol is increased as compared to methods using a recombinant host cell that does not comprise a heterologous polypeptide with KARI activity, a heterologous polynucleotide encoding a polypeptide with KARI activity, reduced or eliminated aldehyde dehydrogenase activity, reduced or eliminated aldehyde oxidase activity, reduced or eliminated acetolactate reductase activity, or a combination thereof. In some embodiments, one or more of the effective rate, effective titer, or effective yield of isobutanol is increased as compared to methods using a recombinant host cell that does not comprise (i) a heterologous polypeptide with KARI activity or a heterologous polynucleotide encoding a polypeptide with KARI activity and (ii) at least one modification that enchances performance of the engineered isobutanol production pathway. In some embodiments, DHMB production, isobutyric acid production, or both is reduced as compared to methods using a recombinant host cell that does not comprise a heterologous polypeptide with KARI activity, a heterologous polynucleotide encoding a polypeptide with KARI activity, reduced or eliminated aldehyde dehydrogenase activity, reduced or eliminated aldehyde oxidase activity, reduced or eliminated acetolactate reductase activity, or a combination thereof. In some embodiments, DHMB production, isobutyric

acid production, or both is reduced as compared to methods using a recombinant host cell that does not comprise (i) a heterologous polypeptide with KARI activity or a heterologous polynucleotide encoding a polypeptide with KARI activity and (ii) at least one modification that enchances performance of the engineered isobutanol production pathway. In some embodiments, the molar ratio of isobutanol to glycerol is greater than 1.

Methods for producing isobutanol also comprise providing a recombinant host cell that produces isobutanol and contacting the host cell with a carbon substrate under conditions whereby isobutanol is produced, wherein at least a portion of the contacting occurs under anaerobic conditions, and wherein the ratio of isobutanol to glycerol produced is greater than 1.

Methods for producing isobutanol also comprise growing a recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate under conditions whereby butanol is produced and removing DHMB from the culture.

Compositions produced by such methods are also provided herein. In some embodiments, the composition comprises isobutanol and a recombinant host cell provided above. In some embodiments, the composition comprises butanol and no more than about 0.5 mM DHMB.

Fermentative compositions are also provided herein. In some embodiments, a fermentative composition comprises the host cell and isobutanol produced according to the methods provided above.

Compositions comprising i) a recombinant yeast capable of producing butanol, ii) butanol, and iii) no more than about 0.5 mM DHMB are also provided.

Methods for producing a recombinant host cell are also provided. Such methods can comprise (a) providing a recombinant host cell comprising a modification in a polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity or aldehyde oxidase activity; and (b) transforming the host cell with a polynucleotide encoding a polypeptide of an isobutanol biosynthetic pathway.

Methods for reducing or eliminating the conversion of isbutyraldehye to isobutyric acid are also provided. Such methods can comprise (a) providing the recombinant host cell of any one of claims 1-52; and (b) subjecting the host cell to conditions wherein the conversion of isbutyraldehye to isobutyric acid is reduced or eliminated compared to methods using a recombinant host cell that does not comprise reduced or eliminated aldehyde dehydrogenase and/or aldehyde oxidase activity.

Certain polypeptides are also provided herein. In some embodiments, the polypeptides comprise at least about 90% identity or at least about 95% identity or at least about 99% identity to SEQ ID NO: 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 417, 419, 421, 423, 425, 427, 429, 431, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 624, 626, 628, 630, 632 or an active fragment thereof and have ketol-acid reductoisomerase activity. In some embodiments, the polypeptides comprise at least about 90% identity or at least about 95% identity or at least about 99% identity to SEQ ID NO: 417, 419, 421, 423, 425, or 427 or an active fragment thereof and have ketol-acid reductoisomerase activity. In some embodiments, a polypeptide comprises a sequence with at least about 90% identity, at least about 95% identity, or at least about 99% identity to SEQ ID NO: 927, 928, 196, 266, 267, 389, 405, 637, 781, 782, 783, 835, 853, 854, 855, 856, 857, or 859.

Polynucleotides encoding such polypeptides and host cells comprising such polynucleotides and polypeptides are also provided.

Methods of converting acetolactate to 2,3-dihydroxyisovalerate are also provided. For example, such methods can comprise (a) providing a polypeptide described above, and (b) contacting the polypeptide with

acetolactate under conditions wherein 2,3-dihydroxyisovalerate is produced.

Recombinant yeast cells are also provided herein. In some embodiments, a recombinant yeast comprises a biosynthetic pathway capable of converting pyruvate to acetolactate, and the yeast produces less than 0.01 moles 2,3-dihydroxy-2-methyl butyrate (DHMB) per mole of sugar consumed. In some embodiments, a recombinant yeast comprises capable of converting pyruvate to acetolactate, and the yeast produces DHMB at a rate of less than about 1.0 mM/hour. In some embodiments, a recombinant yeast comprises a biosynthetic pathway capable of converting pyruvate to acetolactate, and the yeast produces an amount of 2,3-dihydroxy-3-isovalerate (DHIV) that is at least about 1.5 times the amount of DHMB produced.

Methods of identifying a gene involved in DHMB production are also provided. In some embodiments, the methods comprise (i) providing a collection of yeast strains comprising at least two or more gene deletions; (ii) measuring the amount of DHMB produced by individual yeast strains; (iii) selecting a yeast strain that produces no more than about 1.0 mM DHMB/hour; and (iv) identifying the gene that is deleted in the selected yeast strain. In some embodiments, the methods comprise (i) providing a collection of yeast strains that over-express at least two or more genes; (ii) measuring the amount of DHMB produced by individual yeast strains; (iii) selecting a yeast strain that produces at least about 1.0 mM DHMB; and (iv) identifying the gene that is over-expressed in the selected yeast strain.

Methods for the production of butanol are also provided. In some embodiments, the methods comprise (a) growing a recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate under conditions whereby butanol is produced; and b) measuring DHIV and/or DHMB concentration. In some embodiments, the growing and measuring can be performed simultaneously or sequentially and in any order. In some embodiments, the measuring comprises liquid chromatography-mass spectrometry.

Methods for increasing ketol-acid reductoisomerase (KARI) activity are also provided. In some embodiments, the methods comprise (a) providing a composition comprising acetolactate, a KARI enzyme, and an acetolactate reductase enzyme and (b) decreasing DHMB levels. In some embodiments, decreasing DHMB levels is achieved by decreasing acetolactate reductase enzyme activity. In some embodiments, decreasing DHMB levels is achieved by removing DHMB from the composition.

In some embodiments, increasing KARI enzyme productivity in a host cell can comprises culturing a host cell, wherein the host cell comprises a heterologous KARI enzyme and at least one genetic modification that reduces, disrupts, or eliminates acetolactate reductase expression or activity, and wherein the KARI enzyme activity is decreased in the presence of DHMB. In some embodiments, the KARI has at least about 90%, at least about 95%, or at least about 99% identity to *E. coli* or *L. lactis* KARI. In some embodiments, the reduced, disrupted, or eliminated acetolactate reductase expression or activity substantially reduces the presence of DHMB.

Methods for increasing dihydroxyacid dehydratase (DHAD) activity are also provided. In some embodiments, the methods comprise (a) providing a composition comprising dihydroxyisovalerate (DHIV) and a DHAD enzyme and (b) decreasing DHMB levels.

BRIEF DESCRIPTION OF THE FIGURES AND INCORPORATION BY REFERENCE OF THE TABLE FILED ELECTRONICALLY HEREWITH

The invention can be more fully understood from the following detailed description, the Figures, and the accompanying sequence descriptions, which form part of this application.

FIGURE 1- Shows four different isobutanol biosynthetic pathways. The steps labeled "a", "b", "c", "d", "e", "f", "g", "h", "i", "j" and "k" represent the substrate to product conversions described below.

FIGURE 2 depicts an alignment of the amino acid sequences of the KARI from *Pseudomonas fluorescens* ("PF5"; SEQ ID NO: 5) and KARI

2017200439 20 Jan 2017

from *Anaerostipes caccae* ("K9"; SEQ ID NO: 27). The bolded positions are targeted for mutagenesis as described herein.

FIGURE 3A, 3B and 3C depict an alignment of the amino acid sequences of KARI enzymes from *Bifidobacterium angulatum* DSM 20098 ("K1"; SEQ ID NO: 141), *Bifidobacterium dentium* ATCC 27678 ("K2"; SEQ ID NO: 143), *Clostridium beijerinckii* NCIMB 8052 ("K7"; SEQ ID NO: 275), *Anaerostipes caccae* DSM 14662 ("K9"; SEQ ID NO: 27), *Enterococcus gallinarum* EG2 ("K25"SEQ ID NO: 376), *Streptococcus thermophilus* LMD-9 ("K26" SEQ ID NO: 121), *Lactococcus lactis* subsp. cremoris MG1363 ("K29"; SEQ ID NO: 382), *Zymomonas mobilis* ("S2"; SEQ ID NO: 277), and *Lactococcus lactis* ("LTS"; SEQ ID NO: 380). The bolded positions are targeted for mutagenesis as described herein.

FIGURE 4 is a plasmid map of pLH556 (pHR81-PIlv5-Pf5.KARI) vector (SEQ ID NO: 138).

FIGURE 5 shows the specific rate of isobutanol production, Qp, of the two strains, PNY1910 and PNY2242.

FIGURE 6 shows the accumulation of DHIV + DHMB in the culture supernatant during the fermentation time course with PNY1910 (triangles) and PNY2242 (diamonds). (DHMB and DHIV are not distinguished by the HPLC method used.)

FIGURE 7 shows the yield of glycerol, pyruvic acid, 2,3-butanediol (BDO), DHIV/DHMB, α-ketoisovalerate (aKIV), and isobutyric acid (iBuAc). DHIV and DHMB are shown together as these are not distinguished by the HPLC method used.

FIGURE 8 shows a summary of V_{max}/K_M values for K9G9 variants as described in Example 16.

FIGURES 9A, B, and C show isobutanol to glycerol molar yield ratios, isobutanol molar yields, and isobutanol titers for K9 variants as described in Example 19.

FIGURE 10 shows an isobutanol biosynthetic pathway. Step "a" represents the conversion of pyruvate to acetolactate. Step "b" represents the conversion of acetolactate to DHIV. Step "c" represents the conversion of DHIV to KIV. Step "d" represents the conversion of KIV to

2017200439 20 Jan 2017

isobutyraldehyde. Step "e" represents the conversion of isobutyraldehyde to isobutanol. Step "f" represents the conversion of acetolactate to DHMB.

FIGURE 11 shows a phyolgenetic tree of YMR226c homologs from species of ascomycete yeast. A filamentous fungi (*Neurospora crassa*) sequence is included as an outgroup.

FIGURE 12 shows a multiple sequence alignment (MSF Format) of nucleotide sequences of ORFs with homology to YMR226C. The gene names shown correspond to the accession numbers and SEQ ID NOs. given in Table 7. The alignment was produced by AlignX (Vector NTI).

FIGURE 13 shows a graph of the molar yield of DHMB over time.

FIGURE 14 depicts the production of isobutanol and isobutyric acid in yeast strain NYLA84.

Table Z- is a table (filed electronically herewith and incorporated by reference) of the Profile HMM of experimentally verified KARI enzymes listed in Table 1 and as described in US App. Pub. Nos. 2010/0197519 and 2009/0163376, which are herein incorporated by reference in their entireties.

Table 1. Experimentally verified KARI enzymes.

GI Number	Accession	SEQ ID NO:	Microorganism
70732562	YP_262325.1	5	*Pseudomonas fluorescens Pf-5*
15897495	NP_342100.1	1	*Sulfolobus solfataricus P2*
18313972	NP_560639.1	2	*Pyrobaculum aerophilum str. IM2*
76801743	YP_326751.1	7	*Natronomonas pharaonis DSM 2160*
16079881	NP_390707.1	8	*Bacillus subtilis subsp. subtilis str. 168*
19552493	NP_600495.1	9	*Corynebacterium glutamicum ATCC 13032*
6225553	O32414	10	*Phaeospririlum molischianum*
17546794	NP_520196.1	3	*Ralstonia solanacearum GMI1000*
56552037	YP_162876.1	11	*Zymomonas mobilis subsp. mobilis ZM4*
114319705	YP_741388.1	12	*Alkalilimnicola ehrlichei MLHE-1*
57240359	ZP_00368308.1	13	*Campylobacter lari RM2100*
120553816	YP_958167.1	14	*Marinobacter aquaeolei VT8*
71065099	YP_263826.1	15	*Psychrobacter arcticus 273-4*
83648555	YP_436990.1	16	*Hahella chejuensis KCTC 2396*
74318007	YP_315747.1	17	*Thiobacillus denitrificans ATCC 25259*

67159493	ZP_00420011.1	18	*Azotobacter vinelandii AvOP*
66044103	YP_233944.1	19	*Pseudomonas syringae pv. syringae B728a*
28868203	NP_790822.1	20	*Pseudomonas syringae pv. tomato str. DC3000*
26991362	NP_746787.1	21	*Pseudomonas putida KT2440*
104783656	YP_610154.1	22	*Pseudomonas entomophila L48*
146306044	YP_001186509.1	23	*Pseudomonas mendocina ymp*
15599888	NP_253382.1	4	*Pseudomonas aeruginosa PAO1*
42780593	NP_977840.1	24	*Bacillus cereus ATCC 10987*
42781005	NP_978252.1	25	*Bacillus cereus ATCC 10987*
266346	Q01292	6	*Spinacia oleracea*

The eleven positions in the profile HMM representing the columns in the alignment which correspond to the eleven cofactor switching positions in *Pseudomonas fluorescens Pf-5* KARI are identified as positions 24, 33, 47, 50, 52, 53, 61, 80, 115, 156, and 170. Table Z is submitted herewith electronically and is incorporated herein by reference.

The sequences provided in the sequence listing filed herewith (Name: 20120323_CL5367USNA_SEQLIST.txt; Size 2,003,893 bytes; Date of Creation March 23, 2012), is herein incorporated by reference.

Consistent with the World Intellectual Property Organization (WIPO) Standard ST.25 (2009), certain primers given in the sequence listing and herein use N to represent nucleotides a or g or c or t. K is used to represent g or t. M is used to represent a or c.

DETAILED DESCRIPTION OF THE INVENTION

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. In case of conflict, the present application including the definitions will control. Also, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular. All publications, patents and other references mentioned herein are incorporated by reference in their entireties for all purposes as if each individual publication or patent application were specifically and individually indicated to be incorporated

2017200439 20 Jan 2017

by reference, unless only specific sections of patents or patent publications are indicated to be incorporated by reference.

Although methods and materials similar or equivalent to those disclosed herein can be used in practice or testing of the present invention, suitable methods and materials are disclosed below. The materials, methods and examples are illustrative only and are not intended to be limiting. Other features and advantages of the invention will be apparent from the detailed description and from the claims.

The last step in the biosynthesis of isobutanol via a pyruvate-utilizing biosynthetic pathway is the conversion of isobutyraldehyde to isobutanol (Figure 1). A side reaction in this pathway is the conversion of isobutyraldehyde to isobutyric acid which results in reduced amounts of isobutyraldehyde available to convert into isobutanol and reduced isobutanol yield. For an efficient biosynthetic process, there is a need to prevent the conversion of isobutyraldehyde to isobutyric acid such that increased amounts of isobutyraldehyde are available for conversion to isobutanol and isobutanol yields are increased.

Aldehyde dehydrogenases are a family of enzymes that catalyze the oxidation (dehydrogenation) of aldehydes (Wang *et al.*, *J. Bacteriol.* 180:822-30, 1998; Navarro-Avino *et al.*, *Yeast* 15:829-42, 1999; and Saint-Prix *et al.*, *Microbiology* 150:2209-20, 2004). There is a need to identify suitable aldehyde dehydrogenases that can be modified to reduce or eliminate aldehyde dehydrogenase activity, and can reduce or eliminate the conversion of isobutyraldehyde to isobutyric acid, such that increased amounts of isobutyraldehyde are available for conversion to isobutanol and isobutanol yields are increased.

Aldehyde oxidases are a family of enzymes that catalyze the production of carboxylic acids from aldehydes (Nomura *et al., Biosci. Biotechnol. Biochem.* 62:1134-7, 1998; and Johnson *et al.*, *Genetics* 151:1379-1391, 1999). There is a need to identify suitable aldehyde oxidases that can be modified to reduce or eliminate aldehyde oxidase activity and can reduce or eliminate the conversion of isobutyraldehyde to

isobutyric acid, such that increased amounts of isobutyraldehyde are available for conversion to isobutanol and isobutanol yields are increased.

The biosynthesis pathway for the production of butanol in genetically engineered yeast includes the conversion of acetolactate to 2,3-dihydroxy-3-isovalerate (DHIV), which is subsequently converted to butanol. *See* Figure 10. However, a side reaction in this pathway, which decreases the overall production of butanol, is the conversion of acetolactate to 2,3-dihydroxy-2-methylbutyrate (DHMB). In fact, Applicants have discovered that DHMB has inhibitory effects on enzymes (dihydroxyacid dehydratase and ketol-acid reductoisomerase) in an isobutanol production pathway. For an efficient biosynthetic process, there is a need to prevent the conversion of acetolactate to DHMB.

Applicants have solved the stated problems by providing recombinant yeast host cells comprising an isobutanol biosynthetic pathway; and at least one of: i) reduced or eliminated aldehyde dehydrogenase activity ii) reduced or eliminated aldehyde oxidase activity iii) reduced or eliminated acetolactate reductase activity; iv) a heterologous polynucleotide encoding a polypeptide having ketol-acid reductoisomerase activity; and v) a heterologous polypeptide having ketol-acid reductoisomerase activity. Further, Applicants provide methods of producing butanol utilizing such host cells. Such recombinant host cells can be used to increase the production of a product of a biosynthetic pathway (e.g., isobutanol, 1-butanol, or 2-butanol) and/or reduce or eliminate the conversion of pathway intermediates to undesirable byproducts. Applicants have also provided a suitable screening strategy for evaluating various candidate enzymes. The identified enzymes can be altered to enhance the production of a product of a biosynthetic pathway (e.g., isobutanol, 1-butanol, or 2-butanol) and/or reduce or eliminate the conversion of pathway intermediates to undesirable byproducts.

In order to further define this invention, the following terms, abbreviations and definitions are provided.

It will be understood that "derived from" with reference to polypeptides disclosed herein encompasses sequences synthesized

based on the amino acid sequences of the KARIs present in the indicated organisms as well as those cloned directly from the organism's genetic material.

As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having," "contains" or "containing," or any other variation thereof, will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers and are intended to be non-exclusive or open-ended. For example, a composition, a mixture, a process, a method, an article, or an apparatus that comprises a list of elements is not necessarily limited to only those elements but can include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

As used herein, the term "consists of," or variations such as "consist of" or "consisting of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, but that no additional integer or group of integers can be added to the specified method, structure, or composition.

As used herein, the term "consists essentially of," or variations such as "consist essentially of" or "consisting essentially of," as used throughout the specification and claims, indicate the inclusion of any recited integer or group of integers, and the optional inclusion of any recited integer or group of integers that do not materially change the basic or novel properties of the specified method, structure or composition. *See* M.P.E.P. § 2111.03.

Also, the indefinite articles "a" and "an" preceding an element or component of the invention are intended to be nonrestrictive regarding the number of instances, *i.e.*, occurrences of the element or component. Therefore "a" or "an" should be read to include one or at least one, and the

singular word form of the element or component also includes the plural unless the number is obviously meant to be singular.

The term "invention" or "present invention" as used herein is a non-limiting term and is not intended to refer to any single embodiment of the particular invention but encompasses all possible embodiments as described in the claims as presented or as later amended and supplemented, or in the specification.

As used herein, the term "about" modifying the quantity of an ingredient or reactant of the invention employed refers to variation in the numerical quantity that can occur, for example, through typical measuring and liquid handling procedures used for making concentrates or solutions in the real world; through inadvertent error in these procedures; through differences in the manufacture, source, or purity of the ingredients employed to make the compositions or to carry out the methods; and the like. The term "about" also encompasses amounts that differ due to different equilibrium conditions for a composition resulting from a particular initial mixture. Whether or not modified by the term "about", the claims include equivalents to the quantities. In one embodiment, the term "about" means within 10% of the reported numerical value, or within 5% of the reported numerical value.

As used herein, “synergistic” refers to a greater-than-additive effect produced by a combination (i.e., an effect that is greater than the sum of individual effects) or an additive effect when the individual effects are not expected to be additive. The term also refers to the addition of one compound which results in less of another compound being required.

The term "butanol biosynthetic pathway"" as used herein refers to the enzymatic pathway to produce 1-butanol, 2-butanol, or isobutanol. For example, isobutanol biosynthetic pathways are disclosed in U.S. Patent Application Publication No. 2007/0092957, which is incorporated by reference herein.

The term "isobutanol biosynthetic pathway" refers to the enzymatic pathway to produce isobutanol. Certain isobutanol biosynthetic pathways are illustrated in Figure 1 and described herein. From time to time

"isobutanol biosynthetic pathway" is used synonymously with "isobutanol production pathway".

The term "butanol" as used herein refers to 2-butanol, 1-butanol, isobutanol or mixtures thereof. Isobutanol is also known as 2-methyl-1-propanol.

A recombinant host cell comprising an "engineered alcohol production pathway" (such as an engineered butanol or isobutanol production pathway) refers to a host cell containing a modified pathway that produces alcohol in a manner different than that normally present in the host cell. Such differences include production of an alcohol not typically produced by the host cell, or increased or more efficient production.

The term "heterologous biosynthetic pathway" as used herein refers to an enzyme pathway to produce a product in which at least one of the enzymes is not endogenous to the host cell containing the biosynthetic pathway.

The term "extractant" as used herein refers to one or more organic solvents which can be used to extract butanol from a fermentation broth.

The term "effective isobutanol productivity" as used herein refers to the total amount in grams of isobutanol produced per gram of cells.

The term "effective titer" as used herein, refers to the total amount of a particular alcohol (e.g. butanol) produced by fermentation per liter of fermentation medium. The total amount of butanol includes: (i) the amount of butanol in the fermentation medium; (ii) the amount of butanol recovered from the organic extractant; and (iii) the amount of butanol recovered from the gas phase, if gas stripping is used.

The term "effective rate" as used herein, refers to the total amount of butanol produced by fermentation per liter of fermentation medium per hour of fermentation.

The term "effective yield" as used herein, refers to the amount of butanol produced per unit of fermentable carbon substrate consumed by the biocatalyst.

The term "separation" as used herein is synonymous with "recovery" and refers to removing a chemical compound from an initial mixture to obtain the compound in greater purity or at a higher concentration than the purity or concentration of the compound in the initial mixture.

The term "aqueous phase," as used herein, refers to the aqueous phase of a biphasic mixture obtained by contacting a fermentation broth with a water-immiscible organic extractant. In an embodiment of a process described herein that includes fermentative extraction, the term "fermentation broth" then specifically refers to the aqueous phase in biphasic fermentative extraction.

The term "organic phase," as used herein, refers to the non-aqueous phase of a biphasic mixture obtained by contacting a fermentation broth with a water-immiscible organic extractant.

The terms "PDC-," "PDC knockout," or "PDC-KO" as used herein refer to a cell that has a genetic modification to inactivate or reduce expression of a gene encoding pyruvate decarboxylase (PDC) so that the cell substantially or completely lacks pyruvate decarboxylase enzyme activity. If the cell has more than one expressed (active) PDC gene, then each of the active PDC genes can be inactivated or have minimal expression thereby producing a PDC- cell.

The term "polynucleotide" is intended to encompass a singular nucleic acid as well as plural nucleic acids, and refers to a nucleic acid molecule or construct, e.g., messenger RNA (mRNA) or plasmid DNA (pDNA). A polynucleotide can contain the nucleotide sequence of the full-length cDNA sequence, or a fragment thereof, including the untranslated 5' and 3' sequences and the coding sequences. The polynucleotide can be composed of any polyribonucleotide or polydeoxyribonucleotide, which can be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that can be

single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. "Polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

A polynucleotide sequence can be referred to as "isolated," in which it has been removed from its native environment. For example, a heterologous polynucleotide encoding a polypeptide or polypeptide fragment having dihydroxy-acid dehydratase activity contained in a vector is considered isolated for the purposes of the present invention. Further examples of an isolated polynucleotide include recombinant polynucleotides maintained in heterologous host cells or purified (partially or substantially) polynucleotides in solution. Isolated polynucleotides or nucleic acids according to the present invention further include such molecules produced synthetically. An isolated polynucleotide fragment in the form of a polymer of DNA can be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

The term "NAD(P)H consumption assay" refers to an enzyme assay for the determination of the specific activity of the KARI enzyme, involving measuring the disappearance of the KARI cofactor, NAD(P)H, from the enzyme reaction. Such assays are described in Aulabaugh and Schloss, Biochemistry 29: 2824-2830, 1990, which is herein incorporated by reference in its entirety.

The term “NAD(P)H” refers to either NADH or NADPH.

"KARI" is the abbreviation for the enzyme ketol-acid reducto-isomerase.

The term "close proximity" when referring to the position of various amino acid residues of a KARI enzyme with respect to the adenosyl 2'-phosphate of NADPH means amino acids in the three-dimensional model for the structure of the enzyme that are within about 4.5 Å of the phosphorus atom of the adenosyl 2'-phosphate of NADPH bound to the enzyme.

The term "ketol-acid reductoisomerase" (abbreviated "KARI"), and "acetohydroxy acid isomeroreductase" will be used interchangeably and refer to enzymes capable of catalyzing the reaction of (S)-acetolactate to

2,3-dihydroxyisovalerate, classified as EC number EC 1.1.1.86 (*Enzyme Nomenclature* 1992, Academic Press, San Diego). As used herein the term "Class I ketol-acid reductoisomerase enzyme" means the short form that typically has between 330 and 340 amino acid residues, and is distinct from the long form, called class II, that typically has approximately 490 residues. These enzymes are available from a number of sources, including, but not limited to *E. coli* (amino acid SEQ ID NO: 942; GenBank Accession Number NC_000913 REGION: 3955993.3957468), *Vibrio cholerae* (GenBank Accession Number NC_002505 REGION: 157441.158925), *Pseudomonas aeruginosa*, (GenBank Accession Number NC_002516, REGION: 5272455.5273471), and *Pseudomonas fluorescens* (amino acid SEQ ID NO: 943; GenBank Accession Number NC_004129 REGION: 6017379.6018395). KARI enzymes are described for example, in U.S. Patent Nos. 7,910, 342 and 8,129,162 and U.S. Pub, App, No.2010/0197519, all of which are herein incorporated by reference in their entireties.

KARI is found in a variety of organisms and amino acid sequence comparisons across species have revealed that there are 2 types of this enzyme: a short form (class I) found in fungi and most bacteria, and a long form (class II) typical of plants. Class I KARIs typically have between 330-340 amino acid residues. The long form KARI enzymes have about 490 amino acid residues. However, some bacteria such as *Escherichia coli* possess a long form, where the amino acid sequence differs appreciably from that found in plants. KARI is encoded by the ilvC gene and is an essential enzyme for growth of *E. coli* and other bacteria in a minimal medium. Class II KARIs generally consist of a 225-residue N-terminal domain and a 287-residue C-terminal domain. The N-terminal domain, which contains the NADPH-binding site, has an αβstructure and resembles domains found in other pyridine nucleotide-dependent oxidoreductases. The C-terminal domain consists almost entirely of α-helices.

Ketol-acid reductoisomerase (KARI) enzymes are useful in pathways for the production of isobutanol using engineered

microorganisms (US Patents 7,851,188 and 7,993,889, incorporated by reference herein).

A KARI that can utilize NADH can capitalize on the NADH produced by the existing glycolytic and other metabolic pathways in most commonly used microbial cells and can result in improved isobutanol production. Rane et al. (Arch. Biochem. Biophys., 338: 83-89, 1997) discusses cofactor switching of a ketol acid reductoisomerase isolated from *E. coli*. US Appl. Pub. Nos. 2009/0163376 and 2010/0197519 (each of which is herein incorporated by reference it its entirety) describe the generation of KARI enzymes which can use NADH. US Appl. Pub. No. 2010/0143997 (which is herein incorporated by reference in its entirety) describes *E. coli* variants with improved K_M values for NADH.

The terms "ketol-acid reductoisomerase activity" and "KARI activity" refer to the ability to catalyze the substrate to product conversion (S)-acetolactate to 2,3-dihydroxyisovalerate.

The term "acetolactate synthase" refers to an enzyme that catalyzes the conversion of pyruvate to acetolactate and CO_2. Acetolactate has two stereoisomers ((*R*) and (*S*)); the enzyme prefers the (*S*)- isomer, which is made by biological systems. Certain acetolactate synthases are known by the EC number 2.2.1.6 (*Enzyme Nomenclature 1992,* Academic Press, San Diego). These enzymes are available from a number of sources, including, but not limited to, *Bacillus subtilis* (GenBank Nos: CAB15618, Z99122, NCBI (National Center for Biotechnology Information) amino acid sequence, NCBI nucleotide sequence, respectively), *Klebsiella pneumoniae* (GenBank Nos: AAA25079, M73842 and *Lactococcus lactis* (GenBank Nos: AAA25161, L16975).

The term "acetohydroxy acid dehydratase" refers to an enzyme that catalyzes the conversion of 2,3-dihydroxyisovalerate to α-ketoisovalerate. Certain acetohydroxy acid dehydratases are known by the EC number 4.2.1.9. These enzymes are available from a vast array of microorganisms, including, but not limited to, *E. coli* (GenBank Nos: YP_026248, NC_000913, *S. cerevisiae* (GenBank Nos: NP_012550, NC_001142), *M. maripaludis* (GenBank Nos: CAF29874, BX957219), *B.*

subtilis (GenBank Nos: CAB14105, Z99115), *Lactococcus lactis* (SEQ ID NO: 926), and *Streptococcus mutans* (SEQ ID NO: 939).

The term "branched-chain α-keto acid decarboxylase" refers to an enzyme that catalyzes the conversion of α-ketoisovalerate to isobutyraldehyde and CO_2. Certain branched-chain α-keto acid decarboxylases are known by the EC number 4.1.1.72 and are available from a number of sources, including, but not limited to, *Lactococcus lactis* (GenBank Nos: AAS49166, AY548760; CAG34226, AJ746364, *Salmonella typhimurium* (GenBank Nos: NP-461346, NC-003197), *Clostridium acetobutylicum* (GenBank Nos: NP-149189, NC-001988), *Macrococcus caseolyticus* (SEQ ID NO: 940), and *Listeria grayi* (SEQ ID NO: 941).

The term "branched-chain alcohol dehydrogenase" refers to an enzyme that catalyzes the conversion of isobutyraldehyde to isobutanol. Certain branched-chain alcohol dehydrogenases are known by the EC number 1.1.1.265, but can also be classified under other alcohol dehydrogenases (specifically, EC 1.1.1.1 or 1.1.1.2). These enzymes utilize NADH (reduced nicotinamide adenine dinucleotide) and/or NADPH as electron donor and are available from a number of sources, including, but not limited to, *S. cerevisiae* (GenBank Nos: NP_010656, NC_001136; NP_014051, NC_001145), *E. coli* (GenBank No: NP_417484), *C. acetobutylicum* (GenBank Nos: NP_349892, NC_003030), *B. indica* (amino acid SEQ ID NO: 945), *A. xylosoxidans* (amino acid SEQ ID NO: 944).

The term "branched-chain keto acid dehydrogenase" refers to an enzyme that catalyzes the conversion of α-ketoisovalerate to isobutyryl-CoA (isobutyryl-cofactor A), using NAD^+ (nicotinamide adenine dinucleotide) as electron acceptor. Certain branched-chain keto acid dehydrogenases are known by the EC number 1.2.4.4. These branched-chain keto acid dehydrogenases comprise four subunits, and sequences from all subunits are available from a vast array of microorganisms, including, but not limited to, *B. subtilis* (GenBank Nos: CAB14336, Z99116; CAB14335, Z99116; CAB14334, Z99116; and CAB14337,

Z99116) and *Pseudomonas putida* (GenBank Nos: AAA65614, M57613; AAA65615, M57613; AAA65617, M57613; and AAA65618, M57613).

As used herein, "aldehyde dehydrogenase activity" refers to any polypeptide having a biological function of an aldehyde dehydrogenase, including the examples provided herein. Such polypeptides include a polypeptide that catalyzes the oxidation (dehydrogenation) of aldehydes. Such polypeptides include a polypeptide that catalyzes the conversion of isobutyraldehyde to isobutyric acid. Such polypeptides also include a polypeptide that corresponds to Enzyme Commission Numbers EC 1.2.1.3, EC 1.2.1.4 or EC 1.2.1.5. Such polypeptides can be determined by methods well known in the art and disclosed herein.

As used herein, "aldehyde oxidase activity" refers to any polypeptide having a biological function of an aldehyde oxidase, including the examples provided herein. Such polypeptides include a polypeptide that catalyzes carboxylic acids from aldehydes. Such polypeptides include a polypeptide that catalyzes the conversion of isobutyraldehyde to isobutyric acid. Such polypeptides also include a polypeptide that corresponds to Enzyme Commission Number EC 1.2.3.1. Such polypeptides can be determined by methods well known in the art and disclosed herein.

As used herein, "pyruvate decarboxylase activity" refers to the activity of any polypeptide having a biological function of a pyruvate decarboxylase enzyme, including the examples provided herein. Such polypeptides include a polypeptide that catalyzes the conversion of pyruvate to acetaldehyde. Such polypeptides also include a polypeptide that corresponds to Enzyme Commission Number 4.1.1.1. Such polypeptides can be determined by methods well known in the art and disclosed herein. A polypeptide having pyruvate decarboxylate activity can be, by way of example, PDC1, PDC5, PDC6, or any combination thereof.

As used herein, "acetolactate reductase activity" refers to the activity of any polypeptide having the ability to catalyze the conversion of

acetolactate to DHMB. Such polypeptides can be determined by methods well known in the art and disclosed herein.

As used herein, "DHMB" refers to 2,3-dihydroxy-2-methyl butyrate. DHMB includes "fast DHMB," which has the 2S, 3S configuration, and "slow DHMB," which has the 2S, 3R configurate. See Kaneko et al., Phytochemistry 39: 115-120 (1995), which is herein incorporated by reference in its entirety and refers to fast DHMB as angliceric acid and slow DHMB as tigliceric acid.

As used herein, "reduced activity" refers to any measurable decrease in a known biological activity of a polypeptide when compared to the same biological activity of the polypeptide prior to the change resulting in the reduced activity. Such a change can include a modification of a polypeptide or a polynucleotide encoding a polypeptide as described herein. A reduced activity of a polypeptide disclosed herein can be determined by methods well known in the art and disclosed herein.

As used herein, "eliminated activity" refers to the complete abolishment of a known biological activity of a polypeptide when compared to the same biological activity of the polypeptide prior to the change resulting in the eliminated activity. Such a change can include a modification of a polypeptide or a polynucleotide encoding a polypeptide as described herein. An eliminated activity includes a biological activity of a polypeptide that is not measurable when compared to the same biological activity of the polypeptide prior to the change resulting in the eliminated activity. An eliminated activity of a polypeptide disclosed herein can be determined by methods well known in the art and disclosed herein.

The term "carbon substrate" or "fermentable carbon substrate" refers to a carbon source capable of being metabolized by host organisms of the present invention and particularly carbon sources selected from the group consisting of monosaccharides, oligosaccharides, polysaccharides, and one-carbon substrates or mixtures thereof. Non-limiting examples of carbon substrates are provided herein and include, but are not limited to, monosaccharides, oligosaccharides, polysaccharides, ethanol, lactate, succinate, glycerol, carbon dioxide, methanol, glucose, fructose, sucrose,

xylose, arabinose, dextrose, or mixtures thereof. Other carbon substrates can include ethanol, lactate, succinate, or glycerol.

"Fermentation broth" as used herein means the mixture of water, sugars (fermentable carbon sources), dissolved solids (if present), microorganisms producing alcohol, product alcohol and all other constituents of the material held in the fermentation vessel in which product alcohol is being made by the reaction of sugars to alcohol, water and carbon dioxide (CO_2) by the microorganisms present. From time to time, as used herein the term "fermentation medium" and "fermented mixture" can be used synonymously with "fermentation broth".

"Biomass" as used herein refers to a natural product containing a hydrolysable starch that provides a fermentable sugar, including any cellulosic or lignocellulosic material and materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides, disaccharides, and/or monosaccharides. Biomass can also comprise additional components, such as protein and/or lipids. Biomass can be derived from a single source, or biomass can comprise a mixture derived from more than one source. For example, biomass can comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood, and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, rye, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof.

"Feedstock" as used herein means a product containing a fermentable carbon source. Suitable feedstock include, but are not limited to, rye, wheat, corn, sugar cane, and mixtures thereof.

The term "aerobic conditions" as used herein means growth conditions in the presence of oxygen.

The term "microaerobic conditions" as used herein means growth conditions with low levels of oxygen (i.e., below normal atmospheric oxygen levels).

The term "anaerobic conditions" as used herein means growth conditions in the absence of oxygen.

The term "specific activity" as used herein is defined as the units of activity in a given amount of protein. Thus, the specific activity is not directly measured but is calculated by dividing 1) the activity in units/ml of the enzyme sample by 2) the concentration of protein in that sample, so the specific activity is expressed as units/mg, where an enzyme unit is defined as moles of product formed/minute. The specific activity of a sample of pure, fully active enzyme is a characteristic of that enzyme. The specific activity of a sample of a mixture of proteins is a measure of the relative fraction of protein in that sample that is composed of the active enzyme of interest.

The terms "k_{cat}" and "K_M" are known to those skilled in the art and are described in Enzyme Structure and Mechanism, 2nd ed. (Ferst; W.H. Freeman Press, NY, 1985; pp 98-120). K_M , the Michaelis constant, is the concentration of substrate that leads to half-maximal velocity. The term "k_{cat}", often called the "turnover number", is defined as the maximum number of substrate molecules converted to products per active site per unit time, or the number of times the enzyme turns over per unit time. $k_{cat} = V_{max}/[E]$, where [E] is the enzyme concentration (Ferst, *supra*). The terms "total turnover" and "total turnover number" are used herein to refer to the amount of product formed by the reaction of a KARI enzyme with substrate.

The term "catalytic efficiency" is defined as the k_{cat}/K_M of an enzyme. Catalytic efficiency is used to quantify the specificity of an enzyme for a substrate.

The term "isolated nucleic acid molecule", "isolated nucleic acid fragment" and "genetic construct" will be used interchangeably and will mean a polymer of RNA or DNA that is single- or double-stranded,

2017200439 20 Jan 2017

optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA can be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

The term "amino acid" refers to the basic chemical structural unit of a protein or polypeptide. The following abbreviations are used herein to identify specific amino acids:

Amino Acid	Three-Letter Abbreviation	One-Letter Abbreviation
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glycine	Gly	G
Histidine	His	H
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

The term "gene" refers to a nucleic acid fragment that is capable of being expressed as a specific protein, optionally including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and

2017200439 20 Jan 2017

coding sequences that are not found together in nature. Accordingly, a chimeric gene can comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of a microorganism. A "foreign" gene refers to a gene not normally found in the host microorganism, but that is introduced into the host microorganism by gene transfer. Foreign genes can comprise native genes inserted into a non-native microorganism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

As used herein, "native" refers to the form of a polynucleotide, gene, or polypeptide as found in nature with its own regulatory sequences, if present.

As used herein the term "coding sequence" or "coding region" refers to a DNA sequence that encodes for a specific amino acid sequence. "Suitable regulatory sequences" refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences can include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing sites, effector binding sites and stem-loop structures.

As used herein, "endogenous" refers to the native form of a polynucleotide, gene or polypeptide in its natural location in the organism or in the genome of an organism. "Endogenous polynucleotide" includes a native polynucleotide in its natural location in the genome of an organism. "Endogenous gene" includes a native gene in its natural location in the genome of an organism. "Endogenous polypeptide" includes a native polypeptide in its natural location in the organism transcribed and translated from a native polynucleotide or gene in its natural location in the genome of an organism.

The term "heterologous" when used in reference to a polynucleotide, a gene, or a polypeptide refers to a polynucleotide, gene, or polypeptide not normally found in the host organism. "Heterologous" also includes a native coding region, or portion thereof, that is reintroduced into the source organism in a form that is different from the corresponding native gene, *e.g.*, not in its natural location in the organism's genome. The heterologous polynucleotide or gene can be introduced into the host organism by, *e.g.*, gene transfer. A heterologous gene can include a native coding region with non-native regulatory regions that is reintroduced into the native host. For example, a heterologous gene can include a native coding region that is a portion of a chimeric gene including non-native regulatory regions that is reintroduced into the native host. "Heterologous polypeptide" includes a native polypeptide that is reintroduced into the source organism in a form that is different from the corresponding native polypeptide.

A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

As used herein, the term "modification" refers to a change in a polynucleotide disclosed herein that results in reduced or eliminated activity of a polypeptide encoded by the polynucleotide, as well as a change in a polypeptide disclosed herein that results in reduced or eliminated activity of the polypeptide. Such changes can be made by methods well known in the art, including, but not limited to, deleting, mutating (e.g., spontaneous mutagenesis, random mutagenesis, mutagenesis caused by mutator genes, or transposon mutagenesis), substituting, inserting, down-regulating, altering the cellular location, altering the state of the polynucleotide or polypeptide (e.g., methylation, phosphorylation or ubiquitination), removing a cofactor, introduction of an antisense RNA/DNA, introduction of an interfering RNA/DNA, chemical modification, covalent modification, irradiation with UV or X-rays, homologous recombination, mitotic recombination, promoter replacement methods, and/or combinations thereof. Guidance in determining which nucleotides or amino acid residues can be modified, can be found by

comparing the sequence of the particular polynucleotide or polypeptide with that of homologous polynucleotides or polypeptides, e.g., yeast or bacterial, and maximizing the number of modifications made in regions of high homology (conserved regions) or consensus sequences.

The term "recombinant genetic expression element" refers to a nucleic acid fragment that expresses one or more specific proteins, including regulatory sequences preceding (5' non-coding sequences) and following (3' termination sequences) coding sequences for the proteins. A chimeric gene is a recombinant genetic expression element. The coding regions of an operon can form a recombinant genetic expression element, along with an operably linked promoter and termination region.

"Regulatory sequences" refers to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences can include promoters, enhancers, operators, repressors, transcription termination signals, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure.

The term "promoter" refers to a nucleic acid sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters can be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic nucleic acid segments. It is understood by those skilled in the art that different promoters can direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". "Inducible promoters," on the other hand, cause a gene to be expressed when the promoter is induced or turned on by a promoter-

2017200439 20 Jan 2017

specific signal or molecule. It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths can have identical promoter activity. For example, it will be understood that "FBA1 promoter" can be used to refer to a fragment derived from the promoter region of the FBA1 gene.

The term "terminator" as used herein refers to DNA sequences located downstream of a coding sequence. This includes polyadenylation recognition sequences and other sequences encoding regulatory signals capable of affecting mRNA processing or gene expression. The polyadenylation signal is usually characterized by affecting the addition of polyadenylic acid tracts to the 3' end of the mRNA precursor. The 3' region can influence the transcription, RNA processing or stability, or translation of the associated coding sequence. It is recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths can have identical terminator activity. For example, it will be understood that "CYC1 terminator" can be used to refer to a fragment derived from the terminator region of the CYC1 gene.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of effecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression can also refer to translation of mRNA into a polypeptide.

The term "overexpression," as used herein, refers to expression that is higher than endogenous expression of the same or related gene. A heterologous gene is overexpressed if its expression is higher than that of

a comparable endogenous gene. The term overexpression refers to an increase in the level of nucleic acid or protein in a host cell. Thus, overexpression can result from increasing the level of transcription or translation of an endogenous sequence in a host cell or can result from the introduction of a heterologous sequence into a host cell. Overexpression can also result from increasing the stability of a nucleic acid or protein sequence.

As used herein the term "transformation" refers to the transfer of a nucleic acid fragment into the genome of a host microorganism, resulting in genetically stable inheritance. Host microorganisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" microorganisms.

The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements can be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

The term "site-saturation library" refers to a library which contains random substitutions at a specific amino acid position with up to and including all 20 possible amino acids at once.

2017200439 20 Jan 2017

The term "error-prone PCR" refers to adding random copying errors by imposing imperfect or 'sloppy' PCR reaction conditions which generate randomized libraries of mutations in a specific nucleotide sequence.

As used herein the term "codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without affecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The term "codon-optimized" as it refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA. Such optimization includes replacing at least one, or more than one, or a significant number, of codons with one or more codons that are more frequently used in the genes of that organism.

Deviations in the nucleotide sequence that comprise the codons encoding the amino acids of any polypeptide chain allow for variations in the sequence coding for the gene. Since each codon consists of three nucleotides, and the nucleotides comprising DNA are restricted to four specific bases, there are 64 possible combinations of nucleotides, 61 of which encode amino acids (the remaining three codons encode signals ending translation). The "genetic code" which shows which codons encode which amino acids is reproduced herein as Table 2A. As a result, many amino acids are designated by more than one codon. For example, the amino acids alanine and proline are coded for by four triplets, serine and arginine by six, whereas tryptophan and methionine are coded by just one triplet. This degeneracy allows for DNA base composition to vary over a wide range without altering the amino acid sequence of the proteins encoded by the DNA.

2017200439 20 Jan 2017

Table 2A. The Standard Genetic Code

	T	C	A	G
T	TTT Phe (F) TTC " TTA Leu (L) TTG "	TCT Ser (S) TCC " TCA " TCG "	TAT Tyr (Y) TAC " TAA **Stop** TAG **Stop**	TGT Cys (C) TGC TGA **Stop** TGG Trp (W)
C	CTT Leu (L) CTC " CTA " CTG "	CCT Pro (P) CCC " CCA " CCG "	CAT His (H) CAC " CAA Gln (Q) CAG "	CGT Arg (R) CGC " CGA " CGG "
A	ATT Ile (I) ATC " ATA " **ATG** Met (M)	ACT Thr (T) ACC " ACA " ACG "	AAT Asn (N) AAC " AAA Lys (K) AAG "	AGT Ser (S) AGC " AGA Arg (R) AGG "
G	GTT Val (V) GTC " GTA " GTG "	GCT Ala (A) GCC " GCA " GCG "	GAT Asp (D) GAC " GAA Glu (E) GAG "	GGT Gly (G) GGC " GGA " GGG "

Many organisms display a bias for use of particular codons to code for insertion of a particular amino acid in a growing peptide chain. Codon preference, or codon bias, differences in codon usage between organisms, is afforded by degeneracy of the genetic code, and is well documented among many organisms. Codon bias often correlates with the efficiency of translation of messenger RNA (mRNA), which is in turn believed to be dependent on, *inter alia*, the properties of the codons being translated and the availability of particular transfer RNA (tRNA) molecules. The predominance of selected tRNAs in a cell is generally a reflection of the codons used most frequently in peptide synthesis. Accordingly, genes can be tailored for optimal gene expression in a given organism based on codon optimization.

Given the large number of gene sequences available for a wide variety of animal, plant and microbial species, it is possible to calculate the relative frequencies of codon usage. Codon usage tables are readily available, for example, at the "Codon Usage Database" available at www.kazusa.or.jp/codon/ (visited March 20, 2008), and these tables can be adapted in a number of ways. *See* Nakamura, Y., *et al. Nucl. Acids Res. 28*:292 (2000). Codon usage tables for yeast, calculated from

2017200439 20 Jan 2017

GenBank Release 128.0 [15 February 2002], are reproduced below as Table 2B. This table uses mRNA nomenclature, and so instead of thymine (T) which is found in DNA, the tables use uracil (U) which is found in RNA. Table 2B has been adapted so that frequencies are calculated for each amino acid, rather than for all 64 codons.

Table 2B. Codon Usage Table for *Saccharomyces cerevisiae*

Amino Acid	Codon	Number	Frequency per thousand
Phe	UUU	170666	26.1
Phe	UUC	120510	18.4
Leu	UUA	170884	26.2
Leu	UUG	177573	27.2
Leu	CUU	80076	12.3
Leu	CUC	35545	5.4
Leu	CUA	87619	13.4
Leu	CUG	68494	10.5
Ile	AUU	196893	30.1
Ile	AUC	112176	17.2
Ile	AUA	116254	17.8
Met	AUG	136805	20.9
Val	GUU	144243	22.1
Val	GUC	76947	11.8
Val	GUA	76927	11.8
Val	GUG	70337	10.8
Ser	UCU	153557	23.5
Ser	UCC	92923	14.2
Ser	UCA	122028	18.7
Ser	UCG	55951	8.6
Ser	AGU	92466	14.2
Ser	AGC	63726	9.8
Pro	CCU	88263	13.5
Pro	CCC	44309	6.8
Pro	CCA	119641	18.3
Pro	CCG	34597	5.3
Thr	ACU	132522	20.3
Thr	ACC	83207	12.7
Thr	ACA	116084	17.8
Thr	ACG	52045	8.0
Ala	GCU	138358	21.2
Ala	GCC	82357	12.6
Ala	GCA	105910	16.2
Ala	GCG	40358	6.2

Amino Acid	Codon	Number	Frequency per thousand
Tyr	UAU	122728	18.8
Tyr	UAC	96596	14.8
His	CAU	89007	13.6
His	CAC	50785	7.8
Gln	CAA	178251	27.3
Gln	CAG	79121	12.1
Asn	AAU	233124	35.7
Asn	AAC	162199	24.8
Lys	AAA	273618	41.9
Lys	AAG	201361	30.8
Asp	GAU	245641	37.6
Asp	GAC	132048	20.2
Glu	GAA	297944	45.6
Glu	GAG	125717	19.2
Cys	UGU	52903	8.1
Cys	UGC	31095	4.8
Trp	UGG	67789	10.4
Arg	CGU	41791	6.4
Arg	CGC	16993	2.6
Arg	CGA	19562	3.0
Arg	CGG	11351	1.7
Arg	AGA	139081	21.3
Arg	AGG	60289	9.2
Gly	GGU	156109	23.9
Gly	GGC	63903	9.8
Gly	GGA	71216	10.9
Gly	GGG	39359	6.0
Stop	UAA	6913	1.1
Stop	UAG	3312	0.5
Stop	UGA	4447	0.7

By utilizing this or similar tables, one of ordinary skill in the art can apply the frequencies to any given polypeptide sequence, and produce a nucleic acid fragment of a codon-optimized coding region which encodes the polypeptide, but which uses codons optimal for a given species.

Randomly assigning codons at an optimized frequency to encode a given polypeptide sequence, can be done manually by calculating codon frequencies for each amino acid, and then assigning the codons to the

2017200439 20 Jan 2017

polypeptide sequence randomly. Additionally, various algorithms and computer software programs are readily available to those of ordinary skill in the art. For example, the "EditSeq" function in the Lasergene Package, available from DNAstar, Inc., Madison, WI, the backtranslation function in the VectorNTI Suite, available from InforMax, Inc., Bethesda, MD, and the "backtranslate" function in the GCG-Wisconsin Package, available from Accelrys, Inc., San Diego, CA. In addition, various resources are publicly available to codon-optimize coding region sequences, *e.g.*, the "backtranslation" function at www.entelechon.com/bioinformatics/backtranslation.php?lang=eng (visited April 15, 2008) and the "backtranseq" function available at http://bioinfo.pbi.nrc.ca:8090/EMBOSS/index.html (visited July 9, 2002). Constructing a rudimentary algorithm to assign codons based on a given frequency can also easily be accomplished with basic mathematical functions by one of ordinary skill in the art.

Codon-optimized coding regions can be designed by various methods known to those skilled in the art including software packages such as "synthetic gene designer" (userpages.umbc.edu/~wug1/codon/sgd/, visited March 19, 2012).

A polynucleotide or nucleic acid fragment is "hybridizable" to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA molecule, when a single-stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularly Chapter 11 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. Stringency conditions can be adjusted to screen for moderately similar fragments (such as homologous sequences from distantly related organisms), to highly similar fragments (such as genes that duplicate functional enzymes from closely related

organisms). Post hybridization washes determine stringency conditions. One set of conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45 °C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50 °C for 30 min. Another set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60 °C. Another set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65 °C. An additional set of stringent conditions include hybridization at 0.1X SSC, 0.1% SDS, 65 °C and washes with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS, for example.

Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the hybridization, mismatches between bases are possible. The appropriate stringency for hybridizing nucleic acids depends on the length of the nucleic acids and the degree of complementation, variables well known in the art. The greater the degree of similarity or homology between two nucleotide sequences, the greater the value of Tm for hybrids of nucleic acids having those sequences. The relative stability (corresponding to higher Tm) of nucleic acid hybridizations decreases in the following order: RNA:RNA, DNA:RNA, DNA:DNA. For hybrids of greater than 100 nucleotides in length, equations for calculating Tm have been derived (see Sambrook et al., supra, 9.50 9.51). For hybridizations with shorter nucleic acids, i.e., oligonucleotides, the position of mismatches becomes more important, and the length of the oligonucleotide determines its specificity (see Sambrook et al., supra, 11.7 11.8). In one embodiment the length for a hybridizable nucleic acid is at least about 10 nucleotides. In one embodiment, a minimum length for a hybridizable nucleic acid is at least about 15 nucleotides; at least about 20 nucleotides; or the length is at least about 30 nucleotides. Furthermore, the skilled artisan will recognize that the temperature and wash solution salt concentration can be adjusted as necessary according to factors such as length of the probe.

As used herein, the term "polypeptide" is intended to encompass a singular "polypeptide" as well as plural "polypeptides," and refers to a molecule composed of monomers (amino acids) linearly linked by amide bonds (also known as peptide bonds). The term "polypeptide" refers to any chain or chains of two or more amino acids, and does not refer to a specific length of the product. Thus, peptides, dipeptides, tripeptides, oligopeptides, "protein, " "amino acid chain," or any other term used to refer to a chain or chains of two or more amino acids, are included within the definition of "polypeptide," and the term "polypeptide" can be used instead of, or interchangeably with any of these terms. A polypeptide can be derived from a natural biological source or produced by recombinant technology, but is not necessarily translated from a designated nucleic acid sequence. It can be generated in any manner, including by chemical synthesis.

By an "isolated" polypeptide or a fragment, variant, or derivative thereof is intended a polypeptide that is not in its natural milieu. No particular level of purification is required. For example, an isolated polypeptide can be removed from its native or natural environment. Recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for purposed of the invention, as are native or recombinant polypeptides which have been separated, fractionated, or partially or substantially purified by any suitable technique.

As used herein, the terms "variant" and "mutant" are synonymous and refer to a polypeptide differing from a specifically recited polypeptide by one or more amino acid insertions, deletions, mutations, and substitutions, created using, *e.g.*, recombinant DNA techniques, such as mutagenesis. Guidance in determining which amino acid residues can be replaced, added, or deleted without abolishing activities of interest, can be found by comparing the sequence of the particular polypeptide with that of homologous polypeptides, *e.g.*, yeast or bacterial, and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequences.

2017200439 20 Jan 2017

"Engineered polypeptide" as used herein refers to a polypeptide that is synthetic, i.e., differing in some manner from a polypeptide found in nature.

Alternatively, recombinant polynucleotide variants encoding these same or similar polypeptides can be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as silent changes which produce various restriction sites, can be introduced to optimize cloning into a plasmid or viral vector for expression. Mutations in the polynucleotide sequence can be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide. For example, mutations can be used to reduce or eliminate expression of a target protein and include, but are not limited to, deletion of the entire gene or a portion of the gene, inserting a DNA fragment into the gene (in either the promoter or coding region) so that the protein is not expressed or expressed at lower levels, introducing a mutation into the coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the coding region to alter amino acids so that a non-functional or a less enzymatically active protein is expressed.

Amino acid "substitutions" can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements, or they can be the result of replacing one amino acid with an amino acid having different structural and/or chemical properties, *i.e.*, non-conservative amino acid replacements. "Conservative" amino acid substitutions can be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged

2017200439 20 Jan 2017

(acidic) amino acids include aspartic acid and glutamic acid. Alternatively, "non-conservative" amino acid substitutions can be made by selecting the differences in polarity, charge, solubility, hydrophobicity, hydrophilicity, or the amphipathic nature of any of these amino acids. "Insertions" or "deletions" can be within the range of variation as structurally or functionally tolerated by the recombinant proteins. The variation allowed can be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Altschul, S. F., *et al.*, *J. Mol. Biol.*, 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides can be used in sequence-dependent methods of gene identification (*e.g.*, Southern hybridization) and isolation (*e.g.*, *in situ* hybridization of bacterial colonies or bacteriophage plaques). In addition, short oligonucleotides of 12-15 bases can be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence. The instant specification teaches the complete amino acid and nucleotide sequence encoding particular proteins. The skilled artisan, having the benefit of the sequences as reported herein, can now use all or a substantial portion of the disclosed sequences for purposes known to

2017200439 20 Jan 2017

those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

The term "complementary" is used to describe the relationship between nucleotide bases that are capable of hybridizing to one another. For example, with respect to DNA, adenine is complementary to thymine and cytosine is complementary to guanine, and with respect to RNA, adenine is complementary to uracil and cytosine is complementary to guanine.

The term "percent identity", as known in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods, including but not limited to those described in: 1.) Computational Molecular Biology (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) Biocomputing: Informatics and Genome Projects (Smith, D. W., Ed.) Academic: NY (1993); 3.) Computer Analysis of Sequence Data, Part I (Griffin, A. M., and Griffin, H. G., Eds.) Humania: NJ (1994); 4.) Sequence Analysis in Molecular Biology (von Heinje, G., Ed.) Academic (1987); and 5.) Sequence Analysis Primer (Gribskov, M. and Devereux, J., Eds.) Stockton: NY (1991).

Methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. Sequence alignments and percent identity calculations can be performed using the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignments of the sequences are performed using the "Clustal method of alignment" which encompasses several varieties of the algorithm including the "Clustal V method of alignment" corresponding to the alignment method labeled Clustal V

2017200439 20 Jan 2017

(described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. *et al.*, *Comput. Appl. Biosci.*, 8:189-191 (1992)) and found in the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program. Additionally the "Clustal W method of alignment" is available and corresponds to the alignment method labeled Clustal W (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. *et al.*, *Comput. Appl. Biosci.* 8:189-191(1992)) and found in the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). Default parameters for multiple alignment (GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). After alignment of the sequences using the Clustal W program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, such as from other species, wherein such polypeptides have the same or similar function or activity, or in describing the corresponding polynucleotides. Useful examples of percent identities include, but are not limited to: 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 55% to 100% can be useful in describing the present invention, such as 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%,

2017200439 20 Jan 2017

94%, 95%, 96%, 97%, 98% or 99%. Suitable polynucleotide fragments not only have the above homologies but typically comprise a polynucleotide having at least 50 nucleotides, at least 100 nucleotides, at least 150 nucleotides, at least 200 nucleotides, or at least 250 nucleotides. Further, suitable polynucleotide fragments having the above homologies encode a polypeptide having at least 50 amino acids, at least 100 amino acids, at least 150 amino acids, at least 200 amino acids, or at least 250 amino acids.

The term "sequence analysis software" refers to any computer algorithm or software program that is useful for the analysis of nucleotide or amino acid sequences. "Sequence analysis software" can be commercially available or independently developed. Typical sequence analysis software will include, but is not limited to: 1.) the GCG suite of programs (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul *et al., J. Mol. Biol.,* 215:403-410 (1990)); 3.) DNASTAR (DNASTAR, Inc. Madison, WI); 4.) Sequencher (Gene Codes Corporation, Ann Arbor, MI); and 5.) the FASTA program incorporating the Smith-Waterman algorithm (W. R. Pearson, *Comput. Methods Genome Res.,* [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20. Editor(s): Suhai, Sandor. Plenum: New York, NY). Within the context of this application it will be understood that where sequence analysis software is used for analysis, that the results of the analysis will be based on the "default values" of the program referenced, unless otherwise specified. As used herein "default values" will mean any set of values or parameters that originally load with the software when first initialized.

Standard recombinant DNA and molecular cloning techniques are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (hereinafter "Maniatis"); and by Silhavy, T. J., Bennan, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); and by Ausubel, F. M. *et al.*,

2017200439 20 Jan 2017

Current Protocols in Molecular Biology, published by Greene Publishing Assoc. and Wiley-Interscience (1987). Additional methods used here are in *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA). Other molecular tools and techniques are known in the art and include splicing by overlapping extension polymerase chain reaction (PCR) (Yu, et al. (2004) Fungal Genet. Biol. 41:973–981), positive selection for mutations at the *URA*3 locus of *Saccharomyces cerevisiae* (Boeke, J.D. et al. (1984) Mol. Gen. Genet. 197, 345-346; M A Romanos, et al. Nucleic Acids Res. 1991 January 11; 19(1): 187), the cre-lox site-specific recombination system as well as mutant lox sites and FLP substrate mutations (Sauer, B. (1987) Mol Cell Biol 7: 2087-2096; Senecoff, et al. (1988) Journal of Molecular Biology, Volume 201, Issue 2, Pages 405-421; Albert, et al. (1995) The Plant Journal. Volume 7, Issue 4, pages 649–659), "seamless" gene deletion (Akada, et al. (2006) Yeast;23(5):399-405), and gap repair methodology (Ma *et al.*, *Genetics* 58:201-216; 1981).

The genetic manipulations of a recombinant host cell disclosed herein can be performed using standard genetic techniques and screening and can be made in any host cell that is suitable to genetic manipulation (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202).

In embodiments, a recombinant host cell disclosed herein can be any yeast or fungi host useful for genetic modification and recombinant gene expression including those yeast mentioned elsewhere herein, such as in Table 7. In other embodiments, a recombinant host cell can be a member of the genera *Issatchenkia, Zygosaccharomyces, Schizosaccharomyces, Dekkera, Torulopsis, Brettanomyces, Torulaspora, Hanseniaspora, Kluveromyces, Yarrowia*, and some species of *Candida*.

Polypeptides with KARI Activity

In some embodiments, the recombinant host cells and methods provided herein address a need that arises in the microbial production of

2017200439 20 Jan 2017

isobutanol where the KARI enzyme performs a vital role. In the isobutanol biosynthetic pathway shown in Figure 1, the substrate to product conversion of acetolactate to dihydroxyisovalerate (DHIV) is catalyzed by the KARI enzyme. Disclosed in US Appl. Publication No. US2011/0244536 and incorporated by reference, are polypeptides having ketol-acid reductoisomerase activity that are members of the SLSL Clade of KARIs. Polypeptides having KARI activity disclosed therein were found to be effective for isobutanol production. The SLSL Clade of KARIs include those KARI enzymes listed in Table 3.

Table 3. Effective KARIs

Description	SEQ ID NO: Nucleic acid	SEQ ID NO: Amino acid
Staphylococcus capitis SK14	66	67
Staphylococcus epidermidis M23864-W1	68	69
Staphylococcus hominis SK119	134	135
Staphylococcus aureus subsp. *aureus* TCH130	70	71
Staphylococcus warneri L37603	72	73
Staphylococcus epidermidis W23144	74	75
Staphylococcus saprophyticus subsp. Saprophyticus ATCC15305	76	77
Staphylococcus carnosus subsp. Carnosus TM300	78	79
Listeria monocytogenes EGD-e	80	81
Listeria grayi DSM 20601	82	83
Enterococcus casseliflavus EC30	84	85
Enterococcus gallinarum EG2	86	87
Macrococcus caseolyticus JCSC5402	88	89
Streptococcus vestibularis	90	91
Streptococcus mutans UA159	92	93
Streptococcus gordonii str, cgakkus sybstr. CH1	94	95
Streptococcus suis 89/1591	96	97
Streptococcus infantarius subsp. infantarius ATCC BAA-102	98	99
Lactococcus lactis subsp cremoris MG1363	100	101

2017200439 20 Jan 2017

Lactococcus lactis	102	103
Leuconostoc mesenteroides subsp mesenteroides ATCC8293	104	105
Lactobacillus buchneri ATCC 11577	106	107
Staphylococcus haemolyticus JCSC1435	108	109
Staphylococcus epidermidis ATCC12228	110	111
Streptococcus pneumoniae CGSP14	112	113
Streptococcus pneumoniae TIGR4	114	115
Streptococcus sanguinis SK36	116	117
Streptococcus salivarius SK126	118	119
Streptococcus thermophilus LMD-9	120	121
Streptococcus pneumoniae CCRI 1974M2	122	123
Lactococcus lactis subsp. lactis II1403	124	125
Leuconostoc mesenteroides subsp cremoris ATCC19254	126	127
Leuconostoc mesenteroides subsp cremoris	128	129
Lactobacillus brevis subsp. gravesensis ATCC27305	130	131
Lactococcus lactis subsp lactis NCDO2118	132	133

As described and demonstrated herein, Applicants have discovered additional KARI enzymes and variants of the additional KARIs that result in isobutanol production comparable to and/or exceeding that observed with the KARI from *Lactococcus lactis*. Such KARI enzymes and variants result in comparable or higher isobutanol titer and/or higher effective isobutanol productivity when compared to that observed with *Lactococcus lactis* KARI in the same conditions. Accordingly, in embodiments, polypeptides having KARI activity that function in an isobutanol production pathway have isobutanol titer and/or effective isobutanol productivity comparable to or better than that with the *Lactococcus lactis* KARI (SEQ ID NO: 380).

Such polypeptides having KARI activity may thus be suitable for isobutanol production. It will be appreciated that using a combination of structural and sequence information available in the art, polypeptides

2017200439 20 Jan 2017

comprising KARI activity and less than 100% identity to the exemplified sequences can be constructed for use in isobutanol biosynthetic pathways. For example, crystal structures of the *E. coli* KARI enzyme at 2.6 Å resolution have been solved (Tyagi, et al., Protein Sci., 14: 3089-3100, 2005) as has the structure of the *P. aeruginosa* KARI (Ahn, et al., J. Mol. Biol., 328: 505-515, 2003) and the KARI enzyme from spinach (Biou V., et al. The EMBO Journal, 16: 3405-3415, 1997). Furthermore, described herein is a Profile HMM (provided herein; Table Z) prepared using amino acid sequences of 25 KARI proteins with experimentally verified function as outlined in Table 1. The KARIs were from *Pseudomonas fluorescens* Pf-5, *Sulfolobus solfataricus P2*, *Pyrobaculum aerophilum str.* IM2, *Natronomonas pharaonis* DSM 2160, *Bacillus subtilis subsp. subtilis* str. 168, *Corynebacterium glutamicum* ATCC 13032, *Phaeospririlum molischianum*, *Ralstonia solanacearum* GMI1000, *Zymomonas mobilis subsp. mobilis* ZM4, *Alkalilimnicola ehrlichei* MLHE-, *Campylobacter lari* RM2100, *Marinobacter aquaeolei* VT8, *Psychrobacter arcticus* 273-4, *Hahella chejuensis* KCTC 2396, *Thiobacillus denitrificans* ATCC 25259, *Azotobacter vinelandii* AvOP, *Pseudomonas syringae pv. syringae* B728a, *Pseudomonas syringae pv. tomato str. DC3000*, *Pseudomonas putida* KT2440, *Pseudomonas entomophila* L48, *Pseudomonas mendocina ymp, Pseudomonas aeruginosa* PAO1, *Bacillus cereus ATCC 10987*, *Bacillus cereus* ATCC 10987, and *Spinacia oleracea*. Any protein that matches the Profile HMM with an E value of $<10^{-3}$ using hmmsearch program in the HMMER package is expected to be a functional KARI.

Production of isobutanol is believed to utilize the glycolysis pathway present in the host microorganism. During the production of two molecules of pyruvate from glucose during glycolysis, there is net production of two molecules of NADH from NAD+ by the glyceraldehyde-3-phosphate dehydrogenase reaction. During the further production of one molecule of isobutanol from two molecules of pyruvate, there is net consumption of one molecule of NAD(P)H, by the KARI reaction, and one molecule of NAD(P)H by the isobutanol dehydrogenase reaction. The

interconversion of NADH with NADPH is generally slow and inefficient in yeast; thus, NADPH to be consumed is generated by metabolism (for example, by the pentose phosphate pathway) consuming substrate in the process. Meanwhile, the cell strives to maintain homeostasis in the NAD+/NADH ratio, leading to the excess NADH produced in isobutanol production being consumed in wasteful reduction of other metabolic intermediates; e.g., by the production of glycerol (Bakker, *et al.*, 2001. Stoichiometry and compartmentation of NADH metabolism in Saccharomyces cerevisiae. FEMS Microbiol. Rev. 25:15-37.). Thus, an imbalance between NADH produced and NADPH consumed by the isobutanol pathway can lead to a reduction in the molar yield of isobutanol produced from glucose in two ways: 1) unnecessary operation of metabolism to produce NADPH, and 2) wasteful reaction of metabolic intermediates to maintain NAD+/NADH homeostasis. Polypeptides having KARI activity that function well in an isobutanol pathway and have a low K_M for NADH can be used to improve the production of isobutanol.

Also disclosed herein are substitutions to the KARI enzyme sequences provided in Table 3 and in Table 10 to produce variants with varying ability to utilize NADH as a cofactor. Such variants provide alternatives that may be employed to optimize the efficiency of a biosynthetic pathway utilizing KARI, such as an isobutanol biosynthetic pathway, for particular production conditions. Demonstrated in the Examples is isobutanol production under conditions switched from aerobic to anaerobic for variants of the K9 KARI enzyme derived from *Anaerostipes caccae* with differing abilities to utilize NADH. Thus, equipped with this disclosure, one of skill in the art will be able to produce recombinant host cells comprising a SLSL Clade KARI enzyme, or a an *Enterococcus gallinarum, Streptococcus thermophilus Lactococcus lactis subsp. cremoris MG1363*, *Bifidobacterium angulatum, Bifidobacterium dentium,* or *Anaerostipes caccae, Lactococcus lactis* KARI enzyme or a variant or active fragment thereof suited for a range of production conditions.

2017200439 20 Jan 2017

In some embodiments, provided herein is a polypeptide having KARI activity and having at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, or at least about 99% identity to a KARI enzyme of Table 3 or Table 10, or Examples 16, 17, 21 and having a K_M for NADH less than about 300 μM, 100 μM, 50 μM, 20 μM, 10 μM, or 5 μM. In some embodiments, provided herein is an engineered polypeptide having KARI activity and having at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, or at least about 99% identity to a KARI enzyme of Table 3, Table 10 or Examples 16, 17, 21. In some embodiments, such polypeptides have a K_M for NADH less than that of the corresponding native enzyme. In some embodiments, the ratio of K_M for NADH to K_M for NAPDH is less than 0.1, in some embodiments less than 1, in some embodiments less than 2, in some embodiments, less than 4.

KARI enzymes and variants thereof that are particularly suitable for isobutanol production include, but are not limited to, variants of a ketol-acid reductoisomerase from *Anaerostipes caccae* DSM 14662 (SEQ ID NO: 643): "K9G9" (SEQ ID NO: 644) and "K9D3" (SEQ ID NO: 645) which have K_M for NADH lower than that of the native enzyme (SEQ ID NO: 643).

Host cells provided herein may comprise a polypeptide having ketol-acid reductoisomerase activity. In embodiments, such polypeptides have at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, or at least about 99% identity to SEQ ID NO: 643, an active variant thereof; or a KARI derived from *Anaerostipes caccae* DSM 14662, or an active variant thereof. In embodiments, the polypeptides have at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, or at least about 99% identity to SEQ ID NO: 645 or 644. In embodiments, the polypeptides comprise SEQ ID NO: 645or 644.

In some embodiments, polypeptides having KARI activity comprise at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 98%, or at least about 99% identity to the amino acid

20 Jan 2017 2017200439

sequence of SEQ ID NO: 419 [JB4P], 427 [SB2], and all those variants listed in Tables 25 and 26. Such variants provide alternatives for optimizing the efficiency of the isobutanol biosynthetic pathway for particular production conditions. Demonstrated in the Examples is isobutanol production under conditions.

Identification of additional polypeptides having KARI activity

Described in Example 1 is a biodiversity screen of KARI-encoding genes from various bacterial and fungal species which revealed suitable KARIs for isobutanol production. Equipped with this disclosure, one of skill in the art will be readily able to identify additional suitable polypeptides having KARI activity.

The sequences of other polynucleotides, genes and/or polypeptides can be identified in the literature and in bioinformatics databases well known to the skilled person using sequences disclosed herein and available in the art. For example, such sequences can be identified through BLAST searching of publicly available databases with polynucleotide or polypeptide sequences provided herein. In such a method, identities can be based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

Additionally, polynucleotide or polypeptide sequences disclosed herein can be used to identify other KARI homologs in nature. For example, each of the KARI encoding nucleic acid fragments disclosed herein can be used to isolate genes encoding homologous proteins. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent protocols include, but are not limited to (1) methods of nucleic acid hybridization; (2) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies [e.g., polymerase chain reaction (PCR), Mullis et al., U.S. Patent No. 4,683,202; ligase chain reaction (LCR), Tabor et al., Proc. Acad. Sci. USA 82:1074 (1985); or strand displacement amplification (SDA), Walker et al., Proc. Natl. Acad.

2017200439 20 Jan 2017

Sci. U.S.A., 89:392 (1992)]; and (3) methods of library construction and screening by complementation.

It will be appreciated that one of ordinary skill in the art, equipped with this disclosure, can generate active fragments of polypeptides provided herein, for example, by truncating polypeptides provided herein based on sequence alignments at the N-terminus and confirming KARI activity. In embodiments, *Anaerostipes caccae* KARIs and variants thereof provided herein are truncated at the N-terminus. In one embodiment, up to and including the first five amino acids are truncated from a polypeptide provided herein. In embodiments, the polypeptide is SEQ ID NO: 27 or a variant thereof. In one embodiment, a polypeptide having KARI activity comprises SEQ ID NO: 635, 637 (encoded by polynucleotide sequences SEQ ID NO: 636 and 638, respectively), K9_Annabel_SH (SEQ ID NO:862, protein SEQ ID NO:863) and K9_Zeke_SH (SEQ ID NO: 860, protein SEQ ID NO: 861), or any variant listed in Table 40.

Lowering K_M for NADH

As shown in Figure 2 and Examples, mutations in the positions corresponding to 50, 52 and 53, and optionally 47, of the *Pseudomonas fluorescens* KARI in the KARI enzyme from *Anaerostipes caccae* result in KARIs with lowered K_M for NADH as compared to wild-type, verifying that mutations in these positions produce NADH accepting variants of highly effective KARIs. Further mutations of *Anaerostipes caccae* KARI, revealed positions which further lower the K_M for NADH.

As demonstrated herein (see Examples), substitution of amino acids in the phosphate binding region, particularly in two or more positions corresponding to positions 47, 50, 52, and 53 of PF5 KARI (SEQ ID NO: 5) results in lowered K_M for NADH. Therefore, provided herein are polypeptides derived from an organism listed herein, for example, in Tables 3 and 10 having KARI activity and comprising substitutions at at least two of the four positions corresponding to positions 47, 50, 52, and 53 of PF5 KARI as compared to the native amino acid sequence.

2017200439 20 Jan 2017

Provided herein are polypeptides having KARI activity and comprising substitutions in the phosphate binding region. Provided herein are polypeptides having KARI activity and comprising substitutions at positions corresponding to S56 and S58 of K9 KARI (SEQ ID NO: 27). In some embodiments the substitution at the position corresponding to S56 is A. In some embodiments, the substitution at the position corresponding to S58 is D or E. In some embodiments, the substitution at the position corresponding to S53 is Q, E, P, or A. In some embodiments, the substitution at the position corresponding to S56 is V or D. In some embodiments, the substitution at the position corresponding to S58 is D or Q. In embodiments, the polypeptides further comprise a substitution at one or more positions corresponding to I86, N87, N107, T131, or T191 of K9 KARI (SEQ ID NO: 27). In some embodiments, the polypeptides comprise a substitution at at least 2, at least 3, at least 4, or all of the indicated positions. In some embodiments, the substitution at the position corresponding to I86 is T or V. In some embodiments, the substitution at the position corresponding to N87 is P. In some embodiments, the substitution at the position corresponding to N107 is S. In some embodiments, the substitution at the position corresponding to T131 is C, L, A, M or V. In some embodiments, the substitution at the position corresponding to T191 is A, S, D, C, or G.

In embodiments, the polypeptides comprise fewer than 10, 15, or 20 substitutions with respect to the wild-type sequence. In embodiments, the polypeptides match the Profile HMM based on experimentally verified KARIs and given in Table Z with an E value less than $< 10^{-3}$. Sequences can be compared to the profile HMM given in Table Z using hmmsearch (HMMER software package available from Janelia Farm Research Campus, Ashburn, VA).

Additional polypeptides having KARI activity and lowered K_M for NADH can be obtained using methods described and demonstrated herein. For example, a polypeptide having KARI activity can be employed in the construction of a site-saturation gene library as described herein. Kits for construction of such gene libraries are commercially available (for

2017200439 20 Jan 2017

example, from USB Corporation, Cleveland, OH, #78480.) Site-directed mutagenesis can also be carried out using commercially available kits (for example, the QuickChange II XL site directed mutagenesis kit, Catalog # 200524, Stratagene, La Jolla, CA). Primer design for target sites for mutagenesis is well-known in the art, and multiple sequence alignment to identify the target sites is likewise well-known.

Once variants have been generated, KARI activity with NADH or NADPH can be readily assessed using methods known in the art and/or disclosed herein. For example, KARI activity can be determined by measuring the disappearance of the NADPH or NADH from the reaction at 340 nm or by determination of the Michaelis constant via measurement of formation of 2,3-dihydroxyisovalerate using HPLC/MS. Likewise, isobutanol production from a strain comprising variants can be confirmed.

Cofactor specificity

To determine cofactor specificity, V_{max}/K_M ratios can be calculated for each cofactor at saturating acetolactate; those variants with a higher ratio for NADH will react at a higher rate with NADH than NADPH under conditions of equal-molar concentrations of the two cofactors and saturating acetolactate. V_{max} and K_M values for NADH and NADPH can be determined using methods known in the art and/or provided herein (see Example 16). For example, to determine V_{max} and K_M values for NADH and NADPH, the partially purified proteins can be assayed at various concentrations of NADH and NADPH.

As demonstrated herein (see Examples 16 and 18 and Figure 8), substitution of additional amino acids in K9G9 results in variants having increased specificity for NADH. Thus, provided herein are polypeptides comprising substitution at one or more or all of the positions corresponding to K57, Y53, and E74 of K9 KARI (SEQ ID NO: 27). Also provided herein are polypeptides comprising substitutions at one or more or all of the positions corresponding to Y53, K57, E74, N87 and K90, In embodiments, the substitution at the position corresponding to Y53 is F. In embodiments, the substitution at the position corresponding to K57 is E.

In embodiments, the substitution at the position corresponding to E74 is G. In embodiments, the substitution at the position corresponding to N87 is P. In embodiments, the substitution at the position corresponding to K90 is M or L. In embodiments, the variants comprise substitutions of at least one position corresponding to S56 or S58 of SEQ ID NO: 27 and further comprise at least one, at least two, at least three, or more than three further substitution(s) corresponding to positions of SEQ ID NO: 27 identified herein.

In embodiments, the polypeptides comprise fewer than 2, 3, 4, 5, 10, 15, or 20 substitutions with respect to the wild-type sequence. In embodiments, the polypeptides match the Profile HMM based on experimentally verified KARIs and given in Table Z with an E value less than $< 10^{-3}$.

As demonstrated in the Examples, variants of K9SB2 (SEQ ID NO: 427) were generated and screened for variants with reduced NADPH affinity, revealing additional positions for substitution. Thus, in embodiments, polypeptides further comprise substitutions at one or more positions corresponding to F53, G55, A56, W59, F67, I84, L85, Q91, M94, and P135 of SEQ ID NO: 427. In embodiments, the substitution at position G55 is D or C, the substitution at position Q91 is L, the substitution at position A56 is T or V, the substitution at P135 is S, the substitution at position F53 is L, the substitution at position M94 is I, the substitution at position F67 is L or I, the substitution at position W59 is C, the substitution at position I84 is L, and the substitution at position L85 is M.

KARI Structure

Structural information useful in the identification and modification of polypeptides having KARI activity is provided in art, such as in the references described here as well as in the Profile HMM provided herewith in Table Z and in US App. Pub. Nos. 20100197519 and 20090163376, incorporated herein by reference

2017200439 20 Jan 2017

It was reported that phosphate p2' oxygen atoms of NADPH form hydrogen bonds with side chains of Arg162, Ser165 and Ser167 of spinach KARI (Biou V., et al. The EMBO Journal, 16: 3405-3415, 1997). Studies by Ahn et al., (J. Mol. Biol., 328: 505-515, 2003) had identified three NADPH phosphate binding sites (Arg47, Ser50 and Thr52) for *Pseudomonas aeruginosa* (PAO-KARI) following comparing its structure with that of the spinach KARI. The structure of PF5-KARI with bound NADPH, acetolactate and magnesium ions was built based on the crystal structure of *P. aeruginosa* PAO1-KARI (PDB ID 1NP3, Ahn H. J. et al., J. Mol. Biol., 328: 505-515, 2003) which has 92% amino acid sequence homology to PF5 KARI. PAO1-KARI structure is a homo-dodecamer and each dodecamer consists of six homo-dimers with extensive dimer interface. The active site of KARI is located in this dimer interface. The biological assembly is formed by six homo-dimers positioned on the edges of a tetrahedron resulting in a highly symmetrical dodecamer of 23 point group symmetry.

The model of PF5-KARI dimer was built based on the coordinates of monomer A and monomer B of PAO1-KARI and sequence of PF5-KARI using DeepView/Swiss PDB viewer (Guex, N. and Peitsch, M.C., Electrophoresis, 18: 2714-2723, 1997). This model was then imported to program O (Jones, T.A. *et al*, Acta Crystallogr. A 47: 110-119, 1991) on a Silicon Graphics system for further modification.

The structure of PAO1-KARI has no NADPH, substrate or inhibitor or magnesium in the active site. Therefore, the spinach KARI structure (PDB ID 1yve, Biou V. et al., The EMBO Journal, 16: 3405-3415, 1997.), which has magnesium ions, NADPH and inhibitor (N-Hydroxy-N-isopropyloxamate) in the acetolacate binding site, was used to model these molecules in the active site. The plant KARI has very little sequence homology to either PF5- or PAO1 KARI (<20% amino acid identity), however the structures in the active site region of these two KARI enzymes are very similar. To overlay the active site of these two KARI structures, commands LSQ_ext, LSQ_improve, LSQ_mol in the program O were used to line up the active site of monomer A of spinach KARI to

2017200439 20 Jan 2017

the monomer A of PF5 KARI model. The coordinates of NADPH, two magnesium ions and the inhibitor bound in the active site of spinach KARI were extracted and incorporated to molecule A of PF5 KARI. A set of the coordinates of these molecules were generated for monomer B of PF5 KARI by applying the transformation operator from monomer A to monomer B calculated by the program.

Because there is no NADPH in the active site of PAO1 KARI crystal structure, the structures of the phosphate binding loop region in the NADPH binding site (residues 44-45 in PAO1 KARI, 157-170 in spinach KARI) are very different between the two. To model the NADPH bound form, the model of the PF5-KARI phosphate binding loop (44-55) was replaced by that of 1yve (157-170). Any discrepancy of side chains between these two was converted to those in the PF5-KARI sequence using the mutate_replace command in program O, and the conformations of the replaced side-chains were manually adjusted. The entire NADPH/Mg/ inhibitor bound dimeric PF5-KARI model went through one round of energy minimization using program CNX (ACCELRYS San Diego CA , Burnger, A.T. and Warren, G.L., Acta Crystallogr., D 54: 905-921, 1998) after which the inhibitor was replaced by the substrate, acetolactate (AL), in the model.

Isobutanol Production

Host cells provided herein can comprise a polypeptide having ketol-acid reductoisomerase activity. As described and demonstrated herein, Applicants have discovered additional KARI enzymes and variants of the additional KARIs that result in isobutanol production comparable to and/or exceeding that observed with the KARI from *Lactococcus lactis* (see Examples). Accordingly, in embodiments, polypeptides having KARI activity that function in an isobutanol production pathway have effective isobutanol productivity and/or produce isobutanol at a titer comparable to or better than that with the *Lactococcus lactis* KARI (SEQ ID NO: 380). Such polypeptides are thus considered to be useful for isobutanol production, particularly in cells comprising isobutanol production pathways

described herein. In embodiments, polypeptides provided herein have effective isobutanol productivity and/or produce isobutanol at a titer greater than or about equal to that observed with the the *Lactococcus lactis* KARI (SEQ ID NO: 380) under the same conditions. In embodiments, polypeptides provided herein have effective isobutanol productivity greater than about 3 grams per gram of cells, greater than about 4, greater than about 5, or greater than about 6 grams per gram of cells after about 48 hours wherein at least the last about 24 hours of the 48 hours are under anaerobic conditions.

Furthermore, Applicants have discovered that variants of the polypeptides having KARI activity described above, including those with K_M for NADH lower than that of the unsubstituted polypeptide, provide advantages for isobutanol production under anaerobic conditions. While not wishing to be bound by theory, it is believed that such variants provide improved isobutanol production due to more effective use of NADH as reducing equivalents. In embodiments, isobutanol production employing such a variant provides reduced glycerol accumulation. In embodiments, the molar ratio of isobutanol to glycerol is increased for a variant of a polypeptide having KARI activity described above with K_M for NADH lower than that of the unsubstituted polypeptide. In embodiments, the molar ratio of isobutanol to glycerol is greater than 1. In embodiments, the molar ratio of isobutanol to glycerol is greater than 2. In embodiments, the molar ratio is greater than 3. In embodiments, the molar ratio is greater than 4, greater than 5, greater than 6, greater than 7, greater than 8, greater than 9, greater than 10, greater than 12, or greater than 14. In embodiments, the molar ratio is in the range of about 1 to 5, about 1 to 10, about 2 to 8, about 5 to 10, about 5 to 15 about 10 to 15 or about 12 to 15.

As demonstrated in the Examples herein, as the biochemical specificity for the NADH cofactor, as defined by (NADH V_{max}/K_M)/(NADPH V_{max}/K_M) increases, there is an observed increase in the isobutanol/glycerol ratio, suggesting that the altered cofactor specificity led to diminished NADPH utilization and by-product formation.

Modification of Aldehyde Dehydrogenase

In embodiments of the invention, a recombinant host cell can comprise reduced or eliminated aldehyde dehydrogenase activity and an isobutanol biosynthetic pathway wherein the host cell produces butanol. In other embodiments, the recombinant host cell can comprise an isobutanol or a 1-butanol biosynthetic pathway as described further herein. In other embodiments, the isobutanol biosynthetic pathway can comprise a polynucleotide encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of: (a) pyruvate to acetolactate; (b) acetolactate to 2,3-dihydroxyisovalerate; (c) 2,3-dihydroxyisovalerate to 2-ketoisovalerate; (d) 2-ketoisovalerate to isobutyraldehyde; and (e) isobutyraldehyde to isobutanol. In other embodiments, the isobutanol biosynthetic pathway can comprise polynucleotides encoding polypeptides having acetolactate synthase, keto acid reductoisomerase, dihydroxy acid dehydratase, ketoisovalerate decarboxylase, and alcohol dehydrogenase activity. In other embodiments, the recombinant cell comprises a 1-butanol biosynthetic pathway. In other embodiments, the 1-butanol biosynthetic pathway comprises a polynucleotide encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of: (a) acetyl-CoA to acetoacetyl-CoA; (b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA; (c) 3-hydroxybutyryl-CoA to crotonyl-CoA; (d) crotonyl-CoA to butyryl-CoA; (e) butyryl-CoA to butyraldehyde; (f) butyraldehyde to 1-butanol. In other embodiments, the 1-butanol biosynthetic pathway can comprise polynucleotides encoding polypeptides having activity.

In embodiments of the invention, a recombinant host cell can comprise a modification or disruption of a polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity or a modification or disruption of a polypeptide having aldehyde dehydrogenase activity. Many methods for genetic modification and disruption of target genes to reduce or eliminate expression are known to one of ordinary skill in the art and can be used to create a recombinant host cell disclosed herein. In other embodiments, the recombinant host cell can comprise a deletion,

2017200439 20 Jan 2017

mutation, and/or substitution in an endogenous polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity or in an endogenous polypeptide having aldehyde dehydrogenase activity. Such modifications, disruptions, deletions, mutations, and/or substitutions can result in aldehyde dehydrogenase activity that is reduced or eliminated. Modifications that can be used include, but are not limited to, deletion of the entire gene or a portion of the gene encoding an aldehyde dehydrogenase protein, inserting a DNA fragment into the encoding gene (in either the promoter or coding region) so that the protein is not expressed or expressed at lower levels, introducing a mutation into the coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the coding region to alter amino acids so that a non-functional or a less active protein is expressed. In other embodiments, expression of a target gene can be blocked by expression of an antisense RNA or an interfering RNA, and constructs can be introduced that result in cosuppression. In other embodiments, the synthesis or stability of the transcript can be lessened by mutation. In embodiments, the efficiency by which a protein is translated from mRNA can be modulated by mutation. All of these methods can be readily practiced by one skilled in the art making use of the known or identified sequences encoding target proteins.

In other embodiments, DNA sequences surrounding a target aldehyde dehydrogenase coding sequence are also useful in some modification procedures and are available, for example, for yeasts such as *Saccharomyces cerevisiae* in the complete genome sequence coordinated by Genome Project ID9518 of Genome Projects coordinated by NCBI (National Center for Biotechnology Information) with identifying GOPID #13838. An additional non-limiting example of yeast genomic sequences is that of *Candida albicans*, which is included in GPID #10771, #10701 and #16373. Other yeast genomic sequences can be readily found by one of skill in the art in publicly available databases.

In other embodiments, DNA sequences surrounding a target aldehyde dehydrogenase coding sequence can be useful for modification

methods using homologous recombination. In a non-limiting example of this method, aldehyde dehydrogenase gene flanking sequences can be placed bounding a selectable marker gene to mediate homologous recombination whereby the marker gene replaces the aldehyde dehydrogenase gene. In another non-limiting example, partial aldehyde dehydrogenase gene sequences and aldehyde dehydrogenase gene flanking sequences bounding a selectable marker gene can be used to mediate homologous recombination whereby the marker gene replaces a portion of the target aldehyde dehydrogenase gene. In embodiments, the selectable marker can be bounded by site-specific recombination sites, so that following expression of the corresponding site-specific recombinase, the resistance gene is excised from the aldehyde dehydrogenase gene without reactivating the latter. In embodiments, the site-specific recombination leaves behind a recombination site which disrupts expression of the aldehyde dehydrogenase protein. In other embodiments, the homologous recombination vector can be constructed to also leave a deletion in the aldehyde dehydrogenase gene following excision of the selectable marker, as is well known to one skilled in the art.

In other embodiments, deletions can be made to an aldehyde dehydrogenase target gene using mitotic recombination as described by Wach et al. (Yeast, 10:1793-1808; 1994). Such a method can involve preparing a DNA fragment that contains a selectable marker between genomic regions that can be as short as 20 bp, and which bound a target DNA sequence. In other embodiments, this DNA fragment can be prepared by PCR amplification of the selectable marker gene using as primers oligonucleotides that hybridize to the ends of the marker gene and that include the genomic regions that can recombine with the yeast genome. In embodiments, the linear DNA fragment can be efficiently transformed into yeast and recombined into the genome resulting in gene replacement including with deletion of the target DNA sequence (((as disclosed, for example, in Methods in Enzymology, Volume 194, Guide to Yeast Genetics and Molecular and Cell Biology (Part A, 2004, Christine

2017200439 20 Jan 2017

Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, CA)).

Moreover, promoter replacement methods can be used to exchange the endogenous transcriptional control elements allowing another means to modulate expression such as described by Mnaimneh et al., ((2004) Cell 118(1):31-44).

In other embodiments, the aldehyde dehydrogenase target gene encoded activity can be disrupted using random mutagenesis, which can then be followed by screening to identify strains with reduced or substantially eliminated activity. In this type of method, the DNA sequence of the target gene encoding region, or any other region of the genome affecting carbon substrate dependency for growth, need not be known. In embodiments, a screen for cells with reduced aldehyde dehydrogenase activity, or other mutants having reduced aldehyde dehydrogenase activity, can be useful for recombinant host cells of the invention.

Methods for creating genetic mutations are common and well known in the art and can be applied to the exercise of creating mutants. Commonly used random genetic modification methods (reviewed in Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) include spontaneous mutagenesis, mutagenesis caused by mutator genes, chemical mutagenesis, irradiation with UV or X-rays, or transposon mutagenesis.

Chemical mutagenesis of host cells can involve, but is not limited to, treatment with one of the following DNA mutagens: ethyl methanesulfonate (EMS), nitrous acid, diethyl sulfate, or N-methyl-N'-nitro-N-nitroso-guanidine (MNNG). Such methods of mutagenesis have been reviewed in Spencer et al. (Mutagenesis in Yeast, 1996, Yeast Protocols: Methods in Cell and Molecular Biology. Humana Press, Totowa, NJ). In embodiments, chemical mutagenesis with EMS can be performed as disclosed in Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Irradiation with ultraviolet (UV) light or X-rays can also be used to produce random mutagenesis in yeast cells. The primary effect of mutagenesis by UV irradiation is the formation

2017200439 20 Jan 2017

of pyrimidine dimers which disrupt the fidelity of DNA replication. Protocols for UV-mutagenesis of yeast can be found in Spencer et al. (Mutagenesis in Yeast, 1996, Yeast Protocols: Methods in Cell and Molecular Biology. Humana Press, Totowa, NJ). In embodiments, the introduction of a mutator phenotype can also be used to generate random chromosomal mutations in host cells. In embodiments, common mutator phenotypes can be obtained through disruption of one or more of the following genes: PMS1, MAG1, RAD18 or RAD51. In other embodiments, restoration of the non-mutator phenotype can be obtained by insertion of the wildtype allele. In other embodiments, collections of modified cells produced from any of these or other known random mutagenesis processes can be screened for reduced or eliminated aldehyde dehydrogenase activity.

Genomes have been completely sequenced and annotated and are publicly available for the following yeast strains: *Ashbya gossypii* ATCC 10895, *Candida glabrata* CBS 138, *Kluyveromyces lactis* NRRL Y-1140, *Pichia stipitis* CBS 6054, *Saccharomyces cerevisiae* S288c, *Schizosaccharomyces pombe* 972h-, and *Yarrowia lipolytica* CLIB122. Typically BLAST (described above) searching of publicly available databases with known aldehyde dehydrogenase polynucleotide or polypeptide sequences, such as those provided herein, is used to identify aldehyde dehydrogenase-encoding sequences of other host cells, such as yeast cells.

In other embodiments, a polypeptide having aldehyde dehydrogenase activity can catalyze the conversion of isobutyraldehyde to isobutyric acid. In other embodiments, the conversion of isobutyraldehyde to isobutyric acid in a recombinant host cell is reduced or eliminated. In still other embodiments, a polynucleotide, gene or polypeptide having aldehyde dehydrogenase activity can correspond to Enzyme Commission Number EC 1.2.1.3, EC 1.2.1.4, and/or EC 1.2.1.5.

In embodiments, a recombinant host cell of the invention can be *S. cerevisiae,* and a polypeptide having aldehyde dehydrogenase activity can be ALD2, ALD3, ALD4, ALD5, ALD6, or combinations thereof. In other

embodiments, a recombinant host cell can be *Kluyveromyces lactis,* and a polypeptide having aldehyde dehydrogenase activity can be KLLA0F00440, KLLA0E23057, KLLA0D10021, KLLA0D09999G, or combinations thereof. In other embodiments, a recombinant host cell can be *Pichia stipitis*, and a polypeptide having aldehyde dehydrogenase activity can ALD2, ALD3, ALD4, ALD5, ALD7, or combinations thereof. In other embodiments, a recombinant host cell can be *Lactobacillus plantarum,* and a polypeptide having aldehyde dehydrogenase activity can be AldH. In other embodiments, a recombinant host cell can be *E. coli*, and a polypeptide having aldehyde dehydrogenase activity can be aldA, aldB, aldH, or combinations thereof.

In embodiments of the invention, a recombinant host cell can be *S. cerevisiae,* and an endogenous polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity can be ALD2, ALD3, ALD4, ALD5, ALD6, or combinations thereof. In embodiments of the invention, a recombinant host cell can be *S. cerevisiae,* and an endogenous polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity can be ALD6. In other embodiments, a recombinant host cell can be *Kluyveromyces lactis,* and an endogenous polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity can be KLLA0F00440, KLLA0E23057, KLLA0D10021, KLLA0D09999G, or combinations thereof. In other embodiments, a recombinant host cell can be *Pichia stipitis*, and an endogenous polynucleotide or gene encoding a polypeptide having aldehyde dehydrogenase activity can be ALD2, ALD3, ALD4, ALD5, ALD7, or combinations thereof. In embodiments, the polypeptide having aldehyde dehydrogenase activity is a homolog of ALD6 from *Saccharomyces cerevisiae*. *S. cerevisiae* deletion strains containing aldehyde dehydrogenase gene deletions with a kanMX cassette are commercially available from American Type Culture Collection [catalog #4000753].

In other embodiments, a recombinant host cell can be *Lactobacillus plantarum,* and an endogenous polynucleotide encoding a polypeptide

2017200439 20 Jan 2017

having aldehyde dehydrogenase activity can be AldH. In other embodiments, a recombinant host cell can be *E. coli*, and an endogenous polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity can be aldA, aldB, aldH, or combinations thereof.

Examples of aldehyde dehydrogenase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, those of the following Table 4.

Table 4. Aldehyde dehydrogenase target gene coding regions and proteins.

	Nucleic acid SEQ ID NO:	Amino acid SEQ ID NO:
ALD2 from *S. cerevisiae*	732	733
ALD3 from *S. cerevisiae*	734	735
ALD4 from *S. cerevisiae*	736	737
ALD5 from *S. cerevisiae*	738	739
ALD6 from *S. cerevisiae*	740	741
KLLA0F00440 from *Kluyveromyces lactis*	742	743
KLLA0E23057 from *Kluyveromyces lactis*	744	745
KLLA0D10021 from *Kluyveromyces lactis*	746	747
KLLA0D09999 from *Kluyveromyces lactis*	748	749
ALD2 from *Pichia stipits*	750	751
ALD3 from *Pichia stipitis*	752	753
ALD4 from *Pichia stipitis*	754	755
ALD5 from *Pichia stipitis*	756	757
ALD7 from *Pichia stipitis*	758	759
aldA from *E. coli*	760	761
aldB from *E. coli*	762	763
aldH (puuC) from *E. coli*	764	765

Other examples of aldehyde dehydrogenase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, aldehyde dehydrogenase polynucleotides, genes and/or polypeptides having at least about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to any one of the sequences of Table 4, wherein such a polynucleotide or gene encodes, or such a polypeptide has, aldehyde dehydrogenase activity. Still other examples of aldehyde dehydrogenase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to an active variant, fragment or derivative of any one of the sequences of Table 4, wherein such a polynucleotide or gene encodes, or such a polypeptide has, aldehyde dehydrogenase activity.

In embodiments, the sequences of other aldehyde dehydrogenase polynucleotides, genes and/or polypeptides can be identified in the literature and in bioinformatics databases well known to the skilled person using sequences disclosed herein and available in the art. For example, such sequences can be identified through BLAST searching of publicly available databases with known aldehyde dehydrogenase-encoding polynucleotide or polypeptide sequences. In such a method, identities can be based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

Additionally, the aldehyde dehydrogenase polynucleotide or polypeptide sequences disclosed herein or known the art can be used to identify other aldehyde dehydrogenase homologs in nature. For example, each of the aldehyde dehydrogenase encoding nucleic acid fragments disclosed herein can be used to isolate genes encoding homologous proteins. Isolation of homologous genes using sequence-dependent protocols is well known in the art. Examples of sequence-dependent

2017200439 20 Jan 2017

protocols include, but are not limited to (1) methods of nucleic acid hybridization; (2) methods of DNA and RNA amplification, as exemplified by various uses of nucleic acid amplification technologies [*e.g.*, polymerase chain reaction (PCR), Mullis *et al.*, U.S. Patent No. 4,683,202; ligase chain reaction (LCR), Tabor *et al.*, *Proc. Acad. Sci.* USA 82:1074 (1985); or strand displacement amplification (SDA), Walker *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*, 89:392 (1992)]; and (3) methods of library construction and screening by complementation.

Accordingly, it is within the scope of the invention to provide aldehyde dehydrogenase polynucleotides, genes and polypeptides having at least about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to any of the aldehyde dehydrogenase polynucleotides or polypeptides disclosed herein (e.g., SEQ ID NOs: 732-765 of Table 4) Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

The modification of aldehyde dehydrogenase in a recombinant host cell disclosed herein to reduce or eliminate aldehyde dehydrogenase activity can be confirmed using methods known in the art. For example, disruption of a particular aldehyde dehydrogenase could be confirmed with PCR screening using primers internal and external to the aldehyde dehydrogenase gene or by Southern blot using a probe designed to the aldehyde dehydrogenase gene sequence. Alternatively, one could utilize gas chromatography-mass spectroscopy or liquid chromatography to screen strains exposed to isobutyraldehyde for decreased formation of isobutyric acid. Accordingly, provided herein is a method of screening for strains with decreased isobutyric acid formation comprising: a)providing a strain comprising a modification in a polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity and/or a modification in a polynucleotide encoding a polypeptide having aldehyde oxidase activity; b) contacting the cell with isobutyraldehyde; and c) measuring

isobutyric acid formation; wherein isobutyric acid formation is reduced as compared to a control strain without the modification. In some embodiments, the modification is a deletion, mutation, and/or substitution. In some embodiments, the measuring is carried out using gas chromatography-mass spectroscopy. In some embodiments, isobutyric acid is reduced by at least about 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90%. In some embodiments, isobutyric acid formation is substantially eliminated.

Modification of Aldehyde Oxidase

In embodiments of the invention, a recombinant host cell disclosed herein can have a modification or disruption of a polynucleotide, gene or polypeptide encoding aldehyde oxidase. In embodiments, the recombinant host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide or gene encoding a polypeptide having aldehyde oxidase activity, or in an endogenous polypeptide having aldehyde oxidase activity. Such modifications, disruptions, deletions, mutations, and/or substitutions can result in aldehyde oxidase activity that is reduced or eliminated.

In embodiments of the invention, a polypeptide having aldehyde oxidase activity can catalyze the conversion of isobutyraldehyde to isobutyric acid. In other embodiments, the conversion of isobutyraldehyde to isobutyric acid in a recombinant host cell is reduced or eliminated. In other embodiments, a polynucleotide, gene or polypeptide having aldehyde oxidase activity can correspond to Enzyme Commission Number EC 1.2.3.1.

In embodiments, a recombinant host cell of the invention can be *Pichia stipitis* and a polynucleotide, gene or polypeptide having aldehyde oxidase activity can be AOX1 and/or AOX2.

Examples of aldehyde oxidase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, those of the following Table 5.

2017200439 20 Jan 2017

Table 5. Aldehyde oxidase target gene coding regions and proteins.

	Nucleic acid SEQ ID NO:	Amino acid SEQ ID NO:
AOX1 from Pichia stipitis	864	866
AOX2 from Pichia stipitis	867	868

Other examples of aldehyde oxidase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, aldehyde oxidase polynucleotides, genes and/or polypeptides having at least about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to any one of the sequences of Table 5, wherein such a polynucleotide or gene encodes a polypeptide having, or such a polypeptide has, aldehyde oxidase activity. Still other examples of aldehyde oxidase polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to an active variant, fragment or derivative of any one of the sequences of Table 5, wherein such a polynucleotide or gene encodes, or such a polypeptide has, aldehyde oxidase activity.

In embodiments, a polynucleotide, gene and/or polypeptide encoding an aldehyde oxidase sequence disclosed herein or known in the art can be modified, as disclosed above for acetolactate reductase or aldehyde dehydrogenase. In other embodiments, a polynucleotide, gene and/or polypeptide encoding aldehyde oxidase can be used to identify another aldehyde oxidase polynucleotide, gene and/or polypeptide sequence and/or can be used to identify an aldehyde oxidase homolog in other cells, as disclosed above for aldehyde dehydrogenase. Such aldehyde oxidase encoding sequences can be identified, for example, in the literature and/or in bioinformatics databases well known to the skilled person. For example, the identification of an aldehyde oxidase encoding sequence in another cell type using bioinformatics can be accomplished through BLAST (as disclosed above) searching of publicly available

databases with a known hexose kinase encoding DNA and polypeptide sequence, such as any of those provided herein. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

The modification of aldehyde oxidase in a recombinant host cell disclosed herein to reduce or eliminate aldehyde oxidase activity can be confirmed using methods known in the art. For example, disruption of a particular aldehyde oxidase could be confirmed with PCR screening using primers internal and external to the aldehyde oxidase gene or by Southern blot using a probe designed to the aldehyde oxidase gene sequence. Alternatively, one could utilize gas chromatography or other analytical methods to screen strains exposed to isobutyraldehyde for decreased formation of isobutyric acid (as described and demonstrated in the Examples). In some embodiments, isobutyric acid is reduced by at least about 10%, at least 20%, at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90%. In some embodiments, isobutyric acid formation is substantially eliminated.

Applicants have provided recombinant host cells comprising reduced or eliminated aldehyde dehydrogenase and/or aldehyde oxidase activity. In embodiments, a recombinant host cell disclosed herein can further comprise a modification in a polynucleotide encoding a polypeptide having pyruvate decarboxylase activity and/or a modification in a polynucleotide encoding a polypeptide having hexokinase 2 activity. In embodiments, a recombinant host cell of the invention can produce a production of a biosynthetic pathway (e.g., isobutanol), and can comprise a polynucleotide encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of: (a) pyruvate to acetolactate; (b) acetolactate to 2,3-dihydroxyisovalerate; (c) 2,3-dihydroxyisovalerate to 2-ketoisovalerate; (d) 2-ketoisovalerate to isobutyraldehyde; and (e) isobutyraldehyde to isobutanol. In other embodiments, such a recombinant host cell can produce a product of a biosynthetic pathway (e.g., isobutanol) at a yield or amount that is greater

than the yield or amount of the same product produced by a recombinant host cell that does not comprise reduced or eliminated aldehyde dehydrogenase activity and/or aldehyde oxidase activity. In other embodiments, a recombinant host cell of the invention can reduce or eliminate the conversion of isobutyraldehyde to isobutyric acid, and can be used for screening candidate polypeptides having aldehyde dehydrogenase and/or aldehyde oxidase activity. As such, Applicants have also provided methods of increasing the yield or titer of a product of a biosynthetic pathway (e.g., isobutanol), methods for reducing or eliminating the conversion of isobutyraldehyde to isobutyric acid, and methods for screening candidate polypeptides having aldehyde dehydrogenase and/or aldehyde oxidase activity.

In embodiments of the invention, methods of producing a recombinant host cell are provided which comprise (a) providing a recombinant host cell disclosed herein; and (b) transforming said host cell with a polynucleotide encoding a polypeptide of a biosynthetic pathway (e.g., an isobutanol biosynthetic pathway). In other embodiments, methods of producing a recombinant host cell are provided which comprise (a) providing a recombinant host cell comprising a modification in a polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity or in a polypeptide having aldehyde dehydrogenase activity; and (b) transforming said host cell with a polynucleotide encoding a polypeptide of a biosynthetic pathway (e.g., an isobutanol biosynthetic pathway). In other embodiments, methods of producing a recombinant host cell are provided which comprise (a) providing a recombinant host cell comprising a modification in a polynucleotide encoding a polypeptide having aldehyde oxidase activity or in a polypeptide having aldehyde oxidase activity; and (b) transforming said host cell with a polynucleotide encoding a polypeptide of an isobutanol biosynthetic pathway.

In embodiments, methods for reducing or eliminating the conversion of isobutyraldehyde to isobutyric acid are provided which comprise (a) providing a recombinant host cell disclosed herein; and (b)

2017200439 20 Jan 2017

growing said host cell under conditions wherein the conversion of isobutyraldehyde to isobutyric acid is reduced or eliminated compared to a recombinant host cell that does not comprise reduced or eliminated aldehyde dehydrogenase and/or aldehyde oxidase activity. The conversion of isobutyraldehyde to isobutyric acid of a recombinant host cell disclosed herein can be measured by methods known in the art (see, e.g., Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and/or described herein.

Reduction of DHMB

The production of DHMB in a host cell comprising an isobutanol biosynthetic pathway indicates that not all of the pathway substrates are being converted to the desired product. Thus, yield is lowered. In addition, DHMB can have inhibitory effects on product production. For example, DHMB can decrease the activity of enzymes in the biosynthetic pathway or have other inhibitory effects on yeast growth and/or productivity during fermentation. Thus, the methods described herein provide ways of reducing DHMB during fermentation. The methods include both methods of decreasing the production of DHMB and methods of removing DHMB from fermenting compositions.

Decreasing DHMB Production

In some embodiments described herein, a recombinant host cell can comprise reduced or eliminated ability to convert acetolactate to DHMB. The ability of a host cell to convert acetolactate to DHMB can be reduced or eliminated, for example, by a modification or disruption of a polynucleotide or gene encoding a polypeptide having acetolactate reductase activity or a modification or disruption of a polypeptide having acetolactate reductase activity. In other embodiments, the recombinant host cell can comprise a deletion, mutation, and/or substitution in an endogenous polynucleotide or gene encoding a polypeptide having acetolactate reductase activity or in an endogenous polypeptide having acetolactate reductase. Such modifications, disruptions, deletions,

mutations, and/or substitutions can result in acetolactate reductase activity that is reduced, substantially eliminated, or eliminated. In some embodiments of the invention, the product of the biosynthetic pathway is produced at a greater yield or amount compared to the production of the same product in a recombinant host cell that does not comprise reduced or eliminated ability to convert acetolactate to DHMB.

Thus, the product can be a composition comprising butanol that is substantially free of, or free of DHMB. In some embodiments, the composition comprising butanol contains no more than about 5 mM, about 4 mM, about 3 mM, about 2 mM, about 1 mM, about 0.5 mM, about 0.4 mM, about 0.3 mM DHMB, or about 0.2 mM DHMB.

The product can also be a composition comprising 2,3-butanediol (BDO) that is substantially free of, or free of DHMB. In some embodiments, the composition comprising BDO contains no more than about 5 mM, about 4 mM, about 3 mM, about 2 mM, about 1 mM, about 0.5 mM, about 0.4 mM, about 0.3 mM DHMB, or about 0.2 mM DHMB.

Any product of a biosynthetic pathway that involves the conversion of acetolactate to a substrate other than DHMB can be produced with greater effectiveness in a recombinant host cell disclosed herein having the described modification of acetolactate reductase activity. Such products include, but are not limited to, butanol, e.g., isobutanol, 2-butanol, and BDO, and branched chain amino acids.

In some embodiments, the host cell comprises at least one deletion, mutation, and/or substitution in at least one endogenous polynucleotide encoding a polypeptide having acetolactate reductase activity. In some embodiments, the host cell comprises at least one deletion, mutation, and/or substitution in each of at least two endogenous polynucleotides encoding polypeptides having acetolactate reductase activity.

In some embodiments, a polypeptide having acetolactate reductase activity can catalyze the conversion of acetolactate to DHMB. In some embodiments, a polypeptide having acetolactate reductase activity is

2017200439 20 Jan 2017

capable of catalyzing the reduction of acetolactate to 2S,3S-DHMB (fast DHMB) and/or 2S,3R-DHMB (slow DHMB).

Table 6

Polypeptides and polynucleotides having acetolactate reductase activity in *Saccharomyces cerevisiae*

Gene	SEQ ID NO: (nucleic acid, amino acid)
YMR226C	676, 677
YIL074C (Chr 9)	678, 679
YIR036C (Chr 9)	680, 681
YPL061W (ALD6)(Chr 16)	682, 683
YPL088W(Chr 16)	684, 685
YCR105W (ADH7)(Chr 3)	686, 687
YDR541C(Chr 4)	688, 689
YER081 (SER3)(Chr 5)	690, 691
YPL275W (FDH2)(Chr 16)	692, 693
YBR006W (UGA5)(Chr2)	694, 695
YOL059W (Chr 15)	696, 697
YER081W (Chr 5)	869, 870
YOR375C (Chr 15)	871, 872

In some embodiments, the conversion of acetolactate to DHMB in a recombinant host cell is reduced, substantially eliminated, or eliminated. In some embodiments, the polypeptide having acetolactate reductase activity is selected from the group consising of: YMR226C, YER081W, YIL074C, YBR006W, YPL275W, YOL059W, YIR036C, YPL061W, YPL088W, YCR105W, YOR375C, and YDR541C. In some embodiments, the polypeptide having acetolactate reductase activity is a polypeptide comprising a sequence listed in Table 6 or a sequence that is at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% identical to a polypeptide sequence listed in Table 6. In some embodiments, the polypeptide having acetolactate reducatase activity is a polypeptide encoded by a polynucleotide sequence listed in Table 6 or a sequence that is at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% identical to a polynucleotide sequence listed in Table 6.

2017200439 20 Jan 2017

Table 7: Example YMR226C Yeast Homologs

Species	Accession #	SEQ ID NO: (nucleic acid, amino acid)
Saccharomyces paradoxus	AABY01000127	698, 699
Saccharomyces bayanus	AACA01000631	700, 701 MSQGRKAAERLANKTVLITGASA GIGKATALEYLEASNGNMKLILAA RRLEKLEELKKTIDEEFPNAKVH VGQLDITQAEKIKPFIENLPEAFK DIDILINNAGKALGSERVGEIATQ DIQDVFDTNVTALINVTQAVLPIF QAKNSGDIVNLGLGGRQRRIPH RLHLLCFQVCRRCVH*QFEKGT DQHEDQSYLDRAGAG*DRVLTG QIQR**GTS*KRLQGHYAVDGRR RG*LNRIFHFQKAEHRGCRHPDL PHQPSLALPHLSRL* (SEQ ID NO:701) The sequence came from a comparative genomics study using "draft" genome sequences with 7-fold coverage (Kellis et al, Nature 423:241-254 (2003)).
Saccharomyces castellii	AACF01000116	702, 703
Saccharomyces mikatae	AACH01000019	704, 705
Ashbya gossypii	AE016819	706, 707
Candida glabrata	CR380959	708,709
Debaryomyces hansenii	CR382139	710, 711
Scheffersomyces stipitis (formerly *Pichia stipitis)*	XM_001387479	712, 713
Meyerozyma guilliermondii (formerly *Pichia guilliermondii)*	XM_001482184	714, 715
Vanderwaltozyma polyspora (formerly *Kluyveromyces polysporus)*	XM_001645671	716, 717
Candida dubliniensis	XM_002419771	718, 719
Zygosaccharomyces rouxii	XM_002494574	720, 721
Lachancea thermotolerans (formerly *Kluyveromyces thermotolerans)*	XM_002553230	722, 723
Kluyveromyces lactis	XM_451902	724, 725
Saccharomyces kluyveri	SAKL0H04730	726, 727
Yarrowia lipolytica	XM_501554	728, 729
Schizosaccharomyces pombe	NM_001018495	730, 731

In some embodiments, a polypeptide having acetolactate reductase activity is YMR226C or a homolog of YMR226C. Thus, in some embodiments, the polypeptide having acetolactate reducatase activity is a polypeptide comprising a sequence listed in Table 7 or a sequence that is at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% identical to a polypeptide sequence listed in Table 7. In some embodiments, the polypeptide having acetolactate reducatase activity is a polypeptide encoded by a polynucleotide sequence listed in Table 7 or a sequence that is at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, or at least about 99% identical to a polynucleotide sequence listed in Table 7. Acetolactate reductases capable of converting acetolactate to DHMB can be identified, for example, by screening genetically altered yeast for changes in acetolactate consumption, changes in DHMB production, changes in DHIV production, or changes in other downstream product (*e.g.*, butanol) production.

One way of identifying a gene involved in DHMB production comprises measuring the amount of DHMB produced by individual yeast strains in a yeast knock-out library. Knock-out libraries are available, for example, from Open Biosystems® (a division of Thermo Fisher Scientific, Waltham, MA). In this method, a decrease in DHMB production indicates that the gene that has been knocked-out functions to increase DHMB production, and an increase in DHMB production indicates that the gene that has been knocked-out functions to decrease DHMB production.

Two ways that a knockout ("KO") library can be used to identify candidate genes for involvement in DHMB synthesis include: (1) DHMB and DHIV accumulated in the culture during growth from endogenous substrates (acetolactate and NADPH or NADH) can be analyzed in samples from cultures. These samples can be placed in a hot (80-100 ºC) water bath for 10-20 min, or diluted into a solution such as 2% formic acid that will kill and permeabilize the cells. After either treatment, small

2017200439 20 Jan 2017

molecules will be found in the supernatant after centrifugation (5 min, 1100 x g). The DHMB/DHIV ratio of a control strain (e.g., BY4743) can be compared to that of the different KO derivatives, and the gene(s) missing from any strain(s) with exceptionally low DHMB/DHIV ratios can encode acetolactate reductase (ALR). (2) DHMB and/or DHIV formation rates in vitro from exogenous substrates (acetolactate and NADH and/or NADPH) can be measured in timed samples taken from a suspension of permeabilized cells, and inactivated in either of the ways described above. Since the substrates for DHMB and DHIV synthesis are the same, this allows one to measure the relative levels of ALR and KARI activity in the sample.

Another way of identifying a gene involved in DHMB production comprises measuring the amount of DHMB produced by individual yeast strains in a yeast overexpression library. Overexpression libraries are available, for example, from Open Biosystems® (a division of Thermo Fisher Scientific, Waltham, MA). In this method, a decrease in DHMB production indicates that the overexpressed gene functions to decrease DHMB production, and an increase in DHMB production indicates that the overexpressed gene functions to increase DHMB production.

Another way of identifying a gene involved in DHMB production is to biochemically analyze a DHMB-producing yeast strain. For example, DHMB-producing cells can be disrupted. This disruption can be performed at low pH and cold temperatures. The cell lysates can be separated into fractions, e.g., by adding ammonium sulfate or other techniques known to those of skill in the art, and the resulting fractions can be assayed for enzymatic activity. For example, the fractions can be assayed for the ability to convert acetolactate to DHMB. Fractions with enzymatic activity can be treated by methods known in the art to purify and concentrate the enzyme (e.g., dialysis and chromatographic separation). When a sufficient purity and concentration is achieved, the enzyme can be sequenced, and the corresponding gene encoding the acetolactate reductase capable of converting acetolactate to DHMB can be identified.

Furthermore, since the reduction of acetolactate to DHMB occurs in yeast, but does not occur to the same extent in *E. coli*, acetolactate reductases that are expressed in yeast, but not expressed in *E. coli*, can be selected for screening. Selected enzymes can be expressed in yeast or other protein expression systems and screened for the capability to convert acetolactate to DHMB.

Enzymes capable of catalyzing the conversion of acetolactate to DHMB can be screened by assaying for acetolactate levels, by assaying for DHMB levels, by assaying for DHIV levels, or by assaying for any of the downstream products in the conversion of DHIV to butanol, including isobutanol.

DHMB can be measured using any technique known to those of skill in the art. For example, DHMB can be separated and quantified by methods known to those of skill in the art and techniques described in the Examples provided herein. For example, DHMB can be separated and quantified using liquid chromatography-mass spectrometry, liquid chromatography-nuclear magnetic resonance (NMR), thin-layer chromatography, and/or HPLC with UV/Vis detection.

In embodiments, selected acetolactate reductase polynucleotides, genes and/or polypeptides disclosed herein can be modified or disrupted. Many suitable methods are known to those of ordinary skill in the art and include those described for aldehyde dehydrogenase (above).

The modification of acetolactate reductase in a recombinant host cell disclosed herein to reduce or eliminate acetolactate reductase activity can be confirmed using methods known in the art. For example, the presence or absence of an acetolactate reductase-encoding polynucleotide sequence can be determined using PCR screening. A decrease in acetolactate reductase activity can also be determined based on a reduction in conversion of acetolactate to DHMB. A decrease in acetolactate reductase activity can also be determined based on a reduction in DHMB production. A decrease in acetolactate reductase activity can also be determined based on an increase in butanol production.

Thus, in some embodiments, a yeast that is capable of producing butanol produces no more than about 5 mM, about 4 mM, about 3 mM, about 2 mM, about 1 mM, about 0.9 mM, about 0.8 mM., about 0.7 mM, about 0.6 mM, about 0.5 mM, about 0.4 mM or about 0.3 mM DHMB. In some embodiments, a yeast producing butanol produces no more than about about 5 mM, about 4 mM, about 3 mM, about 2 mM, about 1 mM, about 0.9 mM, about 0.8 mM., about 0.7 mM, about 0.6 mM, about 0.5 mM, about 0.4 mM or about 0.3 mM DHMB. In some embodiments, a yeast producing butanol produces no more than about 0.2 mM or 0.2 mM DHMB.

In some embodiments, a yeast capable of producing butanol produces no more than about about 10 mM DHMB when cultured under fermentation conditions for at least about 50 hours. In some embodiments, a yeast capable of producing butanol produces no more than about about 5 mM DHMB when cultured under fermentation conditions for at least about 20 hours, at least about 25 hours, at least about 30 hours, at least about 35 hours, at least about 40 hours, at least about 45 hours, or at least about 50 hours. In some embodiments, a yeast capable of producing butanol produced no more than about 3 mM DHMB when cultured under fermentation conditions for at least about 5 hours, at least about 10 hours, at least about 15 hours, at least about 20 hours, at least about 25 hours, at least about 30 hours, at least about 35 hours, at least about 40 hours, at least about 45 hours, or at least about 50 hours. In some embodiments, a yeast capable of producing butanol produced no more than about 1 mM DHMB when cultured under fermentation conditions for at least about 1 hour, at least about 5 hours, at least about 10 hours, at least about 15 hours, at least about 20 hours, at least about 25 hours, at least about 30 hours, at least about 35 hours, at least about 40 hours, at least about 45 hours, or at least about 50 hours. In some embodiments, a yeast capable of producing butanol produced no more than about 0.5 mM DHMB when cultured under fermentation conditions for at least about 1 hour, at least about 5 hours, at least about 10 hours, at least about 15 hours, at least about 20 hours, at least about 25 hours, at

least about 30 hours, at least about 35 hours, at least about 40 hours, at least about 45 hours, or at least about 50 hours.

In some embodiments, a yeast comprising at least one deletion, mutation, and/or substitution in an endogenous polynucleotide encoding an acetolactate reductase produces no more than about 0.5 times, about 0.4 times, about 0.3 times, about 0.2 times, about 0.1 times, about 0.05 times the amount of DHMB produced by a yeast containing the endogenous polynucleotide encoding an acelotacatate reductase when cultured under fermentation conditions for the same amount of time. In some embodiments, a yeast that is capable of producing butanol produces an amount of DHIV that is at least about 5 mM, at least about 6 mM, at least about 7 mM, at least about 8 mM, at least about 9 mM, or at least about 10 mM.

In some embodiments, a yeast that is capable of producing butanol produces an amount of DHIV that is at least about the amount of DHMB produced. In some embodiments, a yeast that is capable of producing butanol produces an amount of DHIV that is at least about twice, about three times, about five times, about ten times, about 15 times, about 20 times, about 25 times, about 30 times, about 35 times, about 40 times, about 45 times, or about 50 times the amount of DHMB produced. In some embodiments, a yeast that is capable of producing butanol produces DHIV at a rate that is at least about equal to the rate of DHMB production. In some embodiments, a yeast that is capable of producing butanol produces DHIV at a rate that is at least about twice, about three times, about five times, about ten times, about 15 times, about 20 times, about 25 times, about 30 times, about 35 times, about 40 times, about 45 times, or about 50 times the rate of DHMB production.

In some embodiments, a yeast that is capable of producing butanol produces less than 0.010 moles of DHMB per mole of glucose consumed. In some embodiments, a yeast produces less than about 0.009, less than about 0.008, less than about 0.007, less than about 0.006, or less than about 0.005 moles of DHMB per mole of glucose consumed. In some embodiments, a yeast produces less than about 0.004, less than about

0.003, less than about 0.002, or less than about 0.001 moles of DHMB per mole of glucose consumed.

In some embodiments, acetolactate reductase activity is inhibited by chemical means. For example, acetolactate reductase could be inhibited using other known substrates such as those listed in Fujisawa et al. including L-serine, D-serine, 2-methyl-DL-serine, D-threonine, L-allo-threonine, L-3-hydroxyisobutyrate, D-3-hydroxyisobutyrate, 3-hydroxypropionate, L-3-hydroxybutyrate, and D-3-hydroxybutyrate. Biochimica et Biophysica Acta 1645:89-94 (2003), which is herein incorporated by reference in its entirety.

DHMB Removal

In other embodiments described herein, a reduction in DHMB can be achieved by removing DHMB from a fermentation. Thus, fermentations with reduced DHMB concentrations are also described herein. Removal of DHMB can result, for example, in a product of greater purity, or a product requiring less processing to achieve a desired purity. Therefore, compositions comprising products of biosynthetic pathways such as ethanol or butanol with increased purity are also provided.

DHMB can be removed during or after a fermentation process and can be removed by any means known in the art. DHMB can be removed, for example, by extraction into an organic phase or reactive extraction.

In some embodiments, the fermentation broth comprises less than about 0.5 mM DHMB. In some embodiments, the fermentation broth comprises less than about 1.0 mM DHMB after about 5 hours, about 10 hours, about 15 hours, about 20 hours, about 25 hours, about 30 hours, about 35 hours, about 40 hours, about 45 hours, or about 50 hours of fermentation. In some embodiments, the fermentation broth comprises less than about 5.0 mM DHMB after about 20 hours, about 25 hours, about 30 hours, about 35 hours, about 40 hours, about 45 hours, or about 50 hours of fermentation.

Butanol Biosynthetic Pathways

2017200439 20 Jan 2017

Certain suitable isobutanol biosynthetic pathways are disclosed in U.S. Patents 7,851,188 and 7,993,889, each of which is incorporated by reference herein. A diagram of the disclosed isobutanol biosynthetic pathways is provided in Figure 1. As described in U.S. Patent 7,851,188 , steps in an example isobutanol biosynthetic pathway include conversion of:

- pyruvate to acetolactate (*see* Fig. 1, pathway step a therein), as catalyzed for example by acetolactate synthase (ALS),
- acetolactate to 2,3-dihydroxyisovalerate (*see* Fig. 1, pathway step b therein) as catalyzed for example by acetohydroxy acid isomeroreductase (KARI);
- 2,3-dihydroxyisovalerate to 2-ketoisovalerate (*see* Fig. 1, pathway step c therein) as catalyzed for example by acetohydroxy acid dehydratase, also called dihydroxy-acid dehydratase (DHAD);
- 2-ketoisovalerate to isobutyraldehyde (*see* Fig. 1, pathway step d therein) as catalyzed for example by branched-chain 2-keto acid decarboxylase; and
- isobutyraldehyde to isobutanol (*see* Fig. 1, pathway step e therein) as catalyzed for example by branched-chain alcohol dehydrogenase.

In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

– pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;

– acetolactate to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by ketol-acid reductoisomerase;

– 2,3-dihydroxyisovalerate to a-ketoisovalerate, which may be catalyzed, for example, by dihydroxyacid dehydratase;

– α-ketoisovalerate to valine, which may be catalyzed, for example, by transaminase or valine dehydrogenase;

– valine to isobutylamine, which may be catalyzed, for example, by valine decarboxylase;

– isobutylamine to isobutyraldehyde, which may be catalyzed by, for example, omega transaminase; and,

– isobutyraldehyde to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

In another embodiment, the isobutanol biosynthetic pathway comprises the following substrate to product conversions:

– pyruvate to acetolactate, which may be catalyzed, for example, by acetolactate synthase;

– acetolactate to 2,3-dihydroxyisovalerate, which may be catalyzed, for example, by acetohydroxy acid reductoisomerase;

– 2,3-dihydroxyisovalerate to α-ketoisovalerate, which may be catalyzed, for example, by acetohydroxy acid dehydratase;

–α-ketoisovalerate to isobutyryl-CoA, which may be catalyzed, for example, by branched-chain keto acid dehydrogenase;

–isobutyryl-CoA to isobutyraldehyde, which may be catalyzed, for example, by acelylating aldehyde dehydrogenase; and,

– isobutyraldehyde to isobutanol, which may be catalyzed, for example, by a branched-chain alcohol dehydrogenase.

In another embodiment, the isobutanol biosynthetic pathway comprises the substrate to product conversions shown as steps k, g, and e in Figure 1.

Genes and polypeptides that can be used for the substrate to product conversions described above as well as those for additional isobutanol pathways, are described in U.S. Patent Appl. Pub. No. 2007/0092957 and PCT Pub. No. WO 2011/019894, both incorporated by reference herein. US Appl. Pub. Nos. 2011/019894, 2007/0092957, 2010/0081154, which are herein incorporated by reference, describe dihydroxyacid dehydratases including those from *Lactococcus lactis* and *Streptococcus mutans*. Ketoisovalerate decarboxylases include those derived from *Lactococcus lactis*, *Macrococcus caseolyticus* (SEQ ID NO: 542) and *Listeria grayi* (SEQ ID NO: 543). U.S. Patent Appl. Publ. No. 2009/0269823 and U.S. Appl. Publ. No. 2011/0269199, incorporated by reference, describe alcohol dehydrogenases, including those that utilize NADH as a cofactor. Alcohol dehydrogenases include SadB from *Achromobacter xylosoxidans*. Additional alcohol dehydrogenases include

2017200439 20 Jan 2017

horse liver ADH and *Beijerinkia indica* ADH. Alcohol dehydrogenases include those that utilize NADH as a cofactor. In one embodiment a butanol biosynthetic pathway comprises a) a ketol-acid reductoisomerase that has a K_M for NADH less than about 300 μM, less than about 100 μM, less than about 50 μM, less than about 20 μM or less than about 10 μM; b) an alcohol dehydrogenase that utilizes NADH as a cofactor; or c) both a) and b).

WO 2011/019894 and US Appl. Pub. Nos. 2011/019894, 2007/0092957, 2010/0081154, which are herein incorporated by reference in their entireties, describe suitable dihydroxyacid dehydratases. Methods of increasing DHAD activity are described, for example, in U.S. Patent Application Publication No. 2010/0081173 and U.S. Patent Application No. 13/029,558, filed February 17, 2011, which are herein incorporated by reference in their entireties.

Suitable ketol-acid reductoisomerase (KARI) enzymes are described in U.S. Patent Appl. Pub. Nos. 2008/0261230 A1, 2009/0163376, 2010/0197519, 2010/0143997 and 2011/0244536, which are herein incorporated by reference in their entireties. Examples of KARIs disclosed therein are those from *Vibrio cholerae, Pseudomonas aeruginosa* PAO1, and *Pseudomonas fluorescens* PF5. In some embodiments, the KARI enzyme has a specific activity of at least about 0.1 micromoles/min/mg, at least about 0.2 micromoles/min/mg, at least about 0.3 micromoles/min/mg, at least about 0.4 micromoles/min/mg, at least about 0.5 micromoles/min/mg, at least about 0.6 micromoles/min/mg, at least about 0.7 micromoles/min/mg, at least about 0.8 micromoles/min/mg, at least about 0.9 micromoles/min/mg, at least about 1.0 micromoles/min/mg, or at least about 1.1 micromoles/min/mg. Suitable polypeptides to catalyze the substrate to product conversion acetolactate to 2,3-dihydroxyisovalerate include those that that have a KM for NADH less than about 300 μM, less than about 100 μM, less than about 50 μM, less than about 25 μM or less than about 10μM.

In some embodiments, the KARI utilizes NADPH. Methods of measuring NADPH consumption are known in the art. For example, US

2017200439 20 Jan 2017

Published Application No. 2008/0261230, which is herein incorporated by reference in its entirety, provides methods of measuring NADPH consumption. In some embodiments, an NADPH consumption assay is a method that measures the disappearance of the cofactor, NADPH, during the enzymatic conversion of acetolactate to α-β-dihydroxy-isovalerate at 340 nm. The activity is calculated using the molar extinction coefficient of 6220 $M^{-1}cm^{-1}$ for NADPH and is reported as μmole of NADPH consumed per min per mg of total protein in cell extracts (see Aulabaugh and Schloss, Biochemistry 29: 2824-2830, 1990).

In some embodiments, the KARI is capable of utilizing NADH. In some embodiments, the KARI is capable of utilizing NADH under anaerobic conditions. KARI enzymes using NADH are described, for example, in U.S. Patent Application Publication No. 2009/0163376, which is herein incorporated by reference in its entirety.

Additional genes that can be used can be identified by one skilled in the art through bioinformatics or using methods well-known in the art.

Additionally described in U.S. Patent Application Publication No. US 2007/0092957 A1, which is incorporated by reference herein, are construction of chimeric genes and genetic engineering of bacteria and yeast for isobutanol production using the disclosed biosynthetic pathways.

Biosynthetic pathways for the production of 1-butanol that may be used include those described in U.S. Appl. Pub. No. 2008/0182308, which is incorporated herein by reference. In one embodiment, the 1-butanol biosynthetic pathway comprises the following substrate to product conversions:

– a) acetyl-CoA to acetoacetyl-CoA, which may be catalyzed, for example, by acetyl-CoA acetyl transferase;

– b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, which may be catalyzed, for example, by 3-hydroxybutyryl-CoA dehydrogenase;

– c) 3-hydroxybutyryl-CoA to crotonyl-CoA, which may be catalyzed, for example, by crotonase;

– d) crotonyl-CoA to butyryl-CoA, which may be catalyzed, for example, by butyryl-CoA dehydrogenase;

– e) butyryl-CoA to butyraldehyde, which may be catalyzed, for example, by butyraldehyde dehydrogenase; and,

– f) butyraldehyde to 1-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

Biosynthetic pathways for the production of 2-butanol that may be used include those described in U.S. Appl. Pub. No. 2007/0259410 and U.S. Appl. Pub. No. 2009/0155870, which are incorporated herein by reference. In one embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

– a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;

– b) alpha-acetolactate to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;

– c) acetoin to 3-amino-2-butanol, which may be catalyzed, for example, acetonin aminase;

– d) 3-amino-2-butanol to 3-amino-2-butanol phosphate, which may be catalyzed, for example, by aminobutanol kinase;

– e) 3-amino-2-butanol phosphate to 2-butanone, which may be catalyzed, for example, by aminobutanol phosphate phosphorylase; and,

– f) 2-butanone to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

In another embodiment, the 2-butanol biosynthetic pathway comprises the following substrate to product conversions:

– a) pyruvate to alpha-acetolactate, which may be catalyzed, for example, by acetolactate synthase;

– b) alpha-acetolactate to acetoin, which may be catalyzed, for example, by acetolactate decarboxylase;

– c) acetoin to 2,3-butanediol, which may be catalyzed, for example, by butanediol dehydrogenase;

– d) 2,3-butanediol to 2-butanone, which may be catalyzed, for example, by dial dehydratase; and,

– e) 2-butanone to 2-butanol, which may be catalyzed, for example, by butanol dehydrogenase.

In some embodiments of the invention, a recombinant host cell comprises a biosynthetic pathway. The biosynthetic pathway can comprise reduced or eliminated aldehyde dehydrogenase activity and an isobutanol or 1-butanol biosynthetic pathway wherein the pathway comprises the substrate to product conversion pyruvate to acetolactate. In some embodiments, a host cell comprising a biosynthetic pathway capable of converting pyurvate to acetolacatate comprises a polynucleotide encoding a polypeptide having acetolactate synthase activity. For example, the biosynthetic pathway can be a butanol producing pathway or a butanediol producing pathway. The biosynthetic pathway can also be a branched-chain amino acid (e.g., leucine, isoleucine, valine) producing pathway.

In other embodiments, the recombinant host cell can comprise an isobutanol, 1-butanol, or a 2-butanol biosynthetic pathway as described herein. In some embodiments, the butanol biosynthetic pathway is an isobutanol biosynthetic pathway. Production of isobutanol or 2-butanol in a recombinant host cell disclosed herein may benefit from a reduction, substantial elimination or elimination of an acetolactate reductase activity.

Modifications

Functional deletion of the pyruvate decarboxylase gene has been used to increase the availability of pyruvate for utilization in biosynthetic product pathways. For example, U.S. Application Publication No. US 2007/0031950 A1, which is herein incorporated by reference in its entirety, discloses a yeast strain with a disruption of one or more pyruvate decarboxylase genes and expression of a D-lactate dehydrogenase gene, which is used for production of D-lactic acid. U.S.Application Publication No. US 2005/0059136 A1, which is herein incorporated by reference in its entirety, discloses glucose tolerant two carbon source independent (GCSI) yeast strains with no pyruvate decarboxylase activity, which can have an exogenous lactate dehydrogenase gene. Nevoigt and Stahl (*Yeast* 12:1331-1337 (1996)) describe the impact of reduced pyruvate decarboxylase and increased NAD-dependent glycerol-3-phosphate

dehydrogenase in *Saccharomyces cerevisiae* on glycerol yield. U.S. Appl. Pub. No. 2009/0305363, which is herein incorporated by reference in its entirety, discloses increased conversion of pyruvate to acetolactate by engineering yeast for expression of a cytosol-localized acetolactate synthase and substantial elimination of pyruvate decarboxylase activity.

In embodiments of the invention, a recombinant host cell disclosed herein can comprise a modification in an endogenous polynucleotide encoding a polypeptide having pyruvate decarboxylase (PDC) activity or a modification in an endogenous polypeptide having PDC activity. In embodiments, a recombinant host cell disclosed herein can have a modification or disruption of a polynucleotide, gene and/or polypeptide encoding PDC. In embodiments, a recombinant host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide or gene encoding a polypeptide having PDC activity, or in an endogenous polypeptides having PDC activity. Such modifications, disruptions, deletions, mutations, and/or substitutions can result in PDC activity that is reduced or eliminated, resulting, for example, in a PDC knock-out (PDC-KO) phenotype.

In embodiments of the invention, an endogenous pyruvate decarboxylase activity of a recombinant host cell disclosed herein converts pyruvate to acetaldehyde, which can then be converted to ethanol or to acetyl-CoA via acetate. In other embodiments, a recombinant host cell is *Kluyveromyces lactis* containing one gene encoding pyruvate decarboxylase, *Candida glabrata* containing one gene encoding pyruvate decarboxylase, or *Schizosaccharomyces pombe* containing one gene encoding pyruvate decarboxylase.

In other embodiments, a recombinant host cell is *Saccharomyces cerevisiae* containing three isozymes of pyruvate decarboxylase encoded by the PDC1, PDC5, and PDC6 genes, as well as a pyruvate decarboxylase regulatory gene, PDC2. In a non-limiting example in *S. cerevisiae*, the PDC1 and PDC5 genes, or the PDC1, PDC5, and PDC6 genes, are disrupted. In another non-limiting example in *S. cerevisiae*, pyruvate decarboxylase activity can be reduced by disrupting the PDC2

regulatory gene. In another non-limiting example in *S. cerevisiae*, polynucleotides or genes encoding pyruvate decarboxylase proteins such as those having about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to PDC1, PDC2, PDC5 and/or PDC6 can be disrupted.

In embodiments, a polypeptide having PDC activity or a polynucleotide or gene encoding a polypeptide having PDC activity corresponds to Enzyme Commission Number EC 4.1.1.1. In other embodiments, a PDC gene of a recombinant host cell disclosed herein is not active under the fermentation conditions used, and therefore such a gene would not need to be modified or inactivated.

Examples of recombinant host cells with reduced pyruvate decarboxylase activity due to disruption of pyruvate decarboxylase encoding genes have been reported, such as for *Saccharomyces* in Flikweert *et al.* (*Yeast* (1996) 12:247-257), for *Kluyveromyces* in Bianchi *et al.* (*Mol. Microbiol.* (1996) 19(1):27-36), and disruption of the regulatory gene in Hohmann (*Mol. Gen. Genet.* (1993) 241:657-666). *Saccharomyces* strains having no pyruvate decarboxylase activity are available from the ATCC with Accession #200027 and #200028. Examples of PDC polynucleotides, genes and/or polypeptides that can be targeted for modification or inactivation in the recombinant host cells disclosed herein include, but are not limited to, those of the following Table 8.

Table 8. Pyruvate decarboxylase target gene coding regions and proteins.

Description	SEQ ID NO: Nucleic acid	SEQ ID NO: Amino acid
PDC1 pyruvate decarboxylase from *Saccharomyces cerevisiae*	648	649
PDC5 pyruvate decarboxylase from *Saccharomyces cerevisiae*	650	651
PDC6 pyruvate decarboxylase from *Saccharomyces cerevisiae*	652	653
pyruvate decarboxylase from *Candida glabrata*	654	655

PDC1 pyruvate decarboxylase from *Pichia stipitis*	656	657
PDC2 pyruvate decarboxylase from *Pichia stipitis*	658	659
pyruvate decarboxylase from *Kluyveromyces lactis*	660	661
pyruvate decarboxylase from *Yarrowia lipolytica*	662	663
pyruvate decarboxylase from *Schizosaccharomyces pombe*	664	665
pyruvate decarboxylase from *Zygosaccharomyces rouxii*	666	667

Other examples of PDC polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, PDC polynucleotides, genes and/or polypeptides having at least about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to any one of the sequences of Table 8, wherein such a polynucleotide or gene encodes, or such a polypeptide has, PDC activity. Still other examples of PDC polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to an active variant, fragment or derivative of any one of the sequences of Table 8, wherein such a polynucleotide or gene encodes, or such a polypeptide has, PDC activity.

In embodiments, a polynucleotide, gene and/or polypeptide encoding a PDC sequence disclosed herein or known in the art can be modified, as disclosed above for aldehyde dehydrogenase. In other embodiments, a polynucleotide, gene and/or polypeptide encoding PDC can be used to identify another PDC polynucleotide, gene and/or polypeptide sequence or to identify a PDC homolog in other cells, as disclosed above for acetolactate dehydrogenase. Such a PDC encoding sequence can be identified, for example, in the literature and/or in bioinformatics databases well known to the skilled person. For example, the identification of a PDC encoding sequence in other cell types using bioinformatics can be accomplished through BLAST (as described above) searching of publicly available databases with a known PDC encoding

2017200439 20 Jan 2017

DNA and polypeptide sequence, such as those provided herein. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

The modification of PDC in a recombinant host cell disclosed herein to reduce or eliminate PDC activity can be confirmed using methods known in the art. For example, disruption of a particular pyruvate decarboxylase could be confirmed with PCR screening using primers external to the gene sequence, or by Southern blot using a probe designed to the pyruvate decarboxylase gene sequence. Alternatively, one could utilize analytical methods such as gas chromatography or HPLC to screen strains for decreased or eliminated production of acetaldehyde and/or ethanol.

Functional deletion of the hexokinase 2 gene has been used to reduce glucose repression and to increase the availability of pyruvate for utilization in biosynthetic pathways. For example, International Publication No. WO 2000/061722 A1, which is incorporated herein by reference in its entirety discloses the production of yeast biomass by aerobically growing yeast having one or more functionally deleted hexokinase 2 genes or analogs. In addition, Rossell et al. (Yeast Research 8:155-164 (2008)) found that Saccharomyces cerevisiae with a deletion of the hexokinase 2 gene showed 75% reduction in fermentative capacity, defined as the specific rate of carbon dioxide production under sugar-excess and anaerobic conditions. After starvation, the fermentation capacity was similar to that of a strain without the hexokinase 2 gene deletion. Diderich et al. (Applied and Environmental Microbiology 67:1587-1593 (2001)) found that S. cerevisiae with a deletion of the hexokinase 2 gene had lower pyruvate decarboxylase activity.

In embodiments, a recombinant host cell disclosed herein can comprise a modification in an endogenous polynucleotide encoding a polypeptide having hexokinase 2 activity and/or a modification in a polypeptide having hexokinase 2 activity. In embodiments, a recombinant host cell disclosed herein can have a modification or disruption of a

polynucleotide, gene or polypeptide encoding hexokinase 2. In embodiments, a recombinant host cell comprises a deletion, mutation, and/or substitution in an endogenous polynucleotide or gene encoding a polypeptide having hexokinase 2 activity, or an endogenous polypeptide having hexokinase 2 activity. Such modifications, disruptions, deletions, mutations, and/or substitutions can result in hexokinase 2 activity that is reduced or eliminated, resulting, for example, in a hexokinase 2 knockout (HXK2-KO) phenotype. In embodiments, the host cell comprises a modification as described in U.S. Appn. Serial. Nos. 2011/0124060 A1 or 2012/0015416 A1, which are incorporated herein by reference in their entireties.

In embodiments, a polypeptide having hexokinase 2 activity can catalyze the conversion of hexose to hexose-6-phosphate, and/or can catalyze the conversion of D-glucose to D-glucose 6-phosphate, D-fructose to D-fructose 6-phosphate, and/or D-mannose to D-mannose 6-phosphate. In other embodiments, a polynucleotide, gene or polypeptide having hexokinase 2 activity can correspond to Enzyme Commission Number EC 2.7.1.1.

In embodiments of the invention, a recombinant host cell can be *S. cerevisiae* and a polynucleotide, gene or polypeptide having hexokinase 2 activity can be HXK2. In other embodiments, a recombinant host cell can be K. lactis and a polynucleotide, gene or polypeptide having hexokinase 2 activity can be RAG5. In other embodiments, a recombinant host cell can be H. polymorpha and a polynucleotide, gene or polypeptide having hexokinase 2 activity can be HPGLK1. In other embodiments, a recombinant host cell can be S. pombe and a polynucleotide, gene or polypeptide having hexokinase 2 activity can be HXK2.

Examples of hexokinase 2 polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, those of the following Table 9.

Table 9. Hexokinase 2 target gene coding regions and proteins.

2017200439 20 Jan 2017

HXK2 from *S. cerevisiae*	Nucleic acid (SEQ ID NO: 668)
	Amino acid (SEQ ID NO: 669)
RAG5 from *K. lactis*	Nucleic acid (SEQ ID NO: 670):
	Amino acid (SEQ ID NO: 671):
HPGLK1 from *H. polymorpha*	Nucleic acid (SEQ ID NO: 672)
	Amino acid (SEQ ID NO: 673)
HXK2 from *S. pombe*	Nucleic acid (SEQ ID NO: 674):
	Amino acid (SEQ ID NO: 675):

Other examples of hexokinase 2 polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to, hexokinase 2 polynucleotides, genes and/or polypeptides having at least about 70% to about 75%, about 75% to about 80%, about 80% to about 85%, about 85% to about 90%, about 90% to about 95%, about 96%, about 97%, about 98%, or about 99% sequence identity to any one of the sequences of Table 9, wherein such a polynucleotide or gene encodes, or such a polypeptide has, hexokinase 2 activity. Still other examples of hexokinase 2 polynucleotides, genes and polypeptides that can be targeted for modification or inactivation in a recombinant host cell disclosed herein include, but are not limited to an active variant, fragment or derivative of any one of the sequences of Table 9, wherein such a polynucleotide or gene encodes, or such a polypeptide has, hexokinase 2 activity.

In embodiments, a polynucleotide, gene and/or polypeptide encoding a hexokinase 2 sequence disclosed herein or known in the art can be modified or disrupted, as disclosed above for aldehyde dehydrogenase. In other embodiments, a polynucleotide, gene and/or polypeptide encoding hexokinase 2 can be used to identify another

2017200439 20 Jan 2017

hexokinase 2 polynucleotide, gene and/or polypeptide sequence or to identify a hexokinase 2 homolog in other cells, as disclosed above for aldehyde dehydrogenase. Such a hexokinase 2 encoding sequence can be identified, for example, in the literature and/or in bioinformatics databases well known to the skilled person. For example, the identification of a hexokinase 2 encoding sequence in other cell types using bioinformatics can be accomplished through BLAST (as described above) searching of publicly available databases with a known hexokinase 2 encoding DNA and polypeptide sequence, such as those provided herein. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

The modification of hexokinase 2 in a recombinant host cell disclosed herein to reduce or eliminate hexokinase 2 activity can be confirmed using methods known in the art. For example, disruption of hexokinase 2 could be confirmed with PCR screening using primers external to the hexokinase 2 gene, or by Southern blot using a probe designed to the hexokinase 2 gene sequence. Alternatively, one could examine putative hexokinase 2 knockout strains for increased biomass yield on glucose-containing media.

Examples of additional modifications that can be useful in cells provided herein include modifications to reduce glycerol-3-phosphate dehydrogenase activity and/or disruption in at least one gene encoding a polypeptide having pyruvate decarboxylase activity or a disruption in at least one gene encoding a regulatory element controlling pyruvate decarboxylase gene expression as described in U.S. Patent Appl. Pub. No. 2009/0305363 (incorporated herein by reference), modifications to a host cell that provide for increased carbon flux through an Entner-Doudoroff Pathway or reducing equivalents balance as described in U.S. Patent Appl. Pub. No. 2010/0120105 (incorporated herein by reference). Other modifications include integration of at least one polynucleotide encoding a polypeptide that catalyzes a step in a pyruvate-utilizing biosynthetic pathway described in PCT Appn. Pub. No. WO 2012/033832,

2017200439 20 Jan 2017

which is herein incorporated by reference in its entirety. A genetic modification which has the effect of reducing glucose repression wherein the yeast production host cell is pdc- is described in U.S. Appl. Publ No. US 2011/0124060, which is herein incorporated by reference in its entirety.

U.S. Appl. Publ. No. 20120064561A1, which is herein incorporated by reference, discloses recombinant host cells comprising (a) at least one heterologous polynucleotide encoding a polypeptide having dihydroxy-acid dehydratase activity; and (b)(i) at least one deletion, mutation, and/or substitution in an endogenous gene encoding a polypeptide affecting Fe-S cluster biosynthesis; and/or (ii) at least one heterologous polynucleotide encoding a polypeptide affecting Fe-S cluster biosynthesis. In embodiments, the polypeptide affecting Fe-S cluster biosynthesis is encoded by *AFT1*, *AFT2*, *FRA2*, *GRX3*, or *CCC1*. In embodiments, the polypeptide affecting Fe-S cluster biosynthesis is constitutive mutant *AFT1* L99A, *AFT1* L102A, *AFT1* C291F, or *AFT1* C293F.

Additionally, host cells can comprise heterologous polynucleotides encoding a polypeptides with phosphoketolase activity and/or a heterologous polynucleotide encoding a polypeptide with phosphotransacetylase activity such as, for example, those encoded by SEQ ID NOs: 962 and 963, and as described in PCT Appn. Pub. No. WO 2011/159853, which is herein incorporated by reference in its entirety.

Isobutanol and Other Products

In embodiments of the invention, methods for the production of a product of a biosynthetic pathway are provided which comprise (a) providing a recombinant host cell disclosed herein; and (b) growing the host cell under conditions whereby the product of the biosynthetic pathway is produced. In other embodiments, the product is produced as a co-product along with ethanol. In still other embodiments, the product of the biosynthetic pathway is isobutanol.

In other embodiments of the invention, the product of the biosynthetic pathway is produced at a greater yield or amount compared

2017200439 20 Jan 2017

to the production of the same product in a recombinant host cell that does not comprise reduced or eliminated aldehyde dehydrogenase and/or aldehyde oxidase activity and/or acetolactate reductase activity. In embodiments, yield is increased by at least about 2%, at least about 5% or at least about 10%. In embodiments, this greater yield includes production at a yield of greater than about 10% of theoretical, at a yield of greater than about 20% of theoretical, at a yield of greater than about 25% of theoretical, at a yield of greater than about 30% of theoretical, at a yield of greater than about 40% of theoretical, at a yield of greater than about 50% of theoretical, at a yield of greater than about 60% of theoretical, at a yield of greater than about 70% of theoretical, at a yield of greater than about 75% of theoretical, at a yield of greater than about 80% of theoretical at a yield of greater than about 85% of theoretical, at a yield of greater than about 90% of theoretical, at a yield of greater than about 95% of theoretical, at a yield of greater than about 96% of theoretical, at a yield of greater than about 97% of theoretical, at a yield of greater than about 98% of theoretical, at a yield of greater than about 99% of theoretical, or at a yield of about 100% of theoretical. In other embodiments, the product is produced as a co-product along with ethanol. In still other embodiments, the product of the biosynthetic pathway is isobutanol.

Any product of a biosynthetic pathway that has the conversion of isobutyraldehyde to isobutyric acid as a pathway by-product can be produced with greater effectiveness in a recombinant host cell disclosed herein having the described modification of aldehyde dehydrogenase and/or aldehyde oxidase activity. A list of such products includes, but is not limited to, isobutanol.

Microbial Hosts for Isobutanol Production

Microbial hosts for isobutanol production can be selected from bacteria, cyanobacteria, filamentous fungi and yeasts. The microbial host used for butanol production should be tolerant to isobutanol so that the yield is not limited by butanol toxicity. Although butanol-tolerant mutants have been isolated from solventogenic *Clostridia*, little information is

2017200439 20 Jan 2017

available concerning the butanol tolerance of other potentially useful bacterial strains. Most of the studies on the comparison of alcohol tolerance in bacteria suggest that butanol is more toxic than ethanol (de Cavalho, et al., Microsc. Res. Tech., 64: 215-22, 2004) and (Kabelitz, et al., FEMS Microbiol. Lett., 220: 223-227, 2003, Tomas, et al., J. Bacteriol., 186: 2006-2018, 2004) report that the yield of 1-butanol during fermentation in *Clostridium acetobutylicum* can be limited by 1-butanol toxicity. The primary effect of 1-butanol on *Clostridium acetobutylicum* is disruption of membrane functions (Hermann et al., Appl. Environ. Microbiol., 50: 1238-1243, 1985).

The microbial hosts selected for the production of isobutanol should be tolerant to isobutanol and should be able to convert carbohydrates to isobutanol. The criteria for selection of suitable microbial hosts include the following: intrinsic tolerance to isobutanol, high rate of glucose utilization, availability of genetic tools for gene manipulation, and the ability to generate stable chromosomal alterations.

Suitable host strains with a tolerance for isobutanol can be identified by screening based on the intrinsic tolerance of the strain. The intrinsic tolerance of microbes to isobutanol can be measured by determining the concentration of isobutanol that is responsible for 50% inhibition of the growth rate (IC_{50}) when grown in a minimal medium. The IC_{50} values can be determined using methods known in the art. For example, the microbes of interest can be grown in the presence of various amounts of isobutanol and the growth rate monitored by measuring the optical density at 600 nanometers. The doubling time can be calculated from the logarithmic part of the growth curve and used as a measure of the growth rate. The concentration of isobutanol that produces 50% inhibition of growth can be determined from a graph of the percent inhibition of growth versus the isobutanol concentration. In one embodiment, the host strain has an IC_{50} for isobutanol of greater than about 0.5%.

The microbial host for isobutanol production should also utilize glucose at a high rate. Most microbes are capable of metabolizing

2017200439 20 Jan 2017

carbohydrates. However, certain environmental microbes cannot metabolize carbohydrates to high efficiency, and therefore would not be suitable hosts.

The ability to genetically modify the host is essential for the production of any recombinant microorganism. The mode of gene transfer technology can be by electroporation, conjugation, transduction or natural transformation. A broad range of host conjugative plasmids and drug resistance markers are available. The cloning vectors are tailored to the host microorganisms based on the nature of antibiotic resistance markers that can function in that host.

The microbial host also has to be manipulated in order to inactivate competing pathways for carbon flow by deleting various genes. This requires the availability of either transposons to direct inactivation or chromosomal integration vectors. Additionally, the production host should be amenable to chemical mutagenesis so that mutations to improve intrinsic isobutanol tolerance can be obtained.

Based on the criteria described above, suitable microbial hosts for the production of isobutanol include, but are not limited to, members of the genera *Clostridium*, *Zymomonas*, *Escherichia, Salmonella, Rhodococcus, Pseudomonas, Bacillus, Vibrio, Lactobacillus, Enterococcus, Alcaligenes, Klebsiella, Paenibacillus, Arthrobacter, Corynebacterium, Brevibacterium, Pichia*, *Candida*, *Issatchenkia, Hansenula, Kluyveromyces,* and *Saccharomyces.* Suitable hosts include: *Escherichia coli, Alcaligenes eutrophus, Bacillus licheniformis, Paenibacillus macerans, Rhodococcus erythropolis, Pseudomonas putida, Lactobacillus plantarum, Enterococcus faecium, Enterococcus gallinarium, Enterococcus faecalis, Bacillus subtilis* and *Saccharomyces cerevisiae.* In some embodiments, the host cell is *Saccharomyces cerevisiae. S. cerevisiae* yeast are known in the art and are available from a variety of sources, including, but not limited to, American Type Culture Collection (Rockville, MD), Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, LeSaffre, Gert Strand AB, Ferm Solutions, North American Bioproducts, Martrex, and Lallemand. *S. cerevisiae* include, but

2017200439 20 Jan 2017

are not limited to, BY4741, CEN.PK 113-7D, Ethanol Red® yeast, Ferm Pro™ yeast, Bio-Ferm® XR yeast, Gert Strand Prestige Batch Turbo alcohol yeast, Gert Strand Pot Distillers yeast, Gert Strand Distillers Turbo yeast, FerMax™ Green yeast, FerMax™ Gold yeast, Thermosacc® yeast, BG-1, PE-2, CAT-1, CBS7959, CBS7960, and CBS7961.

Construction of Production Host

Recombinant microorganisms containing the necessary genes that will encode the enzymatic pathway for the conversion of a fermentable carbon substrate to butanol can be constructed using techniques well known in the art. In the present invention, genes encoding the enzymes of one of the isobutanol biosynthetic pathways of the invention, for example, acetolactate synthase, acetohydroxy acid isomeroreductase, acetohydroxy acid dehydratase, branched-chain α-keto acid decarboxylase, and branched-chain alcohol dehydrogenase, can be isolated from various sources, as described above.

Methods of obtaining desired genes from a genome are common and well known in the art of molecular biology. For example, if the sequence of the gene is known, suitable genomic libraries can be created by restriction endonuclease digestion and can be screened with probes complementary to the desired gene sequence. Once the sequence is isolated, the DNA can be amplified using standard primer-directed amplification methods such as polymerase chain reaction (U.S. 4,683,202) to obtain amounts of DNA suitable for transformation using appropriate vectors. Tools for codon optimization for expression in a heterologous host are readily available. Some tools for codon optimization are available based on the GC content of the host microorganism.

Once the relevant pathway genes are identified and isolated they can be transformed into suitable expression hosts by means well known in the art. Vectors or cassettes useful for the transformation of a variety of host cells are common and commercially available from companies such as EPICENTRE® (Madison, WI), Invitrogen Corp. (Carlsbad, CA), Stratagene (La Jolla, CA), and New England Biolabs, Inc. (Beverly, MA).

Typically the vector or cassette contains sequences directing transcription and translation of the relevant gene, a selectable marker, and sequences allowing autonomous replication or chromosomal integration. Suitable vectors comprise a region 5' of the gene which harbors transcriptional initiation controls and a region 3' of the DNA fragment which controls transcriptional termination. Both control regions can be derived from genes homologous to the transformed host cell, although it is to be understood that such control regions can also be derived from genes that are not native to the specific species chosen as a production host.

Initiation control regions or promoters, which are useful to drive expression of the relevant pathway coding regions in the desired host cell are numerous and familiar to those skilled in the art. Virtually any promoter capable of driving these genetic elements, including those used in the Examples, is suitable for the present invention including, but not limited to, *CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI* (useful for expression in *Saccharomyces*); *AOX1* (useful for expression in *Pichia*); and *lac, ara, tet, trp*, IP_L, IP_R, *T7, tac, and trc* (useful for expression in *Escherichia coli*, *Alcaligenes*, and *Pseudomonas*) as well as the *amy, apr, npr* promoters and various phage promoters useful for expression in *Bacillus subtilis, Bacillus licheniformis, and Paenibacillus macerans.* For yeast recombinant host cells, a number of promoters can be used in constructing expression cassettes for genes, including, but not limited to, the following constitutive promoters suitable for use in yeast: FBA1, TDH3 (GPD), ADH1, ILV5, and GPM1; and the following inducible promoters suitable for use in yeast: GAL1, GAL10, OLE1, and CUP1. Other yeast promoters include hybrid promoters UAS(PGK1)-FBA1p (SEQ ID NO: 406), UAS(PGK1)-ENO2p (SEQ ID NO: 538), UAS(FBA1)-PDC1p (SEQ ID NO: 539), UAS(PGK1)-PDC1p (SEQ ID NO: 540), and UAS(PGK)-OLE1p (SEQ ID NO: 541).

Promoters, transcriptional terminators, and coding regions can be cloned into a yeast 2 micron plasmid and transformed into yeast cells

2017200439 20 Jan 2017

(Ludwig, et al. Gene, 132: 33-40, 1993; US Appl. Pub. No. 20080261861A1).

Adjusting the amount of gene expression in a given host may be achieved by varying the level of transcription, such as through selection of native or artificial promoters. In addition, techniques such as the use of promoter libraries to achieve desired levels of gene transcription are well known in the art. Such libraries can be generated using techniques known in the art, for example, by cloning of random cDNA fragments in front of gene cassettes (Goh *et al.* (2002) *AEM* 99, 17025), by modulating regulatory sequences present within promoters (Ligr *et al.* (2006) *Genetics* 172, 2113),or by mutagenesis of known promoter sequences (Alper *et al.* (2005) *PNAS*, 12678; Nevoigt *et al.* (2006) *AEM* 72, 5266).

Termination control regions can also be derived from various genes native to the hosts. Optionally, a termination site can be unnecessaryor can be included.

Certain vectors are capable of replicating in a broad range of host bacteria and can be transferred by conjugation. The complete and annotated sequence of pRK404 and three related vectors-pRK437, pRK442, and pRK442(H) are available. These derivatives have proven to be valuable tools for genetic manipulation in Gram-negative bacteria (Scott et al., Plasmid, 50: 74-79, 2003). Several plasmid derivatives of broad-host-range Inc P4 plasmid RSF1010 are also available with promoters that can function in a range of Gram-negative bacteria. Plasmid pAYC36 and pAYC37, have active promoters along with multiple cloning sites to allow for the heterologous gene expression in Gram-negative bacteria.

Chromosomal gene replacement tools are also widely available. For example, a thermosensitive variant of the broad-host-range replicon pWV101 has been modified to construct a plasmid pVE6002 which can be used to effect gene replacement in a range of Gram-positive bacteria (Maguin et al., J. Bacteriol., 174: 5633-5638, 1992). Additionally, in vitro transposomes are available to create random mutations in a variety of genomes from commercial sources such as EPICENTRE®.

2017200439 20 Jan 2017

The expression of a butanol biosynthetic pathway in various microbial hosts is described in more detail below.

Expression of a butanol biosynthetic pathway in *E. coli*

Vectors or cassettes useful for the transformation of *E. coli* are common and commercially available from the companies listed above. For example, the genes of an isobutanol biosynthetic pathway can be isolated from various sources, cloned into a modified pUC19 vector and transformed into *E. coli* NM522.

Expression of a butanol biosynthetic pathway in *Rhodococcus erythropolis*

A series of *E. coli-Rhodococcus* shuttle vectors are available for expression in *R. erythropolis*, including, but not limited to, pRhBR17 and pDA71 (Kostichka et al., Appl. Microbiol. Biotechnol., 62: 61-68, 2003). Additionally, a series of promoters are available for heterologous gene expression in *R. erythropolis* (Nakashima et al., Appl. Environ. Microbiol., 70: 5557-5568, 2004 and Tao et al., Appl. Microbiol. Biotechnol., 68: 346-354, 2005). Targeted gene disruption of chromosomal genes in *R. erythropolis* can be created using the method described by Tao et al., *supra*, and Brans et al. (Appl. Environ. Microbiol., 66: 2029-2036, 2000).

The heterologous genes required for the production of isobutanol, as described above, can be cloned initially in pDA71 or pRhBR71 and transformed into *E. coli*. The vectors can then be transformed into *R. erythropolis* by electroporation, as described by Kostichka et al., *supra*. The recombinants can be grown in synthetic medium containing glucose and the production of isobutanol can be followed using methods known in the art.

Expression of a butanol biosynthetic pathway in *B. subtilis*

Methods for gene expression and creation of mutations in *B. subtilis* are also well known in the art. For example, the genes of an isobutanol biosynthetic pathway can be isolated from various sources, cloned into a modified pUC19 vector and transformed into *Bacillus subtilis* BE1010. Additionally, the five genes of an isobutanol biosynthetic pathway can be split into two operons for expression. The three genes of the pathway (*bubB*, *ilvD*, and *kivD*) can be integrated into the chromosome of *Bacillus*

2017200439 20 Jan 2017

subtilis BE1010 (Payne, et al., J. Bacteriol., 173, 2278-2282, 1991). The remaining two genes (*ilvC* and *bdhB*) can be cloned into an expression vector and transformed into the *Bacillus* strain carrying the integrated isobutanol genes

Expression of a butanol biosynthetic pathway in *B. licheniformis*

Most of the plasmids and shuttle vectors that replicate in *B. subtilis* can be used to transform *B. licheniformis* by either protoplast transformation or electroporation. The genes required for the production of isobutanol can be cloned in plasmids pBE20 or pBE60 derivatives (Nagarajan et al., Gene, 114: 121-126, 1992). Methods to transform *B. licheniformis* are known in the art (Fleming et al. Appl. Environ. Microbiol., 61: 3775-3780, 1995). The plasmids constructed for expression in *B. subtilis* can be transformed into *B. licheniformis* to produce a recombinant microbial host that produces isobutanol.

Expression of a butanol biosynthetic pathway in *Paenibacillus macerans*

Plasmids can be constructed as described above for expression in *B. subtilis* and used to transform *Paenibacillus macerans* by protoplast transformation to produce a recombinant microbial host that produces isobutanol.

Expression of the butanol biosynthetic pathway in *Alcaligenes (Ralstonia) eutrophus*

Methods for gene expression and creation of mutations in *Alcaligenes eutrophus* are known in the art (Taghavi et al., Appl. Environ. Microbiol., 60: 3585-3591, 1994). The genes for an isobutanol biosynthetic pathway can be cloned in any of the broad host range vectors described above, and electroporated to generate recombinants that produce isobutanol. The poly(hydroxybutyrate) pathway in *Alcaligenes* has been described in detail, a variety of genetic techniques to modify the *Alcaligenes eutrophus* genome is known, and those tools can be applied for engineering an isobutanol biosynthetic pathway.

Expression of a butanol biosynthetic pathway in *Pseudomonas putida*

2017200439 20 Jan 2017

Methods for gene expression in *Pseudomonas putida* are known in the art (see for example Ben-Bassat et al., U.S. Patent No. 6,586,229, which is incorporated herein by reference). The butanol pathway genes can be inserted into pPCU18 and this ligated DNA can be electroporated into electrocompetent *Pseudomonas putida* DOT-T1 C5aAR1 cells to generate recombinants that produce isobutanol.

Expression of a butanol biosynthetic pathway in *Saccharomyces cerevisiae*

Methods for gene expression in *Saccharomyces cerevisiae* are known in the art (e.g., Methods in Enzymology, Volume 194, *Guide to* Yeast Genetics and Molecular and Cell Biology, Part A, 2004, Christine Guthrie and Gerald R. Fink, eds., Elsevier Academic Press, San Diego, CA). Expression of genes in yeast typically requires a promoter, followed by the gene of interest, and a transcriptional terminator. A number of yeast promoters, including those used in the Examples herein, can be used in constructing expression cassettes for genes encoding an isobutanol biosynthetic pathway, including, but not limited to constitutive promoters FBA, GPD, ADH1, and GPM, and the inducible promoters GAL1, GAL10, and CUP1. Suitable transcriptional terminators include, but are not limited to FBAt, GPDt, GPMt, ERG10t, GAL1t, CYC1, and ADH1. For example, suitable promoters, transcriptional terminators, and the genes of an isobutanol biosynthetic pathway can be cloned into *E. coli*-yeast shuttle vectors and transformed into yeast cells as described in U.S. App. Pub. No. 20100129886. These vectors allow strain propagation in both *E. coli* and yeast strains. Typically the vector contains a selectable marker and sequences allowing autonomous replication or chromosomal integration in the desired host. Typically used plasmids in yeast are shuttle vectors pRS423, pRS424, pRS425, and pRS426 (American Type Culture Collection, Rockville, MD), which contain an *E. coli* replication origin (e.g., pMB1), a yeast 2μ origin of replication, and a marker for nutritional selection. The selection markers for these four vectors are His3 (vector pRS423), Trp1 (vector pRS424), Leu2 (vector pRS425) and Ura3

2017200439 20 Jan 2017

(vector pRS426). Construction of expression vectors with genes encoding polypeptides of interest can be performed by either standard molecular cloning techniques in *E. coli* or by the gap repair recombination method in yeast.

The gap repair cloning approach takes advantage of the highly efficient homologous recombination in yeast. Typically, a yeast vector DNA is digested (e.g., in its multiple cloning site) to create a "gap" in its sequence. A number of insert DNAs of interest are generated that contain a $\geq$ 21 bp sequence at both the 5' and the 3' ends that sequentially overlap with each other, and with the 5' and 3' terminus of the vector DNA. For example, to construct a yeast expression vector for "Gene X', a yeast promoter and a yeast terminator are selected for the expression cassette. The promoter and terminator are amplified from the yeast genomic DNA, and Gene X is either PCR amplified from its source organism or obtained from a cloning vector comprising Gene X sequence. There is at least a 21 bp overlapping sequence between the 5' end of the linearized vector and the promoter sequence, between the promoter and Gene X, between Gene X and the terminator sequence, and between the terminator and the 3' end of the linearized vector. The "gapped" vector and the insert DNAs are then co-transformed into a yeast strain and plated on the medium containing the appropriate compound mixtures that allow complementation of the nutritional selection markers on the plasmids. The presence of correct insert combinations can be confirmed by PCR mapping using plasmid DNA prepared from the selected cells. The plasmid DNA isolated from yeast (usually low in concentration) can then be transformed into an *E. coli* strain, e.g. *TOP10*, followed by mini preps and restriction mapping to further verify the plasmid construct. Finally the construct can be verified by sequence analysis.

Like the gap repair technique, integration into the yeast genome also takes advantage of the homologous recombination system in yeast. Typically, a cassette containing a coding region plus control elements (promoter and terminator) and auxotrophic marker is PCR-

2017200439 20 Jan 2017

amplified with a high-fidelity DNA polymerase using primers that hybridize to the cassette and contain 40-70 base pairs of sequence homology to the regions 5' and 3' of the genomic area where insertion is desired. The PCR product is then transformed into yeast and plated on medium containing the appropriate compound mixtures that allow selection for the integrated auxotrophic marker. For example, to integrate "Gene X" into chromosomal location "Y", the promoter-coding regionX-terminator construct is PCR amplified from a plasmid DNA construct and joined to an autotrophic marker (such as *URA3*) by either SOE PCR or by common restriction digests and cloning. The full cassette, containing the promoter-coding regionX-terminator-*URA3* region, is PCR amplified with primer sequences that contain 40-70 bp of homology to the regions 5' and 3' of location "Y" on the yeast chromosome. The PCR product is transformed into yeast and selected on growth media lacking uracil. Transformants can be verified either by colony PCR or by direct sequencing of chromosomal DNA.

Expression of a butanol biosynthetic pathway in *Lactobacillus plantarum*

The *Lactobacillus* genus belongs to the *Lactobacillales* family and many plasmids and vectors used in the transformation of *Bacillus subtilis* and *Streptococcus* can be used for *Lactobacillus.* Non-limiting examples of suitable vectors include pAMβ1 and derivatives thereof (Renault et al., Gene 183:175-182, 1996); and (O'Sullivan et al., Gene, 137: 227-231, 1993); pMBB1 and pHW800, a derivative of pMBB1 (Wyckoff et al., Appl. Environ. Microbiol., 62: 1481-1486, 1996); pMG1, a conjugative plasmid (Tanimoto et al., J. Bacteriol., 184: 5800-5804, 2002); pNZ9520 (Kleerebezem et al., Appl. Environ. Microbiol., 63: 4581-4584, 1997); pAM401 (Fujimoto et al., Appl. Environ. Microbiol., 67: 1262-1267, 2001); and pAT392 (Arthur et al., Antimicrob. Agents Chemother., 38: 1899-1903, 1994). Several plasmids from *Lactobacillus plantarum* have also been

2017200439 20 Jan 2017

reported (van Kranenburg R, et al. Appl. Environ. Microbiol., 71: 1223-1230, 2005).

Expression of a butanol biosynthetic pathway in various *Enterococcus* species (*E. faecium, E. gallinarium, and E. faecalis*)

The *Enterococcus* genus belongs to the *Lactobacillales* family and many plasmids and vectors used in the transformation of *Lactobacilli*, *Bacilli* and *Streptococci* species can be used for *Enterococcus* species. Non-limiting examples of suitable vectors include pAMβ1 and derivatives thereof (Renault et al., Gene, 183: 175-182, 1996); and (O'Sullivan et al., Gene, 137: 227-231, 1993); pMBB1 and pHW800, a derivative of pMBB1 (Wyckoff et al. Appl. Environ. Microbiol., 62: 1481-1486, 1996); pMG1, a conjugative plasmid (Tanimoto et al., J. Bacteriol., 184: 5800-5804, 2002); pNZ9520 (Kleerebezem et al., Appl. Environ. Microbiol., 63: 4581-4584, 1997); pAM401 (Fujimoto et al., Appl. Environ. Microbiol., 67: 1262-1267, 2001); and pAT392 (Arthur et al., Antimicrob. Agents Chemother., 38:, 1899-1903, 1994). Expression vectors for *E. faecalis* using the *nisA* gene from *Lactococcus* can also be used (Eichenbaum et al., Appl. Environ. Microbiol., 64: 2763-2769, 1998). Additionally, vectors for gene replacement in the *E. faecium* chromosome can be used (Nallaapareddy et al., Appl. Environ. Microbiol., 72: 334-345, 2006).

Fermentation Media

Fermentation media in the present invention must contain suitable carbon substrates. Suitable substrates can include but are not limited to monosaccharides such as glucose and fructose, oligosaccharides such as lactose, maltose, galactose, sucrose, polysaccharides such as starch or cellulose or mixtures thereof and unpurified mixtures from renewable feedstocks such as cheese whey permeate, cornsteep liquor, sugar beet molasses, and barley malt. Additionally the carbon substrate can also be one-carbon substrates such as carbon dioxide, or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. In addition to one and two carbon substrates methylotrophic microorganisms are also known to utilize a number of other

2017200439 20 Jan 2017

carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylotrophic yeast are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion et al., Microb. Growth C1 Compd., [Int. Symp.], 7th (1993), 415-32. (eds): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Sulter et al., Arch. Microbiol., 153: 485-489, 1990). Hence it is contemplated that the source of carbon utilized in the present invention can encompass a wide variety of carbon containing substrates and will only be limited by the choice of microorganism.

Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention, in some embodiments, the carbon substrates are glucose, fructose, and sucrose, or mixtures of these with C5 sugars such as xylose and/or arabinose for yeasts cells modified to use C5 sugars. Sucrose can be derived from renewable sugar sources such as sugar cane, sugar beets, cassava, sweet sorghum, and mixtures thereof. Glucose and dextrose can be derived from renewable grain sources through saccharification of starch based feedstocks including grains such as corn, wheat, rye, barley, oats, and mixtures thereof. In addition, fermentable sugars can be derived from renewable cellulosic or lignocellulosic biomass through processes of pretreatment and saccharification, as described, for example, in U.S. Patent App. Pub. No. 2007/0031918 A1, which is herein incorporated by reference in its entirety. Biomass refers to any cellulosic or lignocellulosic material and includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides and/or monosaccharides. Biomass can also comprise additional components, such as protein and/or lipid. Biomass can be derived from a single source, or biomass can comprise a mixture derived from more than one source; for example, biomass can comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste,

industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof.

In addition to an appropriate carbon source, fermentation media must contain suitable minerals, salts, cofactors, buffers and other components, known to those skilled in the art, suitable for growth of the cultures and promotion of the enzymatic pathway necessary for butanol production described herein.

Culture Conditions

Typically cells are grown at a temperature in the range of about 20 °C to about 40 °C in an appropriate medium. Suitable growth media in the present invention are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth or Yeast Medium (YM) broth or broth that includes yeast nitrogen base, ammonium sulfate, and dextrose (as the carbon/energy source) or YPD Medium, a blend of peptone, yeast extract, and dextrose in optimal proportions for growing most *Saccharomyces cerevisiae* strains. Other defined or synthetic growth media can also be used, and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, e.g., cyclic adenosine 2',3'-monophosphate (cAMP), can also be incorporated into the fermentation medium.

Suitable pH ranges for the fermentation are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is preferred for the initial condition. Suitable pH ranges for the fermentation of yeast are typically between about pH 3.0 to about pH 9.0. In one embodiment, about pH 5.0 to about

pH 8.0 is used for the initial condition. Suitable pH ranges for the fermentation of other microorganisms are between about pH 3.0 to about pH 7.5. In one embodiment, about pH 4.5 to about pH 6.5 is used for the initial condition.

Fermentations can be performed under aerobic or anaerobic conditions. In one embodiment, anaerobic or microaerobic conditions are used for fermentation.

Industrial Batch and Continuous Fermentations

The present processes may employ a batch method of fermentation. A classical batch fermentation is a closed system where the composition of the medium is set at the beginning of the fermentation and not subject to artificial alterations during the fermentation. Thus, at the beginning of the fermentation the medium is inoculated with the desired microorganism or microorganisms, and fermentation is permitted to occur without adding anything to the system. Typically, however, a "batch" fermentation is batch with respect to the addition of carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems the metabolite and biomass compositions of the system change constantly up to the time the fermentation is stopped. Within batch cultures cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. If untreated, cells in the stationary phase will eventually die. Cells in log phase generally are responsible for the bulk of production of end product or intermediate.

A variation on the standard batch system is the Fed-Batch system. Fed-Batch fermentation processes are also suitable in the present invention and comprise a typical batch system with the exception that the substrate is added in increments as the fermentation progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the medium. Measurement of the actual substrate concentration in Fed-Batch systems is difficult and is therefore estimated

2017200439 20 Jan 2017

on the basis of the changes of measurable factors such as pH, dissolved oxygen and the partial pressure of waste gases such as CO_2. Batch and Fed-Batch fermentations are common and well known in the art and examples can be found in Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition (1989) Sinauer Associates, Inc., Sunderland, MA, or Deshpande, Mukund (Appl. Biochem. Biotechnol., 36: 227, 1992), herein incorporated by reference.

Although the present invention is performed in batch mode it is contemplated that the method would be adaptable to continuous fermentation methods. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned medium is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth.

Continuous fermentation allows for modulation of one factor or any number of factors that affect cell growth or end product concentration. For example, one method will maintain a limiting nutrient such as the carbon source or nitrogen level at a fixed rate and allow all other parameters to moderate. In other systems a number of factors affecting growth can be altered continuously while the cell concentration, measured by medium turbidity, is kept constant. Continuous systems strive to maintain steady state growth conditions and thus the cell loss due to the medium being drawn off must be balanced against the cell growth rate in the fermentation. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, *supra*.

It is contemplated that the present invention can be practiced using either batch, fed-batch or continuous processes and that any known mode of fermentation would be suitable. Additionally, it is contemplated that cells can be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for isobutanol production.

2017200439 20 Jan 2017

Methods for Butanol Isolation from the Fermentation Medium

Bioproduced butanol can be isolated from the fermentation medium using methods known in the art for ABE fermentations (*see, e.g.*, Durre, *Appl. Microbiol. Biotechnol. 49*:639-648 (1998), Groot *et al.*, *Process. Biochem. 27*:61-75 (1992), and references therein). For example, solids can be removed from the fermentation medium by centrifugation, filtration, decantation, or the like. Then, the butanol can be isolated from the fermentation medium using methods such as distillation, azeotropic distillation, liquid-liquid extraction, adsorption, gas stripping, membrane evaporation, or pervaporation.

Because butanol forms a low boiling point, azeotropic mixture with water, distillation can be used to separate the mixture up to its azeotropic composition. Distillation can be used in combination with another separation method to obtain separation around the azeotrope. Methods that can be used in combination with distillation to isolate and purify butanol include, but are not limited to, decantation, liquid-liquid extraction, adsorption, and membrane-based techniques. Additionally, butanol can be isolated using azeotropic distillation using an entrainer (*see, e.g.*, Doherty and Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, New York, 2001).

The butanol-water mixture forms a heterogeneous azeotrope so that distillation can be used in combination with decantation to isolate and purify the butanol. In this method, the butanol containing fermentation broth is distilled to near the azeotropic composition. Then, the azeotropic mixture is condensed, and the butanol is separated from the fermentation medium by decantation. The decanted aqueous phase can be returned to the first distillation column as reflux. The butanol-rich decanted organic phase can be further purified by distillation in a second distillation column.

The butanol can also be isolated from the fermentation medium using liquid-liquid extraction in combination with distillation. In this method, the butanol is extracted from the fermentation broth using liquid-liquid extraction with a suitable solvent. The butanol-containing organic phase is then distilled to separate the butanol from the solvent.

2017200439 20 Jan 2017

Distillation in combination with adsorption can also be used to isolate butanol from the fermentation medium. In this method, the fermentation broth containing the butanol is distilled to near the azeotropic composition and then the remaining water is removed by use of an adsorbent, such as molecular sieves (Aden *et al.*, *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Report NREL/TP-510-32438, National Renewable Energy Laboratory, June 2002).

Additionally, distillation in combination with pervaporation can be used to isolate and purify the butanol from the fermentation medium. In this method, the fermentation broth containing the butanol is distilled to near the azeotropic composition, and then the remaining water is removed by pervaporation through a hydrophilic membrane (Guo *et al.*, *J. Membr. Sci.* 245, 199-210 (2004)).

In situ product removal (ISPR) (also referred to as extractive fermentation) can be used to remove butanol (or other fermentative alcohol) from the fermentation vessel as it is produced, thereby allowing the microorganism to produce butanol at high yields. One method for ISPR for removing fermentative alcohol that has been described in the art is liquid-liquid extraction. In general, with regard to butanol fermentation, for example, the fermentation medium, which includes the microorganism, is contacted with an organic extractant at a time before the butanol concentration reaches a toxic level. The organic extractant and the fermentation medium form a biphasic mixture. The butanol partitions into the organic extractant phase, decreasing the concentration in the aqueous phase containing the microorganism, thereby limiting the exposure of the microorganism to the inhibitory butanol.

Liquid-liquid extraction can be performed, for example, according to the processes described in U.S. Patent Appl. Pub. No. 2009/0305370, the disclosure of which is hereby incorporated in its entirety. U.S. Patent Appl. Pub. No. 2009/0305370 describes methods for producing and recovering butanol from a fermentation broth using liquid-liquid extraction, the

methods comprising the step of contacting the fermentation broth with a water immiscible extractant to form a two-phase mixture comprising an aqueous phase and an organic phase. Typically, the extractant can be an organic extractant selected from the group consisting of saturated, mono-unsaturated, poly-unsaturated (and mixtures thereof) C_{12} to C_{22} fatty alcohols, C_{12} to C_{22} fatty acids, esters of C_{12} to C_{22} fatty acids, C_{12} to C_{22} fatty aldehydes, and mixtures thereof. The extractant(s) for ISPR can be non-alcohol extractants. The ISPR extractant can be an exogenous organic extractant such as oleyl alcohol, behenyl alcohol, cetyl alcohol, lauryl alcohol, myristyl alcohol, stearyl alcohol, 1-undecanol, oleic acid, lauric acid, myristic acid, stearic acid, methyl myristate, methyl oleate, undecanal, lauric aldehyde, 20-methylundecanal, and mixtures thereof.

In some embodiments, the ester can be formed by contacting the alcohol in a fermentation medium with a carboxylic acid (e.g., fatty acids) and a catalyst capable of esterfiying the alcohol with the carboxylic acid, as described in PCT Appn. Pub. No. WO/2011/159998, which is herein incorporated by reference in its entirety. In such embodiments, the carboxylic acid can serve as an ISPR extractant into which the alcohol esters partition. The carboxylic acid can be supplied to the fermentation vessel and/or derived from the biomass supplying fermentable carbon fed to the fermentation vessel. Lipids present in the feedstock can be catalytically hydrolyzed to carboxylic acid, and the same catalyst (e.g., enzymes) can esterify the carboxylic acid with the alcohol. The catalyst can be supplied to the feedstock prior to fermentation, or can be supplied to the fermentation vessel before or contemporaneously with the supplying of the feedstock. When the catalyst is supplied to the fermentation vessel, alcohol esters can be obtained by hydrolysis of the lipids into carboxylic acid and substantially simultaneous esterification of the carboxylic acid with butanol present in the fermentation vessel. Carboxylic acid and/or native oil not derived from the feedstock can also be fed to the fermentation vessel, with the native oil being hydrolyzed into carboxylic acid. Any carboxylic acid not esterified with the alcohol can serve as part of the ISPR extractant. The extractant containing alcohol esters can be

2017200439 20 Jan 2017

separated from the fermentation medium, and the alcohol can be recovered from the extractant. The extractant can be recycled to the fermentation vessel. Thus, in the case of butanol production, for example, the conversion of the butanol to an ester may reduce the free butanol concentration in the fermentation medium, shielding the microorganism from the toxic effect of increasing butanol concentration. In addition, unfractionated grain can be used as feedstock without separation of lipids therein, since the lipids can be catalytically hydrolyzed to carboxylic acid, thereby decreasing the rate of build-up of lipids in the ISPR extractant.

In situ product removal can be carried out in a batch mode or a continuous mode. In a continuous mode of *in situ* product removal, product is continually removed from the reactor. In a batchwise mode of *in situ* product removal, a volume of organic extractant is added to the fermentation vessel and the extractant is not removed during the process. For *in situ* product removal, the organic extractant can contact the fermentation medium at the start of the fermentation forming a biphasic fermentation medium. Alternatively, the organic extractant can contact the fermentation medium after the microorganism has achieved a desired amount of growth, which can be determined by measuring the optical density of the culture. Further, the organic extractant can contact the fermentation medium at a time at which the product alcohol level in the fermentation medium reaches a preselected level. In the case of butanol production according to some embodiments of the present invention, the carboxylic acid extractant can contact the fermentation medium at a time before the butanol concentration reaches a toxic level, so as to esterify the butanol with the carboxylic acid to produce butanol esters and consequently reduce the concentration of butanol in the fermentation vessel. The ester-containing organic phase can then be removed from the fermentation vessel (and separated from the fermentation broth which constitutes the aqueous phase) after a desired effective titer of the butanol esters is achieved. In some embodiments, the ester-containing organic phase is separated from the aqueous phase after fermentation of the

2017200439 20 Jan 2017

available fermentable sugar in the fermentation vessel is substantially complete.

EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various uses and conditions.

GENERAL METHODS:

Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989, by T. J. Silhavy, M. L. Bennan, and L. W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1984, and by Ausubel, F. M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience, N.Y., 1987.

Materials and methods suitable for the maintenance and growth of bacterial cultures are also well known in the art. Techniques suitable for use in the following Examples can be found in Manual of Methods for General Bacteriology, Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds., American Society for Microbiology, Washington, DC., 1994, or by Thomas D. Brock in Biotechnology: A Textbook of Industrial Microbiology, Second Edition, Sinauer Associates, Inc., Sunderland, MA, 1989. All reagents, restriction enzymes and materials used for the growth and maintenance of bacterial cells were obtained from Aldrich Chemicals

2017200439 20 Jan 2017

(Milwaukee, WI), BD Diagnostic Systems (Sparks, MD), Life Technologies (Rockville, MD), or Sigma Chemical Company (St. Louis, MO), unless otherwise specified.

The meaning of abbreviations used is as follows: "Å" means Angstrom, "min" means minute(s), "h" means hour(s), "µl" means microliter(s), "ng/µl" means nano gram per microliter, "pmol/µl" means pico mole per microliter, "ml" means milliliter(s), "L" means liter(s), "g/L" mean gram per liter, "ng" means nano gram, "sec" means second(s), "ml/min" means milliliter per minute(s), "w/v" means weight per volume, "v/v" means volume per volume, "nm" means nanometer(s), "mm" means millimeter(s), "cm" means centimeter(s), "mM" means millimolar, "M" means molar, "g" means gram(s), "µg" means microgram(s), "mg" means milligram(s), "*g*" means the gravitation constant, "rpm" means revolutions per minute, "HPLC" means high performance liquid chromatography, "MS" means mass spectrometry, "HPLC/MS" means high performance liquid chromatography/mass spectrometry, "EDTA" means ethylendiamine-tetraacetic acid, "dNTP" means deoxynucleotide triphosphate, "^{o}C" means degrees Celsius, and "V" means voltage.

<u>High throughput screening assay of gene libraries</u>

High throughput screening of the gene libraries of mutant KARI enzymes was performed as described herein (with the exception of Examples 16 and 21): 10× freezing medium containing 554.4 g/L glycerol, 68 mM of $(NH_4)_2SO_4$, 4 mM $MgSO_4$, 17 mM sodium citrate, 132 mM KH_2PO_4, 36 mM K_2HPO_4 was prepared with molecular pure water and filter-sterilized. Freezing medium was prepared by diluting the 10x freezing medium with the LB medium. An aliquot (200 µl) of the freezing medium was used for each well of the 96-well archive plates (cat #3370, Corning Inc. Corning, NY).

Clones from the LB agar plates were selected and inoculated into the 96-well archive plates containing the freezing medium and grown overnight at 37 ^{o}C without shaking. The archive plates were then stored at -80^{o}C. *E. coli* strain Bw25113 transformed with *pBAD*-HisB (Invitrogen) was always used as the negative control. The positive controls for the

2017200439 20 Jan 2017

libraries in Examples 3, 4, and 5 are the wild type K9-KARI, AB1D3, AB1D3 respectively.

Clones from archive plates were inoculated into the 96-deep well plates. Each well contained 3.0 μl of cells from thawed archive plates, 200 μl of the LB medium containing 100 μg/ml ampicillin and 0.02%(w/v) arabinose as the inducer. Cells were the grown overnight at 37 °C with 80% humidity while shaking (900rpm), harvested by centrifugation (4000 rpm, 5 min at 25 °C). (Eppendorf centrifuge, Brinkmann Instruments, Inc. Westbury, NY) and the cell pellet was stored at -20 °C for later analysis. The assay substrate, (*R*,*S*)-acetolactate, was synthesized as described by Aulabaugh and Schloss (Aulabaugh and Schloss, Biochemistry, 29: 2824-2830, 1990). All other chemicals used in the assay were purchased from Sigma.

The enzymatic conversion of acetolactate to α,β-dihydroxy-isovalerate by KARI was followed by measuring the disappearance of the cofactor, NADPH or NADH, from the reaction at 340 nm using a plate reader (Molecular Device, Sunnyvale, CA). The activity was calculated using the molar extinction coefficient of 6220 $M^{-1}cm^{-1}$ for either NADPH or NADH. The stock solutions used were: K_2HPO_4 (0.2 M); KH_2PO_4 (0.2 M); EDTA (0.5 M); $MgCl_2$ (1.0 M); NADPH (2.0 mM); NADH (2.0 mM) and acetolactate (45 mM). The 100 ml reaction buffer (pH 6.8) containing: 2.0 ml K_2HPO_4, 3.0 ml KH_2PO_4, 4.0 ml $MgCl_2$, 0.1 ml EDTA and 90.9 ml water was prepared.

Frozen cell pellet in deep-well plates and BugBuster were warmed up at room temperature for 30 min at the same time. Each well of 96-well assay plates was filled with 120 μl of the reaction buffer and 20 μl of NADH (2.0 mM). 75 μl of 50% BugBuster (v/v in water) was added to each well after 30 min warm-up and cells were suspended using plate shaker. The plates were incubated at room temperature for 20 min. An aliquot (15 to 25 μl depending the expected activity) of cell lysate was transferred into each well of 96-well assay plates. Absorbance at 340 nm was recorded as background, 16 μl of acetolactate (4.5 mM, diluted with the reaction buffer) was added to each well and mixed with shaking by the plate reader.

2017200439 20 Jan 2017

Absorbance at 340 nm was recorded at 0, and 10 to 30 minutes depending the expected activity after substrate addition. The difference in absorbance (before and after substrate addition) was used to determine the activity of the mutants. Mutants with higher KARI activity compared to the positive control were selected for re-screening.

The number of clones screened for the libraries in Example1, 2 and 3 are about 12,000, 12,000 and 92 respectively. The top performers from each library were re-screened described below as secondary assay.

Secondary assay of active mutants

Cells containing selected mutants identified by high throughput screening (above) were grown overnight, at 37°C, in 3.0 ml of the LB medium containing 100 μg/ml ampicillin and 0.025%(w/v) arabinose as the inducer while shaking at 250 rpm. The cells were then aliquoted into 96 deepwell plates (200 μl per well) and harvested by centrifugation at 4,000 x*g* for 5 min at room temperature. 75 μl of 50% BugBuster (v/v in water) was added to each well and cells were suspended using plate shaker. The plates were incubated at room temperature for 20 min. An aliquot (15 to 25 μl depending the expected activity) of cell lysate was transferred into each well of 96-well assay plates, which contain 120 μl of the reaction buffer and 20 μl of NADH (2.0 mM) per well. Absorbance at 340 nm was recorded as background, 16 μl of acetolactate (4.5 mM, diluted with the reaction buffer) was added to each well and mixed with shaking by the plate reader. Absorbance at 340 nm was recorded at 0, and 5 to 10 minutes depending the expected activity after substrate addition. The difference in absorbance (before and after substrate addition) was used to determine the activity of the mutants. Mutants with higher KARI activity compared to the positive control were selected for further characterization.

Measurement of NADH and NADPH Michaelis Constants

KARI enzyme activity can be routinely measured by NADH or NADPH oxidation as described above, however to measure the Michaelis constant (K_M) for these pyridine nucleotides formation of the 2,3-dihydroxyisovalerate product was measured directly using HPLC/MS.

2017200439 20 Jan 2017

Protein concentration of crude cell extract from Bugbuster lysed cells (as described above) was measured using the BioRad protein assay reagent (BioRad Laboratories, Inc., Hercules, CA 94547). Between 0.2 and 1.0 micrograms of crude extract protein was added to a reaction buffer consisting of 100 mM MOPS KOH, pH 6.8, 10 mM $MgCl_2$, 1mM EDTA, 1 mM glucose-6-phosphate (Sigma-Aldrich), 0.2 Units of *Leuconostoc mesenteroides* glucose-6-phosphate dehydrogenase (Sigma-Aldrich), and various concentrations of NADH or NADPH, to a volume of 90 μL. The reaction was initiated by the addition of 10 μL of *[S]*-acetolactate to a final concentration of 2.5 mM and a final volume of 100 μL. After 10 min incubations at 30°C, the reaction was quenched by withdrawing 50 μL of the reaction mixture and adding it to 150 μL of 0.1 % formic acid. To measure the K_M of NADH and NADPH, the concentrations used were 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3 and 1 mM.

To analyze for 2,3-dihydroxyisovalerate, 2 μL of the formic acid quenched reaction mixture was injected into a Waters Acquity HPLC equipped with Waters SQD mass spectrometer (Waters Corporation, Milford , MA). The chromatography conditions were: flow rate (0.5 ml/min), on a Waters Acquity HSS T3 column (2.1 mm diameter, 100 mm length). Buffer A consisted of 0.1% (v/v) in water, Buffer B was 0.1% formic acid in acetonitrile. The sample was analyzed using 1% buffer B (in buffer A) for 1 min, followed by a linear gradient from 1% buffer B at 1 min to 75% buffer B at 1.5 min. The reaction product, 2,3-dihydroxyisovalerate, was detected by ionization at m/z=133, using electrospay ionization -30 V cone voltage. The amount of product 2,3-dihydroxyisovalerate was calculated by comparison to an authentic standard.

To calculate the K_M for NADH and NADPH, the rate data for DHIV formation measured in assays at a fixed concentration of S-acetolactate (2.5 mM) was fitted to the single substrate Michaelis-Menten equation, using a least-squares regression in Microsoft Excel, assuming saturating acetolactate concentration.

2017200439 20 Jan 2017

Construction of plasmids pYZ058, pLH550, pLH556, and pLH702

pYZ058 (pHR81-P_{CUP1}-AlsS-P_{ILV5}-yeast KARI; SEQ ID NO: 176) was derived from pYZ090 (pHR81-P_{CUP1}-AlsS-P_{ILV5}-lactis KARI; SEQ ID NO: 195). pYZ090 was cut with PmeI and SfiI enzymes, and ligated with a PCR product of yeast KARI. The PCR product was amplified from genomic DNA of *Saccharomyces* cerevisiae BY4741 (Research Genetics Inc.) strain using upper primer 5'-catcatcacagtttaaacagtatgttgaagcaaatcaacttcggtgg-3' (SEQ ID NO: 272) and lower primer 5'- ggacgggccctgcaggccttattggttttctggtctcaactttctgac-3' (SEQ ID NO: 273), and digested with PmeI and SfiI enzymes. pYZ058 was confirmed by sequencing.

pLH550 (pHR81-PCUP1-AlsS-PILV5-Pf5.KARI, SEQ ID NO: 175) was derived from pYZ058 (SEQ ID NO: 176). The wild type Pf5.KARI gene was PCR amplified with OT1349 (5'-catcatcacagtttaaacagtatgaaagttttctacgataaagactgcgacc-3'; SEQ ID NO: 177) and OT1318 (5'-gcacttgataggcctgcagggccttagttcttggctttgtcgacgattttg-3'; SEQ ID NO: 178), digested with PmeI and SfiI enzymes and ligated with pYZ058 vector cut with PmeI and SfiI. The vector generated, pLH550, was confirmed by sequencing.

pLH556 (SEQ ID NO: 138; FIGURE 4) was derived from pLH550 by digesting the vector with SpeI and NotI enzymes, and ligating with a linker annealed from OT1383 (5'-ctagtcaccggtggc-3', SEQ ID NO: 179) and OT1384 (5'-ggccgccaccggtga-3', SEQ ID NO: 180) which contains overhang sequences for SpeI and NotI sites. This cloning step eliminates the AlsS gene and a large fragment of the PCUP1 promoter, with 160 bp residual upstream sequence that is not functional. pLH556 was confirmed by sequencing.

pHR81::ILV5p-K9D3 (pLH702, SEQ ID NO: 181) was derived from pLH556. The K9D3 mutant KARI gene was excised from vector pBAD-K9D3 using PmeI and SfiI enzymes, and ligated with pLH556 at PmeI and SfiI sites, replacing the Pf5.KARI gene with the K9D3 gene. The constructed vector was confirmed by sequencing.

Example 1

2017200439 20 Jan 2017

Construction of yeast isobutanol pathway strains containing various KARI genes

To identify polypeptides having KARI activity and performance in yeast isobutanol production, biodiversity screening of KARI-encoding genes from various bacterial and fungal species was carried out. The KARI genes were codon optimized based on codon preferences of *Saccharomyces cerevisiae* genes where indicated in Table 10. For each KARI gene, a PmeI restriction site and additional 3 bp (AGT) was added to the 5' end with the sequence 5'-GTTTAAACAGT-3' (SEQ ID NO: 136) before the ATG start codon, and a SfiI restriction site was added to the 3' end with the sequence 5'-GGCCCTGCAGGCC-3' (SEQ ID NO: 137). All of the KARI genes were synthesized by GenScript USA Inc. (Piscataway, NJ). Each KARI gene was subcloned into pHR81-P_{CUP1}-AlsS-P_{Ilv5}-Pf5.Ilv5 vector (SEQ ID NO: 175) via the PmeI and SfiI sites (Ilv5 encodes for yeast ketol-acid reductoisomerase). This vector contains two expression cassettes: *Bacillus subtilis* acetolactate synthase (AlsS) gene under the yeast CUP1 promoter, and yeast Ilv5 gene controlled by the Ilv5 promoter. Sequence analysis was performed to confirm the KARI gene sequences.

The pHR81-PCUP1-AlsS-PI_{lv5}-KARI vectors carrying the KARI genes were co-transformed with pLH468 (pRS423-P_{FBA1}-DHAD- P_{TDH3}-kivD-P_{GPM1}-hADH1; SEQ ID NO: 139) into host strain BP1135 (PNY1505; Example 8) (CEN.pk 113-7D delta ura3::loxP delta his3 delta pdc6 delta pdc1::ilvD.Sm delta pdc5::sadB delta gpd2::loxP delta fra2). The yeast transformants were selected on minimum drop-out media plates (SE-Ura-His, 2% ethanol) after 5-7 days at 30°C, and restreaked on SE-Ura-His to obtain cell patches after additional 3 day incubation. The cell patches were used for shake flask inoculation.

Example 2

Screening the KARI diversity collection for isobutanol production

The various KARI genes were evaluated based on their "effective productivities" in yeast. The effective productivity was determined after a

2017200439 20 Jan 2017

certain period of growth under progressively oxygen-limited conditions (e.g. 48 h). The yeast biomass was calculated with the assumption that 1 OD_{600} of yeast cells is equivalent to 0.3 g/L.

The yeast isobutanol pathway strains carrying various KARI genes were inoculated into 10 mL SEG-Ura,His media with 0.2% glucose and 0.2% ethanol, and grown aerobically overnight at 30°C, to about 2 OD. The cultures were centrifuged and a portion of the cells were resuspended in SEG-Ura,His (2% glucose, 1% ethanol) to an initial OD_{600} of 0.4 in 25 mL total volume in a 125 mL shake flask. The shake flasks were closed with a screw-on solid plastid cap, and the cultures were grown under progressively oxygen-limited conditions in the flask under minimal air and oxygen exchange with the outside environment. After 48 h incubation at 30°C, 250 RPM, the cultures were removed for OD_{600} measurement and HPLC analysis to measure isobutanol production.

From the KARI genes screened, as shown below, multiple had comparable or better isobutanol titers than *Lactococcus lactis* KARI. In particular, the K9 (*Anaerostipes caccae* DSM 14662) KARI clone showed a high isobutanol titer and effective isobutanol productivity, as measured after 48 h of growth under progressively oxygen-limited conditions (Table 10).

Table 10. Isobutanol titers and effective productivities from yeast isobutanol production strains carrying various KARI genes measured after 48 h of growth under progressively oxygen-limited conditions in shake flasks at 30°C.

2017200439 20 Jan 2017

KARI clone	SEQ ID NO: (nucleic acid, amino acid) *All nucleic acid seqs except LTS and S2 are codon-optimized	Isobutanol titer (g/L)	Effective Isobutanol Productivity (g/g)	Source Organism
B3K01 ("K1")	140, 141	2.6	4.1	*Bifidobacterium angulatum* DSM 20098
B3K02 ("K2")	142, 143	3.5	3.7	*Bifidobacterium dentium* ATCC 27678
B3K09 ("K9")	26, 27	4.3	5.2	*Anaerostipes caccae* DSM 14662
B3K25 ("K25")	375, 376	3.6	4.4	*Enterococcus gallinarum* EG2
B3K26 ("K26")	381, 382	4.4	3.2	*Streptococcus thermophilus* LMD-9
B3K29 ("K29")	377, 378	4.1	3.3	*Lactococcus lactis* subsp. cremoris MG1363
LTS	379, 380	2.7	3.1	*Lactococcus lactis*
B3K07 ("K7")	274, 275	3.7	2.8	*Clostridium beijerinckii* NCIMB 8052
S2	276, 277	3.6	1.5	*Zymomonas mobilis*

Example 3

KARI enzyme analysis of the yeast isobutanol pathway strains

2017200439 20 Jan 2017

IpOHA (N-isopropyl oxalylhydroxamic acid) is a mimic of a reaction intermediate for the reaction catalyzed by the KARI enzyme. It is a tight binding inhibitor that binds to the active site of the KARI enzyme. The synthesis of IpOHA and its tight binding to KARI from *E. coli* is described in literature (A. Aulabaugh and J. V. Schloss, Biochemistry, 1990, 29, 2824-2830). Its use for active site titration has not been reported before. IpOHA was synthesized from [^{14}C]-oxalate according to literature.

The yeast cultures from Example 2 were harvested and analyzed for KARI enzyme activities. 25 mL of the cultures was pelleted and resuspended in 10 mL of 50 mM Tris-HCl, pH 7.5. The cells were centrifuged again to remove the buffer and the cell pellets are stored at -70C. The cell pellets were resuspended in 1 mL of 50 mM Tris-HCl pH 7.5 and sonicated. Soluble crude cell extracts were used to perform the enzyme assays. A portion of enzyme was incubated with a molar excess of [^{14}C]-IpOHA, and saturating concentrations of NAD(P)H and Mg^{2+}. Because a reversible, dilution-sensitive complex forms first, extract concentrations were kept high, to favor complexation and thus reduce the time taken for tight complex formation. Because it was not known *a priori* how long it would take each KARI to form the tight complex, two time points were taken for each sample to verify that the results agree. At the end of the incubation time, small molecules were separated from protein molecules by ultrafiltration using Microcon® (Millipore Inc., Billerica, MA), and the high molecular weight fraction was counted. The concentration of KARI in the sample in either μM or mg/ml was back-calculated from the ^{14}C dpm, the volumes, and the KARI subunit molecular weight. A fixed-time enzyme assay was run concurrently, and the data were used to calculate U/ml. The specific activity was calculated by dividing U/ml by mg/ml for a given sample. The assumption made was that full activity and the ability to bind IpOHA are strictly correlated. The specific activities of the KARI enzymes thus measured are listed in Table 11. The KARI activity in "Units per mg" represents the activity per milligram of KARI enzyme as quantitated using the IpOHA assay. Total protein concentration was determined by the Bradford method, and the

2017200439 20 Jan 2017

expression level of KARI is calculated by dividing KARI enzyme amount by the amount of total soluble cellular proteins.

Table 11. KARI enzyme activities measured by the IPOHA assay

KARI clone	KARI activity U/mg	KARI % total protein	Organism
B3K01	0.56	21	*Bifidobacterium angulatum* DSM 20098
B3K02	0.44	28	*Bifidobacterium dentium* ATCC 27678
B3K09	2.4	15	*Anaerostipes caccae* DSM 14662
B3K25	1.3	21	*Enterococcus gallinarum* EG2
B3K26	1.5	17	*Streptococcus thermophilus* LMD-9
B3K29	1.6	14	*Lactococcus lactis subsp. cremoris* MG1363
LTS	0.8	23.0	*Lactococcus lactis*

Example 4

Construction of a site-saturation gene library to identify variants utilizing NADH with K_M lower than wild type

To construct the pBAD-based bacterial expression vector for K9 KARI, the K9 KARI gene (synthesized by Genscript, Piscataway, NJ) was subcloned into pBAD-ps-JEA1 vector (SEQ ID NO: 905) via the PmeI and SfiI sites. The ketol-acid reductoIsomerase (KARI) from *Anaerostipes caccae* (called K9-KARI) was used for the library construction. One gene library was constructed using the commercially available kits, T4 polynucleotide kinase (PNK) (USB Corporation, Cleveland, Ohio, #70031Z) and Chang_IT Multiple Mutation Site Directed Mutagenesis Kit (USB Corporation, Cleveland, Ohio, #78480).

The oligonucleotides (K9_56_58_060210f: GAAGGANNKAAANNKTGGAAGAGAGC, SEQ ID NO: 144; and K9_56_58_060210r: GCTCTCTTCCAMNNTTTMNNTCCTTC, SEQ ID NO: 145) were synthesized by Integrated DNA Technologies, Inc (Coralville IA). They were first phosphorylated by T4 PNK. In brief, A 30 μl reaction mixture contained: 3.0 μl of 10x T4 PNK buffer supplied with the

2017200439 20 Jan 2017

kit, 4.0 μl of primer (about 35 μM) , 0.8 μl of 100 mM ATP mix, 0.6 μl T4 PNk and 22 μl of water. The reaction mixture was incubated at 37 °C for 1.0 hr and T4 PNK was then deactivated at 65 °C for 20 min.

The phosphorylated primers were then directly used for the subsequent PCR reaction to introduce the mutations at two sites into K9 KARI wild type using the kit. In brief, a 30 μl reaction mixture contained: 3.0 μl of 10x reaction buffer supplied with the kit, 3.0 μl of phosphorylated forward primer and reverse primer, 2.0 μl of K9 KARI wild type (50 ng/μl), 1.2 μl Chang_IT enzyme and 17.8 μl of water. This reaction mixture was placed into thin well 200 μl-capacity PCR tubes and the following PCR reaction program were used for the PCR: The starting temperature was 95°C for 2 min followed by 30 heating/cooling cycles. Each cycle consisted of 95°C for 30 sec, 55°C for 30 sec, and 68°C for 20 min. At the completion of the temperature cycling, the samples were kept at 68° C for 25 min more, and then held at 4°C for later processing. The PCR reaction was cleaned up using the Zymo DNA clean-up kit (Zymo Research Corporation, Orange CA, #D4004). DNA was eluted out of membrane using 84 μl of water. The DNA template was removed with Dpn I (Promega, Madison WI, #R6231) at 37 °C for 3 hr (reaction mixture: 10 μl of 10x reaction buffer, 1.0 μl BSA, 6.0 μl of Dpn I and 83 μl cleaned PCR DNA). The Dpn I digested DNA was cleaned up again with Zymo DNA clean-up kit and digested again with Dpn I to completely remove the DNA template (reaction mixture: 1.5 μl of 10x reaction buffer, 0.15 μl BSA, 0.85 μl of Dpn I and 83 μl cleaned PCR DNA). The reaction mixture was directly used to transform an electro-competent strain of *E. coli* Bw25113(*ΔilvC*) (described in U.S. Patent No. 8,129,162, which is herein incorporated by reference in its entirety) using a BioRad Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). The transformed clones were streaked on agar plates containing the LB medium and 100 μg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA) and incubated at 37 °C overnight. Clones were screened for activity using NADH. K_M for the variants was measured (Table 12).

2017200439 20 Jan 2017

Table 12

Kinetic Values for K9 KARI Variants in *E. coli* Extracts, as Determined via DHIV Formation Assays

Mutant	SEQ ID NO: (nucleic acid, amino acid)	Mutations	K_M (μM) (NADH)	K_M (μM) (NADPH)
K9 Wt	26, 27		326	0.2
AB1D1	28, 29	S56A	164	1
495B5	30, 31	S56A/S58H	44	4
AB1D3 (also referred to as "K9D3")	32, 33	S56A/S58D	38	9
AB1G9 (also referred to as "K9G9")	34, 35	S56AS58E	47	23

EXAMPLE 5

Construction of site saturation gene libraries to lower K_M for NADH

Based on work with *Pseudomonas fluorescens* KARI (PF5-KARI) positions 24, 33, 61, 80, 156 and 170 were targeted as mutagenesis targets for K9 KARI. Through multiple sequence alignment (MSA) between PF5-KARI and K9 KARI (Fig. 2), the corresponding positions are 30, 39, 67, 86, 162, and 176.

To identify more mutagenesis targets, MSA of existing KARI enzymes (K1, K2, K7, K9, K25, K26, *L. Lactis* and S2), determined to produce isobutanol in a butanologen strain (see other examples) was used to identify more mutagenesis targets. Positions 41, 87, 131, 191, 227, and 246 were selected as mutagenesis targets.

The oligonucleotides targeting positions 30, 39, 41, 67, 86, 87, 131, 162, 176, 191, 227, and 246 were commercially synthesized by Integrated DNA Technologies, Inc (Coralville IA) (Table 13). Eight pairs of oligonucleotides targeting positions 30, 67, 131, 162, 176, 191, 227, and 246 were used to generate Megaprimers using Supermix from Invitrogen (Cat#10572-014, Invitrogen, Carlsbad, CA). For each PCR reaction, a pair

2017200439 20 Jan 2017

of primers, any combination of one forward primer and one reverse primer encoding different positions from those eight pairs of oligonucleotides (e.g. K9_30_101110f and K9_67_101110r), were used. There are total P_8^2 or 56 combinations. A 25 μl reaction mixture contained: 22.5 μl of Supermix solution, 1.0 μl of forward primer and 1.0 μl of reverse primer, 0.5 μl of AB1D3 DNA template (50 ng/μl). The mixture was placed in a thin well 200 μl tube for the PCR reaction in a Mastercycler gradient equipment (Brinkmann Instruments, Inc. Westbury, NY). The following conditions were used for the PCR reaction: The starting temperature was 95°C for 1.0 min followed by 35 heating/cooling cycles. Each cycle consisted of 95°C for 20 sec, 55°C for 20 sec, and 72°C for 1.0 min. At the completion of the temperature cycling, the samples were kept at 72° C for 2.0 min more, and then held awaiting sample recovery at 4°C. The PCR product was cleaned up using a DNA cleaning kit (Cat#D4003, Zymo Research, Orange, CA) as recommended by the manufacturer.

The Megaprimers were then used to generate gene libraries using the QuickChange II XL site directed mutagenesis kit (Catalog # 200524, Stratagene, La Jolla CA). A 25 μl reaction mixture contained: 2.5 μl of 10x reaction buffer, 1.0 μl of 50 ng/μl template, 20.5 μl of Megaprimer, 0.5 μl of 40 mM dNTP mix, 0.5 μl pfu-ultra DNA polymerase. Except for the Megaprimer and the templates, all reagents used here were supplied with the kit indicated above. This reaction mixture was placed in a thin well 200 μl-capacity PCR tube and the following reactions were used for the PCR: The starting temperature was 95°C for 30 sec followed by 25 heating/cooling cycles. Each cycle consisted of 95°C for 30 sec, 55°C for 1 min, and 68°C for 6 min. At the completion of the temperature cycling, the samples were kept at 68° C for 8 min more, and then held at 4°C for later processing. The PCR reaction mixture was processed with Dpn I restriction enzyme same as that used in Example 4.

The oligonucleotides K9_37&39_101110f, K9_37&39_101110r and K9_86&87_101110f, K9_86&87_101110r were directly then used to generate gene libraries using the QuickChange II XL site directed

2017200439 20 Jan 2017

mutagenesis kit. Two 25 μl reaction mixtures for the two oligonucleotide sets. each 25 μl reaction mixture contained: 2.5 μl of 10x reaction buffer, 1.0 μl of 50 ng/μl template, 1.0 μl of forward primer, 1.0 μl reverse primer, 0.5 μl of 40 mM dNTP mix, 0.5 μl pfu-ultra DNA polymerase and 18.5 μl of water. The PCR program and the subsequent Dpn I processing are the same.

The Dpn I processed DNA mixture was cleaned up using Zymo DNA clean-up kit following the manufacturer's protocol. The cleaned-up DNA was used to transform an electro-competent strain of *E. coli* Bw25113 (*ΔilvC*) using a BioRad Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). The transformed clones were streaked on agar plates containing the LB medium and 100 μg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA) and incubated at 37 °C overnight. Clones were screened for improved activity using NADH. K_M for the improved mutants was measured (Table 14).

Table 13: Primers

Targeted position(s) of *K9-KARI*	Primers
30	K9_30_101110f: gactatcgccgttatcggtNNKggttctcaaggtcac SEQ ID NO: 146 K9_30_101110r: GTGACCTTGAGAACCMNNACCGATAACGGCGATAGTC SEQ ID NO: 147
67	K9_67_101110f: gagctgaagaacaaggtNNKgaagtctacaccgctgc SEQ ID NO: 148 K9_67_101110r: GCAGCGGTGTAGACTTCMNNACCTTGTTCTTCAGCTC SEQ ID NO: 149
131	K9_131_101110f: caaaggacgttgatgtcNNKatgatcgctccaaag SEQ ID NO: 150 K9_131_101110r: CTTTGGAGCGATCATMNNGACATCAACGTCCTTT SEQ ID NO: 151
162	K9_162_101110f: gctgtcgaacaagacNNKactggcaaggctttg SEQ ID NO: 152 K9_162_101110r: CAAAGCCTTGCCAGTMNNGTCTTGTTCGACAGC SEQ ID NO: 153
176	K9_176_101110f: gctttggcctacgctttaNNKatcggtggtgctagagc SEQ ID NO: 154 K9_176_101110r:GCTCTAGCACCACCGATMNNTAAAGCGTAGGCCAAAGC SEQ ID NO: 155
191	K9_191_101110f: gaaactaccttcagaNNKgaaactgaaaccgac SEQ ID NO: 156 K9_191_101110r: GTCGGTTTCAGTTTCMNNTCTGAAGGTAGTTTC SEQ ID NO: 157
227	K9_227_101110f: gccggttacgacccaNNKaacgcttacttcgaatg SEQ ID NO: 158 K9_227_101110r: CATTCGAAGTAAGCGTTMNNTGGGTCGTAACCGGC SEQ ID NO: 159
246	K9_246_101110f: gttgacttgatctacNNKtctggtttctccggtatgc SEQ ID NO: 160 K9_246_101110r: GCATACCGGAGAAACCAGAMNNGTAGATCAAGTCAAC SEQ ID NO: 161

20 Jan 2017

2017200439

39, 41	K9_37&39_101110f: gttctcaaggtcacgctNNKgccNNKaatgctaaggaatcc SEQ ID NO: 162 K9_37&39_101110r: GGATTCCTTAGCATTMNNGGCMNNAGCGTGACCTTGAGAAC SEQ ID NO: 163
86, 87	K9_86&87_101110f: gacatcattatgatcttgNNKNNKgatgaaaagcaggc SEQ ID NO: 164 K9_86&87_101110r: GCCTGCTTTTCATCMNNMNNCAAGATCATAATGATGTC SEQ ID NO: 165

Table 14

List of some mutants with their measured K_M values

Mutant	SEQ ID NO: (nucleic acid, amino acid)	Mutations	K_M (μM) (NADH)	K_M (μM) (NADPH)
AO7A9	38, 39	S56A/S58D/I86T/N87P	15	7
AO7B5	36, 37	S56A/S58D/I86V/N87P	8	4
AO7H8	40, 41	S56A/S58D/N87P	8	6
AO7D8	42, 43	S56A/S58D/T131C/T191S	26	6
AO7F7	44, 45	S56A/S58D/T131V/T191A	28	7
AO7H7	46, 47	S56A/S58D/T191S	29	8

EXAMPLE 6

Construction of a combinatorial library to lower K_M for NADH

Based on the mutagenesis results (Example 4), T131L, T131A, T131V, T131M, T131C, T191D, T191C, T191S, and T191G are considered as beneficial mutations to improve K_M for NADH. A combinatorial library to introduce these beneficial mutations into AO7B5 was made.

All oligonucleotdies were synthesized by the Integrated DNA Technologies, Inc (Coralville IA). They were first phosphorylated by T4 PNK. In brief, a 20 μl reaction mixture contained: 2.0 μl of 10x T4 PNK buffer supplied with the kit, 2.85 μl of primer (about 35 μM , 0.6 μl of 100 mM ATP mix, 0.4 μl T4 PNK and 14.15 μl of water. The reaction mixture was incubated at 37 ^{o}C for 1.0 hr and T4 PNK was then deactivated at 65 ^{o}C for 20 min.

2017200439 20 Jan 2017

The phosphorylated primers were then directly used for the subsequent PCR reaction to introduce the mutations at two sites into AO7B5 using the kit. In brief, a 50 μl reaction mixture contained: 5.0 μl of 10x reaction buffer supplied with the kit, 2.5 μl of phosphorylated forward primer (0.5 μl of each forward primer shown in Table 15), 2.5 μl reverse primer (0.625 μl of each forward primer shown at Table 15), 2.5 μl of AO7B5 (50 ng/μl), 2.5 μl Chang_IT enzyme and 35 μl of water. This reaction mixture was placed into thin well 200 μl-capacity PCR tubes and the following PCR reaction program were used for the PCR: The starting temperature was 95°C for 2 min followed by 30 heating/cooling cycles. Each cycle consisted of 95°C for 30 sec, 55°C for 30 sec, and 68°C for 20 min. At the completion of the temperature cycling, the samples were kept at 68° C for 25 min more, and then held at 4°C for later processing. The PCR reaction was cleaned up using the Zymo DNA clean-up kit (Zymo Research Corporation, Orange CA, #D4004). DNA was eluted out of membrane using 84 μl of water. The DNA template was removed with Dpn I (Promega, Madison WI, #R6231) at 37 °C for 3 hr (reaction mixture: 10 μl of 10x reaction buffer, 1.0 μl BSA, 6.0 μl of Dpn I and 83 μl cleaned PCR DNA). The Dpn I digested DNA was cleaned up again with Zymo DNA clean-up kit and digested again with Dpn I to completely remove the DNA template (reaction mixture: 1.5 μl of 10x reaction buffer, 0.15 μl BSA, 0.85 μl of Dpn I and 83 μl cleaned PCR DNA).

The Dpn I processed DNA mixture was cleaned up using Zymo DNA clean-up kit following the manufacturer's protocol. The cleaned-up DNA was used to transform an electro-competent strain of *E. coli* Bw25113 (*ΔilvC*) using a BioRad Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). The transformed clones were streaked on agar plates containing the LB medium and 100 μg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA) and incubated at 37 °C overnight. Clones were screened for improved activity using NADH. K_M for the improved mutants was measured (Table 16).

Table 15: Primers for example 6

2017200439 20 Jan 2017

Targeted position(s) of *K9-KARI*	Primers
131	K9_131L_112210f: ggacgttgatgtcTTGatgatcgctcc SEQ ID NO: 166 K9_131A_112210f: ggacgttgatgtcGCAatgatcgctcc SEQ ID NO: 167 K9_131V_112210f: ggacgttgatgtcGTTatgatcgctcc SEQ ID NO: 168 K9_131M_112210f: ggacgttgatgtcATGatgatcgctcc SEQ ID NO: 169 K9_131C_112210f: ggacgttgatgtcTGAatgatcgctcc SEQ ID NO: 170
191	K9_191D_112210r: GGTTTCAGTTTCGTCTCTGAAGGTAGTTTC SEQ ID NO: 171 K9_191C_112210r: GGTTTCAGTTTCGCATCTGAAGGTAGTTTC SEQ ID NO: 172 K9_191S_112210r: GGTTTCAGTTTCCGATCTGAAGGTAGTTTC SEQ ID NO: 173 K9_191G_112210r: GGTTTCAGTTTCGCCTCTGAAGGTAGTTTC SEQ ID NO: 174

Table 16

List of some mutants with their measured K_M values

Mutant	SEQ ID NO: (nucleic acid, amino acid)	Mutations	K_M (μM) (NADH)	K_M (μM) (NADPH)
AWB9	52, 53	S56A/S58D/I86V/N87P/T131A	10	4
AWC1	54, 55	S56A/S58D/I86V/N87P/T131V	9	3
AWD6	62, 63	S56A/S58D/I86V/N87P/T131V/T191S	5	2
AWD10	64, 65	S56A/S58D/I86V/N87P/T131A/T191C	8	3
AWF4	56, 57	S56A/S58D/I86V/N87P/N107S/T131V	7	3
AWF6	58, 59	S56A/S58D/I86V/N87P/T131V/T191D	7	2
AWG4	50, 51	S56A/S58D/I86V/N87P/T131M	7	3
AWH3	60, 61	S56A/S58D/I86V/N87P/T131V/T191G	6	2
AS6F1	48, 49	S56A/S58D/I86V/N87P/T131M/T191G	4	1

Example 7

Isobutanol production from K9 KARI variants

The following variants of K9 KARI were generated as described above.

2017200439 20 Jan 2017

Table 17.

KARI variants and the corresponding yeast expression vectors

Clone Names	Yeast Vector Name	Nucleotide Point Mutation Locations					Amino Acid Mutation Positions
		166	168	172	173	174	
WT K9 KARI	pHR81-PIlv5-KARI-K9	T	T	T	C	C	S56, S58
AB1G9	pHR81-PIlv5-KARI-K9.G9	G	G	G	A	G	S56A, S58E
495B5	pHR81-PIlv5-KARI-K9.B5	G	T	C	A	T	S56A, S58H
AB1D3	pHR81-PIlv5-KARI-K9.D3	G	T	G	A	T	S56A, S58D
AB1D1	pHR81-PIlv5-KARI-K9.D1	G	T	T	C	C	S56A

The yeast expression plasmids were made by subcloning of the variant KARI genes from *E.coli* vectors (pBAD.KARI) into pHR81-PIlv5-Pf5.KARI vector pLH556 (Figure 4, SEQ ID NO: 138) at PmeI and SfiI sites. Yeast pathway strains were made in PNY2204 host (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-pUC19-loxP-kanMX-loxP-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t; Example 13) by co-transforming the KARI vectors as pathway plasmid #1, and pBP915 (pRS423-P_{FBA1}-DHAD-P_{GPM1}-hADH1; SEQ ID NO: 182) as pathway plasmid #2. Transformants were patched to the same medium containing 2% glucose and 0.1% ethanol as carbon sources. Three patches were tested for isobutanol production under microaerobic conditions in serum vials. A clone that was transformed with pBP915 and the pLH702 plasmid which expresses K9D3 was designated PNY1910.

2017200439 20 Jan 2017

Yeast colonies from the transformation on SE-Ura-His plates appeared after 5-7 days. The colonies were patched onto fresh SE-Ura-His plates, incubate at 30°C for 3 days. The patched cells were inoculated into 25 mL SEG-Ura,His media with 0.2% glucose and 0.2% ethanol and grown aerobically for 1-2 days at 30°C, to 2-3 OD. The cells were centrifuged and re-suspended in 1 mL of SEG-Ura,His media (2% glucose, 0.1% ethanol, 10 mg/L ergosterol, 50 mM MES, pH 5.5, thiamine 30 mg/L, nicotinic acid 30 mg/L). A calculated amount of cells were transferred to 45 mL total volume of the same media for a starting OD=0.2 in a 60 mL serum vial, with the top closed tightly by a crimper. This step was done in the regular bio-hood in air. The serum vials were incubated at 30C, 200 rpm for 2 days. At 48 h, the samples were removed for OD and HPLC analysis of glucose, isobutanol and pathway intermediates. 24 h samples were taken in an anaerobic chamber to maintain the anaerobic condition in the serum vials. In the initial phase of the 48h incubation, the air present in the head space (~15 mL) and the liquid media is consumed by the growing yeast cells. After the oxygen in the head space is consumed, the culture becomes anaerobic. Therefore this experiment includes switching condition from aerobic to oxygen limiting and anaerobic conditions.

Of the four K9 variants, AB1G9 and AB1D3 produced relatively high isobutanol titers, while 495B5 and AB1D1 have lower titer. Wild type K9 KARI strain produced the lowest titer. While not wishing to be bound by theory, it is believed that the lower titer is due to the shifted balance of NADH and NADPH when cells are switched from aerobic to anaerobic conditions. By this rationale, under anaerobic conditions, NADH concentration and availability increased significantly, favoring the variant KARI enzymes that use NADH. Based on the kinetic analysis, AB1G9 ("K9G9") and AB1D3 ("K9D3") mutants have relatively high K_M for NADPH (23 & 9.2 μM), in addition to their relative low K_M for NADH (47 & 38 μM). As comparison, 495B5 and AB1D1's K_M's are 2.5 and 1.1 μM respectively for NADPH, and wt K9's K_M is 0.10 μM. The low NADH K_M of AB1G9 and AB1D3, together with the high NADPH K_M of AB1G9 and AB1D3 may

2017200439 20 Jan 2017

have led to reduced NADPH utilization under anaerobic conditions, and relatively high NADH utilization. As evidence, AB1G9 and AB1D3 have lower glycerol accumulation (isobutanol:glycerol = 3.3) compared to 495B5 and AB1D1 (2-3). The isobutanol:glycerol ratio is for the wild type K9 is 1:1 under the same switched aerobic to anaerobic condition.

Table 18.

Kinetic properties of wild type and variant K9 KARI enzymes, and isobutanol titer and productivity measured from aerobic to anaerobic switch experiment in serum vials.

Clone Names	**K_M (NADPH)**	**NADPH V_{max} (U/mg)**	**K_M (NADH)**	**NADH V_{max} (U/mg)**	**Isobutanol g/L**	**Effective Isobutanol Productivity (g/g cells)**
WT K9 KARI	0.19	2.0	326	1.5	0.9	3.1
AB1G9	23	2.4	47	2.0	3.4	10.5
495B5	3.5	2.5	44	1.9	2.2	9.2
AB1D3	9.2	2.1	38	1.9	3.3	10.8
AB1D1	1.1	2.8	164	2.1	2.3	9.5

Example 8

Construction of *Saccharomyces cerevisiae* strains BP1135 (PNY1505) and PNY1507 and Isobutanol-Producing Derivatives

This example describes construction of *Saccharomyces cerevisiae* strains BP1135 and PNY1507. These strains were derived from PNY1503 (BP1064). PNY1503 was derived from CEN.PK 113-7D (CBS 8340; Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversiry Centre, Netherlands). BP1135 contains an additional deletion of the *FRA2* gene.

2017200439 20 Jan 2017

PNY1507 was derived from BP1135 with additional deletion of the *ADH1* gene, with integration of the *kivD* gene from *Lactococcus lactis*, codon optimized for expression in *Saccharomyces cerevisiae*, into the *ADH1* locus.

Deletions/integrations were created by homologous recombination with PCR fragments containing regions of homology upstream and downstream of the target gene and the *URA3* gene for selection of transformants. The *URA3* gene was removed by homologous recombination to create a scarless deletion/integration.

The scarless deletion/integration procedure was adapted from Akada *et al.*, *Yeast, 23*:399, 2006. In general, the PCR cassette for each deletion/integration was made by combining four fragments, A-B-U-C, and the gene to be integratied by cloning the individual fragments into a plasmid prior to the entire cassette being amplified by PCR for the deletion/integration procedure. The gene to be integrated was included in the cassette between fragments A and B. The PCR cassette contained a selectable/counter-selectable marker, *URA3* (Fragment U), consisting of the native CEN.PK 113-7D URA3 gene, along with the promoter (250 bp upstream of the *URA3* gene) and terminator (150 bp downstream of the *URA3* gene) regions. Fragments A and C (each approximately 100 to 500 bp long) corresponded to the sequence immediately upstream of the target region (Fragment A) and the 3' sequence of the target region (Fragment C). Fragments A and C were used for integration of the cassette into the chromosome by homologous recombination. Fragment B (500 bp long) corresponded to the 500 bp immediately downstream of the target region and was used for excision of the URA3 marker and Fragment C from the chromosome by homologous recombination, as a direct repeat of the sequence corresponding to Fragment B was created upon integration of the cassette into the chromosome.

<u>*FRA2* Deletion</u>

The *FRA2* deletion was designed to delete 250 nucleotides from the 3' end of the coding sequence, leaving the first 113 nucleotides of the

2017200439 20 Jan 2017

FRA2 coding sequence intact. An in-frame stop codon was present 7 nucleotides downstream of the deletion. The four fragments for the PCR cassette for the scarless *FRA2* deletion were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). FRA2 Fragment A was amplified with primer oBP594 (SEQ ID NO: 183) and primer oBP595 (SEQ ID NO: 184), containing a 5' tail with homology to the 5' end of FRA2 Fragment B. FRA2 Fragment B was amplified with primer oBP596 (SEQ ID NO: 185), containing a 5' tail with homology to the 3' end of FRA2 Fragment A, and primer oBP597 (SEQ ID NO:186), containing a 5' tail with homology to the 5' end of FRA2 Fragment U. FRA2 Fragment U was amplified with primer oBP598 (SEQ ID NO: 187), containing a 5' tail with homology to the 3' end of FRA2 Fragment B, and primer oBP599 (SEQ ID NO: 188), containing a 5' tail with homology to the 5' end of FRA2 Fragment C. FRA2 Fragment C was amplified with primer oBP600 (SEQ ID NO:189), containing a 5' tail with homology to the 3' end of FRA2 Fragment U, and primer oBP601 (SEQ ID NO:190). PCR products were purified with a PCR Purification kit (Qiagen). FRA2 Fragment AB was created by overlapping PCR by mixing FRA2 Fragment A and FRA2 Fragment B and amplifying with primers oBP594 (SEQ ID NO:183) and oBP597 (SEQ ID NO:186). FRA2 Fragment UC was created by overlapping PCR by mixing FRA2 Fragment U and FRA2 Fragment C and amplifying with primers oBP598 (SEQ ID NO:187) and oBP601 (SEQ ID NO:190). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The FRA2 ABUC cassette was created by overlapping PCR by mixing FRA2 Fragment AB and FRA2 Fragment UC and amplifying with primers oBP594 (SEQ ID NO:183 and oBP601 (SEQ ID NO:190). The PCR product was purified with a PCR Purification kit (Qiagen).

Competent cells of PNY1503 were made and transformed with the FRA2 ABUC PCR cassette using a Frozen-EZ Yeast Transformation II kit (Zymo Research; Orange, CA). Transformation mixtures were plated on

2017200439 20 Jan 2017

synthetic complete media lacking uracil supplemented with 1% ethanol at 30 °C. Transformants with a *fra2* knockout were screened for by PCR with primers oBP602 (SEQ ID NO:191) and oBP603 (SEQ ID NO:192) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). A correct transformant was grown in YPE (yeast extract, peptone, 1% ethanol) and plated on synthetic complete medium containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the URA3 marker. The deletion and marker removal were confirmed by PCR with primers oBP602 (SEQ ID NO:191) and oBP603 (SEQ ID NO:192) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The absence of the *FRA2* gene from the isolate was demonstrated by a negative PCR result using primers specific for the deleted coding sequence of *FRA2*, oBP605 (SEQ ID NO:193) and oBP606 (SEQ ID NO:194). The correct isolate was selected as strain CEN.PK 113-7D MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ and designated as PNY1505 (BP1135).

This strain was transformed with isobutanol pathway plasmids (pYZ090, SEQ ID NO: 195) and pLH468 (SEQ ID NO: 139), and one clone was designated BP1168 (PNY1506).

pYZ090 (SEQ ID NO: 195) was constructed to contain a chimeric gene having the coding region of the alsS gene from *Bacillus subtilis* (nt position 457-2172) expressed from the yeast CUP1 promoter (nt 2-449) and followed by the CYC1 terminator (nt 2181-2430) for expression of ALS, and a chimeric gene having the coding region of the ilvC gene from *Lactococcus lactis* (nt 3634-4656) expressed from the yeast ILV5 promoter (2433-3626) and followed by the ILV5 terminator (nt 4682-5304) for expression of KARI.

ADH1 Deletion and *kivD LI(y)* Integration

The *ADH1* gene was deleted and replaced with the *kivD* coding region from Lactococcus *lactis* codon optimized for expression in *Saccharomyces cerevisiae*. The scarless cassette for the *ADH1* deletion-

2017200439 20 Jan 2017

kivD_LI(y) integration was first cloned into plasmid pUC19-URA3MCS, as described in U.S. Appln. No. 61/356379, filed June 18, 2010, incorporated herein by reference. The vector is pUC19 based and contains the sequence of the *URA3* gene from *Saccharomyces cerevisiae* CEN.PK 113-7D situated within a multiple cloning site (MCS). pUC19 contains the pMB1 replicon and a gene coding for beta-lactamase for replication and selection in *Escherichia coli*. In addition to the coding sequence for *URA3*, the sequences from upstream (250 bp) and downstream (150 bp) of this gene are present for expression of the *URA3* gene in yeast. The vector can be used for cloning purposes and can be used as a yeast integration vector.

The *kivD* coding region from *Lactococcus lactis* codon optimized for expression in *Saccharomyces cerevisiae* was amplified using pLH468 (SEQ ID NO:139) as template with primer oBP562 (SEQ ID NO:197), containing a PmeI restriction site, and primer oBP563 (SEQ ID NO:198), containing a 5' tail with homology to the 5' end of ADH1 Fragment B. ADH1 Fragment B was amplified from genomic DNA prepared as above with primer oBP564 (SEQ ID NO:199), containing a 5' tail with homology to the 3' end of *kivD_LI(y)*, and primer oBP565 (SEQ ID NO:200), containing a FseI restriction site. PCR products were purified with a PCR Purification kit (Qiagen). kivD_LI(y)-ADH1 Fragment B was created by overlapping PCR by mixing the kivD_LI(y) and ADH1 Fragment B PCR products and amplifying with primers oBP562 (SEQ ID NO:197) and oBP565 (SEQ ID NO:200). The resulting PCR product was digested with PmeI and FseI and ligated with T4 DNA ligase into the corresponding sites of pUC19-URA3MCS after digestion with the appropriate enzymes. ADH1 Fragment A was amplified from genomic DNA with primer oBP505 (SEQ ID NO:201), containing a SacI restriction site, and primer oBP506 (SEQ ID NO:202), containing an AscI restriction site. The ADH1 Fragment A PCR product was digested with SacI and AscI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing kivD_LI(y)-ADH1 Fragment B. ADH1 Fragment C was amplified from genomic DNA with primer oBP507 (SEQ ID NO:203), containing a PacI restriction site, and

2017200439 20 Jan 2017

primer oBP508 (SEQ ID NO:204), containing a SalI restriction site. The ADH1 Fragment C PCR product was digested with PacI and SalI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing ADH1 Fragment A-kivD_LI(y)-ADH1 Fragment B. The hybrid promoter UAS(PGK1)-P_{FBA1} was amplified from vector pRS316-UAS(PGK1)-P_{FBA1}-GUS (SEQ ID NO:209) with primer oBP674 (SEQ ID NO:205), containing an AscI restriction site, and primer oBP675 (SEQ ID NO:206), containing a PmeI restriction site. The UAS(PGK1)-P_{FBA1} PCR product was digested with AscI and PmeI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing kivD_LI(y)-ADH1 Fragments ABC. The entire integration cassette was amplified from the resulting plasmid with primers oBP505 (SEQ ID NO:201) and oBP508 (SEQ ID NO:204) and purified with a PCR Purification kit (Qiagen).

Competent cells of PNY1505 were made and transformed with the ADH1-kivD_LI(y) PCR cassette constructed above using a Frozen-EZ Yeast Transformation II kit (Zymo Research). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 1% ethanol at 30 °C. Transformants were grown in YPE (1% ethanol) and plated on synthetic complete medium containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the URA3 marker. The deletion of *ADH1* and integration of *kivD_LI(y)* were confirmed by PCR with external primers oBP495 (SEQ ID NO:207) and oBP496 (SEQ ID NO:208) and with *kivD_LI(y)* specific primer oBP562 (SEQ ID NO:197) and external primer oBP496 (SEQ ID NO:208) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The correct isolate was selected as strain CEN.PK 113-7D MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1tpdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t and designated as PNY1507 (BP1201). PNY1507 was transformed with isobutanol pathway plasmids pYZ090 (SEQ ID NO:195) and pBP915 (SEQ ID NO: 182) and the resultant strain was named PNY1513.

Construction of the pRS316-UAS(PGK1)-FBA1p-GUS vector

2017200439 20 Jan 2017

To clone a cassette UAS(PGK1)-FBA1p (SEQ ID NO:766, first a 602bp FBA1 promoter (FBA1p) was PCR-amplified from genomic DNA of CEN.PK with primers T-FBA1(SalI) (SEQ ID NO:767) and B-FBA1(SpeI) (SEQ ID NO:768), and cloned into SalI and SpeI sites on the plasmid pWS358-PGK1p-GUS (SEQ ID NO:769) after the PGK1p promoter was removed with a SalI/SpeI digest of the plasmid, yielding pWS358-FBA1p-GUS. The pWS358-PGK1p-GUS plasmid was generated by inserting a PGK1p and beta-glucuronidase gene (GUS) DNA fragments into multiple cloning site of pWS358, which was derived from pRS423 vector (Christianson et al., Gene, 110:119-122, 1992). Secondly, the resulting pWS358-FBA1p-GUS plasmid was digested with SalI and SacI, a DNA fragment containing a FBA1p promoter, GUS gene, and FBAt terminator gel-purified, and cloned into SalI/SacI sites on pRS316 to create pRS316-FBA1p-GUS. Thirdly, a 118bp DNA fragment containing an upstream activation sequence (UAS) located between positions -519 and -402 upstream of the 3-phosphoglycerate kinase (PGK1) open reading frame, namely UAS(PGK1), was PCR-amplified from genomic DNA of CEN.PK with primers T-U/PGK1(KpnI) (SEQ ID NO:770) and B-U/PGK1(SalI) (SEQ ID NO:771). The PCR product was digested with KpnI and SalI and cloned into KpnI/SalI sites on pRS316-FBA1p-GUS to create pRS316-UAS(PGK1)-FBA1p-GUS.

Example 9

Improved recombinant host cells comprising elimination of ALD6

The purpose of this example is to describe methods to modify a yeast host strain for improved production of isobutanol. These modifications include integration of genes encoding isobutyraldehyde reductase activity and elimination of the native genes ALD6 and YMR226c, encoding NADP+-dependent acetaldehyde dehydrogenase and a NADPH-dependent dehydrogenase, respectively.

Construction of *S. cerevisiae* Strain PNY2211

2017200439 20 Jan 2017

PNY2211 was constructed in several steps from *S. cerevisiae* strain PNY1507 (Example 8) as described in the following paragraphs. First PNY1507 was modified to contain a phosophoketolase gene. Next, an acetolactate synthase gene (*alsS*) was added to the strain, using an integration vector targeted to sequences adjacent to the phosphokeloase gene. Finally, homologous recombination was used to remove the phosphoketolase gene and integration vector sequences, resulting in a scarless insertion of *alsS* in the intergenic region between pdc1Δ::ilvD (described in Example 12) and the native TRX1 gene of chromosome XII. The resulting genotype of PNY2211 is MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH| sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t.

A phosphoketolase gene cassette was introduced into PNY1507 by homologous recombination. The integration construct was generated as follows. The plasmid pRS423::CUP1-alsS+FBA-budA (previously described in US2009/0305363, which is herein incorporated by reference in its entirety) was digested with *Not*I and *Xma*I to remove the 1.8 kb FBA-budA sequence, and the vector was religated after treatment with Klenow fragment. Next, the CUP1 promoter was replaced with a TEF1 promoter variant (M4 variant previously described by Nevoigt *et al. Appl. Environ. Microbiol. 72*: 5266-5273 (2006), which is herein incorporated by reference in its entirety) via DNA synthesis and vector construction service from DNA2.0 (Menlo Park, CA). The resulting plasmid, pRS423::TEF(M4)-alsS was cut with *Stu*I and *Mlu*I (removes 1.6 kb portion containing part of the alsS gene and CYC1 terminator), combined with the 4 kb PCR product generated from pRS426::GPD-xpk1+ADH-eutD (SEQ ID NO:383) with primers N1176 (SEQ ID NO:282) and N1177 (SEQ ID NO:283) and an 0.8 kb PCR product DNA (SEQ ID NO: 284) generated from yeast genomic DNA (ENO1 promoter region) with primers N822 (SEQ ID NO:285) and N1178 (SEQ ID NO:286) and transformed into *S. cerevisiae* strain BY4741 (ATCC #201388); gap repair cloning methodology, *see* Ma *et al. Gene 58*:201-216 (1987). Transformants were obtained by plating cells on

2017200439 20 Jan 2017

synthetic complete medium without histidine. Proper assembly of the expected plasmid (pRS423::TEF(M4)-xpk1+ENO1-eutD, SEQ ID NO:293) was confirmed by PCR (primers N821 (SEQ ID NO:287) and N1115 (SEQ ID NO:288)) and by restriction digest (BglII). Two clones were subsequently sequenced. The 3.1 kb TEF(M4)-xpk1 gene was isolated by digestion with *Sac*I and *Not*I and cloned into the pUC19-URA3::ilvD-TRX1 vector (Clone A, cut with *Afl*II). Cloning fragments were treated with Klenow fragment to generate blunt ends for ligation. Ligation reactions were transformed into *E. coli* Stbl3 cells, selecting for ampicillin resistance. Insertion of TEF(M4)-xpk1 was confirmed by PCR (primers N1110 (SEQ ID NO:367) and N1114 (SEQ ID NO:290)). The vector was linearized with *Afl*II and treated with Klenow fragment. The 1.8 kb KpnI-HincII geneticin resistance cassette (SEQ ID NO: 384)was cloned by ligation after Klenow fragment treatment. Ligation reactions were transformed into *E. coli* Stbl3 cells, selecting for ampicillin resistance. Insertion of the geneticin cassette was confirmed by PCR (primers N160SeqF5 (SEQ ID NO:210) and BK468 (SEQ ID NO:368)). The plasmid sequence is provided as SEQ ID NO: 291 (pUC19-URA3::pdc1::TEF(M4)-xpk1::kan).

The resulting integration cassette (pdc1::TEF(M4)-xpk1::KanMX::TRX1) was isolated (*Asc*I and *Nae*I digestion generated a 5.3 kb band that was gel purified) and transformed into PNY1507 using the Zymo Research Frozen-EZ Yeast Transformation Kit (Cat. No. T2001). Transformants were selected by plating on YPE plus 50 µg/ml G418. Integration at the expected locus was confirmed by PCR (primers N886 (SEQ ID NO:211) and N1214 (SEQ ID NO:281)). Next, plasmid pRS423::GAL1p-Cre (SEQ ID NO:271), encoding Cre recombinase, was used to remove the loxP-flanked KanMX cassette. Proper removal of the cassette was confirmed by PCR (primers oBP512 (SEQ ID NO: 337) and N160SeqF5 (SEQ ID NO:210)). Finally, the *als*S integration plasmid described in Example 13, pUC19-kan::pdc1::FBA-alsS::TRX1, clone A) was transformed into this strain using the included geneticin selection marker. Two integrants were tested for acetolactate synthase activity by transformation with plasmids pYZ090ΔalsS (SEQ ID NO:371) and pBP915

2017200439 20 Jan 2017

(SEQ ID NO:182) (transformed using Protocol #2 in Amberg, Burke and Strathern "Methods in Yeast Genetics" (2005)), and evaluation of growth and isobutanol production in glucose-containing media (methods for growth and isobutanol measurement are as follows: All strains were grown in synthetic complete medium, minus histidine and uracil containing 0.3 % glucose and 0.3 % ethanol as carbon sources (10 mL medium in 125 mL vented Erlenmeyer flasks (VWR Cat. No. 89095-260). After overnight incubation (30 °C, 250 rpm in an Innova®40 New Brunswick Scientific Shaker), cultures were diluted back to 0.2 OD (Eppendorf BioPhotometer measurement) in synthetic complete medium containing 2% glucose and 0.05% ethanol (20 ml medium in 125 mL tightly-capped Erlenmeyer flasks (VWR Cat. No. 89095-260)). After 48 hours incubation (30 °C, 250 rpm in an Innova®40 New Brunswick Scientific Shaker), culture supernatants (collected using Spin-X centrifuge tube filter units, Costar Cat. No. 8169) were analyzed by HPLC per methods described in U.S. Appl. Pub. No. 2007/0092957, which is herein incorporated by reference in its entirety) One of the two clones was positive and was named PNY2218.

PNY2218 was treated with Cre recombinase and the resulting clones were screened for loss of the xpk1 gene and pUC19 integration vector sequences by PCR (primers N886 (SEQ ID NO: 211) and N160SeqR5 (SEQ ID NO: 388)). This left only the alsS gene integrated in the pdc1-TRX1 intergenic region after recombination the DNA upstream of xpk1 and the homologous DNA introduced during insertion of the integration vector (a "scarless" insertion since vector, marker gene and loxP sequences are lost). Although this recombination could have occurred at any point, the vector integration appeared to be stable even without geneticin selection and the recombination event was only observed after introduction of the Cre recombinase. One clone was designated PNY2211.

An isolate of PNY2218 containing the plasmids pYZ090ΔalsS and pBP915 was designated PNY2209.

PNY1528 (hADH integrations in PNY2211)

2017200439 20 Jan 2017

Deletions/integrations were created by homologous recombination with PCR products containing regions of homology upstream and downstream of the target region and the *URA3* gene for selection of transformants. The *URA3* gene was removed by homologous recombination to create a scarless deletion/integration.

YPRCΔ15 deletion and horse liver *adh* integration

The YPRCΔ15 locus was deleted and replaced with the horse liver adh gene, codon optimized for expression in Saccharomyces cerevisiae, along with the PDC5 promoter region (538 bp) from Saccharomyces cerevisiae and the ADH1 terminator region (316 bp) from Saccharomyces cerevisiae. The scarless cassette for the YPRCΔ15 deletion- P[PDC5]-adh_HL(y)-ADH1t integration was first cloned into plasmid pUC19-URA3MCS (described in Example 8).

Fragments A-B-U-C were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). YPRCΔ15 Fragment A was amplified from genomic DNA with primer oBP622 (SEQ ID NO: 212), containing a KpnI restriction site, and primer oBP623 (SEQ ID NO: 213), containing a 5' tail with homology to the 5' end of YPRCΔ15 Fragment B. YPRCΔ15 Fragment B was amplified from genomic DNA with primer oBP624 (SEQ ID NO: 214), containing a 5' tail with homology to the 3' end of YPRCΔ15 Fragment A, and primer oBP625 (SEQ ID NO: 215), containing a FseI restriction site. PCR products were purified with a PCR Purification kit (Qiagen). YPRCΔ15 Fragment A - YPRCΔ15 Fragment B was created by overlapping PCR by mixing the YPRCΔ15 Fragment A and YPRCΔ15 Fragment B PCR products and amplifying with primers oBP622 (SEQ ID NO: 212) and oBP625 (SEQ ID NO: 215). The resulting PCR product was digested with KpnI and FseI and ligated with T4 DNA ligase into the corresponding sites of pUC19-URA3MCS after digestion with the appropriate enzymes. YPRCΔ15 Fragment C was amplified from genomic DNA with primer oBP626 (SEQ ID NO: 216), containing a NotI restriction site, and primer oBP627 (SEQ ID NO: 217), containing a PacI restriction

2017200439 20 Jan 2017

site. The YPRCΔ15 Fragment C PCR product was digested with NotI and PacI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing YPRCΔ15 Fragments AB. The PDC5 promoter region was amplified from CEN.PK 113-7D genomic DNA with primer HY21 (SEQ ID NO: 218), containing an AscI restriction site, and primer HY24 (SEQ ID NO: 219), containing a 5' tail with homology to the 5' end of adh_HI(y). adh_HI(y)-ADH1t was amplified from pBP915 (SEQ ID NO: 182) with primers HY25 (SEQ ID NO: 220), containing a 5' tail with homology to the 3' end of P[PDC5], and HY4 (SEQ ID NO: 221), containing a PmeI restriction site. PCR products were purified with a PCR Purification kit (Qiagen). P[PDC5]-adh_HL(y)-ADH1t was created by overlapping PCR by mixing the P[PDC5] and adh_HL(y)-ADH1t PCR products and amplifying with primers HY21 (SEQ ID NO: 218) and HY4 (SEQ ID NO: 221).The resulting PCR product was digested with AscI and PmeI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing YPRCΔ15 Fragments ABC. The entire integration cassette was amplified from the resulting plasmid with primers oBP622 (SEQ ID NO: 212) and oBP627 (SEQ ID NO: 217).

Competent cells of PNY2211 were made and transformed with the *YPRCΔ15* deletion- P[PDC5]-adh_HL(y)-ADH1t integration cassette PCR product using a Frozen-EZ Yeast Transformation II kit (Zymo Research; Orange, CA). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 1% ethanol at 30 °C. Transformants were screened for by PCR with primers URA3-end F (SEQ ID NO: 222) and oBP637 (SEQ ID NO: 224). Correct transformants were grown in YPE (1% ethanol) and plated on synthetic complete medium supplemented with 1% EtOH and containing 5-fluoro-orotic acid (0.1%) at 30 C to select for isolates that lost the *URA3* marker. The deletion of *YPRCΔ15* and integration of P[PDC5]-adh_HL(y)-ADH1t were confirmed by PCR with external primers oBP636 (SEQ ID NO: 223) and oBP637 (SEQ ID NO: 224) using genomic DNA prepared with a YeaStar Genomic DNA kit (Zymo Research). A correct isolate of the following genotype was selected for further modification: CEN.PK 113-7D MATa ura3Δ::loxP his3Δ

2017200439 20 Jan 2017

pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_HI-ADH1t.

Horse liver *adh* integration at *fra2Δ*

The horse liver *adh* gene, codon optimized for expression in *Saccharomyces cerevisiae*, along with the *PDC1* promoter region (870 bp) from *Saccharomyces cerevisiae* and the *ADH1* terminator region (316 bp) from *Saccharomyces cerevisiae*, was integrated into the site of the *fra2* deletion. The scarless cassette for the *fra2Δ*- P[PDC1]-adh_HL(y)-ADH1t integration was first cloned into plasmid pUC19-URA3MCS.

Fragments A-B-U-C were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). fra2Δ Fragment C was amplified from genomic DNA with primer oBP695 (SEQ ID NO: 229), containing a NotI restriction site, and primer oBP696 (SEQ ID NO: 230), containing a PacI restriction site. The fra2Δ Fragment C PCR product was digested with NotI and PacI and ligated with T4 DNA ligase into the corresponding sites of pUC19-URA3MCS. fra2Δ Fragment B was amplified from genomic DNA with primer oBP693 (SEQ ID NO: 227), containing a PmeI restriction site, and primer oBP694 (SEQ ID NO: 228), containing a FseI restriction site. The resulting PCR product was digested with PmeI and FseI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing fra2Δ fragment C after digestion with the appropriate enzymes. fra2Δ Fragment A was amplified from genomic DNA with primer oBP691 (SEQ ID NO: 225), containing BamHI and AsiSI restriction sites, and primer oBP692 (SEQ ID NO: 226), containing AscI and SwaI restriction sites. The fra2Δ fragment A PCR product was digested with BamHI and AscI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing fra2Δ fragments BC after digestion with the appropriate enzymes. The *PDC1* promoter region was amplified from CEN.PK 113-7D

2017200439 20 Jan 2017

genomic DNA with primer HY16 (SEQ ID NO: 231), containing an AscI restriction site, and primer HY19 (SEQ ID NO: 232), containing a 5' tail with homology to the 5' end of adh_HI(y). adh_HI(y)-ADH1t was amplified from pBP915 with primers HY20 (SEQ ID NO: 233), containing a 5' tail with homology to the 3' end of P[PDC1], and HY4 (SEQ ID NO: 221), containing a PmeI restriction site. PCR products were purified with a PCR Purification kit (Qiagen). P[PDC1]-adh_HL(y)-ADH1t was created by overlapping PCR by mixing the P[PDC1] and adh_HL(y)-ADH1t PCR products and amplifying with primers HY16 (SEQ ID NO: 231) and HY4 (SEQ ID NO: 221).The resulting PCR product was digested with AscI and PmeI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing fra2Δ Fragments ABC. The entire integration cassette was amplified from the resulting plasmid with primers oBP691 (SEQ ID NO: 225) and oBP696 (SEQ ID NO: 230).

Competent cells of the PNY2211 variant with adh_HI(y) integrated at *YPRCΔ15*were made and transformed with the *fra2Δ*- P[PDC1]-adh_HL(y)-ADH1t integration cassette PCR product using a Frozen-EZ Yeast Transformation II kit (Zymo Research). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 1% ethanol at 30 °C. Transformants were screened for by PCR with primers URA3-end F (SEQ ID NO: 222) and oBP731 (SEQ ID NO: 235). Correct transformants were grown in YPE (1% ethanol) and plated on synthetic complete medium supplemented with 1% EtOH and containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the *URA3* marker. The integration of P[PDC1]-adh_HL(y)-ADH1t was confirmed by colony PCR with internal primer HY31 (SEQ ID NO: 236) and external primer oBP731 (SEQ ID NO: 235) and PCR with external primers oBP730 (SEQ ID NO: 234) and oBP731 (SEQ ID NO: 235) using genomic DNA prepared with a YeaStar Genomic DNA kit (Zymo Research). A correct isolate of the following genotype was designated PNY1528: CEN.PK 113-7D MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-

2017200439 20 Jan 2017

ADH|adh_HI-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_HI-ADH1t.

PNY2237 (scarless YMR226c deletion)

The gene YMR226c was deleted from *S. cerevisiae* strain PNY1528 by homologous recombination using a PCR amplified 2.0 kb linear scarless deletion cassette. The cassette was constructed from spliced PCR amplified fragments comprised of the URA3 gene, along with its native promoter and terminator as a selectable marker, upstream and downstream homology sequences flanking the YMR226c gene chromosomal locus to promote integration of the deletion cassette and removal of the native intervening sequence and a repeat sequence to promote recombination and removal of the URA3 marker. Forward and reverse PCR primers (N1251 and N1252, SEQ ID NOs: 247 and 248, respectively), amplified a 1,208 bp URA3 expression cassette originating from pLA33 (pUC19::loxP-URA3-loxP (SEQ ID NO: 268)). Forward and reverse primers (N1253 and N1254, SEQ ID NOs: 249 and 250, respectively), amplified a 250 bp downstream homology sequence with a 3' URA3 overlap sequence tag from a genomic DNA preparation of *S. cerevisiae* strain PNY2211 (above). Forward and reverse PCR primers (N1255 and N1256, SEQ ID NOs: 251 and 252, respectively) amplified a 250 bp repeat sequence with a 5' URA3 overlap sequence tag from a genomic DNA preparation of *S. cerevisiae* strain PNY2211. Forward and reverse PCR primers (N1257 and N1258, SEQ ID NOs: 253 and 254, respectively) amplified a 250 bp upstream homology sequence with a 5' repeat overlap sequence tag from a genomic DNA preparation of *S. cerevisiae* strain PNY2211.

Approximately 1.5 μg of the PCR amplified cassette was transformed into strain PNY1528 (above) made competent using the ZYMO Research Frozen Yeast Transformation Kit and the transformation mix plated on SE 1.0% -uracil and incubated at 30^{0}C for selection of cells with an integrated ymr226cΔ::URA3 cassette. Transformants appearing after 72 to 96 hours are subsequently short-streaked on the same medium

2017200439 20 Jan 2017

and incubated at 30^{0}C for 24 to 48 hours. The short-streaks are screened for ymr226cΔ::URA3 by PCR, with a 5' outward facing URA3 deletion cassette-specific internal primer (N1249, SEQ ID NO: 245) paired with a flanking inward facing chromosome-specific primer (N1239, SEQ ID NO: 243) and a 3' outward-facing URA3 deletion cassette-specific primer (N1250, SEQ ID NO: 246) paired with a flanking inward-facing chromosome-specific primer (N1242, SEQ ID NO: 244). A positive PNY1528 ymr226cΔ::URA3 PCR screen resulted in 5' and 3' PCR products of 598 and 726 bp, respectively.

Three positive PNY1528 ymr226cΔ::URA3 clones were picked and cultured overnight in a YPE 1% medium of which 100μL was plated on YPE 1% + 5-FOA for marker removal. Colonies appearing after 24 to 48 hours were PCR screened for marker loss with 5' and 3' chromosome-specific primers (N1239 and N1242). A positive PNY1528 ymr226cΔ markerless PCR screen resulted in a PCR product of 801 bp. Multiple clones were obtained and one was designated PNY2237.

PNY2238 and PNY2243 (ALD6 deletion strains)

A vector was designed to replace the *ALD6* coding sequence with a Cre-lox recyclable URA3 selection marker. Sequences 5' and 3' of ALD6 were amplified by PCR (primer pairs N1179 and N1180 and N1181 and N1182, respectively; SEQ ID NOs: 237, 238, 239, and 240, respectively). After cloning these fragments into TOPO vectors (Invitrogen Cat. No. K2875-J10) and sequencing (M13 forward (SEQ ID NO:269) and reverse (SEQ ID NO:270) primers), the 5' and 3' flanks were cloned into pLA33 (pUC19::loxP::URA3::loxP) (SEQ ID NO:268) at the *EcoR*I and *Sph*I sites, respectively. Each ligation reaction was transformed into *E. coli* Stbl3 cells, which were incubated on LB Amp plates to select for transformants. Proper insertion of sequences was confirmed by PCR (primers M13 forward (SEQ ID NO: 269) and N1180 (SEQ ID NO:238) and M13 reverse (SEQ ID NO:270) and N1181 (SEQ ID NO:239), respectively).

The vector described above (pUC19::ald6Δ::loxP-URA3-loxP) was linearized with *Ahd*I and transformed into PNY1528 and PNY2237 using

2017200439 20 Jan 2017

the standard lithium acetate method (except that incubation of cells with DNA was extended to 2.5h). Transformants were obtained by plating on synthetic complete medium minus uracil that provided 1% ethanol as the carbon source. Patched transformants were screened by PCR to confirm the deletion/integration, using primers N1212 (SEQ ID NO: 241) and N1180 (5' end) (SEQ ID NO: 238) and N1181 (SEQ ID NO: 239) and N1213 (SEQ ID NO: 242) (3' end). A plasmid carrying Cre recombinase (pRS423::GAL1p-Cre = SEQ ID No. 271) was transformed into the strain using histidine marker selection. Transformants were passaged on YPE supplemented with 0.5% galactose. Colonies were screened for resistance to 5-FOA (loss of URA3 marker) and for histidine auxotrophy (loss of the Cre plasmid). Proper removal of the URA3 gene via the flanking loxP sites was confirmed by PCR (primers N1262 and N1263, SEQ ID NOs: 255 and 256, respectively). Additionally, primers internal to the ALD6 gene (N1230 and N1231; SEQ ID NOs: 261 and 262, respectively) were used to insure that no merodiploids were present. Finally, *ald6*Δ::loxP clones were screened by PCR to confirm that a translocation between *ura3*Δ::loxP (N1228 and N1229, SEQ ID NOs: 259 and 260) and *gpd2*Δ::loxP (N1223 and N1225, SEQ ID NOs: 257 and 258) had not occurred. Two positive clones were identified from screening of transformants of PNY1528. Clone B has been designated PNY2243. Three positive clones were identified from screening transformants of PNY2237. Clones E and K were both assessed for isobutanol production at small scale (below). Although statistically identical in most parameters, Clone E was selected (PNY2238) for further development.

Example 10

Isobutanol pathway plasmids

The purpose of this example is to describe construction or modification of isobutanol pathway plasmids for production of isobutanol in host strains.

pYZ067 (SEQ ID NO:374) was constructed to contain the following chimeric genes: 1) the coding region of the ilvD gene from *S. mutans*

2017200439 20 Jan 2017

UA159 with a C-terminal Lumio tag expressed from the yeast FBA1 promoter followed by the FBA1 terminator for expression of dihydroxy acid dehydratase, 2) the coding region for horse liver ADH expressed from the yeast GPM1 promoter followed by the ADH1 terminator for expression of alcohol dehydrogenase, and 3) the coding region of the KivD gene from *Lactococcus lactis* expressed from the yeast TDH3 promoter followed by the TDH3 terminator for expression of ketoisovalerate decarboxylase.

pYZ067ΔkivDΔhADH (SEQ ID NO: 385) was constructed from pYZ067 (SEQ ID NO: 374) by deleting the promoter-gene-terminator cassettes for both *kivD* and *adh*. pYZ067 was digested with BamHI and SacI (New England BioLabs; Ipswich, MA), and the 7934 bp fragment was purified on an agarose gel followed by a Gel Extraction kit (Qiagen; Valencia, CA). The isolated fragment of DNA was treated with DNA Polymerase I, Large (Klenow) Fragment (New England BioLabs; Ipswich, MA) and then self-ligated with T4 DNA ligase and used to transform competent TOP10 *Escherichia coli* (Invitrogen; Carlsbad, CA). Plasmids from transformants were isolated and checked for the proper deletion by sequence analysis. A correct plasmid isolate was designated pYZ067ΔkivDΔhADH.

pYZ067ΔkivDΔilvD (SEQ ID NO: 772) was constructed to contain a chimeric gene having the coding region of the adh gene from horse liver (nt position 3148-2021), codon optimized for expression in Saccharomyces cerevisiae, expressed from the yeast GPM promoter (nt 3916-3160) and followed by the ADH1 terminator (nt 2012-1697) for expression of ADH. pYZ067DkivDDilvD was constructed from pYZ067 by deleting the promoter-gene-terminator cassettes for both kivD and ilvD. pYZ067 was digested with AatII and SacI (New England BioLabs; Ipswich, MA) and the 10196 bp fragment was purified on an agarose gel followed by a Gel Extraction kit (Qiagen; Valencia, CA). The isolated fragment of DNA was treated with DNA Polymerase I, Large (Klenow) Fragment (New England BioLabs; Ipswich, MA) and then self-ligated with T4 DNA ligase. The resulting plasmid was then digested with NgoMIV and BamHI (New England BioLabs; Ipswich, MA) and the 7533 bp fragment was purified on

2017200439 20 Jan 2017

an agarose gel followed by a Gel Extraction kit (Qiagen; Valencia, CA). The isolated fragment of DNA was treated with DNA Polymerase I, Large (Klenow) Fragment (New England BioLabs; Ipswich, MA) and then self-ligated with T4 DNA ligase. Plasmids were isolated and checked for the proper deletions by sequence analysis. A correct plasmid isolate was designated pYZ067DkivDDilvD.

pK9G9.OLE1p.ilvD (SEQ ID NO: 773), derived from pYZ090 (SEQ ID NO: 195), was constructed to contain a chimeric gene having the coding region of the ilvD gene from Streptococcus mutans (nt position 5377-3641) expressed from the yeast OLE1 promoter (nt 5986-5387) and followed by the FBA1 terminator (nt 3632-3320) for expression of DHAD, and a chimeric gene having the coding region of the variant K9G9 of the ilvC gene from Anaerostipes caccae (nucleic acid and amino acid SEQ ID NOs: 774 and 647) (nt 1628-2659) expressed from the yeast ILV5 promoter (nt 427-1620) and followed by the ILV5 terminator (nt 2685-3307) for expression of KARI. Construction of the plasmid was as follows. The chimeric gene from plasmid pYZ067 having the coding region of the ilvD gene from Streptococcus mutans expressed from the yeast FBA1 promoter and followed by the FBA1 terminator was ligated into pYZ090 after digestion with restriction enzymes NgoMIV and BamHI. The alsS coding region and 280 bp from the 3' end of the CUP1 promoter was deleted from the resulting plasmid by digesting with the restriction enzymes SpeI and PacI and self-ligating the resulting large DNA fragment. The yeast FBA1 promoter upstream of ilvD was removed from the resulting plasmid by digesting with the restriction enzymes NgoMIV and PmlI and was replaced with the yeast OLE1 promoter amplified with primers pOLE1-NgoMI (SEQ ID NO: 775) and pOLE1-PmlI (SEQ ID NO: 776). The coding region of the ilvC gene from Lactococcus lactis was deleted from the resulting plasmid by digestion with restriction enzymes PmeI and SfiI followed by gel purification of the large DNA fragment. The coding region of the variant K9G9 ilvC gene (SEQ ID NO: 777) from Anaerostipes caccae was digested out of pLH701 (SEQ ID NO: 778) with

2017200439 20 Jan 2017

PmeI and SfiI and gel purified. The two DNA fragments were ligated to generate pK9G9.OLE1p.ilvD.

EXAMPLE 11

Construction of PNY2240 and PNY2242

Strain PNY2240 was derived from PNY2211 after transformation with plasmids pLH702 (SEQ ID NO: 181) and pBP915 (SEQ ID NO: 182). Transformants were plated on synthetic complete medium without histidine or uracil (1% ethanol as carbon source). Transformants were patched to the same medium containing, instead, 2% glucose and 0.05% ethanol as carbon sources. Three patches were used to inoculate liquid medium (synthetic complete minus uracil with 0.3% glucose and 0.3% ethanol as carbon sources). To test isobutanol production, liquid cultures were sub-cultured into synthetic complete medium minus uracil containing 2% glucose and 0.05% ethanol as carbon sources that also contained BME vitamin mix (Sigma Cat. No. B6891). Cultures were incubated in sealed serum vials (10 ml medium in 15 ml vials) at 30°C with shaking (250 rpm in an Infors Multitron shaker). After 48 hours, culture medium was filtered (Spin-X column) and analyzed by HPLC (as described in US App. Pub. No. 2007/0092957, which is incorporated herein by reference in its entirety). One clone was designated PNY2240.

Strain PNY2242 was derived from PNY2238 after transformation with plasmids pLH702 (SEQ ID NO: 181) and pYZ067ΔkivDΔhADH (described herein above). Transformants were plated on synthetic complete medium without histidine or uracil (1% ethanol as carbon source). Transformants were patched to the same medium containing, instead, 2% glucose and 0.05% ethanol as carbon sources. Three patches were tested for isobutanol production, as described above. All three performed similarly in terms of glucose consumption and isobutanol production. One clone was designated PNY2242 and was further characterized under fermentation conditions, as described herein below.

2017200439 20 Jan 2017

Example 12

Construction of Saccharomyces cerevisiae strain BP1064 (PNY1503)

The strain BP1064 was derived from CEN.PK 113-7D (CBS 8340; Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre, Netherlands) and contains deletions of the following genes: URA3, HIS3, PDC1, PDC5, PDC6, and GPD2. BP1064 was transformed with plasmids pYZ090 (SEQ ID NO:195) and pLH468 (SEQ ID NO:139) to create strain NGCI-070 (BP1083; PNY1504).

Deletions, which completely removed the entire coding sequence, were created by homologous recombination with PCR fragments containing regions of homology upstream and downstream of the target gene and either a G418 resistance marker or URA3 gene for selection of transformants. The G418 resistance marker, flanked by loxP sites, was removed using Cre recombinase (pRS423::PGAL1-cre; SEQ ID NO: 271). The URA3 gene was removed by homologous recombination to create a scarless deletion, or if flanked by loxP sites was removed using Cre recombinase.

URA3 Deletion

To delete the endogenous URA3 coding region, a ura3::loxP-kanMX-loxP cassette was PCR-amplified from pLA54 template DNA (SEQ ID NO:386). pLA54 contains the K. lactis TEF1 promoter and kanMX marker, and is flanked by loxP sites to allow recombination with Cre recombinase and removal of the marker. PCR was done using Phusion DNA polymerase and primers BK505 and BK506 (SEQ ID NOs:294 and 295). The URA3 portion of each primer was derived from the 5' region upstream of the URA3 promoter and 3' region downstream of the coding region such that integration of the loxP-kanMX-loxP marker resulted in replacement of the URA3 coding region. The PCR product was transformed into CEN.PK 113-7D using standard genetic techniques (Methods in Yeast Genetics, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants were selected on

2017200439 20 Jan 2017

YPD containing G418 (100 µg/ml) at 30 °C. Transformants were screened to verify correct integration by PCR using primers LA468 and LA492 (SEQ ID NOs:296 and 297) and designated CEN.PK 113-7D Δura3::kanMX.

HIS3 Deletion

The four fragments for the PCR cassette for the scarless HIS3 deletion were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs; Ipswich, MA) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen; Valencia, CA). HIS3 Fragment A was amplified with primer oBP452 (SEQ ID NO:298) and primer oBP453 (SEQ ID NO:299), containing a 5' tail with homology to the 5' end of HIS3 Fragment B. HIS3 Fragment B was amplified with primer oBP454 (SEQ ID NO:300), containing a 5' tail with homology to the 3' end of HIS3 Fragment A, and primer oBP455 (SEQ ID NO:301), containing a 5' tail with homology to the 5' end of HIS3 Fragment U. HIS3 Fragment U was amplified with primer oBP456 (SEQ ID NO:302), containing a 5' tail with homology to the 3' end of HIS3 Fragment B, and primer oBP457 (SEQ ID NO:303), containing a 5' tail with homology to the 5' end of HIS3 Fragment C. HIS3 Fragment C was amplified with primer oBP458 (SEQ ID NO:304), containing a 5' tail with homology to the 3' end of HIS3 Fragment U, and primer oBP459 (SEQ ID NO:305). PCR products were purified with a PCR Purification kit (Qiagen). HIS3 Fragment AB was created by overlapping PCR by mixing HIS3 Fragment A and HIS3 Fragment B and amplifying with primers oBP452 (SEQ ID NO:298) and oBP455 (SEQ ID NO:301). HIS3 Fragment UC was created by overlapping PCR by mixing HIS3 Fragment U and HIS3 Fragment C and amplifying with primers oBP456 (SEQ ID NO:302) and oBP459 (SEQ ID NO:305). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The HIS3 ABUC cassette was created by overlapping PCR by mixing HIS3 Fragment AB and HIS3 Fragment UC and amplifying with primers oBP452 (SEQ ID NO:298) and oBP459 (SEQ ID NO:305). The PCR product was purified with a PCR Purification kit (Qiagen).

2017200439 20 Jan 2017

Competent cells of CEN.PK 113-7D Δura3::kanMX were made and transformed with the HIS3 ABUC PCR cassette using a Frozen-EZ Yeast Transformation II kit (Zymo Research; Orange, CA). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 2% glucose at 30 °C. Transformants with a his3 knockout were screened for by PCR with primers oBP460 (SEQ ID NO:306) and oBP461 (SEQ ID NO:307) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). A correct transformant was selected as strain CEN.PK 113-7D Δura3::kanMX Δhis3::URA3.

KanMX Marker Removal from the Δura3 Site and URA3 Marker Removal from the Δhis3 Site

The KanMX marker was removed by transforming CEN.PK 113-7D Δura3::kanMX Δhis3::URA3 with pRS423::PGAL1-cre (SEQ ID NO: 271,) using a Frozen-EZ Yeast Transformation II kit (Zymo Research) and plating on synthetic complete medium lacking histidine and uracil supplemented with 2% glucose at 30 °C. Transformants were grown in YP supplemented with 1% galactose at 30 °C for ~6 hours to induce the Cre recombinase and KanMX marker excision and plated onto YPD (2% glucose) plates at 30 °C for recovery. An isolate was grown overnight in YPD and plated on synthetic complete medium containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the URA3 marker. 5-FOA resistant isolates were grown in and plated on YPD for removal of the pRS423::PGAL1-cre plasmid. Isolates were checked for loss of the KanMX marker, URA3 marker, and pRS423::PGAL1-cre plasmid by assaying growth on YPD+G418 plates, synthetic complete medium lacking uracil plates, and synthetic complete medium lacking histidine plates. A correct isolate that was sensitive to G418 and auxotrophic for uracil and histidine was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 and designated as BP857. The deletions and marker removal were confirmed by PCR and sequencing with primers oBP450 (SEQ ID NO:308) and oBP451 (SEQ ID NO:309) for Δura3 and primers oBP460 (SEQ ID NO:306) and oBP461 (SEQ ID NO:307) for Δhis3 using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen).

2017200439 20 Jan 2017

PDC6 Deletion

The four fragments for the PCR cassette for the scarless PDC6 deletion were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). PDC6 Fragment A was amplified with primer oBP440 (SEQ ID NO:310) and primer oBP441 (SEQ ID NO:311), containing a 5' tail with homology to the 5' end of PDC6 Fragment B. PDC6 Fragment B was amplified with primer oBP442 (SEQ ID NO:312), containing a 5' tail with homology to the 3" end of PDC6 Fragment A, and primer oBP443 (SEQ ID NO:313), containing a 5' tail with homology to the 5' end of PDC6 Fragment U. PDC6 Fragment U was amplified with primer oBP444 (SEQ ID NO:314), containing a 5' tail with homology to the 3' end of PDC6 Fragment B, and primer oBP445 (SEQ ID NO:315), containing a 5' tail with homology to the 5' end of PDC6 Fragment C. PDC6 Fragment C was amplified with primer oBP446 (SEQ ID NO:316), containing a 5' tail with homology to the 3' end of PDC6 Fragment U, and primer oBP447 (SEQ ID NO:317). PCR products were purified with a PCR Purification kit (Qiagen). PDC6 Fragment AB was created by overlapping PCR by mixing PDC6 Fragment A and PDC6 Fragment B and amplifying with primers oBP440 (SEQ ID NO:310) and oBP443 (SEQ ID NO:313). PDC6 Fragment UC was created by overlapping PCR by mixing PDC6 Fragment U and PDC6 Fragment C and amplifying with primers oBP444 (SEQ ID NO:314) and oBP447 (SEQ ID NO:317). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The PDC6 ABUC cassette was created by overlapping PCR by mixing PDC6 Fragment AB and PDC6 Fragment UC and amplifying with primers oBP440 (SEQ ID NO:310) and oBP447 (SEQ ID NO:317). The PCR product was purified with a PCR Purification kit (Qiagen).

Competent cells of CEN.PK 113-7D Δura3::loxP Δhis3 were made and transformed with the PDC6 ABUC PCR cassette using a Frozen-EZ Yeast Transformation II kit (Zymo Research). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with

2017200439 20 Jan 2017

2% glucose at 30 °C. Transformants with a pdc6 knockout were screened for by PCR with primers oBP448 (SEQ ID NO:318) and oBP449 (SEQ ID NO:319) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). A correct transformant was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6::URA3.

CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6::URA3 was grown overnight in YPD and plated on synthetic complete medium containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the URA3 marker. The deletion and marker removal were confirmed by PCR and sequencing with primers oBP448 (SEQ ID NO:318) and oBP449 (SEQ ID NO:319) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The absence of the PDC6 gene from the isolate was demonstrated by a negative PCR result using primers specific for the coding sequence of PDC6, oBP554 (SEQ ID NO:320) and oBP555 (SEQ ID NO:321). The correct isolate was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 and designated as BP891.

PDC1 Deletion ilvDSm Integration

The PDC1 gene was deleted and replaced with the ilvD coding region from *Streptococcus mutans* ATCC #700610. The A fragment followed by the ilvD coding region from *Streptococcus mutans* for the PCR cassette for the PDC1 deletion-ilvDSm integration was amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs) and NYLA83 genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). NYLA83 is a strain which carries the PDC1 deletion-ilvDSm integration described in U.S. Patent Application Publication No. 2009/0305363, which is herein incorporated by reference in its entirety. PDC1 Fragment A-ilvDSm (SEQ ID NO:322) was amplified with primer oBP513 (SEQ ID NO:326) and primer oBP515 (SEQ ID NO:327), containing a 5' tail with homology to the 5' end of PDC1 Fragment B. The B, U, and C fragments for the PCR cassette for the PDC1 deletion-ilvDSm integration were amplified using Phusion High Fidelity PCR Master Mix (New England BioLabs) and CEN.PK 113-7D genomic DNA as template, prepared with a Gentra Puregene Yeast/Bact

2017200439 20 Jan 2017

kit (Qiagen). PDC1 Fragment B was amplified with primer oBP516 (SEQ ID NO:328) containing a 5' tail with homology to the 3' end of PDC1 Fragment A-ilvDSm, and primer oBP517 (SEQ ID NO:329), containing a 5' tail with homology to the 5' end of PDC1 Fragment U. PDC1 Fragment U was amplified with primer oBP518 (SEQ ID NO:330), containing a 5' tail with homology to the 3' end of PDC1 Fragment B, and primer oBP519 (SEQ ID NO:331), containing a 5' tail with homology to the 5' end of PDC1 Fragment C. PDC1 Fragment C was amplified with primer oBP520 (SEQ ID NO:332), containing a 5' tail with homology to the 3' end of PDC1 Fragment U, and primer oBP521 (SEQ ID NO:333). PCR products were purified with a PCR Purification kit (Qiagen). PDC1 Fragment A-ilvDSm-B was created by overlapping PCR by mixing PDC1 Fragment A-ilvDSm and PDC1 Fragment B and amplifying with primers oBP513 (SEQ ID NO:326) and oBP517 (SEQ ID NO:329). PDC1 Fragment UC was created by overlapping PCR by mixing PDC1 Fragment U and PDC1 Fragment C and amplifying with primers oBP518 (SEQ ID NO:330) and oBP521 (SEQ ID NO:333). The resulting PCR products were purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The PDC1 A-ilvDSm-BUC cassette (SEQ ID NO:323) was created by overlapping PCR by mixing PDC1 Fragment A-ilvDSm-B and PDC1 Fragment UC and amplifying with primers oBP513 (SEQ ID NO:326) and oBP521 (SEQ ID NO:333). The PCR product was purified with a PCR Purification kit (Qiagen).

Competent cells of CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 were made and transformed with the PDC1 A-ilvDSm-BUC PCR cassette using a Frozen-EZ Yeast Transformation II kit (Zymo Research). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 2% glucose at 30 °C. Transformants with a pdc1 knockout ilvDSm integration were screened for by PCR with primers oBP511 (SEQ ID NO:336) and oBP512 (SEQ ID NO:337) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The absence of the PDC1 gene from the isolate was demonstrated by a negative PCR result using primers specific for the coding sequence of PDC1, oBP550 (SEQ ID NO:338) and oBP551 (SEQ ID NO:339). A

2017200439 20 Jan 2017

correct transformant was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm-URA3.

CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm-URA3 was grown overnight in YPD and plated on synthetic complete medium containing 5-fluoro-orotic acid (0.1%) at 30 °C to select for isolates that lost the URA3 marker. The deletion of PDC1, integration of ilvDSm, and marker removal were confirmed by PCR and sequencing with primers oBP511 (SEQ ID NO:336) and oBP512 (SEQ ID NO:337) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The correct isolate was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm and designated as BP907.

PDC5 Deletion sadB Integration

The PDC5 gene was deleted and replaced with the sadB coding region from *Achromobacter xylosoxidans* (the sadB gene is described in U.S. Patent Appl. No. 2009/0269823, which is herein incorporated by reference in its entirety). A segment of the PCR cassette for the PDC5 deletion-sadB integration was first cloned into plasmid pUC19-URA3MCS.

pUC19-URA3MCS is pUC19 based and contains the sequence of the URA3 gene from *Saccharomyces cerevisiae* situated within a multiple cloning site (MCS). pUC19 contains the pMB1 replicon and a gene coding for beta-lactamase for replication and selection in *Escherichia coli*. In addition to the coding sequence for URA3, the sequences from upstream and downstream of this gene were included for expression of the URA3 gene in yeast. The vector can be used for cloning purposes and can be used as a yeast integration vector.

The DNA encompassing the URA3 coding region along with 250 bp upstream and 150 bp downstream of the URA3 coding region from *Saccharomyces cerevisiae* CEN.PK 113-7D genomic DNA was amplified with primers oBP438 (SEQ ID NO:334), containing BamHI, AscI, PmeI, and FseI restriction sites, and oBP439 (SEQ ID NO:335), containing XbaI, PacI, and NotI restriction sites, using Phusion High-Fidelity PCR Master Mix (New England BioLabs). Genomic DNA was prepared using a Gentra Puregene Yeast/Bact kit (Qiagen). The PCR product and pUC19 (SEQ ID

2017200439 20 Jan 2017

NO:325) were ligated with T4 DNA ligase after digestion with BamHI and XbaI to create vector pUC19-URA3MCS. The vector was confirmed by PCR and sequencing with primers oBP264 (SEQ ID NO:342) and oBP265 (SEQ ID NO:343).

The coding sequence of sadB and PDC5 Fragment B were cloned into pUC19-URA3MCS to create the sadB-BU portion of the PDC5 A-sadB-BUC PCR cassette. The coding sequence of sadB was amplified using pLH468-sadB (SEQ ID NO:359) as template with primer oBP530 (SEQ ID NO:344), containing an AscI restriction site, and primer oBP531 (SEQ ID NO:345), containing a 5' tail with homology to the 5' end of PDC5 Fragment B. PDC5 Fragment B was amplified with primer oBP532 (SEQ ID NO:346), containing a 5' tail with homology to the 3' end of sadB, and primer oBP533 (SEQ ID NO:347), containing a PmeI restriction site. PCR products were purified with a PCR Purification kit (Qiagen). sadB-PDC5 Fragment B was created by overlapping PCR by mixing the sadB and PDC5 Fragment B PCR products and amplifying with primers oBP530 (SEQ ID NO:344) and oBP533 (SEQ ID NO:347). The resulting PCR product was digested with AscI and PmeI and ligated with T4 DNA ligase into the corresponding sites of pUC19-URA3MCS after digestion with the appropriate enzymes. The resulting plasmid was used as a template for amplification of sadB-Fragment B-Fragment U using primers oBP536 (SEQ ID NO:348) and oBP546 (SEQ ID NO:349), containing a 5' tail with homology to the 5' end of PDC5 Fragment C. PDC5 Fragment C was amplified with primer oBP547 (SEQ ID NO:350) containing a 5' tail with homology to the 3' end of PDC5 sadB-Fragment B-Fragment U, and primer oBP539 (SEQ ID NO:351). PCR products were purified with a PCR Purification kit (Qiagen). PDC5 sadB-Fragment B-Fragment U-Fragment C was created by overlapping PCR by mixing PDC5 sadB-Fragment B-Fragment U and PDC5 Fragment C and amplifying with primers oBP536 (SEQ ID NO:348) and oBP539 (SEQ ID NO:351). The resulting PCR product was purified on an agarose gel followed by a Gel Extraction kit (Qiagen). The PDC5 A-sadB-BUC cassette (SEQ ID NO:324) was created by amplifying PDC5 sadB-Fragment B-Fragment U-Fragment C with

2017200439 20 Jan 2017

primers oBP542 (SEQ ID NO:352), containing a 5' tail with homology to the 50 nucleotides immediately upstream of the native PDC5 coding sequence, and oBP539 (SEQ ID NO:351). The PCR product was purified with a PCR Purification kit (Qiagen).

Competent cells of CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm were made and transformed with the PDC5 A-sadB-BUC PCR cassette using a Frozen-EZ Yeast Transformation II kit (Zymo Research). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 1% ethanol (no glucose) at 30C. Transformants with a pdc5 knockout sadB integration were screened for by PCR with primers oBP540 (SEQ ID NO:353) and oBP541 (SEQ ID NO:354) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The absence of the PDC5 gene from the isolate was demonstrated by a negative PCR result using primers specific for the coding sequence of PDC5, oBP552 (SEQ ID NO:355) and oBP553 (SEQ ID NO:356). A correct transformant was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB-URA3.

CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB-URA3 was grown overnight in YPE (1% ethanol) and plated on synthetic complete medium supplemented with ethanol (no glucose) and containing 5-fluoro-orotic acid (0.1%) at 30 C to select for isolates that lost the URA3 marker. The deletion of PDC5, integration of sadB, and marker removal were confirmed by PCR with primers oBP540 (SEQ ID NO:353) and oBP541 (SEQ ID NO:354) using genomic DNA prepared with a Gentra Puregene Yeast/Bact kit (Qiagen). The correct isolate was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB and designated as BP913.

GPD2 Deletion

To delete the endogenous GPD2 coding region, a gpd2::loxP-URA3-loxP cassette (SEQ ID NO:361) was PCR-amplified using loxP-URA3-loxP PCR (SEQ ID NO:360) as template DNA. loxP-URA3-loxP contains the URA3 marker from (ATCC # 77107) flanked by loxP recombinase sites. PCR was done using Phusion DNA polymerase and

2017200439 20 Jan 2017

primers LA512 and LA513 (SEQ ID NOs:340 and 341). The GPD2 portion of each primer was derived from the 5' region upstream of the GPD2 coding region and 3' region downstream of the coding region such that integration of the loxP-URA3-loxP marker resulted in replacement of the GPD2 coding region. The PCR product was transformed into BP913 and transformants were selected on synthetic complete media lacking uracil supplemented with 1% ethanol (no glucose). Transformants were screened to verify correct integration by PCR using primers oBP582 and AA270 (SEQ ID NOs:357 and 358).

The URA3 marker was recycled by transformation with pRS423::PGAL1-cre (SEQ ID NO:271) and plating on synthetic complete media lacking histidine supplemented with 1% ethanol at 30 C. Transformants were streaked on synthetic complete medium supplemented with 1% ethanol and containing 5-fluoro-orotic acid (0.1%) and incubated at 30 C to select for isolates that lost the URA3 marker. 5-FOA resistant isolates were grown in YPE (1% ethanol) for removal of the pRS423::PGAL1-cre plasmid. The deletion and marker removal were confirmed by PCR with primers oBP582 (SEQ ID NO:357) and oBP591 (SEQ ID NO:362). The correct isolate was selected as strain CEN.PK 113-7D Δura3::loxP Δhis3 Δpdc6 Δpdc1::ilvDSm Δpdc5::sadB Δgpd2::loxP and designated as BP1064 (PNY1503).

EXAMPLE 13

Construction of PNY2204 and Isobutanol Pathway Plasmids

The purpose of this example is to describe construction of a vector to enable integration of a gene encoding acetolactate synthase into the naturally occurring intergenic region between the PDC1 and TRX1 coding sequences in Chromosome XII. Strains resulting from the use of this vector are also described.

Construction of integration vector pUC19-kan::pdc1::FBA-alsS::TRX1

The FBA-alsS-CYCt cassette was constructed by moving the 1.7kb BbvCI/PacI fragment from pRS426::GPD::alsS::CYC (described in U.S.

2017200439 20 Jan 2017

Patent No. 7,851,188, which is herein incorporated by reference in its entirety) to pRS426::FBA::ILV5::CYC (described in U.S. Patent No. 7,851,188, which is herein incorporated by reference in its entirety), which had been previously digested with BbvCI/PacI to release the ILV5 gene. Ligation reactions were transformed into *E. coli* TOP10 cells and transformants were screened by PCR using primers N98SeqF1 (SEQ ID NO:363) and N99SeqR2 (SEQ ID NO:365). The FBA-alsS-CYCt cassette was isolated from the vector using BglII and NotI for cloning into pUC19-URA3::ilvD-TRX1 (clone "B") at the AflII site (Klenow fragment was used to make ends compatible for ligation). Transformants containing the alsS cassette in both orientations in the vector were obtained and confirmed by PCR using primers N98SeqF4 (SEQ ID NO:364) and N1111 (SEQ ID NO:366) for configuration "A" and N98SeqF4 (SEQ ID NO:364) and N1110 (SEQ ID NO:367) for configuration "B". A geneticin selectable version of the "A" configuration vector was then made by removing the URA3 gene (1.2 kb *Not*I/*Nae*I fragment) and adding a geneticin cassette. Klenow fragment was used to make all ends compatible for ligation, and transformants were screened by PCR to select a clone with the geneticin resistance gene in the same orientation as the previous URA3 marker using primers BK468 (SEQ ID NO:368) and N160SeqF5 (SEQ ID NO:210). The resulting clone was called pUC19-kan::pdc1::FBA-alsS::TRX1 (clone A)(SEQ ID NO:387).

Construction of *alsS* integrant strains and isobutanol-producing derivatives

The pUC19-kan::pdc1::FBA-alsS integration vector described above was linearized with *Pme*I and transformed into PNY1507 (Example 8). *Pme*I cuts the vector within the cloned pdc1-TRX1 intergenic region and thus leads to targeted integration at that location (Rodney Rothstein, Methods in Enzymology, 1991, volume 194, pp. 281-301). Transformants were selected on YPE plus 50 μg/ml G418. Patched transformants were screened by PCR for the integration event using primers N160SeqF5 (SEQ ID NO:210) and oBP512 (SEQ ID NO:337). Two transformants were tested indirectly for acetolactate synthase function by evaluating the strains ability to make isobutanol. To do this, additional isobutanol

2017200439 20 Jan 2017

pathway genes were supplied on *E. coli*-yeast shuttle vectors (pYZ090ΔalsS and pBP915, described below). One clone was designated as PNY2205. The plasmid-free parent strain was designated PNY2204 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-pUC19-loxP-kanMX-loxP-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ adh1Δ::UAS(PGK1)P[FBA1]-kivD_LI(y)-ADH1t).

Isobutanol pathway plasmids (pYZ090ΔalsS and pBP915)

pYZ090 (SEQ ID NO:195) was digested with *Spe*I and *Not*I to remove most of the CUP1 promoter and all of the *alsS* coding sequence and CYC terminator. The vector was then self-ligated after treatment with Klenow fragment and transformed into *E. coli* Stbl3 cells, selecting for ampicillin resistance. Removal of the DNA region was confirmed for two independent clones by DNA sequencing across the ligation junction by PCR using primer N191 (SEQ ID NO:370). The resulting plasmid was named pYZ090ΔalsS (SEQ ID NO:371). The pLH468 plasmid was constructed for expression of DHAD, KivD and HADH in yeast. pBP915 (SEQ ID NO: 182) was constructed from pLH468 (SEQ ID NO:139) by deleting the *kivD* gene and 957 base pairs of the *TDH3* promoter upstream of *kivD*. pLH468 was digested with SwaI and the large fragment (12896 bp) was purified on an agarose gel followed by a Gel Extraction kit (Qiagen; Valencia, CA). The isolated fragment of DNA was self-ligated with T4 DNA ligase and used to transform electrocompetent TOP10 *Escherichia coli* (Invitrogen; Carlsbad, CA). Plasmids from transformants were isolated and checked for the proper deletion by restriction analysis with the SwaI restriction enzyme. Isolates were also sequenced across the deletion site with primers oBP556 (SEQ ID NO:372) and oBP561 (SEQ ID NO:373). A clone with the proper deletion was designated pBP915 (pLH468ΔkivD)(SEQ ID NO:182).

pYZ090 is based on the pHR81 (ATCC #87541, Manassas, VA) backbone. pYZ090 was constructed to contain a chimeric gene having the coding region of the alsS gene from *Bacillus subtilis* (nt position 457-2172)

20 Jan 2017 2017200439

expressed from the yeast CUP1 promoter (nt 2-449) and followed by the CYC1 terminator (nt 2181-2430) for expression of ALS, and a chimeric gene having the coding region of the ilvC gene from *Lactococcus lactis* (nt 3634-4656) expressed from the yeast ILV5 promoter (2433-3626) and followed by the ILV5 terminator (nt 4682-5304) for expression of KARI.

Example 14.

Isobutanol production-PNY1910 and PNY2242

Methods:

Preparation of inoculum medium

1 L of inoculum medium contained: 6.7 g, Yeast Nitrogen Base w/o amino acids (Difco 0919-15-3); 2.8 g, Yeast Synthetic Drop-out Medium Supplement Without Histidine, Leucine, Tryptophan and Uracil (Sigma Y2001); 20 mL of 1% (w/v) L-Leucine; 4 mL of 1% (w/v) L-Tryptophan; 3 g of ethanol; 10 g of glucose.

Preparation of defined fermentation medium

The volume of broth after inoculation was 800 mL, with the following final composition, per liter: 5 g ammonium sulfate, 2.8 g potassium phosphate monobasic, 1.9 g magnesium sulfate septahydrate, 0.2 mL antifoam (Sigma DF204), Yeast Synthetic Drop-out Medium Supplement without Histidine, Leucine, Tryptophan, and Uracil (Sigma Y2001), 16 mg L-leucine, 4 mg L-tryptophan, 6 mL of a vitamin mixture (in 1 L water, 50 mg biotin, 1 g Ca-pantothenate, 1 g nicotinic acid, 25 g myo-inositol, 1 g thiamine chloride hydrochloride, 1 g pyridoxol hydrochloride, 0.2 g p-aminobenzoic acid) 6 mL of a trace mineral solution (in 1 L water, 15 g EDTA, 4.5 g zinc sulfate heptahydrate, 0.8 g manganese chloride dehydrate, 0.3 g cobalt chloride hexahydrate, 0.3 g copper sulfate pentahydrate, 0.4 g disodium molybdenum dehydrate, 4.5 g calcium chloride dihydrate, 3 g iron sulfate heptahydrate, 1 g boric acid, 0.1 g potassium iodide), 30 mg thiamine HCl, 30 mg nicotinic acid. The pH was adjusted to 5.2 with 2N KOH and glucose added to 10 g/L.

2017200439 20 Jan 2017

Preparation of inoculum

A 125 mL shake flask was inoculated directly from a frozen vial by pipetting the whole vial culture (approx. 1 ml) into 10 mL of the inoculum medium. The flask was incubated at 260 rpm and 30 °C. The strain was grown overnight until OD about 1.0. OD at λ = 600 nm was determined in a Beckman spectrophotometer (Beckman, USA).

Bioreactor experimental design

Fermentations were carried out in 1L Biostat B DCU3 fermenters (Sartorius, USA) with a working volume on 0.8 L. Off-gas composition was monitored by a Prima DB mass spectrometer (Thermo Electron Corp., USA). The temperature was maintained at 30C and pH controlled at 5.2 with 2N KOH throughout the entire fermentation. Directly after inoculation with 80 mL of the inoculum, dO was controlled by agitation at 30%, pH was controlled at 5.25, aeration was controlled at 0.2 L/min. Once OD of approximately 3 was reached, the gas was switched to N2 for anaerobic cultivation. Throughout the fermentation, glucose was maintained in excess (5-20 g/L) by manual additions of a 50% (w/w) solution.

Methods for analyzing cultivation experiments

OD at λ = 600 nm was determined in a spectrophotometer by pipetting a well mixed broth sample into a cuvette (CS500 VWR International, Germany). If biomass concentration of the sample exceeded the linear absorption range of the spectrophotometer (typically OD values from 0.000 to 0.300), the sample was diluted with 0.9% NaCl solution to yield values in the linear range.

Measurements of glucose, isobutanol, and other fermentation by-products in the culture supernatant were carried out by HPLC, using a Bio-Rad Aminex HPX-87H column (Bio-Rad, USA), with refractive index (RI) and a diode array (210 nm) detectors. Chromatographic separation was achieved using 0.01 N H_2SO_4 as the mobile phase with a flow rate of 0.6 mL/min and a column temperature of 40 °C. Isobutanol retention time is

2017200439 20 Jan 2017

32.2 minutes under these conditions. Isobutanol concentration in off-gas samples was determined by mass-spectrometer.

Results

Maximal biomass concentration measured as optical density (OD), volumetric rate of isobutanol production, final isobutanol titer, and isobutanol yield on glucose are presented in the table below. The strain PNY2242 had higher titers and faster rates than the strain PNY1910 and produced isobutanol with higher specific rate and titer. The specific rates are shown in Figure 5. Accumulation of the DHIV + DHMB in the culture supernatant was three times higher with PNY1910 compared to the PNY2242 strain (Figure 6). Yield of glycerol, pyruvic acid, BDO, DHIV+DHMB*, αKIV, and isobutyric acid on glucose is shown in Figure 7.

*DHIV analyzed by HPLC method includes both DHIV and DHMB.

Table 19

Strain	Max. OD600	Rate (g/L/h)	Titer (g/L)	Yield (g/g)
PNY1910	5.0	0.16	10.9	0.25
PNY2242	5.0	0.23	16.1	0.27

Example 15.
Construction of K9G9 error prone PCR library

Error prone PCR of K9G9 was performed to generate a library that can be screened for variants with increases in the Km values for NADPH relative to NADH. Mutagenic PCR of K9G9 was performed with the GeneMorph® II EZClone Domain Mutagenesis Kit (Catalog #200552; Agilent Technologies, Stratagene Products Division, La Jolla, CA). Primers K9G9_EZ_F1 (AAA CAT GGA AGA ATG TAA GAT GGC; SEQ ID NO: 390) and K9G9_EZ_R1 (TCA GTT GTT AAT CAA CTT GTC TTC G; SEQ ID NO: 391) were commercially synthesized by Integrated DNA Technologies, Inc (Coralville IA). Other than the primers, template, and ddH_2O, reagents used here were supplied with the kit indicated above.

2017200439 20 Jan 2017

The mutagenic PCR mixture consisted of 4 µl of pHR81-PIIv5-KARI-K9.G9 (SEQ ID NO: 392) (770 ng/µg), 1.25 µl of each primer (100 ng/ µl stocks), 5 µl of 10× Mutazyme II reaction buffer, 1 µl of 40 mM dNTP mix, 1.5 µl of Mutazyme II DNA polymerase, and 36 µl of ddH_2O. The following conditions were used for the PCR reaction: The starting temperature was 95°C for 2.0 min followed by 30 heating/cooling cycles. Each cycle consisted of 95°C for 30 sec, 48°C for 30 sec, and 72°C for 2.0 min. At the completion of the temperature cycling, the sample was kept at 72° C for 10.0 min more, and then held awaiting sample recovery at 4°C. The reaction product was separated from the template via agarose gel electrophloresis (1% agarose, 1X TBE buffer) and recovered using the StrataPrep® DNA Gel Extraction Kit (Cat# 400766, Agilent Technologies, Stratagene Products Division, La Jolla, CA) as recommended by the manufacturer.

The isolated reaction product was employed as a megaprimer to generate gene libraries in the "EZClone reaction" of the kit indicated above. Other than the megaprimer, template, and ddH_2O, reagents used here were supplied with the kit indicated above.The reaction consisted of 25 µl of the 2× EZClone enzyme mix, 4 µl of megaprimer (125 ng/ µl), 2 µl of K9G9 in a pBAD.KARI vector (25 ng/ µl), 3 µl of EZClone solution, and 16µl of ddH_2O. The following conditions were used for the reaction: The starting temperature was 95°C for 1.0 min followed by 30 heating/cooling cycles. Each cycle consisted of 95°C for 50 sec, 60°C for 50 sec, and 68°C for 10.0 min. At the completion of the temperature cycling, the samples were kept at 72° C for 10.0 min more, and then held awaiting sample recovery at 4°C. 1 µl of the Dpn I (10 U/µl) was added and the mixture was incubated for 4 hours at 37°C.

4 µl of the Dpn I digested "EZClone reaction" product was then transformed into 50 µl XL10-Gold® Ultracompetent *E. coli* cells (provided in the GeneMorph® II EZClone Domain Mutagenesis Kit) as recommended by the manufacturer. The transformants were spread on agar plates containing the LB medium and 100 µg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA), incubated at 37°C overnight, and

2017200439 20 Jan 2017

store at 4°C. These steps were repeated with 4 µl Dpn I digested "EZClone reaction" product and 50 µl cells per transformation for a total of 10 transformations. The resultant library in XL-Gold was scraped off the agar plates with a solution containing M9 salts, combined, diluted into media containing the LB medium and 100 µg/ml ampicillin, and incubated at 37°C overnight. The library DNA was isolated from the cells with the QIAprep Spin Miniprep Kit (Catalog #2706; Qiagen, Valencia, CA) according to the protocol provided by the manufacturer. The amplified library was then used to transform an electro-competent strain of *E. coli* Bw25113 (ΔilvC) using a BioRad Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). The transformed clones were spread on agar plates containing the LB medium and 100 µg/ml ampicillin (#101320-154, Teknova Inc. Hollister, CA) and incubated at 37°C overnight. Clones were employed for high throughput screening as described in Example 16.

Example 16.

Identification of K9G9 variants with increased K_M for NADPH via screening for diminished NADP+ inhibition of NADH activity

The K9G9 library described in Example 15 was screened for variants with reduced $NADP^+$ inhibition of NADH-dependent KARI activity. A K9G9 variant with reduced $NADP^+$ inhibition of activity with NADH can potentially exhibit an increase in the ratio of the K_M for NADPH to the K_M for NADH. With a specific objective to increase K_M for NADPH relative to Km for NADH, the hits from the screen were partially purified and kinetic analyses were performed to determine V_{max} and K_M parameters with NADH and with NADPH.

High throughput screening assay of K9G9 gene library

High throughput screening of the gene libraries of mutant KARI enzymes was performed as described herein: 10× freezing medium containing 554.4 g/L glycerol, 68 mM of $(NH_4)_2SO4$, 4 mM $MgSO_4$, 17 mM sodium citrate, 132 mM KH_2PO_4, 36 mM K_2HPO_4 was prepared with molecular pure water and filter-sterilized. Freezing medium was prepared by diluting the 10x freezing medium with the LB medium. An aliquot (200

2017200439 20 Jan 2017

μL) of the 1x freezing medium was used for each well of the 96-well archive plates (cat #3370, Corning Inc. Corning, NY).

Clones from the LB agar plates were selected and inoculated into the 96-well archive plates containing the freezing medium and grown overnight at 37 oC without shaking. The archive plates were then stored at -80°C. *E. coli* strain Bw25113(ΔilvC), as described in U.S. Patent 8,129,162, transformed with *pBAD*-HisB (Invitrogen) was always used as the negative control. The positive control for the library was K9G9-KARI in *E. coli* strain Bw25113 (ΔilvC), as described in U.S. Patent 8,129,162.

Clones from archive plates were inoculated into the 96-deep well plates. Each well contained 3.0 μl of cells from thawed archive plates, 200 μl of the LB medium containing 100 μg/ml ampicillin and 0.02%(w/v) arabinose as the inducer. Cells were the grown overnight at 37°C with 80% humidity while shaking (900 rpm), harvested by centrifugation (3750 rpm, 5 min at 25°C). (Eppendorf centrifuge, Brinkmann Instruments, Inc. Westbury, NY) and the cell pellet was stored at -20°C for later analysis.

The assay substrate, (*R*,*S*)-acetolactate, was synthesized as described by Aulabaugh and Schloss (Aulabaugh and Schloss, Biochemistry, 29: 2824-2830, 1990). All other chemicals used in the assay were purchased from Sigma. The enzymatic conversion of acetolactate to α,β-dihydroxyisovalerate by KARI was followed by measuring the oxidation of the cofactor, NADH, from the reaction at 340 nm using a plate reader (Saphire 2, Tecan, Mannedorf, Switzerland). The activity was calculated using the molar extinction coefficient of 6220 $M^{-1}cm^{-1}$ NADH.

Frozen cell pellet in deep-well plates and BugBuster (Novagen 71456, Darmstadt, Germany) were warmed up at room temperature for 30 min at the same time. 75 μl of 50% BugBuster (v/v in water) was added to each well after 30 min warm-up and cells were suspended using plate shaker. The plates with cell pellet/50% Bug Buster suspension were incubated at room temperature for 30 min. Cell lysate diluted with 75 μL d.d water, resulting in 0.5X lysate. Assays of the diluted cell free extracts were performed at 30 °C in buffer containing 2.4 mM (R/S)-acetolactate,

2017200439 20 Jan 2017

100 mM HEPES pH 6.8, 100 mM KCl, 10 mM $MgCl_2$, 150 µM NADH, 12.5 µL 0.5 X cell lysate with or without 2.5 mM NADP+.

Identification of K9G9 variants with reduced NADP+ inhibition of NADH KARI activity

The ratio for the measured rate of NADH oxidation in the presence of NADP+ to the measured rate of NADH oxidation in the absence of NADP+ was calculated for each variant and positive control well (2 per plate). The mean and standard deviation of ratios for all of the positive control wells (104 total) were calculated.

A variant well was considered to contain an initial hit if the rate in the absence of NADP+ was greater than 0.1 OD/hr and the rate ratio was both greater than 0.45 (three standard deviations higher than the positive control mean) and less than 1. A total of 521 hits were identified from a pool of 4607 potential variants. These initial hits were consolidated, forming a smaller library for further analysis.

Secondary screening of initial library hit

The consolidated hit library was grown in biological triplicate and cell free extracts were prepared and assayed as described above. Rate ratios were then calculated for the variants and positive controls as above. Final hits that were selected for detailed kinetic analysis met the following criteria: the rate in the absence of NADP+ was greater than 0.6 OD/hr, rate ratio was greater than 0.51 and less than 1, and at least two out of three biological replicates passed the criteria. Seventeen hits were identified for kinetic analysis and streaked out on to LB plates with 100 µg/mL ampicillin added.

Sequence analysis of K9G9 variants

DNA sequencing of the seventeen variants identified from the secondary HTS screening was accomplished by using TempliPhi™ (GE Healthcare) with the primers pBAD-For (ATGCCATAGCATTTTTATCC; SEQ ID NO: 393) and pBAD-Rev (CTGATTTAATCTGTATCAGGCT; SEQ ID NO: 394).

2017200439 20 Jan 2017

Table 20. Amino Acid Substitutions for K9G9 Variants

Variant	Seq	Amino Acid Substitutions
878 C1	873	None identified
879 A7	874	K90M
879 C2	875	H37Q
880 A11	876	A182T, P320Q
880 B4	877	K57E
880 D11	878	K90M, A174V
881 A2	879	K90M, I133V, K282T
881 G3	880	Y53F, E74G
881 G9	881	K90E
882 B12	882	H118R
882 C10	883	G31S, R61S, C121Y, D129N, G183D
882 C7	933	E54G
882 F9	934	K90E, Q160H
882 G6	935	G55A
882 G12	936	V142L, S285Y
883 C4	937	A170V
883 G9	938	L197M, K310M

Kinetic analysis of partially purified variant protein

E. coli strain Bw25113 (ΔilvC), as described in U.S. Patent 8,129,162, was used to express the seventeen variants and positive control K9G9. Strains were grown for 8 hours in 10 mL of LB broth (#46-060-CM, Mediatech, Manassas, VA) containing 100 µg/mL ampicillin at 37°C with shaking in 125 mL baffled, vented filtered lid flasks. 200 µL of this culture was used to inoculate 100 mL LB broth with 100 µg/mL ampicillin and 0.2% (w/v) arabinose added. These cultures were grown for 16 to 18 hours at 37°C with shaking in 500 mL baffled, vented filtered lid flasks. Cells were harvested in a 20 mL and two 40 mL aliquots, supernatants were decanted and the pellets were frozen at -80°C.

To partially purify the protein, the cell pellet corresponding with the 20 mL cell culture harvest was thawed and resuspended in 1 mL Bug Buster Master Mix (Novagen 71456, Darmstadt, Germany). The cell suspension was incubated at room temperature for 15 minutes followed by 15 minute incubation at 60°C to denature the heat liable proteins. Cell debris and denatured proteins were pelleted by centrifugation for 30

2017200439 20 Jan 2017

minutes at 4°C. Supernatant containing the heat stable cytosolic protein, including K9G9 and variants, was recovered and stored at 4°C.

The total protein of the heat stable cytosolic protein fraction was measured by the Bradford Assay using Coomaisse Plus (Thermo Scientific #23238, Rockford, Ill.). BSA was employed as the standard. The concentration of protein was measured by determining the absorbance at 595 nm using a Cary 300 spectrophotometer (Agilent Technologies, Wilmington, DE).

To determine V_{max} and K_M values for NADH and NADPH, the partially purified proteins were assayed at various concentrations of NADH (0, 16.4, 32.8, 65.7, 98.5, 164.3 and 246.5 µM) and NADPH (0, 12.8, 25.6, 51.2, 76.8 and 128 µM). Assays were conducted at 30°C in 100 mM HEPES (pH 6.8), 10 mM $MgCl_2$, 100 mM KCl and 4.8 mM R/*S*-acetolactate. Between 0.1 to 0.35 mg/mL total protein was added to the assay. The rate of conversion of *S*-acetolactate to DHIV was measured via monitoring the oxidation of NAD(P)H at 340 nm using a Cary 300 spectrophotometer (Agilent Technologies, Wilmington, DE). The activity was calculated using the molar extinction coefficient of 6220 $M^{-1}cm^{-1}$. V_{max} and K_m values were calculated by plotting specific activity (U/mg) vs. cofactor concentration and the data were fit to the Michaelis-Menten equation using Kaleidagraph software (Synergy, Reading, PA).

Table 21. Kinetic Values for Partially Purified K9G9 Variants as Determined via NAD(P)H Consumption Assays

Variant	V_{max} NADPH, U/mg	K_m NADPH, µM	V_{max}/K_m NADPH, L/min*mg	V_{max} NADH, U/mg	K_m NADH, µM	V_{max}/K_m NADH, L/min*mg
K9G9	1.53	45.5	0.034	1.09	67.4	0.016
878 C1	0.75	42.2	0.018	0.62	107.8	0.006
879 A7	2.51	546	0.005	1.44	263	0.006
879 C2	1.27	103	0.012	1.23	187	0.007
880 A11	0.72	86.9	0.008	0.51	117	0.004
880 B4	1.23	233	0.005	1.14	133	0.009
880 D11	1.38	130	0.011	1.50	232	0.006
881 A2	0.88	93.5	0.009	1.13	166.8	0.007
881 G3	0.69	99.2	0.007	0.69	61.8	0.011
881 G9	1.03	158	0.007	0.96	310	0.003

2017200439 20 Jan 2017

882 B12	0.87	30.3	0.029	0.49	78.9	0.006
882 C10	0.71	34.1	0.021	0.56	97.9	0.006
882 C7	1.62	45.3	0.036	0.96	75.6	0.013
882 F9	1.39	256	0.005	1.19	335	0.004
882 G6	0.95	47.4	0.020	0.74	98.7	0.007
882 G12	1.06	63.5	0.017	0.75	81.2	0.009
883 C4	1.26	46.8	0.027	0.67	83.9	0.008
883 G9	1.26	38	0.033	1.01	71.9	0.014

Example 17.

Manual recombination of K9 KARI variants via site directed mutagenesis

Site directed mutagenesis of the K9G9 derivatives K9JB4 and K9JG3 (identified in Example 16 as 880 B4 and 881 G3, respectively) was performed to incorporate other amino acid changes described in the examples. The initial step was to add to the N87P substitution, which is described in Example 5. Mutations were introduced into the KARI genes with primers N87PC1 (CTGACATCATTATGATCTTGATCCCAGATGAAAAGCAGGCTACCATG TAC; SEQ ID NO: 395) and N87PC1r (GTACATGGTAGCCTGCTTTTCATCTGGGATCAAGATCATAATGATGT CAG; SEQ ID NO: 396), employing the QuikChange® II Site-Directed Mutagenesis Kit (Catalog #200523; Agilent Technologies, Stratagene Products Division,La Jolla, CA). Except for the primers, templates, and ddH_2O, all reagents used here were supplied with the kit indicated above. Primers were commercially synthesized by Integrated DNA Technologies, Inc (Coralville IA). Templates were K9 KARI variants in *E. coli* vectors (pBAD.KARI). For mutagenesis of K9JB4, the reaction mixture contained 1 µl K9JB4 (50 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1 µl of Pfu Ultra HF DNA polymerase, and 40 µl of ddH_2O. For, the K9JG3 reaction mixture, 1 µl K9JB4 (50 ng/ul) was substituted with 1 µl K9JG3 (50 ng/µl). The following conditions were used for both reactions: The starting temperature was 95°C for 30 sec followed by 16 heating/cooling cycles. Each cycle consisted of 95°C for 30 sec, 55°C for 30 sec, and 68°C for 5.0 min. At the completion of the

2017200439 20 Jan 2017

temperature cycling, the samples held awaiting sample recovery at 4°C. 1 µl of the Dpn I (10 U/µl) was added to each reaction and the mixtures were incubated for 1 hour at 37°C.

2 µl of each mutagenic reaction was transformed into One Shot® TOP10 Chemically Competent *E. coli* (Invitrogen, Catalog #C404003) according to the manufacturer's instructions. The transformants were spread on agar plates containing the LB medium and 100 µg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA) and incubated at 37°C overnight. Multiple transformants were then selected for TempliPhi™ (GE Healthcare) based DNA sequencing employing primers pBAD-For (ATGCCATAGCATTTTTATCC; SEQ ID NO: 393) and pBAD-Rev (CTGATTTAATCTGTATCAGGCT; SEQ ID NO: 394). Transformants with confirmed KARI sequences were inoculated into LB medium containing 100 µg/ml ampicillin and incubated at 33°C with shaking at 225 rpm. Plasmid DNA was isolated from the cells with the QIAprep Spin Miniprep Kit (Catalog #2706; Qiagen, Valencia, CA) according to the protocol provided by the manufacturer. The resultant clones K9JB4P and K9JG3P were derived from K9JB4 and K9JG3, respectively.

Additional site directed mutagenesis was performed as described above with modifications.

Variant K9JA1 was derived from K9JG3P employing primers oK57E1 (GGTTTATTCGAAGGTGCGGAGGAGTGGAAAAGAGCTG; SEQ ID NO: 397) and oK57E1r (CAGCTCTTTTCCACTCCTCCGCACCTTCGAATAAACC; SEQ ID NO: 398). The mutagenesis reaction contained 1 µl K9JG3P (50 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1 µl of PfuUltra HF DNA polymerase, and 40 µl of ddH_2O. Liquid cultures for *E. coli* transformants were incubated at 37°C instead of 33°C.

Variant K9SB2 was derived from K9JB4P employing primers oY53F1 (GTAACGTTATCATTGGTTTATACGAAGGTGCGGAGGAG; SEQ ID NO: 399) and oY53F1r

2017200439 20 Jan 2017

(CTCCTCCGCACCTTCGAATAAACCAATGATAACGTTAC; SEQ ID NO: 400). The mutagenesis reaction contained 1 µl K9JB4P (50 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1 µl of PfuUltra HF DNA polymerase, and 40 µl of ddH2O. Liquid cultures for *E. coli* transformants were incubated at 37 °C instead of 33°C.

Variant K9SB2-K90L was derived from K9SB2 employing primers oK90L1(GATCTTGATCCCAGATGAATTGCAGGCTACCATGTACAAAAA C; SEQ ID NO: 401) and oK90L1r (GTT TTT GTA CAT GGT AGC CTG CAA TTC ATC TGG GAT CAA GAT C; SEQ ID NO: 402). The mutagenesis reaction contained 2.5 µl K9SB2 (50 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1 µl of PfuUltra HF DNA polymerase, and 38.5 µl of ddH_2O. For the heating/cooling cycles, the step of 55°C for 30 sec was increased to 1 min. Liquid cultures for *E. coli* transformants were incubated at 37 °C instead of 33°C.

Variant K9SB2-K90M was derived from K9SB2 employing primers oK90M1 (CTTGATCCCAGATGAAATGCAGGCTACCATGTACAAAAAC; SEQ ID NO: 403) and oK90M1r (GTT TTT GTA CAT GGT AGC CTG CAT TTC ATC TGG GAT CAA G; SEQ ID NO: 404). The mutagenesis reaction contained 2.5 µl K9SB2 (50 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1 µl of PfuUltra HF DNA polymerase, and 38.5 µl of ddH_2O. For the heating/cooling cycles, the step of 55°C for 30 sec was increased to 1 min. Liquid cultures for *E. coli* transformants were incubated at 37 °C instead of 33°C.

Table 22. Amino Acid Substitutions of K9G9 Variants and Combinations

Variant	Amino Acid Seq ID No:	Nucleic Acid SEQ ID NO:	Amino Acid Substitutions
K9JB4	417	418	S56A, K57E, S58E
K9JB4P	419	420	S56A, K57E, S58E, N87P
K9JG3	421	422	Y53F, S56A, S58E, E74G

20 Jan 2017 2017200439

K9JG3P	423	424	Y53F, S56A, S58E, E74G, N87P
K9JA1	425	426	Y53F, S56A, K57E, S58E, E74G, N87P
K9SB2	427	428	Y53F, S56A, K57E, S58E, N87P
K9SB2-K90L	429	430	Y53F, S56A, K57E, S58E, N87P, K90L
K9SB2-K90M	431	432	Y53F, S56A, K57E, S58E, N87P, K90M

Example 18

Kinetic characterization of purified K9G9 derivatives with increased ratios of K_M NADPH to K_M NADH

K9G9 and variants were overexpressed in *E. coli* strain Bw25113 (*ΔilvC*), as described in U.S. Patent 8,129,162, and purified in order to obtain a more accurate determination of cofactor affinity and maximum velocity.

For expression and characterization, *E. coli* plasmids (pBAD.KARI) were used to transform an electro-competent strain of *E. coli* Bw25113 (*ΔilvC*) as described in U.S. Patent 8,129,162, using a BioRad Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). The transformed clones were spread on agar plates containing the LB medium and 100 µg/ml ampicillin (#101320-154, Teknova Inc. Hollister, CA) and incubated at 37°C overnight. A single transformant for each strain was streaked out onto LB plates with 100 µg/mL ampicillin. A single colony from each of these plates was used to inoculate 10 mL LB broth with 100 µg/mL ampicillin. These cultures were grown for 8 hours at 37°C with shaking in 125 mL baffled flasks with vented, filtered lids. 200 µL of this culture was used to inoculate two 500 mL baffled flasks with filtered vented lids containing LB broth with 100 µg/mL ampicillin and 0.2% (w/v) arabinose. The expression cultures were grown for 16 – 18 hours at 37°C with shaking. Cells were harvested in 40 mL aliquots via centrifugation; the supernatant was discarded and cell pellets were frozen at -80°C until purification.

K9G9 and all variants were purified using the same process. Two cell pellets, representing 40 mL cell culture aliquots each, were resuspended in 4 mL Bug Buster Master Mix (Novagen 71456, Darmstadt, Germany) and incubated for 15 minutes at room temperature followed by

2017200439 20 Jan 2017

15 minutes at 60°C. Denatured proteins and cell debris was pelleted by centrifugation at 7,000 rpm for 30 minutes and 4°C. The supernatant was decanted, save and filtered through a Acrodisc 0.2 µm syringe filter (PN4192, Pall, Ann Arbor, MI). K9G9 was purified from the filtered heat treated cell free extract using a GE Healthcare HiLoad 26/60 Superdex 200 gel filtration column (17-1071-01, Buckinghamshire, England). The column was pre-equilibrated with 0.2 CV equilibration with 50 mM HEPES (pH 7.5) 5 mM $MgCl_2$ buffer at a 2.0 mL/min flow rate prior to protein loading. K9G9 and variants were eluted over a 1.5 CV isocratic step consisting of 50 mM HEPES (pH 7.5) 5 mM $MgCl_2$ buffer at a 2.0 mL/min flow rate. Fractions 2.5 mL in volume were collected using a Frac-950 fraction collector (Buckinghamshire, England) in a serpentine pattern. K9G9 and variants all eluted between fractions D5 – E5 or D6 – E4. Fractions were pooled using a 15 mL Amicon Ultra YM-30 spin filter (UFC903008, Millipore, Billercia, MA) and washed with 10 mL 100 mM HEPES (pH 6.8) and 10 mM $MgCl_2$ buffer. Filtrate was discarded and the purified protein was eluted from the membrane using 1 mL buffer containing 100 mM HEPES (pH 6.8) and 10 mM $MgCl_2$.

To determine V_{max} and K_M values for NADH and NADPH, the purified proteins were assayed at various concentrations of NAD(P)H (0 to 1000 µM) coupled with a NAD(P)H regeneration system. Assays were conducted at 30°C in a buffer containing 100mM MOPS, pH 6.8, 10mM MgCl2, 1mM EDTA, 5 mM (R/S)-acetolactate, 1mM glucose-6-phosphate, 3 mU/µL glucose-6-phosphate dehydrogenase. The reaction as quenched after ten minutes with three volumes 0.1% formic acid. DHIV concentration was measured using LC-MS. The rate of conversion of *S*-acetolactate to DHIV was determined by measuring the amount of DHIV produced at a fixed time point. V_{max} and K_m values were calculated by plotting specific activity (U/mg) vs. cofactor concentration and the data were fit to the Michaelis-Menten equation. Measurements of acetolactate Km values (at a fixed concentration of NADH) indicated that the fixed acetolactate concentration employed for the cofactor Km determinations was saturating.

2017200439 20 Jan 2017

Table 23. Kinetic Values for Purified K9G9 Variants as Determined via DHIV Formation Assays

Variant	V_{max} NADPH, U/mg	K_m NADPH, µM	V_{max}/K_m NADPH, L/min*mg	V_{max} NADH, U/mg	K_m NADH, µM	V_{max}/K_m NADH, L/min*mg
K9G9	2.2	24.1	0.091	1.9	78.2	0.024
K9JB4	2.7	249	0.011	3.4	115	0.030
K9JB4P	2	83.2	0.024	2.9	34.1	0.085
K9G3	3.1	113	0.027	2.8	106	0.026
K9G3P	1.8	33.6	0.054	2.1	18.1	0.116
K9JA1	2.6	63.4	0.041	3.4	14	0.243
K9SB2	1.7	44.8	0.038	1.8	11.6	0.155
K9SB2-K90L	2.1	173	0.012	2.4	28.6	0.084
K9SB2-K90M	1.8	245	0.007	2.2	41.3	0.053

Example 19.

Isobutanol production of K9G9 derivatives with increased ratios of K_m NADPH to K_m NADH

The yeast expression plasmids for K9JB4, K9JB4P, K9JG3, K9JG3P, K9JA1, and K9SB2 were made by subcloning of the variant KARI genes from *E.coli* vectors (pBAD.KARI) into pHR81-PIlv5-KARI-K9.G9 at PmeI and SfiI sites. The resultant plasmids together with pHR81-PIlv5-KARI-K9.G9 and pHR81-PIlv5-KARI-K9.D3 (SEQ ID NO: 181) were analyzed for isobutanol production and by-product formation in yeast. Yeast pathway strains were made in PNY2259 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-ADH|adh_Hl-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_Hl-ADH1t ymr226cΔ ald6Δ::loxP; Example 22) host by co-transforming the KARI vectors as pathway plasmid #1, and pBP915 (pRS423-P_{FBA1}-DHAD-P_{GPM1}-hADH1; SEQ ID NO: 182) as pathway plasmid #2. The transformed cells were plated on synthetic medium without histidine or uracil (1% ethanol as carbon source). Three transformants were transferred to fresh plates of the same media. The transformants were tested for isobutanol production under anaerobic conditions in serum vials.

2017200439 20 Jan 2017

Yeast colonies from the transformation on SE-Ura-His plates appeared after 3-5 days. The three colonies from each variant were patched onto fresh SE-Ura-His plates, incubate at 30◦C for 3 days.

Growth media and procedure

Two types of media were used during the growth procedure of yeast strains: an aerobic pre-culture media and an anaerobic culture media. All chemicals were obtained from Sigma unless otherwise noted (St. Louis, MO)

Aerobic pre-culture media (SE-Ura-His): 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine, leucine, tryptophan and uracil, 0.2% ethanol, 0.2% glucose, 0.01% w/v leucine and 0.002% w/v tryptophan.

Anaerobic culture media (SEG-Ura-His): 50 mM MES (pH 5.5, 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine, leucine, tryptophan and uracil, 0.1% ethanol, 3% glucose, 0.01% leucine, 0.002% tryptophan, 30 mg/L nicotinic acid, 30 mg/L thiamine and 10 mg/L ergosterol made up in 50/50 v/v Tween /ethanol solution.

The patched cells were inoculated into 25 mL SEG-Ura,His media with 0.2% glucose and 0.2% ethanol, and grown under progressively oxygen-limited conditions with lid closed for approximately 48 hours at 30◦C with shaking, until a target OD_{600} value of approximately 1.5 to 2 was achieved. OD_{600} values were recorded. Cells were pelleted via centrifugation and the supernatant was discarded. Cell pellets were transferred into a Coy Anaerobic Bag (Grass Lake, MI) where pellets were resuspended in 1.0 mL anaerobic growth media (SEG-Ura-His). The resuspended cell pellets were used to inoculate 30 mL SEG-Ura-His media in 50 mL serum bottles (Wheaton, 223748, Millville, NJ) to a target initial OD_{600} value of 0.2. All anaerobic media, serum vials, stoppers and crimps were allowed to degas in the anaerobic bag for at least 24 hours prior to inoculation. Serum bottles were stoppered, crimped and transferred out of the anaerobic bag and grown at 30°C with shaking at

2017200439 20 Jan 2017

240 rpm. Anaerobic cultures were grown for 24 to 72 hours with a target OD_{600} value of at least 1.2. Additional anaerobic growth steps used the cells from the previous anaerobic culture step as inoculant. Three transformants were evaluated for each variant.

HPLC analysis of yeast strains with K9G9 KARI variants

Samples were taken for HPLC analysis and to obtain OD_{600} values at the end of the anaerobic growth period. HPLC analysis was performed using a Waters 2695 separations unit, 2996 photodiode array detector, and 2414 refractive index detector (Waters, Milford, MA) with a Shodex Sugar SH-G pre-column and Shodex Sugar SH1011 separations column (Shodex, JM Science, Grand Island, NY). Compounds were separated by isocratic elution at 0.01 N sulfuric acid with a flow rate of 0.5 mL/min. Chromatograms were analyzed using the Waters Empower Pro software.

Molar yields for glycerol, isobutanol and the glycerol/isobutanol ratio were determined. Mean and standard deviations were calculated from triplicate analyses for each variant. Student's t-test was then employed to determine if the difference in the values was statistically significant from the K9D3 control values. For the new variants, the increases in K_M values for NADPH relative to K_M for NADH are expected result reduced NADPH utilization. Results reported in the Table below and in Figure 9 indicate that the new variants with increased ratios of K_M NADPH to K_M NADH exhibit higher isobutanol to glycerol ratios relative to K9D3 and K9G9. K9SB2 demonstrated a 35% increase in isoubtanol titer compared to K9D3.

Table 24. K9G9 Variants Kinetic and Isobutanol Data

	Variant	K_m (NADPH)/ K_m (NADH)	Isobutanol/ Glycerol Ratio	Isobutanol Molar Yield	Isobutanol Titer, mM
Experiment 1	K9D3	0.24	1.67 ± 0.02	0.581 ± 0.007	33.9 ± 1.8
	K9JB4	2.2	2.10 ± 0.06	0.603 ± 0.006	35.3 ± 1.3
	K9JG3	1.1	2.07 ± 0.06	0.598 ± 0.004	39.1 ± 1.6
Experiment 2	K9D3	0.24	1.75 ± 0.06	0.586 ± .0100	63.7 ± 0.6
	K9JA1	4.5	2.24 ± 0.04	0.611 ±	78.2 ± 5.4

20 Jan 2017 2017200439

				0.002	
	K9SB2	3.9	2.21 ± 0.10	0.608 ± 0.007	77.2 ± 1.4
Experiment 3	K9G9	0.31	2.10 ± 0.03	0.624 ± 0.011	54.2 ± 1.5
	K9JG3P	1.9	2.56 ± 0.08	0.652 ± 0.009	66.3 ± 4.0
	K9JB4P	2.4	2.54 ± 0.07	0.654 ± 0.006	61.5 ± 3.4

Example 20.
Construction of K9SB2 error prone PCR library

The K9SB2 error prone PCR library was constructed in a similar manner as the K9G9 library with the following modifications. The mutagenic PCR mixture consisted of 9.5 µl K9SB2 in a pBAD.KARI vector (190 ng/ µl), 1.25 µl of primer K9G9_EZ_F1 (100 ng/ µl), 1.25 µl of primer K9G9_EZ_R1 (100 ng/ µl), 5 µl of 10× Mutazyme II reaction buffer, 1 µl of 40 mM dNTP mix, 1.5 µl of Mutazyme II DNA polymerase, and 30.5 µl of ddH2O. The "EZclone reaction" contained 25 µl of the 2× EZClone enzyme mix, 3 µl of megaprimer (K9SB2 mutagenic PCR product, 190 ng/ µl), 2.6 µl of K9SB2 template DNA (19 ng/ µl), 3 µl of EZClone solution 1, and 16µl of ddH_2O. For the Dpn I step, the mixture was incubated for 3 hr at 37°C. Clones were employed for high throughput screening as described in Example 21.

Example 21.
Screening for K9SB2 variants with further increased ratios of Km NADPH to Km NADPH based on increased NADH to NADPH activity ratios

The K9SB2 library described in Example 20 was screened for variants with reduced NADPH affinity. With the specific objective to increase Km for NADPH relative to Km for NADH, the hits from the screen were partially purified and kinetic analyses were performed to determine V_{max} and K_m parameters with NADH and with NADPH.

High throughput screening assay of K9SB2 gene library

Variants were screened using HTS as described in Example 16, with the following exceptions. Assays buffer consisted of 2.4 mM (R/S)-

2017200439 20 Jan 2017

acetolactate,100 mM HEPES pH 6.8, 10 mM $MgCl_2$, 150 µM NADH or 100 µM NADPH and 12.5 µL 0.5 X cell lysate.

The ratio for the measured rate for oxidation of 100 µM NADPH to the measured rate for oxidation of 150 µM NADH was calculated for each variant and positive control well (2 per plate). A variant well was considered to contain an initial hit if the NADH rate was greater than 0.6 OD/hr and the rate ratio (NADPH/NADH) was less than 0.37 (three standard deviations lower than the positive control mean). A total of 218 hits were identified from a pool of 4947 potential variants. These initial hits were consolidated, forming a smaller library for further analysis.

The consolidated initial hit library was grown in biological triplicate and cell free extracts were prepared and assayed as described above. Rate ratios were then calculated for the variants and positive controls as above. Final hits that were selected for detailed kinetic analysis met the following criteria: the NADPH/NADH rate ratio was less than 0.45, the NADH rate was greater than 0.6 OD/hr and at least two out of three biological replicates passed the criteria. 107 variants were identified.

Data were also analyzed to identify variants that had a higher rate of conversion for *S*-acetolactate to DHIV with the NADH cofactor. The average rate and standard deviation of NADH oxidation was calculated for all the positive controls. A variant was considered a potential hit if the rate of NADH oxidation was at least 3 standard deviations higher than the rate of the positive control (2.524 OD/hr). 68 variants were identified and sequence analysis determined that 17 had at least one amino acid substitution. The substitutions T93A and T93I each appeared twice and variants 2017 B12 and D6 have been selected for further analysis.

DNA sequencing of the 107 variants identified from the secondary HTS screening was accomplished by using TempliPhi™ (GE Healthcare) with the primers pBAD-For (ATGCCATAGCATTTTTATCC; SEQ ID NO: 393) and pBAD-Rev (CTGATTTAATCTGTATCAGGCT; SEQ ID NO: 394). 105 sequences were different from the parent and the amino acid substitutions are listed in the first of the following two tables.

2017200439 20 Jan 2017

DNA sequencing of the 68 variants identified from the NADH rate screening was accomplished by using TempliPhi™ (GE Healthcare) with primers pBAD-For (ATGCCATAGCATTTTTATCC; SEQ ID NO: 393) and pBAD-Rev (CTGATTTAATCTGTATCAGGCT; SEQ ID NO: 394). 17 sequences were different form wild type and the amino acid substitutions of the 2 substitutions that appeared repeatedly are listed in the second table below.

Table 25. K9SB2 Variants' Amino Acid Substitutions

Variant	Seq	Amino Acid Substitutions
K9SB2	427	Y53F, S56A, K57E, S58E, N87P
2011 A2	433	Y53F, G55D, S56A, K57E, S58E, N87P
2011 A3	434	Y53F, S56A, K57E, S58E, N87P, M94I
2011 A5	435	Y53F, S56A, K57E, S58E, M83I, N87P, L185M, E217D
2011 A7	436	Y53F, S56A, K57E, S58E, N87P, D98V
2011 A9	437	Y53F, S56A, K57E, S58E, F67I, N87P
2011 A11	438	Y53F, S56A, K57E, S58E, N87P, M94T, K126E, T273A
2011 B1	439	Y53F, S56A, K57E, S58E, N87P, M94I, A279T
2011 B2	440	Y53F, S56A, K57E, S58E, N87P, Q91L
2011 B3	441	Y53L, S56A, K57E, S58E, N87P
2011 B4	442	Y53F, S56A, K57E, S58E, N87P, P135T
2011 B7	443	Y53F, S56A, K57E, S58E, N87P, L185M
2011 B8	444	Y53F, S56A, K57E, S58E, N87P, C233S, F296Y
2011 B10	445	Y53F, S56A, K57E, S58E, N87P, A303D
2011 C1	446	Y53F, S56A, K57E, S58E, E63K, N87P, G251D, K294R
2011 C3	448	Y53F, S56A, K57E, S58E, A72V, N87P, N102Y, F189I, Y245H
2011 C6	449	Y53F, S56A, K57E, S58E, I84F, N87P
2011 C7	450	E13V, Y53F, S56A, K57E, S58E, M94I, N87P, T141I
2011 C8	451	Y53F, S56A, K57E, S58E, A72V, N87P, N102Y, F189I, Y245H
2011 C9	447	Y53F, S56A, K57E, S58E, N87P, E194D
2011 C10	452	Y53F, A56G, K57E, S58E, K60N, N87P
2011 C12	453	Y53F, S56A, K57E, S58E, I84L, N87P, N97T
2011 D1	454	L39M, Y53F, S56A, K57E, S58E, E68G, N87P
2011 D2	455	Y53F, S56A, K57E, S58E, N87P, M94I, V307I
2011 D3	456	Y53F, S56A, K57E, S58E, F67I, N87P
2011 D4	457	Y53F, S56A, K57E, S58E, N87P, P135L, A202V
2011 D5	458	Y53F, S56A, K57E, S58E, N87P, G164S, G199A
2011 D6	500	Y53F, S56A, K57E, S58E, N87P, S247C
2011 D8	459	Y53F, S56A, K57E, S58E, N87P, N116I
2011 D9	460	Y53F, S56A, K57E, S58E, N87P, K90M
2011 D11	461	Y53F, S56A, K57E, S58E, N87P, M94L, T259I
2011 D12	462	Y53F, S56A, K57E, S58E, M83K, N87P

2011 E3	463	Y53F, S56A, K57E, S58E, N87P, I122V, L297W
2011 E4	464	Y53F, S56A, K57E, S58E, N87P, A112S, Q160R
2011 E8	465	Y53F, S56A, K57E, S58E, N87P, V142I, P320L
2011 E11	466	Y53F, S56A, K57E, S58E, N87P, Q91L
2011 F2	467	Y53F, S56V, K57E, S58E, N87P, A210T
2011 F4	468	Y53F, S56A, K57E, S58E, N87P, F189I
2011 F6	469	Y53F, S56G, K57E, S58E, K60N, N87P
2011 F9	470	A41V, Y53F, S56A, K57E, S58E, N87P, S305P
2011 F10	471	Y53F, S56A, K57E, S58E, L85M, N87P
2011 G1	472	E13V, Y53F, S56A, K57E, S58E, M94I, N87P, T141I
2011 G3	473	H35Q, Y53F, S56A, K57E, S58E, N87P
2011 G4	474	Y53F, S56A, K57E, S58E, I84F, N87P
2011 G8	475	A26T, Y53F, S56A, K57E, S58E, N87P, K90M
2011 G9	476	Y53F, S56T, K57E, S58E, N87P
2011 G10	477	Y53F, S56A, K57E, S58E, I84F, N87P
2011 H1	478	Y53F, S56A, K57E, S58E, N87P, M94I, V156L
2011 H5	479	Y53F, S56A, K57E, S58E, I84L, N87P
2011 H7	480	Y53F, S56T, K57E, S58E, N87P
2011 H9	481	Y53F, G55C, S56A, K57E, S58E, N87P
2012 A2	482	Y53F, S56A, K57E, S58E, R61G, I86V, N87P
2012 A7	483	Y53F, S56A, K57E, S58E, T71S, A76V, N87P
2012 A8	484	Y53F, S56A, K57E, S58E, N87P, M212T
2012 A9	485	H35Q, Y53F, S56A, K57E, S58E, A72T, N87P
2012 A10	486	A36T, Y53F, S56A, K57E, S58E, N87P
2012 A11	487	Y53F, S56A, K57E, S58E, N87P, S247T
2012 B3	488	Y53F, S56A, K57E, S58E, N87P, P135S
2012 B4	489	Y53F, S56A, K57E, S58E, W59R, N87P, K278E
2012 B5	490	Y53F, S56V, K57E, S58E, N87P, I234V
2012 B6	491	Y30H, Y53F, S56A, K57E, S58E, N87P
2012 B9	492	I50M, Y53F, S56A, K57E, S58E, N87P
2012 C2	493	Y53L, S56A, K57E, S58E, N87P
2012 C3	494	Y53F, S56A, K57E, S58E, N87P, F115V, T191S, V208I, C209W, F292I
2012 C5	495	Y53F, S56A, K57E, S58E, N87P, M94I
2012 C6	496	Y53F, S56A, K57E, S58E, F67L, N87P
2012 C8	497	Y53F, S56A, K57E, S58E, N87P, M94I, M169T
2012 C10	498	Y53F, S56A, K57E, S58E, F67I, N87P, T276I
2012 D1	499	Y53F, S56A, K57E, S58E, I84F, N87P, M132T
2012 D8	501	Y53F, S56A, K57E, S58E, N87P, P135S
2012 D11	502	K8N, Y53F, S56A, K57E, S58E, N87P, K90M, T141I
2012 E5	503	Y53F, S56A, K57E, S58E, I84L, N87P
2012 E9	504	Y53F, S56A, K57E, S58E, N87P, V142I, T191S, C233S
2012 F1	505	Y53F, S56A, K57E, S58E, N87P, H235Y
2012 F2	506	Y53F, S56V, K57E, S58E, N87P, V232D
2012 F3	507	Y53F, S56A, K57E, S58E, N87P, K90M, V142I, T187S
2012 F4	508	Y53F, S56A, K57E, S58E, N87P, M94I, G149D
2012 F7	509	E13V, Y53F, S56A, K57E, S58E, N87P, M94I, T141I
2012 F10	510	Y53F, S56A, K57E, S58E, Q65H, N87P, F189I

20172004 39 20 Jan 2017

2012 F12	511	Y53F, S56V, K57E, S58E, N87P
2012 G3	512	Y53F, S56T, K57E, S58E, N87P, R190S
2012 G4	513	Y53F, S56A, K57E, S58E, N87P, N102S, V142I
2012 G5	514	Y53F, S56A, K57E, S58E, I84L, N87P
2012 G8	515	Y53F, S56A, K57E, S58E, K77N, N87P, A92V
2012 G9	516	Y53F, S56A, K57E, S58E, N87P, M94I, V307I
2012 G10	517	Y53F, S56A, K57E, S58E, N87P, T195I
2012 G12	518	Y53F, S56A, K57E, S58E, N87P, F309I
2012 H1	519	Y53F, S56A, K57E, S58E, N87P, K90T, A180S
2012 H3	520	Y53F, S56A, K57E, S58E, W59C, N87P
2012 H7	521	Y53F, S56A, K57E, S58E, N87P, M94I, A202T
2012 H9	522	H35N, Y53F, S56A, K57E, S58E, N87P
2012 H11	523	Y53F, S56A, K57E, S58E, A72V, N87P, L211M, I240M
2013 A2	524	Y53F, S56T, K57E, S58E, N87P, Q288H
2013 A4	525	Y53F, S56A, K57E, S58E, L85M, N87P
2013 A5	526	L52S, Y53F, S56A, K57E, S58E, N87P
2013 B2	527	A36T, Y53F, S56A, K57E, S58E, N87P, V203I
2013 B5	528	Y53F, S56A, K57E, S58E, N87P, P135T
2013 B7	529	I9M, Y53F, S56A, K57E, S58E, N87P, K90E
2013 B8	530	Y53F, G55C, S56A, K57E, S58E, N87P
2013 B9	531	Y53F, S56V, K57E, S58E, N87P
2013 B11	532	A38V, Y53F, S56A, K57E, S58E, N87P
2013 C1	533	Y53L, S56A, K57E, S58E, N87P, M237I
2013 C6	534	K23M, Y53F, S56A, K57E, S58E, N87P, E194D
2013 C8	535	Y53F, S56A, K57E, S58E, N87P, P135T
2013 C12	536	Y53F, S56A, K57E, S58E, A72V, N87P, T93S, A176V, H235Y
2013 D1	537	Y53F, S56A, K57E, S58E, N87P, M94R, K310M

Table 26. K9SB2 Variants' Amino Acid Substitutions

Variant	AA Seq ID NO; Nucleic acid SEQ ID NO	Amino Acid Substitutions
K9SB2	427	Y53F, S56A, K57E, S58E, N87P
2017 B12	639; 640	Y53F, S56A, K57E, S58E, N87P, T93I
2017 D6	641; 642	Y53F, S56A, K57E, S58E, N87P, T93A

Kinetic analysis of partially purified K9SB2 variant proteins

E. coli strain Bw25113 (*ΔilvC*), as described in U.S. Patent 8,129,162, was used to express the 107 variants from the secondary HTS

2017200439 20 Jan 2017

screening and positive control K9SB2. Clones from archive plates were inoculated into the 96-deep well plates. Each well contained 3.0 µl of cells from thawed archive plates, 200 µl of the LB medium containing 100 µg/ml ampicillin and 0.02%(w/v) arabinose as the inducer. Cells were the grown overnight at 37°C with 80% humidity while shaking (900 rpm), harvested by centrifugation (4000 rpm, 7 min at 4°C) (75004251, Thermo Scientific, Rockford, IL) and the cell pellet was stored at -80°C for later analysis.

Frozen cell pellets in deep-well plates were thawed at room temperature for 30 minutes at the same time. 75 µl of 50% BugBuster (Novagen 71456, Darmstadt, Germany) (v/v in water) was added to each and cells were suspended using a plate shaker. The cells suspension in 50% Bug Buster was incudbated for 30 minutes at room temperature which was then followed by a 15 minute incubation at 60°C. Cell debris and denatured heat labile proteins were pelleted by centrifugation (4000 rpm, 15 min at 4°C) (75004251, Thermo Scientific, Rockford, IL) and 75 µL of the supernatant was transferred to a flat bottomed 96-well plate (Corning, 3370, Corning, NY) and diluted two-fold with 75 µL 100 mM HEPES (pH 6.8), 100 mM KCl, 10 mM $MgCl_2$.

Total protein was determined by using the Bradford Assay with Coomaisse Plus (Thermo Scientific, #23238, Rockford, IL). BSA was employed as the standard. The concentration of protein was measured by determining the absorbance at 595 nm using a Cary 300 spectrophotometer (Agilent Technologies, Wilmington, DE).

To determine V_{max} and K_M values for NADH and NADPH, the partially purified proteins were assayed at various concentrations of NADH (20, 30, 40, 60, 80, 120, 200 and 300 µM) and NADPH (60, 80, 120, 200, 300 and 400 µM). Assays were conducted at 30°C in 100 mM HEPES (pH 6.8), 10 mM $MgCl_2$, 100 mM KCl and 4.8 mM *R/S*-acetolactate. Between 0.005 to 0.015 mg/mL total protein was added to the assay. The rate of conversion of *S*-acetolactate to DHIV was measured via monitoring the oxidation of NAD(P)H at 340 nm using a Spectramax 384 Plus plate reader (Molecular Devices, Sunnyvale, CA). The activity was calculated using the molar extinction coefficient of 6220 $M^{-1}cm^{-1}$. V_{max} and K_m values

20172004439 20 Jan 2017

were calculated by plotting specific activity (U/mg) vs. cofactor concentration and the data were fit to the Michaelis-Menten equation using Kaleidagraph software (Synergy, Reading, PA).

Table 27. Kinetic Values for Partially Purified K9SB2 Variants as Determined via NAD(P)H Consumption Assays

Variant	V_{max} NADPH, U/mg	K_m NADPH, µM	V_{max}/K_m NADPH, L/min*mg	V_{max} NADH, U/mg	K_m NADH, µM	V_{max}/K_m NADH, L/min*mg
K9SB2	1.79	153	0.012	2.10	87.4	0.024
2011 A2	1.48	897	0.002	1.94	71.2	0.027
2011 A3	1.43	371	0.004	1.33	44.1	0.030
2011 A5	0.93	109	0.009	0.23	17.4	0.017
2011 A7	2.27	334	0.007	1.70	69.2	0.025
2011 A9	2.09	266	0.008	0.40	n/d	n/a
2011 A11	2.21	294	0.008	1.25	25.8	0.048
2011 B1	1.79	421	0.004	64.5	55171	0.001
2011 B2	0.33	254	0.001	0.02	264	0.007
2011 B3	1.87	505	0.004	0.68	225	0.003
2011 B4	0.36	294	0.001	0.26	171	0.002
2011 B7	n/d	1.03	n/a	0.48	25.5	0.019
2011 B8	0.83	109.5	0.008	0.61	28.7	0.021
2011 B10	0.88	171	0.005	0.46	7.53	0.061
2011 C1	1.06	404	0.003	1.51	191	0.008
2011 C3	1.08	844	0.001	2.53	500	0.005
2011 C6	1.19	388	0.003	1.90	189	0.010
2011 C7	0.71	946	0.001	1.95	457	0.004
2011 C8	1.73	1546	0.001	3.73	750	0.005
2011 C9	1.02	123	0.008	2.29	177	0.013
2011 C10	1.02	656	0.002	3.84	899	0.004
2011 C12	2.74	244	0.011	3.08	99	0.031
2011 D1	4.68	501	0.009	3.54	80.4	0.044
2011 D2	2.34	547	0.004	2.18	77.3	0.028
2011 D3	0.05	306	0.0002	0.05	44.3	0.001
2011 D4	0.47	857	0.001	0.44	91.6	0.005
2011 D5	0.75	550	0.001	0.42	57.1	0.007
2011 D6	0.70	200	0.004	0.52	25.11	0.021
2011 D8	0.04	214	0.0002	0.04	38.9	0.001
2011 D9	0.18	407	0.0004	0.16	50.2	0.003
2011 D11	0.78	185	0.004	0.61	15.0	0.041
2011 D12	0.74	190	0.004	0.80	39.0	0.021
2011 E3	0.77	163	0.005	1.54	128	0.012
2011 E4	1.59	270	0.006	3.78	234	0.016
2011 E8	0.91	435	0.002	2.16	252	0.009
2011 E11	5.56	6466	0.001	3.06	511	0.006
2011 F2	0.39	692	0.001	1.79	136	0.013
2011 F4	2.07	242	0.009	2.00	68.3	0.029

2011 F6	1.19	946	0.001	1.35	231	0.006
2011 F9	0.57	269	0.002	0.45	27.1	0.017
2011 F10	1.43	390	0.004	1.55	79.8	0.019
2011 G1	1.31	1533	0.001	0.50	40.1	0.013
2011 G3	0.61	1003	0.001	1.00	283	0.004
2011 G4	1.00	316	0.003	2.13	174	0.012
2011 G8	0.90	482	0.002	1.22	106	0.012
2011 G9	0.30	530	0.001	3.48	549	0.006
2011 G10	0.93	358	0.003	2.13	238	0.009
2011 H1	0.96	218	0.004	1.57	114	0.014
2011 H5	1.17	81.2	0.014	1.85	86.5	0.022
2011 H7	0.20	435	0.001	1.25	173	0.007
2011 H9	1.30	741	0.002	1.55	177	0.009
2012 A2	1.71	264	0.007	1.58	55.4	0.029
2012 A7	1.98	215	0.009	1.87	67.3	0.028
2012 A8	1.19	91.8	0.013	1.24	22.7	0.055
2012 A9	0.44	481	0.001	0.34	38.1	0.009
2012 A10	1.21	340	0.004	1.31	66.7	0.020
2012 A11	1.99	342	0.006	1.37	35.9	0.038
2012 B3	0.88	1214	0.001	0.42	63.8	0.007
2012 B4	4.34	1593	0.003	1.66	95.1	0.018
2012 B5	4.88	7389	0.001	1.19	85.2	0.014
2012 B6	3.72	428	0.009	2.23	51.3	0.044
2012 B9	1.17	523	0.002	0.87	63.7	0.014
2012 C2	4.43	923	0.005	3.16	189	0.017
2012 C3	1.40	203	0.007	1.80	68.9	0.026
2012 C5	1.73	348	0.005	3.66	268	0.014
2012 C6	2.18	234	0.009	2.90	103	0.028
2012 C8	1.53	394	0.004	2.63	194	0.014
2012 C10	1.12	286	0.004	1.41	78.9	0.018
2012 D1	0.72	599	0.001	0.58	64.5	0.009
2012 D8	1.17	1528	0.001	0.47	73.6	0.006
2012 D11	0.43	334	0.001	0.52	27.7	0.019
2012 E5	2.13	257	0.008	4.77	313	0.015
2012 E9	1.07	1326	0.001	n/d	n/d	n/a
2012 F1	1.70	272	0.006	1.48	62.5	0.024
2012 F2	0.39	925	0.0004	0.79	97.8	0.008
2012 F3	1.40	2213	0.0006	0.94	142	0.007
2012 F4	2.70	719	0.004	1.50	72.9	0.021
2012 F7	0.86	840	0.001	0.77	117	0.007
2012 F10	1.92	170	0.011	1.75	27.4	0.064
2012 F12	0.90	1582	0.0006	1.75	117	0.015
2012 G3	1.47	1003	0.002	1.07	127	0.008
2012 G4	0.73	615	0.001	0.63	81.6	0.007
2012 G5	1.92	240	0.008	2.00	83.5	0.024
2012 G8	1.17	315	0.004	0.99	58.7	0.017
2012 G9	2.41	717	0.003	1.37	91.5	0.015

2017200439 20 Jan 2017

2012 G10	1.06	400	0.003	0.71	39.8	0.018
2012 G12	1.58	147	0.011	2.00	70.0	0.029
2012 H1	1.49	195	0.008	1.74	68.9	0.025
2012 H3	14.98	7389	0.002	1.45	99.0	0.015
2012 H7	1.14	246	0.005	1.30	76.2	0.017
2012 H9	0.37	210	0.002	n/d	n/d	n/a
2012 H11	0.65	162	0.004	0.62	32.0	0.019
2013 A2	0.58	285	0.002	0.64	71.0	0.009
2013 A4	0.63	188	0.003	0.86	81.5	0.011
2013 A5	0.61	886	0.001	0.88	210	0.004
2013 B2	0.62	282	0.002	0.71	70.4	0.010
2013 B5	6.68	7389	0.001	0.083	150	0.006
2013 B7	1.22	433	0.003	1.12	79.0	0.014
2013 B8	0.42	90.7	0.005	1.27	191	0.007
2013 B9	5.31	13970	0.0004	1.38	217	0.006
2013 B11	0.48	212	0.002	0.60	63.2	0.010
2013 C1	0.49	149	0.003	0.68	64.8	0.0105
2013 C6	0.54	163	0.003	0.36	24.4	0.015
2013 C8	2.87	3752	0.001	0.70	188	0.004
2013 C12	0.75	495	0.002	0.79	115	0.007
2013 D1	1.31	1608	0.001	0.87	188	0.005

Example 22:

Construction of strain PNY2259

The purpose of this example is to describe the assembly of the constructs used to replace the chromosomal copy of kivD_LI(y) in PNY2238 at the *adh1Δ* locus with kivD_Lg(y).

The deletion/integration was created by homologous recombination with PCR products containing regions of homology upstream and downstream of the target region and the *URA3* gene for selection of transformants. The *URA3* gene was removed by homologous recombination to create a scarless deletion/integration. The plasmid to integrate kivD_Lg(y) was derived from a plasmid constructed to integrate UAS(PGK1)P[FBA1]-kivD_LI(y) into the *ADH1* locus of *Saccharomyces cerevisiae*. Construction of the plasmid used to integrate UAS(PGK1)P[FBA1]-kivD_LI(y) into the *ADH1* locus is described below. The plasmids were constructed in pUC19-URA3MCS.

Construction of the *ADH1* deletion/UAS(PGK1)P[FBA1]-kivD_LI(y) integration plasmid

2017200439 20 Jan 2017

The *kivD* coding region from *Lactococcus lactis* codon optimized for expression in *Saccharomyces cerevisiae*, kivD_LI(y), was amplified using pLH468 (SEQ ID NO: 139) as template with primer oBP562 (SEQ ID NO: 197), containing a PmeI restriction site, and primer oBP563 (SEQ ID NO: 198), containing a 5' tail with homology to the 5' end of ADH1 Fragment B. ADH1 Fragment B was amplified from *Saccharomyces cerevisiae* CEN.PK 113-7D genomic DNA with primer oBP564 (SEQ ID NO: 199), containing a 5' tail with homology to the 3' end of *kivD_LI(y)*, and primer oBP565 (SEQ ID NO: 200), containing a FseI restriction site. PCR products were purified with a PCR Purification kit (Qiagen; Valencia, CA). kivD_LI(y)-ADH1 Fragment B was created by overlapping PCR by mixing the kivD_LI(y) and ADH1 Fragment B PCR products and amplifying with primers oBP562 (SEQ ID NO: 197) and oBP565 (SEQ ID NO: 200). The resulting PCR product was digested with PmeI and FseI and ligated with T4 DNA ligase into the corresponding sites of pUC19-URA3MCS after digestion with the appropriate enzymes. ADH1 Fragment A was amplified from genomic DNA with primer oBP505 (SEQ ID NO: 201), containing a SacI restriction site, and primer oBP506 (SEQ ID NO: 202), containing an AscI restriction site. The ADH1 Fragment A PCR product was digested with SacI and AscI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing kivD_LI(y)-ADH1 Fragment B. ADH1 Fragment C was amplified from genomic DNA with primer oBP507 (SEQ ID NO: 203), containing a PacI restriction site, and primer oBP508 (SEQ ID NO: 204), containing a SalI restriction site. The ADH1 Fragment C PCR product was digested with PacI and SalI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing ADH1 Fragment A-kivD_LI(y)-ADH1 Fragment B. The hybrid promoter UAS(PGK1)-P_{FBA1} (SEQ ID NO: 406) was amplified from vector pRS316- UAS(PGK1)-P_{FBA1}-GUS with primer oBP674 (SEQ ID NO: 205), containing an AscI restriction site, and primer oBP675 (SEQ ID NO: 206), containing a PmeI restriction site. The UAS(PGK1)-P_{FBA1} PCR product was digested with AscI and PmeI and ligated with T4 DNA ligase into the corresponding sites of the plasmid containing kivD_LI(y)-ADH1 Fragments ABC to generate pBP1181.

2017200439 20 Jan 2017

Construction of pBP1716 and pBP1719

kivD_LI(y) was removed from the *ADH1* deletion/UAS(PGK1)P[FBA1]-kivD_LI(y) integration plasmid pBP1181. The plasmid was digested with PmeI and FseI and the large DNA fragment was purified on an agarose gel followed by a gel extraction kit (Qiagen). ADH1 fragment B was amplified from pBP1181 with primer oBP821 (SEQ ID NO: 407), containing a PmeI restriction site, and primer oBP484 (SEQ ID NO: 408), containing a FseI restriction site. The ADH1 fragment B PCR product was digested with PmeI and FseI and ligated with T4 DNA ligase into the corresponding sites of the gel purified large DNA fragment. A PCR fragment corresponding to the 3' 500bp of kivD_LI(y) was cloned into the resulting vector for the targeted deletion of kivD_LI(y) in PNY1528. The fragment was amplified from pBP1181 with primers oBP822 (SEQ ID NO: 409), containing a NotI restriction site, and oBP823 (SEQ ID NO: 410), containing a PacI restriction site. The fragment was digested with NotI and PacI and ligated with T4 DNA ligase into the corresponding sites downstream of URA3 in the above plasmid with the kivD_LI(y) deletion after digestion with the appropriate restriction enzymes. The resulting plasmid was designated pBP1716.

The *kivD* coding region from *Listeria grayi* codon optimized for expression in *Saccharomyces cerevisiae* (SEQ ID NO: 411), kivD_Lg(y), was synthesized by DNA2.0 (Menlo Park, CA). kivD_Lg(y) was amplified with primers oBP828 (SEQ ID NO: 412), containing a PmeI restriction site, and oBP829 (SEQ ID NO: 413) containing a PmeI restriction site. The resulting PCR product was digested with PmeI and ligated with T4 DNA ligase into the corresponding site in pBP1716 after digestion with the appropriate enzyme. The orientation of the cloned gene was checked by PCR with primers FBAp-F (SEQ ID NO: 414) and oBP829 (SEQ ID NO: 413). An isolate with kivD_Lg(y) in the correct orientation was designated pBP1719.

Construction of strain PNY2259

2017200439 20 Jan 2017

The kivD_LI(y) deletion/kivD_Lg(y) integration cassette was amplified from pBP1719 with primers oBP505 (SEQ ID NO: 201) and oBP823 (SEQ ID NO: 410). Competent cells of the PNY2238 were made and transformed with the PCR product using a Frozen-EZ Yeast Transformation II kit (Zymo Research; Orange, CA). Transformation mixtures were plated on synthetic complete media lacking uracil supplemented with 1% ethanol at 30C. Transformant strains were screened by PCR (JumpStart(TM) REDTaq (c) ReadyMix(TM)) using primers Ura3-end F (SEQ ID NO: 222) and HY-50 (SEQ ID NO: 415). Transformants were grown in YPE (1% ethanol) and plated on synthetic complete medium supplemented with 1% EtOH and containing 5-fluoro-orotic acid (0.1%) at 30 C to select for isolates that lost the *URA3* marker. The deletion of kivD_LI(y) and integration of kivD_Lg(y) was confirmed by PCR with primers HY-50 and oBP834 (SEQ ID NO: 416). One correct isolate contained kivD_Lg(y) at the same locus and expressed from the same promoter as kivD_LI(y) in PNY2238 was designated PNY2259.

Example 23

Construction of two site-saturation gene libraries to identify variants with cofactor preference to NADH

In Example 4, primers having the degeneracy codon NNK were used (N represents all 4 nucleotides A, C G and T while K stands for G and T). In this Example, primer mixtures containing primers encoding each individual amino acid change of A, C, D, E, F, G, H, I, L, M, N, P, Q, V, W, or Y for positions 53, 56 and 58 of K9 KARI were employed and substitutions to S, T, K, and R were excluded as non-preferred for these positions. The size of the saturation library targeting the three NADPH phosphate binding sites (53, 56 and 58) of K9 KARI is 4,096 (as compared to 32*32*32 or 32,768 variants using NNK degeneracy code primers as in Example 4).

One library construction method started from position 58. Primer mixtures were first made by mixing all the primers targeting the same

2017200439 20 Jan 2017

positions (for example, the primer mixture, K9_53f, was made by mixing equal mole of all 16 forward primers targeting position 53 (listed in the table below). Similarly, K9_56f and K9_58f were prepared. The shared reverse primer is K9_191G_112210r (SEQ ID NO: 174): GGTTTCAGTTTCGCCTCTGAAGGTAGTTTC (called SR in this example). The mutation at position 58 was first introduced into AS6F1 through PCR. The mutagenesis procedure is similar to the one described in Example 4. In brief, K9_58f and SR were phosphorylated. The phosphorylated primers were then directly used to introduce mutation at position 58 into AS6F1 using USB Change_It kit (USB Corporation, Cleveland, OH, #78480). The template was removed with Dpn I. The cleaned up PCR product (Zymo DNA Clean & Concentrator-5; Zymo Research Corporation, Irvine, CA, Cat #D4003) was transformed into KOBW-3a cells. After overnight growth on LB agar plates at a 37 ^{o}C incubator, all cells were collected and DNA was extracted using the Qiaprep Spin miniprep kit (Qiagen Inc. Valencia, CA, Cat #27106)

The extracted DNA was then used as templates to introduce mutation at position 56 using K9_56f and SR same as the mutagenesis for position 58. At last the mutation at position 53 was similarly introduced. After mutations at all three positions (53, 56 and 58) were introduced into AS6F1, the new library was screened same as the one described in example 4 and some selected mutants are listed in the table below.

The other method began with the position 53. The primer mixtures K9_56r and K9_58r were similarly prepared using primers listed in the table below). The shared forward primer is pBAD_266f: CTCTCTACTGTTTCTCCATACCCG (SEQ ID NO: 634; called SF in this example). The mutation at position 53 was first introduced into AS6F1 through PCR. The mutagenesis procedure is similar to the one described above using AS6F1 as the template and K9_53f and SR as the two PCR primers. The resulted mutated DNA (at position 53) was used as templates and K9_56r and SF were used as the mutagenesis primers to introduced mutation at position 56. At last the mutation at position 58 was similarly introduced using K9_58r and SF. After mutations at all three

20 Jan 2017 2017200439

positions (53, 56 and 58) were introduced into AS6F1, the new library was screened same as above and some selected mutants were listed in the table below.

Table 28. forward mutational primers.

Targeted position(s) of *K9-KARI*	Primers		
53	K9_53F_031411f	GTTATCATCGGATTATTCGAAGGA	(SEQ ID NO: 544)
	K9_53L_031411f	GTTATCATCGGATTATTGGAAGGA	(SEQ ID NO: 545)
	K9_53Y_031411f	GTTATCATCGGATTATATGAAGGA	(SEQ ID NO: 546)
	K9_53C_031411f	GTTATCATCGGATTATGTGAAGGA	(SEQ ID NO: 547)
	K9_53W_031411f	GTTATCATCGGATTATGGGAAGGA	(SEQ ID NO: 548)
	K9_53P_031411f	GTTATCATCGGATTACCAGAAGGA	(SEQ ID NO: 549)
	K9_53H_031411f	GTTATCATCGGATTACATGAAGGA	(SEQ ID NO: 550)
	K9_53Q_031411f	GTTATCATCGGATTACAAGAAGGA	(SEQ ID NO: 551)
	K9_53I_031411f	GTTATCATCGGATTAATTGAAGGA	(SEQ ID NO: 552)
	K9_53M_031411f	GTTATCATCGGATTAATGGAAGGA	(SEQ ID NO: 553)
	K9_53N_031411f	GTTATCATCGGATTAAATGAAGGA	(SEQ ID NO: 554)
	K9_53V_031411f	GTTATCATCGGATTAGTTGAAGGA	(SEQ ID NO: 555)
	K9_53A_031411f	GTTATCATCGGATTAGCTGAAGGA	(SEQ ID NO: 556)
	K9_53D_031411f	GTTATCATCGGATTAGATGAAGGA	(SEQ ID NO: 557)
	K9_53E_031411f	GTTATCATCGGATTAGAAGAAGGA	(SEQ ID NO: 558)
	K9_53G_031411f	GTTATCATCGGATTAGGTGAAGGA	(SEQ ID NO: 559)
56	K9_56F_031411f	GGATTACCTGAAGGATTCAAA	(SEQ ID NO: 560)
	K9_56L_031411f	GGATTACCTGAAGGATTGAAA	(SEQ ID NO: 561)
	K9_56Y_031411f	GGATTACCTGAAGGATATAAA	(SEQ ID NO: 562)
	K9_56C_031411f	GGATTACCTGAAGGATGTAAA	(SEQ ID NO: 563)
	K9_56W_031411f	GGATTACCTGAAGGATGGAAA	(SEQ ID NO: 564)
	K9_56P_031411f	GGATTACCTGAAGGACCAAAA	(SEQ ID NO: 565)
	K9_56H_031411f	GGATTACCTGAAGGACATAAA	(SEQ ID NO: 566)
	K9_56Q_031411f	GGATTACCTGAAGGACAAAAA	(SEQ ID NO: 567)
	K9_56I_031411f	GGATTACCTGAAGGAATTAAA	(SEQ ID NO: 568)
	K9_56M_031411f	GGATTACCTGAAGGAATGAAA	(SEQ ID NO: 569)
	K9_56N_031411f	GGATTACCTGAAGGAAATAAA	(SEQ ID NO: 570)
	K9_56V_031411f	GGATTACCTGAAGGAGTTAAA	(SEQ ID NO: 571)
	K9_56A_031411f	GGATTACCTGAAGGAGCTAAA	(SEQ ID NO: 572)
	K9_56D_031411f	GGATTACCTGAAGGAGATAAA	(SEQ ID NO: 573)
	K9_56E_031411f	GGATTACCTGAAGGAGAAAAA	(SEQ ID NO: 574)
	K9_56G_031411f	GGATTACCTGAAGGAGGTAAA	(SEQ ID NO: 575)
58	K9_58F_051611f	GATTACCTGAAGGATTTAAATTCTGGAAGAGAGC	(SEQ ID NO: 576)
	K9_58L_051611f	GATTACCTGAAGGATTTAAATTGTGGAAGAGAGC	(SEQ ID NO: 577)
	K9_58Y_051611f	GATTACCTGAAGGATTTAAATATTGGAAGAGAGC	(SEQ ID NO: 578)
	K9_58C_051611f	GATTACCTGAAGGATTTAAATGTTGGAAGAGAGC	(SEQ ID NO: 579)
	K9_58W_051611f	GATTACCTGAAGGATTTAAATGGTGGAAGAGAGC	(SEQ ID NO: 580)
	K9_58P_051611f	GATTACCTGAAGGATTTAAACCATGGAAGAGAGC	(SEQ ID NO: 581)
	K9_58H_051611f	GATTACCTGAAGGATTTAAACATTGGAAGAGAGC	(SEQ ID NO: 582)
	K9_58Q_051611f	GATTACCTGAAGGATTTAAACAATGGAAGAGAGC	(SEQ ID NO: 583)
	K9_58I_051611f	GATTACCTGAAGGATTTAAAATTTGGAAGAGAGC	(SEQ ID NO: 584)
	K9_58M_051611f	GATTACCTGAAGGATTTAAAATGTGGAAGAGAGC	(SEQ ID NO: 585)
	K9_58N_051611f	GATTACCTGAAGGATTTAAAAATTGGAAGAGAGC	(SEQ ID NO: 586)
	K9_58V_051611f	GATTACCTGAAGGATTTAAAGTTTGGAAGAGAGC	(SEQ ID NO: 587)
	K9_58A_051611f	GATTACCTGAAGGATTTAAAGCTTGGAAGAGAGC	(SEQ ID NO: 588)
	K9_58D_051611f	GATTACCTGAAGGATTTAAAGATTGGAAGAGAGC	(SEQ ID NO: 589)
	K9_58E_051611f	GATTACCTGAAGGATTTAAAGAATGGAAGAGAGC	(SEQ ID NO: 590)

20 Jan 2017 2017200439

	K9_58G_051611f	GATTACCTGAAGGATTTAAAGGTTGGAAGAGAGC	(SEQ ID NO: 591)

Table 29. forward mutational primers.

Targeted position(s) of *K9-KARI*	Primers		
56	K9_56F_071211r	GCTCTCTTCCATGGTTTGAATCCTTC	(SEQ ID NO: 592)
	K9_56L_071211r	GCTCTCTTCCATGGTTTCAATCCTTC	(SEQ ID NO: 593)
	K9_56Y_071211r	GCTCTCTTCCATGGTTTATATCCTTC	(SEQ ID NO: 594)
	K9_56C_071211r	GCTCTCTTCCATGGTTTACATCCTTC	(SEQ ID NO: 595)
	K9_56W_071211r	GCTCTCTTCCATGGTTTCCATCCTTC	(SEQ ID NO: 596)
	K9_56P_071211r	GCTCTCTTCCATGGTTTTGGTCCTTC	(SEQ ID NO: 597)
	K9_56H_071211r	GCTCTCTTCCATGGTTTATGTCCTTC	(SEQ ID NO: 598)
	K9_56Q_071211r	GCTCTCTTCCATGGTTTTTGTCCTTC	(SEQ ID NO: 599)
	K9_56I_071211r	GCTCTCTTCCATGGTTTAATTCCTTC	(SEQ ID NO: 600)
	K9_56M_071211r	GCTCTCTTCCATGGTTTCATTCCTTC	(SEQ ID NO: 601)
	K9_56N_071211r	GCTCTCTTCCATGGTTTATTTCCTTC	(SEQ ID NO: 602)
	K9_56V_071211r	GCTCTCTTCCATGGTTTAACTCCTTC	(SEQ ID NO: 603)
	K9_56A_071211r	GCTCTCTTCCATGGTTTAGCTCCTTC	(SEQ ID NO: 604)
	K9_56D_071211r	GCTCTCTTCCATGGTTTATCTCCTTC	(SEQ ID NO: 605)
	K9_56E_071211r	GCTCTCTTCCATGGTTTTTCTCCTTC	(SEQ ID NO: 606)
	K9_56G_071211r	GCTCTCTTCCATGGTTTACCTCCTTC	(SEQ ID NO: 607)
58	K9_58F_071211r	GTTCTTCTGCTCTCTTCCAGAATTT	(SEQ ID NO: 608)
	K9_58L_071211r	GTTCTTCTGCTCTCTTCCACAATTT	(SEQ ID NO: 609)
	K9_58Y_071211r	GTTCTTCTGCTCTCTTCCAATATTT	(SEQ ID NO: 610)
	K9_58C_071211r	GTTCTTCTGCTCTCTTCCAACATTT	(SEQ ID NO: 611)
	K9_58W_071211r	GTTCTTCTGCTCTCTTCCACCATTT	(SEQ ID NO: 612)
	K9_58P_071211r	GTTCTTCTGCTCTCTTCCATGGTTT	(SEQ ID NO: 613)
	K9_58H_071211r	GTTCTTCTGCTCTCTTCCAATGTTT	(SEQ ID NO: 614)
	K9_58Q_071211r	GTTCTTCTGCTCTCTTCCATTGTTT	(SEQ ID NO: 615)
	K9_58I_071211r	GTTCTTCTGCTCTCTTCCAAATTTT	(SEQ ID NO: 616)
	K9_58M_071211r	GTTCTTCTGCTCTCTTCCACATTTT	(SEQ ID NO: 617)
	K9_58N_071211r	GTTCTTCTGCTCTCTTCCAATTTTT	(SEQ ID NO: 618)
	K9_58V_071211r	GTTCTTCTGCTCTCTTCCAAACTTT	(SEQ ID NO: 619)
	K9_58A_071211r	GTTCTTCTGCTCTCTTCCAAGCTTT	(SEQ ID NO: 620)
	K9_58D_071211r	GTTCTTCTGCTCTCTTCCAATCTTT	(SEQ ID NO: 621)
	K9_58E_071211r	GTTCTTCTGCTCTCTTCCATTCTTT	(SEQ ID NO: 622)
	K9_58G_071211r	GTTCTTCTGCTCTCTTCCAACCTTT	(SEQ ID NO: 623)

Table 30. List of some mutants with their measured K_M values

Mutant	SEQ ID NO: (nucleic acid, amino acid)	Mutations	K_M (μM) (NADH)	K_M (μM) (NADPH)
K9 Wt	26, 27	--	326	0.2
BI7D12	625, 624	Y53Q, S56V, S58D, I86V, N87P, T131M, T191G	39	196
BI10F1	627, 626	Y53E, S56V, S58D, I86V, N87P, T131M, T191G	74	573
BJ6G6	629, 628	Y53P, S56D, S58Q, I86V, N87P, T131M, T191G	298	672

20 Jan 2017 2017200439

BJ7D6	631, 630	Y53P, S56V, S58E, I86V, N87P, T131M, T191G	37	236
BJ7F7	633, 632	Y53A, S56D, S58Q, I86V, N87P, T131M, T191G	269	762

Example 24

Construction of an ald6Δ strain and isobutanol-producing derivatives

A 5.3 kb (BglII/EcoRV) DNA fragment from pRS426::GPD-xpk1+ADH-eutD (SEQ ID NO:383) containing expression cassettes for xpk1 and eutD genes from Lactobacillus plantarum was added between the ALD6 flanking sequences, at the SnaBI site of the pUC19::ald6D::loxP-URA3-loxP vector described in Example 9, above. The ligation reaction was transformed into *E. coli* Stbl3 cells, which were incubated on LB Amp plates to select for transformants. Insertion of the xpk1-eutD cassette was confirmed by PCR (primers). A positive clone (pUC19::Δald6::URA3::xpkS) was obtained.

The vector described above was linearized with AhdI and transformed into PNY1507 (described herein) cells prepared with the Zymo Research Frozen-EZ Yeast Transformation Kit (Cat. No. T2001) with a modification to manufacturer's protocol that included an additional outgrowth incubation of 2.5 hrs. in 2.0 mL YPE (yeast extract, peptone with 1% ethanol) medium. Transformants were obtained by plating on synthetic complete medium minus uracil that provided 1% ethanol as the carbon source. Patched transformants were screened by PCR to confirm the deletion/integration, using primers N1090 and N1213 (SEQ ID NOs: 779 and 242). A plasmid carrying Cre recombinase (pRS423::GAL1p-Cre; SEQ ID No. 271) was transformed into the strain using histidine marker selection. Transformants were passaged on YPE supplemented with 0.5% galactose. Colonies were screened for resistance to 5-FOA (loss of URA3 marker) and for histidine auxotrophy (loss of the Cre plasmid). Proper removal of the URA3 gene via the flanking loxP sites was confirmed by PCR with primers N1212 and N1214 (SEQ ID NOs: 241 and 281). Finally, the alsS integration plasmid (SEQ ID NO:780) was

2017200439 20 Jan 2017

transformed into this strain using the included geneticin selection marker. Integrants were confirmed using primers N160SeqF5 and oBP512 (SEQ ID NO: 388 and 337).

Plasmids pYZ090ΔalsS and pBP915 (SEQ ID NOs: 371 and 182) were transformed into the strain by lithium acetate transformation (Protocol #2 in "Methods in Yeast Genetics" 2005. Amberg, Burke and Strathern). Transformants were selected by plating on synthetic complete minus histidine and uracil with ethanol as the carbon source. Transformants were patched and then repatched onto synthetic complete minus histidine and uracil with 2% glucose and 0.05% ethanol. Six clones were evaluated for growth and isobutanol production. One of these has been designated PNY2216.

Example 25

YMR226c Deletion from *S. cerevisiae* Strain PNY2211 (Construction of PNY2248)

The gene YMR226c was deleted from *S. cerevisiae* strain PNY2211 (described in Example 9) by homologous recombination using a PCR amplified linear KanMX4-based deletion cassette available in *S. cerevisiae* strain BY4743 ymr226cΔ::KanMX4 (ATCC 4020812). Forward and reverse PCR primers N1237 (SEQ ID NO:784) and N1238 (SEQ ID NO:785), amplified a 2,051 bp ymr226cΔ::KanMX4 deletion cassette from chromosome XIII. The PCR product contained upstream and downstream sequences of 253 and 217 bp, respectively, flanking the ymr226cΔ::KanMX4 deletion cassette, that are 100% homologous to the sequences flanking the native YMR226c locus in strain PNY2211. Recombination and genetic exchange occur at the flanking homologous sequences effectively deleting the YMR226c gene and integrating the ymr226cΔ::KanMX4 deletion cassette.

Approximately 2.0 µg of the PCR amplified product was transformed into strain PNY2211 made competent using the lithium-acetate method previously described in *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202

2017200439 20 Jan 2017

(2005)), and the transformation mix was plated on YPE plus geneticin (50 µg/mL) and incubated at 30 °C for selection of cells with an integrated ymr226cΔ::KanMX4 cassette. Transformants were screened for ymr226cΔ::KanMX4 by PCR, with a 5' outward facing KanMX4 deletion cassette-specific internal primer N1240 (SEQ ID NO:786) paired with a flanking inward facing chromosome-specific primer N1239 (SEQ ID NO:243) and a 3' outward-facing KanMX4 deletion cassette-specific primer N1241 (SEQ ID NO:787) paired with a flanking inward-facing chromosome-specific primer N1242 (SEQ ID NO:244). Positive PNY2211 ymr226cΔ::KanMX4 clones were obtained, one of which was designated PNY2248.

Example 26

Production of Isobutanol with Decreased DHMB Yield in YMR226c Knock-Out

PNY2211 ymr226cΔ::KanMX4 transformants and a non-deletion control (PNY2211 with native YMR226c) were tested for butanol production in glucose medium by first introducing the isobutanol pathway-containing plasmids pYZ090ΔalsS (SEQ ID NO:371) and pBP915 (SEQ ID NO:182) simultaneously by the Quick and Dirty lithium acetate transformation method described in *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2005)). Plasmid selection was based on histidine and uracil auxotrophy on selection plates containing ethanol (synthetic complete medium with 1.0% ethanol -his -ura). After three to five days, several transformants showing the most robust growth were adapted to glucose medium by patching onto SD 2.0% glucose + 0.05% ethanol -his -ura and incubated 48 to 72 hours at 300 °C. Three streaks showing the most robust growth were used to inoculate a 10 mL seed culture in SD 0.2% glucose + 0.2% ethanol -his -ura in 125 mL vented flasks and grown at 30 °C, 250 rpm for approximately 24 hours. Cells were then subcultured into synthetic complete medium with 2% glucose + 0.05% ethanol –his –ura in 125 ml tightly-capped flasks and incubated 48 hours at 30 °C. Culture

2017200439 20 Jan 2017

supernatants collected after inoculation and after 48 hours incubation were analyzed by HPLC to determine production of isobutanol and by LC/MS to quantify DHMB. Controls strains were observed to produce DHMB at a molar yield of 0.03 to 0.07 mole per mole glucose. A peak corresponding to DHMB was not observed in culture supernatants of the ymr226cΔ strains, one of which was designated PNY2249.

Example 27

Identification of Genes that Encode Acetolactate Reductase (ALR) Activity Enzymes Using Yeast Knockout Library

From a knockout ("KO") collection of >6000 yeast strains derived from the strain BY4743, available from Open Biosystems® (a division of Thermo Fisher Scientific, Waltham, MA), 95 candidate dehydrogenase gene knockout strains were chosen. Starter cultures of knockout strains were grown in 96-well deepwell plates (Costar 3960, Corning Inc., Corning NY, or similar) on rich medium YPD, and subcultured at a starting OD 600nm of ~0.3 in medium containing 0.67% Yeast Nitrogen Base, 0.1% casamino acids, 2% glucose, and 0.1 M K^+-MES, pH 5.5. Samples were taken over a 5-day period for DHMB and DHIV measurements. DHIV and the two isomers of DHMB were separated and quantified by liquid chromatography-mass spectrometry ("LC/MS") on a Waters (Milford, MA) AcquityTQD system, using an Atlantis T3 (part #186003539) column. The column was maintained at 30 °C, and the flow rate was 0.5 ml/min. The A mobile phase was 0.1% formic acid in water, and the B mobile phase was 0.1% formic acid in acetonitrile. Each run consisted of 1 min at 99% A, a linear gradient over 1 min to 25% B, followed by 1 min at 99% A. The column effluent was monitored for peaks at m/z = 133 (negative ESI), with cone voltage 32.5V, by Waters ACQ_TQD (s/n QBA688) mass spec detector. The so-called "fast DHMB" typically emerged at 1.10 min, followed by DHIV at 1.2 min, and "slow" DHMB emerged at 1.75 min. Baseline separation was obtained and peak areas for DHIV were converted to μM DHIV concentrations by reference to analyses of standards solutions made from a 1M aqueous stock. These measurements showed that most of the changes in DHMB levels occurred

2017200439 20 Jan 2017

in the first 48-60 hours, so a single sample was collected at about that time in subsequent experiments. In this experiment, fast DHMB was found at much higher levels than slow DHMB, which was not always detectable. The ratio of DHIV to fast DHMB in most cultures was ~ 3, but a strain lacking the YMR226C gene consistently showed very low levels of fast DHMB, and normal DHIV, so that the DHIV/fast DHMB ratio was about 100. This suggested that YMR226Cp is the major ALR in this background.

To confirm that YMR226Cp is the major ALR in this background, the *in vitro* levels of ALR and KARI were tested in the *ymr226c* deletion strain (American Type Culture Collection (ATCC), Manassas VA, ATCC #4020812) and its parent, BY4743 (ATCC #201390; American Type Culture Collection, Manassas VA). Fifty ml tubes containing 6 ml YPD were inoculated from YPD agar plates and allowed to grow overnight (30 ºC, 250 rpm). The cells were pelleted, washed once in water, and resuspended in 1 ml yeast cytoplasm buffer (Van Eunen *et al. FEBS Journal 277*: 749-760 (2010)) containing a yeast protease inhibitor cocktail (Roche, Basel, Switzerland, Cat # 11836170001, used as directed by the vendor, 1 tablet per 10 mls of buffer). Toluene (0.02 ml, Fisher Scientific, Fair Lawn NJ) was added, and the tubes were shaken at top speed for 10 min on a Vortex Genie 2 shaker (Scientific Industries, Bohemia NY, Model G-560) for permeabilization. The tubes were placed in a water bath at 30 ºC, and substrates were added to the following final concentrations: (S)-acetolactate (made enzymatically as described below in Example 29) to 9.4 mM, NADPH (Sigma-Aldrich, St. Louis MO) 0.2 mM plus a NAD(P)H-regeneration system consisting of ~ 10 mM glucose-6-phosphate and 2.5 U/ml *Leuconostoc mesenteroides* glucose-6-phosphate dehydrogenase (Sigma, St. Louis, MO, Cat # G8404). At timed intervals, aliquots (0.15 ml) were added to 0.15 ml aliquots of 2% formic acid to stop the reaction. The samples were then analyzed for DHIV and both isomers of DHMB by LC/MS as described above; only fast DHMB and DHIV were observed. The specific activities of the two enzymes in the two strains are shown in Table 31.

2017200439 20 Jan 2017

Table 31. KARI and ALR Enzyme Activities

Strain	KARI	ALR
BY4743	1.7 mU/mg protein	20 mU/mg
YMR226C deletion strain	2.2 mU/mg protein	0.1 mU/mg

The data suggests that the YMR226C gene product accounted for >99% of the ALR activity.

Example 28

Identification of Genes that Encode Acetolactase Reductase (ALR) Activity Enzymes Using Yeast Overeexpresion Library

From a "Yeast ORF" collection of >5000 transformants of Y258 each with a plasmid carrying a known yeast gene plus a C-terminal tag, under the control of an inducible promoter (Open Biosystems®, a division of Thermo Fisher Scientific, Waltham, MA), ninety-six strains with plasmids containing genes associated with dehydrogenase activity were grown in 96-well format by adaptation of the growth and induction protocol recomended by the vendor (Open Biosystems®). The cells were pelleted and permeabilized with toluene as described above, and a concentrated substrate mix was added to give final concentrations as in Example 27. Timed samples were taken and analyzed for DHIV and both isomers of DHMB. The ratios of the ALR/KARI were calculated and compared. Strains with elevated ratios were candidates for overproduction of ALR activities. When the data for relative rates of fast DHMB and DHIV formation were displayed in a Minitab® (Microsoft Inc., Redmond, WA) boxplot, half the ratios fell between ~9-13, and most of the rest fell within 3 and 19. The exceptions identified as outliers included YER081W, YIL074C, YMR226C, YBR006W, and YOR375C, for which ratios of ALR/KARI fell between 22 and 40. In a similar analysis of relative rates of slow DHMB and DHIV formation, half the ratios fell between 9 and 11, but YMR226C, YPL275W, YER081W, AND YOL059W appeared as outliers, with ratios between 13 and 25. Thus, overexpression of YMR226C and YER081W, increased synthesis of both DHMBs. In addition, YIL074C, YBR006W, and YOR375C increased fast DHMB synthesis, and YPL275W and YOL059W increased slow DHMB synthesis. The genomic DNA

2017200439 20 Jan 2017

sequences (which may include introns) and ORF translation sequences of genes identified in overexpression are provided in Table 6.

Example 29

Inhibition of KARI by DHMB

Enzymatic Production of (S)-Acetolactate

(S)-acetolactate was used as a starting material for DHMB synthesis. (S)-acetolactate was made enzymatically, as follows. An *E. coli* TOP10 strain (Invitrogen, Carlsbad, CA) modified to express Klebsiella BudB (previously described in US Patent No. 7,851,188, which is herein incorporated by reference in its entirety; see Example 9 of that patent) under IPTG control was used as a source of enzyme. It was grown in 200-1000 ml culture volumes. For example, 200 ml was grown in Luria Broth (Mediatech, Manassas, VA) containing 0.1 mg/ml Ampicillin (Sigma, St. Louis, MO) in a 0.5 L conical flask, which was shaken at 250 rpm at 37 °C. At OD 600 ~0.4, isopropylthiogalactoside (Sigma, St. Louis, MO) was added to 0.4 mM, and growth was continued for 2 hours before the cells were collected by centrifugation, yielding ~ 1 g wet weight cells. Likewise, partial purifications were conducted at scales from ~0.5 to 5 g wet cells. For example, ~0.5 g cells were suspended in 2.5 ml buffer containing 25 mM Na-MES pH 6, broken by sonication at 0 °C, and clarified by centrifugation. Crude extract was supplemented with 0.1 mM thiamin pyrophosphate, 10 mM MgCl2, and 1 mM EDTA (all from Sigma, St. Louis, MO). Next, 0.07 ml of 10% w/v aqueous streptomycin sulfate (Sigma, St. Louis, MO) was added and the sample was heated in a 56 °C water bath for 20 min. It was clarified by centrifugation, and ammonium sulfate was added to 50% of saturation. The mixture was centrifuged, and the pellet was brought up in 0.5 ml 25 mM Na-MES, pH 6.2, and used without further characterization. Acetolactate syntheses were also conducted at various scales. A large preparation was conducted as follows: 5.5 g sodium pyruvate was dissolved in 25 mM Na-MES, pH 6.2, to ~ 45 ml and supplemented with 10 mM MgCl2, 1 mM thiamin

2017200439 20 Jan 2017

pyrophosphate, 1 mM EDTA (all from (Sigma, St. Louis, MO), 25 mM sodium acetate (Fisher Scientific, Fair Lawn NJ), and 0.25 ml of a BudB preparation. The mixture was stirred under a pH meter at room temperature. As the reaction proceeded, CO2 was evolved, and the pH rose. Pyruvic acid (Alfa, Ward Hill, MA) was added slowly via peristaltic pump to keep the pH between 6 and 7. As the pH rises, the enzyme reaction slows, but if it is allowed to fall below 6, decarboxylation of acetolactic acid becomes a problem. When the reaction was complete, the mixture was stored at -80 °C.

Synthesis of DHMB

DHMB was synthesized chemically from (S)-acetolactate. Three ml of a crude acetolactate preparation at ~0.8 M at pH ~8 was treated with 1.2 equiv $NaBH_4$ (Aldrich Chemical Co, Milwaukee, WI). The reaction was allowed to sit at room temperature overnight before being divided in two and desalted in two portions on a 60 cm x 1 cm diameter column of Biogel P-2 (Bio-Rad, Hercules, CA) using water as the mobile phase. The fractions containing mixed DHMBs were concentrated by rotary evaporation and adjusted to pH 2.2 with sulfuric acid.

The diastereomers of DHMB were separated using an HPLC system (consisting of an LKB 2249 pump and gradient controller (LKB, now a division of General Electric, Chalfont St Giles, UK) and a Hewlett-Packard (now Agilent, Santa Clara, CA) 1040A UV/vis detector) with a Waters Atlantis T3 (5 um, 4.6 x 150 mm) run at room temperature in 0.2% aqueous formic acid, pH 2.5, at a flow rate of 0.3 mL/min, with UV detection at 215 nm. "Fast" DHMB was eluted at 8.1 min and "slow" DHMB was eluted at 13.7 min. DHIV was not present. The pooled fractions were taken nearly to dryness, and coevaporated with toluene to remove residual formic acid. The residue was then dissolved in water and made basic with triethylamine (Fisher, Fair Lawn, NJ).

Concentration Determination and Absolute Structure of DHMB

The concentration of purified DHMB solutions was determined as follows. The concentration was estimated based on the mmol acetolactate used in the $NaBH_4$ reduction. To portions of the DHMBs, a known quantity

2017200439 20 Jan 2017

of sodium benzoate (made by dissolving solid benzoic acid (ACS grade, Fisher Scientific, Fair Lawn, NJ) in aqueous NaOH)) was added to give two-component mixtures in (approximately) equimolar amounts. A similar sample of DHIV was also prepared from the solid sodium salt obtained via custom synthesis (Albany Molecular Research, Albany NY). The samples were coevaporated several times with D_2O (Aldrich, Milwaukee, WI) and redissolved in D_2O. Integrated proton NMR spectra were obtained and used to determine the mole ratio of DHIV or DHMB to benzoate. Comparison of the NMR spectra of the DHMBs with the literature spectra for the free acids in $CDCl_3$ (Kaneko *et al.*, *Phytochemistry 39*: 115-120 (1995)) showed that fast DHMB was the erythro isomer. Since enzymatically synthesized acetolactate has the (S) configuration at C-2, the fast DHMB has the 2S, 3S configuration. Slow DHMB has the threo 2S, 3R configuration.

Dilutions of the NMR samples were also analyzed by LC/MS using separately prepared benzoic acid solutions as standards. Benzoic acid, DHIV, and the two isomers of DHMB were separated and quantified by LC/MS on a Waters (Milford, MA) AcquityTQD system, using an Atlantis T3 (part #186003539) column, as described above. Benzoic acid was detected at m/z=121 (negative ESI), and emerged at 2.05 min. The concentration of benzoate in the mixtures was within experimental uncertainty of the expected value. The experiment also showed that either isomer of DHMB had ~80% of the sensitivity of DHIV in LC/MS *(i.e.*, MS peak area observed /nmol injected) throughout the response range of the instrument. Thus, if a DHIV standard is used to quantify DHMB found in cell extracts or in enzymatic reactions, the apparent DHMB concentrations need to be multiplied by 1.25.

Measuring Inhibition of KARI by DHMB

Purified KARI encoded by genes either from Lactococcus lactis (SEQ ID NO: 864), a derivative of Pseudomonas fluorescens KARI known as JEA1 (SEQ ID NO: 799; U.S. Appl. Pub No. 2010/0197519, which is herein incorporated by reference in its entirety), or a variant of Anaerostipes caccae KARI known as K9D3 (SEQ ID NO:788), were tested

2017200439 20 Jan 2017

for their sensitivity to DHMB inhibition in spectrophotometric assays in a Shimadzu (Kyoto, Japan) UV160U instrument with a TCC240A temperature control unit, set at 30 ºC. The buffer was 0.1 M K+ Hepes, pH 6.8, containing 10 mM $MgCl_2$ and 1 mM EDTA. NADPH was present at 0.2 mM, and racemic acetolactate was present at either 3mM or 0.725 mM (S) isomer. The rate of NADPH oxidation in the presence and absence of either fast or slow DHMB was measured. V_{max} for each sample was calculated from the observed rate and the known acetolactate K_M using the Michaelis-Menten equation. A volumetric K_i was estimated for each measurement in the presence of DHMB using the Michaelis-Menten equation as modified for competitive inhibition vs. acetolactate (the K_M term in the Michaelis-Menten equation is multiplied by $(1+[I]/ K_i)$, and the equation is solved for K_i. The results were converted to mM upon completion of the NMR experiment and are shown in Table 32.

Table 32. K_I Values for KARI Inhibition by DHMB Isomers

Strain	Fast DHMB	Slow DHMB
JEA1	0.23 mM	0.23 mM
K9D3	0.3 mM	0.2 mM
L. lactis	2.8 mM	2.3 mM

Example 30

Inhibition of DHAD by DHMB

Purified dihydroxyacid dehydratase (DHAD) from *Staphococcus mutans* was tested for inhibition of conversion of dihydroxyisovalerate (DHIV) to 2-ketoisovalerate (2-KIV) by DHMB by using a modification of a colorimetric assay as described by Szamosi *et al., Plant Phys. 101*: 999-1004 (1993). The assay took place in a 2 mL Eppendorf tube placed in a heating block maintained at 30 °C. The assay mixture had a final volume of 0.8 mL containing 100 mM Hepes-KOH buffer, pH 6.8, 10 mM $MgCl_2$, 0.5-10 mM DHIV, 0-40 mM DHMB, and 18 µg DHAD. The assay was initiated by adding a 10x concentrated stock of substrate. Samples were removed (0.35 mL) at times 0.1 and 30 minutes, and the reaction was stopped by mixing into 0.35 mL 0.1 N HCl with 0.05% 2,4-dinitrophenylhydrazine (Aldrich) in a second Eppendorf tube. After incubating 30 minutes at room temperature, 0.35 mL of 4N NaOH was

2017200439 20 Jan 2017

added to the mixture, mixed, and centrifuged at 15,000xG for 2 minutes in a centrifuge (Beckman-Coulter Microfuge 18). The absorbance of the solution at 540 nm was then measured in a 1 cm pathlength cuvette using a Cary 300 Bio UV-Vis spectrophometer (Varian). Based on a standard curve using authentic 2-KIV (Fluka), 1 OD absorbance at 540 nm is produced by 0.28 mM 2-KIV. The rate of 2-KIV formation was measured in the presence and absence of either fast or slow DHMB. Both forms of DHMB behaved liked competitive inhibitors of DHIV. Their inhibition constants (Ki) were calculated from the Michaelis-Menten equation for simple competitive inhibition: $v = S*V_{max}/(S+K_M*(1+I/K_i))$, where v is the measured rate of 2-KIV formation, S is the initial concentration of DHIV, V_{max} is the maximum rate calculated from the observed rate at 10 mM DHIV and no DHMB, K_M is a previously measured constant of 0.5 mM, and I is the concentration of DHMB. The fast and slow isomers of DHMB had calculated inhibition constants of 7 mM and 5 mM, respectively.

Example 31

Identification of YMR226C Homologs

Homologs of the YMR226C gene of *Saccharomyces cerevisiae* were sought by BLAST searches of the GenBank non-redundant nucleotide database (blast.ncbi.nlm.nih.gov/Blast.cgi), the Fungal Genomes BLAST Search Tool at the Saccharomyces Genome Database (www.yeastgenome.org/cgi-bin/blast-fungal.pl), and the BLAST Tool of the Genolevures Project (genolevures.org/blast.html#). Unique sequences from 18 yeast species showing high sequence identity to YMR226C were identified, and the complete ORF for these genes was recovered from the accessioned record in the associated database. The polypeptide sequences encoded by these ORFs were determined by the Translation feature of Vector NTI (Invitrogen, Carlsbad CA). The polynucleotide and polypeptide sequences are shown below in Table 33. The yeast species, nucleotide database accession number, and DNA and protein sequences are given in the Table. The *S. kluyveri* sequence is in the Genolevures database under the accession number given; the others are in GenBank. The percent identities between the sequences are shown in Table 34.

2017200439 20 Jan 2017

The 18 ORFs were aligned using AlignX (Vector NTI; the gene encoding a putative NADP+-dependent dehydrogenase from *Neurospora crassa* (XM_957621, identified in the GenBank BLAST search using the YMR226C nucleotide sequence) was used as an outgroup. The resulting phylogenetic tree is shown in Figure 11, and a sequence alignment is shown in Figure 12.

The sequence identity of these homologs to YMR226C ranges from a minimum of 55% (*Yarrowia lipolytica* and *Schizosaccharomyces pombe*) to a maximum of 90% (*S. paradoxus*). A BLAST search also revealed a cDNA from *S. pastorianus* (accession number CJ997537) with 92% sequence identity over 484 base pairs, but since this species is a hybrid between *S. bayanus* (whose YMR226C homolog shows 82% identity to the *S. cerevisiae* sequence), and because only a partial ORF sequence was available, this sequence was not included in the comparison. When the YMR226C sequence from the canonical laboratory strain S288C was compared with the sequences from 12 other strains of *S. cerevisiae*, only 4 single-nucleotide polymorphisms are found (sequence identity 99.5%), indicating that this is a highly-conserved gene in that species.

Table 33: YMR226C Yeast Homologs

Species	Accession # (Date of database accession)	Nucleic acid SEQ ID NO:	Amino acid SEQ ID NO:
Saccharomyces paradoxus	AABY01000127 (3/7/2011)	698	699
Saccharomyces bayanus	AACA01000631 (3/7/2011)	700	701
Saccharomyces castellii	AACF01000116 (3/7/2011)	702	703
Saccharomyces mikatae	AACH01000019 (3/7/2011)	704	705
Ashbya gossypii	AE016819 (3/7/2011)	706	707
Candida glabrata	CR380959 (3/7/2011)	708	709
Debaryomyces hansenii	CR382139 (3/7/2011)	710	711
Scheffersomyces	XM_001387479	712	713

20172004439 20 Jan 2017

Species	Accession # (Date of database accession)	Nucleic acid SEQ ID NO:	Amino acid SEQ ID NO:
stipitis (formerly *Pichia stipitis)*	(3/7/2011)		
Meyerozyma guilliermondii (formerly *Pichia guilliermondii)*	XM_001482184 (3/7/2011)	714	715
Vanderwaltozyma polyspora (formerly *Kluyveromyces polysporus)*	XM_001645671 (3/4/2011)	716	717
Candida dubliniensis	XM_002419771 (3/7/2011)	718	719
Zygosaccharomyces rouxii	XM_002494574 (3/7/2011)	720	721
Lachancea thermotolerans (formerly *Kluyveromyces thermotolerans)*	XM_002553230 (3/4/2011)	722	723
Kluyveromyces lactis	XM_451902 (3/4/2011)	724	725
Saccharomyces kluyveri	SAKL0H04730 (3/7/2011)	726	727
Yarrowia lipolytica	XM_501554 (3/8/2011)	728	729
Schizosaccharomyces pombe	NM_001018495 (3/8/2011)	730	731

Table 34: YMR226C Homolog Percent Identity

Species	Sm	Sb	Sca	Ag	Dh	Ss	Mg	Cd	Cg	Vp	Sk	Kl	Lt	Zr	Sce	Sp	Yl	Nc
Spa	88	82	70	64	62	62	58	57	67	68	68	69	68	68	90	55	55	56
Sm		82	70	64	60	62	58	56	67	69	68	70	68	69	86	57	56	57
Sb			71	63	59	62	58	53	67	66	68	70	69	67	82	56	56	58
Sca				60	62	61	60	59	65	69	69	71	64	70	69	57	53	54
Ag					56	60	57	54	59	61	62	62	62	62	63	54	55	55
Dh						64	62	61	61	63	62	61	59	63	62	57	57	53
Ss							68	64	61	62	62	64	62	63	62	56	58	58
Mg								60	57	58	60	60	59	62	59	57	57	56
Cd									57	62	59	60	54	60	58	57	53	49
Cg										69	70	68	67	67	66	55	56	55
Vp											71	72	67	70	71	58	52	51
Sk												77	71	72	69	53	54	54
Kl													71	72	71	56	52	54
Lt														69	69	53	60	58
Zr															69	58	55	55
Sce																55	55	56
Spo																	58	60

20 Jan 2017 2017200439

Yl																	61
Nc																	

Table 34 Key: *Saccharomyces paradoxus* ("Spa"); *Saccharomyces mikatae* ("Sm"); *Saccharomyces bayanus* ("Sb"); *Saccharomyces castellii* ("Sca"); *Ashbya gossypii* ("Ag"); *Debaryomyces hansenii* ("Dh"); *Scheffersomyces stipitis* ("Ss"); *Meyerozyma guilliermondii* ("Mg"); *Candida dubliniensis* ("Cd"); *Candida glabrata* ("Cg"); *Vanderwaltozyma polyspora* ("Vp"); *Saccharomyces kluyveri* ("Sk"); *Kluyveromyces lactis* ("Kl"); *Lachancea thermotolerans* ("Lt"); *Zygosaccharomyces rouxii* ("Zr"); *Saccharomyces cerevisiae* ("Sce"); *Schizosaccharomyces pombe* ("Spo"); *Yarrowia lipolytica* ("Yl"); *Neurospora crassa* ("Nc")

Example 32 (Prophetic)

Screening of aldehyde dehydrogenases from *S. cerevisiae* for ability to convert isobutyraldehyde to isobutyric acid using enzymatic assays

This example demonstrates a method to determine which endogenous aldehyde dehydrogenases from *S. cerevisiae* can enzymatically convert isobutyraldehyde to isobutyric acid.

S. cerevisiae strains containing individual disruptions in the known aldehyde dehydrogenase enzymes are obtained from ATCC: BY4741 Δ*ald2::kanMX4* (ATCC #4000753); BY4741 Δ*ald3::kanMX4* (ATCC #4000752); BY4741 Δ*ald4::kanMX4* (ATCC #4001671); BY4741 Δ*ald5::kanMX4* (ATCC #4000213); and BY4741 Δ*ald6::kanMX4* (ATCC #4002767).

The deletion strains above are first grown overnight in tubes containing 5 ml YPD media at 30 °C. The 5 ml overnight cultures are transferred into 100 ml of medium in a 500 ml flask and incubated at 30 °C shaking at 220 rpm. The cultures are harvested when they reach 1 to 2 O.D. at 600 nm. The samples are washed with 10 ml of 20 mM Tris (pH 7.5) and then are resuspended in 1 ml of the same Tris buffer. The samples are transferred into 2.0 ml tubes containing 0.1 mm silica (Lysing Matrix B, MP biomedicals). The cells are then broken in a bead-beater (BIO101). The supernatant is obtained by centrifugation in a microfuge at

2017200439 20 Jan 2017

13,000 rpm at 4 °C for 30 minutes. Typically, 0.06 to 0.1 mg of crude extract protein is used in a single assay. Protein in the crude extracts was determined by Bradford assay with Coomassie stain.

Aldehyde dehydrogenase activity is measured following a protocol given by Sigma-Aldrich and by Bostian and Betts (Bostian, K.A. and Betts, G.F. (1978) Biochemical Journal 173, 773-786). Crude extracts from the deletion strains above and commercially available aldehyde dehydrogenase are tested using this method.

An alternative assay is to add isobutyraldehyde at concentrations from 1 to 30 mM to approximately 0.1 mg of crude extract protein in a sealed glass GC vials which are incubated at 30 °C for 30 minutes. The extracts are then centrifuged through 0.22 μm spin filters (Corning, Cat #8169) at 3000 rpm for 3 minutes, and the filtrate is transferred to a GC vial for GC-MS analysis. Isobutyraldehyde and isobutyric acid are detected.

The GC method utilizes an Agilent 7890 GC equipped with a 5975 mass spectrometer for detection, and a DB-1701 column (30 m x 0.25 mm ID, 0.25 μm film) from Agilent (Santa Clara, CA). The carrier gas is helium at a constant flow rate of 1.0 mL/min; injector split is 1:10 at 250 °C; oven temperature is 40 °C for 1 min, 40 °C to 120 °C at 10 °C/min, and 120 °C to 240 °C at 30 °C/min. MS detection is used in full scan mode for identification and quantitation of isobutyraldehyde and isobutyric acid. Calibrated standard curves are generated for the following compounds: isobutyraldehyde, isobutyric acid, and isobutanol.

Example 33

Construction of expression vectors for isobutanol pathway gene expression in *S. cerevisiae*

pLH475-JEA1 construction

The pLH475-JEA1 plasmid (SEQ ID NO: 419) was constructed for expression of ALS and KARI in yeast. pLH475-JEA1 is a pHR81 vector (ATCC #87541) containing the following chimeric genes:1) the CUP1 promoter (SEQ ID NO: 789), acetolactate synthase coding region from

2017200439 20 Jan 2017

Bacillus subtilis (AlsS; SEQ ID NO: 790; protein SEQ ID NO: 791) and CYC1 terminator 2 (SEQ ID NO: 792); 2) an ILV5 promoter (SEQ ID NO: 793), Pf5.IlvC-JEA1 coding region and ILV5 terminator (SEQ ID NO: 794); and 3) the FBA1 promoter (SEQ ID NO: 795), *S. cerevisiae* KARI coding region (ILV5; SEQ ID NO: 796; protein SEQ ID NO: 797) and CYC1 terminator (SEQ ID NO: 798).

The Pf5.IlvC-JEA1 coding region is a sequence encoding KARI derived from *Pseudomonas fluorescens* but containing mutations, that was described in U.S. Patent Application Publication Nos. 2009/0163376 and 2010/0197519, which are herein incorporated by reference in their entireties. The Pf5.IlvC-JEA1 encoded KARI (nucleic acid and amino acid SEQ ID NOs: 799 and 800) has the amino acid changes as compared to the natural *Pseudomonas fluorescens* KARI.

.

Expression Vector pLH468

The pLH468 plasmid (SEQ ID NO:139) was constructed for expression of DHAD, KivD, and HADH in yeast.

Coding regions for *L. lactis k*etoisovalerate decarboxylase (KivD) and Horse liver alcohol dehydrogenase (HADH) were synthesized by DNA2.0 based on codons that were optimized for expression in *Saccharomyces cerevisiae* (SEQ ID NOs: 801 and 802, respectively) and provided in plasmids pKivDy-DNA2.0 and pHadhy-DNA2.0. The encoded proteins are SEQ ID NOs: 803 and 804, respectively. Individual expression vectors for KivD and HADH were constructed. To assemble pLH467 (pRS426::P_{TDH3}-*kivD*y-*TDH3*t), vector pNY8; also named pRS426.GPD-ald-GPDt, described in US Patent App. Pub. US2008/0182308, Example 17, which is herein incorporated by reference) was digested with *AscI* and *SfiI* enzymes, thus excising the *GPD* promoter and the ald coding region. A *TDH3* promoter fragment (SEQ ID NO: 805) from pNY8 was PCR amplified to add an AscI site at the 5' end, and an SpeI site at the 3' end, using 5' primer OT1068 and 3' primer OT1067 (SEQ ID NOs: 806 and 807). The AscI/SfiI digested pNY8 vector fragment was ligated with the TDH3 promoter PCR product digested with AscI and

2017200439 20 Jan 2017

SpeI, and the SpeI-SfiI fragment containing the codon optimized kivD coding region isolated from the vector pKivD-DNA2.0. The triple ligation generated vector pLH467 (pRS426::PTDH3-kivDy-TDH3t). pLH467 (SEQ ID NO: 808) was verified by restriction mapping and sequencing.

pLH435 (pRS425::PGPM1-Hadhy-ADH1t) was derived from vector pRS425::GPM-sadB (SEQ ID NO: 809) which is described in U.S. Appl. Pub. No. 2009/0305363, Example 3, which is herein incorporated by reference. pRS425::GPM-sadB is the pRS425 vector (ATCC #77106) with a chimeric gene containing the GPM1 promoter (SEQ ID NO: 810), coding region from a butanol dehydrogenase of *Achromobacter xylosoxidans* (sadB; DNA SEQ ID NO: 811; protein SEQ ID NO: 812), and ADH1 terminator (SEQ ID NO: 444). pRS425::GPMp-sadB contains *BbvI* and *PacI* sites at the 5' and 3' ends of the sadB coding region, respectively. A *NheI* site was added at the 5' end of the sadB coding region by site-directed mutagenesis using primers OT1074 and OT1075 (SEQ ID NO: 813 and 814) to generate vector pRS425-GPMp-sadB-NheI, which was verified by sequencing. pRS425::P_{GPM1}-*sadB-Nhe*I was digested with *NheI* and *PacI* to drop out the *sadB* coding region, and ligated with the *NheI-PacI* fragment containing the codon optimized HADH coding region from vector pHadhy-DNA2.0 to create pLH435 (SEQ ID NO: 815).

To combine KivD and HADH expression cassettes in a single vector, yeast vector pRS411 (ATCC # 87474) was digested with *SacI* and *NotI*, and ligated with the *SacI-SalI* fragment from pLH467 that contains the P_{TDH3}-*kivD*y-*TDH3*t cassette together with the *SalI-NotI* fragment from pLH435 that contains the P_{GPM1}-H*adh*y-*ADH1*t cassette in a triple ligation reaction. This yielded the vector pRS411::P_{TDH3}-*kivD*y-P_{GPM1}-H*adh*y (pLH441, SEQ ID NO: 816), which was verified by restriction mapping.

In order to generate a co-expression vector for all three genes in the lower isobutanol pathway: *ilvD*, *kivD*y and H*adh*y, we used pRS423 FBA ilvD(Strep) (SEQ ID NO: 817), as the source of the IlvD gene. This shuttle vector contains an F1 origin of replication (nt 1423 to 1879) for maintenance in *E. coli* and a 2 micron origin (nt 8082 to 9426) for replication in yeast. The vector has an FBA1 promoter (nt 2111 to 3108;

2017200439 20 Jan 2017

SEQ ID NO: 795) and FBA1 terminator (nt 4861 to 5860; SEQ ID NO: 818). In addition, it carries the His marker (nt 504 to 1163) for selection in yeast and ampicillin resistance marker (nt 7092 to 7949) for selection in *E. coli*. The ilvD coding region (nt 3116 to 4828; SEQ ID NO: 819; protein SEQ ID NO: 820) from *Streptococcus mutans* UA159 (ATCC #700610) is between the FBA1 promoter and FBA1 terminator forming a chimeric gene for expression. In addition there is a lumio tag fused to the ilvD coding region (nt 4829-4849).

The first step was to linearize pRS423 FBA ilvD(Strep) (also called pRS423-FBA(SpeI)-IlvD(*Streptococcus mutans*)-Lumio) with *SacI* and *SacII* (with *SacII* site blunt ended using T4 DNA polymerase), to give a vector with total length of 9,482 bp. The second step was to isolate the *kivD*y-h*ADH*y cassette from pLH441 with *SacI* and *KpnI* (with *KpnI* site blunt ended using T4 DNA polymerase), which gives a 6,063 bp fragment. This fragment was ligated with the 9,482 bp vector fragment from pRS423-FBA(SpeI)-IlvD(*Streptococcus mutans*)-Lumio. This generated vector pLH468 (pRS423::P_{FBA1}-*ilvD*(Strep)Lumio-*FBA1*t-P_{TDH3}-*kivD*y-*TDH3*t-P_{GPM1}-h*adh*y-*ADH1*t), which was confirmed by restriction mapping and sequencing.

Example 34

Construction of *S. cerevisiae* host strain containing modifications in pyruvate decarboxylase and hexokinase 2

This example describes insertion-inactivation of endogenous PDC1, PDC5, and PDC6 genes of *S. cerevisiae*. PDC1, PDC5, and PDC6 genes encode the three major isozymes of pyruvate decarboxylase. Hexokinase 2, which is responsible for phosphorylation of glucose and transcriptional repression, is also inactivated. The resulting PDC/HXK2 inactivation strain was used as a host for expression vectors pLH475-JEA1 and pLH468.

Construction of pdc6:: P_{GPM1}-sadB integration cassette and PDC6 deletion:

2017200439 20 Jan 2017

A *pdc6::*P_{GPM1}*-sadB-ADH1t-URA3*r integration cassette was made by joining the *GPM-sadB-ADH*t segment (SEQ ID NO: 821) from pRS425::*GPM-sadB* (described in U.S. Patent App. Pub. No. 2009/0305363, incorporated by reference) to the *URA3*r gene from pUC19-URA3r. pUC19-URA3r (SEQ ID NO: 822) contains the *URA3* marker from pRS426 (ATCC # 77107) flanked by 75 bp homologous repeat sequences to allow homologous recombination *in vivo* and removal of the *URA3* marker. The two DNA segments were joined by SOE PCR (as described by Horton *et al.* (1989) *Gene* 77:61-68) using as template pRS425::*GPM-sadB* and pUC19-URA3r plasmid DNAs, with Phusion DNA polymerase (New England Biolabs Inc., Beverly, MA; catalog no. F-540S) and primers 114117-11A through 114117-11D (SEQ ID NOs: 823, 824, 825, 826), and 114117-13A and 114117-13B (SEQ ID NOs: 827 and 828).

The outer primers for the SOE PCR (114117-13A and 114117-13B) contained 5' and 3' ~50 bp regions homologous to regions upstream and downstream of the *PDC6* promoter and terminator, respectively. The completed cassette PCR fragment was transformed into BY4700 (ATCC # 200866) and transformants were maintained on synthetic complete media lacking uracil and supplemented with 2% glucose at 30 °C using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202). Transformants were screened by PCR using primers 112590-34G and 112590-34H (SEQ ID NOs: 829 and 830), and 112590-34F and 112590-49E (SEQ ID NOs: 831 and 832) to verify integration at the *PDC*6 locus with deletion of the *PDC6* coding region. The *URA3r* marker was recycled by plating on synthetic complete media supplemented with 2% glucose and 5-FOA at 30°C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD -URA media to verify the absence of growth. The resulting identified strain has the genotype: BY4700 *pdc6::*P_{GPM1}*-sadB-ADH1*t.

Construction of pdc1:: P_{PDC1}-ilvD integration cassette and PDC1 deletion:

A *pdc1::* P_{PDC1}*-ilvD-FBA1t-URA3*r integration cassette (SEQ ID NO: 833) was made by joining the *ilvD-FBA1t* segment from pLH468 to the

2017200439 20 Jan 2017

*URA3*r gene from pUC19-URA3r by SOE PCR (as described by Horton *et al.* (1989) *Gene* 77:61-68) using as template pLH468 and pUC19-URA3r plasmid DNAs, with Phusion DNA polymerase (New England Biolabs Inc., Beverly, MA; catalog no. F-540S) and primers 114117-27A through 114117-27D (SEQ ID NOs: 823, 824, 825, 826).

The outer primers for the SOE PCR (114117-27A and 114117-27D) contained 5' and 3' ~50 bp regions homologous to regions downstream of the *PDC1* promoter and downstream of the *PDC1* coding sequence. The completed cassette PCR fragment was transformed into BY4700 *pdc6*::P_{GPM1}-*sadB-ADH1*t and transformants were maintained on synthetic complete media lacking uracil and supplemented with 2% glucose at 30 °C using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202). Transformants were screened by PCR using primers 114117-36D and 135 (SEQ ID NOs 834 and 835), and primers 112590-49E and 112590-30F (SEQ ID NOs 832 and 836) to verify integration at the *PDC*1 locus with deletion of the *PDC1* coding sequence. The *URA3r* marker was recycled by plating on synthetic complete media supplemented with 2% glucose and 5-FOA at 30 °C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD -URA media to verify the absence of growth. The resulting identified strain "NYLA67" has the genotype: BY4700 *pdc6*:: P_{GPM1}-*sadB-ADH1*t *pdc1*:: P_{PDC1}-*ilvD-FBA1t.*

HIS3 deletion

To delete the endogenous HIS3 coding region, a *his3::URA3r2* cassette was PCR-amplified from URA3r2 template DNA (SEQ ID NO: 837). URA3r2 contains the *URA3* marker from pRS426 (ATCC # 77107) flanked by 500 bp homologous repeat sequences to allow homologous recombination *in vivo* and removal of the *URA3* marker. PCR was done using Phusion DNA polymerase and primers 114117-45A and 114117-45B (SEQ ID NOs: 838 and 839) which generated a ~2.3 kb PCR product. The *HIS3* portion of each primer was derived from the 5' region upstream of the *HIS3* promoter and 3' region downstream of the coding region such

20 Jan 2017 2017200439

that integration of the *URA3r2* marker results in replacement of the *HIS3* coding region. The PCR product was transformed into NYLA67 using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants were selected on synthetic complete media lacking uracil and supplemented with 2% glucose at 30 °C. Transformants were screened to verify correct integration by replica plating of transformants onto synthetic complete media lacking histidine and supplemented with 2% glucose at 30 °C. The *URA3r* marker was recycled by plating on synthetic complete media supplemented with 2% glucose and 5-FOA at 30 °C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD-URA media to verify the absence of growth. The resulting identified strain, called NYLA73, has the genotype: BY4700 *pdc6::*P_{GPM1}*-sadB-ADH1*t *pdc1::* P_{PDC1}*-ilvD-FBA1t* *Δhis3*.

Deletion of HXK2:

A *hxk2::*URA3r cassette was PCR-amplified from URA3r2 template (described above) using Phusion DNA polymerase and primers 384 and 385 (SEQ ID NOs: 840and 841) which generated a ~2.3 kb PCR product. The *HXK2* portion of each primer was derived from the 5' region upstream of the *HXK2* promoter and 3' region downstream of the coding region such that integration of the *URA3*r2 marker results in replacement of the *HXK2* coding region. The PCR product was transformed into NYLA73 using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants were selected on synthetic complete media lacking uracil and supplemented with 2% glucose at 30 °C. Transformants were screened by PCR to verify correct integration at the *HXK2* locus with replacement of the *HXK2* coding region using primers N869 and N871 (SEQ ID NOs: 842 and 843). The *URA3r2* marker was recycled by plating on synthetic complete media supplemented with 2% glucose and 5-FOA at 30 °C following standard protocols. Marker removal was confirmed by patching colonies from the 5-FOA plates onto SD -URA media to verify the

2017200439 20 Jan 2017

absence of growth, and by PCR to verify correct marker removal using primers N946 and N947 (SEQ ID NOs: 844and 845). The resulting identified strain named NYLA83 has the genotype: BY4700 *pdc6::* P_{GPM1}-*sadB-ADH1*t *pdc1::* P_{PDC1}-*ilvD-FBA1t Δhis3 Δhxk2.*

Construction of pdc5::kanMX integration cassette and PDC5 deletion

A *pdc5::kanMX4* cassette was PCR-amplified from strain YLR134W chromosomal DNA (ATCC No. 4034091) using Phusion DNA polymerase and primers PDC5::KanMXF and PDC5::KanMXR (SEQ ID NOs: 846 and 847) which generated a ~2.2 kb PCR product. The *PDC5* portion of each primer was derived from the 5' region upstream of the *PDC5* promoter and 3' region downstream of the coding region such that integration of the *kanMX4* marker results in replacement of the *PDC5* coding region. The PCR product was transformed into NYLA83, and transformants were selected and screened as described above. The identified correct transformants named NYLA84 have the genotype: BY4700 *pdc6::* P_{GPM1}-*sadB-ADH1*t *pdc1::* P_{PDC1}-*ilvD-FBA1t Δhis3 Δhxk2 pdc5::kanMX4.* Plasmid vectors pLH468 and pLH475-JEA1 were simultaneously transformed into strain NYLA84 (BY4700 *pdc6::*P_{GPM1}-*sadB-ADH1*t *pdc1::*P_{PDC1}-*ilvD-FBA1t Δhis3 Δhxk2 pdc5::kanMX4*) using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) and the resulting strain was maintained on synthetic complete media lacking histidine and uracil, and supplemented with 1% ethanol at 30 °C.

Example 35 (Prophetic)

Construction of *S. cerevisiae* host strain containing modifications in pyruvate decarboxylase, hexokinase 2, and aldehyde dehydrogenase.

This example describes inactivation of one or more aldehyde dehydrogenases that abolish or reduce formation of isobutryic acid. Disruption of *ALD2* is used as an example, but the method could be used for disruption of any aldehyde dehydrogenase. The resulting NYLA84-derived strain is used as a host for expression vectors pLH475-JEA1 and pLH468.

2017200439 20 Jan 2017

Deletion of *ALD2:*

A *ald2::*URA3r cassette is PCR-amplified from URA3r2 template (described above) using Phusion DNA polymerase and primers T001 and T002 (SEQ ID NOs: 848 and 849) which generates a ~2.3 kb PCR product. The *ALD2* portion of each primer is derived from the 5' sequence and 3' sequence of the *ALD2* gene such that integration of the *URA3*r2 marker results in replacement of the *ALD2* coding region. The PCR product is transformed into NYLA84 using standard genetic techniques (*Methods in Yeast Genetics*, 2005, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 201-202) and transformants are selected on synthetic complete media lacking uracil and supplemented with 1% ethanol at 30 °C. Transformants are screened by PCR to verify correct integration at the *ALD2* locus with replacement of the *ALD2* coding region using primers T003 and T004 (SEQ ID NOs: 850 and 851). The *URA3r2* marker is recycled by plating on synthetic complete media supplemented with 1% ethanol and 5-FOA at 30 °C following standard protocols. Marker removal is confirmed by patching colonies from the 5-FOA plates onto SE -URA media to verify the absence of growth, and by PCR to verify correct marker removal using primers T004 and T005 (SEQ ID NOs: 851 and 852). The resulting identified strain is named NYLA84 Δald2 and is of the genotype: BY4700 *pdc6::* P_{GPM1}*-sadB-ADH1*t *pdc1::* P_{PDC1}*-ilvD-FBA1t Δhis3 Δhxk2 Δald2.*

Example 36

Production of isobutanol in recombinant *S. cerevisiae* in NYLA84

HPLC Method

Analysis for glucose and fermentation by-product composition is well known to those skilled in the art. For example, one high performance liquid chromatography (HPLC) method utilizes a Shodex SH-1011 column with a Shodex SH-G guard column (both available from Waters Corporation, Milford, MA), with refractive index (RI) detection. Chromatographic separation is achieved using 0.01 M H_2SO_4 as the

2017200439 20 Jan 2017

mobile phase with a flow rate of 0.5 mL/min and a column temperature of 50 °C. Isobutanol retention time is 47.6 minutes.

Production of isobutanol in recombinant *S. cerevisiae* in NYLA84

The purpose of this example is to describe the production of isobutanol in the yeast strain NYLA84. The yeast strain comprises deletions of PDC1, PDC5, and PDC6, genes encoding three isozymes of pyruvate decarboxylase, and deletion of HXK2 encoding hexokinase 2. The strain also contains constructs for heterologous expression of AlsS (acetolactate synthase), KARI (keto acid reductoisomerase), DHAD (dihydroxy acid dehydratase), KivD (ketoisovalerate decarboxylase), SadB (secondary alcohol dehydrogenase), and HADH (horse liver alcohol dehydrogenase).

Strain construction

Plasmids pLH468 and pLH475-JEA1 were introduced into NYLA84, by standard PEG/lithium acetate-mediated transformation methods. Transformants were selected on synthetic complete medium lacking glucose, histidine and uracil. Ethanol (1% v/v) was used as the carbon source. After three days, transformants were patched to synthetic complete medium lacking histidine and uracil supplemented with both 2% glucose and 0.5% ethanol as carbon sources. Freezer vials were made by dilution of log-phase cultures with 45% glycerol to a final glycerol concentration of 15% (w/v).

Production of isobutanol

Freezer vials were used to inoculate 80 ml of synthetic complete medium lacking histidine and uracil supplemented with both 2% glucose and 0.5% ethanol as carbon sources.

Fermentation conditions:

A 1 liter fermenter was prepared and sterilized with 0.4L water. After cooling, filter sterilized medium was added to give the following final concentrations in 800 mL post-inoculation:

Medium (final concentration):

6.7 g/L, Yeast Nitrogen Base w/o amino acids (Difco)

2017200439 20 Jan 2017

2.8 g/L, Yeast Synthetic Drop-out Medium Supplement Without Histidine, Leucine, Tryptophan and Uracil (Sigma Y2001)

20 mL/L of 1% (w/v) L-Leucine

4 mL/L of 1% (w/v) L-Tryptophan

1 mL/L ergosterol/tween/ethanol solution

0.2 mL/L Sigma DF204

10 g/L glucose

The fermenter was set to control at pH 5.5 with KOH, 30% dO, temperature 30 °C, and airflow of 0.2 SLPM (or, 0.25 vvm). At inoculation, the airflow was set to 0.01 SLPM initially, then increased to 0.2 SLPM once growth was established. Glucose was maintained at 5-15 g/L throughout by manual addition.

Air was continuously supplied to the fermentor at 0.25 vvm. Continuous aeration led to significant stripping of isobutanol from the aqueous phase of the fermentor. To quantify the loss of isobutanol due to stripping, the off-gas from the fermentor was directly sent to a mass spectrometer (Prima dB mass spectrometer, Thermo Electron Corp., Madison, WI) to quantify the amount of isobutanol in the gas stream. The isobutanol peaks at mass to charge ratios of 74 or 42 were monitored continuously to quantify the amount of isobutanol in the gas stream. Glucose and organic acids in the aqueous phase were monitored during the fermentation using HPLC. Glucose was also monitored quickly using a glucose analyzer (YSI, Inc., Yellow Springs, OH). Isobutanol in the aqueous phase was quantified by HPLC as described in HPLC Method herein above after the aqueous phase was removed periodically from the fermentor. The effective titer, corrected for the isobutanol lost due to stripping, was 7.5 g/L. The titer of isobutyric acid was 1.28 g/L (Figure 14).

Example 37 (Prophetic)

Production of isobutanol in recombinant *S. cerevisiae* in NYLA84 Δ*ald2*.

The purpose of this example is to describe the production of isobutanol in the yeast strain NYLA84. Disruption of *ALD2* is used as an

2017200439 20 Jan 2017

example, but the example could be used for disruption of any aldehyde dehydrogenase. The NYLA84 Δ*ald2* yeast strain comprises deletions of *PDC1, PDC5,* and *PDC6,* genes encoding three isozymes of pyruvate decarboxylase, deletion of *HXK2* encoding hexokinase 2, and deletion of *ALD2* encoding aldehyde dehydrogenase. The strain also contains constructs for heterologous expression of AlsS (acetolactate synthase), KARI (keto acid reductoisomerase), DHAD (dihydroxy acid dehydratase), KivD (ketoisovalerate decarboxylase), and SadB (secondary alcohol dehydrogenase).

Strain construction

Plasmids pLH468 and pLH475-JEA1 are introduced into NYLA84 Δ*ald2*, by standard PEG/lithium acetate-mediated transformation methods. Transformants are selected on synthetic complete medium lacking glucose, histidine and uracil. Ethanol (1% v/v) is used as the carbon source. After three days, transformants are patched to synthetic complete medium lacking histidine and uracil supplemented with both 2% glucose and 0.5% ethanol as carbon sources. Freezer vials are made by dilution of log-phase cultures with 45% glycerol to a final glycerol concentration of 15% (w/v).

Production of isobutanol

Freezer vials are used to inoculate 80 ml of synthetic complete medium lacking histidine and uracil supplemented with both 0.25% glucose and 0.5% ethanol as carbon sources. A 1 liter fermenter is prepared and sterilized with 0.4L water. After cooling, filter sterilized medium is added to give the following final concentrations in 800 mL post-inoculation:

Medium (final concentration):

6.7 g/L, Yeast Nitrogen Base w/o amino acids (Difco)

2.8 g/L, Yeast Synthetic Drop-out Medium Supplement Without Histidine, Leucine, Tryptophan and Uracil (Sigma Y2001)

20 mL/L of 1% (w/v) L-Leucine

4 mL/L of 1% (w/v) L-Tryptophan

1 mL/L ergosterol/tween/ethanol solution

2017200439 20 Jan 2017

0.2 mL/L Sigma DF204

10 g/L glucose

The fermenter is set to control at pH 5.5 with KOH, 30% dO, temperature 30 °C, and airflow of 0.2 SLPM (or, 0.25 vvm). At inoculation, the airflow is set to 0.01 SLPM initially, then increased to 0.2 SLPM once growth is established. Glucose is maintained at 5-15 g/L throughout by manual addition.

To quantify the loss of isobutanol due to stripping, the off-gas from the fermentor is directly sent to a mass spectrometer (Prima dB mass spectrometer, Thermo Electron Corp., Madison, WI) to quantify the amount of isobutanol in the gas stream. The isobutanol peaks at mass to charge ratios of 74 or 42 are monitored continuously to quantify the amount of isobutanol in the gas stream. Glucose and organic acids in the aqueous phase are monitored during the fermentation using HPLC. Glucose is also monitored quickly using a glucose analyzer (YSI, Inc., Yellow Springs, OH). Isobutanol and isobutyric acid in the aqueous phase are quantified by HPLC after the aqueous phase was removed periodically from the fermentor.

Example 38

Analysis of isobutyric acid production

Strains were inoculated into synthetic complete medium, minus histidine and uracil, supplemented with 0.05% ethanol. Once cultures had reach stationary phase they were subcultured into fresh medium (starting OD = 0.2). For PNY2205 (Example 13), medium was supplemented with 0.05% ethanol to satisfy the C2 requirement observed in PDC KO yeast that do not have the phosphoketolase pathway. For PNY2209 (Example 13), cells were subcultured into media with and without ethanol. For the ald6Δ clones (PNY2216 and isogenic clones, Example 24), medium without ethanol was used. Cultures were grown in screw capped shake flasks (20 ml medium in 125 ml flasks). Culture supernatants were collected immediately after inoculation and again after 48 hours. These culture supernatants were analyzed by HPLC (described in

2017200439 20 Jan 2017

US20070092957, incorporated herein by reference) to determine glucose consumption and isobutyric acid concentration.

Table 35. Molar yield of isobutyric acid in ald6Δ strains.

Strain	Isobutyric acid molar yield (mole/mole glucose consumed)
PNY2205*	0.09
PNY2209*	0.07
PNY2209	0.09
PNY2216 and 5 other isogenic clones	0.03

*Indicates culture medium was supplemented with 0.05% ethanol

Example 39

Increased Production of Isobutanol and Decreased By-Product Production in New Strains

The purpose of this example is to describe small-scale culturing experiments with the newly constructed strains using plasmid pK9G9.OLE1p.ilvD to supply the remaining isobutanol pathway genes. New host strains PNY1528, PNY2243, PNY2237 and PNY2238 were transformed with plasmid pK9G9.OLE1p.ilvD and tested for isobutanol production. Transformants were obtained by selection on synthetic complete medium minus uracil with 1% ethanol as the carbon source. Transformants were patched on synthetic complete medium minus uracil with 2% glucose and 0.05% ethanol as carbon sources. Patches were used to inoculate liquid medium (synthetic complete minus uracil with 0.3% glucose and 0.3% ethanol as carbon sources). To test isobutanol production, liquid cultures were sub-cultured into synthetic complete medium minus uracil containing 2% glucose and 0.05% ethanol as carbon sources that also contained BME vitamin mix (Sigma Cat. No. B6891). Cultures were incubated in sealed serum vials (10 ml medium in 15 ml vials) at 30°C with shaking (250 rpm in an Infors Multitron shaker). After 48 hours, culture medium was filtered (Spin-X column) and analyzed by HPLC (as described previously in US App. Pub. No. 20070092957). Molar

2017200439 20 Jan 2017

yield of the pathway by-product isobutyrate was decreased by 50% in strains carrying the *ald6*Δ. Strains based on PNY2238 were found to have higher glucose consumption and isobutanol titer (clone K results shown in the table below).

Table 36. Molar Yield of isobutyric acid in the ald6Δ strains

Strain	Relevant phenotype	Isobutyric acid molar yield (mole/mole glucose consumed)
PNY2205*	ALD6+, C2-dependent	0.09
PNY2209*	ALD6+, C2-independent	0.07
PNY2209	ALD6+, C2-independent	0.09
PNY2216 and 5 other isogenic clones	ALD6-, C2-independent	0.03

* indicates the culture medium was supplemented with 0.05% ethanol

All strains contained plasmid pK9G9.OLE1p.ilvD. For PNY2237- and PNY2238-derived strains, the data presented are an average of two biological replicates.

Example 40

Construction of a site-saturation gene library and screening the isobutanol production of the resultant variants in PNY2259

The reverse primer mixture (called K9_309r in this example) containing primers encoding all 20 individual amino acid changes at position 309 (Table 37) and the forward primer K9_131T_080211f: GGACTTGATGTCACTATGATC (called K9_131Tf in this example) were employed to create a single-site saturation library targeting the position of 309 of K9 KARI. A plasmid containing variant K9SB2 (pHR81-ILV5p.K9SB2.TEF1(M4)p.ilvD (SEQ ID NO: 930, also called "pLH744").

In brief, a megaprimer was prepared through a regular PCR. The megaprimer PCR mixture consisted of 45 μl of SuperMix (Invitrogen, Carlsbad, CA, #10572063), 2.0 μl K9_131Tf (20 μM), 2.0 μl K9_309r (20 μM) and 1.0 μl template (50 ng/μl). The PCR program for making the megaprimer is: the starting temperature was 95°C for 1.0 min followed by 35 heating/cooling cycles. Each cycle consisted of 95°C for 20 sec, 55°C

2017200439 20 Jan 2017

for 20 sec, and 72°C for 1.0 min. The megaprimer was then used to introduce mutation into K9SB2 using the same procedure as shown in Example 5. Clones from the single-site saturation library were sequenced.

The resultant unique variants together with pLH744 were analyzed for isobutanol production and by-product formation in yeast (triple for each mutant). Yeast pathway strains were made in PNY2259 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-ADH|adh_HI-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_HI-ADH1t ymr226cΔ ald6Δ::loxP; Example 22) host by transforming the KARI vectors. The transformed cells were plated on synthetic medium without histidine or uracil (1% ethanol as carbon source). Three transformants were transferred to fresh plates of the same media. The transformants were tested for isobutanol production under anaerobic conditions in 48-well plates (Axygen, Union City, CA #391-05-061).

Yeast colonies from the transformation on SE-Ura plates appeared after 5-7 days. The three colonies from each variant were patched onto fresh SE-Ura plates, and incubated at 30◦C for 3 days.

Growth media and procedure

Two types of media were used during the growth procedure of yeast strains: an aerobic pre-culture media and an anaerobic culture media. All chemicals were obtained from Sigma unless otherwise noted (St. Louis, MO)

Aerobic pre-culture media (SE-Ura): 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine, leucine, tryptophan and uracil, 0.2% ethanol, 0.2% glucose, 0.01% w/v leucine, 0.002% w/v histidine, and 0.002% w/v tryptophan.

Anaerobic culture media (SEG-Ura-His): 50 mM MES (pH 5.5, 6.7 g/L yeast nitrogen base without amino acids (Difco, 291940, Sparks, MD), 1.4 g/L yeast synthetic drop-out medium supplement without histidine,

2017200439 20 Jan 2017

leucine, tryptophan and uracil, 0.1% ethanol, 3% glucose, 0.01% leucine, 0.002% w/v histidine, 0.002% tryptophan, 30 mg/L nicotinic acid, 30 mg/L thiamine and 10 mg/L ergosterol made up in 50/50 v/v Tween /ethanol solution.

The patched cells were inoculated into 48-well plates. Each well contains 1.5 ml aerobic media. The plates were covered with permeable foils and grown at 30◦C with shaking overnight. The cell density (OD_{600}) was then measured. The amount of cells to make 1.5 ml of cell suspension (in anaerobic media) with the final OD_{600}=0.2 for each well were calculated and 1.5 ml cell suspension was prepared with anaerobic media and added into each well. Oxygen in 48-well plates was removed using an anaerobic chamber following the manufacturer's protocol (Coy Laboratory Products Inc. Grass Lake, MI). Cells were then grown at 30◦C with shaking for two days. The cell density (OD_{600}) was then measured. The best grown mutants underwent the same two-day growth in 48-well plates for isobutanol production measurement. After two days' anaerobic growth, the cell density (OD_{600}) was then measured. Cells were centrifuged at 4,000 g for 5 min and the supernatant was collected for the isobutanol measurement using LC/MS.

<u>LC/MS analysis of yeast strains with K9 KARI mutants</u>

Samples were taken for LCMS analysis at the end of the anaerobic growth period. LCMS analysis was performed using a Waters AcQuity UPLC separations unit and AcQuity TQD triple quad mass spectrometer (Waters, Milford, MA) with a Waters AcQuity UPLC HSS T3 separations column (Waters, Milford, MA). Compounds were separated using a reverse phase gradient of water (+ 0.1% formic acid) and acetonitrile (+ 0.1% formic acid) starting with 99% aqueous and ending with 99% organic, at a flow rate of 0.5 mL/min. Chromatograms were analyzed using Waters Masslynx 4.1 software (Waters, Milford, MA). Micro molar yields for isobutanol were calculated using Waters Quanlynx software (Waters, Milford, MA) using a calibration curve of triplicate analyses of standards.

Table 37. Reverse Primers

2017200439 20 Jan 2017

Targeted position(s) of *K9-KARI*	Primers
309	K9_309I_111711r: CTTTCTCATAGCCTTAATGTGGAC(SEQ ID NO: 884) K9_309L_111711r: CTTTCTCATAGCCTTTAAGTGGAC(SEQ ID NO: 885) K9_309Y_111711r: CTTTCTCATAGCCTTATAGTGGAC(SEQ ID NO: 886) K9_309C_111711r: CTTTCTCATAGCCTTACAGTGGAC(SEQ ID NO: 887) K9_309W_111711r: CTTTCTCATAGCCTTCCAGTGGAC(SEQ ID NO: 888) K9_309P_111711r: CTTTCTCATAGCCTTTGGGTGGAC(SEQ ID NO: 889) K9_309H_111711r: CTTTCTCATAGCCTTATGGTGGAC(SEQ ID NO: 890) K9_309Q_111711r: CTTTCTCATAGCCTTTTGGTGGAC(SEQ ID NO: 891) K9_309M_111711r: CTTTCTCATAGCCTTCATGTGGAC(SEQ ID NO: 892) K9_309N_111711r: CTTTCTCATAGCCTTATTGTGGAC(SEQ ID NO: 893) K9_309V_111711r: CTTTCTCATAGCCTTAACGTGGAC(SEQ ID NO: 894) K9_309A_111711r: CTTTCTCATAGCCTTAGCGTGGAC(SEQ ID NO: 895) K9_309D_111711r: CTTTCTCATAGCCTTATCGTGGAC(SEQ ID NO: 896) K9_309E_111711r: CTTTCTCATAGCCTTTTCGTGGAC(SEQ ID NO: 897) K9_309G_111711r: CTTTCTCATAGCCTTACCGTGGAC(SEQ ID NO: 898) K9_309S_111711r: CTTTCTCATAGCCTTAGAGTGGAC(SEQ ID NO: 899) K9_309T_111711r: CTTTCTCATAGCCTTAGTGTGGAC(SEQ ID NO: 900) K9_309R_111711r: CTTTCTCATAGCCTTTCTGTGGAC(SEQ ID NO: 901) K9_309K_111711r: CTTTCTCATAGCCTTCTTGTGGAC(SEQ ID NO: 902)

Table 38. List of some variants with OD_{600} and isobutanol titer (mM) after two days anaerobic growth

Variant	Amino Acid Seq ID No:	Repeat	OD_{600}	Isobutanol titer (mM)
K9C4	927	1	0.5502	76.99
		2	0.6578	84.05
		3	0.7301	98.91
K9C8	928	1	0.6887	116.57
		2	0.5309	78.77
		3	0.6859	102.49
K9SB2	427	1	0.6314	88.39
		2	0.5977	81.18
		3	0.2325	44.60

Example 41

Construction of K9SB2_SH (K9SB2_short), a truncated version of K9SB2

A gene encoding a version of K9SB2 lacking the first five N-terminal amino acids (MEECK) was prepared by PCR with the Phusion® High-Fidelity PCR Kit (Catalog #E0553L, New England Biolabs). Primers Kshort1

20 Jan 2017 2017200439

(AAACCGGTTTAAACAGTATGGCTAAGATTTACTACCAAGAAGACTG; SEQ ID NO: 903) and YGrev (TTCTGTTTTATCAGACCGCTTC; SEQ ID NO: 904) were synthesized by Integrated DNA Technologies, Inc (Coralville IA). Other than the primers, dNTP mix (Cat#18427-013, Invitrogen, Carlsbad, CA), template, and ddH_2O, reagents used here were supplied with the kit indicated above. The PCR mixture consisted of 1 µl of K9SB2 in a pBAD.KARI plasmid (20 ng/µl; plasmid preparation described in Example 17; SEQ ID NO: 428), 2 µl of each primer (20 µM stocks), 10 µl of 5× Phusion HF Buffer, 2 µl of 5 mM dNTP mix, 0.5 µl polymerase, and 28 µl of ddH_2O. The following conditions were used for the PCR reaction: The starting temperature was 98°C for 2 min followed by 30 heating/cooling cycles. Each cycle consisted of 98°C for 10 sec, 48°C for 30 sec, and 72°C for 1.5 min. At the completion of the temperature cycling, the sample was kept at 72° C for 10 min more, and then held awaiting sample recovery at 4°C. The reaction product was separated from the template via agarose gel electrophloresis (1% agarose, 1X TBE buffer) and recovered using the Illustra GFX™ PCR DNA and Gel Band Purification Kit (Cat# 28-9034-70, GE Healthcare Sciences, Piscataway, NJ) as recommended by the manufacturer. The purified PCR product was digested with PmeI and SfiI and ligated into the corresponding sites of pBAD-ps-JEA1 (SEQ ID NO: 905) and the sequence of the resultant construct K9SB2_SH in pBAD.KARI (SEQ ID NO: 929) was confirmed.

Example 42

Generation of yeast expression plasmids for studies of K9SB2 and other KARI enzymes

Construction of pHR81-ILV5p-K9SB2-TEF(M4)-IlvD (pLH744; SEQ ID NO: 930):

The K9SB2 gene was excised from pHR81-ILV5p-K9SB2 by PmeI and SfiI digestion, and ligated with pHR81-ILV5p-K9D3-TEF(M4)-IlvD (pBP2090) at PmeI and SfiI sites. The ligated vector was named pHR81-ILV5p-K9SB2-TEF(M4)-IlvD (pLH744), and this vector was confirmed by sequencing.

Construction of K9SB2_SH_DHAD (SEQ ID NO: 931):

20 Jan 2017 2017200439

The digested K9SB2_SH PCR product (described in example 41) was ligated into PmeI and SfiI sites of pLH744 and confirmed by sequencing.

Construction of K9SB2_SH_81 (SEQ ID NO: 932)

The digested K9SB2_SH PCR product (described in example 41) was ligated into PmeI and SfiI sites of pHR81-PIlv5-KARI-K9.G9 and confirmed by sequencing.

Construction of plasmids for yeast expression of K9SB2 derivatives

The yeast expression plasmids for variants listed in Table 39 (described in examples 17 and 21) were generated by subcloning the corresponding KARI genes from pBAD.KARI plasmids into the PmeI and SfiI sites of plasmid pLH744 and plasmid pHR81-PIlv5-KARI-K9.G9. The constructs were confirmed by sequencing.

Table 39: K9SB2 derivatives subcloned into yeast expression plasmids

K9SB2 Derivative	**Alternate name**	**Amino Acid Sequence ID**	**Nucleic Acid Sequence ID**
K9SB2-K90M	K9_David	431	432
K9SB2-G55D	K9_Eliza	433	946
K9SB2-Q91L	K9_Frank	440	947
K9SB2-A303D	K9_Grace	445	948
K9SB2-M94-V307I	K9_Ingrid	455	949
K9SB2-F67I	K9_Jarvis	437	950
K9SB2-A56G-K90N	K9_Kelly	452	951
K9SB2-G55C	K9_Norman	481	952
K9SB2-P135S	K9_Ophelia	488	953
K9SB2-F53L	K9_Pat	441	954
K9SB2-Q94I	K9_Quentin	495	955
K9SB2-F67L	K9_Ralph	496	956
K9SB2-K8N-K90M-T141I	K9_Sophia	502	957
K9SB2-E13V-M94I-T141I	K9_Tiberius	509	958

2017200439 20 Jan 2017

K9SB2-A56V	K9_Ursala	511	959
K9SB2-I84L	K9_Victor	514	960
K9SB2-W59C	K9_Watson	520	961
K9SB2-T93A	K9_Xavier	641	642

Construction of plasmids for yeast expression of truncated versions of K9SB2 derivatives

N-terminally truncated versions of K9SB2 derivatives were prepared as described in example 41, with modifications. The primer Kshort1 was replaced with the 5'-phophorylated primer KPSH1 (AAACAGTATG GCT AAG ATT TAC TAC CAA GAA GAC TG; SEQ ID NO: 906), which was synthesized by Integrated DNA Technologies, Inc (Coralville IA). K9SB2 in the PCR was replaced by pooled mixtures of the variants (pBAD.KARI plasmids) listed in Table 39. The purified PCR products were digested with SfiI and ligated into the PmeI and SfiI sites of pHR81-PIIv5-KARI-K9.G9. DNA sequencing with TempliPhi™ (GE Healthcare) and primers pHR81-F (ACACCCAGTATTTTCCCTTTCC) and pHR81- Rev (CTA GTG TAC AGA TGT ATG TCG G) (SEQ ID NOs: 924 and 925) was performed to identify each truncated derivative. The identified variants are indicated in Table 40. The coding sequences for the truncated versions were subsequently subcloned into the PmeI and SfiI sites of plasmid pLH744 and confirmed by sequencing.

Table 40. Truncated versions of K9SB2 derivatives

Variant	**Amino Acid Sequence ID**	**Nucleic Acid Sequence ID**
K9_David_SH	196	263
K9_Eliza_SH	266	907
K9_Frank_SH	267	908
K9_Grace_SH	389	909
K9_Ingrid_SH	405	910
K9_Jarvis_SH	781	911
K9_Kelly_SH	782	912
K9_Norman_SH	783	913
K9_Ophelia_SH	835	914
K9_Pat_SH	853	915
K9_Quentin_SH	854	916
K9_Ralph_SH	855	917
K9_Ursala_SH	856	918

2017200439 20 Jan 2017

K9_Watson_SH	857	858
K9_Xavier_SH	859	919

Construction of K9_Zeke_SH (K9SB2-K90M-T93A) and K9_Annabel_SH (K9SB2-K90M-T93I)

K9_Zeke_SH (SEQ ID NO: 860, protein SEQ ID NO: 861) and K9_Annabel_SH (SEQ ID NO: 862, protein SEQ ID NO: 863) were derived from K9_David_SH via site directed mutagenesis employing the employing the QuikChange® Lightning Site-Directed Mutagenesis Kit (Catalog #210518; Agilent Technologies, Stratagene Products Division,La Jolla, CA). Except for the primers, templates, and ddH_2O, all reagents used here were supplied with the kit indicated above. Primers were synthesized by Integrated DNA Technologies, Inc (Coralville IA).

For K9_Zeke_SH, primers employed were oMT93A (GATCCCAGATGAAATGCAGGCTGCCATGTACAAAAACGACATCG; SEQ ID NO: 920)) and oMT93Arev (CGATGTCGTTTTTGTACATGGCAGCCTGCATTTCATCTGGGATC; SEQ ID NO: 921). The reaction mixture contained 1 µl K9_David_SH in the pHR81-PIlv5-KARI-K9.G9-derived vector (20 ng/µl), 1 µl of each primer (150 ng/ul), 5 µl of 10× reaction buffer, 1 µl of dNTP mix, 1.5 µl QuikSolution, 1 ul QuikChange Lightning enzyme, and 38.5 µl ddH_2O. For K9_Annabel_SH, the primers were replaced with oMT93I (GATCCCAGATGAAATGCAGGCTATCATGTACAAAAACGACATCG; SEQ ID NO: 922) and oMT93Irev (CGATGTCGTTTTTGTACATGATAGCCTGCATTTCATCTGGGATC; SEQ ID NO: 923).

The following conditions were used for both reactions: The starting temperature was 95°C for 2 min followed by 18 heating/cooling cycles. Each cycle consisted of 95°C for 20 sec, 60°C for 10 sec, and 68°C for 6 min. At the completion of the temperature cycling, the samples held awaiting sample recovery at 4°C. 2 µl of the Dpn I was added to each reaction and the mixtures were incubated for 1 hour at 37°C. 2 µl of each mutagenic reaction was transformed into One Shot® Stbl3™ Chemically Competent *E. coli* (Invitrogen, Catalog #C7373-03) according to the

2017200439 20 Jan 2017

manufacturer's instructions. The transformants were spread on agar plates containing the LB medium and 100 µg/ml ampicillin (Cat#L1004, Teknova Inc. Hollister, CA) and incubated at 37°C overnight. Multiple transformants were then selected for TempliPhi™ DNA sequencing ™ (GE Healthcare), which employed primers pHR81-F (ACACCCAGTATTTTCCCTTTCC; SEQ ID NO: 924) and pHR81- Rev (CTA GTG TAC AGA TGT ATG TCG G; SEQ ID NO: 925). Variants with the confirmed sequences (K9_Zeke_SH and K9_Annabel_SH in pHR81-PIlv5-KARI-K9.G9 derived vectors) were subcloned into the PmeI and SfiI sites of pLH744 (SEQ ID NO: 930).

Example 43

Isobutanol production of K9SB2 and derivatives grown under anaerobic conditions in 48-well plates

The yeast expression plasmids for K9SB2, K9SB2_SH, and K9SB2-T93A, prepared with vector derived from pLH744 as described in Example 42, were employed to evaluate isobutanol production in yeast grown under anaerobic conditions in a 48-well plate. Isobutanol production strains were made in host PNY2259 (MATa ura3Δ::loxP his3Δ pdc6Δ pdc1Δ::P[PDC1]-DHAD|ilvD_Sm-PDC1t-P[FBA1]-ALS|alsS_Bs-CYC1t pdc5Δ::P[PDC5]-ADH|sadB_Ax-PDC5t gpd2Δ::loxP fra2Δ::P[PDC1]-ADH|adh_Hl-ADH1t adh1Δ::UAS(PGK1)P[FBA1]-kivD_Lg(y)-ADH1t yprcΔ15Δ::P[PDC5]-ADH|adh_Hl-ADH1t ymr226cΔ ald6Δ::loxP) by transforming the plasmids containing the coding sequences for the KARI variants and plating on synthetic medium without uracil (1% ethanol as carbon source). Yeast colonies from the transformation on SE-Ura plates appeared after 3-5 days of incubation at 30°C. At least three colonies from each variant were patched onto fresh SE-Ura plates and incubated at 30°C.

Yeast Cultivation Conditions:

Aerobic cultivation medium: SE -Ura medium with 2 g/l ethanol.

2017200439 20 Jan 2017

Anaerobic cultivation medium: SEG –Ura with 30 g/l glucose and 1 g/l ethanol, supplemented with 10 mg/l ergosterol, 50 mM MES buffer (pH 5.5), 30 mg/l thiamine, and 30 mg/l nicotinic acid.

48-well plates: Axygen catalog # P-5ML-48-C-S, 5 ml/well total volume, culture volume of 1.5 ml/well.

Plates were covered with a permeable adhesive film (VWR ; catalog number 60941-086) for aerobic cultivation. Plates were shaken at 225 rpm at 30 °C. For anaerobic cultivation, freshly inoculated plates covered with permeable film were purged of oxygen by equilibration in an anaerobic chamber for 2 hours. The plate covers were then exchanged for adhesive aluminum covers (VWR ; catalog number 89049-034) and each plate was placed into an airtight plastic box (Mitsubishi Gas Chemical America, Inc; New York, NY; Catalog 50-25) along with a fresh oxygen scavenger pack (Mitsubishi Gas Chemical America, Inc; New York, NY; Catalog 10-01). The entire assembly (plate(s) and oxygen scavenger pack inside a sealed airtight plastic box) was removed from the anaerobic chamber and shaken at 225 rpm at 30 °C.

Experimental Protocol

Single yeast colonies on SE –Ura agar plates were streaked onto fresh SE –Ura agar plates and incubated at 30 °C until dense patches of cells had grown. Liquid precultures in 48-well plates were inoculated with loops of these cells for initial aerobic cultivation. After shaking overnight, the OD_{600} of each culture well was measured by transferring 0.15 ml of each well into a flat-bottom 96-well plate and measuring the absorbance of each well at 600 nm with a Molecular Devices (Sunnyvale, CA) plate reader. A linear transformation based on an experimentally-determined calibration line was applied to these plate reader-measured optical densities to convert them into comparable absorbance values for a cuvette-based spectrophotometer.

A calculated portion of each aerobic preculture well was inoculated into the corresponding well of a fresh 48-well plate with 1.5 ml of the SEG –Ura medium, to achieve an initial OD_{600} (in cuvette spectrophotometer absorbance units) of 0.2. In the process of inoculating the fresh plate, the

20 Jan 2017 2017200439

aerobic preculture plate was centrifuged, the supernatant was removed from each well, and the cells in each well were resuspended in fresh SEG –Ura medium. This anaerobic cultivation plate was shaken for 3 days. The isobutanol concentration in the culture supernatants was measured by HPLC (Table 41).

Table 41. Isobutanol titers reached in the first anaerobic passaging cycle

Variant	# wells	Mean Isobutanol Titer (mM)	Standard Deviation of Isobutanol Titer (mM)
K9SB2	16	35.2	11.9
K9SB2_SH	8	67.4	12.2
K9SB2-T93A	8	40.8	9.1

A follow-up anaerobic cultivation was initiated from the first anaerobic cultivation as follows: A calculated portion of each anaerobic culture well was inoculated into the corresponding well of a fresh 48-well plate with 1.5 ml of the SEG –Ura medium, to achieve an initial OD_{600} (in cuvette spectrophotometer units) of 0.2. In the process of inoculating the fresh plate, the growth plate was centrifuged, the supernatant was removed from each well, and the cells in each well were resuspended in fresh SEG –Ura medium, in order to minimize carryover of metabolites from one cultivation to the next. The follow-up (second-cycle) anaerobic cultivation plate was shaken for 2 days. The isobutanol concentration in the culture supernatants was measured by HPLC (Table 42).

Table 42. Isobutanol titers reached in the second anaerobic passaging cycle

Variant	# wells	Mean Isobutanol Titer (mM)	Standard Deviation of Isobutanol Titer (mM)
K9SB2	16	67.6	10.8
K9SB2_SH	8	85.7	9.2
K9SB2-T93A	8	76.3	16.8

Example 44

Kinetic characterization of K9G9, K9SB2, and K9SB2_SH

2017200439 20 Jan 2017

K9G9, K9SB2, K9SB2_SH were overexpressed via pBAD.KARI plasmids in *E. coli* strain Bw25113 (Δ*ilvC*) and purified for detailed measurement of the kinetic parameters.

Expression, purification and cofactor kinetic analyses were performed as described in example 18, with the following modifications. Expression cultures were grown in 20 mL of LB with 100 µg/mL ampicillin and 0.2% (w/v) arabinose in a 125 mL vented cap flask. The expression media was inoculated with 1/10 volume of 8 hour culture. Expression cultures were grown for 18 hours for K9 SB2 and 24 hours for K9SB2_SH.

Table 43. Kinetic Parameters Comparing Reactions with NADH and NADPH

Variant	SEQ ID NO:	V_{max} NADPH, U/mg	K_m NADPH, µM	V_{max}/K_m NADPH, L/min*mg	V_{max} NADH, U/mg	K_m NADH, µM	V_{max}/K_m NADH, L/min*mg
K9G9	644	2.2	24.1	0.091	1.9	78.2	0.024
K9SB2	427	1.7	44.8	0.038	1.8	11.6	0.155
K9SB2_SH	637	1.7	109.9	0.015	1.7	13.3	0.128

Example 45
Isobutanol production of K9SB2 and derivatives

Isobutanol production analysis in 48-well plates was performed, as described, in example 42, with the following modifications. Aerobic and anaerobic cultivation media are the same as those described in example 19, but with 0.01% w/v histidine added. OD_{600} values of aerobic pre-cultures were measured using a Cary 300 spectrophotometer (Agilent Technology, Wilmington, DE). A Heraeus Multifug X1R with a M-20 rotor (Thermo Scientific, Waltham, MA) was used to pellet the aerobic pre-culture cells and the spent cultivation media was discarded. The plates with cell pellets were transferred into the Coy Anaerobic Bag (Grass Lake, MI) and 100 µL of anaerobic cultivation media was added to each pellet. The anerobic cultivation media was allowed to degas for at least 24 hours prior to the 48-well plate receiving 1.5 mL aliquots; this process was performed inside the anaerobic bag. The anaerobic culture plate was inoculated with the appropriate volume of cell resuspension and covered

20 Jan 2017 2017200439

with an adhesive aluminum foil. Plates were placed into a MCG 2.5L AnaeroPack system (MCG, Japan). The box was sealed and removed from the anaerobic bag and placed in a 30°C incubator for 80 hours with shaking at 220 rpm. Three transformants were analyzed per variant, six transformations were analyzed for K9D3 and K92B2 (Table 44).

Table 44. Isobutanol Titers

Variant	SEQ ID NO:	Isobutanol Titer, mM
K9D3	645	76.4 ± 9.7
K9SB2	427	93.8 ± 3.3
K9_Frank	440	27.8 ± 0.5
K9_Grace	445	89.4 ± 15.5
K9_Ingrid	455	88.3 ± 5.4
K9_Jarvis	437	91.8 ± 5.8
K9_Kelly	452	40.9 ± 4.5
K9_Norman	481	66.2 ± 10.6
K9_Ophelia	488	28.3 ± 5.3
K9_Pat	441	77.1 ± 25.6
K9_Quentin	495	80.1 ± 10.8
K9_Ralph	496	82.1 ± 21.3
K9_Sophia	502	25.3 ± 13.4
K9_Tiberius	509	11.4 ± 10.4
K9_Ursala	511	57.3 ± 26.2
K9_Victor	514	93.0 ± 11.4

Isobutanol production analysis in serum vials was performed for select variants as described in Example 19, with the following modifications. KARI variants were expressed in yeast from plasmids derived from pLH744, prepared as described in Example 42. Histidine was added to both the aerobic pre-culture and anaerobic growth media to a final concentration of 0.01% w/v. Three or four transformants were analyzed per variant (Table 45, 46, and 47).

Table 45. Isobutanol Titers Experiment 1

Variant	SEQ ID NO:	Passage	Hours	Isobutanol Titer, mM	Isobutanol/ Glycerol	
K9SB2	427	1	75	28.2 ± 6.85	2.11 ± 0.16	
K9SB2_SH	637	1	75	87.9 ± 1.04	2.49 ± 0.04	
K9SB2	427	2	48	85.0 ± 4.35	2.67 ± 0.06	
K9SB2_SH	637	2	48	97.4 ± 4.79	2.49 ± 0.07	

Table 46. Isobutanol Titers Experiment 2

2017200439 20 Jan 2017

Variant	SEQ ID NO:	Passage	Hours	Isobutanol Titer, mM	Isobutanol/ Glycerol
K9SB2	427	1	44	30.2 ± 4.1	2.30 ± 0.15
K9SB2_SH	637	1	44	40.3 ± 5.6	2.41 ± 0.07
K9_David	431	1	44	36.2 ± 3.5	2.40 ± 0.10
K9_David_SH	196	1	44	41.9 ± 6.4	2.37 ± 0.04
K9_Grace	445	1	44	40.2 ± 5.5	2.30 ± 0.05
K9_Pat	441	1	44	29.0 ± 3.1	2.35 ± 0.10
K9SB2	427	2	46	59.2 ± 0.6	2.61 ± 0.03
K9SB2_SH	637	2	46	74.8 ± 3.0	2.58 ± 0.03
K9_David	431	2	46	63.8 ± 0.6	2.65 ± 0.06
K9_David_SH	196	2	46	76.7 ± 1.5	2.53 ± 0.04
K9_Grace	445	2	46	62.3 ± 3.1	2.52 ± 0.07
K9_Pat	441	2	46	50.1 ± 2.1	2.48 ± 0.07

Table 47 Isobutanol Titers

Variant	SEQ ID NO:	Passage	Hours	Isobutanol Titer, mM
K9SB2	427	2	50	66.2 ± 1.8
Annabel_SH	862	2	50	74.4 ± 2.4
Zeke_SH	860	2	50	76.8 ± 8.3

Where the terms “comprise”, “comprises”, “comprised” or “comprising” are used in this specification, they are to be interpreted as specifying the presence of the stated features, integers, steps or components referred to, but not to preclude the presence or addition of one or more other feature, integer, step, component or group thereof.

HMMER2.0 [2.2g] — File format version: a unique identifier for this save file format.
NAME Functionally Verified KARIs — Name of the profile HMM
LENG 354 — Model length: the number of match states in the model.

ALPH Amino — Symbol alphbet: This determines the symbol alphabet and the size of the symbol emission probability distributions. If Amino, the alphabet size is set to 20 and the symbol alphabet to ``ACDEFGHIKLMNPQRSTVWY" (alphabetic order).

MAP yes — Map annotation flag: If set to yes, each line of data for the match state/consensus column in the main section of the file is followed by an extra number. This number gives the index of the alignment column that the match state was made from. This information provides a ``map" of the match states (1..*M*) onto the columns of the alignment (1..*alen*). It is used for quickly aligning the model back to the original alignment, e.g. when using hmmalign -mapali.

COM hmmbuild -n Functionally Verified KARIs exp-KARI.hmm exp-KARI_mod.aln — Command line for every HMMER command that modifies the save file: This one means that hmmbuild (default patrameters) was applied to generate the save file.

COM hmmcalibrate exp-KARI.hmm — Command line for every HMMER command that modifies the save file: This one means that hmmcalibrate (default parametrs) was applied to the save profile.
NSEQ 25 — Sequence number: the number of sequences the HMM was trained on
DATE Mon Dec 8 17:34:51 2008 — Creation date: When was the save file was generated.

XT -8455 -4 -1000 -1000 -8455 -4 -8455 -4 — Eight ``special" transitions for controlling parts of the algorithm-specific parts of the Plan7 model. The null probability used to convert these back to model probabilities is 1.0. The order of the eight fields is N->B, N->N, E->C, E->J, C->T, C->C, J->B, J->J.
NULT -4 -8455 — The transition probability distribution for the null model (single G state).

NULE 595 -1558 85 338 -294 453 -1158 197 249 902 -1085 -142 -21 -313 45 531 201 384 -1998 -644 — The extreme value distribution parameters μ and lambda respectively; both floating point values. These values are set when the model is calibrated with hmmcalibrate. They are used to determine E-values of bit scores.
EVD -333.712708 0.110102

HMM	A	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	Y	Position in alignment
	m->m	m->i	m->d	i->m	i->i	d->m	d->d	b->m	m->e												
	-650	*	-1463																		
1(Q)	-648	-1356	-136	-44	-1453	-1166	-219	-1455	321	-1417	-911	-227	-1496	3263	122	-643	-684	-1239	-1542	-1030	**7100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-701	-1378	-650	*												
2(M)	-4231	-3929	-5216	-5402	-3438	-4370	-4528	-3232	-5113	-2613	5320	-5052	-4790	-4977	-4823	-4692	-4459	-3629	-4103	-4017	**7200%**
-	-147	-501	232	42	-382	397	104	-625	209	-467	-722	276	396	44	95	361	121	-368	-296	-251	
-	-3303	-3318	-325	-3473	-136	-701	-1378	*	*												
3(F)	-1308	-1104	-2227	-2120	3516	-2093	-244	-196	-1891	64	66	-1626	-2278	-1503	-1798	-1617	-1350	-389	305	1335	**8600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-943	-1060	*	*												
4(A)	1616	-1744	1125	33	-2015	-1540	-262	-1686	937	-1765	-911	-252	-1658	154	-383	-488	640	-3	-2038	-1421	**8700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-901	-7402	-1125	-894	-1115	-2352	-314	*	*												
5(C)	-346	2578	1084	-712	2092	-1540	-384	-167	-624	-482	125	-731	-1705	-451	-883	-631	-338	-50	-774	-133	**8800%**
-	-149	-500	235	43	-381	398	106	-626	210	-466	-721	275	394	45	96	359	118	-369	-295	-249	
-	-1009	-1006	-7567	-131	-3527	-1916	-444	*	*												
6(S)	800	-586	-1937	-1415	-821	-1740	-954	1279	-1204	-584	19	-1258	-1964	-1013	-1358	1715	-476	1117	-1320	-938	**9000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-17	-6953	-7995	-894	-1115	-146	-3378	*	*												
7(K)	-956	-2411	-803	501	-2743	-1919	-558	-2483	2435	-2420	-1502	57	-2010	1146	458	829	224	-2040	-2577	-1913	**9100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
8(V)	-2472	-2010	-5089	-4702	-2534	-4789	-4391	2241	-4574	-151	-1318	-4442	-4600	-4417	-4628	-4080	-82	3023	-3952	-3510	**9200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

9(Y)	-4673	-3685	-5210	-5505	2423	-5069	-1332	-3424	-5065	-392	-2838	-3726	-4920	-3835	-4458	-4313	-4533	-3643	-581	4349	**9300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
10(Y)	-2170	-2625	-2489	-2097	-1555	-2986	-1481	-2628	906	-2674	-2098	-2051	-3206	-1513	-1078	-2258	1039	-2435	-2009	4185	**9400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
11(D)	-2498	-4412	3500	1042	-4581	-2437	-1765	-4500	733	-4361	-3682	515	-2961	-1429	-2799	-2158	-2558	-3974	-4550	-3541	**9500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
12(K)	11	-2371	348	819	-2692	-535	-527	-2443	2294	-2387	-1461	590	-1960	-68	904	-67	-837	-1993	-2554	-1871	**9600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
13(D)	-2663	-4633	3700	580	-4789	-2487	-1872	-4738	731	-4578	-3963	-1073	-3046	-1551	-2987	-2292	-2742	-4201	-4759	-3709	**9700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
14(C)	2503	3193	-4266	-3818	-2010	-3276	-2896	762	-3517	-1437	-1051	-3233	-3509	-3212	-3411	-2499	-1792	1507	-2796	-2431	**9800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
15(D)	-1363	-2905	2748	542	-3202	-2072	-920	-2977	290	-2912	-2023	1270	-2294	-489	-1186	53	1116	-2518	-3086	-2349	**9900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
16(L)	-1268	-1113	-3338	-540	-1057	-2827	-1716	569	-2409	2299	-236	-2381	-2862	-2089	-2316	-232	-1213	1306	-1645	-1304	**10000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
17(S)	-1350	-2877	588	1045	-3189	-496	-920	-2963	-628	-2901	-2011	1860	-2289	-489	-1184	2139	190	-2503	-3077	-2343	**10100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	*	*												
18(G)	-454	-832	-968	-1110	-2112	3143	-1211	-2091	-1317	-2264	-1691	-978	-1499	-1202	-1421	-646	-774	-1550	-1916	-1919	**10200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
19(H)	-898	-1313	-545	-482	-320	-1336	4297	-1552	-160	-1493	-1035	-579	-1675	-363	-322	-934	-951	-1354	-725	107	**10300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
20(D)	-872	-1812	3234	432	-2215	-967	-433	-2172	-569	-2269	-1704	99	-1453	-184	-1141	-728	-973	-1814	-2146	-1646	**10400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
21(E)	-766	-1695	521	2831	-2050	-1029	-293	-1804	-118	-1919	-1331	69	-1441	-4	-527	-653	-814	-1512	-1988	-1505	**10500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
22(Y)	-1337	-1229	-1681	-1596	1268	-1957	121	-918	-1294	-769	-585	-1229	-2163	-1111	-1301	-1443	-1359	-932	592	3932	**10600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-109	-3775	*	*												
23(I)	-2294	-1931	-4749	-4227	-1724	-4227	-3320	2306	-3952	1990	-634	-3878	94	-3538	-3812	-3411	-2247	1576	-2891	-2629	**10700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
24(I)	-2801	-2299	-5406	-5003	-2108	-5164	-4649	3051	-4886	1593	-869	-4829	-4788	-4454	-4829	-4493	-2764	1435	-3781	-3585	10800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
25(K)	-234	-2632	306	-500	-3007	-2141	-719	-2712	2540	-2619	-1730	-778	-2231	2257	968	-1109	-1152	-2288	-2738	-2136	10900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
26(G)	-2184	-3900	796	392	-4174	2903	-1580	-4030	-1636	-3937	-3173	-967	-2810	-1	-2362	1069	-2220	-3530	-4130	-3229	11000%
-	-149	-501	233	42	-375	399	104	-625	210	-463	-722	276	396	44	96	358	116	-371	-296	-251	
-	-155	-3318	-9181	-3674	-118	-701	-1378	*	*												
27(K)	-3243	-3775	-4129	-2558	-4750	-3647	-1490	-4021	3681	-3617	-2982	-2368	-3580	-1076	1318	-3119	-2876	-3817	-3395	-3374	12600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
28(K)	-1684	-2925	-1665	-979	-3407	-2535	-923	-3021	2737	-2865	-2032	202	-2582	1301	804	-1564	1681	-2645	-2905	-2448	12700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
29(V)	-2623	-2122	-5300	-4990	-2769	-5101	-5131	2388	-4945	-1532	-1474	-4790	-4868	-4890	-5101	-4482	-2619	3219	-4505	-3990	12800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
30(A)	3309	-1828	-4057	-4294	-4382	656	-3657	-4147	-4169	-4428	-3497	-2821	-2904	-3694	-3937	-1470	59	-2957	-4610	-4522	12900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
31(V)	-2625	-2122	-5304	-4993	-2772	-5111	-5142	2881	-4950	-1532	-1474	-4796	-4873	-4896	-5108	-4492	-2621	2896	-4512	-3997	13000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
32(I)	-2790	-2287	-5403	-5009	-2155	-5170	-4698	3324	-4899	1175	-912	-4835	-4802	-4495	-4860	-4506	-2757	1192	-3838	-3622	13100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
33(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	13200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
34(Y)	-4838	-3766	-5229	-5579	1502	-5108	-1300	-3726	-5134	-3040	-3131	-3723	-4963	-3861	-4500	-4356	-4689	-3881	2986	4507	13300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
35(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	13400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
36(S)	-1473	-2007	-3647	-3780	-3430	-2363	-3314	228	-3616	-3373	-2876	-2840	-3093	-3395	-3541	3475	-1885	-2307	-3927	-3474	13500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
37(Q)	-4589	-4392	-3927	-4146	-5099	-4221	-4099	-5973	-3840	-5564	-5304	-4230	-4693	4575	-3826	-4704	-4772	-5612	-4577	-4751	13600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

38(G)	677	-2128	-3838	-4171	-4647	3536	-3816	-4506	-4340	-4749	-3857	-3009	-3149	-3871	-4137	-1784	-2005	-3297	-4725	-4735	13700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
39(H)	-2667	-3375	-2682	-2114	-3744	-3201	4738	-3782	-445	-3553	-2886	-2112	866	-1265	1506	-2614	-2557	-3469	-3282	-2908	13800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
40(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	13900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
41(H)	-3103	-3404	-2950	-2573	-783	-3679	4549	-3407	-1372	-3071	-2715	-2454	-3764	2546	-1428	-2990	-2976	-3308	2269	-295	14000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
42(A)	3357	-1795	-4134	-4277	-4057	-2118	-3548	-3549	-4035	-4024	-3192	-2817	-2900	-3608	-3823	217	-1660	-276	-4363	-4211	14100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
43(Q)	-1061	1950	-2044	-1475	-1236	-2372	-1154	-789	-1218	1062	1123	743	-2446	2895	-1441	-1392	-1005	-693	-1678	-1278	14200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
44(N)	-4000	-4117	-3389	-3749	-5073	-3911	-4123	-6022	-4503	-5797	-5419	4397	-4479	-4255	-4592	-4115	-4312	-5371	-4650	-4731	14300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
45(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	14400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
46(R)	-1731	-3015	275	-931	-3487	-2518	-973	-3116	2321	-2955	-2123	224	-2603	256	2808	-1596	-1613	-2730	-2995	-2515	14500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
47(D)	-2896	-4843	3855	944	-5037	-2600	-2082	-5082	-2528	-4903	-4373	-1209	-3196	-1786	-3536	-2501	-3007	-4517	-5004	-3956	14600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
48(S)	-1536	-2212	-2363	-2679	-4293	-2279	-3082	-4365	-3331	-4524	-3676	288	-3026	-2967	-3497	3508	-1962	-3259	-4477	-4066	14700%
-	-148	-500	232	44	-381	398	105	-627	211	-465	-721	275	393	45	95	360	118	-370	-295	-250	
-	-155	-3318	-9181	-2405	-302	-701	-1378	*	*												
49(G)	-2521	-3968	1232	-911	-4849	3373	-2126	-4854	-2535	-4752	-4136	-53	-3115	-1836	-3440	-2284	-2716	-4157	-4880	-3914	15400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
50(V)	-2767	-2324	-5232	-4770	396	-4827	-3784	-36	-4546	848	-611	-4472	-4518	-3980	-4367	-4081	-2716	3323	-3037	-2660	15500%
-	-148	-500	233	43	-381	399	106	-626	211	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-148	-3381	-9181	-203	-2928	-701	-1378	*	*												
51(D)	-1684	-3285	2735	2014	-3554	-2196	-1177	-3350	92	-3279	-2427	692	-2505	-770	-1595	-1483	-1666	332	-3460	-2676	15700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
52(V)	-3122	-2888	-5092	-5160	-3522	-4180	-4687	-905	-5060	-2626	-2570	-4662	-4579	-4940	-4923	-4013	-3297	3796	-4414	-4190	15800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
53(V)	369	366	-3075	-2462	-883	-2557	-1420	1415	378	-757	-117	-2098	-2610	-1809	-2037	-1630	1166	2145	-1385	-343	**15900%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
54(V)	-2624	-2122	-5302	-4991	-2772	-5108	-5139	2623	-4948	-1533	-1475	-4794	-4871	-4894	-5106	-4488	-2620	3088	-4511	-3996	**16000%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
55(G)	929	-2107	-3852	-4182	-4633	3492	-3809	-4486	-4335	-4732	-3835	-2997	-3132	-3863	-4127	-1761	-1982	-3275	-4720	-4725	**16100%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
56(L)	-3427	-2938	-5791	-5325	-1449	-5374	-4410	-543	-5063	3041	-255	-5207	-4820	-4126	-4691	-4757	-3351	883	-3184	-3234	**16200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
57(R)	-3040	-3724	-3266	82	-4620	-3470	-1396	-3905	804	-3529	-2874	-2133	-3439	-978	3800	-2894	-2709	-3682	-3353	-3267	**16300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
58(K)	31	-2412	-803	1532	-2743	-1920	-559	-2483	1772	-2421	-1503	-556	1229	727	1079	-566	-893	-2041	-2579	-1915	**16400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
59(G)	-2671	-4661	1614	587	-4832	3103	-1901	-4803	-2269	-4648	-4047	421	-3049	-1587	-3230	-2297	-2766	-4245	-4850	-3752	**16500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
60(S)	-1499	-2308	-1932	-1859	-4006	1604	-2121	-3754	1362	-3793	-2945	-1833	-2827	-1794	-1902	2738	-1771	-2970	-3910	-3479	**16600%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-247	-8139	-2699	-894	-1115	-701	-1378	*	*													
61(K)	1362	-2232	-619	-98	-2567	-427	-435	-2309	1599	-2265	-1349	1101	-1861	886	-512	833	-740	-1868	-2441	-1767	**16700%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-9	-7900	-8943	-894	-1115	-344	-2238	*	*													
62(S)	-1288	-1904	-3742	-4011	-4384	-2155	-3593	-4209	-3996	-4479	-3573	-2789	-2948	-3606	-3832	3517	228	-3028	-4600	-4451	**16800%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
63(W)	726	-873	-3261	-2634	1926	-2567	-1425	660	-2252	-701	-68	-2174	-2617	-1898	18	-1648	-972	983	4091	-958	**16900%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
64(E)	1527	-2404	212	1636	-2722	-1878	-556	-2474	1241	-2419	-1497	350	-1985	-100	-659	96	70	-2025	-2589	-1903	**17000%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
65(K)	-8	-2242	-895	770	-2502	-1963	-609	-2192	2589	22	-1353	-631	-2052	692	-617	-889	-906	-361	-2455	-1836	**17100%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
66(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	**17200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													

67(Q)	-1006	-2441	-869	1767	-2780	-1965	-586	-2510	1702	-2445	-1534	-603	-2052	1923	873	-888	236	-630	-2596	-1949	**17300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
68(A)	1489	-2393	167	1234	-2711	-1873	-547	-2462	895	-2408	-1485	1161	-1977	-90	-648	666	141	-2014	-2577	-1892	**17400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
69(D)	2104	-2898	2124	985	-3163	-2096	1397	-2935	-693	-2897	-2025	-723	-2329	-543	-1250	-1245	-1368	-2501	-3087	-530	**17500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
70(G)	-2294	-2898	-2521	-2885	-4852	3641	-3456	-5042	-3796	-5094	-4356	365	-3545	-3376	-4005	-2451	-2700	-3996	-4706	-4575	**17600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
71(F)	-2596	-2266	-4685	-4188	3199	-4136	1018	505	-3812	1986	-405	-3595	-3961	-3157	-3524	-3277	-2509	-1337	-1621	-840	**17700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
72(K)	47	-2348	338	950	-2668	-1854	-512	-507	1721	-2364	-1438	-490	-1947	672	436	806	687	-1970	-2533	-1851	**17800%**
-	-149	-500	232	46	-381	399	105	-627	210	-466	-721	277	393	45	95	359	119	-370	-295	-250	
-	-155	-3318	-9181	-2159	-366	-701	-1378	*	*												
73(V)	-1810	-1639	-4149	-3689	-1869	-3417	-2822	29	-3369	320	-897	-3230	112	-3099	-3291	-2619	-767	3269	-2708	-2354	**18400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
74(K)	847	-1093	-2131	-1554	304	127	-1127	-637	1445	645	1186	-1534	-2401	-1174	-1547	-764	-172	-528	-1519	1413	**18500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
75(T)	-1284	-2794	1526	1290	-3096	-2041	1289	-2863	-548	-2808	-1914	-668	-2242	-427	-1095	1451	1827	-2411	-2986	-2264	**18600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
76(V)	-1089	-957	-3143	-2535	-943	-2618	-1496	1052	-2198	-792	1859	-146	686	-1884	-2111	-1695	945	2346	-1458	-1106	**18700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
77(W)	1606	-2321	-752	612	-2628	-323	-527	-2366	1480	-2331	-1416	-510	-789	-73	421	23	-829	-1936	2212	-1843	**18800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
78(E)	-1509	-3540	1372	3127	-3861	-120	-1391	-3685	-1319	-3605	-2787	-900	-2659	-1005	-1976	-400	-655	-3194	-3790	-2957	**18900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
79(A)	3390	-1868	-4092	-4341	-4332	-2153	-3680	-3942	-4157	-4333	-3471	-2869	-2948	-3730	-3919	-1525	931	-2894	-4580	-4483	**19000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
80(V)	2003	-1721	-4449	-3995	-2160	-3763	-3240	1342	-3745	-1435	-1124	-3561	-3855	-3494	-3700	-2979	-58	2574	-3091	-2698	**19100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
81(K)	1714	-2501	-959	446	-2858	-2043	-654	-2574	1964	-2506	-1609	-689	-2135	-203	1088	428	-1032	-2148	-2652	-2027	**19200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
82(W)	265	-2347	815	432	-2663	634	-519	-2410	619	-2361	-1438	-495	-1952	1955	-609	-382	147	-1966	2858	-1853	19300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
83(A)	3391	-1860	-3998	-4279	-4411	-2128	-3684	-4207	-4197	-4490	-3565	-2837	-2929	-3729	-3959	706	-1718	-3001	-4636	-4534	19400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
84(D)	-2747	-4795	3813	396	-4912	-2496	-1935	-4905	-2324	-4735	-4166	-1079	-3082	603	-3296	-2353	-2844	-4347	-4929	-3809	19500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
85(V)	-2717	-2220	-5338	-4951	-2254	-5099	-4670	1963	-4844	1553	-1011	-4759	-4771	-4509	-4836	-4427	-2688	2741	-3899	-3628	19600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
86(V)	-2635	-2129	-5306	-4970	-2652	-5125	-5011	2554	-4915	-354	-1368	-4781	-4852	-4798	-5038	-4487	-2622	3019	-4355	-3902	19700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
87(M)	-1340	-1208	-3317	-2708	-968	-2860	-1708	577	-2346	932	4131	-2382	-2878	250	-2265	-228	-1278	-506	-1629	-1313	19800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
88(I)	-2566	-2177	-5017	-4470	669	-4496	-3487	2791	-4191	1116	1394	-4156	-4228	-3615	-3972	-3687	-2499	1692	-2860	-2711	19900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
89(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	20000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
90(I)	1212	-1286	-3846	-3262	-1360	-3195	-2166	1616	-2918	1031	-493	-2824	-3211	-2583	-2782	-2308	1598	1299	-2020	-1668	20100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
91(P)	-1614	-2214	-3396	-3710	-4516	-2407	-3618	-4516	-3976	-4705	-3849	-2890	3993	-3625	-3900	666	-2068	-3354	-4610	-4474	20200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
92(D)	-4580	-4701	4174	-3014	-5700	-3967	-3905	-6376	-4478	-6024	-5744	-3355	-4501	-3870	-4926	-4440	-4750	-5894	-4922	-5231	20300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
93(E)	-1123	-2199	-983	2715	-2589	-2046	-942	-2250	-625	-2356	1979	-870	-2250	-554	-1093	463	932	-1902	-2660	-2064	20400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
94(H)	399	-1137	-2012	-14	1582	-2306	1600	246	-1252	190	-325	-1456	-2374	1474	-1479	-94	-905	896	-1557	-1158	20500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
95(Q)	-2742	-3142	-2766	-2681	-2790	-3344	-2460	-160	-1802	-2456	-2353	-2682	-3710	4317	-1866	-2894	-2844	-2559	-3295	-2711	20600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

96(A)	1981	-2315	-809	-268	-2645	-531	-579	-2374	232	-2350	-1445	-567	1217	711	445	447	-874	-1951	-2540	-1883	**20700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
97(D)	491	-2351	1394	1381	-2671	-1854	1062	-2421	1010	-2367	-1440	-489	-1947	1017	362	-760	250	-623	-2535	-1852	**20800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
98(V)	-2039	-1706	-4456	-3939	-1846	-3939	-3049	1986	-3656	1460	-826	804	-3870	-3351	-3565	-3105	-2000	2330	-2796	-2442	**20900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
99(Y)	-4840	-3766	-5230	-5581	1898	-5109	-1300	-3727	-5135	-3041	-3132	-3723	-4964	-3861	-4501	-4357	-4690	-3883	3325	4377	**21000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
100(E)	-163	-2353	-734	1681	-2674	-1859	888	-2422	1668	-792	-1443	777	-1952	890	286	-766	238	-1975	-2536	-1856	**21100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-257	-8139	-2649	-894	-1115	-701	-1378	*	*												
101(E)	1017	-2763	862	2042	-3060	-1913	-775	-2836	-495	-2773	-1886	1956	-2143	-136	-1056	265	-1185	-2377	-2948	-2207	**21200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-9	-7891	-8933	-894	-1115	-338	-2261	*	*												
102(E)	-944	-2422	863	2138	-2740	-436	-567	-2493	894	-2437	-1515	-518	-1994	1767	-673	109	-885	-1023	-2605	-1917	**21300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
103(I)	-2660	-2156	-5316	-4965	-2520	-5119	-4900	3165	-4894	297	-1251	-4775	-4828	-4705	-4975	-4470	-2642	2240	-4202	-3814	**21400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
104(E)	1068	-2341	-760	2003	628	-1887	876	-2380	1240	-2347	-1436	-529	-1983	881	-618	-804	-855	-1954	-2530	-1862	**21500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
105(P)	-343	-3144	1561	442	-3538	-489	-1216	-3329	-1038	-3274	-2420	-848	2974	-812	-1635	469	-1644	-2849	-3462	-2693	**21600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
106(N)	-1173	-2375	-814	827	-2376	-2071	1767	-2279	-479	-2336	-1509	3151	-2218	-415	-957	-1093	-1120	-198	-2486	647	**21700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
107(M)	-3415	-2890	-5826	-5252	-1352	-5488	-4282	1361	-5022	2621	2728	-5181	-4778	-4005	-4613	-4776	-3292	69	-3071	-3194	**21800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
108(K)	-1941	-3098	-1997	-1232	-3650	-2740	-1025	-3210	3059	-3010	-2208	499	-2766	1457	1261	-1817	-90	-2858	-3002	-2622	**21900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
109(P)	1129	-2426	-740	964	-2747	-1913	-589	-2491	1139	-2440	-1525	-552	1941	1446	-655	-480	-913	-2050	-2610	-1935	**22000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
110(G)	-2276	-2907	-2347	-2709	-4832	3554	-3349	-5005	-3678	-5053	-4315	1193	-3507	-3243	-3937	-2418	-2674	-3974	-4703	-4521	**22100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
111(A)	1730	-2349	958	-198	-2661	-1868	-535	-2405	927	-2362	-1444	414	-1966	788	-630	790	-840	-303	-2540	-1863	**22200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
112(T)	1350	-1149	-14	-1461	-1155	-2314	-1111	758	-1275	-1024	1167	-1475	-2388	-1111	-1501	334	1843	354	-1581	-1182	**22300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
113(L)	-3333	-2796	-5806	-5293	-1506	-5535	-4502	1096	-5103	2935	-282	-5232	-4857	-4172	-4762	-4864	-3236	506	-3264	-3351	**22400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
114(A)	1769	-1525	158	-857	-1603	148	-891	-1181	-752	187	-712	-1040	-2228	660	-1135	-1111	-913	1305	-1913	-1442	**22500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
115(F)	-4110	-3437	-5436	-5431	4216	-5143	-2159	-1742	-5074	563	-1124	-4290	-4871	-3987	-4561	-4547	-4016	-2374	-1356	-292	**22600%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
116(A)	3091	-1829	-3998	-4219	-4413	119	-3637	-4216	-4134	-4469	-3523	-2798	-2896	-3656	-3927	1514	-1679	-2983	-4632	-4539	**22700%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
117(H)	-5197	-4539	-4720	-5009	-4036	-4506	5435	-6314	-4911	-5786	-5667	-4954	-4960	-5011	-4732	-5391	-5395	-6022	-4063	-3641	**22800%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
118(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**22900%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
119(F)	-4044	-3387	-5534	-5444	4093	-5246	-2370	-1514	-5107	1089	-868	-4443	-4880	-3998	-4592	-4639	-3934	-2200	-1536	-523	**23000%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
120(N)	885	-1899	-2020	-1781	-2956	-2135	-1925	-2602	-1809	3	-2052	3468	-2633	-1676	-2141	413	-1437	-2139	-3194	-2737	**23100%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
121(I)	-2673	-2169	-5324	-4969	-2477	-5123	-4876	3293	-4893	358	-1211	-4780	-4824	-4681	-4961	-4472	-2653	1969	-4158	-3791	**23200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
122(H)	-3381	-3705	-3197	-3491	-4166	638	5216	-5496	-3798	-5304	-4811	-3481	-4185	-3770	-3879	-3508	-3702	-4793	-4170	-3759	**23300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
123(Y)	-4816	-3757	-5210	-5549	3410	-5097	2153	-3719	-5105	-3041	-3127	-3715	-4955	-3851	-4483	-4344	-4669	-3870	-547	3677	**23400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
124(G)	-1065	-2519	948	-272	-2820	1844	998	-2566	972	-284	-1622	1553	-2090	-229	-802	-938	-1011	-2133	-2708	-2021	**23500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													

125(Q)	-412	-2285	-2466	-2186	-2068	-2877	-2019	-1589	-1588	1526	-1121	-2187	-3153	3585	-1718	-2137	-1964	-1719	-2789	-2414	23600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
126(I)	-2254	-1916	-4813	-4439	-2466	-4221	-3932	3248	-4248	-1515	-1324	-4044	-4259	-4063	-4255	230	-2280	2003	-3673	-3237	23700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
127(K)	-888	-2234	334	1172	-2504	-1881	-546	-93	1370	-300	-1337	465	-1974	655	-646	-794	-827	1255	-2448	-1798	23800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
128(P)	715	-1925	-3618	-3897	-4464	653	-3594	-4274	-4053	-4520	-3596	-2770	3775	-3593	-3911	-1550	-1770	-3067	-4647	-4548	23900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
129(P)	479	-2398	-1173	-637	-2915	-2106	-848	-2610	-289	-2586	-1713	-884	2238	1247	2195	51	-1147	-2184	-2757	-2174	24000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
130(A)	1787	-2663	1377	529	-2976	-1992	-762	-2736	1785	-2680	-1776	-623	-2161	-319	-936	297	-1120	-2285	-2853	-2146	24100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	*	*												
131(F)	-1308	-1104	-2227	-2120	3516	-2093	-244	-196	-1891	64	66	-1626	-2278	-1503	-1798	-1617	-1350	-389	305	1335	24200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
132(P)	-603	-937	-997	-1058	-1832	-1041	-1092	-1737	-1074	-1874	-1416	-992	3539	-1065	-1192	-789	-866	-1383	-1765	-1661	24300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
133(K)	-804	-1483	-564	-230	-1920	-1335	-101	-1605	2889	-1630	-1021	-349	-1569	232	698	-786	-759	-1358	-1637	-1317	24400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-109	-3775	*	*												
134(D)	-2405	-4159	3349	-651	-4260	-261	-1744	-4307	-1947	-4207	-3514	2151	-2936	-1416	-2754	-2102	-2471	-3802	-4324	637	24500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
135(I)	-2047	-1713	-4504	-3983	-1821	-3943	-3061	2461	-3697	1581	-797	-3593	-3873	-3371	-3587	-342	-2009	1904	-2784	-2441	24600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
136(D)	-2024	-3444	3495	-680	-3868	-2331	-1596	-44	-1632	-3675	-2915	685	-2782	-1248	-2305	478	-2088	-3196	-3911	-3098	24700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
137(V)	-3122	-2888	-5092	-5160	-3522	-4180	-4687	-905	-5060	-2626	-2570	-4662	-4579	-4940	-4923	-4013	-3297	3796	-4414	-4190	24800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
138(I)	-53	-875	-3230	-2609	1867	-393	-1422	2613	-2236	-723	-81	-2157	-2608	-1885	-2086	-1633	276	-271	-1325	844	24900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
139(M)	315	-2345	-4754	-4279	-1396	-4001	-3301	-697	-3877	816	4676	-3879	-3994	-3361	-3676	-3242	-2531	-1114	-2746	-2629	25000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
140(V)	-2623	-2122	-5301	-4990	-2770	-5102	-5132	2415	-4945	-1532	-1474	-4791	-4869	-4890	-5102	-4483	-2619	3206	-4506	-3991	**25100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
141(A)	3405	2528	-4529	-4796	-4340	-2257	-3851	-3901	-4447	-4351	-3532	-3057	-3052	-3976	-4112	-1643	-1844	-2929	-4572	-4519	**25200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
142(P)	-4853	-4392	-5213	-5573	-5853	-4408	-5077	-6679	-5780	-6281	-6067	-5357	4310	-5648	-5396	-5166	-5194	-6092	-4900	-5786	**25300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
143(K)	-4484	-4357	-4380	-3992	-5413	-4236	-3307	-5555	3994	-5171	-4707	-3921	-4535	-3079	-2169	-4529	-4408	-5264	-4403	-4729	**25400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
144(G)	2167	-1833	-3963	-4199	-4430	2715	-3642	-4236	-4146	-4489	-3540	-2795	-2898	-3661	-3939	910	-1682	-2994	-4647	-4556	**25500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
145(P)	-2604	-2948	-4094	-4235	-3544	-3269	-3767	-3353	-3912	-3066	2095	-3659	4036	-3912	-3822	-2883	-2963	-3249	-4027	-3787	**25600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
146(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**25700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
147(H)	-2569	-3440	-1867	-1702	-3820	-2996	4731	-3830	634	-3639	-2963	-1838	1551	-1305	-748	-2470	-2510	-3482	-3434	-2990	**25800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
148(T)	194	-1498	-3255	-2899	-2240	-2226	-2291	-1754	-2652	1634	-1430	-2330	-2747	-2399	-2684	567	2687	-1484	-2682	-2351	**25900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
149(V)	-3122	-2888	-5092	-5160	-3522	-4180	-4687	-905	-5060	-2626	-2570	-4662	-4579	-4940	-4923	-4013	-3297	3796	-4414	-4190	**26000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
150(R)	-4845	-4446	-5107	-4682	-5507	-4412	-3791	-5946	-2789	-5502	-5118	-4521	-4754	-3672	4219	-4989	-4832	-5644	-4538	-4993	**26100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
151(R)	-962	-2395	-777	1012	-2721	76	1031	-2459	-142	-2413	-1501	-560	-2018	-128	2308	1224	-66	-2023	-2585	-1919	**26200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
152(E)	-902	-2032	-899	2078	-2228	-1934	-611	-1897	-259	-221	-1156	520	-2024	816	-736	-858	1303	-287	-2295	537	**26300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
153(Y)	-4820	-3765	-5219	-5565	3303	-5093	-1317	-3703	-5127	-3017	-3111	-3732	-4959	-3868	-4500	-4356	-4679	-3867	-565	4052	**26400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

154(V)	129	-1901	-989	821	-2060	-1969	-654	-1704	498	-52	-1037	-703	-2057	695	-796	443	-344	1871	-2192	-1626	**26500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
155(Q)	576	-2355	344	1156	-2675	-1856	-515	-508	1502	-2370	-1444	571	-1949	1878	419	-764	-822	-1976	-2538	-1856	**26600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
156(G)	-3239	-3889	516	-2361	-5355	3646	-3337	-5629	-3818	-5498	-4951	-2619	-3905	-3187	-4377	-3211	-3532	-4837	-4895	-4826	**26700%**
-	-149	-500	232	44	-381	399	105	-627	211	-466	-721	277	393	45	95	359	117	-368	-295	-250	
-	-155	-3318	-9181	-2159	-366	-701	-1378	*	*												
157(G)	753	-2516	-789	488	-2848	2300	-672	-2582	596	-2529	-1627	481	-2112	-224	471	-962	-1024	-2149	-2694	-2033	**27300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
158(G)	-52	-2212	-3792	-4133	-4698	3627	-3843	-4580	-4356	-4812	-3937	-3058	-3216	-3901	-4170	-1874	-2095	-3384	-4734	-4766	**27400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
159(V)	-2485	-2030	-5123	-4769	-2667	-4797	-4593	2349	-4661	-1545	-1424	-4502	-4648	-4554	-4752	-4115	825	2986	-4159	-3678	**27500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
160(P)	-2541	-3139	-2413	-2753	-4726	-2991	-3342	-5055	-3527	-5058	-4393	1199	4031	-3244	-3757	-2665	-2911	-4148	-4583	-4362	**27600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
161(C)	1577	3078	1357	-656	-2664	-219	891	-2359	-617	-2434	-1576	-891	-2199	-545	-1093	872	-1043	-1946	-2701	-2102	**27700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
162(L)	-2140	-2404	-3995	-3997	-2053	-3121	-3283	-1687	-3689	3041	-1200	-3360	-3626	-3433	-3567	480	-2414	-1973	-3145	-2755	**27800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
163(I)	-2527	-2092	-5072	-4613	2047	-4674	-3911	2668	-4413	117	-841	-4323	-4439	-4023	-4320	-3916	-2488	2176	-3342	-3040	**27900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
164(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	**28000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
165(V)	-2623	-2122	-5301	-4990	-2770	-5102	-5132	2426	-4946	-1532	-1474	-4791	-4869	-4891	-5102	-4483	-2619	3200	-4506	-3991	**28100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
166(H)	-495	-2631	903	-2051	722	-3242	3753	-2386	-2056	-2342	-1863	-2047	-3330	-1815	-2362	-2318	-2233	-2272	-981	3315	**28200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
167(Q)	-4589	-4392	-3927	-4146	-5099	-4221	-4099	-5973	-3840	-5564	-5304	-4230	-4693	4575	-3826	-4704	-4772	-5612	-4577	-4751	**28300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
168(D)	-2873	-4605	3943	-902	-4948	-2633	-2157	-5087	-2604	-4922	-4387	428	-3235	-1872	-3575	-2522	-3009	-4491	-4932	-3946	**28400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
169(A)	1776	-1612	-1274	138	-1698	-2092	-816	-1295	150	-1526	-780	-943	1212	-538	-1013	-1056	-894	1275	-1968	635	**28500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
170(S)	-1545	-2420	1001	-1518	-4049	-2206	-2206	-3839	-2264	-3938	-3103	-1627	-2814	-1909	-2758	2666	2313	-3045	-4143	-3560	**28600%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
171(G)	-2999	-3461	-2978	-3207	-5161	3669	-3454	-5283	15	-5188	-4565	-3174	-3946	-3312	-3140	-3128	-3317	-4492	-4622	-4749	**28700%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
172(N)	-887	-2349	-736	911	-2668	-1860	-518	-2415	711	-2363	-1440	1880	-1954	436	799	671	1230	-539	-2534	-1855	**28800%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
173(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	**28900%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
174(K)	-1368	-2597	-1358	-720	-2953	-2298	1933	-2603	1975	851	-1706	-958	-2360	859	1012	-1273	1094	-2245	-2690	-2165	**29000%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
175(D)	-27	-2613	2320	1049	-2923	-1962	1973	-2684	544	-2624	-1712	1408	-2115	345	-869	-969	-1058	-2231	-2793	-2087	**29100%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
176(V)	-1096	-938	-3279	-2658	-899	-2643	-1513	1265	1006	1388	-124	-2232	-2688	-1950	-2151	-1725	-311	1669	-1425	827	**29200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
177(A)	3342	-1826	-4064	-4295	-4368	111	-3651	-4129	-4159	-4412	-3484	-2820	-2903	-3689	-3929	-1469	409	-2949	-4599	-4508	**29300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
178(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	**29400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
179(S)	2216	-1831	-3961	-4157	-4409	656	-3608	-4213	-4076	-4462	-3514	-2782	-2892	-3613	-3895	2686	-1675	-2982	-4623	-4523	**29500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
180(Y)	-3634	-3050	-4918	-4872	36	-4597	-1405	223	-4437	-250	-1998	-3545	-4494	-3522	-4004	-3782	-3536	-2621	2928	4349	**29600%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
181(A)	3391	-1860	-3998	-4279	-4411	-2128	-3684	-4207	-4197	-4490	-3565	-2837	-2929	-3729	-3959	706	-1718	-3001	-4636	-4534	**29700%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
182(K)	370	307	113	-351	-2280	-1925	-615	324	1888	-2040	-1194	-632	-2022	-205	-736	1541	14	-1628	-2330	-1726	**29800%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													

183(G)	2572	-2028	-3934	-4246	-4575	2752	-3783	-4406	-4316	-4661	-3751	-2958	-3070	-3837	-4092	-1679	-1898	-3191	-4701	-4686	29900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
184(I)	-2178	-1808	-4630	-4153	-2094	-4190	-3417	3121	-3909	311	-1023	698	-4099	-3656	-3864	-3388	-2148	1742	-3147	-2761	30000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
185(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	30100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
186(G)	1392	2751	-4353	-4536	-4308	2864	-3681	-4084	-4233	-4354	-3425	-2859	-2890	-3744	-3957	712	-1656	-2914	-4553	-4470	30200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
187(G)	855	-1822	-3738	-3769	-4188	2507	-3358	-3950	-3667	-4196	-3283	-2667	986	-3302	-3621	-1441	2334	-2867	-4408	-4236	30300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
188(R)	-3706	-3692	-4490	-3846	1391	-4057	-2273	-3795	-1906	-3355	-3181	-3458	-4263	-2675	3948	-3768	-3671	-3813	-2293	-1328	30400%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-155	-8139	-3345	-894	-1115	-701	-1378	*	*												
189(A)	2844	-1670	-3814	-3873	-4048	-1958	-3316	-3787	-3686	-4061	-3156	-2598	-2734	-3301	-3572	1088	1907	-2713	-4286	-4136	30500%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-7992	-9034	-894	-1115	-1303	-750	*	*												
190(G)	-4176	-3995	-4855	-5222	-5686	3828	-4823	-6386	-5533	-6087	-5741	-4896	-4606	-5312	-5178	-4461	-4560	-5613	-4754	-5635	30600%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-7992	-9034	-894	-1115	-422	-1980	*	*												
191(V)	-2496	-2036	-5139	-4788	-2680	-4825	-4637	2522	-4686	-1549	-1431	-4527	-4668	-4584	-4783	-4148	629	2919	-4191	-3706	30700%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
192(I)	-2760	-2307	-5270	-4785	1172	-4884	-3992	3346	-4576	662	-572	-4539	-4535	-4008	-4400	-4134	-2703	757	-3223	-3041	30800%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
193(E)	454	-3086	-801	3279	-4125	-2346	-1868	-3919	-1868	-3932	-3156	-1291	1208	-1533	-2450	-1842	-2097	-3305	-4121	-3395	30900%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
194(T)	-2738	-3109	-4509	-4810	-4918	-3305	-4346	-4865	-4769	-5072	-4551	-3987	-3998	-4580	-4545	-2999	4033	-4113	-4684	-4915	31000%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
195(T)	-1323	-1975	-2766	-2978	-4207	-2152	-3105	-4004	-3295	-4229	-3354	134	-2901	-2988	-3387	-113	3742	-2969	-4405	-4119	31100%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
196(F)	-4110	-3437	-5436	-5431	4216	-5143	-2159	-1742	-5074	563	-1124	-4290	-4871	-3987	-4561	-4547	-4016	-2374	-1356	-292	31200%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
197(K)	-111	-2844	-1470	1220	-3294	-2415	-860	-2939	2448	-2798	-1945	-1056	-2475	767	2054	-1421	-1432	-2544	-2858	-2356	31300%
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
198(E)	545	-3735	1715	2880	-3981	-2308	-1442	-3818	-1408	-3725	-2924	-909	-2713	1281	-2087	-1777	-2041	-3331	-3909	-3046	**31400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
199(E)	-4574	-4665	-2714	3919	-5655	-3995	-3886	-6219	-4238	-5898	-5604	-3415	-4513	-3838	-4570	-4456	-4726	-5786	-4878	-5197	**31500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
200(T)	-1211	1331	-3446	-2962	-1399	-2495	-1930	-869	-2610	-1374	-769	-2403	-2815	-2309	-2526	-1685	3305	195	-1933	1430	**31600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
201(E)	-1941	-3222	-921	3293	-3618	-2473	-1316	-3186	916	-3225	-2465	-1147	-2749	-923	-1064	-1790	-1902	-171	-3398	-2802	**31700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
202(T)	-1286	-1898	-3764	-4016	-4329	-2157	-3572	-4120	-3959	-4408	-3517	-2789	-2947	-3588	-3797	697	3756	-2989	-4554	-4399	**31800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
203(D)	-4580	-4701	4174	-3014	-5700	-3967	-3905	-6376	-4478	-6024	-5744	-3355	-4501	-3870	-4926	-4440	-4750	-5894	-4922	-5231	**31900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
204(L)	-3705	-3122	-6060	-5527	-1359	-5814	-4569	1065	-5292	3069	-146	-5564	-4963	-4163	-4828	-5215	-3571	-1279	-3159	-3298	**32000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
205(F)	-3777	-3220	-5271	-5259	4268	-4892	-2120	417	-4916	-1143	-1314	-4142	-4753	-3956	-4473	-4270	-3740	-1899	-1349	-269	**32100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
206(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**32200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
207(E)	-4574	-4665	-2714	3919	-5655	-3995	-3886	-6219	-4238	-5898	-5604	-3415	-4513	-3838	-4570	-4456	-4726	-5786	-4878	-5197	**32300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
208(Q)	-3157	-3746	-3170	-2450	-4497	-3515	-1763	-4161	-443	-3809	-3189	-2392	-3620	4200	1284	-3063	-2944	-3900	-3556	-3420	**32400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
209(A)	2672	-1334	-3318	-2853	-1740	371	-2072	483	-2577	-1549	-928	-2359	-2798	-2295	-2567	-1629	191	932	-2245	-1899	**32500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
210(V)	-2620	-2125	-5293	-4983	-2756	-5076	-5100	1877	-4932	-1522	-1466	-4777	-4855	-4870	-5082	-4456	-2619	3416	-4480	-3969	**32600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
211(L)	-4414	-3800	-5638	-5628	-2290	-4980	-4628	-1886	-5423	3316	-1236	-5514	-4997	-4750	-5002	-5379	-4399	-2629	-3665	-3690	**32700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

212(C)	-2243	5044	-4840	-4445	-1998	-3905	-3598	-31	-4138	449	-930	-3902	-4040	-3778	-4010	-3184	-2306	1347	-3209	-2883	**32800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
213(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**32900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
214(G)	677	-2128	-3838	-4171	-4647	3536	-3816	-4506	-4340	-4749	-3857	-3009	-3149	-3871	-4137	-1784	-2005	-3297	-4725	-4735	**33000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
215(V)	378	724	-3707	-3104	-1180	-2986	-1919	1210	-2734	1302	-359	-2627	-3014	-2382	-2566	-2089	1123	1949	-1773	-1423	**33100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
216(M)	-948	-1407	-1515	156	-1452	-2164	1677	-1030	-821	-1302	1976	-1113	-2245	-718	-1173	773	1715	1332	-1794	-1343	**33200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
217(E)	1397	-2528	-725	2286	-2932	240	-791	-2681	328	-2645	-1744	-674	-2162	-351	-939	545	-1095	-2227	-2828	-2143	**33300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
218(L)	-3705	-3122	-6060	-5527	-1359	-5814	-4569	1065	-5292	3069	-146	-5564	-4963	-4163	-4828	-5215	-3571	-1279	-3159	-3298	**33400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
219(V)	-2600	-2108	-5251	-4894	-2568	-5025	-4783	2479	-4810	-1354	1358	-4683	-4772	-4654	-4895	-4362	-2584	3018	-4181	-3758	**33500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
220(K)	-1633	-2905	-1573	706	-3375	-2487	-900	-3003	2925	-2849	-2008	-1128	-2541	1714	784	-1509	-105	-2617	-2894	-2418	**33600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
221(A)	2352	2066	-2593	-2000	-947	-2434	-1271	-486	-32	-832	714	-1817	-2498	-1493	-1792	-1483	274	453	-1419	501	**33700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
222(G)	224	-1905	-3562	-3696	-3684	3361	-3297	-3220	-3625	81	-2886	-2733	-2977	-3326	-3545	-1606	-1763	-2574	-4068	-3810	**33800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
223(F)	-4781	-3756	-5207	-5542	4341	-5070	-1342	-3653	-5111	-2971	-3065	-3743	-4949	-3874	-4496	-4351	-4650	-3829	-591	1725	**33900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
224(E)	-2413	-4114	221	3465	-4392	-2485	-1689	-4248	-1608	-4112	-3396	-1094	-2951	-1336	871	-2119	-2441	-3763	-4239	-3395	**34000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
225(T)	-1461	-1864	-3139	-2645	-2659	-2483	-2136	-1734	-1646	-2359	-1761	-2298	-2936	-1995	920	-1748	3354	967	-2989	-2654	**34100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
226(L)	-3831	-3266	-5314	-5148	-673	-5068	-2476	-1443	-4706	3059	-789	-4359	-4756	-3864	-4314	-4462	-3729	-2115	-1672	1736	**34200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
227(V)	-1819	-1960	-4426	-4359	-2977	-3037	-3800	-439	-4098	-2145	-1909	-3451	-3600	-3879	-4011	-2397	2510	2999	-3974	-3608	**34300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
228(E)	-2863	-4790	1397	3563	-4990	-2594	-2061	-5021	-2476	-4848	-4298	-1204	-3182	-1761	-3454	-2476	-2968	-4462	-4951	-3920	**34400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
229(A)	2686	-1916	-1440	275	-2529	-292	-1240	-2176	-998	-2345	1183	-1184	-2355	-905	-1425	512	-1158	-1817	-2697	-2174	**34500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
230(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**34600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
231(Y)	-4099	-3483	-4921	-5048	-109	-4705	-1565	-2914	-4494	-2334	2010	-3723	-4707	-3735	-4111	-4065	-4068	-3172	-847	4618	**34700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
232(Q)	1711	-2410	-772	934	-2739	-1925	-604	-2477	-171	-2433	-1524	-574	-2035	2086	345	902	-923	-2042	-2608	-1941	**34800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
233(P)	-3403	-4071	-1922	817	-5220	-3359	-3173	-5423	-3337	-5281	-4771	-2564	4045	-2989	-3760	-3320	-3624	-4817	-4763	-4636	**34900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
234(E)	-2870	-4786	1265	3587	-4993	-2600	-2068	-5026	-2483	-4852	-4303	-1212	-3188	-1768	-3458	-2484	-2974	-4467	-4950	-3925	**35000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
235(M)	-3089	-2618	-5526	-4976	-1443	-5128	-4045	653	-4735	1429	4269	-4803	-4610	-3915	-4423	-4378	-2995	1140	-3054	-3074	**35100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
236(A)	3631	-2768	-4492	-4815	-4888	-2992	-4271	-4781	-4818	-5025	-4365	-3727	-3728	-4477	-4545	-2567	-2762	-3852	-4724	-4942	**35200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
237(Y)	-4797	-3764	-5203	-5543	1114	-5069	-1339	-3694	-5111	-3013	-3107	-3741	-4951	-3876	-4497	-4354	-4666	-3859	-588	4723	**35300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
238(F)	-3828	-3605	-4146	-4086	4292	-4207	-2060	-3492	774	-3071	-3005	-3556	-4434	-3287	-2952	-3868	-3858	-3571	-1593	-494	**35400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
239(E)	-2775	-4471	-511	3582	-4815	-2610	-2057	-4863	-2317	-4711	-4124	1234	-3182	-1755	-3103	-2442	-2884	-4306	-4753	-3820	**35500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
240(C)	-1407	5023	-4397	-4323	-3016	-2468	-3398	-1251	-3952	-2540	-2082	-3044	-3133	-3588	-3744	-1803	1473	1125	-3677	-3390	**35600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

241(L)	-3370	-2847	-5795	-5233	-1386	-5465	-4298	708	-5010	2859	1349	-5155	-4777	-4026	-4622	-4755	-3254	814	-3103	-3213	**35700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
242(H)	-2519	-4224	-445	946	-4287	-2505	4583	-4377	-1764	-4237	-3571	2007	-3009	-1439	-2361	-2209	-2569	-3893	-4267	-3353	**35800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
243(E)	-3177	2571	-2701	3711	-4851	-3438	-3479	-4765	-3558	-4932	-4406	-3081	-4005	-3370	-3802	-3269	-3451	-4260	-4554	-4524	**35900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
244(L)	-85	-1333	-3893	-3280	-1111	-3185	-2083	1066	-2910	2310	1961	-2823	-3170	-2501	-2721	-2289	-113	436	-1859	-1558	**36000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
245(K)	-2513	-3173	-2941	-2370	-4402	-3094	-1824	-3895	3666	-3734	-3068	-2271	-3377	-1447	-616	-2537	751	-3485	-3625	-3471	**36100%**
-	-149	-500	233	43	-381	400	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-155	-3318	-9181	-196	-2974	-701	-1378	*	*												
246(L)	-3571	-3023	-5954	-5375	-1321	-5646	-4390	-632	-5138	2962	1671	-5358	-4852	-4044	-4689	-4963	-3436	742	-3082	-3239	**36300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
247(I)	-2980	-2484	-5473	-5109	-1958	-5196	-4587	3728	-4915	267	-781	-4933	-4833	-4427	-4799	-4598	-2949	-64	-3627	-3397	**36400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
248(V)	-1685	-1668	-4095	-3732	-2081	-3082	-2893	-227	-3402	-1488	1383	-3123	-3419	-3145	-3320	367	-1807	3332	-2874	-2504	**36500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
249(D)	-2963	-4569	3864	-1039	-4953	-2751	-2187	-4998	767	-4822	-4260	-1424	-3314	-1891	-3072	-2624	-3060	-4467	-4770	-3962	**36600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
250(L)	-2768	-2715	-4842	-4633	-1675	-3998	-3790	-1038	-4207	3056	-562	-4150	-4179	-3740	-3989	-3399	699	-1545	-3154	-3067	**36700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
251(M)	-2822	-2356	-5342	-4861	-1759	-4985	-4151	2587	-4663	173	4005	-4649	-4601	-4076	-4487	-4251	-2764	766	-3321	-3210	**36800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
252(Y)	-4562	-3630	-5142	-5401	1516	-4992	-1300	-3544	-4968	-2963	-2986	-3671	-4868	-3786	-4393	381	-4432	-3662	2413	4375	**36900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
253(E)	-1959	-3457	-568	3135	-3841	-2347	-1410	-3622	-1165	-3547	-2751	-1000	-2716	1879	-1666	-1757	692	-3162	-3709	-2960	**37000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
254(G)	-347	-2818	-1215	201	-3253	2635	-921	-2921	1474	-2822	-1972	-1002	-2459	-486	658	-1397	-1435	-2521	-2923	-2378	**37100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
255(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**37200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
256(I)	-2042	-1769	-4321	-3740	-1316	-3753	-2668	3134	-3389	1017	1999	194	-3653	-2958	-3221	-2884	-1980	-344	-2325	-2073	**37300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
257(A)	1914	-1640	-1237	128	-1748	-661	-793	-1355	-577	-49	-817	-905	-2155	-498	-993	1037	162	-1149	-2002	624	**37400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
258(N)	365	-3809	1001	557	-4083	1196	-1518	-3930	-1535	-3838	-3055	3219	-2763	-1148	-2243	-1845	-2131	-3433	-4027	-3144	**37500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
259(M)	-3656	-3159	-5816	-5350	-1349	-5421	-4248	-822	-4928	948	4920	-5248	-4838	-4039	-4539	-4860	-3558	-1557	-3044	-3030	**37600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
260(R)	-1614	-1949	-2260	-1663	-886	-2765	-1089	-1596	-1089	360	-1133	2239	-2814	-1215	2408	-1789	-1535	-1468	1995	1546	**37700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
261(Y)	-1548	-2973	568	-509	-2846	-2207	1172	-2986	441	-2941	-2110	1645	-2446	-659	-1264	-1389	-1509	-2582	-2924	3695	**37800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
262(S)	279	-1844	-3877	-4131	-4448	136	-3634	-4260	-4132	-4511	-3561	-2782	-2903	-3648	-3936	3391	-1692	-3009	-4661	-4566	**37900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
263(I)	-2653	-2149	-5311	-4961	-2538	-5114	-4905	2957	-4891	332	-1267	-4770	-4827	-4713	-4978	-4466	-2636	2521	-4218	-3821	**38000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
264(S)	-2212	-2711	-4019	-4348	-4697	-2899	-4045	-4988	-4527	-5102	-4364	-3492	-3638	-4203	-4355	3681	-2664	-3902	-4616	-4605	**38100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
265(N)	-2725	-4778	2906	990	-4896	-2486	-1922	-4885	-2320	-4718	-4140	3045	-3069	-1612	-3311	-2334	-2821	-4325	-4919	-3794	**38200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
266(T)	-2061	-3396	-596	903	-4071	-2367	-1713	-3874	-1685	-3859	-3103	2125	-2850	-1369	-2277	-1897	3157	-3351	-4055	-3271	**38300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
267(A)	3410	-2035	-3979	-4290	-4573	659	-3798	-4400	-4332	-4660	-3754	-2978	-3080	-3858	-4099	-1690	-1908	-3194	-4697	-4687	**38400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
268(E)	-2118	-3486	-1036	2964	-3935	-2588	-1284	-3596	1878	-3417	-2636	-1209	-2827	1323	-773	-1923	-2032	-3199	-3455	-2917	**38500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
269(Y)	-4524	-3618	-5100	-5310	1910	-4972	-1299	-3522	634	-2951	-2965	-3649	-4847	-3741	-4299	-4199	-4391	-3637	2997	4211	**38600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-155	-8139	-3345	-894	-1115	-701	-1378	*	*												

270(G)	-4176	-3995	-4855	-5222	-5686	3828	-4823	-6386	-5533	-6087	-5741	-4896	-4606	-5312	-5178	-4461	-4560	-5613	-4754	-5635	**38700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-7992	-9034	-894	-1115	-422	-1980	*	*												
271(D)	-2710	-4705	3025	1828	-4880	1758	-1932	-4863	-2320	-4703	-4115	-1084	-3073	-1621	-3297	-2330	-2809	-4301	-4894	-3793	**38800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
272(Y)	-2497	-2175	-4651	-4137	2447	-4046	-2215	255	-3766	892	1558	-3537	-3886	-3109	-3468	-3181	-2410	-1282	-1615	3508	**38900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
273(V)	-1425	-1250	-3480	-2894	-1283	-3035	-1955	691	-2570	-1	-429	-2568	-3060	942	-2514	-2129	1701	2578	-1907	-1553	**39000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
274(T)	516	-1643	-1918	-1401	-2170	-2112	-1387	-1759	-1234	-2016	-1265	-1442	-2421	-1149	887	1341	2345	822	-2454	-2006	**39100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
275(G)	677	-2128	-3838	-4171	-4647	3536	-3816	-4506	-4340	-4749	-3857	-3009	-3149	-3871	-4137	-1784	-2005	-3297	-4725	-4735	**39200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
276(P)	-992	-2210	343	-359	-2447	-1960	-675	-2143	533	-2204	-1351	-651	2813	-260	-802	465	-939	-873	-2467	1093	**39300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
277(R)	-1214	-2548	-1097	1072	175	-2145	-716	-2587	848	-2528	-1653	-795	-2228	-273	2862	417	-1133	-2191	-2671	-2084	**39400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
278(V)	289	-2035	-5133	-4789	-2692	-4777	-4639	2142	-4689	-1561	-1443	-4511	-4649	-4585	-4784	-4102	-2487	3125	-4202	-3717	**39500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
279(I)	-2265	-1919	-4828	-4452	-2473	-4254	-3954	3155	-4265	-1516	-1326	-4066	-4279	-4082	-4274	226	-2288	2182	-3688	-3250	**39600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
280(D)	-1731	-3162	2329	-550	-3318	-2239	-1273	-3221	-1145	-3214	-2403	2295	-2573	-899	-1742	-1561	1851	-2804	-3366	1327	**39700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
281(E)	1097	-2699	1227	2368	-2994	-2011	-796	-2753	381	-2704	-1808	-640	-2190	-358	-992	-1065	-1162	-2310	-2885	1152	**39800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
282(E)	-166	-2372	859	1835	-2692	-1861	1182	-2444	490	-2388	-1462	590	478	1356	-620	116	-837	-1994	-2555	-1871	**39900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
283(T)	228	-1688	-3655	-3444	-3179	-2145	-2891	-2611	-3215	-3076	-2328	-2557	-2821	-2916	-3214	2251	2366	1397	-3549	-3270	**40000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
284(K)	-2991	-3623	-3848	-2331	-4472	-3512	-1356	-3763	2942	-3413	1860	-2188	-3430	-937	2705	-2874	-2650	-3556	-3267	-3189	**40100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
285(E)	443	-2370	732	1690	-2691	-1863	-526	-2442	1639	-2385	-1460	-498	-1960	1120	437	106	-836	-1992	-2552	-1869	**40200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
286(A)	1871	-2286	-843	814	-2570	-1928	-578	269	1096	-2279	-1391	-586	-2019	-136	1056	-62	-881	-1889	-2488	-1849	**40300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
287(M)	-3656	-3159	-5816	-5350	-1349	-5421	-4248	-822	-4928	948	4920	-5248	-4838	-4039	-4539	-4860	-3558	-1557	-3044	-3030	**40400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
288(K)	-1646	-2891	-1591	287	-3346	-526	-912	-2971	2831	-2832	-1997	-1146	-2554	-472	1762	-1527	-1524	-2597	-2885	1245	**40500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
289(E)	-172	-2394	367	2205	-2713	-487	-545	-2465	-134	-2409	-1485	1305	-1975	663	831	-795	72	-2015	-2577	-1891	**40600%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
290(C)	1574	3024	-4584	-4122	-2155	-3932	-3330	1746	-3870	-1406	-1109	-3691	-3957	-3613	-3805	-3144	-2046	2342	-3118	-2720	**40700%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
291(L)	-187	-2175	-4307	-3889	-898	-3779	-2344	-944	-3485	2855	-476	-3390	-3782	-3025	-3298	-2972	-2345	-1269	-1846	1565	**40800%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
292(K)	862	-2347	143	1211	-2665	-1855	873	-2414	1692	-2362	917	-492	-1949	889	-603	-763	783	-1968	-2532	-1851	**40900%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
293(D)	-2148	-3878	2790	1765	-4119	-2356	-1511	-3962	-1467	-3852	-3075	24	-2782	-1139	2142	-1879	-2163	-3473	-4024	-3155	**41000%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
294(I)	-2630	-2131	-5302	-4991	-2737	-5092	-5106	3464	-4941	-1495	-1447	-4789	-4862	-4869	-5086	-4473	-2627	2071	-4467	-3968	**41100%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
295(Q)	346	-3134	-1818	-1401	-3862	-2760	-1314	-3433	1329	-3271	-2513	-1545	-2936	3817	-430	-2018	-2031	-3060	-3278	-2908	**41200%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
296(S)	-1354	-2895	1712	354	-3192	-2068	-914	-2967	-621	-2903	-2012	1817	-2288	724	-1177	1978	-96	-2508	-3076	-2340	**41300%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
297(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**41400%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													
298(E)	-437	-2374	-769	2013	-2697	-1895	-552	-2438	623	-2389	-1472	-536	-1991	-97	777	829	1488	-1999	-2559	-1890	**41500%**	
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249		
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*													

299(F)	-4347	-3577	-4619	543	3858	-4820	-1320	-3438	-4609	-2900	-2894	-3532	-4750	-3628	-4209	-4078	-4237	-3546	-601	2917	**41600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
300(A)	2827	-1603	-4068	-3628	-2047	-3165	-2823	1205	-3349	-1486	-1089	-3103	-3432	-3073	-3298	-2387	197	1100	-2812	-2447	**41700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
301(K)	-2364	-3363	-2464	163	-4038	-3041	-1188	-3501	2928	-3234	-2487	-1693	-3047	-758	2488	-458	-2152	-3194	-3159	-2895	**41800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
302(M)	-893	-2361	740	1780	-2680	-537	-524	-2429	930	-2376	1895	-498	-1958	-66	722	775	-833	-1982	-2545	-1864	**41900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
303(W)	-2965	-2553	-4795	-4482	3045	-4315	-1779	-1426	-4093	92	-987	-3564	-4179	-3330	-3743	-3475	-2878	664	4754	-84	**42000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
304(I)	-2621	-2136	-5246	-4859	-2364	-4987	-4577	3052	-4747	-250	1043	-4641	-4713	-4488	-4769	-4301	-2597	2384	-3934	-3596	**42100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
305(L)	684	-1319	-1741	761	-1375	-2223	-1037	-927	-1042	1693	-533	-1287	-2325	-915	-1349	767	-152	84	-1756	-1329	**42200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
306(E)	364	-4165	621	3314	-4393	-2398	-1686	-4282	-1836	-4169	-3446	728	-2893	-1339	-2626	-2045	-2402	-3767	-4362	-3401	**42300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
307(N)	-939	979	-1235	-681	-1738	1352	935	-1357	-549	-1572	-816	2186	-2155	-482	165	-1022	-880	30	-1990	1446	**42400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
308(Q)	667	-2393	-773	511	-2721	-316	-544	-2465	1211	-2405	-1484	584	-1989	2135	1528	-812	-868	-2019	-2566	-1895	**42500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
309(A)	2050	-1857	-1081	-526	-2013	-2012	188	-1645	-385	-1213	-1011	486	-2103	-346	1657	-953	306	-261	-2169	-1624	**42600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
310(G)	-1848	-2469	-2089	-2292	1262	2944	-2358	-3563	-2904	-3628	-3017	2347	-3167	-2550	-3185	-1986	-2162	-3006	-2814	-1874	**42700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
311(Y)	475	1019	-1606	-1042	225	-2192	-935	-946	-891	-1226	1222	357	-2267	1002	1577	-1172	-888	-800	-1730	2446	**42800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
312(P)	-87	-2372	-1362	756	-3738	-2205	-2007	-3445	-1924	-3575	-2761	-1555	3598	-1697	-2362	-1566	286	-2803	-3832	-3277	**42900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	*	*												
313(K)	-804	-1483	-564	-230	-1920	-1335	-101	-1605	2889	-1630	-1021	-349	-1569	232	698	-786	-759	-1358	-1637	-1317	**43000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-38	-5840	-6882	-894	-1115	-3098	-179	*	*												
314(E)	-766	-1695	521	2831	-2050	-1029	-293	-1804	-118	-1919	-1331	69	-1441	-4	-527	-653	-814	-1512	-1988	-1505	**43100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-109	-3775	*	*												
315(T)	-942	-2382	-739	1086	-2714	151	-581	-2459	-171	-2415	-1499	414	-2004	-128	839	1365	1730	-2017	-2592	-1915	**43200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
316(M)	-2196	-1920	-4499	-3891	1726	-3822	-2504	-645	-3523	1973	3030	-3442	-3673	-2938	-3257	-2944	-2114	326	-2014	1662	**43300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
317(H)	-883	-2314	-747	517	-2618	-1863	1714	-647	1272	-2322	-1408	1011	472	-69	433	-772	1411	-299	-2507	-1836	**43400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
318(A)	2474	-2397	-816	-367	-2797	-273	-722	-2529	555	-2507	-1610	592	-2110	837	-805	-138	-1006	-2092	-2699	-2039	**43500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
319(M)	-154	-986	-2485	-337	1024	-375	-1232	325	-444	867	1235	-1752	-2474	-1419	1020	-1455	670	-411	831	535	**43600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
320(R)	-1311	-2432	-1349	-724	-2724	-2272	-799	-2361	613	-644	1079	976	597	-382	2908	-1246	-1219	-2044	-2579	-2061	**43700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
321(R)	897	-2364	-833	905	-2678	-1930	-568	-2405	1293	-2366	1485	-575	-2020	-117	2045	95	-893	-1984	-2543	-1892	**43800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
322(N)	505	-2300	-750	523	-2598	121	-525	-594	485	95	-1395	1720	-1957	348	224	551	-821	-1910	-2497	-1828	**43900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
323(E)	444	-2266	-766	1488	-2551	-1871	-533	-2276	889	-2266	1474	1478	-1963	682	-629	-781	12	-279	-2472	476	**44000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
324(N)	1511	-1770	-728	-204	-2244	-1781	-426	282	601	-2121	-1133	1769	-1827	1398	447	-748	-766	-1713	-1779	-1595	**44100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
325(N)	1053	-3109	1756	1735	-3404	-2143	-1074	-3191	-846	-3124	-2254	2158	-2417	-657	-1430	-1361	197	-2727	-3303	-2540	**44200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
326(H)	-2064	-3071	-1245	-1267	-3262	-2570	4711	-3611	-1060	-3528	-2812	-1479	-2961	1288	-1287	670	-2133	-3161	-3310	-2601	**44300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
327(Q)	-891	-2294	332	948	-2585	-739	-537	-2316	-131	-138	-1391	-518	1404	2021	-634	-785	619	-561	-2495	-1830	**44400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												

328(I)	-1632	-1661	-2846	-86	-1626	-2983	-1663	3240	-1327	-191	-722	-2160	-3016	-1615	822	-2097	-1556	-560	-2157	-1805	**44500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
329(E)	-2734	-3605	-1382	3593	-3624	-2986	-2317	-3317	-2175	-3167	1983	-1898	-3440	-2054	-2556	-2649	-2820	-3234	-3920	-3370	**44600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-2336	-8139	-325	-894	-1115	-701	-1378	*	*												
330(W)	-1530	-1265	-2068	-1964	479	-1810	-483	-1181	-1470	-968	-802	-1648	-2104	-1454	-1405	-1757	-1583	-1218	5462	838	**44700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-38	-5840	-6882	-894	-1115	-109	-3775	*	*												
331(K)	8	-2067	-905	437	-2275	-1941	-611	-307	2031	-2031	-1189	-636	-2030	1425	-709	-864	-350	1337	-2323	-1722	**44800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
332(V)	-2445	-2012	-5067	-4708	-2628	-4682	-4459	1586	-4580	-1533	-1402	-4414	-4575	-4456	-4653	-3990	1117	3227	-4065	-3597	**44900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
333(G)	-4435	-4203	-5092	-5462	-5893	3834	-5028	-6627	-5765	-6297	-5970	-5141	-4804	-5546	-5385	-4727	-4815	-5862	-4924	-5849	**45000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
334(E)	1477	-2744	-762	2410	-3114	-2113	-884	-2850	-445	-2792	-1912	619	-2297	-451	1346	-1191	-1280	-2412	-2952	-2288	**45100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
335(K)	-1204	-2643	366	1309	-2998	-2086	-724	-2722	2626	-2637	-1741	-718	-2200	1198	862	-1073	-1133	-2287	-2770	-2133	**45200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
336(L)	-3571	-3023	-5954	-5375	-1321	-5646	-4390	-632	-5138	2962	1671	-5358	-4852	-4044	-4689	-4963	-3436	742	-3082	-3239	**45300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
337(R)	-4845	-4446	-5107	-4682	-5507	-4412	-3791	-5946	-2789	-5502	-5118	-4521	-4754	-3672	4219	-4989	-4832	-5644	-4538	-4993	**45400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
338(E)	943	-2422	1002	1200	-2741	377	-572	-2493	982	-2439	-1517	-522	-1998	-117	-681	1129	-534	-2044	-2609	-1921	**45500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
339(M)	-3391	-2886	-5774	-5202	-1338	-5407	-4210	-576	-4943	1721	4369	-5109	-4742	-3963	-4548	-4689	-3273	771	-3038	-3137	**45600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
340(M)	-1812	-1601	-3997	-3375	696	-3401	-2222	-516	-2983	441	4360	-2991	-3334	-2559	641	-2503	-1740	404	-1934	-1661	**45700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8139	-9181	-894	-1115	-701	-1378	*	*												
341(P)	102	-1789	-1729	-1313	428	-2112	-1410	-2021	-1236	-2235	-1470	-1394	3205	695	-1607	703	-1207	-1700	-2613	-2121	**45800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-259	-8139	-2637	-894	-1115	-701	-1378	*	*												
342(W)	-3486	-3022	-4312	-4121	1891	-4228	-1173	-2749	310	-2389	-2250	-3147	-4206	-2971	-3005	-3484	-3396	-2818	5644	611	**45900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	

-	-239	-7889	-2749	-894	-1115	-1590	-583	*	*												
343(I)	-2220	-1737	-4860	-4531	-2271	-4617	-4497	3348	-4448	-1059	-1008	-4311	-4407	-4340	-4559	-3974	-2216	2150	-3915	-3445	**46000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-114	-7660	-3820	-894	-1115	-1149	-865	*	*												
344(A)	1699	-2218	532	-33	-2555	-553	-413	-2304	1212	-2260	-1348	582	-1822	37	-536	966	-724	-1859	-2439	-1755	**46100%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-10	-7753	-8795	-894	-1115	-897	-1111	*	*												
345(A)	1523	-2068	-769	-231	-2383	1040	-522	-2084	928	-2120	-1245	-517	-1922	1335	-611	251	-768	-123	-2361	-1732	**46200%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-9	-7949	-8991	-894	-1115	-1432	-668	*	*												
346(N)	-1650	-3264	1760	-348	-3555	-230	-1131	-3362	324	-3285	-2448	2847	1624	-733	-1607	-1438	-1645	-2887	-3466	-2662	**46300%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-193	-7842	-3049	-894	-1115	-1432	-668	*	*												
347(K)	150	-2932	-2433	-1483	-3710	-2747	-907	-3141	3369	-2897	-2178	-1487	-2774	-488	1175	-1994	-1888	-2848	-2822	-2610	**46400%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-11	-7660	-8702	-894	-1115	-1824	-479	*	*												
348(L)	-740	-922	-1768	154	-921	-2070	-829	1384	247	1472	-100	-1202	-2134	-805	485	-1085	-677	471	-1340	-944	**46500%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-11	-7660	-8702	-894	-1115	-943	-1059	*	*												
349(V)	138	-1046	-3186	-2599	-1089	-2803	-1711	589	645	945	-236	-2305	-2830	-2012	-2236	-1892	-1154	2537	-1701	-1344	**46600%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-9	-7842	-8885	-894	-1115	-380	-2109	*	*												
350(D)	-2086	-3722	2888	1158	-4001	1601	-1510	-3811	-1573	-3773	-3018	-904	-2733	-1154	-2294	-1833	-2125	575	-3993	-3117	**46700%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8046	-9088	-894	-1115	-701	-1378	*	*												
351(K)	-78	-2308	398	-137	-2626	-483	542	-2374	2441	-2323	-1400	-447	531	720	-562	-724	290	-1929	-2493	-1812	**46800%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-8	-8046	-9088	-894	-1115	-701	-1378	*	*												
352(D)	898	-1911	1604	-335	-2089	-1892	-573	-1746	489	-144	811	-609	34	-182	-708	541	-74	-281	-2188	-1605	**46900%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-10	-7745	-8787	-894	-1115	-701	-1378	*	*												
353(K)	-1676	-2544	-1956	-1135	-2823	-2531	-800	-2469	2955	-2426	-1683	-1240	-2548	-407	1469	-1628	-1517	-119	-2497	1619	**47000%**
-	-149	-500	233	43	-381	399	106	-626	210	-466	-720	275	394	45	96	359	117	-369	-294	-249	
-	-11	-7649	-8691	-894	-1115	-701	-1378	*	*												
354(N)	-1074	-1925	-1092	-1190	-3523	-1781	-1760	-3345	-1701	-3432	-2589	3429	1327	-1460	-2119	1182	-1362	-2558	-3601	-3046	**47100%**
-	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	
-	*	*	*	*	*	*	*	*	0												

2017200439 20 Jan 2017

The Claims defining the invention are as follows:

1. A recombinant host cell comprising an engineered isobutanol production pathway and
    a. at least one of
        i. a heterologous polypeptide with ketol-acid reductoisomerase (KARI) activity selected from the group consisting of:
            1. a polypeptide having at least about 90% identity to a KARI enzyme derived from *Bifidobacterium angulatum, Bifidobacterium dentium, Zymomonas mobilis, Clostridium beijerinckii* or *Anaerostipes caccae*, or an active fragment thereof;
            2. a polypeptide having at least about 90% identity or at least about 95% identity to SEQ ID NO: 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, or 65; or
        ii. a heterologous polynucleotide encoding the heterologous polypeptide with KARI activity of a); and
    b. at least one host cell modification that enhances performance of the engineered isobutanol production pathway.

2. The recombinant host cell of claim 1 wherein the combination of a) and b) results in a synergistic increase in isobutanol production pathway performance.

3. The recombinant host cell of claim 1 or claim 2 wherein the host cell modification comprises reduced or eliminated aldehyde dehydrogenase expression or activity and reduced or eliminated acetolactate reductase expression or activity.

4. The recombinant host cell of claim 3 comprising reduced or eliminated aldehyde dehydrogenase expression or activity and reduced or eliminated acetolactate reductase expression or activity and wherein the heterologous polypeptide with KARI activity has a K_M for NADH less than 300 μM.

5. The recombinant host cell of claim 1 wherein the heterologous polypeptide with KARI activity has at least about 90% or at least about 95% identity to SEQ ID NO: 27, 29, 141, 143, 275, or 277.

6. The recombinant host cell of any one of claims 1-5, wherein the heterologous polypeptide with KARI activity comprises substitutions in amino acids corresponding to S56 and S58 of SEQ ID NO: 27.

7. The recombinant host cell of claim 6, wherein the heterologous polypeptide with KARI activity further comprises a substitution of one or more of the amino acids corresponding to I86, N87, T131, or T191 of SEQ ID NO: 27.

8. The recombinant host cell of any one of claims 1-7, wherein the host cell has an effective isobutanol productivity of at least about 3, at least about 4, or at least about 5 grams per gram of cells after about 48 hours, wherein at least the last about 24 of the 48 hours are under anaerobic conditions.

9. The recombinant host cell of any one of claims 1-8, wherein the heterologous polypeptide with KARI activity has a K_M for NADH less than about 350, less than about 100, less than about 50, or less than about 10 μM at pH 6.8.

10. The recombinant host cell of claim 1, wherein the heterologous polypeptide with KARI activity comprises an amino acid substitution

at one or more of the positions corresponding to amino acids A41, S56, S58, I87, T131, T191, R227, or Q246 of SEQ ID NO: 27.

11. The recombinant host cell of any one of claims 1-10, wherein said host cell modification comprises at least one of a deletion, mutation, or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity.

12. The recombinant host cell of claim 11, wherein said polypeptide having aldehyde dehydrogenase activity corresponds to Enzyme Commission Number EC 1.21.3, EC 1.2.1.4, EC1.2.1.5, or a combination thereof.

13. The recombinant host cell of any one of claims 11-12, wherein said host cell is *S. cerevisiae* and said polypeptide having aldehyde dehydrogenase activity is ALD2, ALD3, ALD4, ALD5, ALD6 or a homolog thereof.

14. The recombinant host cell of any one of claims 11-12, wherein said host cell is *K. lactis* and said polypeptide having aldehyde dehydrogenase activity is KLLA0F00440, KLLA0E23057, KLLA0D10021, or KLLA0D09999G.

15. The recombinant host cell of any one of claims 11-12, wherein said host cell is *P. stipitis* and said polypeptide having aldehyde dehydrogenase activity is ALD2, ALD3, ALD4, ALD5, or ALD7.

16. The recombinant host cell of any one of claims 11-12, wherein said host cell is *Lactobacillus plantarum* and said polypeptide having aldehyde dehydrogenase activity is AldH.

17. The recombinant host cell of any one of claims 11-12, wherein said host cell is *E. coli* and said polypeptide having aldehyde dehydrogenase activity is aldA, aldB, or aldH.

18. The recombinant host cell of any one of claims 1-17, wherein said host cell comprises at least one or more of a deletion, mutation, or substitution in an endogenous polynucleotide encoding a polypeptide having aldehyde oxidase activity.

19. The recombinant host cell of claim 18, wherein said polypeptide having aldehyde oxidase activity corresponds to Enzyme Commission Number EC 1.2.3.1

20. The recombinant host cell of claim 19, wherein said polypeptide having aldehyde oxidase activity is AOX1,AOX2, or a combination thereof.

21. The recombinant host cell of any one of claims 1-20, wherein said host cell comprises at least one or more deletion, mutation, or substitution in an endogenous polynucleotide encoding a polypeptide having acetolactate reductase activity.

22. The recombinant host cell of claim 21, wherein the endogenous polynucleotide is selected from the group consisting of SEQ ID NO: 676, SEQ ID NO: 678, SEQ ID NO: 680, SEQ ID NO: 682, SEQ ID NO: 684, SEQ ID NO: 686, SEQ ID NO: 688, SEQ ID NO: 690, SEQ ID NO: 692, SEQ ID NO: 694, SEQ ID NO: 696, SEQ ID NO:702, SEQ ID NO: 704, SEQ ID NO: 706, SEQ ID NO: 708, SEQ ID NO: 710, SEQ ID NO: 712, SEQ ID NO: 714, SEQ ID NO: 716, SEQ ID NO: 718, SEQ ID NO: 720, SEQ ID NO: 722, SEQ ID NO: 724, SEQ ID NO:726 , SEQ ID NO:728, and SEQ ID NO: 730

23. The recombinant host cell of claim 21, wherein the polypeptide having acetolactate reductase activity is YMR226C.

24. The recombinant host cell of any one of claims 21-23, wherein the host cell is a yeast host cell.

2017200439 20 Jan 2017

25. The recombinant host cell of claim 24, wherein the yeast is *Saccharomyces*, *Schizosaccharomyces, Hansenula*, *Candida*, *Kluyveromyces*, *Yarrowia, Issatchenkia,* or *Pichia.*

26. The recombinant host cell of claim 1 or 25, wherein the host cell is *Saccharomyces cerevisiae.*

27. The recombinant host cell of any one of claims 1-26, wherein the engineered isobutanol production pathway comprises the following substrate to product conversions:

a. pyruvate to acetolactate
b. acetolactate to 2,3-dihydroxyisovalerate
c. 2,3-dihydroxyisovalerate to 2-ketoisovalerate
d. 2-ketoisovalerate to isobutyraldehyde; and
e. isobutyraldehyde to isobutanol

wherein each substrate to product conversions is catalyzed by an enzyme that is recombinantly expressed by the host cell.

28. The recombinant host cell of claim 27, wherein all of the substrate to product conversions are catalyzed by enzymes heterologous to the host cell.

29. The recombinant host cell of claim 28, wherein at least one heterologous polynucleotide encoding at least one of the enzymes heterologous to the host cell is chromosomally integrated into the host cell.

30. The recombinant host cell of any one of claims 27-29, wherein the substrate to product conversions are catalyzed by enzymes substantially localized to the cytosol.

2017200439 20 Jan 2017

31. The recombinant host cell of any one of claims 27-30, wherein the substrate to product conversion for isobutyraldehyde to isobutanol is catalyzed by an alcohol dehydrogenase enzyme which utilizes NADH as a cofactor.

32. The recombinant host cell of any one of claims 1-31, wherein said host cell further comprises reduced, disrupted or eliminated ability to convert acetolactate to 2,3-dihydroxy-2-methylbutyrate.

33. The recombinant host cell of any one of claims 1-32, wherein the host cell is yeast and has reduced, disrupted or eliminated pyruvate decarboxylase expression or activity.

34. The recombinant host cell of claim 33, wherein the host cell has reduced, disrupted, or eliminated PDC1, PDC5, or PDC6 activity or a combination thereof.

35. The recombinant host cell of any one of claims 1-34, wherein the host cell has reduced or eliminated NAD-dependent glycerol-3-phosphate dehydrogenase expression or activity.

36. The recombinant host cell of claim 35, wherein the host cell has reduced GPD2 expression or activity.

37. The recombinant host cell of any one of claims 1-36, wherein the host cell has reduced or eliminated FRA2 expression or activity.

38. The recombinant host cell of any one claims 1-37, wherein the host cell produces isobutanol under anaerobic conditions and wherein the molar ratio of isobutanol to glycerol is greater than 1.

2017200439 20 Jan 2017

39. The recombinant host cell of any one of claims 1-38, wherein the polypeptide having KARI activity matches the profile HMM provided in Table Z with a profile HMM E value of $<10^{-3}$.

40. The recombinant host cell of any one of claims 1-39, wherein said host cell produces isobutanol at a yield greater than about 25%, about 50%, about 75%, or about 90% of theoretical yield.

41. The recombinant host cell of any one of claims 1-40, wherein isobutanol is produced and wherein ethanol is also produced.

42. A method for producing isobutanol comprising:
    a. providing the recombinant host cell of any one of claims 1-41, and
    b. contacting the host cell of a) with a carbon substrate under conditions whereby isobutanol is produced.

43. The method of claim 42, wherein at least a portion of the contacting occurs under anaerobic conditions.

44. The method of claim 42 or 43, wherein the contacting occurs in the presence of an extractant.

45. The method of claim 44, wherein the contacting occurs in the presence of a sufficient quantity of organic extractant to form a two-phase system comprising an aqueous phase and an organic phase.

46. The method of any one of claims 42-45, wherein one or more of the effective rate, effective titer, or effective yield of isobutanol is increased as compared to a recombinant host cell that does not comprise (i) at least one of a heterologous polypeptide with KARI activity or a heterologous polynucleotide encoding a polypeptide

2017200439 20 Jan 2017

with KARI activity and (ii) at least one modification that enchances performace of the engineered isobutanol production pathway.

47. The method of any one of claims 42-46, wherein DHMB production, isobutyric acid production, or both is reduced as compared to a recombinant host cell that does not comprise (i) at least one of a heterologous polypeptide with KARI activity ora heterologous polynucleotide encoding a polypeptide with KARI activity and (ii) at least one modification that enhances performance of the engineered isobutanol production pathway.

48. The method any one of claims 42-47 wherein the molar ratio of isobutanol to glycerol is greater than 1.

49. A fermentative composition comprising the host cell and isobutanol produced according to the method of any one of claims 42-28.

50. A method for producing isobutanol comprising:
    a. providing a recombinant host cell which produces isobutanol
    b. contacting the host cell of a) with a carbon substrate under conditions whereby isobutanol is produced;
    wherein at least a portion of the contacting occurs under anaerobic conditions; and
    wherein the ratio of isobutanol to glycerol produced is greater than 1.

51. A method for the production of butanol comprising:
    a. growing a recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate under conditions whereby butanol is produced; and
    b. removing DHMB from the culture.

52. A composition produced by the method of claim 51, wherein the composition comprises butanol and no more than about 0.5 mM DHMB.

53. A composition comprising i) a recombinant yeast capable of producing butanol, ii) butanol, and iii) no more than about 0.5 mM DHMB.

54. A method for producing a recombinant host cell comprising:
    a. providing a recombinant host cell comprising a modification in a polynucleotide encoding a polypeptide having aldehyde dehydrogenase activity or aldehyde oxidase activity; and
    b. transforming the host cell with a polynucleotide encoding a polypeptide of an isobutanol biosynthetic pathway.

55. A method for reducing or eliminating the conversion of isbutyraldehye to isobutyric acid comprising:
    a. providing the recombinant host cell of any one of claims 1-41; and
    b. subjecting the host cell to conditions wherein the conversion of isbutyraldehye to isobutyric acid is reduced or eliminated compared to a recombinant host cell that does not comprise reduced or eliminated aldehyde dehydrogenase and/or aldehyde oxidase activity.

56. A composition comprising isobutanol and a recombinant host cell of any one of claims 1-41.

57. A polypeptide comprising at least about 90% identity or at least about 95% identity or at least about 99% identity to SEQ ID NO: 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 417, 419, 421, 423, 425, 427, 429, 431, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450,

451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 624, 626, 628, 630, 632 or an active fragment thereof wherein said polypeptide has ketol-acid reductoisomerase activity.

58. A polypeptide comprising at least about 90% identity or at least about 95% identity or at least about 99% identity to SEQ ID NO: 417, 419, 421, 423, 425, or 427 or an active fragment thereof wherein said polypeptide has ketol-acid reductoisomerase activity.

59. A polynucleotide encoding the polypeptide of claim 57 or 58.

60. A method of converting acetolactate to 2,3-dihydroxyisovalerate comprising:

2017200439 20 Jan 2017

a. providing a polypeptide of claim 57 or 58;
b. contacting the polypeptide of a) with acetolactate under conditions wherein 2,3-dihydroxyisovalerate is produced.

61. A recombinant host cell comprising a heterologous polypeptide of claim 57, claim 58 or a polynucleotide of claim 59.

62. A recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate, wherein said yeast produces less than 0.01 moles 2,3-dihydroxy-2-methyl butyrate (DHMB) per mole of sugar consumed.

63. A recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate, wherein said yeast produces DHMB at a rate of less than about 1.0 mM/hour.

64. A recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate, wherein said yeast produces an amount of 2,3-dihydroxy-3-isovalerate (DHIV) that is at least about 1.5 times the amount of DHMB produced.

65. A method for the production of butanol comprising:
a) growing a recombinant yeast comprising a biosynthetic pathway capable of converting pyruvate to acetolactate under conditions whereby butanol is produced; and
b) measuring DHIV and/or DHMB concentration;
wherein steps a) and b) can be performed simultaneously or sequentially and in any order.

66. The method of claim 65, wherein the measuring comprises liquid chromatography-mass spectrometry.

67. A method for increasing ketol-acid reductoisomerase (KARI) activity comprising a) providing a composition comprising acetolactate, a KARI enzyme, and an acetolactate reductase enzyme and b) decreasing DHMB levels.

68. The method of claim 67, wherein said decreasing DHMB levels is achieved by decreasing acetolactate reductase enzyme activity.

69. The method of claim 67, wherein said decreasing DHMB levels is achieved by removing DHMB from the composition.

70. A method for increasing dihydroxyacid dehydratase (DHAD) activity comprising a) providing a composition comprising dihydroxyisovalerate (DHIV) and a DHAD enzyme and b) decreasing DHMB levels.

71. A method of increasing KARI enzyme productivity in a host cell, wherein the host cell comprises an engineered isobutanol production pathway that comprises a heterologous KARI enzyme and at least one genetic modification that reduces, disrupts, or eliminates acetolactate reductase expression or activity, and wherein the KARI enzyme activity is decreased in the presence of DHMB.

72. The method of claim 71, wherein the KARI has at least about 90%, at least about 95%, or at least about 99% identity to *E. coli* or *L. lactis* KARI.

2017200439 20 Jan 2017

73. The method of claim 71 or 72, wherein the reduced, disrupted, or eliminated acetolactate reductase expression or activity substantially reduces the presence of DHMB.

74. The recombinant host cell of claim 1 comprising a heterologous polynucleotide encoding a polypeptide having KARI activity comprising SEQ ID NO: 33 or 35 or an active fragment thereof.

2017200439 20 Jan 2017

Figure 1

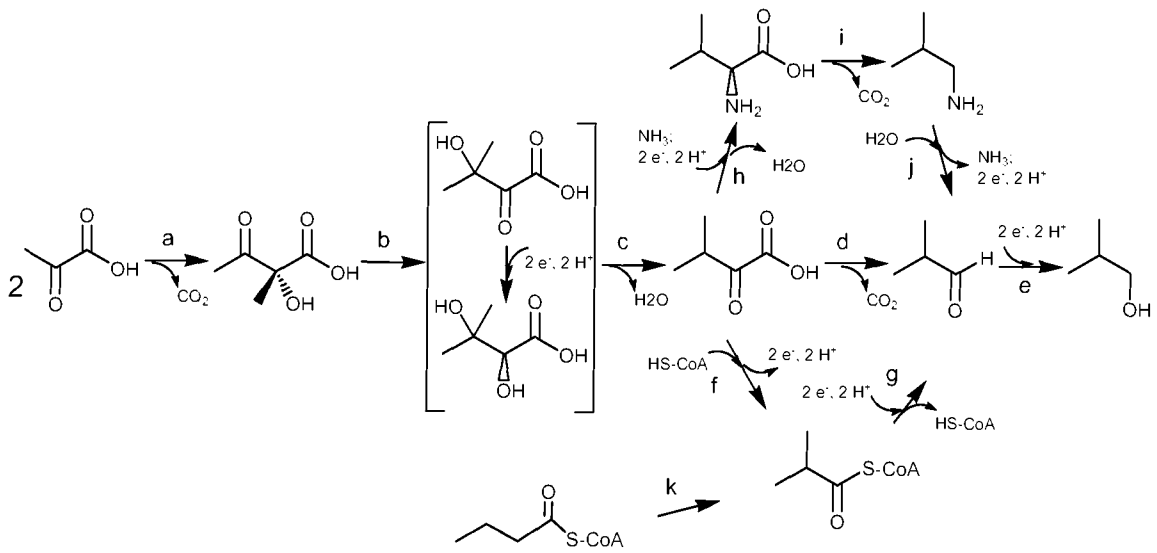


FIGURE 2.

```
K9    (1) MEECKMAKIYYQEDCNLSLLDGKTIAVIGYGSQGHAHALNAKESGCNVII
PF5   (1) ------MKVFYDKDCDLSIIQGKKVAIIGYGSQGHAQACNLKDSGVDVTV

K9   (51) GLYEGSKSWKRAEEQGFEVYTAAEAAKKADIIMILINDEKQATMYKNDIE
PF5  (45) GLRKGSATVAKAEAHGLKVTDVAAAVAGADLVMILTPDEFQSQLYKNEIE

K9  (101) PNLEAGNMLMFAHGFNIHFGCIVPPKDVDVTMIAPKGPGHTVRSEYEEGK
PF5  (95) PNIKKGATLAFSHGFAIHYNQVVPRADLDVIMIAPKAPGHTVRSEFVKGG

K9  (151) GVPCLVAVEQDATGKALDMALAYALAIGGARAGVLETTFRTETETDLFGE
PF5 (145) GIPDLIAIYQDASGNAKNVALSYAAGVGGGRTGIIETTFKDETETDLFGE

K9  (201) QAVLCGGVCALMQAGFETLVEAGYDPRNAYFECIHEMKLIVDLIYQSGFS
PF5 (195) QAVLCGGTVELVKAGFETLVEAGYAPEMAYFECLHELKLIVDLMYEGGIA

K9  (251) GMRYSISNTAEYGDYITGPKIITEDTKKAMKKILSDIQDGTFAKDFLVDM
PF5 (245) NMNYSISNNAEYGEYVTGPEVINAESRQAMRNALKRIQDGEYAKMFISEG

K9  (301) SDAGSQVHFKAMRKLASEHPAEVVGEEIRSLYSWSDEDKLINN---
PF5 (295) --ATGYPSMTAKRRNNAAHGIEIIGEQLRSMMPWIGANKIVDKAKN
```

FIGURE 3A.

```
          K9     (1) MEECKMAKIYYQEDCNLSLLDGKTIAVIGYGSQGHAHALNAKESG-CNVI
          K1     (1) ----MAATIWYEKDADLSVFEGKKVAVIGYGSQGHAHALNLRDSG-VDVV
          K2     (1) ----MAAQIWYENDGDLSVLDGKKVAIIGYGSQGHAHALNLRDSG-VDVV
          K7     (1) -----MPKMYYEKDTDLNLLRGKKVAIIGYGSQGHAHALNLHESG-VDVV
          S2     (1) -QEEELMKVYYDSDADLGLIKSKRIAILGYGSQGHAHAQNLRDSGVAEVA
         K25     (1) -----MVKVYYNEVTDN--LGDKTIAIIGYGSQGHAHAQNLRDTG-HNVV
         K26     (1) ----MAVQMEYEKDVKVPALDGKKIAVIGYGSQGHAHSQNLRDTG-HDVI
   L. Lactis     (1) ----MAVTMYYEDDVEVSALAGKQIAVIGYGSQGHAHAQNLRDSG-HNVI

          K9    (50) IGLYEGSKSWKRAEEQGFEVYTAAEAAKKADIIMILINDEKQATMYKNDI
          K1    (46) VGLRPTSKSVEYAKEQGLEVQSVADATAEADVVMILLPDQYQAAVYKSEI
          K2    (46) VGLRPTSKSVEQAKEQGLEVKSVPEAAAEADIIMILAPDQYQRTIWANDI
          K7    (45) VGLYNGSKSWAKAEAAGLQVATVADAAKAADLIMILLPDEKQAKIYNEEI
          S2    (50) IALRPDSASVKKAQDAGFKVLTNAEAAKWADILMILAPDEHQAAIYAEDL
         K25    (43) IGIRAG-KSAENAKNDGFTVLPVSEAVQQANVVMILAPDEIQGELYEKEI
         K26    (46) IGVRPG-KSFDKAKEDGFDTYTVAEATKLADVIMILAPDEIQQELYEAEI
   L. Lactis    (46) IGVRHG-KSFDKAKEDGFETFEVGEAVAKADVIMVLAPDELQQSIYEEDI

          K9   (100) EPNLEAGNMLMFAHGFNIHFGCIVPPKDVDVTMIAPKGPGHTVRSEYEEG
          K1    (96) EPNLKPGAALAFAHGFNIHYGYIKPSEDHPVFMVAPKGPGHIVRREYTNG
          K2    (96) EPNIKPGAAIAFAHGFNIHYGYIKPSEDHPVFMVAPKGPGHIVRREYANG
          K7    (95) APNLEEGNALVFAHGFNIHFGQVVPPSYVDVFMVAPKGPGHLVRRTYTEG
          S2   (100) KDNLRPGSAIAFAHGLNIHFGLIEPRKDIDVFMIAPKGPGHTVRSEYVRG
         K25    (92) APNLSAGDALAFAHGFNIHFEVITPPKEVDVFLVAPKGPGHLVRRTFSEG
         K26    (95) APNLEAGNAVGFAHGFNIHFEFIKVPADVDVFMCAPKGPGHLVRRTFEEG
   L. Lactis    (95) KPNLKAGSALGFAHGFNIHFGYIKVPEDVDVFMVAPKAPGHLVRRTYTEG

          K9   (150) KGVPCLVAVEQDATGKALDMALAYALAIGGARAGVLETTFRTETETDLFG
          K1   (146) RGVPVVVAVEQDPDGKTWPLCLAYAKALGALRAGAIKTTFTEETETDLFG
          K2   (146) RGVPVVVAVEQDPRGDAWDITLAYAKALGALRAGAIKTTFKEETETDLFG
          K7   (145) AGVPCLIAVHQDATGKAKQYALAYANGIGGARAGVLETTFKDETETDLFG
```

FIGURE 3B.

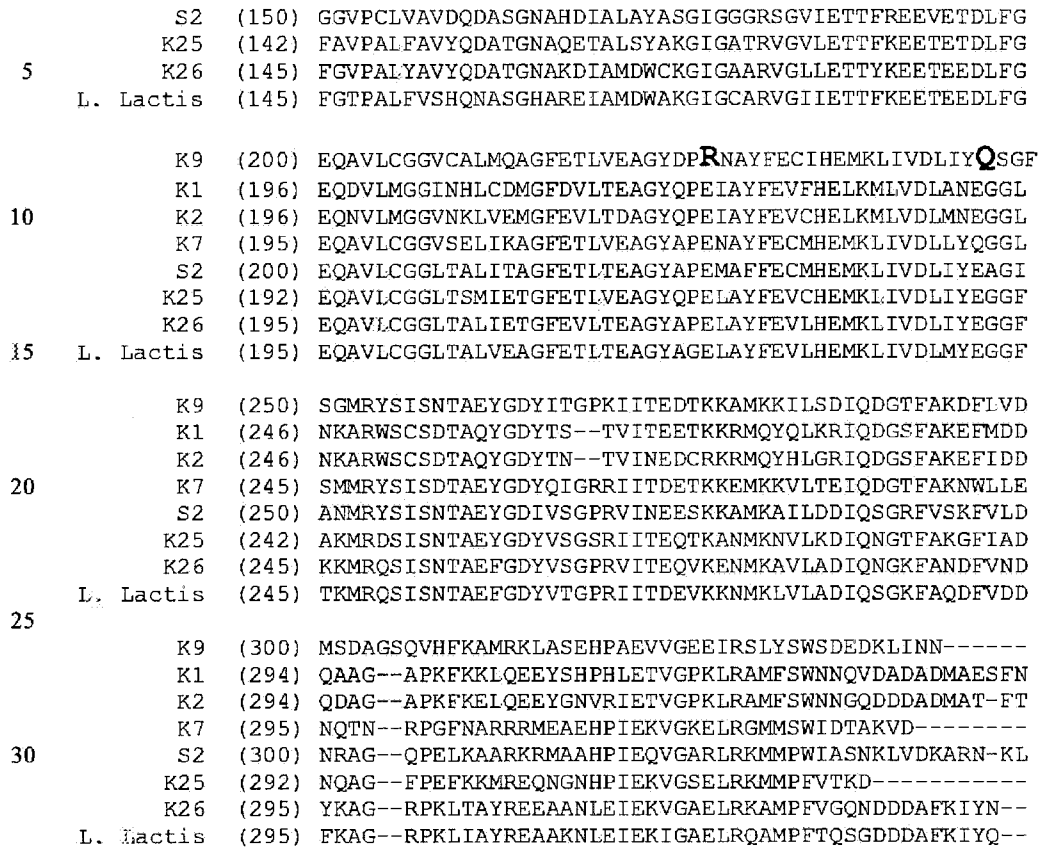
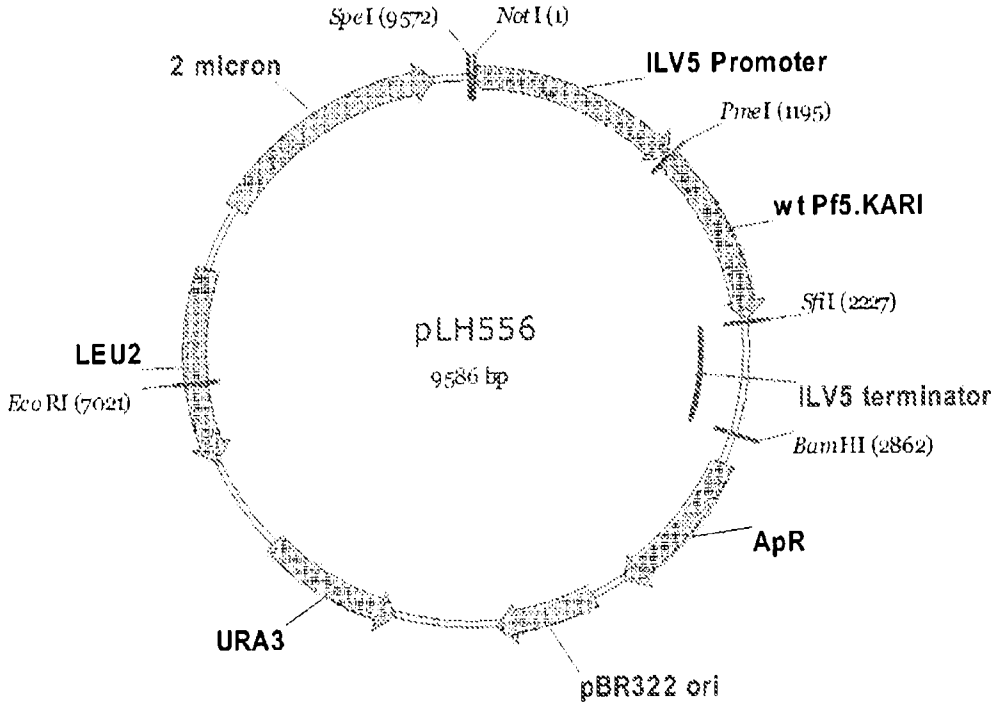

S2 (150) GGVPCLVAVDQDASGNAHDIALAYASGIGGGRSGVIETTFREEVETDLFG
K25 (142) FAVPALFAVYQDATGNAQETALSYAKGIGATRVGVLETTFKEETETDLFG
K26 (145) FGVPALYAVYQDATGNAKDIAMDWCKGIGAARVGLLETTYKEETEEDLFG
L. Lactis (145) FGTPALFVSHQNASGHAREIAMDWAKGIGCARVGIIETTFKEETEEDLFG
K9 (200) EQAVLCGGVCALMQAGFETLVEAGYDPRNAYFECIHEMKLIVDLIYQSGF
K1 (196) EQDVLMGGINHLCDMGFDVLTEAGYQPEIAYFEVFHELKMLVDLANEGGL
K2 (196) EQNVLMGGVNKLVEMGFEVLTDAGYQPEIAYFEVCHELKMLVDLMNEGGL
K7 (195) EQAVLCGGVSELIKAGFETLVEAGYAPENAYFECMHEMKLIVDLLYQGGL
S2 (200) EQAVLCGGLTALITAGFETLTEAGYAPEMAFFECMHEMKLIVDLIYEAGI
K25 (192) EQAVLCGGLTSMIETGFETLVEAGYQPELAYFEVCHEMKLIVDLIYEGGF
K26 (195) EQAVLCGGLTALIETGFEVLTEAGYAPELAYFEVLHEMKLIVDLIYEGGF
L. Lactis (195) EQAVLCGGLTALVEAGFETLTEAGYAGELAYFEVLHEMKLIVDLMYEGGF
K9 (250) SGMRYSISNTAEYGDYITGPKIITEDTKKAMKKILSDIQDGTFAKDFLVD
K1 (246) NKARWSCSDTAQYGDYTS--TVITEETKKRMQYQLKRIQDGSFAKEFMDD
K2 (246) NKARWSCSDTAQYGDYTN--TVINEDCRKRMQYHLGRIQDGSFAKEFIDD
K7 (245) SMMRYSISDTAEYGDYQIGRRIITDETKKEMKKVLTEIQDGTFAKNWLLE
S2 (250) ANMRYSISNTAEYGDIVSGPRVINEESKKAMKAILDDIQSGRFVSKFVLD
K25 (242) AKMRDSISNTAEYGDYVSGSRIITEQTKANMKNVLKDIQNGTFAKGFIAD
K26 (245) KKMRQSISNTAEFGDYVSGPRVITEQVKENMKAVLADIQNGKFANDFVND
L. Lactis (245) TKMRQSISNTAEFGDYVTGPRIITDEVKKNMKLVLADIQSGKFAQDFVDD
K9 (300) MSDAGSQVHFKAMRKLASEHPAEVVGEEIRSLYSWSDEDKLINN------
K1 (294) QAAG--APKFKKLQEEYSHPHLETVGPKLRAMFSWNNQVDADADMAESFN
K2 (294) QDAG--APKFKELQEEYGNVRIETVGPKLRAMFSWNNGQDDDADMAT-FT
K7 (295) NQTN--RPGFNARRRMEAEHPIEKVGKELRGMMSWIDTAKVD--------
S2 (300) NRAG--QPELKAARKRMAAHPIEQVGARLRKMMPWIASNKLVDKARN-KL
K25 (292) NQAG--FPEFKKMREQNGNHPIEKVGSELRKMMPFVTKD----------
K26 (295) YKAG--RPKLTAYREEAANLEIEKVGAELRKAMPFVGQNDDDAFKIYN--
L. Lactis (295) FKAG--RPKLIAYREAAKNLEIEKIGAELRQAMPFTQSGDDDAFKIYQ--

FIGURE 3C.

```
          K9  (344) ---------
          K1  (342) GKIARTQVQ
          K2  (341) GKIARSQVQ
          K7  (335) ---------
          S2  (347) GCF------
         K25  (329) ---------
         K26  (341) ---------
L. Lactis     (341) ---------
```

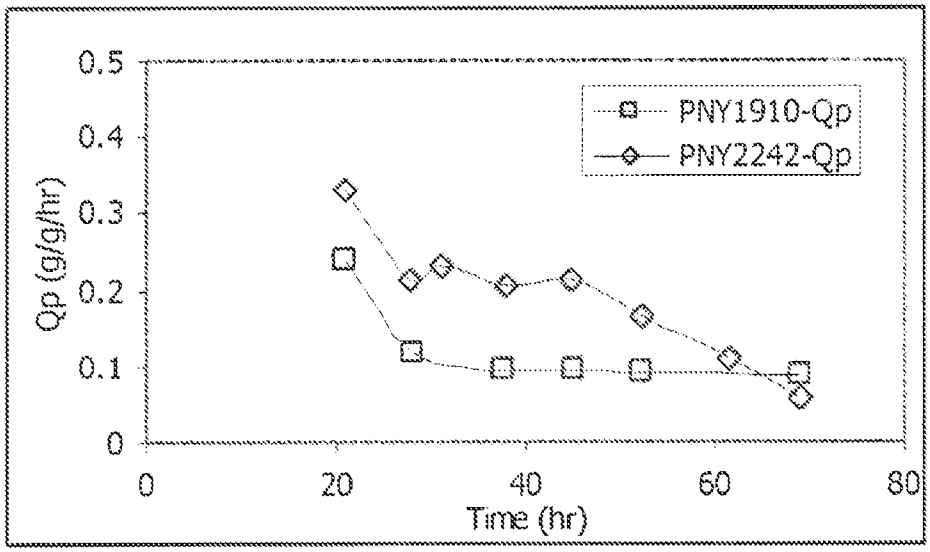
20 Jan 2017 2017200439

FIGURE4. (pLH556)



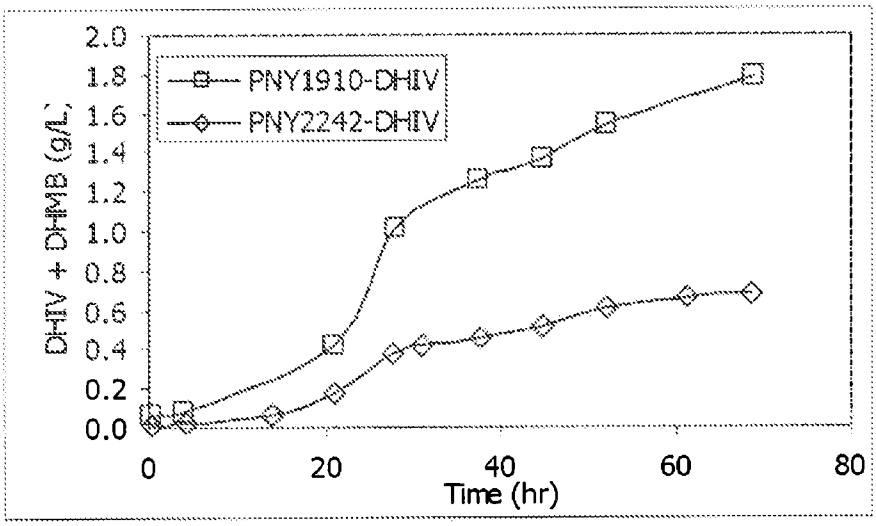
20 Jan 2017 2017200439

FIGURE 5



20 Jan 2017 2017200439

FIGURE 6.



20 Jan 2017
2017200439

FIGURE 7

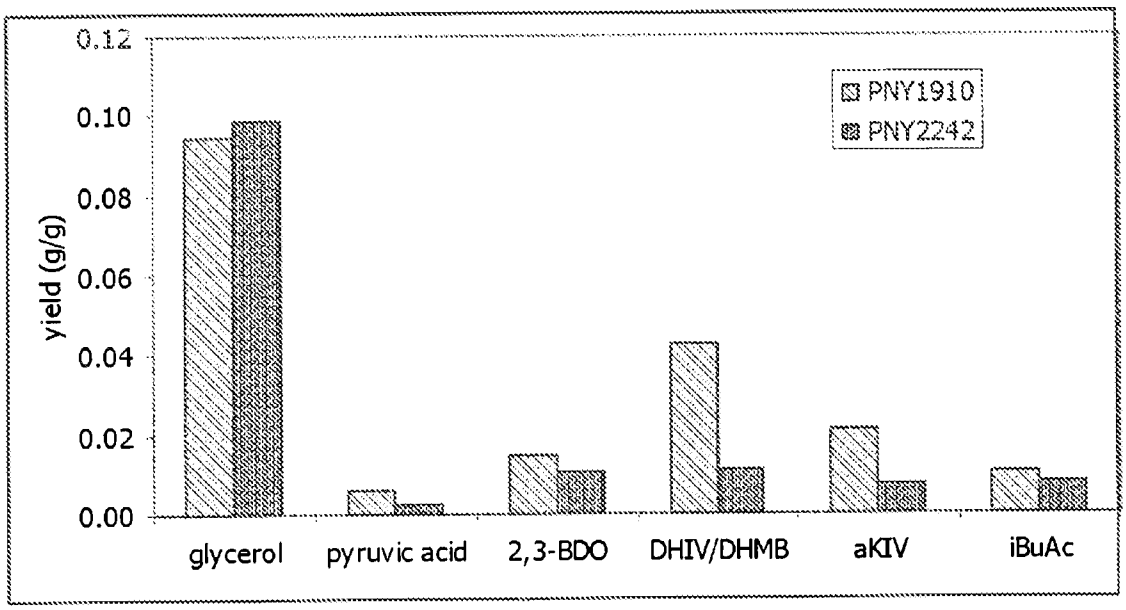


FIGURE 8

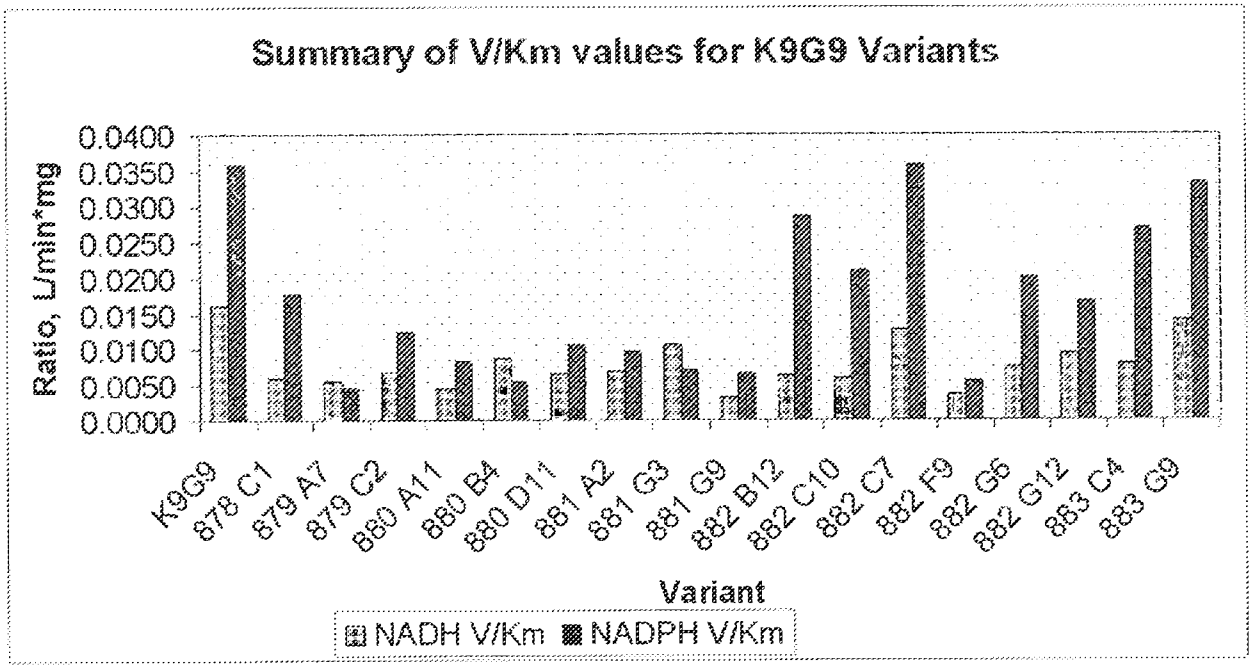

Summary of V/Km values for K9G9 Variants
Ratio, L/min*mg
0.0400
0.0350
0.0300
0.0250
0.0200
0.0150
0.0100
0.0050
0.0000
K9G9
878 C1
879 A7
879 C2
880 A11
880 B4
880 D11
881 A2
881 G3
881 G9
882 B12
882 C10
882 C7
882 F9
882 G6
882 G12
883 C4
883 G9
Variant
NADH V/Km
NADPH V/Km

FIGURE 9A

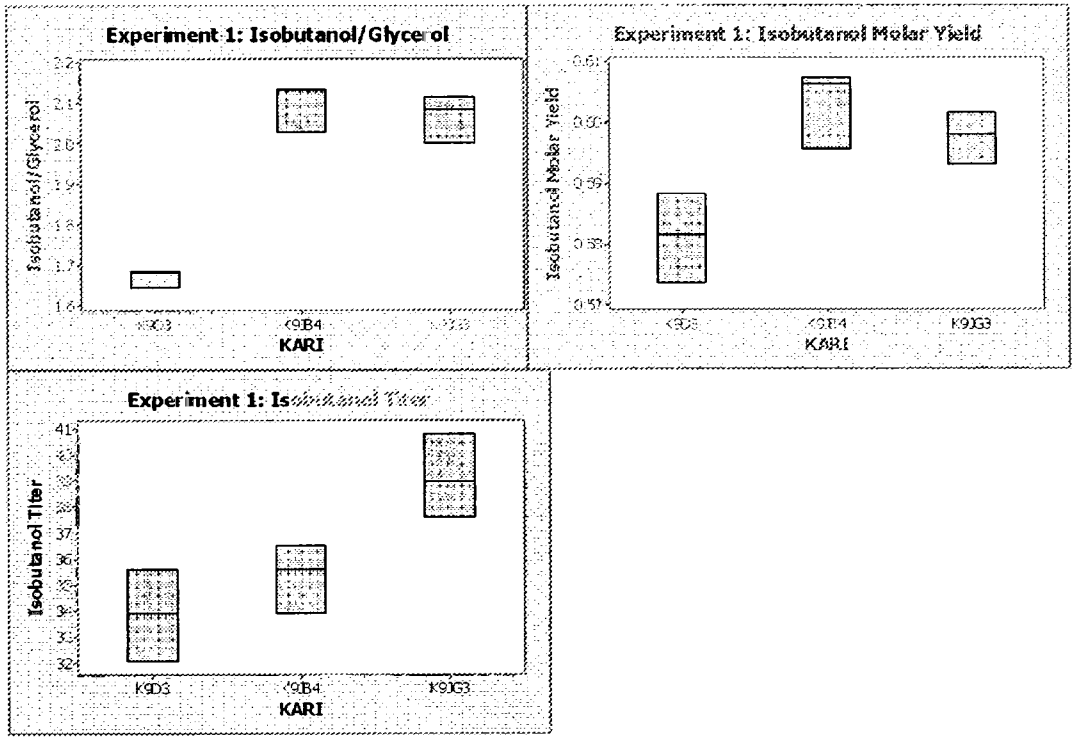

Experiment 1: Isobutanol/Glycerol
Isobutanol/Glycerol
K9D3
K9B4
K9G3
KARI
Experiment 1: Isobutanol Molar Yield
Isobutanol Molar Yield
K9D3
K9B4
K9G3
KARI
Experiment 1: Isobutanol Titer
Isobutanol Titer
41
37
36
32
K9D3
K9B4
K9G3
KARI

FIGURE 9B

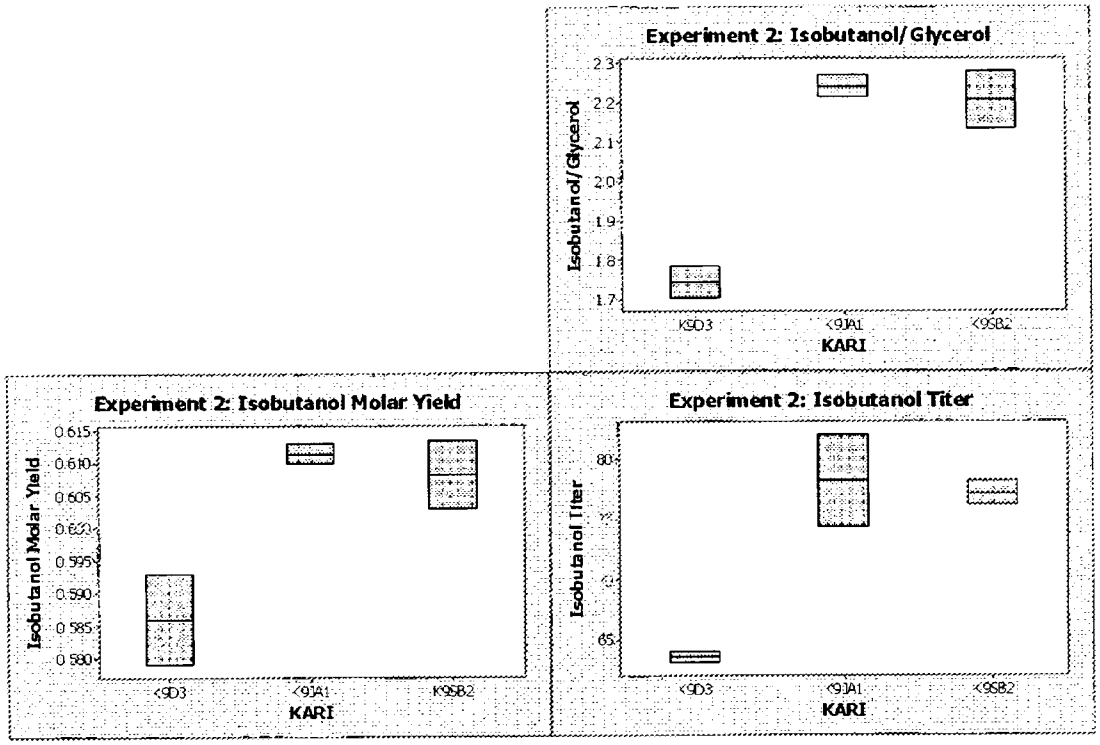
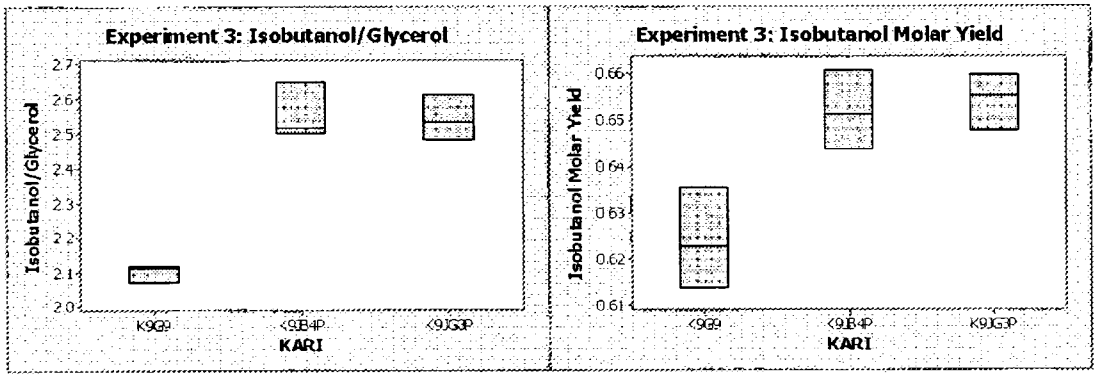
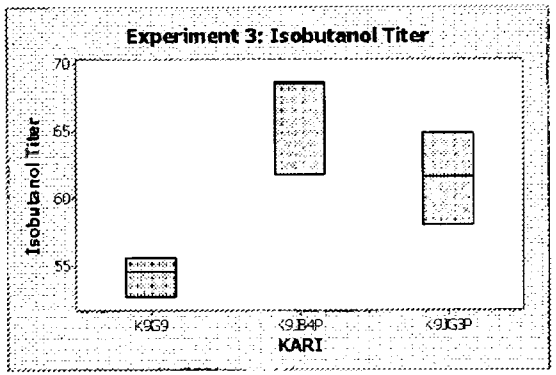

Experiment 2: Isobutanol/Glycerol
Isobutanol/Glycerol
2.3
2.2
2.1
2.0
1.9
1.8
1.7
K9D3
K9A1
K9B2
KARI
Experiment 2: Isobutanol Molar Yield
Isobutanol Molar Yield
0.615
0.610
0.605
0.600
0.595
0.590
0.585
0.580
K9D3
K9A1
K9B2
KARI
Experiment 2: Isobutanol Titer
Isobutanol Titer
80
65
K9D3
K9A1
K9B2
KARI

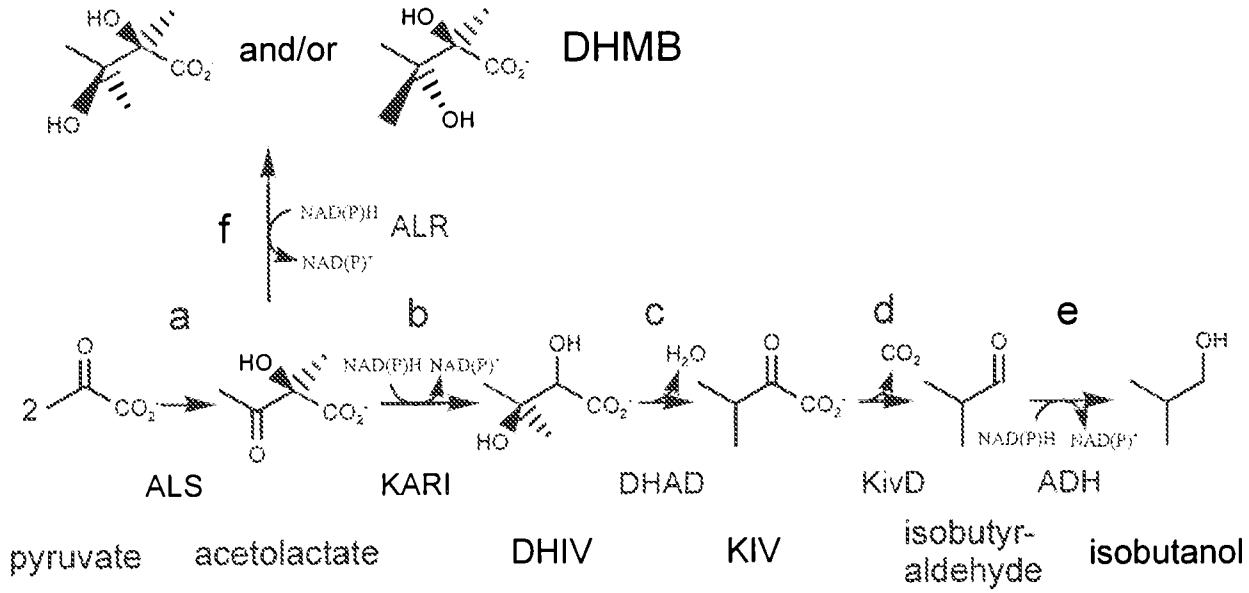
FIGURE 9C


Experiment 3: Isobutanol/Glycerol
Isobutanol/Glycerol
2.7
2.6
2.5
2.4
2.3
2.2
2.1
2.0
K9G9
<9JB4P
<9JG3P
KARI
Experiment 3: Isobutanol Molar Yield
Isobutanol Molar Yield
0.66
0.65
0.64
0.63
0.62
0.61
<9G9
<9JB4P
K9JG3P
KARI


Experiment 3: Isobutanol Titer
Isobutanol Titer
70
65
60
55
K9G9
<9JB4P
<9JG3P
KARI

2017200439 20 Jan 2017

Figure 10



20 Jan 2017 2017200439

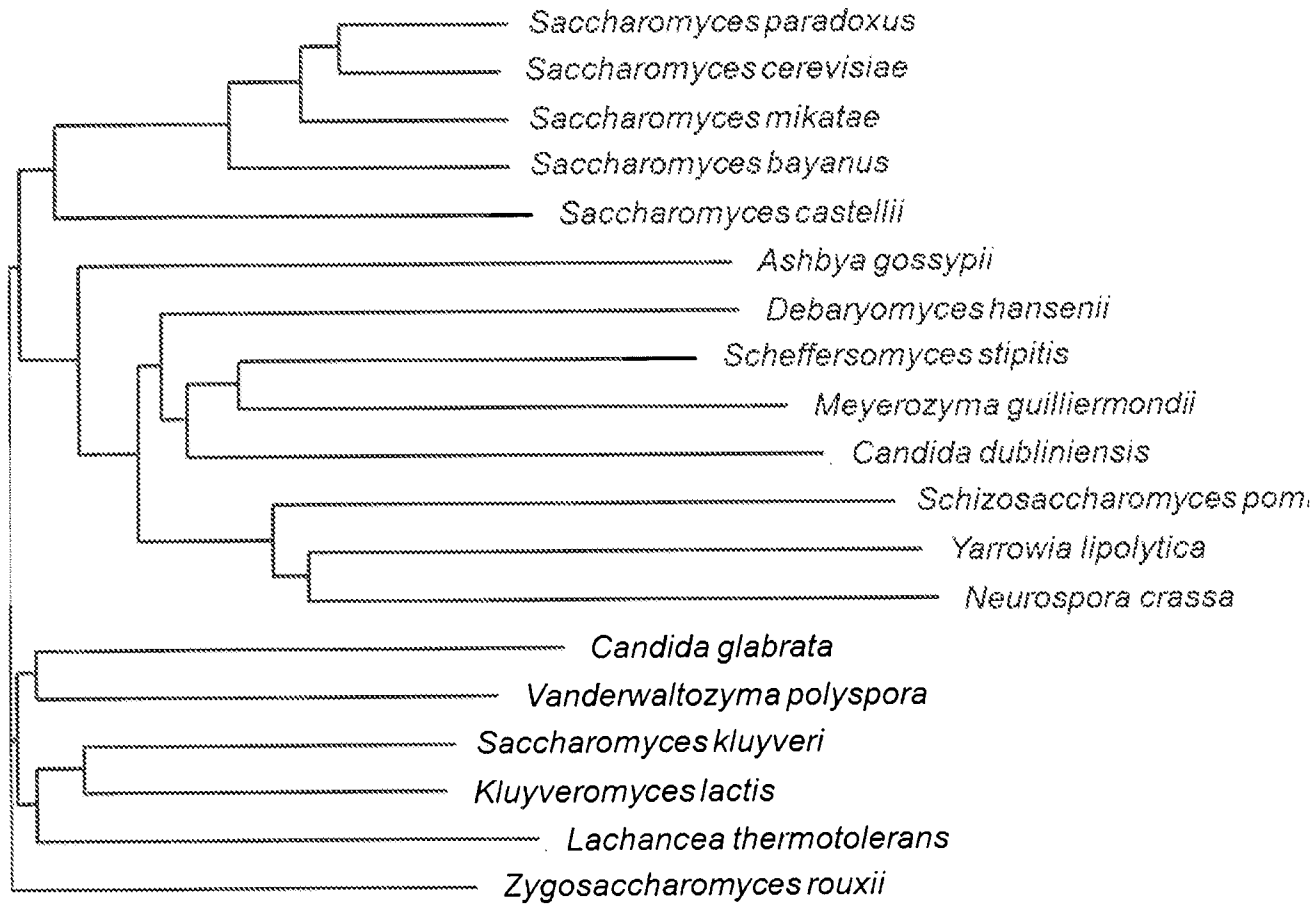
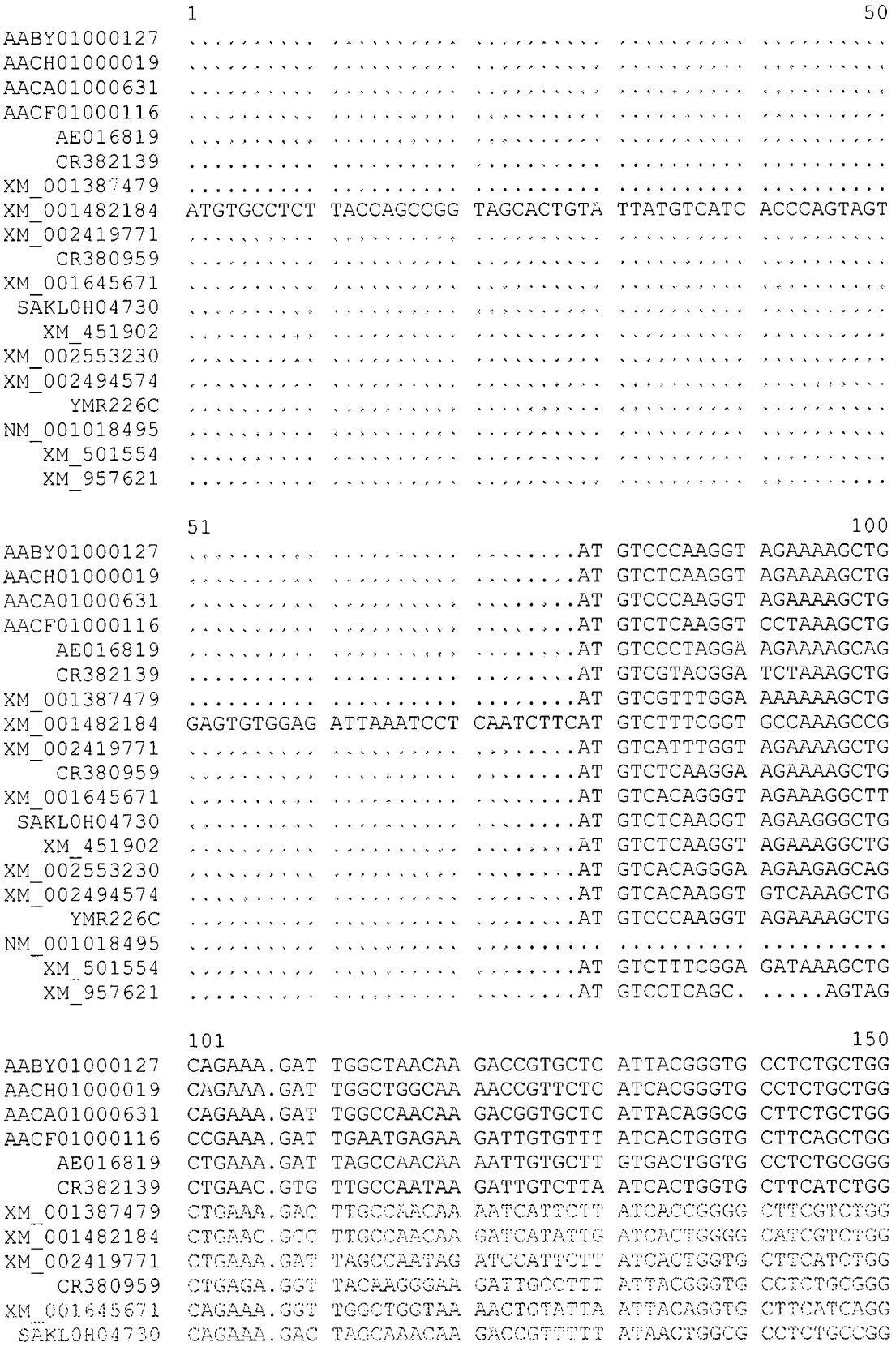


Figure 11

Figure 12


1 50
AABY01000127 ....
AACH01000019 ....
AACA01000631 ....
AACF01000116 ....
AE016819 ....
CR382139 ....
XM_001387479 ....
XM_001482184 ATGTGCCTCT TACCAGCCGG TAGCACTGTA TTATGTCATC ACCCAGTAGT
XM_002419771 ....
CR380959 ....
XM_001645671 ....
SAKL0H04730 ....
XM_451902 ....
XM_002553230 ....
XM_002494574 ....
YMR226C ....
NM_001018495 ....
XM_501554 ....
XM_957621 ....
51 100
AABY01000127 ....AT GTCCCAAGGT AGAAAAGCTG
AACH01000019 ....AT GTCTCAAGGT AGAAAAGCTG
AACA01000631 ....AT GTCCCAAGGT AGAAAAGCTG
AACF01000116 ....AT GTCTCAAGGT CCTAAAGCTG
AE016819 ....AT GTCCCTAGGA AGAAAAGCAG
CR382139 ....AT GTCGTACGGA TCTAAAGCTG
XM_001387479 ....AT GTCGTTTGGA AAAAAAGCTG
XM_001482184 GAGTGTGGAG ATTAAATCCT CAATCTTCAT GTCTTTCGGT GCCAAAGCCG
XM_002419771 ....AT GTCATTTGGT AGAAAAGCTG
CR380959 ....AT GTCTCAAGGA AGAAAAGCTG
XM_001645671 ....AT GTCACAGGGT AGAAAGGCTT
SAKL0H04730 ....AT GTCTCAAGGT AGAAGGGCTG
XM_451902 ....AT GTCTCAAGGT AGAAAGGCTG
XM_002553230 ....AT GTCACAGGGA AGAAGAGCAG
XM_002494574 ....AT GTCACAAGGT GTCAAAGCTG
YMR226C ....AT GTCCCAAGGT AGAAAAGCTG
NM_001018495 ....
XM_501554 ....AT GTCTTTCGGA GATAAAGCTG
XM_957621 ....AT GTCCTCAGC. .....AGTAG
101 150
AABY01000127 CAGAAA.GAT TGGCTAACAA GACCGTGCTC ATTACGGGTG CCTCTGCTGG
AACH01000019 CAGAAA.GAT TGGCTGGCAA AACCGTTCTC ATCACGGGTG CCTCTGCTGG
AACA01000631 CAGAAA.GAT TGGCCAACAA GACGGTGCTC ATTACAGGCG CTTCTGCTGG
AACF01000116 CCGAAA.GAT TGAATGAGAA GATTGTGTTT ATCACTGGTG CTTCAGCTGG
AE016819 CTGAAA.GAT TAGCCAACAA AATTGTGCTT GTGACTGGTG CCTCTGCGGG
CR382139 CTGAAC.GTG TTGCCAATAA GATTGTCTTA ATCACTGGTG CTTCATCTGG
XM_001387479 CTGAAA.GAC TTGCCAACAA AATCATTCTT ATCACCGGGG CTTCGTCTGG
XM_001482184 CTGAAC.GCC TTGCCAACAA GATCATATTG ATCACTGGGG CATCGTCTGG
XM_002419771 CTGAAA.GAT TAGCCAATAG ATCCATTCTT ATCACTGGTG CTTCATCTGG
CR380959 CTGAGA.GGT TACAAGGGAA GATTGCCTTT ATTACGGGTG CCTCTGCGGG
XM_001645671 CAGAAA.GGT TGGCTGGTAA AACTGTATTA ATTACAGGTG CTTCATCAGG
SAKL0H04730 CAGAAA.GAC TAGCAAACAA GACCGTTTTT ATAACTGGCG CCTCTGCCGG

20 Jan 2017
2017200439

```
   XM_451902  CTGAAA.GAT TGCAAAACAA GACAATTTTC ATTACCGGTG CTTCTGCAGG
XM_002553230  CTGAAA.GAC TGGCAGGAAA GACTGTCTTC ATCACAGGCG CATCAGCCGG
XM_002494574  CTGAAA.GAC TAGCTGGTAA GACTGTATTC ATTACAGGTG CTTCTGCAGG
     YMR226C  CAGAAA.GAT TGGCTAAGAA GACTGTCCTC ATTACAGGTG CATCTGCTGG
NM_001018495  ATGAGCCGTT TGGATGGAAA AACGATTTTA ATCACTGGTG CCTCTTCTGG
   XM_501554  CTGCTC.GAC TTGCGGGCAA GACCGTCTTT GTTACCGGCG CCTCGTCCGG
   XM_957621  CCAAGC.GCC TCGCGGGCAA GACCATTGTC ATAACCGGAG CGTCCTCAGG

              151                                                200
AABY01000127  TATTGGTAAG GCCACCGCAT TAGAGTATTT G...GAGGCA TCCAATGGTG
AACH01000019  TATTGGCAAA GCCACTGCAT TAGAGTATTT G...GAGGCA TCCAATGGCG
AACA01000631  TATTGGTAAG GCCACCGCAT TGGAGTATTT G...GAAGCA TCCAATGGAA
AACF01000116  TATTGGGCAA GCCACCGCTT TGGAATACAT G...GATGCG TCGAACGGTA
    AE016819  CATTGGCCGT GCTACAGCCA TTAACTATGC A...GACGCG ACGGACGGGG
    CR382139  AATTGGTGAA GCAACTGCCA AAGAAATTGC A...TCAGCC GCTAATGGCA
XM_001387479  TATTGGTGAA GCTACAGCTA GAGAGTTTGC A...TCTGCT GCCAATGGGA
XM_001482184  TATAGGCGAG GCTACCGCCA GAGAATTCGC T...GCTGCT GCCAATGGAA
XM_002419771  GATTGGTGAA GCATGTGCTA AAGTTTTCGC T...GAAGCA TCTAATGGTC
    CR380959  CATCGGTAAA GCTACAGCCA TTGAGTATTT G...GATGCT TCCAATGGTA
XM_001645671  GATTGGGAAA GCCACTGCAT TAGAATATCT A...GATGCC TCCAATGGTC
 SAKL0H04730  CATTGGCCAA GCTACTGCTT TGGAATACTG T...GATGCT TCTAACGGTA
   XM_451902  TATTGGTCAA GCCACAGCAT TGGAATATCT A...GATGCT GCTAACGGTA
XM_002553230  TATCGGTCAG GCCACTGCGC AAGAATACCT G...GAAGCA TCCGAAGGCA
XM_002494574  TATCGGTCAA GCAACTGCAA AGGAATATTT G...GATGCA TCCAATGGTC
     YMR226C  TATTGGTAAG GCGACCGCAT TAGAGTACTT G...GAGGCA TCCAATGGTG
NM_001018495  AATTGGAAAA AGCACTGCTT TTGAAATTGC C...AAAGTT GCCAAAG...
   XM_501554  CATTGGCCAG GCCACTGTTC TCGCTCTAGC C...GAAGCT GCCAAGGGCG
   XM_957621  AATCGGGCGT AGCACAGCCT TCGAGTTTGC GCGGACCGCG CCCAACCACG

              201                                                250
AABY01000127  ATATGAAACT GGTCTTAGCT GCTAGAAGAT TAGAAAAGCT CGAGGAATTA
AACH01000019  ATATGAAATT AATCTTAGCC GCTAGAAGAT TAGAAAAGCT CGAGGAATTG
AACA01000631  ACATGAAACT GATCTTGGCT GCGAGGAGAT TGGAGAAGCT AGAGGAGCTG
AACF01000116  CTGTGAAATT GGTTCTAGTT GCCAGAAGAT TGGAGAAATT ACAACAATTG
    AE016819  CAATCAAGTT GATTTTGGTG GCAAGACGCG CAGAAAAGCT CACCAGCTTG
    CR382139  ATTTAAAATT AGTGTTGTGT GCTAGACGAA AAGAAAAGTT GGATAATTTA
XM_001387479  ATATCAGATT GATTTTGACA GCCAGAAGAA AAGAAAAGTT GGCTCAATTG
XM_001482184  AAATTCTGTT GATTTTGACC GCTCGGAGAG AAGACAAACT CAAGTCTCTC
XM_002419771  AAGTTAAATT AGTTTTAGGA GCAAGAAGAA AAGAACGATT AGTTAAATTA
    CR380959  GTGTGAAGCT AGTTCTTGGT GCACGTAGAA TGGAGAAATT GGAGGAGTTG
XM_001645671  ATATGAAGTT AATTTTAGTT GCAAGAAGAT TAGAAAAATT GCAAGAGTTG
 SAKL0H04730  AAATAAACTT GGTGTTAAGT GCCAGAAGGC TGGAAAAATT GCAAGAGTTA
   XM_451902  ATGTCAAATT GATCTTAGCA GCAAGAAGGT TGGCTAAGTT GGAAGAATTG
XM_002553230  AAATCAAGTT GATCCTTGCA GCAAGAAGAC TCGACAAGCT GGAGGAAATC
XM_002494574  AAATTAAATT GATCTTGGCT GCAAGAAGAT TAGAGAAATT ACACGAGTTT
     YMR226C  ATATGAAACT GATCTTGGCT GCTAGAAGAT TAGAAAAGCT CGAGGAATTG
NM_001018495  ...TAAAACT TATTTTGGCT GCTCGCAGAT TTTCTACCGT TGAAGAAATT
   XM_501554  ACCTCAAGTT TGTGCTTGCT GCCCGACGAA CCGACCGTCT GGACGAGCTC
   XM_957621  GCCTCAAGCT GATCTTGACG GCCCGCCGCG TCGATGCCCT GGAGCAGATT

              251                                                300
AABY01000127  AAGAAAACTA TTGATCAGG. AGTTTCCA.. AACGCCAAAG TTCATGTGGC
AACH01000019  AAGAAGACTA TCGATGAAG. AGTTTCCA.. AACGCAAAGG TCCATGTGAC
AACA01000631  AAGAAGACCA TCGACGAGG. AGTTTCCC.. AATGCAAAGG TTCACGTTGG
AACF01000116  AAGGAAGTCA TTGAGGCAA. AATACCCT.. AAGAGTAAAG TCTATATTGG
    AE016819  AAACAGGAGA TCGA.AAGCA AGTATCCC.. AACGCCAAGA TCCATGTCGG
    CR382139  TCT.AAAGAA TTGACTGACA AATATTCA.. TCCATCAAGG TTCATGTTGC
```

2017200439 20 Jan 2017

```
XM_001387479  TCA.GACTCA TTGACCAAGG AATTTCCA.. ACTATCAAAA TCCATTCTGC
XM_001482184  TCG.CAACAA TTGAGCCTCA TTTACCCG.. CAAATTAAAA TCCATTCTGC
XM_002419771  TCT.GATACT TTAATTAAAC AATATCCT.. AATATTAAAA TTCATCATGA
    CR380959  AAG.AAGGAA TTGCTGGCTC AATATCCT.. GATGCAAAGA TTCATATAGG
XM_001645671  AAGGAAACAA TTTGTAAAG. AATATCCA.. GAATCTAAGG TTCATGTTGA
 SAKL0H04730  AAGGACAAAA TCACCAAGG. AGTATCCT.. GAAGCCAAGG TTTATATTGG
   XM_451902  AAGGAAAAAA TCAATGCTG. AATACCCA.. CAAGCTAAAG TATATATCGG
XM_002553230  AAAGCCAAGG TTTCTAAAG. ACTTCCCT.. GAAGCACAGG TGCATATCGG
XM_002494574  AAAGAACAAA CTAC.AAAGA GTTACCCA.. AGCGCTCAAG TCCACATTGG
     YMR226C  AAGAAGACCA TTGATCAAG. AGTTTCCA.. AACGCAAAAG TTCATGTGGC
NM_001018495  GCAAAGGAGT TAGA.ATCGA AATAT..... GAAGTATCGG TTCTTCCTCT
   XM_501554  AAGAAGAAGC TGGAGACCG. ACTACAAG.. GGTATCCAGG TGCTGCCTTT
   XM_957621  GCCAAGGAGA TCCGCCAAGA GGTTGGCGAG GGCGTGCAAG TGCTGCCTGT

                    301                                           350
AABY01000127  CCAACTGGAT ATCACTCAAG CAGAAAAGAT CAAGCCCTTT ATTGAGAATT
AACH01000019  CAAACTGGAC ATCACACAGA CAGAAAAGAT CAAGCCCTTT ATTGAAAACT
AACA01000631  CCAACTGGAT ATCACACAGG CCGAGAAGAT CAAGCCCTTC ATTGAAAACT
AACF01000116  GAAGTTGGAT GTGACAGAGC TTGAGACCAT TCAACCATTC TTGGATAATC
    AE016819  ACAATTGGAT GTGACCCAAC TGGACCAGAT CCGCCCATTT TTGGAGGGAC
    CR382139  TCAACTAGAT GTATCTAAGC TCGAGACTAT CAAGCCATTT ATCAATGATT
XM_001387479  CAAATTGGAT GTGACCGAAC ATGATGGCAT CAAGCCTTTC ATTTCTGGTT
XM_001482184  TCGTCTTGAT GTCTCTGAGT TTTCGTCACT TAAGCCGTTC ATTACTGGGT
XM_002419771  TTTTTTGGAT GTTACTATTA AAGATTCAAT TTCAAAATTC ATTGCTGGAA
    CR380959  TAAACTGGAT GTTACAGACT TTGAAAACGT CAAGCAGTTT TTGGCTGACT
XM_001645671  AGAATTAGAT ATTTCTGATA TTAATAGAAT CCCAGAATTT ATTGCAAAAT
 SAKL0H04730  TGTGCTTGAT GTGACCGAAA CGGAAAAAAT CAAACCATTC TTGGATGGTT
   XM_451902  TCAATTGGAC GTCACTGAAA CTGAGAAGAT TCAACCTTTC ATTGATAACT
XM_002553230  CCAGCTAGAT GTGACTCAGA CGGACAAAAT CCAGCCTTTT GTCGACAATT
XM_002494574  TAAATTGGAC GTCACTGCAA TTGACACCAT AAAACCATTT TTGGATAAAT
     YMR226C  CCAGCTGGAT ATCACTCAAG CAGAAAAAAT CAAGCCCTTC ATTGAAAACT
NM_001018495  TAAATTGGAT GTTTCTGATT TGAAGTCTAT TCCTGGGGTA ATTGAGTCAT
   XM_501554  CAAGCTGGAC GTGTCCAAGG TCGAGGAGAC CGAGAACATT GTGTCCAAGC
   XM_957621  CAAGCTGGAC GTCAGCCAGC CGGAGGAGGT CAGGGGGTTT GTTGGGAACT

                    351                                           400
AABY01000127  TGCCAAAGGA GTTCAAAGAC ATTGACATTT TGGTGAACAA CGCTGGGAAG
AACH01000019  TGCCAGAGGA GTTCAAAGAC ATTGATATTC TGGTGAACAA CGCTGGTAAG
AACA01000631  TGCCGGAGGC ATTCAAGGAT ATTGACATCC TGATAAACAA TGCCGGCAAA
AACF01000116  TTCCTGAGGA ATTTAAGGAT ATTGATATCT TGATTAATAA TGCCGGGAAG
    AE016819  TACCTGAGGA GTTCCGAGAC ATTGATATTT TAATTAACAA CGCAGGTAAG
    CR382139  TACCGAAAGA ATTCTCTGAC GTGGATGTAT TAGTCAACAA TGCAGGCTTG
XM_001387479  TACCCAAGGA TTTCGCCGAC ATCGATGTGT TGATCAACAA TGCTGGAAAA
XM_001482184  TGCCAAAGGA TTTTGCTAGC ATCGACGTTT TGGTGAATAA TGCGGGGAAA
XM_002419771  TTCCTCATGA ATTTG..AAC CT.GATGTAT TAATTAATAA TAGTGGTAAA
    CR380959  TGCCAGAAGA GTTCAAGGAC ATCGACATCC TGATCAATAA CGCTGGTAAA
XM_001645671  TACCTGAAGA ATTCAAAGAT ATTGATATAT TGATTAACAA TGCAGGTAAA
 SAKL0H04730  TACCAGAAGA ATTTAAAGAT ATTGACATCT TGATCAATAA TGCAGGCAAA
   XM_451902  TGCCGGAAGA ATTCAAGGAT ATCGATATTT TGATTAACAA TGCCGGTAAA
XM_002553230  TGCCCGAAGA GTTCAAAGAC ATCGACATCC TGATCAACAA CGCGGGCAAG
XM_002494574  TACCAAAGGA ATTTCAAGAT ATCGATATTT TGATCAACAA TGCCGGTAAG
     YMR226C  TGCCACAAGA GTTCAAGGAT ATTGACATTC TGGTGAACAA TGCCGGAAAG
NM_001018495  TGCCAAAGGA ATTTGCTGAT ATCGATGTCT TGATTAATAA TGCTGGACTT
   XM_501554  TGCCCAAGGA GTTTTCCGAG GTGGACGTGC TTATCAACAA CGCCGGCATG
   XM_957621  TGCCCGAGGA GTGGAGGGAT ATCCATGTGC TGGTCAATAA TGCTGGTCTC

                    401                                           450
```

```
AABY01000127  GCCCTTGGTA CCGACCGTGT GGGGGAGATT GCAACACAAG ATATCCAGGA
AACH01000019  GCTCTTGGTA CGGACCGTGT TGGGGAGATT GATACACAGG ACGTCCAGGA
AACA01000631  GCCCTGGGCT CCGAACGTGT CGGGGAAATT GCCACACAGG ACATCCAGGA
AACF01000116  GCATTAGGTT CCGATCGTGT AGGTGATATT GATATAAAAG ATGTGAAGGG
    AE016819  GCCCTCGGCA CTGAGAGGGT GGGGGAAATC TCGATGGACG ATATCCAGGA
    CR382139  GCTTTGGGCC GTGATGAAGT TGGAACCATT GACACAGATG ATATGTTATC
XM_001387479  GCTCTTGGAA AAGCATCTGT TGGTGAAATC AGTGACAGTG ATATCCAAGG
XM_001482184  GCATTGGGAA GAGCCAATGT TGGTGAAATT TCCCAAGAGG AAATCAATGG
XM_002419771  GCCTTGGGGA AAGAAGAAGT TGGAGAATTG AAAGATGAAG ATATTACGGA
    CR380959  GCGTTGGGGT CTGACAAAGT TGGAGACATT GACCCTGAGG ATATCGCAGG
XM_001645671  GCATTAGGAA GTGATACTAT TGGTAATATC GAGAATGAGG ATATTAAAGG
 SAKL0H04730  GCGTTAGGCT CTGATCCTGT TGGTACCATC AAAACTGAAG ATATTGAAGG
   XM_451902  GCTTTGGGAT CTGATGTTGT CGGTACCATC AGTAGCGAGG ACATCAAAGG
XM_002553230  GCGCTCGGAT CCGACCCCGT GGGCACAATC GACCCCAATG ATATTCAAGG
XM_002494574  GCATTAGGTA CTGATAAAGT TGGTGATATT GCAGATGAAG ACGTGGAAGG
     YMR226C  GCTCTTGGCA GTGACCGTGT GGGCCAGATC GCAACGGAGG ATATCCAGGA
NM_001018495  GCTCTAGGTA CCGATAAAGT CATTGATCTT AATATTGATG ACGCCGTTAC
   XM_501554  GTCCACGGCA CCGAAAAGGT TGGCTCCATC AACCAGAACG ACATTGAGAT
   XM_957621  GTAAAAGG.. .....A..GC TCCCTCCATC GCCGAAGAAG ACATCAACGT

                451                                              500
AABY01000127  TGTGTTTGAC ACCAACGTCA CAGCTTTAAT TAATATCACT CAAGCTGTGC
AACH01000019  CGTGTTCGAC ACCAACGTCT CGGCTTTGAT TAATGTCACA CAGGCTGTTC
AACA01000631  CGTGTTCGAC ACCAACGTCA CGGCGTTGAT CAACGTCACG CAAGCAGTGC
AACF01000116  AATGATGGAT ACCAATGTCT TGGGGTTGAT CAATGTGACG CAAGCTGTGT
    AE016819  GGTTTTCAAC ACTAATGTTA TCGGCTTGGT GCACTTGACT CAGGAGGTTC
    CR382139  GATGTTTCAA ACTAATGTTT TAGGGTTAAT TACCATCACA CAGGCTGTTT
XM_001387479  CATGATGCAA ACGAATGTCT TGGGACTCAT CAACATGACT CAGGCTGTGA
XM_001482184  CATGTTCCAT ACCAATGTTC TTGGGTTGAT AAACTTAACT CAGGAGGTGT
XM_002419771  AATGTTTGAT ACTAATGTCA TTGGAGTCAT TCGTATGACT CAAGCAGTTT
    CR380959  AATGGTTAAC ACCAACGTCC TTGCATTGAT CAATTTAACA CAATTGTTGT
XM_001645671  TATGTTTGAG ACTAACGTTT TTGGATTAAT CTGTTTAACA CAAGCTGTAC
 SAKL0H04730  AATGATCAAC ACCAATGTCT TAGCTCTTAT CAATATTACT CAAGCTGTCT
   XM_451902  TATGATAGAT ACTAACGTTG TTGCCCTTAT CAACGTTACC CAAGCTGTTT
XM_002553230  CATGATCCAG ACTAACGTTA TCGGGCTTAT AAATGTTACC CAAGCCGTTC
XM_002494574  TATGTTCGAC ACCAATGTCT TGGGGTTAAT CAAAGTTACT CAAGCTGTTT
     YMR226C  CGTGTTTGAC ACCAACGTCA CGGCTTTAAT CAATATCACA CAAGCTGTAC
NM_001018495  CATGATTACT ACCAATGTTC TTGGTATGAT GGCTATGACT CGTGCGGTTC
   XM_501554  CATGTTCCAC ACAAACGTGC TCGGACTCAT TTCTGTCACT CAGCAGTTTG
   XM_957621  CATGTTTGCC ACCAACGTCA CCGGCCTGAT CAACATGACC CAAGCCATCC

                501                                              550
AABY01000127  TGCCCATTTT TCAAGCCAAG AACTCAG... .........G GGATATTGTG
AACH01000019  TGCCCATCTT CCAAGCTAAG AACTCAG... .........G GGATATTGTG
AACA01000631  TGCCAATTTT CCAAGCCAAG AACTCAG... .........G GGACATCGTC
AACF01000116  TGCACATTTT CCAAAAGAAG AACTCCG... .........G TGATATTGTG
    AE016819  TACCTATTAT GAAAGCCAAG AATTCCG... .........G GGACATTGTC
    CR382139  TGCCAATCAT GAAAAAGAAG AACAGCG... .........G AGATGTTGTT
XM_001387479  TTCCCATTTT TAAGGCTAAA AATTCTG... .........G AGATATCGTC
XM_001482184  TACCCATCTT CAAAAAGAAA AATGCTG... .........G AGATATTGTG
XM_002419771  TACCTTTACT TAAAAAAAAA CCTTATG... .........C TGATGTGGTT
    CR380959  TGCCATTATT CAAGAAGAAG AACAGTG... .........G TGATATCGTC
XM_001645671  TTCCAATATT CAAGGCTAAA AATGGTG... .........G TGATATTGTC
 SAKL0H04730  TGCCAATCTT CAAAGCCAAG AATTTCG... .........G TGATATCGTA
   XM_451902  TGCCTATTTT CAAAGCAAAG AATTCCG... .........G TGACATCGTT
XM_002553230  TGCCCATCTT CAAGGCCAAA AACTCTG... .........G TGATATCGTG
XM_002494574  TACCTATCTT CAAAAGAAAA AATTCTG... .........G TGATGTCGTT
```

2017200439 20 Jan 2017

```
     YMR226C  TGCCCATATT CCAAGCCAAG AATTCAG... .........G AGATATTGTA
NM_001018495  TTCCTATATT CTACAGCAAA AACAAGG... .........G TGATATTTTG
   XM_501554  TCGGCGAGAT GCGAAAGCGA AACAAGG... .........G CGACATTGTC
   XM_957621  TGCCCATCTT CAAGGCCCGC GGCTCCGAGG GTGGTTCGGG AGACATTGTG

                551                                                600
AABY01000127  AACTTGGG.T TCGGTGGCTG GCAGGGATGC ATACCCAACG GGTTCCATCT
AACH01000019  AACTTGGG.C TCGGTAGCTG GCAGAGATGC ATACCCAACG GGCTCCATCT
AACA01000631  AACTTGGGGC TCGGTGGCCG GCAGAGACGC ATACCCCACA GGCTCCATCT
AACF01000116  AACTTAGG.T TCAGTTGCTG GAAGAGATGC ATACCCAACA GGGTCCATTT
    AE016819  AATGTTGG.G TCGATTGCCG GCCGCGAAGC CTACCCTGGT GGCTCTATTT
    CR382139  AATATAGG.T TCAATTGCTG GAAGAGACTC TTACCCTGGA GGTGGAATTT
XM_001387479  AACATCGG.T TCGATTGCTG GAAGAGACCC TTACCCTGGT GGATCGATCT
XM_001482184  AACATTGG.C TCAGTGGCCG GTAGAGAACC TTACCCTGGA GGTGCAGTAT
XM_002419771  TTCATTGG.A AGTATTGCTG GACGTGTTCC TTATAAAAAT GGAGGTGGTT
    CR380959  AACTTGGG.A TCGATTGCTG GTAGAGACGC ATACCCAACG GGTGCTATAT
XM_001645671  AATTTAGG.G TCAATTGCAG GCATAGAAGC TTACCCAACA GGATCTATAT
 SAKL0H04730  AACTTGGG.G TCTGTCGCTG GTAGAGATGC TTATCCAACC GGTGCAATCT
   XM_451902  AACTTAGG.T TCTGTTGCCG GTAGAGATGC ATATCCAACT GGTTCTATCT
XM_002553230  AACCTGGG.T TCTGTCGCTG GCAGAGAAGC TTACCCTACA GGATCTATTT
XM_002494574  AACATTAG.T TCGGTTGCTG GTAGAGAGGC ATACCCAGGT GGTTCCATTT
     YMR226C  AATTTGGG.T TCAATCGCTG GCAGAGACGC ATACCCAACA GGTTCTATCT
NM_001018495  AACGTTGG.C AGTATTGCCG GCAGAGAATC ATACGTAGGC GGCTCCGTTT
   XM_501554  AACATTGG.C TCCATCGCCG GACGAGAGCC CTACGTTGGA GGAGGAATCT
   XM_957621  AACATCGG.T TCTATTGCCG GGAGAGAACC ATACGCGGGA GGCTCCATCT

                601                                                650
AABY01000127  ATTGTGCCTC CAAGTTTGCC GTGGGGGCGT TCACTGATAG TTTAAGAAAG
AACH01000019  ATTGTGCATC TAAGTTTGCC GTCGGGGCTT TCACTGAGAG TTTGAGAATG
AACA01000631  ACTGTGCTTC CAAGTTTGCC GTCGGTGCGT TCACTGACAG TTTGAGAAAG
AACF01000116  ACTGTGCTTC TAAATTTGCC GTGAGGGCCT TTACTGAAAG TTTGAGAAGG
    AE016819  ACTGTGCCAC GAAACATGCG GTCAAGGCTT TCACCAGGGC CATGCGGAAG
    CR382139  ACTGTCCAAC TAAGGCAAGT GTCAAGTCGT TTTCGCAAGT TTTAAGAAAG
XM_001387479  ACTGTGCCTC CAAGGCTGCT GTTAAGTTCT TCTCGCATTC TTTGAGAAAG
XM_001482184  ACTGTGCTTC AAAGGCAGCA GTTAACTACT TTTCTCATTC TTTGAGAAAG
XM_002419771  ATTGTGCATC TAAAGCTGCT GTTCGTAGTT TCACCGATAC ATTTAGAAAA
    CR380959  ACTGTGCAAC AAAACATGCT GTCAGGGCAT TCACACAATC CTTAAGGAAG
XM_001645671  ATTGTGCAAC TAAATTTGCA GTTAAAGCAT TCACTGAAAG TTTAAGAAAG
 SAKL0H04730  ACTGTGCTAG CAAACATGCA GTCAGAGCCT TCACTCAAAG TTTGAGGAAG
   XM_451902  ATTGTGCTTC GAAGCATGCT GTCAGAGCGT TCACTCAGTC TTTGAGAAAA
XM_002553230  ACTGCGCTAC GAAGCACGCG GTGCGTGCTT TCACCCAGAG CCTGCGCAAG
XM_002494574  ACTGTGCTAC TAAACACGCT GTTAAGGCAT TCACTGAAAG TTTGCGTAAG
     YMR226C  ATTGTGCCTC TAAGTTTGCC GTGGGGGCGT TCACTGATAG TTTGAGAAAG
NM_001018495  ACTGCTCTAC CAAGTCTGCC CTTGCTCAAT TCACTTCCGC TTTGCGTAAG
   XM_501554  ACTGTGCCAC CAAGGCCGCC GTGCGATCTT TCACTGAGAC TCTCCGAAAA
   XM_957621  ACTGCGCCAC CAAGGCTGCT GTGCGCAGCT TCACCGACGC GCTGCGCAAG

                651                                                700
AABY01000127  GAGCTTATCA ACACCAAGAT CAGAGTCATC CTAATCGCAC CAGGGCTAGT
AACH01000019  GAACTTATAA ACACTAAGAT TAGAGTCATT CTAATTGCAC CAGGGTTAGT
AACA01000631  GAACTGATCA ACACGAAGAT CAGAGTTATC TTGATCGCGC CGGGGCTGGT
AACF01000116  GAATTAATTA ATACCAAGAT TAGGGTGATA TTGATAGCCC CGGGTATCGT
    AE016819  GAGCTCATTA GCACCAAGAT CCGGGTCTTC GAAATTGCGC CGGGCTCTGT
    CR382139  GAATTGATTA GCACCAAGAT TAGAGTTCTT GAGGTTGACC CTGGTAATGT
XM_001387479  GAACTCATTA ACACCAGAAT CAGAGTTTTG GAAGTTGATC CAGGTGCTGT
XM_001482184  GAAACTATCA ATTCCAAAAT CAGGGTCATG GAGGTGGATC CTGGGGCAGT
XM_002419771  GAAACTATTA ATACTGGTAT TAGAGTCATT GAAGTTGATC CAGGTGCAGT
```

2017200439 20 Jan 2017

```
    CR380959  GAATTGATCA ACACCGACAT TAGAGTAATT GAAATTGCTC CTGGTATGGT
XM_001645671  GAATTGATTA ATACAAAGAT CAGAGTTATT GAAATTGCAC CAGGTATGGT
 SAKL0H04730  GAATTGGTGA ACACCAATAT CAGAGTGATT GAAATTGCTC CGGGTAATGT
   XM_451902  GAATTAATCA ATACTGGTAT TAGGGTCATT GAGATTGCTC CAGGTAATGT
XM_002553230  GAACTGATCA ACACAAACAT CAGGGTTATT GAGGTCGCTC CAGGTAACGT
XM_002494574  GAATTAGTCG ATACAAAAAT CAGAGTCATG AGTATTGATC CTGGTAATGT
     YMR226C  GAGCTCATCA ACACTAAAAT TAGAGTCATT CTAATTGCAC CAGGGCTAGT
NM_001018495  GAGACTATTG ACACTCGCAT TCGTATTATG GAGGTTGATC CTGGCTTGGT
   XM_501554  GAGAACATCG ACACTCGAAT CCGAGTCATT GAGGTTGATC CTGGAGCCGT
   XM_957621  GAGCTGATCG CCACGCGCAT CCGTGTCATG GAGATTGACC CTGGCCAGGT

                701                                                750
AABY01000127  CGAAACTGAA TTTTCACTGG TTAGATACAG AGGCAACGAG GAGCAAGCCA
AACH01000019  CGAAACTGAG TTTTCCCTGG TTAGATACAG AGGTAACGAA GAACAAGCCA
AACA01000631  TGAGACCGAG TTCTCACTGG TCAGATACAG AGGTAATGAG GAACAAGCTA
AACF01000116  CGAAACTGAA TTCTCAGTTG TTAGATACAA GGGTGATAAT GAGCGTGCTA
    AE016819  AGAAACGGAA TTCTCCATGG TTCGTATGCG CGGTAACGAA GAGAATGCCA
    CR382139  TGAAACTGAA TTTTCAAATG TCAGATTCAA GGGCGATATG GAAAAGGCAA
XM_001387479  GTTGACCGAG TTCTCTTTGG TTCGTTTCCA CGGTGATCAG GGAGCTGCTG
XM_001482184  AGAGACAGAG TTCTCGTTGG TTCGTTTTGG CGGTGATGCC GAGGCTGCGA
XM_002419771  ACTTACTGAA TTTAGTGTTG TTCGTTATAA AGGTGACACT GATGCTGCCG
    CR380959  CGAAACCGAG TTTTCTGTGG TCAGGTACAA AGGTGACAAG TCCAAAGCAG
XM_001645671  TAACACTGAA TTTTCTGTAA TTAGATATAA AGGTGACCAA GAAAAGGCAG
 SAKL0H04730  TGAAACCGAG TTTTCCTTAG TTAGATATAA AGGTGATACG GACCGTGCTA
   XM_451902  CGAAACTGAG TTCTCTCTAG TTAGATACAA GGGTGATGCC GATCGTGCTA
XM_002553230  GGAGACCGAG TTTTCTCTGG TTAGATACAA GGGCGACTCT GAGAAAGCCA
XM_002494574  AGAGACCGAA TTTTCTATGG TTAGATTCCG TGGTGATACA GAAAAGGCAA
     YMR226C  CGAGACTGAA TTTTCACTAG TTAGATACAG AGGTAACGAG GAACAAGCCA
NM_001018495  CGAAACCGAA TTCAGCGTTG TGAGATTCCA CGGAGACAAA CAAAAGGCTG
   XM_501554  TGAGACCGAG TTCTCCGTCG TGCGATTCCG AGGAGACAAG TCCAAGGCCG
   XM_957621  CGAGACCGAG TTCAGCGTGG TGAGGTTCTA TGGTGATAAG AACAAGGCTG

                751                                                800
AABY01000127  AGAATGTCTA CAA.GGACAC CACCCCATTA ATGGCTGATG ACGTGGCTGA
AACH01000019  AGAATGTTTA CAA.GGACAC CACTCCGTTG ATGGCCGATG ACGTGGCTGA
AACA01000631  AAAACGTCTA CAA.GGACAC TACGCCGTTG ATGGCCGACG ACGTGGCTGA
AACF01000116  AATCTGTCTA CGATGGAGTT CACCCCTTGG AAG.CAGACG ACGTAGCAGA
    AE016819  AGAAAGTGTA CCA.GGGATT TGAACCCCTA GATGGTGATG ATATCGCTGA
    CR382139  AGCTGGTTTA CGC.GGGTAC TGAACCATTA TTATCCGAAG ACGTAGCTGA
XM_001387479  ATGCTGTTTA TGA.AGGTAC CCAACCTTTG GATGCCTCTG ATATCGCAGA
XM_001482184  AAAAGGTGTA TGA.GGGAAC CGAGCCTTTG GGCCCAGAAG ATATTGCGGA
XM_002419771  ATGCTGTTTA TAC.TGGTAC TGAACCATTA ACACCAGAAG ATGTTGCTGA
    CR380959  ACGACGTCTA CAG.AGGTAC AACACCACTA TATGCCGATG ATATCGCGGA
XM_001645671  ATAAAGTTTA TGA.AAACAC TACTCCTTTA TATGCAGATG ACATCGCTGA
 SAKL0H04730  AAAAGGTTTA TGA.AGGTAC TAACCCATTA TATGCAGATG ACATTGCAGA
   XM_451902  AACAGGTTTA CAA.AGGTAC TACTCCTCTA TATGCAGATG ACATTGCTGA
XM_002553230  AGAAGGTTTA CGA.AGGCAC ACAACCCCTT TACGCTGACG ATATCGCAGA
XM_002494574  AGAAGGTTTA CCA.AGACAC TGTCCCATTA TATGCAGATG ACATTGCAGA
     YMR226C  AGAATGTTTA CAA.GGATAC TACCCCATTG ATGGCTGATG ACGTGGCTGA
NM_001018495  ATAATGTTTA CAA.AAATAG TGAGCCTTTG ACACCCGAAG ACATTGCTGA
   XM_501554  ACGCTGTTTA CGC.TGGAAC CGAGCCTCTG GTCGCTGACG ATATTGCCGA
   XM_957621  ATGCCGTCTA TGC.CGGTGT CGATCCCTTG ACGCCCGATG ATATCGCAGA

                801                                                850
AABY01000127  TTTGATCGTG TACGCAACTT CCAGGAAACA AAACACTGTA ATTGCAGACA
AACH01000019  TTTGATTGTG TATGCGACTT CAAGGAAGCA GAACACTGTA ATTGCAGACA
AACA01000631  CTTAATCGTA TATTCCACTT CCAGAAAGCA GAACACCGTG GTTGCCGACA
```

2017200439 20 Jan 2017

```
AACF01000116  TTTAATTGTA TACACCACTT CAAGAAAACA GAACACAGTA ATTGCTGACA
    AE016819  TACAATTGTC TATGCCACAT CCAGAAGATC CAACACCGTA GTTGCAGAGA
    CR382139  GGTTGTCGTA TTCGGACTTA CAAGAAAGCA AAATACCGTT ATTGCTGAGA
XM_001387479  AGTTATCGTG TTTGGTATCA CCAGAAAGCA GAACACCGTC ATAGCCGAAA
XM_001482184  AATCATTGTG TTTGCTGTGT CGAGAAAAGC CAAAACTGTC ATTGCGGAAA
XM_002419771  AGTGGTTGTT TTTGCATCTT CAAGAAAACA AAATACCGTT ATTGCTGATA
    CR380959  TTTGATTGTG TACTCTACCA GCAGAAAGCC AAACATGGTG GTAGCAGATG
XM_001645671  TTTGATAGTT TACACCACTT CTAGAAAATC GAATACCGTT ATCGCTGATG
 SAKL0H04730  CCTTATTGTG TATGCTACTT CTAGAAAGCC TAATACTGTC ATTGCGGATG
   XM_451902  CTTGATCGTT TATGCCACTT CAAGAAAACC TAATACTGTC ATCGCTGATG
XM_002553230  CCTAATCGTT TACGCAACCT CGAGAAAACC AAACACCGTC ATCGCGGACG
XM_002494574  TTTAATCGTC TATGCAACCT CTAGAAAGCA AAACACTGTC ATTGCTGACA
     YMR226C  TCTGATCGTC TATGCAACTT CCAGAAAACA AAATACTGTA ATTGCAGACA
NM_001018495  GGTGATTCTT TTTGCCCTCA CTCGCAGAGA AAACGTCGTT ATTGCCGATA
   XM_501554  GTTCATCACC TACACTCTCA CTCGACGAGA GAATGTCGTC ATTGCCGATA
   XM_957621  GATCGTGGTG TTCGTCGTAA CACGACGGGA GAACGTTGTT GTTGCTGATA

                851                                                900
AABY01000127  CGCTAATCTT TCCAACCAAC CAAGCATCGC CTCACCACAT CTTCCGTGGA
AACH01000019  CACTAATCTT TCCTACCAAC CAAGCGTCAC CTTACCATAT CTTTCGCGGG
AACA01000631  CCCTGATCTT CCCCACCAAC CAAGCCTCGC CCTACCACAT CTTTCGCGGT
AACF01000116  CTTTGATATT CCCAACCTCT CAAGGTTCCG CATTCCACGT CCATCGCGAT
    AE016819  TGGTCGTTTA CCCATCCGCG CAAGGTTCTC TGTACGATAC TCACCGCAAC
    CR382139  CATTAGTCTT TTCAACCAAT CAAGCCAGCT CATCTCACTT ATACCGTGAA
XM_001387479  CCTTGGTATT CCCAAGTCAC CAGGCTTCTG CCTCTCATGT TTACAAGGCT
XM_001482184  CTTTGGTGTT TCCTACCCAT CAGGCTGGAG CAGTTCATGT TCATAGAGGG
XM_002419771  CTTTGATTTT CCCAAATCAT CAAGCTTCTC CAGATCATGT TTATAGAAAA
    CR380959  TCCTGGTCTT CCCAACACAC CAGGCATCGG CTTCGCACAT CTACAGGGGC
XM_001645671  TTTTGGTATT CCCAACATGC CAAGCTTCTG CATCCCATAT CTATCGTGGA
 SAKL0H04730  TTTTGGTTTT TGCTTCAAAC CAAGCATCCC CTTACCATAT CTATCGCGGT
   XM_451902  TTTTGGTATT TGCTTCCAAC CAAGCATCTC CTTACCACAT TTACCGTGGC
XM_002553230  TTTTGGTTTT CGCTTCGAAC CAGGCTTCGC CTTACCACAT TTACCGTGGT
XM_002494574  CTTTGATCTT CTCTTCTAAC CAGGCATCAC CATACCACCT CTACAGAGGC
     YMR226C  CTTTAATCTT TCCAACAAAC CAAGCGTCAC CTCATCATAT CTTCCGTGGA
NM_001018495  CACTTGTTTT CCCATCCCAT CAAGGTGGTG CCAATCATGT GTACAGAAAG
   XM_501554  CTCTCATTTT CCCCAACCAC CAGGCTTCTC CTACTCACGT CTACCGAAAG
   XM_957621  CGTTGGTCTT CCCTAGCCAT CAGGCTGGCG CTGGTATTAT GCACCGCAAG

                901                  921
AABY01000127  TGA..... ..........  . (SEQ ID NO:41)
AACH01000019  TGA..... ..........  . (SEQ ID NO:47)
AACA01000631  TAA..... ..........  . (SEQ ID NO:43)
AACF01000116  TAA..... ..........  . (SEQ ID NO:45)
    AE016819  TAA..... ..........  . (SEQ ID NO:49)
    CR382139  AGCGATAAAT AA......  . (SEQ ID NO:53)
XM_001387479  CCTAAGTAG. ..........  . (SEQ ID NO:55)
XM_001482184  CCGCTTGAGT GA......  . (SEQ ID NO:57)
XM_002419771  CCTAATTAA. ..........  . (SEQ ID NO:61)
    CR380959  GACTAA.... ..........  . (SEQ ID NO:51)
XM_001645671  TAA..... ..........  . (SEQ ID NO:59)
 SAKL0H04730  GACTAA.... ..........  . (SEQ ID NO:69)
   XM_451902  GAATAG.... ..........  . (SEQ ID NO:67)
XM_002553230  TAG..... ..........  . (SEQ ID NO:65)
XM_002494574  TCTCAAGACA AAACCAATTG A (SEQ ID NO:63)
     YMR226C  TAA..... ..........  . (SEQ ID NO:24)
NM_001018495  CAAGCGTAG. ..........  . (SEQ ID NO:73)
   XM_501554  AACTGA.... ..........  . (SEQ ID NO:71)
```

XM_957621 TCGACATGA. . . . . (SEQ ID NO:75)

2017200439 20 Jan 2017

Figure 13

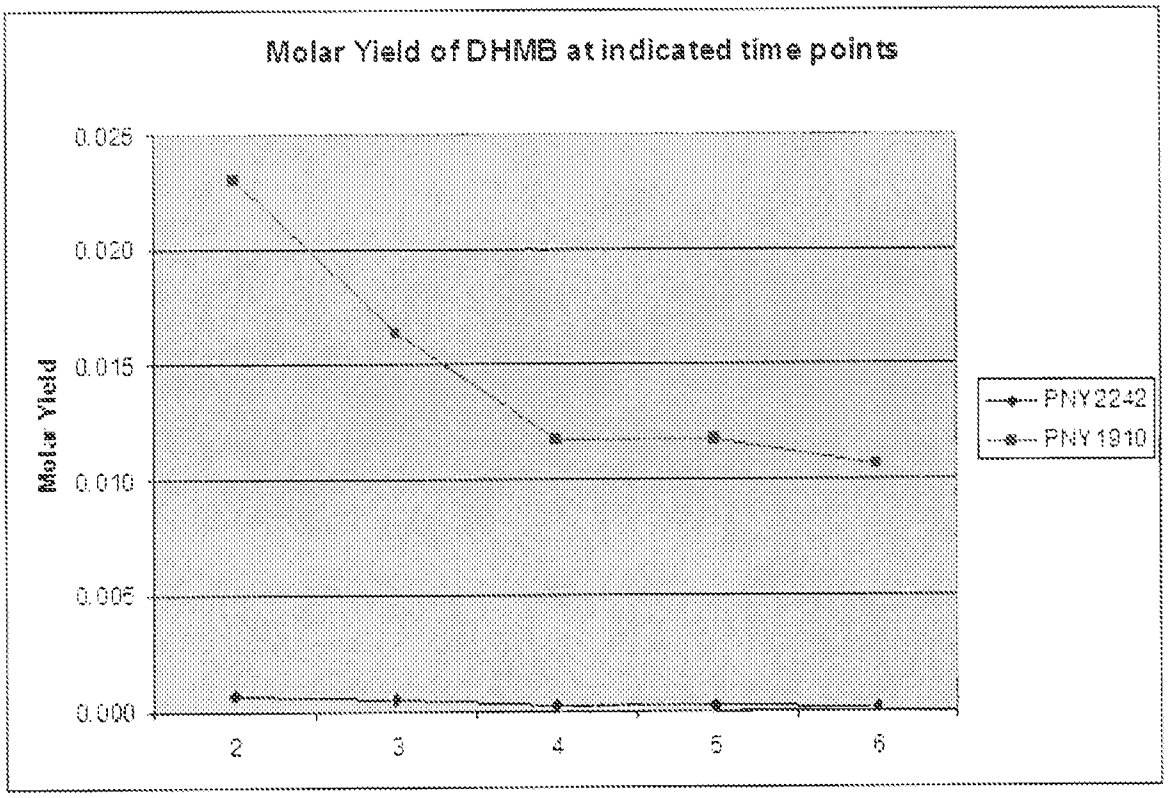
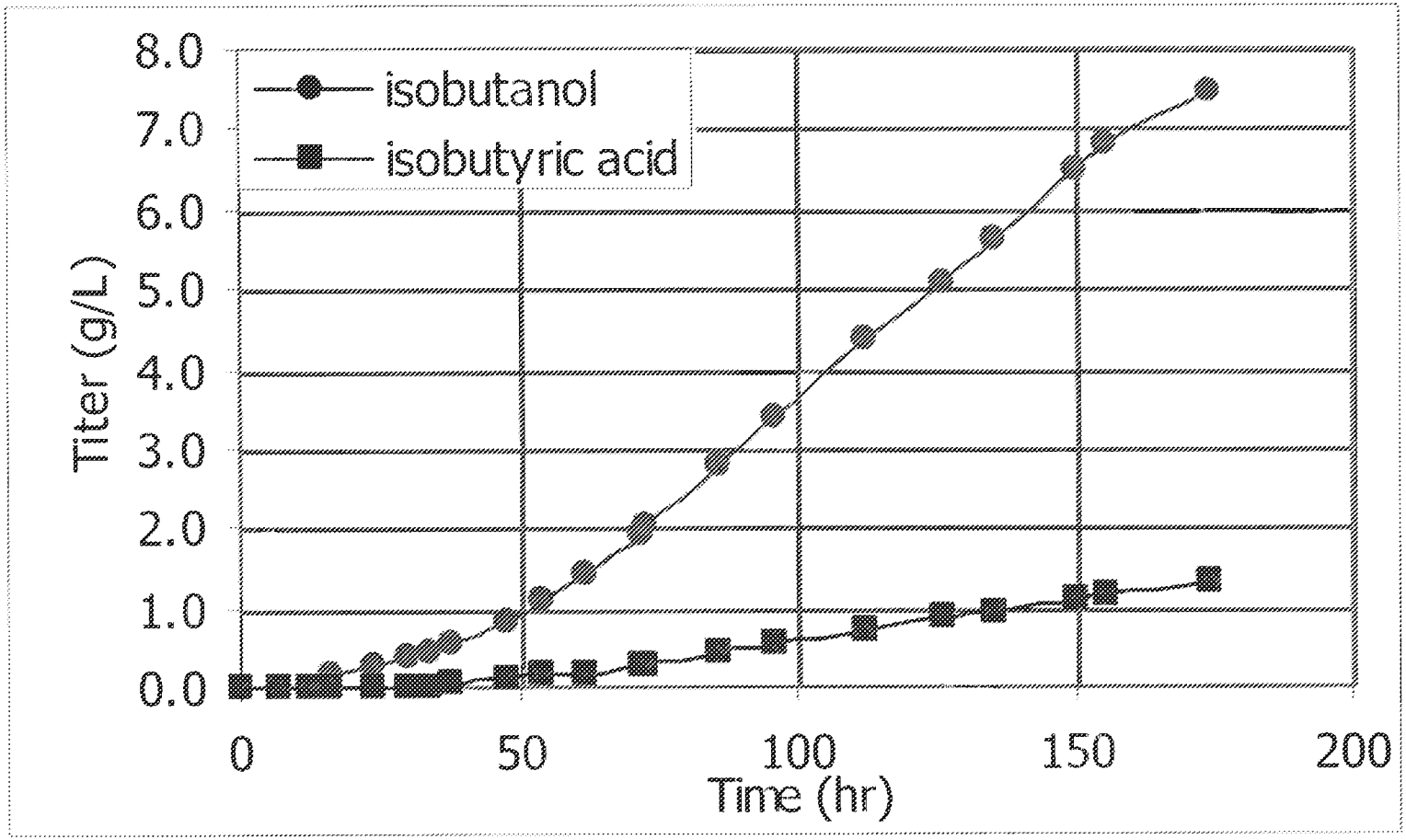


Figure 14


8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
Titer (g/L)
isobutanol
isobutyric acid
0
50
100
150
200
Time (hr)

20 Jan 2017 2017200439

SEQUENCE LISTING

<110> BUTAMAX(TM) Advanced Biofuels, LLC

<120> Host Cells and Methods for Production of Isobutanol

<130> CL5367USNA

<150> 61/472,484
<151> 2011-04-06

<150> 61/467,261
<151> 2011-03-24

<150> 61/472,487
<151> 2011-04-06

<150> 61/476,271
<151> 2011-03-24

<150> 61/570,513
<151> 2011-12-14

<150> 61/467,249
<151> 2011-03-24

<150> 61/472,497
<151> 2011-04-06

<150> 61/472,474
<151> 2011-04-06

<160> 965

<170> PatentIn version 3.5

<210> 1
<211> 335
<212> PRT
<213> Sulfolobus solfataricus

<400> 1

Met Lys Cys Thr Ser Lys Ile Tyr Thr Asp Asn Asp Ala Asn Leu Asp
1 5 10 15

Leu Ile Lys Gly Lys Arg Ile Ala Val Leu Gly Tyr Gly Ser Gln Gly
20 25 30

Arg Ala Trp Ala Gln Asn Leu Arg Asp Ser Gly Leu Asn Val Val Val
35 40 45

Gly Leu Glu Arg Glu Gly Lys Ser Trp Glu Leu Ala Lys Ser Asp Gly
50 55 60

Ile Thr Pro Leu His Thr Lys Asp Ala Val Lys Asp Ala Asp Ile Ile
65 70 75 80

Ile Phe Leu Val Pro Asp Met Val Gln Arg Thr Leu Trp Leu Glu Ser
85 90 95

Val Gln Pro Tyr Met Lys Lys Gly Ala Asp Leu Val Phe Ala His Gly
100 105 110

Phe Asn Ile His Tyr Lys Leu Ile Asp Pro Pro Lys Asp Ser Asp Val

```
        115                 120                 125

Tyr Met Ile Ala Pro Lys Gly Pro Gly Pro Thr Val Arg Glu Tyr Tyr
    130                 135                 140

Lys Ala Gly Gly Gly Val Pro Ala Leu Val Ala Val His Gln Asp Val
145                 150                 155                 160

Ser Gly Thr Ala Leu His Lys Ala Leu Ala Ile Ala Lys Gly Ile Gly
                165                 170                 175

Ala Thr Arg Ala Gly Val Ile Pro Thr Thr Phe Lys Glu Glu Thr Glu
            180                 185                 190

Thr Asp Leu Phe Gly Glu Gln Val Ile Leu Val Gly Gly Ile Met Glu
        195                 200                 205

Leu Met Arg Ala Ala Phe Glu Thr Leu Val Glu Glu Gly Tyr Gln Pro
    210                 215                 220

Glu Val Ala Tyr Phe Glu Thr Ile Asn Glu Leu Lys Met Leu Val Asp
225                 230                 235                 240

Leu Val Tyr Glu Lys Gly Ile Ser Gly Met Leu Lys Ala Val Ser Asp
                245                 250                 255

Thr Ala Lys Tyr Gly Gly Met Thr Val Gly Lys Phe Val Ile Asp Glu
            260                 265                 270

Ser Val Arg Lys Arg Met Lys Glu Ala Leu Gln Arg Ile Lys Ser Gly
        275                 280                 285

Lys Phe Ala Glu Glu Trp Val Glu Glu Tyr Gly Arg Gly Met Pro Thr
    290                 295                 300

Val Val Asn Gly Leu Ser Asn Val Gln Asn Ser Leu Glu Glu Lys Ile
305                 310                 315                 320

Gly Asn Gln Leu Arg Asp Leu Val Gln Lys Gly Lys Pro Lys Ser
                325                 330                 335


<210>  2
<211>  328
<212>  PRT
<213>  Pyrobaculum aerophilum

<400>  2

Met Ala Lys Ile Tyr Thr Asp Arg Glu Ala Ser Leu Glu Pro Leu Lys
1               5                   10                  15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ile Gln Gly Arg Ala Gln
            20                  25                  30

Ala Leu Asn Leu Arg Asp Ser Gly Leu Glu Val Ile Ile Gly Leu Arg
        35                  40                  45
```

```
Arg Gly Gly Lys Ser Trp Glu Leu Ala Thr Ser Glu Gly Phe Arg Val
    50                  55                  60

Tyr Glu Ile Gly Glu Ala Val Arg Lys Ala Asp Val Ile Leu Val Leu
65                  70                  75                  80

Ile Pro Asp Met Glu Gln Pro Lys Val Trp Gln Glu Gln Ile Ala Pro
                85                  90                  95

Asn Leu Lys Glu Gly Val Val Val Asp Phe Ala His Gly Phe Asn Val
            100                 105                 110

His Phe Gly Leu Ile Lys Pro Pro Lys Asn Ile Asp Val Ile Met Val
        115                 120                 125

Ala Pro Lys Ala Pro Gly Lys Ala Val Arg Glu Glu Tyr Leu Ala Gly
    130                 135                 140

Arg Gly Val Pro Ala Leu Val Ala Val Tyr Gln Asp Tyr Ser Gly Ser
145                 150                 155                 160

Ala Leu Lys Tyr Ala Leu Ala Leu Ala Lys Gly Ile Gly Ala Thr Arg
                165                 170                 175

Ala Gly Val Ile Glu Thr Thr Phe Ala Glu Glu Thr Glu Thr Asp Leu
            180                 185                 190

Ile Gly Glu Gln Ile Val Leu Val Gly Gly Leu Met Glu Leu Ile Lys
        195                 200                 205

Lys Gly Phe Glu Val Leu Val Glu Met Gly Tyr Gln Pro Glu Val Ala
    210                 215                 220

Tyr Phe Glu Val Leu Asn Glu Ala Lys Leu Ile Met Asp Leu Ile Trp
225                 230                 235                 240

Gln Arg Gly Ile Tyr Gly Met Leu Asn Gly Val Ser Asp Thr Ala Lys
                245                 250                 255

Tyr Gly Gly Leu Thr Val Gly Pro Arg Val Ile Asp Glu Asn Val Lys
            260                 265                 270

Arg Lys Met Lys Glu Ala Ala Met Arg Val Lys Ser Gly Glu Phe Ala
        275                 280                 285

Lys Glu Trp Val Glu Glu Tyr Asn Arg Gly Ala Pro Thr Leu Arg Lys
    290                 295                 300

Leu Met Glu Glu Ala Arg Thr His Pro Ile Glu Lys Val Gly Glu Glu
305                 310                 315                 320

Met Arg Lys Leu Leu Phe Gly Pro
                325
```

20 Jan 2017 2017200439

<210> 3
<211> 338
<212> PRT
<213> Ralstonia solanacearum

<400> 3

```
Met Lys Val Phe Tyr Asp Lys Asp Ala Asp Leu Ser Leu Ile Lys Gly
1               5                  10                  15

Lys Asn Val Thr Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Leu Asn Leu Asn Asp Ser Gly Val Lys Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Asn Gly Ala Ser Trp Asn Lys Ala Val Asn Ala Gly Leu Gln Val Lys
    50                  55                  60

Glu Val Ala Glu Ala Val Lys Asp Ala Asp Val Val Met Ile Leu Leu
65                  70                  75                  80

Pro Asp Glu Gln Ile Ala Asp Val Tyr Lys Asn Glu Val His Gly Asn
                85                  90                  95

Ile Lys Gln Gly Ala Ala Leu Ala Phe Ala His Gly Phe Asn Val His
            100                 105                 110

Tyr Gly Ala Val Ile Pro Arg Ala Asp Leu Asp Val Ile Met Val Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Gly Thr Tyr Ala Gln Gly Gly
    130                 135                 140

Gly Val Pro His Leu Ile Ala Val His Gln Asp Lys Ser Gly Ser Ala
145                 150                 155                 160

Arg Asp Ile Ala Leu Ser Tyr Ala Thr Ala Asn Gly Gly Gly Arg Ala
                165                 170                 175

Gly Ile Ile Glu Thr Asn Phe Arg Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Ile Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Ile Tyr Glu
225                 230                 235                 240

Gly Gly Ile Gly Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255
```

```
Gly Glu Tyr Val Thr Gly Pro Arg Val Val Thr Ala Glu Thr Lys Gln
            260                 265                 270

Ala Met Lys Gln Cys Leu His Asp Ile Gln Thr Gly Glu Tyr Ala Lys
        275                 280                 285

Ser Phe Leu Leu Glu Asn Lys Ala Gly Ala Pro Thr Leu Ile Ser Arg
    290                 295                 300

Arg Arg Leu Thr Ala Asp His Gln Ile Glu Gln Val Gly Ala Lys Leu
305                 310                 315                 320

Arg Ala Met Met Pro Trp Ile Ala Lys Asn Lys Leu Val Asp Gln Ser
                325                 330                 335

Lys Asn
```

<210> 4
<211> 338
<212> PRT
<213> Pseudomonas aeruginosa

<400> 4

```
Met Arg Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Ser
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Ala
    50                  55                  60

Asp Val Lys Thr Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Glu Glu Ile Glu Pro Asn
                85                  90                  95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160
```

```
Lys Asn Val Ala Leu Ser Tyr Ala Cys Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Cys Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Ala
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Thr Glu Gly Ala Ala Asn Tyr Pro Ser Met Thr Ala Tyr
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Pro Ile Glu Gln Ile Gly Glu Lys Leu
305                 310                 315                 320

Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Ser
                325                 330                 335

Lys Asn
```

<210> 5
<211> 338
<212> PRT
<213> Pseudomonas fluorescens

<400> 5

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
```

65 70 75 80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
85 90 95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
100 105 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145 150 155 160

Lys Asn Val Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Thr
165 170 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
210 215 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
245 250 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
260 265 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
275 280 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
290 295 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305 310 315 320

Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
325 330 335

Lys Asn

<210> 6

20 Jan 2017 2017200439

<211> 338
<212> PRT
<213> Spinacia oleracea

<400> 6

```
Met Arg Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                  10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Ser
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Ala
    50                  55                  60

Asp Val Lys Thr Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Glu Glu Ile Glu Pro Asn
                85                  90                  95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Cys Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Cys Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Ala
            260                 265                 270
```

```
Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Thr Glu Gly Ala Ala Asn Tyr Pro Ser Met Thr Ala Tyr
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Pro Ile Glu Gln Ile Gly Glu Lys Leu
305                 310                 315                 320

Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Ser
                325                 330                 335

Lys Asn
```

<210> 7
<211> 331
<212> PRT
<213> Natronomonas pharaanis

<400> 7

```
Met Thr Asp Ala Thr Ile Tyr Tyr Asp Asp Asp Ala Glu Ser Thr Val
1               5                   10                  15

Leu Asp Asp Lys Thr Val Ala Val Leu Gly Tyr Gly Ser Gln Gly His
            20                  25                  30

Ala His Ala Gln Asn Leu Asp Asp Ser Gly Val Asp Val Val Val Gly
        35                  40                  45

Leu Arg Glu Asp Ser Ser Ser Arg Ser Ala Ala Glu Ala Asp Gly Leu
    50                  55                  60

Asp Val Ala Thr Pro Arg Gly Ala Ala Glu Gln Ala Asp Leu Val Ser
65                  70                  75                  80

Val Leu Val Pro Asp Thr Val Gln Pro Ala Val Tyr Glu Gln Ile Glu
                85                  90                  95

Asp Val Leu Gln Pro Gly Asp Thr Leu Gln Phe Ala His Gly Phe Asn
            100                 105                 110

Ile His Tyr Gly Gln Ile Glu Pro Ser Glu Asp Val Asn Val Thr Met
        115                 120                 125

Val Ala Pro Lys Ser Pro Gly His Leu Val Arg Arg Asn Tyr Glu Asn
    130                 135                 140

Asp Glu Gly Thr Pro Gly Leu Leu Ala Val Tyr Gln Asp Pro Ser Gly
145                 150                 155                 160

Glu Ala His Asp Leu Gly Leu Ala Tyr Ala Lys Ala Ile Gly Cys Thr
                165                 170                 175
```

```
Arg Ala Gly Val Val Glu Thr Thr Phe Arg Glu Glu Thr Glu Thr Asp
            180                 185                 190

Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Thr Ser Leu Val
        195                 200                 205

Lys Thr Gly Tyr Glu Thr Leu Val Asp Ala Gly Tyr Ser Pro Glu Met
    210                 215                 220

Ala Tyr Phe Glu Cys Leu Asn Glu Leu Lys Leu Ile Val Asp Leu Met
225                 230                 235                 240

Tyr Glu Gly Gly Asn Ser Glu Met Trp Asp Ser Val Ser Asp Thr Ala
                245                 250                 255

Glu Tyr Gly Gly Leu Thr Arg Gly Asp Arg Ile Val Asp Asp His Ala
            260                 265                 270

Arg Glu Lys Met Glu Glu Val Leu Glu Glu Val Gln Asn Gly Thr Phe
        275                 280                 285

Ala Arg Glu Trp Ile Ser Glu Asn Gln Ala Gly Arg Pro Ser Tyr Lys
    290                 295                 300

Gln Leu Arg Ala Ala Glu Lys Asn His Asp Ile Glu Ala Val Gly Glu
305                 310                 315                 320

Asp Leu Arg Ala Leu Phe Ala Trp Gly Asp Asp
                325                 330
```

<210> 8
<211> 342
<212> PRT
<213> Bacillus subtilis

<400> 8

```
Met Val Lys Val Tyr Tyr Asn Gly Asp Ile Lys Glu Asn Val Leu Ala
1               5                   10                  15

Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Leu Lys Glu Ser Gly Val Asp Val Ile Val Gly Val Arg
        35                  40                  45

Gln Gly Lys Ser Phe Thr Gln Ala Gln Glu Asp Gly His Lys Val Phe
    50                  55                  60

Ser Val Lys Glu Ala Ala Ala Gln Ala Glu Ile Ile Met Val Leu Leu
65                  70                  75                  80

Pro Asp Glu Gln Gln Gln Lys Val Tyr Glu Ala Glu Ile Lys Asp Glu
                85                  90                  95
```

```
Leu Thr Ala Gly Lys Ser Leu Val Phe Ala His Gly Phe Asn Val His
            100                 105                 110

Phe His Gln Ile Val Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Gln Gly Ala
    130                 135                 140

Gly Val Pro Ala Leu Phe Ala Ile Tyr Gln Asp Val Thr Gly Glu Ala
145                 150                 155                 160

Arg Asp Lys Ala Leu Ala Tyr Ala Lys Gly Ile Gly Gly Ala Arg Ala
                165                 170                 175

Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Glu Gly Leu Ala Gly Met Arg Tyr Ser Ile Ser Asp Thr Ala Gln Trp
                245                 250                 255

Gly Asp Phe Val Ser Gly Pro Arg Val Val Asp Ala Lys Val Lys Glu
            260                 265                 270

Ser Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys
        275                 280                 285

Glu Trp Ile Val Glu Asn Gln Val Asn Arg Pro Arg Phe Asn Ala Ile
    290                 295                 300

Asn Ala Ser Glu Asn Glu His Gln Ile Glu Val Val Gly Arg Lys Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Val Lys Gln Gly Lys Lys Lys Glu Ala Val
                325                 330                 335

Val Ser Val Ala Gln Asn
            340
```

<210> 9
<211> 338
<212> PRT
<213> Corynebacterium glutamicun

<400> 9

```
Met Ala Ile Glu Leu Leu Tyr Asp Ala Asp Ala Asp Leu Ser Leu Ile
```

```
1               5                   10                  15

Gln Gly Arg Lys Val Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala
            20                  25                  30

His Ser Gln Asn Leu Arg Asp Ser Gly Val Glu Val Val Ile Gly Leu
        35                  40                  45

Arg Glu Gly Ser Lys Ser Ala Glu Lys Ala Lys Glu Ala Gly Phe Glu
    50                  55                  60

Val Lys Thr Thr Ala Glu Ala Ala Ala Trp Ala Asp Val Ile Met Leu
65                  70                  75                  80

Leu Ala Pro Asp Thr Ser Gln Ala Glu Ile Phe Thr Asn Asp Ile Glu
                85                  90                  95

Pro Asn Leu Asn Ala Gly Asp Ala Leu Leu Phe Gly His Gly Leu Asn
            100                 105                 110

Ile His Phe Asp Leu Ile Lys Pro Ala Asp Asp Ile Ile Val Gly Met
        115                 120                 125

Val Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Gln Phe Val Asp
    130                 135                 140

Gly Lys Gly Val Pro Cys Leu Ile Ala Val Asp Gln Asp Pro Thr Gly
145                 150                 155                 160

Thr Ala Gln Ala Leu Thr Leu Ser Tyr Ala Ala Ala Ile Gly Gly Ala
                165                 170                 175

Arg Ala Gly Val Ile Pro Thr Thr Phe Glu Ala Glu Thr Val Thr Asp
            180                 185                 190

Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Glu Glu Leu Val
        195                 200                 205

Lys Val Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Glu Pro Glu Met
    210                 215                 220

Ala Tyr Phe Glu Val Leu His Glu Leu Lys Leu Ile Val Asp Leu Met
225                 230                 235                 240

Phe Glu Gly Gly Ile Ser Asn Met Asn Tyr Ser Val Ser Asp Thr Ala
                245                 250                 255

Glu Phe Gly Gly Tyr Leu Ser Gly Pro Arg Val Ile Asp Ala Asp Thr
            260                 265                 270

Lys Ser Arg Met Lys Asp Ile Leu Thr Asp Ile Gln Asp Gly Thr Phe
        275                 280                 285

Thr Lys Arg Leu Ile Ala Asn Val Glu Asn Gly Asn Thr Glu Leu Glu
```

290 295 300

Gly Leu Arg Ala Ser Tyr Asn Asn His Pro Ile Glu Glu Thr Gly Ala
305 310 315 320

Lys Leu Arg Asp Leu Met Ser Trp Val Lys Val Asp Ala Arg Ala Glu
325 330 335

Thr Ala

<210> 10
<211> 339
<212> PRT
<213> Phaeospririlum molischianum

<220>
<221> misc_feature
<222> (310)..(310)
<223> Xaa can be any naturally occurring amino acid

<400> 10

Met Arg Val Tyr Tyr Asp Arg Asp Ala Asp Val Asn Leu Ile Lys Ser
1 5 10 15

Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Val
20 25 30

Leu Asn Leu Arg Asp Ser Gly Val Lys Asp Val Ala Val Ala Leu Arg
35 40 45

Pro Gly Ser Ala Ser Ile Lys Lys Ala Glu Ala Glu Gly Leu Lys Val
50 55 60

Leu Thr Pro Ala Glu Ala Ala Ala Trp Ala Asp Val Val Met Ile Leu
65 70 75 80

Thr Pro Asp Glu Leu Gln Ala Asp Leu Tyr Lys Ser Glu Leu Ala Ala
85 90 95

Asn Leu Lys Pro Gly Ala Ala Leu Val Phe Ala His Gly Leu Ala Ile
100 105 110

His Phe Lys Leu Ile Glu Ala Arg Ala Asp Leu Asp Val Phe Met Val
115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Gly Glu Tyr Leu Lys Gly
130 135 140

Gly Gly Val Pro Cys Leu Val Ala Val Ala Gln Asn Pro Thr Gly Asn
145 150 155 160

Ala Leu Glu Leu Ala Leu Ser Tyr Ala Ser Ala Ile Gly Gly Gly Arg
165 170 175

```
Ser Gly Ile Ile Glu Thr Thr Phe Arg Glu Glu Cys Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Val Val Leu Cys Gly Gly Leu Ser Lys Leu Ile Gln
        195                 200                 205

Tyr Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala
    210                 215                 220

Tyr Phe Glu Cys Leu His Glu Val Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Val Thr Gly Ser Arg Ile Ile Thr Glu Ala Thr Lys
            260                 265                 270

Ala Glu Met Lys Arg Val Leu Ala Asp Ile Gln Ser Gly Arg Phe Val
        275                 280                 285

Arg Asp Trp Met Leu Glu Cys Lys Ala Gly Gln Pro Ser Phe Lys Ala
    290                 295                 300

Thr Arg Arg Ile Gln Xaa Glu His Val Ile Glu Val Val Gly Glu Lys
305                 310                 315                 320

Leu Arg Gly Met Met Pro Trp Ile Ser Lys Asn Lys Leu Val Asp Lys
                325                 330                 335

Ala Arg Asn
```

<210> 11
<211> 339
<212> PRT
<213> Zymomonas mobilis

<400> 11

```
Met Lys Val Tyr Tyr Asp Ser Asp Ala Asp Leu Gly Leu Ile Lys Ser
1               5                   10                  15

Lys Lys Ile Ala Ile Leu Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Gln Asn Leu Arg Asp Ser Gly Val Ala Glu Val Ala Ile Ala Leu Arg
        35                  40                  45

Pro Asp Ser Ala Ser Val Lys Lys Ala Gln Asp Ala Gly Phe Lys Val
    50                  55                  60

Leu Thr Asn Ala Glu Ala Ala Lys Trp Ala Asp Ile Leu Met Ile Leu
65                  70                  75                  80

Ala Pro Asp Glu His Gln Ala Ala Ile Tyr Ala Glu Asp Leu Lys Asp
```

```
                85                  90                  95

Asn Leu Arg Pro Gly Ser Ala Ile Ala Phe Ala His Gly Leu Asn Ile
            100                 105                 110

His Phe Gly Leu Ile Glu Pro Arg Lys Asp Ile Asp Val Phe Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Val Arg Gly
    130                 135                 140

Gly Gly Val Pro Cys Leu Val Ala Val Asp Gln Asp Ala Ser Gly Asn
145                 150                 155                 160

Ala His Asp Ile Ala Leu Ala Tyr Ala Ser Gly Ile Gly Gly Gly Arg
                165                 170                 175

Ser Gly Val Ile Glu Thr Thr Phe Arg Glu Glu Val Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Thr
        195                 200                 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala
    210                 215                 220

Phe Phe Glu Cys Met His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Ala Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Ile Val Ser Gly Pro Arg Val Ile Asn Glu Glu Ser Lys
            260                 265                 270

Lys Ala Met Lys Ala Ile Leu Asp Asp Ile Gln Ser Gly Arg Phe Val
        275                 280                 285

Ser Lys Phe Val Leu Asp Asn Arg Ala Gly Gln Pro Glu Leu Lys Ala
    290                 295                 300

Ala Arg Lys Arg Met Ala Ala His Pro Ile Glu Gln Val Gly Ala Arg
305                 310                 315                 320

Leu Arg Lys Met Met Pro Trp Ile Ala Ser Asn Lys Leu Val Asp Lys
                325                 330                 335

Ala Arg Asn
```

<210> 12
<211> 338
<212> PRT
<213> Alkalilimnicola ehrlichei

<400> 12

Met Gln Val Tyr Tyr Asp Lys Asp Ala Asp Leu Ser Ile Ile Gln Gly
1 5 10 15

Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30

Asn Asn Leu Lys Glu Ser Gly Val Asp Val Val Val Gly Leu Arg Glu
35 40 45

Gly Ser Ser Ser Ala Ala Lys Ala Gln Lys Ala Gly Leu Ala Val Ala
50 55 60

Ser Ile Glu Asp Ala Ala Ala Gln Ala Asp Val Val Met Ile Leu Ala
65 70 75 80

Pro Asp Glu His Gln Ala Val Ile Tyr His Asn Gln Ile Ala Pro Asn
85 90 95

Val Lys Pro Gly Ala Ala Ile Ala Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Gly Gln Ile Gln Pro Ala Ala Asp Leu Asp Val Ile Met Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Ser Thr Tyr Val Glu Gly Gly
130 135 140

Gly Val Pro Ser Leu Ile Ala Ile His Gln Asp Ala Thr Gly Lys Ala
145 150 155 160

Lys Asp Ile Ala Leu Ser Tyr Ala Ser Ala Asn Gly Gly Gly Arg Ala
165 170 175

Gly Val Ile Glu Thr Ser Phe Arg Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile Thr Ser Leu Ile Gln Ala
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
210 215 220

Phe Glu Cys Leu His Glu Thr Lys Leu Ile Val Asp Leu Leu Tyr Gln
225 230 235 240

Gly Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
245 250 255

Gly Asp Phe Thr Arg Gly Pro Arg Val Ile Asn Glu Glu Ser Arg Glu
260 265 270

Ala Met Arg Glu Ile Leu Ala Glu Ile Gln Glu Gly Glu Phe Ala Arg
275 280 285

```
Glu Phe Val Leu Glu Asn Gln Ala Gly Cys Pro Thr Leu Thr Ala Arg
    290                 295                 300

Arg Arg Leu Ala Ala Glu His Glu Ile Glu Val Val Gly Glu Arg Leu
305                 310                 315                 320

Arg Gly Met Met Pro Trp Ile Asn Ala Asn Lys Leu Val Asp Lys Asp
                325                 330                 335

Lys Asn
```

<210> 13
<211> 340
<212> PRT
<213> Campylobacter lari

<400> 13

```
Met Ala Val Ser Ile Tyr Tyr Asp Lys Asp Cys Asp Ile Asn Leu Ile
1               5                   10                  15

Lys Ser Lys Lys Val Ala Ile Ile Gly Phe Gly Ser Gln Gly His Ala
            20                  25                  30

His Ala Met Asn Leu Arg Asp Ser Gly Val Glu Val Ile Ile Gly Leu
        35                  40                  45

Lys Glu Gly Gly Gln Ser Trp Ala Lys Ala Gln Lys Ala Asn Phe Ile
    50                  55                  60

Val Lys Ser Val Lys Glu Ala Thr Lys Glu Ala Asp Leu Ile Met Ile
65                  70                  75                  80

Leu Ala Pro Asp Glu Ile Gln Ser Glu Ile Phe Asn Glu Glu Ile Lys
                85                  90                  95

Pro Glu Leu Lys Ala Gly Lys Thr Leu Ala Phe Ala His Gly Phe Asn
            100                 105                 110

Ile His Tyr Gly Gln Ile Val Ala Pro Lys Gly Ile Asp Val Ile Met
        115                 120                 125

Ile Ala Pro Lys Ala Pro Gly His Thr Val Arg His Glu Phe Ser Ile
    130                 135                 140

Gly Gly Gly Thr Pro Cys Leu Ile Ala Ile His Gln Asp Glu Ser Lys
145                 150                 155                 160

Asn Ala Lys Asn Leu Ala Leu Ser Tyr Ala Ser Ala Ile Gly Gly Gly
                165                 170                 175

Arg Thr Gly Ile Ile Glu Thr Thr Phe Lys Ala Glu Thr Glu Thr Asp
            180                 185                 190
```

```
Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Ser Ala Leu Ile
        195                 200                 205

Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Glu Pro Glu Met
    210                 215                 220

Ala Tyr Phe Glu Cys Leu His Glu Met Lys Leu Ile Val Asp Leu Ile
225                 230                 235                 240

Tyr Gln Gly Gly Ile Ala Asp Met Arg Tyr Ser Val Ser Asn Thr Ala
                245                 250                 255

Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Lys Glu Thr
            260                 265                 270

Lys Glu Ala Met Lys Gly Val Leu Lys Asp Ile Gln Asn Gly Ser Phe
        275                 280                 285

Ala Lys Asp Phe Ile Leu Glu Arg Arg Ala Asn Phe Ala Arg Met His
    290                 295                 300

Ala Glu Arg Lys Leu Met Asn Asp Ser Leu Ile Glu Lys Thr Gly Arg
305                 310                 315                 320

Glu Leu Arg Ala Met Met Pro Trp Ile Ser Ala Lys Lys Leu Val Asp
                325                 330                 335

Lys Asp Lys Asn
            340
```

<210> 14
<211> 338
<212> PRT
<213> Marinobacter aquaeolei

<400> 14

```
Met Gln Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Leu Gly Phe Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Val Val Gly Leu Arg Ala
        35                  40                  45

Gly Ser Ser Ser Ile Ala Lys Ala Glu Ala Tyr Gly Leu Lys Thr Ser
    50                  55                  60

Asp Val Ala Ser Ala Val Ala Ser Ala Asp Val Val Met Val Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Ala Gln Leu Tyr Arg Glu Glu Ile Glu Pro Asn
                85                  90                  95
```

```
Leu Lys Gln Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Ile Val Pro Arg Lys Asp Leu Asp Val Ile Met Val Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Thr Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Phe Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Ile Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Glu
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Ser Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Leu Asn Tyr Pro Ser Met Thr Ala Arg
    290                 295                 300

Arg Arg Gln Asn Ala Ala His Glu Ile Glu Thr Val Gly Glu Lys Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ser Ala Asn Lys Ile Val Asp Lys Asp
                325                 330                 335

Lys Asn
```

<210> 15
<211> 338
<212> PRT
<213> Psychrobacter arcticus

<400> 15

```
Met Asn Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Val Gln Gly
```

```
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Leu Asn Leu Gln Asp Ser Asn Val Asp Val Thr Val Gly Leu Arg Ala
        35                  40                  45

Asp Ser Gly Ser Trp Lys Lys Ala Glu Asn Ala Gly Leu Lys Val Ala
    50                  55                  60

Glu Val Glu Glu Ala Val Lys Ala Ala Asp Ile Ile Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Lys Glu Leu Tyr Asn Asp Val Ile Glu Pro Asn
                85                  90                  95

Ile Lys Gln Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Ile Pro Arg Ser Asp Leu Asp Val Ile Met Val Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Ala Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Gln Ala
145                 150                 155                 160

Lys Gln Leu Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Ser
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Met
        195                 200                 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asp Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Glu
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Ser Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Asn Tyr Pro Ser Met Thr Ala Arg
```

290 295 300

Arg Arg Asn Asn Ala Glu His Gln Ile Glu Ile Thr Gly Ala Lys Leu
305 310 315 320

Arg Gly Met Met Pro Trp Ile Gly Gly Asn Lys Ile Ile Asp Lys Asp
325 330 335

Lys Asn

<210> 16
<211> 338
<212> PRT
<213> Hahella chejuensis

<400> 16

Met Gln Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30

Asn Asn Leu Lys Asp Ser Gly Val Asp Val Cys Val Gly Leu Arg Lys
35 40 45

Gly Ser Gly Ser Trp Ala Lys Ala Glu Asn Ala Gly Leu Ala Val Lys
50 55 60

Glu Val Ala Glu Ala Val Ala Gly Ala Asp Val Val Met Ile Leu Thr
65 70 75 80

Pro Asp Glu Phe Gln Ala Gln Leu Tyr Lys Ser Glu Ile Glu Pro Asn
85 90 95

Leu Lys Ser Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
100 105 110

Tyr Asn Gln Ile Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
130 135 140

Gly Ile Pro Asp Leu Ile Ala Ile Phe Gln Asp Ala Ser Gly Ser Ala
145 150 155 160

Lys Asp Leu Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
165 170 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Ala
195 200 205

```
Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Asp Gln Ser Arg Ala
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ala Glu Gly Ala His Asn Tyr Pro Ser Met Thr Ala Tyr
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Pro Ile Glu Gln Val Gly Glu Lys Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ala Ser Asn Lys Ile Val Asp Lys Ser
                325                 330                 335

Lys Asn
```

<210> 17
<211> 338
<212> PRT
<213> Thiobacillus denitrificans

<400> 17

```
Met Lys Val Tyr Tyr Asp Lys Asp Ala Asp Leu Ser Leu Ile Lys Gln
1               5                   10                  15

Arg Lys Val Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Asn Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Ala Leu Arg Pro
        35                  40                  45

Gly Ser Ala Ser Ala Lys Lys Ala Glu Asn Ala Gly Leu Thr Val Lys
    50                  55                  60

Ser Val Pro Glu Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Ser Arg Leu Tyr Arg Asp Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Gln Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
            100                 105                 110
```

20 Jan 2017 2017200439

```
Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Lys Ala
145                 150                 155                 160

Lys Glu Thr Ala Leu Ser Tyr Ala Ser Ala Ile Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Asp Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Val Lys Val Ile Asn Glu Gln Ser Arg Ala
            260                 265                 270

Ala Met Lys Glu Cys Leu Ala Asn Ile Gln Asn Gly Ala Tyr Ala Lys
        275                 280                 285

Arg Phe Ile Leu Glu Gly Gln Ala Asn Tyr Pro Glu Met Thr Ala Trp
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gln Ile Glu Val Val Gly Ala Lys Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ala Ala Asn Lys Leu Val Asp His Ser
                325                 330                 335

Lys Asn
```

<210> 18
<211> 338
<212> PRT
<213> Azotobacter vinelandii

<400> 18

```
Met Lys Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Ser
1               5                   10                  15
```

2017200439 20 Jan 2017

```
Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Tyr Val Gly Leu Arg Ala
        35                  40                  45

Gly Ser Ala Ser Val Ala Lys Ala Glu Ala His Gly Leu Thr Val Lys
    50                  55                  60

Ser Val Lys Asp Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Arg Leu Tyr Lys Asp Glu Ile Glu Pro Asn
                85                  90                  95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Arg Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Val Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Leu Ala Leu Ser Tyr Ala Cys Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Cys Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Phe Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Gln Ser Arg Gln
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Thr Glu Gly Ala Ala Asn Tyr Pro Ser Met Thr Ala Tyr
    290                 295                 300
```

```
Arg Arg Asn Asn Ala Ala His Gln Ile Glu Val Val Gly Glu Lys Leu
305                 310                 315                 320

Arg Thr Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Thr
                325                 330                 335

Lys Asn
```

<210> 19
<211> 338
<212> PRT
<213> Pseudomonas syringae

<400> 19

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Ser Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Val Glu Pro Asn
                85                  90                  95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Val Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
```

```
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Lys Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Asp
                325                 330                 335

Lys Asn
```

<210> 20
<211> 338
<212> PRT
<213> Pseudomonas syringae

<400> 20

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Ser Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Val Glu Pro Asn
                85                  90                  95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125
```

```
Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Val Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Thr Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Glu His Gly Ile Glu Val Ile Gly Glu Lys Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Asp
                325                 330                 335

Lys Asn
```

<210> 21
<211> 338
<212> PRT
<213> Pseudomonas putida

<400> 21

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30
```

```
Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Ala
    50                  55                  60

Asp Val Ala Thr Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Ala Leu Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ser Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Glu Ser Arg Lys
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Asn Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305                 310                 315                 320
```

```
Arg Ser Met Met Pro Trp Ile Ser Ala Asn Lys Ile Val Asp Lys Thr
                325                 330                 335

Lys Asn
```

<210> 22
<211> 338
<212> PRT
<213> Pseudomonas entomophila

<400> 22

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Ile Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Thr Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Gln Leu Tyr Lys Gln Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220
```

```
Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Glu Glu Ser Arg Lys
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Asn Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Ser Ala Asn Lys Ile Val Asp Lys Thr
                325                 330                 335

Lys Asn
```

<210> 23
<211> 338
<212> PRT
<213> Pseudomonas mendocina

<400> 23

```
Met Lys Val Tyr Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Ile Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Ser Ala Val Ala Ala Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80

Pro Asp Glu Phe Gln Gly Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Thr Glu Phe Val Lys Gly Gly
```

```
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Val Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ser Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Val Ile Gly Glu Gln Leu
305                 310                 315                 320

Arg Ala Met Met Pro Trp Ile Ala Ala Asn Lys Ile Val Asp Lys Thr
                325                 330                 335

Lys Asn


<210>  24
<211>  336
<212>  PRT
<213>  Bacillus cereus

<400>  24

Met Ala Lys Val Tyr Tyr Glu Lys Asp Val Thr Val Asn Val Leu Lys
1               5                   10                  15

Glu Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Gln Asn Leu Arg Asp Asn Gly Phe Asp Val Val Val Gly Leu Arg
        35                  40                  45
```

```
Lys Gly Lys Ser Trp Asp Lys Ala Lys Glu Asp Gly Phe Ser Val Tyr
    50                  55                  60

Thr Val Ala Glu Ala Ala Lys Gln Ala Asp Val Val Met Ile Leu Leu
65                  70                  75                  80

Pro Asp Glu Leu Gln Pro Glu Val Tyr Glu Ala Glu Ile Ala Pro Asn
                85                  90                  95

Leu Gln Ala Gly Asn Ser Leu Val Phe Ala His Gly Phe Asn Val His
            100                 105                 110

Phe Asp Gln Val Lys Pro Pro Ala Asn Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Ser Glu Gly Gly
    130                 135                 140

Ala Val Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Val Ala
145                 150                 155                 160

Thr Glu Lys Ala Leu Ser Tyr Ala Asp Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Thr Ala Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Asp Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Leu Glu Asn Met Arg Tyr Ser Val Ser Asp Thr Ala Gln Trp
                245                 250                 255

Gly Asp Phe Val Ser Gly Pro Arg Val Val Thr Glu Asp Thr Lys Lys
            260                 265                 270

Ala Met Gly Thr Val Leu Ala Glu Ile Gln Asp Gly Thr Phe Ala Arg
        275                 280                 285

Gly Trp Ile Ala Glu His Lys Ala Gly Arg Pro Asn Phe His Ala Thr
    290                 295                 300

Asn Glu Lys Glu Asn Glu His Glu Ile Glu Val Val Gly Arg Lys Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Val Gln Pro Arg Val Lys Val Gly Met Lys
                325                 330                 335
```

<210> 25
<211> 335
<212> PRT
<213> Bacillus cereus

<400> 25

Met Lys Thr Tyr Tyr Glu Lys Asp Ala Asn Val Glu Leu Leu Lys Gly
1 5 10 15

Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
20 25 30

Gln Asn Leu Arg Asp Ser Gly Val Glu Val Val Val Gly Val Arg Pro
35 40 45

Gly Lys Ser Phe Glu Val Ala Lys Thr Asp Gly Phe Glu Val Met Ser
50 55 60

Val Ser Glu Ala Val Arg Thr Ala Gln Val Val Gln Met Leu Leu Pro
65 70 75 80

Asp Glu Gln Gln Ala His Val Tyr Lys Ala Gly Val Glu Glu Asn Leu
85 90 95

Arg Glu Gly Gln Met Leu Leu Phe Ser His Gly Phe Asn Ile His Phe
100 105 110

Gly Gln Ile Asn Pro Pro Ser Tyr Val Asp Val Ala Met Val Ala Pro
115 120 125

Lys Ser Pro Gly His Leu Val Arg Arg Val Phe Gln Glu Gly Asn Gly
130 135 140

Val Pro Ala Leu Val Ala Val His Gln Asp Ala Thr Gly Thr Ala Leu
145 150 155 160

His Val Ala Leu Ala Tyr Ala Lys Gly Val Gly Cys Thr Arg Ala Gly
165 170 175

Val Ile Glu Thr Thr Phe Gln Glu Glu Thr Glu Thr Asp Leu Phe Gly
180 185 190

Glu Gln Thr Val Leu Cys Gly Gly Val Thr Ala Leu Val Lys Ala Gly
195 200 205

Phe Glu Thr Leu Thr Glu Gly Gly Tyr Arg Pro Glu Ile Ala Tyr Phe
210 215 220

Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu Gly
225 230 235 240

Gly Leu Thr Asn Met Arg His Ser Ile Ser Asp Thr Ala Glu Phe Gly
245 250 255

20 Jan 2017 2017200439

```
Asp Tyr Val Thr Gly Ser Arg Ile Val Thr Asp Glu Thr Lys Lys Glu
            260                 265                 270

Met Lys Arg Val Leu Thr Glu Ile Gln Gln Gly Glu Phe Ala Lys Lys
        275                 280                 285

Trp Ile Leu Glu Asn Gln Ala Gly Arg Pro Thr Tyr Asn Ala Met Lys
    290                 295                 300

Lys Ala Glu Gln Asn His Gln Leu Glu Lys Val Gly Ala Glu Leu Arg
305                 310                 315                 320

Glu Met Met Ser Trp Ile Asp Ala Pro Lys Glu Leu Val Lys Lys
                325                 330                 335
```

<210> 26
<211> 1032
<212> DNA
<213> Anaerostipes caccae

<400> 26

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggttctaa gtcctggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca      960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg     1020

attaacaact ga                                                         1032
```

<210> 27
<211> 343
<212> PRT
<213> Anaerostipes caccae

<400> 27

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
```

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

20 Jan 2017 2017200439

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 28
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AB1D1 variant

<400> 28
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgaa gtcctggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 29
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AB1D1 variant

<400> 29

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Ser Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 30
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-495B5 variant

<400> 30
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa gcattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 31
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI 495B5 variant

<400> 31

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys His Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 32
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AB1D3 variant

<400> 32
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 33
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AB1D3 variant

<400> 33

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 34
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AB1G9 variant

<400> 34
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 35
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AB1G9 variant

<400> 35

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 36
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-variant AO7B5

<400> 36
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc tgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 37
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AO7B5 variant

<400> 37

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 38
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AO7A9 variant

<400> 38
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgacgcc tgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 39
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AO7A9 variant

<400> 39

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Thr Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 40
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AO7H8 variant

<400> 40
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 41
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AO7H8 variant

<400> 41

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 42
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AO7D8 variant

<400> 42
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc tgtatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga tcggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 43
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AO7D8 variant

<400> 43

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Cys Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 44
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI- AO7F7 variant

<400> 44
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga gctgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 45
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AO7F7 variant

<400> 45

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ala Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 46
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AO7H7 variant

<400> 46
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga agtgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 47
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae-KARI A07H7 variant

<400> 47

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 48
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae AS6F1 variant

<400> 48
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 49
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI-AS6F1 variant

<400> 49

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 50
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae AWG4 variant

<400> 50
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 51
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AW4G variant

<400> 51

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 52
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae AWB9 variant

<400> 52
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gcaatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 53
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWB9 variant

<400> 53

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Ala Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 54
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWC1 variant

<400> 54
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 55
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWC1 variant

<400> 55

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 56
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWF4 variant

<400> 56
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtag catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 57
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWF4 variant

<400> 57

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Ser Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 58
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWF6 variant

<400> 58
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga gacgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 59
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWF6 variant

<400> 59

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Asp Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 60
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWH3 variant

<400> 60
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggcgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 61
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWH3 variant

<400> 61

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 62
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWD6 variant

<400> 62
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gttatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga tcggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 63
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWD6 variant

<400> 63

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Val Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 64
<211> 1029
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWD10 variant

<400> 64
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgctaa ggattggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc gcaatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga tgcgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 65
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI AWD10 variant

<400> 65

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn

```
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Ala Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Cys Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
```

290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 66
<211> 1002
<212> DNA
<213> Staphylococcus capitis

<400> 66
atgacaaaag tttattatga tcaatcagtt gaaaaagatg ctttacaagg taaaaaaatt 60

gcagtaattg gttacggttc acaaggacat gcacacgctc aaaatttaaa agacaatggt 120

tatgatgtag tcgttggtat tcgacctggt cattctttcg atagagctaa agaagatgga 180

tttgatgttt atccagtgaa tgaggctact aaacaagcag atgtagtaat gattttatta 240

cctgatgaaa tccaaggtaa tgtatataaa aatgaaattg aacctaactt agaagcaggc 300

aatgcattag cttttgcgca cggatttaac attcactttg gagttattca acccctgct 360

gatgtagatg tattcttagt agcacctaaa ggtcctggtc atttagtaag acgtaccttt 420

gtagaaggta ctgctgtacc agcattattc ggtgtacaac aagatgattc aggtcatgca 480

cgtgatatcg cactaagcta tgccaaaggt attggtgcaa cacgtgctgg tgttattgaa 540

actacattca aagaagaaac tgttactgat ttattcggtg aacaagctgt actttgcggg 600

ggtatccata aattaattca aagtggtttt gaaactttag tagaagctgg gtatcaaaaa 660

gagttagcat acttcgaagt attacatgaa atgaaattaa tcgttgattt aatgtacgaa 720

ggcggtatgg aaaacgttcg ctactctatt tcaaatacag ctgaatttgg tgattatgta 780

tctggaccac gtgttatcac tccagaagtt aaaaacaaca tgaaaactgt attagaagat 840

attcaaaacg gtaactttgc taatcgtttt gttaaagata atgaaaatgg ctttaaagaa 900

ttttaccaat tacgtgaaca acaacatggt catgaaatcg aagcggttgg tcgtgaatta 960

agaaaaatga tgccatttat taaagctaaa agtattcaaa aa 1002

<210> 67
<211> 334
<212> PRT
<213> Staphylococcus capitis

<400> 67

Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln
1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Val Gly Ile Arg

```
        35                  40                  45

Pro Gly His Ser Phe Asp Arg Ala Lys Glu Asp Gly Phe Asp Val Tyr
    50                  55                  60

Pro Val Asn Glu Ala Thr Lys Gln Ala Asp Val Val Met Ile Leu Leu
65                  70                  75                  80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Thr
    130                 135                 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Asp Ser Gly His Ala
145                 150                 155                 160

Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Val Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Glu Val Lys Asn
            260                 265                 270

Asn Met Lys Thr Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ala Asn
        275                 280                 285

Arg Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Gln Leu
    290                 295                 300

Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
305                 310                 315                 320

Arg Lys Met Met Pro Phe Ile Lys Ala Lys Ser Ile Gln Lys
```

325 330

<210> 68
<211> 1002
<212> DNA
<213> Staphylococcus epidermidis

<400> 68
atgacaaaag tttattatga tggatctgta actaaagacg cattacaagg taaaaaaatt 60

gcagtaatcg gatatggctc acaaggacat gcccacgcgc aaaacttaaa agatagtggc 120

tatgatgtca tcatcggtat ccgtcctgga cactcatacg ataaagctaa aaaagatggt 180

ttcgatgtgt atccagtaag tgaagctaca aaacaagcag atgtcgttat ggttttatta 240

ccagatgaaa tccaaggtaa cgtatacaaa aatgaaatcg cacctaactt agaagctggt 300

aatgccctag ctttcgccca cggttttaac attcactttg gtgttattca acctcctgct 360

gatgtagatg tatttttagt agctcctaaa ggaccaggtc accttgtaag acgtacattc 420

gttgaaggca gtgcagtacc ttcattattt ggtgttcaac aagatgctac tggaaatgca 480

cgtgatatcg ctttaagcta tgccaaaggt attggtgcta cacgtgctgg tgttattgaa 540

acaacattta aagaagaaac tgaaactgat ttattcggtg agcaagctgt actttgtggt 600

ggtatccata aattaattca aagtggtttt gaaactttag tagaagctgg ttatcaaaaa 660

gaattagctt acttcgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa 720

ggcggtatgg aaaacgttcg ctattctatt tcaaatactg ccgaatttgg tgattatgta 780

tctggaccac gtgtcatcac tccagaagtt aaaaacaaca tgaaaactgt attagaagat 840

atccaaaacg gtaactttgc taatcgtttt gttaaagata atgaaaatgg ctttaaagaa 900

ttctatcaat tacgtgaaca acaacatggt catgaaattg aagcggtagg tcgtgagctt 960

cgtgaaatga tgccatttat taaatctaaa agcattgaaa aa 1002

<210> 69
<211> 334
<212> PRT
<213> Staphylococcus epidermidis

<400> 69

Met Thr Lys Val Tyr Tyr Asp Gly Ser Val Thr Lys Asp Ala Leu Gln
1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Ser Gly Tyr Asp Val Ile Ile Gly Ile Arg
35 40 45

Pro Gly His Ser Tyr Asp Lys Ala Lys Lys Asp Gly Phe Asp Val Tyr
50 55 60

Pro Val Ser Glu Ala Thr Lys Gln Ala Asp Val Val Met Val Leu Leu
65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Ala Pro Asn

```
                85                  90                  95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser
    130                 135                 140

Ala Val Pro Ser Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Asn Ala
145                 150                 155                 160

Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Glu Val Lys Asn
            260                 265                 270

Asn Met Lys Thr Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ala Asn
        275                 280                 285

Arg Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Gln Leu
    290                 295                 300

Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
                325                 330
```

<210> 70
<211> 1002
<212> DNA
<213> Staphylococcus aureus

<400> 70
atgacaacag tttattatga tcaagacgta aaaacggacg ctttacaagg caaaaaaatt 60

gcagtagtag gttatggatc acaaggccac gcgcatgcac aaaacttaaa agacaatgga 120

tatgatgtag tcatcggtat tcgcccaggt cgttcttttg acaaggctaa agaagatgga 180

tttgatgtgt tccctgttgc agaagcagtt aagcaagctg atgtaattat ggtgctatta 240

cctgatgaaa ttcaaggtga tgtatacaaa aacgaaattg aaccaaattt agaaaaacat 300

aatgtgcttg catttgctca tggctttaac attcattttg gtgttattca accaccagct 360

gatgttgatg tatttttagt agctcctaaa ggaccgggtc atttagttag acgtacattt 420

gttgaaggtt ctgctgtacc atcactattt ggtattcaac aagacgcttc aggtcaagca 480

cgtaatattg ctttaagtta tgcaaaaggt attggtgcaa ctcgtgcagg tgttattgaa 540

acaacattta aagaagaaac tgagacagat ttatttggtg aacaagcagt actttgcggt 600

ggtgtatcga aattaattca aagtggcttt gaaacattag tagaagcggg ttatcaacca 660

gaattagctt attttgaagt attacatgaa atgaaattaa tcgttgattt gatgtatgaa 720

ggcggtatgg aaaatgtacg ttactcaatt tcaaatactg ctgaatttgg tgactatgtt 780

tcaggaccac gtgttatcac accagatgtt aaagaaaata tgaaagctgt attaactgat 840

atccaaaatg gtaacttcag taatcgcttt atcgaagaca ataaaaatgg attcaaagaa 900

ttttataaat tacgcgaaga acaacatggt catcaaattg aaaaagttgg tcgtgaatta 960

cgcgaaatga tgccttttat taaatctaaa agcattgaaa aa 1002

<210> 71
<211> 334
<212> PRT
<213> Staphylococcus aureus

<400> 71

Met Thr Thr Val Tyr Tyr Asp Gln Asp Val Lys Thr Asp Ala Leu Gln
1 5 10 15

Gly Lys Lys Ile Ala Val Val Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg
35 40 45

Pro Gly Arg Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Val Phe
50 55 60

Pro Val Ala Glu Ala Val Lys Gln Ala Asp Val Ile Met Val Leu Leu
65 70 75 80

Pro Asp Glu Ile Gln Gly Asp Val Tyr Lys Asn Glu Ile Glu Pro Asn
85 90 95

Leu Glu Lys His Asn Val Leu Ala Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser

130 135 140

Ala Val Pro Ser Leu Phe Gly Ile Gln Gln Asp Ala Ser Gly Gln Ala
145 150 155 160

Arg Asn Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Lys Leu Ile Gln Ser
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Glu
260 265 270

Asn Met Lys Ala Val Leu Thr Asp Ile Gln Asn Gly Asn Phe Ser Asn
275 280 285

Arg Phe Ile Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Lys Leu
290 295 300

Arg Glu Glu Gln His Gly His Gln Ile Glu Lys Val Gly Arg Glu Leu
305 310 315 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
325 330

<210> 72
<211> 1002
<212> DNA
<213> Staphylococcus warneri

<400> 72
atgacaacag tttattatga tgaatcagta aaaacagacg ctttagaagg taaaaaaatt 60

gctgtgattg gatatggttc acaaggccat gctcacgcac aaaacttaaa agataacgga 120

tatgatgttg ttatcggtat ccgtccagga cgctcttttg ataaagctaa ggaaaatggt 180

ttcgatgtat atcctgtggc agaagcaacg aaacaagctg acgttatcat ggttttaatt 240

ccagatgaaa tccaaggtga tgtatataaa aatgaaattg aacctaattt agaagctggt 300

aacgcacttg catttgctca tggtttcaat attcactttg gtgttattca acccctgca 360

gatgtcgatg ttttcttagt tgcaccaaaa ggaccaggtc acttagttag acgtacattt 420

gaagaaggta gtgcagtgcc atctttattt ggtgttcaac aagatgctac aggtcatgca 480

agagatattg cattaagtta tgcaaaaggt atcggatcta caagagctgg tgtaattgaa 540

acaacattta aagaagaaac agaaacggat ttatttggtg aacaagctgt actatgtggc 600

ggtatccata aattaattca aagtggattt gaaacattag ttgaagctgg ttaccaacca 660

gaattagcct attttgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa 720

ggcggtatgg ataatgtccg ttattcaatt tctaatactg cagagtatgg tgactatgta 780

tctggacgtc gtatcatcac tccagatgtc aaagacaata tgaaagctgt attaacagat 840

attcaaaacg gtaaatttgc taatagtttt gttgaagata ataaaaatgg atttaaagaa 900

ttctataaaa tgcgtgaaga acaagctggt catccaattg aaaaagtggg ccgtaaatta 960

cgtgatatga tgccatttat taaatcaaaa agcattgaaa ag 1002

<210> 73
<211> 334
<212> PRT
<213> Staphylococcus warneri

<400> 73

Met Thr Thr Val Tyr Tyr Asp Glu Ser Val Lys Thr Asp Ala Leu Glu
1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg
35 40 45

Pro Gly Arg Ser Phe Asp Lys Ala Lys Glu Asn Gly Phe Asp Val Tyr
50 55 60

Pro Val Ala Glu Ala Thr Lys Gln Ala Asp Val Ile Met Val Leu Ile
65 70 75 80

Pro Asp Glu Ile Gln Gly Asp Val Tyr Lys Asn Glu Ile Glu Pro Asn
85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Gly Val Ile Gln Pro Pro Ala Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly Ser
130 135 140

Ala Val Pro Ser Leu Phe Gly Val Gln Gln Asp Ala Thr Gly His Ala
145 150 155 160

Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ser Thr Arg Ala
165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe

```
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Asp Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
                245                 250                 255

Gly Asp Tyr Val Ser Gly Arg Arg Ile Ile Thr Pro Asp Val Lys Asp
            260                 265                 270

Asn Met Lys Ala Val Leu Thr Asp Ile Gln Asn Gly Lys Phe Ala Asn
        275                 280                 285

Ser Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Lys Met
    290                 295                 300

Arg Glu Glu Gln Ala Gly His Pro Ile Glu Lys Val Gly Arg Lys Leu
305                 310                 315                 320

Arg Asp Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
                325                 330
```

<210> 74
<211> 1002
<212> DNA
<213> Staphylococcus epidermidis

<400> 74

```
atgacaaaag tatattacga tgaaacagta actcaagatg cattacaagg taaaaaaatt       60

gctgtcattg gttatggctc acaaggacat gcacatgcac aaaatttaaa ggacaatggt      120

tatgatgtag tcattggtct acgtccagga cgatcattta ataaagctaa agaagatgga      180

tttgatgttt atacggtaag tgatgctaca caacaagctg atgtagtgat ggtactattg      240

cctgatgaaa ttcaaggtaa agtatataac gaggaaatta aaccacattt agaaaaagga      300

aatgctttag cattcgcaca cggttttaat atccacttca gtgttatcga accacctagt      360

gatgtcgatg tctttttagt agcacctaaa ggaccaggtc atttagttag acgtacattt      420

gttgaaggaa gtgccgtacc agcattattt ggtgttcaac aagatgctac aggccaagca      480

agaaacattg ctttaagcta cgcaaaaggc attggtgcta ctcgtgccgg ggtcattgaa      540

acgacattta aagaagaaac tgaaacagat ttattcggtg aacaagctgt tctttgtgga      600

ggagtttcca aattaattca gagtggtttc gaaacacttg tagaagcagg ttaccaacct      660

gaattagctt attttgaagt cttacacgaa atgaaattaa ttgttgattt aatgtatgaa      720

ggcggaatgg aaaacgttcg ttattctatc tctaacactg ctgaatttgg cgactatgtt      780

tctggaccaa gagtaattac acctgatgtt aaagaaaata tgaaaaaagt acttgaagat      840
```

attcaaaatg gtaactttag ccgtagattt gttgaagata acaaaaatgg ctttaaagaa 900

ttctatcaat tacgtgaaga tcaacatggt catcaaattg aacaagttgg acgtgaatta 960

agagaaatga tgccattcat taaatctaaa agtattgaaa aa 1002

<210> 75
<211> 334
<212> PRT
<213> Staphylococcus epidermidis

<400> 75

Met Thr Lys Val Tyr Tyr Asp Glu Thr Val Thr Gln Asp Ala Leu Gln
1 5 10 15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Leu Arg
35 40 45

Pro Gly Arg Ser Phe Asn Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr
50 55 60

Thr Val Ser Asp Ala Thr Gln Gln Ala Asp Val Val Met Val Leu Leu
65 70 75 80

Pro Asp Glu Ile Gln Gly Lys Val Tyr Asn Glu Glu Ile Lys Pro His
85 90 95

Leu Glu Lys Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Ser Val Ile Glu Pro Pro Ser Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser
130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala
145 150 155 160

Arg Asn Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Lys Leu Ile Gln Ser
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu

```
225                 230                 235                 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Glu
            260                 265                 270

Asn Met Lys Lys Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ser Arg
        275                 280                 285

Arg Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Gln Leu
    290                 295                 300

Arg Glu Asp Gln His Gly His Gln Ile Glu Gln Val Gly Arg Glu Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
                325                 330
```

<210> 76
<211> 1002
<212> DNA
<213> Staphylococcus saprophyticus

<400> 76

```
atgacaacag tttattatga ccaatcagtt acgaaagacg ctttacaagg taaaaaagta       60

gcaatcatag gctatggttc tcaaggacac gcacatgctc aaaatttaaa agacaacggt      120

tatgatgtca ttatcggaat cagacctgga tgttcttttg acaaagctaa agatgatggc      180

ttcgaagtat atcctgtaga tgaagcagca aaacaagctg atgtcatcat ggtgttatta      240

ccagatgaaa ttcaaggcca agtatataaa gaagaaattg aaccaaactt agaagcaaac      300

aatgcattag tattcgcgca tggttttaat attcatttcg gtgttattca accaccagaa      360

aacgtagatg tattcttagt agcgcctaaa ggacctggac atttagtacg tcgtacattt      420

gctgaaggaa gcgcagtccc tgccctattc gcagttgaac aagatccgag tggtgaagct      480

agagatttag cattaagcta tgctaaaggt atcggtgcaa cacgtgcagg tgtattagaa      540

acatcattta aagaagaaac agaaacagat ttattcggtg aacaagcagt gctttgtggt      600

ggtacgacta aattagtaca atctggtttc gaaacgttag tagaagcagg ttaccaacca      660

gaaattgcat actttgaagt attgcatgaa atgaaattga ttgttgattt aatgtatgaa      720

ggcggtatgg aaaatatgcg ctattcaatt tcaaatacag ctgaatttgg tgactatgtt      780

tctggaccac gtattatcac accggatgtt aaagataata tgaaagctgt attagatgat      840

attcaaaaag gaaacttcag tgatcgattc attaaagata atcaaaataa ttttgaagaa      900

ttccataaat taagagaaga acaacatggt catcaaatcg aagcggttgg tagagaactc      960

cgtgatatga tgccattcat caaatctaag agcattgaaa aa                        1002
```

<210> 77
<211> 334
<212> PRT
<213> Staphylococcus saprophyticus

<400> 77

```
Met Thr Thr Val Tyr Tyr Asp Gln Ser Val Thr Lys Asp Ala Leu Gln
1               5                  10                  15

Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Ile Ile Gly Ile Arg
        35                  40                  45

Pro Gly Cys Ser Phe Asp Lys Ala Lys Asp Asp Gly Phe Glu Val Tyr
    50                  55                  60

Pro Val Asp Glu Ala Ala Lys Gln Ala Asp Val Ile Met Val Leu Leu
65                  70                  75                  80

Pro Asp Glu Ile Gln Gly Gln Val Tyr Lys Glu Glu Ile Glu Pro Asn
                85                  90                  95

Leu Glu Ala Asn Asn Ala Leu Val Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Gly Val Ile Gln Pro Pro Glu Asn Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Ala Glu Gly Ser
    130                 135                 140

Ala Val Pro Ala Leu Phe Ala Val Glu Gln Asp Pro Ser Gly Glu Ala
145                 150                 155                 160

Arg Asp Leu Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Leu Glu Thr Ser Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Thr Lys Leu Val Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Ile Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Ile Ile Thr Pro Asp Val Lys Asp
            260                 265                 270

Asn Met Lys Ala Val Leu Asp Asp Ile Gln Lys Gly Asn Phe Ser Asp
```

275 280 285

Arg Phe Ile Lys Asp Asn Gln Asn Asn Phe Glu Glu Phe His Lys Leu
290 295 300

Arg Glu Glu Gln His Gly His Gln Ile Glu Ala Val Gly Arg Glu Leu
305 310 315 320

Arg Asp Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
325 330

<210> 78
<211> 1002
<212> DNA
<213> Staphylococcus carnosus

<400> 78
atgacaacag tttattatga cgaaacagta aaagaagatg ctttaaaagg taaaaaaatt 60

gcagtagttg gttatggttc tcaaggccat gctcatgcac aaaacttaaa agacaatggt 120

tatgacgttg taatcggtat ccgcccaggt aattctttcg ataaagctaa aaaagacggc 180

tttgatgtat atccagttga tgaagcagtg aaacaagctg atgtagtaat ggtattatta 240

cctgatgaaa tccaaggtaa agtatatgaa gaagaaatcg caccaaactt agaagctggc 300

aacgctttag cttttgcaca tggcttcaac atccactttg atgttatcca accacctgaa 360

gatgttgatg tattcttagt tgcccctaaa ggaccaggac atttagtaag acgtacattt 420

actgaaggaa gctcagtacc tgccctattc ggtgtagaac aagatgcttc tggcaatgct 480

ttcaacttag ctttaagtta tgcaaaaggt atcggagcaa caaaagctgg ggttatcgaa 540

actacattta aagaagaaac tgaaactgat ttatttggtg aacaagctgt actttgcggc 600

ggtgtaacaa gattaatcca aagtggattt gaaacattag ttgaagctgg ttaccaacca 660

gaattagctt attttgaagt attacatgaa atgaaattaa tcgttgattt aatgtatgaa 720

ggcggaatgg aaacaatgcg ttattctatc tcaaatacag cagaatacgg cgattatgtt 780

tctggtccac gtgtcatcac tgatgatgtt aaacaaaaca tgaaagctgt cttagaagat 840

atccaaaacg gtaaattcgc aaacagtttc attgaagata ataaaaatgg tttcaaagag 900

ttctacaaat tgcgtgaaga agcacaagat catcaaattc aagaagttgg taaagaactt 960

cgtgaaatga tgccttttgt tgaaaataaa actattgaaa aa 1002

<210> 79
<211> 334
<212> PRT
<213> Staphylococcus carnosus

<400> 79

Met Thr Thr Val Tyr Tyr Asp Glu Thr Val Lys Glu Asp Ala Leu Lys
1 5 10 15

Gly Lys Lys Ile Ala Val Val Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Ile Arg

```
        35                  40                  45

Pro Gly Asn Ser Phe Asp Lys Ala Lys Lys Asp Gly Phe Asp Val Tyr
    50                  55                  60

Pro Val Asp Glu Ala Val Lys Gln Ala Asp Val Val Met Val Leu Leu
65                  70                  75                  80

Pro Asp Glu Ile Gln Gly Lys Val Tyr Glu Glu Ile Ala Pro Asn
                85                  90                  95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Asp Val Ile Gln Pro Pro Glu Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly Ser
    130                 135                 140

Ser Val Pro Ala Leu Phe Gly Val Glu Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Phe Asn Leu Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Lys Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Thr Arg Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Thr Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Asp Asp Val Lys Gln
            260                 265                 270

Asn Met Lys Ala Val Leu Glu Asp Ile Gln Asn Gly Lys Phe Ala Asn
        275                 280                 285

Ser Phe Ile Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Lys Leu
    290                 295                 300

Arg Glu Glu Ala Gln Asp His Gln Ile Gln Glu Val Gly Lys Glu Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Val Glu Asn Lys Thr Ile Glu Lys
```

325 330

<210> 80
<211> 993
<212> DNA
<213> Listeria monocytogenes

<400> 80
atgacaaaag tttattatga agatgcagta aaaaacaacg cactagaagg taaaacagta 60

gcagtaatcg ggtacggttc gcaaggtcac gcacattctc aaaatctacg tgacaatggc 120

aataacgtta ttatcggcat tcgcgaagga aaatctgccg aatctgctag aaacgatggc 180

tttgatgttt attctgttag cgaagccgct gataaagctg atgtcatcat gattcttttg 240

ccagatgaaa cacaaggcga aacatacgaa aatgaaatta aacctaacct aaaagctggt 300

aattcacttg ttttcgctca cggttttaac attcattttg acgtaattaa tcctccaagc 360

gatgtagatg tttttctagt agctccaaaa ggtcctggtc acttagttcg ccgcacattt 420

gttgaaggtg gcgctgttcc ttccctattc gctatctatc aagatgccac tggaaacgca 480

cgcgacacag ccctttccta tgcaaaaggt attggcgcaa ctcgtgctgg cgttatcgaa 540

accactttca aagaagaaac cgaaaccgat ctatttggcg aacaagcagt tctttgtggg 600

ggtgcgactc accttatcca agctggtttc gaaacacttg tagaggctgg ctaccaacca 660

gaacttgctt attttgaagt attacatgaa atgaaactaa ttgttgattt gatgtatgaa 720

ggtggtatgg aaaaaatgcg ccactcgatc tccaatacag cagaatatgg tgattatgtt 780

tccggtcctc gagttgttac agccgatacg aaaaaagcaa tgaaagaagt acttaccgac 840

attcaaaatg gtaactttgc taaatctttc atcaatgaca acaaaaatgg ttttaaagaa 900

ttccatagaa tgcgtaaaga acaacaaggt catcaaatcg aaaaagttgg tgcagaactt 960

cgcgaaatga tgccatttgt caaaccacaa cat 993

<210> 81
<211> 331
<212> PRT
<213> Listeria monocytogenes

<400> 81

Met Thr Lys Val Tyr Tyr Glu Asp Ala Val Lys Asn Asn Ala Leu Glu
1 5 10 15

Gly Lys Thr Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ser Gln Asn Leu Arg Asp Asn Gly Asn Asn Val Ile Ile Gly Ile Arg
35 40 45

Glu Gly Lys Ser Ala Glu Ser Ala Arg Asn Asp Gly Phe Asp Val Tyr
50 55 60

Ser Val Ser Glu Ala Ala Asp Lys Ala Asp Val Ile Met Ile Leu Leu
65 70 75 80

Pro Asp Glu Thr Gln Gly Glu Thr Tyr Glu Asn Glu Ile Lys Pro Asn

85 90 95

Leu Lys Ala Gly Asn Ser Leu Val Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Asp Val Ile Asn Pro Pro Ser Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Gly
130 135 140

Ala Val Pro Ser Leu Phe Ala Ile Tyr Gln Asp Ala Thr Gly Asn Ala
145 150 155 160

Arg Asp Thr Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Thr His Leu Ile Gln Ala
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Met Glu Lys Met Arg His Ser Ile Ser Asn Thr Ala Glu Tyr
245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Val Thr Ala Asp Thr Lys Lys
260 265 270

Ala Met Lys Glu Val Leu Thr Asp Ile Gln Asn Gly Asn Phe Ala Lys
275 280 285

Ser Phe Ile Asn Asp Asn Lys Asn Gly Phe Lys Glu Phe His Arg Met
290 295 300

Arg Lys Glu Gln Gln Gly His Gln Ile Glu Lys Val Gly Ala Glu Leu
305 310 315 320

Arg Glu Met Met Pro Phe Val Lys Pro Gln His
325 330

<210> 82
<211> 993
<212> DNA
<213> Listeria grayi

<400> 82
atgacaaaag tatattatga agatgctgta aaagaaaatg cactaacggg gaagaaagtt 60

gccgttatcg gttatggatc acaaggacat gcgcattcgc aaaatctgcg tgacaatggc 120

aatgacgtta tcatcggtat ccgcgaagga cgttctaaag aaagtgcgaa aaacgacggc 180

ttcgatgttt actccgttgc cgaagcgagt aaacaagcag atgtcatcat gatccttcta 240

ccagacgaaa tccaaggaga tacatacgaa aaagaaatca gcccaaactt ggaagcagga 300

aatgccattg ttttcgccca tggtttcaac atccattttg atgtgatcaa acctcctaca 360

gatgtagatg ttttcctagt tgcaccaaaa ggccccggac atctagtaag acgtaccttt 420

gtagaaggcg gtgccgtccc ttccctcttc gccgtccacc aagatgccac tggcaaagcg 480

caggacctcg ccctttctta tgccaaagga atcggagcta cacgtgccgg cgttatcgaa 540

acaactttca aagaagaaac agaaacggat ctattcggtg aacaagctgt cctatgcggc 600

ggtgccactc accttatcca agcaggcttt gaaacacttg tagaagcagg ctatcaaccg 660

gaactcgctt acttcgaagt actgcatgag atgaaactga tcgtagacct gatgtacgaa 720

ggcggaatgg aaaaaatgcg ctattcgatc tccaatacag cagaatatgg cgactatgta 780

tccggtccgc gcatcgtgac cagcgatacg aagaaagcga tgaaagaagt attgaccgat 840

atccaaaatg gtaattttgc caaatcgttt atcgatgaca ataaaaacgg cttcaaagaa 900

tttcaccgaa tgcgcaaaga acagcaaggt catcaaatcg aagcagtcgg caaaaatctc 960

cgcgaaatga tgccgttcgt aaatccaaga aaa 993

<210> 83
<211> 331
<212> PRT
<213> Listeria grayi

<400> 83

Met Thr Lys Val Tyr Tyr Glu Asp Ala Val Lys Glu Asn Ala Leu Thr
1 5 10 15

Gly Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ser Gln Asn Leu Arg Asp Asn Gly Asn Asp Val Ile Ile Gly Ile Arg
35 40 45

Glu Gly Arg Ser Lys Glu Ser Ala Lys Asn Asp Gly Phe Asp Val Tyr
50 55 60

Ser Val Ala Glu Ala Ser Lys Gln Ala Asp Val Ile Met Ile Leu Leu
65 70 75 80

Pro Asp Glu Ile Gln Gly Asp Thr Tyr Glu Lys Glu Ile Ser Pro Asn
85 90 95

Leu Glu Ala Gly Asn Ala Ile Val Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Asp Val Ile Lys Pro Pro Thr Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Gly

```
    130                 135                 140

Ala Val Pro Ser Leu Phe Ala Val His Gln Asp Ala Thr Gly Lys Ala
145                 150                 155                 160

Gln Asp Leu Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ala Thr His Leu Ile Gln Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Lys Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Ile Val Thr Ser Asp Thr Lys Lys
            260                 265                 270

Ala Met Lys Glu Val Leu Thr Asp Ile Gln Asn Gly Asn Phe Ala Lys
        275                 280                 285

Ser Phe Ile Asp Asp Asn Lys Asn Gly Phe Lys Glu Phe His Arg Met
    290                 295                 300

Arg Lys Glu Gln Gln Gly His Gln Ile Glu Ala Val Gly Lys Asn Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Val Asn Pro Arg Lys
                325                 330
```

<210> 84
<211> 987
<212> DNA
<213> Enterococcus casseliflavus

<400> 84

```
atggtaaaag tatattacaa cgaggaagtt acagataatt taggagacaa gaagattgcg       60

atcgtcggct acggctctca agggcacgct catgcgcaaa atttacggga cacaggtcat      120

gaagtcttta tcggcattcg tgaaggaaaa tcagcggaag ccgcaagaaa cgatggattt      180

gaggttttcc cagtagctga agccgttcaa aaagcggatg tagtcatgat cttggcaccc      240

gatgagatcc aaggatcgtt gtataaaaat gagatcgaac cgaatctttc tgcgggcaat      300

gcattggcat ttgctcatgg tttcaatatt cattttgatg tgatcgcgcc accaaaagag      360

gtggatgttt ttctagttgc gccaaaagga ccaggtcatt tggtgcgtcg aacctttacc      420

gaagggtttg cggtaccagc cttgtttgca gtttatcaag atgcgactgg taacgcacaa      480
```

gcaacggctt tatcctatgc aaaaggcatt ggtgccacac gggtcggcgt tttggaaacg 540

acctttaaag aagagaccga aaccgatctt ttcggtgagc aagccgtact ttgtggcggt 600

ttgaccagta tgatcgaagc gggctttgaa acattggtag aagcggggta ccaaccggaa 660

ttggcttatt ttgaagtttg tcatgaattg aaactgatcg ttgatttgat ttatgaagga 720

ggcttcaaaa aaatgcgcca ttcgatctcg aatacggctg agtatggtga ctacgtttct 780

ggtgctcggg tgatcactgc acaagccaaa gaaaacatga aagccgtatt gacagacatc 840

caaaacggcc aatttgcgaa aggattcatt gaagacaatc aaaacggctt ccctgaattc 900

cataaaatgc gtgaggaaaa cggcaatcat caaattgaaa aagtcggcag cgagttacgg 960

aaaatgatgc cgtttgttgc gaaagat 987

<210> 85
<211> 329
<212> PRT
<213> Enterococcus casseliflavus

<400> 85

Met Val Lys Val Tyr Tyr Asn Glu Glu Val Thr Asp Asn Leu Gly Asp
1 5 10 15

Lys Lys Ile Ala Ile Val Gly Tyr Gly Ser Gln Gly His Ala His Ala
20 25 30

Gln Asn Leu Arg Asp Thr Gly His Glu Val Phe Ile Gly Ile Arg Glu
35 40 45

Gly Lys Ser Ala Glu Ala Ala Arg Asn Asp Gly Phe Glu Val Phe Pro
50 55 60

Val Ala Glu Ala Val Gln Lys Ala Asp Val Val Met Ile Leu Ala Pro
65 70 75 80

Asp Glu Ile Gln Gly Ser Leu Tyr Lys Asn Glu Ile Glu Pro Asn Leu
85 90 95

Ser Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe
100 105 110

Asp Val Ile Ala Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro
115 120 125

Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly Phe Ala
130 135 140

Val Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln
145 150 155 160

Ala Thr Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Val Gly
165 170 175

Val Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly

```
            180                 185                 190

Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Ala Gly
        195                 200                 205

Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe
    210                 215                 220

Glu Val Cys His Glu Leu Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly
225                 230                 235                 240

Gly Phe Lys Lys Met Arg His Ser Ile Ser Asn Thr Ala Glu Tyr Gly
                245                 250                 255

Asp Tyr Val Ser Gly Ala Arg Val Ile Thr Ala Gln Ala Lys Glu Asn
            260                 265                 270

Met Lys Ala Val Leu Thr Asp Ile Gln Asn Gly Gln Phe Ala Lys Gly
        275                 280                 285

Phe Ile Glu Asp Asn Gln Asn Gly Phe Pro Glu Phe His Lys Met Arg
    290                 295                 300

Glu Glu Asn Gly Asn His Gln Ile Glu Lys Val Gly Ser Glu Leu Arg
305                 310                 315                 320

Lys Met Met Pro Phe Val Ala Lys Asp
                325
```

<210> 86
<211> 984
<212> DNA
<213> Enterococcus gallinarum

<400> 86

```
atggtaaaag tatattataa cgaagtgaca gacaatttag gagataagac tattgcaatc       60

attggttacg gttctcaagg acatgcccat gcacagaatt tacgggatac aggacacaac      120

gtcgtgatcg ggattcgtgc aggaaaatca gctgagaacg caaagaatga tggatttacc      180

gtgcttccag tttctgaagc tgtacaacaa gcaaatgttg tgatgatctt agctcctgat      240

gagatccaag gagaacttta cgagaaggaa attgcgccta atctttctgc aggggatgca      300

ttagcctttg ctcatggctt taatattcat tttgaggtca tcacgcctcc aaaagaggtg      360

gatgtcttcc tggtggcacc taaaggacca ggccatttgg tgcgccggac ctttagcgaa      420

ggatttgctg tccctgcgtt gtttgctgtt tatcaagatg ctactggcaa tgcacaagag      480

acagccttgt cctatgccaa aggaatcggg gctacaagag tcggtgtact agaaacaacg      540

ttcaaggaag agacggaaac agacttgttt ggggaacaag ccgtattatg tggcggactg      600

acaagtatga tcgagactgg gtttgaaacg ctggtagaag ccggctatca gccagaacta      660

gcctattttg aagtttgtca cgaaatgaaa ctgattgttg acttgattta tgaaggagga      720

ttcgccaaga tgcgtgattc aatctccaat acagcagaat atggggatta tgtctctggt      780

tcacgaatca ttaccgaaca aacaaaagcc aatatgaaga atgtcttgaa agatattcaa      840
```

aacggcacgt ttgccaaagg ttttattgcg gataatcaag ctggttttcc ggaattcaaa 900

aaaatgcggg aacaaaacgg gaatcatcca atcgaaaaag ttgggtcaga attaagaaaa 960

atgatgcctt ttgttacgaa agat 984

<210> 87
<211> 328
<212> PRT
<213> Enterococcus gallinarum

<400> 87

Met Val Lys Val Tyr Tyr Asn Glu Val Thr Asp Asn Leu Gly Asp Lys
1 5 10 15

Thr Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln
20 25 30

Asn Leu Arg Asp Thr Gly His Asn Val Val Ile Gly Ile Arg Ala Gly
35 40 45

Lys Ser Ala Glu Asn Ala Lys Asn Asp Gly Phe Thr Val Leu Pro Val
50 55 60

Ser Glu Ala Val Gln Gln Ala Asn Val Val Met Ile Leu Ala Pro Asp
65 70 75 80

Glu Ile Gln Gly Glu Leu Tyr Glu Lys Glu Ile Ala Pro Asn Leu Ser
85 90 95

Ala Gly Asp Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe Glu
100 105 110

Val Ile Thr Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro Lys
115 120 125

Gly Pro Gly His Leu Val Arg Arg Thr Phe Ser Glu Gly Phe Ala Val
130 135 140

Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln Glu
145 150 155 160

Thr Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Val Gly Val
165 170 175

Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly Glu
180 185 190

Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Thr Gly Phe
195 200 205

Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe Glu
210 215 220

Val Cys His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly Gly

225 230 235 240

Phe Ala Lys Met Arg Asp Ser Ile Ser Asn Thr Ala Glu Tyr Gly Asp
245 250 255

Tyr Val Ser Gly Ser Arg Ile Ile Thr Glu Gln Thr Lys Ala Asn Met
260 265 270

Lys Asn Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys Gly Phe
275 280 285

Ile Ala Asp Asn Gln Ala Gly Phe Pro Glu Phe Lys Lys Met Arg Glu
290 295 300

Gln Asn Gly Asn His Pro Ile Glu Lys Val Gly Ser Glu Leu Arg Lys
305 310 315 320

Met Met Pro Phe Val Thr Lys Asp
325

<210> 88
<211> 990
<212> DNA
<213> Macrococcus caseolyticus

<400> 88
atggctaaag tatattatga taacgatgta aataaggaat acttaaaaaa taagaaaatt 60

gcagtattag ggtacggctc tcaaggacac gcgcatgcat tgaatcttaa ggacagtggc 120

tatgacgtag tcatcggagt acgtgaaggg cagtccagaa caaaggcaga agaagatgga 180

tttcaagcat atgacgtcag tgaagctgtg gaacgtgcgg atgttaccgt cgtccttatg 240

ccggacgaag tacagcagcg tgtgttcaac gaggaaatcg tacctcatct caaaccacaa 300

tcggctctag ttttcgctca tggatttaat gtgcatttcg gtgcaataaa agctcctgac 360

gatgttgatg tcttcttagt cgcaccaaaa ggaccaggac atcttgttcg acgcgaatat 420

gttaagggca gtgctgtccc tgcgttattt gcagtcgatc aggatgtaac tggtgatgca 480

aaggcccttg cactgaacta tgcacacggc atcggtgcaa cgcgcgccgg cgtcatcgaa 540

acgaccttca aggaagaaac agagacagac ctttttggtg agcaagcggt actctgcggc 600

ggtatcacaa agctcatcca ctacggtttt gaaacactga cggaagcagg ttatcagcca 660

gaactcgcct atttcgaagt actgcatgaa atgaagctga ttgtggacct tatgtacgag 720

ggcggtatgg agaagatgag acattccatc tcaaataccg ctgagtttgg agattatgtc 780

tctggtgcgc gcgttattac ccccgaagtt aaggaaaata tgaaacaggt attagctgat 840

atccagtctg gcacattcag caggcagttt atcgaagaca ataataatgg atttgagtct 900

tttaaatcta tgagaagtac ccaggcagga catcctatag agaaagtcgg tgcagaccta 960

agaatgatga tgccgtttat caataataac 990

<210> 89
<211> 330
<212> PRT
<213> Macrococcus caseolyticus

<400> 89

```
Met Ala Lys Val Tyr Tyr Asp Asn Asp Val Asn Lys Glu Tyr Leu Lys
1               5                  10                  15

Asn Lys Lys Ile Ala Val Leu Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Leu Lys Asp Ser Gly Tyr Asp Val Val Ile Gly Val Arg
        35                  40                  45

Glu Gly Gln Ser Arg Thr Lys Ala Glu Glu Asp Gly Phe Gln Ala Tyr
    50                  55                  60

Asp Val Ser Glu Ala Val Glu Arg Ala Asp Val Thr Val Val Leu Met
65                  70                  75                  80

Pro Asp Glu Val Gln Gln Arg Val Phe Asn Glu Glu Ile Val Pro His
                85                  90                  95

Leu Lys Pro Gln Ser Ala Leu Val Phe Ala His Gly Phe Asn Val His
            100                 105                 110

Phe Gly Ala Ile Lys Ala Pro Asp Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Glu Tyr Val Lys Gly Ser
    130                 135                 140

Ala Val Pro Ala Leu Phe Ala Val Asp Gln Asp Val Thr Gly Asp Ala
145                 150                 155                 160

Lys Ala Leu Ala Leu Asn Tyr Ala His Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile Thr Lys Leu Ile His Tyr
        195                 200                 205

Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Lys Met Arg His Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Ala Arg Val Ile Thr Pro Glu Val Lys Glu
            260                 265                 270

Asn Met Lys Gln Val Leu Ala Asp Ile Gln Ser Gly Thr Phe Ser Arg
```

275 280 285

Gln Phe Ile Glu Asp Asn Asn Asn Gly Phe Glu Ser Phe Lys Ser Met
290 295 300

Arg Ser Thr Gln Ala Gly His Pro Ile Glu Lys Val Gly Ala Asp Leu
305 310 315 320

Arg Met Met Met Pro Phe Ile Asn Asn Asn
325 330

<210> 90
<211> 1020
<212> DNA
<213> Streptococcus vestibularis

<400> 90
atggcagttc aaatggaata tgaaaaagac gtaaaagtag cagcacttga cggtaaaaaa 60

atcgccgtta tcggttacgg ttcacaaggt catgcacatg cacaaaactt gcgtgattca 120

ggtcacgatg tgatcatcgg tgttcgtcca ggtaaatcat ttgataaagc taaagaagat 180

ggttttgata cttacacagt agcagaagca actaaattgg ctgatatcat tatggttttg 240

gcaccagatg aaattcaaaa agatatctat aaagatgaaa ttgagccaaa cttgagtgca 300

ggcaaagctc ttggttttgc acacggattt aacatccact tcgaattcat caaggttcct 360

aaagatgttg atgtctttat ggttgctcct aaaggaccag gtcacttggt acgccgtaca 420

tatacagaag gctttggtgt tccagcactt tacgcagtat accaagatgc tacaggcaat 480

gctaaagaca tcgctatgga ctgggctaaa ggtattggtt ctgcacgtgt tggattactt 540

gtaacaacat ttaaagaaga aactgaagaa gatttgtttg gtgagcaagc ggttcttatg 600

ggtgggttga cacatcttat tgaagctggt tttgaagtct tgactgaagc aggctatgca 660

ccacaattgg cttactttga agtgcttcac gaaatgaaat tgatcgttga cttgatctac 720

gaaggtggat tcaagaaaat gcgtcaatca tgttcaaata ctgcagaatt tggtgacttt 780

gtaactggtc cacgtgtcat cggtccagaa gtcaaagaaa acatgaaagc tgcccttgcg 840

gatatccaat caggtaaatt tgcacgtgaa ttcgttgaag accatgatgc tggcttccca 900

cgtttgaaag cttaccgtaa agaagctgaa gagcttgaaa ttgaaaagat tggtgctgaa 960

ttgcgtaaag caatgccatt cgttggtcaa aatgacgatg acgcattcaa aatctataac 1020

<210> 91
<211> 340
<212> PRT
<213> Streptococcus vestibularis

<400> 91

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val

```
        35                  40                  45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
    50                  55                  60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Ile Ile Met Val Leu
65                  70                  75                  80

Ala Pro Asp Glu Ile Gln Lys Asp Ile Tyr Lys Asp Glu Ile Glu Pro
                85                  90                  95

Asn Leu Ser Ala Gly Lys Ala Leu Gly Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Glu Phe Ile Lys Val Pro Lys Asp Val Asp Val Phe Met Val
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
    130                 135                 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145                 150                 155                 160

Ala Lys Asp Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Ser Ala Arg
                165                 170                 175

Val Gly Leu Leu Val Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Met Gly Gly Leu Thr His Leu Ile Glu
        195                 200                 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Gln Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Cys Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Phe Val Thr Gly Pro Arg Val Ile Gly Pro Glu Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Ala Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Arg Glu Phe Val Glu Asp His Asp Ala Gly Phe Pro Arg Leu Lys Ala
    290                 295                 300

Tyr Arg Lys Glu Ala Glu Glu Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
```

325 330 335

Lys Ile Tyr Asn
340

<210> 92
<211> 1020
<212> DNA
<213> Streptococcus mutans

<400> 92
atggcagtag aaatgttata tgaagcagat gtaaaagtag cagcacttga tggtaagaag 60

attgctgtta ttggttatgg atcacaagga catgctcatg cccaaaactt gagagattca 120

ggacacgatg ttatcatcgg tgtgcgtcat gggaaatcgt ttgacaaggc taaagaagat 180

ggttttgata cctatgaagt gggtgaagct actaaattgg ctgatattat catggtttta 240

gcaccggatg aaattcaaaa agatatctac aaagatgaaa ttgctcctaa tttaagtgct 300

ggtaaggctc ttggttttgc tcatggtttt aatatccatt tcggttatat taaggcacca 360

gaagatgtag atgttttcat ggttgcccct aaaggaccgg gccatcttgt acgtcgtact 420

tacacagaag gatttggagt accttcttta tatgctgttt atcaaaatcc aactggtaat 480

gctgaaaata ttgctttaga ttgggctaag ggtattggtt ctgcgcgtgt aggtctcctt 540

gtaaccactt tcaaagaaga aacggaagaa gacctttttg gtgaacaagc ggttcttatg 600

ggtggtttaa ctcaccttat tgaagcaggt tttgaagtgc ttactgaagc agggtatgct 660

ccacaattgg cttattttga agtgcttcat gaaatgaaat tgattgtaga ccttatctat 720

gaaggtggtt tcaagaaaat gcgtcaatca tgttcaaata ccgctgaatt tggtgacttt 780

gtgacaggcc cacgtgttat cggtccagaa gtaaaagaaa acatgaaggc tgctcttgct 840

gatattcaat caggtaaatt tgcccgtgaa ttcgtggaag atcatgatgc tggcttccca 900

cgtttgaaag ctttccgtaa agaagctgaa ggactcgaaa tcgaaaaaat tggtgcagaa 960

cttcgtaaag caatgccatt tgttaatcaa aacgatgacg atgctttcaa aatttataac 1020

<210> 93
<211> 340
<212> PRT
<213> Streptococcus mutans

<400> 93

Met Ala Val Glu Met Leu Tyr Glu Ala Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Glu Val Gly Glu Ala Thr Lys Leu Ala Asp Ile Ile Met Val Leu

65 70 75 80

Ala Pro Asp Glu Ile Gln Lys Asp Ile Tyr Lys Asp Glu Ile Ala Pro
85 90 95

Asn Leu Ser Ala Gly Lys Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Lys Ala Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Val Pro Ser Leu Tyr Ala Val Tyr Gln Asn Pro Thr Gly Asn
145 150 155 160

Ala Glu Asn Ile Ala Leu Asp Trp Ala Lys Gly Ile Gly Ser Ala Arg
165 170 175

Val Gly Leu Leu Val Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Met Gly Gly Leu Thr His Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Gln Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Cys Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Phe Val Thr Gly Pro Arg Val Ile Gly Pro Glu Val Lys
260 265 270

Glu Asn Met Lys Ala Ala Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Arg Glu Phe Val Glu Asp His Asp Ala Gly Phe Pro Arg Leu Lys Ala
290 295 300

Phe Arg Lys Glu Ala Glu Gly Leu Glu Ile Glu Lys Ile Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Asn Gln Asn Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Asn
340

<210> 94

<211> 1020
<212> DNA
<213> Streptococcus gordonii

<400> 94
atggcagttc aaatggaata cgaaaaagat gtaaaagtag cagcgcttga cggtaagaaa 60

attgccgtaa tcggctatgg ttcacaaggt catgctcatg cacaaaatct gcgtgataca 120

ggccatgatg tgattatcgg tgttcgccct ggtaagtcat tcgataaggc taaagaagat 180

ggctttgata cttacacagt agcagaggca gctaagttgg ctgatgtaat catgattttg 240

gcgccagatg aaattcaaca agaactctac gaagcagaaa tcgctccaaa cttggaagct 300

ggtaatgcag ttggatttgc ccatggtttc aatatccatt ttgaatttat caaagttcct 360

gctgatgtag atgtctttat gtgtgcccct aaaggtcctg gtcacttggt tcgtcgtact 420

tttgaagaag gatttggtgt tccggctctt tatgcagtct accaagatgc aacaggaaat 480

gcgaaaaaca ttgctatgga ctggtgtaaa ggtgttggtg cagctcgtgt tggtttgctt 540

gaaacaactt acaaagaaga aacagaagaa gatctctttg gtgaacaagc tgtcctttgt 600

ggtggtttga ctgcccttat cgaagcaggt tttgaagtct tgacagaagc aggttacgct 660

cctgaattgg cttactttga agttcttcat gaaatgaaat tgatcgttga cttgatctac 720

gagggtggat tcaagaaaat gcgccaatca atttcaaata ctgctgaata tggtgattac 780

gtatcaggcc cacgtgtaat tactgagcaa gtcaaagaaa acatgaaggc tgttttggca 840

gacattcaaa atggtaaatt tgcaaatgac tttgtaaatg actacaaggc tggccgtcca 900

aaactcactg cttaccgtga acaagcagcc aaccttgaaa tcgaaaaagt aggtgctgaa 960

ttgcgtaaag caatgccatt tgttggtaaa aatgacgacg acgcatttaa aatctacaac 1020

<210> 95
<211> 340
<212> PRT
<213> Streptococcus gordonii

<400> 95

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Thr Val Ala Glu Ala Ala Lys Leu Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asn Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
260 265 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
275 280 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
290 295 300

Tyr Arg Glu Gln Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Gly Lys Asn Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Asn
340

<210> 96
<211> 1020
<212> DNA
<213> Streptococcus suis

<400> 96
atgacagtaa caatgcaata tgaaaaagat gtaacagtgg cagcacttga cggtaaacgt 60

atcgccgtta tcggttatgg ttcacaaggt catgcccatg ctcaaaactt gcgtgataca 120

ggacacgatg tcatcatcgg tgtgcgtgca ggtaagtcat ttgacaaggc aaaagaagac 180

ggttttgaaa cttttgaagt agcagaggca gcaaaacaag ctgatgtcat tatgattttg 240

gctccagacg aaattcaagc agatttgtat gagaacgaag ttgctccaaa cttggaagct 300

ggaaatgcgc ttggttttgc tcatggtttc aacgttcact ttgaatttat caaggtacca 360

gcagatgtgg atgtcttcat gtgtgcacca aaaggtccag gtcacttggt ccgccgtacc 420

tttgaagaag gttttggagt accagccctc tatgcggtct accaagatgc aactggcaat 480

gcaaaacata ttgcgatgga ctgggcaaaa ggtgttggtt cagcccgtgt tggtcttttg 540

gaaacaactt tcaaagaaga gacagaagaa gatttgtttg gtgagcaagc tgttctctgc 600

ggtggtttga cagcccttat gcaggcaggt tttgaagtct tgacagaagc tggttatgca 660

ccagaattgg cttatttcga agttctccac gagatgaaac tcatcgttga ccttgtctac 720

gaaggtggct ttaagaaaat gcgccaatct atttcaaaca ctgctgaatt tggtgactat 780

gtatcaggtc cacgcgtgat tacagaccaa gtcaaagaaa acatgaaggc agtccttgca 840

gatattcaat ctggtaaatt tgcaaatgac tttgtaaatg actataaagc aggccgtcca 900

cgtatggaag cttatcgtaa agaagcagag aatcttgaaa tcgaaaaagt aggtgcagaa 960

ctccgcaaag caatgccatt tgtcggacgc aatgatgacg acgcattcaa aatctataat 1020

<210> 97
<211> 340
<212> PRT
<213> Streptococcus suis

<400> 97

Met Thr Val Thr Met Gln Tyr Glu Lys Asp Val Thr Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Arg Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Ala Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Ala Glu Ala Ala Lys Gln Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Ala Asp Leu Tyr Glu Asn Glu Val Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Leu Gly Phe Ala His Gly Phe Asn Val
100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly

```
    130                 135                 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145                 150                 155                 160

Ala Lys His Ile Ala Met Asp Trp Ala Lys Gly Val Gly Ser Ala Arg
                165                 170                 175

Val Gly Leu Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Val Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Asp Gln Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Arg Met Glu Ala
    290                 295                 300

Tyr Arg Lys Glu Ala Glu Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Arg Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 98
<211> 1020
<212> DNA
<213> Streptococcus infantarius

<400> 98
atggcagtaa caatggaata cgaaaaagac gtaaaagtag cagctcttga tggtaaaaaa 60

attgccgtta ttggttatgg atcacaaggt catgctcatg ctcaaaactt gcgtgactca 120

ggtcacgatg ttatcattgg ggttcgccat ggtaaatcat tcgataaagc aaaagaagat 180

ggatttgata cttatgaagt agcagaagca acaaaacttg ctgatgttat catggtattg 240

gctcctgatg aaatccaagc taaactttat gctgaagaaa tcgctccaaa ccttgaagct 300

ggtaacgctc ttggatttgc acatggtttt aatatccgtt ttggatacat taaagctcca 360

gaaacagtag atgtcttcat gtgtgctcct aaaggaccag gtcaccttgt tcgtcgtact 420

tacacagaag gatttggtgt accagcactt tacgctgttt accaagatgc tactggtaat 480

gctaaagaca tcgcaatgga ttggtctaaa ggtatcggtg ctgcacgtgt tggacttctt 540

gaaacaacat ttaaagaaga aactgaagaa gacctctttg gtgaacaagc agtactttgt 600

ggtggtttaa ctgctcttat cgaagctggt tttgaagttc ttactgaagc tggctatgct 660

ccagaattgg cttactttga agttcttcat gaaatgaaac ttatcgttga ccttatctac 720

gaaggtggat tcaagaaaat gcgtcaatca atttcaaata cagctgaatt tggtgactac 780

gtatctggac cacgtgttat cactaaagat gttaaagaaa atatgaaagc tgttcttgct 840

gatatccaat caggtaaatt cgctgaagat tttgttaacg actaccaagc aggtcgtcca 900

aaacttgaag cataccgtaa agaagctgca gctcttgaaa ttgaaaaagt gggtgctgaa 960

cttcgtaaag caatgccttt tgttaaccaa aacgatgacg atgcattcaa aatttataac 1020

<210> 99
<211> 340
<212> PRT
<213> Streptococcus infantarius

<400> 99

Met Ala Val Thr Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Glu Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Ala Lys Leu Tyr Ala Glu Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

Arg Phe Gly Tyr Ile Lys Ala Pro Glu Thr Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Ser Lys Gly Ile Gly Ala Ala Arg

```
                165                 170                 175

Val Gly Leu Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
        195                 200                 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Lys Asp Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Glu Asp Phe Val Asn Asp Tyr Gln Ala Gly Arg Pro Lys Leu Glu Ala
    290                 295                 300

Tyr Arg Lys Glu Ala Ala Ala Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Asn Gln Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 100
<211> 1020
<212> DNA
<213> Lactococcus lactis

<400> 100

```
atggcagtta caatgtatta tgaagaagat gtagaagtag ccgcactcgc gggtaagaaa       60

atcgcagtga ttggatatgg ctcacaagga cacgctcatg cacaaaactt gcgtgattct      120

ggtcatgatg tgattattgg tgtccgtcag gggaaatctt ttgataaagc aaaagaagat      180

ggttttgaaa catttgaagt aggagaagca gtagctaaag ctgacgtcat tatggttctg      240

gcacctgatg aacttcaaca atctatttat gaagaggaca taaaaccaaa tttgaaagca      300

ggttcagcac ttggttttgc ccatggtttc aatattcatt ttggctatat tgaagttcca      360

gaagatgttg atgtcttcat ggttgcgcca aaagcgccgg gacatctcgt tcggcggact      420

tttaccgaag gtttcggaac gccagctttg ttcgtttcgc atcaaaatgc cactggtcat      480

gcgcgtgaaa ttgccatgga ctgggccaaa ggaattggct gtgcccgtgt cggtatcatt      540

gaaacaactt tcaaagaaga aacagaagaa gatttgtttg gcgaacaggc cgtgctttgt      600
```

ggcggtttga cagcacttgt tgaagctggt tttgaaacac tgacagaagc tggatatgct 660

ggcgaattgg cttactttga agtgctgcat gaaatgaaat tgattgttga ccttatgtac 720

gaaggtggtt tcactaaaat gcgtcagtca atctcaaaca ctgccgaatt tggtgattat 780

gtgactggac cacgcattat tactgacgaa gttaaaaaga atatgaaact cgtgttggct 840

gacattcaat caggaaaatt tgcgcaagat ttcgttgatg atttcaaagc tggacgtcca 900

aaattaactg cttatcgtga agcagctaaa aatctggaaa ttgaaaaaat cggtgcagaa 960

ctacgtaaag caatgccatt tacacaatct ggtgatgacg acgcctttaa aatttatcaa 1020

<210> 101
<211> 340
<212> PRT
<213> Lactococcus lactis

<400> 101

Met Ala Val Thr Met Tyr Tyr Glu Glu Asp Val Glu Val Ala Ala Leu
1 5 10 15

Ala Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Gln Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Glu Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Thr Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu

195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225 230 235 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
260 265 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Thr Ala
290 295 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Gln
340

<210> 102
<211> 1032
<212> DNA
<213> Lactococcus lactis

<400> 102
atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa 60

attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct 120

ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat 180

ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg 240

gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300

ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360

gaagacgttg acgtctttat ggttgcgcct aaggctccag gtcaccttgt ccgtcggact 420

tatactgaag gttttggtac accagctttg tttgtttcac accaaaatgc aagtggtcat 480

gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattggtt gtgctcgagt gggaattatt 540

gaaacaactt ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt 600

ggaggtttga cagcacttgt tgaagccggt tttgaaacac tgacagaagc tggatacgct 660

ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720

gaaggtggtt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780

gtgactggtc cacggattat tactgacgaa gttaaaaaga atatgaagct tgttttggct 840

gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900

aaattaatag cctatcgcga agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag 960

cacgtcaagc aatgccattc acacaatctg gtgatgacga tgcctttaaa atctatcagt 1020

aatttctctt at 1032

<210> 103
<211> 344
<212> PRT
<213> Lactococcus lactis

<400> 103

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220

20 Jan 2017 2017200439

```
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225                 230                 235                 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
            260                 265                 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala
    290                 295                 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

His Val Lys Gln Cys His Ser His Asn Leu Val Met Thr Met Pro Leu
                325                 330                 335

Lys Ser Ile Ser Asn Phe Ser Tyr
            340
```

<210> 104
<211> 1041
<212> DNA
<213> Leuconcstoc mesenteroides

<400> 104

```
atgactacaa aaatgtttta tgataaagat attgatacaa caccattgga aaacaaaaag       60

attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct acgcgactca      120

ggatttgatg tcatcatggg attgcgccca ggaaaatctt ttgatagtgc taaaaaagat      180

gggttcgaag tttactcagc tgctgaagca acggcacaag cggatgttgt tatgatggaa      240

acccctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca      300

ggttcttacc ttggattttc tcatgggttc aacattgttt atggtttgat taagcccaat      360

gctgatatta atgtcatgat cattgcgcca aagggtccag gaaacattga acgtcgccaa      420

ttcgttgaag ggggtgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat      480

acagctgaag tggccaaggc ctatgcgaaa ggtattggtt caggtcgcgc aggtatcttg      540

gaaacgactt ttgaagaaga aacaacagaa gacttgtttg gtgaacaagc tgtactttgt      600

ggtggcttaa cacaattgat tgaggcagga tttaacacat tggtggaagc aggttactca      660

cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga tttgatcttt      720

gaaggtggtt tcgagaagat gagacacgat tgctcaaata cttgtgaata tggtgaaatg      780

ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag      840

gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag      900

gatttggaaa agatgcggac agagtataag agcggtttgt acgagcaaac aggtaagaag      960

gttcgtgcta tgatgccatg gatttcagat gcagataagt actcgacagc agcagatact     1020
```

gagcaattct cagcagcaaa g 1041

<210> 105
<211> 347
<212> PRT
<213> Leuconostoc mesenteroides

<400> 105

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Thr Pro Leu
1 5 10 15

Glu Asn Lys Lys Ile Ala Val Ile Gly Tyr Gly Ala Gln Gly His Ala
20 25 30

Gln Ala Asn Asn Leu Arg Asp Ser Gly Phe Asp Val Ile Met Gly Leu
35 40 45

Arg Pro Gly Lys Ser Phe Asp Ser Ala Lys Lys Asp Gly Phe Glu Val
50 55 60

Tyr Ser Ala Ala Glu Ala Thr Ala Gln Ala Asp Val Val Met Met Glu
65 70 75 80

Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro
85 90 95

Asn Leu Lys Ala Gly Ser Tyr Leu Gly Phe Ser His Gly Phe Asn Ile
100 105 110

Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile
115 120 125

Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly
130 135 140

Gly Gly Ile Pro Ser Leu Tyr Gly Val His Gln Asp Pro Thr Gly Asp
145 150 155 160

Thr Ala Glu Val Ala Lys Ala Tyr Ala Lys Gly Ile Gly Ser Gly Arg
165 170 175

Ala Gly Ile Leu Glu Thr Thr Phe Glu Glu Glu Thr Thr Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu
195 200 205

Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala
210 215 220

Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe
225 230 235 240

Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu

245 250 255

Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys
260 265 270

Gln Gly Met Arg Asp Val Leu Lys Asp Ile Gln Asp Gly Thr Tyr Ala
275 280 285

Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys
290 295 300

Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys
305 310 315 320

Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr
325 330 335

Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys
340 345

<210> 106
<211> 1044
<212> DNA
<213> Lactobacillus buchneri

<400> 106
atgagtgtag aaatgttgta tgacaaggat gttaccacca attatcttca aggaaagaaa 60

attgctttta tcggttatgg atcacaaggt catgcccaag ctaacaactt aagagattca 120

ggttatgacg ttatcgttgg tgttcgccca ggacaatcat ttgaaaatgc caaagttgat 180

ggatttgatg tatacacacc agccgaagca gctcgtcgtg cagactggat tcaaatgtta 240

acgcctgatg aagttatggg cgacgtttat aaaaatgaaa ttgcacctaa cctagaggaa 300

ggcaatgtat taggcttttc acatggcttc aacattcatt ataaagaaat cgtaccacca 360

gcaaacgttg atgttgttat gatggcacct aaaggcccag gtaatctttg tcgtcgtaca 420

tacgttgaag gctccggtgt tccagcatta tatggctatt tccaagacta ttcaggtcat 480

gccgaagatt tatccaagga attcgccaaa ggaaatggtg cagcacgtgc cggactattg 540

aagacaacct tcaaagaaga aactgaagaa gatttgtttg gtgagcaaaa cgtccttatg 600

ggtggtgtta ccgcgcttat cgaaaccggt tatgaagtct taaccgaagc aggatactca 660

ccacagttgg catacttcga agtggatcac gaaatgaaat tgatctgtga ccttatctac 720

gaaggtgggt tcaacaagat gtatgctgac tgctcgaaca cttctgaata tggttcatat 780

gtagttggtc caaaggtcgt tggtaaggaa tccaagcaag ctatgaaaga cgctttgaag 840

cggattcaag acggttcatt tgctaaagaa ttcatggatg actaccgtaa cggcttcaag 900

aagctttaca agatgcgtga acgttcagct aactcacttc tttcacgtgt tggtgctgat 960

cttcgtgatc acatgtcatt tgttggtgaa gccgacaagt acagtacacc tactgaagaa 1020

aaagcagaag cagaagctga aaag 1044

<210> 107
<211> 348

20 Jan 2017
2017200439

<212> PRT
<213> Lactobacillus buchneri

<400> 107

```
Met Ser Val Glu Met Leu Tyr Asp Lys Asp Val Thr Thr Asn Tyr Leu
1               5                  10                  15

Gln Gly Lys Lys Ile Ala Phe Ile Gly Tyr Gly Ser Gln Gly His Ala
            20                  25                  30

Gln Ala Asn Asn Leu Arg Asp Ser Gly Tyr Asp Val Ile Val Gly Val
        35                  40                  45

Arg Pro Gly Gln Ser Phe Glu Asn Ala Lys Val Asp Gly Phe Asp Val
    50                  55                  60

Tyr Thr Pro Ala Glu Ala Ala Arg Arg Ala Asp Trp Ile Gln Met Leu
65                  70                  75                  80

Thr Pro Asp Glu Val Met Gly Asp Val Tyr Lys Asn Glu Ile Ala Pro
                85                  90                  95

Asn Leu Glu Glu Gly Asn Val Leu Gly Phe Ser His Gly Phe Asn Ile
            100                 105                 110

His Tyr Lys Glu Ile Val Pro Pro Ala Asn Val Asp Val Val Met Met
        115                 120                 125

Ala Pro Lys Gly Pro Gly Asn Leu Cys Arg Arg Thr Tyr Val Glu Gly
    130                 135                 140

Ser Gly Val Pro Ala Leu Tyr Gly Tyr Phe Gln Asp Tyr Ser Gly His
145                 150                 155                 160

Ala Glu Asp Leu Ser Lys Glu Phe Ala Lys Gly Asn Gly Ala Ala Arg
                165                 170                 175

Ala Gly Leu Leu Lys Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Asn Val Leu Met Gly Gly Val Thr Ala Leu Ile Glu
        195                 200                 205

Thr Gly Tyr Glu Val Leu Thr Glu Ala Gly Tyr Ser Pro Gln Leu Ala
    210                 215                 220

Tyr Phe Glu Val Asp His Glu Met Lys Leu Ile Cys Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Asn Lys Met Tyr Ala Asp Cys Ser Asn Thr Ser Glu
                245                 250                 255

Tyr Gly Ser Tyr Val Val Gly Pro Lys Val Val Gly Lys Glu Ser Lys
            260                 265                 270
```

20 Jan 2017
2017200439

```
Gln Ala Met Lys Asp Ala Leu Lys Arg Ile Gln Asp Gly Ser Phe Ala
        275                 280                 285

Lys Glu Phe Met Asp Asp Tyr Arg Asn Gly Phe Lys Lys Leu Tyr Lys
    290                 295                 300

Met Arg Glu Arg Ser Ala Asn Ser Leu Leu Ser Arg Val Gly Ala Asp
305                 310                 315                 320

Leu Arg Asp His Met Ser Phe Val Gly Glu Ala Asp Lys Tyr Ser Thr
                325                 330                 335

Pro Thr Glu Glu Lys Ala Glu Ala Glu Ala Glu Lys
            340                 345
```

<210> 108
<211> 1002
<212> DNA
<213> Staphylococcus haemolyticus

<400> 108

```
atgactaaag tttattacga tcaatcagtt gagaaagatg cattacaagg taaaaaaatt       60

gcaattatcg gttacggttc acaaggacac gcgcatgcac aaaaccttaa agacaacggt      120

tatgacgtta ttgttggtat tcgccctggt cattcttttg ataaagctaa agaagatggc      180

ttcgatgttt atccagttgc agaagctgtt aaacaagcag atgtcatcat ggtgttattg      240

ccagatgaaa tccaaggaaa tgtctataaa aatgaaattg aacctaactt agaagctgga      300

aatgccctag catttgcgca tggttttaat atccactttg aagttatcaa accacctaaa      360

gatgtggatg tcttcttagt agctcctaaa ggacctggtc atttagtaag acgtaccttt      420

gttgaaggaa ctgccgtacc agcactattt ggtgtacaac aagatgctac tggtcaagca      480

cgcgatattt cattaagtta tgcgaaaggt atcggagcta cgcgtgcagg tgtaattgaa      540

acaacattta aagaagaaac tgaaacagat ttattcggtg agcaagcagt actttgtggt      600

ggtattcata aattaatcca aagtggcttt gaaactttag ttgaagcggg ctatcaaaaa      660

gaattagcat acttcgaagt attacatgaa atgaagttaa ttgttgattt aatgtatgaa      720

ggcggtatgg agaacgttcg ttactcaatt tctaacacag cagagtatgg tgactatgtt      780

tcaggaccta gagtcatcac tcctgatgtt aaagacaata tgaaagcagt attaaaagat      840

attcaaaatg gtaattttgc aaatagcttt gttaaagaca atgaaaatgg ctttaaagaa      900

ttctataaat tacgtgaaca acaacacggt catgaaatcg aagcagttgg tcgcgaatta      960

agaaaaatga tgccattcat taaatctaaa agtattcaaa aa                        1002
```

<210> 109
<211> 334
<212> PRT
<213> Staphylococcus haemolyticus

<400> 109

```
Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln
1               5                   10                  15
```

```
Gly Lys Lys Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Ile Val Gly Ile Arg
        35                  40                  45

Pro Gly His Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr
    50                  55                  60

Pro Val Ala Glu Ala Val Lys Gln Ala Asp Val Ile Met Val Leu Leu
65                  70                  75                  80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Glu Val Ile Lys Pro Pro Lys Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Thr
    130                 135                 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala
145                 150                 155                 160

Arg Asp Ile Ser Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Tyr
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Asp
            260                 265                 270

Asn Met Lys Ala Val Leu Lys Asp Ile Gln Asn Gly Asn Phe Ala Asn
        275                 280                 285

Ser Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Lys Leu
    290                 295                 300
```

20 Jan 2017 2017200439

```
Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
305                 310                 315                 320

Arg Lys Met Met Pro Phe Ile Lys Ser Lys Ser Ile Gln Lys
                325                 330
```

<210> 110
<211> 1002
<212> DNA
<213> Staphylococcus epidermidis

<400> 110

```
atgacaaaag tatattacga cgaaacagta actcaggatg cattacaagg taaaaaaatt       60

gctgtcattg gttatggctc acaaggacat gcacatgcac aaaatttaaa ggacaatggt      120

tatgatgtag tcattggcct gcgtccagga cgatcattta ataaagctaa agaagatgga      180

tttgatgttt atacggtaag tgaagcaaca caacaagcag atgtagtgat ggtactattg      240

cctgatgaaa ttcaaggtga agtatataac aaggaaatta aaccatattt agaaaaagga      300

aatgctttag cattcgcaca cggttttaat atccatttca gtgttatcga accacctagt      360

gatgtcgatg tctttttagt agcacctaaa ggacctggtc atttagttag acgtacattt      420

gttgaaggaa gtgccgtacc agcattattt ggtgttcaac aagatgctac aggccaagct      480

agaaacattg ctttaagcta cgcaaaaggc attggtgcta ctcgtgccgg ggtcattgaa      540

acgacattta aagaagaaac tgaaacagat ttattcggtg aacaagctgt actttgtgga      600

ggagtttcca aattaattca gagtggattc gaaacacttg tggaagcagg ttaccaacct      660

gaattagctt attttgaagt cttacacgaa atgaaattaa ttgttgattt aatgtatgaa      720

ggcggaatgg aaaatgtccg ttattctatc tctaacactg ctgaatttgg cgactatgtt      780

tctggaccaa gagtaattac acctaatgtt aaagaaaata tgaaaaaagt acttgaagat      840

attcaaaatg gtaactttag ccgtagattt gttgaagata acaaaaatgg ctttaaagaa      900

ttctatcaat tacgtgaaga tcaacatggt catcaaattg aacaagttgg acgtgaatta      960

agagaaatga tgccattcat taaatctaaa agtattgaaa aa                        1002
```

<210> 111
<211> 334
<212> PRT
<213> Staphylococcus epidermidis

<400> 111

```
Met Thr Lys Val Tyr Tyr Asp Glu Thr Val Thr Gln Asp Ala Leu Gln
1               5                   10                  15

Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Val Ile Gly Leu Arg
        35                  40                  45

Pro Gly Arg Ser Phe Asn Lys Ala Lys Glu Asp Gly Phe Asp Val Tyr
    50                  55                  60
```

```
Thr Val Ser Glu Ala Thr Gln Gln Ala Asp Val Val Met Val Leu Leu
65                  70                  75                  80

Pro Asp Glu Ile Gln Gly Glu Val Tyr Asn Lys Glu Ile Lys Pro Tyr
                85                  90                  95

Leu Glu Lys Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
            100                 105                 110

Phe Ser Val Ile Glu Pro Pro Ser Asp Val Asp Val Phe Leu Val Ala
        115                 120                 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser
    130                 135                 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly Gln Ala
145                 150                 155                 160

Arg Asn Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
                165                 170                 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Lys Leu Ile Gln Ser
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr
    210                 215                 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
                245                 250                 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asn Val Lys Glu
            260                 265                 270

Asn Met Lys Lys Val Leu Glu Asp Ile Gln Asn Gly Asn Phe Ser Arg
        275                 280                 285

Arg Phe Val Glu Asp Asn Lys Asn Gly Phe Lys Glu Phe Tyr Gln Leu
    290                 295                 300

Arg Glu Asp Gln His Gly His Gln Ile Glu Gln Val Gly Arg Glu Leu
305                 310                 315                 320

Arg Glu Met Met Pro Phe Ile Lys Ser Lys Ser Ile Glu Lys
                325                 330
```

<210> 112
<211> 1021
<212> DNA

20 Jan 2017 2017200439

<213> Streptococcus pneumoniae

<400> 112
atgacagttc aaatggaata tgaaaaagat gttaaagtag cagcacttga cggtaaaaaa 60

atcgccgtta tcggttatgg ttcacaaggg catgcgcatg ctcaaaactt gcgtgattca 120

ggtcgtgacg ttattatcgg tgtacgtcca ggtaaatctt ttgataaagc aaaagaagat 180

ggatttgata cttacacagt aacagaagct actaagttgg ctgatgttat catgatcttg 240

gcgccagacg aaattcaaca agaattgtac gaagcagaaa tcgctccaaa cttggaagct 300

ggaaacgcag ttggatttgc ccatggtttc aacatccact ttgaatttat caaagttcct 360

gcggatgtag atgtcttcat gtgtgctcct aaaggaccag gacacttggt acgtcgtact 420

tacgaagaag gatttggtgt tccagctctt tatgcagtat accaagatgc aacaggaaat 480

gctaaaaaca ttgctatgga ctggtgtaaa ggtgttggag cggctcgtgt aggtcttctt 540

gaaacaactt acaaagaaga aactgaagaa gatttgtttg gtgaacaagc tgtactttgt 600

ggtggtttga ctgcccttat cgaagcaggt ttcgaagtct tgacagaagc aggttacgct 660

ccagaattgg cttactttga agttcttcac gaaatgaaat tgatcgttga cttgatctac 720

gaaggtggat tcaagaaaat gcgtcaatct atttcaaaca ctgctgaata cggtgactat 780

gtatcaggtc cacgtgtaat cactgaacaa gttaaagaaa atatgaaggc tgtcttggca 840

gacatccaaa atggtaaatt tgcaaatgac tttgtaaatg actataaagc tggacgtcca 900

aaattgactg cttaccgtga acaagcagct aaccttgaaa ttgaaaaagt tggtgcagaa 960

ttgcgtaaag caatgccatt cgttggtaaa aacgatgatg atgcattcaa aatctataac 1020

t 1021

<210> 113
<211> 340
<212> PRT
<213> Streptococcus pneumoniae

<400> 113

Met Thr Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly Arg Asp Val Ile Ile Gly Val
35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Thr Val Thr Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile

100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asn Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
260 265 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
275 280 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
290 295 300

Tyr Arg Glu Gln Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305 310 315 320

Leu Arg Lys Ala Met Pro Phe Val Gly Lys Asn Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Asn
340

<210> 114
<211> 1008
<212> DNA
<213> Streptococcus pneumoniae

<400> 114
atggaatatg aaaaagatgt taaagtagca gcacttgacg gtaaaaaaat cgccgttatc 60

ggttatggtt cacaagggca tgcgcatgct caaaacttgc gtgattcagg tcgtgacgtt 120

attatcggtg tacgtccagg taaatctttt gataaagcaa aagaagatgg atttgatact 180
tacacagtag cagaagctac taagttggct gatgttatca tgatcttggc gccagacgaa 240
attcaacaag aattgtacga agcagaaatc gctccaaact tggaagctgg aaacgcagtt 300
ggatttgccc atggtttcaa catccacttt gaatttatca aagttcctgc ggatgtagat 360
gtcttcatgt gtgctcctaa aggaccagga cacttggtac gtcgtactta cgaagaagga 420
tttggtgttc agctcttta tgcagtatac caagatgcaa caggaaatgc taaaaacatt 480
gctatggact ggtgtaaagg tgttggagcg gctcgtgtag gtcttcttga aacaacttac 540
aaagaagaaa ctgaagaaga tttgtttggt gaacaagctg tactttgtgg tggtttgact 600
gcccttatcg aagcaggttt cgaagtcttg acagaagcag gttacgctcc agaattggct 660
tactttgaag ttcttcacga aatgaaattg atcgttgact tgatctacga aggtggattc 720
aagaaaatgc gtcaatctat ttcaaacact gctgaatacg gtgactatgt atcaggtcca 780
cgtgtaatca ctgaacaagt taaagaaaat atgaaggctg tcttggcaga catccaaaat 840
ggtaaatttg caaatgactt tgtaaatgac tataaagctg gacgtccaaa attgactgct 900
taccgtgaac aagcagctaa ccttgaaatt gaaaaagttg gtgcagaatt gcgtaaagca 960
atgccattcg ttggtaaaaa cgacgatgat gcattcaaaa tctataac 1008

<210> 115
<211> 336
<212> PRT
<213> Streptococcus pneumoniae

<400> 115

Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu Asp Gly Lys Lys
1 5 10 15

Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln Asn
20 25 30

Leu Arg Asp Ser Gly Arg Asp Val Ile Ile Gly Val Arg Pro Gly Lys
35 40 45

Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr Tyr Thr Val Ala
50 55 60

Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu Ala Pro Asp Glu
65 70 75 80

Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro Asn Leu Glu Ala
85 90 95

Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile His Phe Glu Phe
100 105 110

Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys Ala Pro Lys Gly
115 120 125

Pro Gly His Leu Val Arg Arg Thr Tyr Glu Glu Gly Phe Gly Val Pro

```
    130                 135                 140

Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Lys Asn Ile
145                 150                 155                 160

Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg Val Gly Leu Leu
                165                 170                 175

Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu Phe Gly Glu Gln
            180                 185                 190

Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu Ala Gly Phe Glu
        195                 200                 205

Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala Tyr Phe Glu Val
    210                 215                 220

Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly Gly Phe
225                 230                 235                 240

Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu Tyr Gly Asp Tyr
                245                 250                 255

Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys Glu Asn Met Lys
            260                 265                 270

Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala Asn Asp Phe Val
        275                 280                 285

Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala Tyr Arg Glu Gln
    290                 295                 300

Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu Leu Arg Lys Ala
305                 310                 315                 320

Met Pro Phe Val Gly Lys Asn Asp Asp Asp Ala Phe Lys Ile Tyr Asn
                325                 330                 335
```

<210> 116
<211> 1020
<212> DNA
<213> Streptococcus sanguinis

<400> 116

```
atggcagtaa caatggaata cgaaaaagat gttaaagtag cagcgcttga cggtaaaaaa       60

atcgccgtta tcggttatgg ttcacaagga catgcgcatg cacaaaactt gcgtgataca      120

ggccacgatg tgattatcgg tgttcgtcct ggtaagtcat ttgacaaggc taaagaagat      180

ggctttgata cttatacagt agcagaagca gctaaattag ctgatgtcat catgattttg      240

gctccagatg aaatccaaca agatctctat gaagcagaaa tcgctccaaa cttggaagct      300

ggaaacgcag ttgggtttgc tcatggtttc aacatccatt ttgagtttat caaagttcct      360

gccgatgtag atgtctttat gtgtgctcct aaaggccctg gtcacttggt tcgccgtact      420

ttcgaagaag gttttggtgt accagctctg tatgcagtct accaagacgc tactggaaat      480
```

gccaaagaca tcgcaatgga ctggtgtaaa ggtgttggtt cagctcgtgt tggacttctt 540

gaaacaactt ataaagaaga aacagaagaa gatctctttg gtgaacaagc tgtgctttgt 600

ggtggcttga ctgccctgat tgaagcagga tttgaagttc tgactgaagc aggttatgcg 660

ccagaattgg cttactttga agtgctgcac gaaatgaaac tgattgtaga cttgatctat 720

gaaggtggct tcaagaagat gcgccaatct atctcaaata cagctgaata tggtgactat 780

gtatctggtc cgcgcgtcat tactgagcaa gtcaaagaaa acatgaaggc agtcttggct 840

gatatccaaa acggtaaatt tgcgaacgac ttcgttgatg attacaaggc tggtcgtcca 900

aaactcactg cttaccgcga gcaagctgct aatctggaga ttgaaaaggt cggtgcagaa 960

ttgcgtaaag caatgccatt cgttggtaaa aacgacgacg atgctttcaa gatttataat 1020

<210> 117
<211> 340
<212> PRT
<213> Streptococcus sanguinis

<400> 117

Met Ala Val Thr Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Thr Val Ala Glu Ala Ala Lys Leu Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Asp Leu Tyr Glu Ala Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ser Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu

```
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
        195                 200                 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
        275                 280                 285

Asn Asp Phe Val Asp Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Gln Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Lys Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 118
<211> 1020
<212> DNA
<213> Streptococcus salivarius

<400> 118

```
atggcagttc aaatggaata tgaaaaagac gtaaaagtag cagcacttga cggtaaaaaa       60

atcgccgtta tcggttacgg ttcacaaggt catgcacatg cacaaaactt gcgtgataca      120

ggtcacgatg tgattatcgg tgttcgtcca ggtaaatcat ttgataaagc taaagaagat      180

ggttttgaca cttacacagt agcagaagca actaaattgg ctgatgttat catgattttg      240

gctccagacg aaatccaaca agaactttac gaagcagaaa ttgctccaaa ccttgaagct      300

ggtaatgctg ttggttttgc acacggattt aacatccact tcgaattcat caaagttcct      360

aaagatgttg atgtcttcat gtgtgctcct aaaggaccag gtcacttggt acgtcgtaca      420

tttgaagaag gatttggtgt tcctgcactt tatgcagtat accaagatgc tacaggtaac      480

gctaaagata tcgctatgga ctggtgtaaa ggtgttggtg cagcacgtgt tggtcttctt      540

gaaacaacat acaaagaaga aactgaagaa gatttgttcg gtgaacaagc agttctttgt      600

ggtggtttga ctgcccttat cgaagctggt tttgaagtct tgactgaagc tggatatgct      660

ccagaattgg cttactttga agtgcttcac gaaatgaaat tgatcgttga cttgatttac      720
```

gaaggtggat tcaagaaaat gcgtcaatca atttctaaca ctgctgaatt cggtgactac 780

gtatcaggtc cacgtgttat cactgaacaa gtgaaagaaa acatgaaagc agttcttgct 840

gacatccaaa acggtaaatt cgctaacgac ttcgttaacg actacaaagc tggacgtcca 900

aaacttactg cttaccgtga agaagctgct aaccttgaaa tcgaaaaagt tggtgctgaa 960

ttgcgtaaag caatgccatt cgttggtcaa aacgacgatg acgcattcaa aatctataac 1020

<210> 119
<211> 340
<212> PRT
<213> Streptococcus salivarius

<400> 119

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Ala Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Glu Phe Ile Lys Val Pro Lys Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
195 200 205

Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala

```
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
        275                 280                 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 120
<211> 1020
<212> DNA
<213> Streptococcus thermophilus

<400> 120
atggcagttc aaatggaata tgaaaaagac gtaaaagtac cagcacttga tggtaaaaaa 60

atcgccgtta tcggttacgg ttcacaaggt cacgctcact cacaaaactt gcgtgataca 120

ggtcacgatg ttatcatcgg tgttcgtcca ggtaaatcat ttgataaagc taaagaagat 180

ggttttgata cttacacagt agcagaagca actaaattgg ctgatgttat catgattttg 240

gctccagacg aaatccaaca agaactttac gaagcagaaa tcgctccaaa tcttgaagct 300

ggtaacgctg ttggttttgc tcacggtttc aacatccact ttgaatttat taaagtacca 360

gcagatgttg atgtatttat gtgtgcacct aaaggaccag gtcacttggt acgtcgtaca 420

tttgaagaag gttttggtgt acctgccctt tacgcagtat accaagatgc cacaggtaac 480

gctaaagaca tcgctatgga ctggtgtaaa ggtattggtg cagcacgtgt aggtcttctt 540

gaaacaacat acaaagaaga aactgaagaa gatttgtttg gtgaacaagc agttctttgt 600

ggtggtttga ctgcccttat cgaaactggt tttgaagtct tgactgaagc tggatatgca 660

ccagaattgg cttactttga agtacttcac gaaatgaaat taatcgttga cttgatttac 720

gaaggtggat tcaagaaaat gcgtcaatca atttctaaca ctgctgaatt cggtgactat 780

gtatcaggtc cacgtgttat cactgaacaa gttaaagaaa acatgaaagc agttcttgct 840

gatatccaaa atggtaaatt cgctaatgac ttcgttaacg actacaaagc tggacgtcca 900

aaacttactg cttaccgtga agaagctgct aaccttgaaa tcgaaaaagt tggtgctgaa 960

ttgcgtaaag caatgccatt cgttggtcaa aacgacgatg atgcattcaa aatctataac 1020

<210> 121
<211> 340
<212> PRT
<213> Streptococcus thermophilus

<400> 121

Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Pro Ala Leu
1 5 10 15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ser Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
50 55 60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu
65 70 75 80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
85 90 95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly
130 135 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145 150 155 160

Ala Lys Asp Ile Ala Met Asp Trp Cys Lys Gly Ile Gly Ala Ala Arg
165 170 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
195 200 205

Thr Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu

```
                245                 250                 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
            260                 265                 270

Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
        275                 280                 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 122
<211> 975
<212> DNA
<213> Streptococcus pneumoniae

<400> 122

```
ttgacggtaa aaaaatacgc cgttatcggt tatggttcac aagggcatgc gcatgctcaa       60

aacttgcgtg attcaggtcg tgacgttatt ataggtgtac gtccaggtaa atcttttgat      120

aaagcaaaag aagatggatt tgatacttac acagtagcag aagctactaa gttggctgat      180

gttatcatga tcttggcgcc agacgaaatt caacaagaat tgtacgaagc agaaatcgct      240

ccaaacttgg aagctggaaa cgcagttgga tttgcccatg gtttcaacat ccactttgaa      300

tttatcaaag ttcctgcgga tgtagatgtc ttcatgtgtg ctcctaaagg accaggacac      360

ttggtacgtc gtacttacga agaaggattt ggtgttccag ctctttatgc agtataccaa      420

gatgcaacag gaaatgctaa aaacattgct atggactggt gtaaaggtgt tggagcggct      480

cgtgtaggtc ttcttgaaac aacttacaaa gaagaaactg aagaagattt gtttggtgaa      540

caagctgtac tttgtggtgg tttgactgcc cttatcgaag caggtttcga agtcttgaca      600

gaagcaggtt acgctccaga attggcttac tttgaagttc ttcacgaaat gaaattgatc      660

gttgacttga tctacgaagg tggattcaag aaaatgcgtc aatctatttc aaacactgct      720

gaatacggtg actatgtatc aggtccacgt gtaatcactg aacaagttaa agaaaatatg      780

aaggctgtct tggcagacat ccaaaatggt aaatttgcaa atgactttgt aaatgactat      840

aaagctggac gtccaaaatt gactgcttac cgtgaacaag cagctaacct tgaaattgaa      900

aaagttggtg cagaattgcg taaagcaatg ccattcgttg gtaaaaacga cgatgatgca      960

ttcaaaatct ataac                                                       975
```

<210> 123
<211> 325
<212> PRT
<213> Streptococcus pneumoniae

<400> 123

```
Met Thr Val Lys Lys Tyr Ala Val Ile Gly Tyr Gly Ser Gln Gly His
1               5                  10                  15

Ala His Ala Gln Asn Leu Arg Asp Ser Gly Arg Asp Val Ile Ile Gly
            20                  25                  30

Val Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp
        35                  40                  45

Thr Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile
    50                  55                  60

Leu Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala
65                  70                  75                  80

Pro Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn
                85                  90                  95

Ile His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met
            100                 105                 110

Cys Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Glu Glu
        115                 120                 125

Gly Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly
    130                 135                 140

Asn Ala Lys Asn Ile Ala Met Asp Trp Cys Lys Gly Val Gly Ala Ala
145                 150                 155                 160

Arg Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp
                165                 170                 175

Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile
            180                 185                 190

Glu Ala Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu
        195                 200                 205

Ala Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile
    210                 215                 220

Tyr Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala
225                 230                 235                 240

Glu Tyr Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val
                245                 250                 255

Lys Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe
            260                 265                 270

Ala Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr
```

275 280 285

Ala Tyr Arg Glu Gln Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala
290 295 300

Glu Leu Arg Lys Ala Met Pro Phe Val Gly Lys Asn Asp Asp Asp Ala
305 310 315 320

Phe Lys Ile Tyr Asn
325

<210> 124
<211> 1020
<212> DNA
<213> Lactococcus lactis

<400> 124
atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa 60

attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct 120

ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat 180

ggctttgaaa catttgaagt aggagaagcg gtagctaaag ctgatgttat tatggttttg 240

gcgccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300

ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360

gaagacgttg acgtctttat ggttgcacct aaggctccag gtcaccttgt ccgtcggact 420

tatactgaag gttttggtac accagctttg tttgtttcac accaaaatgc aagtggtcat 480

gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattggtt gtgctcgagt gggaattatt 540

gaaacaacct ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt 600

ggaggtttga cagcacttgt tgaagccggt tttgaaacac tgacagaagc tggatacgct 660

ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720

gaaggtggtt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780

gtgactggtc caaggattat tactgacgca gttaaaaaga atatgaagct tgttttggct 840

gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900

aaattaacag cctatcgcga agctgctaaa aatcttgaaa ttgaaaaaat tggggcagaa 960

ttacgtaaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020

<210> 125
<211> 340
<212> PRT
<213> Lactococcus lactis

<400> 125

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val

```
        35                  40                  45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
    50                  55                  60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65                  70                  75                  80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
                85                  90                  95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
        115                 120                 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
    130                 135                 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
145                 150                 155                 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
                165                 170                 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
        195                 200                 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225                 230                 235                 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Ala Val Lys
            260                 265                 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
```

325 330 335

Lys Ile Tyr Gln
340

<210> 126
<211> 1041
<212> DNA
<213> Leuconostoc mesenteroides

<400> 126
atgactacaa aaatgtttta tgacaaagat attgatacaa aaccattgga aaacaaaaag 60

attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct tcgcgactca 120

ggatttgatg tcatcatggg attgcgccca ggaaaatatt ttgatagtgc taaaaaagat 180

gggttcgaag tttactcagc tgctgaagca acggcacaag cggattttgt tatgatggaa 240

acccctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca 300

ggttcttaca ttggattttc tcatgggttc aacattgttt atggtttgat taagcccaat 360

gctgatatta atgtcatgat catcgcgcca aagggtccag gaaacattga acgtcgccaa 420

ttcgttgaag gggttgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat 480

acagctgaag tggccaaggc ctatgcgaaa ggtattggtt caggtcgcgc aggtatcttg 540

gaaacgactt ttgaagaaga aacaacagaa aacttgtttg gtgaacaggc tgtactttgt 600

ggtggcttaa cacaattgat tgaggcagga tttaacacat tggtggaagc aggttactca 660

cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga tttgatcttt 720

gaaggtggtt tcgagaagat gagacatgat tgctcaaata cttgtgaata tggtgaaatg 780

ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag 840

gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag 900

gatttggaaa agatgcggac agagtataag agcggtttgt acgagcaaac aggtaagaag 960

gttcgtgcta tgatgccatg gatttcagat gcagataagt actcgacagc agcagatact 1020

gagcaattct cagcagcaaa g 1041

<210> 127
<211> 347
<212> PRT
<213> Leuconostoc mesenteroides

<400> 127

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Lys Pro Leu
1 5 10 15

Glu Asn Lys Lys Ile Ala Val Ile Gly Tyr Gly Ala Gln Gly His Ala
20 25 30

Gln Ala Asn Asn Leu Arg Asp Ser Gly Phe Asp Val Ile Met Gly Leu
35 40 45

Arg Pro Gly Lys Tyr Phe Asp Ser Ala Lys Lys Asp Gly Phe Glu Val
50 55 60

```
Tyr Ser Ala Ala Glu Ala Thr Ala Gln Ala Asp Phe Val Met Met Glu
65                  70                  75                  80

Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro
                85                  90                  95

Asn Leu Lys Ala Gly Ser Tyr Ile Gly Phe Ser His Gly Phe Asn Ile
            100                 105                 110

Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly
    130                 135                 140

Val Gly Ile Pro Ser Leu Tyr Gly Val His Gln Asp Pro Thr Gly Asp
145                 150                 155                 160

Thr Ala Glu Val Ala Lys Ala Tyr Ala Lys Gly Ile Gly Ser Gly Arg
                165                 170                 175

Ala Gly Ile Leu Glu Thr Thr Phe Glu Glu Glu Thr Thr Glu Asn Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu
        195                 200                 205

Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe
225                 230                 235                 240

Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu
                245                 250                 255

Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys
            260                 265                 270

Gln Gly Met Arg Asp Val Leu Lys Asp Ile Gln Asp Gly Thr Tyr Ala
        275                 280                 285

Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys
    290                 295                 300

Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys
305                 310                 315                 320

Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr
                325                 330                 335

Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys
            340                 345
```

<210> 128
<211> 1041
<212> DNA
<213> Leuconostoc mesenteroides

<400> 128
atgactacaa aaatgtttta tgacaaagat attgatacaa aaccattgga aaacaaaaag 60

attgcagtta ttggctatgg tgcacaaggg catgctcaag cgaataatct tcgcgtctca 120

ggatttgatg tcatcatggg attgcgccca ggaaaatctt ttgatagtgt taaaaaagat 180

gggttcgaag tttactcagc tgctgaagca acggcacaag cggatgttgt tatgatggaa 240

acccctgacg aattacaagc ggcagtttgg gagaaagaag ttgagcctaa ccttaaggca 300

ggttctgacc ttggattttc tcatgggttc aacattgttt atggtttgat taagcccaat 360

gctgatatta atgtcatgat cattgcgcca aagggtccag gaaacattga acgtcgccaa 420

ttcgttgaag ggggtgggat tccttctttg tatggtgttc atcaagatcc tacgggtgat 480

acagctgaag tggccaaggc ctatgcgaaa ggtattggtt caggttgcgc aggtatcttg 540

gaaacgactt ttgaagaaga aacaacagaa gacttgtttg gtgaacaagc tgtactttgt 600

ggtggcttaa cacaattgat tgaggcagga tttaacacat tggtggaagc aggttactca 660

cctgaattgg cttatttcga aacatctcat gaaatgaaga tgattgttga tttgatcttt 720

gaaggtggtt tcgagaagat gagacatgat tgctcaaata cttgtgaata tggtgaaatg 780

ttgaacggac cacgtatcat cactgaagaa tcaaagcaag gaatgcgtga tgttttgaag 840

gatatccaag atggtactta tgctaagaag tggttggcag aatacaattc tggtttgaag 900

gatttggaaa agatgcggac agagtataag agcggtttgt acgagcaaac aggtaagaag 960

gttcgtgcta tgatgccatg gatttcagat gcagataagt actcgacagc agcagatact 1020

gagcaattct cagcagcaaa g 1041

<210> 129
<211> 347
<212> PRT
<213> Leuconostoc mesenteroides

<400> 129

Met Thr Thr Lys Met Phe Tyr Asp Lys Asp Ile Asp Thr Lys Pro Leu
1 5 10 15

Glu Asn Lys Lys Ile Ala Val Ile Gly Tyr Gly Ala Gln Gly His Ala
20 25 30

Gln Ala Asn Asn Leu Arg Val Ser Gly Phe Asp Val Ile Met Gly Leu
35 40 45

Arg Pro Gly Lys Ser Phe Asp Ser Val Lys Lys Asp Gly Phe Glu Val
50 55 60

Tyr Ser Ala Ala Glu Ala Thr Ala Gln Ala Asp Val Val Met Met Glu
65 70 75 80

Thr Pro Asp Glu Leu Gln Ala Ala Val Trp Glu Lys Glu Val Glu Pro

```
                85                  90                  95

Asn Leu Lys Ala Gly Ser Asp Leu Gly Phe Ser His Gly Phe Asn Ile
            100                 105                 110

Val Tyr Gly Leu Ile Lys Pro Asn Ala Asp Ile Asn Val Met Ile Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly Asn Ile Glu Arg Arg Gln Phe Val Glu Gly
    130                 135                 140

Gly Gly Ile Pro Ser Leu Tyr Gly Val His Gln Asp Pro Thr Gly Asp
145                 150                 155                 160

Thr Ala Glu Val Ala Lys Ala Tyr Ala Lys Gly Ile Gly Ser Gly Cys
                165                 170                 175

Ala Gly Ile Leu Glu Thr Thr Phe Glu Glu Glu Thr Thr Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Gln Leu Ile Glu
        195                 200                 205

Ala Gly Phe Asn Thr Leu Val Glu Ala Gly Tyr Ser Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Thr Ser His Glu Met Lys Met Ile Val Asp Leu Ile Phe
225                 230                 235                 240

Glu Gly Gly Phe Glu Lys Met Arg His Asp Cys Ser Asn Thr Cys Glu
                245                 250                 255

Tyr Gly Glu Met Leu Asn Gly Pro Arg Ile Ile Thr Glu Glu Ser Lys
            260                 265                 270

Gln Gly Met Arg Asp Val Leu Lys Asp Ile Gln Asp Gly Thr Tyr Ala
        275                 280                 285

Lys Lys Trp Leu Ala Glu Tyr Asn Ser Gly Leu Lys Asp Leu Glu Lys
    290                 295                 300

Met Arg Thr Glu Tyr Lys Ser Gly Leu Tyr Glu Gln Thr Gly Lys Lys
305                 310                 315                 320

Val Arg Ala Met Met Pro Trp Ile Ser Asp Ala Asp Lys Tyr Ser Thr
                325                 330                 335

Ala Ala Asp Thr Glu Gln Phe Ser Ala Ala Lys
            340                 345
```

<210> 130
<211> 1038
<212> DNA
<213> Lactococcus brevis

<400> 130
atgagtgtag aaatgttgta tgacaaggat gttaccacca attatcttca aggaaagaaa 60

attgctttta tcggttatgg atctcaaggc catgcccaag ctaacaactt aagagattca 120

ggttatgacg ttatcgttgg tgttcgccca ggacaatcat ttgaaaatgc caaaatcgat 180

ggattcgatg tttacacacc agccgaagca gctcgtcgtg cagactggat tcaaatgtta 240

acgcctgatg aagttatggg tgatgtttat aaaaatgaaa ttgcacctaa ccttgaggaa 300

ggcaatgtat taggcttttc acatggcttc aacattcatt ataaagaaat cgtaccacca 360

gcaaacgttg atgttgttat gatggcacct aaaggcccag gtaatctttg tcgtcgtaca 420

tacgttgaag gctccggtgt tccagcatta tatggttatt tccaagacta ttcaggccat 480

gccgaagatt tatccaagga attcgccaaa ggaaatggtg cagcacgtgc cggactattg 540

aagacaacct tcaaagaaga aactgaagaa gatttgtttg gtgagcaaaa cgtccttatg 600

ggtggtgtta ccgcgctcat tgaaaccggt tatgaagtct taaccgaagc aggatactca 660

ccacagttgg catactttga agtagatcat gaaatgaaat tgatctgtga ccttatctac 720

gaaggtgggt tcaacaagat gtatgctgac tgctcgaaca cttctgaata tggttcatat 780

gtagttggtc caaaggtcgt tggtaaggaa tccaagcaag ctatgaaaga cgctttgaag 840

cggattcaag acggttcatt tgctaaagaa ttcatggatg actaccgtaa cggcttcaag 900

aagctttaca agatgcgtga acgttcagct aactcacttc tttcacgtgt tggtgctgat 960

cttcgtgatc acatgtcatt tgttggtgaa gccgacaagt acagtacacc tactgaagaa 1020

aaagcagaag ctgaaaag 1038

<210> 131
<211> 346
<212> PRT
<213> Lactococcus brevis

<400> 131

Met Ser Val Glu Met Leu Tyr Asp Lys Asp Val Thr Thr Asn Tyr Leu
1 5 10 15

Gln Gly Lys Lys Ile Ala Phe Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

Gln Ala Asn Asn Leu Arg Asp Ser Gly Tyr Asp Val Ile Val Gly Val
35 40 45

Arg Pro Gly Gln Ser Phe Glu Asn Ala Lys Ile Asp Gly Phe Asp Val
50 55 60

Tyr Thr Pro Ala Glu Ala Ala Arg Arg Ala Asp Trp Ile Gln Met Leu
65 70 75 80

Thr Pro Asp Glu Val Met Gly Asp Val Tyr Lys Asn Glu Ile Ala Pro
85 90 95

Asn Leu Glu Glu Gly Asn Val Leu Gly Phe Ser His Gly Phe Asn Ile
100 105 110

```
His Tyr Lys Glu Ile Val Pro Pro Ala Asn Val Asp Val Val Met Met
        115                 120                 125

Ala Pro Lys Gly Pro Gly Asn Leu Cys Arg Arg Thr Tyr Val Glu Gly
    130                 135                 140

Ser Gly Val Pro Ala Leu Tyr Gly Tyr Phe Gln Asp Tyr Ser Gly His
145                 150                 155                 160

Ala Glu Asp Leu Ser Lys Glu Phe Ala Lys Gly Asn Gly Ala Ala Arg
                165                 170                 175

Ala Gly Leu Leu Lys Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Asn Val Leu Met Gly Gly Val Thr Ala Leu Ile Glu
        195                 200                 205

Thr Gly Tyr Glu Val Leu Thr Glu Ala Gly Tyr Ser Pro Gln Leu Ala
    210                 215                 220

Tyr Phe Glu Val Asp His Glu Met Lys Leu Ile Cys Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Asn Lys Met Tyr Ala Asp Cys Ser Asn Thr Ser Glu
                245                 250                 255

Tyr Gly Ser Tyr Val Val Gly Pro Lys Val Val Gly Lys Glu Ser Lys
            260                 265                 270

Gln Ala Met Lys Asp Ala Leu Lys Arg Ile Gln Asp Gly Ser Phe Ala
        275                 280                 285

Lys Glu Phe Met Asp Asp Tyr Arg Asn Gly Phe Lys Lys Leu Tyr Lys
    290                 295                 300

Met Arg Glu Arg Ser Ala Asn Ser Leu Leu Ser Arg Val Gly Ala Asp
305                 310                 315                 320

Leu Arg Asp His Met Ser Phe Val Gly Glu Ala Asp Lys Tyr Ser Thr
                325                 330                 335

Pro Thr Glu Glu Lys Ala Glu Ala Glu Lys
            340                 345
```

<210> 132
<211> 1020
<212> DNA
<213> Lactococcus lactis

<400> 132

```
atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa       60

attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct      120

ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat      180
```

ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg 240

gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca 300

ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca 360

gaagacgttg acgtctttat ggttgcgcct aaggctccag gtcaccttgt ccgtcggact 420

tatactgaag gttttggtac accagctttg tttgtttcac accaaaatgc aagtggtcat 480

gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattggtt gtgctcgagt gggaattatt 540

gaaacaactt ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt 600

ggaggtttga cagcacttgt tgaagccggt tttgaaacac tgacagaagc tggatacgct 660

ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat 720

gaaggtggtt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat 780

gtgactggtc cacggattat tactgacgaa gttaaaaaga atatgaagct tgttttggct 840

gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca 900

aaattaatag cctatcgcga agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag 960

ctacgtcaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020

<210> 133
<211> 340
<212> PRT
<213> Lactococcus lactis

<400> 133

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

```
Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
145                 150                 155                 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
                165                 170                 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
        195                 200                 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225                 230                 235                 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
            260                 265                 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala
    290                 295                 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Gln
            340
```

<210> 134
<211> 1002
<212> DNA
<213> Staphylococcus hominis

<400> 134
atgacaaaag tttattatga tcaatcagta gaaaaagatg cactacaagg taaaaaaatc 60

gctattatcg gttacggttc tcaaggccat gcacatgctc aaaaccttaa agataacggt 120

tacgacgtaa gagttggtat tcgtccaggt gattcattta ataaagctaa agaagatgga 180

ttagatgttt atccagtggc tgaagcagtt aaacaagcag acgttattat ggtattactt 240

cctgatgaaa ttcaaggtaa tgtatataaa aatgaaattg ctccaaattt agaagctggt 300

aatgcattag cttttgcaca tggatttaat attcatttta acgtcatcga accacctaaa 360

gatgtagatg tcttcttagt agcacctaaa ggtcctggtc atctagtaag acgtacattt 420

gttgaaggtt cagctgtgcc agcattattt ggtgttcaac aagatgctac aggtcatgca 480

cgtgacattg ctttaagtta tgcaaaaggt attggtgcta cacgtgctgg tgttattgaa 540

acaacattta aagaggaaac tgaaactgac ttattcggtg aacaagctgt actttgtggg 600

ggaattcata aactgattca aagtggtttc gagacattag tggaagctgg atatcaaaaa 660

gaattagcat attttgaagt actacatgaa atgaaactta tcgttgattt aatgtatgaa 720

ggcggtatgg aaaatgtacg ttattcaatt tcaaatactg ctgaatttgg tgactacgta 780

tctggtccac gtgtcattac tccagatgtt aaagacaata tgaaagaagt acttaaagat 840

attcaaaatg gtaatttcgc taatcgcttt gttaaagata acgaaaatgg atttaaagaa 900

ttttatgaat tgcgtgaaca acaacacggt catgaaattg aagcagtggg tcgtgaacta 960

agaaaaatga tgccattcat taaatctaaa agcattcaaa aa 1002

<210> 135
<211> 334
<212> PRT
<213> Staphylococcus hominis

<400> 135

Met Thr Lys Val Tyr Tyr Asp Gln Ser Val Glu Lys Asp Ala Leu Gln
1 5 10 15

Gly Lys Lys Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Gln Asn Leu Lys Asp Asn Gly Tyr Asp Val Arg Val Gly Ile Arg
35 40 45

Pro Gly Asp Ser Phe Asn Lys Ala Lys Glu Asp Gly Leu Asp Val Tyr
50 55 60

Pro Val Ala Glu Ala Val Lys Gln Ala Asp Val Ile Met Val Leu Leu
65 70 75 80

Pro Asp Glu Ile Gln Gly Asn Val Tyr Lys Asn Glu Ile Ala Pro Asn
85 90 95

Leu Glu Ala Gly Asn Ala Leu Ala Phe Ala His Gly Phe Asn Ile His
100 105 110

Phe Asn Val Ile Glu Pro Pro Lys Asp Val Asp Val Phe Leu Val Ala
115 120 125

Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Val Glu Gly Ser
130 135 140

Ala Val Pro Ala Leu Phe Gly Val Gln Gln Asp Ala Thr Gly His Ala
145 150 155 160

Arg Asp Ile Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Ala
165 170 175

Gly Val Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe
180 185 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Ile His Lys Leu Ile Gln Ser
195 200 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Gln Lys Glu Leu Ala Tyr
210 215 220

Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr Glu
225 230 235 240

Gly Gly Met Glu Asn Val Arg Tyr Ser Ile Ser Asn Thr Ala Glu Phe
245 250 255

Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Pro Asp Val Lys Asp
260 265 270

Asn Met Lys Glu Val Leu Lys Asp Ile Gln Asn Gly Asn Phe Ala Asn
275 280 285

Arg Phe Val Lys Asp Asn Glu Asn Gly Phe Lys Glu Phe Tyr Glu Leu
290 295 300

Arg Glu Gln Gln His Gly His Glu Ile Glu Ala Val Gly Arg Glu Leu
305 310 315 320

Arg Lys Met Met Pro Phe Ile Lys Ser Lys Ser Ile Gln Lys
325 330

<210> 136
<211> 11
<212> DNA
<213> Artificial sequence

<220>
<223> PmeI restriction site and additional 3 bp (AGT)

<400> 136
gtttaaacag t 11

<210> 137
<211> 13
<212> DNA
<213> Artificial sequence

<220>
<223> SfiI restriction site

<400> 137
ggccctgcag gcc 13

<210> 138
<211> 9586
<212> DNA
<213> Artificial sequence

<220>
<223> pLH556

<400> 138
ggccgcacct ggtaaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa 60
aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac 120
agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat 180
ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa 240
tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta 300
cccgacggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat 360
aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata 420
tacaaaatat cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta 480
gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtgaaca ccttatataa 540
cttagcccgt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag 600
acgtgataaa aatttacttt ctaactcttc tcacgctgcc cctatctgtt cttccgctct 660
accgtgagaa ataaagcatc gagtacggca gttcgctgtc actgaactaa aacaataagg 720
ctagttcgaa tgatgaactt gcttgctgtc aaacttctga gttgccgctg atgtgacact 780
gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca 840
atagagctca gtaggagtca gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg 900
taagttgtgc gcgtgcacat ttcgcccgtt cccgctcatc ttgcagcagg cggaaatttt 960
catcacgctg taggacgcaa aaaaaaaata attaatcgta caagaatctt ggaaaaaaaa 1020
ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acacccagta 1080
ttttcccttt ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta 1140
ttcctaggag ttatattttt ttaccctacc agcaatataa gtaaaaaact gtttaaacag 1200
tatgaaagtt ttctacgata aagactgcga cctgtcgatc atccaaggta agaaagttgc 1260
catcatcggc tacggttccc agggccacgc tcaagcatgc aacctgaagg attccggcgt 1320
agacgtgact gttggcctgc gtaaaggctc ggctaccgtt gccaaggctg aagcccacgg 1380
cttgaaagtg accgacgttg ctgcagccgt tgccggtgcc gacttggtca tgatcctgac 1440
cccggacgag ttccagtccc agctgtacaa gaacgaaatc gagccgaaca tcaagaaggg 1500
cgccactctg gccttctccc acggcttcgc gatccactac aaccaggttg tgcctcgtgc 1560
cgacctcgac gtgatcatga tcgcgccgaa ggctccaggc cacaccgtac gttccgagtt 1620
cgtcaagggc ggtggtattc ctgacctgat cgcgatctac caggacgctt ccggcaacgc 1680
caagaacgtt gccctgtcct acgccgcagg cgtgggcggc ggccgtaccg gcatcatcga 1740
aaccaccttc aaggacgaga ctgaaaccga cctgttcggt gagcaggctg ttctgtgtgg 1800
cggtaccgtc gagctggtca aagccggttt cgaaaccctg gttgaagctg gctacgctcc 1860
agaaatggcc tacttcgagt gcctgcacga actgaagctg atcgttgacc tcatgtacga 1920
aggcggtatc gccaacatga actactcgat ctccaacaac gctgaatacg gcgagtacgt 1980
gactggtcca gaagtcatca acgccgaatc ccgtcaggcc atgcgcaatg ctctgaagcg 2040
catccaggac ggcgaatacg cgaagatgtt catcagcgaa ggcgctaccg gctacccatc 2100

gatgaccgcc aagcgtcgta acaacgctgc tcacggtatc gaaatcatcg gcgagcaact 2160
gcgctcgatg atgccttgga tcggtgccaa caaaatcgtc gacaaagcca agaactaagg 2220
ccctgcaggc ctatcaagtg ctggaaactt tttctcttgg aatttttgca acatcaagtc 2280
atagtcaatt gaattgaccc aatttcacat ttaagatttt tttttttca tccgacatac 2340
atctgtacac taggaagccc tgtttttctg aagcagcttc aaatatatat attttttaca 2400
tatttattat gattcaatga acaatctaat taaatcgaaa acaagaaccg aaacgcgaat 2460
aaataattta tttagatggt gacaagtgta taagtcctca tcgggacagc tacgatttct 2520
ctttcggttt tggctgagct actggttgct gtgacgcagc ggcattagcg cggcgttatg 2580
agctaccctc gtggcctgaa agatggcggg aataaagcgg aactaaaaat tactgactga 2640
gccatattga ggtcaatttg tcaactcgtc aagtcacgtt tggtggacgg ccccttttcca 2700
acgaatcgta tatactaaca tgcgcgcgct tcctatatac acatatacat atatatatat 2760
atatatatgt gtgcgtgtat gtgtacacct gtatttaatt tccttactcg cgggtttttc 2820
ttttttctca attcttggct tcctctttct cgagcggacc ggatcctccg cggtgccggc 2880
agatctattt aaatggcgcg ccgacgtcag gtggcacttt tcggggaaat gtgcgcggaa 2940
cccctatttg tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac 3000
cctgataaat gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg 3060
tcgcccttat tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc 3120
tggtgaaagt aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg 3180
atctcaacag cggtaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga 3240
gcacttttaa agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc 3300
aactcggtcg ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag 3360
aaaagcatct tacggatggc atgacagtaa gagaattatg cagtgctgcc ataaccatga 3420
gtgataacac tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg 3480
ctttttttgca caacatgggg gatcatgtaa ctcgccttga tcgttgggaa ccggagctga 3540
atgaagccat accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt 3600
tgcgcaaact attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact 3660
ggatggaggc ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt 3720
ttattgctga taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg 3780
ggccagatgg taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta 3840
tggatgaacg aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac 3900
tgtcagacca agtttactca tatatacttt agattgattt aaaacttcat ttttaattta 3960
aaaggatcta ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt 4020
tttcgttcca ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt 4080
tttttctgcg cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt 4140
gtttgccgga tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc 4200
agataccaaa tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg 4260

tagcaccgcc tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg 4320
ataagtcgtg tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggt 4380
cgggctgaac gggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac 4440
tgagatacct acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg 4500
acaggtatcc ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg 4560
gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat 4620
ttttgtgatg ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt 4680
tacggttcct ggccttttgc tggccttttg ctcacatgtt ctttcctgcg ttatcccctg 4740
attctgtgga taaccgtatt accgccttg agtgagctga taccgctcgc cgcagccgaa 4800
cgaccgagcg cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc 4860
ctctccccgc gcgttggccg attcattaat gcagctggca cgacaggttt cccgactgga 4920
aagcgggcag tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg 4980
ctttacactt tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc 5040
acacaggaaa cagctatgac catgattacg ccaagctttt tctttccaat ttttttttt 5100
tcgtcattat aaaaatcatt acgaccgaga ttcccgggta ataactgata taattaaatt 5160
gaagctctaa tttgtgagtt tagtatacat gcatttactt ataatacagt tttttagttt 5220
tgctggccgc atcttctcaa atatgcttcc cagcctgctt ttctgtaacg ttcaccctct 5280
accttagcat cccttccctt tgcaaatagt cctcttccaa caataataat gtcagatcct 5340
gtagagacca catcatccac ggttctatac tgttgaccca atgcgtctcc cttgtcatct 5400
aaacccacac cgggtgtcat aatcaaccaa tcgtaacctt catctcttcc acccatgtct 5460
ctttgagcaa taaagccgat aacaaaatct ttgtcgctct tcgcaatgtc aacagtaccc 5520
ttagtatatt ctccagtaga tagggagccc ttgcatgaca attctgctaa catcaaaagg 5580
cctctaggtt cctttgttac ttcttctgcc gcctgcttca aaccgctaac aatacctggg 5640
cccaccacac cgtgtgcatt cgtaatgtct gcccattctg ctattctgta tacacccgca 5700
gagtactgca atttgactgt attaccaatg tcagcaaatt ttctgtcttc gaagagtaaa 5760
aaattgtact tggcggataa tgcctttagc ggcttaactg tgccctccat ggaaaaatca 5820
gtcaagatat ccacatgtgt ttttagtaaa caaattttgg gacctaatgc ttcaactaac 5880
tccagtaatt ccttggtggt acgaacatcc aatgaagcac acaagtttgt ttgcttttcg 5940
tgcatgatat taaatagctt ggcagcaaca ggactaggat gagtagcagc acgttcctta 6000
tatgtagctt tcgacatgat ttatcttcgt ttcctgcagg tttttgttct gtgcagttgg 6060
gttaagaata ctgggcaatt tcatgtttct tcaacactac atatgcgtat atataccaat 6120
ctaagtctgt gctccttcct tcgttcttcc ttctgttcgg agattaccga atcaaaaaaa 6180
tttcaaggaa accgaaatca aaaaaaagaa taaaaaaaaa atgatgaatt gaaaagcttg 6240
catgcctgca ggtcgactct agtatactcc gtctactgta cgatacactt ccgctcaggt 6300
ccttgtcctt taacgaggcc ttaccactct tttgttactc tattgatcca gctcagcaaa 6360
ggcagtgtga tctaagattc tatcttcgcg atgtagtaaa actagctaga ccgagaaaga 6420

gactagaaat gcaaaaggca cttctacaat ggctgccatc attattatcc gatgtgacgc 6480
tgcatttttt ttttttttt ttttttttt ttttttttt ttttttttt ttttttgtac 6540
aaatatcata aaaaaagaga atctttttaa gcaaggattt tcttaacttc ttcggcgaca 6600
gcatcaccga cttcggtggt actgttggaa ccacctaaat caccagttct gatacctgca 6660
tccaaaacct ttttaactgc atcttcaatg gctttacctt cttcaggcaa gttcaatgac 6720
aatttcaaca tcattgcagc agacaagata gtggcgatag ggttgacctt attctttggc 6780
aaatctggag cggaaccatg gcatggttcg tacaaaccaa atgcggtgtt cttgtctggc 6840
aaagaggcca aggacgcaga tggcaacaaa cccaaggagc ctgggataac ggaggcttca 6900
tcggagatga tatcaccaaa catgttgctg gtgattataa taccatttag gtgggttggg 6960
ttcttaacta ggatcatggc ggcagaatca atcaattgat gttgaacttt caatgtaggg 7020
aattcgttct tgatggtttc ctccacagtt tttctccata atcttgaaga ggccaaaaca 7080
ttagctttat ccaaggacca aataggcaat ggtggctcat gttgtagggc catgaaagcg 7140
gccattcttg tgattctttg cacttctgga acggtgtatt gttcactatc ccaagcgaca 7200
ccatcaccat cgtcttcctt tctcttacca aagtaaatac ctcccactaa ttctctaaca 7260
acaacgaagt cagtaccttt agcaaattgt ggcttgattg gagataagtc taaaagagag 7320
tcggatgcaa agttacatgg tcttaagttg gcgtacaatt gaagttcttt acggattttt 7380
agtaaacctt gttcaggtct aacactaccg gtaccccatt taggaccacc cacagcacct 7440
aacaaaacgg catcagcctt cttggaggct tccagcgcct catctggaag tggaacacct 7500
gtagcatcga tagcagcacc accaattaaa tgattttcga aatcgaactt gacattggaa 7560
cgaacatcag aaatagcttt aagaacctta atggcttcgg ctgtgatttc ttgaccaacg 7620
tggtcacctg gcaaaacgac gatcttctta ggggcagaca ttacaatggt atatccttga 7680
aatatatata aaaaaaaaaa aaaaaaaaaa aaaaaaaaat gcagcttctc aatgatattc 7740
gaatacgctt tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc 7800
gtacttgtta cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag 7860
atagtatatt tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat 7920
gtatttcggt tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat 7980
ccccggttca ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag 8040
catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa 8100
agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac 8160
gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca 8220
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta 8280
ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt 8340
ttctaacaaa gcatcttaga ttacttttttt tctcctttgt gcgctctata atgcagtctc 8400
ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat 8460
tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc 8520
tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt 8580

gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta 8640
tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt 8700
attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat 8760
actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa 8820
ggtggatggg taggttatat agggatatag cacagagata tatagcaaag agatactttt 8880
gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg 8940
tttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag 9000
ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa 9060
cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac 9120
ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt 9180
atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct 9240
cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta 9300
gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt 9360
cctttgatat tggatcatat gcatagtacc gagaaactag aggatctccc attaccgaca 9420
tttgggcgct atacgtgcat atgttcatgt atgtatctgt atttaaaaca cttttgtatt 9480
atttttcctc atatatgtgt ataggtttat acggatgatt taattattac ttcaccaccc 9540
tttatttcag gctgatatct tagccttgtt actagtcacc ggtggc 9586

<210> 139
<211> 15539
<212> DNA
<213> Artificial sequence

<220>
<223> Plasmid

<400> 139
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360
ttttttttt ccacctagcg gatgactctt ttttttttctt agcgattggc attatcacat 420
aatgaattat acattatata aagtaatgtg atttcttcga agaatatact aaaaaatgag 480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca 540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660
attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat 720
tccggctggt cgctaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt 840

ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcggtg 900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020
attaccctcc acgttgattg tctgcgaggc aagaatgatc atcaccgtag tgagagtgcg 1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgat 1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg 1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg 1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1440
aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 1620
aaaaccgtct atcagggcga tggcccacta cgtgaaccat caccctaatc aagtttttg 1680
gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagcccccg atttagagct 1740
tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc 1800
gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt 1860
aatgcgccgc tacagggcgc gtccattcgc cattcaggct gcgcaactgt tgggaagggc 1920
gcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 1980
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgag 2040
cgcgcgtaat acgactcact atagggcgaa ttgggtaccg ggcccccccct cgaggtcgac 2100
ggcgcgccac tggtagagag cgactttgta tgccccaatt gcgaaacccg cgatatcctt 2160
ctcgattctt tagtacccga ccaggacaag gaaaaggagg tcgaaacgtt tttgaagaaa 2220
caagaggaac tacacggaag ctctaaagat ggcaaccagc cagaaactaa gaaaatgaag 2280
ttgatggatc caactggcac cgctggcttg aacaacaata ccagccttcc aacttctgta 2340
aataacggcg gtacgccagt gccaccagta ccgttacctt tcggtatacc tcctttcccc 2400
atgtttccaa tgcccttcat gcctccaacg gctactatca caaatcctca tcaagctgac 2460
gcaagcccta agaaatgaat aacaatactg acagtactaa ataattgcct acttggcttc 2520
acatacgttg catacgtcga tatagataat aatgataatg acagcaggat tatcgtaata 2580
cgtaatagct gaaaatctca aaaatgtgtg ggtcattacg taaataatga taggaatggg 2640
attcttctat ttttcctttt tccattctag cagccgtcgg gaaaacgtgg catcctctct 2700
ttcgggctca attggagtca cgctgccgtg agcatcctct ctttccatat ctaacaactg 2760
agcacgtaac caatggaaaa gcatgagctt agcgttgctc caaaaaagta ttggatggtt 2820
aataccattt gtctgttctc ttctgacttt gactcctcaa aaaaaaaaat ctacaatcaa 2880
cagatcgctt caattacgcc ctcacaaaaa ctttttttcct tcttcttcgc ccacgttaaa 2940
ttttatccct catgttgtct aacggatttc tgcacttgat ttattataaa aagacaaaga 3000

cataatactt ctctatcaat ttcagttatt gttcttcctt gcgttattct tctgttcttc 3060
tttttctttt gtcatatata accataacca agtaatacat attcaaacta gtatgactga 3120
caaaaaaact cttaaagact taagaaatcg tagttctgtt tacgattcaa tggttaaatc 3180
acctaatcgt gctatgttgc gtgcaactgg tatgcaagat gaagactttg aaaaacctat 3240
cgtcggtgtc atttcaactt gggctgaaaa cacaccttgt aatatccact tacatgactt 3300
tggtaaacta gccaaagtcg gtgttaagga agctggtgct tggccagttc agttcggaac 3360
aatcacggtt tctgatggaa tcgccatggg aacccaagga atgcgtttct ccttgacatc 3420
tcgtgatatt attgcagatt ctattgaagc agccatggga ggtcataatg cggatgcttt 3480
tgtagccatt ggcggttgtg ataaaaacat gcccggttct gttatcgcta tggctaacat 3540
ggatatccca gccatttttg cttacggcgg aacaattgca cctggtaatt tagacggcaa 3600
agatatcgat ttagtctctg tctttgaagg tgtcggccat tggaaccacg gcgatatgac 3660
caaagaagaa gttaaagctt tggaatgtaa tgcttgtccc ggtcctggag gctgcggtgg 3720
tatgtatact gctaacacaa tggcgacagc tattgaagtt ttgggactta gccttccggg 3780
ttcatcttct cacccggctg aatccgcaga aaagaaagca gatattgaag aagctggtcg 3840
cgctgttgtc aaaatgctcg aaatgggctt aaaaccttct gacattttaa cgcgtgaagc 3900
ttttgaagat gctattactg taactatggc tctgggaggt tcaaccaact caacccttca 3960
cctcttagct attgcccatg ctgctaatgt ggaattgaca cttgatgatt tcaatacttt 4020
ccaagaaaaa gttcctcatt tggctgattt gaaaccttct ggtcaatatg tattccaaga 4080
cctttacaag gtcggagggg taccagcagt tatgaaatat ctccttaaaa atggcttcct 4140
tcatggtgac cgtatcactt gtactggcaa aacagtcgct gaaaatttga aggcttttga 4200
tgatttaaca cctggtcaaa aggttattat gccgcttgaa aatcctaaac gtgaagatgg 4260
tccgctcatt attctccatg gtaacttggc tccagacggt gccgttgcca aagtttctgg 4320
tgtaaaagtg cgtcgtcatg tcggtcctgc taaggtcttt aattctgaag aagaagccat 4380
tgaagctgtc ttgaatgatg atattgttga tggtgatgtt gttgtcgtac gttttgtagg 4440
accaaagggc ggtcctggta tgcctgaaat gctttccctt tcatcaatga ttgttggtaa 4500
agggcaaggt gaaaaagttg cccttctgac agatggccgc ttctcaggtg gtacttatgg 4560
tcttgtcgtg ggtcatatcg ctcctgaagc acaagatggc ggtccaatcg cctacctgca 4620
aacaggagac atagtcacta ttgaccaaga cactaaggaa ttacactttg atatctccga 4680
tgaagagtta aaacatcgtc aagagaccat tgaattgcca ccgctctatt cacgcggtat 4740
ccttggtaaa tatgctcaca tcgtttcgtc tgcttctagg ggagccgtaa cagacttttg 4800
gaagcctgaa gaaactggca aaaaatgttg tcctggttgc tgtggttaag cggccgcgtt 4860
aattcaaatt aattgatata gtttttttaat gagtattgaa tctgtttaga aataatggaa 4920
tattattttt atttatttat ttatattatt ggtcggctct tttcttctga aggtcaatga 4980
caaaatgata tgaaggaaat aatgatttct aaaattttac aacgtaagat atttttacaa 5040
aagcctagct catcttttgt catgcactat tttactcacg cttgaaatta acggccagtc 5100
cactgcggag tcatttcaaa gtcatcctaa tcgatctatc gtttttgata gctcattttg 5160

gagttcgcga ttgtcttctg ttattcacaa ctgttttaat ttttatttca ttctggaact 5220
cttcgagttc tttgtaaagt ctttcatagt agcttacttt atcctccaac atatttaact 5280
tcatgtcaat ttcggctctt aaattttcca catcatcaag ttcaacatca tcttttaact 5340
tgaatttatt ctctagctct tccaaccaag cctcattgct ccttgattta ctggtgaaaa 5400
gtgatacact ttgcgcgcaa tccaggtcaa aactttcctg caaagaattc accaatttct 5460
cgacatcata gtacaatttg ttttgttctc ccatcacaat ttaatatacc tgatggattc 5520
ttatgaagcg ctgggtaatg gacgtgtcac tctacttcgc ctttttccct actcctttta 5580
gtacggaaga caatgctaat aaataagagg gtaataataa tattattaat cggcaaaaaa 5640
gattaaacgc caagcgttta attatcagaa agcaaacgtc gtaccaatcc ttgaatgctt 5700
cccaattgta tattaagagt catcacagca acatattctt gttattaaat taattattat 5760
tgatttttga tattgtataa aaaaaccaaa tatgtataaa aaaagtgaat aaaaaatacc 5820
aagtatggag aaatatatta gaagtctata cgttaaacca cccgggcccc ccctcgaggt 5880
cgacggtatc gataagcttg atatcgaatt cctgcagccc gggggatcca ctagttctag 5940
agcggccgct ctagaactag taccacaggt gttgtcctct gaggacataa aatacacacc 6000
gagattcatc aactcattgc tggagttagc atatctacaa ttgggtgaaa tggggagcga 6060
tttgcaggca tttgctcggc atgccggtag aggtgtggtc aataagagcg acctcatgct 6120
atacctgaga aagcaacctg acctacagga aagagttact caagaataag aattttcgtt 6180
ttaaaaccta agagtcactt taaaatttgt atacacttat ttttttata acttatttaa 6240
taataaaaat cataaatcat aagaaattcg cttactctta attaatcaaa aagttaaaat 6300
tgtacgaata gattcaccac ttcttaacaa atcaaaccct tcattgattt tctcgaatgg 6360
caatacatgt gtaattaaag gatcaagagc aaacttcttc gccataaagt cggcaacaag 6420
ttttggaaca ctatccttgc tcttaaaacc gccaaatata gctcccttcc atgtacgacc 6480
gcttagcaac agcataggat tcatcgacaa attttgtgaa tcaggaggaa cacctacgat 6540
cacactgact ccatatgcct cttgacagca ggacaacgca gttaccatag tatcaagacg 6600
gcctataact tcaaaagaga aatcaactcc accgtttgac atttcagtaa ggacttcttg 6660
tattggtttc ttataatctt gagggttaac acattcagta gccccgacct ccttagcttt 6720
tgcaaatttg tccttattga tgtctacacc tataatcctc gctgcgcctg cagctttaca 6780
ccccataata acgcttagtc ctactcctcc taaaccgaat actgcacaag tcgaaccctg 6840
tgtaaccttt gcaactttaa ctgcggaacc gtaaccggtg gaaaatccgc accctatcaa 6900
gcaaactttt tccagtggtg aagctgcatc gattttagcg acagatatct cgtccaccac 6960
tgtgtattgg gaaaatgtag aagtaccaag gaaatggtgt ataggtttcc ctctgcatgt 7020
aaatctgctt gtaccatcct gcatagtacc tctaggcata gacaaatcat ttttaaggca 7080
gaaattaccc tcaggatgtt tgcagactct acacttacca cattgaggag tgaacagtgg 7140
gatcacttta tcaccaggac gaacagtggt aacaccttca cctatggatt caacgattcc 7200
ggcagcctcg tgtcccgcga ttactggcaa aggagtaact agagtgccac tcaccacatg 7260
gtcgtcggat ctacagattc cggtggcaac catcttgatt ctaacctcgt gtgcttttgg 7320

tggcgctact tctacttctt ctatgctaaa cggctttttc tcttcccaca aaactgccgc 7380
tttacactta ataactttac cggctgttga catcctcagc tagctattgt aatatgtgtg 7440
tttgtttgga ttattaagaa gaataattac aaaaaaaatt acaaaggaag gtaattacaa 7500
cagaattaag aaaggacaag aaggaggaag agaatcagtt cattatttct tctttgttat 7560
ataacaaacc caagtagcga tttggccata cattaaaagt tgagaaccac cctccctggc 7620
aacagccaca actcgttacc attgttcatc acgatcatga aactcgctgt cagctgaaat 7680
ttcacctcag tggatctctc tttttattct tcatcgttcc actaaccttt ttccatcagc 7740
tggcagggaa cggaaagtgg aatcccattt agcgagcttc ctcttttctt caagaaaaga 7800
cgaagcttgt gtgtgggtgc gcgcgctagt atctttccac attaagaaat ataccataaa 7860
ggttacttag acatcactat ggctatatat atatatatat atatatgtaa cttagcacca 7920
tcgcgcgtgc atcactgcat gtgttaaccg aaaagtttgg cgaacacttc accgacacgg 7980
tcatttagat ctgtcgtctg cattgcacgt cccttagcct taaatcctag gcgggagcat 8040
tctcgtgtaa ttgtgcagcc tgcgtagcaa ctcaacatag cgtagtctac ccagtttttc 8100
aagggtttat cgttagaaga ttctcccttt tcttcctgct cacaaatctt aaagtcatac 8160
attgcacgac taaatgcaag catgcggatc ccccgggctg caggaattcg atatcaagct 8220
tatcgatacc gtcgactggc cattaatctt tcccatatta gatttcgcca agccatgaaa 8280
gttcaagaaa ggtctttaga cgaattaccc ttcatttctc aaactggcgt caagggatcc 8340
tggtatggtt ttatcgtttt atttctggtt cttatagcat cgttttggac ttctctgttc 8400
ccattaggcg gttcaggagc cagcgcagaa tcattctttg aaggatactt atcctttcca 8460
attttgattg tctgttacgt tggacataaa ctgtatacta gaaattggac tttgatggtg 8520
aaactagaag atatggatct tgataccggc agaaaacaag tagatttgac tcttcgtagg 8580
gaagaaatga ggattgagcg agaaacatta gcaaaaagat ccttcgtaac aagattttta 8640
catttctggt gttgaaggga aagatatgag ctatacagcg gaatttccat atcactcaga 8700
ttttgttatc taatttttc cttcccacgt ccgcgggaat ctgtgtatat tactgcatct 8760
agatatatgt tatcttatct tggcgcgtac atttaatttt caacgtattc tataagaaat 8820
tgcgggagtt tttttcatgt agatgatact gactgcacgc aaatataggc atgatttata 8880
ggcatgattt gatggctgta ccgataggaa cgctaagagt aacttcagaa tcgttatcct 8940
ggcggaaaaa attcatttgt aaactttaaa aaaaaaagcc aatatcccca aaattattaa 9000
gagcgcctcc attattaact aaaatttcac tcagcatcca caatgtatca ggtatctact 9060
acagatatta catgtggcga aaaagacaag aacaatgcaa tagcgcatca agaaaaaaca 9120
caaagctttc aatcaatgaa tcgaaaatgt cattaaaata gtatataaat tgaaactaag 9180
tcataaagct ataaaaagaa aatttattta aatgcaagat ttaaagtaaa ttcacggccc 9240
tgcaggcctc agctcttgtt ttgttctgca aataacttac ccatcttttt caaaacttta 9300
ggtgcaccct cctttgctag aataagttct atccaataca tcctatttgg atctgcttga 9360
gcttctttca tcacggatac gaattcattt tctgttctca caattttgga cacaactctg 9420
tcttccgttg ccccgaaact ttctggcagt tttgagtaat tccacatagg aatgtcatta 9480

taactctggt tcggaccatg aatttccctc tcaaccgtgt aaccatcgtt attaatgata 9540
aagcagattg ggtttatctt ctctctaatg gctagtccta attcttggac agtcagttgc 9600
aatgatccat ctccgataaa caataaatgt ctagattctt tatctgcaat ttggctgcct 9660
agagctgcgg ggaaagtgta tcctatagat ccccacaagg gttgaccaat aaaatgtgat 9720
ttcgatttca gaaatataga tgaggcaccg aagaaagaag tgccttgttc agccacgatc 9780
gtctcattac tttgggtcaa attttcgaca gcttgccaca gtctatcttg tgacaacagc 9840
gcgttagaag gtacaaaatc ttcttgcttt ttatctatgt acttgccttt atattcaatt 9900
tcggacaagt caagaagaga tgatatcagg gattcgaagt cgaaattttg gattctttcg 9960
ttgaaaattt taccttcatc gatattcaag gaaatcattt tattttcatt aagatggtga 10020
gtaaatgcac ccgtactaga atcggtaagc tttacaccca acataagaat aaaatcagca 10080
gattccacaa attccttcaa gtttggctct gacagagtac cgttgtaaat ccccaaaaat 10140
gagggcaatg cttcatcaac agatgattta ccaaagttca aagtagtaat aggtaactta 10200
gtctttgaaa taaactgagt aacagtcttc tctaggccga acgatataat ttcatggcct 10260
gtgattacaa ttggtttctt ggcattcttc agactttcct gtattttgtt cagaatctct 10320
tgatcagatg tattcgacgt ggaattttcc ttcttaagag gcaaggatgg tttttcagcc 10380
ttagcggcag ctacatctac aggtaaattg atgtaaaccg gctttctttc ctttagtaag 10440
gcagacaaca ctctatcaat ttcaacagtt gcattctcgg ctgtcaataa agtcctggca 10500
gcagtaaccg gttcgtgcat cttcataaag tgcttgaaat caccatcagc caacgtatgg 10560
tgaacaaact taccttcgtt ctgcactttc gaggtaggag atcccacgat ctcaacaaca 10620
ggcaggttct cagcatagga gcccgctaag ccattaactg cggataattc gccaacacca 10680
aatgtagtca agaatgccgc agccttttc gttcttgcgt acccgtcggc catataggag 10740
gcatttaact cattagcatt tcccacccat ttcatatctt tgtgtgaaat aatttgatct 10800
agaaattgca aattgtagtc acctggtact ccgaatattt cttctatacc taattcgtgt 10860
aatctgtcca acagatagtc acctactgta tacattttgt ttactagttt atgtgtgttt 10920
attcgaaact aagttcttgg tgttttaaaa ctaaaaaaaa gactaactat aaaagtagaa 10980
tttaagaagt ttaagaaata gatttacaga attacaatca atacctaccg tctttatata 11040
cttattagtc aagtagggga ataatttcag ggaactggtt tcaacctttt ttttcagctt 11100
tttccaaatc agagagagca gaaggtaata gaaggtgtaa gaaaatgaga tagatacatg 11160
cgtgggtcaa ttgccttgtg tcatcattta ctccaggcag gttgcatcac tccattgagg 11220
ttgtgcccgt tttttgcctg tttgtgcccc tgttctctgt agttgcgcta agagaatgga 11280
cctatgaact gatggttggt gaagaaaaca atattttggt gctgggattc ttttttttc 11340
tggatgccag cttaaaaagc gggctccatt atatttagtg gatgccagga ataaactgtt 11400
cacccagaca cctacgatgt tatatattct gtgtaacccg ccccctattt tgggcatgta 11460
cgggttacag cagaattaaa aggctaattt tttgactaaa taaagttagg aaaatcacta 11520
ctattaatta tttacgtatt ctttgaaatg gcagtattga taatgataaa ctcgaactga 11580
aaaagcgtgt tttttattca aaatgattct aactccctta cgtaatcaag gaatcttttt 11640

gccttggcct ccgcgtcatt aaacttcttg ttgttgacgc taacattcaa cgctagtata 11700
tattcgtttt tttcaggtaa gttcttttca acgggtctta ctgatgaggc agtcgcgtct 11760
gaacctgtta agaggtcaaa tatgtcttct tgaccgtacg tgtcttgcat gttattagct 11820
ttgggaattt gcatcaagtc ataggaaaat ttaaatcttg gctctcttgg gctcaaggtg 11880
acaaggtcct cgaaaatagg gcgcgcccca ccgcggtgga gctccagctt ttgttccctt 11940
tagtgagggt taattgcgcg cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat 12000
tgttatccgc tcacaattcc acacaacata cgagccggaa gcataaagtg taaagcctgg 12060
ggtgcctaat gagtgagcta actcacatta attgcgttgc gctcactgcc cgctttccag 12120
tcgggaaacc tgtcgtgcca gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 12180
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 12240
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 12300
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 12360
gccgcgttgc tggcgttttt ccataggctc cgccccctg acgagcatca caaaaatcga 12420
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 12480
ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 12540
tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 12600
gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 12660
tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 12720
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 12780
ttcttgaagt ggtggcctaa ctacggctac actagaagaa cagtatttgg tatctgcgct 12840
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 12900
accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 12960
tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 13020
cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat 13080
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 13140
caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 13200
gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt 13260
gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 13320
ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct 13380
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt 13440
gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc 13500
tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt 13560
agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg 13620
gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg 13680
actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct 13740
tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc 13800

attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt 13860
tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt 13920
tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg 13980
aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta tcagggttat 14040
tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg 14100
cgcacatttc cccgaaaagt gccacctgaa cgaagcatct gtgcttcatt ttgtagaaca 14160
aaaatgcaac gcgagagcgc taatttttca aacaaagaat ctgagctgca tttttacaga 14220
acagaaatgc aacgcgaaag cgctatttta ccaacgaaga atctgtgctt cattttttgta 14280
aaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag ctgcattttt 14340
acagaacaga aatgcaacgc gagagcgcta ttttaccaac aaagaatcta tacttctttt 14400
ttgttctaca aaaatgcatc ccgagagcgc tatttttcta acaaagcatc ttagattact 14460
ttttttctcc tttgtgcgct ctataatgca gtctcttgat aacttttttgc actgtaggtc 14520
cgttaaggtt agaagaaggc tactttggtg tctattttct cttccataaa aaaagcctga 14580
ctccacttcc cgcgtttact gattactagc gaagctgcgg gtgcattttt tcaagataaa 14640
ggcatccccg attatattct ataccgatgt ggattgcgca tactttgtga acagaaagtg 14700
atagcgttga tgattcttca ttggtcagaa aattatgaac ggtttcttct attttgtctc 14760
tatatactac gtataggaaa tgtttacatt ttcgtattgt tttcgattca ctctatgaat 14820
agttcttact acaatttttt tgtctaaaga gtaatactag agataaacat aaaaaatgta 14880
gaggtcgagt ttagatgcaa gttcaaggag cgaaaggtgg atgggtaggt tatataggga 14940
tatagcacag agatatatag caaagagata cttttgagca atgtttgtgg aagcggtatt 15000
cgcaatattt tagtagctcg ttacagtccg gtgcgttttt ggtttttga aagtgcgtct 15060
tcagagcgct tttggttttc aaaagcgctc tgaagttcct atactttcta gagaatagga 15120
acttcggaat aggaacttca aagcgtttcc gaaaacgagc gcttccgaaa atgcaacgcg 15180
agctgcgcac atacagctca ctgttcacgt cgcacctata tctgcgtgtt gcctgtatat 15240
atatatacat gagaagaacg gcatagtgcg tgtttatgct taaatgcgta cttatatgcg 15300
tctatttatg taggatgaaa ggtagtctag tacctcctgt gatattatcc cattccatgc 15360
ggggtatcgt atgcttcctt cagcactacc ctttagctgt tctatatgct gccactcctc 15420
aattggatta gtctcatcct tcaatgctat catttccttt gatattggat catactaaga 15480
aaccattatt atcatgacat taacctataa aaataggcgt atcacgaggc cctttcgtc 15539

<210> 140
<211> 1053
<212> DNA
<213> Bifidobacterium angulatum

<400> 140
atggccgcta ctatctggta cgaaaaggat gctgatttgt ccgtcttcga aggtaagaag 60
gtcgccgtta tcggttacgg ttcccaaggt cacgctcatg ccttaaactt gagagactct 120
ggtgttgatg ttgtcgttgg tttgcgtcca acctctaagt ccgtcgaata cgctaaagaa 180
caaggtttgg aagttcagtc tgtcgctgac gccactgctg aagccgatgt cgttatgatc 240

ttgctgccag accaatacca ggctgccgtt tacaagtccg aaattgaacc aaacttgaaa 300

ccaggtgctg ccttggcttt cgcccacggt ttcaacatcc attacggtta cattaagcca 360

tctgaagatc acccagtttt catggtcgct ccaaagggtc caggtcacat cgtcagacgt 420

gaatacacca acggtagagg tgttccagtc gttgtcgctg tcgaacaaga cccagatggt 480

aaaacttggc cactgtgttt agcttacgcc aaggctttgg gtgctttgag agctggtgct 540

atcaagacca ctttcaccga agaaaccgaa accgacttgt ttggtgaaca agatgtttta 600

atgggtggta tcaaccactt gtgcgacatg ggtttcgatg tcttgactga agccggttac 660

caaccagaaa ttgcttactt cgaagttttt cacgaattga agatgttggt tgacttggcc 720

aacgaaggtg gtttgaacaa ggctagatgg tcttgttccg acaccgctca atacggtgac 780

tacacctcta ctgttatcac cgaagaaacc aagaagagaa tgcaatacca gttgaagaga 840

atccaagatg gttccttcgc taaagaattc atggacgatc aagctgccgg tgctccaaag 900

ttcaagaagt tgcaggaaga atactctcac ccacacttgg aaaccgttgg tccaaagttg 960

agagctatgt tctcctggaa caatcaagtc gacgccgacg ccgacatggc tgaatccttc 1020

aacggtaaga tcgctagaac ccaagtccaa taa 1053

<210> 141
<211> 350
<212> PRT
<213> Bifidobacterium angulatum

<400> 141

Met Ala Ala Thr Ile Trp Tyr Glu Lys Asp Ala Asp Leu Ser Val Phe
1 5 10 15

Glu Gly Lys Lys Val Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Leu Asn Leu Arg Asp Ser Gly Val Asp Val Val Val Gly Leu
35 40 45

Arg Pro Thr Ser Lys Ser Val Glu Tyr Ala Lys Glu Gln Gly Leu Glu
50 55 60

Val Gln Ser Val Ala Asp Ala Thr Ala Glu Ala Asp Val Val Met Ile
65 70 75 80

Leu Leu Pro Asp Gln Tyr Gln Ala Ala Val Tyr Lys Ser Glu Ile Glu
85 90 95

Pro Asn Leu Lys Pro Gly Ala Ala Leu Ala Phe Ala His Gly Phe Asn
100 105 110

Ile His Tyr Gly Tyr Ile Lys Pro Ser Glu Asp His Pro Val Phe Met
115 120 125

Val Ala Pro Lys Gly Pro Gly His Ile Val Arg Arg Glu Tyr Thr Asn
130 135 140

```
Gly Arg Gly Val Pro Val Val Val Ala Val Glu Gln Asp Pro Asp Gly
145                 150                 155                 160

Lys Thr Trp Pro Leu Cys Leu Ala Tyr Ala Lys Ala Leu Gly Ala Leu
                165                 170                 175

Arg Ala Gly Ala Ile Lys Thr Thr Phe Thr Glu Glu Thr Glu Thr Asp
            180                 185                 190

Leu Phe Gly Glu Gln Asp Val Leu Met Gly Gly Ile Asn His Leu Cys
        195                 200                 205

Asp Met Gly Phe Asp Val Leu Thr Glu Ala Gly Tyr Gln Pro Glu Ile
    210                 215                 220

Ala Tyr Phe Glu Val Phe His Glu Leu Lys Met Leu Val Asp Leu Ala
225                 230                 235                 240

Asn Glu Gly Gly Leu Asn Lys Ala Arg Trp Ser Cys Ser Asp Thr Ala
                245                 250                 255

Gln Tyr Gly Asp Tyr Thr Ser Thr Val Ile Thr Glu Glu Thr Lys Lys
            260                 265                 270

Arg Met Gln Tyr Gln Leu Lys Arg Ile Gln Asp Gly Ser Phe Ala Lys
        275                 280                 285

Glu Phe Met Asp Asp Gln Ala Ala Gly Ala Pro Lys Phe Lys Lys Leu
    290                 295                 300

Gln Glu Glu Tyr Ser His Pro His Leu Glu Thr Val Gly Pro Lys Leu
305                 310                 315                 320

Arg Ala Met Phe Ser Trp Asn Asn Gln Val Asp Ala Asp Ala Asp Met
                325                 330                 335

Ala Glu Ser Phe Asn Gly Lys Ile Ala Arg Thr Gln Val Gln
            340                 345                 350
```

<210> 142
<211> 1050
<212> DNA
<213> Bifidobacterium dentium

<400> 142

```
atggccgctc aaatctggta cgaaaacgat ggtgacttgt ccgtcttgga tggtaagaag       60

gtcgctatca ttggttacgg ttctcaaggt cacgcccatg ctttgaactt gagagactct      120

ggtgttgatg ttgtcgttgg tctgcgtcca acctctaagt ccgtcgaaca agctaaagaa      180

cagggtttgg aagtcaagtc cgtcccagaa gctgccgctg aagctgacat cattatgatc      240

ttggccccag atcaatacca gagaactatt tgggctaacg acatcgaacc aaatattaag      300

ccaggtgccg ctatcgcctt cgctcacggt ttcaacatcc attacggtta cattaagcca      360

tctgaagatc acccagtttt catggtcgct ccaaaaggtc caggtcacat tgttagacgt      420
```

gaatacgcta acggtagagg tgttccagtc gttgtcgctg tcgaacaaga cccaagaggt 480

gacgcttggg atatcacctt ggcctacgct aaggctttgg gtgctttgag agctggtgcc 540

attaagacca ccttcaagga agaaaccgaa accgacttgt ttggtgaaca aaacgtttta 600

atgggtggtg ttaacaagtt ggtcgaaatg ggtttcgaag tcttgaccga cgctggttac 660

caaccagaaa tcgcctactt cgaagtttgt cacgaattga agatgttagt cgatttgatg 720

aacgaaggtg gtttgaacaa ggctagatgg tcttgttccg acactgccca atacggtgac 780

tacaccaaca ccgttatcaa cgaagactgc agaaagcgta tgcaatacca cttgggtaga 840

atccaagacg gttctttcgc taaggaattc atcgacgatc aagatgctgg tgccccaaag 900

ttcaaggaat tgcaggaaga atacggtaac gttagaatcg aaactgtcgg tccaaagttg 960

cgtgctatgt tctcctggaa caatggtcaa gacgacgacg ccgacatggc taccttcact 1020

ggtaagatcg ctagatccca agtccaataa 1050

<210> 143
<211> 349
<212> PRT
<213> Bifidobacterium dentium

<400> 143

Met Ala Ala Gln Ile Trp Tyr Glu Asn Asp Gly Asp Leu Ser Val Leu
1 5 10 15

Asp Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Leu Asn Leu Arg Asp Ser Gly Val Asp Val Val Val Gly Leu
35 40 45

Arg Pro Thr Ser Lys Ser Val Glu Gln Ala Lys Glu Gln Gly Leu Glu
50 55 60

Val Lys Ser Val Pro Glu Ala Ala Ala Glu Ala Asp Ile Ile Met Ile
65 70 75 80

Leu Ala Pro Asp Gln Tyr Gln Arg Thr Ile Trp Ala Asn Asp Ile Glu
85 90 95

Pro Asn Ile Lys Pro Gly Ala Ala Ile Ala Phe Ala His Gly Phe Asn
100 105 110

Ile His Tyr Gly Tyr Ile Lys Pro Ser Glu Asp His Pro Val Phe Met
115 120 125

Val Ala Pro Lys Gly Pro Gly His Ile Val Arg Arg Glu Tyr Ala Asn
130 135 140

Gly Arg Gly Val Pro Val Val Val Ala Val Glu Gln Asp Pro Arg Gly
145 150 155 160

Asp Ala Trp Asp Ile Thr Leu Ala Tyr Ala Lys Ala Leu Gly Ala Leu

165 170 175

Arg Ala Gly Ala Ile Lys Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp
180 185 190

Leu Phe Gly Glu Gln Asn Val Leu Met Gly Gly Val Asn Lys Leu Val
195 200 205

Glu Met Gly Phe Glu Val Leu Thr Asp Ala Gly Tyr Gln Pro Glu Ile
210 215 220

Ala Tyr Phe Glu Val Cys His Glu Leu Lys Met Leu Val Asp Leu Met
225 230 235 240

Asn Glu Gly Gly Leu Asn Lys Ala Arg Trp Ser Cys Ser Asp Thr Ala
245 250 255

Gln Tyr Gly Asp Tyr Thr Asn Thr Val Ile Asn Glu Asp Cys Arg Lys
260 265 270

Arg Met Gln Tyr His Leu Gly Arg Ile Gln Asp Gly Ser Phe Ala Lys
275 280 285

Glu Phe Ile Asp Asp Gln Asp Ala Gly Ala Pro Lys Phe Lys Glu Leu
290 295 300

Gln Glu Glu Tyr Gly Asn Val Arg Ile Glu Thr Val Gly Pro Lys Leu
305 310 315 320

Arg Ala Met Phe Ser Trp Asn Asn Gly Gln Asp Asp Asp Ala Asp Met
325 330 335

Ala Thr Phe Thr Gly Lys Ile Ala Arg Ser Gln Val Gln
340 345

<210> 144
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56_58_060210f

<220>
<221> misc_feature
<222> (7)..(8)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (13)..(14)
<223> n is a, c, g, or t

<400> 144
gaaggannka aannktggaa gagagc 26

<210> 145
<211> 26

<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56_58_060210r

<220>
<221> misc_feature
<222> (13)..(14)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 145
gctctcttcc amnnttmnn tccttc 26

<210> 146
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> K9_30_101110f primer

<220>
<221> misc_feature
<222> (20)..(21)
<223> n is a, c, g, or t

<400> 146
gactatcgcc gttatcggtn nkggttctca aggtcac 37

<210> 147
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_30_101110r

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<400> 147
gtgaccttga gaaccmnnac cgataacggc gatagtc 37

<210> 148
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_67_101110f

<220>
<221> misc_feature
<222> (18)..(19)
<223> n is a, c, g, or t

<400> 148

gagctgaaga acaaggtnnk gaagtctaca ccgctgc 37

<210> 149
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_67_101110r

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 149
gcagcggtgt agacttcmnn accttgttct tcagctc 37

<210> 150
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131_101110f

<220>
<221> misc_feature
<222> (18)..(19)
<223> n is a, c, g, or t

<400> 150
caaaggacgt tgatgtcnnk atgatcgctc caaag 35

<210> 151
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131_101110r

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<400> 151
ctttggagcg atcatmnnga catcaacgtc cttt 34

<210> 152
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_162_101110f

<220>
<221> misc_feature
<222> (16)..(17)
<223> n is a, c, g, or t

<400> 152

gctgtcgaac aagacnnkac tggcaaggct ttg 33

<210> 153
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_162_101110r

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<400> 153
caaagccttg ccagtmnngt cttgttcgac agc 33

<210> 154
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_176_101110f

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 154
gctttggcct acgctttann katcggtggt gctagagc 38

<210> 155
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_176_101110r

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 155
gctctagcac caccgatmnn taaagcgtag gccaaagc 38

<210> 156
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_191_101110f

<220>
<221> misc_feature
<222> (16)..(17)
<223> n is a, c, g, or t

<400> 156

gaaactacct tcagannkga aactgaaacc gac 33

<210> 157
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> K9_191_101110r

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<400> 157
gtcggtttca gtttcmnntc tgaaggtagt ttc 33

<210> 158
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_227_101110f

<220>
<221> misc_feature
<222> (16)..(17)
<223> n is a, c, g, or t

<400> 158
gccggttacg acccannkaa cgcttacttc gaatg 35

<210> 159
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_227_101110r

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 159
cattcgaagt aagcgttmnn tgggtcgtaa ccggc 35

<210> 160
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_246_101110f

<220>
<221> misc_feature
<222> (16)..(17)
<223> n is a, c, g, or t

<400> 160

gttgacttga tctacnnktc tggtttctcc ggtatgc 37

<210> 161
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_246_101110r

<220>
<221> misc_feature
<222> (21)..(22)
<223> n is a, c, g, or t

<400> 161
gcataccgga gaaaccagam nngtagatca agtcaac 37

<210> 162
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_37&39_101110f

<220>
<221> misc_feature
<222> (18)..(19)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (24)..(25)
<223> n is a, c, g, or t

<400> 162
gttctcaagg tcacgctnnk gccnnkaatg ctaaggaatc c 41

<210> 163
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_37&39_101110r

<220>
<221> misc_feature
<222> (17)..(18)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (23)..(24)
<223> n is a, c, g, or t

<400> 163
ggattcctta gcattmnngg cmnnagcgtg accttgagaa c 41

<210> 164
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_86&87_101110f

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (22)..(23)
<223> n is a, c, g, or t

<400> 164
gacatcatta tgatcttgnn knnkgatgaa aagcaggc 38

<210> 165
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_86&87_101110r

<220>
<221> misc_feature
<222> (16)..(17)
<223> n is a, c, g, or t

<220>
<221> misc_feature
<222> (19)..(20)
<223> n is a, c, g, or t

<400> 165
gcctgctttt catcmnnmnn caagatcata atgatgtc 38

<210> 166
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131L_112210f

<400> 166
ggacgttgat gtcttgatga tcgctcc 27

<210> 167
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131A_112210f

<400> 167
ggacgttgat gtcgcaatga tcgctcc 27

<210> 168
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131V_112210f

<400> 168
ggacgttgat gtcgttatga tcgctcc 27

<210> 169
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131M_112210f

<400> 169
ggacgttgat gtcatgatga tcgctcc 27

<210> 170
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_131C_112210f

<400> 170
ggacgttgat gtctgaatga tcgctcc 27

<210> 171
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_191D_112210r

<400> 171
ggtttcagtt tcgtctctga aggtagtttc 30

<210> 172
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_191C_112210r

<400> 172
ggtttcagtt tcgcatctga aggtagtttc 30

<210> 173
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_191S_112210r

<400> 173
ggtttcagtt tccgatctga aggtagtttc 30

<210> 174
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_191G_112210r

20 Jan 2017
2017200439

<400> 174
ggtttcagtt tcgcctctga aggtagtttc 30

<210> 175
<211> 11864
<212> DNA
<213> Artificial sequence

<220>
<223> pHR81-PCUP1-AlsS-PILV5-Pf5.KARI (pLH550) plasmid

<400> 175
catcatcaca gtttaaacag tatgaaagtt ttctacgata aagactgcga cctgtcgatc 60
atccaaggta agaaagttgc catcatcggc tacggttccc agggccacgc tcaagcatgc 120
aacctgaagg attccggcgt agacgtgact gttggcctgc gtaaaggctc ggctaccgtt 180
gccaaggctg aagcccacgg cttgaaagtg accgacgttg ctgcagccgt tgccggtgcc 240
gacttggtca tgatcctgac cccggacgag ttccagtccc agctgtacaa gaacgaaatc 300
gagccgaaca tcaagaaggg cgccactctg gccttctccc acggcttcgc gatccactac 360
aaccaggttg tgcctcgtgc cgacctcgac gtgatcatga tcgcgccgaa ggctccaggc 420
cacaccgtac gttccgagtt cgtcaagggc ggtggtattc ctgacctgat cgcgatctac 480
caggacgctt ccggcaacgc caagaacgtt gccctgtcct acgccgcagg cgtgggcggc 540
ggccgtaccg gcatcatcga aaccaccttc aaggacgaga ctgaaaccga cctgttcggt 600
gagcaggctg ttctgtgtgg cggtaccgtc gagctggtca aagccggttt cgaaaccctg 660
gttgaagctg gctacgctcc agaaatggcc tacttcgagt gcctgcacga actgaagctg 720
atcgttgacc tcatgtacga aggcggtatc gccaacatga actactcgat ctccaacaac 780
gctgaatacg gcgagtacgt gactggtcca gaagtcatca acgccgaatc ccgtcaggcc 840
atgcgcaatg ctctgaagcg catccaggac ggcgaatacg cgaagatgtt catcagcgaa 900
ggcgctaccg gctacccatc gatgaccgcc aagcgtcgta acaacgctgc tcacggtatc 960
gaaatcatcg gcgagcaact gcgctcgatg atgccttgga tcggtgccaa caaaatcgtc 1020
gacaaagcca agaactaagg ccctgcaggc cagaggaaaa taatatcaag tgctggaaac 1080
tttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac 1140
atttaagatt ttttttttt catccgacat acatctgtac actaggaagc cctgtttttc 1200
tgaagcagct tcaaatatat atattttta catatttatt atgattcaat gaacaatcta 1260
attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg 1320
tataagtcct catcgggaca gctacgattt ctctttcggt tttggctgag ctactggttg 1380
ctgtgacgca gcggcattag cgcggcgtta tgagctaccc tcgtggcctg aaagatggcg 1440
ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg 1500
tcaagtcacg tttggtggac ggccccttc caacgaatcg tatatactaa catgcgcgcg 1560
cttcctatat acacatatac atatatatat atatatatat gtgtgcgtgt atgtgtacac 1620
ctgtatttaa tttccttact cgcgggtttt tctttttct caattcttgg cttcctcttt 1680
ctcgagcgga ccggatcctc cgcggtgccg gcagatctat ttaaatggcg cgccgacgtc 1740

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 1800
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 1860
aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 1920
ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 1980
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 2040
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 2100
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 2160
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 2220
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 2280
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 2340
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 2400
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 2460
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 2520
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 2580
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 2640
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga 2700
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 2760
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga 2820
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 2880
agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 2940
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 3000
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgta 3060
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 3240
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 3300
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 3360
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 3420
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 3480
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 3540
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 3600
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 3720
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 3780
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 3900

cgccaagctt tttctttcca attttttttt tttcgtcatt ataaaaatca ttacgaccga 3960
gattcccggg taataactga tataattaaa ttgaagctct aatttgtgag tttagtatac 4020
atgcatttac ttataataca gtttttttagt tttgctggcc gcatcttctc aaatatgctt 4080
cccagcctgc ttttctgtaa cgttcaccct ctaccttagc atcccttccc tttgcaaata 4140
gtcctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat 4200
actgttgacc caatgcgtct cccttgtcat ctaaacccac accgggtgtc ataatcaacc 4260
aatcgtaacc ttcatctctt ccacccatgt ctctttgagc aataaagccg ataacaaaat 4320
ctttgtcgct cttcgcaatg tcaacagtac ccttagtata ttctccagta gatagggagc 4380
ccttgcatga caattctgct aacatcaaaa ggcctctagg ttcctttgtt acttcttctg 4440
ccgcctgctt caaaccgcta acaatacctg ggcccaccac accgtgtgca ttcgtaatgt 4500
ctgcccattc tgctattctg tatacacccg cagagtactg caatttgact gtattaccaa 4560
tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta 4620
gcggcttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta 4680
aacaaatttt gggacctaat gcttcaacta actccagtaa ttccttggtg gtacgaacat 4740
ccaatgaagc acacaagttt gtttgctttt cgtgcatgat attaaatagc ttggcagcaa 4800
caggactagg atgagtagca gcacgttcct tatatgtagc tttcgacatg atttatcttc 4860
gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa tactgggcaa tttcatgttt 4920
cttcaacact acatatgcgt atatatacca atctaagtct gtgctccttc cttcgttctt 4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaaag 5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg caggtcgact ctagtatact 5100
ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact 5160
cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatcttcg 5220
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 5280
atggctgcca tcattattat ccgatgtgac gctgcatttt ttttttttt ttttttttt 5340
ttttttttt ttttttttt ttttttttgt acaaatatca taaaaaaaga gaatcttttt 5400
aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 5460
aaccacctaa atcaccagtt ctgatacctg catccaaaac ctttttaact gcatcttcaa 5520
tggctttacc ttcttcaggc aagttcaatg acaatttcaa catcattgca gcagacaaga 5580
tagtggcgat agggttgacc ttattctttg gcaaatctgg agcggaacca tggcatggtt 5640
cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 5700
aacccaagga gcctgggata acggaggctt catcggagat gatatcacca aacatgttgc 5760
tggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat 5820
caatcaattg atgttgaact ttcaatgtag ggaattcgtt cttgatggtt tcctccacag 5880
tttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca 5940
atggtggctc atgttgtagg gccatgaaag cggccattct tgtgattctt tgcacttctg 6000
gaacggtgta ttgttcacta tcccaagcga caccatcacc atcgtcttcc tttctcttac 6060

caaagtaaat acctcccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt 6120
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 6180
tggcgtacaa ttgaagttct ttacggattt ttagtaaacc ttgttcaggt ctaacactac 6240
cggtacccca tttaggacca cccacagcac ctaacaaaac ggcatcagcc ttcttggagg 6300
cttccagcgc ctcatctgga agtggaacac ctgtagcatc gatagcagca ccaccaatta 6360
aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 6420
taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 6480
taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 6600
tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat tgaattttga 6660
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 6720
tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 6780
atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct 6840
tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa 6900
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa 7020
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg cattttttaca 7080
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttg 7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt 7200
tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact gtaggtccgt 7260
taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc 7320
cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttttca agataaaggc 7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat 7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt 7560
tcttactaca attttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag 7620
gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat atagggatat 7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtattcgc 7740
aatattttag tagctcgtta cagtccggtg cgtttttggt tttttgaaag tgcgtcttca 7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact 7860
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc 7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata 7980
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 8040
atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat tccatgcggg 8100
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat 8160
tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat atgcatagta 8220

ccgagaaact agaggatctc ccattaccga catttgggcg ctatacgtgc atatgttcat 8280
gtatgtatct gtatttaaaa cacttttgta ttatttttcc tcatatatgt gtataggttt 8340
atacggatga tttaattatt acttcaccac cctttatttc aggctgatat cttagccttg 8400
ttactagtta gaaaaagaca tttttgctgt cagtcactgt caagagattc ttttgctggc 8460
atttcttcta gaagcaaaaa gagcgatgcg tcttttccgc tgaaccgttc cagcaaaaaa 8520
gactaccaac gcaatatgga ttgtcagaat catataaaag agaagcaaat aactccttgt 8580
cttgtatcaa ttgcattata atatcttctt gttagtgcaa tatcatatag aagtcatcga 8640
aatagatatt aagaaaaaca aactgtacaa tcaatcaatc aatcatcgct gaggatgttg 8700
acaaaagcaa caaaagaaca aaaatccctt gtgaaaaaca gaggggcgga gcttgttgtt 8760
gattgcttag tggagcaagg tgtcacacat gtatttggca ttccaggtgc aaaaattgat 8820
gcggtatttg acgctttaca agataaagga cctgaaatta tcgttgcccg gcacgaacaa 8880
aacgcagcat tcatggccca agcagtcggc cgtttaactg gaaaaccggg agtcgtgtta 8940
gtcacatcag gaccgggtgc ctctaacttg gcaacaggcc tgctgacagc gaacactgaa 9000
ggagaccctg tcgttgcgct tgctggaaac gtgatccgtg cagatcgttt aaaacggaca 9060
catcaatctt tggataatgc ggcgctattc cagccgatta caaaatacag tgtagaagtt 9120
caagatgtaa aaaatatacc ggaagctgtt acaaatgcat ttaggatagc gtcagcaggg 9180
caggctgggg ccgcttttgt gagctttccg caagatgttg tgaatgaagt cacaaatacg 9240
aaaaacgtgc gtgctgttgc agcgccaaaa ctcggtcctg cagcagatga tgcaatcagt 9300
gcggccatag caaaaatcca aacagcaaaa cttcctgtcg ttttggtcgg catgaaaggc 9360
ggaagaccgg aagcaattaa agcggttcgc aagcttttga aaaaggttca gcttccattt 9420
gttgaaacat atcaagctgc cggtaccctt tctagagatt tagaggatca atattttggc 9480
cgtatcggtt tgttccgcaa ccagcctggc gatttactgc tagagcaggc agatgttgtt 9540
ctgacgatcg gctatgaccc gattgaatat gatccgaaat tctggaatat caatggagac 9600
cggacaatta tccatttaga cgagattatc gctgacattg atcatgctta ccagcctgat 9660
cttgaattga tcggtgacat tccgtccacg atcaatcata tcgaacacga tgctgtgaaa 9720
gtggaatttg cagagcgtga gcagaaaatc ctttctgatt taaaacaata tatgcatgaa 9780
ggtgagcagg tgcctgcaga ttggaaatca gacagagcgc accctcttga aatcgttaaa 9840
gagttgcgta atgcagtcga tgatcatgtt acagtaactt gcgatatcgg ttcgcacgcc 9900
atttggatgt cacgttattt ccgcagctac gagccgttaa cattaatgat cagtaacggt 9960
atgcaaacac tcggcgttgc gcttccttgg gcaatcggcg cttcattggt gaaaccggga 10020
gaaaaagtgg tttctgtctc tggtgacggc ggtttcttat tctcagcaat ggaattagag 10080
acagcagttc gactaaaagc accaattgta cacattgtat ggaacgacag cacatatgac 10140
atggttgcat tccagcaatt gaaaaaatat aaccgtacat ctgcggtcga tttcggaaat 10200
atcgatatcg tgaaatatgc ggaaagcttc ggagcaactg gcttgcgcgt agaatcacca 10260
gaccagctgg cagatgttct gcgtcaaggc atgaacgctg aaggtcctgt catcatcgat 10320
gtcccggttg actacagtga taacattaat ttagcaagtg acaagcttcc gaaagaattc 10380

ggggaactca tgaaaacgaa agctctctag ttaattaatc atgtaattag ttatgtcacg 10440
cttacattca cgccctcccc ccacatccgc tctaaccgaa aaggaaggag ttagacaacc 10500
tgaagtctag gtccctattt attttttat agttatgtta gtattaagaa cgttatttat 10560
atttcaaatt tttctttttt ttctgtacag acgcgtgtac gcatgtaaca ttatactgaa 10620
aaccttgctt gagaaggttt tgggacgctc gaaggcttta atttgcgggc ggccgcacct 10680
ggtaaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa aagaccagag 10740
tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac agacatcttt 10800
ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat ttggctaaga 10860
acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa tggagttaat 10920
cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta cccgacggga 10980
aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat aatactagag 11040
ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata tacaaaatat 11100
cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta gtaccattcc 11160
tcagagaaga ggtatacata acaagaaaat cgcgtgaaca ccttatataa cttagcccgt 11220
tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag acgtgataaa 11280
aatttacttt ctaactcttc tcacgctgcc cctatctgtt cttccgctct accgtgagaa 11340
ataaagcatc gagtacggca gttcgctgtc actgaactaa aacaataagg ctagttcgaa 11400
tgatgaactt gcttgctgtc aaacttctga gttgccgctg atgtgacact gtgacaataa 11460
attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca atagagctca 11520
gtaggagtca gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg taagttgtgc 11580
gcgtgcacat ttcgcccgtt cccgctcatc ttgcagcagg cggaaatttt catcacgctg 11640
taggacgcaa aaaaaaaata attaatcgta caagaatctt ggaaaaaaaa ttgaaaaatt 11700
ttgtataaaa gggatgacct aacttgactc aatggctttt acacccagta ttttcccttt 11760
ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta ttcctaggag 11820
ttatattttt ttaccctacc agcaatataa gtaaaaaact gttt 11864

<210> 176
<211> 11871
<212> DNA
<213> Artificial sequence

<220>
<223> pYZ058 (pHR81-PCUP1-AlsS-PILV5-yeast KARI plasmid

<400> 176
tcccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60
aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120
ttacttcacc acccttttatt tcaggctgat atcttagcct tgttactagt tagaaaaaga 180
cattttgct gtcagtcact gtcaagagat tcttttgctg gcatttcttc tagaagcaaa 240
aagagcgatg cgtcttttcc gctgaaccgt tccagcaaaa aagactacca acgcaatatg 300
gattgtcaga atcatataaa agagaagcaa ataactcctt gtcttgtatc aattgcatta 360

taatatcttc ttgttagtgc aatatcatat agaagtcatc gaaatagata ttaagaaaaa 420
caaactgtac aatcaatcaa tcaatcatcg ctgaggatgt tgacaaaagc aacaaaagaa 480
caaaaatccc ttgtgaaaaa cagaggggcg gagcttgttg ttgattgctt agtggagcaa 540
ggtgtcacac atgtatttgg cattccaggt gcaaaaattg atgcggtatt tgacgcttta 600
caagataaag gacctgaaat tatcgttgcc cggcacgaac aaaacgcagc attcatggcc 660
caagcagtcg gccgtttaac tggaaaaccg ggagtcgtgt tagtcacatc aggaccgggt 720
gcctctaact tggcaacagg cctgctgaca gcgaacactg aaggagaccc tgtcgttgcg 780
cttgctggaa acgtgatccg tgcagatcgt ttaaaacgga cacatcaatc tttggataat 840
gcggcgctat tccagccgat tacaaaatac agtgtagaag ttcaagatgt aaaaaatata 900
ccggaagctg ttacaaatgc atttaggata gcgtcagcag ggcaggctgg ggccgctttt 960
gtgagctttc cgcaagatgt tgtgaatgaa gtcacaaata cgaaaaacgt gcgtgctgtt 1020
gcagcgccaa aactcggtcc tgcagcagat gatgcaatca gtgcggccat agcaaaaatc 1080
caaacagcaa aacttcctgt cgttttggtc ggcatgaaag gcggaagacc ggaagcaatt 1140
aaagcggttc gcaagctttt gaaaaaggtt cagcttccat ttgttgaaac atatcaagct 1200
gccggtaccc tttctagaga tttagaggat caatattttg gccgtatcgg tttgttccgc 1260
aaccagcctg gcgatttact gctagagcag gcagatgttg ttctgacgat cggctatgac 1320
ccgattgaat atgatccgaa attctggaat atcaatggag accggacaat tatccattta 1380
gacgagatta tcgctgacat tgatcatgct taccagcctg atcttgaatt gatcggtgac 1440
attccgtcca cgatcaatca tatcgaacac gatgctgtga aagtggaatt tgcagagcgt 1500
gagcagaaaa tcctttctga tttaaaacaa tatatgcatg aaggtgagca ggtgcctgca 1560
gattggaaat cagacagagc gcaccctctt gaaatcgtta aagagttgcg taatgcagtc 1620
gatgatcatg ttacagtaac ttgcgatatc ggttcgcacg ccatttggat gtcacgttat 1680
ttccgcagct acgagccgtt aacattaatg atcagtaacg gtatgcaaac actcggcgtt 1740
gcgcttcctt gggcaatcgg cgcttcattg gtgaaaccgg gagaaaaagt ggtttctgtc 1800
tctggtgacg gcggtttctt attctcagca atggaattag agacagcagt tcgactaaaa 1860
gcaccaattg tacacattgt atggaacgac agcacatatg acatggttgc attccagcaa 1920
ttgaaaaaat ataaccgtac atctgcggtc gatttcggaa atatcgatat cgtgaaatat 1980
gcggaaagct tcggagcaac tggcttgcgc gtagaatcac cagaccagct ggcagatgtt 2040
ctgcgtcaag gcatgaacgc tgaaggtcct gtcatcatcg atgtcccggt tgactacagt 2100
gataacatta atttagcaag tgacaagctt ccgaaagaat tcggggaact catgaaaacg 2160
aaagctctct agttaattaa tcatgtaatt agttatgtca cgcttacatt cacgccctcc 2220
ccccacatcc gctctaaccg aaaaggaagg agttagacaa cctgaagtct aggtccctat 2280
ttatttttt atagttatgt tagtattaag aacgttattt atatttcaaa tttttctttt 2340
ttttctgtac agacgcgtgt acgcatgtaa cattatactg aaaaccttgc ttgagaaggt 2400
tttgggacgc tcgaaggctt taatttgcgg gcggccgcac ctggtaaaac ctctagtgga 2460
gtagtagatg taatcaatga agcggaagcc aaaagaccag agtagaggcc tatagaagaa 2520

actgcgatac cttttgtgat ggctaaacaa acagacatct ttttatatgt ttttacttct 2580
gtatatcgtg aagtagtaag tgataagcga atttggctaa gaacgttgta agtgaacaag 2640
ggacctcttt tgcctttcaa aaaaggatta aatggagtta atcattgaga tttagttttc 2700
gttagattct gtatccctaa ataactccct tacccgacgg gaaggcacaa aagacttgaa 2760
taatagcaaa cggccagtag ccaagaccaa ataatactag agttaactga tggtcttaaa 2820
caggcattac gtggtgaact ccaagaccaa tatacaaaat atcgataagt tattcttgcc 2880
caccaattta aggagcctac atcaggacag tagtaccatt cctcagagaa gaggtataca 2940
taacaagaaa atcgcgtgaa caccttatat aacttagccc gttattgagc taaaaaacct 3000
tgcaaaattt cctatgaata agaatacttc agacgtgata aaaatttact ttctaactct 3060
tctcacgctg cccctatctg ttcttccgct ctaccgtgag aaataaagca tcgagtacgg 3120
cagttcgctg tcactgaact aaaacaataa ggctagttcg aatgatgaac ttgcttgctg 3180
tcaaacttct gagttgccgc tgatgtgaca ctgtgacaat aaattcaaac cggttatagc 3240
ggtctcctcc ggtaccggtt ctgccacctc caatagagct cagtaggagt cagaacctct 3300
gcggtggctg tcagtgactc atccgcgttt cgtaagttgt gcgcgtgcac atttcgcccg 3360
ttcccgctca tcttgcagca ggcggaaatt ttcatcacgc tgtaggacgc aaaaaaaaaa 3420
taattaatcg tacaagaatc ttggaaaaaa aattgaaaaa ttttgtataa aagggatgac 3480
ctaacttgac tcaatggctt ttacacccag tattttccct ttccttgttt gttacaatta 3540
tagaagcaag acaaaaacat atagacaacc tattcctagg agttatattt ttttacccta 3600
ccagcaatat aagtaaaaaa ctgtttaaac agtatgttga agcaaatcaa cttcggtggt 3660
actgttgaaa ccgtctacga aagagctgac tggccaagag aaaagttgtt ggactacttc 3720
aagaacgaca cttttgcttt gatcggttac ggttcccaag gttacggtca aggtttgaac 3780
ttgagagaca acggtttgaa cgttatcatt ggtgtccgta aagatggtgc ttcttggaag 3840
gctgccatcg aagacggttg ggttccaggc aagaacttgt tcactgttga agatgctatc 3900
aagagaggta gttacgttat gaacttgttg tccgatgccg ctcaatcaga aacctggcct 3960
gctatcaagc cattgttgac caagggtaag actttgtact tctcccacgg tttctcccca 4020
gtcttcaagg acttgactca cgttgaacca ccaaaggact tagatgttat cttggttgct 4080
ccaaagggtt ccggtagaac tgtcagatct ttgttcaagg aaggtcgtgg tattaactct 4140
tcttacgccg tctggaacga tgtcaccggt aaggctcacg aaaaggccca agctttggcc 4200
gttgccattg gttccggtta cgtttaccaa accactttcg aaagagaagt caactctgac 4260
ttgtacggtg aaagaggttg tttaatgggt ggtatccacg gtatgttctt ggctcaatac 4320
gacgtcttga gagaaaacgg tcactcccca tctgaagctt tcaacgaaac cgtcgaagaa 4380
gctacccaat ctctataccc attgatcggt aagtacggta tggattacat gtacgatgct 4440
tgttccacca ccgccagaag aggtgctttg gactggtacc caatcttcaa gaatgctttg 4500
aagcctgttt tccaagactt gtacgaatct accaagaacg gtaccgaaac caagagatct 4560
ttggaattca actctcaacc tgactacaga gaaaagctag aaaaggaatt agacaccatc 4620
agaaacatgg aaatctggaa ggttggtaag gaagtcagaa agttgagacc agaaaaccaa 4680

taaggccctg caggcctatc aagtgctgga aacttttct cttggaattt ttgcaacatc 4740
aagtcatagt caattgaatt gacccaattt cacatttaag attttttttt tttcatccga 4800
catacatctg tacactagga agccctgttt ttctgaagca gcttcaaata tatatatttt 4860
ttacatattt attatgattc aatgaacaat ctaattaaat cgaaaacaag aaccgaaacg 4920
cgaataaata atttatttag atggtgacaa gtgtataagt cctcatcggg acagctacga 4980
tttctctttc ggttttggct gagctactgg ttgctgtgac gcagcggcat tagcgcggcg 5040
ttatgagcta ccctcgtggc ctgaaagatg gcgggaataa agcggaacta aaaattactg 5100
actgagccat attgaggtca atttgtcaac tcgtcaagtc acgtttggtg gacggcccct 5160
ttccaacgaa tcgtatatac taacatgcgc gcgcttccta tatacacata tacatatata 5220
tatatatata tatgtgtgcg tgtatgtgta cacctgtatt taatttcctt actcgcgggt 5280
ttttcttttt tctcaattct tggcttcctc tttctcgagc ggaccggatc ctccgcggtg 5340
ccggcagatc tatttaaatg gcgcgccgac gtcaggtggc acttttcggg gaaatgtgcg 5400
cggaacccct atttgtttat ttttctaaat acattcaaat atgtatccgc tcatgagaca 5460
ataaccctga taaatgcttc aataatattg aaaaaggaag agtatgagta ttcaacattt 5520
ccgtgtcgcc cttattccct tttttgcggc attttgcctt cctgtttttg ctcacccaga 5580
aacgctggtg aaagtaaaag atgctgaaga tcagttgggt gcacgagtgg gttacatcga 5640
actggatctc aacagcggta agatccttga gagttttcgc cccgaagaac gttttccaat 5700
gatgagcact tttaaagttc tgctatgtgg cgcggtatta tcccgtattg acgccgggca 5760
agagcaactc ggtcgccgca tacactattc tcagaatgac ttggttgagt actcaccagt 5820
cacagaaaag catcttacgg atggcatgac agtaagagaa ttatgcagtg ctgccataac 5880
catgagtgat aacactgcgg ccaacttact tctgacaacg atcggaggac cgaaggagct 5940
aaccgctttt ttgcacaaca tgggggatca tgtaactcgc cttgatcgtt gggaaccgga 6000
gctgaatgaa gccataccaa acgacgagcg tgacaccacg atgcctgtag caatggcaac 6060
aacgttgcgc aaactattaa ctggcgaact acttactcta gcttcccggc aacaattaat 6120
agactggatg gaggcggata aagttgcagg accacttctg cgctcggccc ttccggctgg 6180
ctggtttatt gctgataaat ctggagccgg tgagcgtggg tctcgcggta tcattgcagc 6240
actggggcca gatggtaagc cctcccgtat cgtagttatc tacacgacgg ggagtcaggc 6300
aactatggat gaacgaaata gacagatcgc tgagataggt gcctcactga ttaagcattg 6360
gtaactgtca gaccaagttt actcatatat actttagatt gatttaaaac ttcatttta 6420
atttaaaagg atctaggtga agatcctttt tgataatctc atgaccaaaa tcccttaacg 6480
tgagttttcg ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga 6540
tccttttttt ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggt 6600
ggtttgtttg ccggatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag 6660
agcgcagata ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa 6720
ctctgtagca ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag 6780
tggcgataag tcgtgtctta ccgggttgga ctcaagacga tagttaccgg ataaggcgca 6840

gcggtcgggc tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac 6900
cgaactgaga tacctacagc gtgagctatg agaaagcgcc acgcttcccg aagggagaaa 6960
ggcggacagg tatccggtaa gcggcagggt cggaacagga gagcgcacga gggagcttcc 7020
agggggaaac gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg 7080
tcgatttttg tgatgctcgt caggggggcg gagcctatgg aaaaacgcca gcaacgcggc 7140
ctttttacgg ttcctggcct tttgctggcc ttttgctcac atgttctttc ctgcgttatc 7200
ccctgattct gtggataacc gtattaccgc ctttgagtga gctgataccg ctcgccgcag 7260
ccgaacgacc gagcgcagcg agtcagtgag cgaggaagcg gaagagcgcc caatacgcaa 7320
accgcctctc cccgcgcgtt ggccgattca ttaatgcagc tggcacgaca ggtttcccga 7380
ctggaaagcg ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc 7440
ccaggcttta cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca 7500
atttcacaca ggaaacagct atgaccatga ttacgccaag ctttttcttt ccaatttttt 7560
ttttttcgtc attataaaaa tcattacgac cgagattccc gggtaataac tgatataatt 7620
aaattgaagc tctaatttgt gagtttagta tacatgcatt tacttataat acagtttttt 7680
agttttgctg gccgcatctt ctcaaatatg cttcccagcc tgcttttctg taacgttcac 7740
cctctacctt agcatccctt ccctttgcaa atagtcctct tccaacaata ataatgtcag 7800
atcctgtaga gaccacatca tccacggttc tatactgttg acccaatgcg tctcccttgt 7860
catctaaacc cacaccgggt gtcataatca accaatcgta accttcatct cttccaccca 7920
tgtctctttg agcaataaag ccgataacaa aatctttgtc gctcttcgca atgtcaacag 7980
tacccttagt atattctcca gtagataggg agcccttgca tgacaattct gctaacatca 8040
aaaggcctct aggttccttt gttacttctt ctgccgcctg cttcaaaccg ctaacaatac 8100
ctgggcccac cacaccgtgt gcattcgtaa tgtctgccca ttctgctatt ctgtatacac 8160
ccgcagagta ctgcaatttg actgtattac caatgtcagc aaattttctg tcttcgaaga 8220
gtaaaaaatt gtacttggcg gataatgcct ttagcggctt aactgtgccc tccatggaaa 8280
aatcagtcaa gatatccaca tgtgttttta gtaaacaaat tttgggacct aatgcttcaa 8340
ctaactccag taattccttg gtggtacgaa catccaatga agcacacaag tttgtttgct 8400
tttcgtgcat gatattaaat agcttggcag caacaggact aggatgagta gcagcacgtt 8460
ccttatatgt agctttcgac atgatttatc ttcgtttcct gcaggttttt gttctgtgca 8520
gttgggttaa gaatactggg caatttcatg tttcttcaac actacatatg cgtatatata 8580
ccaatctaag tctgtgctcc ttccttcgtt cttccttctg ttcggagatt accgaatcaa 8640
aaaaatttca aggaaaccga aatcaaaaaa aagaataaaa aaaaaatgat gaattgaaaa 8700
gcttgcatgc ctgcaggtcg actctagtat actccgtcta ctgtacgata cacttccgct 8760
caggtccttg tcctttaacg aggccttacc actcttttgt tactctattg atccagctca 8820
gcaaaggcag tgtgatctaa gattctatct tcgcgatgta gtaaaactag ctagaccgag 8880
aaagagacta gaaatgcaaa aggcacttct acaatggctg ccatcattat tatccgatgt 8940
gacgctgcat tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 9000

tgtacaaata tcataaaaaa agagaatctt tttaagcaag gattttctta acttcttcgg 9060
cgacagcatc accgacttcg gtggtactgt tggaaccacc taaatcacca gttctgatac 9120
ctgcatccaa aacctttta actgcatctt caatggcttt accttcttca ggcaagttca 9180
atgacaattt caacatcatt gcagcagaca agatagtggc gatagggttg accttattct 9240
ttggcaaatc tggagcggaa ccatggcatg gttcgtacaa accaaatgcg gtgttcttgt 9300
ctggcaaaga ggccaaggac gcagatggca acaaacccaa ggagcctggg ataacggagg 9360
cttcatcgga gatgatatca ccaaacatgt tgctggtgat tataatacca tttaggtggg 9420
ttgggttctt aactaggatc atggcggcag aatcaatcaa ttgatgttga actttcaatg 9480
tagggaattc gttcttgatg gtttcctcca cagtttttct ccataatctt gaagaggcca 9540
aaacattagc tttatccaag gaccaaatag gcaatggtgg ctcatgttgt agggccatga 9600
aagcggccat tcttgtgatt ctttgcactt ctggaacggt gtattgttca ctatcccaag 9660
cgacaccatc accatcgtct tcctttctct taccaaagta aatacctccc actaattctc 9720
taacaacaac gaagtcagta cctttagcaa attgtggctt gattggagat aagtctaaaa 9780
gagagtcgga tgcaaagtta catggtctta agttggcgta caattgaagt tctttacgga 9840
tttttagtaa accttgttca ggtctaacac taccggtacc ccatttagga ccacccacag 9900
cacctaacaa aacggcatca gccttcttgg aggcttccag cgcctcatct ggaagtggaa 9960
cacctgtagc atcgatagca gcaccaccaa ttaaatgatt ttcgaaatcg aacttgacat 10020
tggaacgaac atcagaaata gctttaagaa ccttaatggc ttcggctgtg atttcttgac 10080
caacgtggtc acctggcaaa acgacgatct tcttaggggc agacattaca atggtatatc 10140
cttgaaatat atataaaaaa aaaaaaaaaa aaaaaaaaaa aaaatgcagc ttctcaatga 10200
tattcgaata cgctttgagg agatacagcc taatatccga caaactgttt tacagattta 10260
cgatcgtact tgttacccat cattgaattt tgaacatccg aacctgggag ttttccctga 10320
aacagatagt atatttgaac ctgtataata atatatagtc tagcgcttta cggaagacaa 10380
tgtatgtatt tcggttcctg gagaaactat tgcatctatt gcataggtaa tcttgcacgt 10440
cgcatccccg gttcattttc tgcgtttcca tcttgcactt caatagcata tctttgttaa 10500
cgaagcatct gtgcttcatt ttgtagaaca aaaatgcaac gcgagagcgc taatttttca 10560
aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgaaag cgctatttta 10620
ccaacgaaga atctgtgctt cattttgta aaacaaaaat gcaacgcgag agcgctaatt 10680
tttcaaacaa agaatctgag ctgcattttt acagaacaga aatgcaacgc gagagcgcta 10740
ttttaccaac aaagaatcta tacttctttt ttgttctaca aaaatgcatc ccgagagcgc 10800
tatttttcta acaaagcatc ttagattact ttttttctcc tttgtgcgct ctataatgca 10860
gtctcttgat aacttttgc actgtaggtc cgttaaggtt agaagaaggc tactttggtg 10920
tctattttct cttccataaa aaaagcctga ctccacttcc cgcgtttact gattactagc 10980
gaagctgcgg gtgcattttt tcaagataaa ggcatccccg attatattct ataccgatgt 11040
ggattgcgca tactttgtga acagaaagtg atagcgttga tgattcttca ttggtcagaa 11100
aattatgaac ggtttcttct attttgtctc tatatactac gtataggaaa tgtttacatt 11160

ttcgtattgt tttcgattca ctctatgaat agttcttact acaatttttt tgtctaaaga 11220

gtaatactag agataaacat aaaaaatgta gaggtcgagt ttagatgcaa gttcaaggag 11280

cgaaaggtgg atgggtaggt tatataggga tatagcacag agatatatag caaagagata 11340

cttttgagca atgtttgtgg aagcggtatt cgcaatattt tagtagctcg ttacagtccg 11400

gtgcgttttt ggtttttga aagtgcgtct tcagagcgct tttggttttc aaaagcgctc 11460

tgaagttcct atactttcta gagaatagga acttcggaat aggaacttca aagcgtttcc 11520

gaaaacgagc gcttccgaaa atgcaacgcg agctgcgcac atacagctca ctgttcacgt 11580

cgcacctata tctgcgtgtt gcctgtatat atatatacat gagaagaacg gcatagtgcg 11640

tgtttatgct taaatgcgta cttatatgcg tctatttatg taggatgaaa ggtagtctag 11700

tacctcctgt gatattatcc cattccatgc ggggtatcgt atgcttcctt cagcactacc 11760

ctttagctgt tctatatgct gccactcctc aattggatta gtctcatcct tcaatgctat 11820

catttccttt gatattggat catatgcata gtaccgagaa actagaggat c 11871

<210> 177
<211> 52
<212> DNA
<213> Artificial sequence

<220>
<223> Primer OT1349

<400> 177
catcatcaca gtttaaacag tatgaaagtt ttctacgata aagactgcga cc 52

<210> 178
<211> 51
<212> DNA
<213> Artificial sequence

<220>
<223> Primer OT1318

<400> 178
gcacttgata ggcctgcagg gccttagttc ttggctttgt cgacgatttt g 51

<210> 179
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Primer OT1383

<400> 179
ctagtcaccg gtggc 15

<210> 180
<211> 15
<212> DNA
<213> Artificial sequence

<220>
<223> Primer OT1384

<400> 180
ggccgccacc ggtga 15

20 Jan 2017 2017200439

<210> 181
<211> 9612
<212> DNA
<213> Artificial sequence

<220>
<223> pHR81::ILV5p-K9D3 (pLH702) plasmid

<400> 181
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 60

cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 120

cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgctaagga 180

ttggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 240

ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 300

catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca 360

tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 420

aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 480

cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 540

tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 600

cggtgaacaa gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 660

cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 720

gttgatcgtt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 780

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 840

ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt 900

tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 960

acacccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 1020

caagttgatt aacaactgat attttcctct ggccctgcag gcctatcaag tgctggaaac 1080

tttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac 1140

atttaagatt ttttttttt catccgacat acatctgtac actaggaagc cctgtttttc 1200

tgaagcagct tcaaatatat atatttttta catatttatt atgattcaat gaacaatcta 1260

attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg 1320

tataagtcct catcgggaca gctacgattt ctctttcggt tttggctgag ctactggttg 1380

ctgtgacgca gcggcattag cgcggcgtta tgagctaccc tcgtggcctg aaagatggcg 1440

ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg 1500

tcaagtcacg tttggtggac ggccccttc caacgaatcg tatatactaa catgcgcgcg 1560

cttcctatat acacatatac atatatatat atatatatat gtgtgcgtgt atgtgtacac 1620

ctgtatttaa tttccttact cgcgggtttt tctttttttct caattcttgg cttcctcttt 1680

ctcgagcgga ccggatcctc cgcggtgccg gcagatctat ttaaatggcg cgccgacgtc 1740

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 1800

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 1860

aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 1920
ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 1980
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 2040
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 2100
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 2160
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 2220
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 2280
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 2340
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 2400
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 2460
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 2520
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 2580
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 2640
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga 2700
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 2760
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga 2820
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 2880
agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 2940
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 3000
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgta 3060
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 3240
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 3300
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 3360
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 3420
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 3480
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 3540
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 3600
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 3720
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 3780
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 3900
cgccaagctt tttctttcca attttttttt tttcgtcatt ataaaaatca ttacgaccga 3960
gattcccggg taataactga tataattaaa ttgaagctct aatttgtgag tttagtatac 4020

atgcatttac ttataataca gtttttagt tttgctggcc gcatcttctc aaatatgctt 4080
cccagcctgc ttttctgtaa cgttcaccct ctaccttagc atcccttccc tttgcaaata 4140
gtcctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat 4200
actgttgacc caatgcgtct cccttgtcat ctaaacccac accgggtgtc ataatcaacc 4260
aatcgtaacc ttcatctctt ccacccatgt ctctttgagc aataaagccg ataacaaaat 4320
ctttgtcgct cttcgcaatg tcaacagtac ccttagtata ttctccagta gatagggagc 4380
ccttgcatga caattctgct aacatcaaaa ggcctctagg ttcctttgtt acttcttctg 4440
ccgcctgctt caaaccgcta acaatacctg ggcccaccac accgtgtgca ttcgtaatgt 4500
ctgcccattc tgctattctg tatacacccg cagagtactg caatttgact gtattaccaa 4560
tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta 4620
gcggcttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta 4680
aacaaatttt gggacctaat gcttcaacta actccagtaa ttccttggtg gtacgaacat 4740
ccaatgaagc acacaagttt gtttgctttt cgtgcatgat attaaatagc ttggcagcaa 4800
caggactagg atgagtagca gcacgttcct tatatgtagc tttcgacatg atttatcttc 4860
gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa tactgggcaa tttcatgttt 4920
cttcaacact acatatgcgt atatatacca atctaagtct gtgctccttc cttcgttctt 4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaaag 5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg caggtcgact ctagtatact 5100
ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact 5160
cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatcttcg 5220
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 5280
atggctgcca tcattattat ccgatgtgac gctgcatttt ttttttttt ttttttttt 5340
ttttttttt ttttttttt ttttttttgt acaaatatca taaaaaaaga gaatcttttt 5400
aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 5460
aaccacctaa atcaccagtt ctgatacctg catccaaaac ctttttaact gcatcttcaa 5520
tggctttacc ttcttcaggc aagttcaatg acaatttcaa catcattgca gcagacaaga 5580
tagtggcgat agggttgacc ttattctttg gcaaatctgg agcggaacca tggcatggtt 5640
cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 5700
aacccaagga gcctgggata acggaggctt catcggagat gatatcacca aacatgttgc 5760
tggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat 5820
caatcaattg atgttgaact ttcaatgtag ggaattcgtt cttgatggtt tcctccacag 5880
tttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca 5940
atggtggctc atgttgtagg gccatgaaag cggccattct tgtgattctt tgcacttctg 6000
gaacggtgta ttgttcacta tcccaagcga caccatcacc atcgtcttcc tttctcttac 6060
caaagtaaat acctcccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt 6120
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 6180

tggcgtacaa ttgaagttct ttacggattt ttagtaaacc ttgttcaggt ctaacactac 6240
cggtacccca tttaggacca cccacagcac ctaacaaaac ggcatcagcc ttcttggagg 6300
cttccagcgc ctcatctgga agtggaacac ctgtagcatc gatagcagca ccaccaatta 6360
aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 6420
taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 6480
taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 6600
tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat tgaattttga 6660
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 6720
tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 6780
atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct 6840
tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa 6900
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa 7020
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg cattttttaca 7080
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttg 7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt 7200
tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact gtaggtccgt 7260
taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc 7320
cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttca agataaaggc 7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat 7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt 7560
tcttactaca attttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag 7620
gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat atagggatat 7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtattcgc 7740
aatattttag tagctcgtta cagtccggtg cgtttttggt tttttgaaag tgcgtcttca 7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact 7860
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc 7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata 7980
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 8040
atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat tccatgcggg 8100
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat 8160
tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat atgcatagta 8220
ccgagaaact agaggatctc ccattaccga catttgggcg ctatacgtgc atatgttcat 8280
gtatgtatct gtatttaaaa cacttttgta ttatttttcc tcatatatgt gtataggttt 8340

atacggatga tttaattatt acttcaccac cctttatttc aggctgatat cttagccttg 8400
ttactagtca ccggtggcgg ccgcacctgg taaaacctct agtggagtag tagatgtaat 8460
caatgaagcg gaagccaaaa gaccagagta gaggcctata gaagaaactg cgataccttt 8520
tgtgatggct aaacaaacag acatcttttt atatgttttt acttctgtat atcgtgaagt 8580
agtaagtgat aagcgaattt ggctaagaac gttgtaagtg aacaagggac ctcttttgcc 8640
tttcaaaaaa ggattaaatg gagttaatca ttgagattta gttttcgtta gattctgtat 8700
ccctaaataa ctcccttacc cgacgggaag gcacaaaaga cttgaataat agcaaacggc 8760
cagtagccaa gaccaaataa tactagagtt aactgatggt cttaaacagg cattacgtgg 8820
tgaactccaa gaccaatata caaaatatcg ataagttatt cttgcccacc aatttaagga 8880
gcctacatca ggacagtagt accattcctc agagaagagg tatacataac aagaaaatcg 8940
cgtgaacacc ttatataact tagcccgtta ttgagctaaa aaaccttgca aaatttccta 9000
tgaataagaa tacttcagac gtgataaaaa tttactttct aactcttctc acgctgcccc 9060
tatctgttct tccgctctac cgtgagaaat aaagcatcga gtacggcagt tcgctgtcac 9120
tgaactaaaa caataaggct agttcgaatg atgaacttgc ttgctgtcaa acttctgagt 9180
tgccgctgat gtgacactgt gacaataaat tcaaaccggt tatagcggtc tcctccggta 9240
ccggttctgc cacctccaat agagctcagt aggagtcaga acctctgcgg tggctgtcag 9300
tgactcatcc gcgtttcgta agttgtgcgc gtgcacattt cgcccgttcc cgctcatctt 9360
gcagcaggcg gaaattttca tcacgctgta ggacgcaaaa aaaaaataat taatcgtaca 9420
agaatcttgg aaaaaaaatt gaaaaatttt gtataaaagg gatgacctaa cttgactcaa 9480
tggcttttac acccagtatt ttccctttcc ttgtttgtta caattataga agcaagacaa 9540
aaacatatag acaacctatt cctaggagtt atatttttt accctaccag caatataagt 9600
aaaaaactgt tt 9612

<210> 182
<211> 12896
<212> DNA
<213> artificial sequence

<220>
<223> pBP915 plasmid

<400> 182
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360
ttttttttt ccacctagcg gatgactctt ttttttttct agcgattggc attatcacat 420
aatgaattat acattatata aagtaatgtg atttcttcga agaatatact aaaaaatgag 480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca 540

aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660
attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat 720
tccggctggt cgctaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt 840
ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcggtg 900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020
attaccctcc acgttgattg tctgcgaggc aagaatgatc atcaccgtag tgagagtgcg 1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgat 1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg 1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg 1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1440
aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 1620
aaaaccgtct atcagggcga tggcccacta cgtgaaccat caccctaatc aagtttttg 1680
gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag ggagcccccg atttagagct 1740
tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc 1800
gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt 1860
aatgcgccgc tacagggcgc gtccattcgc cattcaggct gcgcaactgt tgggaagggc 1920
gcggtgcggg cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga 1980
ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgag 2040
cgcgcgtaat acgactcact atagggcgaa ttgggtaccg ggcccccccct cgaggtcgac 2100
ggcgcgccac tggtagagag cgactttgta tgccccaatt gcgaaacccg cgatatcctt 2160
ctcgattctt tagtacccga ccaggacaag gaaaaggagg tcgaaacgtt tttgaagaaa 2220
caagaggaac tacacggaag ctctaaagat ggcaaccagc cagaaactaa gaaaatgaag 2280
ttgatggatc caactggcac cgctggcttg aacaacaata ccagccttcc aacttctgta 2340
aataacggcg gtacgccagt gccaccagta ccgttacctt tcggtatacc tcctttcccc 2400
atgtttccaa tgcccttcat gcctccaacg gctactatca caaatcctca tcaagctgac 2460
gcaagcccta agaaatgaat aacaatactg acagtactaa ataattgcct acttggcttc 2520
acatacgttg catacgtcga tatagataat aatgataatg acagcaggat tatcgtaata 2580
cgtaatagct gaaaatctca aaaatgtgtg ggtcattacg taaataatga taggaatggg 2640
attcttctat ttttcctttt tccattctag cagccgtcgg gaaaacgtgg catcctctct 2700

ttcgggctca attggagtca cgctgccgtg agcatcctct ctttccatat ctaacaactg 2760
agcacgtaac caatggaaaa gcatgagctt agcgttgctc caaaaaagta ttggatggtt 2820
aataccattt gtctgttctc ttctgacttt gactcctcaa aaaaaaaaat ctacaatcaa 2880
cagatcgctt caattacgcc ctcacaaaaa ctttttttcct tcttcttcgc ccacgttaaa 2940
ttttatccct catgttgtct aacggatttc tgcacttgat ttattataaa aagacaaaga 3000
cataatactt ctctatcaat ttcagttatt gttcttcctt gcgttattct tctgttcttc 3060
tttttctttt gtcatatata accataacca agtaatacat attcaaacta gtatgactga 3120
caaaaaaact cttaaagact taagaaatcg tagttctgtt tacgattcaa tggttaaatc 3180
acctaatcgt gctatgttgc gtgcaactgg tatgcaagat gaagactttg aaaaacctat 3240
cgtcggtgtc atttcaactt gggctgaaaa cacaccttgt aatatccact tacatgactt 3300
tggtaaacta gccaaagtcg gtgttaagga agctggtgct tggccagttc agttcggaac 3360
aatcacggtt tctgatggaa tcgccatggg aacccaagga atgcgtttct ccttgacatc 3420
tcgtgatatt attgcagatt ctattgaagc agccatggga ggtcataatg cggatgcttt 3480
tgtagccatt ggcggttgtg ataaaaacat gcccggttct gttatcgcta tggctaacat 3540
ggatatccca gccatttttg cttacggcgg aacaattgca cctggtaatt tagacggcaa 3600
agatatcgat ttagtctctg tctttgaagg tgtcggccat tggaaccacg gcgatatgac 3660
caaagaagaa gttaaagctt tggaatgtaa tgcttgtccc ggtcctggag gctgcggtgg 3720
tatgtatact gctaacacaa tggcgacagc tattgaagtt ttgggactta gccttccggg 3780
ttcatcttct cacccggctg aatccgcaga aaagaaagca gatattgaag aagctggtcg 3840
cgctgttgtc aaaatgctcg aaatgggctt aaaaccttct gacattttaa cgcgtgaagc 3900
ttttgaagat gctattactg taactatggc tctgggaggt tcaaccaact caacccttca 3960
cctcttagct attgcccatg ctgctaatgt ggaattgaca cttgatgatt tcaatacttt 4020
ccaagaaaaa gttcctcatt tggctgattt gaaaccttct ggtcaatatg tattccaaga 4080
cctttacaag gtcggagggg taccagcagt tatgaaatat ctccttaaaa atggcttcct 4140
tcatggtgac cgtatcactt gtactggcaa aacagtcgct gaaaatttga aggcttttga 4200
tgatttaaca cctggtcaaa aggttattat gccgcttgaa aatcctaaac gtgaagatgg 4260
tccgctcatt attctccatg gtaacttggc tccagacggt gccgttgcca aagtttctgg 4320
tgtaaaagtg cgtcgtcatg tcggtcctgc taaggtcttt aattctgaag aagaagccat 4380
tgaagctgtc ttgaatgatg atattgttga tggtgatgtt gttgtcgtac gttttgtagg 4440
accaaagggc ggtcctggta tgcctgaaat gctttccctt tcatcaatga ttgttggtaa 4500
agggcaaggt gaaaaagttg cccttctgac agatggccgc ttctcaggtg gtacttatgg 4560
tcttgtcgtg ggtcatatcg ctcctgaagc acaagatggc ggtccaatcg cctacctgca 4620
aacaggagac atagtcacta ttgaccaaga cactaaggaa ttacactttg atatctccga 4680
tgaagagtta aaacatcgtc aagagaccat tgaattgcca ccgctctatt cacgcggtat 4740
ccttggtaaa tatgctcaca tcgtttcgtc tgcttctagg ggagccgtaa cagacttttg 4800
gaagcctgaa gaaactggca aaaaatgttg tcctggttgc tgtggttaag cggccgcgtt 4860

```
aattcaaatt aattgatata gtttttaat gagtattgaa tctgtttaga aataatggaa     4920
tattattttt atttatttat ttatattatt ggtcggctct tttcttctga aggtcaatga     4980
caaaatgata tgaaggaaat aatgatttct aaaattttac aacgtaagat atttttacaa     5040
aagcctagct catcttttgt catgcactat tttactcacg cttgaaatta acggccagtc     5100
cactgcggag tcatttcaaa gtcatcctaa tcgatctatc gtttttgata gctcattttg     5160
gagttcgcga ttgtcttctg ttattcacaa ctgttttaat ttttatttca ttctggaact     5220
cttcgagttc tttgtaaagt ctttcatagt agcttacttt atcctccaac atatttaact     5280
tcatgtcaat ttcggctctt aaattttcca catcatcaag ttcaacatca tcttttaact     5340
tgaatttatt ctctagctct tccaaccaag cctcattgct ccttgattta ctggtgaaaa     5400
gtgatacact ttgcgcgcaa tccaggtcaa aactttcctg caaagaattc accaatttct     5460
cgacatcata gtacaatttg ttttgttctc ccatcacaat ttaatatacc tgatggattc     5520
ttatgaagcg ctgggtaatg gacgtgtcac tctacttcgc ctttttccct actcctttta     5580
gtacggaaga caatgctaat aaataagagg gtaataataa tattattaat cggcaaaaaa     5640
gattaaacgc caagcgttta attatcagaa agcaaacgtc gtaccaatcc ttgaatgctt     5700
cccaattgta tattaagagt catcacagca acatattctt gttattaaat taattattat     5760
tgatttttga tattgtataa aaaaaccaaa tatgtataaa aaaagtgaat aaaaaatacc     5820
aagtatggag aaatatatta gaagtctata cgttaaacca cccgggcccc ccctcgaggt     5880
cgacggtatc gataagcttg atatcgaatt cctgcagccc gggggatcca ctagttctag     5940
agcggccgct ctagaactag taccacaggt gttgtcctct gaggacataa aatacacacc     6000
gagattcatc aactcattgc tggagttagc atatctacaa ttgggtgaaa tggggagcga     6060
tttgcaggca tttgctcggc atgccggtag aggtgtggtc aataagagcg acctcatgct     6120
atacctgaga aagcaacctg acctacagga aagagttact caagaataag aattttcgtt     6180
ttaaaaccta agagtcactt taaaatttgt atacacttat ttttttata acttatttaa     6240
taataaaaat cataaatcat aagaaattcg cttactctta attaatcaaa aagttaaaat     6300
tgtacgaata gattcaccac ttcttaacaa atcaaaccct tcattgattt tctcgaatgg     6360
caatacatgt gtaattaaag gatcaagagc aaacttcttc gccataaagt cggcaacaag     6420
ttttggaaca ctatccttgc tcttaaaacc gccaaatata gctcccttcc atgtacgacc     6480
gcttagcaac agcataggat tcatcgacaa attttgtgaa tcaggaggaa cacctacgat     6540
cacactgact ccatatgcct cttgacagca ggacaacgca gttaccatag tatcaagacg     6600
gcctataact tcaaaagaga aatcaactcc accgtttgac atttcagtaa ggacttcttg     6660
tattggtttc ttataatctt gagggttaac acattcagta gccccgacct ccttagcttt     6720
tgcaaatttg tccttattga tgtctacacc tataatcctc gctgcgcctg cagctttaca     6780
ccccataata acgcttagtc ctactcctcc taaaccgaat actgcacaag tcgaaccctg     6840
tgtaaccttt gcaactttaa ctgcggaacc gtaaccggtg gaaaatccgc accctatcaa     6900
gcaaactttt tccagtggtg aagctgcatc gattttagcg acagatatct cgtccaccac     6960
tgtgtattgg gaaaatgtag aagtaccaag gaaatggtgt ataggtttcc ctctgcatgt     7020
```

aaatctgctt gtaccatcct gcatagtacc tctaggcata gacaaatcat ttttaaggca 7080
gaaattaccc tcaggatgtt tgcagactct acacttacca cattgaggag tgaacagtgg 7140
gatcacttta tcaccaggac gaacagtggt aacaccttca cctatggatt caacgattcc 7200
ggcagcctcg tgtcccgcga ttactggcaa aggagtaact agagtgccac tcaccacatg 7260
gtcgtcggat ctacagattc cggtggcaac catcttgatt ctaacctcgt gtgcttttgg 7320
tggcgctact tctacttctt ctatgctaaa cggctttttc tcttcccaca aaactgccgc 7380
tttacactta ataactttac cggctgttga catcctcagc tagctattgt aatatgtgtg 7440
tttgtttgga ttattaagaa gaataattac aaaaaaaatt acaaaggaag gtaattacaa 7500
cagaattaag aaaggacaag aaggaggaag agaatcagtt cattatttct tctttgttat 7560
ataacaaacc caagtagcga tttggccata cattaaaagt tgagaaccac cctccctggc 7620
aacagccaca actcgttacc attgttcatc acgatcatga aactcgctgt cagctgaaat 7680
ttcacctcag tggatctctc tttttattct tcatcgttcc actaaccttt ttccatcagc 7740
tggcagggaa cggaaagtgg aatcccattt agcgagcttc ctcttttctt caagaaaaga 7800
cgaagcttgt gtgtgggtgc gcgcgctagt atctttccac attaagaaat ataccataaa 7860
ggttacttag acatcactat ggctatatat atatatatat atatatgtaa cttagcacca 7920
tcgcgcgtgc atcactgcat gtgttaaccg aaaagtttgg cgaacacttc accgacacgg 7980
tcatttagat ctgtcgtctg cattgcacgt cccttagcct taaatcctag gcgggagcat 8040
tctcgtgtaa ttgtgcagcc tgcgtagcaa ctcaacatag cgtagtctac ccagtttttc 8100
aagggtttat cgttagaaga ttctcccttt tcttcctgct cacaaatctt aaagtcatac 8160
attgcacgac taaatgcaag catgcggatc ccccgggctg caggaattcg atatcaagct 8220
tatcgatacc gtcgactggc cattaatctt tcccatatta gatttcgcca agccatgaaa 8280
gttcaagaaa ggtctttaga cgaattaccc ttcatttctc aaactggcgt caagggatcc 8340
tggtatggtt ttatcgtttt atttctggtt cttatagcat cgttttggac ttctctgttc 8400
ccattaggcg gttcaggagc cagcgcagaa tcattctttg aaggatactt atcctttcca 8460
attttgattg tctgttacgt tggacataaa ctgtatacta gaaattggac tttgatggtg 8520
aaactagaag atatggatct tgataccggc agaaaacaag tagatttgac tcttcgtagg 8580
gaagaaatga ggattgagcg agaaacatta gcaaaaagat ccttcgtaac aagattttta 8640
catttctggt gttgaaggga aagatatgag ctatacagcg gaatttccat atcactcaga 8700
ttttgttatc taatttttc cttcccacgt ccgcgggaat ctgtgtatat tactgcatct 8760
agatatatgt tatcttatct tggcgcgtac atttaatttt caacgtattc tataagaaat 8820
tgcgggagtt tttttcatgt agatgatact gactgcacgc aaatataggc atgatttata 8880
ggcatgattt gatggctgta ccgataggaa cgctaagagt aacttcagaa tcgttatcct 8940
ggcggaaaaa attcatttgt aaactttaaa aaaaaaagcc aatatcccca aaattattaa 9000
gagcgcctcc attattaact aaaatttcac tcagcatcca caatgtatca ggtatctact 9060
acagatatta catgtggcga aaaagacaag aacaatgcaa tagcgcatca agaaaaaaca 9120
caaagctttc aatcaatgaa tcgaaaatgt cattaaaata gtatataaat tgaaactaag 9180

tcataaagct ataaaaagaa aatttattta aatcttggct ctcttgggct caaggtgaca 9240
aggtcctcga aaatagggcg cgccccaccg cggtggagct ccagcttttg ttccctttag 9300
tgagggttaa ttgcgcgctt ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt 9360
tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctggggt 9420
gcctaatgag tgagctaact cacattaatt gcgttgcgct cactgcccgc tttccagtcg 9480
ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg 9540
cgtattgggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 9600
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 9660
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 9720
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc 9780
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 9840
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt 9900
ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg 9960
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 10020
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 10080
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 10140
ttgaagtggt ggcctaacta cggctacact agaagaacag tatttggtat ctgcgctctg 10200
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 10260
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct 10320
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 10380
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 10440
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 10500
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc 10560
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct 10620
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca 10680
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt 10740
aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg caacgttgtt 10800
gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc 10860
ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa agcggttagc 10920
tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc actcatggtt 10980
atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt ttctgtgact 11040
ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc 11100
ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt gctcatcatt 11160
ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag atccagttcg 11220
atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac cagcgtttct 11280
gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa 11340

tgttgaatac tcatactctt ccttttttcaa tattattgaa gcatttatca gggttattgt 11400
ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc 11460
acatttcccc gaaaagtgcc acctgaacga agcatctgtg cttcattttg tagaacaaaa 11520
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 11580
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa 11640
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg cattttttaca 11700
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttttg 11760
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt 11820
tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact gtaggtccgt 11880
taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc 11940
cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttttca agataaaggc 12000
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 12060
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat 12120
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt 12180
tcttactaca attttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag 12240
gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat atagggatat 12300
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtattcgc 12360
aatattttag tagctcgtta cagtccggtg cgtttttggt tttttgaaag tgcgtcttca 12420
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact 12480
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc 12540
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata 12600
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 12660
atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat tccatgcggg 12720
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat 12780
tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat actaagaaac 12840
cattattatc atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 12896

<210> 183
<211> 22
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP594

<400> 183
agctgtctcg tgttgtgggt tt 22

<210> 184
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP595

<400> 184
cttaataata gaacaatatc atcctttacg ggcatcttat agtgtcgtt 49

<210> 185
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP596

<400> 185
gcgccaacga cactataaga tgcccgtaaa ggatgatatt gttctatta 49

<210> 186
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP597

<400> 186
tatggaccct gaaaccacag ccacattgca acgacgacaa tgccaaacc 49

<210> 187
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP598

<400> 187
tccttggttt ggcattgtcg tcgttgcaat gtggctgtgg tttcagggt 49

<210> 188
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP599

<400> 188
atcctctcgc ggagtccctg ttcagtaaag gccatgaagc tttttcttt 49

<210> 189
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP600

<400> 189
attggaaaga aaaagcttca tggcctttac tgaacaggga ctccgcgag 49

<210> 190
<211> 22
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP601

<400> 190
tcataccaca atcttagacc at 22

<210> 191
<211> 21
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP602

<400> 191
tgttcaaacc cctaaccaac c 21

<210> 192
<211> 22
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP603

<400> 192
tgttcccaca atctattacc ta 22

<210> 193
<211> 21
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP605

<400> 193
tactgaacag ggactccgcg a 21

<210> 194
<211> 21
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP606

<400> 194
tcataccaca atcttagacc a 21

<210> 195
<211> 11856
<212> DNA
<213> Artificial sequence

<220>
<223> Plasmid pYZ090

<400> 195
tcccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60

aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120

ttacttcacc acccttttatt tcaggctgat atcttagcct tgttactagt tagaaaaaga 180

catttttgct gtcagtcact gtcaagagat tcttttgctg gcatttcttc tagaagcaaa 240

aagagcgatg cgtcttttcc gctgaaccgt tccagcaaaa aagactacca acgcaatatg 300

gattgtcaga atcatataaa agagaagcaa ataactcctt gtcttgtatc aattgcatta 360
taatatcttc ttgttagtgc aatatcatat agaagtcatc gaaatagata ttaagaaaaa 420
caaactgtac aatcaatcaa tcaatcatcg ctgaggatgt tgacaaaagc aacaaaagaa 480
caaaaatccc ttgtgaaaaa cagaggggcg gagcttgttg ttgattgctt agtggagcaa 540
ggtgtcacac atgtatttgg cattccaggt gcaaaaattg atgcggtatt tgacgcttta 600
caagataaag gacctgaaat tatcgttgcc cggcacgaac aaaacgcagc attcatggcc 660
caagcagtcg gccgtttaac tggaaaaccg ggagtcgtgt tagtcacatc aggaccgggt 720
gcctctaact tggcaacagg cctgctgaca gcgaacactg aaggagaccc tgtcgttgcg 780
cttgctggaa acgtgatccg tgcagatcgt ttaaaacgga cacatcaatc tttggataat 840
gcggcgctat tccagccgat tacaaaatac agtgtagaag ttcaagatgt aaaaaatata 900
ccggaagctg ttacaaatgc atttaggata gcgtcagcag ggcaggctgg ggccgctttt 960
gtgagctttc cgcaagatgt tgtgaatgaa gtcacaaata cgaaaaacgt gcgtgctgtt 1020
gcagcgccaa aactcggtcc tgcagcagat gatgcaatca gtgcggccat agcaaaaatc 1080
caaacagcaa aacttcctgt cgttttggtc ggcatgaaag gcggaagacc ggaagcaatt 1140
aaagcggttc gcaagctttt gaaaaaggtt cagcttccat ttgttgaaac atatcaagct 1200
gccggtaccc tttctagaga tttagaggat caatattttg gccgtatcgg tttgttccgc 1260
aaccagcctg gcgatttact gctagagcag gcagatgttg ttctgacgat cggctatgac 1320
ccgattgaat atgatccgaa attctggaat atcaatggag accggacaat tatccattta 1380
gacgagatta tcgctgacat tgatcatgct taccagcctg atcttgaatt gatcggtgac 1440
attccgtcca cgatcaatca tatcgaacac gatgctgtga aagtggaatt tgcagagcgt 1500
gagcagaaaa tcctttctga tttaaaacaa tatatgcatg aaggtgagca ggtgcctgca 1560
gattggaaat cagacagagc gcaccctctt gaaatcgtta aagagttgcg taatgcagtc 1620
gatgatcatg ttacagtaac ttgcgatatc ggttcgcacg ccatttggat gtcacgttat 1680
ttccgcagct acgagccgtt aacattaatg atcagtaacg gtatgcaaac actcggcgtt 1740
gcgcttcctt gggcaatcgg cgcttcattg gtgaaaccgg gagaaaaagt ggtttctgtc 1800
tctggtgacg gcggtttctt attctcagca atggaattag agacagcagt tcgactaaaa 1860
gcaccaattg tacacattgt atggaacgac agcacatatg acatggttgc attccagcaa 1920
ttgaaaaaat ataaccgtac atctgcggtc gatttcggaa atatcgatat cgtgaaatat 1980
gcggaaagct tcggagcaac tggcttgcgc gtagaatcac cagaccagct ggcagatgtt 2040
ctgcgtcaag gcatgaacgc tgaaggtcct gtcatcatcg atgtcccggt tgactacagt 2100
gataacatta atttagcaag tgacaagctt ccgaaagaat tcggggaact catgaaaacg 2160
aaagctctct agttaattaa tcatgtaatt agttatgtca cgcttacatt cacgccctcc 2220
ccccacatcc gctctaaccg aaaaggaagg agttagacaa cctgaagtct aggtccctat 2280
ttatttttt atagttatgt tagtattaag aacgttattt atatttcaaa tttttctttt 2340
ttttctgtac agacgcgtgt acgcatgtaa cattatactg aaaaccttgc ttgagaaggt 2400
tttgggacgc tcgaaggctt taatttgcgg gcggccgcac ctggtaaaac ctctagtgga 2460

gtagtagatg taatcaatga agcggaagcc aaaagaccag agtagaggcc tatagaagaa 2520
actgcgatac cttttgtgat ggctaaacaa acagacatct ttttatatgt ttttacttct 2580
gtatatcgtg aagtagtaag tgataagcga atttggctaa gaacgttgta agtgaacaag 2640
ggacctcttt tgcctttcaa aaaaggatta aatggagtta atcattgaga tttagttttc 2700
gttagattct gtatccctaa ataactccct tacccgacgg gaaggcacaa aagacttgaa 2760
taatagcaaa cggccagtag ccaagaccaa ataatactag agttaactga tggtcttaaa 2820
caggcattac gtggtgaact ccaagaccaa tatacaaaat atcgataagt tattcttgcc 2880
caccaattta aggagcctac atcaggacag tagtaccatt cctcagagaa gaggtataca 2940
taacaagaaa atcgcgtgaa caccttatat aacttagccc gttattgagc taaaaaacct 3000
tgcaaaattt cctatgaata agaatacttc agacgtgata aaaatttact ttctaactct 3060
tctcacgctg cccctatctg ttcttccgct ctaccgtgag aaataaagca tcgagtacgg 3120
cagttcgctg tcactgaact aaaacaataa ggctagttcg aatgatgaac ttgcttgctg 3180
tcaaacttct gagttgccgc tgatgtgaca ctgtgacaat aaattcaaac cggttatagc 3240
ggtctcctcc ggtaccggtt ctgccacctc caatagagct cagtaggagt cagaacctct 3300
gcggtggctg tcagtgactc atccgcgttt cgtaagttgt gcgcgtgcac atttcgcccg 3360
ttcccgctca tcttgcagca ggcggaaatt ttcatcacgc tgtaggacgc aaaaaaaaaa 3420
taattaatcg tacaagaatc ttggaaaaaa aattgaaaaa ttttgtataa aagggatgac 3480
ctaacttgac tcaatggctt ttacacccag tattttccct ttccttgttt gttacaatta 3540
tagaagcaag acaaaaacat atagacaacc tattcctagg agttatattt ttttacccta 3600
ccagcaatat aagtaaaaaa ctgtttaaac agtatggcag ttacaatgta ttatgaagat 3660
gatgtagaag tatcagcact tgctggaaag caaattgcag taatcggtta tggttcacaa 3720
ggacatgctc acgcacagaa tttgcgtgat tctggtcaca acgttatcat tggtgtgcgc 3780
cacggaaaat cttttgataa agcaaaagaa gatggctttg aaacatttga agtaggagaa 3840
gcagtagcta aagctgatgt tattatggtt ttggcaccag atgaacttca acaatccatt 3900
tatgaagagg acatcaaacc aaacttgaaa gcaggttcag cacttggttt tgctcacgga 3960
tttaatatcc attttggcta tattaaagta ccagaagacg ttgacgtctt tatggttgcg 4020
cctaaggctc caggtcacct tgtccgtcgg acttatactg aaggttttgg tacaccagct 4080
ttgtttgttt cacaccaaaa tgcaagtggt catgcgcgtg aaatcgcaat ggattgggcc 4140
aaaggaattg gttgtgctcg agtgggaatt attgaaacaa cttttaaaga agaaacagaa 4200
gaagatttgt ttggagaaca agctgttcta tgtggaggtt tgacagcact tgttgaagcc 4260
ggttttgaaa cactgacaga agctggatac gctggcgaat tggcttactt tgaagttttg 4320
cacgaaatga aattgattgt tgacctcatg tatgaaggtg gttttactaa aatgcgtcaa 4380
tccatctcaa atactgctga gtttggcgat tatgtgactg gtccacggat tattactgac 4440
gaagttaaaa agaatatgaa gcttgttttg gctgatattc aatctggaaa atttgctcaa 4500
gatttcgttg atgacttcaa agcggggcgt ccaaaattaa tagcctatcg cgaagctgca 4560
aaaaatcttg aaattgaaaa aattggggca gagctacgtc aagcaatgcc attcacacaa 4620

tctggtgatg acgatgcctt taaaatctat cagtaaggcc ctgcaggcca gaggaaaata 4680
atatcaagtg ctggaaactt tttctcttgg aatttttgca acatcaagtc atagtcaatt 4740
gaattgaccc aatttcacat ttaagatttt tttttttca tccgacatac atctgtacac 4800
taggaagccc tgtttttctg aagcagcttc aaatatatat atttttaca tatttattat 4860
gattcaatga acaatctaat taaatcgaaa acaagaaccg aaacgcgaat aaataattta 4920
tttagatggt gacaagtgta taagtcctca tcgggacagc tacgatttct ctttcggttt 4980
tggctgagct actggttgct gtgacgcagc ggcattagcg cggcgttatg agctaccctc 5040
gtggcctgaa agatggcggg aataaagcgg aactaaaaat tactgactga gccatattga 5100
ggtcaatttg tcaactcgtc aagtcacgtt tggtggacgg ccccttcca acgaatcgta 5160
tatactaaca tgcgcgcgct tcctatatac acatatacat atatatatat atatatatgt 5220
gtgcgtgtat gtgtacacct gtatttaatt tccttactcg cgggtttttc ttttttctca 5280
attcttggct tcctctttct cgagcggacc ggatcctccg cggtgccggc agatctattt 5340
aaatggcgcg ccgacgtcag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg 5400
tttatttttc taaatacatt caaatatgta tccgctcatg agacaataac cctgataaat 5460
gcttcaataa tattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat 5520
tccctttttt gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt 5580
aaaagatgct gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag 5640
cggtaagatc cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa 5700
agttctgcta tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg 5760
ccgcatacac tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct 5820
tacggatggc atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac 5880
tgcggccaac ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttgca 5940
caacatgggg gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat 6000
accaaacgac gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact 6060
attaactggc gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc 6120
ggataaagtt gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga 6180
taaatctgga gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg 6240
taagccctcc cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg 6300
aaatagacag atcgctgaga taggtgcctc actgattaag cattggtaac tgtcagacca 6360
agtttactca tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta 6420
ggtgaagatc ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca 6480
ctgagcgtca gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg 6540
cgtaatctgc tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga 6600
tcaagagcta ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa 6660
tactgttctt ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc 6720
tacatacctc gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg 6780

tcttaccggg ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac 6840
gggggttcg tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct 6900
acagcgtgag ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc 6960
ggtaagcggc agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg 7020
gtatctttat agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg 7080
ctcgtcaggg gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct 7140
ggccttttgc tggccttttg ctcacatgtt ctttcctgcg ttatcccctg attctgtgga 7200
taaccgtatt accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg 7260
cagcgagtca gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc 7320
gcgttggccg attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag 7380
tgagcgcaac gcaattaatg tgagttagct cactcattag gcaccccagg ctttacactt 7440
tatgcttccg gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa 7500
cagctatgac catgattacg ccaagctttt tctttccaat ttttttttt tcgtcattat 7560
aaaaatcatt acgaccgaga ttcccgggta ataactgata taattaaatt gaagctctaa 7620
tttgtgagtt tagtatacat gcatttactt ataatacagt tttttagttt tgctggccgc 7680
atcttctcaa atatgcttcc cagcctgctt ttctgtaacg ttcaccctct accttagcat 7740
cccttccctt tgcaaatagt cctcttccaa caataataat gtcagatcct gtagagacca 7800
catcatccac ggttctatac tgttgaccca atgcgtctcc cttgtcatct aaacccacac 7860
cgggtgtcat aatcaaccaa tcgtaacctt catctcttcc acccatgtct ctttgagcaa 7920
taaagccgat aacaaaatct ttgtcgctct tcgcaatgtc aacagtaccc ttagtatatt 7980
ctccagtaga tagggagccc ttgcatgaca attctgctaa catcaaaagg cctctaggtt 8040
cctttgttac ttcttctgcc gcctgcttca aaccgctaac aatacctggg cccaccacac 8100
cgtgtgcatt cgtaatgtct gcccattctg ctattctgta tacacccgca gagtactgca 8160
atttgactgt attaccaatg tcagcaaatt ttctgtcttc gaagagtaaa aaattgtact 8220
tggcggataa tgcctttagc ggcttaactg tgccctccat ggaaaaatca gtcaagatat 8280
ccacatgtgt ttttagtaaa caaattttgg gacctaatgc ttcaactaac tccagtaatt 8340
ccttggtggt acgaacatcc aatgaagcac acaagtttgt ttgcttttcg tgcatgatat 8400
taaatagctt ggcagcaaca ggactaggat gagtagcagc acgttcctta tatgtagctt 8460
tcgacatgat ttatcttcgt ttcctgcagg tttttgttct gtgcagttgg gttaagaata 8520
ctgggcaatt tcatgtttct tcaacactac atatgcgtat atataccaat ctaagtctgt 8580
gctccttcct tcgttcttcc ttctgttcgg agattaccga atcaaaaaaa tttcaaggaa 8640
accgaaatca aaaaaaagaa taaaaaaaaa atgatgaatt gaaaagcttg catgcctgca 8700
ggtcgactct agtatactcc gtctactgta cgatacactt ccgctcaggt ccttgtcctt 8760
taacgaggcc ttaccactct tttgttactc tattgatcca gctcagcaaa ggcagtgtga 8820
tctaagattc tatcttcgcg atgtagtaaa actagctaga ccgagaaaga gactagaaat 8880
gcaaaaggca cttctacaat ggctgccatc attattatcc gatgtgacgc tgcatttttt 8940

```
ttttttttt ttttttttt ttttttttt ttttttttt ttttttgtac aaatatcata     9000
aaaaaagaga atcttttaa gcaaggattt tcttaacttc ttcggcgaca gcatcaccga     9060
cttcggtggt actgttggaa ccacctaaat caccagttct gatacctgca tccaaaacct     9120
ttttaactgc atcttcaatg gctttacctt cttcaggcaa gttcaatgac aatttcaaca     9180
tcattgcagc agacaagata gtggcgatag ggttgacctt attctttggc aaatctggag     9240
cggaaccatg gcatggttcg tacaaaccaa atgcggtgtt cttgtctggc aaagaggcca     9300
aggacgcaga tggcaacaaa cccaaggagc ctgggataac ggaggcttca tcggagatga     9360
tatcaccaaa catgttgctg gtgattataa taccatttag gtgggttggg ttcttaacta     9420
ggatcatggc ggcagaatca atcaattgat gttgaacttt caatgtaggg aattcgttct     9480
tgatggtttc ctccacagtt tttctccata atcttgaaga ggccaaaaca ttagctttat     9540
ccaaggacca aataggcaat ggtggctcat gttgtagggc catgaaagcg gccattcttg     9600
tgattctttg cacttctgga acggtgtatt gttcactatc ccaagcgaca ccatcaccat     9660
cgtcttcctt tctcttacca aagtaaatac ctcccactaa ttctctaaca acaacgaagt     9720
cagtaccttt agcaaattgt ggcttgattg gagataagtc taaaagagag tcggatgcaa     9780
agttacatgg tcttaagttg gcgtacaatt gaagttcttt acggattttt agtaaacctt     9840
gttcaggtct aacactaccg gtaccccatt taggaccacc cacagcacct aacaaaacgg     9900
catcagcctt cttggaggct tccagcgcct catctggaag tggaacacct gtagcatcga     9960
tagcagcacc accaattaaa tgattttcga aatcgaactt gacattggaa cgaacatcag    10020
aaatagcttt aagaacctta atggcttcgg ctgtgatttc ttgaccaacg tggtcacctg    10080
gcaaaacgac gatcttctta ggggcagaca ttacaatggt atatccttga aatatatata    10140
aaaaaaaaaa aaaaaaaaaa aaaaaaaaat gcagcttctc aatgatattc gaatacgctt    10200
tgaggagata cagcctaata tccgacaaac tgttttacag atttacgatc gtacttgtta    10260
cccatcattg aattttgaac atccgaacct gggagttttc cctgaaacag atagtatatt    10320
tgaacctgta taataatata tagtctagcg ctttacggaa gacaatgtat gtatttcggt    10380
tcctggagaa actattgcat ctattgcata ggtaatcttg cacgtcgcat ccccggttca    10440
ttttctgcgt ttccatcttg cacttcaata gcatatcttt gttaacgaag catctgtgct    10500
tcattttgta gaacaaaaat gcaacgcgag agcgctaatt tttcaaacaa agaatctgag    10560
ctgcattttt acagaacaga aatgcaacgc gaaagcgcta ttttaccaac gaagaatctg    10620
tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc taatttttca aacaaagaat    10680
ctgagctgca tttttacaga acagaaatgc aacgcgagag cgctatttta ccaacaaaga    10740
atctatactt ctttttttgtt ctacaaaaat gcatcccgag agcgctattt ttctaacaaa    10800
gcatcttaga ttacttttttt tctcctttgt gcgctctata atgcagtctc ttgataactt    10860
tttgcactgt aggtccgtta aggttagaag aaggctactt tggtgtctat tttctcttcc    10920
ataaaaaaag cctgactcca cttcccgcgt ttactgatta ctagcgaagc tgcgggtgca    10980
ttttttcaag ataaaggcat ccccgattat attctatacc gatgtggatt gcgcatactt    11040
tgtgaacaga aagtgatagc gttgatgatt cttcattggt cagaaaatta tgaacggttt    11100
```

cttctatttt gtctctatat actacgtata ggaaatgttt acattttcgt attgttttcg 11160

attcactcta tgaatagttc ttactacaat ttttttgtct aaagagtaat actagagata 11220

aacataaaaa atgtagaggt cgagtttaga tgcaagttca aggagcgaaa ggtggatggg 11280

taggttatat agggatatag cacagagata tatagcaaag agatactttt gagcaatgtt 11340

tgtggaagcg gtattcgcaa tattttagta gctcgttaca gtccggtgcg tttttggttt 11400

tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag cgctctgaag ttcctatact 11460

ttctagagaa taggaacttc ggaataggaa cttcaaagcg tttccgaaaa cgagcgcttc 11520

cgaaaatgca acgcgagctg cgcacataca gctcactgtt cacgtcgcac ctatatctgc 11580

gtgttgcctg tatatatata tacatgagaa gaacggcata gtgcgtgttt atgcttaaat 11640

gcgtacttat atgcgtctat ttatgtagga tgaaaggtag tctagtacct cctgtgatat 11700

tatcccattc catgcggggt atcgtatgct tccttcagca ctacccttta gctgttctat 11760

atgctgccac tcctcaattg gattagtctc atccttcaat gctatcattt cctttgatat 11820

tggatcatat gcatagtacc gagaaactag aggatc 11856

<210> 196
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> K9_David_SH

<400> 196

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
130 135 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
165 170 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 197
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP562

<400> 197
aattgtttaa acatgtatac agtaggtgac tatc 34

<210> 198
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP563

<400> 198
aatcataaat cataagaaat tcgcttatca gctcttgttt tgttctgca 49

<210> 199
<211> 49
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP564

<400> 199
ttatttgcag aacaaaacaa gagctgataa gcgaatttct tatgattta 49

<210> 200
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP565

<400> 200
aattggccgg ccaaaaaaag catgcacgta taca 34

<210> 201
<211> 32
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP505

<400> 201
aattgagctc actgtagccc tagacttgat ag 32

<210> 202
<211> 35
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP506

<400> 202
aattggcgcg cctgtatatg agatagttga ttgta 35

<210> 203
<211> 35
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP507

<400> 203
aattttaatt aagtctaggt tctttggctg ttcaa 35

<210> 204
<211> 33
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP508

<400> 204
aattgtcgac tttagaagtg tcaacaacgt atc 33

<210> 205
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP674

<400> 205
aattggcgcg ccaattaccg tcgctcgtga tttg 34

<210> 206
<211> 34
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP675

<400> 206
aattgtttaa acttgaatat gtattacttg gtta 34

<210> 207
<211> 22
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP495

<400> 207
ggagatatac aatagaacag at 22

<210> 208
<211> 22
<212> DNA
<213> artificial sequence

<220>
<223> primer oBP496

<400> 208
tagcaatggg gtttttttca gt 22

<210> 209
<211> 8366
<212> DNA
<213> Artificial Sequence

<220>
<223> pRS316-UAS(PGK1)-PFBA1-GUS

<400> 209
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accacgcttt tcaattcaat tcatcatttt ttttttattc tttttttga tttcggtttc 240

tttgaaattt ttttgattcg gtaatctccg aacagaagga agaacgaagg aaggagcaca 300

gacttagatt ggtatatata cgcatatgta gtgttgaaga aacatgaaat tgcccagtat 360
tcttaaccca actgcacaga acaaaaacct gcaggaaacg aagataaatc atgtcgaaag 420
ctacatataa ggaacgtgct gctactcatc ctagtcctgt tgctgccaag ctatttaata 480
tcatgcacga aaagcaaaca aacttgtgtg cttcattgga tgttcgtacc accaaggaat 540
tactggagtt agttgaagca ttaggtccca aaatttgttt actaaaaaca catgtggata 600
tcttgactga tttttccatg gagggcacag ttaagccgct aaaggcatta tccgccaagt 660
acaatttttt actcttcgaa gacagaaaat ttgctgacat tggtaataca gtcaaattgc 720
agtactctgc gggtgtatac agaatagcag aatgggcaga cattacgaat gcacacggtg 780
tggtgggccc aggtattgtt agcggtttga agcaggcggc agaagaagta acaaaggaac 840
ctagaggcct tttgatgtta gcagaattgt catgcaaggg ctccctatct actggagaat 900
atactaaggg tactgttgac attgcgaaga gcgacaaaga ttttgttatc ggctttattg 960
ctcaaagaga catgggtgga agagatgaag gttacgattg gttgattatg acacccggtg 1020
tgggtttaga tgacaaggga gacgcattgg gtcaacagta tagaaccgtg gatgatgtgg 1080
tctctacagg atctgacatt attattgttg gaagaggact atttgcaaag ggaagggatg 1140
ctaaggtaga gggtgaacgt tacagaaaag caggctggga agcatatttg agaagatgcg 1200
gccagcaaaa ctaaaaaact gtattataag taaatgcatg tatactaaac tcacaaatta 1260
gagcttcaat ttaattatat cagttattac cctgcggtgt gaaataccgc acagatgcgt 1320
aaggagaaaa taccgcatca ggaaattgta aacgttaata ttttgttaaa attcgcgtta 1380
aatttttgtt aaatcagctc atttttaac caataggccg aaatcggcaa aatcccttat 1440
aaatcaaaag aatagaccga gatagggttg agtgttgttc cagtttggaa caagagtcca 1500
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 1560
ccactacgtg aaccatcacc ctaatcaagt tttttggggt cgaggtgccg taaagcacta 1620
aatcggaacc ctaaagggag cccccgattt agagcttgac ggggaaagcc ggcgaacgtg 1680
gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg 1740
gtcacgctgc gcgtaaccac cacacccgcc gcgcttaatg cgccgctaca gggcgcgtcg 1800
cgccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg 1860
ctattacgcc agctggcgaa gggggggatgt gctgcaaggc gattaagttg ggtaacgcca 1920
gggttttccc agtcacgacg ttgtaaaacg acggccagtg aattgtaata cgactcacta 1980
tagggcgaat tggagctcca ccgcggtggt ttaacgtata gacttctaat atatttctcc 2040
atacttggta ttttttattc actttttta tacatatttg gttttttat acaatatcaa 2100
aaatcaataa taattaattt aataacaaga atatgttgct gtgatgactc ttaatataca 2160
attgggaagc attcaaggat tggtacgacg tttgctttct gataattaaa cgcttggcgt 2220
ttaatctttt ttgccgatta ataatattat tattaccctc ttatttatta gcattgtctt 2280
ccgtactaaa aggagtaggg aaaaaggcga agtagagtga cacgtccatt acccagcgct 2340
tcataagaat ccatcaggta tattaaattg tgatgggaga acaaaacaaa ttgtactatg 2400
atgtcgagaa attggtgaat tctttgcagg aaagttttga cctggattgc gcgcaaagtg 2460

tatcactttt caccagtaaa tcaaggagca atgaggcttg gttggaagag ctagagaata 2520
aattcaagtt aaaagatgat gttgaacttg atgatgtgga aaatttaaga gccgaaattg 2580
acatgaagtt aaatatgttg gaggataaag taagctacta tgaaagactt tacaaagaac 2640
tcgaagagtt ccagaatgaa ataaaaatta aaacagttgt gaataacaga agacaatcgc 2700
gaactccaaa atgagctatc aaaaacgata gatcgattag gatgactttg aaatgactcc 2760
gcagtggact ggccgttaat ttcaagcgtg agtaaaatag tgcatgacaa aagatgagct 2820
aggcttttgt aaaaatatct tacgttgtaa aattttagaa atcattattt ccttcatatc 2880
attttgtcat tgaccttcag aagaaaagag ccgaccaata atataaataa ataaataaaa 2940
ataatattcc attatttcta aacagattca atactcatta aaaaactata tcaattaatt 3000
tgaattaacg cggccgctct agttaattaa tcattgtttg cctccctgct gcggtttttc 3060
accgaagttc atgccagtcc agcgtttttg cagcagaaaa gccgccgact tcggtttgcg 3120
gtcgcgagtg aagatccctt tcttgttacc gccaacgcgc aatatgcctt gcgaggtcgc 3180
aaaatcggcg aaattccata cctgttcacc gacgacggcg ctgacgcgat caaagacgcg 3240
gtgatacata tccagccatg cacactgata ctcttcactc cacatgtcgg tgtacattga 3300
gtgcagcccg gctaacgtat ccacgccgta ttcggtgatg ataatcggct gatgcagttt 3360
ctcctgccag gccagaagtt ctttttccag taccttctct gccgtttcca aatcgccgct 3420
ttggacatac catccgtaat aacggttcag gcacagcaca tcaaagagat cgctgatggt 3480
atcggtgtga gcgtcgcaga acattacatt gacgcaggtg atcggacgcg tcgggtcgag 3540
tttacgcgtt gcttccgcca gtggcgcgaa atattcccgt gcaccttgcg gacgggtatc 3600
cggttcgttg gcaatactcc acatcaccac gcttgggtgg tttttgtcac gcgctatcag 3660
ctctttaatc gcctgtaagt gcgcttgctg agtttccccg ttgactgcct cttcgctgta 3720
cagttctttc ggcttgttgc ccgcttcgaa accaatgcct aaagagaggt taaagccgac 3780
agcagcagtt tcatcaatca ccacgatgcc atgttcatct gcccagtcga gcatctcttc 3840
agcgtaaggg taatgcgagg tacggtagga gttggcccca atccagtcca ttaatgcgtg 3900
gtcgtgcacc atcagcacgt tatcgaatcc tttgccacgc aagtccgcat cttcatgacg 3960
accaaagcca gtaaagtaga acggtttgtg gttaatcagg aactgttcgc ccttcactgc 4020
cactgaccgg atgccgacgc gaagcgggta gatatcacac tctgtctggc ttttggctgt 4080
gacgcacagt tcatagagat aaccttcacc cggttgccag aggtgcggat tcaccacttg 4140
caaagtcccg ctagtgcctt gtccagttgc aaccacctgt tgatccgcat cacgcagttc 4200
aacgctgaca tcaccattgg ccaccacctg ccagtcaaca gacgcgtggt tacagtcttg 4260
cgcgacatgc gtcaccacgg tgatatcgtc cacccaggtg ttcggcgtgg tgtagagcat 4320
tacgctgcga tggattccgg catagttaaa gaaatcatgg aagtaagact gctttttctt 4380
gccgttttcg tcggtaatca ccattcccgg cgggatagtc tgccagttca gttcgttgtt 4440
cacacaaacg gtgatacgta cacttttccc ggcaataaca tacggcgtga catcggcttc 4500
aaatggagta tagccgccct gatgctccat cacttcctga ttattgaccc acactttgcc 4560
gtaatgagtg accgcatcga aacgcagcac gatacgctgg cctgcccaac ctttcggtat 4620

aaagacttcg cgctgatacc agacgttgcc cgcataatta cgaatatctg catcggcgaa 4680
ctgatcgtta aaactgcctg gcacagcaat tgcccggctt tcttgtaacg cgctttccca 4740
ccaacgctga tcaattccac agttttcgcg atccagactg aatgcccaca ggccgtcgag 4800
ttttttgatt tcacgggttg gggtttctac aggacgtacc atactagttt gaatatgtat 4860
tacttggtta tggttatata tgacaaaaga aaaagaagaa cagaagaata acgcaaggaa 4920
gaacaataac tgaaattgat agagaagtat tatgtctttg tctttttata ataaatcaag 4980
tgcagaaatc cgttagacaa catgagggat aaaatttaac gtgggcgaag aagaaggaaa 5040
aaagtttttg tgagggcgta attgaagcga tctgttgatt gtagattttt ttttttgag 5100
gagtcaaagt cagaagagaa cagacaaatg gtattaacca tccaatactt ttttggagca 5160
acgctaagct catgcttttc cattggttac gtgctcagtt gttagatatg gaaagagagg 5220
atgctcacgg cagcgtgact ccaattgagc ccgaaagaga ggatgccacg ttttcccgac 5280
ggctgctaga atggaaaaag gaaaaataga agaatcccat tcctatcatt atttacgtaa 5340
tgacccacac atttttgaga ttttcaacta ttacgtatta cgataatcct gctgtcatta 5400
tcattattat ctatatcgac gtatgcaacg tatgtgaagc caagtaggcg tcgacaggac 5460
cttgttgtgt gacgaaattg gaagctgcaa tcaataggaa gacaggaagt cgagcgtgtc 5520
tgggtttttt cagttttgtt ctttttgcaa acaaatcacg agcgacggta attggtaccc 5580
agcttttgtt ccctttagtg agggttaatt ccgagcttgg cgtaatcatg gtcatagctg 5640
tttcctgtgt gaaattgtta tccgctcaca attccacaca acataggagc cggaagcata 5700
aagtgtaaag cctggggtgc ctaatgagtg aggtaactca cattaattgc gttgcgctca 5760
ctgcccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc 5820
gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 5880
cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 5940
tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 6000
aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctcggccc ccctgacgag 6060
catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 6120
caggcgttcc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 6180
ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcaatg ctcacgctgt 6240
aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaacccccc 6300
gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 6360
cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 6420
ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 6480
tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 6540
tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 6600
cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 6660
tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 6720
tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 6780

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 6840
cgttcatcca tagttgcctg actgcccgtc gtgtagataa ctacgatacg ggagggctta 6900
ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 6960
tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 7020
gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 7080
agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt 7140
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg 7200
tgaaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 7260
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 7320
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg 7380
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 7440
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 7500
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt 7560
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 7620
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 7680
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 7740
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgggtcctt ttcatcacgt 7800
gctataaaaa taattataat ttaaattttt taatataaat atataaatta aaaatagaaa 7860
gtaaaaaaag aaattaaaga aaaaatagtt tttgttttcc gaagatgtaa aagactctag 7920
ggggatcgcc aacaaatact acctttttatc ttgctcttcc tgctctcagg tattaatgcc 7980
gaattgtttc atcttgtctg tgtagaagac cacacacgaa aatcctgtga ttttacattt 8040
tacttatcgt taatcgaatg tatatctatt taatctgctt ttcttgtcta ataaatatat 8100
atgtaaagta cgctttttgt tgaaattttt taaacctttg tttatttttt tttcttcatt 8160
ccgtaactct tctaccttct ttatttactt tctaaaatcc aaatacaaaa cataaaaata 8220
aataaacaca gagtaaattc ccaaattatt ccatcattaa aagatacgag gcgcgtgtaa 8280
gttacaggca agcgatccgt cctaagaaac cattattatc atgacattaa cctataaaaa 8340
taggcgtatc acgaggccct ttcgtc 8366

<210> 210
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> primer N160SeqF5

<400> 210
cctgaagtct aggtccctat tt 22

<210> 211
<211> 20
<212> DNA
<213> Artificial sequence

20 Jan 2017 2017200439

<220>
<223> N886 primer

<400> 211
caatgattgt tggtaaaggg 20

<210> 212
<211> 32
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP622

<400> 212
aattggtacc ccaaaaggaa tattgggtca ga 32

<210> 213
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP623

<400> 213
ccattgttta aacggcgcgc cggatccttt gcgaaaccct atgctctgt 49

<210> 214
<211> 49
<212> DNA
<213> Primer oBP624

<400> 214
gcaaaggatc cggcgcgccg tttaaacaat ggaaggtcgg gatgagcat 49

<210> 215
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP625

<400> 215
aattggccgg cctacgtaac attctgtcaa ccaa 34

<210> 216
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP626

<400> 216
aattgcggcc gcttcatata tgacgtaata aaat 34

<210> 217
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP627

<400> 217
aattttaatt aatttttttt cttggaatca gtac 34

<210> 218
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY21

<400> 218
ttaaggcgcg cctatttgta atacgtatac gaattccttc 40

<210> 219
<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY24

<400> 219
acttaataac tttaccggct gttgacattt tgttcttctt gttattgtat tgtgtt 56

<210> 220
<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY25

<400> 220
aacacaatac aataacaaga agaacaaaat gtcaacagcc ggtaaagtta ttaagt 56

<210> 221
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY4

<400> 221
ggaagtttaa acaccacagg tgttgtcctc tgaggacata 40

<210> 222
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> Primer URA3-end F

<400> 222
gcatatttga gaagatgcgg ccagcaaaac 30

<210> 223
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP636

<400> 223
catttttttc cctctaagaa gc 22

<210> 224
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP637

<400> 224
tttttgcaca gttaaactac cc 22

<210> 225
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP691

<400> 225
aattggatcc gcgatcgcga cgttctctcc gttgttcaaa 40

<210> 226
<211> 41
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP692

<400> 226
aattggcgcg ccatttaaat atatatgtat atatataaca c 41

<210> 227
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP693

<400> 227
aattgtttaa acaaaggatg atattgttct atta 34

<210> 228
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP694

<400> 228
aattggccgg ccgcaacgac gacaatgcca aac 33

<210> 229
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP695

<400> 229
aattgcggcc gcatgacagg tgaaagaatt gaaa 34

<210> 230
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP696

<400> 230
aattttaatt aaacgggcat cttatagtgt cgtt 34

<210> 231
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY16

<400> 231
ttaaggcgcg ccccgcacgc cgaaatgcat gcaagtaacc 40

<210> 232
<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY19

<400> 232
acttaataac tttaccggct gttgacattt tgattgattt gactgtgtta ttttgc 56

<210> 233
<211> 56
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY20

<400> 233
gcaaaataac acagtcaaat caatcaaaat gtcaacagcc ggtaaagtta ttaagt 56

<210> 234
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP730

<400> 234
ttgctccaaa gagatgtctt ta 22

<210> 235
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP731

20 Jan 2017 2017200439

<400> 235
tgttcccaca atctattacc ta 22

<210> 236
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY31

<400> 236
gccgacttta tggcgaagaa gtttgctctt gatc 34

<210> 237
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> N1179 primer

<400> 237
gcgaattctg atctgatgcg ctttgcatat ctcatatt 38

<210> 238
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> N1180 primer

<400> 238
gcggtaccgc cgtttcggct tggatacgcc atatgcaa 38

<210> 239
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> N1181 primer

<400> 239
gccgatcgtg taccaacctg catttctttc cgtcatat 38

<210> 240
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> N1182 primer

<400> 240
gcgcatgcgt gtacctttg attggaattg gttcgcag 38

<210> 241
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1212 primer

<400> 241
ttctatgtta aattttaacg atgtagacac 30

<210> 242
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1213 primer

<400> 242
atggcgaaat ggcagtactc gggggaatgc 30

<210> 243
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1239

<400> 243
gacgaggata atgtgcatga acggg 25

<210> 244
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1242

<400> 244
gtttcgttgc gttggtggaa tggc 24

<210> 245
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1249

<400> 245
gctttatgga ccctgaaacc acagccac 28

<210> 246
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1250

<400> 246
ggaaagaaaa agcttcatgg ccttgcg 27

<210> 247
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1251

<400> 247
ccaatgtggc tgtggtttca ggg 23

<210> 248
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1252

<400> 248
gcaaggccat gaagcttttt ctttcc 26

<210> 249
<211> 62
<212> DNA
<213> Primer N1253

<400> 249
ggaaagaaaa agcttcatgg ccttgcgctc atcaacacta aaattagagt cattctaatt 60
gc 62

<210> 250
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1254

<400> 250
ttatccacgg aagatatgat gaggtgacg 29

<210> 251
<211> 64
<212> DNA
<213> Primer N1255

<400> 251
ggaccctgaa accacagcca cattggtacc ttccttgcca aagggattaa acctgacaat 60
tacc 64

<210> 252
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1256

<400> 252
acctagctaa actaagtaaa tctgtatac 29

<210> 253
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1257

<400> 253
gatgatgcta tttggtgcag agggtgatg 29

<210> 254
<211> 60
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1258

<400> 254
cagatttact tagtttagct aggtgtatca aaatacgttc tcaatgttct atttcccgcc 60

<210> 255
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1262

<400> 255
cacgtaaggg catgatagaa ttgg 24

<210> 256
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1263

<400> 256
ggatatagca gttgttgtac actagc 26

<210> 257
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> primer N1223

<400> 257
caaggaatta ccatcaccgt cacc 24

<210> 258
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> primer N1225

<400> 258
cttgtgtata atgataaatt ggttggg 27

<210> 259
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1228

<400> 259
caacggttca tcatctcatg gatctgc 27

<210> 260
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1229

<400> 260
gttacttggt tctggcgagg tattgg 26

<210> 261
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1230

<400> 261
gtaagaaggt tggtattcca gctgg 25

<210> 262
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N1231

<400> 262
ctgtttcgat accagaacct agaccg 26

<210> 263
<211> 1017
<212> DNA
<213> Artificial sequence

<220>
<223> K9_David_SH

<400> 263
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaatgcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 264
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Primer pOLE1-NgoMI

<400> 264
cgtggccggc cgtgcaaaaa ctaactccga g 31

<210> 265
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer pOLE1-PmlI

<400> 265
gtactcacgt gtttgttgta atgttttagt gctg 34

<210> 266
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> K9_Eliza_SH

<400> 266

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Asp Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

20 Jan 2017 2017200439

```
His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 267
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> K9_Frank_SH

<400> 267

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                  10                  15
```

```
Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Leu Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300
```

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 268
<211> 4236
<212> DNA
<213> artificial sequence

<220>
<223> Synthetic construct

<400> 268
gatccgcatt gcggattacg tattctaatg ttcagataac ttcgtatagc atacattata 60

cgaagttatg cagattgtac tgagagtgca ccataccaca gcttttcaat tcaattcatc 120

attttttttt tattcttttt tttgatttcg gtttctttga aatttttttg attcggtaat 180

ctccgaacag aaggaagaac gaaggaagga gcacagactt agattggtat atatacgcat 240

atgtagtgtt gaagaaacat gaaattgccc agtattctta acccaactgc acagaacaaa 300

aacctgcagg aaacgaagat aaatcatgtc gaaagctaca tataaggaac gtgctgctac 360

tcatcctagt cctgttgctg ccaagctatt taatatcatg cacgaaaagc aaacaaactt 420

gtgtgcttca ttggatgttc gtaccaccaa ggaattactg gagttagttg aagcattagg 480

tcccaaaatt tgtttactaa aaacacatgt ggatatcttg actgattttt ccatggaggg 540

cacagttaag ccgctaaagg cattatccgc caagtacaat tttttactct tcgaagacag 600

aaaatttgct gacattggta atacagtcaa attgcagtac tctgcgggtg tatacagaat 660

agcagaatgg gcagacatta cgaatgcaca cggtgtggtg ggcccaggta ttgttagcgg 720

tttgaagcag gcggcagaag aagtaacaaa ggaacctaga ggccttttga tgttagcaga 780

attgtcatgc aagggctccc tatctactgg agaatatact aagggtactg ttgacattgc 840

gaagagcgac aaagattttg ttatcggctt tattgctcaa agagacatgg gtggaagaga 900

tgaaggttac gattggttga ttatgacacc cggtgtgggt ttagatgaca agggagacgc 960

attgggtcaa cagtatagaa ccgtggatga tgtggtctct acaggatctg acattattat 1020

tgttggaaga ggactatttg caaagggaag ggatgctaag gtagagggtg aacgttacag 1080

aaaagcaggc tgggaagcat atttgagaag atgcggccag caaaactaaa aaactgtatt 1140

ataagtaaat gcatgtatac taaactcaca aattagagct tcaatttaat tatatcagtt 1200

attaccctat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcatcagga 1260

aattgtaaac gttaatattt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt 1320

ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat 1380

agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa 1440

cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccta 1500

atcaagataa cttcgtatag catacattat acgaagttat ccagtgatga tacaacgagt 1560
tagccaaggt gaattcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg 1620
ttacccaact taatcgcctt gcagcacatc ccccttttcgc cagctggcgt aatagcgaag 1680
aggcccgcac cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa tggcgcctga 1740
tgcggtattt tctccttacg catctgtgcg gtatttcaca ccgcatatgg tgcactctca 1800
gtacaatctg ctctgatgcc gcatagttaa gccagccccg acacccgcca acacccgctg 1860
acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct gtgaccgtct 1920
ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg agacgaaagg 1980
gcctcgtgat acgcctattt ttataggtta atgtcatgat aataatggtt tcttagacgt 2040
caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac 2100
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 2160
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 2220
tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc 2280
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 2340
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 2400
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 2460
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 2520
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc 2580
tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg 2640
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 2700
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 2760
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 2820
cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 2880
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 2940
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 3000
agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 3060
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atccttttg 3120
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 3180
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 3240
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc 3300
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc cttctagtgt 3360
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 3420
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 3480
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac 3540
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 3600
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg 3660

gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 3720

tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca gggggggcgga 3780

gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 3840

ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct 3900

ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 3960

aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt 4020

aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 4080

atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 4140

tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 4200

acgccaagct tgcatgcctg caggtcgact ctagag 4236

<210> 269
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> M13 forward primer

<400> 269
gtaaaacgac ggccagt 17

<210> 270
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> M13 reverse primer

<400> 270
cacacaggaa acagctatga cc 22

<210> 271
<211> 7523
<212> DNA
<213> Artificial sequence

<220>
<223> Construct

<400> 271
ccagcttttg ttccctttag tgagggttaa ttgcgcgctt ggcgtaatca tggtcatagc 60

tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatagga gccggaagca 120

taaagtgtaa agcctggggt gcctaatgag tgaggtaact cacattaatt gcgttgcgct 180

cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga atcggccaac 240

gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc actgactcgc 300

tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt 360

tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg 420

ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc cccctgacg 480

agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat 540

accaggcgtt tccccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta 600
ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcat agctcacgct 660
gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc 720
ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa 780
gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg 840
taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag 900
tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt 960
gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta 1020
cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc 1080
agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca 1140
cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa 1200
cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat 1260
ttcgttcatc catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct 1320
taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt 1380
tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat 1440
ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta 1500
atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg 1560
gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt 1620
tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg 1680
cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg 1740
taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc 1800
ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa 1860
ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac 1920
cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt 1980
ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg 2040
gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa tattattgaa 2100
gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata 2160
aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgaacga agcatctgtg 2220
cttcattttg tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg 2280
agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tattttacca acgaagaatc 2340
tgtgcttcat ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaaga 2400
atctgagctg catttttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa 2460
gaatctatac ttctttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca 2520
aagcatctta gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac 2580
tttttgcact gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt 2640
ccataaaaaa agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg 2700

```
catttttca agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac     2760
tttgtgaaca gaaagtgata gcgttgatga ttcttcattg gtcagaaaat tatgaacggt     2820
ttcttctatt ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt     2880
cgattcactc tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga     2940
taaacataaa aaatgtagag gtcgagttta gatgcaagtt caaggagcga aaggtggatg     3000
ggtaggttat atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg     3060
tttgtggaag cggtattcgc aatattttag tagctcgtta cagtccggtg cgtttttggt     3120
tttttgaaag tgcgtcttca gagcgctttt ggttttcaaa agcgctctga agttcctata     3180
ctttctagag aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct     3240
tccgaaaatg caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct     3300
gcgtgttgcc tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa     3360
atgcgtactt atatgcgtct atttatgtag gatgaaaggt agtctagtac ctcctgtgat     3420
attatcccat tccatgcggg gtatcgtatg cttccttcag cactaccctt tagctgttct     3480
atatgctgcc actcctcaat tggattagtc tcatccttca atgctatcat ttcctttgat     3540
attggatcat ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca     3600
cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc     3660
tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg     3720
gcgcgtcagc gggtgttggc gggtgtcggg gctggcttaa ctatgcggca tcagagcaga     3780
ttgtactgag agtgcaccat aaattcccgt tttaagagct tggtgagcgc taggagtcac     3840
tgccaggtat cgtttgaaca cggcattagt cagggaagtc ataacacagt cctttcccgc     3900
aattttcttt ttctattact cttggcctcc tctagtacac tctatatttt tttatgcctc     3960
ggtaatgatt ttcatttttt tttttcccct agcggatgac tctttttttt tcttagcgat     4020
tggcattatc acataatgaa ttatacatta tataaagtaa tgtgatttct tcgaagaata     4080
tactaaaaaa tgagcaggca agataaacga aggcaaagat gacagagcag aaagccctag     4140
taaagcgtat tacaaatgaa accaagattc agattgcgat ctctttaaag ggtggtcccc     4200
tagcgataga gcactcgatc ttcccagaaa aagaggcaga agcagtagca gaacaggcca     4260
cacaatcgca agtgattaac gtccacacag gtatagggtt tctggaccat atgatacatg     4320
ctctggccaa gcattccggc tggtcgctaa tcgttgagtg cattggtgac ttacacatag     4380
acgaccatca caccactgaa gactgcggga ttgctctcgg tcaagctttt aaagaggccc     4440
tactggcgcg tggagtaaaa aggtttggat caggatttgc gcctttggat gaggcacttt     4500
ccagagcggt ggtagatctt tcgaacaggc cgtacgcagt tgtcgaactt ggtttgcaaa     4560
gggagaaagt aggagatctc tcttgcgaga tgatcccgca ttttcttgaa agctttgcag     4620
aggctagcag aattaccctc cacgttgatt gtctgcgagg caagaatgat catcaccgta     4680
gtgagagtgc gttcaaggct cttgcggttg ccataagaga agccacctcg cccaatggta     4740
ccaacgatgt tccctccacc aaaggtgttc ttatgtagtg acaccgatta tttaaagctg     4800
cagcatacga tatatataca tgtgtatata tgtataccta tgaatgtcag taagtatgta     4860
```

20 Jan 2017 2017200439

tacgaacagt atgatactga agatgacaag gtaatgcatc attctatacg tgtcattctg 4920
aacgaggcgc gctttccttt tttctttttg ctttttcttt ttttttctct tgaactcgac 4980
ggatctatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggaaa 5040
ttgtaaacgt taatattttg ttaaaattcg cgttaaattt ttgttaaatc agctcatttt 5100
ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag accgagatag 5160
ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg gactccaacg 5220
tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgaacca tcaccctaat 5280
caagtttttt ggggtcgagg tgccgtaaag cactaaatcg gaaccctaaa gggagccccc 5340
gatttagagc ttgacgggga aagccggcga acgtggcgag aaaggaaggg aagaaagcga 5400
aaggagcggg cgctagggcg ctggcaagtg tagcggtcac gctgcgcgta accaccacac 5460
ccgccgcgct taatgcgccg ctacagggcg cgtcgcgcca ttcgccattc aggctgcgca 5520
actgttggga agggcgatcg gtgcgggcct cttcgctatt acgccagctg gcgaaagggg 5580
gatgtgctgc aaggcgatta agttgggtaa cgccagggtt ttcccagtca cgacgttgta 5640
aaacgacggc cagtgagcgc gcgtaatacg actcactata gggcgaattg ggtaccgggc 5700
ccccctcga ggtattagaa gccgccgagc gggcgacagc cctccgacgg aagactctcc 5760
tccgtgcgtc ctcgtcttca ccggtcgcgt tcctgaaacg cagatgtgcc tcgcgccgca 5820
ctgctccgaa caataaagat tctacaatac tagcttttat ggttatgaag aggaaaaatt 5880
ggcagtaacc tggccccaca aaccttcaaa ttaacgaatc aaattaacaa ccataggatg 5940
ataatgcgat tagtttttta gccttatttc tggggtaatt aatcagcgaa gcgatgattt 6000
ttgatctatt aacagatata taaatggaaa agctgcataa ccactttaac taatactttc 6060
aacattttca gtttgtatta cttcttattc aaatgtcata aaagtatcaa caaaaaattg 6120
ttaatatacc tctatacttt aacgtcaagg agaaaaatgt ccaatttact gcccgtacac 6180
caaaatttgc ctgcattacc ggtcgatgca acgagtgatg aggttcgcaa gaacctgatg 6240
gacatgttca gggatcgcca ggcgttttct gagcatacct ggaaaatgct tctgtccgtt 6300
tgccggtcgt gggcggcatg gtgcaagttg aataaccgga aatggtttcc cgcagaacct 6360
gaagatgttc gcgattatct tctatatctt caggcgcgcg gtctggcagt aaaaactatc 6420
cagcaacatt tgggccagct aaacatgctt catcgtcggt ccgggctgcc acgaccaagt 6480
gacagcaatg ctgtttcact ggttatgcgg cggatccgaa aagaaaacgt tgatgccggt 6540
gaacgtgcaa aacaggctct agcgttcgaa cgcactgatt tcgaccaggt tcgttcactc 6600
atggaaaata gcgatcgctg ccaggatata cgtaatctgg catttctggg gattgcttat 6660
aacaccctgt tacgtatagc cgaaattgcc aggatcaggg ttaaagatat ctcacgtact 6720
gacggtggga gaatgttaat ccatattggc agaacgaaaa cgctggttag caccgcaggt 6780
gtagagaagg cacttagcct gggggtaact aaactggtcg agcgatggat ttccgtctct 6840
ggtgtagctg atgatccgaa taactacctg ttttgccggg tcagaaaaaa tggtgttgcc 6900
gcgccatctg ccaccagcca gctatcaact cgcgccctgg aagggatttt tgaagcaact 6960
catcgattga tttacggcgc taaggatgac tctggtcaga gatacctggc ctggtctgga 7020

cacagtgccc gtgtcggagc cgcgcgagat atggcccgcg ctggagtttc aataccggag 7080

atcatgcaag ctggtggctg gaccaatgta aatattgtca tgaactatat ccgtaacctg 7140

gatagtgaaa caggggcaat ggtgcgcctg ctggaagatg gcgattagga gtaagcgaat 7200

ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag tgtatacaaa 7260

ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa ctctttcctg 7320

taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca cacctctacc 7380

ggcatgccga gcaaatgcct gcaaatcgct ccccatttca cccaattgta gatatgctaa 7440

ctccagcaat gagttgatga atctcggtgt gtattttatg tcctcagagg acaacacctg 7500

tggtccgcca ccgcggtgga gct 7523

<210> 272
<211> 47
<212> DNA
<213> Artificial sequence

<220>
<223> upper primer for pYZ058 construction

<400> 272
catcatcaca gtttaaacag tatgttgaag caaatcaact tcggtgg 47

<210> 273
<211> 48
<212> DNA
<213> Artificial sequence

<220>
<223> lower primer for pYZ058 construction

<400> 273
ggacgggccc tgcaggcctt attggttttc tggtctcaac tttctgac 48

<210> 274
<211> 1005
<212> DNA
<213> Clostridium beijerinckii

<400> 274
atgccaaaga tgtactacga aaaggacact gatttgaact tgttgagagg taagaaggtc 60

gccattatcg gttacggttc ccaaggtcac gctcatgcct tgaacttgca cgaatctggt 120

gtcgacgttg tcgttggttt gtacaacggt tctaagtcct gggctaaggc tgaagctgct 180

ggtctgcaag tcgccaccgt tgctgacgct gccaaggctg ccgatttgat catgattttg 240

ctgccagatg aaaagcaagc taaaatctac aacgaagaaa ttgctccaaa cttggaagaa 300

ggtaatgcct tggttttcgc tcacggtttc aacatccatt tcggtcaagt cgttccacca 360

tcctacgtcg acgttttcat ggtcgctcca aagggtccag gtcacttggt tagacgtacc 420

tacactgaag gtgctggtgt cccatgtttg atcgctgttc atcaagatgc cactggtaaa 480

gctaaacagt acgctttggc ctacgctaac ggtattggtg gtgccagagc tggtgtcttg 540

gaaaccacct tcaaggacga aaccgaaacc gacttgttcg gtgaacaagc tgtcttatgt 600

ggtggtgttt ctgaattgat caaggctggt tttgaaacct tggtcgaagc cggttacgct 660

ccagaaaacg cttacttcga atgcatgcac gaaatgaagt tgattgttga cttgttgtac 720

caaggtggtt tatccatgat gagatactct atctccgaca ctgctgaata cggtgactac 780

cagattggta gaagaatcat taccgacgaa accaagaagg aaatgaagaa ggttttgacc 840

gaaatccaag atggtacttt cgctaagaac tggctgttgg aaaaccagac caatagacca 900

ggtttcaacg ccagaagaag aatggaagct gaacacccaa tcgaaaaggt tggtaaggaa 960

ttgagaggta tgatgtcctg gatcgacacc gctaaggttg attaa 1005

<210> 275
<211> 334
<212> PRT
<213> Clostridium beijerinckii

<400> 275

Met Pro Lys Met Tyr Tyr Glu Lys Asp Thr Asp Leu Asn Leu Leu Arg
1 5 10 15

Gly Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Leu His Glu Ser Gly Val Asp Val Val Val Gly Leu Tyr
35 40 45

Asn Gly Ser Lys Ser Trp Ala Lys Ala Glu Ala Ala Gly Leu Gln Val
50 55 60

Ala Thr Val Ala Asp Ala Ala Lys Ala Ala Asp Leu Ile Met Ile Leu
65 70 75 80

Leu Pro Asp Glu Lys Gln Ala Lys Ile Tyr Asn Glu Glu Ile Ala Pro
85 90 95

Asn Leu Glu Glu Gly Asn Ala Leu Val Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Gln Val Val Pro Pro Ser Tyr Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Ala Gly Val Pro Cys Leu Ile Ala Val His Gln Asp Ala Thr Gly Lys
145 150 155 160

Ala Lys Gln Tyr Ala Leu Ala Tyr Ala Asn Gly Ile Gly Gly Ala Arg
165 170 175

Ala Gly Val Leu Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Ser Glu Leu Ile Lys
195 200 205

```
Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Met His Glu Met Lys Leu Ile Val Asp Leu Leu Tyr
225                 230                 235                 240

Gln Gly Gly Leu Ser Met Met Arg Tyr Ser Ile Ser Asp Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Gln Ile Gly Arg Arg Ile Ile Thr Asp Glu Thr Lys
            260                 265                 270

Lys Glu Met Lys Lys Val Leu Thr Glu Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asn Trp Leu Leu Glu Asn Gln Thr Asn Arg Pro Gly Phe Asn Ala
    290                 295                 300

Arg Arg Arg Met Glu Ala Glu His Pro Ile Glu Lys Val Gly Lys Glu
305                 310                 315                 320

Leu Arg Gly Met Met Ser Trp Ile Asp Thr Ala Lys Val Asp
                325                 330
```

<210> 276
<211> 1020
<212> DNA
<213> Zymomonas mobilis

<400> 276

```
atgaaagttt actacgattc tgacgccgac ctgggcctca tcaaatctaa gaaaattgcc        60

atcctcggct acggttccca gggccacgct catgcacaga acctgcgcga ttccggcgta       120

gcggaagtgg ctattgccct gcgtccagac tcggcttccg ttaagaaggc tcaagacgcc       180

ggcttcaaag tgctcaccaa tgctgaagcc gctaagtggg ccgacatcct catgatcctg       240

gccccggacg agcaccaggc cgcgatctac gcggaggacc tcaaggacaa cctccgcccg       300

ggctccgcta tcgccttcgc ccacggcctc aacatccact tcggcctgat tgagcctcgt       360

aaggacatcg acgtgttcat gatcgcgccg aagggtccag gccacaccgt acgttccgag       420

tacgtccgcg gcggtggtgt tccttgcctg gtcgcggtcg accaggacgc ttccggcaac       480

gcccacgaca ttgccctggc ctacgcctca ggcatcggcg gcggccgttc cggcgtcatc       540

gaaaccacct tccgcgagga ggttgaaacc gacctgttcg gtgagcaggc tgttctgtgt       600

ggcggtctca ccgcgctgat cacagccggt ttcgaaaccc tgactgaagc tggctacgct       660

ccagaaatgg ccttcttcga gtgcatgcac gaaatgaagc tgatcgttga cctcatctac       720

gaagccggta tcgccaacat gcgctactcg atctccaaca ccgctgaata cggcgacatc       780

gtgtctggtc cacgcgtcat caacgaggaa tccaaaaagg ccatgaaggc tattctggac       840

gacatccaga gcggccgctt cgtgtcgaag ttcgtcctcg acaaccgcgc cggccagcca       900

gagttgaagg ccgcgcgtaa acgcatggct gctcacccta tcgaacaggt cggcgcgcgc       960

ctgcgcaaga tgatgccttg gatcgcttcc aacaaactcg tcgacaaagc ccgcaactaa      1020
```

<210> 277
<211> 339
<212> PRT
<213> Zymomonas mobilis

<400> 277

```
Met Lys Val Tyr Tyr Asp Ser Asp Ala Asp Leu Gly Leu Ile Lys Ser
1               5                  10                  15

Lys Lys Ile Ala Ile Leu Gly Tyr Gly Ser Gln Gly His Ala His Ala
            20                  25                  30

Gln Asn Leu Arg Asp Ser Gly Val Ala Glu Val Ala Ile Ala Leu Arg
        35                  40                  45

Pro Asp Ser Ala Ser Val Lys Lys Ala Gln Asp Ala Gly Phe Lys Val
    50                  55                  60

Leu Thr Asn Ala Glu Ala Ala Lys Trp Ala Asp Ile Leu Met Ile Leu
65                  70                  75                  80

Ala Pro Asp Glu His Gln Ala Ala Ile Tyr Ala Glu Asp Leu Lys Asp
                85                  90                  95

Asn Leu Arg Pro Gly Ser Ala Ile Ala Phe Ala His Gly Leu Asn Ile
            100                 105                 110

His Phe Gly Leu Ile Glu Pro Arg Lys Asp Ile Asp Val Phe Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Val Arg Gly
    130                 135                 140

Gly Gly Val Pro Cys Leu Val Ala Val Asp Gln Asp Ala Ser Gly Asn
145                 150                 155                 160

Ala His Asp Ile Ala Leu Ala Tyr Ala Ser Gly Ile Gly Gly Gly Arg
                165                 170                 175

Ser Gly Val Ile Glu Thr Thr Phe Arg Glu Glu Val Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Thr
        195                 200                 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Pro Glu Met Ala
    210                 215                 220

Phe Phe Glu Cys Met His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Ala Gly Ile Ala Asn Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255
```

```
Tyr Gly Asp Ile Val Ser Gly Pro Arg Val Ile Asn Glu Glu Ser Lys
            260                 265                 270

Lys Ala Met Lys Ala Ile Leu Asp Asp Ile Gln Ser Gly Arg Phe Val
        275                 280                 285

Ser Lys Phe Val Leu Asp Asn Arg Ala Gly Gln Pro Glu Leu Lys Ala
    290                 295                 300

Ala Arg Lys Arg Met Ala Ala His Pro Ile Glu Gln Val Gly Ala Arg
305                 310                 315                 320

Leu Arg Lys Met Met Pro Trp Ile Ala Ser Asn Lys Leu Val Asp Lys
                325                 330                 335

Ala Arg Asn
```

<210> 278
<211> 267
<212> PRT
<213> Saccharomyces cerevisiae

<400> 278

```
Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Lys Lys Thr Val
1               5                   10                  15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
            20                  25                  30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Ile Leu Ala Ala Arg
        35                  40                  45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Gln Glu Phe
    50                  55                  60

Pro Asn Ala Lys Val His Val Ala Gln Leu Asp Ile Thr Gln Ala Glu
65                  70                  75                  80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Gln Glu Phe Lys Asp Ile
                85                  90                  95

Asp Ile Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Arg Val
            100                 105                 110

Gly Gln Ile Ala Thr Glu Asp Ile Gln Asp Val Phe Asp Thr Asn Val
        115                 120                 125

Thr Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Ile Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly
```

165 170 175

Ala Phe Thr Asp Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg
180 185 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val
195 200 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr
210 215 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225 230 235 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
245 250 255

Asn Gln Ala Ser Pro His His Ile Phe Arg Gly
260 265

<210> 279
<211> 500
<212> PRT
<213> Saccharomyces cerevisiae

<400> 279

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu
1 5 10 15

Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn
20 25 30

Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro
35 40 45

Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp
50 55 60

Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu
65 70 75 80

Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu
85 90 95

Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala
100 105 110

Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
115 120 125

Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn
130 135 140

Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr Met Asn Phe Thr Thr Leu
145 150 155 160

```
Glu Pro Ile Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Ile
                165                 170                 175

Met Met Leu Ala Trp Lys Ile Ala Pro Ala Leu Ala Met Gly Asn Val
            180                 185                 190

Cys Ile Leu Lys Pro Ala Ala Val Thr Pro Leu Asn Ala Leu Tyr Phe
        195                 200                 205

Ala Ser Leu Cys Lys Lys Val Gly Ile Pro Ala Gly Val Val Asn Ile
    210                 215                 220

Val Pro Gly Pro Gly Arg Thr Val Gly Ala Ala Leu Thr Asn Asp Pro
225                 230                 235                 240

Arg Ile Arg Lys Leu Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Ser
                245                 250                 255

Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu
            260                 265                 270

Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys
        275                 280                 285

Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln
    290                 295                 300

Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp
305                 310                 315                 320

Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val
                325                 330                 335

Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
            340                 345                 350

Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu
        355                 360                 365

Gly Ala Lys Ile Leu Thr Gly Gly Glu Lys Val Gly Asp Lys Gly Tyr
    370                 375                 380

Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile
385                 390                 395                 400

Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys
                405                 410                 415

Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu
            420                 425                 430

Gly Ser Gly Ile Glu Thr Glu Ser Leu Ser Thr Gly Leu Lys Val Ala
        435                 440                 445
```

Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe
450 455 460

Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
465 470 475 480

Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
485 490 495

Arg Ile Lys Leu
500

<210> 280
<211> 804
<212> DNA
<213> Saccharomyces cerevisiae

<400> 280
atgtcccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt 60

gcatctgctg gtattggtaa ggcgaccgca ttagagtact tggaggcatc caatggtgat 120

atgaaactga tcttggctgc tagaagatta gaaaagctcg aggaattgaa gaagaccatt 180

gatcaagagt ttccaaacgc aaaagttcat gtggcccagc tggatatcac tcaagcagaa 240

aaaatcaagc ccttcattga aaacttgcca caagagttca aggatattga cattctggtg 300

aacaatgccg gaaaggctct tggcagtgac cgtgtgggcc agatcgcaac ggaggatatc 360

caggacgtgt ttgacaccaa cgtcacggct ttaatcaata tcacacaagc tgtactgccc 420

atattccaag ccaagaattc aggagatatt gtaaatttgg gttcaatcgc tggcagagac 480

gcatacccaa caggttctat ctattgtgcc tctaagtttg ccgtgggggc gttcactgat 540

agtttgagaa aggagctcat caacactaaa attagagtca ttctaattgc accagggcta 600

gtcgagactg aattttcact agttagatac agaggtaacg aggaacaagc caagaatgtt 660

tacaaggata ctaccccatt gatggctgat gacgtggctg atctgatcgt ctatgcaact 720

tccagaaaac aaaatactgt aattgcagac actttaatct ttccaacaaa ccaagcgtca 780

cctcatcata tcttccgtgg ataa 804

<210> 281
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1214 primer

<400> 281
aaaaaggggg tagaagcgtt ttgaagctat 30

<210> 282
<211> 62
<212> DNA
<213> Artificial sequence

<220>
<223> N1176 primer

<400> 282
gcatagcaat ctaatctaag ttccagctga ggatgacaac agattactca tcaccagcat 60

at 62

<210> 283
<211> 62
<212> DNA
<213> Artificial sequence

<220>
<223> N1117 primer

<400> 283
atcaacacac aaacactaaa tcaaagctga ggatggattt atttgagtca ttagcacaaa 60

aa 62

<210> 284
<211> 800
<212> DNA
<213> Saccharomyces cerevisiae

<400> 284
ccactagtca gatgccgcgg gcacttgagc acctcatgca cagcaataac acaacacaat 60

ggttagtagc aacctgaatt cggtcattga tgcatgcatg tgccgtgaag cgggacaacc 120

agaaaagtcg tctataaatg ccggcacgtg cgatcatcgt ggcggggttt taagagtgca 180

tatcacaaat tgtcgcatta ccgcggaacc gccagatatt cattacttga cgcaaaagcg 240

tttgaaataa tgacgaaaaa gaaggaagaa aaaaaaagaa aaataccgct tctaggcggg 300

ttatctactg atccgagctt ccactaggat agcacccaaa cacctgcata tttggacgac 360

ctttacttac accaccaaaa accactttcg cctctcccgc ccctgataac gtccactaat 420

tgagcgatta cctgagcggt cctcttttgt ttgcagcatg agacttgcat actgcaaatc 480

gtaagtagca acgtctcaag gtcaaaactg tatggaaacc ttgtcacctc acttaattct 540

agctagccta ccctgcaagt caagaggtct ccgtgattcc tagccacctc aaggtatgcc 600

tctccccgga aactgtggcc ttttctggca cacatgatct ccacgatttc aacatataaa 660

tagcttttga taatggcaat attaatcaaa tttattttac ttctttcttg taacatctct 720

cttgtaatcc cttattcctt ctagctattt ttcataaaaa accaagcaac tgcttatcaa 780

cacacaaaca ctaaatcaaa 800

<210> 285
<211> 44
<212> DNA
<213> Artificial sequence

<220>
<223> N822 primer

<400> 285
cgcctcagct ttgatttagt gtttgtgtgt tgataagcag ttgc 44

<210> 286
<211> 63
<212> DNA
<213> Artificial sequence

<220>
<223> N1178 primer

<400> 286
ggtatcgata agcttgatat cgaattcctg cgcccgggcc actagtcaga tgccgcgggc 60

act 63

<210> 287
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> N821 primer

<400> 287
cgcccgggcc actagtcaga tgccgcgggc acttgagc 38

<210> 288
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1115 primer

<400> 288
tttttgtgct aatgactcaa ataaatccat 30

<210> 289
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1110 primer

<400> 289
gcgatttaat ctctaattat tagttaaagt 30

<210> 290
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> N1114 primer

<400> 290
atatgctggt gatgagtaat ctgttgtcat 30

<210> 291
<211> 9220
<212> DNA
<213> Artificial sequence

<220>
<223> pUC19-URA3::pdc1::TEF(M4)-xpk1::kan Plasmid

<400> 291
ccgcattgcg gattacgtat tctaatgttc agataacttc gtataatgta tgctatacga 60

agttatcgaa cagagaaact aaatccacat taattgagag ttctatctat tagaaaatgc 120

aaactccaac taaatgggaa aacagataac ctcttttatt ttttttaat gtttgatatt 180

cgagtctttt tcttttgtta ggtttatatt catcatttca atgaataaaa gaagcttctt 240
attttggttg caaagaatga aaaaaaagga ttttttcata cttctaaagc ttcaattata 300
accaaaaatt ttataaatga agagaaaaaa tctagtagta tcaagttaaa cttagaaaaa 360
ctcatcgagc atcaaatgaa actgcaattt attcatatca ggattatcaa taccatattt 420
ttgaaaaagc cgtttctgta atgaaggaga aaactcaccg aggcagttcc ataggatggc 480
aagatcctgg tatcggtctg cgattccgac tcgtccaaca tcaatacaac ctattaattt 540
cccctcgtca aaaataaggt tatcaagtga gaaatcacca tgagtgacga ctgaatccgg 600
tgagaatggc aaaagcttat gcatttcttt ccagacttgt tcaacaggcc agccattacg 660
ctcgtcatca aaatcactcg catcaaccaa accgttattc attcgtgatt gcgcctgagc 720
gagacgaaat acgcgatcgc tgttaaaagg acaattacaa acaggaatcg aatgcaaccg 780
gcgcaggaac actgccagcg catcaacaat attttcacct gaatcaggat attcttctaa 840
tacctggaat gctgttttgc cggggatcgc agtggtgagt aaccatgcat catcaggagt 900
acggataaaa tgcttgatgg tcggaagagg cataaattcc gtcagccagt ttagtctgac 960
catctcatct gtaacatcat tggcaacgct acctttgcca tgtttcagaa acaactctgg 1020
cgcatcgggc ttcccataca atcgatagat tgtcgcacct gattgcccga cattatcgcg 1080
agcccattta tacccatata aatcagcatc catgttggaa tttaatcgcg gcctcgaaac 1140
gtgagtcttt tccttaccca tctcgagttt taatgttact tctcttgcag ttagggaact 1200
ataatgtaac tcaaaataag attaaacaaa ctaaaataaa aagaagttat acagaaaaac 1260
ccatataaac cagtactaat ccataataat aatacacaaa aaaactatca aataaaacca 1320
gaaaacagat tgaatagaaa aatttttcg atctcctttt atattcaaaa ttcgatatat 1380
gaaaaaggga actctcagaa aatcaccaaa tcaatttaat tagatttttc ttttccttct 1440
agcgttggaa agaaaaattt ttctttttt ttttagaaat gaaaaatttt tgccgtagga 1500
atcaccgtat aaaccctgta taaacgctac tctgttcacc tgtgtaggct atgattgacc 1560
cagtgttcat tgttattgcg agagagcggg agaaaagaac cgatacaaga gatccatgct 1620
ggtatagttg tctgtccaac actttgatga acttgtagga cgatgatgtg tatttagacg 1680
agtacgtgtg tgactattaa gtagttatga tagagaggtt tgtacggtgt gttctgtgta 1740
attcgattga gaaaatggtt atgaatccct agataacttc gtataatgta tgctatacga 1800
agttatccag tgatgataca acgagttagc caaggtgggg gatcctctag agtcttaagg 1860
ccgcccgcaa attaaagcct tcgagcgtcc caaaaccttc tcaagcaagg ttttcagtat 1920
aatgttacat gcgtacacgc gtttgtacag aaaaaaaaga aaaatttgaa atataaataa 1980
cgttcttaat actaacataa ctattaaaaa aaataaatag ggacctagac ttcaggttgt 2040
ctaactcctt ccttttcggt tagagcggat gtgggaggag ggcgtgaatg taagcgtgac 2100
ataactaatt acatgattaa ttaattattt taaacccttc cattgccaat cattaacttc 2160
tggcaagtca gttccggcat cccggatata ggcattgtgt ttagcaagca tattatccat 2220
ggattgaacg aaggccgcac cagtgttttc cattgctggt tgcgccgcaa ttgccgactt 2280
agctaagtcg aagcggtcca tctggttcat gacccgtacg tcgaatggtg tggtaatatc 2340

accattttca cggtaaccgt ggacgtataa gttatggttg tgacgatcaa agaagatgtc 2400
acgaactaag tcttcgtaac cgtggaaagc aaagaccact ggtttgtcct tagtaaagta 2460
atggtcaaac tcagcatctg acaagccccg cggatccttt tcaggactac gtaacttcaa 2520
gatgtcgacc acgttcacga aacgaatctt catctctggg aaactgtcgt gtagtaattg 2580
gatggcagcc aacgtttcaa gcgttggttc cgtcccagca gctgcaaaga caatgtctgg 2640
ttcgctacct tggtccgtac ttgcccaatc aatgatacca agaccattgt caactaattg 2700
cttagcttct tcaatgctga accattgttg acgtgggtgt tttgacgtaa ccacgtagtt 2760
gatcttttct tggctccgga aaatgacgtc accgacagct aataacgtgt tggcatcggc 2820
tggtaaatat tcacgaatgt attctggttt cttttcggcc aaatgagtta atgcacctgg 2880
atcttggtgg gtataaccat tatggtcttg ttggaataca gttgaagccg cgataatgtt 2940
aagtgatggg tacttttttac gccaatcaag ttcattggct ttacgtaacc acttgaagtg 3000
ttgcgtcaac attgagtcca caacgcgtag gaaggcttca taactggcaa ataacccatg 3060
acgtccagtt aagacgtaac cttctaacca accttcagct tggtgttcag ataactgagc 3120
atctaagacc cggccagctg gtgcttcata ttggtcacta tctggatgaa tgtcttccat 3180
ccattgacga ttagtggttt cgaagacacc atataaacgg ttagacatgg tttcatcagg 3240
tccgaacaac cggaagttat caggattttt cttgatgaca tcccgcaaat agtctgacca 3300
aacgatcata tcttgcttaa cattcgcgcc ttctttggac gtatcgaccg cataatcacg 3360
gaagtttggt aagttcaagg ctttcggatc gaccccacca ttggtgattg ggttagcagc 3420
catccgactg tccccagtag gaataatttc tttaatatca tccttcaaag agccatcttc 3480
attgaagagt tcttttggtt gatatgattc gagccaatca actaaagcat ccgcatgttc 3540
catgtcattt tgatcaacag gaatcggaat ttgatgagca cggaatgaac cttcgatctt 3600
atcaccgtcc catgacttcg gaccagtcca gcccttaggt gcgcggaaga cgatcattgg 3660
ccatactggc aatgttgcat cgttattttc gcgagcatgc ttctggattg ccttgatctt 3720
ttcaacggct tcatccatgg ccttagctaa ggctgggtga acctttcag gatcgtcacc 3780
ttcaacgaag attggttccc aattcatgct ttcgaagtat tccttaatct tagcatcaga 3840
agtccgacca aaaatcgttg gattagaaat cttaaaacca tttaagttca agattggtaa 3900
aacagccccg tcgttgattg ggttaatgaa cttcgttgat tgccatgaag ttgctaatgg 3960
acccgtttcg gattccccat caccaacaac aaccgcggcg atttcgtcag gattgtcaag 4020
aattgcccca accccgtgtg aaattgagta accaagttcg ccaccttcgt ggattgaacc 4080
gggtgtttca ggtgccgcat gggaagcaac cccacctggg aatgagaatt gcttgaagag 4140
cttttgcatc ccttcaacat cctgcgtaat ttctggataa atatcggtgt aagtaccgtc 4200
aaggtaagag tttgaaacca tcacttgacc accatgacct ggaccttcaa cgtagaacat 4260
cttcaaaccg tacttgttga tgacccggtt aagatgagca tagataaagt tttgaccggc 4320
aatcgtcccc cagtgaccaa ttggatgaac cttaacgtca ctggccttca atggccgttg 4380
taatagtgga ttatctttta aataaagttg accaactgat aagtagttgg cagcacgcca 4440
gtacttatca acttttgca aatatgctgg tgatgagtaa tctgttgtca tcctcagctg 4500

gaacttagat tagattgcta tgctttctct ctaacgagca agaagtaaaa aaagttgtaa 4560
tagaacaaga aaaatgaaac tgaagcttga gaaattgaag accgtttatt agcttaaata 4620
tcaatgggag gtcatcgaaa gagaaaaaaa tcaagaaaga aactctcaag aaaaagaaac 4680
gtgataaaaa tttttattgc ctctctcgac gaagagaaag aaacgaggcg gtccctttt 4740
tcttttccaa acctttagta cgggtaatta gcgacaccct agaggaagaa agaggggaaa 4800
tttagtatgc tgtgcttggg tgtcttgaag tggtacggcg atgcgcggag tccgagaaaa 4860
tctggaagag taaaaagggg gtagaagcgt tttgaagcta tccgcggtgg ttaagcctaa 4920
ccaggccaat tcaacagact gtcggcaact tcttgtctgg tctttccatg gtaagtgaca 4980
gtgcagtaat aatatgaacc aatttatttt tcgttacata aaaatgctta taaaacttta 5040
actaataatt agagattaaa tcgcaaacgg ccggccaatg tggctgtggt ttcagggtcc 5100
ataaagcttt tcaattcatc ttttttttt ttgttctttt ttttgattcc ggtttctttg 5160
aaatttttt gattcggtaa tctccgagca gaaggaagaa cgaaggaagg agcacagact 5220
tagattggta tatatacgca tatgtggtgt tgaagaaaca tgaaattgcc cagtattctt 5280
aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt cgaaagctac 5340
atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaagctat ttaatatcat 5400
gcacgaaaag caaacaaact tgtgtgcttc attggatgtt cgtaccacca aggaattact 5460
ggagttagtt gaagcattag gtcccaaaat ttgtttacta aaaacacatg tggatatctt 5520
gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg ccaagtacaa 5580
ttttttactc ttcgaagaca gaaaatttgc tgacattggt aatacagtca aattgcagta 5640
ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac acggtgtggt 5700
gggcccaggt attgttagcg gtttgaagca ggcggcggaa gaagtaacaa aggaacctag 5760
aggccttttg atgttagcag aattgtcatg caagggctcc ctagctactg gagaatatac 5820
taagggtact gttgacattg cgaagagcga caaagatttt gttatcggct ttattgctca 5880
aagagacatg ggtggaagag atgaaggtta cgattggttg attatgacac ccggtgtggg 5940
tttagatgac aagggagacg cattgggtca acagtataga accgtggatg atgtggtctc 6000
tacaggatct gacattatta ttgttggaag aggactattt gcaaagggaa gggatgctaa 6060
ggtagagggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa gatgcggcca 6120
gcaaaactaa aaaactgtat tataagtaaa tgcatgtata ctaaactcac aaattagagc 6180
ttcaatttaa ttatatcagt tattacccgg gaatctcggt cgtaatgatt tctataatga 6240
cgaaaaaaaa aaaattggaa agaaaaagct tcatggcctt gcggccgctt aattaatcta 6300
gagtcgacct gcaggcatgc aagcttggcg taatcatggt catagctgtt tcctgtgtga 6360
aattgttatc cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc 6420
tggggtgcct aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc 6480
cagtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc 6540
ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 6600
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 6660

ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 6720
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 6780
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 6840
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 6900
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 6960
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccccgt tcagcccgac 7020
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 7080
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 7140
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc 7200
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 7260
accaccgctg gtagcggtgg ttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 7320
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 7380
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 7440
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 7500
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 7560
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 7620
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac 7680
cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 7740
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 7800
gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc 7860
agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg 7920
gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc 7980
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 8040
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 8100
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc 8160
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 8220
agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc 8280
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca 8340
cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt 8400
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aatagggggtt 8460
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 8520
ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt cggtgatgac 8580
ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct gtaagcggat 8640
gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg tcggggctgg 8700
cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatgcg gtgtgaaata 8760
ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc attcgccatt caggctgcgc 8820

aactgttggg aagggcgatc ggtgcgggcc tcttcgctat tacgccagct ggcgaaaggg 8880

ggatgtgctg caaggcgatt aagttgggta acgccagggt tttcccagtc acgacgttgt 8940

aaaacgacgg ccagtgaatt cgagctcggt acccggggat ccggcgcgcc gttttatttg 9000

tatcgaggtg tctagtcttc tattacacta atgcagtttc agggttttgg aaaccacact 9060

gtttaaacag tgttccttaa tcaaggatac ctctttttt ttccttggtt ccactaattc 9120

atcggttttt tttttggaag acatcttttc caacgaaaag aatatacata tcgtttaaga 9180

gaaattctcc aaatttgtaa agaagcggac ccagacttaa 9220

<210> 292
<211> 38
<212> DNA
<213> artificial sequence

<220>
<223> primer BK468

<400> 292
gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 293
<211> 10934
<212> DNA
<213> Artificial sequence

<220>
<223> pRS423::TEF(M4)-xpk1+ENO1-eutD

<400> 293
ggtggagctc cagcttttgt tccctttagt gagggttaat tgcgcgcttg gcgtaatcat 60

ggtcatagct gtttcctgtg tgaaattgtt atccgctcac aattccacac aacataggag 120

ccggaagcat aaagtgtaaa gcctggggtg cctaatgagt gaggtaactc acattaattg 180

cgttgcgctc actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa 240

tcggccaacg cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca 300

ctgactcgct gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg 360

taatacggtt atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc 420

agcaaaaggc caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc 480

cccctgacga gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac 540

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 600

tgccgcttac cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 660

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 720

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca 780

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 840

cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 900

gaaggacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 960

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 1020

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 1080

ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 1140
ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 1200
atgagtaaac ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga 1260
tctgtctatt tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac 1320
gggagggctt accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg 1380
ctccagattt atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg 1440
caactttatc cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt 1500
cgccagttaa tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct 1560
cgtcgtttgg tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat 1620
cccccatgtt gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta 1680
agttggccgc agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca 1740
tgccatccgt aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat 1800
agtgtatgcg gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac 1860
atagcagaac tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa 1920
ggatcttacc gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt 1980
cagcatcttt tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg 2040
caaaaaaggg aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat 2100
attattgaag catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt 2160
agaaaaataa acaaataggg gttccgcgca catttccccg aaaagtgcca cctgaacgaa 2220
gcatctgtgc ttcattttgt agaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca 2280
aagaatctga gctgcatttt tacagaacag aaatgcaacg cgaaagcgct attttaccaa 2340
cgaagaatct gtgcttcatt tttgtaaaac aaaaatgcaa cgcgagagcg ctaatttttc 2400
aaacaaagaa tctgagctgc atttttacag aacagaaatg caacgcgaga gcgctatttt 2460
accaacaaag aatctatact tctttttgt tctacaaaaa tgcatcccga gagcgctatt 2520
tttctaacaa agcatcttag attacttttt ttctcctttg tgcgctctat aatgcagtct 2580
cttgataact ttttgcactg taggtccgtt aaggttagaa gaaggctact ttggtgtcta 2640
ttttctcttc cataaaaaaa gcctgactcc acttcccgcg tttactgatt actagcgaag 2700
ctgcgggtgc atttttcaa gataaaggca tccccgatta tattctatac cgatgtggat 2760
tgcgcatact ttgtgaacag aaagtgatag cgttgatgat tcttcattgg tcagaaaatt 2820
atgaacggtt tcttctattt tgtctctata tactacgtat aggaaatgtt tacattttcg 2880
tattgttttc gattcactct atgaatagtt cttactacaa ttttttgtc taaagagtaa 2940
tactagagat aaacataaaa aatgtagagg tcgagtttag atgcaagttc aaggagcgaa 3000
aggtggatgg gtaggttata tagggatata gcacagagat atatagcaaa gagatacttt 3060
tgagcaatgt ttgtggaagc ggtattcgca atattttagt agctcgttac agtccggtgc 3120
gtttttggtt ttttgaaagt gcgtcttcag agcgcttttg gttttcaaaa gcgctctgaa 3180
gttcctatac tttctagaga ataggaactt cggaatagga acttcaaagc gtttccgaaa 3240

acgagcgctt ccgaaaatgc aacgcgagct gcgcacatac agctcactgt tcacgtcgca 3300
cctatatctg cgtgttgcct gtatatatat atacatgaga agaacggcat agtgcgtgtt 3360
tatgcttaaa tgcgtactta tatgcgtcta tttatgtagg atgaaaggta gtctagtacc 3420
tcctgtgata ttatcccatt ccatgcgggg tatcgtatgc ttccttcagc actacccttt 3480
agctgttcta tatgctgcca ctcctcaatt ggattagtct catccttcaa tgctatcatt 3540
tcctttgata ttggatcatc taagaaacca ttattatcat gacattaacc tataaaaata 3600
ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac 3660
acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag 3720
cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat 3780
cagagcagat tgtactgaga gtgcaccata aattcccgtt ttaagagctt ggtgagcgct 3840
aggagtcact gccaggtatc gtttgaacac ggcattagtc agggaagtca taacacagtc 3900
ctttcccgca attttctttt tctattactc ttggcctcct ctagtacact ctatattttt 3960
ttatgcctcg gtaatgattt tcatttttt ttttccccta gcggatgact ctttttttt 4020
cttagcgatt ggcattatca cataatgaat tatacattat ataaagtaat gtgatttctt 4080
cgaagaatat actaaaaaat gagcaggcaa gataaacgaa ggcaaagatg acagagcaga 4140
aagccctagt aaagcgtatt acaaatgaaa ccaagattca gattgcgatc tctttaaagg 4200
gtggtcccct agcgatagag cactcgatct tcccagaaaa agaggcagaa gcagtagcag 4260
aacaggccac acaatcgcaa gtgattaacg tccacacagg tatagggttt ctggaccata 4320
tgatacatgc tctggccaag cattccggct ggtcgctaat cgttgagtgc attggtgact 4380
tacacataga cgaccatcac accactgaag actgcgggat tgctctcggt caagctttta 4440
aagaggccct actggcgcgt ggagtaaaaa ggtttggatc aggatttgcg cctttggatg 4500
aggcactttc cagagcggtg gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg 4560
gtttgcaaag ggagaaagta ggagatctct cttgcgagat gatcccgcat tttcttgaaa 4620
gctttgcaga ggctagcaga attaccctcc acgttgattg tctgcgaggc aagaatgatc 4680
atcaccgtag tgagagtgcg ttcaaggctc ttgcggttgc cataagagaa gccacctcgc 4740
ccaatggtac caacgatgtt ccctccacca aaggtgttct tatgtagtga caccgattat 4800
ttaaagctgc agcatacgat atatatacat gtgtatatat gtatacctat gaatgtcagt 4860
aagtatgtat acgaacagta tgatactgaa gatgacaagg taatgcatca ttctatacgt 4920
gtcattctga acgaggcgcg ctttcctttt ttctttttgc tttttctttt tttttctctt 4980
gaactcgacg gatctatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc 5040
atcaggaaat tgtaaacgtt aatattttgt taaaattcgc gttaaatttt tgttaaatca 5100
gctcattttt taaccaatag gccgaaatcg gcaaaatccc ttataaatca aaagaataga 5160
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 5220
actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 5280
caccctaatc aagtttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 5340
ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 5400

agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt agcggtcacg ctgcgcgtaa 5460
ccaccacacc cgccgcgctt aatgcgccgc tacagggcgc gtcgcgccat tcgccattca 5520
ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagctgg 5580
cgaaaggggg atgtgctgca aggcgattaa gttgggtaac gccagggttt tcccagtcac 5640
gacgttgtaa aacgacggcc agtgagcgcg cgtaatacga ctcactatag ggcgaattgg 5700
gtaccgggcc ccccctcgag gtcgacggta tcgataagct tgatatcgaa ttcctgcgcc 5760
cgggccacta gtcagatgcc gcgggcactt gagcacctca tgcacagcaa taacacaaca 5820
caatggttag tagcaacctg aattcggtca ttgatgcatg catgtgccgt gaagcgggac 5880
aaccagaaaa gtcgtctata aatgccggca cgtgcgatca tcgtggcggg gttttaagag 5940
tgcatatcac aaattgtcgc attaccgcgg aaccgccaga tattcattac ttgacgcaaa 6000
agcgtttgaa ataatgacga aaaagaagga agaaaaaaaa agaaaaatac cgcttctagg 6060
cgggttatct actgatccga gcttccacta ggatagcacc caaacacctg catatttgga 6120
cgacctttac ttacaccacc aaaaaccact ttcgcctctc ccgcccctga taacgtccac 6180
taattgagcg attacctgag cggtcctctt ttgtttgcag catgagactt gcatactgca 6240
aatcgtaagt agcaacgtct caaggtcaaa actgtatgga aaccttgtca cctcacttaa 6300
ttctagctag cctaccctgc aagtcaagag gtctccgtga ttcctagcca cctcaaggta 6360
tgcctctccc cggaaactgt ggccttttct ggcacacatg atctccacga tttcaacata 6420
taaatagctt ttgataatgg caatattaat caaatttatt ttacttcttt cttgtaacat 6480
ctctcttgta atcccttatt ccttctagct atttttcata aaaaaccaag caactgctta 6540
tcaacacaca aacactaaat caaagctgag gatggattta tttgagtcat tagcacaaaa 6600
aattactggt aaagatcaaa caattgtttt ccctgaagga actgaacccc gaattgtcgg 6660
tgcggcagcg cgattagctg cagacggctt ggttaagccg attgttttag gtgcaacgga 6720
caaagttcag gctgtggcta acgatttgaa tgcggattta acaggcgttc aagtccttga 6780
tcctgcgaca tacccggctg aagataagca agcaatgctt gatgccctcg ttgaacggcg 6840
gaaaggtaag aatacgccag aacaagcggc taaaatgctg gaagatgaaa actactttgg 6900
cacgatgctc gtttatatgg gcaaagcgga tgggatggtt tcaggtgcaa tccatccaac 6960
tggtgatacg gtacggccag cgttacaaat tattaagacc aagcccggtt cacaccgaat 7020
ctcgggtgca tttatcatgc aaaagggtga ggaacgctac gtctttgctg actgtgccat 7080
caatattgat cccgatgccg atacgttagc ggaaattgcc actcagagtg cggctactgc 7140
taaggtcttc gatattgacc cgaaagttgc gatgctcagc ttctcaacta agggttcggc 7200
taagggtgaa atggtcacta aagtgcaaga agcaacggcc aaggcgcaag ctgctgaacc 7260
ggaattggct atcgatggtg aacttcaatt tgacgcggcc ttcgttgaaa aagttggttt 7320
gcaaaaggct cctggttcca aagtagctgg tcatgccaat gtctttgtat ttccagagct 7380
tcagtctggt aatattggct ataagattgc gcaacgattt ggtcattttg aagcggtggg 7440
tcctgtcttg caaggcctga acaagccggt ctccgacttg tcacgtggat gcagtgaaga 7500
agacgtttat aaggttgcga ttattacagc agcccaagga ttagcttaat taattaagag 7560

20 Jan 2017 2017200439

taagcgaatt tcttatgatt tatgattttt attattaaat aagttataaa aaaaataagt 7620
gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt cttgagtaac 7680
tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgctct tattgaccac 7740
acctctaccg gcatgccgag caaatgcctg caaatcgctc cccatttcac ccaattgtag 7800
atatgctaac tccagcaatg agttgatgaa tctcggtgtg tattttatgt cctcagagga 7860
caacacctgt ggtactagtt ctagagcggc cgcccgcaaa ttaaagcctt cgagcgtccc 7920
aaaaccttct caagcaaggt tttcagtata atgttacatg cgtacacgcg tttgtacaga 7980
aaaaaaagaa aaatttgaaa tataaataac gttcttaata ctaacataac tattaaaaaa 8040
aataaatagg gacctagact tcaggttgtc taactccttc cttttcggtt agagcggatg 8100
tgggaggagg gcgtgaatgt aagcgtgaca taactaatta catgattaat taattatttt 8160
aaacccttcc attgccaatc attaacttct ggcaagtcag ttccggcatc ccggatatag 8220
gcattgtgtt tagcaagcat attatccatg gattgaacga aggccgcacc agtgttttcc 8280
attgctggtt gcgccgcaat tgccgactta gctaagtcga agcggtccat ctggttcatg 8340
acccgtacgt cgaatggtgt ggtaatatca ccattttcac ggtaaccgtg gacgtataag 8400
ttatggttgt gacgatcaaa gaagatgtca cgaactaagt cttcgtaacc gtggaaagca 8460
aagaccactg gtttgtcctt agtaaagtaa tggtcaaact cagcatctga caagccccgc 8520
ggatcctttt caggactacg taacttcaag atgtcgacca cgttcacgaa acgaatcttc 8580
atctctggga aactgtcgtg tagtaattgg atggcagcca acgtttcaag cgttggttcc 8640
gtcccagcag ctgcaaagac aatgtctggt tcgctacctt ggtccgtact tgcccaatca 8700
atgataccaa gaccattgtc aactaattgc ttagcttctt caatgctgaa ccattgttga 8760
cgtgggtgtt ttgacgtaac cacgtagttg atcttttctt ggctccggaa aatgacgtca 8820
ccgacagcta ataacgtgtt ggcatcggct ggtaaatatt cacgaatgta ttctggtttc 8880
ttttcggcca aatgagttaa tgcacctgga tcttggtggg tataaccatt atggtcttgt 8940
tggaatacag ttgaagccgc gataatgtta agtgatgggt acttttacg ccaatcaagt 9000
tcattggctt tacgtaacca cttgaagtgt tgcgtcaaca ttgagtccac aacgcgtagg 9060
aaggcttcat aactggcaaa taacccatga cgtccagtta agacgtaacc ttctaaccaa 9120
ccttcagctt ggtgttcaga taactgagca tctaagaccc ggccagctgg tgcttcatat 9180
tggtcactat ctggatgaat gtcttccatc cattgacgat tagtggtttc gaagacacca 9240
tataaacggt tagacatggt ttcatcaggt ccgaacaacc ggaagttatc aggatttttc 9300
ttgatgacat cccgcaaata gtctgaccaa acgatcatat cttgcttaac attcgcgcct 9360
tctttggacg tatcgaccgc ataatcacgg aagtttggta agttcaaggc tttcggatcg 9420
accccaccat tggtgattgg gttagcagcc atccgactgt ccccagtagg aataatttct 9480
ttaatatcat ccttcaaaga gccatcttca ttgaagagtt cttttggttg atatgattcg 9540
agccaatcaa ctaaagcatc cgcatgttcc atgtcatttt gatcaacagg aatcggaatt 9600
tgatgagcac ggaatgaacc ttcgatctta tcaccgtccc atgacttcgg accagtccag 9660
cccttaggtg cgcggaagac gatcattggc catactggca atgttgcatc gttattttcg 9720

cgagcatgct tctggattgc cttgatcttt tcaacggctt catccatggc cttagctaag 9780

gctgggtgaa ccttttcagg atcgtcacct tcaacgaaga ttggttccca attcatgctt 9840

tcgaagtatt ccttaatctt agcatcagaa gtccgaccaa aaatcgttgg attagaaatc 9900

ttaaaaccat ttaagttcaa gattggtaaa acagccccgt cgttgattgg gttaatgaac 9960

ttcgttgatt gccatgaagt tgctaatgga cccgtttcgg attccccatc accaacaaca 10020

accgcggcga tttcgtcagg attgtcaaga attgccccaa ccccgtgtga aattgagtaa 10080

ccaagttcgc caccttcgtg gattgaaccg ggtgtttcag gtgccgcatg ggaagcaacc 10140

ccacctggga atgagaattg cttgaagagc ttttgcatcc cttcaacatc ctgcgtaatt 10200

tctggataaa tatcggtgta agtaccgtca aggtaagagt ttgaaaccat cacttgacca 10260

ccatgacctg gaccttcaac gtagaacatc ttcaaaccgt acttgttgat gacccggtta 10320

agatgagcat agataaagtt ttgaccggca atcgtccccc agtgaccaat tggatgaacc 10380

ttaacgtcac tggccttcaa tggccgttgt aatagtggat tatcttttaa ataaagttga 10440

ccaactgata agtagttggc agcacgccag tacttatcaa ctttttgcaa atatgctggt 10500

gatgagtaat ctgttgtcat cctcagctgg aacttagatt agattgctat gctttctctc 10560

taacgagcaa gaagtaaaaa aagttgtaat agaacaagaa aaatgaaact gaagcttgag 10620

aaattgaaga ccgtttatta gcttaaatat caatgggagg tcatcgaaag agaaaaaaat 10680

caagaaagaa actctcaaga aaaagaaacg tgataaaaat ttttattgcc tctctcgacg 10740

aagagaaaga aacgaggcgg tccctttttt cttttccaaa cctttagtac gggtaattag 10800

cgacacccta gaggaagaaa gaggggaaat ttagtatgct gtgcttgggt gtcttgaagt 10860

ggtacggcga tgcgcggagt ccgagaaaat ctggaagagt aaaaaggggg tagaagcgtt 10920

ttgaagctat ccgc 10934

<210> 294
<211> 80
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 294
ttccggtttc tttgaaattt ttttgattcg gtaatctccg agcagaagga gcattgcgga 60

ttacgtattc taatgttcag 80

<210> 295
<211> 81
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 295
gggtaataac tgatataatt aaattgaagc tctaatttgt gagtttagta caccttggct 60

aactcgttgt atcatcactg g 81

<210> 296
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 296
gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 297
<211> 31
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 297
gctaaattcg agtgaaacac aggaagacca g 31

<210> 298
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 298
ttctcgacgt gggccttttt cttg 24

<210> 299
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 299
tgcagcttta aataatcggt gtcactactt tgccttcgtt tatcttgcc 49

<210> 300
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 300
gagcaggcaa gataaacgaa ggcaaagtag tgacaccgat tatttaaag 49

<210> 301
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 301
tatggaccct gaaaccacag ccacattgta accaccacga cggttgttg 49

<210> 302
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 302
tttagcaaca accgtcgtgg tggttacaat gtggctgtgg tttcagggt 49

<210> 303
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 303
ccagaaaccc tatacctgtg tggacgtaag gccatgaagc tttttcttt 49

<210> 304
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 304
attggaaaga aaaagcttca tggccttacg tccacacagg tatagggtt 49

<210> 305
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 305
cataagaaca cctttggtgg ag 22

<210> 306
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 306
aggattatca ttcataagtt tc 22

<210> 307
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 307
ttcttggagc tgggacatgt ttg 23

<210> 308
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 308
tgatgatatt tcataaataa tg 22

<210> 309
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 309
atgcgtccat ctttacagtc ctg 23

<210> 310
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 310
tacgtacgga ccaatcgaag tg 22

<210> 311
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 311
aattcgtttg agtacactac taatggcttt gttggcaata tgtttttgc 49

<210> 312
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 312
atatagcaaa aacatattgc caacaaagcc attagtagtg tactcaaac 49

<210> 313
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 313
tatggaccct gaaaccacag ccacattctt gttatttata aaaagacac 49

<210> 314
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 314
ctcccgtgtc tttttataaa taacaagaat gtggctgtgg tttcagggt 49

<210> 315
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 315
taccgtaggc gtccttagga aagatagaag gccatgaagc tttttcttt 49

<210> 316
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 316
attggaaaga aaaagcttca tggccttcta tctttcctaa ggacgccta 49

<210> 317
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 317
ttattgtttg gcatttgtag c 21

<210> 318
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 318
ccaagcatct cataaaccta tg 22

<210> 319
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 319
tgtgcagatg cagatgtgag ac 22

<210> 320
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 320
agttattgat accgtac 17

<210> 321
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 321
cgagataccg taggcgtcc 19

<210> 322
<211> 2237
<212> DNA
<213> Artificial sequence

<220>
<223> Fragment

<400> 322
ttatgtatgc tcttctgact tttcgtgtga tgaggctcgt ggaaaaaatg aataatttat 60

gaatttgaga acaattttgt gttgttacgg tattttacta tggaataatc aatcaattga 120

ggattttatg caaatatcgt ttgaatattt ttccgaccct ttgagtactt ttcttcataa 180

ttgcataata ttgtccgctg ccccttttc tgttagacgg tgtcttgatc tacttgctat 240

cgttcaacac caccttattt tctaactatt tttttttag ctcatttgaa tcagcttatg 300

gtgatggcac atttttgcat aaacctagct gtcctcgttg aacataggaa aaaaaaatat 360

ataaacaagg ctctttcact ctccttgcaa tcagatttgg gtttgttccc tttattttca 420

tatttcttgt catattcctt tctcaattat tattttctac tcataacctc acgcaaaata 480

acacagtcaa atcaatcaaa atgactgaca aaaaaactct taaagactta agaaatcgta 540

gttctgttta cgattcaatg gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta 600

tgcaagatga agactttgaa aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca 660

caccttgtaa tatccactta catgactttg gtaaactagc caaagtcggt gttaaggaag 720

ctggtgcttg gccagttcag ttcggaacaa tcacggtttc tgatggaatc gccatgggaa 780

cccaaggaat gcgtttctcc ttgacatctc gtgatattat tgcagattct attgaagcag 840

ccatgggagg tcataatgcg gatgcttttg tagccattgg cggttgtgat aaaaacatgc 900

ccggttctgt tatcgctatg gctaacatgg atatcccagc cattttgct tacggcggaa 960

caattgcacc tggtaattta gacggcaaag atatcgattt agtctctgtc tttgaaggtg 1020

tcggccattg gaaccacggc gatatgacca aagaagaagt taaagctttg gaatgtaatg 1080

cttgtcccgg tcctggaggc tgcggtggta tgtatactgc taacacaatg gcgacagcta 1140

ttgaagtttt gggacttagc cttccgggtt catcttctca cccggctgaa tccgcagaaa 1200

agaaagcaga tattgaagaa gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa 1260
aaccttctga cattttaacg cgtgaagctt ttgaagatgc tattactgta actatggctc 1320
tgggaggttc aaccaactca acccttcacc tcttagctat tgcccatgct gctaatgtgg 1380
aattgacact tgatgatttc aatactttcc aagaaaaagt tcctcatttg gctgatttga 1440
aaccttctgg tcaatatgta ttccaagacc tttacaaggt cggaggggta ccagcagtta 1500
tgaaatatct ccttaaaaat ggcttccttc atggtgaccg tatcacttgt actggcaaaa 1560
cagtcgctga aaatttgaag gcttttgatg atttaacacc tggtcaaaag gttattatgc 1620
cgcttgaaaa tcctaaacgt gaagatggtc cgctcattat tctccatggt aacttggctc 1680
cagacggtgc cgttgccaaa gtttctggtg taaaagtgcg tcgtcatgtc ggtcctgcta 1740
aggtctttaa ttctgaagaa gaagccattg aagctgtctt gaatgatgat attgttgatg 1800
gtgatgttgt tgtcgtacgt tttgtaggac caaagggcgg tcctggtatg cctgaaatgc 1860
tttccctttc atcaatgatt gttggtaaag ggcaaggtga aaaagttgcc cttctgacag 1920
atggccgctt ctcaggtggt acttatggtc ttgtcgtggg tcatatcgct cctgaagcac 1980
aagatggcgg tccaatcgcc tacctgcaaa caggagacat agtcactatt gaccaagaca 2040
ctaaggaatt acactttgat atctccgatg aagagttaaa acatcgtcaa gagaccattg 2100
aattgccacc gctctattca cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg 2160
cttctagggg agccgtaaca gacttttgga agcctgaaga aactggcaaa aaatgagcga 2220
tttaatctct aattatt 2237

<210> 323
<211> 4420
<212> DNA
<213> Artificial sequence

<220>
<223> Cassette

<400> 323
ttatgtatgc tcttctgact tttcgtgtga tgaggctcgt ggaaaaaatg aataatttat 60
gaatttgaga acaattttgt gttgttacgg tattttacta tggaataatc aatcaattga 120
ggattttatg caaatatcgt ttgaatattt ttccgaccct ttgagtactt ttcttcataa 180
ttgcataata ttgtccgctg ccccttttc tgttagacgg tgtcttgatc tacttgctat 240
cgttcaacac caccttattt tctaactatt tttttttag ctcatttgaa tcagcttatg 300
gtgatggcac atttttgcat aaacctagct gtcctcgttg aacataggaa aaaaaaatat 360
ataaacaagg ctctttcact ctccttgcaa tcagatttgg gtttgttccc tttattttca 420
tatttcttgt catattcctt tctcaattat tattttctac tcataacctc acgcaaaata 480
acacagtcaa atcaatcaaa atgactgaca aaaaaactct taaagactta agaaatcgta 540
gttctgttta cgattcaatg gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta 600
tgcaagatga agactttgaa aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca 660
caccttgtaa tatccactta catgactttg gtaaactagc caaagtcggt gttaaggaag 720
ctggtgcttg gccagttcag ttcggaacaa tcacggtttc tgatggaatc gccatgggaa 780

cccaaggaat gcgtttctcc ttgacatctc gtgatattat tgcagattct attgaagcag 840
ccatgggagg tcataatgcg gatgcttttg tagccattgg cggttgtgat aaaaacatgc 900
ccggttctgt tatcgctatg gctaacatgg atatcccagc cattttgct tacggcggaa 960
caattgcacc tggtaattta gacggcaaag atatcgattt agtctctgtc tttgaaggtg 1020
tcggccattg gaaccacggc gatatgacca aagaagaagt taaagctttg gaatgtaatg 1080
cttgtcccgg tcctggaggc tgcggtggta tgtatactgc taacacaatg gcgacagcta 1140
ttgaagtttt gggacttagc cttccgggtt catcttctca cccggctgaa tccgcagaaa 1200
agaaagcaga tattgaagaa gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa 1260
aaccttctga cattttaacg cgtgaagctt ttgaagatgc tattactgta actatggctc 1320
tgggaggttc aaccaactca acccttcacc tcttagctat tgcccatgct gctaatgtgg 1380
aattgacact tgatgatttc aatactttcc aagaaaaagt tcctcatttg gctgatttga 1440
aaccttctgg tcaatatgta ttccaagacc tttacaaggt cggaggggta ccagcagtta 1500
tgaaatatct ccttaaaaat ggcttccttc atggtgaccg tatcacttgt actggcaaaa 1560
cagtcgctga aaatttgaag gcttttgatg atttaacacc tggtcaaaag gttattatgc 1620
cgcttgaaaa tcctaaacgt gaagatggtc cgctcattat tctccatggt aacttggctc 1680
cagacggtgc cgttgccaaa gtttctggtg taaaagtgcg tcgtcatgtc ggtcctgcta 1740
aggtctttaa ttctgaagaa gaagccattg aagctgtctt gaatgatgat attgttgatg 1800
gtgatgttgt tgtcgtacgt tttgtaggac caaagggcgg tcctggtatg cctgaaatgc 1860
tttccctttc atcaatgatt gttggtaaag ggcaaggtga aaaagttgcc cttctgacag 1920
atggccgctt ctcaggtggt acttatggtc ttgtcgtggg tcatatcgct cctgaagcac 1980
aagatggcgg tccaatcgcc tacctgcaaa caggagacat agtcactatt gaccaagaca 2040
ctaaggaatt acactttgat atctccgatg aagagttaaa acatcgtcaa gagaccattg 2100
aattgccacc gctctattca cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg 2160
cttctagggg agccgtaaca gacttttgga agcctgaaga aactggcaaa aaatgagcga 2220
tttaatctct aattattagt taaagtttta taagcatttt tatgtaacga aaaataaatt 2280
ggttcatatt attactgcac tgtcacttac catggaaaga ccagacaaga agttgccgac 2340
agtctgttga attggcctgg ttaggcttaa gtctgggtcc gcttctttac aaatttggag 2400
aatttctctt aaacgatatg tatattcttt tcgttggaaa agatgtcttc caaaaaaaaa 2460
accgatgaat tagtggaacc aaggaaaaaa aaagaggtat ccttgattaa ggaacactgt 2520
ttaaacagtg tggtttccaa aaccctgaaa ctgcattagt gtaatagaag actagacacc 2580
tcgatacaaa taatggttac tcaattcaaa actgccagcg aattcgactc tgcaattgct 2640
caagacaagc tagttgtcgt agatttctac gccacttggt gcggtccatg taaaatgatt 2700
gctccaatga ttgaaaaatg tggctgtggt ttcagggtcc ataaagcttt tcaattcatc 2760
ttttttttt ttgttctttt ttttgattcc ggtttctttg aaatttttt gattcggtaa 2820
tctccgagca gaaggaagaa cgaaggaagg agcacagact tagattggta tatatacgca 2880
tatgtggtgt tgaagaaaca tgaaattgcc cagtattctt aacccaactg cacagaacaa 2940

aaacctgcag gaaacgaaga taaatcatgt cgaaagctac atataaggaa cgtgctgcta 3000

ctcatcctag tcctgttgct gccaagctat ttaatatcat gcacgaaaag caaacaaact 3060

tgtgtgcttc attggatgtt cgtaccacca aggaattact ggagttagtt gaagcattag 3120

gtcccaaaat ttgtttacta aaaacacatg tggatatctt gactgatttt tccatggagg 3180

gcacagttaa gccgctaaag gcattatccg ccaagtacaa ttttttactc ttcgaagaca 3240

gaaaatttgc tgacattggt aatacagtca aattgcagta ctctgcgggt gtatacagaa 3300

tagcagaatg ggcagacatt acgaatgcac acggtgtggt gggcccaggt attgttagcg 3360

gtttgaagca ggcggcggaa gaagtaacaa aggaacctag aggccttttg atgttagcag 3420

aattgtcatg caagggctcc ctagctactg gagaatatac taagggtact gttgacattg 3480

cgaagagcga caaagatttt gttatcggct ttattgctca aagagacatg ggtggaagag 3540

atgaaggtta cgattggttg attatgacac ccggtgtggg tttagatgac aagggagacg 3600

cattgggtca acagtataga accgtggatg atgtggtctc tacaggatct gacattatta 3660

ttgttggaag aggactattt gcaaagggaa gggatgctaa ggtagagggt gaacgttaca 3720

gaaaagcagg ctgggaagca tatttgagaa gatgcggcca gcaaaactaa aaaactgtat 3780

tataagtaaa tgcatgtata ctaaactcac aaattagagc ttcaatttaa ttatatcagt 3840

tattacccgg gaatctcggt cgtaatgatt tctataatga cgaaaaaaaa aaaattggaa 3900

agaaaaagct tcatggcctt ccactttccc aaacaacacc tacggtatct ctcaagtctt 3960

atggggttcc attggtttca ccactggtgc taccttgggt gctgctttcg ctgctgaaga 4020

aattgatcca aagaagagag ttatcttatt cattggtgac ggttctttgc aattgactgt 4080

tcaagaaatc tccaccatga tcagatgggg cttgaagcca tacttgttcg tcttgaacaa 4140

cgatggttac accattgaaa agttgattca cggtccaaag gctcaataca acgaaattca 4200

aggttgggac cacctatcct tgttgccaac tttcggtgct aaggactatg aaacccacag 4260

agtcgctacc accggtgaat gggacaagtt gacccaagac aagtctttca acgacaactc 4320

taagatcaga atgattgaaa tcatgttgcc agtcttcgat gctccacaaa acttggttga 4380

acaagctaag ttgactgctg ctaccaacgc taagcaataa 4420

<210> 324
<211> 3521
<212> DNA
<213> Artificial sequence

<220>
<223> Cassette

<400> 324
aaggaaataa agcaaataac aataacacca ttattttaat ttttttctta ttactgtcgc 60

taacacctgt atggttgcaa ccaggtgaga atccttctga tgcatacttt atgcgtttat 120

gcgttttgcg ccccttggaa aaaaattgat tctcatcgta aatgcatact acatgcgttt 180

atgggaaaag cctccatatc caaaggtcgc gtttctttta gaaaaactaa tacgtaaacc 240

tgcattaagg taagattata tcagaaaatg tgttgcaaga aatgcattat gcaatttttt 300

gattatgaca atctctcgaa agaaatttca tatgatgaga cttgaataat gcagcggcgc 360

20 Jan 2017 2017200439

```
ttgctaaaag aacttgtata taagagctgc cattctcgat caatatactg tagtaagtcc      420
tttcctctct ttcttattac acttatttca cataatcaat ctcaaagaga acaacacaat      480
acaataacaa gaagaacaaa atgaaagctc tggtttatca cggtgaccac aagatctcgc      540
ttgaagacaa gcccaagccc acccttcaaa agcccacgga tgtagtagta cgggttttga      600
agaccacgat ctgcggcacg gatctcggca tctacaaagg caagaatcca gaggtcgccg      660
acgggcgcat cctgggccat gaaggggtag gcgtcatcga ggaagtgggc gagagtgtca      720
cgcagttcaa gaaaggcgac aaggtcctga tttcctgcgt cacttcttgc ggctcgtgcg      780
actactgcaa gaagcagctt tactcccatt gccgcgacgg cgggtggatc ctgggttaca      840
tgatcgatgg cgtgcaggcc gaatacgtcc gcatcccgca tgccgacaac agcctctaca      900
agatccccca gacaattgac gacgaaatcg ccgtcctgct gagcgacatc ctgcccaccg      960
gccacgaaat cggcgtccag tatgggaatg tccagccggg cgatgcggtg gctattgtcg     1020
gcgcgggccc cgtcggcatg tccgtactgt tgaccgccca gttctactcc ccctcgacca     1080
tcatcgtgat cgacatggac gagaatcgcc tccagctcgc caaggagctc ggggcaacgc     1140
acaccatcaa ctccggcacg gagaacgttg tcgaagccgt gcataggatt gcggcagagg     1200
gagtcgatgt tgcgatcgag gcggtgggca taccggcgac ttgggacatc tgccaggaga     1260
tcgtcaagcc cggcgcgcac atcgccaacg tcggcgtgca tggcgtcaag gttgacttcg     1320
agattcagaa gctctggatc aagaacctga cgatcaccac gggactggtg aacacgaaca     1380
cgacgcccat gctgatgaag gtcgcctcga ccgacaagct tccgttgaag aagatgatta     1440
cccatcgctt cgagctggcc gagatcgagc acgcctatca ggtattcctc aatggcgcca     1500
aggagaaggc gatgaagatc atcctctcga acgcaggcgc tgcctgagct aattaacata     1560
aaactcatga ttcaacgttt gtgtattttt ttacttttga aggttataga tgtttaggta     1620
aataattggc atagatatag ttttagtata ataaatttct gatttggttt aaaatatcaa     1680
ctatttttt tcacatatgt tcttgtaatt acttttctgt cctgtcttcc aggttaaaga     1740
ttagcttcta atattttagg tggtttatta tttaatttta tgctgattaa tttatttact     1800
tgtttaaacg gccggccaat gtggctgtgg tttcagggtc cataaagctt ttcaattcat     1860
ctttttttt tttgttcttt tttttgattc cggtttcttt gaaatttttt tgattcggta     1920
atctccgagc agaaggaaga acgaaggaag gagcacagac ttagattggt atatatacgc     1980
atatgtggtg ttgaagaaac atgaaattgc ccagtattct taacccaact gcacagaaca     2040
aaaacctgca ggaaacgaag ataaatcatg tcgaaagcta catataagga acgtgctgct     2100
actcatccta gtcctgttgc tgccaagcta tttaatatca tgcacgaaaa gcaaacaaac     2160
ttgtgtgctt cattggatgt tcgtaccacc aaggaattac tggagttagt tgaagcatta     2220
ggtcccaaaa tttgtttact aaaaacacat gtggatatct tgactgattt ttccatggag     2280
ggcacagtta agccgctaaa ggcattatcc gccaagtaca atttttact cttcgaagac     2340
agaaaatttg ctgacattgg taatacagtc aaattgcagt actctgcggg tgtatacaga     2400
atagcagaat gggcagacat tacgaatgca cacggtgtgg tgggcccagg tattgttagc     2460
ggtttgaagc aggcggcgga agaagtaaca aaggaaccta gaggcctttt gatgttagca     2520
```

gaattgtcat gcaagggctc cctagctact ggagaatata ctaagggtac tgttgacatt 2580
gcgaagagcg acaaagattt tgttatcggc tttattgctc aaagagacat gggtggaaga 2640
gatgaaggtt acgattggtt gattatgaca cccggtgtgg gtttagatga caagggagac 2700
gcattgggtc aacagtatag aaccgtggat gatgtggtct ctacaggatc tgacattatt 2760
attgttggaa gaggactatt tgcaaaggga agggatgcta aggtagaggg tgaacgttac 2820
agaaaagcag gctgggaagc atatttgaga agatgcggcc agcaaaacta aaaaactgta 2880
ttataagtaa atgcatgtat actaaactca caaattagag cttcaattta attatatcag 2940
ttattacccg ggaatctcgg tcgtaatgat ttctataatg acgaaaaaaa aaaaattgga 3000
aagaaaaagc ttcatggcct tctactttcc caacagatgt atacgctatc gtccaagtct 3060
tgtggggttc cattggtttc acagtcggcg ctctattggg tgctactatg gccgctgaag 3120
aacttgatcc aaagaagaga gttattttat tcattggtga cggttctcta caattgactg 3180
ttcaagaaat ctctaccatg attagatggg gtttgaagcc atacattttt gtcttgaata 3240
acaacggtta caccattgaa aaattgattc acggtcctca tgccgaatat aatgaaattc 3300
aaggttggga ccacttggcc ttattgccaa cttttggtgc tagaaactac gaaacccaca 3360
gagttgctac cactggtgaa tgggaaaagt tgactcaaga caaggacttc caagacaact 3420
ctaagattag aatgattgaa gttatgttgc cagtctttga tgctccacaa aacttggtta 3480
aacaagctca attgactgcc gctactaacg ctaaacaata a 3521

<210> 325
<211> 2686
<212> DNA
<213> Artificial sequence

<220>
<223> Plasmid

<400> 325
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
attcgccatt caggctgcgc aactgttggg aagggcgatc ggtgcgggcc tcttcgctat 300
tacgccagct ggcgaaaggg ggatgtgctg caaggcgatt aagttgggta acgccagggt 360
tttcccagtc acgacgttgt aaaacgacgg ccagtgaatt cgagctcggt acccggggat 420
cctctagagt cgacctgcag gcatgcaagc ttggcgtaat catggtcata gctgtttcct 480
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 540
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc 600
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 660
agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 720
gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 780
gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 840

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga cgagcatcac 900

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 960

tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 1020

ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 1080

ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 1140

cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac 1200

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 1260

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 1320

atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 1380

aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 1440

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 1500

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 1560

cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 1620

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 1680

tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct 1740

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 1800

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 1860

atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 1920

cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct 1980

tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 2040

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 2100

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 2160

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 2220

agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa 2280

gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 2340

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 2400

accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg 2460

gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg aagcatttat 2520

cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaaata 2580

ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tctaagaaac cattattatc 2640

atgacattaa cctataaaaa taggcgtatc acgaggccct ttcgtc 2686

<210> 326
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 326
ttatgtatgc tcttctgact tttc 24

<210> 327
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 327
aataattaga gattaaatcg ctcatttttt gccagtttct tcaggcttc 49

<210> 328
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 328
agcctgaaga aactggcaaa aaatgagcga tttaatctct aattattag 49

<210> 329
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 329
tatggaccct gaaaccacag ccacattttt caatcattgg agcaatcat 49

<210> 330
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 330
taaaatgatt gctccaatga ttgaaaaatg tggctgtggt ttcagggtc 49

<210> 331
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 331
accgtaggtg ttgtttggga aagtggaagg ccatgaagct ttttctttc 49

<210> 332
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 332
ttggaaagaa aaagcttcat ggccttccac tttcccaaac aacacctac 49

<210> 333
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 333
ttattgctta gcgttggtag cag 23

<210> 334
<211> 57
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 334
aattggatcc ggcgcgccgt ttaaacggcc ggccaatgtg gctgtggttt cagggtc 57

<210> 335
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 335
aatttctaga ttaattaagc ggccgcaagg ccatgaagct ttttctttc 49

<210> 336
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 336
tttttggtgg ttccggcttc c 21

<210> 337
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 337
aaagttggca tagcggaaac tt 22

<210> 338
<211> 16
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 338
gtcattgaca ccatct 16

<210> 339
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 339
agagataccg taggtgttg 19

<210> 340
<211> 90
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 340
gtattttggt agattcaatt ctctttccct ttccttttcc ttcgctcccc ttccttatca 60

gcattgcgga ttacgtattc taatgttcag 90

<210> 341
<211> 90
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 341
ttggttgggg gaaaaagagg caacaggaaa gatcagaggg ggaggggggg ggagagtgtc 60

accttggcta actcgttgta tcatcactgg 90

<210> 342
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 342
tcggtgcggg cctcttcgct a 21

<210> 343
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 343
aatgtgagtt agctcactca t 21

<210> 344
<211> 33
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 344
aattggcgcg ccatgaaagc tctggtttat cac 33

<210> 345
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 345
tgaatcatga gttttatgtt aattagctca ggcagcgcct gcgttcgag 49

<210> 346
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 346
atcctctcga acgcaggcgc tgcctgagct aattaacata aaactcatg 49

<210> 347
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 347
aattgtttaa acaagtaaat aaattaatca gcat 34

<210> 348
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 348
acacaataca ataacaagaa gaacaaaatg aaagctctgg tttatcacg 49

<210> 349
<211> 49
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 349
agcgtataca tctgttggga aagtagaagg ccatgaagct ttttctttc 49

<210> 350
<211> 49
<212> DNA
<213> Artificial sequence

20 Jan 2017
2017200439

<220>
<223> Primer

<400> 350
ttggaaagaa aaagcttcat ggccttctac tttcccaaca gatgtatac 49

<210> 351
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 351
ttattgttta gcgttagtag cg 22

<210> 352
<211> 72
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 352
cataatcaat ctcaaagaga acaacacaat acaataacaa gaagaacaaa atgaaagctc 60

tggtttatca cg 72

<210> 353
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 353
taggcataat caccgaagaa g 21

<210> 354
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 354
aaaatggtaa gcagctgaaa g 21

<210> 355
<211> 17
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 355
agttgttaga actgttg 17

<210> 356
<211> 19

<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 356
gacgatagcg tatacatct 19

<210> 357
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 357
cttagcctct agccatagcc at 22

<210> 358
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Primer

<400> 358
ttagttttgc tggccgcatc ttc 23

<210> 359
<211> 15456
<212> DNA
<213> Artificial sequence

<220>
<223> Template

<400> 359
aaagagtaat actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca 60

aggagcgaaa ggtggatggg taggttatat agggatatag cacagagata tatagcaaag 120

agatactttt gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca 180

gtccggtgcg ttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag 240

cgctctgaag ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg 300

tttccgaaaa cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt 360

cacgtcgcac ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata 420

gtgcgtgttt atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag 480

tctagtacct cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca 540

ctacccttta gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat 600

gctatcattt cctttgatat tggatcatac taagaaacca ttattatcat gacattaacc 660

tataaaaata ggcgtatcac gaggcccttt cgtctcgcgc gtttcggtga tgacggtgaa 720

aacctctgac acatgcagct cccggagacg gtcacagctt gtctgtaagc ggatgccggg 780

agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg ggtgtcgggg ctggcttaac 840

tatgcggcat cagagcagat tgtactgaga gtgcaccata aattcccgtt ttaagagctt 900

ggtgagcgct aggagtcact gccaggtatc gtttgaacac ggcattagtc agggaagtca 960
taacacagtc ctttcccgca attttctttt tctattactc ttggcctcct ctagtacact 1020
ctatattttt ttatgcctcg gtaatgattt tcatttttt ttttccacct agcggatgac 1080
tctttttttt tcttagcgat tggcattatc acataatgaa ttatacatta tataaagtaa 1140
tgtgatttct tcgaagaata tactaaaaaa tgagcaggca agataaacga aggcaaagat 1200
gacagagcag aaagccctag taaagcgtat tacaaatgaa accaagattc agattgcgat 1260
ctctttaaag ggtggtcccc tagcgataga gcactcgatc ttcccagaaa aagaggcaga 1320
agcagtagca gaacaggcca cacaatcgca agtgattaac gtccacacag gtatagggtt 1380
tctggaccat atgatacatg ctctggccaa gcattccggc tggtcgctaa tcgttgagtg 1440
cattggtgac ttacacatag acgaccatca caccactgaa gactgcggga ttgctctcgg 1500
tcaagctttt aaagaggccc taggggccgt gcgtggagta aaaaggtttg gatcaggatt 1560
tgcgcctttg gatgaggcac tttccagagc ggtggtagat ctttcgaaca ggccgtacgc 1620
agttgtcgaa cttggtttgc aaagggagaa agtaggagat ctctcttgcg agatgatccc 1680
gcattttctt gaaagctttg cagaggctag cagaattacc ctccacgttg attgtctgcg 1740
aggcaagaat gatcatcacc gtagtgagag tgcgttcaag gctcttgcgg ttgccataag 1800
agaagccacc tcgcccaatg gtaccaacga tgttccctcc accaaaggtg ttcttatgta 1860
gtgacaccga ttatttaaag ctgcagcata cgatatatat acatgtgtat atatgtatac 1920
ctatgaatgt cagtaagtat gtatacgaac agtatgatac tgaagatgac aaggtaatgc 1980
atcattctat acgtgtcatt ctgaacgagg cgcgctttcc ttttttcttt ttgctttttc 2040
ttttttttte tcttgaactc gacggatcta tgcggtgtga aataccgcac agatgcgtaa 2100
ggagaaaata ccgcatcagg aaattgtaag cgttaatatt ttgttaaaat tcgcgttaaa 2160
tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa tcccttataa 2220
atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttggaaca agagtccact 2280
attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc 2340
actacgtgaa ccatcaccct aatcaagttt tttggggtcg aggtgccgta aagcactaaa 2400
tcggaaccct aaagggagcc cccgatttag agcttgacgg ggaaagccgg cgaacgtggc 2460
gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcggt 2520
cacgctgcgc gtaaccacca cacccgccgc gcttaatgcg ccgctacagg gcgcgtccat 2580
tcgccattca ggctgcgcaa ctgttgggaa gggcgcggtg cgggcctctt cgctattacg 2640
ccagctggcg aaaggggga gtgctgcaag gcgattaagt tgggtaacgc cagggttttc 2700
ccagtcacga cgttgtaaaa cgacggccag tgagcgcgcg taatacgact cactataggg 2760
cgaattgggt accgggcccc ccctcgaggt cgacggcgcg ccactggtag agagcgactt 2820
tgtatgcccc aattgcgaaa cccgcgatat ccttctcgat tctttagtac ccgaccagga 2880
caaggaaaag gaggtcgaaa cgtttttgaa gaaacaagag gaactacacg gaagctctaa 2940
agatggcaac cagccagaaa ctaagaaaat gaagttgatg gatccaactg gcaccgctgg 3000
cttgaacaac aataccagcc ttccaacttc tgtaaataac ggcggtacgc cagtgccacc 3060

agtaccgtta cctttcggta tacctccttt ccccatgttt ccaatgccct tcatgcctcc 3120
aacggctact atcacaaatc ctcatcaagc tgacgcaagc cctaagaaat gaataacaat 3180
actgacagta ctaaataatt gcctacttgg cttcacatac gttgcatacg tcgatataga 3240
taataatgat aatgacagca ggattatcgt aatacgtaat agctgaaaat ctcaaaaatg 3300
tgtgggtcat tacgtaaata atgataggaa tgggattctt ctatttttcc ttttccattt 3360
ctagcagccg tcgggaaaac gtggcatcct ctctttcggg ctcaattgga gtcacgctgc 3420
cgtgagcatc ctctctttcc atatctaaca actgagcacg taaccaatgg aaaagcatga 3480
gcttagcgtt gctccaaaaa agtattggat ggttaatacc atttgtctgt tctcttctga 3540
ctttgactcc tcaaaaaaaa aaatctacaa tcaacagatc gcttcaatta cgccctcaca 3600
aaaacttttt tccttcttct tcgcccacgt taaattttat ccctcatgtt gtctaacgga 3660
tttctgcact tgatttatta taaaaagaca aagacataat acttctctat caatttcagt 3720
tattgttctt ccttgcgtta ttcttctgtt cttctttttc ttttgtcata tataaccata 3780
accaagtaat acatattcaa actagtatga ctgacaaaaa aactcttaaa gacttaagaa 3840
atcgtagttc tgtttacgat tcaatggtta aatcacctaa tcgtgctatg ttgcgtgcaa 3900
ctggtatgca agatgaagac tttgaaaaac ctatcgtcgg tgtcatttca acttgggctg 3960
aaaacacacc ttgtaatatc cacttacatg actttggtaa actagccaaa gtcggtgtta 4020
aggaagctgg tgcttggcca gttcagttcg gaacaatcac ggtttctgat ggaatcgcca 4080
tgggaaccca aggaatgcgt ttctccttga catctcgtga tattattgca gattctattg 4140
aagcagccat gggaggtcat aatgcggatg cttttgtagc cattggcggt tgtgataaaa 4200
acatgcccgg ttctgttatc gctatggcta acatggatat cccagccatt tttgcttacg 4260
gcggaacaat tgcacctggt aatttagacg gcaaagatat cgatttagtc tctgtctttg 4320
aaggtgtcgg ccattggaac cacggcgata tgaccaaaga agaagttaaa gctttggaat 4380
gtaatgcttg tcccggtcct ggaggctgcg gtggtatgta tactgctaac acaatggcga 4440
cagctattga agttttggga cttagccttc cgggttcatc ttctcacccg gctgaatccg 4500
cagaaaagaa agcagatatt gaagaagctg gtcgcgctgt tgtcaaaatg ctcgaaatgg 4560
gcttaaaacc ttctgacatt ttaacgcgtg aagcttttga agatgctatt actgtaacta 4620
tggctctggg aggttcaacc aactcaaccc ttcacctctt agctattgcc catgctgcta 4680
atgtggaatt gacacttgat gatttcaata ctttccaaga aaaagttcct catttggctg 4740
atttgaaacc ttctggtcaa tatgtattcc aagaccttta caaggtcgga ggggtaccag 4800
cagttatgaa atatctcctt aaaaatggct tccttcatgg tgaccgtatc acttgtactg 4860
gcaaaacagt cgctgaaaat ttgaaggctt ttgatgattt aacacctggt caaaaggtta 4920
ttatgccgct tgaaaatcct aaacgtgaag atggtccgct cattattctc catggtaact 4980
tggctccaga cggtgccgtt gccaaagttt ctggtgtaaa agtgcgtcgt catgtcggtc 5040
ctgctaaggt ctttaattct gaagaagaag ccattgaagc tgtcttgaat gatgatattg 5100
ttgatggtga tgttgttgtc gtacgttttg taggaccaaa gggcggtcct ggtatgcctg 5160
aaatgctttc cctttcatca atgattgttg gtaaagggca aggtgaaaaa gttgcccttc 5220

tgacagatgg ccgcttctca ggtggtactt atggtcttgt cgtgggtcat atcgctcctg 5280
aagcacaaga tggcggtcca atcgcctacc tgcaaacagg agacatagtc actattgacc 5340
aagacactaa ggaattacac tttgatatct ccgatgaaga gttaaaacat cgtcaagaga 5400
ccattgaatt gccaccgctc tattcacgcg gtatccttgg taaatatgct cacatcgttt 5460
cgtctgcttc taggggagcc gtaacagact tttggaagcc tgaagaaact ggcaaaaaat 5520
gttgtcctgg ttgctgtggt taagcggccg cgttaattca aattaattga tatagttttt 5580
taatgagtat tgaatctgtt tagaaataat ggaatattat ttttatttat ttatttatat 5640
tattggtcgg ctcttttctt ctgaaggtca atgacaaaat gatatgaagg aaataatgat 5700
ttctaaaatt ttacaacgta agatattttt acaaaagcct agctcatctt ttgtcatgca 5760
ctattttact cacgcttgaa attaacggcc agtccactgc ggagtcattt caaagtcatc 5820
ctaatcgatc tatcgttttt gatagctcat tttggagttc gcgattgtct tctgttattc 5880
acaactgttt taatttttat ttcattctgg aactcttcga gttctttgta aagtctttca 5940
tagtagctta ctttatcctc caacatattt aacttcatgt caatttcggc tcttaaattt 6000
tccacatcat caagttcaac atcatctttt aacttgaatt tattctctag ctcttccaac 6060
caagcctcat tgctccttga tttactggtg aaaagtgata cactttgcgc gcaatccagg 6120
tcaaaacttt cctgcaaaga attcaccaat ttctcgacat catagtacaa tttgttttgt 6180
tctcccatca caatttaata tacctgatgg attcttatga agcgctgggt aatggacgtg 6240
tcactctact tcgccttttt ccctactcct tttagtacgg aagacaatgc taataaataa 6300
gagggtaata ataatattat taatcggcaa aaaagattaa acgccaagcg tttaattatc 6360
agaaagcaaa cgtcgtacca atccttgaat gcttcccaat tgtatattaa gagtcatcac 6420
agcaacatat tcttgttatt aaattaatta ttattgattt ttgatattgt ataaaaaaac 6480
caaatatgta taaaaaaagt gaataaaaaa taccaagtat ggagaaatat attagaagtc 6540
tatacgttaa accacccggg ccccccctcg aggtcgacgg tatcgataag cttgatatcg 6600
aattcctgca gcccggggga tccactagtt ctagagcggc cgctctagaa ctagtaccac 6660
aggtgttgtc ctctgaggac ataaaataca caccgagatt catcaactca ttgctggagt 6720
tagcatatct acaattgggt gaaatgggga gcgatttgca ggcatttgct cggcatgccg 6780
gtagaggtgt ggtcaataag agcgacctca tgctatacct gagaaagcaa cctgacctac 6840
aggaaagagt tactcaagaa taagaatttt cgttttaaaa cctaagagtc actttaaaat 6900
ttgtatacac ttatttttt tataacttat ttaataataa aaatcataaa tcataagaaa 6960
ttcgcttact cttaattaat caggcagcgc ctgcgttcga gaggatgatc ttcatcgcct 7020
tctccttggc gccattgagg aatacctgat aggcgtgctc gatctcggcc agctcgaagc 7080
gatgggtaat catcttcttc aacggaagct tgtcggtcga ggcgaccttc atcagcatgg 7140
gcgtcgtgtt cgtgttcacc agtcccgtgg tgatcgtcag gttcttgatc cagagcttct 7200
gaatctcgaa gtcaaccttg acgccatgca cgccgacgtt ggcgatgtgc gcgccgggct 7260
tgacgatctc ctggcagatg tcccaagtcg ccggtatgcc caccgcctcg atcgcaacat 7320
cgactccctc tgccgcaatc ctatgcacgg cttcgacaac gttctccgtg ccggagttga 7380

tggtgtgcgt tgccccgagc tccttggcga gctggaggcg attctcgtcc atgtcgatca 7440
cgatgatggt cgagggggag tagaactggg cggtcaacag tacggacatg ccgacggggc 7500
ccgcgccgac aatagccacc gcatcgcccg gctggacatt cccatactgg acgccgattt 7560
cgtggccggt gggcaggatg tcgctcagca ggacggcgat ttcgtcgtca attgtctggg 7620
ggatcttgta gaggctgttg tcggcatgcg ggatgcggac gtattcggcc tgcacgccat 7680
cgatcatgta acccaggatc cacccgccgt cgcggcaatg ggagtaaagc tgcttcttgc 7740
agtagtcgca cgagccgcaa gaagtgacgc aggaaatcag gaccttgtcg cctttcttga 7800
actgcgtgac actctcgccc acttcctcga tgacgcctac cccttcatgg cccaggatgc 7860
gcccgtcggc gacctctgga ttcttgcctt tgtagatgcc gagatccgtg ccgcagatcg 7920
tggtcttcaa aacccgtact actacatccg tgggcttttg aagggtgggc ttgggcttgt 7980
cttcaagcga gatcttgtgg tcaccgtgat aaaccagagc tttcatcctc agctattgta 8040
atatgtgtgt ttgtttggat tattaagaag aataattaca aaaaaaatta caaaggaagg 8100
taattacaac agaattaaga aaggacaaga aggaggaaga gaatcagttc attatttctt 8160
ctttgttata taacaaaccc aagtagcgat ttggccatac attaaaagtt gagaaccacc 8220
ctccctggca acagccacaa ctcgttacca ttgttcatca cgatcatgaa actcgctgtc 8280
agctgaaatt tcacctcagt ggatctctct ttttattctt catcgttcca ctaacctttt 8340
tccatcagct ggcagggaac ggaaagtgga atcccattta gcgagcttcc tcttttcttc 8400
aagaaaagac gaagcttgtg tgtgggtgcg cgcgctagta tctttccaca ttaagaaata 8460
taccataaag gttacttaga catcactatg gctatatata tatatatata tatatatgta 8520
acttagcacc atcgcgcgtg catcactgca tgtgttaacc gaaaagtttg gcgaacactt 8580
caccgacacg gtcatttaga tctgtcgtct gcattgcacg tcccttagcc ttaaatccta 8640
ggcgggagca ttctcgtgta attgtgcagc ctgcgtagca actcaacata gcgtagtcta 8700
cccagttttt caagggttta tcgttagaag attctccctt ttcttcctgc tcacaaatct 8760
taaagtcata cattgcacga ctaaatgcaa gcatgcggat cccccgggct gcaggaattc 8820
gatatcaagc ttatcgatac cgtcgactgg ccattaatct ttcccatatt agatttcgcc 8880
aagccatgaa agttcaagaa aggtctttag acgaattacc cttcatttct caaactggcg 8940
tcaagggatc ctggtatggt tttatcgttt tatttctggt tcttatagca tcgttttgga 9000
cttctctgtt cccattaggc ggttcaggag ccagcgcaga atcattcttt gaaggatact 9060
tatcctttcc aattttgatt gtctgttacg ttggacataa actgtatact agaaattgga 9120
ctttgatggt gaaactagaa gatatggatc ttgataccgg cagaaaacaa gtagatttga 9180
ctcttcgtag ggaagaaatg aggattgagc gagaaacatt agcaaaaaga tccttcgtaa 9240
caagattttt acatttctgg tgttgaaggg aaagatatga gctatacagc ggaatttcca 9300
tatcactcag attttgttat ctaatttttt ccttcccacg tccgcgggaa tctgtgtata 9360
ttactgcatc tagatatatg ttatcttatc ttggcgcgta catttaattt tcaacgtatt 9420
ctataagaaa ttgcgggagt ttttttcatg tagatgatac tgactgcacg caaatatagg 9480
catgatttat aggcatgatt tgatggctgt accgatagga acgctaagag taacttcaga 9540

```
atcgttatcc tggcggaaaa aattcatttg taaactttaa aaaaaaaagc caatatcccc  9600
aaaattatta agagcgcctc cattattaac taaaatttca ctcagcatcc acaatgtatc  9660
aggtatctac tacagatatt acatgtggcg aaaaagacaa gaacaatgca atagcgcatc  9720
aagaaaaaac acaaagcttt caatcaatga atcgaaaatg tcattaaaat agtatataaa  9780
ttgaaactaa gtcataaagc tataaaaaga aaatttattt aaatgcaaga tttaaagtaa  9840
attcacggcc ctgcaggcct cagctcttgt tttgttctgc aaataactta cccatctttt  9900
tcaaaacttt aggtgcaccc tcctttgcta gaataagttc tatccaatac atcctatttg  9960
gatctgcttg agcttctttc atcacggata cgaattcatt ttctgttctc acaattttgg 10020
acacaactct gtcttccgtt gccccgaaac tttctggcag ttttgagtaa ttccacatag 10080
gaatgtcatt ataactctgg ttcggaccat gaatttccct ctcaaccgtg taaccatcgt 10140
tattaatgat aaagcagatt gggtttatct tctctctaat ggctagtcct aattcttgga 10200
cagtcagttg caatgatcca tctccgataa acaataaatg tctagattct ttatctgcaa 10260
tttggctgcc tagagctgcg gggaaagtgt atcctataga tccccacaag ggttgaccaa 10320
taaaatgtga tttcgatttc agaaatatag atgaggcacc gaagaaagaa gtgccttgtt 10380
cagccacgat cgtctcatta ctttgggtca aattttcgac agcttgccac agtctatctt 10440
gtgacaacag cgcgttagaa ggtacaaaat cttcttgctt tttatctatg tacttgcctt 10500
tatattcaat ttcggacaag tcaagaagag atgatatcag ggattcgaag tcgaaatttt 10560
ggattctttc gttgaaaatt ttaccttcat cgatattcaa ggaaatcatt ttattttcat 10620
taagatggtg agtaaatgca cccgtactag aatcggtaag ctttacaccc aacataagaa 10680
taaaatcagc agattccaca aattccttca agtttggctc tgacagagta ccgttgtaaa 10740
tccccaaaaa tgagggcaat gcttcatcaa cagatgattt accaaagttc aaagtagtaa 10800
taggtaactt agtctttgaa ataaactgag taacagtctt ctctaggccg aacgatataa 10860
tttcatggcc tgtgattaca attggtttct tggcattctt cagactttcc tgtattttgt 10920
tcagaatctc ttgatcagat gtattcgacg tggaattttc cttcttaaga ggcaaggatg 10980
gtttttcagc cttagcggca gctacatcta caggtaaatt gatgtaaacc ggctttcttt 11040
cctttagtaa ggcagacaac actctatcaa tttcaacagt tgcattctcg gctgtcaata 11100
aagtcctggc agcagtaacc ggttcgtgca tcttcataaa gtgcttgaaa tcaccatcag 11160
ccaacgtatg gtgaacaaac ttaccttcgt tctgcacttt cgaggtagga gatcccacga 11220
tctcaacaac aggcaggttc tcagcatagg agcccgctaa gccattaact gcggataatt 11280
cgccaacacc aaatgtagtc aagaatgccg cagccttttt cgttcttgcg tacccgtcgg 11340
ccatatagga ggcatttaac tcattagcat ttcccaccca tttcatatct ttgtgtgaaa 11400
taatttgatc tagaaattgc aaattgtagt cacctggtac tccgaatatt tcttctatac 11460
ctaattcgtg taatctgtcc aacagatagt cacctactgt atacattttg tttactagtt 11520
tatgtgtgtt tattcgaaac taagttcttg gtgttttaaa actaaaaaaa agactaacta 11580
taaaagtaga atttaagaag tttaagaaat agatttacag aattacaatc aatacctacc 11640
gtctttatat acttattagt caagtagggg aataatttca gggaactggt ttcaaccttt 11700
```

ttttcagct ttttccaaat cagagagagc agaaggtaat agaaggtgta agaaaatgag 11760
atagatacat gcgtgggtca attgccttgt gtcatcattt actccaggca ggttgcatca 11820
ctccattgag gttgtgcccg ttttttgcct gtttgtgccc ctgttctctg tagttgcgct 11880
aagagaatgg acctatgaac tgatggttgg tgaagaaaac aatattttgg tgctgggatt 11940
ctttttttt ctggatgcca gcttaaaaag cgggctccat tatatttagt ggatgccagg 12000
aataaactgt tcacccagac acctacgatg ttatatattc tgtgtaaccc gccccctatt 12060
ttgggcatgt acgggttaca gcagaattaa aaggctaatt ttttgactaa ataaagttag 12120
gaaaatcact actattaatt atttacgtat tctttgaaat ggcagtattg ataatgataa 12180
actcgaactg aaaaagcgtg ttttttattc aaaatgattc taactccctt acgtaatcaa 12240
ggaatctttt tgccttggcc tccgcgtcat taaacttctt gttgttgacg ctaacattca 12300
acgctagtat atattcgttt ttttcaggta agttcttttc aacgggtctt actgatgagg 12360
cagtcgcgtc tgaacctgtt aagaggtcaa atatgtcttc ttgaccgtac gtgtcttgca 12420
tgttattagc tttgggaatt tgcatcaagt cataggaaaa tttaaatctt ggctctcttg 12480
ggctcaaggt gacaaggtcc tcgaaaatag ggcgcgcccc accgcggtgg agctccagct 12540
tttgttccct ttagtgaggg ttaattgcgc gcttggcgta atcatggtca tagctgtttc 12600
ctgtgtgaaa ttgttatccg ctcacaattc cacacaacat acgagccgga agcataaagt 12660
gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg cgctcactgc 12720
ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg 12780
ggagaggcgg tttgcgtatt gggcgctctt ccgcttcctc gctcactgac tcgctgcgct 12840
cggtcgttcg gctgcggcga gcggtatcag ctcactcaaa ggcggtaata cggttatcca 12900
cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga 12960
accgtaaaaa ggccgcgttg ctggcgtttt tccataggct ccgccccccct gacgagcatc 13020
acaaaaatcg acgctcaagt cagaggtggc gaaacccgac aggactataa agataccagg 13080
cgtttccccc tggaagctcc ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat 13140
acctgtccgc ctttctccct tcgggaagcg tggcgctttc tcatagctca cgctgtaggt 13200
atctcagttc ggtgtaggtc gttcgctcca agctgggctg tgtgcacgaa ccccccgttc 13260
agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaacccg gtaagacacg 13320
acttatcgcc actggcagca gccactggta acaggattag cagagcgagg tatgtaggcg 13380
gtgctacaga gttcttgaag tggtggccta actacggcta cactagaaga acagtatttg 13440
gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc tcttgatccg 13500
gcaaacaaac caccgctggt agcggtggtt tttttgtttg caagcagcag attacgcgca 13560
gaaaaaaagg atctcaagaa gatcctttga tcttttctac ggggtctgac gctcagtgga 13620
acgaaaactc acgttaaggg attttggtca tgagattatc aaaaaggatc ttcacctaga 13680
tccttttaaa ttaaaaatga agttttaaat caatctaaag tatatatgag taaacttggt 13740
ctgacagtta ccaatgctta atcagtgagg cacctatctc agcgatctgt ctatttcgtt 13800
catccatagt tgcctgactc cccgtcgtgt agataactac gatacgggag ggcttaccat 13860

ctggccccag tgctgcaatg ataccgcgag acccacgctc accggctcca gatttatcag 13920
caataaacca gccagccgga agggccgagc gcagaagtgg tcctgcaact ttatccgcct 13980
ccatccagtc tattaattgt tgccgggaag ctagagtaag tagttcgcca gttaatagtt 14040
tgcgcaacgt tgttgccatt gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg 14100
cttcattcag ctccggttcc caacgatcaa ggcgagttac atgatccccc atgttgtgca 14160
aaaaagcggt tagctccttc ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt 14220
tatcactcat ggttatggca gcactgcata attctcttac tgtcatgcca tccgtaagat 14280
gcttttctgt gactggtgag tactcaacca agtcattctg agaatagtgt atgcggcgac 14340
cgagttgctc ttgcccggcg tcaatacggg ataataccgc gccacatagc agaactttaa 14400
aagtgctcat cattggaaaa cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt 14460
tgagatccag ttcgatgtaa cccactcgtg cacccaactg atcttcagca tcttttactt 14520
tcaccagcgt ttctgggtga gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa 14580
gggcgacacg gaaatgttga atactcatac tcttcctttt tcaatattat tgaagcattt 14640
atcagggtta ttgtctcatg agcggataca tatttgaatg tatttagaaa aataaacaaa 14700
taggggttcc gcgcacattt ccccgaaaag tgccacctga acgaagcatc tgtgcttcat 14760
tttgtagaac aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc 14820
atttttacag aacagaaatg caacgcgaaa gcgctatttt accaacgaag aatctgtgct 14880
tcatttttgt aaaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca aagaatctga 14940
gctgcatttt tacagaacag aaatgcaacg cgagagcgct attttaccaa caaagaatct 15000
atacttcttt tttgttctac aaaaatgcat cccgagagcg ctatttttct aacaaagcat 15060
cttagattac tttttttctc ctttgtgcgc tctataatgc agtctcttga taacttttg 15120
cactgtaggt ccgttaaggt tagaagaagg ctactttggt gtctattttc tcttccataa 15180
aaaaagcctg actccacttc ccgcgtttac tgattactag cgaagctgcg ggtgcatttt 15240
ttcaagataa aggcatcccc gattatattc tataccgatg tggattgcgc atactttgtg 15300
aacagaaagt gatagcgttg atgattcttc attggtcaga aaattatgaa cggtttcttc 15360
tattttgtct ctatatacta cgtataggaa atgtttacat tttcgtattg ttttcgattc 15420
actctatgaa tagttcttac tacaattttt ttgtct 15456

<210> 360
<211> 1559
<212> DNA
<213> Artificial sequence

<220>
<223> Template

<400> 360
gcattgcgga ttacgtattc taatgttcag taccgttcgt ataatgtatg ctatacgaag 60
ttatgcagat tgtactgaga gtgcaccata ccacctttc aattcatcat ttttttttta 120
ttctttttt tgatttcggt ttccttgaaa ttttttgat tcggtaatct ccgaacagaa 180
ggaagaacga aggaaggagc acagacttag attggtatat atacgcatat gtagtgttga 240

agaaacatga aattgcccag tattcttaac ccaactgcac agaacaaaaa cctgcaggaa 300
acgaagataa atcatgtcga aagctacata taaggaacgt gctgctactc atcctagtcc 360
tgttgctgcc aagctattta atatcatgca cgaaaagcaa acaaacttgt gtgcttcatt 420
ggatgttcgt accaccaagg aattactgga gttagttgaa gcattaggtc ccaaaatttg 480
tttactaaaa acacatgtgg atatcttgac tgatttttcc atggagggca cagttaagcc 540
gctaaaggca ttatccgcca agtacaattt tttactcttc gaagacagaa aatttgctga 600
cattggtaat acagtcaaat tgcagtactc tgcgggtgta tacagaatag cagaatgggc 660
agacattacg aatgcacacg gtgtggtggg cccaggtatt gttagcggtt tgaagcaggc 720
ggcagaagaa gtaacaaagg aacctagagg ccttttgatg ttagcagaat tgtcatgcaa 780
gggctcccta tctactggag aatatactaa gggtactgtt gacattgcga agagcgacaa 840
agattttgtt atcggcttta ttgctcaaag agacatgggt ggaagagatg aaggttacga 900
ttggttgatt atgacacccg gtgtgggttt agatgacaag ggagacgcat tgggtcaaca 960
gtatagaacc gtggatgatg tggtctctac aggatctgac attattattg ttggaagagg 1020
actatttgca aagggaaggg atgctaaggt agagggtgaa cgttacagaa aagcaggctg 1080
ggaagcatat ttgagaagat gcggccagca aaactaaaaa actgtattat aagtaaatgc 1140
atgtatacta aactcacaaa ttagagcttc aatttaatta tatcagttat taccctatgc 1200
ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggaaa ttgtaaacgt 1260
taatattttg ttaaaattcg cgttaaattt ttgttaaatc agctcatttt ttaaccaata 1320
ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag accgagatag ggttgagtgt 1380
tgttccagtt tggaacaaga gtccactatt aaagaacgtg gactccaacg tcaaagggcg 1440
aaaaaccgtc tatcagggcg atggcccact acgtgaacca tcaccctaat caagataact 1500
tcgtataatg tatgctatac gaacggtacc agtgatgata caacgagtta gccaaggtg 1559

<210> 361
<211> 1685
<212> DNA
<213> Artificial sequence

<220>
<223> Cassette

<400> 361
gtattttggt agattcaatt ctctttccct ttccttttcc ttcgctcccc ttccttatca 60
gcattgcgga ttacgtattc taatgttcag ataacttcgt atagcataca ttatacgaag 120
ttatgcagat tgtactgaga gtgcaccata ccacagcttt tcaattcaat tcatcatttt 180
ttttttattc tttttttttga tttcggtttc tttgaaattt ttttgattcg gtaatctccg 240
aacagaagga agaacgaagg aaggagcaca gacttagatt ggtatatata cgcatatgta 300
gtgttgaaga aacatgaaat tgcccagtat tcttaaccca actgcacaga acaaaaacct 360
gcaggaaacg aagataaatc atgtcgaaag ctacatataa ggaacgtgct gctactcatc 420
ctagtcctgt tgctgccaag ctatttaata tcatgcacga aaagcaaaca aacttgtgtg 480
cttcattgga tgttcgtacc accaaggaat tactggagtt agttgaagca ttaggtccca 540

aaatttgttt actaaaaaca catgtggata tcttgactga tttttccatg gagggcacag 600
ttaagccgct aaaggcatta tccgccaagt acaatttttt actcttcgaa gacagaaaat 660
ttgctgacat tggtaataca gtcaaattgc agtactctgc gggtgtatac agaatagcag 720
aatgggcaga cattacgaat gcacacggtg tggtgggccc aggtattgtt agcggtttga 780
agcaggcggc agaagaagta acaaaggaac ctagaggcct tttgatgtta gcagaattgt 840
catgcaaggg ctccctatct actggagaat atactaaggg tactgttgac attgcgaaga 900
gcgacaaaga ttttgttatc ggctttattg ctcaaagaga catgggtgga agagatgaag 960
gttacgattg gttgattatg acacccggtg tgggtttaga tgacaaggga gacgcattgg 1020
gtcaacagta tagaaccgtg gatgatgtgg tctctacagg atctgacatt attattgttg 1080
gaagaggact atttgcaaag ggaagggatg ctaaggtaga gggtgaacgt tacagaaaag 1140
caggctggga agcatatttg agaagatgcg gccagcaaaa ctaaaaaact gtattataag 1200
taaatgcatg tatactaaac tcacaaatta gagcttcaat ttaattatat cagttattac 1260
cctatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat caggaaattg 1320
taaacgttaa tattttgtta aaattcgcgt taaatttttg ttaaatcagc tcatttttta 1380
accaataggc cgaaatcggc aaaatccctt ataaatcaaa agaatagacc gagatagggt 1440
tgagtgttgt tccagtttgg aacaagagtc cactattaaa gaacgtggac tccaacgtca 1500
aagggcgaaa aaccgtctat cagggcgatg gcccactacg tgaaccatca ccctaatcaa 1560
gataacttcg tatagcatac attatacgaa gttatccagt gatgatacaa cgagttagcc 1620
aaggtgacac tctccccccc cctccccctc tgatctttcc tgttgcctct ttttccccca 1680
accaa 1685

<210> 362
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP591

<400> 362
cccattaata tactattgag a 21

<210> 363
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> primer N98SeqF1

<400> 363
cgtgttagtc acatcaggac 20

<210> 364
<211> 20
<212> DNA
<213> Artificial sequence

<220>

<223> primer N98SeqF4

<400> 364
ggtttctgtc tctggtgacg 20

<210> 365
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> primer N99SeqR2

<400> 365
catcgactgc attacgcaac tc 22

<210> 366
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> primer N1111

<400> 366
tatttgtatc gaggtgtcta gtcttctatt 30

<210> 367
<211> 30
<212> DNA
<213> Artificial sequence

<220>
<223> primer N1110

<400> 367
gcgatttaat ctctaattat tagttaaagt 30

<210> 368
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> BK468 Primer

<400> 368
gcctcgagtt ttaatgttac ttctcttgca gttaggga 38

<210> 369
<211> 7589
<212> DNA
<213> Artificial sequence

<220>
<223> pUC19-kan::pdc1::FBA-alsS::TRX1

<400> 369
tatttgtatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttggaaac 60

cacactgttt aaacagtgtt ccttaatcaa ggatacctct ttttttttcc ttggttccac 120

taattcatcg gttttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180

ttaagagaaa ttctccaaat ttgtaaagaa gcggacccag acttaaggcc gcccgcaaat 240

taaagccttc gagcgtccca aaaccttctc aagcaaggtt ttcagtataa tgttacatgc 300

gtacacgcgt ctgtacagaa aaaaaagaaa aatttgaaat ataaataacg ttcttaatac 360
taacataact ataaaaaaat aaatagggac ctagacttca ggttgtctaa ctccttcctt 420
ttcggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480
gattaattaa ctagagagct ttcgttttca tgagttcccc gaattctttc ggaagcttgt 540
cacttgctaa attaatgtta tcactgtagt caaccgggac atcgatgatg acaggacctt 600
cagcgttcat gccttgacgc agaacatctg ccagctggtc tggtgattct acgcgcaagc 660
cagttgctcc gaagctttcc gcatatttca cgatatcgat atttccgaaa tcgaccgcag 720
atgtacggtt atatttttc aattgctgga atgcaaccat gtcatatgtg ctgtcgttcc 780
atacaatgtg tacaattggt gcttttagtc gaactgctgt ctctaattcc attgctgaga 840
ataagaaacc gccgtcacca gagacagaaa ccacttttc tcccggtttc accaatgaag 900
cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg 960
ttaacggctc gtagctgcgg aaataacgtg acatccaaat ggcgtgcgaa ccgatatcgc 1020
aagttactgt aacatgatca tcgactgcat tacgcaactc tttaacgatt tcaagagggt 1080
gcgctctgtc tgatttccaa tctgcaggca cctgctcacc ttcatgcata tattgtttta 1140
aatcagaaag gattttctgc tcacgctctg caaattccac tttcacagca tcgtgttcga 1200
tatgattgat cgtggacgga atgtcaccga tcaattcaag atcaggctgg taagcatgat 1260
caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga 1320
atttcggatc atattcaatc gggtcatagc cgatcgtcag aacaacatct gcctgctcta 1380
gcagtaaatc gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatcctcta 1440
aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt 1500
tcaaaagctt gcgaaccgct ttaattgctt ccggtcttcc gcctttcatg ccgaccaaaa 1560
cgacaggaag ttttgctgtt tggatttttg ctatggccgc actgattgca tcatctgctg 1620
caggaccgag ttttggcgct gcaacagcac gcacgttttt cgtatttgtg acttcattca 1680
caacatcttg cggaaagctc acaaaagcgg ccccagcctg ccctgctgac gctatcctaa 1740
atgcatttgt aacagcttcc ggtatatttt ttacatcttg aacttctaca ctgtattttg 1800
taatcggctg gaatagcgcc gcattatcca aagattgatg tgtccgtttt aaacgatctg 1860
cacggatcac gtttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca 1920
ggcctgttgc caagttagag gcacccggtc ctgatgtgac taacacgact cccggttttc 1980
cagttaaacg gccgactgct tgggccatga atgctgcgtt ttgttcgtgc cgggcaacga 2040
taatttcagg tcctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa 2100
tgccaaatac atgtgtgaca ccttgctcca ctaagcaatc aacaacaagc tccgcccctc 2160
tgtttttcac aagggatttt tgttcttttg ttgcttttgt caacatcctc agctctagat 2220
ttgaatatgt attacttggt tatggttata tatgacaaaa gaaaaagaag aacagaagaa 2280
taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta 2340
taataaatca agtgcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga 2400
agaagaagga aaaaagtttt tgtgagggcg taattgaagc gatctgttga ttgtagattt 2460

```
ttttttttg aggagtcaaa gtcagaagag aacagacaaa tggtattaac catccaatac     2520
ttttttggag caacgctaag ctcatgcttt tccattggtt acgtgctcag ttgttagata     2580
tggaaagaga ggatgctcac ggcagcgtga ctccaattga gcccgaaaga gaggatgcca     2640
cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca     2700
ttatttacgt aatgacccac acatttttga gattttcaac tattacgtat tacgataatc     2760
ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaagtagg     2820
caattattta gtactgtcag tattgttatt catttcagat cttaagccta accaggccaa     2880
ttcaacagac tgtcggcaac ttcttgtctg gtctttccat ggtaagtgac agtgcagtaa     2940
taatatgaac caatttattt ttcgttacat aaaaatgctt ataaaacttt aactaataat     3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggct aactcgttgt     3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac     3120
cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca taactactta     3180
atagtcacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt     3240
ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc     3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac     3360
agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa     3420
tttttctttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc     3480
tgagagttcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt     3540
ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt     3600
agtactggtt tatatgggtt tttctgtata acttcttttt attttagttt gtttaatctt     3660
attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa ctcgagatgg     3720
gtaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat gctgatttat     3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt     3840
atgggaagcc cgatgcgcca gagttgtttc tgaaacatgg caaaggtagc gttgccaatg     3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct cttccgacca     3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca     4020
aaacagcatt ccaggtatta gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc     4080
tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct tttaacagcg     4140
atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg gttgatgcga     4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata     4260
agcttttgcc attctcaccg gattcagtcg tcactcatgg tgatttctca cttgataacc     4320
ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc ggaatcgcag     4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct ccttcattac     4440
agaaacggct tttcaaaaa tatggtattg ataatcctga tatgaataaa ttgcagtttc     4500
atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt tttctcttca     4560
tttataaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc tttttttca      4620
```

ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa 4680
caaaagaaaa agactcgaat atcaaacatt aaaaaaaaat aaaagaggtt atctgttttc 4740
ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt 4800
ttctctgttc gataacttcg tatagcatac attatacgaa gttatctgaa cattagaata 4860
cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct 4920
tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 4980
acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 5040
tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 5100
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 5160
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 5220
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 5280
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 5340
ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 5400
acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 5460
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 5520
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 5580
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 5640
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 5700
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 5760
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 5820
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 5880
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 5940
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 6000
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 6060
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 6120
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 6180
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6240
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 6300
gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 6360
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 6420
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 6480
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 6540
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 6600
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 6660
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 6720
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 6780

gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840

ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 6960

tccttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140

cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc 7200

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260

gcgcgtcagc gggtgttggc gggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320

ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 7380

accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg cgattaagtt 7500

gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560

tcggtacccg gggatccggc gcgccgttt 7589

<210> 370
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> primer N191

<400> 370
atccgcggat agatctccca ttaccgacat ttgggcgc 38

<210> 371
<211> 9593
<212> DNA
<213> Artificial sequence

<220>
<223> Plasmid pYZ090DalsS

<400> 371
ggccgcacct ggtaaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa 60

aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac 120

agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat 180

ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa 240

tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta 300

cccgacggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat 360

aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata 420

tacaaaatat cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta 480

gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtgaaca ccttatataa 540

cttagcccgt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag 600

acgtgataaa aatttacttt ctaactcttc tcacgctgcc cctatctgtt cttccgctct 660

accgtgagaa ataaagcatc gagtacggca gttcgctgtc actgaactaa aacaataagg 720
ctagttcgaa tgatgaactt gcttgctgtc aaacttctga gttgccgctg atgtgacact 780
gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca 840
atagagctca gtaggagtca gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg 900
taagttgtgc gcgtgcacat ttcgcccgtt cccgctcatc ttgcagcagg cggaaatttt 960
catcacgctg taggacgcaa aaaaaaaata attaatcgta caagaatctt ggaaaaaaaa 1020
ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acacccagta 1080
ttttcccttt ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta 1140
ttcctaggag ttatattttt ttaccctacc agcaatataa gtaaaaaact gtttaaacag 1200
tatggcagtt acaatgtatt atgaagatga tgtagaagta tcagcacttg ctggaaagca 1260
aattgcagta atcggttatg gttcacaagg acatgctcac gcacagaatt tgcgtgattc 1320
tggtcacaac gttatcattg gtgtgcgcca cggaaaatct tttgataaag caaaagaaga 1380
tggctttgaa acatttgaag taggagaagc agtagctaaa gctgatgtta ttatggtttt 1440
ggcaccagat gaacttcaac aatccattta tgaagaggac atcaaaccaa acttgaaagc 1500
aggttcagca cttggttttg ctcacggatt taatatccat tttggctata ttaaagtacc 1560
agaagacgtt gacgtcttta tggttgcgcc taaggctcca ggtcaccttg tccgtcggac 1620
ttatactgaa ggttttggta caccagcttt gtttgtttca caccaaaatg caagtggtca 1680
tgcgcgtgaa atcgcaatgg attgggccaa aggaattggt tgtgctcgag tgggaattat 1740
tgaaacaact tttaaagaag aaacagaaga agatttgttt ggagaacaag ctgttctatg 1800
tggaggtttg acagcacttg ttgaagccgg ttttgaaaca ctgacagaag ctggatacgc 1860
tggcgaattg gcttactttg aagttttgca cgaaatgaaa ttgattgttg acctcatgta 1920
tgaaggtggt tttactaaaa tgcgtcaatc catctcaaat actgctgagt ttggcgatta 1980
tgtgactggt ccacggatta ttactgacga agttaaaaag aatatgaagc ttgttttggc 2040
tgatattcaa tctggaaaat ttgctcaaga tttcgttgat gacttcaaag cggggcgtcc 2100
aaaattaata gcctatcgcg aagctgcaaa aaatcttgaa attgaaaaaa ttggggcaga 2160
gctacgtcaa gcaatgccat tcacacaatc tggtgatgac gatgccttta aaatctatca 2220
gtaaggccct gcaggccaga ggaaaataat atcaagtgct ggaaactttt tctcttggaa 2280
tttttgcaac atcaagtcat agtcaattga attgacccaa tttcacattt aagatttttt 2340
ttttttcatc cgacatacat ctgtacacta ggaagccctg tttttctgaa gcagcttcaa 2400
atatatatat tttttacata tttattatga ttcaatgaac aatctaatta aatcgaaaac 2460
aagaaccgaa acgcgaataa ataatttatt tagatggtga caagtgtata agtcctcatc 2520
gggacagcta cgatttctct ttcggttttg gctgagctac tggttgctgt gacgcagcgg 2580
cattagcgcg gcgttatgag ctaccctcgt ggcctgaaag atggcgggaa taaagcggaa 2640
ctaaaaatta ctgactgagc catattgagg tcaatttgtc aactcgtcaa gtcacgtttg 2700
gtggacggcc cctttccaac gaatcgtata tactaacatg cgcgcgcttc ctatatacac 2760
atatacatat atatatatat atatatgtgt gcgtgtatgt gtacacctgt atttaatttc 2820

cttactcgcg ggtttttctt ttttctcaat tcttggcttc ctctttctcg agcggaccgg 2880
atcctccgcg gtgccggcag atctatttaa atggcgcgcc gacgtcaggt ggcacttttc 2940
ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc 3000
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 3060
gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc cttcctgttt 3120
ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 3180
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 3240
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta 3300
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 3360
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 3420
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 3480
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 3540
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 3600
tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 3660
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 3720
cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 3780
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 3840
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 3900
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 3960
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 4020
aaatcccttа acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 4080
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 4140
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 4200
ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc 4260
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag 4320
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 4380
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 4440
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 4500
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 4560
cgagggagct tccagggggа aacgcctggt atctttatag tcctgtcggg tttcgccacc 4620
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 4680
ccagcaacgc ggcctttttа cggttcctgg ccttttgctg gccttttgct cacatgttct 4740
ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata 4800
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 4860
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 4920
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 4980

ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg 5040
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcttttc 5100
tttccaattt ttttttttc gtcattataa aaatcattac gaccgagatt cccgggtaat 5160
aactgatata attaaattga agctctaatt tgtgagttta gtatacatgc atttacttat 5220
aatacagttt tttagttttg ctggccgcat cttctcaaat atgcttccca gcctgctttt 5280
ctgtaacgtt caccctctac cttagcatcc cttccctttg caaatagtcc tcttccaaca 5340
ataataatgt cagatcctgt agagaccaca tcatccacgg ttctatactg ttgacccaat 5400
gcgtctccct tgtcatctaa acccacaccg ggtgtcataa tcaaccaatc gtaaccttca 5460
tctcttccac ccatgtctct ttgagcaata aagccgataa caaaatcttt gtcgctcttc 5520
gcaatgtcaa cagtaccctt agtatattct ccagtagata gggagccctt gcatgacaat 5580
tctgctaaca tcaaaaggcc tctaggttcc tttgttactt cttctgccgc ctgcttcaaa 5640
ccgctaacaa tacctgggcc caccacaccg tgtgcattcg taatgtctgc ccattctgct 5700
attctgtata cacccgcaga gtactgcaat ttgactgtat taccaatgtc agcaaatttt 5760
ctgtcttcga agagtaaaaa attgtacttg gcggataatg cctttagcgg cttaactgtg 5820
ccctccatgg aaaaatcagt caagatatcc acatgtgttt ttagtaaaca aattttggga 5880
cctaatgctt caactaactc cagtaattcc ttggtggtac gaacatccaa tgaagcacac 5940
aagtttgttt gcttttcgtg catgatatta aatagcttgg cagcaacagg actaggatga 6000
gtagcagcac gttccttata tgtagctttc gacatgattt atcttcgttt cctgcaggtt 6060
tttgttctgt gcagttgggt taagaatact gggcaatttc atgtttcttc aacactacat 6120
atgcgtatat ataccaatct aagtctgtgc tccttccttc gttcttcctt ctgttcggag 6180
attaccgaat caaaaaaatt tcaaggaaac cgaaatcaaa aaaaagaata aaaaaaaaat 6240
gatgaattga aaagcttgca tgcctgcagg tcgactctag tatactccgt ctactgtacg 6300
atacacttcc gctcaggtcc ttgtccttta acgaggcctt accactcttt tgttactcta 6360
ttgatccagc tcagcaaagg cagtgtgatc taagattcta tcttcgcgat gtagtaaaac 6420
tagctagacc gagaaagaga ctagaaatgc aaaaggcact tctacaatgg ctgccatcat 6480
tattatccga tgtgacgctg catttttttt ttttttttt ttttttttt ttttttttt 6540
ttttttttt ttttgtacaa atatcataaa aaaagagaat ctttttaagc aaggattttc 6600
ttaacttctt cggcgacagc atcaccgact tcggtggtac tgttggaacc acctaaatca 6660
ccagttctga tacctgcatc caaaaccttt ttaactgcat cttcaatggc tttaccttct 6720
tcaggcaagt tcaatgacaa tttcaacatc attgcagcag acaagatagt ggcgataggg 6780
ttgaccttat tctttggcaa atctggagcg gaaccatggc atggttcgta caaaccaaat 6840
gcggtgttct tgtctggcaa agaggccaag gacgcagatg gcaacaaacc caaggagcct 6900
gggataacgg aggcttcatc ggagatgata tcaccaaaca tgttgctggt gattataata 6960
ccatttaggt gggttgggtt cttaactagg atcatggcgg cagaatcaat caattgatgt 7020
tgaactttca atgtagggaa ttcgttcttg atggtttcct ccacagtttt tctccataat 7080
cttgaagagg ccaaaacatt agctttatcc aaggaccaaa taggcaatgg tggctcatgt 7140

tgtagggcca tgaaagcggc cattcttgtg attctttgca cttctggaac ggtgtattgt 7200
tcactatccc aagcgacacc atcaccatcg tcttccttc tcttaccaaa gtaaatacct 7260
cccactaatt ctctaacaac aacgaagtca gtacctttag caaattgtgg cttgattgga 7320
gataagtcta aaagagagtc ggatgcaaag ttacatggtc ttaagttggc gtacaattga 7380
agttctttac ggatttttag taaaccttgt tcaggtctaa cactaccggt accccattta 7440
ggaccaccca cagcacctaa caaaacggca tcagccttct tggaggcttc cagcgcctca 7500
tctggaagtg gaacacctgt agcatcgata gcagcaccac caattaaatg attttcgaaa 7560
tcgaacttga cattggaacg aacatcagaa atagctttaa gaaccttaat ggcttcggct 7620
gtgatttctt gaccaacgtg gtcacctggc aaaacgacga tcttcttagg ggcagacatt 7680
acaatggtat atccttgaaa tatatataaa aaaaaaaaaa aaaaaaaaaa aaaaaaatgc 7740
agcttctcaa tgatattcga atacgctttg aggagataca gcctaatatc cgacaaactg 7800
ttttacagat ttacgatcgt acttgttacc catcattgaa ttttgaacat ccgaacctgg 7860
gagttttccc tgaaacagat agtatatttg aacctgtata ataatatata gtctagcgct 7920
ttacggaaga caatgtatgt atttcggttc ctggagaaac tattgcatct attgcatagg 7980
taatcttgca cgtcgcatcc ccggttcatt ttctgcgttt ccatcttgca cttcaatagc 8040
atatctttgt taacgaagca tctgtgcttc attttgtaga acaaaaatgc aacgcgagag 8100
cgctaatttt tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga 8160
aagcgctatt ttaccaacga agaatctgtg cttcattttt gtaaaacaaa aatgcaacgc 8220
gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa 8280
cgcgagagcg ctattttacc aacaaagaat ctatacttct tttttgttct acaaaaatgc 8340
atcccgagag cgctatttttt ctaacaaagc atcttagatt actttttttc tcctttgtgc 8400
gctctataat gcagtctctt gataactttt tgcactgtag gtccgttaag gttagaagaa 8460
ggctactttg gtgtctattt tctcttccat aaaaaaagcc tgactccact tcccgcgttt 8520
actgattact agcgaagctg cgggtgcatt ttttcaagat aaaggcatcc ccgattatat 8580
tctataccga tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct 8640
tcattggtca gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg 8700
aaatgtttac attttcgtat tgttttcgat tcactctatg aatagttctt actacaattt 8760
ttttgtctaa agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg 8820
caagttcaag gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata 8880
tagcaaagag atacttttga gcaatgtttg tggaagcggt attcgcaata ttttagtagc 8940
tcgttacagt ccggtgcgtt tttggttttt tgaaagtgcg tcttcagagc gcttttggtt 9000
ttcaaaagcg ctctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact 9060
tcaaagcgtt tccgaaaacg agcgcttccg aaaatgcaac gcgagctgcg cacatacagc 9120
tcactgttca cgtcgcacct atatctgcgt gttgcctgta tatatatata catgagaaga 9180
acggcatagt gcgtgtttat gcttaaatgc gtacttatat gcgtctattt atgtaggatg 9240
aaaggtagtc tagtacctcc tgtgatatta tcccattcca tgcggggtat cgtatgcttc 9300

cttcagcact acccttagc tgttctatat gctgccactc ctcaattgga ttagtctcat 9360

ccttcaatgc tatcatttcc tttgatattg gatcatatgc atagtaccga gaaactagag 9420

gatctcccat taccgacatt tgggcgctat acgtgcatat gttcatgtat gtatctgtat 9480

ttaaaacact tttgtattat ttttcctcat atatgtgtat aggtttatac ggatgattta 9540

attattactt caccaccctt tatttcaggc tgatatctta gccttgttac tag 9593

<210> 372
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> primer oBP556

<400> 372
tattttcgag gaccttgtca cc 22

<210> 373
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> oBP561

<400> 373
tggccattaa tctttcccat attag 25

<210> 374
<211> 13114
<212> DNA
<213> artificial sequence

<220>
<223> constructed plasmid

<400> 374
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240

gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300

ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360

ttttttttt ccacctagcg gatgactctt ttttttttctt agcgattggc attatcacat 420

aatgaattat acattatata aagtaatgtg atttcttcga agaatatact aaaaaatgag 480

caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca 540

aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600

tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660

attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat 720

tccggctggt cgctaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780

actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt 840

ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcggtg 900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020
attaccctcc acgttgattg tctgcgaggc aagaatgatc atcaccgtag tgagagtgcg 1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgat 1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg 1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg 1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1440
aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 1620
aaaaccgtct atcagggcga tggcccacta cgtggccggc ttcacatacg ttgcatacgt 1680
cgatatagat aataatgata atgacagcag gattatcgta atacgtaata gctgaaaatc 1740
tcaaaaatgt gtgggtcatt acgtaaataa tgataggaat gggattcttc tatttttcct 1800
ttttccattc tagcagccgt cgggaaaacg tggcatcctc tctttcgggc tcaattggag 1860
tcacgctgcc gtgagcatcc tctctttcca tatctaacaa ctgagcacgt aaccaatgga 1920
aaagcatgag cttagcgttg ctccaaaaaa gtattggatg gttaatacca tttgtctgtt 1980
ctcttctgac tttgactcct caaaaaaaaa aatctacaat caacagatcg cttcaattac 2040
gccctcacaa aaacttttt ccttcttctt cgcccacgtt aaattttatc cctcatgttg 2100
tctaacggat ttctgcactt gatttattat aaaaagacaa agacataata cttctctatc 2160
aatttcagtt attgttcttc cttgcgttat tcttctgttc ttctttttct tttgtcatat 2220
ataaccataa ccaagtaata catattcaaa cacgtgagta tgactgacaa aaaaactctt 2280
aaagacttaa gaaatcgtag ttctgtttac gattcaatgg ttaaatcacc taatcgtgct 2340
atgttgcgtg caactggtat gcaagatgaa gactttgaaa aacctatcgt cggtgtcatt 2400
tcaacttggg ctgaaaacac accttgtaat atccacttac atgactttgg taaactagcc 2460
aaagtcggtg ttaaggaagc tggtgcttgg ccagttcagt tcggaacaat cacggtttct 2520
gatggaatcg ccatgggaac ccaaggaatg cgtttctcct tgacatctcg tgatattatt 2580
gcagattcta ttgaagcagc catgggaggt cataatgcgg atgcttttgt agccattggc 2640
ggttgtgata aaaacatgcc cggttctgtt atcgctatgg ctaacatgga tatcccagcc 2700
atttttgctt acggcggaac aattgcacct ggtaatttag acggcaaaga tatcgattta 2760
gtctctgtct ttgaaggtgt cggccattgg aaccacggcg atatgaccaa agaagaagtt 2820
aaagctttgg aatgtaatgc ttgtcccggt cctggaggct gcggtggtat gtatactgct 2880
aacacaatgg cgacagctat tgaagttttg ggacttagcc ttccgggttc atcttctcac 2940
ccggctgaat ccgcagaaaa gaaagcagat attgaagaag ctggtcgcgc tgttgtcaaa 3000

atgctcgaaa tgggcttaaa accttctgac attttaacgc gtgaagcttt tgaagatgct 3060
attactgtaa ctatggctct gggaggttca accaactcaa cccttcacct cttagctatt 3120
gcccatgctg ctaatgtgga attgacactt gatgatttca atactttcca agaaaaagtt 3180
cctcatttgg ctgatttgaa accttctggt caatatgtat tccaagacct ttacaaggtc 3240
ggaggggtac cagcagttat gaaatatctc cttaaaaatg gcttccttca tggtgaccgt 3300
atcacttgta ctggcaaaac agtcgctgaa aatttgaagg cttttgatga tttaacacct 3360
ggtcaaaagg ttattatgcc gcttgaaaat cctaaacgtg aagatggtcc gctcattatt 3420
ctccatggta acttggctcc agacggtgcc gttgccaaag tttctggtgt aaaagtgcgt 3480
cgtcatgtcg gtcctgctaa ggtctttaat tctgaagaag aagccattga agctgtcttg 3540
aatgatgata ttgttgatgg tgatgttgtt gtcgtacgtt ttgtaggacc aaagggcggt 3600
cctggtatgc ctgaaatgct ttccctttca tcaatgattg ttggtaaagg gcaaggtgaa 3660
aaagttgccc ttctgacaga tggccgcttc tcaggtggta cttatggtct tgtcgtgggt 3720
catatcgctc ctgaagcaca agatggcggt ccaatcgcct acctgcaaac aggagacata 3780
gtcactattg accaagacac taaggaatta cactttgata tctccgatga agagttaaaa 3840
catcgtcaag agaccattga attgccaccg ctctattcac gcggtatcct tggtaaatat 3900
gctcacatcg tttcgtctgc ttctagggga gccgtaacag acttttggaa gcctgaagaa 3960
actggcaaaa aatgttgtcc tggttgctgt ggttaagcgg ccgcgttaat tcaaattaat 4020
tgatatagtt ttttaatgag tattgaatct gtttagaaat aatggaatat tatttttatt 4080
tatttattta tattattggt cggctctttt cttctgaagg tcaatgacaa aatgatatga 4140
aggaaataat gatttctaaa attttacaac gtaagatatt tttacaaaag cctagctcat 4200
cttttgtcat gcactatttt actcacgctt gaaattaacg gccagtccac tgcggagtca 4260
tttcaaagtc atcctaatcg atctatcgtt tttgatagct cattttggag ttcgcgagga 4320
tccactagtt ctagagcggc cgctctagaa ctagtaccac aggtgttgtc ctctgaggac 4380
ataaaataca caccgagatt catcaactca ttgctggagt tagcatatct acaattgggt 4440
gaaatgggga gcgatttgca ggcatttgct cggcatgccg gtagaggtgt ggtcaataag 4500
agcgacctca tgctatacct gagaaagcaa cctgacctac aggaaagagt tactcaagaa 4560
taagaatttt cgttttaaaa cctaagagtc actttaaaat ttgtatacac ttatttttt 4620
tataacttat ttaataataa aaatcataaa tcataagaaa ttcgcttact cttaattaat 4680
caaaaagtta aaattgtacg aatagattca ccacttctta acaaatcaaa cccttcattg 4740
attttctcga atggcaatac atgtgtaatt aaaggatcaa gagcaaactt cttcgccata 4800
aagtcggcaa caagttttgg aacactatcc ttgctcttaa aaccgccaaa tatagctccc 4860
ttccatgtac gaccgcttag caacagcata ggattcatcg acaaattttg tgaatcagga 4920
ggaacaccta cgatcacact gactccatat gcctcttgac agcaggacaa cgcagttacc 4980
atagtatcaa gacggcctat aacttcaaaa gagaaatcaa ctccaccgtt tgacatttca 5040
gtaaggactt cttgtattgg tttcttataa tcttgagggt taacacattc agtagccccg 5100
acctccttag cttttgcaaa tttgtcctta ttgatgtcta cacctataat cctcgctgcg 5160

cctgcagctt tacaccccat aataacgctt agtcctactc ctcctaaacc gaatactgca 5220
caagtcgaac cctgtgtaac ctttgcaact ttaactgcgg aaccgtaacc ggtggaaaat 5280
ccgcacccta tcaagcaaac tttttccagt ggtgaagctg catcgatttt agcgacagat 5340
atctcgtcca ccactgtgta ttgggaaaat gtagaagtac caaggaaatg gtgtataggt 5400
ttccctctgc atgtaaatct gcttgtacca tcctgcatag tacctctagg catagacaaa 5460
tcatttttaa ggcagaaatt accctcagga tgtttgcaga ctctacactt accacattga 5520
ggagtgaaca gtgggatcac tttatcacca ggacgaacag tggtaacacc ttcacctatg 5580
gattcaacga ttccggcagc ctcgtgtccc gcgattactg gcaaaggagt aactagagtg 5640
ccactcacca catggtcgtc ggatctacag attccggtgg caaccatctt gattctaacc 5700
tcgtgtgctt ttggtggcgc tacttctact tcttctatgc taaacggctt tttctcttcc 5760
cacaaaactg ccgctttaca cttaataact ttaccggctg ttgacatcct cagctagcta 5820
ttgtaatatg tgtgtttgtt tggattatta agaagaataa ttacaaaaaa aattacaaag 5880
gaaggtaatt acaacagaat taagaaagga caagaaggag gaagagaatc agttcattat 5940
ttcttctttg ttatataaca aacccaagta gcgatttggc catacattaa aagttgagaa 6000
ccaccctccc tggcaacagc cacaactcgt taccattgtt catcacgatc atgaaactcg 6060
ccgtcagctg aaatttcacc tcagtggatc tctcttttta ttcttcatcg ttccactaac 6120
ctttttccat cagctggcag ggaacggaaa gtggaatccc atttagcgag cttcctcttt 6180
tcttcaagaa aagacgaagc ttgtgtgtgg gtgcgcgcgc tagtatcttt ccacattaag 6240
aaatatacca taaaggttac ttagacatca ctatggctat atatatatat atatatatat 6300
gtaacttagc accatcgcgc gtgcatcact gcatgtgtta accgaaaagt ttggcgaaca 6360
cttcaccgac acggtcattt agatctgtcg tctgcattgc acgtccctta gccttaaatc 6420
ctaggcggga gcattctcgt gtaattgtgc agcctgcgta gcaactcaac atagcgtagt 6480
ctacccagtt tttcaagggt ttatcgttag aagattctcc cttttcttcc tgctcacaaa 6540
tcttaaagtc atacattgca cgactaaatg caagcgacgt cagggaaaga tatgagctat 6600
acagcggaat ttccatatca ctcagatttt gttatctaat tttttccttc ccacgtccgc 6660
gggaatctgt gtatattact gcatctagat atatgttatc ttatcttggc gcgtacattt 6720
aattttcaac gtattctata agaaattgcg ggagtttttt tcatgtagat gatactgact 6780
gcacgcaaat ataggcatga tttataggca tgatttgatg gctgtaccga taggaacgct 6840
aagagtaact tcagaatcgt tatcctggcg gaaaaaattc atttgtaaac tttaaaaaaa 6900
aaagccaata tccccaaaat tattaagagc gcctccatta ttaactaaaa tttcactcag 6960
catccacaat gtatcaggta tctactacag atattacatg tggcgaaaaa gacaagaaca 7020
atgcaatagc gcatcaagaa aaaacacaaa gctttcaatc aatgaatcga aaatgtcatt 7080
aaaatagtat ataaattgaa actaagtcat aaagctataa aaagaaaatt tatttaaatg 7140
caagatttaa agtaaattca cggccctgca ggcctcagct cttgttttgt tctgcaaata 7200
acttacccat ctttttcaaa actttaggtg caccctcctt tgctagaata agttctatcc 7260
aatacatcct atttggatct gcttgagctt ctttcatcac ggatacgaat tcattttctg 7320

ttctcacaat tttggacaca actctgtctt ccgttgcccc gaaactttct ggcagttttg 7380
agtaattcca cataggaatg tcattataac tctggttcgg accatgaatt tccctctcaa 7440
ccgtgtaacc atcgttatta atgataaagc agattgggtt tatcttctct ctaatggcta 7500
gtcctaattc ttggacagtc agttgcaatg atccatctcc gataaacaat aaatgtctag 7560
attctttatc tgcaatttgg ctgcctagag ctgcggggaa agtgtatcct atagatcccc 7620
acaagggttg accaataaaa tgtgatttcg atttcagaaa tatagatgag gcaccgaaga 7680
aagaagtgcc ttgttcagcc acgatcgtct cattactttg ggtcaaattt tcgacagctt 7740
gccacagtct atcttgtgac aacagcgcgt tagaaggtac aaaatcttct tgctttttat 7800
ctatgtactt gcctttatat tcaatttcgg acaagtcaag aagagatgat atcagggatt 7860
cgaagtcgaa attttggatt ctttcgttga aaattttacc ttcatcgata ttcaaggaaa 7920
tcattttatt ttcattaaga tggtgagtaa atgcacccgt actagaatcg gtaagcttta 7980
cacccaacat aagaataaaa tcagcagatt ccacaaattc cttcaagttt ggctctgaca 8040
gagtaccgtt gtaaatcccc aaaaatgagg gcaatgcttc atcaacagat gatttaccaa 8100
agttcaaagt agtaataggt aacttagtct ttgaaataaa ctgagtaaca gtcttctcta 8160
ggccgaacga tataatttca tggcctgtga ttacaattgg tttcttggca ttcttcagac 8220
tttcctgtat tttgttcaga atctcttgat cagatgtatt cgacgtggaa ttttccttct 8280
taagaggcaa ggatggtttt tcagccttag cggcagctac atctacaggt aaattgatgt 8340
aaaccggctt tctttccttt agtaaggcag acaacactct atcaatttca acagttgcat 8400
tctcggctgt caataaagtc ctggcagcag taaccggttc gtgcatcttc ataaagtgct 8460
tgaaatcacc atcagccaac gtatggtgaa caaacttacc ttcgttctgc actttcgagg 8520
taggagatcc cacgatctca acaacaggca ggttctcagc ataggagccc gctaagccat 8580
taactgcgga taattcgcca acaccaaatg tagtcaagaa tgccgcagcc tttttcgttc 8640
ttgcgtaccc gtcggccata taggaggcat ttaactcatt agcatttccc acccatttca 8700
tatctttgtg tgaaataatt tgatctagaa attgcaaatt gtagtcacct ggtactccga 8760
atatttcttc tatacctaat tcgtgtaatc tgtccaacag atagtcacct actgtataca 8820
tgtttaaact ttgtttacta gtttatgtgt gtttattcga aactaagttc ttggtgtttt 8880
aaaactaaaa aaaagactaa ctataaaagt agaatttaag aagtttaaga aatagattta 8940
cagaattaca atcaatacct accgtcttta tatacttatt agtcaagtag gggaataatt 9000
tcagggaact ggtttcaacc tttttttca gcttttcca aatcagagag agcagaaggt 9060
aatagaaggt gtaagaaaat gagatagata catgcgtggg tcaattgcct tgtgtcatca 9120
tttactccag gcaggttgca tcactccatt gaggttgtgc ccgtttttg cctgtttgtg 9180
cccctgttct ctgtagttgc gctaagagaa tggacctatg aactgatggt tggtgaagaa 9240
aacaatattt tggtgctggg attctttttt tttctggatg ccagcttaaa aagcgggctc 9300
cattatattt agtggatgcc aggaataaac tgttcaccca gacacctacg atgttatata 9360
ttctgtgtaa cccgcccccc attttgggca tgtacgggtt acagcagaat taaaaggcta 9420
atttttgac taaataaagt taggaaaatc actactatta attatttacg tattctttga 9480

aatggcagta ttggagctcc agcttttgtt ccctttagtg agggttaatt gcgcgcttgg 9540
cgtaatcatg gtcatagctg tttcctgtgt gaaattgtta tccgctcaca attccacaca 9600
acatacgagc cggaagcata aagtgtaaag cctggggtgc ctaatgagtg agctaactca 9660
cattaattgc gttgcgctca ctgcccgctt tccagtcggg aaacctgtcg tgccagctgc 9720
attaatgaat cggccaacgc gcggggagag gcggtttgcg tattgggcgc tcttccgctt 9780
cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact 9840
caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatgtgag 9900
caaaaggcca gcaaaaggcc aggaaccgta aaaaggccgc gttgctggcg tttttccata 9960
ggctccgccc ccctgacgag catcacaaaa atcgacgctc aagtcagagg tggcgaaacc 10020
cgacaggact ataaagatac caggcgtttc cccctggaag ctccctcgtg cgctctcctg 10080
ttccgaccct gccgcttacc ggatacctgt ccgcctttct cccttcggga agcgtggcgc 10140
tttctcatag ctcacgctgt aggtatctca gttcggtgta ggtcgttcgc tccaagctgg 10200
gctgtgtgca cgaaccccc gttcagcccg accgctgcgc cttatccggt aactatcgtc 10260
ttgagtccaa cccggtaaga cacgacttat cgccactggc agcagccact ggtaacagga 10320
ttagcagagc gaggtatgta ggcggtgcta cagagttctt gaagtggtgg cctaactacg 10380
gctacactag aagaacagta tttggtatct gcgctctgct gaagccagtt accttcggaa 10440
aaagagttgg tagctcttga tccggcaaac aaaccaccgc tggtagcggt ggtttttttg 10500
tttgcaagca gcagattacg cgcagaaaaa aaggatctca agaagatcct ttgatctttt 10560
ctacggggtc tgacgctcag tggaacgaaa actcacgtta agggattttg gtcatgagat 10620
tatcaaaaag gatcttcacc tagatccttt taaattaaaa atgaagtttt aaatcaatct 10680
aaagtatata tgagtaaact tggtctgaca gttaccaatg cttaatcagt gaggcaccta 10740
tctcagcgat ctgtctattt cgttcatcca tagttgcctg actccccgtc gtgtagataa 10800
ctacgatacg ggagggctta ccatctggcc ccagtgctgc aatgataccg cgagacccac 10860
gctcaccggc tccagattta tcagcaataa accagccagc cggaagggcc gagcgcagaa 10920
gtggtcctgc aactttatcc gcctccatcc agtctattaa ttgttgccgg gaagctagag 10980
taagtagttc gccagttaat agtttgcgca acgttgttgc cattgctaca ggcatcgtgg 11040
tgtcacgctc gtcgtttggt atggcttcat tcagctccgg ttcccaacga tcaaggcgag 11100
ttacatgatc ccccatgttg tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg 11160
tcagaagtaa gttggccgca gtgttatcac tcatggttat ggcagcactg cataattctc 11220
ttactgtcat gccatccgta agatgctttt ctgtgactgg tgagtactca accaagtcat 11280
tctgagaata gtgtatgcgg cgaccgagtt gctcttgccc ggcgtcaata cgggataata 11340
ccgcgccaca tagcagaact ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa 11400
aactctcaag gatcttaccg ctgttgagat ccagttcgat gtaacccact cgtgcaccca 11460
actgatcttc agcatctttt actttcacca gcgtttctgg gtgagcaaaa acaggaaggc 11520
aaaatgccgc aaaaaaggga ataagggcga cacggaaatg ttgaatactc atactcttcc 11580
tttttcaata ttattgaagc atttatcagg gttattgtct catgagcgga tacatatttg 11640

aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 11700
ctgaacgaag catctgtgct tcattttgta gaacaaaaat gcaacgcgag agcgctaatt 11760
tttcaaacaa agaatctgag ctgcattttt acagaacaga aatgcaacgc gaaagcgcta 11820
ttttaccaac gaagaatctg tgcttcattt ttgtaaaaca aaaatgcaac gcgagagcgc 11880
taatttttca aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgagag 11940
cgctatttta ccaacaaaga atctatactt ctttttttgtt ctacaaaaat gcatcccgag 12000
agcgctattt ttctaacaaa gcatcttaga ttacttttttt tctcctttgt gcgctctata 12060
atgcagtctc ttgataactt tttgcactgt aggtccgtta aggttagaag aaggctactt 12120
tggtgtctat tttctcttcc ataaaaaaag cctgactcca cttcccgcgt ttactgatta 12180
ctagcgaagc tgcgggtgca ttttttcaag ataaaggcat ccccgattat attctatacc 12240
gatgtggatt gcgcatactt tgtgaacaga aagtgatagc gttgatgatt cttcattggt 12300
cagaaaatta tgaacggttt cttctatttt gtctctatat actacgtata ggaaatgttt 12360
acattttcgt attgttttcg attcactcta tgaatagttc ttactacaat ttttttgtct 12420
aaagagtaat actagagata aacataaaaa atgtagaggt cgagtttaga tgcaagttca 12480
aggagcgaaa ggtggatggg taggttatat agggatatag cacagagata tatagcaaag 12540
agatactttt gagcaatgtt tgtggaagcg gtattcgcaa tattttagta gctcgttaca 12600
gtccggtgcg tttttggttt tttgaaagtg cgtcttcaga gcgcttttgg ttttcaaaag 12660
cgctctgaag ttcctatact ttctagagaa taggaacttc ggaataggaa cttcaaagcg 12720
tttccgaaaa cgagcgcttc cgaaaatgca acgcgagctg cgcacataca gctcactgtt 12780
cacgtcgcac ctatatctgc gtgttgcctg tatatatata tacatgagaa gaacggcata 12840
gtgcgtgttt atgcttaaat gcgtacttat atgcgtctat ttatgtagga tgaaaggtag 12900
tctagtacct cctgtgatat tatcccattc catgcggggt atcgtatgct tccttcagca 12960
ctacccttta gctgttctat atgctgccac tcctcaattg gattagtctc atccttcaat 13020
gctatcattt cctttgatat tggatcatac taagaaacca ttattatcat gacattaacc 13080
tataaaaata ggcgtatcac gaggcccttt cgtc 13114

<210> 375
<211> 987
<212> DNA
<213> Enterococcus gallinarum

<400> 375
atggttaagg tctactacaa cgaagttact gacaacttgg gtgacaagac tatcgctatc 60
atcggttacg gttctcaagg tcacgcccac gctcaaaact tgagagacac cggtcataat 120
gttgtcatcg gtatcagagc tggtaaatct gccgaaaacg ctaagaacga tggtttcact 180
gttttgccag tctccgaagc tgttcaacag gccaacgttg tcatgatctt ggctccagac 240
gaaattcagg gtgaattgta cgaaaaggaa atcgctccaa acttgtctgc tggtgacgct 300
ttagccttcg ctcacggttt caacatccat ttcgaagtca ttaccccacc aaaagaagtt 360
gacgttttct tggttgctcc aaagggtcca ggtcacttgg tcagacgtac cttctctgaa 420

ggtttcgccg ttccagcttt gttcgccgtc taccaagatg ctaccggtaa cgcccaggaa 480

actgctttgt cctacgccaa gggtatcggt gctaccagag ttggtgtctt ggaaaccacc 540

ttcaaggaag aaaccgaaac cgacttgttc ggtgaacaag ctgttttgtg tggtggtctg 600

acttctatga ttgaaaccgg tttcgaaact ctggtcgaag ctggttacca accagaattg 660

gcctacttcg aagtttgcca cgaaatgaag ttgatcgtcg atttgatcta cgaaggtggt 720

ttcgctaaga tgagagactc tatctccaac accgccgaat acggtgacta cgtttctggt 780

tctagaatca ttaccgaaca aactaaggct aacatgaaaa atgtcttgaa ggacatccag 840

aacggtacct ttgctaaggg tttcattgct gataaccaag ccggtttccc agaattcaag 900

aagatgagag aacagaacgg taatcaccca atcgaaaagg tcggttctga attgagaaag 960

atgatgccat tcgttactaa ggactga 987

<210> 376
<211> 328
<212> PRT
<213> Enterococcus gallinarum

<400> 376

Met Val Lys Val Tyr Tyr Asn Glu Val Thr Asp Asn Leu Gly Asp Lys
1 5 10 15

Thr Ile Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala His Ala Gln
20 25 30

Asn Leu Arg Asp Thr Gly His Asn Val Val Ile Gly Ile Arg Ala Gly
35 40 45

Lys Ser Ala Glu Asn Ala Lys Asn Asp Gly Phe Thr Val Leu Pro Val
50 55 60

Ser Glu Ala Val Gln Gln Ala Asn Val Val Met Ile Leu Ala Pro Asp
65 70 75 80

Glu Ile Gln Gly Glu Leu Tyr Glu Lys Glu Ile Ala Pro Asn Leu Ser
85 90 95

Ala Gly Asp Ala Leu Ala Phe Ala His Gly Phe Asn Ile His Phe Glu
100 105 110

Val Ile Thr Pro Pro Lys Glu Val Asp Val Phe Leu Val Ala Pro Lys
115 120 125

Gly Pro Gly His Leu Val Arg Arg Thr Phe Ser Glu Gly Phe Ala Val
130 135 140

Pro Ala Leu Phe Ala Val Tyr Gln Asp Ala Thr Gly Asn Ala Gln Glu
145 150 155 160

Thr Ala Leu Ser Tyr Ala Lys Gly Ile Gly Ala Thr Arg Val Gly Val
165 170 175

```
Leu Glu Thr Thr Phe Lys Glu Glu Thr Glu Thr Asp Leu Phe Gly Glu
            180                 185                 190

Gln Ala Val Leu Cys Gly Gly Leu Thr Ser Met Ile Glu Thr Gly Phe
        195                 200                 205

Glu Thr Leu Val Glu Ala Gly Tyr Gln Pro Glu Leu Ala Tyr Phe Glu
    210                 215                 220

Val Cys His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr Glu Gly Gly
225                 230                 235                 240

Phe Ala Lys Met Arg Asp Ser Ile Ser Asn Thr Ala Glu Tyr Gly Asp
                245                 250                 255

Tyr Val Ser Gly Ser Arg Ile Ile Thr Glu Gln Thr Lys Ala Asn Met
            260                 265                 270

Lys Asn Val Leu Lys Asp Ile Gln Asn Gly Thr Phe Ala Lys Gly Phe
        275                 280                 285

Ile Ala Asp Asn Gln Ala Gly Phe Pro Glu Phe Lys Lys Met Arg Glu
    290                 295                 300

Gln Asn Gly Asn His Pro Ile Glu Lys Val Gly Ser Glu Leu Arg Lys
305                 310                 315                 320

Met Met Pro Phe Val Thr Lys Asp
                325
```

<210> 377
<211> 1023
<212> DNA
<213> Lactococcus lactis

<400> 377

```
atggctgtca ctatgtacta cgaagaagac gttgaagtcg ccgccttggc cggtaagaag       60

atcgccgtta tcggttacgg ttcccaaggt cacgctcatg cccaaaactt gagagactct      120

ggtcacgatg ttatcattgg tgttagacag ggtaaatctt tcgacaaggc taaagaagat      180

ggtttcgaaa cctttgaagt tggtgaagct gtcgccaagg ctgacgttat catggtcttg      240

gctccagatg aattgcaaca gtctatctac gaagaagaca ttaagccaaa cttgaaggct      300

ggttctgctt tgggtttcgc tcacggtttc aacatccatt tcggttacat tgaagttcca      360

gaagacgttg atgtcttcat ggtcgcccca aaggctccag gtcacttggt tagacgtacc      420

ttcaccgaag gtttcggtac cccagctttg ttcgtctctc accaaaacgc cactggtcat      480

gctagagaaa tcgccatgga ctgggctaag ggtattggtt gtgctagagt tggtatcatt      540

gaaaccactt tcaaggaaga aaccgaagaa gatttgttcg gtgaacaagc tgtcctgtgc      600

ggtggtttaa ctgccttggt tgaagctggt ttcgaaacct tgactgaagc cggttacgct      660

ggtgaattgg cttacttcga agttttacac gaaatgaagt tgatcgttga cttgatgtac      720

gaaggtggtt tcaccaagat gagacaatct atttccaaca ctgctgaatt cggtgactac      780
```

gtcaccggtc caagaatcat cactgatgaa gttaagaaga acatgaagtt ggtcctggcc 840

gacatccaat ccggtaaatt tgctcaggat ttcgttgacg attttaaggc cggtagacca 900

aagttgaccg cttaccgtga agctgccaag aacttggaaa tcgaaaagat cggtgctgaa 960

ttgagaaagg ctatgccatt cacccaatcc ggtgacgacg acgctttcaa gatttaccaa 1020

taa 1023

<210> 378
<211> 340
<212> PRT
<213> Lactococcus lactis

<400> 378

Met Ala Val Thr Met Tyr Tyr Glu Glu Asp Val Glu Val Ala Ala Leu
1 5 10 15

Ala Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asp Val Ile Ile Gly Val
35 40 45

Arg Gln Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Glu Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Phe Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Thr Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220

```
Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225                 230                 235                 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
            260                 265                 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Gln
            340
```

<210> 379
<211> 1023
<212> DNA
<213> Lactococcus lactis

<400> 379

```
atggcagtta caatgtatta tgaagatgat gtagaagtat cagcacttgc tggaaagcaa       60

attgcagtaa tcggttatgg ttcacaagga catgctcacg cacagaattt gcgtgattct      120

ggtcacaacg ttatcattgg tgtgcgccac ggaaaatctt ttgataaagc aaaagaagat      180

ggctttgaaa catttgaagt aggagaagca gtagctaaag ctgatgttat tatggttttg      240

gcaccagatg aacttcaaca atccatttat gaagaggaca tcaaaccaaa cttgaaagca      300

ggttcagcac ttggttttgc tcacggattt aatatccatt ttggctatat taaagtacca      360

gaagacgttg acgtctttat ggttgcgcct aaggctccag gtcaccttgt ccgtcggact      420

tatactgaag gttttggtac accagctttg tttgtttcac accaaaatgc aagtggtcat      480

gcgcgtgaaa tcgcaatgga ttgggccaaa ggaattggtt gtgctcgagt gggaattatt      540

gaaacaactt ttaaagaaga aacagaagaa gatttgtttg gagaacaagc tgttctatgt      600

ggaggtttga cagcacttgt tgaagccggt tttgaaacac tgacagaagc tggatacgct      660

ggcgaattgg cttactttga agttttgcac gaaatgaaat tgattgttga cctcatgtat      720

gaaggtggtt ttactaaaat gcgtcaatcc atctcaaata ctgctgagtt tggcgattat      780

gtgactggtc cacggattat tactgacgaa gttaaaaaga atatgaagct tgttttggct      840

gatattcaat ctggaaaatt tgctcaagat ttcgttgatg acttcaaagc ggggcgtcca      900

aaattaatag cctatcgcga agctgcaaaa aatcttgaaa ttgaaaaaat tggggcagag      960
```

ctacgtcaag caatgccatt cacacaatct ggtgatgacg atgcctttaa aatctatcag 1020

taa 1023

<210> 380
<211> 340
<212> PRT
<213> Lactococcus lactis

<400> 380

Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
1 5 10 15

Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
20 25 30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val
35 40 45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
50 55 60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65 70 75 80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
85 90 95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
115 120 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
130 135 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
145 150 155 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
165 170 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
195 200 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
210 215 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225 230 235 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
245 250 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
260 265 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
275 280 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala
290 295 300

Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305 310 315 320

Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
325 330 335

Lys Ile Tyr Gln
340

<210> 381
<211> 1023
<212> DNA
<213> Streptococcus thermophilus

<400> 381
atggccgttc aaatggaata cgaaaaggac gttaaggtcc cagccttgga tggtaagaag 60

atcgccgtta tcggttacgg ttctcaaggt cacgctcatt ctcaaaactt gagagacacc 120

ggtcacgatg ttatcattgg tgttagacca ggtaaatctt tcgacaaggc taaagaagac 180

ggttttgata cctacactgt cgctgaagcc actaagttgg ccgacgttat catgattttg 240

gctccagatg aaatccaaca ggaattgtac gaagctgaaa ttgccccaaa cttggaagct 300

ggtaatgccg tcggtttcgc tcacggtttc aacatccatt tcgaattcat caaggttcca 360

gccgacgttg atgtcttcat gtgtgctcca aagggtccag gtcacttggt cagacgtacc 420

ttcgaagaag gttttggtgt cccagctttg tacgccgttt accaagatgc tactggtaac 480

gccaaggaca tcgctatgga ttggtgtaaa ggtattggtg ctgccagagt tggtttgttg 540

gaaaccactt acaaggaaga aaccgaagaa gacttgttcg gtgaacaggc tgtcttatgc 600

ggtggtttaa ccgctttgat cgaaactggt ttcgaagttt tgactgaagc tggttacgcc 660

ccagaattgg cttacttcga agtcctgcac gaaatgaagt tgatcgttga cttgatctac 720

gaaggtggtt tcaagaagat gagacaatct atctccaaca ccgctgaatt cggtgactac 780

gtttctggtc caagagtcat cactgaacaa gttaaggaaa acatgaaggc tgttttggct 840

gacattcaga atggtaaatt cgctaacgac ttcgttaacg attacaaggc cggtagacca 900

aagttgaccg cttaccgtga agaagctgcc aacttggaaa tcgaaaaggt cggtgctgaa 960

ttgagaaagg ctatgccatt cgttggtcaa aacgacgacg acgctttcaa gatttacaac 1020

tga 1023

<210> 382

<211> 340
<212> PRT
<213> Streptococcus thermophilus

<400> 382

```
Met Ala Val Gln Met Glu Tyr Glu Lys Asp Val Lys Val Pro Ala Leu
1               5                  10                  15

Asp Gly Lys Lys Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
            20                  25                  30

His Ser Gln Asn Leu Arg Asp Thr Gly His Asp Val Ile Ile Gly Val
        35                  40                  45

Arg Pro Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Asp Thr
    50                  55                  60

Tyr Thr Val Ala Glu Ala Thr Lys Leu Ala Asp Val Ile Met Ile Leu
65                  70                  75                  80

Ala Pro Asp Glu Ile Gln Gln Glu Leu Tyr Glu Ala Glu Ile Ala Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Ala Val Gly Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Glu Phe Ile Lys Val Pro Ala Asp Val Asp Val Phe Met Cys
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Leu Val Arg Arg Thr Phe Glu Glu Gly
    130                 135                 140

Phe Gly Val Pro Ala Leu Tyr Ala Val Tyr Gln Asp Ala Thr Gly Asn
145                 150                 155                 160

Ala Lys Asp Ile Ala Met Asp Trp Cys Lys Gly Ile Gly Ala Ala Arg
                165                 170                 175

Val Gly Leu Leu Glu Thr Thr Tyr Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Ile Glu
        195                 200                 205

Thr Gly Phe Glu Val Leu Thr Glu Ala Gly Tyr Ala Pro Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Glu Gly Gly Phe Lys Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Ser Gly Pro Arg Val Ile Thr Glu Gln Val Lys
            260                 265                 270
```

```
Glu Asn Met Lys Ala Val Leu Ala Asp Ile Gln Asn Gly Lys Phe Ala
        275                 280                 285

Asn Asp Phe Val Asn Asp Tyr Lys Ala Gly Arg Pro Lys Leu Thr Ala
    290                 295                 300

Tyr Arg Glu Glu Ala Ala Asn Leu Glu Ile Glu Lys Val Gly Ala Glu
305                 310                 315                 320

Leu Arg Lys Ala Met Pro Phe Val Gly Gln Asn Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Asn
            340
```

<210> 383
<211> 11017
<212> DNA
<213> Artificial sequence

<220>
<223> pRS426::GPD-xpk1+ADH-eutD

<400> 383

```
gatcccccgg gctgcaggaa ttcgatatca agcttatcga taccgtcgac ctcgaggggg       60
ggcccggtac ccaattcgcc ctatagtgag tcgtattacg cgcgctcact ggccgtcgtt      120
ttacaacgtc gtgactggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat      180
ccccctttcg ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag      240
ttgcgcagcc tgaatggcga atggcgcgac gcgccctgta gcggcgcatt aagcgcggcg      300
ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct      360
ttcgctttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat      420
cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt      480
gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg      540
acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac      600
cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta      660
aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca      720
atttcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac cgcatagggt      780
aataactgat ataattaaat tgaagctcta atttgtgagt ttagtataca tgcatttact      840
tataatacag ttttttagtt ttgctggccg catcttctca aatatgcttc ccagcctgct      900
tttctgtaac gttcaccctc taccttagca tcccttccct ttgcaaatag tcctcttcca      960
acaataataa tgtcagatcc tgtagagacc acatcatcca cggttctata ctgttgaccc     1020
aatgcgtctc ccttgtcatc taaacccaca ccgggtgtca taatcaacca atcgtaacct     1080
tcatctcttc cacccatgtc tctttgagca ataaagccga taacaaaatc tttgtcgctc     1140
ttcgcaatgt caacagtacc cttagtatat tctccagtag atagggagcc cttgcatgac     1200
aattctgcta acatcaaaag gcctctaggt tcctttgtta cttcttctgc cgcctgcttc     1260
```

aaaccgctaa caatacctgg gcccaccaca ccgtgtgcat tcgtaatgtc tgcccattct 1320
gctattctgt atacacccgc agagtactgc aatttgactg tattaccaat gtcagcaaat 1380
tttctgtctt cgaagagtaa aaaattgtac ttggcggata atgcctttag cggcttaact 1440
gtgccctcca tggaaaaatc agtcaagata tccacatgtg tttttagtaa acaaattttg 1500
ggacctaatg cttcaactaa ctccagtaat tccttggtgg tacgaacatc caatgaagca 1560
cacaagtttg tttgcttttc gtgcatgata ttaaatagct tggcagcaac aggactagga 1620
tgagtagcag cacgttcctt atatgtagct ttcgacatga tttatcttcg tttcctgcag 1680
gtttttgttc tgtgcagttg ggttaagaat actgggcaat ttcatgtttc ttcaacacta 1740
catatgcgta tatataccaa tctaagtctg tgctccttcc ttcgttcttc cttctgttcg 1800
gagattaccg aatcaaaaaa atttcaaaga aaccgaaatc aaaaaaaaga ataaaaaaaa 1860
aatgatgaat tgaattgaaa agctgtggta tggtgcactc tcagtacaat ctgctctgat 1920
gccgcatagt taagccagcc ccgacacccg ccaacacccg ctgacgcgcc ctgacgggct 1980
tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg tctccgggag ctgcatgtgt 2040
cagaggtttt caccgtcatc accgaaacgc gcgagacgaa agggcctcgt gatacgccta 2100
tttttatagg ttaatgtcat gataataatg gtttcttagt atgatccaat atcaaaggaa 2160
atgatagcat tgaaggatga gactaatcca attgaggagt ggcagcatat agaacagcta 2220
aagggtagtg ctgaaggaag catacgatac cccgcatgga atgggataat atcacaggag 2280
gtactagact acctttcatc ctacataaat agacgcatat aagtacgcat ttaagcataa 2340
acacgcacta tgccgttctt ctcatgtata tatatataca ggcaacacgc agatataggt 2400
gcgacgtgaa cagtgagctg tatgtgcgca gctcgcgttg cattttcgga agcgctcgtt 2460
ttcggaaacg ctttgaagtt cctattccga agttcctatt ctctagaaag tataggaact 2520
tcagagcgct tttgaaaacc aaaagcgctc tgaagacgca ctttcaaaaa accaaaaacg 2580
caccggactg taacgagcta ctaaaatatt gcgaataccg cttccacaaa cattgctcaa 2640
aagtatctct ttgctatata tctctgtgct atatccctat ataacctacc catccacctt 2700
tcgctccttg aacttgcatc taaactcgac ctctacattt tttatgttta tctctagtat 2760
tactctttag acaaaaaaat tgtagtaaga actattcata gagtgaatcg aaaacaatac 2820
gaaaatgtaa acatttccta tacgtagtat atagagacaa aatagaagaa accgttcata 2880
attttctgac caatgaagaa tcatcaacgc tatcactttc tgttcacaaa gtatgcgcaa 2940
tccacatcgg tatagaatat aatcggggat gcctttatct tgaaaaaatg cacccgcagc 3000
ttcgctagta atcagtaaac gcgggaagtg gagtcaggct ttttttatgg aagagaaaat 3060
agacaccaaa gtagccttct tctaacctta acggacctac agtgcaaaaa gttatcaaga 3120
gactgcatta tagagcgcac aaaggagaaa aaaagtaatc taagatgctt tgttagaaaa 3180
atagcgctct cgggatgcat ttttgtagaa caaaaaagaa gtatagattc tttgttggta 3240
aaatagcgct ctcgcgttgc atttctgttc tgtaaaaatg cagctcagat tctttgtttg 3300
aaaaattagc gctctcgcgt tgcatttttg ttttacaaaa atgaagcaca gattcttcgt 3360
tggtaaaata gcgctttcgc gttgcatttc tgttctgtaa aaatgcagct cagattcttt 3420

gtttgaaaaa ttagcgctct cgcgttgcat ttttgttcta caaaatgaag cacagatgct 3480
tcgttcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta 3540
aatacattca aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata 3600
ttgaaaaagg aagagtatga gtattcaaca tttccgtgtc gcccttattc cctttttgc 3660
ggcattttgc cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga 3720
agatcagttg ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct 3780
tgagagtttt cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg 3840
tggcgcggta ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta 3900
ttctcagaat gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat 3960
gacagtaaga gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt 4020
acttctgaca acgatcggag gaccgaagga gctaaccgct tttttgcaca acatggggga 4080
tcatgtaact cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga 4140
gcgtgacacc acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga 4200
actacttact ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc 4260
aggaccactt ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc 4320
cggtgagcgt gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg 4380
tatcgtagtt atctacacga cggggagtca ggcaactatg gatgaacgaa atagacagat 4440
cgctgagata ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata 4500
tatactttag attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct 4560
ttttgataat ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga 4620
ccccgtagaa aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg 4680
cttgcaaaca aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc 4740
aactcttttt ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct 4800
agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc 4860
tctgctaatc ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt 4920
ggactcaaga cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg 4980
cacacagccc agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagct 5040
atgagaaagc gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag 5100
ggtcggaaca ggagagcgca cgagggagct tccaggggga aacgcctggt atctttatag 5160
tcctgtcggg tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg 5220
gcggagccta tggaaaaacg ccagcaacgc ggcctttta cggttcctgg ccttttgctg 5280
gccttttgct cacatgttct ttcctgcgtt atcccctgat tctgtggata accgtattac 5340
cgcctttgag tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt 5400
gagcgaggaa gcggaagagc gcccaatacg caaaccgcct ctccccgcgc gttggccgat 5460
tcattaatgc agctggcacg acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc 5520
aattaatgtg agttacctca ctcattaggc accccaggct ttacacttta tgcttccggc 5580

20 Jan 2017 2017200439

tcctatgttg tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca 5640
tgattacgcc aagcgcgcaa ttaaccctca ctaaagggaa caaaagctgg agctccaccg 5700
cggatagatc tagttcgagt ttatcattat caatactgcc atttcaaaga atacgtaaat 5760
aattaatagt agtgattttc ctaactttat ttagtcaaaa aattagcctt ttaattctgc 5820
tgtaacccgt acatgcccaa aatagggggc gggttacaca gaatatataa catcgtaggt 5880
gtctgggtga acagtttatt cctggcatcc actaaatata atggagcccg ctttttaagc 5940
tggcatccag aaaaaaaaag aatcccagca ccaaaatatt gttttcttca ccaaccatca 6000
gttcataggt ccattctctt agcgcaacta cagagaacag gggcacaaac aggcaaaaaa 6060
cgggcacaac ctcaatggag tgatgcaacc tgcctggagt aaatgatgac acaaggcaat 6120
tgacccacgc atgtatctat ctcattttct tacaccttct attaccttct gctctctctg 6180
atttggaaaa agctgaaaaa aaaggttgaa accagttccc tgaaattatt cccctacttg 6240
actaataagt atataaagac ggtaggtatt gattgtaatt ctgtaaatct atttcttaaa 6300
cttcttaaat tctactttta tagttagtct ttttttttagt tttaaaacac caagaactta 6360
gtttcgaata aacacacata aacgctgagg atgacaacag attactcatc accagcatat 6420
ttgcaaaaag ttgataagta ctggcgtgct gccaactact tatcagttgg tcaactttat 6480
ttaaaagata atccactatt acaacggcca ttgaaggcca gtgacgttaa ggttcatcca 6540
attggtcact gggggacgat tgccggtcaa aactttatct atgctcatct taaccgggtc 6600
atcaacaagt acggtttgaa gatgttctac gttgaaggtc caggtcatgg tggtcaagtg 6660
atggtttcaa actcttacct tgacggtact tacaccgata tttatccaga aattacgcag 6720
gatgttgaag ggatgcaaaa gctcttcaag caattctcat tcccaggtgg ggttgcttcc 6780
catgcggcac ctgaaacacc cggttcaatc cacgaaggtg gcgaacttgg ttactcaatt 6840
tcacacgggg ttggggcaat tcttgacaat cctgacgaaa tcgccgcggt tgttgttggt 6900
gatggggaat ccgaaacggg tccattagca acttcatggc aatcaacgaa gttcattaac 6960
ccaatcaacg acggggctgt tttaccaatc ttgaacttaa atggttttaa gatttctaat 7020
ccaacgattt ttggtcggac ttctgatgct aagattaagg aatacttcga aagcatgaat 7080
tgggaaccaa tcttcgttga aggtgacgat cctgaaaagg ttcacccagc cttagctaag 7140
gccatggatg aagccgttga aaagatcaag gcaatccaga agcatgctcg cgaaaataac 7200
gatgcaacat tgccagtatg gccaatgatc gtcttccgcg cacctaaggg ctggactggt 7260
ccgaagtcat gggacggtga taagatcgaa ggttcattcc gtgctcatca aattccgatt 7320
cctgttgatc aaaatgacat ggaacatgcg gatgctttag ttgattggct cgaatcatat 7380
caaccaaaag aactcttcaa tgaagatggc tctttgaagg atgatattaa agaaattatt 7440
cctactgggg acagtcggat ggctgctaac ccaatcacca atggtggggt cgatccgaaa 7500
gccttgaact taccaaactt ccgtgattat gcggtcgata cgtccaaaga aggcgcgaat 7560
gttaagcaag atatgatcgt ttggtcagac tatttgcggg atgtcatcaa gaaaaatcct 7620
gataacttcc ggttgttcgg acctgatgaa accatgtcta accgtttata tggtgtcttc 7680
gaaaccacta atcgtcaatg gatggaagac attcatccag atagtgacca atatgaagca 7740

20 Jan 2017 2017200439

ccagctggcc gggtcttaga tgctcagtta tctgaacacc aagctgaagg ttggttagaa 7800
ggttacgtct taactggacg tcatgggtta tttgccagtt atgaagcctt cctacgcgtt 7860
gtggactcaa tgttgacgca acacttcaag tggttacgta aagccaatga acttgattgg 7920
cgtaaaaagt acccatcact taacattatc gcggcttcaa ctgtattcca acaagaccat 7980
aatggttata cccaccaaga tccaggtgca ttaactcatt tggccgaaaa gaaaccagaa 8040
tacattcgtg aatatttacc agccgatgcc aacacgttat tagctgtcgg tgacgtcatt 8100
ttccggagcc aagaaaagat caactacgtg gttacgtcaa aacacccacg tcaacaatgg 8160
ttcagcattg aagaagctaa gcaattagtt gacaatggtc ttggtatcat tgattgggca 8220
agtacggacc aaggtagcga accagacatt gtctttgcag ctgctgggac ggaaccaacg 8280
cttgaaacgt tggctgccat ccaattacta cacgacagtt tcccagagat gaagattcgt 8340
ttcgtgaacg tggtcgacat cttgaagtta cgtagtcctg aaaaggatcc gcggggcttg 8400
tcagatgctg agtttgacca ttactttact aaggacaaac cagtggtctt tgctttccac 8460
ggttacgaag acttagttcg tgacatcttc tttgatcgtc acaaccataa cttatacgtc 8520
cacggttacc gtgaaaatgg tgatattacc acaccattcg acgtacgggt catgaaccag 8580
atggaccgct tcgacttagc taagtcggca attgcggcgc aaccagcaat ggaaaacact 8640
ggtgcggcct tcgttcaatc catggataat atgcttgcta aacacaatgc ctatatccgg 8700
gatgccggaa ctgacttgcc agaagttaat gattggcaat ggaagggttt aaaataatta 8760
attaatcatg taattagtta tgtcacgctt acattcacgc cctcctccca catccgctct 8820
aaccgaaaag gaaggagtta gacaacctga agtctaggtc cctatttatt ttttttaata 8880
gttatgttag tattaagaac gttatttata tttcaaattt ttctttttt tctgtacaaa 8940
cgcgtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg 9000
aaggctttaa tttgcgggcg gccgctctag aactagtacc acaggtgttg tcctctgagg 9060
acataaaata cacaccgaga ttcatcaact cattgctgga gttagcatat ctacaattgg 9120
gtgaaatggg gagcgatttg caggcatttg ctcggcatgc cggtagaggt gtggtcaata 9180
agagcgacct catgctatac ctgagaaagc aacctgacct acaggaaaga gttactcaag 9240
aataagaatt ttcgttttaa aacctaagag tcactttaaa atttgtatac acttatttt 9300
tttataactt atttaataat aaaaatcata aatcataaga aattcgctta ctcttaatta 9360
attaagctaa tccttgggct gctgtaataa tcgcaacctt ataaacgtct tcttcactgc 9420
atccacgtga caagtcggag accggcttgt tcaggccttg caagacagga cccaccgctt 9480
caaaatgacc aaatcgttgc gcaatcttat agccaatatt accagactga agctctggaa 9540
atacaaagac attggcatga ccagctactt tggaaccagg agccttttgc aaaccaactt 9600
tttcaacgaa ggccgcgtca aattgaagtt caccatcgat agccaattcc ggttcagcag 9660
cttgcgcctt ggccgttgct tcttgcactt tagtgaccat ttcaccctta gccgaaccct 9720
tagttgagaa gctgagcatc gcaactttcg ggtcaatatc gaagacctta gcagtagccg 9780
cactctgagt ggcaatttcc gctaacgtat cggcatcggg atcaatattg atggcacagt 9840
cagcaaagac gtagcgttcc tcacccttt gcatgataaa tgcacccgag attcggtgtg 9900

aaccgggctt ggtcttaata atttgtaacg ctggccgtac cgtatcacca gttggatgga 9960

ttgcacctga aaccatccca tccgctttgc ccatataaac gagcatcgtg ccaaagtagt 10020

tttcatcttc cagcatttta gccgcttgtt ctggcgtatt cttacctttc cgccgttcaa 10080

cgagggcatc aagcattgct tgcttatctt cagccgggta tgtcgcagga tcaaggactt 10140

gaacgcctgt taaatccgca ttcaaatcgt tagccacagc ctgaactttg tccgttgcac 10200

ctaaaacaat cggcttaacc aagccgtctg cagctaatcg cgctgccgca ccgacaattc 10260

ggggttcagt tccttcaggg aaaacaattg tttgatcttt accagtaatt ttttgtgcta 10320

atgactcaaa taaatccatc ctcagcgaga tagttgattg tatgcttggt atagcttgaa 10380

atattgtgca gaaaaagaaa caaggaagaa agggaacgag aacaatgacg aggaaacaaa 10440

agattaataa ttgcaggtct atttatactt gatagcaaga cagcaaactt ttttttattt 10500

caaattcaag taactggaag gaaggccgta taccgttgct cattagagag tagtgtgcgt 10560

gaatgaagga aggaaaaagt ttcgtgtgct tcgagatacc cctcatcagc tctggaacaa 10620

cgacatctgt tggtgctgtc tttgtcgtta atttttttcct ttagtgtctt ccatcatttt 10680

tttgtcattg cggatatggt gagacaacaa cgggggagag agaaaagaaa aaaaaagaaa 10740

agaagttgca tgcgcctatt attacttcaa tagatggcaa atggaaaaag ggtagtgaaa 10800

cttcgatatg atgatggcta tcaagtctag ggctacagta ttagttcgtt atgtaccacc 10860

atcaatgagg cagtgtaatt ggtgtagtct tgtttagccc attatgtctt gtctggtatc 10920

tgttctattg tatatctccc ctccgccacc tacatgttag ggagaccaac gaaggtatta 10980

taggaatccc gatgtatggg tttggttgcc agaaaag 11017

<210> 384
<211> 1862
<212> DNA
<213> Artificial sequence

<220>
<223> 1.8kb KpnI-HincII geneticin resistant cassette

<400> 384
gtcgactcta gaggatcccc caccttggct aactcgttgt atcatcactg gataacttcg 60

tatagcatac attatacgaa gttatctagg gattcataac cattttctca atcgaattac 120

acagaacaca ccgtacaaac ctctctatca taactactta atagtcacac acgtactcgt 180

ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt ggacagacaa ctataccagc 240

atggatctct tgtatcggtt cttttctccc gctctctcgc aataacaatg aacactgggt 300

caatcatagc ctacacaggt gaacagagta gcgtttatac agggtttata cggtgattcc 360

tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa tttttctttc caacgctaga 420

aggaaaagaa aaatctaatt aaattgattt ggtgattttc tgagagttcc ctttttcata 480

tatcgaattt tgaatataaa aggagatcga aaaaattttt ctattcaatc tgttttctgg 540

ttttatttga tagtttttt gtgtattatt attatggatt agtactggtt tatatgggtt 600

tttctgtata acttcttttt attttagttt gtttaatctt attttgagtt acattatagt 660

tccctaactg caagagaagt aacattaaaa ctcgagatgg gtaaggaaaa gactcacgtt 720

tcgaggccgc gattaaattc caacatggat gctgatttat atgggtataa atgggctcgc 780
gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt atgggaagcc cgatgcgcca 840
gagttgtttc tgaaacatgg caaaggtagc gttgccaatg atgttacaga tgagatggtc 900
agactaaact ggctgacgga atttatgcct cttccgacca tcaagcattt tatccgtact 960
cctgatgatg catggttact caccactgcg atccccggca aaacagcatt ccaggtatta 1020
gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc tggcagtgtt cctgcgccgg 1080
ttgcattcga ttcctgtttg taattgtcct tttaacagcg atcgcgtatt tcgtctcgct 1140
caggcgcaat cacgaatgaa taacggtttg gttgatgcga gtgattttga tgacgagcgt 1200
aatggctggc ctgttgaaca agtctggaaa gaaatgcata agcttttgcc attctcaccg 1260
gattcagtcg tcactcatgg tgatttctca cttgataacc ttatttttga cgaggggaaa 1320
ttaataggtt gtattgatgt tggacgagtc ggaatcgcag accgatacca ggatcttgcc 1380
atcctatgga actgcctcgg tgagttttct ccttcattac agaaacggct ttttcaaaaa 1440
tatggtattg ataatcctga tatgaataaa ttgcagtttc atttgatgct cgatgagttt 1500
ttctaagttt aacttgatac tactagattt tttctcttca tttataaaat ttttggttat 1560
aattgaagct ttagaagtat gaaaaaatcc ttttttttca ttctttgcaa ccaaaataag 1620
aagcttcttt tattcattga aatgatgaat ataaacctaa caaaagaaaa agactcgaat 1680
atcaaacatt aaaaaaaaat aaaagaggtt atctgttttc ccatttagtt ggagtttgca 1740
ttttctaata gatagaactc tcaattaatg tggatttagt ttctctgttc gataacttcg 1800
tatagcatac attatacgaa gttatctgaa cattagaata cgtaatccgc aatgcgggta 1860
cc 1862

<210> 385
<211> 7938
<212> DNA
<213> Artificial sequence

<220>
<223> YZ067DkivDDhADH

<400> 385
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360
ttttttttt ccacctagcg gatgactctt ttttttttctt agcgattggc attatcacat 420
aatgaattat acattatata aagtaatgtg atttcttcga agaatatact aaaaaatgag 480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca 540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac 600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg 660

20 Jan 2017 2017200439

attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat 720
tccggctggt cgctaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc 780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt 840
ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcggtg 900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta 960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga 1020
attaccctcc acgttgattg tctgcgaggc aagaatgatc atcaccgtag tgagagtgcg 1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt 1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgat 1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta 1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg 1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg 1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt 1440
aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag 1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt 1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga 1620
aaaaccgtct atcagggcga tggcccacta cgtggccggc ttcacatacg ttgcatacgt 1680
cgatatagat aataatgata atgacagcag gattatcgta atacgtaata gctgaaaatc 1740
tcaaaaatgt gtgggtcatt acgtaaataa tgataggaat gggattcttc tatttttcct 1800
ttttccattc tagcagccgt cgggaaaacg tggcatcctc tctttcgggc tcaattggag 1860
tcacgctgcc gtgagcatcc tctctttcca tatctaacaa ctgagcacgt aaccaatgga 1920
aaagcatgag cttagcgttg ctccaaaaaa gtattggatg gttaatacca tttgtctgtt 1980
ctcttctgac tttgactcct caaaaaaaaa aatctacaat caacagatcg cttcaattac 2040
gccctcacaa aaacttttt ccttcttctt cgcccacgtt aaattttatc cctcatgttg 2100
tctaacggat ttctgcactt gatttattat aaaaagacaa agacataata cttctctatc 2160
aatttcagtt attgttcttc cttgcgttat tcttctgttc ttctttttct tttgtcatat 2220
ataaccataa ccaagtaata catattcaaa cacgtgagta tgactgacaa aaaaactctt 2280
aaagacttaa gaaatcgtag ttctgtttac gattcaatgg ttaaatcacc taatcgtgct 2340
atgttgcgtg caactggtat gcaagatgaa gactttgaaa aacctatcgt cggtgtcatt 2400
tcaacttggg ctgaaaacac accttgtaat atccacttac atgactttgg taaactagcc 2460
aaagtcggtg ttaaggaagc tggtgcttgg ccagttcagt tcggaacaat cacggtttct 2520
gatggaatcg ccatgggaac ccaaggaatg cgtttctcct tgacatctcg tgatattatt 2580
gcagattcta ttgaagcagc catgggaggt cataatgcgg atgcttttgt agccattggc 2640
ggttgtgata aaaacatgcc cggttctgtt atcgctatgg ctaacatgga tatcccagcc 2700
atttttgctt acggcggaac aattgcacct ggtaatttag acggcaaaga tatcgattta 2760
gtctctgtct ttgaaggtgt cggccattgg aaccacggcg atatgaccaa agaagaagtt 2820

aaagctttgg aatgtaatgc ttgtcccggt cctggaggct gcggtggtat gtatactgct 2880
aacacaatgg cgacagctat tgaagttttg ggacttagcc ttccgggttc atcttctcac 2940
ccggctgaat ccgcagaaaa gaaagcagat attgaagaag ctggtcgcgc tgttgtcaaa 3000
atgctcgaaa tgggcttaaa accttctgac attttaacgc gtgaagcttt tgaagatgct 3060
attactgtaa ctatggctct gggaggttca accaactcaa cccttcacct cttagctatt 3120
gcccatgctg ctaatgtgga attgacactt gatgatttca atactttcca agaaaaagtt 3180
cctcatttgg ctgatttgaa accttctggt caatatgtat tccaagacct ttacaaggtc 3240
ggaggggtac cagcagttat gaaatatctc cttaaaaatg gcttccttca tggtgaccgt 3300
atcacttgta ctggcaaaac agtcgctgaa aatttgaagg cttttgatga tttaacacct 3360
ggtcaaaagg ttattatgcc gcttgaaaat cctaaacgtg aagatggtcc gctcattatt 3420
ctccatggta acttggctcc agacggtgcc gttgccaaag tttctggtgt aaaagtgcgt 3480
cgtcatgtcg gtcctgctaa ggtctttaat tctgaagaag aagccattga agctgtcttg 3540
aatgatgata ttgttgatgg tgatgttgtt gtcgtacgtt ttgtaggacc aaagggcggt 3600
cctggtatgc ctgaaatgct ttccctttca tcaatgattg ttggtaaagg gcaaggtgaa 3660
aaagttgccc ttctgacaga tggccgcttc tcaggtggta cttatggtct tgtcgtgggt 3720
catatcgctc ctgaagcaca agatggcggt ccaatcgcct acctgcaaac aggagacata 3780
gtcactattg accaagacac taaggaatta cactttgata tctccgatga agagttaaaa 3840
catcgtcaag agaccattga attgccaccg ctctattcac gcggtatcct tggtaaatat 3900
gctcacatcg tttcgtctgc ttctagggga gccgtaacag acttttggaa gcctgaagaa 3960
actggcaaaa aatgttgtcc tggttgctgt ggttaagcgg ccgcgttaat tcaaattaat 4020
tgatatagtt ttttaatgag tattgaatct gtttagaaat aatggaatat tatttttatt 4080
tatttattta tattattggt cggctctttt cttctgaagg tcaatgacaa aatgatatga 4140
aggaaataat gatttctaaa attttacaac gtaagatatt tttacaaaag cctagctcat 4200
cttttgtcat gcactatttt actcacgctt gaaattaacg gccagtccac tgcggagtca 4260
tttcaaagtc atcctaatcg atctatcgtt tttgatagct cattttggag ttcgcgagga 4320
tcccagcttt tgttcccttt agtgagggtt aattgcgcgc ttggcgtaat catggtcata 4380
gctgtttcct gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag 4440
cataaagtgt aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg 4500
ctcactgccc gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca 4560
acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc 4620
gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 4680
gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 4740
ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccctga 4800
cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 4860
ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 4920
taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 4980

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 5040
ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 5100
aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 5160
tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac 5220
agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 5280
ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 5340
tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc 5400
tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt 5460
cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta 5520
aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct 5580
atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg 5640
cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga 5700
tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt 5760
atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt 5820
taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt 5880
tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat 5940
gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc 6000
cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc 6060
cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat 6120
gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag 6180
aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt 6240
accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc 6300
ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa 6360
gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc aatattattg 6420
aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa 6480
taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgaac gaagcatctg 6540
tgcttcattt tgtagaacaa aaatgcaacg cgagagcgct aatttttcaa acaaagaatc 6600
tgagctgcat ttttacagaa cagaaatgca acgcgaaagc gctattttac caacgaagaa 6660
tctgtgcttc atttttgtaa aacaaaaatg caacgcgaga gcgctaattt ttcaaacaaa 6720
gaatctgagc tgcattttta cagaacagaa atgcaacgcg agagcgctat tttaccaaca 6780
aagaatctat acttcttttt tgttctacaa aaatgcatcc cgagagcgct atttttctaa 6840
caaagcatct tagattactt tttttctcct ttgtgcgctc tataatgcag tctcttgata 6900
actttttgca ctgtaggtcc gttaaggtta gaagaaggct actttggtgt ctattttctc 6960
ttccataaaa aaagcctgac tccacttccc gcgtttactg attactagcg aagctgcggg 7020
tgcatttttt caagataaag gcatccccga ttatattcta taccgatgtg gattgcgcat 7080
actttgtgaa cagaaagtga tagcgttgat gattcttcat tggtcagaaa attatgaacg 7140

gtttcttcta ttttgtctct atatactacg tataggaaat gtttacattt tcgtattgtt 7200

ttcgattcac tctatgaata gttcttacta caatttttt gtctaaagag taatactaga 7260

gataaacata aaaaatgtag aggtcgagtt tagatgcaag ttcaaggagc gaaaggtgga 7320

tgggtaggtt atatagggat atagcacaga gatatatagc aaagagatac ttttgagcaa 7380

tgtttgtgga agcggtattc gcaatatttt agtagctcgt tacagtccgg tgcgtttttg 7440

gtttttgaa agtgcgtctt cagagcgctt ttggttttca aaagcgctct gaagttccta 7500

tactttctag agaataggaa cttcggaata ggaacttcaa agcgtttccg aaaacgagcg 7560

cttccgaaaa tgcaacgcga gctgcgcaca tacagctcac tgttcacgtc gcacctatat 7620

ctgcgtgttg cctgtatata tatatacatg agaagaacgg catagtgcgt gtttatgctt 7680

aaatgcgtac ttatatgcgt ctatttatgt aggatgaaag gtagtctagt acctcctgtg 7740

atattatccc attccatgcg gggtatcgta tgcttccttc agcactaccc tttagctgtt 7800

ctatatgctg ccactcctca attggattag tctcatcctt caatgctatc atttcctttg 7860

atattggatc atactaagaa accattatta tcatgacatt aacctataaa aataggcgta 7920

tcacgaggcc ctttcgtc 7938

<210> 386
<211> 4586
<212> DNA
<213> Artificial sequence

<220>
<223> pLA54 template DNA

<400> 386
gggtaccgag ctcgaattca ctggccgtcg ttttacaacg tcgtgactgg gaaaaccctg 60

gcgttaccca acttaatcgc cttgcagcac atccccctttt cgccagctgg cgtaatagcg 120

aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc gaatggcgcc 180

tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tggtgcactc 240

tcagtacaat ctgctctgat gccgcatagt taagccagcc ccgacacccg ccaacacccg 300

ctgacgcgcc ctgacgggct tgtctgctcc cggcatccgc ttacagacaa gctgtgaccg 360

tctccgggag ctgcatgtgt cagaggtttt caccgtcatc accgaaacgc gcgagacgaa 420

agggcctcgt gatacgccta tttttatagg ttaatgtcat gataataatg gtttcttaga 480

cgtcaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta tttttctaaa 540

tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt 600

gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg 660

cattttgcct tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag 720

atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg 780

agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg 840

gcgcggtatt atcccgtatt gacgccgggc aagagcaact cggtcgccgc atacactatt 900

ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga 960

cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac 1020

ttctgacaac gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc 1080
atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc 1140
gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac 1200
tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag 1260
gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg 1320
gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta 1380
tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg 1440
ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata 1500
tactttagat tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt 1560
ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc 1620
ccgtagaaaa gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct 1680
tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa 1740
ctcttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag 1800
tgtagccgta gttaggccac acttcaaga actctgtagc accgcctaca tacctcgctc 1860
tgctaatcct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg 1920
actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca 1980
cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat 2040
gagaaagcgc cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg 2100
tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc 2160
ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc 2220
ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc 2280
cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg 2340
cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga 2400
gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc 2460
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 2520
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc 2580
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 2640
attacgccaa gcttgcatgc ctgcaggtcg actctagagg atccccgcat tgcggattac 2700
gtattctaat gttcagataa cttcgtatag catacattat acgaagttat ctagggattc 2760
ataaccattt tctcaatcga attacacaga acacaccgta caaacctctc tatcataact 2820
acttaatagt cacacacgta ctcgtctaaa tacacatcat cgtcctacaa gttcatcaaa 2880
gtgttggaca gacaactata ccagcatgga tctcttgtat cggttctttt ctcccgctct 2940
ctcgcaataa caatgaacac tgggtcaatc atagcctaca caggtgaaca gagtagcgtt 3000
tatacagggt ttatacggtg attcctacgg caaaaatttt tcatttctaa aaaaaaaaag 3060
aaaaattttt ctttccaacg ctagaaggaa aagaaaaatc taattaaatt gatttggtga 3120
ttttctgaga gttccctttt tcatatatcg aattttgaat ataaaaggag atcgaaaaaa 3180

ttttctatt caatctgttt tctggtttta tttgatagtt tttttgtgta ttattattat 3240
ggattagtac tggtttatat gggtttttct gtataacttc tttttatttt agtttgttta 3300
atcttatttt gagttacatt atagttccct aactgcaaga gaagtaacat taaaactcga 3360
gatgggtaag gaaaagactc acgtttcgag gccgcgatta aattccaaca tggatgctga 3420
tttatatggg tataaatggg ctcgcgataa tgtcgggcaa tcaggtgcga caatctatcg 3480
attgtatggg aagcccgatg cgccagagtt gtttctgaaa catggcaaag gtagcgttgc 3540
caatgatgtt acagatgaga tggtcagact aaactggctg acggaattta tgcctcttcc 3600
gaccatcaag cattttatcc gtactcctga tgatgcatgg ttactcacca ctgcgatccc 3660
cggcaaaaca gcattccagg tattagaaga atatcctgat tcaggtgaaa atattgttga 3720
tgcgctggca gtgttcctgc gccggttgca ttcgattcct gtttgtaatt gtccttttaa 3780
cagcgatcgc gtatttcgtc tcgctcaggc gcaatcacga atgaataacg gtttggttga 3840
tgcgagtgat tttgatgacg agcgtaatgg ctggcctgtt gaacaagtct ggaaagaaat 3900
gcataagctt ttgccattct caccggattc agtcgtcact catggtgatt tctcacttga 3960
taaccttatt tttgacgagg ggaaattaat aggttgtatt gatgttggac gagtcggaat 4020
cgcagaccga taccaggatc ttgccatcct atggaactgc ctcggtgagt tttctccttc 4080
attacagaaa cggctttttc aaaaatatgg tattgataat cctgatatga ataaattgca 4140
gtttcatttg atgctcgatg agtttttcta agtttaactt gatactacta gatttttttct 4200
cttcatttat aaaatttttg gttataattg aagctttaga agtatgaaaa aatccttttt 4260
tttcattctt tgcaaccaaa ataagaagct tcttttattc attgaaatga tgaatataaa 4320
cctaacaaaa gaaaaagact cgaatatcaa acattaaaaa aaaataaaag aggttatctg 4380
ttttcccatt tagttggagt ttgcattttc taatagatag aactctcaat taatgtggat 4440
ttagtttctc tgttcgtttt tttttgtttt gttctcactg tatttacatt tctatttagt 4500
atttagttat tcatataatc tataacttcg tatagcatac attatacgaa gttatccagt 4560
gatgatacaa cgagttagcc aaggtg 4586

<210> 387
<211> 7589
<212> DNA
<213> Artificial sequence

<220>
<223> pUC19-kan::pdc1::FBA-alsS::TRX1

<400> 387
tatttgtatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttggaaac 60
cacactgttt aaacagtgtt ccttaatcaa ggatacctct ttttttttcc ttggttccac 120
taattcatcg gttttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180
ttaagagaaa ttctccaaat ttgtaaagaa gcggacccag acttaaggcc gcccgcaaat 240
taaagccttc gagcgtccca aaaccttctc aagcaaggtt ttcagtataa tgttacatgc 300
gtacacgcgt ctgtacagaa aaaaaagaaa aatttgaaat ataaataacg ttcttaatac 360
taacataact ataaaaaaat aaatagggac ctagacttca ggttgtctaa ctccttcctt 420

ttcggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480
gattaattaa ctagagagct ttcgttttca tgagttcccc gaattctttc ggaagcttgt 540
cacttgctaa attaatgtta tcactgtagt caaccgggac atcgatgatg acaggacctt 600
cagcgttcat gccttgacgc agaacatctg ccagctggtc tggtgattct acgcgcaagc 660
cagttgctcc gaagctttcc gcatatttca cgatatcgat atttccgaaa tcgaccgcag 720
atgtacggtt atatttttc aattgctgga atgcaaccat gtcatatgtg ctgtcgttcc 780
atacaatgtg tacaattggt gcttttagtc gaactgctgt ctctaattcc attgctgaga 840
ataagaaacc gccgtcacca gagacagaaa ccacttttc tcccggtttc accaatgaag 900
cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg 960
ttaacggctc gtagctgcgg aaataacgtg acatccaaat ggcgtgcgaa ccgatatcgc 1020
aagttactgt aacatgatca tcgactgcat tacgcaactc tttaacgatt tcaagagggt 1080
gcgctctgtc tgatttccaa tctgcaggca cctgctcacc ttcatgcata tattgtttta 1140
aatcagaaag gattttctgc tcacgctctg caaattccac tttcacagca tcgtgttcga 1200
tatgattgat cgtggacgga atgtcaccga tcaattcaag atcaggctgg taagcatgat 1260
caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga 1320
atttcggatc atattcaatc gggtcatagc cgatcgtcag aacaacatct gcctgctcta 1380
gcagtaaatc gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatcctcta 1440
aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt 1500
tcaaaagctt gcgaaccgct ttaattgctt ccggtcttcc gcctttcatg ccgaccaaaa 1560
cgacaggaag ttttgctgtt tggatttttg ctatggccgc actgattgca tcatctgctg 1620
caggaccgag ttttggcgct gcaacagcac gcacgttttt cgtatttgtg acttcattca 1680
caacatcttg cggaaagctc acaaaagcgg ccccagcctg ccctgctgac gctatcctaa 1740
atgcatttgt aacagcttcc ggtatatttt ttacatcttg aacttctaca ctgtattttg 1800
taatcggctg gaatagcgcc gcattatcca aagattgatg tgtccgtttt aaacgatctg 1860
cacggatcac gtttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca 1920
ggcctgttgc caagttagag gcacccggtc ctgatgtgac taacacgact cccggttttc 1980
cagttaaacg gccgactgct tgggccatga atgctgcgtt ttgttcgtgc cgggcaacga 2040
taatttcagg tcctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa 2100
tgccaaatac atgtgtgaca ccttgctcca ctaagcaatc aacaacaagc tccgcccctc 2160
tgtttttcac aagggatttt tgttcttttg ttgcttttgt caacatcctc agctctagat 2220
ttgaatatgt attacttggt tatggttata tatgacaaaa gaaaaagaag aacagaagaa 2280
taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta 2340
taataaatca agtgcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga 2400
agaagaagga aaaaagtttt tgtgagggcg taattgaagc gatctgttga ttgtagattt 2460
ttttttttg aggagtcaaa gtcagaagag aacagacaaa tggtattaac catccaatac 2520
ttttttggag caacgctaag ctcatgcttt tccattggtt acgtgctcag ttgttagata 2580

tggaaagaga ggatgctcac ggcagcgtga ctccaattga gcccgaaaga gaggatgcca 2640
cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca 2700
ttatttacgt aatgacccac acatttttga gattttcaac tattacgtat tacgataatc 2760
ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaagtagg 2820
caattattta gtactgtcag tattgttatt catttcagat cttaagccta accaggccaa 2880
ttcaacagac tgtcggcaac ttcttgtctg gtctttccat ggtaagtgac agtgcagtaa 2940
taatatgaac caatttattt ttcgttacat aaaaatgctt ataaaacttt aactaataat 3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggct aactcgttgt 3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac 3120
cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca taactactta 3180
atagtcacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt 3240
ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc 3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac 3360
agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa 3420
tttttctttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc 3480
tgagagttcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt 3540
ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt 3600
agtactggtt tatatgggtt tttctgtata acttcttttt attttagttt gtttaatctt 3660
attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa ctcgagatgg 3720
gtaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat gctgatttat 3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt 3840
atgggaagcc cgatgcgcca gagttgtttc tgaaacatgg caaaggtagc gttgccaatg 3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct cttccgacca 3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca 4020
aaacagcatt ccaggtatta gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc 4080
tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct tttaacagcg 4140
atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg gttgatgcga 4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata 4260
agcttttgcc attctcaccg gattcagtcg tcactcatgg tgatttctca cttgataacc 4320
ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc ggaatcgcag 4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct ccttcattac 4440
agaaacggct ttttcaaaaa tatggtattg ataatcctga tatgaataaa ttgcagtttc 4500
atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt tttctcttca 4560
tttataaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc ttttttttca 4620
ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa 4680
caaaagaaaa agactcgaat atcaaacatt aaaaaaaaat aaaagaggtt atctgttttc 4740

ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt 4800
ttctctgttc gataacttcg tatagcatac attatacgaa gttatctgaa cattagaata 4860
cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct 4920
tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 4980
acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 5040
tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 5100
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 5160
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 5220
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 5280
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 5340
ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 5400
acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 5460
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 5520
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 5580
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 5640
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 5700
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 5760
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 5820
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 5880
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 5940
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 6000
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 6060
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 6120
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 6180
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6240
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 6300
gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 6360
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 6420
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 6480
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 6540
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 6600
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 6660
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 6720
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 6780
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840
ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 6960

tccttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140

cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc 7200

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260

gcgcgtcagc gggtgttggc gggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320

ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 7380

accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg cgattaagtt 7500

gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560

tcggtacccg gggatccggc gcgccgttt 7589

<210> 388
<211> 19
<212> DNA
<213> Artificial sequence

<220>
<223> N160SeqR5 primer

<400> 388
tgagcccgaa agagaggat 19

<210> 389
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> k9_Grace_SH

<400> 389

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

```
Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Asp Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 390
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> primer K9G9_EZ_F1

<400> 390
aaacatggaa gaatgtaaga tggc 24

<210> 391
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9G9_EZ_R1

<400> 391
tcagttgtta atcaacttgt cttcg 25

<210> 392
<211> 9613
<212> DNA
<213> Artificial sequence

<220>
<223> pHR81-PIlv5-KARI-K9.G9

<400> 392
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 60

cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 120

cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgcgaagga 180

gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 240

ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 300

catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca 360

tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 420

aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 480

cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 540

tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 600

cggtgaacaa gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 660

cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 720

gttgatcgtt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 780

cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 840

ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt 900

tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 960

acacccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 1020

caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 1080

ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 1140

catttaagat ttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt 1200

ctgaagcagc ttcaaatata tatatttttt acatatttat tatgattcaa tgaacaatct 1260

aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt 1320

gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggctga gctactggtt 1380

gctgtgacgc agcggcatta gcgcggcgtt atgagctacc ctcgtggcct gaaagatggc 1440

gggaataaag cggaactaaa aattactgac tgagccatat tgaggtcaat ttgtcaactc 1500
gtcaagtcac gtttggtgga cggccccttt ccaacgaatc gtatatacta acatgcgcgc 1560
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca 1620
cctgtattta atttccttac tcgcgggttt ttctttttc tcaattcttg gcttcctctt 1680
tctcgagcgg accggatcct ccgcggtgcc ggcagatcta tttaaatggc gcgccgacgt 1740
caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac 1800
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 1860
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 1920
tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc 1980
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 2040
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 2100
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 2160
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 2220
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc 2280
tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg 2340
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 2400
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 2460
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 2520
cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 2580
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 2640
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 2700
agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 2760
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atccttttg 2820
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 2880
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 2940
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc 3000
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt 3060
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 3120
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 3180
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac 3240
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 3300
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg 3360
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 3420
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca gggggggcgga 3480
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 3540
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct 3600

ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 3660
aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt 3720
aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 3780
atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 3840
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 3900
acgccaagct ttttctttcc aatttttttt ttttcgtcat tataaaaatc attacgaccg 3960
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata 4020
catgcattta cttataatac agtttttag ttttgctggc cgcatcttct caaatatgct 4080
tcccagcctg cttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat 4140
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta 4200
tactgttgac ccaatgcgtc tcccttgtca tctaaaccca caccgggtgt cataatcaac 4260
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa 4320
tctttgtcgc tcttcgcaat gtcaacagta cccttagtat attctccagt agatagggag 4380
cccttgcatg acaattctgc taacatcaaa aggcctctag gttcctttgt tacttcttct 4440
gccgcctgct tcaaaccgct aacaatacct gggcccacca caccgtgtgc attcgtaatg 4500
tctgcccatt ctgctattct gtatacaccc gcagagtact gcaatttgac tgtattacca 4560
atgtcagcaa attttctgtc ttcgaagagt aaaaaattgt acttggcgga taatgccttt 4620
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt 4680
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca 4740
tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag cttggcagca 4800
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt 4860
cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt 4920
tcttcaacac tacatatgcg tatatatacc aatctaagtc tgtgctcctt ccttcgttct 4980
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaa 5040
gaataaaaaa aaaatgatga attgaaaagc ttgcatgcct gcaggtcgac tctagtatac 5100
tccgtctact gtacgataca cttccgctca ggtccttgtc ctttaacgag gccttaccac 5160
tcttttgtta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc 5220
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac 5280
aatggctgcc atcattatta tccgatgtga cgctgcattt ttttttttt ttttttttt 5340
ttttttttt ttttttttt ttttttttg tacaaatatc ataaaaaaag agaatctttt 5400
taagcaagga ttttcttaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg 5460
gaaccaccta aatcaccagt tctgatacct gcatccaaaa ccttttaac tgcatcttca 5520
atggctttac cttcttcagg caagttcaat gacaatttca acatcattgc agcagacaag 5580
atagtggcga tagggttgac cttattcttt ggcaaatctg gagcggaacc atggcatggt 5640
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac 5700
aaacccaagg agcctgggat aacggaggct tcatcggaga tgatatcacc aaacatgttg 5760

ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggcggcagaa 5820
tcaatcaatt gatgttgaac tttcaatgta gggaattcgt tcttgatggt ttcctccaca 5880
gtttttctcc ataatcttga agaggccaaa acattagctt tatccaagga ccaaataggc 5940
aatggtggct catgttgtag ggccatgaaa gcggccattc ttgtgattct ttgcacttct 6000
ggaacggtgt attgttcact atcccaagcg acaccatcac catcgtcttc ctttctctta 6060
ccaaagtaaa tacctcccac taattctcta acaacaacga agtcagtacc tttagcaaat 6120
tgtggcttga ttggagataa gtctaaaaga gagtcggatg caaagttaca tggtcttaag 6180
ttggcgtaca attgaagttc tttacggatt tttagtaaac cttgttcagg tctaacacta 6240
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatcagc cttcttggag 6300
gcttccagcg cctcatctgg aagtggaaca cctgtagcat cgatagcagc accaccaatt 6360
aaatgatttt cgaaatcgaa cttgacattg gaacgaacat cagaaatagc tttaagaacc 6420
ttaatggctt cggctgtgat ttcttgacca acgtggtcac ctggcaaaac gacgatcttc 6480
ttaggggcag acattacaat ggtatatcct tgaaatatat ataaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa aatgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta 6600
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg 6660
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat 6720
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg 6780
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc 6840
ttgcacttca atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa 6900
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac 6960
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa 7020
acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac 7080
agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaatctata cttctttttt 7140
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt 7200
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtccg 7260
ttaaggttag aagaaggcta ctttggtgtc tattttctct tccataaaaa aagcctgact 7320
ccacttcccg cgtttactga ttactagcga agctgcgggt gcatttttc aagataaagg 7380
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat 7440
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta 7500
tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag 7560
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga 7620
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata 7680
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg 7740
caatatttta gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc 7800
agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac 7860
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag 7920

20172004439 20 Jan 2017

ctgcgcacat acagctcact gttcacgtcg cacctatatc tgcgtgttgc ctgtatatat 7980

atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc 8040

tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg 8100

ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa 8160

ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt 8220

accgagaaac tagaggatct cccattaccg acatttgggc gctatacgtg catatgttca 8280

tgtatgtatc tgtatttaaa acacttttgt attatttttc ctcatatatg tgtataggtt 8340

tatacggatg atttaattat tacttcacca ccctttattt caggctgata tcttagcctt 8400

gttactagtc accggtggcg gccgcacctg gtaaaacctc tagtggagta gtagatgtaa 8460

tcaatgaagc ggaagccaaa agaccagagt agaggcctat agaagaaact gcgatacctt 8520

ttgtgatggc taaacaaaca gacatctttt tatatgtttt tacttctgta tatcgtgaag 8580

tagtaagtga taagcgaatt tggctaagaa cgttgtaagt gaacaaggga cctcttttgc 8640

ctttcaaaaa aggattaaat ggagttaatc attgagattt agttttcgtt agattctgta 8700

tccctaaata actcccttac ccgacgggaa ggcacaaaag acttgaataa tagcaaacgg 8760

ccagtagcca agaccaaata atactagagt taactgatgg tcttaaacag gcattacgtg 8820

gtgaactcca agaccaatat acaaaatatc gataagttat tcttgcccac caatttaagg 8880

agcctacatc aggacagtag taccattcct cagagaagag gtatacataa caagaaaatc 8940

gcgtgaacac cttatataac ttagcccgtt attgagctaa aaaaccttgc aaaatttcct 9000

atgaataaga atacttcaga cgtgataaaa atttactttc taactcttct cacgctgccc 9060

ctatctgttc ttccgctcta ccgtgagaaa taaagcatcg agtacggcag ttcgctgtca 9120

ctgaactaaa acaataaggc tagttcgaat gatgaacttg cttgctgtca aacttctgag 9180

ttgccgctga tgtgacactg tgacaataaa ttcaaaccgg ttatagcggt ctcctccggt 9240

accggttctg ccacctccaa tagagctcag taggagtcag aacctctgcg gtggctgtca 9300

gtgactcatc cgcgtttcgt aagttgtgcg cgtgcacatt tcgcccgttc ccgctcatct 9360

tgcagcaggc ggaaattttc atcacgctgt aggacgcaaa aaaaaaataa ttaatcgtac 9420

aagaatcttg gaaaaaaaat tgaaaaattt tgtataaaag ggatgaccta acttgactca 9480

atggctttta cacccagtat tttccctttc cttgtttgtt acaattatag aagcaagaca 9540

aaaacatata gacaacctat tcctaggagt tatatttttt taccctacca gcaatataag 9600

taaaaaactg ttt 9613

<210> 393
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Primer pBAD-For

<400> 393
atgccatagc atttttatcc 20

<210> 394
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Primer pBAD-Rev

<400> 394
ctgatttaat ctgtatcagg ct 22

<210> 395
<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N87PC1

<400> 395
ctgacatcat tatgatcttg atcccagatg aaaagcaggc taccatgtac 50

<210> 396
<211> 50
<212> DNA
<213> Artificial sequence

<220>
<223> Primer N87PC1r

<400> 396
gtacatggta gcctgctttt catctgggat caagatcata atgatgtcag 50

<210> 397
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK57E1

<400> 397
ggtttattcg aaggtgcgga ggagtggaaa agagctg 37

<210> 398
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK57E1r

<400> 398
cagctctttt ccactcctcc gcaccttcga ataaacc 37

<210> 399
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oY53F1

<400> 399
gtaacgttat cattggttta tacgaaggtg cggaggag 38

<210> 400
<211> 38
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oY53F1r

<400> 400
ctcctccgca ccttcgaata aaccaatgat aacgttac 38

<210> 401
<211> 43
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK90L1

<400> 401
gatcttgatc ccagatgaat tgcaggctac catgtacaaa aac 43

<210> 402
<211> 42
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK90L1r

<400> 402
gtttttgtac atggtagcct gcaattcatc tgggatcaag at 42

<210> 403
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK90M1

<400> 403
cttgatccca gatgaaatgc aggctaccat gtacaaaaac 40

<210> 404
<211> 40
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oK90M1r

<400> 404
gtttttgtac atggtagcct gcatttcatc tgggatcaag 40

<210> 405
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> K9_Ingrid_SH

<400> 405

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

```
Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Ile His Phe
    290                 295                 300
```

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 406
<211> 719
<212> DNA
<213> Artificial sequence

<220>
<223> Promoter UAS(PGK1)-PFBA1

<400> 406
aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgc 120

ctacttggct tcacatacgt tgcatacgtc gatatagata ataatgataa tgacagcagg 180

attatcgtaa tacgtaatag ttgaaaatct caaaaatgtg tgggtcatta cgtaaataat 240

gataggaatg ggattcttct atttttcctt tttccattct agcagccgtc gggaaaacgt 300

ggcatcctct ctttcgggct caattggagt cacgctgccg tgagcatcct ctctttccat 360

atctaacaac tgagcacgta accaatggaa aagcatgagc ttagcgttgc tccaaaaaag 420

tattggatgg ttaataccat ttgtctgttc tcttctgact ttgactcctc aaaaaaaaaa 480

aatctacaat caacagatcg cttcaattac gccctcacaa aaacttttt ccttcttctt 540

cgcccacgtt aaattttatc cctcatgttg tctaacggat ttctgcactt gatttattat 600

aaaaagacaa agacataata cttctctatc aatttcagtt attgttcttc cttgcgttat 660

tcttctgttc ttctttttct tttgtcatat ataaccataa ccaagtaata catattcaa 719

<210> 407
<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP821

<400> 407
aattgtttaa actaagcgaa tttcttatga tttatg 36

<210> 408
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP484

<400> 408
aattggccgg ccaaaaaaag catgcacgta taca 34

<210> 409
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP822

<400> 409
aattgcggcc gcgtgcctca tctatatttc tgaaatc 37

<210> 410
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP823

<400> 410
aattttaatt aatcagctct tgttttgttc tgcaaat 37

<210> 411
<211> 1647
<212> DNA
<213> Artificial sequence

<220>
<223> Listeria grayi ketoisovalerate decarboxylase coding sequence--codon optimized for expression in Saccharomyces cerevisiae

<400> 411
atgtataccg taggacagta cttggtagat agactagaag agattggtat cgataaggtt 60

ttcggtgtgc caggggatta caatttgact tttctagatt acattcaaaa tcacgaagga 120

ctttcctggc aagggaatac taatgaacta aacgcagcat atgcagcaga tggctacgcc 180

cgtgaaagag gcgtatcagc tcttgttact acattcggag tgggtgaact gtcagccatt 240

aacggaacag ctggtagttt tgcagaacaa gtccctgtca tccacatcgt gggttctcca 300

actatgaatg tgcaatccaa caaaaagctg gttcatcatt ccttaggaat gggtaacttt 360

cataacttta gtgaaatggc taaggaagtc actgccgcta caaccatgct tactgaagag 420

aatgcagctt cagagatcga cagagtatta gaaacagcct tgttggaaaa gaggccagta 480

tacatcaatc ttccaattga tatagctcat aaagcaatag ttaaacctgc aaaagcacta 540

caaacagaga aatcatctgg tgagagagag gcacaacttg cagaaatcat actatcacac 600

ttagaaaagg ccgctcaacc tatcgtaatc gccggtcatg agatcgcccg tttccagata 660

agagaaagat ttgaaaactg gataaaccaa acaaagttgc cagtaaccaa tttggcatat 720

ggcaaaggct ctttcaatga agagaacgaa catttcattg gtacctatta cccagctttt 780

tctgacaaaa acgttctgga ttacgttgac aatagtgact tcgttttaca ttttggtggg 840

aaaatcattg acaattctac ctcctcattt tctcaaggct ttaagactga aaacacttta 900

accgctgcaa atgacatcat tatgctgcca gatgggtcta cttactctgg gatttctctt 960

aacggtcttt tggcagagct ggaaaaacta aactttactt ttgctgatac tgctgctaaa 1020

caagctgaat tagctgtttt cgaaccacag gccgaaacac cactaaagca agacagattt 1080

caccaagctg ttatgaactt tttgcaagct gatgatgtgt tggtcactga gcaggggaca 1140

tcatctttcg gtttgatgtt ggcacctctg aaaaagggta tgaatttgat cagtcaaaca 1200

ttatggggct ccataggata cacattacct gctatgattg gttcacaaat tgctgcccca 1260

gaaaggagac acattctatc catcggtgat ggatcttttc aactgacagc acaggaaatg 1320

tccaccatct tcagagagaa attgacacca gtgatattca ttatcaataa cgatggctat 1380

acagtcgaaa gagccatcca tggagaggat gagagttaca atgatatacc aacttggaac 1440

ttgcaattag ttgctgaaac atttggtggt gatgccgaaa ctgtcgacac tcacaacgtt 1500

ttcacagaaa cagacttcgc taatacttta gctgctatcg atgctactcc tcaaaaagca 1560

catgtcgttg aagttcatat ggaacaaatg gatatgccag aatcattgag acagattggc 1620

ttagccttat ctaagcaaaa ctcttaa 1647

<210> 412
<211> 36
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP828

<400> 412
aattgtttaa acatgtatac cgtaggacag tacttg 36

<210> 413
<211> 35
<212> DNA
<213> Artificial sequence

<220>
<223> oBP829

<400> 413
aattgtttaa acttaagagt tttgcttaga taagg 35

<210> 414
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer FBAp-F

<400> 414
ctacaatcaa cagatcgctt caattacgc 29

<210> 415
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Primer HY-50

<400> 415
gagagttcac tgtttaacaa tagcaatgg 29

<210> 416
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer oBP834

<400> 416
tgctgatact gctgctaaac aagctg 26

<210> 417
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes KARI variant-K9JB4

<400> 417

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 418
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes KARI variant-K9JB4

<400> 418

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 419
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes KARI variant-K9JB4P

<400> 419

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 420
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes KARI variant K9JB4P

<400> 420

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 421
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9JG3

<400> 421

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Gly Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 422
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant K9JG3

<400> 422

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgaa ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg gagctgctaa gaaggctgac      240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 423
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9JG3P

<400> 423

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Gly Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 424
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-K9JG3P

<400> 424

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgaa ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg gagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 425
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9JA1

<400> 425

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Gly Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 426
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant K9JA1

<400> 426

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg gagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 427
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant- K9SB2

<400> 427

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 428
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9SB2

<400> 428

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 429
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9SB2-K90L

<400> 429

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Leu Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 430
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes KARI variant K9SB2-K90L

<400> 430

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaattg caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 431
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9SB2-K90M

<400> 431

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 432
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-K9SB2-K90M

<400> 432

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaatg caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct      540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa      600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt      660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc      720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct      780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg      840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg      900
```

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 433
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant- 2011_A2

<400> 433

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Asp Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 434
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_A3

<400> 434

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  435
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_A5

<400>  435

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly Asn Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  436
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant 2011_A7

<400>  436

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Val Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 437
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_A9

<400> 437

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Ile Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 438
<211> 343
<212> PRT
<213> Artificial sequence

<220>

20172000439 20 Jan 2017

<223> Anaerostipes caccae KARI variant-2011_A11

<400> 438

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Thr Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Glu Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

```
Ala Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 439
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B1

<400> 439

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175
```

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Thr Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 440
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B2

<400> 440

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Leu Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

20 Jan 2017
2017200439

<210> 441
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B3

<400> 441

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Leu Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255
```

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 442
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B4

<400> 442

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Thr Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140
```

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  443
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_B7

<400>  443

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Met Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

Glu Asp Lys Leu Ile Asn Asn
340

<210> 444
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B8

<400> 444

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Ser Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Tyr Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 445
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_B10

<400> 445

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Asp Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  446
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_C1

<400>  446

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

20172000439 20 Jan 2017

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Lys Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Asp Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Arg Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 447
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant 2011_C9

<400> 447

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Asp Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205
```

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 448
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C3

<400> 448

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Val Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95
```

```
Asn Asp Ile Glu Pro Tyr Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile His Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 449
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C6

<400> 449

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285
```

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 450
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C7

<400> 450

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175
```

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 451
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C8

<400> 451

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Val Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Tyr Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile His Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 452
<211> 343

20 Jan 2017 2017200439

<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C10

<400> 452

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Gly Glu Glu Trp Asn Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255
```

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 453
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_C12

<400> 453

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Thr Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 454
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D1

<400> 454

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Met Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Gly Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

Glu Asp Lys Leu Ile Asn Asn
340

<210> 455
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D2

<400> 455

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

```
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Ile His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 456
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D3

<400> 456

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Ile Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 457
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D4

<400> 457

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Leu Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Val Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 458
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D5

<400> 458

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Ser Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Ala Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  459
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_D8

<400>  459

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Thr Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 460
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D9

<400> 460

20172000439 20 Jan 2017

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285
```

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 461
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_D11

<400> 461

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Leu Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

```
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Ile Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  462
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_D12

<400>  462

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Lys Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 463
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_E3

<400> 463

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Val Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

```
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Trp Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 464
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_E4

<400> 464

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ser
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Arg
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  465
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_E8

<400>  465

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Leu
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 466
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_E11

<400> 466

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Leu Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240
```

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 467
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_F2

<400> 467

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Thr Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 468
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_F4

<400> 468

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 469
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_F6

<400> 469

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Gly Glu Glu Trp Asn Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 470
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_F9

<400> 470

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Val Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Pro Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 471
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_F10

<400> 471

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15
```

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Met Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  472
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_G1

<400>  472

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 473
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_G3

<400> 473

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly Gln Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 474
<211> 343
<212> PRT
<213> Artificial sequence

<220>

20172000439 20 Jan 2017

<223> Anaerostipes caccae KARI variant-2011_G4

<400> 474

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 475
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_G8

<400> 475

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Asn Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Val Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 476
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_G9

<400> 476

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Thr Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

20 Jan 2017
2017200439

<210> 477
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_G10

<400> 477

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255
```

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 478
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_H1

<400> 478

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140
```

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Leu Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  479
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2011_H5

<400>  479

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

Glu Asp Lys Leu Ile Asn Asn
340

<210> 480
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_H7

<400> 480

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Thr Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 481
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2011_H9

<400> 481

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Cys Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  482
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_A2

<400>  482

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Gly Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 483
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_A7

<400> 483

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Ser Ala Ala Glu Ala Val Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205
```

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 484
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_A8

<400> 484

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95
```

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Thr Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 485
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_A9

<400> 485

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly Gln Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Thr Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 486
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_A10

<400> 486

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Thr His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175
```

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 487
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_A11

<400> 487

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Thr Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 488
<211> 343

<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_B3

<400> 488

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Ser Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255
```

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 489
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_B4

<400> 489

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Arg Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Glu Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  490
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_B5

<400>  490

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Val His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

2017200439 20 Jan 2017

Glu Asp Lys Leu Ile Asn Asn
340

<210> 491
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_B6

<400> 491

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly His Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

```
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 492
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_B9

<400> 492

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Met Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 493
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_C2

<400> 493

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Leu Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

```
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 494
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_C3

<400> 494

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Val Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Ile
        195                 200                 205
```

```
Trp Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Ile Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  495
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_C5

<400>  495

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  496
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant

<400>  496
```

20172000439 20 Jan 2017

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Leu Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285
```

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 497
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_C8

<400> 497

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Thr Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

```
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 498
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_C10

<400> 498

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Ile Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Ile Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 499
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_D1

<400> 499

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Phe Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Thr Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 500
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant D6_2011

<400> 500

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Cys Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 501
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_D8

<400> 501

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Ser Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 502
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_D11

<400> 502

```
Met Glu Glu Cys Lys Met Ala Asn Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240
```

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 503
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_E5

<400> 503

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 504
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_E9

<400> 504

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

20 Jan 2017 2017200439

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Ile Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Met Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Asp Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 505
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_F1

<400> 505

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile Tyr Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 506
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_F2

<400> 506

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

20172000439 20 Jan 2017

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Asp Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  507
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_F3

<400>  507

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Met Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Ser Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  508
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_F4

<400>  508

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Asp Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 509
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_F7

<400> 509

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Val Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
85 90 95

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Ile Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 510
<211> 343
<212> PRT
<213> Artificial sequence

<220>

20172000439 20 Jan 2017

<223> Anaerostipes caccae KARI variant-2012_F10

<400> 510

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

His Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Ile Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 511
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_F12

<400> 511

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  512
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_G3

<400>  512

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Thr Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Ser Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 513
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_G4

<400> 513

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Ser Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 514
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_G5

<400> 514

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Leu Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140
```

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  515
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_G8

<400>  515

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Thr Ala Glu Ala Ala Asn Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Val Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Met Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Met
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

20 Jan 2017 2017200439

Glu Asp Lys Leu Ile Asn Asn
340

<210> 516
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_G9

<400> 516

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Ile His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 517
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_G10

<400> 517

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Ile Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  518
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2012_G12

<400>  518

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

20172000439 20 Jan 2017

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

Ser Gln Val His Ile Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 519
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_H1

<400> 519

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Thr Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ser Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 520
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_H3

<400> 520

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Cys Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95
```

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 521
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_H7

<400> 521

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Thr Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  522
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant

<400>  522

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Ile Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175
```

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Ser Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Ser Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 523
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2012_H11

<400> 523

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Val Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Met Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Met
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 524
<211> 343

<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_A2

<400> 524

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Thr Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255
```

```
Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile His
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 525
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant 2013_A4

<400> 525

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Met Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 526
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_A5

<400> 526

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Ser Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

Glu Asp Lys Leu Ile Asn Asn
340

<210> 527
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_B2

<400> 527

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Thr His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Ile Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

```
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 528
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_B5

<400> 528

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Thr Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 529
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_B7

<400> 529

```
Met Glu Glu Cys Lys Met Ala Lys Met Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Glu Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

```
Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  530
<211>  343
<212>  PRT
<213>  Artificial sequence

<220>
<223>  Anaerostipes caccae KARI variant-2013_B8

<400>  530

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Cys Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205
```

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 531
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_B9

<400> 531

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 532
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_B11

<400> 532

20172000439 20 Jan 2017

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Val Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285
```

```
Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 533
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_C1

<400> 533

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Leu Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

```
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Ile Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 534
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_C6

<400> 534

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Met Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80
```

```
Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Asp Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 535
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_C8

<400> 535

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Thr Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

```
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 536
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_C12

<400> 536

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Thr Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ser Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Val
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile Tyr Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 537
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-2013_D1

<400> 537

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Arg Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Met Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 538
<211> 636
<212> DNA
<213> Artificial sequence

<220>
<223> Promoter UAS(PGK1)-ENO2p

<400> 538
aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgc 120

ggcgttatgt cactaacgac gtgcaccaac ttgcggaaag tggaatcccg ttccaaaact 180

ggcatccact aattgataca tctacacacc gcacgccttt tttctgaagc ccactttcgt 240

ggactttgcc atatgcaaaa ttcatgaagt gtgataccaa gtcagcatac acctcactag 300

ggtagtttct ttggttgtat tgatcatttg gttcatcgtg gttcattaat ttttttttctc 360

cattgctttc tggctttgat cttactatca tttggatttt tgtcgaaggt tgtagaattg 420

tatgtgacaa gtggcaccaa gcatatataa aaaaaaaaag cattatcttc ctaccagagt 480

tgattgttaa aaacgtattt atagcaaacg caattgtaat taattcttat tttgtatctt 540

ttcttccctt gtctcaatct tttattttta ttttattttt cttttcttag tttctttcat 600

aacaccaagc aactaatact ataacataca ataata 636

<210> 539
<211> 979
<212> DNA
<213> Artificial sequence

<220>
<223> UAS(FBA1)-PDC1p

<400> 539
tgacagcagg attatcgtaa tacgtaatag ttgaaaatct caaaaatgtg tgggtcatta 60

cgtaaataat gataggaatg ggattcttct atttttcctt tttccattct cgcacgccga 120

aatgcatgca agtaacctat tcaaagtaat atctcataca tgtttcatga gggtaacaac 180

atgcgactgg gtgagcatat gttccgctga tgtgatgtgc aagataaaca agcaaggcag 240

aaactaactt cttcttcatg taataaacac accccgcgtt tatttaccta tctctaaact 300

tcaacacctt atatcataac taatatttct tgagataagc acactgcacc cataccttcc 360

ttaaaaacgt agcttccagt ttttggtggt tccggcttcc ttcccgattc cgcccgctaa 420

acgcatattt ttgttgcctg gtggcatttg caaaatgcat aacctatgca tttaaaagat 480

tatgtatgct cttctgactt ttcgtgtgat gaggctcgtg gaaaaaatga ataatttatg 540

aatttgagaa caattttgtg ttgttacggt attttactat ggaataatca atcaattgag 600

gattttatgc aaatatcgtt tgaatatttt tccgaccctt tgagtacttt tcttcataat 660

tgcataatat tgtccgctgc ccctttttct gttagacggt gtcttgatct acttgctatc 720

gttcaacacc accttatttt ctaactattt ttttttttagc tcatttgaat cagcttatgg 780

tgatggcaca tttttgcata aacctagctg tcctcgttga acataggaaa aaaaaatata 840

taaacaaggc tctttcactc tccttgcaat cagatttggg tttgttccct ttattttcat 900

atttcttgtc atattccttt ctcaattatt attttctact cataacctca cgcaaaataa 960

cacagtcaaa tcaatcaaa 979

<210> 540
<211> 994
<212> DNA
<213> Artificial sequence

<220>
<223> Promoter UAS(PGK1)-PDC1p

<400> 540
aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgt 120

cgacccgcac gccgaaatgc atgcaagtaa cctattcaaa gtaatatctc atacatgttt 180

catgagggta acaacatgcg actgggtgag catatgttcc gctgatgtga tgtgcaagat 240

aaacaagcaa ggcagaaact aacttcttct tcatgtaata aacacacccc gcgtttattt 300

acctatctct aaacttcaac accttatatc ataactaata tttcttgaga taagcacact 360

gcacccatac cttccttaaa aacgtagctt ccagtttttg gtggttccgg cttccttccc 420

gattccgccc gctaaacgca tatttttgtt gcctggtggc atttgcaaaa tgcataacct 480

atgcatttaa aagattatgt atgctcttct gacttttcgt gtgatgaggc tcgtggaaaa 540

aatgaataat ttatgaattt gagaacaatt ttgtgttgtt acggtatttt actatggaat 600

aatcaatcaa ttgaggattt tatgcaaata tcgtttgaat atttttccga ccctttgagt 660

acttttcttc ataattgcat aatattgtcc gctgcccctt tttctgttag acggtgtctt 720

gatctacttg ctatcgttca acaccacctt attttctaac tatttttttt ttagctcatt 780

tgaatcagct tatggtgatg gcacattttt gcataaacct agctgtcctc gttgaacata 840

ggaaaaaaaa atatataaac aaggctcttt cactctcctt gcaatcagat ttgggtttgt 900

tccctttatt ttcatatttc ttgtcatatt cctttctcaa ttattatttt ctactcataa 960

cctcacgcaa aataacacag tcaaatcaat caaa 994

<210> 541
<211> 731
<212> DNA
<213> Artificial sequence

<220>
<223> Promoter UAS(PGK)-OLE1p

<400> 541
aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc 60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgt 120

cgaccggggc gcccgtgcaa aaactaactc cgagcccggg catgtcccgg gttagcgggc 180

ccaacaaagg cgcttatctg gtgggcttcc gtagaagaaa aaaagctgtt gagcgagcta 240

tttcgggtat cccagccttc tctgcagacc gccccagttg gcttggctct ggtgctgttc 300

gttagcatca catcgcctgt gacaggcaga ggtaataacg gcttaaggtt ctcttcgcat 360

agtcggcagc tttctttcgg acgttgaaca ctcaacaaac cttatctagt gcccaaccag 420

gtgtgcttct acgagtcttg ctcactcaga cacacctatc cctattgtta cggctatggg 480

gatggcacac aaaggtggaa ataatagtag ttaacaatat atgcagcaaa tcatcggctc 540

ctggctcatc gagtcttgca aatcagcata tacatatata tatgggggca gatcttgatt 600

catttattgt tctatttcca tctttcctac ttctgtttcc gtttatattt tgtattacgt 660

agaatagaac atcatagtaa tagatagttg tggtgatcat attataaaca gcactaaaac 720

attacaacaa a 731

<210> 542
<211> 546
<212> PRT
<213> Macrococcus caseolyticus

<400> 542

Met Lys Gln Arg Ile Gly Gln Tyr Leu Ile Asp Ala Leu His Val Asn
1 5 10 15

Gly Val Asp Lys Ile Phe Gly Val Pro Gly Asp Phe Thr Leu Ala Phe
20 25 30

Leu Asp Asp Ile Ile Arg His Asp Asn Val Glu Trp Val Gly Asn Thr
35 40 45

Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Val Asn
50 55 60

Gly Leu Ala Ala Val Ser Thr Thr Phe Gly Val Gly Glu Leu Ser Ala
65 70 75 80

Val Asn Gly Ile Ala Gly Ser Tyr Ala Glu Arg Val Pro Val Ile Lys
85 90 95

Ile Ser Gly Gly Pro Ser Ser Val Ala Gln Gln Glu Gly Arg Tyr Val
100 105 110

His His Ser Leu Gly Glu Gly Ile Phe Asp Ser Tyr Ser Lys Met Tyr
115 120 125

Ala His Ile Thr Ala Thr Thr Thr Ile Leu Ser Val Asp Asn Ala Val
130 135 140

Asp Glu Ile Asp Arg Val Ile His Cys Ala Leu Lys Glu Lys Arg Pro
145 150 155 160

Val His Ile His Leu Pro Ile Asp Val Ala Leu Thr Glu Ile Glu Ile
165 170 175

Pro His Ala Pro Lys Val Tyr Thr His Glu Ser Gln Asn Val Asp Ala
180 185 190

Tyr Ile Gln Ala Val Glu Lys Lys Leu Met Ser Ala Lys Gln Pro Val
195 200 205

```
Ile Ile Ala Gly His Glu Ile Asn Ser Phe Lys Leu His Glu Gln Leu
    210                 215                 220

Glu Gln Phe Val Asn Gln Thr Asn Ile Pro Val Ala Gln Leu Ser Leu
225                 230                 235                 240

Gly Lys Ser Ala Phe Asn Glu Glu Asn Glu His Tyr Leu Gly Ile Tyr
                245                 250                 255

Asp Gly Lys Ile Ala Lys Glu Asn Val Arg Glu Tyr Val Asp Asn Ala
            260                 265                 270

Asp Val Ile Leu Asn Ile Gly Ala Lys Leu Thr Asp Ser Ala Thr Ala
        275                 280                 285

Gly Phe Ser Tyr Lys Phe Asp Thr Asn Asn Ile Ile Tyr Ile Asn His
    290                 295                 300

Asn Asp Phe Lys Ala Glu Asp Val Ile Ser Asp Asn Val Ser Leu Ile
305                 310                 315                 320

Asp Leu Val Asn Gly Leu Asn Ser Ile Asp Tyr Arg Asn Glu Thr His
                325                 330                 335

Tyr Pro Ser Tyr Gln Arg Ser Asp Met Lys Tyr Glu Leu Asn Asp Ala
            340                 345                 350

Pro Leu Thr Gln Ser Asn Tyr Phe Lys Met Met Asn Ala Phe Leu Glu
        355                 360                 365

Lys Asp Asp Ile Leu Leu Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
    370                 375                 380

Tyr Asp Leu Ser Leu Tyr Lys Gly Asn Gln Phe Ile Gly Gln Pro Leu
385                 390                 395                 400

Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ser Leu Leu Gly Ser Gln Leu
                405                 410                 415

Ala Asp Met His Arg Arg Asn Ile Leu Leu Ile Gly Asp Gly Ser Leu
            420                 425                 430

Gln Leu Thr Val Gln Ala Leu Ser Thr Met Ile Arg Lys Asp Ile Lys
        435                 440                 445

Pro Ile Ile Phe Val Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Leu
    450                 455                 460

Ile His Gly Met Glu Glu Pro Tyr Asn Asp Ile Gln Met Trp Asn Tyr
465                 470                 475                 480

Lys Gln Leu Pro Glu Val Phe Gly Gly Lys Asp Thr Val Lys Val His
                485                 490                 495
```

```
Asp Ala Lys Thr Ser Asn Glu Leu Lys Thr Val Met Asp Ser Val Lys
            500                 505                 510

Ala Asp Lys Asp His Met His Phe Ile Glu Val His Met Ala Val Glu
        515                 520                 525

Asp Ala Pro Lys Lys Leu Ile Asp Ile Ala Lys Ala Phe Ser Asp Ala
    530                 535                 540

Asn Lys
545
```

<210> 543
<211> 548
<212> PRT
<213> Listeria grayi

<400> 543

```
Met Tyr Thr Val Gly Gln Tyr Leu Val Asp Arg Leu Glu Glu Ile Gly
1               5                   10                  15

Ile Asp Lys Val Phe Gly Val Pro Gly Asp Tyr Asn Leu Thr Phe Leu
            20                  25                  30

Asp Tyr Ile Gln Asn His Glu Gly Leu Ser Trp Gln Gly Asn Thr Asn
        35                  40                  45

Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Glu Arg Gly
    50                  55                  60

Val Ser Ala Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Ile
65                  70                  75                  80

Asn Gly Thr Ala Gly Ser Phe Ala Glu Gln Val Pro Val Ile His Ile
                85                  90                  95

Val Gly Ser Pro Thr Met Asn Val Gln Ser Asn Lys Lys Leu Val His
            100                 105                 110

His Ser Leu Gly Met Gly Asn Phe His Asn Phe Ser Glu Met Ala Lys
        115                 120                 125

Glu Val Thr Ala Ala Thr Thr Met Leu Thr Glu Glu Asn Ala Ala Ser
    130                 135                 140

Glu Ile Asp Arg Val Leu Glu Thr Ala Leu Leu Glu Lys Arg Pro Val
145                 150                 155                 160

Tyr Ile Asn Leu Pro Ile Asp Ile Ala His Lys Ala Ile Val Lys Pro
                165                 170                 175

Ala Lys Ala Leu Gln Thr Glu Lys Ser Ser Gly Glu Arg Glu Ala Gln
            180                 185                 190
```

```
Leu Ala Glu Ile Ile Leu Ser His Leu Glu Lys Ala Ala Gln Pro Ile
        195                 200                 205

Val Ile Ala Gly His Glu Ile Ala Arg Phe Gln Ile Arg Glu Arg Phe
    210                 215                 220

Glu Asn Trp Ile Asn Gln Thr Lys Leu Pro Val Thr Asn Leu Ala Tyr
225                 230                 235                 240

Gly Lys Gly Ser Phe Asn Glu Glu Asn Glu His Phe Ile Gly Thr Tyr
                245                 250                 255

Tyr Pro Ala Phe Ser Asp Lys Asn Val Leu Asp Tyr Val Asp Asn Ser
            260                 265                 270

Asp Phe Val Leu His Phe Gly Gly Lys Ile Ile Asp Asn Ser Thr Ser
        275                 280                 285

Ser Phe Ser Gln Gly Phe Lys Thr Glu Asn Thr Leu Thr Ala Ala Asn
    290                 295                 300

Asp Ile Ile Met Leu Pro Asp Gly Ser Thr Tyr Ser Gly Ile Ser Leu
305                 310                 315                 320

Asn Gly Leu Leu Ala Glu Leu Glu Lys Leu Asn Phe Thr Phe Ala Asp
                325                 330                 335

Thr Ala Ala Lys Gln Ala Glu Leu Ala Val Phe Glu Pro Gln Ala Glu
            340                 345                 350

Thr Pro Leu Lys Gln Asp Arg Phe His Gln Ala Val Met Asn Phe Leu
        355                 360                 365

Gln Ala Asp Asp Val Leu Val Thr Glu Gln Gly Thr Ser Ser Phe Gly
    370                 375                 380

Leu Met Leu Ala Pro Leu Lys Lys Gly Met Asn Leu Ile Ser Gln Thr
385                 390                 395                 400

Leu Trp Gly Ser Ile Gly Tyr Thr Leu Pro Ala Met Ile Gly Ser Gln
                405                 410                 415

Ile Ala Ala Pro Glu Arg Arg His Ile Leu Ser Ile Gly Asp Gly Ser
            420                 425                 430

Phe Gln Leu Thr Ala Gln Glu Met Ser Thr Ile Phe Arg Glu Lys Leu
        435                 440                 445

Thr Pro Val Ile Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg
    450                 455                 460

Ala Ile His Gly Glu Asp Glu Ser Tyr Asn Asp Ile Pro Thr Trp Asn
465                 470                 475                 480
```

Leu Gln Leu Val Ala Glu Thr Phe Gly Gly Asp Ala Glu Thr Val Asp
485 490 495

Thr His Asn Val Phe Thr Glu Thr Asp Phe Ala Asn Thr Leu Ala Ala
500 505 510

Ile Asp Ala Thr Pro Gln Lys Ala His Val Val Glu Val His Met Glu
515 520 525

Gln Met Asp Met Pro Glu Ser Leu Arg Gln Ile Gly Leu Ala Leu Ser
530 535 540

Lys Gln Asn Ser
545

<210> 544
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53F_031411f

<400> 544
gttatcatcg gattattcga agga 24

<210> 545
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53L_031411f

<400> 545
gttatcatcg gattattgga agga 24

<210> 546
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53Y_031411f

<400> 546
gttatcatcg gattatatga agga 24

<210> 547
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53C_031411f

<400> 547
gttatcatcg gattatgtga agga 24

<210> 548
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53W_031411f

<400> 548
gttatcatcg gattatggga agga 24

<210> 549
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53P_031411f

<400> 549
gttatcatcg gattaccaga agga 24

<210> 550
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53H_031411f

<400> 550
gttatcatcg gattacatga agga 24

<210> 551
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53Q_031411f

<400> 551
gttatcatcg gattacaaga agga 24

<210> 552
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53I_031411f

<400> 552
gttatcatcg gattaattga agga 24

<210> 553
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53M_031411f

<400> 553
gttatcatcg gattaatgga agga 24

<210> 554
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53N_031411f

<400> 554
gttatcatcg gattaaatga agga 24

<210> 555
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53V_031411f

<400> 555
gttatcatcg gattagttga agga 24

<210> 556
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53A_031411f

<400> 556
gttatcatcg gattagctga agga 24

<210> 557
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53D_031411f

<400> 557
gttatcatcg gattagatga agga 24

<210> 558
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53E_031411f

<400> 558
gttatcatcg gattagaaga agga 24

<210> 559
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_53G_031411f

<400> 559
gttatcatcg gattaggtga agga 24

<210> 560
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56F_031411f

<400> 560
ggattacctg aaggattcaa a 21

<210> 561
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56L_031411f

<400> 561
ggattacctg aaggattgaa a 21

<210> 562
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56Y_031411f

<400> 562
ggattacctg aaggatataa a 21

<210> 563
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56C_031411f

<400> 563
ggattacctg aaggatgtaa a 21

<210> 564
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56W_031411f

<400> 564
ggattacctg aaggatggaa a 21

<210> 565
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56P_031411f

<400> 565
ggattacctg aaggaccaaa a 21

<210> 566
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56H_031411f

<400> 566
ggattacctg aaggacataa a 21

<210> 567
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56Q_031411f

<400> 567
ggattacctg aaggacaaaa a 21

<210> 568
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56I_031411f

<400> 568
ggattacctg aaggaattaa a 21

<210> 569
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56M_031411f

<400> 569
ggattacctg aaggaatgaa a 21

<210> 570
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56N_031411f

<400> 570
ggattacctg aaggaaataa a 21

<210> 571
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56V_031411f

<400> 571
ggattacctg aaggagttaa a 21

<210> 572
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56A_031411f

<400> 572
ggattacctg aaggagctaa a 21

<210> 573
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56D_031411f

<400> 573
ggattacctg aaggagataa a 21

<210> 574
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56E_031411f

<400> 574
ggattacctg aaggagaaaa a 21

<210> 575
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56G_031411f

<400> 575
ggattacctg aaggaggtaa a 21

<210> 576
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58F_051611f

<400> 576
gattacctga aggatttaaa ttctggaaga gagc 34

<210> 577
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58L_051611f

<400> 577
gattacctga aggatttaaa ttgtggaaga gagc 34

<210> 578
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58Y_051611f

<400> 578
gattacctga aggatttaaa tattggaaga gagc 34

<210> 579
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58C_051611f

<400> 579
gattacctga aggatttaaa tgttggaaga gagc 34

<210> 580
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58W_051611f

<400> 580
gattacctga aggatttaaa tggtggaaga gagc 34

<210> 581
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58P_051611f

<400> 581
gattacctga aggatttaaa ccatggaaga gagc 34

<210> 582
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58H_051611f

<400> 582
gattacctga aggatttaaa cattggaaga gagc 34

<210> 583
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58Q_051611f

<400> 583
gattacctga aggatttaaa caatggaaga gagc 34

<210> 584
<211> 34
<212> DNA
<213> Artificial sequence

20 Jan 2017 2017200439

<220>
<223> Primer K9_58I_051611f

<400> 584
gattacctga aggatttaaa atttggaaga gagc 34

<210> 585
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58M_051611f

<400> 585
gattacctga aggatttaaa atgtggaaga gagc 34

<210> 586
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58N_051611f

<400> 586
gattacctga aggatttaaa aattggaaga gagc 34

<210> 587
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58V_051611f

<400> 587
gattacctga aggatttaaa gtttggaaga gagc 34

<210> 588
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58A_051611f

<400> 588
gattacctga aggatttaaa gcttggaaga gagc 34

<210> 589
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58D_051611f

<400> 589
gattacctga aggatttaaa gattggaaga gagc 34

<210> 590
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58E_051611f

<400> 590
gattacctga aggatttaaa gaatggaaga gagc 34

<210> 591
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58G_051611f

<400> 591
gattacctga aggatttaaa ggttggaaga gagc 34

<210> 592
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56F_071211r

<400> 592
gctctcttcc atggtttgaa tccttc 26

<210> 593
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56L_071211r

<400> 593
gctctcttcc atggtttcaa tccttc 26

<210> 594
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56Y_071211r

<400> 594
gctctcttcc atggtttata tccttc 26

<210> 595
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56C_071211r

<400> 595
gctctcttcc atggtttaca tccttc 26

<210> 596
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56W_071211r

<400> 596
gctctcttcc atggtttcca tccttc 26

<210> 597
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56P_071211r

<400> 597
gctctcttcc atggttttgg tccttc 26

<210> 598
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56H_071211r

<400> 598
gctctcttcc atggtttatg tccttc 26

<210> 599
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56Q_071211r

<400> 599
gctctcttcc atggtttttg tccttc 26

<210> 600
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56I_071211r

<400> 600
gctctcttcc atggtttaat tccttc 26

<210> 601
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56M_071211r

<400> 601
gctctcttcc atggtttcat tccttc 26

<210> 602
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56N_071211r

<400> 602
gctctcttcc atggtttatt tccttc 26

<210> 603
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56V_071211r

<400> 603
gctctcttcc atggtttaac tccttc 26

<210> 604
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56A_071211r

<400> 604
gctctcttcc atggtttagc tccttc 26

<210> 605
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56D_071211r

<400> 605
gctctcttcc atggtttatc tccttc 26

<210> 606
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56E_071211r

<400> 606
gctctcttcc atggtttttc tccttc 26

<210> 607
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_56G_071211r

<400> 607
gctctcttcc atggtttacc tccttc 26

<210> 608
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58F_071211r

<400> 608
gttcttctgc tctcttccag aattt 25

<210> 609
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58L_071211r

<400> 609
gttcttctgc tctcttccac aattt 25

<210> 610
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58Y_071211r

<400> 610
gttcttctgc tctcttccaa tattt 25

<210> 611
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58C_071211r

<400> 611
gttcttctgc tctcttccaa cattt 25

<210> 612
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58W_071211r

<400> 612
gttcttctgc tctcttccac cattt 25

<210> 613
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58P_071211r

<400> 613
gttcttctgc tctcttccat ggttt 25

<210> 614
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58H_071211r

<400> 614
gttcttctgc tctcttccaa tgttt 25

<210> 615
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58Q_071211r

<400> 615
gttcttctgc tctcttccat tgttt 25

<210> 616
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58I_071211r

<400> 616
gttcttctgc tctcttccaa atttt 25

<210> 617
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58M_071211r

<400> 617
gttcttctgc tctcttccac atttt 25

<210> 618
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58N_071211r

<400> 618
gttcttctgc tctcttccaa ttttt 25

<210> 619
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58V_071211r

<400> 619
gttcttctgc tctcttccaa acttt 25

<210> 620
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58A_071211r

<400> 620
gttcttctgc tctcttccaa gcttt 25

<210> 621
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58D_071211r

<400> 621
gttcttctgc tctcttccaa tcttt 25

<210> 622
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58E_071211r

<400> 622
gttcttctgc tctcttccat tcttt 25

<210> 623
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Primer K9_58G_071211r

<400> 623
gttcttctgc tctcttccaa ccttt 25

<210> 624
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BI7D12

<400> 624

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Gln Glu Gly Val Lys Asp Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 625
<211> 1029

20172000439 20 Jan 2017

<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant- BI7D12

<400> 625
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttacaag aaggtgttaa ggactggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 626
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BI10F1

<400> 626

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Glu Glu Gly Val Lys Asp Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 627
<211> 1029

20172000439 20 Jan 2017

<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-BI10F1

<400> 627
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttagaag aaggtgttaa ggactggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 628
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BJ6G6

<400> 628

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Pro Glu Gly Asp Lys Gln Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 629
<211> 1029

<212> DNA
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BJ6G6

<400> 629
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttacctg aaggtgataa gcactggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 630
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BJ7D6

<400> 630

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Pro Glu Gly Val Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 631
<211> 1029

20172000439 20 Jan 2017

<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-BJ7D6

<400> 631
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttacctg aaggtgttaa ggactggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 632
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-BJ7F7

<400> 632

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Ala Glu Gly Asp Lys Gln Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Val Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Met Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Gly Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 633
<211> 1029

20172000439 20 Jan 2017

<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-BJ7F7

<400> 633
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttagctg aaggtgataa gcactggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttggtgcc ggatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc atgatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga ggggaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 634
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Primer pBAD_266f

<400> 634
ctctctactg tttctccata cccg 24

<210> 635
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> N-terminal truncated Anaerostipes caccae KARI variant

<400> 635

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

```
Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Tyr
        35                  40                  45

Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320
```

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 636
<211> 1017
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for N-terminal truncated Anaerostipes caccae KARI variant

<400> 636
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt atacgaaggt tctaagtcct ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcaacgatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 637
<211> 338
<212> PRT
<213> Artificial sequence

<220>
<223> N-terminal truncated Anaerostipes caccae KARI K9SB2 variant

<400> 637

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

```
Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320
```

20172000439 20 Jan 2017

```
Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 638
<211> 1017
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for N-terminal truncated Anaerostipes caccae K9SB2 KARI variant

<400> 638

```
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc       60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt      120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa      180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg      240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc      300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca      360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa      420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag      480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg      540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt      600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac      660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac      720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac      780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct      840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc      900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt      960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga        1017
```

<210> 639
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9SB2-T93I

<400> 639

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30
```

20 Jan 2017 2017200439

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ile Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

20 Jan 2017 2017200439

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 640
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant-K9SB2-T93I

<400> 640
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctatca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagga agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 641
<211> 343
<212> PRT
<213> Artificial sequence

<220>
<223> Anaerostipes caccae KARI variant-K9SB2-T93A

<400> 641

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

20 Jan 2017 2017200439

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Ala Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

20172000439 20 Jan 2017

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 642
<211> 1032
<212> DNA
<213> Artificial sequence

<220>
<223> coding sequence for Anaerostipes caccae KARI variant K9SB2-T93A

<400> 642
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctgcca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccgggt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 643
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae DSM 14662

<400> 643

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

```
Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ser Lys Ser Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320
```

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 644
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae KARI AB1G9 variant

<400> 644

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  645
<211>  343
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Anaerostipes caccae KARI-AB1D3 variant

<400>  645

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110
```

```
His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 646
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae KARI- K9D3 variant

<400> 646

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 647
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae KARI-K9G9 variant

<400> 647

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190
```

```
Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 648
<211> 1689
<212> DNA
<213> Artificial Sequence

<220>
<223> PDC1 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 648

```
atgtctgaaa ttactttggg taaatatttg ttcgaaagat taaagcaagt caacgttaac       60

accgttttcg gtttgccagg tgacttcaac ttgtccttgt tggacaagat ctacgaagtt      120

gaaggtatga gatgggctgg taacgccaac gaattgaacg ctgcttacgc cgctgatggt      180

tacgctcgta tcaagggtat gtcttgtatc atcaccacct tcggtgtcgg tgaattgtct      240

gctttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgttttgca cgttgttggt      300

gtcccatcca tctctgctca agctaagcaa ttgttgttgc accacacctt gggtaacggt      360

gacttcactg ttttccacag aatgtctgcc aacatttctg aaaccactgc tatgatcact      420

gacattgcta ccgccccagc tgaaattgac agatgtatca gaaccactta cgtcacccaa      480

agaccagtct acttaggttt gccagctaac ttggtcgact tgaacgtccc agctaagttg      540

ttgcaaactc caattgacat gtctttgaag ccaaacgatg ctgaatccga aaaggaagtc      600

attgacacca tcttggcttt ggtcaaggat gctaagaacc cagttatctt ggctgatgct      660

tgttgttcca gacacgacgt caaggctgaa actaagaagt tgattgactt gactcaattc      720
```

ccagctttcg tcaccccaat gggtaagggt tccattgacg aacaacaccc aagatacggt 780

ggtgtttacg tcggtacctt gtccaagcca gaagttaagg aagccgttga atctgctgac 840

ttgattttgt ctgtcggtgc tttgttgtct gatttcaaca ccggttcttt ctcttactct 900

tacaagacca agaacattgt cgaattccac tccgaccaca tgaagatcag aaacgccact 960

ttcccaggtg tccaaatgaa attcgttttg caaaagttgt tgaccactat tgctgacgcc 1020

gctaagggtt acaagccagt tgctgtccca gctagaactc cagctaacgc tgctgtccca 1080

gcttctaccc cattgaagca agaatggatg tggaaccaat tgggtaactt cttgcaagaa 1140

ggtgatgttg tcattgctga aaccggtacc tccgctttcg gtatcaacca aaccactttc 1200

ccaaacaaca cctacggtat ctctcaagtc ttatggggtt ccattggttt caccactggt 1260

gctaccttgg gtgctgcttt cgctgctgaa gaaattgatc caaagaagag agttatctta 1320

ttcattggtg acggttcttt gcaattgact gttcaagaaa tctccaccat gatcagatgg 1380

ggcttgaagc catacttgtt cgtcttgaac aacgatggtt acaccattga aaagttgatt 1440

cacggtccaa aggctcaata caacgaaatt caaggttggg accacctatc cttgttgcca 1500

actttcggtg ctaaggacta tgaaacccac agagtcgcta ccaccggtga atgggacaag 1560

ttgacccaag acaagtcttt caacgacaac tctaagatca gaatgattga aatcatgttg 1620

ccagtcttcg atgctccaca aaacttggtt gaacaagcta agttgactgc tgctaccaac 1680

gctaagcaa 1689

<210> 649
<211> 563
<212> PRT
<213> Artificial Sequence

<220>
<223> PDC1 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 649

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1 5 10 15

Val Asn Val Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Glu Gly Met Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
50 55 60

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
85 90 95

His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
100 105 110

```
Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Ala Thr
    130                 135                 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Val Thr Gln
145                 150                 155                 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Asn Val
                165                 170                 175

Pro Ala Lys Leu Leu Gln Thr Pro Ile Asp Met Ser Leu Lys Pro Asn
            180                 185                 190

Asp Ala Glu Ser Glu Lys Glu Val Ile Asp Thr Ile Leu Ala Leu Val
        195                 200                 205

Lys Asp Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Cys Ser Arg
    210                 215                 220

His Asp Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe
225                 230                 235                 240

Pro Ala Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln His
                245                 250                 255

Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Pro Glu Val
            260                 265                 270

Lys Glu Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Ile Val Glu Phe His Ser Asp His Met Lys Ile Arg Asn Ala Thr
305                 310                 315                 320

Phe Pro Gly Val Gln Met Lys Phe Val Leu Gln Lys Leu Leu Thr Thr
                325                 330                 335

Ile Ala Asp Ala Ala Lys Gly Tyr Lys Pro Val Ala Val Pro Ala Arg
            340                 345                 350

Thr Pro Ala Asn Ala Ala Val Pro Ala Ser Thr Pro Leu Lys Gln Glu
        355                 360                 365

Trp Met Trp Asn Gln Leu Gly Asn Phe Leu Gln Glu Gly Asp Val Val
    370                 375                 380

Ile Ala Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Thr Phe
385                 390                 395                 400
```

Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
405 410 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
420 425 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
435 440 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
450 455 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
465 470 475 480

His Gly Pro Lys Ala Gln Tyr Asn Glu Ile Gln Gly Trp Asp His Leu
485 490 495

Ser Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Thr His Arg Val
500 505 510

Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Ser Phe Asn
515 520 525

Asp Asn Ser Lys Ile Arg Met Ile Glu Ile Met Leu Pro Val Phe Asp
530 535 540

Ala Pro Gln Asn Leu Val Glu Gln Ala Lys Leu Thr Ala Ala Thr Asn
545 550 555 560

Ala Lys Gln

<210> 650
<211> 1689
<212> DNA
<213> Artificial Sequence

<220>
<223> PDC5 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 650
atgtctgaaa taaccttagg taaatattta tttgaaagat tgagccaagt caactgtaac 60

accgtcttcg gtttgccagg tgactttaac ttgtctcttt tggataagct ttatgaagtc 120

aaaggtatga gatgggctgg taacgctaac gaattgaacg ctgcctatgc tgctgatggt 180

tacgctcgta tcaagggtat gtcctgtatt attaccacct tcggtgttgg tgaattgtct 240

gctttgaatg gtattgccgg ttcttacgct gaacatgtcg gtgttttgca cgttgttggt 300

gttccatcca tctcttctca agctaagcaa ttgttgttgc atcatacctt gggtaacggt 360

gacttcactg ttttccacag aatgtctgcc aacatttctg aaaccactgc catgatcact 420

gatattgcta acgctccagc tgaaattgac agatgtatca gaaccaccta cactacccaa 480

agaccagtct acttgggttt gccagctaac ttggttgact tgaacgtccc agccaagtta 540

ttggaaactc caattgactt gtctttgaag ccaaacgacg ctgaagctga agctgaagtt 600
gttagaactg ttgttgaatt gatcaaggat gctaagaacc cagttatctt ggctgatgct 660
tgtgcttcta gacatgatgt caaggctgaa actaagaagt tgatggactt gactcaattc 720
ccagtttacg tcaccccaat gggtaagggt gctattgacg aacaacaccc aagatacggt 780
ggtgtttacg ttggtacctt gtctagacca gaagttaaga aggctgtaga atctgctgat 840
ttgatattgt ctatcggtgc tttgttgtct gatttcaata ccggttcttt ctcttactcc 900
tacaagacca aaaatatcgt tgaattccac tctgaccaca tcaagatcag aaacgccacc 960
ttcccaggtg ttcaaatgaa atttgccttg caaaaattgt tggatgctat tccagaagtc 1020
gtcaaggact acaaacctgt tgctgtccca gctagagttc caattaccaa gtctactcca 1080
gctaacactc caatgaagca agaatggatg tggaaccatt tgggtaactt cttgagagaa 1140
ggtgatattg ttattgctga aaccggtact tccgccttcg gtattaacca aactactttc 1200
ccaacagatg tatacgctat cgtccaagtc ttgtggggtt ccattggttt cacagtcggc 1260
gctctattgg gtgctactat ggccgctgaa gaacttgatc caaagaagag agttatttta 1320
ttcattggtg acggttctct acaattgact gttcaagaaa tctctaccat gattagatgg 1380
ggtttgaagc catacatttt tgtcttgaat aacaacggtt acaccattga aaaattgatt 1440
cacggtcctc atgccgaata taatgaaatt caaggttggg accacttggc cttattgcca 1500
acttttggtg ctagaaacta cgaaacccac agagttgcta ccactggtga atgggaaaag 1560
ttgactcaag acaaggactt ccaagacaac tctaagatta gaatgattga agttatgttg 1620
ccagtctttg atgctccaca aaacttggtt aaacaagctc aattgactgc cgctactaac 1680
gctaaacaa 1689

<210> 651
<211> 563
<212> PRT
<213> Artificial Sequence

<220>
<223> PDC5 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 651

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Ser Gln
1 5 10 15

Val Asn Cys Asn Thr Val Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Leu Tyr Glu Val Lys Gly Met Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
50 55 60

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

```
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
                85                  90                  95

His Val Val Gly Val Pro Ser Ile Ser Ser Gln Ala Lys Gln Leu Leu
            100                 105                 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Ala Asn
    130                 135                 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Thr Thr Gln
145                 150                 155                 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Asn Val
                165                 170                 175

Pro Ala Lys Leu Leu Glu Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn
            180                 185                 190

Asp Ala Glu Ala Glu Ala Glu Val Val Arg Thr Val Val Glu Leu Ile
        195                 200                 205

Lys Asp Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Ala Ser Arg
    210                 215                 220

His Asp Val Lys Ala Glu Thr Lys Lys Leu Met Asp Leu Thr Gln Phe
225                 230                 235                 240

Pro Val Tyr Val Thr Pro Met Gly Lys Gly Ala Ile Asp Glu Gln His
                245                 250                 255

Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Arg Pro Glu Val
            260                 265                 270

Lys Lys Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Ile Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Ile Val Glu Phe His Ser Asp His Ile Lys Ile Arg Asn Ala Thr
305                 310                 315                 320

Phe Pro Gly Val Gln Met Lys Phe Ala Leu Gln Lys Leu Leu Asp Ala
                325                 330                 335

Ile Pro Glu Val Val Lys Asp Tyr Lys Pro Val Ala Val Pro Ala Arg
            340                 345                 350

Val Pro Ile Thr Lys Ser Thr Pro Ala Asn Thr Pro Met Lys Gln Glu
        355                 360                 365
```

```
Trp Met Trp Asn His Leu Gly Asn Phe Leu Arg Glu Gly Asp Ile Val
    370                 375                 380

Ile Ala Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Thr Phe
385                 390                 395                 400

Pro Thr Asp Val Tyr Ala Ile Val Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Phe Thr Val Gly Ala Leu Leu Gly Ala Thr Met Ala Ala Glu Glu Leu
            420                 425                 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Ile Phe Val Leu Asn Asn Asn Gly Tyr Thr Ile Glu Lys Leu Ile
465                 470                 475                 480

His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Gly Trp Asp His Leu
                485                 490                 495

Ala Leu Leu Pro Thr Phe Gly Ala Arg Asn Tyr Glu Thr His Arg Val
            500                 505                 510

Ala Thr Thr Gly Glu Trp Glu Lys Leu Thr Gln Asp Lys Asp Phe Gln
        515                 520                 525

Asp Asn Ser Lys Ile Arg Met Ile Glu Val Met Leu Pro Val Phe Asp
    530                 535                 540

Ala Pro Gln Asn Leu Val Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
545                 550                 555                 560

Ala Lys Gln
```

<210> 652
<211> 1599
<212> DNA
<213> Artificial Sequence

<220>
<223> PDC6 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 652

```
atgtctgaaa ttactcttgg aaaatactta tttgaaagat tgaagcaagt taatgttaac       60

accatttttg ggctaccagg cgacttcaac ttgtccctat tggacaagat ttacgaggta      120

gatggattga gatgggctgg taatgcaaat gagctgaacg ccgcctatgc cgccgatggt      180

tacgcacgca tcaagggttt atctgtgctg gtaactactt ttggcgtagg tgaattatcc      240

gccttgaatg gtattgcagg atcgtatgca gaacacgtcg gtgtactgca tgttgttggt      300

gtcccctcta tctccgctca ggctaagcaa ttgttgttgc atcatacctt gggtaacggt      360
```

gattttaccg tttttcacag aatgtccgcc aatatctcag aaactacatc aatgattaca 420
gacattgcta cagccccttc agaaatcgat aggttgatca ggacaacatt tataacacaa 480
aggcctagct acttggggtt gccagcgaat ttggtagatc taaaggttcc tggttctctt 540
ttggaaaaac cgattgatct atcattaaaa cctaacgatc ccgaagctga aaaggaagtt 600
attgataccg tactagaatt gatccagaat tcgaaaaacc ctgttatact atcggatgcc 660
tgtgcttcta ggcacaacgt taaaaaagaa acccagaagt taattgattt gacgcaattc 720
ccagcttttg tgacacctct aggtaaaggg tcaatagatg aacagcatcc cagatatggc 780
ggtgtttatg tgggaacgct gtccaaacaa gacgtgaaac aggccgttga gtcggctgat 840
ttgatccttt cggtcggtgc tttgctctct gattttaaca caggttcgtt ttcctactcc 900
tacaagacta aaaatgtagt ggagtttcat tccgattacg taaaggtgaa gaacgctacg 960
ttcctcggtg tacaaatgaa atttgcacta caaaacttac tgaaggttat tcccgatgtt 1020
gttaagggct acaagagcgt tcccgtacca accaaaactc ccgcaaacaa aggtgtacct 1080
gctagcacgc ccttgaaaca agagtggttg tggaacgaat tgtccaaatt cttgcaagaa 1140
ggtgatgtta tcatttccga gaccggcacg tctgccttcg gtatcaatca aactatcttt 1200
cctaaggacg cctacggtat ctcgcaggtg ttgtgggggt ccatcggttt tacaacagga 1260
gcaactttag gtgctgcctt tgccgctgag gagattgacc ccaacaagag agtcatctta 1320
ttcataggtg acgggtcttt gcagttaacc gtccaagaaa tctccaccat gatcagatgg 1380
gggttaaagc cgtatctttt tgtccttaac aacgacggct acactatcga aaagctgatt 1440
catgggcctc acgcagagta caacgaaatc cagacctggg atcacctcgc cctgttgccc 1500
gcatttggtg cgaaaaagta cgaaaatcac aagatcgcca ctacgggtga gtgggatgcc 1560
ttaaccactg attcagagtt ccagaaaaac tcggtgatc 1599

<210> 653
<211> 533
<212> PRT
<213> Artificial Sequence

<220>
<223> PDC6 pyruvate decarboxylase from Saccharomyces cerevisiae

<400> 653

Met Ser Glu Ile Thr Leu Gly Lys Tyr Leu Phe Glu Arg Leu Lys Gln
1 5 10 15

Val Asn Val Asn Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Asp Gly Leu Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
50 55 60

Lys Gly Leu Ser Val Leu Val Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

```
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
                85                  90                  95

His Val Val Gly Val Pro Ser Ile Ser Ala Gln Ala Lys Gln Leu Leu
            100                 105                 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

Ser Ala Asn Ile Ser Glu Thr Thr Ser Met Ile Thr Asp Ile Ala Thr
    130                 135                 140

Ala Pro Ser Glu Ile Asp Arg Leu Ile Arg Thr Thr Phe Ile Thr Gln
145                 150                 155                 160

Arg Pro Ser Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val
                165                 170                 175

Pro Gly Ser Leu Leu Glu Lys Pro Ile Asp Leu Ser Leu Lys Pro Asn
            180                 185                 190

Asp Pro Glu Ala Glu Lys Glu Val Ile Asp Thr Val Leu Glu Leu Ile
        195                 200                 205

Gln Asn Ser Lys Asn Pro Val Ile Leu Ser Asp Ala Cys Ala Ser Arg
    210                 215                 220

His Asn Val Lys Lys Glu Thr Gln Lys Leu Ile Asp Leu Thr Gln Phe
225                 230                 235                 240

Pro Ala Phe Val Thr Pro Leu Gly Lys Gly Ser Ile Asp Glu Gln His
                245                 250                 255

Pro Arg Tyr Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Gln Asp Val
            260                 265                 270

Lys Gln Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Val Val Glu Phe His Ser Asp Tyr Val Lys Val Lys Asn Ala Thr
305                 310                 315                 320

Phe Leu Gly Val Gln Met Lys Phe Ala Leu Gln Asn Leu Leu Lys Val
                325                 330                 335

Ile Pro Asp Val Val Lys Gly Tyr Lys Ser Val Pro Val Pro Thr Lys
            340                 345                 350

Thr Pro Ala Asn Lys Gly Val Pro Ala Ser Thr Pro Leu Lys Gln Glu
        355                 360                 365
```

```
Trp Leu Trp Asn Glu Leu Ser Lys Phe Leu Gln Glu Gly Asp Val Ile
    370                 375                 380

Ile Ser Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Ile Phe
385                 390                 395                 400

Pro Lys Asp Ala Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

Asp Pro Asn Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Lys Leu Ile
465                 470                 475                 480

His Gly Pro His Ala Glu Tyr Asn Glu Ile Gln Thr Trp Asp His Leu
                485                 490                 495

Ala Leu Leu Pro Ala Phe Gly Ala Lys Lys Tyr Glu Asn His Lys Ile
            500                 505                 510

Ala Thr Thr Gly Glu Trp Asp Ala Leu Thr Thr Asp Ser Glu Phe Gln
        515                 520                 525

Lys Asn Ser Val Ile
    530
```

<210> 654
<211> 1692
<212> DNA
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Candida glabrata

<400> 654

```
atgtctgaga ttactttggg tagatacttg ttcgagagat tgaaccaagt cgacgttaag       60

accatcttcg gtttgccagg tgacttcaac ttgtccctat tggacaagat ctacgaagtt      120

gaaggtatga gatgggctgg taacgctaac gaattgaacg ctgcttacgc tgctgacggt      180

tacgctagaa tcaagggtat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct      240

gccttgaacg gtattgccgg ttcttacgct gaacacgtcg gtgtcttgca cgtcgtcggt      300

gtcccatcca tctcctctca agctaagcaa ttgttgttgc accacacctt gggtaacggt      360

gacttcactg tcttccacag aatgtccgct aacatctctg agaccaccgc tatggtcact      420

gacatcgcta ccgctccagc tgagatcgac agatgtatca gaaccaccta catcacccaa      480

agaccagtct acttgggtct accagctaac ttggtcgacc taaaggtccc agccaagctt      540
```

ttggaaaccc caattgactt gtccttgaag ccaaacgacc cagaagccga aactgaagtc 600
gttgacaccg tcttggaatt gatcaaggct gctaagaacc cagttatctt ggctgatgct 660
tgtgcttcca gacacgacgt caaggctgaa accaagaagt tgattgacgc cactcaattc 720
ccatccttcg ttaccccaat gggtaagggt tccatcgacg aacaacaccc aagattcggt 780
ggtgtctacg tcggtacctt gtccagacca gaagttaagg aagctgttga atccgctgac 840
ttgatcttgt ctgtcggtgc tttgttgtcc gatttcaaca ctggttcttt ctcttactct 900
tacaagacca agaacatcgt cgaattccac tctgactaca tcaagatcag aaacgctacc 960
ttcccaggtg tccaaatgaa gttcgctttg caaaagttgt tgaacgccgt cccagaagct 1020
atcaagggtt acaagccagt ccctgtccca gctagagtcc cagaaaacaa gtcctgtgac 1080
ccagctaccc cattgaagca agaatggatg tggaaccaag tttccaagtt cttgcaagaa 1140
ggtgatgttg ttatcactga aaccggtacc tccgcttttg gtatcaacca aaccccattc 1200
ccaaacaacg cttacggtat ctcccaagtt ctatggggtt ccatcggttt caccaccggt 1260
gcttgtttgg gtgccgcttt cgctgctgaa gaaatcgacc caaagaagag agttatcttg 1320
ttcattggtg acggttcttt gcaattgact gtccaagaaa tctccaccat gatcagatgg 1380
ggcttgaagc catacttgtt cgtcttgaac aacgacggtt acaccatcga aagattgatt 1440
cacggtgaaa aggctggtta caacgacatc caaaactggg accacttggc tctattgcca 1500
accttcggtg ctaaggacta cgaaaaccac agagtcgcca ccaccggtga atgggacaag 1560
ttgacccaag acaaggaatt caacaagaac tccaagatca gaatgatcga agttatgttg 1620
ccagttatgg acgctccaac ttccttgatt gaacaagcta agttgaccgc ttccatcaac 1680
gctaagcaag aa 1692

<210> 655
<211> 564
<212> PRT
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Candida glabrata

<400> 655

Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Asn Gln
1 5 10 15

Val Asp Val Lys Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Glu Gly Met Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Ile
50 55 60

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

20 Jan 2017 2017200439

```
Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
                85                  90                  95

His Val Val Gly Val Pro Ser Ile Ser Ser Gln Ala Lys Gln Leu Leu
            100                 105                 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Val Thr Asp Ile Ala Thr
    130                 135                 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Ile Thr Gln
145                 150                 155                 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Lys Val
                165                 170                 175

Pro Ala Lys Leu Leu Glu Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn
            180                 185                 190

Asp Pro Glu Ala Glu Thr Glu Val Val Asp Thr Val Leu Glu Leu Ile
        195                 200                 205

Lys Ala Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Ala Ser Arg
    210                 215                 220

His Asp Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Ala Thr Gln Phe
225                 230                 235                 240

Pro Ser Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln His
                245                 250                 255

Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Arg Pro Glu Val
            260                 265                 270

Lys Glu Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Ile Val Glu Phe His Ser Asp Tyr Ile Lys Ile Arg Asn Ala Thr
305                 310                 315                 320

Phe Pro Gly Val Gln Met Lys Phe Ala Leu Gln Lys Leu Leu Asn Ala
                325                 330                 335

Val Pro Glu Ala Ile Lys Gly Tyr Lys Pro Val Pro Val Pro Ala Arg
            340                 345                 350

Val Pro Glu Asn Lys Ser Cys Asp Pro Ala Thr Pro Leu Lys Gln Glu
        355                 360                 365
```

```
Trp Met Trp Asn Gln Val Ser Lys Phe Leu Gln Glu Gly Asp Val Val
    370                 375                 380

Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr Pro Phe
385                 390                 395                 400

Pro Asn Asn Ala Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Phe Thr Thr Gly Ala Cys Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
465                 470                 475                 480

His Gly Glu Lys Ala Gly Tyr Asn Asp Ile Gln Asn Trp Asp His Leu
                485                 490                 495

Ala Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Asn His Arg Val
            500                 505                 510

Ala Thr Thr Gly Glu Trp Asp Lys Leu Thr Gln Asp Lys Glu Phe Asn
        515                 520                 525

Lys Asn Ser Lys Ile Arg Met Ile Glu Val Met Leu Pro Val Met Asp
    530                 535                 540

Ala Pro Thr Ser Leu Ile Glu Gln Ala Lys Leu Thr Ala Ser Ile Asn
545                 550                 555                 560

Ala Lys Gln Glu
```

<210> 656
<211> 1788
<212> DNA
<213> Artificial Sequence

<220>
<223> PDC1 pyruvate decarboxylase from Pichia stipitis

<400> 656

```
atggctgaag tctcattagg aagatatctc ttcgagagat tgtaccaatt gcaagtgcag       60

accatcttcg gtgtccctgg tgatttcaac ttgtcgcttt tggacaagat ctacgaagtg      120

gaagatgccc atggcaagaa ttcgtttaga tgggctggta atgccaacga attgaatgca      180

tcgtacgctg ctgacggtta ctcgagagtc aagcgtttag ggtgtttggt cactaccttt      240

ggtgtcggtg aattgtctgc tttgaatggt attgccggtt cttatgccga acatgttggt      300

ttgcttcatg tcgtaggtgt tccatcgatt tcctcgcaag ctaagcaatt gttacttcac      360
```

20172004439 20 Jan 2017

cacactttgg gtaatggtga tttcactgtt ttccatagaa tgtccaacaa catttctcag 420

accacagcct ttatctccga tatcaactcg gctccagctg aaattgatag atgtatcaga 480

gaggcctacg tcaaacaaag accagtttat atcgggttac cagctaactt agttgatttg 540

aatgttccgg cctctttgct tgagtctcca atcaacttgt cgttggaaaa gaacgaccca 600

gaggctcaag atgaagtcat tgactctgtc ttagacttga tcaaaaagtc gctgaaccca 660

atcatcttgg tcgatgcctg tgcctcgaga catgactgta aggctgaagt tactcagttg 720

attgaacaaa cccaattccc agtatttgtc actccaatgg gtaaaggtac cgttgatgag 780

ggtggtgtag acggagaatt gttagaagat gatcctcatt tgattgccaa ggtcgctgct 840

aggttgtctg ctggcaagaa cgctgcctct agattcggag gtgtttatgt cggaaccttg 900

tcgaagcccg aagtcaagga cgctgtagag agtgcagatt tgattttgtc tgtcggtgcc 960

cttttgtctg atttcaacac tggttcattt tcctactcct acagaaccaa gaacatcgtc 1020

gaattccatt ctgattacac taagattaga caagccactt tcccaggtgt gcagatgaag 1080

gaagccttgc aagaattgaa caagaaagtt tcatctgctg ctagtcacta tgaagtcaag 1140

cctgtgccca agatcaagtt ggccaataca ccagccacca gagaagtcaa gttaactcag 1200

gaatggttgt ggaccagagt gtcttcgtgg ttcagagaag gtgatattat tatcaccgaa 1260

accggtacat cctccttcgg tatagttcaa tccagattcc caaacaacac catcggtatc 1320

tcccaagtat tgtggggttc tattggtttc tctgttggtg ccactttggg tgctgccatg 1380

gctgcccaag aactcgaccc taacaagaga accatcttgt ttgttggaga tggttctttg 1440

caattgaccg ttcaggaaat ctccaccata atcagatggg gtaccacacc ttacctttc 1500

gtgttgaaca atgacggtta caccatcgag cgtttgatcc acggtgtaaa tgcctcatat 1560

aatgacatcc aaccatggca aaacttggaa atcttgccta ctttctcggc caagaactac 1620

gacgctgtga gaatctccaa catcggagaa gcagaagata tcttgaaaga caaggaattc 1680

ggaaagaact ccaagattag attgatagaa gtcatgttac caagattgga tgcaccatct 1740

aaccttgcca aacaagctgc cattacagct gccaccaacg ccgaagct 1788

<210> 657
<211> 596
<212> PRT
<213> Artificial Sequence

<220>
<223> PDC1 pyruvate decarboxylase from Pichia stipitis

<400> 657

Met Ala Glu Val Ser Leu Gly Arg Tyr Leu Phe Glu Arg Leu Tyr Gln
1 5 10 15

Leu Gln Val Gln Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Ile Tyr Glu Val Glu Asp Ala His Gly Lys Asn Ser
35 40 45

2017200439 20 Jan 2017

```
Phe Arg Trp Ala Gly Asn Ala Asn Glu Leu Asn Ala Ser Tyr Ala Ala
    50                  55                  60

Asp Gly Tyr Ser Arg Val Lys Arg Leu Gly Cys Leu Val Thr Thr Phe
65                  70                  75                  80

Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala
                85                  90                  95

Glu His Val Gly Leu Leu His Val Val Gly Val Pro Ser Ile Ser Ser
            100                 105                 110

Gln Ala Lys Gln Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe
        115                 120                 125

Thr Val Phe His Arg Met Ser Asn Asn Ile Ser Gln Thr Thr Ala Phe
    130                 135                 140

Ile Ser Asp Ile Asn Ser Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg
145                 150                 155                 160

Glu Ala Tyr Val Lys Gln Arg Pro Val Tyr Ile Gly Leu Pro Ala Asn
                165                 170                 175

Leu Val Asp Leu Asn Val Pro Ala Ser Leu Leu Glu Ser Pro Ile Asn
            180                 185                 190

Leu Ser Leu Glu Lys Asn Asp Pro Glu Ala Gln Asp Glu Val Ile Asp
        195                 200                 205

Ser Val Leu Asp Leu Ile Lys Lys Ser Ser Asn Pro Ile Ile Leu Val
    210                 215                 220

Asp Ala Cys Ala Ser Arg His Asp Cys Lys Ala Glu Val Thr Gln Leu
225                 230                 235                 240

Ile Glu Gln Thr Gln Phe Pro Val Phe Val Thr Pro Met Gly Lys Gly
                245                 250                 255

Thr Val Asp Glu Gly Gly Val Asp Gly Glu Leu Leu Glu Asp Asp Pro
            260                 265                 270

His Leu Ile Ala Lys Val Ala Ala Arg Leu Ser Ala Gly Lys Asn Ala
        275                 280                 285

Ala Ser Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Lys Pro Glu
    290                 295                 300

Val Lys Asp Ala Val Glu Ser Ala Asp Leu Ile Leu Ser Val Gly Ala
305                 310                 315                 320

Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Arg Thr
                325                 330                 335
```

```
Lys Asn Ile Val Glu Phe His Ser Asp Tyr Thr Lys Ile Arg Gln Ala
            340                 345                 350

Thr Phe Pro Gly Val Gln Met Lys Glu Ala Leu Gln Glu Leu Asn Lys
        355                 360                 365

Lys Val Ser Ser Ala Ala Ser His Tyr Glu Val Lys Pro Val Pro Lys
    370                 375                 380

Ile Lys Leu Ala Asn Thr Pro Ala Thr Arg Glu Val Lys Leu Thr Gln
385                 390                 395                 400

Glu Trp Leu Trp Thr Arg Val Ser Ser Trp Phe Arg Glu Gly Asp Ile
                405                 410                 415

Ile Ile Thr Glu Thr Gly Thr Ser Ser Phe Gly Ile Val Gln Ser Arg
            420                 425                 430

Phe Pro Asn Asn Thr Ile Gly Ile Ser Gln Val Leu Trp Gly Ser Ile
        435                 440                 445

Gly Phe Ser Val Gly Ala Thr Leu Gly Ala Ala Met Ala Ala Gln Glu
    450                 455                 460

Leu Asp Pro Asn Lys Arg Thr Ile Leu Phe Val Gly Asp Gly Ser Leu
465                 470                 475                 480

Gln Leu Thr Val Gln Glu Ile Ser Thr Ile Ile Arg Trp Gly Thr Thr
                485                 490                 495

Pro Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu
            500                 505                 510

Ile His Gly Val Asn Ala Ser Tyr Asn Asp Ile Gln Pro Trp Gln Asn
        515                 520                 525

Leu Glu Ile Leu Pro Thr Phe Ser Ala Lys Asn Tyr Asp Ala Val Arg
    530                 535                 540

Ile Ser Asn Ile Gly Glu Ala Glu Asp Ile Leu Lys Asp Lys Glu Phe
545                 550                 555                 560

Gly Lys Asn Ser Lys Ile Arg Leu Ile Glu Val Met Leu Pro Arg Leu
                565                 570                 575

Asp Ala Pro Ser Asn Leu Ala Lys Gln Ala Ala Ile Thr Ala Ala Thr
            580                 585                 590

Asn Ala Glu Ala
        595

<210>  658
<211>  1707
<212>  DNA
<213>  Artificial Sequence
```

20172004 39 20 Jan 2017

<220>
<223> PDC2 pyruvate decarboxylase from Pichia stipitis

<400> 658
atggtatcaa cctacccaga atcagaggtt actctaggaa ggtacctctt tgagcgactc 60

caccaattga aagtggacac cattttcggc ttgccgggtg acttcaacct ttccttattg 120

gacaaagtgt atgaagttcc ggatatgagg tgggctggaa atgccaacga attgaatgct 180

gcctatgctg ccgatggtta ctccagaata aagggattgt cttgcttggt cacaactttt 240

ggtgttggtg aattgtctgc tttaaacgga gttggtggtg cctatgctga acacgtagga 300

cttctacatg tcgttggagt tccatccata tcgtcacagg ctaaacagtt gttgctccac 360

cataccttgg gtaatggtga cttcactgtt tttcacagaa tgtccaatag catttctcaa 420

actacagcat ttctctcaga tatctctatt gcaccaggtc aaatagatag atgcatcaga 480

gaagcatatg ttcatcagag accagtttat gttggtttac cggcaaatat ggttgatctc 540

aaggttcctt ctagtctctt agaaactcca attgatttga aattgaaaca aaatgatcct 600

gaagctcaag aagttgttga aacagtcctg aagttggtgt cccaagctac aaaccccatt 660

atcttggtag acgcttgtgc cctcagacac aattgcaaag aggaagtcaa acaattggtt 720

gatgccacta attttcaagt ctttacaact ccaatgggta aatctggtat ctccgaatct 780

catccaagat tgggcggtgt ctatgtcggg acaatgtcga gtcctcaagt caaaaaagcc 840

gttgaaaatg ccgatcttat actatctgtt ggttcgttgt tatcggactt caatacaggt 900

tcattttcat actcctacaa gacgaagaat gttgttgaat tccactctga ctatatgaaa 960

atcagacagg ccaccttccc aggagttcaa atgaaagaag ccttgcaaca gttgataaaa 1020

agggtctctt cttacatcaa tccaagctac attcctactc gagttcctaa aaggaaacag 1080

ccattgaaag ctccatcaga agctcctttg acccaagaat atttgtggtc taaagtatcc 1140

ggctggttta gagagggtga tattatcgta accgaaactg gtacatctgc tttcggaatt 1200

attcaatccc attttcccag caacactatc ggtatatccc aagtcttgtg gggctcaatt 1260

ggtttcacag taggtgcaac agttggtgct gccatggcag cccaggaaat cgaccctagc 1320

aggagagtaa ttttgttcgt cggtgatggt tcattgcagt tgacggttca ggaaatctct 1380

acgttgtgta aatgggattg taacaatact tatctttacg tgttgaacaa tgatggttac 1440

actatagaaa ggttgatcca cggcaaaagt gccagctaca acgatataca gccttggaac 1500

catttatcct tgcttcgctt attcaatgct aagaaatacc aaaatgtcag agtatcgact 1560

gctggagaat tggactcttt gttctctgat aagaaatttg cttctccaga taggataaga 1620

atgattgagg tgatgttatc gagattggat gcaccagcaa atcttgttgc tcaagcaaag 1680

ttgtctgaac gggtaaacct tgaaaat 1707

<210> 659
<211> 569
<212> PRT
<213> Artificial Sequence

<220>
<223> PDC2 pyruvate decarboxylase from Pichia stipitis

<400> 659

```
Met Val Ser Thr Tyr Pro Glu Ser Glu Val Thr Leu Gly Arg Tyr Leu
1               5                  10                  15

Phe Glu Arg Leu His Gln Leu Lys Val Asp Thr Ile Phe Gly Leu Pro
            20                  25                  30

Gly Asp Phe Asn Leu Ser Leu Leu Asp Lys Val Tyr Glu Val Pro Asp
        35                  40                  45

Met Arg Trp Ala Gly Asn Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala
    50                  55                  60

Asp Gly Tyr Ser Arg Ile Lys Gly Leu Ser Cys Leu Val Thr Thr Phe
65                  70                  75                  80

Gly Val Gly Glu Leu Ser Ala Leu Asn Gly Val Gly Gly Ala Tyr Ala
                85                  90                  95

Glu His Val Gly Leu Leu His Val Val Gly Val Pro Ser Ile Ser Ser
            100                 105                 110

Gln Ala Lys Gln Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe
        115                 120                 125

Thr Val Phe His Arg Met Ser Asn Ser Ile Ser Gln Thr Thr Ala Phe
    130                 135                 140

Leu Ser Asp Ile Ser Ile Ala Pro Gly Gln Ile Asp Arg Cys Ile Arg
145                 150                 155                 160

Glu Ala Tyr Val His Gln Arg Pro Val Tyr Val Gly Leu Pro Ala Asn
                165                 170                 175

Met Val Asp Leu Lys Val Pro Ser Ser Leu Leu Glu Thr Pro Ile Asp
            180                 185                 190

Leu Lys Leu Lys Gln Asn Asp Pro Glu Ala Gln Glu Val Val Glu Thr
        195                 200                 205

Val Leu Lys Leu Val Ser Gln Ala Thr Asn Pro Ile Ile Leu Val Asp
    210                 215                 220

Ala Cys Ala Leu Arg His Asn Cys Lys Glu Glu Val Lys Gln Leu Val
225                 230                 235                 240

Asp Ala Thr Asn Phe Gln Val Phe Thr Thr Pro Met Gly Lys Ser Gly
                245                 250                 255

Ile Ser Glu Ser His Pro Arg Leu Gly Gly Val Tyr Val Gly Thr Met
            260                 265                 270

Ser Ser Pro Gln Val Lys Lys Ala Val Glu Asn Ala Asp Leu Ile Leu
        275                 280                 285
```

```
Ser Val Gly Ser Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr
    290                 295                 300

Ser Tyr Lys Thr Lys Asn Val Val Glu Phe His Ser Asp Tyr Met Lys
305                 310                 315                 320

Ile Arg Gln Ala Thr Phe Pro Gly Val Gln Met Lys Glu Ala Leu Gln
                325                 330                 335

Gln Leu Ile Lys Arg Val Ser Ser Tyr Ile Asn Pro Ser Tyr Ile Pro
            340                 345                 350

Thr Arg Val Pro Lys Arg Lys Gln Pro Leu Lys Ala Pro Ser Glu Ala
        355                 360                 365

Pro Leu Thr Gln Glu Tyr Leu Trp Ser Lys Val Ser Gly Trp Phe Arg
    370                 375                 380

Glu Gly Asp Ile Ile Val Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile
385                 390                 395                 400

Ile Gln Ser His Phe Pro Ser Asn Thr Ile Gly Ile Ser Gln Val Leu
                405                 410                 415

Trp Gly Ser Ile Gly Phe Thr Val Gly Ala Thr Val Gly Ala Ala Met
            420                 425                 430

Ala Ala Gln Glu Ile Asp Pro Ser Arg Arg Val Ile Leu Phe Val Gly
        435                 440                 445

Asp Gly Ser Leu Gln Leu Thr Val Gln Glu Ile Ser Thr Leu Cys Lys
    450                 455                 460

Trp Asp Cys Asn Asn Thr Tyr Leu Tyr Val Leu Asn Asn Asp Gly Tyr
465                 470                 475                 480

Thr Ile Glu Arg Leu Ile His Gly Lys Ser Ala Ser Tyr Asn Asp Ile
                485                 490                 495

Gln Pro Trp Asn His Leu Ser Leu Leu Arg Leu Phe Asn Ala Lys Lys
            500                 505                 510

Tyr Gln Asn Val Arg Val Ser Thr Ala Gly Glu Leu Asp Ser Leu Phe
        515                 520                 525

Ser Asp Lys Lys Phe Ala Ser Pro Asp Arg Ile Arg Met Ile Glu Val
    530                 535                 540

Met Leu Ser Arg Leu Asp Ala Pro Ala Asn Leu Val Ala Gln Ala Lys
545                 550                 555                 560

Leu Ser Glu Arg Val Asn Leu Glu Asn
                565
```

20 Jan 2017 2017200439

<210> 660
<211> 1692
<212> DNA
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Kluyveromyces lactis

<400> 660
atgtctgaaa ttacattagg tcgttacttg ttcgaaagat taaagcaagt cgaagttcaa 60

accatctttg gtctaccagg tgatttcaac ttgtccctat tggacaatat ctacgaagtc 120

ccaggtatga gatgggctgg taatgccaac gaattgaacg ctgcttacgc tgctgatggt 180

tacgccagat taaagggtat gtcctgtatc atcaccacct tcggtgtcgg tgaattgtct 240

gctttgaacg gtattgccgg ttcttacgct gaacacgttg gtgtcttgca cgttgtcggt 300

gttccatccg tctcttctca agctaagcaa ttgttgttgc accacacctt gggtaacggt 360

gacttcactg ttttccacag aatgtgctcc aacatttctg aaaccactgc tatgatcacc 420

gatatcaaca ctgccccagc tgaaatcgac agatgtatca gaaccactta cgtttcccaa 480

agaccagtct acttgggttt gccagctaac ttggtcgact tgactgtccc agcttctttg 540

ttggacactc caattgattt gagcttgaag ccaaatgacc cagaagccga agaagaagtc 600

atcgaaaacg tcttgcaact gatcaaggaa gctaagaacc cagttatctt ggctgatgct 660

tgttgttcca gacacgatgc caaggctgag accaagaagt tgatcgactt gactcaattc 720

ccagccttcg ttaccccaat gggtaagggt tccattgacg aaaagcaccc aagattcggt 780

ggtgtctacg tcggtaccct atcttctcca gctgtcaagg aagccgttga atctgctcac 840

ttggttctat cggtcggtgc tctattgtcc gatttcaaca ctggttcttt ctcttactct 900

tacaagacca agaacattgt cgaattccac tctgactaca ccaagatcag aaggcctacc 960

ttcccaggtg tccaaatgaa gttcgcttta caaaaattgt tgactaaggt tgccgatgct 1020

gctaagggtt acaagccagt tccagttcca tctgaaccag aacacaacga agatgtcgct 1080

gactccactc cattgaagca agaatgggtc tggactcaag tcggtgaatt cttgagagaa 1140

ggtgatgttg ttatcactga aaccggtacc tctgccttcg gtatcaacca aactcatttc 1200

ccaaacaaca catacggtat ctctcaagtt ttatggggtt ccattggttt caccactggt 1260

gctaccttgg gtgctgcctt cgctgccgaa gaaattgatc caaagaagag agttatctta 1320

ttcattggtg acggttcttt gcaattgact gttcaagaaa tctccaccat gatcagatgg 1380

ggcttgaagc catacttgtt cgtattgaac aacgacggtt acaccattga aagattgatt 1440

cacggtgaaa ccgctcaata caactgtatc caaaactggc aacacttgga attattgcca 1500

actttcggtg ccaaggacta cgaagctgtc agagtttcca ccactggtga atggaacaag 1560

ttgaccactg acgaaaagtt ccaagacaac accagaatca gattgatcga agttatgttg 1620

ccaactatgg atgctccatc taacttggtt aagcaagctc aattgactgc tgcatccaac 1680

gctaagaact aa 1692

<210> 661
<211> 563

2017200439 20 Jan 2017

<212> PRT
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Kluyveromyces lactis

<400> 661

```
Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Lys Gln
1               5                  10                  15

Val Glu Val Gln Thr Ile Phe Gly Leu Pro Gly Asp Phe Asn Leu Ser
            20                  25                  30

Leu Leu Asp Asn Ile Tyr Glu Val Pro Gly Met Arg Trp Ala Gly Asn
        35                  40                  45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Leu
    50                  55                  60

Lys Gly Met Ser Cys Ile Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65                  70                  75                  80

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
                85                  90                  95

His Val Val Gly Val Pro Ser Val Ser Ser Gln Ala Lys Gln Leu Leu
            100                 105                 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
        115                 120                 125

Ser Ser Asn Ile Ser Glu Thr Thr Ala Met Ile Thr Asp Ile Asn Thr
    130                 135                 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Thr Thr Tyr Val Ser Gln
145                 150                 155                 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Leu Thr Val
                165                 170                 175

Pro Ala Ser Leu Leu Asp Thr Pro Ile Asp Leu Ser Leu Lys Pro Asn
            180                 185                 190

Asp Pro Glu Ala Glu Glu Glu Val Ile Glu Asn Val Leu Gln Leu Ile
        195                 200                 205

Lys Glu Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Cys Ser Arg
    210                 215                 220

His Asp Ala Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe
225                 230                 235                 240

Pro Ala Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Lys His
                245                 250                 255
```

```
Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Ser Pro Ala Val
            260                 265                 270

Lys Glu Ala Val Glu Ser Ala Asp Leu Val Leu Ser Val Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Ile Val Glu Phe His Ser Asp Tyr Thr Lys Ile Arg Ser Ala Thr
305                 310                 315                 320

Phe Pro Gly Val Gln Met Lys Phe Ala Leu Gln Lys Leu Leu Thr Lys
                325                 330                 335

Val Ala Asp Ala Ala Lys Gly Tyr Lys Pro Val Pro Val Pro Ser Glu
            340                 345                 350

Pro Glu His Asn Glu Ala Val Ala Asp Ser Thr Pro Leu Lys Gln Glu
        355                 360                 365

Trp Val Trp Thr Gln Val Gly Glu Phe Leu Arg Glu Gly Asp Val Val
    370                 375                 380

Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr His Phe
385                 390                 395                 400

Pro Asn Asn Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Phe Thr Thr Gly Ala Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
465                 470                 475                 480

His Gly Glu Thr Ala Gln Tyr Asn Cys Ile Gln Asn Trp Gln His Leu
                485                 490                 495

Glu Leu Leu Pro Thr Phe Gly Ala Lys Asp Tyr Glu Ala Val Arg Val
            500                 505                 510

Ser Thr Thr Gly Glu Trp Asn Lys Leu Thr Thr Asp Glu Lys Phe Gln
        515                 520                 525

Asp Asn Thr Arg Ile Arg Leu Ile Glu Val Met Leu Pro Thr Met Asp
    530                 535                 540
```

20 Jan 2017 2017200439

Ala Pro Ser Asn Leu Val Lys Gln Ala Gln Leu Thr Ala Ala Thr Asn
545 550 555 560

Ala Lys Asn

<210> 662
<211> 1716
<212> DNA
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Yarrowia lipolytica

<400> 662
atgagcgact ccgaacccca aatggtcgac ctgggcgact atctctttgc ccgattcaag 60

cagctaggcg tggactccgt ctttggagtg cccggcgact tcaacctcac cctgttggac 120

cacgtgtaca atgtcgacat gcggtgggtt gggaacacaa acgagctgaa tgccggctac 180

tcggccgacg gctactcccg ggtcaagcgg ctggcatgtc ttgtcaccac ctttggcgtg 240

ggagagctgt ctgccgtggc tgctgtggca ggctcgtacg ccgagcatgt gggcgtggtg 300

catgttgtgg gcgttcccag cacctctgct gagaacaagc atctgctgct gcaccacaca 360

ctcggtaacg gcgacttccg ggtctttgcc cagatgtcca aactcatctc cgagtacacc 420

caccatattg aggaccccag cgaggctgcc gacgtaatcg acaccgccat ccgaatcgcc 480

tacacccacc agcggcccgt ttacattgct gtgccctcca acttctccga ggtcgatatt 540

gccgaccagg ctagactgga tacccccctg gacctttcgc tgcagcccaa cgaccccgag 600

agccagtacg aggtgattga ggagatttgc tcgcgtatca aggccgccaa gaagcccgtg 660

attctcgtcg acgcctgcgc ttcgcgatac agatgtgtgg acgagaccaa ggagctggcc 720

aagatcacca actttgccta ctttgtcact cccatgggta agggttctgt ggacgaggat 780

actgaccggt acggaggaac atacgtcgga tcgctgactg ctcctgctac tgccgaggtg 840

gttgagacag ctgatctcat catctccgta ggagctcttc tgtcggactt caacaccggt 900

tccttctcgt actcctactc caccaaaaac gtggtggaat tgcattcgga ccacgtcaaa 960

atcaagtccg ccacctacaa caacgtcggc atgaaaatgc tgttcccgcc cctgctcgaa 1020

gccgtcaaga aactggttgc cgagacccct gactttgcat ccaaggctct ggctgttccc 1080

gacaccactc ccaagatccc cgaggtaccc gatgatcaca ttacgaccca ggcatggctg 1140

tggcagcgtc tcagttactt tctgaggccc accgacatcg tggtcaccga gaccggaacc 1200

tcgtcctttg gaatcatcca gaccaagttc ccccacaacg tccgaggtat ctcgcaggtg 1260

ctgtggggct ctattggata ctcggtggga gcagcctgtg gagcctccat tgctgcacag 1320

gagattgacc cccagcagcg agtgattctg tttgtgggcg acggctctct tcagctgacg 1380

gtgaccgaga tctcgtgcat gatccgcaac aacgtcaagc cgtacatttt tgtgctcaac 1440

aacgacggct acaccatcga gaggctcatt cacggcgaaa acgcctcgta caacgatgtg 1500

cacatgtgga agtactccaa gattctcgac acgttcaacg ccaaggccca cgagtcgatt 1560

gtggtcaaca ccaagggcga gatggacgct ctgttcgaca acgaagagtt tgccaagccc 1620

gacaagatcc ggctcattga ggtcatgtgc gacaagatgg acgcgcctgc ctcgttgatc 1680

2017200439 20 Jan 2017

aagcaggctg agctctctgc caagaccaac gtttag 1716

<210> 663
<211> 571
<212> PRT
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Yarrowia lipolytica

<400> 663

Met Ser Asp Ser Glu Pro Gln Met Val Asp Leu Gly Asp Tyr Leu Phe
1 5 10 15

Ala Arg Phe Lys Gln Leu Gly Val Asp Ser Val Phe Gly Val Pro Gly
20 25 30

Asp Phe Asn Leu Thr Leu Leu Asp His Val Tyr Asn Val Asp Met Arg
35 40 45

Trp Val Gly Asn Thr Asn Glu Leu Asn Ala Gly Tyr Ser Ala Asp Gly
50 55 60

Tyr Ser Arg Val Lys Arg Leu Ala Cys Leu Val Thr Thr Phe Gly Val
65 70 75 80

Gly Glu Leu Ser Ala Val Ala Ala Val Ala Gly Ser Tyr Ala Glu His
85 90 95

Val Gly Val Val His Val Val Gly Val Pro Ser Thr Ser Ala Glu Asn
100 105 110

Lys His Leu Leu Leu His His Thr Leu Gly Asn Gly Asp Phe Arg Val
115 120 125

Phe Ala Gln Met Ser Lys Leu Ile Ser Glu Tyr Thr His His Ile Glu
130 135 140

Asp Pro Ser Glu Ala Ala Asp Val Ile Asp Thr Ala Ile Arg Ile Ala
145 150 155 160

Tyr Thr His Gln Arg Pro Val Tyr Ile Ala Val Pro Ser Asn Phe Ser
165 170 175

Glu Val Asp Ile Ala Asp Gln Ala Arg Leu Asp Thr Pro Leu Asp Leu
180 185 190

Ser Leu Gln Pro Asn Asp Pro Glu Ser Gln Tyr Glu Val Ile Glu Glu
195 200 205

Ile Cys Ser Arg Ile Lys Ala Ala Lys Lys Pro Val Ile Leu Val Asp
210 215 220

Ala Cys Ala Ser Arg Tyr Arg Cys Val Asp Glu Thr Lys Glu Leu Ala
225 230 235 240

```
Lys Ile Thr Asn Phe Ala Tyr Phe Val Thr Pro Met Gly Lys Gly Ser
                245                 250                 255

Val Asp Glu Asp Thr Asp Arg Tyr Gly Gly Thr Tyr Val Gly Ser Leu
            260                 265                 270

Thr Ala Pro Ala Thr Ala Glu Val Val Glu Thr Ala Asp Leu Ile Ile
        275                 280                 285

Ser Val Gly Ala Leu Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr
    290                 295                 300

Ser Tyr Ser Thr Lys Asn Val Val Glu Leu His Ser Asp His Val Lys
305                 310                 315                 320

Ile Lys Ser Ala Thr Tyr Asn Asn Val Gly Met Lys Met Leu Phe Pro
                325                 330                 335

Pro Leu Leu Glu Ala Val Lys Lys Leu Val Ala Glu Thr Pro Asp Phe
            340                 345                 350

Ala Ser Lys Ala Leu Ala Val Pro Asp Thr Thr Pro Lys Ile Pro Glu
        355                 360                 365

Val Pro Asp Asp His Ile Thr Thr Gln Ala Trp Leu Trp Gln Arg Leu
    370                 375                 380

Ser Tyr Phe Leu Arg Pro Thr Asp Ile Val Val Thr Glu Thr Gly Thr
385                 390                 395                 400

Ser Ser Phe Gly Ile Ile Gln Thr Lys Phe Pro His Asn Val Arg Gly
                405                 410                 415

Ile Ser Gln Val Leu Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Ala
            420                 425                 430

Cys Gly Ala Ser Ile Ala Ala Gln Glu Ile Asp Pro Gln Gln Arg Val
        435                 440                 445

Ile Leu Phe Val Gly Asp Gly Ser Leu Gln Leu Thr Val Thr Glu Ile
    450                 455                 460

Ser Cys Met Ile Arg Asn Asn Val Lys Pro Tyr Ile Phe Val Leu Asn
465                 470                 475                 480

Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile His Gly Glu Asn Ala Ser
                485                 490                 495

Tyr Asn Asp Val His Met Trp Lys Tyr Ser Lys Ile Leu Asp Thr Phe
            500                 505                 510

Asn Ala Lys Ala His Glu Ser Ile Val Val Asn Thr Lys Gly Glu Met
        515                 520                 525
```

Asp Ala Leu Phe Asp Asn Glu Glu Phe Ala Lys Pro Asp Lys Ile Arg
530 535 540

Leu Ile Glu Val Met Cys Asp Lys Met Asp Ala Pro Ala Ser Leu Ile
545 550 555 560

Lys Gln Ala Glu Leu Ser Ala Lys Thr Asn Val
565 570

<210> 664
<211> 1716
<212> DNA
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Schizosaccharomyces pombe

<400> 664
atgagtgggg atattttagt cggtgaatat ctattcaaaa ggcttgaaca attaggggtc 60

aagtccattc ttggtgttcc aggagatttc aatttagctc tacttgactt aattgagaaa 120

gttggagatg agaaatttcg ttgggttggc aataccaatg agttgaatgg tgcttatgcc 180

gctgatggtt atgctcgtgt taatggtctt tcagccattg ttacaacgtt cggcgtggga 240

gagctttccg ctattaatgg agtggcaggt tcttatgcgg agcatgtccc agtagttcat 300

attgttggaa tgccttccac aaaggtgcaa gatactggag ctttgcttca tcatacttta 360

ggagatggag actttcgcac tttcatggat atgtttaaga aagtttctgc ctacagtata 420

atgatcgata acggaaacga tgcagctgaa aagatcgatg aagccttgtc gatttgttat 480

aaaaaggcta ggcctgttta cattggtatt ccttctgatg ctggctactt caaagcatct 540

tcatcaaatc ttgggaaaag actaaagctc gaggaggata ctaacgatcc agcagttgag 600

caagaagtca tcaatcatat ctcggaaatg gttgtcaatg caaagaaacc agtgatttta 660

attgacgctt gtgctgtaag acatcgtgtc gttccagaag tacatgagct gattaaattg 720

acccatttcc ctacatatgt aactcccatg ggtaaatctg caattgacga aacttcgcaa 780

ttttttgacg gcgtttatgt tggttcaatt tcagatcctg aagttaaaga cagaattgaa 840

tccactgatc tgttgctatc catcggtgct ctcaaatcag actttaacac gggttccttc 900

tcttaccacc tcagccaaaa gaatgccgtt gagtttcatt cagaccacat gcgcattcga 960

tatgctcttt atccaaatgt agccatgaag tatattcttc gcaaactgtt gaaagtactt 1020

gatgcttcta tgtgtcattc caaggctgct cctaccattg gctacaacat caagcctaag 1080

catgcggaag gatattcttc caacgagatt actcattgct ggttttggcc taaatttagt 1140

gaatttttga agccccgaga tgttttgatc accgagactg gaactgcaaa ctttggtgtc 1200

cttgattgca ggtttccaaa ggatgtaaca gccatttccc aggtattatg gggatctatt 1260

ggatactccg ttggtgcaat gtttggtgct gttttggccg tccacgattc taaagagccc 1320

gatcgtcgta ccattcttgt agtaggtgat ggatccttac aactgacgat tacagagatt 1380

tcaacctgca ttcgccataa cctcaaacca attattttca taattaacaa cgacggttac 1440

accattgagc gtttaattca tggtttgcat gctagctata acgaaattaa cactaaatgg 1500

ggctaccaac agattcccaa gtttttcgga gctgctgaaa accacttccg cacttactgt 1560

gttaaaactc ctactgacgt tgaaaagttg tttagcgaca aggagtttgc aaatgcagat 1620

gtcattcaag tagttgagct tgtaatgcct atgttggatg cacctcgtgt cctagttgag 1680

caagccaagt tgacgtctaa gatcaataag caatga 1716

<210> 665
<211> 571
<212> PRT
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Schizosaccharomyces pombe

<400> 665

Met Ser Gly Asp Ile Leu Val Gly Glu Tyr Leu Phe Lys Arg Leu Glu
1 5 10 15

Gln Leu Gly Val Lys Ser Ile Leu Gly Val Pro Gly Asp Phe Asn Leu
20 25 30

Ala Leu Leu Asp Leu Ile Glu Lys Val Gly Asp Glu Lys Phe Arg Trp
35 40 45

Val Gly Asn Thr Asn Glu Leu Asn Gly Ala Tyr Ala Ala Asp Gly Tyr
50 55 60

Ala Arg Val Asn Gly Leu Ser Ala Ile Val Thr Thr Phe Gly Val Gly
65 70 75 80

Glu Leu Ser Ala Ile Asn Gly Val Ala Gly Ser Tyr Ala Glu His Val
85 90 95

Pro Val Val His Ile Val Gly Met Pro Ser Thr Lys Val Gln Asp Thr
100 105 110

Gly Ala Leu Leu His His Thr Leu Gly Asp Gly Asp Phe Arg Thr Phe
115 120 125

Met Asp Met Phe Lys Lys Val Ser Ala Tyr Ser Ile Met Ile Asp Asn
130 135 140

Gly Asn Asp Ala Ala Glu Lys Ile Asp Glu Ala Leu Ser Ile Cys Tyr
145 150 155 160

Lys Lys Ala Arg Pro Val Tyr Ile Gly Ile Pro Ser Asp Ala Gly Tyr
165 170 175

Phe Lys Ala Ser Ser Ser Asn Leu Gly Lys Arg Leu Lys Leu Glu Glu
180 185 190

Asp Thr Asn Asp Pro Ala Val Glu Gln Glu Val Ile Asn His Ile Ser
195 200 205

```
Glu Met Val Val Asn Ala Lys Lys Pro Val Ile Leu Ile Asp Ala Cys
    210                 215                 220

Ala Val Arg His Arg Val Val Pro Glu Val His Glu Leu Ile Lys Leu
225                 230                 235                 240

Thr His Phe Pro Thr Tyr Val Thr Pro Met Gly Lys Ser Ala Ile Asp
                245                 250                 255

Glu Thr Ser Gln Phe Phe Asp Gly Val Tyr Val Gly Ser Ile Ser Asp
            260                 265                 270

Pro Glu Val Lys Asp Arg Ile Glu Ser Thr Asp Leu Leu Leu Ser Ile
        275                 280                 285

Gly Ala Leu Lys Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr His Leu
    290                 295                 300

Ser Gln Lys Asn Ala Val Glu Phe His Ser Asp His Met Arg Ile Arg
305                 310                 315                 320

Tyr Ala Leu Tyr Pro Asn Val Ala Met Lys Tyr Ile Leu Arg Lys Leu
                325                 330                 335

Leu Lys Val Leu Asp Ala Ser Met Cys His Ser Lys Ala Ala Pro Thr
            340                 345                 350

Ile Gly Tyr Asn Ile Lys Pro Lys His Ala Glu Gly Tyr Ser Ser Asn
        355                 360                 365

Glu Ile Thr His Cys Trp Phe Trp Pro Lys Phe Ser Glu Phe Leu Lys
    370                 375                 380

Pro Arg Asp Val Leu Ile Thr Glu Thr Gly Thr Ala Asn Phe Gly Val
385                 390                 395                 400

Leu Asp Cys Arg Phe Pro Lys Asp Val Thr Ala Ile Ser Gln Val Leu
                405                 410                 415

Trp Gly Ser Ile Gly Tyr Ser Val Gly Ala Met Phe Gly Ala Val Leu
            420                 425                 430

Ala Val His Asp Ser Lys Glu Pro Asp Arg Arg Thr Ile Leu Val Val
        435                 440                 445

Gly Asp Gly Ser Leu Gln Leu Thr Ile Thr Glu Ile Ser Thr Cys Ile
    450                 455                 460

Arg His Asn Leu Lys Pro Ile Ile Phe Ile Ile Asn Asn Asp Gly Tyr
465                 470                 475                 480

Thr Ile Glu Arg Leu Ile His Gly Leu His Ala Ser Tyr Asn Glu Ile
                485                 490                 495
```

Asn Thr Lys Trp Gly Tyr Gln Gln Ile Pro Lys Phe Phe Gly Ala Ala
500 505 510

Glu Asn His Phe Arg Thr Tyr Cys Val Lys Thr Pro Thr Asp Val Glu
515 520 525

Lys Leu Phe Ser Asp Lys Glu Phe Ala Asn Ala Asp Val Ile Gln Val
530 535 540

Val Glu Leu Val Met Pro Met Leu Asp Ala Pro Arg Val Leu Val Glu
545 550 555 560

Gln Ala Lys Leu Thr Ser Lys Ile Asn Lys Gln
565 570

<210> 666
<211> 1689
<212> DNA
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Zygosaccharomyces rouxii

<400> 666
atgtctgaaa ttactctagg tcgttacttg ttcgaaagat taaagcaagt tgacactaac 60

accatcttcg gtgttccagg tgacttcaac ttgtccttgt tggacaaggt ctacgaagtg 120

caaggtctaa gatgggctgg taacgctaac gaattgaacg ctgcctacgc tgctgacggt 180

tacgccagag ttaagggttt ggctgctttg atcaccacct tcggtgtcgg tgaattgtct 240

gctttgaacg gtattgcagg ttcttacgct gaacacgttg gtgttttgca cattgttggt 300

gttccatctg tctcttctca agctaagcaa ttgttgttgc accacacctt gggtaacggt 360

gacttcactg ttttccacag aatgtccgcc aacatctctg aaaccaccgc tatgttgacc 420

gacatcactg ctgctccagc tgaaattgac cgttgcatca gagttgctta cgtcaaccaa 480

agaccagtct acttgggtct accagctaac ttggttgacc aaaaggtccc agcttctttg 540

ttgaacactc caattgatct atctctaaag gagaacgacc cagaagctga aaccgaagtt 600

gttgacaccg ttttggaatt gatcaaggaa gctaagaacc cagttatctt ggctgatgct 660

tgctgctcca gacacgacgt caaggctgaa accaagaagt tgatcgactt gactcaattc 720

ccatctttcg ttactcctat gggtaagggt tccatcgacg aacaaaaccc aagattcggt 780

ggtgtctacg tcggtactct atccagccca gaagttaagg aagctgttga atctgctgac 840

ttggttctat ctgtcggtgc tctattgtcc gatttcaaca ctggttcttt ctcttactct 900

tacaagacca agaacgttgt tgaattccac tctgaccaca tcaagatcag aaacgctacc 960

ttcccaggtg ttcaaatgaa attcgttttg aagaaactat tgcaagctgt cccagaagct 1020

gtcaagaact acaagccagg tccagtccca gctccgccat ctccaaacgc tgaagttgct 1080

gactctacca ccttgaagca agaatggtta tggagacaag tcggtagctt cttgagagaa 1140

ggtgatgttg ttattaccga aactggtacc tctgctttcg gtatcaacca aactcacttc 1200

cctaaccaaa cttacggtat ctctcaagtc ttgtggggtt ctattggtta caccactggt 1260

tccactttgg gtgctgcctt cgctgctgaa gaaattgacc ctaagaagag agttatcttg 1320

ttcattggtg acggttctct acaattgacc gttcaagaaa tctccaccat gatcagatgg 1380

ggtctaaagc catacttgtt cgttttgaac aacgatggtt acaccattga aagattgatt 1440

cacggtgaaa ccgctgaata caactgtatc caaccatgga agcacttgga attgttgaac 1500

accttcggtg ccaaggacta cgaaaaccac agagtctcca ctgtcggtga atggaacaag 1560

ttgactcaag atccaaaatt caacgaaaac tctagaatta gaatgatcga agttatgctt 1620

gaagtcatgg acgctccatc ttctttggtc gctcaagctc aattgaccgc tgctactaac 1680

gctaagcaa 1689

<210> 667
<211> 563
<212> PRT
<213> Artificial Sequence

<220>
<223> pyruvate decarboxylase from Zygosaccharomyces rouxii

<400> 667

Met Ser Glu Ile Thr Leu Gly Arg Tyr Leu Phe Glu Arg Leu Lys Gln
1 5 10 15

Val Asp Thr Asn Thr Ile Phe Gly Val Pro Gly Asp Phe Asn Leu Ser
20 25 30

Leu Leu Asp Lys Val Tyr Glu Val Gln Gly Leu Arg Trp Ala Gly Asn
35 40 45

Ala Asn Glu Leu Asn Ala Ala Tyr Ala Ala Asp Gly Tyr Ala Arg Val
50 55 60

Lys Gly Leu Ala Ala Leu Ile Thr Thr Phe Gly Val Gly Glu Leu Ser
65 70 75 80

Ala Leu Asn Gly Ile Ala Gly Ser Tyr Ala Glu His Val Gly Val Leu
85 90 95

His Ile Val Gly Val Pro Ser Val Ser Ser Gln Ala Lys Gln Leu Leu
100 105 110

Leu His His Thr Leu Gly Asn Gly Asp Phe Thr Val Phe His Arg Met
115 120 125

Ser Ala Asn Ile Ser Glu Thr Thr Ala Met Leu Thr Asp Ile Thr Ala
130 135 140

Ala Pro Ala Glu Ile Asp Arg Cys Ile Arg Val Ala Tyr Val Asn Gln
145 150 155 160

Arg Pro Val Tyr Leu Gly Leu Pro Ala Asn Leu Val Asp Gln Lys Val
165 170 175

Pro Ala Ser Leu Leu Asn Thr Pro Ile Asp Leu Ser Leu Lys Glu Asn
180 185 190

```
Asp Pro Glu Ala Glu Thr Glu Val Val Asp Thr Val Leu Glu Leu Ile
        195                 200                 205

Lys Glu Ala Lys Asn Pro Val Ile Leu Ala Asp Ala Cys Cys Ser Arg
    210                 215                 220

His Asp Val Lys Ala Glu Thr Lys Lys Leu Ile Asp Leu Thr Gln Phe
225                 230                 235                 240

Pro Ser Phe Val Thr Pro Met Gly Lys Gly Ser Ile Asp Glu Gln Asn
                245                 250                 255

Pro Arg Phe Gly Gly Val Tyr Val Gly Thr Leu Ser Ser Pro Glu Val
            260                 265                 270

Lys Glu Ala Val Glu Ser Ala Asp Leu Val Leu Ser Val Gly Ala Leu
        275                 280                 285

Leu Ser Asp Phe Asn Thr Gly Ser Phe Ser Tyr Ser Tyr Lys Thr Lys
    290                 295                 300

Asn Val Val Glu Phe His Ser Asp His Ile Lys Ile Arg Asn Ala Thr
305                 310                 315                 320

Phe Pro Gly Val Gln Met Lys Phe Val Leu Lys Lys Leu Leu Gln Ala
                325                 330                 335

Val Pro Glu Ala Val Lys Asn Tyr Lys Pro Gly Pro Val Pro Ala Pro
            340                 345                 350

Pro Ser Pro Asn Ala Glu Val Ala Asp Ser Thr Thr Leu Lys Gln Glu
        355                 360                 365

Trp Leu Trp Arg Gln Val Gly Ser Phe Leu Arg Glu Gly Asp Val Val
    370                 375                 380

Ile Thr Glu Thr Gly Thr Ser Ala Phe Gly Ile Asn Gln Thr His Phe
385                 390                 395                 400

Pro Asn Gln Thr Tyr Gly Ile Ser Gln Val Leu Trp Gly Ser Ile Gly
                405                 410                 415

Tyr Thr Thr Gly Ser Thr Leu Gly Ala Ala Phe Ala Ala Glu Glu Ile
            420                 425                 430

Asp Pro Lys Lys Arg Val Ile Leu Phe Ile Gly Asp Gly Ser Leu Gln
        435                 440                 445

Leu Thr Val Gln Glu Ile Ser Thr Met Ile Arg Trp Gly Leu Lys Pro
    450                 455                 460

Tyr Leu Phe Val Leu Asn Asn Asp Gly Tyr Thr Ile Glu Arg Leu Ile
465                 470                 475                 480
```

```
His Gly Glu Thr Ala Glu Tyr Asn Cys Ile Gln Pro Trp Lys His Leu
                485                 490                 495

Glu Leu Leu Asn Thr Phe Gly Ala Lys Asp Tyr Glu Asn His Arg Val
            500                 505                 510

Ser Thr Val Gly Glu Trp Asn Lys Leu Thr Gln Asp Pro Lys Phe Asn
        515                 520                 525

Glu Asn Ser Arg Ile Arg Met Ile Glu Val Met Leu Glu Val Met Asp
    530                 535                 540

Ala Pro Ser Ser Leu Val Ala Gln Ala Gln Leu Thr Ala Ala Thr Asn
545                 550                 555                 560

Ala Lys Gln
```

<210> 668
<211> 1461
<212> DNA
<213> Artificial Sequence

<220>
<223> HXK2 from S. cerevisiae

<400> 668

```
atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcca       60

aaggaattga tgcaacaaat tgagaatttt gaaaaaattt tcactgttcc aactgaaact      120

ttacaagccg ttaccaagca cttcatttcc gaattggaaa agggtttgtc caagaagggt      180

ggtaacattc caatgattcc aggttgggtt atggatttcc caactggtaa ggaatccggt      240

gatttcttgg ccattgattt gggtggtacc aacttgagag ttgtcttagt caagttgggc      300

ggtgaccgta cctttgacac cactcaatct aagtacagat taccagatgc tatgagaact      360

actcaaaatc cagacgaatt gtgggaattt attgccgact ctttgaaagc ttttattgat      420

gagcaattcc cacaaggtat ctctgagcca attccattgg gtttcacctt ttctttccca      480

gcttctcaaa acaaaatcaa tgaaggtatc ttgcaaagat ggactaaagg ttttgatatt      540

ccaaacattg aaaaccacga tgttgttcca atgttgcaaa agcaaatcac taagaggaat      600

atcccaattg aagttgttgc tttgataaac gacactaccg gtactttggt tgcttcttac      660

tacactgacc cagaaactaa gatgggtgtt atcttcggta ctggtgtcaa tggtgcttac      720

tacgatgttt gttccgatat cgaaaagcta caaggaaaac tatctgatga cattccacca      780

tctgctccaa tggccatcaa ctgtgaatac ggttccttcg ataatgaaca tgtcgttttg      840

ccaagaacta aatacgatat caccattgat gaagaatctc caagaccagg ccaacaaacc      900

tttgaaaaaa tgtcttctgg ttactactta ggtgaaattt tgcgtttggc cttgatggac      960

atgtacaaac aaggtttcat cttcaagaac caagacttgt ctaagttcga caagcctttc     1020

gtcatggaca cttcttaccc agccagaatc gaggaagatc cattcgagaa cctagaagat     1080

accgatgact tgttccaaaa tgagttcggt atcaacacta ctgttcaaga acgtaaattg     1140
```

atcagacgtt tatctgaatt gattggtgct agagctgcta gattgtccgt ttgtggtatt 1200

gctgctatct gtcaaaagag aggttacaag accggtcaca tcgctgcaga cggttccgtt 1260

tacaacagat acccaggttt caaagaaaag gctgccaatg ctttgaagga catttacggc 1320

tggactcaaa cctcactaga cgactaccca atcaagattg ttcctgctga agatggttcc 1380

ggtgctggtg ccgctgttat tgctgctttg gcccaaaaaa gaattgctga aggtaagtcc 1440

gttggtatca tcggtgctta a 1461

<210> 669
<211> 486
<212> PRT
<213> Artificial Sequence

<220>
<223> HXK2 from S. cerevisiae

<400> 669

Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met
1 5 10 15

Ala Asp Val Pro Lys Glu Leu Met Gln Gln Ile Glu Asn Phe Glu Lys
20 25 30

Ile Phe Thr Val Pro Thr Glu Thr Leu Gln Ala Val Thr Lys His Phe
35 40 45

Ile Ser Glu Leu Glu Lys Gly Leu Ser Lys Lys Gly Gly Asn Ile Pro
50 55 60

Met Ile Pro Gly Trp Val Met Asp Phe Pro Thr Gly Lys Glu Ser Gly
65 70 75 80

Asp Phe Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu
85 90 95

Val Lys Leu Gly Gly Asp Arg Thr Phe Asp Thr Thr Gln Ser Lys Tyr
100 105 110

Arg Leu Pro Asp Ala Met Arg Thr Thr Gln Asn Pro Asp Glu Leu Trp
115 120 125

Glu Phe Ile Ala Asp Ser Leu Lys Ala Phe Ile Asp Glu Gln Phe Pro
130 135 140

Gln Gly Ile Ser Glu Pro Ile Pro Leu Gly Phe Thr Phe Ser Phe Pro
145 150 155 160

Ala Ser Gln Asn Lys Ile Asn Glu Gly Ile Leu Gln Arg Trp Thr Lys
165 170 175

Gly Phe Asp Ile Pro Asn Ile Glu Asn His Asp Val Val Pro Met Leu
180 185 190

```
Gln Lys Gln Ile Thr Lys Arg Asn Ile Pro Ile Glu Val Val Ala Leu
        195                 200                 205

Ile Asn Asp Thr Thr Gly Thr Leu Val Ala Ser Tyr Tyr Thr Asp Pro
    210                 215                 220

Glu Thr Lys Met Gly Val Ile Phe Gly Thr Gly Val Asn Gly Ala Tyr
225                 230                 235                 240

Tyr Asp Val Cys Ser Asp Ile Glu Lys Leu Gln Gly Lys Leu Ser Asp
                245                 250                 255

Asp Ile Pro Pro Ser Ala Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser
            260                 265                 270

Phe Asp Asn Glu His Val Val Leu Pro Arg Thr Lys Tyr Asp Ile Thr
        275                 280                 285

Ile Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Thr Phe Glu Lys Met
    290                 295                 300

Ser Ser Gly Tyr Tyr Leu Gly Glu Ile Leu Arg Leu Ala Leu Met Asp
305                 310                 315                 320

Met Tyr Lys Gln Gly Phe Ile Phe Lys Asn Gln Asp Leu Ser Lys Phe
                325                 330                 335

Asp Lys Pro Phe Val Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Glu
            340                 345                 350

Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Gln Asn Glu
        355                 360                 365

Phe Gly Ile Asn Thr Thr Val Gln Glu Arg Lys Leu Ile Arg Arg Leu
    370                 375                 380

Ser Glu Leu Ile Gly Ala Arg Ala Ala Arg Leu Ser Val Cys Gly Ile
385                 390                 395                 400

Ala Ala Ile Cys Gln Lys Arg Gly Tyr Lys Thr Gly His Ile Ala Ala
                405                 410                 415

Asp Gly Ser Val Tyr Asn Arg Tyr Pro Gly Phe Lys Glu Lys Ala Ala
            420                 425                 430

Asn Ala Leu Lys Asp Ile Tyr Gly Trp Thr Gln Thr Ser Leu Asp Asp
        435                 440                 445

Tyr Pro Ile Lys Ile Val Pro Ala Glu Asp Gly Ser Gly Ala Gly Ala
    450                 455                 460

Ala Val Ile Ala Ala Leu Ala Gln Lys Arg Ile Ala Glu Gly Lys Ser
465                 470                 475                 480
```

20 Jan 2017 2017200439

Val Gly Ile Ile Gly Ala
485

<210> 670
<211> 1458
<212> DNA
<213> Artificial Sequence

<220>
<223> RAG5 from K. lactis

<400> 670
atggttcgtt taggtccaaa gaagcctcca gccagaaagg ggtccatggc agatgtgcca 60

gctaatttga tggaacaaat ccacggtttg gaaactttgt tcaccgtctc ttcagaaaaa 120

atgagaagca ttgtcaagca tttcatcagt gaattggaca aaggtttgtc caaaaagggt 180

ggtaacattc ctatgattcc aggttgggtt gttgagtatc caactggtaa ggaaactggt 240

gatttcttag ctcttgattt gggtggtacc aacttgagag ttgtgttggt taaattgggt 300

ggtaatcatg atttcgacac cactcaaaac aagtacagat taccagacca tttgagaact 360

ggtacttctg aacaattgtg gtcatttatt gcaaagtgtt tgaaggaatt cgtcgatgaa 420

tggtacccag atggtgtttc tgaaccattg ccattgggtt tcactttctc ataccctgca 480

tctcaaaaga agatcaattc cggtgtgttg caacgttgga ccaagggttt cgatattgaa 540

ggtgttgaag gtcacgatgt tgttccaatg ctacaagaac agattgaaaa gctgaatatc 600

ccaatcaatg tcgttcgatt gatcaacgat accactggta ccttggttgc ctctttgtac 660

actgatcctc aaactaagat gggtatcatt atcggtactg gtgtcaacgg tgcttactac 720

gatgttgttt ctggtattga gaaattggaa ggtttgttgc cagaagatat cggtccagat 780

tctccaatgg caatcaactg tgaatatggt tccttcgata acgaacattt ggtgttgcca 840

agaaccaaat acgatgttat aatcgatgaa gaatctccaa gaccaggtca acaagctttc 900

gaaaagatga cttctggtta ctatctaggt gaaatcatgc gtctagtact attggacttg 960

tacgacagtg gtttcatctt taaggaccaa gatatctcca agttgaaaga ggcttacgtc 1020

atggacacca gttatccatc taagatcgaa gatgatccat tcgaaaactt ggaagacact 1080

gacgatctgt tcaagactaa cttgaacatc gaaactaccg ttgttgagag aaagttgatt 1140

agaaaattag ccgaattggt cggaacaaga gctgcaagat tgactgtttg tggtgtttct 1200

gctatctgtg acaagagagg ctacaagact gctcacattg cagctgatgg ttctgtcttc 1260

aacagatacc caggttacaa ggaaaaggcc gctcaagcct tgaaggatat ctacaactgg 1320

gatgtcgaaa agatggaaga ccacccaatc caattggtgg ctgctgaaga tggttccggt 1380

gttggtgctg ctatcattgc ttgtttgact caaaagagat tggctgccgg taagtctgtt 1440

ggtattaaag gcgaatag 1458

<210> 671
<211> 485
<212> PRT
<213> Artificial Sequence

<220>
<223> RAG5 from K. lactis

<400> 671

```
Met Val Arg Leu Gly Pro Lys Lys Pro Pro Ala Arg Lys Gly Ser Met
1               5                  10                  15

Ala Asp Val Pro Ala Asn Leu Met Glu Gln Ile His Gly Leu Glu Thr
            20                  25                  30

Leu Phe Thr Val Ser Ser Glu Lys Met Arg Ser Ile Val Lys His Phe
        35                  40                  45

Ile Ser Glu Leu Asp Lys Gly Leu Ser Lys Lys Gly Gly Asn Ile Pro
    50                  55                  60

Met Ile Pro Gly Trp Val Val Glu Tyr Pro Thr Gly Lys Glu Thr Gly
65                  70                  75                  80

Asp Phe Leu Ala Leu Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu
                85                  90                  95

Val Lys Leu Gly Gly Asn His Asp Phe Asp Thr Thr Gln Asn Lys Tyr
            100                 105                 110

Arg Leu Pro Asp His Leu Arg Thr Gly Thr Ser Glu Gln Leu Trp Ser
        115                 120                 125

Phe Ile Ala Lys Cys Leu Lys Glu Phe Val Asp Glu Trp Tyr Pro Asp
    130                 135                 140

Gly Val Ser Glu Pro Leu Pro Leu Gly Phe Thr Phe Ser Tyr Pro Ala
145                 150                 155                 160

Ser Gln Lys Lys Ile Asn Ser Gly Val Leu Gln Arg Trp Thr Lys Gly
                165                 170                 175

Phe Asp Ile Glu Gly Val Glu Gly His Asp Val Val Pro Met Leu Gln
            180                 185                 190

Glu Gln Ile Glu Lys Leu Asn Ile Pro Ile Asn Val Val Arg Leu Ile
        195                 200                 205

Asn Asp Thr Thr Gly Thr Leu Val Ala Ser Leu Tyr Thr Asp Pro Gln
    210                 215                 220

Thr Lys Met Gly Ile Ile Ile Gly Thr Gly Val Asn Gly Ala Tyr Tyr
225                 230                 235                 240

Asp Val Val Ser Gly Ile Glu Lys Leu Glu Gly Leu Leu Pro Glu Asp
                245                 250                 255

Ile Gly Pro Asp Ser Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser Phe
            260                 265                 270

Asp Asn Glu His Leu Val Leu Pro Arg Thr Lys Tyr Asp Val Ile Ile
        275                 280                 285
```

```
Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Ala Phe Glu Lys Met Thr
    290                 295                 300

Ser Gly Tyr Tyr Leu Gly Glu Ile Met Arg Leu Val Leu Leu Asp Leu
305                 310                 315                 320

Tyr Asp Ser Gly Phe Ile Phe Lys Asp Gln Asp Ile Ser Lys Leu Lys
                325                 330                 335

Glu Ala Tyr Val Met Asp Thr Ser Tyr Pro Ser Lys Ile Glu Asp Asp
            340                 345                 350

Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Lys Thr Asn Leu
        355                 360                 365

Asn Ile Glu Thr Thr Val Val Glu Arg Lys Leu Ile Arg Lys Leu Ala
    370                 375                 380

Glu Leu Val Gly Thr Arg Ala Ala Arg Leu Thr Val Cys Gly Val Ser
385                 390                 395                 400

Ala Ile Cys Asp Lys Arg Gly Tyr Lys Thr Ala His Ile Ala Ala Asp
                405                 410                 415

Gly Ser Val Phe Asn Arg Tyr Pro Gly Tyr Lys Glu Lys Ala Ala Gln
            420                 425                 430

Ala Leu Lys Asp Ile Tyr Asn Trp Asp Val Glu Lys Met Glu Asp His
        435                 440                 445

Pro Ile Gln Leu Val Ala Ala Glu Asp Gly Ser Gly Val Gly Ala Ala
    450                 455                 460

Ile Ile Ala Cys Leu Thr Gln Lys Arg Leu Ala Ala Gly Lys Ser Val
465                 470                 475                 480

Gly Ile Lys Gly Glu
                485
```

<210> 672
<211> 1416
<212> DNA
<213> Artificial Sequence

<220>
<223> HPGLK1 from H. polymorpha

<400> 672

```
atgagtttgg atactgaagt cgataagatt gtgtcggagt ttgccgtcac ccaggagaca       60

ctccaaaagg gtgtggagcg tttcattgag cttgcaactg ccggactgaa tagtgatgag      120

gacaagtatg gtctgccaat gatcccaact tttgttacct ccatcccaac cggtaaagag      180

aagggcattc tttttgccgc agacttggga ggaaccaatt tcagagtttg ctctgttgcc      240

ttgaacggag atcacacttt caaactgatc cagcagaagt cacatattcc tgccgaactg      300
```

atgacctcca cctcggacga attgttttcg tatcttgcaa gcaaggtcaa gaatttctta 360
gagactcatc atgaaggggc tgttacttct acaggaagcc agaaattcaa gatgggtttc 420
actttcagtt tccctgtctc gcagaccgcc ttaaacgccg gtactttgct aagatggacc 480
aagggattca atattccgga tactgttggt caagaggttg tttctctatt ccaaatgcat 540
ttagacgccc aggaaattcc tgttactgtg tctgccctgt ccaacgatac tgtgggaacc 600
cttcttgcaa gatcctacac gggttccaat aaggagggca ctactgttct aggatgcatc 660
ttcggaacgg gaacaaacgg tgcttacaac gagaagctcg agaatatcaa gaagcttccg 720
gccgaggtga gagagaagct gaaggctcaa ggtgtcaccc acatggtcat taatactgaa 780
tggggttcct tcgataacca gctcaaggtt ttgccaaata cgaagtatga cgctcaagtt 840
gacgaactta ccggcaataa gggcttccac atgtttgaaa agcgtgtttc cggaatgttc 900
ttgggtgaga ttctgagaca tattttggtc gaccttcact ctaagggagt gctatttact 960
cagtacgcca gctacgaatc cctgccccac agattgagga cgccgtggga tctggactct 1020
gaggttctct cactgattga gatcgacgaa tccaccaatt tgcaggccac tgagctgtct 1080
ttgaaacagg cattgagact gccaactact actgaggaga gacttgctat tcaaaaactt 1140
actcgtgctg tggccaagag atctgcctat cttgctgcta ttcctattgc tgctattcta 1200
cacatgaccg agtcttttaa gggccacaac gttgaggtgg acgttggagc agacgggtct 1260
gtggttgagt tctaccctgg attcagaact atgatgagag acgccattgc gcagacgcag 1320
ataggtgcca aaggagagag aagactgcac attaacattg ccaaagacgg ctcatctgtg 1380
ggcgctgcat tgtgcgcatt aagcgagaaa gactaa 1416

<210> 673
<211> 471
<212> PRT
<213> Artificial Sequence

<220>
<223> HPGLK1 from H. polymorpha

<400> 673

Met Ser Leu Asp Thr Glu Val Asp Lys Ile Val Ser Glu Phe Ala Val
1 5 10 15

Thr Gln Glu Thr Leu Gln Lys Gly Val Glu Arg Phe Ile Glu Leu Ala
20 25 30

Thr Ala Gly Leu Asn Ser Asp Glu Asp Lys Tyr Gly Leu Pro Met Ile
35 40 45

Pro Thr Phe Val Thr Ser Ile Pro Thr Gly Lys Glu Lys Gly Ile Leu
50 55 60

Phe Ala Ala Asp Leu Gly Gly Thr Asn Phe Arg Val Cys Ser Val Ala
65 70 75 80

Leu Asn Gly Asp His Thr Phe Lys Leu Ile Gln Gln Lys Ser His Ile
85 90 95

```
Pro Ala Glu Leu Met Thr Ser Thr Ser Asp Glu Leu Phe Ser Tyr Leu
            100                 105                 110

Ala Ser Lys Val Lys Asn Phe Leu Glu Thr His His Glu Gly Ala Val
        115                 120                 125

Thr Ser Thr Gly Ser Gln Lys Phe Lys Met Gly Phe Thr Phe Ser Phe
    130                 135                 140

Pro Val Ser Gln Thr Ala Leu Asn Ala Gly Thr Leu Leu Arg Trp Thr
145                 150                 155                 160

Lys Gly Phe Asn Ile Pro Asp Thr Val Gly Gln Glu Val Val Ser Leu
                165                 170                 175

Phe Gln Met His Leu Asp Ala Gln Glu Ile Pro Val Thr Val Ser Ala
            180                 185                 190

Leu Ser Asn Asp Thr Val Gly Thr Leu Leu Ala Arg Ser Tyr Thr Gly
        195                 200                 205

Ser Asn Lys Glu Gly Thr Thr Val Leu Gly Cys Ile Phe Gly Thr Gly
    210                 215                 220

Thr Asn Gly Ala Tyr Asn Glu Lys Leu Glu Asn Ile Lys Lys Leu Pro
225                 230                 235                 240

Ala Glu Val Arg Glu Lys Leu Lys Ala Gln Gly Val Thr His Met Val
                245                 250                 255

Ile Asn Thr Glu Trp Gly Ser Phe Asp Asn Gln Leu Lys Val Leu Pro
            260                 265                 270

Asn Thr Lys Tyr Asp Ala Gln Val Asp Glu Leu Thr Gly Asn Lys Gly
        275                 280                 285

Phe His Met Phe Glu Lys Arg Val Ser Gly Met Phe Leu Gly Glu Ile
    290                 295                 300

Leu Arg His Ile Leu Val Asp Leu His Ser Lys Gly Val Leu Phe Thr
305                 310                 315                 320

Gln Tyr Ala Ser Tyr Glu Ser Leu Pro His Arg Leu Arg Thr Pro Trp
                325                 330                 335

Asp Leu Asp Ser Glu Val Leu Ser Leu Ile Glu Ile Asp Glu Ser Thr
            340                 345                 350

Asn Leu Gln Ala Thr Glu Leu Ser Leu Lys Gln Ala Leu Arg Leu Pro
        355                 360                 365

Thr Thr Thr Glu Glu Arg Leu Ala Ile Gln Lys Leu Thr Arg Ala Val
    370                 375                 380
```

```
Ala Lys Arg Ser Ala Tyr Leu Ala Ala Ile Pro Ile Ala Ala Ile Leu
385                 390                 395                 400

His Met Thr Glu Ser Phe Lys Gly His Asn Val Glu Val Asp Val Gly
                405                 410                 415

Ala Asp Gly Ser Val Val Glu Phe Tyr Pro Gly Phe Arg Thr Met Met
            420                 425                 430

Arg Asp Ala Ile Ala Gln Thr Gln Ile Gly Ala Lys Gly Glu Arg Arg
        435                 440                 445

Leu His Ile Asn Ile Ala Lys Asp Gly Ser Ser Val Gly Ala Ala Leu
    450                 455                 460

Cys Ala Leu Ser Glu Lys Asp
465                 470
```

<210> 674
<211> 1634
<212> DNA
<213> Artificial Sequence

<220>
<223> HXK2 from S. pombe

<400> 674

```
aacacttttc gcctcacttg cgaatctacg aaaggaatat ataggtggtt caccccttt       60

cttttcattt cgtgttttta atagttattt acatcaacag agataactat ttctgttaac      120

gatttttttt cccacttgtt ttcttccttt tttggtgaat tttaattaat ttataataag      180

caatggaggc taattttcaa caagctgtta aaaagttagt caatgacttt gaatacccta      240

ccgagtcctt gagagaggcc gttaaggagt ttgacgaatt acgtcaaaag ggtttacaaa      300

agaatggtga ggtgcttgct atggctcctg cctttatctc tacccttccc accggcgctg      360

aaactggtga cttcttggcc cttgactttg gtggtaccaa cttgcgtgtt tgttggatcc      420

aacttctcgg tgacggcaag tatgagatga agcacagcaa gtccgtcttg ccccgtgaat      480

gcgttcgtaa cgagtctgtt aagcccatca ttgactttat gagtgaccat gttgagcttt      540

tcatcaagga gcacttccct tccaagtttg gctgccctga ggaggaatac cttcctatgg      600

gtttcacctt ttcttatccc gccaaccaag tttccatcac cgagagctac ttgcttcgtt      660

ggaccaaggg tcttaacatt cctgaggcca tcaacaagga ctttgcccaa tttttgactg      720

aaggtttcaa ggctcgtaac cttcctatta gaatcgaggc tgtcatcaac gataccgtcg      780

gtactctcgt tacccgtgct tatacttcaa aggagagcga cacctttatg ggtatcattt      840

tcggaaccgg taccaacggt gcttacgtcg agcaaatgaa ccaaattccc aagcttgctg      900

gcaagtgtac tggtgatcat atgcttatca acatggaatg gggagcaact gatttctctt      960

gccttcactc cactcgttat gatttacttc ttgatcatga tactcccaat gctggtcgtc     1020

aaatctttga gaagcgcgtt ggtggtatgt atctcggtga gcttttccgc cgtgccttat     1080

tccacttgat caaggtttac aacttcaacg aaggtatttt ccctccttcc attactgatg     1140
```

cttggtcttt ggaaacttct gttctttcca gaatgatggt tgaacgttct gctgagaatg 1200

ttcgtaacgt tcttagtaca ttcaagttcc gtttccgcag cgacgaagag gctttgtacc 1260

tttgggatgc tgctcatgca attggccgtc gtgctgctcg tatgtctgcc gttcccattg 1320

cttctttgta tctttctacc ggccgcgctg gtaagaagag tgatgttggt gttgatggtt 1380

ctttagtcga acactatcct cactttgttg acatgctccg tgaagccttg cgtgagctta 1440

tcggtgataa cgaaaaattg atttccattg gtattgccaa ggatggcagt ggtattggtg 1500

ccgctctttg cgccctccaa gctgttaagg aaaagaaagg cttggcctaa atcatgttag 1560

atgtctgtta gctttttttg aattgtacgt agaaatgagc atgtaaatat gaaattgctt 1620

tttaacagct ttta 1634

<210> 675
<211> 455
<212> PRT
<213> Artificial Sequence

<220>
<223> HXK2 from S. pombe

<400> 675

Met Glu Ala Asn Phe Gln Gln Ala Val Lys Lys Leu Val Asn Asp Phe
1 5 10 15

Glu Tyr Pro Thr Glu Ser Leu Arg Glu Ala Val Lys Glu Phe Asp Glu
20 25 30

Leu Arg Gln Lys Gly Leu Gln Lys Asn Gly Glu Val Leu Ala Met Ala
35 40 45

Pro Ala Phe Ile Ser Thr Leu Pro Thr Gly Ala Glu Thr Gly Asp Phe
50 55 60

Leu Ala Leu Asp Phe Gly Gly Thr Asn Leu Arg Val Cys Trp Ile Gln
65 70 75 80

Leu Leu Gly Asp Gly Lys Tyr Glu Met Lys His Ser Lys Ser Val Leu
85 90 95

Pro Arg Glu Cys Val Arg Asn Glu Ser Val Lys Pro Ile Ile Asp Phe
100 105 110

Met Ser Asp His Val Glu Leu Phe Ile Lys Glu His Phe Pro Ser Lys
115 120 125

Phe Gly Cys Pro Glu Glu Glu Tyr Leu Pro Met Gly Phe Thr Phe Ser
130 135 140

Tyr Pro Ala Asn Gln Val Ser Ile Thr Glu Ser Tyr Leu Leu Arg Trp
145 150 155 160

Thr Lys Gly Leu Asn Ile Pro Glu Ala Ile Asn Lys Asp Phe Ala Gln
165 170 175

```
Phe Leu Thr Glu Gly Phe Lys Ala Arg Asn Leu Pro Ile Arg Ile Glu
            180                 185                 190

Ala Val Ile Asn Asp Thr Val Gly Thr Leu Val Thr Arg Ala Tyr Thr
        195                 200                 205

Ser Lys Glu Ser Asp Thr Phe Met Gly Ile Ile Phe Gly Thr Gly Thr
    210                 215                 220

Asn Gly Ala Tyr Val Glu Gln Met Asn Gln Ile Pro Lys Leu Ala Gly
225                 230                 235                 240

Lys Cys Thr Gly Asp His Met Leu Ile Asn Met Glu Trp Gly Ala Thr
                245                 250                 255

Asp Phe Ser Cys Leu His Ser Thr Arg Tyr Asp Leu Leu Leu Asp His
            260                 265                 270

Asp Thr Pro Asn Ala Gly Arg Gln Ile Phe Glu Lys Arg Val Gly Gly
        275                 280                 285

Met Tyr Leu Gly Glu Leu Phe Arg Arg Ala Leu Phe His Leu Ile Lys
    290                 295                 300

Val Tyr Asn Phe Asn Glu Gly Ile Phe Pro Pro Ser Ile Thr Asp Ala
305                 310                 315                 320

Trp Ser Leu Glu Thr Ser Val Leu Ser Arg Met Met Val Glu Arg Ser
                325                 330                 335

Ala Glu Asn Val Arg Asn Val Leu Ser Thr Phe Lys Phe Arg Phe Arg
            340                 345                 350

Ser Asp Glu Glu Ala Leu Tyr Leu Trp Asp Ala Ala His Ala Ile Gly
        355                 360                 365

Arg Arg Ala Ala Arg Met Ser Ala Val Pro Ile Ala Ser Leu Tyr Leu
    370                 375                 380

Ser Thr Gly Arg Ala Gly Lys Lys Ser Asp Val Gly Val Asp Gly Ser
385                 390                 395                 400

Leu Val Glu His Tyr Pro His Phe Val Asp Met Leu Arg Glu Ala Leu
                405                 410                 415

Arg Glu Leu Ile Gly Asp Asn Glu Lys Leu Ile Ser Ile Gly Ile Ala
            420                 425                 430

Lys Asp Gly Ser Gly Ile Gly Ala Ala Leu Cys Ala Leu Gln Ala Val
        435                 440                 445

Lys Glu Lys Lys Gly Leu Ala
    450                 455
```

<210> 676
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> YMR226C

<400> 676
atgtcccaag gtagaaaagc tgcagaaaga ttggctaaga agactgtcct cattacaggt 60

gcatctgctg gtattggtaa ggcgaccgca ttagagtact tggaggcatc caatggtgat 120

atgaaactga tcttggctgc tagaagatta gaaaagctcg aggaattgaa gaagaccatt 180

gatcaagagt ttccaaacgc aaaagttcat gtggcccagc tggatatcac tcaagcagaa 240

aaaatcaagc ccttcattga aaacttgcca caagagttca aggatattga cattctggtg 300

aacaatgccg gaaaggctct tggcagtgac cgtgtgggcc agatcgcaac ggaggatatc 360

caggacgtgt ttgacaccaa cgtcacggct ttaatcaata tcacacaagc tgtactgccc 420

atattccaag ccaagaattc aggagatatt gtaaatttgg gttcaatcgc tggcagagac 480

gcatacccaa caggttctat ctattgtgcc tctaagtttg ccgtgggggc gttcactgat 540

agtttgagaa aggagctcat caacactaaa attagagtca ttctaattgc accagggcta 600

gtcgagactg aattttcact agttagatac agaggtaacg aggaacaagc caagaatgtt 660

tacaaggata ctaccccatt gatggctgat gacgtggctg atctgatcgt ctatgcaact 720

tccagaaaac aaaatactgt aattgcagac actttaatct ttccaacaaa ccaagcgtca 780

cctcatcata tcttccgtgg ataa 804

<210> 677
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> YMR226C

<400> 677

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Lys Lys Thr Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Ile Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Gln Glu Phe
50 55 60

Pro Asn Ala Lys Val His Val Ala Gln Leu Asp Ile Thr Gln Ala Glu
65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Gln Glu Phe Lys Asp Ile
85 90 95

```
Asp Ile Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Arg Val
            100                 105                 110

Gly Gln Ile Ala Thr Glu Asp Ile Gln Asp Val Phe Asp Thr Asn Val
        115                 120                 125

Thr Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Ile Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly
                165                 170                 175

Ala Phe Thr Asp Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg
            180                 185                 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr
    210                 215                 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
                245                 250                 255

Asn Gln Ala Ser Pro His His Ile Phe Arg Gly
            260                 265
```

<210> 678
<211> 1410
<212> DNA
<213> Artificial Sequence

<220>
<223> YIL074C (Chr 9)

<400> 678

```
atgtcttatt cagctgccga taatttacaa gattcattcc aacgtgccat gaacttttct       60

ggctctcctg gtgcagtctc aacctcacca actcagtcat ttatgaacac actacctcgt      120

cgtgtaagca ttacaaagca accaaaggct ttaaaacctt tttctactgg tgacatgaat      180

attctactgt tggaaaatgt caatgcaact gcaatcaaaa tcttcaagga tcagggttac      240

caagtagagt tccacaagtc ttctctacct gaggatgaat tgattgaaaa aatcaaagac      300

gtacacgcta tcggtataag atccaaaact agattgactg aaaaaatact acagcatgcc      360

aggaatctag tttgtattgg ttgtttttgc ataggtacca atcaagtaga cctaaaatat      420

gccgctagta aaggtattgc tgttttcaat tcgccattct ccaattcaag atccgtagca      480

gaattggtaa ttggtgagat cattagttta gcaagacaat taggtgatag atccattgaa      540
```

ctgcatacag gtacatggaa taaagtcgct gctaggtgtt gggaagtaag aggaaaaact 600
ctcggtatta ttgggtatgg tcacattggt tcgcaattat cagttcttgc agaagctatg 660
ggcctgcatg tgctatacta tgatatcgtg acaattatgg ccttaggtac tgccagacaa 720
gtttctacat tagatgaatt gttgaataaa tctgattttg taacactaca tgtaccagct 780
actccagaaa ctgaaaaaat gttatctgct ccacaattcg ctgctatgaa ggacggggct 840
tatgttatta atgcctcaag aggtactgtc gtggacattc catctctgat ccaagccgtc 900
aaggccaaca aaattgcagg tgctgcttta gatgtttatc cacatgaacc agctaagaac 960
ggtgaaggtt catttaacga tgaacttaac agctggactt ctgagttggt ttcattacca 1020
aatataatcc tgacaccaca tattggtggc tctacagaag aagctcaaag ttcaatcggt 1080
attgaggtgg ctactgcatt gtccaaatac atcaatgaag gtaactctgt cggttctgtg 1140
aacttcccag aagtcagttt gaagtctttg gactacgatc aagagaacac agtacgtgtc 1200
ttgtatattc atcgtaacgt tcctggtgtt ttgaagaccg ttaatgatat cttatccgat 1260
cataatatcg agaaacagtt ttctgattct cacggcgaga tcgcttatct aatggcagac 1320
atctcttctg ttaatcaaag tgaaatcaag gatatatatg aaaagttgaa ccaaacttct 1380
gccaaagttt ccatcaggtt attatactaa 1410

<210> 679
<211> 469
<212> PRT
<213> Artificial Sequence

<220>
<223> YIL074C (Chr 9)

<400> 679

Met Ser Tyr Ser Ala Ala Asp Asn Leu Gln Asp Ser Phe Gln Arg Ala
1 5 10 15

Met Asn Phe Ser Gly Ser Pro Gly Ala Val Ser Thr Ser Pro Thr Gln
20 25 30

Ser Phe Met Asn Thr Leu Pro Arg Arg Val Ser Ile Thr Lys Gln Pro
35 40 45

Lys Ala Leu Lys Pro Phe Ser Thr Gly Asp Met Asn Ile Leu Leu Leu
50 55 60

Glu Asn Val Asn Ala Thr Ala Ile Lys Ile Phe Lys Asp Gln Gly Tyr
65 70 75 80

Gln Val Glu Phe His Lys Ser Ser Leu Pro Glu Asp Glu Leu Ile Glu
85 90 95

Lys Ile Lys Asp Val His Ala Ile Gly Ile Arg Ser Lys Thr Arg Leu
100 105 110

Thr Glu Lys Ile Leu Gln His Ala Arg Asn Leu Val Cys Ile Gly Cys
115 120 125

```
Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Lys Tyr Ala Ala Ser Lys
    130                 135                 140

Gly Ile Ala Val Phe Asn Ser Pro Phe Ser Asn Ser Arg Ser Val Ala
145                 150                 155                 160

Glu Leu Val Ile Gly Glu Ile Ile Ser Leu Ala Arg Gln Leu Gly Asp
                165                 170                 175

Arg Ser Ile Glu Leu His Thr Gly Thr Trp Asn Lys Val Ala Ala Arg
            180                 185                 190

Cys Trp Glu Val Arg Gly Lys Thr Leu Gly Ile Ile Gly Tyr Gly His
        195                 200                 205

Ile Gly Ser Gln Leu Ser Val Leu Ala Glu Ala Met Gly Leu His Val
    210                 215                 220

Leu Tyr Tyr Asp Ile Val Thr Ile Met Ala Leu Gly Thr Ala Arg Gln
225                 230                 235                 240

Val Ser Thr Leu Asp Glu Leu Leu Asn Lys Ser Asp Phe Val Thr Leu
                245                 250                 255

His Val Pro Ala Thr Pro Glu Thr Glu Lys Met Leu Ser Ala Pro Gln
            260                 265                 270

Phe Ala Ala Met Lys Asp Gly Ala Tyr Val Ile Asn Ala Ser Arg Gly
        275                 280                 285

Thr Val Val Asp Ile Pro Ser Leu Ile Gln Ala Val Lys Ala Asn Lys
    290                 295                 300

Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn
305                 310                 315                 320

Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu
                325                 330                 335

Val Ser Leu Pro Asn Ile Ile Leu Thr Pro His Ile Gly Gly Ser Thr
            340                 345                 350

Glu Glu Ala Gln Ser Ser Ile Gly Ile Glu Val Ala Thr Ala Leu Ser
        355                 360                 365

Lys Tyr Ile Asn Glu Gly Asn Ser Val Gly Ser Val Asn Phe Pro Glu
    370                 375                 380

Val Ser Leu Lys Ser Leu Asp Tyr Asp Gln Glu Asn Thr Val Arg Val
385                 390                 395                 400

Leu Tyr Ile His Arg Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp
                405                 410                 415
```

Ile Leu Ser Asp His Asn Ile Glu Lys Gln Phe Ser Asp Ser His Gly
420 425 430

Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asn Gln Ser Glu
435 440 445

Ile Lys Asp Ile Tyr Glu Lys Leu Asn Gln Thr Ser Ala Lys Val Ser
450 455 460

Ile Arg Leu Leu Tyr
465

<210> 680
<211> 792
<212> DNA
<213> Artificial Sequence

<220>
<223> YIR036C (Chr 9)

<400> 680
atgggcaagg ttattttgat tacaggtgcc tcccgtggga ttggcctgca attggtgaaa 60

actgttatcg aagaggacga tgaatgcatc gtctacggcg tagcaagaac ggaagctggt 120

ctgcagtctt tgcaaagaga atacggtgca gacaaatttg tctatcgtgt cctcgacatc 180

acggacaggt ctcgaatgga agcgttggtg gaggaaatcc ggcaaaagca tggaaaactg 240

gacggtattg tcgcaaatgc ggggatgcta gaaccggtga agtccatctc ccagtccaac 300

tccgaacacg acatcaagca gtgggaacgg ctgttcgatg tgaacttttt cagcattgtc 360

tctttggtgg cactgtgttt acccctcttg aagagctcgc catttgtagg caacattgtc 420

ttcgtcagct ctggagccag tgtgaaacca tataacggat ggtcggcgta cggctgctcg 480

aaagccgcat taaaccactt tgccatggac attgccagtg aagagcccag tgataaagtg 540

cgtgccgtgt gtattgcacc gggcgtcgtt gacacgcaga tgcagaaaga tattagggaa 600

acattgggtc ctcagggcat gacacccaag gctctcgaga ggtttactca attgtacaag 660

acttcgtcac tgctggaccc aaaggtgcct gcggcggtac tagcgcaact cgtcctgaaa 720

ggtattcccg actctttgaa cggtcaatat ctccgctaca acgatgagcg actggggccg 780

gtgcagggct ag 792

<210> 681
<211> 263
<212> PRT
<213> Artificial Sequence

<220>
<223> YIR036C (Chr 9)

<400> 681

Met Gly Lys Val Ile Leu Ile Thr Gly Ala Ser Arg Gly Ile Gly Leu
1 5 10 15

Gln Leu Val Lys Thr Val Ile Glu Glu Asp Asp Glu Cys Ile Val Tyr
20 25 30

```
Gly Val Ala Arg Thr Glu Ala Gly Leu Gln Ser Leu Gln Arg Glu Tyr
        35                  40                  45

Gly Ala Asp Lys Phe Val Tyr Arg Val Leu Asp Ile Thr Asp Arg Ser
    50                  55                  60

Arg Met Glu Ala Leu Val Glu Glu Ile Arg Gln Lys His Gly Lys Leu
65                  70                  75                  80

Asp Gly Ile Val Ala Asn Ala Gly Met Leu Glu Pro Val Lys Ser Ile
                85                  90                  95

Ser Gln Ser Asn Ser Glu His Asp Ile Lys Gln Trp Glu Arg Leu Phe
            100                 105                 110

Asp Val Asn Phe Phe Ser Ile Val Ser Leu Val Ala Leu Cys Leu Pro
        115                 120                 125

Leu Leu Lys Ser Ser Pro Phe Val Gly Asn Ile Val Phe Val Ser Ser
    130                 135                 140

Gly Ala Ser Val Lys Pro Tyr Asn Gly Trp Ser Ala Tyr Gly Cys Ser
145                 150                 155                 160

Lys Ala Ala Leu Asn His Phe Ala Met Asp Ile Ala Ser Glu Glu Pro
                165                 170                 175

Ser Asp Lys Val Arg Ala Val Cys Ile Ala Pro Gly Val Val Asp Thr
            180                 185                 190

Gln Met Gln Lys Asp Ile Arg Glu Thr Leu Gly Pro Gln Gly Met Thr
        195                 200                 205

Pro Lys Ala Leu Glu Arg Phe Thr Gln Leu Tyr Lys Thr Ser Ser Leu
    210                 215                 220

Leu Asp Pro Lys Val Pro Ala Ala Val Leu Ala Gln Leu Val Leu Lys
225                 230                 235                 240

Gly Ile Pro Asp Ser Leu Asn Gly Gln Tyr Leu Arg Tyr Asn Asp Glu
                245                 250                 255

Arg Leu Gly Pro Val Gln Gly
            260
```

<210> 682
<211> 1503
<212> DNA
<213> Artificial Sequence

<220>
<223> YPL061W (ALD6)(Chr 16)

<400> 682
atgactaagc tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttg 60

acatacgagc aaccaaccgg tctattcatt aacaacaagt ttatgaaagc tcaagacggt 120
aagacctatc ccgtcgaaga tccttccact gaaaacaccg tttgtgaggt ctcttctgcc 180
accactgaag atgttgaata tgctatcgaa tgtgccgacc gtgctttcca cgacactgaa 240
tgggctaccc aagacccaag agaaagaggc cgtctactaa gtaagttggc tgacgaattg 300
gaaagccaaa ttgacttggt ttcttccatt gaagctttgg acaatggtaa aactttggcc 360
ttagcccgtg gggatgttac cattgcaatc aactgtctaa gagatgctgc tgcctatgcc 420
gacaaagtca acggtagaac aatcaacacc ggtgacggct acatgaactt caccacctta 480
gagccaatcg gtgtctgtgg tcaaattatt ccatggaact ttccaataat gatgttggct 540
tggaagatcg ccccagcatt ggccatgggt aacgtctgta tcttgaaacc cgctgctgtc 600
acacctttaa atgccctata ctttgcttct ttatgtaaga aggttggtat tccagctggt 660
gtcgtcaaca tcgttccagg tcctggtaga actgttggtg ctgctttgac caacgaccca 720
agaatcagaa agctggcttt taccggttct acagaagtcg gtaagagtgt tgctgtcgac 780
tcttctgaat ctaacttgaa gaaaatcact ttggaactag gtggtaagtc cgcccatttg 840
gtctttgacg atgctaacat taagaagact ttaccaaatc tagtaaacgg tattttcaag 900
aacgctggtc aaatttgttc ctctggttct agaatttacg ttcaagaagg tatttacgac 960
gaactattgg ctgctttcaa ggcttacttg gaaaccgaaa tcaaagttgg taatccattt 1020
gacaaggcta acttccaagg tgctatcact aaccgtcaac aattcgacac aattatgaac 1080
tacatcgata tcggtaagaa agaaggcgcc aagatcttaa ctggtggcga aaaagttggt 1140
gacaagggtt acttcatcag accaaccgtt ttctacgatg ttaatgaaga catgagaatt 1200
gttaaggaag aaatttttgg accagttgtc actgtcgcaa agttcaagac tttagaagaa 1260
ggtgtcgaaa tggctaacag ctctgaattc ggtctaggtt ctggtatcga aacagaatct 1320
ttgagcacag gtttgaaggt ggccaagatg ttgaaggccg gtaccgtctg gatcaacaca 1380
tacaacgatt ttgactccag agttccattc ggtggtgtta agcaatctgg ttacggtaga 1440
gaaatgggtg aagaagtcta ccatgcatac actgaagtaa aagctgtcag aattaagttg 1500
taa 1503

<210> 683
<211> 500
<212> PRT
<213> Artificial Sequence

<220>
<223> YPL061W (ALD6)(Chr 16)

<400> 683

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu
1 5 10 15

Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn
20 25 30

Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro
35 40 45

```
Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp
    50                  55                  60

Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu
65                  70                  75                  80

Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu
                85                  90                  95

Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala
            100                 105                 110

Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
        115                 120                 125

Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn
    130                 135                 140

Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr Met Asn Phe Thr Thr Leu
145                 150                 155                 160

Glu Pro Ile Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Ile
                165                 170                 175

Met Met Leu Ala Trp Lys Ile Ala Pro Ala Leu Ala Met Gly Asn Val
            180                 185                 190

Cys Ile Leu Lys Pro Ala Ala Val Thr Pro Leu Asn Ala Leu Tyr Phe
        195                 200                 205

Ala Ser Leu Cys Lys Lys Val Gly Ile Pro Ala Gly Val Val Asn Ile
    210                 215                 220

Val Pro Gly Pro Gly Arg Thr Val Gly Ala Ala Leu Thr Asn Asp Pro
225                 230                 235                 240

Arg Ile Arg Lys Leu Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Ser
                245                 250                 255

Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu
            260                 265                 270

Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys
        275                 280                 285

Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln
    290                 295                 300

Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp
305                 310                 315                 320

Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val
                325                 330                 335
```

```
Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
            340                 345                 350

Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu
        355                 360                 365

Gly Ala Lys Ile Leu Thr Gly Gly Glu Lys Val Gly Asp Lys Gly Tyr
    370                 375                 380

Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile
385                 390                 395                 400

Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys
                405                 410                 415

Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu
            420                 425                 430

Gly Ser Gly Ile Glu Thr Glu Ser Leu Ser Thr Gly Leu Lys Val Ala
        435                 440                 445

Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe
    450                 455                 460

Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
465                 470                 475                 480

Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
                485                 490                 495

Arg Ile Lys Leu
            500
```

<210> 684
<211> 1029
<212> DNA
<213> Artificial Sequence

<220>
<223> YPL088W (Chr 16)

<400> 684

```
atggttttag ttaagcaggt aagactcggt aactcaggtc ttaagatatc accgatagtg       60

ataggatgta tgtcatacgg gtccaagaaa tgggcggact gggtcataga ggacaagacc      120

caaattttca agattatgaa gcattgttac gataaaggtc ttcgtacttt tgacacagca      180

gatttttatt ctaatggttt gagtgaaaga ataattaagg agtttctgga gtactacagt      240

ataaagagag aaacggtggt gattatgacc aaaatttact tcccagttga tgaaacgctt      300

gatttgcatc ataacttcac tttaaatgaa tttgaagaat tggacttgtc caaccagcgg      360

ggtttatcca gaaagcatat aattgctggt gtcgagaact ctgtgaaaag actgggcaca      420

tatatagacc ttttacaaat tcacagatta gatcatgaaa cgccaatgaa agagatcatg      480

aaggcattga atgatgttgt tgaagcgggc cacgttagat acattggggc ttcgagtatg      540
```

ttggcaactg aatttgcaga actgcagttc acagccgata aatatggctg gtttcagttc 600

atttcttcgc agtcttacta caatttgctc tatcgtgaag atgaacgcga attgattcct 660

tttgccaaaa gacacaatat tggtttactt ccatggtctc ctaacgcacg aggcatgttg 720

actcgtcctc tgaaccaaag cacggacagg attaagagtg atccaacttt caagtcgtta 780

catttggata atctcgaaga agaacaaaag gaaattataa atcgtgtgga aaaggtgtcg 840

aaggacaaaa aagtctcgat ggctatgctc tccattgcat gggttttgca taaaggatgt 900

caccctattg tgggattgaa cactacagca agagtagacg aagcgattgc cgcactacaa 960

gtaactctaa cagaagaaga gataaagtac ctcgaggagc cctacaaacc ccagaggcaa 1020

agatgttaa 1029

<210> 685
<211> 342
<212> PRT
<213> Artificial Sequence

<220>
<223> YPL088W(Chr 16)

<400> 685

Met Val Leu Val Lys Gln Val Arg Leu Gly Asn Ser Gly Leu Lys Ile
1 5 10 15

Ser Pro Ile Val Ile Gly Cys Met Ser Tyr Gly Ser Lys Lys Trp Ala
20 25 30

Asp Trp Val Ile Glu Asp Lys Thr Gln Ile Phe Lys Ile Met Lys His
35 40 45

Cys Tyr Asp Lys Gly Leu Arg Thr Phe Asp Thr Ala Asp Phe Tyr Ser
50 55 60

Asn Gly Leu Ser Glu Arg Ile Ile Lys Glu Phe Leu Glu Tyr Tyr Ser
65 70 75 80

Ile Lys Arg Glu Thr Val Val Ile Met Thr Lys Ile Tyr Phe Pro Val
85 90 95

Asp Glu Thr Leu Asp Leu His His Asn Phe Thr Leu Asn Glu Phe Glu
100 105 110

Glu Leu Asp Leu Ser Asn Gln Arg Gly Leu Ser Arg Lys His Ile Ile
115 120 125

Ala Gly Val Glu Asn Ser Val Lys Arg Leu Gly Thr Tyr Ile Asp Leu
130 135 140

Leu Gln Ile His Arg Leu Asp His Glu Thr Pro Met Lys Glu Ile Met
145 150 155 160

Lys Ala Leu Asn Asp Val Val Glu Ala Gly His Val Arg Tyr Ile Gly
165 170 175

```
Ala Ser Ser Met Leu Ala Thr Glu Phe Ala Glu Leu Gln Phe Thr Ala
            180                 185                 190

Asp Lys Tyr Gly Trp Phe Gln Phe Ile Ser Ser Gln Ser Tyr Tyr Asn
        195                 200                 205

Leu Leu Tyr Arg Glu Asp Glu Arg Glu Leu Ile Pro Phe Ala Lys Arg
    210                 215                 220

His Asn Ile Gly Leu Leu Pro Trp Ser Pro Asn Ala Arg Gly Met Leu
225                 230                 235                 240

Thr Arg Pro Leu Asn Gln Ser Thr Asp Arg Ile Lys Ser Asp Pro Thr
                245                 250                 255

Phe Lys Ser Leu His Leu Asp Asn Leu Glu Glu Glu Gln Lys Glu Ile
            260                 265                 270

Ile Asn Arg Val Glu Lys Val Ser Lys Asp Lys Lys Val Ser Met Ala
        275                 280                 285

Met Leu Ser Ile Ala Trp Val Leu His Lys Gly Cys His Pro Ile Val
    290                 295                 300

Gly Leu Asn Thr Thr Ala Arg Val Asp Glu Ala Ile Ala Ala Leu Gln
305                 310                 315                 320

Val Thr Leu Thr Glu Glu Glu Ile Lys Tyr Leu Glu Glu Pro Tyr Lys
                325                 330                 335

Pro Gln Arg Gln Arg Cys
            340
```

<210> 686
<211> 1086
<212> DNA
<213> Artificial Sequence

<220>
<223> YCR105W (ADH7)(Chr 3)

<400> 686

```
atgctttacc cagaaaaatt tcagggcatc ggtatttcca acgcaaagga ttggaagcat       60

cctaaattag tgagttttga cccaaaaccc tttggcgatc atgacgttga tgttgaaatt      120

gaagcctgtg gtatctgcgg atctgatttt catatagccg ttggtaattg gggtccagtc      180

ccagaaaatc aaatccttgg acatgaaata attggccgcg tggtgaaggt tggatccaag      240

tgccacactg gggtaaaaat cggtgaccgt gttggtgttg gtgcccaagc cttggcgtgt      300

tttgagtgtg aacgttgcaa aagtgacaac gagcaatact gtaccaatga ccacgttttg      360

actatgtgga ctccttacaa ggacggctac atttcacaag gaggctttgc ctcccacgtg      420

aggcttcatg aacactttgc tattcaaata ccagaaaata ttccaagtcc gctagccgct      480

ccattattgt gtggtggtat tacagttttc tctccactac taagaaatgg ctgtggtcca      540
```

ggtaagaggg taggtattgt tggcatcggt ggtattgggc atatggggat tctgttggct 600

aaagctatgg gagccgaggt ttatgcgttt tcgcgaggcc actccaagcg ggaggattct 660

atgaaactcg gtgctgatca ctatattgct atgttggagg ataaaggctg gacagaacaa 720

tactctaacg ctttggacct tcttgtcgtt tgctcatcat ctttgtcgaa agttaatttt 780

gacagtatcg ttaagattat gaagattgga ggctccatcg tttcaattgc tgctcctgaa 840

gttaatgaaa agcttgtttt aaaaccgttg ggcctaatgg gagtatcaat ctcaagcagt 900

gctatcggat ctaggaagga aatcgaacaa ctattgaaat tagtttccga aaagaatgtc 960

aaaatatggg tggaaaaact tccgatcagc gaagaaggcg tcagccatgc ctttacaagg 1020

atggaaagcg gagacgtcaa atacagattt actttggtcg attatgataa gaaattccat 1080

aaatag 1086

<210> 687
<211> 361
<212> PRT
<213> Artificial Sequence

<220>
<223> YCR105W (ADH7)(Chr 3)

<400> 687

Met Leu Tyr Pro Glu Lys Phe Gln Gly Ile Gly Ile Ser Asn Ala Lys
1 5 10 15

Asp Trp Lys His Pro Lys Leu Val Ser Phe Asp Pro Lys Pro Phe Gly
20 25 30

Asp His Asp Val Asp Val Glu Ile Glu Ala Cys Gly Ile Cys Gly Ser
35 40 45

Asp Phe His Ile Ala Val Gly Asn Trp Gly Pro Val Pro Glu Asn Gln
50 55 60

Ile Leu Gly His Glu Ile Ile Gly Arg Val Val Lys Val Gly Ser Lys
65 70 75 80

Cys His Thr Gly Val Lys Ile Gly Asp Arg Val Gly Val Gly Ala Gln
85 90 95

Ala Leu Ala Cys Phe Glu Cys Glu Arg Cys Lys Ser Asp Asn Glu Gln
100 105 110

Tyr Cys Thr Asn Asp His Val Leu Thr Met Trp Thr Pro Tyr Lys Asp
115 120 125

Gly Tyr Ile Ser Gln Gly Gly Phe Ala Ser His Val Arg Leu His Glu
130 135 140

His Phe Ala Ile Gln Ile Pro Glu Asn Ile Pro Ser Pro Leu Ala Ala
145 150 155 160

```
Pro Leu Leu Cys Gly Gly Ile Thr Val Phe Ser Pro Leu Leu Arg Asn
                165                 170                 175

Gly Cys Gly Pro Gly Lys Arg Val Gly Ile Val Gly Ile Gly Gly Ile
            180                 185                 190

Gly His Met Gly Ile Leu Leu Ala Lys Ala Met Gly Ala Glu Val Tyr
        195                 200                 205

Ala Phe Ser Arg Gly His Ser Lys Arg Glu Asp Ser Met Lys Leu Gly
    210                 215                 220

Ala Asp His Tyr Ile Ala Met Leu Glu Asp Lys Gly Trp Thr Glu Gln
225                 230                 235                 240

Tyr Ser Asn Ala Leu Asp Leu Leu Val Val Cys Ser Ser Ser Leu Ser
                245                 250                 255

Lys Val Asn Phe Asp Ser Ile Val Lys Ile Met Lys Ile Gly Gly Ser
            260                 265                 270

Ile Val Ser Ile Ala Ala Pro Glu Val Asn Glu Lys Leu Val Leu Lys
        275                 280                 285

Pro Leu Gly Leu Met Gly Val Ser Ile Ser Ser Ser Ala Ile Gly Ser
    290                 295                 300

Arg Lys Glu Ile Glu Gln Leu Leu Lys Leu Val Ser Glu Lys Asn Val
305                 310                 315                 320

Lys Ile Trp Val Glu Lys Leu Pro Ile Ser Glu Glu Gly Val Ser His
                325                 330                 335

Ala Phe Thr Arg Met Glu Ser Gly Asp Val Lys Tyr Arg Phe Thr Leu
            340                 345                 350

Val Asp Tyr Asp Lys Lys Phe His Lys
        355                 360
```

<210> 688
<211> 1035
<212> DNA
<213> Artificial Sequence

<220>
<223> YDR541C (Chr 4)

<400> 688

```
atgtctaata cagttctagt ttctggcgct tcaggtttta ttgccttgca tatcctgtca       60

caattgttaa aacaagatta taaggttatt ggaactgtga gatcccatga aaaagaagca      120

aaattgctaa gacaatttca acataaccct aatttaactt tagaaattgt tccggacatt      180

tctcatccaa atgctttcga taaggttctg cagaaacgtg gacgtgagat taggtatgtt      240

ctacacacgg cctctccttt tcattatgat actaccgaat atgaaaaaga cttattgatt      300

cccgcgttag aaggtacaaa aaacatccta aattctatca agaaatatgc agcagacact      360
```

gtagagcgtg ttgttgtgac ttcttcttgt actgctatta taacccttgc aaagatggac 420

gatcccagtg tggtttttac agaagagagt tggaacgaag caacctggga aagctgtcaa 480

attgatggga taaatgctta ctttgcatcc aagaagtttg ctgaaaaggc tgcctgggag 540

ttcacaaaag agaatgaaga tcacatcaaa ttcaaactaa caacagtcaa cccttctctt 600

ctttttggtc ctcaactttt cgatgaagat gtgcatggcc atttgaatac ttcttgcgaa 660

atgatcaatg gcctaattca taccccagta aatgccagtg ttcctgattt tcattccatt 720

tttattgatg taagggatgt ggccctagct catctgtatg ctttccagaa ggaaaatacc 780

gcgggtaaaa gattagtggt aactaacggt aaatttggaa accaagatat cctggatatt 840

ttgaacgaag attttccaca attaagaggt ctcattcctt tgggtaagcc tggcacaggt 900

gatcaagtca ttgaccgcgg ttcaactaca gataatagtg caacgaggaa aatacttggc 960

tttgagttca gaagtttaca cgaaagtgtc catgatactg ctgcccaaat tttgaagaag 1020

cagaacagat tatga 1035

<210> 689
<211> 344
<212> PRT
<213> Artificial Sequence

<220>
<223> YDR541C (Chr 4)

<400> 689

Met Ser Asn Thr Val Leu Val Ser Gly Ala Ser Gly Phe Ile Ala Leu
1 5 10 15

His Ile Leu Ser Gln Leu Leu Lys Gln Asp Tyr Lys Val Ile Gly Thr
20 25 30

Val Arg Ser His Glu Lys Glu Ala Lys Leu Leu Arg Gln Phe Gln His
35 40 45

Asn Pro Asn Leu Thr Leu Glu Ile Val Pro Asp Ile Ser His Pro Asn
50 55 60

Ala Phe Asp Lys Val Leu Gln Lys Arg Gly Arg Glu Ile Arg Tyr Val
65 70 75 80

Leu His Thr Ala Ser Pro Phe His Tyr Asp Thr Thr Glu Tyr Glu Lys
85 90 95

Asp Leu Leu Ile Pro Ala Leu Glu Gly Thr Lys Asn Ile Leu Asn Ser
100 105 110

Ile Lys Lys Tyr Ala Ala Asp Thr Val Glu Arg Val Val Val Thr Ser
115 120 125

Ser Cys Thr Ala Ile Ile Thr Leu Ala Lys Met Asp Asp Pro Ser Val
130 135 140

```
Val Phe Thr Glu Glu Ser Trp Asn Glu Ala Thr Trp Glu Ser Cys Gln
145                 150                 155                 160

Ile Asp Gly Ile Asn Ala Tyr Phe Ala Ser Lys Lys Phe Ala Glu Lys
                165                 170                 175

Ala Ala Trp Glu Phe Thr Lys Glu Asn Glu Asp His Ile Lys Phe Lys
            180                 185                 190

Leu Thr Thr Val Asn Pro Ser Leu Leu Phe Gly Pro Gln Leu Phe Asp
        195                 200                 205

Glu Asp Val His Gly His Leu Asn Thr Ser Cys Glu Met Ile Asn Gly
    210                 215                 220

Leu Ile His Thr Pro Val Asn Ala Ser Val Pro Asp Phe His Ser Ile
225                 230                 235                 240

Phe Ile Asp Val Arg Asp Val Ala Leu Ala His Leu Tyr Ala Phe Gln
                245                 250                 255

Lys Glu Asn Thr Ala Gly Lys Arg Leu Val Val Thr Asn Gly Lys Phe
            260                 265                 270

Gly Asn Gln Asp Ile Leu Asp Ile Leu Asn Glu Asp Phe Pro Gln Leu
        275                 280                 285

Arg Gly Leu Ile Pro Leu Gly Lys Pro Gly Thr Gly Asp Gln Val Ile
    290                 295                 300

Asp Arg Gly Ser Thr Thr Asp Asn Ser Ala Thr Arg Lys Ile Leu Gly
305                 310                 315                 320

Phe Glu Phe Arg Ser Leu His Glu Ser Val His Asp Thr Ala Ala Gln
                325                 330                 335

Ile Leu Lys Lys Gln Asn Arg Leu
            340
```

<210> 690
<211> 1410
<212> DNA
<213> Artificial Sequence

<220>
<223> YER081 (SER3)(Chr 5)

<400> 690

```
atgacaagca ttgacattaa caacttacaa aataccttтc aacaagctat gaatatgagc       60

ggctccccag gcgctgtttg tacttcacct acgcaatctt tcatgaatac cgttccacag      120

cgcttgaatg ctgtaaagca cccaaaaatt ttgaagcctt tctcaacggg tgatatgaag      180

attttactat tagaaaacgt taatcaaact gctattacaa tcttcgaaga gcaaggttac      240

caagtcgaat tctataaatc ttcattgccc gaggaagagt tgatcgaaaa gatcaaggac      300

gttcatgcta ttggtatcag atcaaagact agattaactt caaatgtctt acaacatgcg      360
```

aagaatctgg tttgtattgg ttgtttctgt atcggtacca accaagttga cttagactac 420

gctaccagca gaggtattgc tgttttcaac tcgcctttct ccaactcaag atcagtagca 480

gaattggtca tcgctgaaat cattagttta gcaagacaac taggtgatag atctatcgaa 540

ttacataccg gtacatggaa taaggttgct gctagatgtt gggaggtaag aggaaaaact 600

cttggtatta ttgggtacgg tcacattggt tcccaattat cagttcttgc agaagctatg 660

ggtttgcatg tgttgtacta cgatattgta actatcatgg ccttgggtac tgccagacaa 720

gtttctacat tagatgaatt gttgaataaa tctgattttg tgacactaca tgtaccagct 780

actcctgaaa ctgaaaaaat gttatctgcc ccacaatttg ctgctatgaa ggatggcgct 840

tatgttatta atgcttcaag aggtactgtc gtggacattc catctttgat ccaagccgtg 900

aaagccaaca aaattgcagg tgctgctttg gatgtttatc cacatgaacc agctaagaac 960

ggtgaaggtt catttaacga tgagctaaat agctggactt ctgaattagt ttcattacca 1020

aatatcatct tgacaccaca cattggtggc tctaccgaag aagcccaaag ctcaatcggt 1080

attgaagtgg ctaccgcatt gtccaaatac atcaatgaag gtaactctgt cggttcagtc 1140

aacttcccag aagtggcatt gaaatcattg tcttacgacc aagagaacac tgtgcgtgtg 1200

ttatacattc accaaaatgt accaggtgtt ttgaagaccg tcaatgatat tttatcgaac 1260

cataacatcg aaaagcaatt ttccgattca aatggtgaaa ttgcttattt aatggctgat 1320

atctcttctg ttgaccaaag cgatattaaa gatatttatg aacaactaaa tcaaacctct 1380

gctaagatct caattagatt gctatattaa 1410

<210> 691
<211> 469
<212> PRT
<213> Artificial Sequence

<220>
<223> YER081 (SER3)(Chr 5)

<400> 691

Met Thr Ser Ile Asp Ile Asn Asn Leu Gln Asn Thr Phe Gln Gln Ala
1 5 10 15

Met Asn Met Ser Gly Ser Pro Gly Ala Val Cys Thr Ser Pro Thr Gln
20 25 30

Ser Phe Met Asn Thr Val Pro Gln Arg Leu Asn Ala Val Lys His Pro
35 40 45

Lys Ile Leu Lys Pro Phe Ser Thr Gly Asp Met Lys Ile Leu Leu Leu
50 55 60

Glu Asn Val Asn Gln Thr Ala Ile Thr Ile Phe Glu Glu Gln Gly Tyr
65 70 75 80

Gln Val Glu Phe Tyr Lys Ser Ser Leu Pro Glu Glu Glu Leu Ile Glu
85 90 95

```
Lys Ile Lys Asp Val His Ala Ile Gly Ile Arg Ser Lys Thr Arg Leu
            100                 105                 110

Thr Ser Asn Val Leu Gln His Ala Lys Asn Leu Val Cys Ile Gly Cys
        115                 120                 125

Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asp Tyr Ala Thr Ser Arg
    130                 135                 140

Gly Ile Ala Val Phe Asn Ser Pro Phe Ser Asn Ser Arg Ser Val Ala
145                 150                 155                 160

Glu Leu Val Ile Ala Glu Ile Ile Ser Leu Ala Arg Gln Leu Gly Asp
                165                 170                 175

Arg Ser Ile Glu Leu His Thr Gly Thr Trp Asn Lys Val Ala Ala Arg
            180                 185                 190

Cys Trp Glu Val Arg Gly Lys Thr Leu Gly Ile Ile Gly Tyr Gly His
        195                 200                 205

Ile Gly Ser Gln Leu Ser Val Leu Ala Glu Ala Met Gly Leu His Val
    210                 215                 220

Leu Tyr Tyr Asp Ile Val Thr Ile Met Ala Leu Gly Thr Ala Arg Gln
225                 230                 235                 240

Val Ser Thr Leu Asp Glu Leu Leu Asn Lys Ser Asp Phe Val Thr Leu
                245                 250                 255

His Val Pro Ala Thr Pro Glu Thr Glu Lys Met Leu Ser Ala Pro Gln
            260                 265                 270

Phe Ala Ala Met Lys Asp Gly Ala Tyr Val Ile Asn Ala Ser Arg Gly
        275                 280                 285

Thr Val Val Asp Ile Pro Ser Leu Ile Gln Ala Val Lys Ala Asn Lys
    290                 295                 300

Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn
305                 310                 315                 320

Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu
                325                 330                 335

Val Ser Leu Pro Asn Ile Ile Leu Thr Pro His Ile Gly Gly Ser Thr
            340                 345                 350

Glu Glu Ala Gln Ser Ser Ile Gly Ile Glu Val Ala Thr Ala Leu Ser
        355                 360                 365

Lys Tyr Ile Asn Glu Gly Asn Ser Val Gly Ser Val Asn Phe Pro Glu
    370                 375                 380
```

```
Val Ala Leu Lys Ser Leu Ser Tyr Asp Gln Glu Asn Thr Val Arg Val
385                 390                 395                 400

Leu Tyr Ile His Gln Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp
                405                 410                 415

Ile Leu Ser Asn His Asn Ile Glu Lys Gln Phe Ser Asp Ser Asn Gly
            420                 425                 430

Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asp Gln Ser Asp
        435                 440                 445

Ile Lys Asp Ile Tyr Glu Gln Leu Asn Gln Thr Ser Ala Lys Ile Ser
    450                 455                 460

Ile Arg Leu Leu Tyr
465
```

<210> 692
<211> 711
<212> DNA
<213> Artificial Sequence

<220>
<223> YPL275W (FDH2)(Chr 16)

<400> 692

```
atggtggtca tcaataagca attaatggtg agtgggatat tgccggcgtg gctaaaaaat       60

gagtatgatc tggaagacaa aataatttca acggtaggtg ccggtagaat tggatatagg      120

gttctggaaa gattggtcgc atttaatccg aagaagttac tgtactacga ctaccaggaa      180

ctacctgcgg aagcaatcaa tagattgaac gaggccagca agcttttcaa tggcagaggt      240

gatattgttc agagagtaga gaaattggag gatatggttg ctcagtcaga tgttgttacc      300

atcaactgtc cattgcacaa ggactcaagg ggtttattca ataaaaagct tatttcccac      360

atgaaagatg gtgcatactt ggtgaatacc gctagaggtg ctatttgtgt cgcagaagat      420

gttgccgagg cagtcaagtc tggtaaattg gctggctatg gtggtgatgt ctgggataag      480

caaccagcac caaaagacca tccctggagg actatggaca ataaggacca cgtgggaaac      540

gcaatgactg ttcatatcag tggcacatct ctgcatgctc aaaagaggta cgctcaggga      600

gtaaagaaca tcctaaatag ttacttttcc aaaaagtttg attaccgtcc acaggatatt      660

attgtgcaga atggttctta tgccaccaga gcttatggac agaagaaata a               711
```

<210> 693
<211> 236
<212> PRT
<213> Artificial Sequence

<220>
<223> YPL275W (FDH2)(Chr 16)

<400> 693

```
Met Val Val Ile Asn Lys Gln Leu Met Val Ser Gly Ile Leu Pro Ala
1               5                   10                  15
```

```
Trp Leu Lys Asn Glu Tyr Asp Leu Glu Asp Lys Ile Ile Ser Thr Val
            20                  25                  30

Gly Ala Gly Arg Ile Gly Tyr Arg Val Leu Glu Arg Leu Val Ala Phe
        35                  40                  45

Asn Pro Lys Lys Leu Leu Tyr Tyr Asp Tyr Gln Glu Leu Pro Ala Glu
    50                  55                  60

Ala Ile Asn Arg Leu Asn Glu Ala Ser Lys Leu Phe Asn Gly Arg Gly
65                  70                  75                  80

Asp Ile Val Gln Arg Val Glu Lys Leu Glu Asp Met Val Ala Gln Ser
                85                  90                  95

Asp Val Val Thr Ile Asn Cys Pro Leu His Lys Asp Ser Arg Gly Leu
            100                 105                 110

Phe Asn Lys Lys Leu Ile Ser His Met Lys Asp Gly Ala Tyr Leu Val
        115                 120                 125

Asn Thr Ala Arg Gly Ala Ile Cys Val Ala Glu Asp Val Ala Glu Ala
    130                 135                 140

Val Lys Ser Gly Lys Leu Ala Gly Tyr Gly Gly Asp Val Trp Asp Lys
145                 150                 155                 160

Gln Pro Ala Pro Lys Asp His Pro Trp Arg Thr Met Asp Asn Lys Asp
                165                 170                 175

His Val Gly Asn Ala Met Thr Val His Ile Ser Gly Thr Ser Leu His
            180                 185                 190

Ala Gln Lys Arg Tyr Ala Gln Gly Val Lys Asn Ile Leu Asn Ser Tyr
        195                 200                 205

Phe Ser Lys Lys Phe Asp Tyr Arg Pro Gln Asp Ile Ile Val Gln Asn
    210                 215                 220

Gly Ser Tyr Ala Thr Arg Ala Tyr Gly Gln Lys Lys
225                 230                 235
```

<210> 694
<211> 1494
<212> DNA
<213> Artificial Sequence

<220>
<223> YBR006W (UGA5)(Chr2)

<400> 694

```
atgactttga gtaagtattc taaaccaact ctaaacgacc ctaatttatt cagagaatct       60

ggttatattg acggaaaatg ggttaagggc actgacgaag tttttgaggt ggtagaccct      120

gcttccggcg aaatcatagc aagagttccc gaacaaccag tctccgtggt tgaggaagcg      180

attgatgttg cctatgaaac tttcaagacg tacaagaata caacaccaag agagagggca      240
```

```
aagtggctca gaaacatgta taacttaatg cttgaaaatt tggatgatct ggcaaccatc      300
attactttag aaaatggtaa agctctaggg gaagctaaag gagaaatcaa atacgcggct      360
tcgtattttg agtggtacgc cgaggaagca ccccgtttat atggtgctac tattcaaccc      420
ttgaaccctc acaacagagt attcacaatt aggcaacctg ttggtgtatg cggtataatt      480
tgtccatgga attttccgag cgccatgatc acgagaaagg ccgccgctgc tttagctgtg      540
ggctgcacag tagtcatcaa gccagactct caaacgccgc tatctgcttt agcaatggca      600
tatttggctg aaaaggcagg ctttcccaag ggttcgttta atgttattct ttcacatgcc      660
aacacaccaa agcttggtaa aacattatgt gaatcaccaa aagtcaagaa agttactttt      720
actggttcta caaacgtcgg taaaatcttg atgaaacaat cttcttctac tttgaagaaa      780
ctgtcttttg agctgggtgg taacgcccct ttcatagtct ttgaggatgc cgatttggat      840
caagccttgg aacaagccat ggcttgtaaa tttaggggtt tgggtcaaac atgtgtgtgc      900
gcaaatagac tttacgttca ctcatccata attgataaat ttgcgaaatt actcgcggag      960
agggtcaaaa aattcgtaat tggccatggt ttggacccaa aaactacaca tggttgtgtc     1020
attaactcca gcgctattga aaaagttgaa agacataaac aggatgccat tgataaggga     1080
gcaaaagttg tgcttgaagg tggacgttta actgagttag gtcctaactt ttatgctcca     1140
gtaattttgt cacacgttcc ctcaacagct attgtttcca aggaggagac ttttggtcca     1200
ttatgtccaa tcttttcttt tgatactatg gaagaagttg tcggatatgc taatgatact     1260
gagtttggtt tagcagcata tgtcttttct aaaaatgtca acactttata cactgtgtct     1320
gaagctttgg aaactggtat ggtttcatgt aatacaggtg ttttttcgga ttgttctata     1380
ccatttggtg gtgttaaaga gtcaggattt ggaagagaag gttcgctata tggtattgaa     1440
gattacactg ttttgaagac catcacaatt gggaatttgc caaacagcat ttaa           1494
```

<210> 695
<211> 497
<212> PRT
<213> Artificial Sequence

<220>
<223> YBR006W (UGA5)(Chr2)

<400> 695

```
Met Thr Leu Ser Lys Tyr Ser Lys Pro Thr Leu Asn Asp Pro Asn Leu
1               5                  10                  15

Phe Arg Glu Ser Gly Tyr Ile Asp Gly Lys Trp Val Lys Gly Thr Asp
            20                  25                  30

Glu Val Phe Glu Val Val Asp Pro Ala Ser Gly Glu Ile Ile Ala Arg
        35                  40                  45

Val Pro Glu Gln Pro Val Ser Val Val Glu Glu Ala Ile Asp Val Ala
    50                  55                  60

Tyr Glu Thr Phe Lys Thr Tyr Lys Asn Thr Thr Pro Arg Glu Arg Ala
65                  70                  75                  80
```

```
Lys Trp Leu Arg Asn Met Tyr Asn Leu Met Leu Glu Asn Leu Asp Asp
                85                  90                  95

Leu Ala Thr Ile Ile Thr Leu Glu Asn Gly Lys Ala Leu Gly Glu Ala
            100                 105                 110

Lys Gly Glu Ile Lys Tyr Ala Ala Ser Tyr Phe Glu Trp Tyr Ala Glu
        115                 120                 125

Glu Ala Pro Arg Leu Tyr Gly Ala Thr Ile Gln Pro Leu Asn Pro His
    130                 135                 140

Asn Arg Val Phe Thr Ile Arg Gln Pro Val Gly Val Cys Gly Ile Ile
145                 150                 155                 160

Cys Pro Trp Asn Phe Pro Ser Ala Met Ile Thr Arg Lys Ala Ala Ala
                165                 170                 175

Ala Leu Ala Val Gly Cys Thr Val Val Ile Lys Pro Asp Ser Gln Thr
            180                 185                 190

Pro Leu Ser Ala Leu Ala Met Ala Tyr Leu Ala Glu Lys Ala Gly Phe
        195                 200                 205

Pro Lys Gly Ser Phe Asn Val Ile Leu Ser His Ala Asn Thr Pro Lys
    210                 215                 220

Leu Gly Lys Thr Leu Cys Glu Ser Pro Lys Val Lys Lys Val Thr Phe
225                 230                 235                 240

Thr Gly Ser Thr Asn Val Gly Lys Ile Leu Met Lys Gln Ser Ser Ser
                245                 250                 255

Thr Leu Lys Lys Leu Ser Phe Glu Leu Gly Gly Asn Ala Pro Phe Ile
            260                 265                 270

Val Phe Glu Asp Ala Asp Leu Asp Gln Ala Leu Glu Gln Ala Met Ala
        275                 280                 285

Cys Lys Phe Arg Gly Leu Gly Gln Thr Cys Val Cys Ala Asn Arg Leu
    290                 295                 300

Tyr Val His Ser Ser Ile Ile Asp Lys Phe Ala Lys Leu Leu Ala Glu
305                 310                 315                 320

Arg Val Lys Lys Phe Val Ile Gly His Gly Leu Asp Pro Lys Thr Thr
                325                 330                 335

His Gly Cys Val Ile Asn Ser Ser Ala Ile Glu Lys Val Glu Arg His
            340                 345                 350

Lys Gln Asp Ala Ile Asp Lys Gly Ala Lys Val Val Leu Glu Gly Gly
        355                 360                 365
```

Arg Leu Thr Glu Leu Gly Pro Asn Phe Tyr Ala Pro Val Ile Leu Ser
370 375 380

His Val Pro Ser Thr Ala Ile Val Ser Lys Glu Glu Thr Phe Gly Pro
385 390 395 400

Leu Cys Pro Ile Phe Ser Phe Asp Thr Met Glu Glu Val Val Gly Tyr
405 410 415

Ala Asn Asp Thr Glu Phe Gly Leu Ala Ala Tyr Val Phe Ser Lys Asn
420 425 430

Val Asn Thr Leu Tyr Thr Val Ser Glu Ala Leu Glu Thr Gly Met Val
435 440 445

Ser Cys Asn Thr Gly Val Phe Ser Asp Cys Ser Ile Pro Phe Gly Gly
450 455 460

Val Lys Glu Ser Gly Phe Gly Arg Glu Gly Ser Leu Tyr Gly Ile Glu
465 470 475 480

Asp Tyr Thr Val Leu Lys Thr Ile Thr Ile Gly Asn Leu Pro Asn Ser
485 490 495

Ile

<210> 696
<211> 1323
<212> DNA
<213> Artificial Sequence

<220>
<223> YOL059W (Chr 15)

<400> 696
atgcttgctg tcagaagatt aacaagatac acattcctta agcgaacgca tccggtgtta 60

tatactcgtc gtgcatataa aattttgcct tcaagatcta ctttcctaag aagatcatta 120

ttacaaacac aactgcactc aaagatgact gctcatacta atatcaaaca gcacaaacac 180

tgtcatgagg accatcctat cagaagatcg gactctgccg tgtcaattgt acatttgaaa 240

cgtgcgccct tcaaggttac agtgattggt tctggtaact gggggaccac catcgccaaa 300

gtcattgcgg aaaacacaga attgcattcc catatcttcg agccagaggt gagaatgtgg 360

gtttttgatg aaaagatcgg cgacgaaaat ctgacggata tcataaatac aagacaccag 420

aacgttaaat atctacccaa tattgacctg ccccataatc tagtggccga tcctgatctt 480

ttacactcca tcaagggtgc tgacatcctt gttttcaaca tccctcatca atttttacca 540

aacatagtca aacaattgca aggccacgtg gcccctcatg taagggccat ctcgtgtcta 600

aaagggttcg agttgggctc caagggtgtg caattgctat cctcctatgt tactgatgag 660

ttaggaatcc aatgtggcgc actatctggt gcaaacttgg caccggaagt ggccaaggag 720

cattggtccg aaaccaccgt ggcttaccaa ctaccaaagg attatcaagg tgatggcaag 780

gatgtagatc ataagatttt gaaattgctg ttccacagac cttacttcca cgtcaatgtc 840

atcgatgatg ttgctggtat atccattgcc ggtgccttga agaacgtcgt ggcacttgca 900

tgtggtttcg tagaaggtat gggatggggt aacaatgcct ccgcagccat tcaaaggctg 960

ggtttaggtg aaattatcaa gttcggtaga atgtttttcc cagaatccaa agtcgagacc 1020

tactatcaag aatccgctgg tgttgcagat ctgatcacca cctgctcagg cggtagaaac 1080

gtcaaggttg ccacatacat ggccaagacc ggtaagtcag ccttggaagc agaaaaggaa 1140

ttgcttaacg gtcaatccgc ccaagggata atcacatgca gagaagttca cgagtggcta 1200

caaacatgtg agttgaccca agaattccca ttattcgagg cagtctacca gatagtctac 1260

aacaacgtcc gcatggaaga cctaccggag atgattgaag agctagacat cgatgacgaa 1320

tag 1323

<210> 697
<211> 440
<212> PRT
<213> Artificial Sequence

<220>
<223> YOL059W (Chr 15)

<400> 697

Met Leu Ala Val Arg Arg Leu Thr Arg Tyr Thr Phe Leu Lys Arg Thr
1 5 10 15

His Pro Val Leu Tyr Thr Arg Arg Ala Tyr Lys Ile Leu Pro Ser Arg
20 25 30

Ser Thr Phe Leu Arg Arg Ser Leu Leu Gln Thr Gln Leu His Ser Lys
35 40 45

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp
50 55 60

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys
65 70 75 80

Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr
85 90 95

Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile
100 105 110

Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp
115 120 125

Glu Asn Leu Thr Asp Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr
130 135 140

Leu Pro Asn Ile Asp Leu Pro His Asn Leu Val Ala Asp Pro Asp Leu
145 150 155 160

```
Leu His Ser Ile Lys Gly Ala Asp Ile Leu Val Phe Asn Ile Pro His
                165                 170                 175

Gln Phe Leu Pro Asn Ile Val Lys Gln Leu Gln Gly His Val Ala Pro
            180                 185                 190

His Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Leu Gly Ser Lys
        195                 200                 205

Gly Val Gln Leu Leu Ser Ser Tyr Val Thr Asp Glu Leu Gly Ile Gln
    210                 215                 220

Cys Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Lys Glu
225                 230                 235                 240

His Trp Ser Glu Thr Thr Val Ala Tyr Gln Leu Pro Lys Asp Tyr Gln
                245                 250                 255

Gly Asp Gly Lys Asp Val Asp His Lys Ile Leu Lys Leu Leu Phe His
            260                 265                 270

Arg Pro Tyr Phe His Val Asn Val Ile Asp Asp Val Ala Gly Ile Ser
        275                 280                 285

Ile Ala Gly Ala Leu Lys Asn Val Val Ala Leu Ala Cys Gly Phe Val
    290                 295                 300

Glu Gly Met Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Leu
305                 310                 315                 320

Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser
                325                 330                 335

Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile
            340                 345                 350

Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
        355                 360                 365

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly
    370                 375                 380

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
385                 390                 395                 400

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Leu Phe Glu Ala Val Tyr
                405                 410                 415

Gln Ile Val Tyr Asn Asn Val Arg Met Glu Asp Leu Pro Glu Met Ile
            420                 425                 430

Glu Glu Leu Asp Ile Asp Asp Glu
        435                 440
```

20 Jan 2017 2017200439

<210> 698
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces paradoxus AABY01000127

<400> 698
atgtcccaag gtagaaaagc tgcagaaaga ttggctaaca agaccgtgct cattacgggt 60

gcctctgctg gtattggtaa ggccaccgca ttagagtatt tggaggcatc caatggtgat 120

atgaaactgg tcttagctgc tagaagatta gaaaagctcg aggaattaaa gaaaactatt 180

gatcaggagt ttccaaacgc caaagttcat gtggcccaac tggatatcac tcaagcagaa 240

aagatcaagc cctttattga gaatttgcca aaggagttca aagacattga cattttggtg 300

aacaacgctg ggaaggccct tggtaccgac cgtgtggggg agattgcaac acaagatatc 360

caggatgtgt ttgacaccaa cgtcacagct ttaattaata tcactcaagc tgtgctgccc 420

atttttcaag ccaagaactc aggggatatt gtgaacttgg gttcggtggc tggcagggat 480

gcatacccaa cgggttccat ctattgtgcc tccaagtttg ccgtgggggc gttcactgat 540

agtttaagaa aggagcttat caacaccaag atcagagtca tcctaatcgc accagggcta 600

gtcgaaactg aattttcact ggttagatac agaggcaacg aggagcaagc caagaatgtc 660

tacaaggaca ccaccccatt aatggctgat gacgtggctg atttgatcgt gtacgcaact 720

tccaggaaac aaaacactgt aattgcagac acgctaatct ttccaaccaa ccaagcatcg 780

cctcaccaca tcttccgtgg atga 804

<210> 699
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces paradoxus AABY01000127

<400> 699

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn Lys Thr Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Val Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Gln Glu Phe
50 55 60

Pro Asn Ala Lys Val His Val Ala Gln Leu Asp Ile Thr Gln Ala Glu
65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Lys Glu Phe Lys Asp Ile
85 90 95

```
Asp Ile Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Thr Asp Arg Val
            100                 105                 110

Gly Glu Ile Ala Thr Gln Asp Ile Gln Asp Val Phe Asp Thr Asn Val
        115                 120                 125

Thr Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly
                165                 170                 175

Ala Phe Thr Asp Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg
            180                 185                 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr
    210                 215                 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
                245                 250                 255

Asn Gln Ala Ser Pro His His Ile Phe Arg Gly
            260                 265
```

<210> 700
<211> 805
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces bayanus AACA01000631

<400> 700

```
atgtcccaag gtagaaaagc tgcagaaaga ttggccaaca agacggtgct cattacaggc        60

gcttctgctg gtattggtaa ggccaccgca ttggagtatt tggaagcatc caatggaaac       120

atgaaactga tcttggctgc gaggagattg gagaagctag aggagctgaa gaagaccatc       180

gacgaggagt ttcccaatgc aaaggttcac gttggccaac tggatatcac acaggccgag       240

aagatcaagc ccttcattga aaacttgccg gaggcattca aggatattga catcctgata       300

aacaatgccg gcaaagccct gggctccgaa cgtgtcgggg aaattgccac acaggacatc       360

caggacgtgt tcgacaccaa cgtcacggcg ttgatcaacg tcacgcaagc agtgctgcca       420

attttccaag ccaagaactc aggggacatc gtcaacttgg ggctcggtgg ccggcagaga       480

cgcatacccc acaggctcca tctactgtgc ttccaagttt gccgtcggtg cgttcactga       540

cagtttgaga aaggaactga tcaacacgaa gatcagagtt atcttgatcg cgccggggct       600
```

ggttgagacc gagttctcac tggtcagata cagaggtaat gaggaacaag ctaaaaacgt 660

ctacaaggac actacgccgt tgatggccga cgacgtggct gacttaatcg tatattccac 720

ttccagaaag cagaacaccg tggttgccga caccctgatc ttccccacca accaagcctc 780

gccctaccac atctttcgcg gttaa 805

<210> 701
<211> 262
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces bayanus AACA01000631

<400> 701

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn Lys Thr Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asn Met Lys Leu Ile Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Glu Glu Phe
50 55 60

Pro Asn Ala Lys Val His Val Gly Gln Leu Asp Ile Thr Gln Ala Glu
65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Glu Ala Phe Lys Asp Ile
85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Glu Arg Val
100 105 110

Gly Glu Ile Ala Thr Gln Asp Ile Gln Asp Val Phe Asp Thr Asn Val
115 120 125

Thr Ala Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
130 135 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Leu Gly Gly Arg Gln Arg
145 150 155 160

Arg Ile Pro His Arg Leu His Leu Leu Cys Phe Gln Val Cys Arg Arg
165 170 175

Cys Val His Gln Phe Glu Lys Gly Thr Asp Gln His Glu Asp Gln Ser
180 185 190

Tyr Leu Asp Arg Ala Gly Ala Gly Asp Arg Val Leu Thr Gly Gln Ile
195 200 205

Gln Arg Gly Thr Ser Lys Arg Leu Gln Gly His Tyr Ala Val Asp Gly
210 215 220

Arg Arg Arg Gly Leu Asn Arg Ile Phe His Phe Gln Lys Ala Glu His
225 230 235 240

Arg Gly Cys Arg His Pro Asp Leu Pro His Gln Pro Ser Leu Ala Leu
245 250 255

Pro His Leu Ser Arg Leu
260

<210> 702
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces castellii AACF01000116

<400> 702
atgtctcaag gtcctaaagc tgccgaaaga ttgaatgaga agattgtgtt tatcactggt 60

gcttcagctg gtattgggca agccaccgct ttggaataca tggatgcgtc gaacggtact 120

gtgaaattgg ttctagttgc cagaagattg gagaaattac aacaattgaa ggaagtcatt 180

gaggcaaaat accctaagag taaagtctat attgggaagt tggatgtgac agagcttgag 240

accattcaac cattcttgga taatcttcct gaggaattta aggatattga tatcttgatt 300

aataatgccg ggaaggcatt aggttccgat cgtgtaggtg atattgatat aaaagatgtg 360

aagggaatga tggataccaa tgtcttgggg ttgatcaatg tgacgcaagc tgtgttgcac 420

attttccaaa agaagaactc cggtgatatt gtgaacttag gttcagttgc tggaagagat 480

gcatacccaa cagggtccat ttactgtgct tctaaatttg ccgtgagggc ctttactgaa 540

agtttgagaa gggaattaat taataccaag attagggtga tattgatagc cccgggtatc 600

gtcgaaactg aattctcagt tgttagatac aagggtgata atgagcgtgc taaatctgtc 660

tacgatggag ttcacccctt ggaagcagac gacgtagcag atttaattgt atacaccact 720

tcaagaaaac agaacacagt aattgctgac actttgatat tcccaacctc tcaaggttcc 780

gcattccacg tccatcgcga ttaa 804

<210> 703
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces castellii AACF01000116

<400> 703

Met Ser Gln Gly Pro Lys Ala Ala Glu Arg Leu Asn Glu Lys Ile Val
1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Leu Glu
20 25 30

```
Tyr Met Asp Ala Ser Asn Gly Thr Val Lys Leu Val Leu Val Ala Arg
        35                  40                  45

Arg Leu Glu Lys Leu Gln Gln Leu Lys Glu Val Ile Glu Ala Lys Tyr
    50                  55                  60

Pro Lys Ser Lys Val Tyr Ile Gly Lys Leu Asp Val Thr Glu Leu Glu
65                  70                  75                  80

Thr Ile Gln Pro Phe Leu Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile
                85                  90                  95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Arg Val
            100                 105                 110

Gly Asp Ile Asp Ile Lys Asp Val Lys Gly Met Met Asp Thr Asn Val
        115                 120                 125

Leu Gly Leu Ile Asn Val Thr Gln Ala Val Leu His Ile Phe Gln Lys
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Arg
                165                 170                 175

Ala Phe Thr Glu Ser Leu Arg Arg Glu Leu Ile Asn Thr Lys Ile Arg
            180                 185                 190

Val Ile Leu Ile Ala Pro Gly Ile Val Glu Thr Glu Phe Ser Val Val
        195                 200                 205

Arg Tyr Lys Gly Asp Asn Glu Arg Ala Lys Ser Val Tyr Asp Gly Val
    210                 215                 220

His Pro Leu Glu Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Thr Thr
225                 230                 235                 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
                245                 250                 255

Ser Gln Gly Ser Ala Phe His Val His Arg Asp
            260                 265
```

<210> 704
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces mikatae AACH01000019

<400> 704
atgtctcaag gtagaaaagc tgcagaaaga ttggctggca aaaccgttct catcacgggt 60

gcctctgctg gtattggcaa agccactgca ttagagtatt tggaggcatc caatggcgat 120

atgaaattaa tcttagccgc tagaagatta gaaaagctcg aggaattgaa gaagactatc 180

gatgaagagt ttccaaacgc aaaggtccat gtgaccaaac tggacatcac acagacagaa 240

aagatcaagc cctttattga aaacttgcca gaggagttca aagacattga tattctggtg 300

aacaacgctg gtaaggctct tggtacggac cgtgttgggg agattgatac acaggacgtc 360

caggacgtgt tcgacaccaa cgtctcggct ttgattaatg tcacacaggc tgttctgccc 420

atcttccaag ctaagaactc aggggatatt gtgaacttgg gctcggtagc tggcagagat 480

gcatacccaa cgggctccat ctattgtgca tctaagtttg ccgtcggggc tttcactgag 540

agtttgagaa tggaacttat aaacactaag attagagtca ttctaattgc accagggtta 600

gtcgaaactg agttttccct ggttagatac agaggtaacg aagaacaagc caagaatgtt 660

tacaaggaca ccactccgtt gatggccgat gacgtggctg atttgattgt gtatgcgact 720

tcaaggaagc agaacactgt aattgcagac acactaatct ttcctaccaa ccaagcgtca 780

ccttaccata tctttcgcgg gtga 804

<210> 705
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces mikatae AACH01000019

<400> 705

Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Ala Gly Lys Thr Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Glu Ala Ser Asn Gly Asp Met Lys Leu Ile Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu Glu Glu Leu Lys Lys Thr Ile Asp Glu Glu Phe
50 55 60

Pro Asn Ala Lys Val His Val Thr Lys Leu Asp Ile Thr Gln Thr Glu
65 70 75 80

Lys Ile Lys Pro Phe Ile Glu Asn Leu Pro Glu Glu Phe Lys Asp Ile
85 90 95

Asp Ile Leu Val Asn Asn Ala Gly Lys Ala Leu Gly Thr Asp Arg Val
100 105 110

Gly Glu Ile Asp Thr Gln Asp Val Gln Asp Val Phe Asp Thr Asn Val
115 120 125

Ser Ala Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Gln Ala
130 135 140

```
Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys Phe Ala Val Gly
                165                 170                 175

Ala Phe Thr Glu Ser Leu Arg Met Glu Leu Ile Asn Thr Lys Ile Arg
            180                 185                 190

Val Ile Leu Ile Ala Pro Gly Leu Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Arg Gly Asn Glu Glu Gln Ala Lys Asn Val Tyr Lys Asp Thr
    210                 215                 220

Thr Pro Leu Met Ala Asp Asp Val Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Thr
                245                 250                 255

Asn Gln Ala Ser Pro Tyr His Ile Phe Arg Gly
            260                 265
```

<210> 706
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Ashbya gossypii AE016819

<400> 706

```
atgtccctag gaagaaaagc agctgaaaga ttagccaaca aaattgtgct tgtgactggt       60

gcctctgcgg gcattggccg tgctacagcc attaactatg cagacgcgac ggacggggca      120

atcaagttga ttttggtggc aagacgcgca gaaaagctca ccagcttgaa acaggagatc      180

gaaagcaagt atcccaacgc caagatccat gtcggacaat tggatgtgac ccaactggac      240

cagatccgcc cattttgga gggactacct gaggagttcc gagacattga tattttaatt      300

aacaacgcag gtaaggccct cggcactgag agggtggggg aaatctcgat ggacgatatc      360

caggaggttt tcaacactaa tgttatcggc ttggtgcact tgactcagga ggttctacct      420

attatgaaag ccaagaattc cggggacatt gtcaatgttg ggtcgattgc cggccgcgaa      480

gcctaccctg gtggctctat ttactgtgcc acgaaacatg cggtcaaggc tttcaccagg      540

gccatgcgga aggagctcat tagcaccaag atccgggtct tcgaaattgc gccgggctct      600

gtagaaacgg aattctccat ggttcgtatg cgcggtaacg aagagaatgc caagaaagtg      660

taccagggat ttgaacccct agatggtgat gatatcgctg atacaattgt ctatgccaca      720

tccagaagat ccaacaccgt agttgcagag atggtcgttt acccatccgc gcaaggttct      780

ctgtacgata ctcaccgcaa ctaa                                             804
```

<210> 707
<211> 267

2017200439 20 Jan 2017

<212> PRT
<213> Artificial Sequence

<220>
<223> Ashbya gossypii AE016819

<400> 707

```
Met Ser Leu Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn Lys Ile Val
1               5                  10                  15

Leu Val Thr Gly Ala Ser Ala Gly Ile Gly Arg Ala Thr Ala Ile Asn
            20                  25                  30

Tyr Ala Asp Ala Thr Asp Gly Ala Ile Lys Leu Ile Leu Val Ala Arg
        35                  40                  45

Arg Ala Glu Lys Leu Thr Ser Leu Lys Gln Glu Ile Glu Ser Lys Tyr
    50                  55                  60

Pro Asn Ala Lys Ile His Val Gly Gln Leu Asp Val Thr Gln Leu Asp
65                  70                  75                  80

Gln Ile Arg Pro Phe Leu Glu Gly Leu Pro Glu Glu Phe Arg Asp Ile
                85                  90                  95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Thr Glu Arg Val
            100                 105                 110

Gly Glu Ile Ser Met Asp Asp Ile Gln Glu Val Phe Asn Thr Asn Val
        115                 120                 125

Ile Gly Leu Val His Leu Thr Gln Glu Val Leu Pro Ile Met Lys Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Val Gly Ser Ile Ala Gly Arg Glu
145                 150                 155                 160

Ala Tyr Pro Gly Gly Ser Ile Tyr Cys Ala Thr Lys His Ala Val Lys
                165                 170                 175

Ala Phe Thr Arg Ala Met Arg Lys Glu Leu Ile Ser Thr Lys Ile Arg
            180                 185                 190

Val Phe Glu Ile Ala Pro Gly Ser Val Glu Thr Glu Phe Ser Met Val
        195                 200                 205

Arg Met Arg Gly Asn Glu Glu Asn Ala Lys Lys Val Tyr Gln Gly Phe
    210                 215                 220

Glu Pro Leu Asp Gly Asp Asp Ile Ala Asp Thr Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Arg Ser Asn Thr Val Val Ala Glu Met Val Val Tyr Pro Ser
                245                 250                 255
```

```
Ala Gln Gly Ser Leu Tyr Asp Thr His Arg Asn
            260                 265
```

<210> 708
<211> 807
<212> DNA
<213> Artificial Sequence

<220>
<223> Candida glabrata CR380959

<400> 708

```
atgtctcaag gaagaaaagc tgctgagagg ttacaaggga agattgcctt tattacgggt       60

gcctctgcgg gcatcggtaa agctacagcc attgagtatt tggatgcttc caatggtagt      120

gtgaagctag ttcttggtgc acgtagaatg gagaaattgg aggagttgaa gaaggaattg      180

ctggctcaat atcctgatgc aaagattcat ataggtaaac tggatgttac agactttgaa      240

aacgtcaagc agtttttggc tgacttgcca gaagagttca aggacatcga catcctgatc      300

aataacgctg gtaaagcgtt ggggtctgac aaagttggag acattgaccc tgaggatatc      360

gcaggaatgg ttaacaccaa cgtccttgca ttgatcaatt taacacaatt gttgttgcca      420

ttattcaaga agaagaacag tggtgatatc gtcaacttgg gatcgattgc tggtagagac      480

gcatacccaa cgggtgctat atactgtgca acaaaacatg ctgtcagggc attcacacaa      540

tccttaagga aggaattgat caacaccgac attagagtaa ttgaaattgc tcctggtatg      600

gtcgaaaccg agttttctgt ggtcaggtac aaaggtgaca agtccaaagc agacgacgtc      660

tacagaggta caacaccact atatgccgat gatatcgcgg atttgattgt gtactctacc      720

agcagaaagc caaacatggt ggtagcagat gtcctggtct tcccaacaca ccaggcatcg      780

gcttcgcaca tctacagggg cgactaa                                          807
```

<210> 709
<211> 268
<212> PRT
<213> Artificial Sequence

<220>
<223> Candida glabrata CR380959

<400> 709

```
Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Gln Gly Lys Ile Ala
1               5                  10                  15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Lys Ala Thr Ala Ile Glu
            20                  25                  30

Tyr Leu Asp Ala Ser Asn Gly Ser Val Lys Leu Val Leu Gly Ala Arg
        35                  40                  45

Arg Met Glu Lys Leu Glu Glu Leu Lys Lys Glu Leu Leu Ala Gln Tyr
    50                  55                  60

Pro Asp Ala Lys Ile His Ile Gly Lys Leu Asp Val Thr Asp Phe Glu
65                  70                  75                  80
```

```
Asn Val Lys Gln Phe Leu Ala Asp Leu Pro Glu Glu Phe Lys Asp Ile
                85                  90                  95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Lys Val
            100                 105                 110

Gly Asp Ile Asp Pro Glu Asp Ile Ala Gly Met Val Asn Thr Asn Val
        115                 120                 125

Leu Ala Leu Ile Asn Leu Thr Gln Leu Leu Leu Pro Leu Phe Lys Lys
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Ile Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ala Ile Tyr Cys Ala Thr Lys His Ala Val Arg
                165                 170                 175

Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Asp Ile Arg
            180                 185                 190

Val Ile Glu Ile Ala Pro Gly Met Val Glu Thr Glu Phe Ser Val Val
        195                 200                 205

Arg Tyr Lys Gly Asp Lys Ser Lys Ala Asp Asp Val Tyr Arg Gly Thr
    210                 215                 220

Thr Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ser Thr
225                 230                 235                 240

Ser Arg Lys Pro Asn Met Val Val Ala Asp Val Leu Val Phe Pro Thr
                245                 250                 255

His Gln Ala Ser Ala Ser His Ile Tyr Arg Gly Asp
            260                 265
```

<210> 710
<211> 813
<212> DNA
<213> Artificial Sequence

<220>
<223> Debaryomyces hansenii CR382139

<400> 710

```
atgtcgtacg gatctaaagc tgctgaacgt gttgccaata agattgtctt aatcactggt       60

gcttcatctg gaattggtga agcaactgcc aaagaaattg catcagccgc taatggcaat      120

ttaaaattag tgttgtgtgc tagacgaaaa gaaaagttgg ataatttatc taaagaattg      180

actgacaaat attcatccat caaggttcat gttgctcaac tagatgtatc taagctcgag      240

actatcaagc catttatcaa tgatttaccg aaagaattct ctgacgtgga tgtattagtc      300

aacaatgcag gcttggcttt gggccgtgat gaagttggaa ccattgacac agatgatatg      360

ttatcgatgt ttcaaactaa tgttttaggg ttaattacca tcacacaggc tgttttgcca      420

atcatgaaaa agaagaacag cggagatgtt gttaatatag gttcaattgc tggaagagac      480
```

tcttaccctg gaggtggaat ttactgtcca actaaggcaa gtgtcaagtc gttttcgcaa 540

gttttaagaa aggaattgat tagcaccaag attagagttc ttgaggttga ccctggtaat 600

gttgaaactg aattttcaaa tgtcagattc aagggcgata tggaaaaggc aaagctggtt 660

tacgcgggta ctgaaccatt attatccgaa gacgtagctg aggttgtcgt attcggactt 720

acaagaaagc aaaataccgt tattgctgag acattagtct tttcaaccaa tcaagccagc 780

tcatctcact tataccgtga aagcgataaa taa 813

<210> 711
<211> 270
<212> PRT
<213> Artificial Sequence

<220>
<223> Debaryomyces hansenii CR382139

<400> 711

Met Ser Tyr Gly Ser Lys Ala Ala Glu Arg Val Ala Asn Lys Ile Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ala Thr Ala Lys Glu
20 25 30

Ile Ala Ser Ala Ala Asn Gly Asn Leu Lys Leu Val Leu Cys Ala Arg
35 40 45

Arg Lys Glu Lys Leu Asp Asn Leu Ser Lys Glu Leu Thr Asp Lys Tyr
50 55 60

Ser Ser Ile Lys Val His Val Ala Gln Leu Asp Val Ser Lys Leu Glu
65 70 75 80

Thr Ile Lys Pro Phe Ile Asn Asp Leu Pro Lys Glu Phe Ser Asp Val
85 90 95

Asp Val Leu Val Asn Asn Ala Gly Leu Ala Leu Gly Arg Asp Glu Val
100 105 110

Gly Thr Ile Asp Thr Asp Asp Met Leu Ser Met Phe Gln Thr Asn Val
115 120 125

Leu Gly Leu Ile Thr Ile Thr Gln Ala Val Leu Pro Ile Met Lys Lys
130 135 140

Lys Asn Ser Gly Asp Val Val Asn Ile Gly Ser Ile Ala Gly Arg Asp
145 150 155 160

Ser Tyr Pro Gly Gly Gly Ile Tyr Cys Pro Thr Lys Ala Ser Val Lys
165 170 175

Ser Phe Ser Gln Val Leu Arg Lys Glu Leu Ile Ser Thr Lys Ile Arg
180 185 190

20172000439 20 Jan 2017

```
Val Leu Glu Val Asp Pro Gly Asn Val Glu Thr Glu Phe Ser Asn Val
        195                 200                 205

Arg Phe Lys Gly Asp Met Glu Lys Ala Lys Leu Val Tyr Ala Gly Thr
    210                 215                 220

Glu Pro Leu Leu Ser Glu Asp Val Ala Glu Val Val Val Phe Gly Leu
225                 230                 235                 240

Thr Arg Lys Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Ser Thr
                245                 250                 255

Asn Gln Ala Ser Ser Ser His Leu Tyr Arg Glu Ser Asp Lys
            260                 265                 270
```

<210> 712
<211> 810
<212> DNA
<213> Artificial Sequence

<220>
<223> Scheffersomyces stipitis XM_001387479

<400> 712

```
atgtcgtttg gaaaaaaagc tgctgaaaga cttgccaaca aaatcattct tatcaccggg       60

gcttcgtctg gtattggtga agctacagct agagagtttg catctgctgc caatgggaat      120

atcagattga ttttgacagc cagaagaaaa gaaaagttgg ctcaattgtc agactcattg      180

accaaggaat ttccaactat caaaatccat tctgccaaat tggatgtgac cgaacatgat      240

ggcatcaagc ctttcatttc tggtttaccc aaggatttcg ccgacatcga tgtgttgatc      300

aacaatgctg gaaaagctct tggaaaagca tctgttggtg aaatcagtga cagtgatatc      360

caaggcatga tgcaaacgaa tgtcttggga ctcatcaaca tgactcaggc tgtgattccc      420

atttttaagg ctaaaaattc tggagatatc gtcaacatcg gttcgattgc tggaagagac      480

ccttaccctg gtggatcgat ctactgtgcc tccaaggctg ctgttaagtt cttctcgcat      540

tctttgagaa aggaactcat taacaccaga atcagagttt tggaagttga tccaggtgct      600

gtgttgaccg agttctcttt ggttcgtttc cacggtgatc agggagctgc tgatgctgtt      660

tatgaaggta cccaaccttt ggatgcctct gatatcgcag aagttatcgt gtttggtatc      720

accagaaagc agaacaccgt catagccgaa accttggtat tcccaagtca ccaggcttct      780

gcctctcatg tttacaaggc tcctaagtag                                       810
```

<210> 713
<211> 269
<212> PRT
<213> Artificial Sequence

<220>
<223> Scheffersomyces stipitis XM_001387479

<400> 713

```
Met Ser Phe Gly Lys Lys Ala Ala Glu Arg Leu Ala Asn Lys Ile Ile
1               5                   10                  15
```

```
Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ala Thr Ala Arg Glu
            20                  25                  30

Phe Ala Ser Ala Ala Asn Gly Asn Ile Arg Leu Ile Leu Thr Ala Arg
        35                  40                  45

Arg Lys Glu Lys Leu Ala Gln Leu Ser Asp Ser Leu Thr Lys Glu Phe
    50                  55                  60

Pro Thr Ile Lys Ile His Ser Ala Lys Leu Asp Val Thr Glu His Asp
65                  70                  75                  80

Gly Ile Lys Pro Phe Ile Ser Gly Leu Pro Lys Asp Phe Ala Asp Ile
                85                  90                  95

Asp Val Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Lys Ala Ser Val
            100                 105                 110

Gly Glu Ile Ser Asp Ser Asp Ile Gln Gly Met Met Gln Thr Asn Val
        115                 120                 125

Leu Gly Leu Ile Asn Met Thr Gln Ala Val Ile Pro Ile Phe Lys Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Ile Gly Ser Ile Ala Gly Arg Asp
145                 150                 155                 160

Pro Tyr Pro Gly Gly Ser Ile Tyr Cys Ala Ser Lys Ala Ala Val Lys
                165                 170                 175

Phe Phe Ser His Ser Leu Arg Lys Glu Leu Ile Asn Thr Arg Ile Arg
            180                 185                 190

Val Leu Glu Val Asp Pro Gly Ala Val Leu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Phe His Gly Asp Gln Gly Ala Ala Asp Ala Val Tyr Glu Gly Thr
    210                 215                 220

Gln Pro Leu Asp Ala Ser Asp Ile Ala Glu Val Ile Val Phe Gly Ile
225                 230                 235                 240

Thr Arg Lys Gln Asn Thr Val Ile Ala Glu Thr Leu Val Phe Pro Ser
                245                 250                 255

His Gln Ala Ser Ala Ser His Val Tyr Lys Ala Pro Lys
            260                 265
```

<210> 714
<211> 891
<212> DNA
<213> Artificial Sequence

<220>
<223> Meyerozyma guilliermondii XM_001482184

<400> 714
atgtgcctct taccagccgg tagcactgta ttatgtcatc acccagtagt gagtgtggag 60

attaaatcct caatcttcat gtctttcggt gccaaagccg ctgaacgcct tgccaacaag 120

atcatattga tcactggggc atcgtctggt ataggcgagg ctaccgccag agaattcgct 180

gctgctgcca atggaaaaat tctgttgatt ttgaccgctc ggagagaaga caaactcaag 240

tctctctcgc aacaattgag cctcatttac ccgcaaatta aaatccattc tgctcgtctt 300

gatgtctctg agttttcgtc acttaagccg ttcattactg ggttgccaaa ggattttgct 360

agcatcgacg ttttggtgaa taatgcgggg aaagcattgg gaagagccaa tgttggtgaa 420

atttcccaag aggaaatcaa tggcatgttc cataccaatg ttcttgggtt gataaactta 480

actcaggagg tgttacccat cttcaaaaag aaaaatgctg gagatattgt gaacattggc 540

tcagtggccg gtagagaacc ttaccctgga ggtgcagtat actgtgcttc aaaggcagca 600

gttaactact tttctcattc tttgagaaag gaaactatca attccaaaat cagggtcatg 660

gaggtggatc ctggggcagt agagacagag ttctcgttgg ttcgttttgg cggtgatgcc 720

gaggctgcga aaaaggtgta tgagggaacc gagcctttgg gcccagaaga tattgcggaa 780

atcattgtgt ttgctgtgtc gagaaaagcc aaaactgtca ttgcggaaac tttggtgttt 840

cctacccatc aggctggagc agttcatgtt catagagggc cgcttgagtg a 891

<210> 715
<211> 296
<212> PRT
<213> Artificial Sequence

<220>
<223> Meyerozyma guilliermondii XM_001482184

<400> 715

Met Cys Leu Leu Pro Ala Gly Ser Thr Val Leu Cys His His Pro Val
1 5 10 15

Val Ser Val Glu Ile Lys Ser Ser Ile Phe Met Ser Phe Gly Ala Lys
20 25 30

Ala Ala Glu Arg Leu Ala Asn Lys Ile Ile Leu Ile Thr Gly Ala Ser
35 40 45

Ser Gly Ile Gly Glu Ala Thr Ala Arg Glu Phe Ala Ala Ala Ala Asn
50 55 60

Gly Lys Ile Leu Leu Ile Leu Thr Ala Arg Arg Glu Asp Lys Leu Lys
65 70 75 80

Ser Leu Ser Gln Gln Leu Ser Leu Ile Tyr Pro Gln Ile Lys Ile His
85 90 95

Ser Ala Arg Leu Asp Val Ser Glu Phe Ser Ser Leu Lys Pro Phe Ile
100 105 110

Thr Gly Leu Pro Lys Asp Phe Ala Ser Ile Asp Val Leu Val Asn Asn
115 120 125

```
Ala Gly Lys Ala Leu Gly Arg Ala Asn Val Gly Glu Ile Ser Gln Glu
    130                 135                 140

Glu Ile Asn Gly Met Phe His Thr Asn Val Leu Gly Leu Ile Asn Leu
145                 150                 155                 160

Thr Gln Glu Val Leu Pro Ile Phe Lys Lys Lys Asn Ala Gly Asp Ile
                165                 170                 175

Val Asn Ile Gly Ser Val Ala Gly Arg Glu Pro Tyr Pro Gly Gly Ala
            180                 185                 190

Val Tyr Cys Ala Ser Lys Ala Ala Val Asn Tyr Phe Ser His Ser Leu
        195                 200                 205

Arg Lys Glu Thr Ile Asn Ser Lys Ile Arg Val Met Glu Val Asp Pro
    210                 215                 220

Gly Ala Val Glu Thr Glu Phe Ser Leu Val Arg Phe Gly Gly Asp Ala
225                 230                 235                 240

Glu Ala Ala Lys Lys Val Tyr Glu Gly Thr Glu Pro Leu Gly Pro Glu
                245                 250                 255

Asp Ile Ala Glu Ile Ile Val Phe Ala Val Ser Arg Lys Ala Lys Thr
            260                 265                 270

Val Ile Ala Glu Thr Leu Val Phe Pro Thr His Gln Ala Gly Ala Val
        275                 280                 285

His Val His Arg Gly Pro Leu Glu
    290                 295
```

<210> 716
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Vanderwaltozyma polyspora XM_001645671

<400> 716

```
atgtcacagg gtagaaaggc ttcagaaagg ttggctggta aaactgtatt aattacaggt       60

gcttcatcag ggattgggaa agccactgca ttagaatatc tagatgcctc caatggtcat      120

atgaagttaa ttttagttgc aagaagatta gaaaaattgc aagagttgaa ggaaacaatt      180

tgtaaagaat atccagaatc taaggttcat gttgaagaat tagatatttc tgatattaat      240

agaatcccag aatttattgc aaaattacct gaagaattca aagatattga tatattgatt      300

aacaatgcag gtaaagcatt aggaagtgat actattggta atatcgagaa tgaggatatt      360

aaaggtatgt ttgagactaa cgtttttgga ttaatctgtt taacacaagc tgtacttcca      420

atattcaagg ctaaaaatgg tggtgatatt gtcaatttag ggtcaattgc aggcatagaa      480

gcttacccaa caggatctat atattgtgca actaaatttg cagttaaagc attcactgaa      540
```

agtttaagaa aggaattgat taatacaaag atcagagtta ttgaaattgc accaggtatg 600

gttaacactg aattttctgt aattagatat aaaggtgacc aagaaaaggc agataaagtt 660

tatgaaaaca ctactccttt atatgcagat gacatcgctg atttgatagt ttacaccact 720

tctagaaaat cgaataccgt tatcgctgat gttttggtat tcccaacatg ccaagcttct 780

gcatcccata tctatcgtgg ataa 804

<210> 717
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> Vanderwaltozyma polyspora XM_001645671

<400> 717

Met Ser Gln Gly Arg Lys Ala Ser Glu Arg Leu Ala Gly Lys Thr Val
1 5 10 15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Lys Ala Thr Ala Leu Glu
20 25 30

Tyr Leu Asp Ala Ser Asn Gly His Met Lys Leu Ile Leu Val Ala Arg
35 40 45

Arg Leu Glu Lys Leu Gln Glu Leu Lys Glu Thr Ile Cys Lys Glu Tyr
50 55 60

Pro Glu Ser Lys Val His Val Glu Glu Leu Asp Ile Ser Asp Ile Asn
65 70 75 80

Arg Ile Pro Glu Phe Ile Ala Lys Leu Pro Glu Glu Phe Lys Asp Ile
85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Thr Ile
100 105 110

Gly Asn Ile Glu Asn Glu Asp Ile Lys Gly Met Phe Glu Thr Asn Val
115 120 125

Phe Gly Leu Ile Cys Leu Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
130 135 140

Lys Asn Gly Gly Asp Ile Val Asn Leu Gly Ser Ile Ala Gly Ile Glu
145 150 155 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Thr Lys Phe Ala Val Lys
165 170 175

Ala Phe Thr Glu Ser Leu Arg Lys Glu Leu Ile Asn Thr Lys Ile Arg
180 185 190

Val Ile Glu Ile Ala Pro Gly Met Val Asn Thr Glu Phe Ser Val Ile
195 200 205

```
Arg Tyr Lys Gly Asp Gln Glu Lys Ala Asp Lys Val Tyr Glu Asn Thr
    210                 215                 220

Thr Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Thr Thr
225                 230                 235                 240

Ser Arg Lys Ser Asn Thr Val Ile Ala Asp Val Leu Val Phe Pro Thr
                245                 250                 255

Cys Gln Ala Ser Ala Ser His Ile Tyr Arg Gly
            260                 265
```

<210> 718
<211> 807
<212> DNA
<213> Artificial Sequence

<220>
<223> Candida dubliniensis XM_002419771

<400> 718

```
atgtcatttg gtagaaaagc tgctgaaaga ttagccaata gatccattct tatcactggt       60

gcttcatctg ggattggtga agcatgtgct aaagttttcg ctgaagcatc taatggtcaa      120

gttaaattag ttttaggagc aagaagaaaa gaacgattag ttaaattatc tgatacttta      180

attaaacaat atcctaatat taaaattcat catgattttt tggatgttac tattaaagat      240

tcaatttcaa aattcattgc tggaattcct catgaatttg aacctgatgt attaattaat      300

aatagtggta aagccttggg gaaagaagaa gttggagaat tgaaagatga agatattacg      360

gaaatgtttg atactaatgt cattggagtc attcgtatga ctcaagcagt tttaccttta      420

cttaaaaaaa aaccttatgc tgatgtggtt ttcattggaa gtattgctgg acgtgttcct      480

tataaaaatg gaggtggtta ttgtgcatct aaagctgctg ttcgtagttt caccgataca      540

tttagaaaag aaactattaa tactggtatt agagtcattg aagttgatcc aggtgcagta      600

cttactgaat ttagtgttgt tcgttataaa ggtgacactg atgctgccga tgctgtttat      660

actggtactg aaccattaac accagaagat gttgctgaag tggttgtttt tgcatcttca      720

agaaaacaaa ataccgttat tgctgatact ttgattttcc caaatcatca agcttctcca      780

gatcatgttt atagaaaacc taattaa                                          807
```

<210> 719
<211> 268
<212> PRT
<213> Artificial Sequence

<220>
<223> Candida dubliniensis XM_002419771

<400> 719

```
Met Ser Phe Gly Arg Lys Ala Ala Glu Arg Leu Ala Asn Arg Ser Ile
1               5                   10                  15

Leu Ile Thr Gly Ala Ser Ser Gly Ile Gly Glu Ala Cys Ala Lys Val
            20                  25                  30
```

```
Phe Ala Glu Ala Ser Asn Gly Gln Val Lys Leu Val Leu Gly Ala Arg
        35                  40                  45

Arg Lys Glu Arg Leu Val Lys Leu Ser Asp Thr Leu Ile Lys Gln Tyr
    50                  55                  60

Pro Asn Ile Lys Ile His His Asp Phe Leu Asp Val Thr Ile Lys Asp
65                  70                  75                  80

Ser Ile Ser Lys Phe Ile Ala Gly Ile Pro His Glu Phe Glu Pro Asp
                85                  90                  95

Val Leu Ile Asn Asn Ser Gly Lys Ala Leu Gly Lys Glu Glu Val Gly
            100                 105                 110

Glu Leu Lys Asp Glu Asp Ile Thr Glu Met Phe Asp Thr Asn Val Ile
        115                 120                 125

Gly Val Ile Arg Met Thr Gln Ala Val Leu Pro Leu Leu Lys Lys Lys
    130                 135                 140

Pro Tyr Ala Asp Val Val Phe Ile Gly Ser Ile Ala Gly Arg Val Pro
145                 150                 155                 160

Tyr Lys Asn Gly Gly Gly Tyr Cys Ala Ser Lys Ala Ala Val Arg Ser
                165                 170                 175

Phe Thr Asp Thr Phe Arg Lys Glu Thr Ile Asn Thr Gly Ile Arg Val
            180                 185                 190

Ile Glu Val Asp Pro Gly Ala Val Leu Thr Glu Phe Ser Val Val Arg
        195                 200                 205

Tyr Lys Gly Asp Thr Asp Ala Ala Asp Ala Val Tyr Thr Gly Thr Glu
    210                 215                 220

Pro Leu Thr Pro Glu Asp Val Ala Glu Val Val Phe Ala Ser Ser
225                 230                 235                 240

Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Pro Asn His
                245                 250                 255

Gln Ala Ser Pro Asp His Val Tyr Arg Lys Pro Asn
            260                 265
```

<210> 720
<211> 822
<212> DNA
<213> Artificial Sequence

<220>
<223> Zygosaccharomyces rouxii XM_002494574

<400> 720
atgtcacaag gtgtcaaagc tgctgaaaga ctagctggta agactgtatt cattacaggt 60

gcttctgcag gtatcggtca agcaactgca aaggaatatt tggatgcatc caatggtcaa 120

attaaattga tcttggctgc aagaagatta gagaaattac acgagtttaa agaacaaact 180

acaaagagtt acccaagcgc tcaagtccac attggtaaat tggacgtcac tgcaattgac 240

accataaaac catttttgga taaattacca aaggaatttc aagatatcga tattttgatc 300

aacaatgccg gtaaggcatt aggtactgat aaagttggtg atattgcaga tgaagacgtg 360

gaaggtatgt tcgacaccaa tgtcttgggg ttaatcaaag ttactcaagc tgttttacct 420

atcttcaaaa gaaaaaattc tggtgatgtc gttaacatta gttcggttgc tggtagagag 480

gcatacccag gtggttccat ttactgtgct actaaacacg ctgttaaggc attcactgaa 540

agtttgcgta aggaattagt cgatacaaaa atcagagtca tgagtattga tcctggtaat 600

gtagagaccg aattttctat ggttagattc cgtggtgata cagaaaaggc aaagaaggtt 660

taccaagaca ctgtcccatt atatgcagat gacattgcag atttaatcgt ctatgcaacc 720

tctagaaagc aaaacactgt cattgctgac actttgatct tctcttctaa ccaggcatca 780

ccataccacc tctacagagg ctctcaagac aaaaccaatt ga 822

<210> 721
<211> 273
<212> PRT
<213> Artificial Sequence

<220>
<223> Zygosaccharomyces rouxii XM_002494574

<400> 721

Met Ser Gln Gly Val Lys Ala Ala Glu Arg Leu Ala Gly Lys Thr Val
1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Lys Glu
20 25 30

Tyr Leu Asp Ala Ser Asn Gly Gln Ile Lys Leu Ile Leu Ala Ala Arg
35 40 45

Arg Leu Glu Lys Leu His Glu Phe Lys Glu Gln Thr Thr Lys Ser Tyr
50 55 60

Pro Ser Ala Gln Val His Ile Gly Lys Leu Asp Val Thr Ala Ile Asp
65 70 75 80

Thr Ile Lys Pro Phe Leu Asp Lys Leu Pro Lys Glu Phe Gln Asp Ile
85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Thr Asp Lys Val
100 105 110

Gly Asp Ile Ala Asp Glu Asp Val Glu Gly Met Phe Asp Thr Asn Val
115 120 125

Leu Gly Leu Ile Lys Val Thr Gln Ala Val Leu Pro Ile Phe Lys Arg
130 135 140

```
Lys Asn Ser Gly Asp Val Val Asn Ile Ser Ser Val Ala Gly Arg Glu
145                 150                 155                 160

Ala Tyr Pro Gly Gly Ser Ile Tyr Cys Ala Thr Lys His Ala Val Lys
                165                 170                 175

Ala Phe Thr Glu Ser Leu Arg Lys Glu Leu Val Asp Thr Lys Ile Arg
            180                 185                 190

Val Met Ser Ile Asp Pro Gly Asn Val Glu Thr Glu Phe Ser Met Val
        195                 200                 205

Arg Phe Arg Gly Asp Thr Glu Lys Ala Lys Lys Val Tyr Gln Asp Thr
    210                 215                 220

Val Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Gln Asn Thr Val Ile Ala Asp Thr Leu Ile Phe Ser Ser
                245                 250                 255

Asn Gln Ala Ser Pro Tyr His Leu Tyr Arg Gly Ser Gln Asp Lys Thr
            260                 265                 270

Asn
```

<210> 722
<211> 804
<212> DNA
<213> Artificial Sequence

<220>
<223> Lachancea thermotolerans XM_002553230

<400> 722

```
atgtcacagg gaagaagagc agctgaaaga ctggcaggaa agactgtctt catcacaggc       60

gcatcagccg gtatcggtca ggccactgcg caagaatacc tggaagcatc cgaaggcaaa      120

atcaagttga tccttgcagc aagaagactc gacaagctgg aggaaatcaa agccaaggtt      180

tctaaagact tccctgaagc acaggtgcat atcggccagc tagatgtgac tcagacggac      240

aaaatccagc cttttgtcga caatttgccc gaagagttca aagacatcga catcctgatc      300

aacaacgcgg gcaaggcgct cggatccgac cccgtgggca caatcgaccc caatgatatt      360

caaggcatga tccagactaa cgttatcggg cttataaatg ttacccaagc cgttctgccc      420

atcttcaagg ccaaaaactc tggtgatatc gtgaacctgg gttctgtcgc tggcagagaa      480

gcttacccta caggatctat ttactgcgct acgaagcacg cggtgcgtgc tttcacccag      540

agcctgcgca aggaactgat caacacaaac atcagggtta ttgaggtcgc tccaggtaac      600

gtggagaccg agttttctct ggttagatac aagggcgact ctgagaaagc caagaaggtt      660

tacgaaggca cacaacccct ttacgctgac gatatcgcag acctaatcgt ttacgcaacc      720

tcgagaaaac caaacaccgt catcgcggac gttttggttt tcgcttcgaa ccaggcttcg      780
```

2017200439 20 Jan 2017

ccttaccaca tttaccgtgg ttag 804

<210> 723
<211> 267
<212> PRT
<213> Artificial Sequence

<220>
<223> Lachancea thermotolerans XM_002553230

<400> 723

```
Met Ser Gln Gly Arg Arg Ala Ala Glu Arg Leu Ala Gly Lys Thr Val
1               5                  10                  15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Gln Glu
            20                  25                  30

Tyr Leu Glu Ala Ser Glu Gly Lys Ile Lys Leu Ile Leu Ala Ala Arg
        35                  40                  45

Arg Leu Asp Lys Leu Glu Glu Ile Lys Ala Lys Val Ser Lys Asp Phe
    50                  55                  60

Pro Glu Ala Gln Val His Ile Gly Gln Leu Asp Val Thr Gln Thr Asp
65                  70                  75                  80

Lys Ile Gln Pro Phe Val Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile
                85                  90                  95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Pro Val
            100                 105                 110

Gly Thr Ile Asp Pro Asn Asp Ile Gln Gly Met Ile Gln Thr Asn Val
        115                 120                 125

Ile Gly Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Glu
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Thr Lys His Ala Val Arg
                165                 170                 175

Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Asn Ile Arg
            180                 185                 190

Val Ile Glu Val Ala Pro Gly Asn Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Lys Gly Asp Ser Glu Lys Ala Lys Lys Val Tyr Glu Gly Thr
    210                 215                 220

Gln Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240
```

```
Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser
                245                 250                 255

Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly
            260                 265
```

<210> 724
<211> 807
<212> DNA
<213> Artificial Sequence

<220>
<223> Kluyveromyces lactis XM_451902

<400> 724

```
atgtctcaag gtagaaaggc tgctgaaaga ttgcaaaaca agacaatttt cattaccggt       60

gcttctgcag gtattggtca agccacagca ttggaatatc tagatgctgc taacggtaat      120

gtcaaattga tcttagcagc aagaaggttg gctaagttgg aagaattgaa ggaaaaaatc      180

aatgctgaat acccacaagc taaagtatat atcggtcaat tggacgtcac tgaaactgag      240

aagattcaac ctttcattga taacttgccg gaagaattca aggatatcga tattttgatt      300

aacaatgccg gtaaagcttt gggatctgat gttgtcggta ccatcagtag cgaggacatc      360

aaaggtatga tagatactaa cgttgttgcc cttatcaacg ttacccaagc tgttttgcct      420

attttcaaag caaagaattc cggtgacatc gttaacttag gttctgttgc cggtagagat      480

gcatatccaa ctggttctat ctattgtgct tcgaagcatg ctgtcagagc gttcactcag      540

tctttgagaa aagaattaat caatactggt attagggtca ttgagattgc tccaggtaat      600

gtcgaaactg agttctctct agttagatac aagggtgatg ccgatcgtgc taaacaggtt      660

tacaaaggta ctactcctct atatgcagat gacattgctg acttgatcgt ttatgccact      720

tcaagaaaac ctaatactgt catcgctgat gttttggtat ttgcttccaa ccaagcatct      780

ccttaccaca tttaccgtgg cgaatag                                          807
```

<210> 725
<211> 268
<212> PRT
<213> Artificial Sequence

<220>
<223> Kluyveromyces lactis XM_451902

<400> 725

```
Met Ser Gln Gly Arg Lys Ala Ala Glu Arg Leu Gln Asn Lys Thr Ile
1               5                   10                  15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Leu Glu
            20                  25                  30

Tyr Leu Asp Ala Ala Asn Gly Asn Val Lys Leu Ile Leu Ala Ala Arg
        35                  40                  45

Arg Leu Ala Lys Leu Glu Glu Leu Lys Glu Lys Ile Asn Ala Glu Tyr
    50                  55                  60
```

```
Pro Gln Ala Lys Val Tyr Ile Gly Gln Leu Asp Val Thr Glu Thr Glu
65                  70                  75                  80

Lys Ile Gln Pro Phe Ile Asp Asn Leu Pro Glu Glu Phe Lys Asp Ile
                85                  90                  95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Val Val
            100                 105                 110

Gly Thr Ile Ser Ser Glu Asp Ile Lys Gly Met Ile Asp Thr Asn Val
        115                 120                 125

Val Ala Leu Ile Asn Val Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
    130                 135                 140

Lys Asn Ser Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145                 150                 155                 160

Ala Tyr Pro Thr Gly Ser Ile Tyr Cys Ala Ser Lys His Ala Val Arg
                165                 170                 175

Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Ile Asn Thr Gly Ile Arg
            180                 185                 190

Val Ile Glu Ile Ala Pro Gly Asn Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Lys Gly Asp Ala Asp Arg Ala Lys Gln Val Tyr Lys Gly Thr
    210                 215                 220

Thr Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser
                245                 250                 255

Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly Glu
            260                 265
```

<210> 726
<211> 807
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces kluyveri SAKL0H04730

<400> 726

```
atgtctcaag gtagaagggc tgcagaaaga ctagcaaaca agaccgtttt tataactggc       60

gcctctgccg gcattggcca agctactgct ttggaatact gtgatgcttc taacggtaaa      120

ataaacttgg tgttaagtgc cagaaggctg gaaaaattgc aagagttaaa ggacaaaatc      180

accaaggagt atcctgaagc caaggtttat attggtgtgc ttgatgtgac cgaaacggaa      240

aaaatcaaac cattcttgga tggtttacca gaagaattta aagatattga catcttgatc      300
```

aataatgcag gcaaagcgtt aggctctgat cctgttggta ccatcaaaac tgaagatatt 360

gaaggaatga tcaacaccaa tgtcttagct cttatcaata ttactcaagc tgtcttgcca 420

atcttcaaag ccaagaattt cggtgatatc gtaaacttgg ggtctgtcgc tggtagagat 480

gcttatccaa ccggtgcaat ctactgtgct agcaaacatg cagtcagagc cttcactcaa 540

agtttgagga aggaattggt gaacaccaat atcagagtga ttgaaattgc tccgggtaat 600

gttgaaaccg agttttcctt agttagatat aaaggtgata cggaccgtgc taaaaaggtt 660

tatgaaggta ctaacccatt atatgcagat gacattgcag accttattgt gtatgctact 720

tctagaaagc ctaatactgt cattgcggat gttttggttt ttgcttcaaa ccaagcatcc 780

ccttaccata tctatcgcgg tgactaa 807

<210> 727
<211> 268
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces kluyveri SAKL0H04730

<400> 727

Met Ser Gln Gly Arg Arg Ala Ala Glu Arg Leu Ala Asn Lys Thr Val
1 5 10 15

Phe Ile Thr Gly Ala Ser Ala Gly Ile Gly Gln Ala Thr Ala Leu Glu
20 25 30

Tyr Cys Asp Ala Ser Asn Gly Lys Ile Asn Leu Val Leu Ser Ala Arg
35 40 45

Arg Leu Glu Lys Leu Gln Glu Leu Lys Asp Lys Ile Thr Lys Glu Tyr
50 55 60

Pro Glu Ala Lys Val Tyr Ile Gly Val Leu Asp Val Thr Glu Thr Glu
65 70 75 80

Lys Ile Lys Pro Phe Leu Asp Gly Leu Pro Glu Glu Phe Lys Asp Ile
85 90 95

Asp Ile Leu Ile Asn Asn Ala Gly Lys Ala Leu Gly Ser Asp Pro Val
100 105 110

Gly Thr Ile Lys Thr Glu Asp Ile Glu Gly Met Ile Asn Thr Asn Val
115 120 125

Leu Ala Leu Ile Asn Ile Thr Gln Ala Val Leu Pro Ile Phe Lys Ala
130 135 140

Lys Asn Phe Gly Asp Ile Val Asn Leu Gly Ser Val Ala Gly Arg Asp
145 150 155 160

Ala Tyr Pro Thr Gly Ala Ile Tyr Cys Ala Ser Lys His Ala Val Arg
165 170 175

```
Ala Phe Thr Gln Ser Leu Arg Lys Glu Leu Val Asn Thr Asn Ile Arg
            180                 185                 190

Val Ile Glu Ile Ala Pro Gly Asn Val Glu Thr Glu Phe Ser Leu Val
        195                 200                 205

Arg Tyr Lys Gly Asp Thr Asp Arg Ala Lys Lys Val Tyr Glu Gly Thr
    210                 215                 220

Asn Pro Leu Tyr Ala Asp Asp Ile Ala Asp Leu Ile Val Tyr Ala Thr
225                 230                 235                 240

Ser Arg Lys Pro Asn Thr Val Ile Ala Asp Val Leu Val Phe Ala Ser
                245                 250                 255

Asn Gln Ala Ser Pro Tyr His Ile Tyr Arg Gly Asp
            260                 265
```

<210> 728
<211> 807
<212> DNA
<213> Artificial Sequence

<220>
<223> Yarrowia lipolytica XM_501554

<400> 728
atgtctttcg gagataaagc tgctgctcga cttgcgggca agaccgtctt tgttaccggc 60

gcctcgtccg gcattggcca ggccactgtt ctcgctctag ccgaagctgc caagggcgac 120

ctcaagtttg tgcttgctgc ccgacgaacc gaccgtctgg acgagctcaa gaagaagctg 180

gagaccgact acaagggtat ccaggtgctg cctttcaagc tggacgtgtc caaggtcgag 240

gagaccgaga acattgtgtc caagctgccc aaggagtttt ccgaggtgga cgtgcttatc 300

aacaacgccg gcatggtcca cggcaccgaa aaggttggct ccatcaacca gaacgacatt 360

gagatcatgt tccacacaaa cgtgctcgga ctcatttctg tcactcagca gtttgtcggc 420

gagatgcgaa agcgaaacaa gggcgacatt gtcaacattg gctccatcgc cggacgagag 480

ccctacgttg gaggaggaat ctactgtgcc accaaggccg ccgtgcgatc tttcactgag 540

actctccgaa aagagaacat cgacactcga atccgagtca ttgaggttga tcctggagcc 600

gttgagaccg agttctccgt cgtgcgattc cgaggagaca agtccaaggc cgacgctgtt 660

tacgctggaa ccgagcctct ggtcgctgac gatattgccg agttcatcac ctacactctc 720

actcgacgag agaatgtcgt cattgccgat actctcattt tccccaacca ccaggcttct 780

cctactcacg tctaccgaaa gaactga 807

<210> 729
<211> 268
<212> PRT
<213> Artificial Sequence

<220>
<223> Yarrowia lipolytica XM_501554

<400> 729

Met Ser Phe Gly Asp Lys Ala Ala Ala Arg Leu Ala Gly Lys Thr Val
1 5 10 15

Phe Val Thr Gly Ala Ser Ser Gly Ile Gly Gln Ala Thr Val Leu Ala
20 25 30

Leu Ala Glu Ala Ala Lys Gly Asp Leu Lys Phe Val Leu Ala Ala Arg
35 40 45

Arg Thr Asp Arg Leu Asp Glu Leu Lys Lys Lys Leu Glu Thr Asp Tyr
50 55 60

Lys Gly Ile Gln Val Leu Pro Phe Lys Leu Asp Val Ser Lys Val Glu
65 70 75 80

Glu Thr Glu Asn Ile Val Ser Lys Leu Pro Lys Glu Phe Ser Glu Val
85 90 95

Asp Val Leu Ile Asn Asn Ala Gly Met Val His Gly Thr Glu Lys Val
100 105 110

Gly Ser Ile Asn Gln Asn Asp Ile Glu Ile Met Phe His Thr Asn Val
115 120 125

Leu Gly Leu Ile Ser Val Thr Gln Gln Phe Val Gly Glu Met Arg Lys
130 135 140

Arg Asn Lys Gly Asp Ile Val Asn Ile Gly Ser Ile Ala Gly Arg Glu
145 150 155 160

Pro Tyr Val Gly Gly Gly Ile Tyr Cys Ala Thr Lys Ala Ala Val Arg
165 170 175

Ser Phe Thr Glu Thr Leu Arg Lys Glu Asn Ile Asp Thr Arg Ile Arg
180 185 190

Val Ile Glu Val Asp Pro Gly Ala Val Glu Thr Glu Phe Ser Val Val
195 200 205

Arg Phe Arg Gly Asp Lys Ser Lys Ala Asp Ala Val Tyr Ala Gly Thr
210 215 220

Glu Pro Leu Val Ala Asp Asp Ile Ala Glu Phe Ile Thr Tyr Thr Leu
225 230 235 240

Thr Arg Arg Glu Asn Val Val Ile Ala Asp Thr Leu Ile Phe Pro Asn
245 250 255

His Gln Ala Ser Pro Thr His Val Tyr Arg Lys Asn
260 265

<210> 730
<211> 780

<212> DNA
<213> Artificial Sequence

<220>
<223> Schizosaccharomyces pombe NM_001018495

<400> 730
atgagccgtt tggatggaaa aacgatttta atcactggtg cctcttctgg aattggaaaa 60

agcactgctt ttgaaattgc caaagttgcc aaagtaaaac ttattttggc tgctcgcaga 120

ttttctaccg ttgaagaaat tgcaaaggag ttagaatcga aatatgaagt atcggttctt 180

cctcttaaat tggatgtttc tgatttgaag tctattcctg gggtaattga gtcattgcca 240

aaggaatttg ctgatatcga tgtcttgatt aataatgctg gacttgctct aggtaccgat 300

aaagtcattg atcttaatat tgatgacgcc gttaccatga ttactaccaa tgttcttggt 360

atgatggcta tgactcgtgc ggttcttcct atattctaca gcaaaaacaa gggtgatatt 420

ttgaacgttg gcagtattgc cggcagagaa tcatacgtag gcggctccgt ttactgctct 480

accaagtctg cccttgctca attcacttcc gctttgcgta aggagactat tgacactcgc 540

attcgtatta tggaggttga tcctggcttg gtcgaaaccg aattcagcgt tgtgagattc 600

cacggagaca aacaaaaggc tgataatgtt tacaaaaata gtgagccttt gacacccgaa 660

gacattgctg aggtgattct ttttgccctc actcgcagag aaaacgtcgt tattgccgat 720

acacttgttt tcccatccca tcaaggtggt gccaatcatg tgtacagaaa gcaagcgtag 780

<210> 731
<211> 259
<212> PRT
<213> Artificial Sequence

<220>
<223> Schizosaccharomyces pombe NM_001018495

<400> 731

Met Ser Arg Leu Asp Gly Lys Thr Ile Leu Ile Thr Gly Ala Ser Ser
1 5 10 15

Gly Ile Gly Lys Ser Thr Ala Phe Glu Ile Ala Lys Val Ala Lys Val
20 25 30

Lys Leu Ile Leu Ala Ala Arg Arg Phe Ser Thr Val Glu Glu Ile Ala
35 40 45

Lys Glu Leu Glu Ser Lys Tyr Glu Val Ser Val Leu Pro Leu Lys Leu
50 55 60

Asp Val Ser Asp Leu Lys Ser Ile Pro Gly Val Ile Glu Ser Leu Pro
65 70 75 80

Lys Glu Phe Ala Asp Ile Asp Val Leu Ile Asn Asn Ala Gly Leu Ala
85 90 95

Leu Gly Thr Asp Lys Val Ile Asp Leu Asn Ile Asp Asp Ala Val Thr
100 105 110

```
Met Ile Thr Thr Asn Val Leu Gly Met Met Ala Met Thr Arg Ala Val
        115                 120                 125

Leu Pro Ile Phe Tyr Ser Lys Asn Lys Gly Asp Ile Leu Asn Val Gly
    130                 135                 140

Ser Ile Ala Gly Arg Glu Ser Tyr Val Gly Gly Ser Val Tyr Cys Ser
145                 150                 155                 160

Thr Lys Ser Ala Leu Ala Gln Phe Thr Ser Ala Leu Arg Lys Glu Thr
                165                 170                 175

Ile Asp Thr Arg Ile Arg Ile Met Glu Val Asp Pro Gly Leu Val Glu
            180                 185                 190

Thr Glu Phe Ser Val Val Arg Phe His Gly Asp Lys Gln Lys Ala Asp
        195                 200                 205

Asn Val Tyr Lys Asn Ser Glu Pro Leu Thr Pro Glu Asp Ile Ala Glu
    210                 215                 220

Val Ile Leu Phe Ala Leu Thr Arg Arg Glu Asn Val Val Ile Ala Asp
225                 230                 235                 240

Thr Leu Val Phe Pro Ser His Gln Gly Gly Ala Asn His Val Tyr Arg
                245                 250                 255

Lys Gln Ala
```

<210> 732
<211> 1521
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD2 from S. cerevisiae

<400> 732

```
atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg       60

ctagggttgt ttatcaacaa tgagttttgt ccatcatcag atggaaagac catcgaaact      120

gtgaacccag ctactggcga accgataaca tccttccaag cagctaacga aaaggatgta      180

gacaaagctg tgaaagctgc cagggctgct tttgataacg tttggtcgaa gacatcttct      240

gagcaacgtg gtatttatct ttcaaactta ttaaaactta ttgaggagga gcaagacaca      300

cttgccgcat tagagacttt agacgctgga aagccttacc attcaaatgc caaaggtgat      360

ttggcacaaa ttttacagct taccagatat tttgctgggt ccgctgataa gtttgacaaa      420

ggtgcaacca taccattgac ttttaacaag tttgcatata ctctaaaagt tccttttggc      480

gttgttgctc aaatcgttcc atggaattat cctctagcta tggcttgttg gaaattgcaa      540

ggtgccttag cagccggtaa cacggttatc atcaaacctg ctgagaatac ctctctatct      600

ctactttatt ttgctacttt aattaaaaaa gcaggttttc cacctggtgt tgtcaatatc      660

gttcctggtt atggatcact tgtaggccaa gccctagcat ctcacatgga tatcgacaaa      720
```

atatctttta cgggaagcac caaggtcggt ggatttgtgt tggaagcttc cggccaatcg 780

aaccttaaag acgttacact agaatgcggt ggtaagtctc ctgctctcgt atttgaagat 840

gcagaccttg ataaggctat cgattggata gcagctggca ttttctacaa ttcaggacag 900

aattgtaccg caaactcaag agtttatgtt caaagttcga tctacgacaa gtttgttgaa 960

aagtttaaag aaactgcaaa gaaggagtgg gatgttgcag gaaaatttga tccgtttgat 1020

gagaaatgca tcgttggtcc agttatatca agtacacagt atgaccgcat caaaagttac 1080

atagaacgtg gtaaaaggga ggaaaagttg gacatgttcc agacctctga atttcctatt 1140

ggtggagcta aaggctactt cattccccca accatcttca ctgatgtccc gcaaacatcg 1200

aaactgttac aggatgagat atttggcccg gttgtggttg ttagcaagtt cacaaattat 1260

gatgacgctc tgaagctggc taatgatact tgctacgggc tcgcctctgc ggtcttcaca 1320

aaagatgtca agaaagcgca catgtttgct cgcgatatta aagcaggaac tgtttggatc 1380

aactcatcta acgatgaaga tgttaccgtt ccttttggcg ggtttaaaat gagtggtatt 1440

ggtagagaac tggggcaaag tggtgttgat acctatcttc aaacaaaagc agttcacata 1500

aatctctctt tggacaacta a 1521

<210> 733
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD2 from S. cerevisiae

<400> 733

Met Pro Thr Leu Tyr Thr Asp Ile Glu Ile Pro Gln Leu Lys Ile Ser
1 5 10 15

Leu Lys Gln Pro Leu Gly Leu Phe Ile Asn Asn Glu Phe Cys Pro Ser
20 25 30

Ser Asp Gly Lys Thr Ile Glu Thr Val Asn Pro Ala Thr Gly Glu Pro
35 40 45

Ile Thr Ser Phe Gln Ala Ala Asn Glu Lys Asp Val Asp Lys Ala Val
50 55 60

Lys Ala Ala Arg Ala Ala Phe Asp Asn Val Trp Ser Lys Thr Ser Ser
65 70 75 80

Glu Gln Arg Gly Ile Tyr Leu Ser Asn Leu Leu Lys Leu Ile Glu Glu
85 90 95

Glu Gln Asp Thr Leu Ala Ala Leu Glu Thr Leu Asp Ala Gly Lys Pro
100 105 110

Tyr His Ser Asn Ala Lys Gly Asp Leu Ala Gln Ile Leu Gln Leu Thr
115 120 125

```
Arg Tyr Phe Ala Gly Ser Ala Asp Lys Phe Asp Lys Gly Ala Thr Ile
    130                 135                 140

Pro Leu Thr Phe Asn Lys Phe Ala Tyr Thr Leu Lys Val Pro Phe Gly
145                 150                 155                 160

Val Val Ala Gln Ile Val Pro Trp Asn Tyr Pro Leu Ala Met Ala Cys
                165                 170                 175

Trp Lys Leu Gln Gly Ala Leu Ala Ala Gly Asn Thr Val Ile Ile Lys
            180                 185                 190

Pro Ala Glu Asn Thr Ser Leu Ser Leu Leu Tyr Phe Ala Thr Leu Ile
        195                 200                 205

Lys Lys Ala Gly Phe Pro Pro Gly Val Val Asn Ile Val Pro Gly Tyr
    210                 215                 220

Gly Ser Leu Val Gly Gln Ala Leu Ala Ser His Met Asp Ile Asp Lys
225                 230                 235                 240

Ile Ser Phe Thr Gly Ser Thr Lys Val Gly Gly Phe Val Leu Glu Ala
                245                 250                 255

Ser Gly Gln Ser Asn Leu Lys Asp Val Thr Leu Glu Cys Gly Gly Lys
            260                 265                 270

Ser Pro Ala Leu Val Phe Glu Asp Ala Asp Leu Asp Lys Ala Ile Asp
        275                 280                 285

Trp Ile Ala Ala Gly Ile Phe Tyr Asn Ser Gly Gln Asn Cys Thr Ala
    290                 295                 300

Asn Ser Arg Val Tyr Val Gln Ser Ser Ile Tyr Asp Lys Phe Val Glu
305                 310                 315                 320

Lys Phe Lys Glu Thr Ala Lys Lys Glu Trp Asp Val Ala Gly Lys Phe
                325                 330                 335

Asp Pro Phe Asp Glu Lys Cys Ile Val Gly Pro Val Ile Ser Ser Thr
            340                 345                 350

Gln Tyr Asp Arg Ile Lys Ser Tyr Ile Glu Arg Gly Lys Arg Glu Glu
        355                 360                 365

Lys Leu Asp Met Phe Gln Thr Ser Glu Phe Pro Ile Gly Gly Ala Lys
    370                 375                 380

Gly Tyr Phe Ile Pro Pro Thr Ile Phe Thr Asp Val Pro Gln Thr Ser
385                 390                 395                 400

Lys Leu Leu Gln Asp Glu Ile Phe Gly Pro Val Val Val Val Ser Lys
                405                 410                 415
```

```
Phe Thr Asn Tyr Asp Asp Ala Leu Lys Leu Ala Asn Asp Thr Cys Tyr
            420                 425                 430

Gly Leu Ala Ser Ala Val Phe Thr Lys Asp Val Lys Lys Ala His Met
        435                 440                 445

Phe Ala Arg Asp Ile Lys Ala Gly Thr Val Trp Ile Asn Ser Ser Asn
    450                 455                 460

Asp Glu Asp Val Thr Val Pro Phe Gly Gly Phe Lys Met Ser Gly Ile
465                 470                 475                 480

Gly Arg Glu Leu Gly Gln Ser Gly Val Asp Thr Tyr Leu Gln Thr Lys
                485                 490                 495

Ala Val His Ile Asn Leu Ser Leu Asp Asn
            500                 505
```

<210> 734
<211> 1521
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD3 from S. cerevisiae

<400> 734

```
atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg       60

ctagggttgt ttatcaacaa tgagttttgt ccatcatcag atggaaagac catcgaaact      120

gtgaacccag ctactggcga accgataaca tccttccaag cagctaacga aaaggatgta      180

gacaaagctg tgaaagctgc cagggctgct tttgataacg tttggtcgaa gacatcttct      240

gagcaacgtg gtatttatct ttcaaactta ttaaaactta ttgaggagga gcaagacaca      300

cttgccgcat tagagacttt agacgctggt aagcctttcc attccaatgc taaacaagac      360

ttagcccaga ttatagaact tacaagatac tatgcggggg cggtcgacaa gttcaatatg      420

ggtgaaacca ttccattgac ttttaacaag tttgcatata ctctaaaagt tccttttggc      480

gttgttgctc aaatcgttcc atggaattat cctctagcta tggcttgtag aaaaatgcaa      540

ggtgccttag cggccggtaa cacggttatc atcaaacctg ctgaaaatac ctctctatct      600

ctactttatt ttgctacttt aattaaaaaa gcaggttttc cacctggtgt tgtcaatgtc      660

attcctggtt atggttccgt tgtggggaaa gctttaggaa cccacatgga tatcgacaaa      720

atatctttta cgggaagtac taaggttggc ggctcagtat tggaagcttc cggccaatcg      780

aaccttaagg atatcacact agaatgcggt ggtaagtctc ctgctcttgt atttgaagat      840

gcagaccttg ataaggctat agaatgggta gcaaatggta tttttttaa ttcgggacag       900

atctgcactg caaactcaag agtttatgtt caaagttcga tctacgacaa gtttgttgaa      960

aagtttaaag aaactgcaaa gaaggagtgg gatgttgcag gaaaatttga tccgtttgat     1020

gagaaatgca tcgttggtcc agttatatca agtacacagt atgaccgcat caaaagttac     1080

atagaacgtg gtaaaaagga ggaaaagttg gacatgttcc agacctctga atttcctatt     1140

ggtggagcta aaggctactt cattccccca accatcttca ctgatgtacc agaaacatct     1200
```

aagttgctgc gtgatgaaat atttggcccg gttgtggttg ttagcaagtt cacaaattat 1260

gatgacgctc tgaagctggc taatgatact tgctacgggc tcgcctctgc ggtcttcacc 1320

aaagatgtca agaaagcgca catgtttgct cgcgatatta aagcaggaac tgtttggatc 1380

aatcaaacca atcaagaaga agctaaagtt ccttttggcg gatttaagat gagtggtatt 1440

ggtagagaat caggcgacac cggcgttgat aactatttac aaataaaatc agtccatgtg 1500

gatctttcat tggataaata a 1521

<210> 735
<211> 506
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD3 from S. cerevisiae

<400> 735

Met Pro Thr Leu Tyr Thr Asp Ile Glu Ile Pro Gln Leu Lys Ile Ser
1 5 10 15

Leu Lys Gln Pro Leu Gly Leu Phe Ile Asn Asn Glu Phe Cys Pro Ser
20 25 30

Ser Asp Gly Lys Thr Ile Glu Thr Val Asn Pro Ala Thr Gly Glu Pro
35 40 45

Ile Thr Ser Phe Gln Ala Ala Asn Glu Lys Asp Val Asp Lys Ala Val
50 55 60

Lys Ala Ala Arg Ala Ala Phe Asp Asn Val Trp Ser Lys Thr Ser Ser
65 70 75 80

Glu Gln Arg Gly Ile Tyr Leu Ser Asn Leu Leu Lys Leu Ile Glu Glu
85 90 95

Glu Gln Asp Thr Leu Ala Ala Leu Glu Thr Leu Asp Ala Gly Lys Pro
100 105 110

Phe His Ser Asn Ala Lys Gln Asp Leu Ala Gln Ile Ile Glu Leu Thr
115 120 125

Arg Tyr Tyr Ala Gly Ala Val Asp Lys Phe Asn Met Gly Glu Thr Ile
130 135 140

Pro Leu Thr Phe Asn Lys Phe Ala Tyr Thr Leu Lys Val Pro Phe Gly
145 150 155 160

Val Val Ala Gln Ile Val Pro Trp Asn Tyr Pro Leu Ala Met Ala Cys
165 170 175

Arg Lys Met Gln Gly Ala Leu Ala Ala Gly Asn Thr Val Ile Ile Lys
180 185 190

```
Pro Ala Glu Asn Thr Ser Leu Ser Leu Leu Tyr Phe Ala Thr Leu Ile
        195                 200                 205

Lys Lys Ala Gly Phe Pro Pro Gly Val Val Asn Val Ile Pro Gly Tyr
    210                 215                 220

Gly Ser Val Val Gly Lys Ala Leu Gly Thr His Met Asp Ile Asp Lys
225                 230                 235                 240

Ile Ser Phe Thr Gly Ser Thr Lys Val Gly Gly Ser Val Leu Glu Ala
                245                 250                 255

Ser Gly Gln Ser Asn Leu Lys Asp Ile Thr Leu Glu Cys Gly Gly Lys
            260                 265                 270

Ser Pro Ala Leu Val Phe Glu Asp Ala Asp Leu Asp Lys Ala Ile Glu
        275                 280                 285

Trp Val Ala Asn Gly Ile Phe Phe Asn Ser Gly Gln Ile Cys Thr Ala
    290                 295                 300

Asn Ser Arg Val Tyr Val Gln Ser Ser Ile Tyr Asp Lys Phe Val Glu
305                 310                 315                 320

Lys Phe Lys Glu Thr Ala Lys Lys Glu Trp Asp Val Ala Gly Lys Phe
                325                 330                 335

Asp Pro Phe Asp Glu Lys Cys Ile Val Gly Pro Val Ile Ser Ser Thr
            340                 345                 350

Gln Tyr Asp Arg Ile Lys Ser Tyr Ile Glu Arg Gly Lys Lys Glu Glu
        355                 360                 365

Lys Leu Asp Met Phe Gln Thr Ser Glu Phe Pro Ile Gly Gly Ala Lys
    370                 375                 380

Gly Tyr Phe Ile Pro Pro Thr Ile Phe Thr Asp Val Pro Glu Thr Ser
385                 390                 395                 400

Lys Leu Leu Arg Asp Glu Ile Phe Gly Pro Val Val Val Val Ser Lys
                405                 410                 415

Phe Thr Asn Tyr Asp Asp Ala Leu Lys Leu Ala Asn Asp Thr Cys Tyr
            420                 425                 430

Gly Leu Ala Ser Ala Val Phe Thr Lys Asp Val Lys Lys Ala His Met
        435                 440                 445

Phe Ala Arg Asp Ile Lys Ala Gly Thr Val Trp Ile Asn Gln Thr Asn
    450                 455                 460

Gln Glu Glu Ala Lys Val Pro Phe Gly Gly Phe Lys Met Ser Gly Ile
465                 470                 475                 480
```

20172000439 20 Jan 2017

Gly Arg Glu Ser Gly Asp Thr Gly Val Asp Asn Tyr Leu Gln Ile Lys
485 490 495

Ser Val His Val Asp Leu Ser Leu Asp Lys
500 505

<210> 736
<211> 1560
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD4 from S. cerevisiae

<400> 736
atgttcagta gatctacgct ctgcttaaag acgtctgcat cctccattgg gagacttcaa 60

ttgagatatt tctcacacct tcctatgaca gtgcctatca agctgcccaa tgggttggaa 120

tatgagcaac caacggggtt gttcatcaac aacaagtttg ttccttctaa acagaacaag 180

accttcgaag tcattaaccc ttccacggaa gaagaaatat gtcatattta tgaaggtaga 240

gaggacgatg tggaagaggc cgtgcaggcc gccgaccgtg ccttctctaa tgggtcttgg 300

aacggtatcg accctattga caggggtaag gctttgtaca ggttagccga attaattgaa 360

caggacaagg atgtcattgc ttccatcgag actttggata acggtaaagc tatctcttcc 420

tcgagaggag atgttgattt agtcatcaac tatttgaaat cttctgctgg ctttgctgat 480

aaaattgatg gtagaatgat tgatactggt agaacccatt tttcttacac taagagacag 540

cctttgggtg tttgtgggca gattattcct tggaatttcc cactgttgat gtgggcctgg 600

aagattgccc ctgctttggt caccggtaac accgtcgtgt tgaagactgc cgaatccacc 660

ccattgtccg ctttgtatgt gtctaaatac atcccacagg cgggtattcc acctggtgtg 720

atcaacattg tatccgggtt tggtaagatt gtgggtgagg ccattacaaa ccatccaaaa 780

atcaaaaagg ttgccttcac agggtccacg gctacgggta gacacattta ccagtccgca 840

gccgcaggct tgaaaaaagt gactttggag ctgggtggta aatcaccaaa cattgtcttc 900

gcggacgccg agttgaaaaa agccgtgcaa aacattatcc ttggtatcta ctacaattct 960

ggtgaggtct gttgtgcggg ttcaagggtg tatgttgaag aatctattta cgacaaattc 1020

attgaagagt tcaaagccgc ttctgaatcc atcaaggtgg gcgacccatt cgatgaatct 1080

actttccaag gtgcacaaac ctctcaaatg caactaaaca aaatcttgaa atacgttgac 1140

attggtaaga atgaaggtgc tactttgatt accggtggtg aaagattagg tagcaagggt 1200

tacttcatta agccaactgt ctttggtgac gttaaggaag acatgagaat tgtcaaagag 1260

gaaatctttg gccctgttgt cactgtaacc aaattcaaat ctgccgacga agtcattaac 1320

atggcgaacg attctgaata cgggttggct gctggtattc acacctctaa tattaatacc 1380

gccttaaaag tggctgatag agttaatgcg ggtacggtct ggataaacac ttataacgat 1440

ttccaccacg cagttccttt cggtgggttc aatgcatctg gtttgggcag ggaaatgtct 1500

gttgatgctt tacaaaacta cttgcaagtt aaagcggtcc gtgccaaatt ggacgagtaa 1560

<210> 737
<211> 519

<212> PRT
<213> Artificial Sequence

<220>
<223> ALD4 from S. cerevisiae

<400> 737

```
Met Phe Ser Arg Ser Thr Leu Cys Leu Lys Thr Ser Ala Ser Ser Ile
1               5                  10                  15

Gly Arg Leu Gln Leu Arg Tyr Phe Ser His Leu Pro Met Thr Val Pro
            20                  25                  30

Ile Lys Leu Pro Asn Gly Leu Glu Tyr Glu Gln Pro Thr Gly Leu Phe
        35                  40                  45

Ile Asn Asn Lys Phe Val Pro Ser Lys Gln Asn Lys Thr Phe Glu Val
    50                  55                  60

Ile Asn Pro Ser Thr Glu Glu Glu Ile Cys His Ile Tyr Glu Gly Arg
65                  70                  75                  80

Glu Asp Asp Val Glu Glu Ala Val Gln Ala Ala Asp Arg Ala Phe Ser
                85                  90                  95

Asn Gly Ser Trp Asn Gly Ile Asp Pro Ile Asp Arg Gly Lys Ala Leu
            100                 105                 110

Tyr Arg Leu Ala Glu Leu Ile Glu Gln Asp Lys Asp Val Ile Ala Ser
        115                 120                 125

Ile Glu Thr Leu Asp Asn Gly Lys Ala Ile Ser Ser Ser Arg Gly Asp
    130                 135                 140

Val Asp Leu Val Ile Asn Tyr Leu Lys Ser Ser Ala Gly Phe Ala Asp
145                 150                 155                 160

Lys Ile Asp Gly Arg Met Ile Asp Thr Gly Arg Thr His Phe Ser Tyr
                165                 170                 175

Thr Lys Arg Gln Pro Leu Gly Val Cys Gly Gln Ile Ile Pro Trp Asn
            180                 185                 190

Phe Pro Leu Leu Met Trp Ala Trp Lys Ile Ala Pro Ala Leu Val Thr
        195                 200                 205

Gly Asn Thr Val Val Leu Lys Thr Ala Glu Ser Thr Pro Leu Ser Ala
    210                 215                 220

Leu Tyr Val Ser Lys Tyr Ile Pro Gln Ala Gly Ile Pro Pro Gly Val
225                 230                 235                 240

Ile Asn Ile Val Ser Gly Phe Gly Lys Ile Val Gly Glu Ala Ile Thr
                245                 250                 255
```

```
Asn His Pro Lys Ile Lys Lys Val Ala Phe Thr Gly Ser Thr Ala Thr
            260                 265                 270

Gly Arg His Ile Tyr Gln Ser Ala Ala Ala Gly Leu Lys Lys Val Thr
        275                 280                 285

Leu Glu Leu Gly Gly Lys Ser Pro Asn Ile Val Phe Ala Asp Ala Glu
    290                 295                 300

Leu Lys Lys Ala Val Gln Asn Ile Ile Leu Gly Ile Tyr Tyr Asn Ser
305                 310                 315                 320

Gly Glu Val Cys Cys Ala Gly Ser Arg Val Tyr Val Glu Glu Ser Ile
                325                 330                 335

Tyr Asp Lys Phe Ile Glu Glu Phe Lys Ala Ala Ser Glu Ser Ile Lys
            340                 345                 350

Val Gly Asp Pro Phe Asp Glu Ser Thr Phe Gln Gly Ala Gln Thr Ser
        355                 360                 365

Gln Met Gln Leu Asn Lys Ile Leu Lys Tyr Val Asp Ile Gly Lys Asn
    370                 375                 380

Glu Gly Ala Thr Leu Ile Thr Gly Gly Glu Arg Leu Gly Ser Lys Gly
385                 390                 395                 400

Tyr Phe Ile Lys Pro Thr Val Phe Gly Asp Val Lys Glu Asp Met Arg
                405                 410                 415

Ile Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Thr Lys Phe
            420                 425                 430

Lys Ser Ala Asp Glu Val Ile Asn Met Ala Asn Asp Ser Glu Tyr Gly
        435                 440                 445

Leu Ala Ala Gly Ile His Thr Ser Asn Ile Asn Thr Ala Leu Lys Val
    450                 455                 460

Ala Asp Arg Val Asn Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp
465                 470                 475                 480

Phe His His Ala Val Pro Phe Gly Gly Phe Asn Ala Ser Gly Leu Gly
                485                 490                 495

Arg Glu Met Ser Val Asp Ala Leu Gln Asn Tyr Leu Gln Val Lys Ala
            500                 505                 510

Val Arg Ala Lys Leu Asp Glu
        515
```

<210> 738
<211> 1563
<212> DNA
<213> Artificial Sequence

20 Jan 2017 2017200439

<220>
<223> ALD5 from S. cerevisiae

<400> 738
atgctttctc gcacaagagc tgcagctccg aattccagaa tattcactag aagcttgtta 60

cgtctttatt ctcaagcacc attacgcgtt ccaattactc ttccaaatgg tttcacctac 120

gaacagccaa cagggttatt catcaatggt gaatttgttg cctcgaagca aaagaaaacg 180

tttgacgtga tcaatccatc taacgaagaa aagataacaa ctgtatacaa ggctatggaa 240

gatgatgttg atgaagccgt tgcagcggct aaaaaagctt ttgaaacgaa gtggtctatt 300

gtagagccgg aggttcgcgc taaagcttta ttcaatctcg ctgacttggt tgagaaacac 360

caagaaacac tggctgccat tgagtcaatg gataatggta agtcattgtt ttgtgcgcgc 420

ggtgacgtcg ctttagtatc taaatacttg cgttcttgcg gtggttgggc agataaaatc 480

tacggtaacg ttattgacac aggtaaaaac cattttacct actcaattaa ggaaccatta 540

ggcgtttgcg gccaaataat cccttggaac ttccctttat tgatgtggtc atggaaaatt 600

gggcctgctc tggctacagg taacaccgtc gtattgaaac ccgctgaaac aacaccttta 660

tctgcccttt tcgcttccca gttgtgtcag gaagcaggca tacccgctgg tgtagtcaat 720

atccttccgg gttccggtag agttgttgga gaaagattga gtgcacaccc agacgtgaag 780

aagattgctt ttacaggctc tactgccacc ggccgccata ttatgaaggt cgctgccgat 840

actgtcaaga aagtcacttt ggagctggga ggtaaatcac caaatattgt gtttgctgac 900

gctgatctag ataaagccgt caagaacatt gccttcggta ttttttacaa ctctggtgaa 960

gtttgctgcg ctggttccag aatatacatt caagatacag tatacgagga ggtgttgcaa 1020

aaactaaagg attacaccga gtcactaaag gtcggtgacc catttgatga ggaagttttc 1080

caaggtgctc aaacatctga caaacagctg cataaaattt tagactatgt cgatgtagca 1140

aaatcagagg gggctcgtct tgtgactgga ggggccagac atggcagtaa aggttatttt 1200

gtcaagccaa cagtgtttgc tgatgtcaaa ggagatatga gaattgttaa ggaggaagtg 1260

tttggtccca ttgtaactgt atccaagttt tctactgttg atgaagtgat tgctatggca 1320

aatgattctc aatatgggtt agccgcaggt attcacacta acgatattaa caaggctgtt 1380

gatgtgtcca aaagagtgaa agctggtact gtttggataa atacctataa caacttccac 1440

caaaatgttc ctttcggtgg cttcggccag tcaggtattg gccgtgaaat gggtgaggct 1500

gctttaagta actacactca aacaaaatct gtcagaattg ccattgacaa gccaattcgt 1560

tga 1563

<210> 739
<211> 520
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD5 from S. cerevisiae

<400> 739

Met Leu Ser Arg Thr Arg Ala Ala Ala Pro Asn Ser Arg Ile Phe Thr
1 5 10 15

```
Arg Ser Leu Leu Arg Leu Tyr Ser Gln Ala Pro Leu Arg Val Pro Ile
            20                  25                  30

Thr Leu Pro Asn Gly Phe Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile
        35                  40                  45

Asn Gly Glu Phe Val Ala Ser Lys Gln Lys Lys Thr Phe Asp Val Ile
    50                  55                  60

Asn Pro Ser Asn Glu Glu Lys Ile Thr Thr Val Tyr Lys Ala Met Glu
65                  70                  75                  80

Asp Asp Val Asp Glu Ala Val Ala Ala Ala Lys Lys Ala Phe Glu Thr
                85                  90                  95

Lys Trp Ser Ile Val Glu Pro Glu Val Arg Ala Lys Ala Leu Phe Asn
            100                 105                 110

Leu Ala Asp Leu Val Glu Lys His Gln Glu Thr Leu Ala Ala Ile Glu
        115                 120                 125

Ser Met Asp Asn Gly Lys Ser Leu Phe Cys Ala Arg Gly Asp Val Ala
    130                 135                 140

Leu Val Ser Lys Tyr Leu Arg Ser Cys Gly Gly Trp Ala Asp Lys Ile
145                 150                 155                 160

Tyr Gly Asn Val Ile Asp Thr Gly Lys Asn His Phe Thr Tyr Ser Ile
                165                 170                 175

Lys Glu Pro Leu Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro
            180                 185                 190

Leu Leu Met Trp Ser Trp Lys Ile Gly Pro Ala Leu Ala Thr Gly Asn
        195                 200                 205

Thr Val Val Leu Lys Pro Ala Glu Thr Thr Pro Leu Ser Ala Leu Phe
    210                 215                 220

Ala Ser Gln Leu Cys Gln Glu Ala Gly Ile Pro Ala Gly Val Val Asn
225                 230                 235                 240

Ile Leu Pro Gly Ser Gly Arg Val Val Gly Glu Arg Leu Ser Ala His
                245                 250                 255

Pro Asp Val Lys Lys Ile Ala Phe Thr Gly Ser Thr Ala Thr Gly Arg
            260                 265                 270

His Ile Met Lys Val Ala Ala Asp Thr Val Lys Lys Val Thr Leu Glu
        275                 280                 285

Leu Gly Gly Lys Ser Pro Asn Ile Val Phe Ala Asp Ala Asp Leu Asp
    290                 295                 300
```

```
Lys Ala Val Lys Asn Ile Ala Phe Gly Ile Phe Tyr Asn Ser Gly Glu
305                 310                 315                 320

Val Cys Cys Ala Gly Ser Arg Ile Tyr Ile Gln Asp Thr Val Tyr Glu
                325                 330                 335

Glu Val Leu Gln Lys Leu Lys Asp Tyr Thr Glu Ser Leu Lys Val Gly
            340                 345                 350

Asp Pro Phe Asp Glu Glu Val Phe Gln Gly Ala Gln Thr Ser Asp Lys
        355                 360                 365

Gln Leu His Lys Ile Leu Asp Tyr Val Asp Val Ala Lys Ser Glu Gly
    370                 375                 380

Ala Arg Leu Val Thr Gly Gly Ala Arg His Gly Ser Lys Gly Tyr Phe
385                 390                 395                 400

Val Lys Pro Thr Val Phe Ala Asp Val Lys Gly Asp Met Arg Ile Val
                405                 410                 415

Lys Glu Glu Val Phe Gly Pro Ile Val Thr Val Ser Lys Phe Ser Thr
            420                 425                 430

Val Asp Glu Val Ile Ala Met Ala Asn Asp Ser Gln Tyr Gly Leu Ala
        435                 440                 445

Ala Gly Ile His Thr Asn Asp Ile Asn Lys Ala Val Asp Val Ser Lys
    450                 455                 460

Arg Val Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asn Phe His
465                 470                 475                 480

Gln Asn Val Pro Phe Gly Gly Phe Gly Gln Ser Gly Ile Gly Arg Glu
                485                 490                 495

Met Gly Glu Ala Ala Leu Ser Asn Tyr Thr Gln Thr Lys Ser Val Arg
            500                 505                 510

Ile Ala Ile Asp Lys Pro Ile Arg
        515                 520


<210>  740
<211>  1503
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  ALD6 from S. cerevisiae

<400>  740
atgactaagc tacactttga cactgctgaa ccagtcaaga tcacacttcc aaatggtttg       60

acatacgagc aaccaaccgg tctattcatt aacaacaagt ttatgaaagc tcaagacggt      120

aagacctatc ccgtcgaaga tccttccact gaaaaacaccg tttgtgaggt ctcttctgcc      180
```

accactgaag atgttgaata tgctatcgaa tgtgccgacc gtgctttcca cgacactgaa 240

tgggctaccc aagacccaag agaaagaggc cgtctactaa gtaagttggc tgacgaattg 300

gaaagccaaa ttgacttggt ttcttccatt gaagctttgg acaatggtaa aactttggcc 360

ttagcccgtg gggatgttac cattgcaatc aactgtctaa gagatgctgc tgcctatgcc 420

gacaaagtca acggtagaac aatcaacacc ggtgacggct acatgaactt caccacctta 480

gagccaatcg gtgtctgtgg tcaaattatt ccatggaact ttccaataat gatgttggct 540

tggaagatcg ccccagcatt ggccatgggt aacgtctgta tcttgaaacc cgctgctgtc 600

acacctttaa atgccctata ctttgcttct ttatgtaaga aggttggtat tccagctggt 660

gtcgtcaaca tcgttccagg tcctggtaga actgttggtg ctgctttgac caacgaccca 720

agaatcagaa agctggcttt taccggttct acagaagtcg gtaagagtgt tgctgtcgac 780

tcttctgaat ctaacttgaa gaaaatcact ttggaactag gtggtaagtc cgcccatttg 840

gtctttgacg atgctaacat taagaagact ttaccaaatc tagtaaacgg tattttcaag 900

aacgctggtc aaatttgttc ctctggttct agaatttacg ttcaagaagg tatttacgac 960

gaactattgg ctgctttcaa ggcttacttg gaaaccgaaa tcaaagttgg taatccattt 1020

gacaaggcta acttccaagg tgctatcact aaccgtcaac aattcgacac aattatgaac 1080

tacatcgata tcggtaagaa agaaggcgcc aagatcttaa ctggtggcga aaaagttggt 1140

gacaagggtt acttcatcag accaaccgtt ttctacgatg ttaatgaaga catgagaatt 1200

gttaaggaag aaatttttgg accagttgtc actgtcgcaa agttcaagac tttagaagaa 1260

ggtgtcgaaa tggctaacag ctctgaattc ggtctaggtt ctggtatcga aacagaatct 1320

ttgagcacag gtttgaaggt ggccaagatg ttgaaggccg gtaccgtctg gatcaacaca 1380

tacaacgatt ttgactccag agttccattc ggtggtgtta agcaatctgg ttacggtaga 1440

gaaatgggtg aagaagtcta ccatgcatac actgaagtaa aagctgtcag aattaagttg 1500

taa 1503

<210> 741
<211> 500
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD6 from S. cerevisiae

<400> 741

Met Thr Lys Leu His Phe Asp Thr Ala Glu Pro Val Lys Ile Thr Leu
1 5 10 15

Pro Asn Gly Leu Thr Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn
20 25 30

Lys Phe Met Lys Ala Gln Asp Gly Lys Thr Tyr Pro Val Glu Asp Pro
35 40 45

Ser Thr Glu Asn Thr Val Cys Glu Val Ser Ser Ala Thr Thr Glu Asp
50 55 60

```
Val Glu Tyr Ala Ile Glu Cys Ala Asp Arg Ala Phe His Asp Thr Glu
65                  70                  75                  80

Trp Ala Thr Gln Asp Pro Arg Glu Arg Gly Arg Leu Leu Ser Lys Leu
                85                  90                  95

Ala Asp Glu Leu Glu Ser Gln Ile Asp Leu Val Ser Ser Ile Glu Ala
            100                 105                 110

Leu Asp Asn Gly Lys Thr Leu Ala Leu Ala Arg Gly Asp Val Thr Ile
        115                 120                 125

Ala Ile Asn Cys Leu Arg Asp Ala Ala Ala Tyr Ala Asp Lys Val Asn
    130                 135                 140

Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr Met Asn Phe Thr Thr Leu
145                 150                 155                 160

Glu Pro Ile Gly Val Cys Gly Gln Ile Ile Pro Trp Asn Phe Pro Ile
                165                 170                 175

Met Met Leu Ala Trp Lys Ile Ala Pro Ala Leu Ala Met Gly Asn Val
            180                 185                 190

Cys Ile Leu Lys Pro Ala Ala Val Thr Pro Leu Asn Ala Leu Tyr Phe
        195                 200                 205

Ala Ser Leu Cys Lys Lys Val Gly Ile Pro Ala Gly Val Val Asn Ile
    210                 215                 220

Val Pro Gly Pro Gly Arg Thr Val Gly Ala Ala Leu Thr Asn Asp Pro
225                 230                 235                 240

Arg Ile Arg Lys Leu Ala Phe Thr Gly Ser Thr Glu Val Gly Lys Ser
                245                 250                 255

Val Ala Val Asp Ser Ser Glu Ser Asn Leu Lys Lys Ile Thr Leu Glu
            260                 265                 270

Leu Gly Gly Lys Ser Ala His Leu Val Phe Asp Asp Ala Asn Ile Lys
        275                 280                 285

Lys Thr Leu Pro Asn Leu Val Asn Gly Ile Phe Lys Asn Ala Gly Gln
    290                 295                 300

Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp
305                 310                 315                 320

Glu Leu Leu Ala Ala Phe Lys Ala Tyr Leu Glu Thr Glu Ile Lys Val
                325                 330                 335

Gly Asn Pro Phe Asp Lys Ala Asn Phe Gln Gly Ala Ile Thr Asn Arg
            340                 345                 350
```

```
Gln Gln Phe Asp Thr Ile Met Asn Tyr Ile Asp Ile Gly Lys Lys Glu
        355                 360                 365

Gly Ala Lys Ile Leu Thr Gly Gly Glu Lys Val Gly Asp Lys Gly Tyr
    370                 375                 380

Phe Ile Arg Pro Thr Val Phe Tyr Asp Val Asn Glu Asp Met Arg Ile
385                 390                 395                 400

Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr Val Ala Lys Phe Lys
                405                 410                 415

Thr Leu Glu Glu Gly Val Glu Met Ala Asn Ser Ser Glu Phe Gly Leu
            420                 425                 430

Gly Ser Gly Ile Glu Thr Glu Ser Leu Ser Thr Gly Leu Lys Val Ala
        435                 440                 445

Lys Met Leu Lys Ala Gly Thr Val Trp Ile Asn Thr Tyr Asn Asp Phe
    450                 455                 460

Asp Ser Arg Val Pro Phe Gly Gly Val Lys Gln Ser Gly Tyr Gly Arg
465                 470                 475                 480

Glu Met Gly Glu Glu Val Tyr His Ala Tyr Thr Glu Val Lys Ala Val
                485                 490                 495

Arg Ile Lys Leu
            500
```

<210> 742
<211> 1569
<212> DNA
<213> Artificial Sequence

<220>
<223> KLLA0F00440 from Kluyveromyces lactis

<400> 742

```
atgttattac gtactaccac tgctggtact agactagctt tcaagtcaca aactgcggca       60

ttcgcacatg ctttacgtta ctattcgcat ttaccattga aggtcaaggt tgatttgcca      120

aatggtattt catacgaaca accaactggt ttgttcatca ataaccaatt tgttccttcc      180

aagcaaggga agacttttga agtgttgagt ccttcaaccg aagaagaaat tactcatgtt      240

tacgaaggtc gggaggaaga cgtggatgcg gctgttgagg ctgcagacaa agctttcaag      300

tccggtgtct ggtctaaggt tgacccacta gtcagatctc agtgtttgag caagttagct      360

gattacattg aaaaggattt ggataccatt gcttctatcg aatctctgga caatggtaaa      420

gctatctcat ctgctaaagg tgatgttcaa ctagttgtta actacctaag atcagctgct      480

ggctacgctg ataaggtgga tggtagagtc attgatactg gttcgactca tttctcttac      540

acagttagag aaccattagg tgtctgtggt caaattatcc catggaactt cccattgttg      600

atgtgggctt ggaagattgc tccagctctt gctaccggta atactgttgt cttgaagacc      660
```

gctgaatcca cacctttgtc tgcactatac gtttctcaat acattccaaa agcgggtatc 720

ccagacggtg ttgtgaacat cgtttctggt ttcggtaaaa ttgttggtga ggctattaca 780

actcatccaa agattaagaa ggttgcattc actggttcca ctgctactgg tcgtcatatc 840

taccaaaatg ccggtgtcaa tttgaagaag gtcacattgg agctaggtgg taagtcacca 900

aacgttgttt tcgctgatgc aaacttgaaa tccgctgttc aaaactcttt gttgggtatc 960

tactataact ccggtgaagt ctgctgtgct ggttccaggt tgtatgttga agaatctatt 1020

tatgacgatt tcttgaagga aattaagtct gcttcagaag ctctaaaggt tggtgatcct 1080

ttcgatgcag aaactttcca gggtgctcaa acttcgcaaa atcaattgga caagatcttg 1140

aaatatgtgg agattggtaa gaatgaaggt gctactttaa ttactggtgg tgaaagaatt 1200

ggtgacaagg gttactttgt tagaccaact attttcgctg atgttgaaga aagcatgaga 1260

atcgttaagg aagaaatctt cggccctgtt gtcactgtca ccaagtttaa aaccattgat 1320

gaattgattg aaaaggccaa cgactctgaa tatggattgg ctgcaggaat tcacacatct 1380

aacatcaata ctgccatcca ggtttcaaac agattgaatg ctggtaccgt ctggattaac 1440

acatacaacg acttccacca ctgtgttccg ttcggtggtt tcaatgcttc tggtattggt 1500

agagaaatgg gttctgaagc tctagacaat tacactcaag tcaaggctgt ccgtgttaaa 1560

ttggagtaa 1569

<210> 743
<211> 522
<212> PRT
<213> Artificial Sequence

<220>
<223> KLLA0F00440 from Kluyveromyces lactis

<400> 743

Met Leu Leu Arg Thr Thr Thr Ala Gly Thr Arg Leu Ala Phe Lys Ser
1 5 10 15

Gln Thr Ala Ala Phe Ala His Ala Leu Arg Tyr Tyr Ser His Leu Pro
20 25 30

Leu Lys Val Lys Val Asp Leu Pro Asn Gly Ile Ser Tyr Glu Gln Pro
35 40 45

Thr Gly Leu Phe Ile Asn Asn Gln Phe Val Pro Ser Lys Gln Gly Lys
50 55 60

Thr Phe Glu Val Leu Ser Pro Ser Thr Glu Glu Glu Ile Thr His Val
65 70 75 80

Tyr Glu Gly Arg Glu Glu Asp Val Asp Ala Ala Val Glu Ala Ala Asp
85 90 95

Lys Ala Phe Lys Ser Gly Val Trp Ser Lys Val Asp Pro Leu Val Arg
100 105 110

```
Ser Gln Cys Leu Ser Lys Leu Ala Asp Tyr Ile Glu Lys Asp Leu Asp
        115                 120                 125

Thr Ile Ala Ser Ile Glu Ser Leu Asp Asn Gly Lys Ala Ile Ser Ser
    130                 135                 140

Ala Lys Gly Asp Val Gln Leu Val Val Asn Tyr Leu Arg Ser Ala Ala
145                 150                 155                 160

Gly Tyr Ala Asp Lys Val Asp Gly Arg Val Ile Asp Thr Gly Ser Thr
                165                 170                 175

His Phe Ser Tyr Thr Val Arg Glu Pro Leu Gly Val Cys Gly Gln Ile
            180                 185                 190

Ile Pro Trp Asn Phe Pro Leu Leu Met Trp Ala Trp Lys Ile Ala Pro
        195                 200                 205

Ala Leu Ala Thr Gly Asn Thr Val Val Leu Lys Thr Ala Glu Ser Thr
    210                 215                 220

Pro Leu Ser Ala Leu Tyr Val Ser Gln Tyr Ile Pro Lys Ala Gly Ile
225                 230                 235                 240

Pro Asp Gly Val Val Asn Ile Val Ser Gly Phe Gly Lys Ile Val Gly
                245                 250                 255

Glu Ala Ile Thr Thr His Pro Lys Ile Lys Lys Val Ala Phe Thr Gly
            260                 265                 270

Ser Thr Ala Thr Gly Arg His Ile Tyr Gln Asn Ala Gly Val Asn Leu
        275                 280                 285

Lys Lys Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Asn Val Val Phe
    290                 295                 300

Ala Asp Ala Asn Leu Lys Ser Ala Val Gln Asn Ser Leu Leu Gly Ile
305                 310                 315                 320

Tyr Tyr Asn Ser Gly Glu Val Cys Cys Ala Gly Ser Arg Leu Tyr Val
                325                 330                 335

Glu Glu Ser Ile Tyr Asp Asp Phe Leu Lys Glu Ile Lys Ser Ala Ser
            340                 345                 350

Glu Ala Leu Lys Val Gly Asp Pro Phe Asp Ala Glu Thr Phe Gln Gly
        355                 360                 365

Ala Gln Thr Ser Gln Asn Gln Leu Asp Lys Ile Leu Lys Tyr Val Glu
    370                 375                 380

Ile Gly Lys Asn Glu Gly Ala Thr Leu Ile Thr Gly Gly Glu Arg Ile
385                 390                 395                 400
```

```
Gly Asp Lys Gly Tyr Phe Val Arg Pro Thr Ile Phe Ala Asp Val Glu
                405                 410                 415

Glu Ser Met Arg Ile Val Lys Glu Glu Ile Phe Gly Pro Val Val Thr
            420                 425                 430

Val Thr Lys Phe Lys Thr Ile Asp Glu Leu Ile Glu Lys Ala Asn Asp
        435                 440                 445

Ser Glu Tyr Gly Leu Ala Ala Gly Ile His Thr Ser Asn Ile Asn Thr
    450                 455                 460

Ala Ile Gln Val Ser Asn Arg Leu Asn Ala Gly Thr Val Trp Ile Asn
465                 470                 475                 480

Thr Tyr Asn Asp Phe His His Cys Val Pro Phe Gly Gly Phe Asn Ala
                485                 490                 495

Ser Gly Ile Gly Arg Glu Met Gly Ser Glu Ala Leu Asp Asn Tyr Thr
            500                 505                 510

Gln Val Lys Ala Val Arg Val Lys Leu Glu
        515                 520
```

<210> 744
<211> 1524
<212> DNA
<213> Artificial Sequence

<220>
<223> KLLA0E23057 from Kluyveromyces lactis

<400> 744

```
atgtcctcta caattgctga gaaattgaac ctcaagatcg tcgaacaaga cgctgttagc       60

atcactttgc caaacggttt gacttaccaa caaccaactg gtttgttcat caacaatcag      120

ttcatcaagt ctcaagacgg taagactttg aaggttgaaa acccatctac tgaggaaatc      180

attgtcgaag tccaatctgc tacttctcaa gacgtcgagt acgccgttga agctgccgat      240

gctgctttca actccgaatg gtctactatg gacccaaaaa agcgtggttc tttgttgttt      300

aagttggctg acttgattga agctcaaaag gaattgattg cttctatcga atctgctgac      360

aacggtaaga ctttggccct agccagaggt gatgttggtt tggtcattga ctacatcaga      420

tctgctgctg gttatgctga caagttgggt ggtagaacta tcaacactgg tgatggttac      480

gctaacttca cttacaagga acctctaggt gtctgtggtc aaatcatccc atggaacttc      540

ccattgatga tgctttcttg gaagatcgcc cctgctttgg ttgctggtaa caccgttatc      600

ttgaagccag cttccccaac cccattgaac gctttgttct ttgcttcttt gtgtaaggaa      660

gcaggtatcc cagctggtgt cgttaacatc gttccaggtc caggtagatc cgttggtgac      720

accatcacca accatccaaa aattagaaag attgccttca ctggttccac tgacattggt      780

agagacgttg ctatcaaggc tgcccaatct aacttgaaga aggtcacctt ggaattgggt      840

ggtaaatccg ctcatttggt ctttgaagat gccaacatta agaagactat tccaaacttg      900

gtcaacggta ttttcaagaa tgctggtcaa atttgttcct ctggttccag aatctatgtc      960
```

caagacacca tctacgatca actattgtct gaattcaaga cttacctgga aactgaaatt 1020

aaggtcggtt ccccattcga tgaatctaac ttccaagctg ctatcaacaa caaggctcaa 1080

ttcgaaacta tcttgaacta catcgacatc ggtaagaagg aaggtgcttc tatcttgact 1140

ggtggtgaaa gagtaggcaa caagggttac ttcattaaac caactgtatt ctacaacgtt 1200

aaggaagata tgagaatcgt caaggaagaa atctttggtc ctgtcgtcac catctccaag 1260

ttctctactg tcgacgaagc tgtcgctttg gctaacgact ccgaattcgg tttgggtgct 1320

ggtatcgaaa ctgaaaacat ctccgttgcc ttgaaggtcg caaagagact aaaagctggt 1380

accgtctgga tcaacactta caacgatttc gacgctgccg ttccattcgg tggttacaag 1440

caatctggtt acggtagaga aatgggtgaa gaagctttcg aatcttacac tcaaatcaag 1500

gccgtcagga tcaagttgga ttaa 1524

<210> 745
<211> 507
<212> PRT
<213> Artificial Sequence

<220>
<223> KLLA0E23057 from Kluyveromyces lactis

<400> 745

Met Ser Ser Thr Ile Ala Glu Lys Leu Asn Leu Lys Ile Val Glu Gln
1 5 10 15

Asp Ala Val Ser Ile Thr Leu Pro Asn Gly Leu Thr Tyr Gln Gln Pro
20 25 30

Thr Gly Leu Phe Ile Asn Asn Gln Phe Ile Lys Ser Gln Asp Gly Lys
35 40 45

Thr Leu Lys Val Glu Asn Pro Ser Thr Glu Glu Ile Ile Val Glu Val
50 55 60

Gln Ser Ala Thr Ser Gln Asp Val Glu Tyr Ala Val Glu Ala Ala Asp
65 70 75 80

Ala Ala Phe Asn Ser Glu Trp Ser Thr Met Asp Pro Lys Lys Arg Gly
85 90 95

Ser Leu Leu Phe Lys Leu Ala Asp Leu Ile Glu Ala Gln Lys Glu Leu
100 105 110

Ile Ala Ser Ile Glu Ser Ala Asp Asn Gly Lys Thr Leu Ala Leu Ala
115 120 125

Arg Gly Asp Val Gly Leu Val Ile Asp Tyr Ile Arg Ser Ala Ala Gly
130 135 140

Tyr Ala Asp Lys Leu Gly Gly Arg Thr Ile Asn Thr Gly Asp Gly Tyr
145 150 155 160

20 Jan 2017 2017200439

```
Ala Asn Phe Thr Tyr Lys Glu Pro Leu Gly Val Cys Gly Gln Ile Ile
                165                 170                 175

Pro Trp Asn Phe Pro Leu Met Met Leu Ser Trp Lys Ile Ala Pro Ala
            180                 185                 190

Leu Val Ala Gly Asn Thr Val Ile Leu Lys Pro Ala Ser Pro Thr Pro
        195                 200                 205

Leu Asn Ala Leu Phe Phe Ala Ser Leu Cys Lys Glu Ala Gly Ile Pro
    210                 215                 220

Ala Gly Val Val Asn Ile Val Pro Gly Pro Gly Arg Ser Val Gly Asp
225                 230                 235                 240

Thr Ile Thr Asn His Pro Lys Ile Arg Lys Ile Ala Phe Thr Gly Ser
                245                 250                 255

Thr Asp Ile Gly Arg Asp Val Ala Ile Lys Ala Ala Gln Ser Asn Leu
            260                 265                 270

Lys Lys Val Thr Leu Glu Leu Gly Gly Lys Ser Ala His Leu Val Phe
        275                 280                 285

Glu Asp Ala Asn Ile Lys Lys Thr Ile Pro Asn Leu Val Asn Gly Ile
    290                 295                 300

Phe Lys Asn Ala Gly Gln Ile Cys Ser Ser Gly Ser Arg Ile Tyr Val
305                 310                 315                 320

Gln Asp Thr Ile Tyr Asp Gln Leu Leu Ser Glu Phe Lys Thr Tyr Leu
                325                 330                 335

Glu Thr Glu Ile Lys Val Gly Ser Pro Phe Asp Glu Ser Asn Phe Gln
            340                 345                 350

Ala Ala Ile Asn Asn Lys Ala Gln Phe Glu Thr Ile Leu Asn Tyr Ile
        355                 360                 365

Asp Ile Gly Lys Lys Glu Gly Ala Ser Ile Leu Thr Gly Gly Glu Arg
    370                 375                 380

Val Gly Asn Lys Gly Tyr Phe Ile Lys Pro Thr Val Phe Tyr Asn Val
385                 390                 395                 400

Lys Glu Asp Met Arg Ile Val Lys Glu Glu Ile Phe Gly Pro Val Val
                405                 410                 415

Thr Ile Ser Lys Phe Ser Thr Val Asp Glu Ala Val Ala Leu Ala Asn
            420                 425                 430

Asp Ser Glu Phe Gly Leu Gly Ala Gly Ile Glu Thr Glu Asn Ile Ser
        435                 440                 445
```

```
Val Ala Leu Lys Val Ala Lys Arg Leu Lys Ala Gly Thr Val Trp Ile
    450                 455                 460

Asn Thr Tyr Asn Asp Phe Asp Ala Ala Val Pro Phe Gly Gly Tyr Lys
465                 470                 475                 480

Gln Ser Gly Tyr Gly Arg Glu Met Gly Glu Glu Ala Phe Glu Ser Tyr
                485                 490                 495

Thr Gln Ile Lys Ala Val Arg Ile Lys Leu Asp
            500                 505
```

<210> 746
<211> 1512
<212> DNA
<213> Artificial Sequence

<220>
<223> KLLA0D10021 from Kluyveromyces lactis

<400> 746

```
atgtccttta ccagcatcaa actagcagat ggtactgagt accaacaacc tttaggttta       60

ttcattaaca atgaatatgt tgaagcacat ggttcggaga ctttaacatc atataaccca      120

agttctggtg aagaaatctg caaagttcaa tgtgcagacc caaagattga tgtcgatttg      180

gctgttaagg ctgcaagaga cgcttttcct ggttggaaag cgactccagc tatagaaaag      240

agagacttgt tctacaaatt ggccggcttg ttggaaagag atattgacat attaagtaaa      300

atcgagtccc tggacactgg aaagccaatt gccaacaatg ccaagtacga catactagaa      360

gtggttgagg tcttaaagta ctatgctgga tggatcgata aggcaggtgg tcaatctttc      420

gctccaagca gtgaaaagtt agcttatact ttacatcagc cattaggcgt ggttggatgt      480

attatcccat tcaactatcc attggctatg atgtcttgga aatggactgc tatcgccact      540

gggaacacaa ctgtttttaa atctgcggat cagactccat tatccatcct atatttcgcc      600

aaattggtga aggaagcagg gttccctgca ggtgttttca atgtcttatc cggtacaggt      660

gcctctgttg gtagtgccat tgttaaacac actggtgtag acaagattgc attcacaggt      720

tccactgcag ttggtcaaat tatccaacaa caatctgcag taaacttgaa acctatcaca      780

ttggaatgtg gtggtaaatc tgcgaacatg atctttgaag attgtgattt agaacaggcc      840

gtgaaatgga gtgcatgggg tattttcaac aatatgggtc aaatctgttc cggtacttca      900

agagtcttcg tgcaagaatc tatttacgag gaatttgtag cacaattggc tgaccatgtc      960

aagaaaactt atatcgtggg tgctccttca gatgatgccg tcatcgtcgg acctcaagtc     1020

tccaaaaaac aacaagaacg tgtccttgaa tacattaaga tcggtatcga agaaggatgc     1080

aaagtgccat tgggtggtct tggagcacct gagccaattg cgaatgatcc agcattagct     1140

aacggttact acgtcaagcc aaccatttta gcagatgtca aagcgtcgta cagaatcgcc     1200

aaggaagaaa tctttggtcc cgtcatagtc gtaggctctt tcaaggacta caaagaagct     1260

ttagacattg ccaacgattc agactacggt cttggttcgg caattttctc aaaagacatc     1320

gtaacatgtc acaaatttgc ccaggacgtg gaagcaggta tggtctggat caactcttct     1380

aacgatgttg acatcaacgt cccattcggt ggtgtaaaaa tgtccggtca tggccgtgaa     1440
```

ttaggtgaat acggtttagc taacttcact aacgttaaat cagttcatgt taacctagga 1500

agaaagttgt aa 1512

<210> 747
<211> 503
<212> PRT
<213> Artificial Sequence

<220>
<223> KLLA0D10021 from Kluyveromyces lactis

<400> 747

Met Ser Phe Thr Ser Ile Lys Leu Ala Asp Gly Thr Glu Tyr Gln Gln
1 5 10 15

Pro Leu Gly Leu Phe Ile Asn Asn Glu Tyr Val Glu Ala His Gly Ser
20 25 30

Glu Thr Leu Thr Ser Tyr Asn Pro Ser Ser Gly Glu Glu Ile Cys Lys
35 40 45

Val Gln Cys Ala Asp Pro Lys Ile Asp Val Asp Leu Ala Val Lys Ala
50 55 60

Ala Arg Asp Ala Phe Pro Gly Trp Lys Ala Thr Pro Ala Ile Glu Lys
65 70 75 80

Arg Asp Leu Phe Tyr Lys Leu Ala Gly Leu Leu Glu Arg Asp Ile Asp
85 90 95

Ile Leu Ser Lys Ile Glu Ser Leu Asp Thr Gly Lys Pro Ile Ala Asn
100 105 110

Asn Ala Lys Tyr Asp Ile Leu Glu Val Val Glu Val Leu Lys Tyr Tyr
115 120 125

Ala Gly Trp Ile Asp Lys Ala Gly Gly Gln Ser Phe Ala Pro Ser Ser
130 135 140

Glu Lys Leu Ala Tyr Thr Leu His Gln Pro Leu Gly Val Val Gly Cys
145 150 155 160

Ile Ile Pro Phe Asn Tyr Pro Leu Ala Met Met Ser Trp Lys Trp Thr
165 170 175

Ala Ile Ala Thr Gly Asn Thr Thr Val Phe Lys Ser Ala Asp Gln Thr
180 185 190

Pro Leu Ser Ile Leu Tyr Phe Ala Lys Leu Val Lys Glu Ala Gly Phe
195 200 205

Pro Ala Gly Val Phe Asn Val Leu Ser Gly Thr Gly Ala Ser Val Gly
210 215 220

```
Ser Ala Ile Val Lys His Thr Gly Val Asp Lys Ile Ala Phe Thr Gly
225                 230                 235                 240

Ser Thr Ala Val Gly Gln Ile Ile Gln Gln Gln Ser Ala Val Asn Leu
                245                 250                 255

Lys Pro Ile Thr Leu Glu Cys Gly Gly Lys Ser Ala Asn Met Ile Phe
            260                 265                 270

Glu Asp Cys Asp Leu Glu Gln Ala Val Lys Trp Ser Ala Trp Gly Ile
        275                 280                 285

Phe Asn Asn Met Gly Gln Ile Cys Ser Gly Thr Ser Arg Val Phe Val
    290                 295                 300

Gln Glu Ser Ile Tyr Glu Glu Phe Val Ala Gln Leu Ala Asp His Val
305                 310                 315                 320

Lys Lys Thr Tyr Ile Val Gly Ala Pro Ser Asp Asp Ala Val Ile Val
                325                 330                 335

Gly Pro Gln Val Ser Lys Lys Gln Gln Glu Arg Val Leu Glu Tyr Ile
            340                 345                 350

Lys Ile Gly Ile Glu Glu Gly Cys Lys Val Pro Leu Gly Gly Leu Gly
        355                 360                 365

Ala Pro Glu Pro Ile Ala Asn Asp Pro Ala Leu Ala Asn Gly Tyr Tyr
    370                 375                 380

Val Lys Pro Thr Ile Leu Ala Asp Val Lys Ala Ser Tyr Arg Ile Ala
385                 390                 395                 400

Lys Glu Glu Ile Phe Gly Pro Val Ile Val Val Gly Ser Phe Lys Asp
                405                 410                 415

Tyr Lys Glu Ala Leu Asp Ile Ala Asn Asp Ser Asp Tyr Gly Leu Gly
            420                 425                 430

Ser Ala Ile Phe Ser Lys Asp Ile Val Thr Cys His Lys Phe Ala Gln
        435                 440                 445

Asp Val Glu Ala Gly Met Val Trp Ile Asn Ser Ser Asn Asp Val Asp
    450                 455                 460

Ile Asn Val Pro Phe Gly Gly Val Lys Met Ser Gly His Gly Arg Glu
465                 470                 475                 480

Leu Gly Glu Tyr Gly Leu Ala Asn Phe Thr Asn Val Lys Ser Val His
                485                 490                 495

Val Asn Leu Gly Arg Lys Leu
            500
```

20172000439 20 Jan 2017

<210> 748
<211> 1506
<212> DNA
<213> Artificial Sequence

<220>
<223> KLLA0D09999 from Kluyveromyces lactis

<400> 748
atgtcattgt acgaaacggt taacattcca cagttgaaca ttacatatga gcagccattg 60

gggttgttca tcaacaacga gtttgtcaag tccagtgacg gtaagaaaat tgaaagcatc 120

aacccaagta ctgaggaatc gattgtctct ttttttgcag cttctgaaaa agatgtcaat 180

agtgctgtta cagcagcaag agatgctttt gaaaacaatt ggtccaagac atcaggcgaa 240

cagagaggaa ttcttttgtc gaagttattt gatttgattg agagagataa gttattgcta 300

gctgcgattg agacgttaga ttcggggaag ccttaccact ctaatagtct tgctgatctt 360

gaacaaatct tgcaattaac tcgttactat gctggcggtg ctgataagtt caccggaggg 420

gaatatattc cattaaacca agacaagtat gcgtatacca tgaaggtacc ttatggtgtt 480

gttgcacaaa ttgttccatg gaattatccc ttggctatgg ctagctggaa aatccaatcg 540

tgtttggctg ctggtaacac tattgttatc aagccagcag agaacacttc tttgtctttg 600

ctttacttcg cacaacttgt caaagaggct ggctttcctc caggtgtgtt gaatatctta 660

ccaggttatg gttccgttac gggaactgca atggctagcc acactgacat agacaagatt 720

gcgtttactg gttcgaccgc tgttggacaa aaagtcatgt cacttgctgc ccaatcaaac 780

ttgaaagctg ttactttgga atgcggtggt aagtctccat gtctaatttt tgctgatgcc 840

aaccttgatg aagccattcc ttgggttgct ggaggtatct tctataactc agggcagaat 900

tgtacagcaa actcgaggat ctacgtccac gaggatattt acgaagagtt tttaactaaa 960

tttgaagaat acgctaaatc aaactggaag ttcggagcta aattcgatcc tttcgataaa 1020

gactgtactc tagggcctgt tatttcaaaa gtccaatatg atcgtatcca atcctatgtt 1080

gaacatggta aagcctctga gaaactaggc ttcaaggagt ttgccgaata tccttctaag 1140

ggttatttta tcccacctac tatattttca gatgttccaa actcttcgaa actctctcaa 1200

gaagaaattt ttggaccggt gggtgttgtg gcatcatttt caacttatga ggaagccatc 1260

aagttggcaa atgactcctc ttacggtttg gcttcttgtg ttttctctga aaatgcccga 1320

acaacagcac aatttgctcg cgatattagg gcaggtaccg tttggatcaa ttcttccaat 1380

gacgaagaaa tttctgttcc atttggtggg ttcaaaatgt ccggtattgg tcgtgaattg 1440

gggcgtaacg gtgttgactc ttatttgcaa acaaaagctg ttcacgtaaa tattgcaaag 1500

ctgtga 1506

<210> 749
<211> 501
<212> PRT
<213> Artificial Sequence

<220>
<223> KLLA0D09999 from Kluyveromyces lactis

<400> 749

2017200439 20 Jan 2017

```
Met Ser Leu Tyr Glu Thr Val Asn Ile Pro Gln Leu Asn Ile Thr Tyr
1               5                  10                  15

Glu Gln Pro Leu Gly Leu Phe Ile Asn Asn Glu Phe Val Lys Ser Ser
            20                  25                  30

Asp Gly Lys Lys Ile Glu Ser Ile Asn Pro Ser Thr Glu Glu Ser Ile
        35                  40                  45

Val Ser Phe Phe Ala Ala Ser Glu Lys Asp Val Asn Ser Ala Val Thr
    50                  55                  60

Ala Ala Arg Asp Ala Phe Glu Asn Asn Trp Ser Lys Thr Ser Gly Glu
65                  70                  75                  80

Gln Arg Gly Ile Leu Leu Ser Lys Leu Phe Asp Leu Ile Glu Arg Asp
                85                  90                  95

Lys Leu Leu Leu Ala Ala Ile Glu Thr Leu Asp Ser Gly Lys Pro Tyr
            100                 105                 110

His Ser Asn Ser Leu Ala Asp Leu Glu Gln Ile Leu Gln Leu Thr Arg
        115                 120                 125

Tyr Tyr Ala Gly Gly Ala Asp Lys Phe Thr Gly Gly Glu Tyr Ile Pro
    130                 135                 140

Leu Asn Gln Asp Lys Tyr Ala Tyr Thr Met Lys Val Pro Tyr Gly Val
145                 150                 155                 160

Val Ala Gln Ile Val Pro Trp Asn Tyr Pro Leu Ala Met Ala Ser Trp
                165                 170                 175

Lys Ile Gln Ser Cys Leu Ala Ala Gly Asn Thr Ile Val Ile Lys Pro
            180                 185                 190

Ala Glu Asn Thr Ser Leu Ser Leu Leu Tyr Phe Ala Gln Leu Val Lys
        195                 200                 205

Glu Ala Gly Phe Pro Pro Gly Val Leu Asn Ile Leu Pro Gly Tyr Gly
    210                 215                 220

Ser Val Thr Gly Thr Ala Met Ala Ser His Thr Asp Ile Asp Lys Ile
225                 230                 235                 240

Ala Phe Thr Gly Ser Thr Ala Val Gly Gln Lys Val Met Ser Leu Ala
                245                 250                 255

Ala Gln Ser Asn Leu Lys Ala Val Thr Leu Glu Cys Gly Gly Lys Ser
            260                 265                 270

Pro Cys Leu Ile Phe Ala Asp Ala Asn Leu Asp Glu Ala Ile Pro Trp
        275                 280                 285
```

Val Ala Gly Gly Ile Phe Tyr Asn Ser Gly Gln Asn Cys Thr Ala Asn
290 295 300

Ser Arg Ile Tyr Val His Glu Asp Ile Tyr Glu Glu Phe Leu Thr Lys
305 310 315 320

Phe Glu Glu Tyr Ala Lys Ser Asn Trp Lys Phe Gly Ala Lys Phe Asp
325 330 335

Pro Phe Asp Lys Asp Cys Thr Leu Gly Pro Val Ile Ser Lys Val Gln
340 345 350

Tyr Asp Arg Ile Gln Ser Tyr Val Glu His Gly Lys Ala Ser Glu Lys
355 360 365

Leu Gly Phe Lys Glu Phe Ala Glu Tyr Pro Ser Lys Gly Tyr Phe Ile
370 375 380

Pro Pro Thr Ile Phe Ser Asp Val Pro Asn Ser Ser Lys Leu Ser Gln
385 390 395 400

Glu Glu Ile Phe Gly Pro Val Gly Val Val Ala Ser Phe Ser Thr Tyr
405 410 415

Glu Glu Ala Ile Lys Leu Ala Asn Asp Ser Ser Tyr Gly Leu Ala Ser
420 425 430

Cys Val Phe Ser Glu Asn Ala Arg Thr Thr Ala Gln Phe Ala Arg Asp
435 440 445

Ile Arg Ala Gly Thr Val Trp Ile Asn Ser Ser Asn Asp Glu Glu Ile
450 455 460

Ser Val Pro Phe Gly Gly Phe Lys Met Ser Gly Ile Gly Arg Glu Leu
465 470 475 480

Gly Arg Asn Gly Val Asp Ser Tyr Leu Gln Thr Lys Ala Val His Val
485 490 495

Asn Ile Ala Lys Leu
500

<210> 750
<211> 1506
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD2 from Pichia stipits

<400> 750
ctacaacttg ttgcccaagt taacgtgaat ggcctttgcc tgggtgtaga tggacaatcc 60

gtattcaccc aattctctac ctacaccaga catcttgaca ccaccgaatg gaatgtggac 120

atcagagtcg ttggaagagt tgatccacac cataccagct tgtatttcag cagccatagt 180

gtgggccacg gtgatgtcct tggtgaagat ggcagcaccc aaaccgtatt gggtctggtt 240

ggcgtaagtg atgacttctt cgtctgacga gaacttaccc accactacaa cagggccaaa 300
tatttcctcg ttgacgattc tcatctcagg cttgatgtca gcaaaaatag ttggcttgat 360
gtagaatcct ttggataatt caccctccga gttcttttca cctcccaaaa gcaagcgggc 420
accttctttc ttaccgattt cgatgtagct caacacttta tcacgctgaa ccttcgaaac 480
ttgtggacca acgactgctt cagaatcgaa catactaccc tgtttgtagg cttctttcac 540
gtgtgcagtg tattcctcca agaacttgtc gtagacagat tcgtgaacgt acactctgga 600
agtggacgta cagatctgac cttggttact catgataccg atggcagccc acttgacagc 660
ttgttccaag tcagcgtctt cacggatgat caaaggcgac ttaccaccac actctaaggt 720
cacagccttc aagttagaag cagcagcctg ttgtataagc ttaccagtag cagtagaacc 780
agtgaaagcg atcttgtcga catctaaatg ggatgaaaga gccttaccgg caacagcacc 840
gaagccggag atgatgttaa ccacaccggc tgggaaacca gcgtccttga agaggcgagc 900
tacgtacaag agagacaatg gtgtaatttc tgaagttttc aacacaacta cgttaccagc 960
agccaaagct ggagccaact tccaggctgc catagccaat gggtagttcc atgggatgat 1020
ctgaccacaa acaccaagtg gttcatggat ggtgtaggcc aacttctttg gatcgttctg 1080
gatcacctta ccctggattt tgtcagccca gccagcaaag tatctgaaac aactgataca 1140
ttcgtcaata tcgtatacgg cgttttgttc cttggtctta ccagagtccc aggcttcaat 1200
ggcaccaatc aagtctcttt gggcttcaaa ggcatcagcg accttgttca acaagattcc 1260
tctgtccaca ccggtgacct tcttccagtc cttaaatgcg gctctggcag ccttgacggc 1320
gatatcgaca tcttcttcct cggcacagta aacggtacca ttgacttcac tggtagatgg 1380
gttgatggag tccaacgtct ttccggaaac agacttgaca aactcgttgt tgatgaacaa 1440
accagtaggc tgtgtgaatg tgttgccgtc tggaagggta accaggtatt ctaatgggag 1500
agacat 1506

<210> 751
<211> 501
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD2 from Pichia stipits

<400> 751

Met Ser Leu Pro Leu Glu Tyr Ser Val Thr Leu Pro Asp Gly Asn Thr
1 5 10 15

Phe Thr Gln Pro Thr Gly Leu Phe Ile Asn Asn Glu Phe Val Lys Ser
20 25 30

Val Ser Gly Lys Thr Leu Asp Ser Ile Asn Pro Ser Thr Ser Glu Val
35 40 45

Asn Gly Thr Val Tyr Cys Ala Glu Glu Glu Asp Val Asp Ile Ala Val
50 55 60

```
Lys Ala Ala Arg Ala Ala Phe Lys Asp Trp Lys Lys Val Thr Gly Val
65                  70                  75                  80

Asp Arg Gly Ile Leu Leu Asn Lys Val Ala Asp Ala Phe Glu Ala Gln
                85                  90                  95

Arg Asp Leu Ile Gly Ala Ile Glu Ala Trp Asp Ser Gly Lys Thr Lys
            100                 105                 110

Glu Gln Asn Ala Val Tyr Asp Ile Asp Glu Cys Ile Ser Cys Phe Arg
        115                 120                 125

Tyr Phe Ala Gly Trp Ala Asp Lys Ile Gln Gly Lys Val Ile Gln Asn
    130                 135                 140

Asp Pro Lys Lys Leu Ala Tyr Thr Ile His Glu Pro Leu Gly Val Cys
145                 150                 155                 160

Gly Gln Ile Ile Pro Trp Asn Tyr Pro Leu Ala Met Ala Ala Trp Lys
                165                 170                 175

Leu Ala Pro Ala Leu Ala Ala Gly Asn Val Val Val Leu Lys Thr Ser
            180                 185                 190

Glu Ile Thr Pro Leu Ser Leu Leu Tyr Val Ala Arg Leu Phe Lys Asp
        195                 200                 205

Ala Gly Phe Pro Ala Gly Val Val Asn Ile Ile Ser Gly Phe Gly Ala
    210                 215                 220

Val Ala Gly Lys Ala Leu Ser Ser His Leu Asp Val Asp Lys Ile Ala
225                 230                 235                 240

Phe Thr Gly Ser Thr Ala Thr Gly Lys Leu Ile Gln Gln Ala Ala Ala
                245                 250                 255

Ser Asn Leu Lys Ala Val Thr Leu Glu Cys Gly Gly Lys Ser Pro Leu
            260                 265                 270

Ile Ile Arg Glu Asp Ala Asp Leu Glu Gln Ala Val Lys Trp Ala Ala
        275                 280                 285

Ile Gly Ile Met Ser Asn Gln Gly Gln Ile Cys Thr Ser Thr Ser Arg
    290                 295                 300

Val Tyr Val His Glu Ser Val Tyr Asp Lys Phe Leu Glu Glu Tyr Thr
305                 310                 315                 320

Ala His Val Lys Glu Ala Tyr Lys Gln Gly Ser Met Phe Asp Ser Glu
                325                 330                 335

Ala Val Val Gly Pro Gln Val Ser Lys Val Gln Arg Asp Lys Val Leu
            340                 345                 350
```

Ser Tyr Ile Glu Ile Gly Lys Lys Glu Gly Ala Arg Leu Leu Leu Gly
355 360 365

Gly Glu Lys Asn Ser Glu Gly Glu Leu Ser Lys Gly Phe Tyr Ile Lys
370 375 380

Pro Thr Ile Phe Ala Asp Ile Lys Pro Glu Met Arg Ile Val Asn Glu
385 390 395 400

Glu Ile Phe Gly Pro Val Val Val Val Gly Lys Phe Ser Ser Asp Glu
405 410 415

Glu Val Ile Thr Tyr Ala Asn Gln Thr Gln Tyr Gly Leu Gly Ala Ala
420 425 430

Ile Phe Thr Lys Asp Ile Thr Val Ala His Thr Met Ala Ala Glu Ile
435 440 445

Gln Ala Gly Met Val Trp Ile Asn Ser Ser Asn Asp Ser Asp Val His
450 455 460

Ile Pro Phe Gly Gly Val Lys Met Ser Gly Val Gly Arg Glu Leu Gly
465 470 475 480

Glu Tyr Gly Leu Ser Ile Tyr Thr Gln Ala Lys Ala Ile His Val Asn
485 490 495

Leu Gly Asn Lys Leu
500

<210> 752
<211> 1434
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD3 from Pichia stipitis

<400> 752
atgtccgcta ttagtagaac attccccagt attatcgagg gaaaggattt ccacagcaat 60

gaaaaacatc cagtatactc ccatgtcacc cagaaagaag caatccacta tttctcctac 120

ttaactgata tcaagaaggc agtttcgaaa atcgctgctg atgctgacga agggtttgaa 180

gaatggtcgt ccatggccta tcaagaaaga gtgaagattt tcgagaaggc tgctgcttta 240

cttgctgaaa gaagagaaga gttgatagct tctcacaaaa atatcggagg tcctacctgg 300

ttttcccatg tgaatgctga tgaaatcatc tcgcaattga aggaatatac ttcacttctc 360

tctagaccta ctggtttagt agcccaatct gctcattctg atctcgcact cgtcgtcaag 420

cagccactag gtcctgtcct cgccattgct ccctggaacg ctcctgttct tttggcaggt 480

agagccatag tggctccgtt ggctgcgggc tgttcggtca tcctaaaagc ttctgaaaag 540

gctccagaat ccgcatacct tgttgtgaag accttcattg atgctggtat cccgtcaaaa 600

gcattgcaat tggtcttcat caaaccagat gacaatccag aattcatcaa ctccatcttg 660

gacactggtt tgatcaaaaa ggtcaacttc actggctcta caatcgtggg caagaagatc 720

gctgaagctg ccagtaaaca cttagtacca tatcttatgg aattaggcgg aaagaatgtg 780

tctattgttg aaaaggatgc tgacttggtg agggccgttg aaactataat ctggagttcg 840

tggtcgcaca aaggtcaaat atgcatgagt actgacaagg tcttcgttga tgaaagcatc 900

tacgacaaat tcgttgctca attaaaagta tccgccaatg agatcgtcaa ggaccccgac 960

tacgcaattt ctcaaagaga tattacattc aagagaaacc ttgttaagtt ggttaagaat 1020

gcattagatt tgggtgcaaa tttgatattt ggtaaattaa atgaccattt ggacagcggt 1080

tccttcagtc cattgatctt ggaaaatgtc acttcaaaca tgttgcttga ttctaccgaa 1140

tcattcggac ctttgttcgc tgtatataag tattcagata caatcaaact tgtcaaggaa 1200

ttaaacagag ctgattatgg attgaaggcc tcaatttggt cccaaaatgt tttgcaagca 1260

ttggaaacag ctaaaaaaat ccacgtaggt ggtgtacata tcaatagttc tacgattcac 1320

gacgaagcga ctctaccaca tggcggtgtt aagtcaagtg gtgctggaag attcaactcc 1380

atatggggta ttgacgattt ttccattacc aagacaatta ctcttagtca gtaa 1434

<210> 753
<211> 477
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD3 from Pichia stipitis

<400> 753

Met Ser Ala Ile Ser Arg Thr Phe Pro Ser Ile Ile Glu Gly Lys Asp
1 5 10 15

Phe His Ser Asn Glu Lys His Pro Val Tyr Ser His Val Thr Gln Lys
20 25 30

Glu Ala Ile His Tyr Phe Ser Tyr Leu Thr Asp Ile Lys Lys Ala Val
35 40 45

Ser Lys Ile Ala Ala Asp Ala Asp Glu Gly Phe Glu Glu Trp Ser Ser
50 55 60

Met Ala Tyr Gln Glu Arg Val Lys Ile Phe Glu Lys Ala Ala Ala Leu
65 70 75 80

Leu Ala Glu Arg Arg Glu Glu Leu Ile Ala Ser His Lys Asn Ile Gly
85 90 95

Gly Pro Thr Trp Phe Ser His Val Asn Ala Asp Glu Ile Ile Ser Gln
100 105 110

Leu Lys Glu Tyr Thr Ser Leu Leu Ser Arg Pro Thr Gly Leu Val Ala
115 120 125

Gln Ser Ala His Ser Asp Leu Ala Leu Val Val Lys Gln Pro Leu Gly
130 135 140

```
Pro Val Leu Ala Ile Ala Pro Trp Asn Ala Pro Val Leu Leu Ala Gly
145                 150                 155                 160

Arg Ala Ile Val Ala Pro Leu Ala Ala Gly Cys Ser Val Ile Leu Lys
                165                 170                 175

Ala Ser Glu Lys Ala Pro Glu Ser Ala Tyr Leu Val Val Lys Thr Phe
            180                 185                 190

Ile Asp Ala Gly Ile Pro Ser Lys Ala Leu Gln Leu Val Phe Ile Lys
        195                 200                 205

Pro Asp Asp Asn Pro Glu Phe Ile Asn Ser Ile Leu Asp Thr Gly Leu
    210                 215                 220

Ile Lys Lys Val Asn Phe Thr Gly Ser Thr Ile Val Gly Lys Lys Ile
225                 230                 235                 240

Ala Glu Ala Ala Ser Lys His Leu Val Pro Tyr Leu Met Glu Leu Gly
                245                 250                 255

Gly Lys Asn Val Ser Ile Val Glu Lys Asp Ala Asp Leu Val Arg Ala
            260                 265                 270

Val Glu Thr Ile Ile Trp Ser Ser Trp Ser His Lys Gly Gln Ile Cys
        275                 280                 285

Met Ser Thr Asp Lys Val Phe Val Asp Glu Ser Ile Tyr Asp Lys Phe
    290                 295                 300

Val Ala Gln Leu Lys Val Ser Ala Asn Glu Ile Val Lys Asp Pro Asp
305                 310                 315                 320

Tyr Ala Ile Ser Gln Arg Asp Ile Thr Phe Lys Arg Asn Leu Val Lys
                325                 330                 335

Leu Val Lys Asn Ala Leu Asp Leu Gly Ala Asn Leu Ile Phe Gly Lys
            340                 345                 350

Leu Asn Asp His Leu Asp Ser Gly Ser Phe Ser Pro Leu Ile Leu Glu
        355                 360                 365

Asn Val Thr Ser Asn Met Leu Leu Asp Ser Thr Glu Ser Phe Gly Pro
    370                 375                 380

Leu Phe Ala Val Tyr Lys Tyr Ser Asp Thr Ile Lys Leu Val Lys Glu
385                 390                 395                 400

Leu Asn Arg Ala Asp Tyr Gly Leu Lys Ala Ser Ile Trp Ser Gln Asn
                405                 410                 415

Val Leu Gln Ala Leu Glu Thr Ala Lys Lys Ile His Val Gly Gly Val
            420                 425                 430
```

20172000439 20 Jan 2017

```
His Ile Asn Ser Ser Thr Ile His Asp Glu Ala Thr Leu Pro His Gly
        435                 440                 445

Gly Val Lys Ser Ser Gly Ala Gly Arg Phe Asn Ser Ile Trp Gly Ile
    450                 455                 460

Asp Asp Phe Ser Ile Thr Lys Thr Ile Thr Leu Ser Gln
465                 470                 475
```

<210> 754
<211> 1434
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD4 from Pichia stipitis

<400> 754

```
atggcaactg cttcaacaac tttctcgagc attatcgaag gtaaggtttt ctacactgaa       60
ggaaaacatc ctgtctactc acacgataac catgaggaaa taattcacca cttttcttac      120
ttgactgaca tcaagaagtc agtcccccaa attgcagctg cagctgagca aggttttgag      180
gaatggtcat ccttcgcata tgcagaaaaa gtgaagattt ttgacaaagc tgctatcttg      240
cttactgaaa ggagagacga attgatcgcc tcgcacaaga acatcggagg tcctacatgg      300
ttctctcacg tcaacgcaga tgagtccgtc tcacaattga aggaatacac tggacttctc      360
tcaagaccta cgggtgttgt agctcagtct attcacactg accttgcact tactgtaaga      420
caaccaattg gcccagttct tgctatttct ccttggaatg cacctgttgt attagcaagt      480
agagccatct tggctccatt ggctgctggt tgttctgtta tcttgaaggc ttccgagaag      540
gctccagaat ctgcctacct catagtcaga accttcattg atgcaggagt accacctaag      600
gcattgcaat tagtatttat caagcccgag gacaacgcta ccttcattag ctcgattttg      660
cacactggtt caatcagaaa gatcaacttc actggatcaa ctctcgttgg gaagaaaatt      720
gcagaagttg ccagtaagta cttagtgcca tatcttatgg aattaggagg gaagaacgtc      780
tctattgttg aaaaggatgc tgacttaacc aaggcagctg gtagtattat ctggagctcg      840
tggtcccaca aaggtcaaat atgtatgagc acagacaagg tttttgttga agattctata      900
tacgaccact ttattgccca attaaaagag ctggctgcgg atatcgtgaa agaccctgac      960
taccaaatat ctcaaagaga tattacattc aaaagaaagc ttgttgagtt ggtaaaggat     1020
gccctagaca agggtgctaa cttggtgttt ggtgaactca attctgatgt caacagcagc     1080
tcattcagtc cattgatatt agagaatgtc acttcagaca tgttgcttga ttccactgaa     1140
tccttcggac ctttgttttc tgtgcataag tatgccaata cgaacaaact agtcaaggaa     1200
ttaaacagag ttgactatgg gttgaaggct tctatttggt cccataatgt tttggaagcc     1260
ttagaaaccg caaagaagat ccatgtcggt ggtgtacata tcaatagtcc aactatccat     1320
gatgaaccaa ctcttccaca cggtggtgtt aaatcaagtg gaactggaag attcaactcg     1380
acatggggaa ttgatgactt ttctatcaca aaaactatta ccctaagcca gtga           1434
```

<210> 755
<211> 477

2017200439 20 Jan 2017

<212> PRT
<213> Artificial Sequence

<220>
<223> ALD4 from Pichia stipitis

<400> 755

```
Met Ala Thr Ala Ser Thr Thr Phe Ser Ser Ile Ile Glu Gly Lys Val
1               5                  10                  15

Phe Tyr Thr Glu Gly Lys His Pro Val Tyr Ser His Asp Asn His Glu
            20                  25                  30

Glu Ile Ile His His Phe Ser Tyr Leu Thr Asp Ile Lys Lys Ser Val
        35                  40                  45

Pro Gln Ile Ala Ala Ala Ala Glu Gln Gly Phe Glu Glu Trp Ser Ser
    50                  55                  60

Phe Ala Tyr Ala Glu Lys Val Lys Ile Phe Asp Lys Ala Ala Ile Leu
65                  70                  75                  80

Leu Thr Glu Arg Arg Asp Glu Leu Ile Ala Ser His Lys Asn Ile Gly
                85                  90                  95

Gly Pro Thr Trp Phe Ser His Val Asn Ala Asp Glu Ser Val Ser Gln
            100                 105                 110

Leu Lys Glu Tyr Thr Gly Leu Leu Ser Arg Pro Thr Gly Val Val Ala
        115                 120                 125

Gln Ser Ile His Thr Asp Leu Ala Leu Thr Val Arg Gln Pro Ile Gly
    130                 135                 140

Pro Val Leu Ala Ile Ser Pro Trp Asn Ala Pro Val Val Leu Ala Ser
145                 150                 155                 160

Arg Ala Ile Leu Ala Pro Leu Ala Ala Gly Cys Ser Val Ile Leu Lys
                165                 170                 175

Ala Ser Glu Lys Ala Pro Glu Ser Ala Tyr Leu Ile Val Arg Thr Phe
            180                 185                 190

Ile Asp Ala Gly Val Pro Pro Lys Ala Leu Gln Leu Val Phe Ile Lys
        195                 200                 205

Pro Glu Asp Asn Ala Thr Phe Ile Ser Ser Ile Leu His Thr Gly Ser
    210                 215                 220

Ile Arg Lys Ile Asn Phe Thr Gly Ser Thr Leu Val Gly Lys Lys Ile
225                 230                 235                 240

Ala Glu Val Ala Ser Lys Tyr Leu Val Pro Tyr Leu Met Glu Leu Gly
                245                 250                 255
```

```
Gly Lys Asn Val Ser Ile Val Glu Lys Asp Ala Asp Leu Thr Lys Ala
            260                 265                 270

Ala Gly Ser Ile Ile Trp Ser Ser Trp Ser His Lys Gly Gln Ile Cys
        275                 280                 285

Met Ser Thr Asp Lys Val Phe Val Glu Asp Ser Ile Tyr Asp His Phe
    290                 295                 300

Ile Ala Gln Leu Lys Glu Ser Ala Ala Asp Ile Val Lys Asp Pro Asp
305                 310                 315                 320

Tyr Gln Ile Ser Gln Arg Asp Ile Thr Phe Lys Arg Lys Leu Val Glu
                325                 330                 335

Leu Val Lys Asp Ala Leu Asp Lys Gly Ala Asn Leu Val Phe Gly Glu
            340                 345                 350

Leu Asn Ser Asp Val Asn Ser Ser Ser Phe Ser Pro Leu Ile Leu Glu
        355                 360                 365

Asn Val Thr Ser Asp Met Leu Leu Asp Ser Thr Glu Ser Phe Gly Pro
    370                 375                 380

Leu Phe Ser Val His Lys Tyr Ala Asn Thr Asn Lys Leu Val Lys Glu
385                 390                 395                 400

Leu Asn Arg Val Asp Tyr Gly Leu Lys Ala Ser Ile Trp Ser His Asn
                405                 410                 415

Val Leu Glu Ala Leu Glu Thr Ala Lys Lys Ile His Val Gly Gly Val
            420                 425                 430

His Ile Asn Ser Pro Thr Ile His Asp Glu Pro Thr Leu Pro His Gly
        435                 440                 445

Gly Val Lys Ser Ser Gly Thr Gly Arg Phe Asn Ser Thr Trp Gly Ile
    450                 455                 460

Asp Asp Phe Ser Ile Thr Lys Thr Ile Thr Leu Ser Gln
465                 470                 475


<210>  756
<211>  1488
<212>  DNA
<213>  Artificial Sequence

<220>
<223>  ALD5 from Pichia stipitis

<400>  756
atgtccttgc cattgttcgt tcccatcaaa ttgccaaacg gcactaccta tgagcagcct       60

accggtttgt tcatcaacaa cgagtttgtc cagtccaagt ccaagaagac ctttggaact      120

gtttctcctt ctactgaaga agagatcacc caagtctacg aagccttctc cgaagatatc      180

gatgatgccg ttgaagctgc tactgctgct ttccactcat cgtggtctac ttccgaccct      240
```

caagtcagaa tgaaggttct ctacaaattg gctgacttga tcgacgaaca cgccgacact 300

ttggcccaca ttgaagctct tgataacggt aagtcgttga tgtgttccaa gggtgacgtt 360

gctttgactg ccgcctactt cagatcctgt gctggctgga ccgacaagat taagggttct 420

gttatcgaaa ctggtgacac ccactttaac tacaccagaa gagaaccaat tggtgtttgt 480

ggccagatca tcccatggaa cttcccattg ttgatggctt cctggaagtt aggaccagtc 540

ctttgtaccg gttgcaccac tgttttgaag actgctgaat cgactccttt gtctgctctt 600

tacttggctt ctttgatcaa ggaagccggt gctccacctg gtgttgtcaa tgttgtttcc 660

ggtttcggtc caactgccgg tgctccaatc tcttctcacc caaagatcaa gaaggttgcc 720

ttcactggtt ctaccgctac cggtagacac atcatgaagg ctgctgctga atccaacttg 780

aagaaggtca ctttggaatt gggtggtaag tctccaaaca ttgtctttga cgatgctgac 840

gttaagtcca ccatccaaca cttggttact ggtatcttct acaacactgg tgaagtgtgc 900

tgtgccggtt ccagaatcta cgtccaggaa ggtatctacg acaagattgt ttccgagttc 960

aagaatgcag ccgaatcact taagatcggt gacccattca aggaagacac cttcatgggt 1020

gcccaaacct ctcaattgca attggacaag atcttgaagt acatcgatat cggtaagaag 1080

gaaggtgcca ccgttatcac tggtggtgaa agattcggta acaagggtta ctttatcaag 1140

ccaaccatct tcggagatgt caaggaagac caccagatcg tcagggatga aatcttcggt 1200

cctgttgtca ccatcaccaa gttcaagacc gtcgaagaag tcatcgcctt ggccaacgac 1260

tccgagtacg gtttggctgc tggtgttcac accactaacc tcagcactgc catcagtgtt 1320

tccaacaaga tcaattccgg taccatttgg gtcaacacct acaatgactt ccacccaatg 1380

gttccattcg gaggttactc ccaatctggt attggtagag aaatgggaga agaagccttg 1440

gacaactaca cccaggtcaa ggcagtcaga attggtctta gccagtaa 1488

<210> 757
<211> 495
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD5 from Pichia stipitis

<400> 757

Met Ser Leu Pro Leu Phe Val Pro Ile Lys Leu Pro Asn Gly Thr Thr
1 5 10 15

Tyr Glu Gln Pro Thr Gly Leu Phe Ile Asn Asn Glu Phe Val Gln Ser
20 25 30

Lys Ser Lys Lys Thr Phe Gly Thr Val Ser Pro Ser Thr Glu Glu Glu
35 40 45

Ile Thr Gln Val Tyr Glu Ala Phe Ser Glu Asp Ile Asp Asp Ala Val
50 55 60

Glu Ala Ala Thr Ala Ala Phe His Ser Ser Trp Ser Thr Ser Asp Pro
65 70 75 80

```
Gln Val Arg Met Lys Val Leu Tyr Lys Leu Ala Asp Leu Ile Asp Glu
                85                  90                  95

His Ala Asp Thr Leu Ala His Ile Glu Ala Leu Asp Asn Gly Lys Ser
            100                 105                 110

Leu Met Cys Ser Lys Gly Asp Val Ala Leu Thr Ala Ala Tyr Phe Arg
        115                 120                 125

Ser Cys Ala Gly Trp Thr Asp Lys Ile Lys Gly Ser Val Ile Glu Thr
    130                 135                 140

Gly Asp Thr His Phe Asn Tyr Thr Arg Arg Glu Pro Ile Gly Val Cys
145                 150                 155                 160

Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met Ala Ser Trp Lys
                165                 170                 175

Leu Gly Pro Val Leu Cys Thr Gly Cys Thr Thr Val Leu Lys Thr Ala
            180                 185                 190

Glu Ser Thr Pro Leu Ser Ala Leu Tyr Leu Ala Ser Leu Ile Lys Glu
        195                 200                 205

Ala Gly Ala Pro Pro Gly Val Val Asn Val Val Ser Gly Phe Gly Pro
    210                 215                 220

Thr Ala Gly Ala Pro Ile Ser Ser His Pro Lys Ile Lys Lys Val Ala
225                 230                 235                 240

Phe Thr Gly Ser Thr Ala Thr Gly Arg His Ile Met Lys Ala Ala Ala
                245                 250                 255

Glu Ser Asn Leu Lys Lys Val Thr Leu Glu Leu Gly Gly Lys Ser Pro
            260                 265                 270

Asn Ile Val Phe Asp Asp Ala Asp Val Lys Ser Thr Ile Gln His Leu
        275                 280                 285

Val Thr Gly Ile Phe Tyr Asn Thr Gly Glu Val Cys Cys Ala Gly Ser
    290                 295                 300

Arg Ile Tyr Val Gln Glu Gly Ile Tyr Asp Lys Ile Val Ser Glu Phe
305                 310                 315                 320

Lys Asn Ala Ala Glu Ser Leu Lys Ile Gly Asp Pro Phe Lys Glu Asp
                325                 330                 335

Thr Phe Met Gly Ala Gln Thr Ser Gln Leu Gln Leu Asp Lys Ile Leu
            340                 345                 350

Lys Tyr Ile Asp Ile Gly Lys Lys Glu Gly Ala Thr Val Ile Thr Gly
        355                 360                 365
```

```
Gly Glu Arg Phe Gly Asn Lys Gly Tyr Phe Ile Lys Pro Thr Ile Phe
    370                 375                 380

Gly Asp Val Lys Glu Asp His Gln Ile Val Arg Asp Glu Ile Phe Gly
385                 390                 395                 400

Pro Val Val Thr Ile Thr Lys Phe Lys Thr Val Glu Glu Val Ile Ala
                405                 410                 415

Leu Ala Asn Asp Ser Glu Tyr Gly Leu Ala Ala Gly Val His Thr Thr
            420                 425                 430

Asn Leu Ser Thr Ala Ile Ser Val Ser Asn Lys Ile Asn Ser Gly Thr
        435                 440                 445

Ile Trp Val Asn Thr Tyr Asn Asp Phe His Pro Met Val Pro Phe Gly
    450                 455                 460

Gly Tyr Ser Gln Ser Gly Ile Gly Arg Glu Met Gly Glu Glu Ala Leu
465                 470                 475                 480

Asp Asn Tyr Thr Gln Val Lys Ala Val Arg Ile Gly Leu Ser Gln
                485                 490                 495
```

<210> 758
<211> 1680
<212> DNA
<213> Artificial Sequence

<220>
<223> ALD7 from Pichia stipitis

<400> 758

```
atgactcctc cttcagcaaa caccctcaag gactccgttt cgtccgtcga atcgttcaca       60

aagatcgaca aggccacccc ctcggcttcc accaaggctt ctgagtccat cttgaagtac      120

actgctttgg acgacattcc tgttggagtc aagaagctta ctgactcgtt ccacaagaat      180

ggaaagactc actccatcca gtacagattg aaccagttga gaaacctcta ttttgtcatc      240

aaggataacc aggatgccat ctgcgatgcc ctcttcaagg atttcggcag agtgcctact      300

gagtcgcaga tcttggaaat cgatggaagc ttgaacgagt tggtgcacac catggccaac      360

ttgcacaact ggttgaaacc cgaacctgtc aaggacttgc ccattacatt gaagctgaat      420

cctatctaca ttgaaagaat tccctacgga gttgtgttga tcatctctcc cttcaactac      480

ccattcttcc tctcattttc agccatcatt ggggccattg ctgctggtaa cgtggttgtg      540

ttcaagccct ccgaattgac cccacacttt tcgcagttgt tcactaactt gttgtctgag      600

gctcttgatg atgacctttt gttcaccata aacggttcga ttcccgagac cacgaaggcc      660

ttggaacaga agtacgacaa gatcatgtac accggtaata acgctgtagg aaccattatc      720

gccaagaagg ctgccgagac tttaactcct gtcatcttag agttgggagg aaagtcacca      780

gcgtttgtct tggacgatct ttctgaaaag gagttgaccg tcgtagccag aagaatcgca      840

tggggtagat ttaccaatgc tggtcaaacc tgtgtggccg tggactatgt tttggctcat      900
```

gactctatca aagagagatt ggttcgcgag atcgtcaagg tggtcaagga ggagttctat 960

ccagaactta acaaagacaa caaggacttc acccacatta tccacgaccg tgctgtatcc 1020

aacttatcca agatcatcaa gaccaccaag ggtaagattg ttgttggagg tgatgtagac 1080

gaagcctcga ggtacgtggc tccaactgtc gtggacaatg tagactggga cgactctacc 1140

atgaagggcg agatctttgg tcctatcttg cctatcttgt cgtacagctc tttggatgaa 1200

gctttgtcaa gattgcagag tcgtcatgac actcctttgg cacagtatat cttcactggt 1260

ggatcgactt cccgtgctaa gaaccctaaa cttaacaaga tctcacagca gatcagatcc 1320

ggtggtgctg tcattaacga tgtcttgatg cacgtagcct taactaatgc cccattcgga 1380

ggtattggta gttccggttc gggatcgtac catggttggt tttctttccg tgcatttact 1440

catgagagaa ccaccatgga acaaaagttg tggaacgact ttgttttgaa ggctagatac 1500

cctcctttca cagaaaagaa ccagaagttg gtcagtgctt cacagacaga ctacaacggt 1560

aaggtctggt ttgatagaca aggggacgtc agagtcaagg gtccatcggg cttgttcagc 1620

acttggactt ctgttgctgg tgtcgcagcc ttgatctact actttgttgg taatttgtag 1680

<210> 759
<211> 559
<212> PRT
<213> Artificial Sequence

<220>
<223> ALD7 from Pichia stipitis

<400> 759

Met Thr Pro Pro Ser Ala Asn Thr Leu Lys Asp Ser Val Ser Ser Val
1 5 10 15

Glu Ser Phe Thr Lys Ile Asp Lys Ala Thr Pro Ser Ala Ser Thr Lys
20 25 30

Ala Ser Glu Ser Ile Leu Lys Tyr Thr Ala Leu Asp Asp Ile Pro Val
35 40 45

Gly Val Lys Lys Leu Thr Asp Ser Phe His Lys Asn Gly Lys Thr His
50 55 60

Ser Ile Gln Tyr Arg Leu Asn Gln Leu Arg Asn Leu Tyr Phe Val Ile
65 70 75 80

Lys Asp Asn Gln Asp Ala Ile Cys Asp Ala Leu Phe Lys Asp Phe Gly
85 90 95

Arg Val Pro Thr Glu Ser Gln Ile Leu Glu Ile Asp Gly Ser Leu Asn
100 105 110

Glu Leu Val His Thr Met Ala Asn Leu His Asn Trp Leu Lys Pro Glu
115 120 125

Pro Val Lys Asp Leu Pro Ile Thr Leu Lys Ser Asn Pro Ile Tyr Ile
130 135 140

```
Glu Arg Ile Pro Tyr Gly Val Val Leu Ile Ile Ser Pro Phe Asn Tyr
145                 150                 155                 160

Pro Phe Phe Leu Ser Phe Ser Ala Ile Ile Gly Ala Ile Ala Ala Gly
                165                 170                 175

Asn Val Val Val Phe Lys Pro Ser Glu Leu Thr Pro His Phe Ser Gln
            180                 185                 190

Leu Phe Thr Asn Leu Leu Ser Glu Ala Leu Asp Asp Asp Leu Leu Phe
        195                 200                 205

Thr Ile Asn Gly Ser Ile Pro Glu Thr Thr Lys Ala Leu Glu Gln Lys
    210                 215                 220

Tyr Asp Lys Ile Met Tyr Thr Gly Asn Asn Ala Val Gly Thr Ile Ile
225                 230                 235                 240

Ala Lys Lys Ala Ala Glu Thr Leu Thr Pro Val Ile Leu Glu Leu Gly
                245                 250                 255

Gly Lys Ser Pro Ala Phe Val Leu Asp Asp Leu Ser Glu Lys Glu Leu
            260                 265                 270

Thr Val Val Ala Arg Arg Ile Ala Trp Gly Arg Phe Thr Asn Ala Gly
        275                 280                 285

Gln Thr Cys Val Ala Val Asp Tyr Val Leu Ala His Asp Ser Ile Lys
    290                 295                 300

Glu Arg Leu Val Arg Glu Ile Val Lys Val Val Lys Glu Glu Phe Tyr
305                 310                 315                 320

Pro Glu Leu Asn Lys Asp Asn Lys Asp Phe Thr His Ile Ile His Asp
                325                 330                 335

Arg Ala Val Ser Asn Leu Ser Lys Ile Ile Lys Thr Thr Lys Gly Lys
            340                 345                 350

Ile Val Val Gly Gly Asp Val Asp Glu Ala Ser Arg Tyr Val Ala Pro
        355                 360                 365

Thr Val Val Asp Asn Val Asp Trp Asp Asp Ser Thr Met Lys Gly Glu
    370                 375                 380

Ile Phe Gly Pro Ile Leu Pro Ile Leu Ser Tyr Ser Ser Leu Asp Glu
385                 390                 395                 400

Ala Leu Ser Arg Leu Gln Ser Arg His Asp Thr Pro Leu Ala Gln Tyr
                405                 410                 415

Ile Phe Thr Gly Gly Ser Thr Ser Arg Ala Lys Asn Pro Lys Leu Asn
            420                 425                 430
```

Lys Ile Ser Gln Gln Ile Arg Ser Gly Gly Ala Val Ile Asn Asp Val
435 440 445

Leu Met His Val Ala Leu Thr Asn Ala Pro Phe Gly Gly Ile Gly Ser
450 455 460

Ser Gly Ser Gly Ser Tyr His Gly Trp Phe Ser Phe Arg Ala Phe Thr
465 470 475 480

His Glu Arg Thr Thr Met Glu Gln Lys Leu Trp Asn Asp Phe Val Leu
485 490 495

Lys Ala Arg Tyr Pro Pro Phe Thr Glu Lys Asn Gln Lys Leu Val Ser
500 505 510

Ala Ser Gln Thr Asp Tyr Asn Gly Lys Val Trp Phe Asp Arg Gln Gly
515 520 525

Asp Val Arg Val Lys Gly Pro Ser Gly Leu Phe Ser Thr Trp Thr Ser
530 535 540

Val Ala Gly Val Ala Ala Leu Ile Tyr Tyr Phe Val Gly Asn Leu
545 550 555

<210> 760
<211> 1440
<212> DNA
<213> Artificial Sequence

<220>
<223> aldA from E. coli

<400> 760
atgtcagtac ccgttcaaca tcctatgtat atcgatggac agtttgttac ctggcgtgga 60

gacgcatgga ttgatgtggt aaaccctgct acagaggctg tcatttcccg catacccgat 120

ggtcaggccg aggatgcccg taaggcaatc gatgcagcag aacgtgcaca accagaatgg 180

gaagcgttgc ctgctattga acgcgccagt tggttgcgca aaatctccgc cgggatccgc 240

gaacgcgcca gtgaaatcag tgcgctgatt gttgaagaag ggggcaagat ccagcagctg 300

gctgaagtcg aagtggcttt tactgccgac tatatcgatt acatggcgga gtgggcacgg 360

cgttacgagg gcgagattat tcaaagcgat cgtccaggag aaaatattct tttgtttaaa 420

cgtgcgcttg gtgtgactac cggcattctg ccgtggaact tcccgttctt cctcattgcc 480

cgcaaaatgg ctcccgctct tttgaccggt aataccatcg tcattaaacc tagtgaattt 540

acgccaaaca atgcgattgc attcgccaaa atcgtcgatg aaataggcct tccgcgcggc 600

gtgtttaacc ttgtactggg gcgtggtgaa accgttgggc aagaactggc gggtaaccca 660

aaggtcgcaa tggtcagtat gacaggcagc gtctctgcag gtgagaagat catggcgact 720

gcggcgaaaa acatcaccaa agtgtgtctg gaattggggg gtaaagcacc agctatcgta 780

atggacgatg ccgatcttga actggcagtc aaagccatcg ttgattcacg cgtcattaat 840

agtgggcaag tgtgtaactg tgcagaacgt gtttatgtac agaaaggcat ttatgatcag 900

ttcgtcaatc ggctgggtga agcgatgcag gcggttcaat ttggtaaccc cgctgaacgc 960

aacgacattg cgatggggcc gttgattaac gccgcggcgc tggaaagggt cgagcaaaaa 1020

gtggcgcgcg cagtagaaga aggggcgaga gtggcgttcg gtggcaaagc ggtagagggg 1080

aaaggatatt attatccgcc gacattgctg ctggatgttc gccaggaaat gtcgattatg 1140

catgaggaaa cctttggccc ggtgctgcca gttgtcgcat ttgacacgct ggaagatgct 1200

atctcaatgg ctaatgacag tgattacggc ctgacctcat caatctatac ccaaaatctg 1260

aacgtcgcga tgaaagccat taaagggctg aagtttggtg aaacttacat caaccgtgaa 1320

aacttcgaag ctatgcaagg cttccacgcc ggatggcgta aatccggtat tggcggcgca 1380

gatggtaaac atggcttgca tgaatatctg cagacccagg tggtttattt acagtcttaa 1440

<210> 761
<211> 479
<212> PRT
<213> Artificial Sequence

<220>
<223> aldA from E. coli

<400> 761

Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val
1 5 10 15

Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu
20 25 30

Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys
35 40 45

Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro
50 55 60

Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg
65 70 75 80

Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile Val Glu Glu Gly Gly Lys
85 90 95

Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile
100 105 110

Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln
115 120 125

Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly
130 135 140

Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala
145 150 155 160

Arg Lys Met Ala Pro Ala Leu Leu Thr Gly Asn Thr Ile Val Ile Lys
165 170 175

```
Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile Ala Phe Ala Lys Ile Val
            180                 185                 190

Asp Glu Ile Gly Leu Pro Arg Gly Val Phe Asn Leu Val Leu Gly Arg
        195                 200                 205

Gly Glu Thr Val Gly Gln Glu Leu Ala Gly Asn Pro Lys Val Ala Met
    210                 215                 220

Val Ser Met Thr Gly Ser Val Ser Ala Gly Glu Lys Ile Met Ala Thr
225                 230                 235                 240

Ala Ala Lys Asn Ile Thr Lys Val Cys Leu Glu Leu Gly Gly Lys Ala
                245                 250                 255

Pro Ala Ile Val Met Asp Asp Ala Asp Leu Glu Leu Ala Val Lys Ala
            260                 265                 270

Ile Val Asp Ser Arg Val Ile Asn Ser Gly Gln Val Cys Asn Cys Ala
        275                 280                 285

Glu Arg Val Tyr Val Gln Lys Gly Ile Tyr Asp Gln Phe Val Asn Arg
    290                 295                 300

Leu Gly Glu Ala Met Gln Ala Val Gln Phe Gly Asn Pro Ala Glu Arg
305                 310                 315                 320

Asn Asp Ile Ala Met Gly Pro Leu Ile Asn Ala Ala Ala Leu Glu Arg
                325                 330                 335

Val Glu Gln Lys Val Ala Arg Ala Val Glu Glu Gly Ala Arg Val Ala
            340                 345                 350

Phe Gly Gly Lys Ala Val Glu Gly Lys Gly Tyr Tyr Tyr Pro Pro Thr
        355                 360                 365

Leu Leu Leu Asp Val Arg Gln Glu Met Ser Ile Met His Glu Glu Thr
    370                 375                 380

Phe Gly Pro Val Leu Pro Val Val Ala Phe Asp Thr Leu Glu Asp Ala
385                 390                 395                 400

Ile Ser Met Ala Asn Asp Ser Asp Tyr Gly Leu Thr Ser Ser Ile Tyr
                405                 410                 415

Thr Gln Asn Leu Asn Val Ala Met Lys Ala Ile Lys Gly Leu Lys Phe
            420                 425                 430

Gly Glu Thr Tyr Ile Asn Arg Glu Asn Phe Glu Ala Met Gln Gly Phe
        435                 440                 445

His Ala Gly Trp Arg Lys Ser Gly Ile Gly Gly Ala Asp Gly Lys His
    450                 455                 460
```

20 Jan 2017 2017200439

Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser
465 470 475

<210> 762
<211> 1539
<212> DNA
<213> Artificial Sequence

<220>
<223> aldB from E. coli

<400> 762
atgaccaata atccccccttc agcacagatt aagcccggcg agtatggttt ccccctcaag 60

ttaaaagccc gctatgacaa ctttattggc ggcgaatggg tagcccctgc cgacggcgag 120

tattaccaga atctgacgcc ggtgaccggg cagctgctgt gcgaagtggc gtcttcgggc 180

aaacgagaca tcgatctggc gctggatgct gcgcacaaag tgaaagataa atgggcgcac 240

acctcggtgc aggatcgtgc ggcgattctg tttaagattg ccgatcgaat ggaacaaaac 300

ctcgagctgt tagcgacagc tgaaacctgg gataacggca aacccattcg cgaaaccagt 360

gctgcggatg taccgctggc gattgaccat ttccgctatt tcgcctcgtg tattcgggcg 420

caggaaggtg ggatcagtga agttgatagc gaaaccgtgg cctatcattt ccatgaaccg 480

ttaggcgtgg tggggcagat tatcccgtgg aacttcccgc tgctgatggc gagctggaaa 540

atggctcccg cgctggcggc gggcaactgt gtggtgctga aacccgcacg tcttaccccg 600

ctttctgtac tgctgctaat ggaaattgtc ggtgatttac tgccgccggg cgtggtgaac 660

gtggtcaatg gcgcaggtgg ggtaattggc gaatatctgg cgacctcgaa acgcatcgcc 720

aaagtggcgt ttaccggctc aacggaagtg ggccaacaaa ttatgcaata cgcaacgcaa 780

aacattattc cggtgacgct ggagttgggc ggtaagtcgc caaatatctt ctttgctgat 840

gtgatggatg aagaagatgc ctttttcgat aaagcgctgg aaggctttgc actgtttgcc 900

tttaaccagg gcgaagtttg cacctgtccg agtcgtgctt tagtgcagga atctatctac 960

gaacgcttta tggaacgcgc catccgccgt gtcgaaagca ttcgtagcgg taacccgctc 1020

gacagcgtga cgcaaatggg cgcgcaggtt tctcacgggc aactggaaac catcctcaac 1080

tacattgata tcggtaaaaa agagggcgct gacgtgctca caggcgggcg gcgcaagctg 1140

ctggaaggtg aactgaaaga cggctactac ctcgaaccga cgattctgtt tggtcagaac 1200

aatatgcggg tgttccagga ggagattttt ggcccggtgc tggcggtgac caccttcaaa 1260

acgatggaag aagcgctgga gctggcgaac gatacgcaat atggcctggg cgcgggcgtc 1320

tggagccgca acggtaatct ggcctataag atggggcgcg gcatacaggc tgggcgcgtg 1380

tggaccaact gttatcacgc ttacccggca catgcggcgt ttggtggcta caaacaatca 1440

ggtatcggtc gcgaaaccca caagatgatg ctggagcatt accagcaaac caagtgcctg 1500

ctggtgagct actcggataa accgttgggg ctgttctga 1539

<210> 763
<211> 512
<212> PRT
<213> Artificial Sequence

<220>
<223> aldB from E. coli

<400> 763

```
Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly
1               5                  10                  15

Phe Pro Leu Lys Leu Lys Ala Arg Tyr Asp Asn Phe Ile Gly Gly Glu
            20                  25                  30

Trp Val Ala Pro Ala Asp Gly Glu Tyr Tyr Gln Asn Leu Thr Pro Val
        35                  40                  45

Thr Gly Gln Leu Leu Cys Glu Val Ala Ser Ser Gly Lys Arg Asp Ile
    50                  55                  60

Asp Leu Ala Leu Asp Ala Ala His Lys Val Lys Asp Lys Trp Ala His
65                  70                  75                  80

Thr Ser Val Gln Asp Arg Ala Ala Ile Leu Phe Lys Ile Ala Asp Arg
                85                  90                  95

Met Glu Gln Asn Leu Glu Leu Leu Ala Thr Ala Glu Thr Trp Asp Asn
            100                 105                 110

Gly Lys Pro Ile Arg Glu Thr Ser Ala Ala Asp Val Pro Leu Ala Ile
        115                 120                 125

Asp His Phe Arg Tyr Phe Ala Ser Cys Ile Arg Ala Gln Glu Gly Gly
    130                 135                 140

Ile Ser Glu Val Asp Ser Glu Thr Val Ala Tyr His Phe His Glu Pro
145                 150                 155                 160

Leu Gly Val Val Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met
                165                 170                 175

Ala Ser Trp Lys Met Ala Pro Ala Leu Ala Ala Gly Asn Cys Val Val
            180                 185                 190

Leu Lys Pro Ala Arg Leu Thr Pro Leu Ser Val Leu Leu Leu Met Glu
        195                 200                 205

Ile Val Gly Asp Leu Leu Pro Pro Gly Val Val Asn Val Val Asn Gly
    210                 215                 220

Ala Gly Gly Val Ile Gly Glu Tyr Leu Ala Thr Ser Lys Arg Ile Ala
225                 230                 235                 240

Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Gln Gln Ile Met Gln
                245                 250                 255

Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys
            260                 265                 270
```

Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe
275 280 285

Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly
290 295 300

Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr
305 310 315 320

Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser
325 330 335

Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His
340 345 350

Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu
355 360 365

Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu
370 375 380

Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn
385 390 395 400

Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val
405 410 415

Thr Thr Phe Lys Thr Met Glu Glu Ala Leu Glu Leu Ala Asn Asp Thr
420 425 430

Gln Tyr Gly Leu Gly Ala Gly Val Trp Ser Arg Asn Gly Asn Leu Ala
435 440 445

Tyr Lys Met Gly Arg Gly Ile Gln Ala Gly Arg Val Trp Thr Asn Cys
450 455 460

Tyr His Ala Tyr Pro Ala His Ala Ala Phe Gly Gly Tyr Lys Gln Ser
465 470 475 480

Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
485 490 495

Thr Lys Cys Leu Leu Val Ser Tyr Ser Asp Lys Pro Leu Gly Leu Phe
500 505 510

<210> 764
<211> 1488
<212> DNA
<213> Artificial Sequence

<220>
<223> aldH(puuC) from E. coli

<400> 764
atgaattttc atcatctggc ttactggcag gataaagcgt taagtctcgc cattgaaaac 60

cgcttattta ttaacggtga atatactgct gcggcggaaa atgaaacctt tgaaaccgtt 120

gatccggtca cccaggcacc gctggcgaaa attgcccgcg gcaagagcgt cgatatcgac 180

cgtgcgatga gcgcagcacg cggcgtattt gaacgcggcg actggtcact ctcttctccg 240

gctaaacgta aagcggtact gaataaactc gccgatttaa tggaagccca cgccgaagag 300

ctggcactgc tggaaactct cgacaccggc aaaccgattc gtcacagtct gcgtgatgat 360

attcccggcg cggcgcgcgc cattcgctgg tacgccgaag cgatcgacaa agtgtatggc 420

gaagtggcga ccaccagtag ccatgagctg gcgatgatcg tgcgtgaacc ggtcggcgtg 480

attgccgcca tcgtgccgtg gaacttcccg ctgttgctga cttgctggaa actcggcccg 540

gcgctggcgg cgggaaacag cgtgattcta aaaccgtctg aaaaatcacc gctcagtgcg 600

attcgtctcg cggggctggc gaaagaagca ggcttgccgg atggtgtgtt gaacgtggtg 660

acgggttttg gtcatgaagc cgggcaggcg ctgtcgcgtc ataacgatat cgacgccatt 720

gcctttaccg gttcaacccg taccgggaaa cagctgctga aagatgcggg cgacagcaac 780

atgaaacgcg tctggctgga agcgggcggc aaaagcgcca acatcgtttt cgctgactgc 840

ccggatttgc aacaggcggc aagcgccacc gcagcaggca ttttctacaa ccagggacag 900

gtgtgcatcg ccggaacgcg cctgttgctg gaagagagca tcgccgatga attcttagcc 960

ctgttaaaac agcaggcgca aaactggcag ccgggccatc cacttgatcc cgcaaccacc 1020

atgggcacct taatcgactg cgcccacgcc gactcggtcc atagctttat tcgggaaggc 1080

gaaagcaaag ggcaactgtt gttggatggc cgtaacgccg ggctggctgc cgccatcggc 1140

ccgaccatct ttgtggatgt ggacccgaat gcgtccttaa gtcgcgaaga gattttcggt 1200

ccggtgctgg tggtcacgcg tttcacatca gaagaacagg cgctacagct tgccaacgac 1260

agccagtacg gccttggcgc ggcggtatgg acgcgcgacc tctcccgcgc gcaccgcatg 1320

agccgacgcc tgaaagccgg ttccgtcttc gtcaataact acaacgacgg cgatatgacc 1380

gtgccgtttg gcggctataa gcagagcggc aacggtcgcg acaaatccct gcatgccctt 1440

gaaaaattca ctgaactgaa aaccatctgg ataagcctgg aggcctga 1488

<210> 765
<211> 495
<212> PRT
<213> Artificial Sequence

<220>
<223> aldH(puuC) from E. coli

<400> 765

Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu
1 5 10 15

Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala
20 25 30

Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu
35 40 45

2017200439 20 Jan 2017

```
Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser
    50                  55                  60

Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro
65                  70                  75                  80

Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala
                85                  90                  95

His Ala Glu Glu Leu Ala Leu Leu Glu Thr Leu Asp Thr Gly Lys Pro
            100                 105                 110

Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile
        115                 120                 125

Arg Trp Tyr Ala Glu Ala Ile Asp Lys Val Tyr Gly Glu Val Ala Thr
    130                 135                 140

Thr Ser Ser His Glu Leu Ala Met Ile Val Arg Glu Pro Val Gly Val
145                 150                 155                 160

Ile Ala Ala Ile Val Pro Trp Asn Phe Pro Leu Leu Leu Thr Cys Trp
                165                 170                 175

Lys Leu Gly Pro Ala Leu Ala Ala Gly Asn Ser Val Ile Leu Lys Pro
            180                 185                 190

Ser Glu Lys Ser Pro Leu Ser Ala Ile Arg Leu Ala Gly Leu Ala Lys
        195                 200                 205

Glu Ala Gly Leu Pro Asp Gly Val Leu Asn Val Val Thr Gly Phe Gly
    210                 215                 220

His Glu Ala Gly Gln Ala Leu Ser Arg His Asn Asp Ile Asp Ala Ile
225                 230                 235                 240

Ala Phe Thr Gly Ser Thr Arg Thr Gly Lys Gln Leu Leu Lys Asp Ala
                245                 250                 255

Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser
            260                 265                 270

Ala Asn Ile Val Phe Ala Asp Cys Pro Asp Leu Gln Gln Ala Ala Ser
        275                 280                 285

Ala Thr Ala Ala Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala
    290                 295                 300

Gly Thr Arg Leu Leu Leu Glu Glu Ser Ile Ala Asp Glu Phe Leu Ala
305                 310                 315                 320

Leu Leu Lys Gln Gln Ala Gln Asn Trp Gln Pro Gly His Pro Leu Asp
                325                 330                 335
```

```
Pro Ala Thr Thr Met Gly Thr Leu Ile Asp Cys Ala His Ala Asp Ser
            340                 345                 350

Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu
        355                 360                 365

Asp Gly Arg Asn Ala Gly Leu Ala Ala Ala Ile Gly Pro Thr Ile Phe
    370                 375                 380

Val Asp Val Asp Pro Asn Ala Ser Leu Ser Arg Glu Glu Ile Phe Gly
385                 390                 395                 400

Pro Val Leu Val Val Thr Arg Phe Thr Ser Glu Glu Gln Ala Leu Gln
                405                 410                 415

Leu Ala Asn Asp Ser Gln Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg
            420                 425                 430

Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser
        435                 440                 445

Val Phe Val Asn Asn Tyr Asn Asp Gly Asp Met Thr Val Pro Phe Gly
    450                 455                 460

Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu
465                 470                 475                 480

Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala
                485                 490                 495
```

<210> 766
<211> 725
<212> DNA
<213> Artificial Sequence

<220>
<223> UAS (PKG1)-FAB1

<400> 766

```
aattaccgtc gctcgtgatt tgtttgcaaa aagaacaaaa ctgaaaaaac ccagacacgc       60

tcgacttcct gtcttcctat tgattgcagc ttccaatttc gtcacacaac aaggtcctgt      120

cgacgcctac ttggcttcac atacgttgca tacgtcgata tagataataa tgataatgac      180

agcaggatta tcgtaatacg taatagttga aaatctcaaa aatgtgtggg tcattacgta      240

aataatgata ggaatgggat tcttctattt ttccttttc cattctagca gccgtcggga      300

aaacgtggca tcctctcttt cgggctcaat tggagtcacg ctgccgtgag catcctctct      360

ttccatatct aacaactgag cacgtaacca atggaaaagc atgagcttag cgttgctcca      420

aaaaagtatt ggatggttaa taccatttgt ctgttctctt ctgactttga ctcctcaaaa      480

aaaaaaaatc tacaatcaac agatcgcttc aattacgccc tcacaaaaac ttttttcctt      540

cttcttcgcc cacgttaaat tttatccctc atgttgtcta acggatttct gcacttgatt      600

tattataaaa agacaaagac ataatacttc tctatcaatt tcagttattg ttcttccttg      660

cgttattctt ctgttcttct ttttcttttg tcatatataa ccataaccaa gtaatacata      720
```

ttcaa 725

<210> 767
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> T-FBA1(SalI) primer

<400> 767
attctacgta cgtcgacgcc tacttggctt cacatacgtt gcatacg 47

<210> 768
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> B-FBA1(SpeI) primer

<400> 768
gtatcaaata ctagtttgaa tatgtattac ttggttatgg ttatatatg 49

<210> 769
<211> 10494
<212> DNA
<213> Artificial Sequence

<220>
<223> pWS358-PGK1p-GUS plasmid

<400> 769
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240

gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300

ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360

ttttttttt cccctagcgg atgactcttt ttttttctta gcgattggca ttatcacata 420

atgaattata cattatataa agtaatgtga tttcttcgaa gaatatacta aaaaatgagc 480

aggcaagata aacgaaggca aagatgacag agcagaaagc cctagtaaag cgtattacaa 540

atgaaaccaa gattcagatt gcgatctctt taaagggtgg tcccctagcg atagagcact 600

cgatcttccc agaaaaagag gcagaagcag tagcagaaca ggccacacaa tcgcaagtga 660

ttaacgtcca cacaggtata gggtttctgg accatatgct agggattcat aaccattttc 720

tcaatcgaat tacacagaac acaccgtaca aacctctcta tcataactac ttaatagtca 780

cacacgtact cgtctaaata cacatcatcg tcctacaagt tcatcaaagt gttggacaga 840

caactatacc agcatggatc tcttgtatcg gttcttttct cccgctctct cgcaataaca 900

atgaacactg ggtcaatcat agcctacaca ggtgaacaga gtagcgttta tacagggttt 960

atacggtgat tcctacggca aaaatttttc atttctaaaa aaaaaaagaa aaatttttct 1020

ttccaacgct agaaggaaaa gaaaaatcta attaaattga tttggtgatt ttctgagagt 1080

tccctttttc atatatcgaa ttttgaatat aaaaggagat cgaaaaaatt tttctattca 1140
atctgttttc tggttttatt tgatagtttt tttgtgtatt attattatgg attagtactg 1200
gtttatatgg gtttttctgt ataacttctt tttattttag tttgtttaat cttattttga 1260
gttacattat agttccctaa ctgcaagaga agtaacatta aaactcgaga tgggtaagga 1320
aaagactcac gtttcgaggc cgcgattaaa ttccaacatg gatgctgatt tatatgggta 1380
taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa 1440
gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca atgatgttac 1500
agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca 1560
ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg gcaaaacagc 1620
attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt 1680
gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt 1740
atttcgtctc gctcaggcgc aatcacgaat gaataacggt ttggttgatg cgagtgattt 1800
tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataagctttt 1860
gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt 1920
tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata 1980
ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg 2040
gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat 2100
gctcgatgag tttttctaag tttaacttga tactactaga ttttttctct tcatttataa 2160
aatttttggt tataattgaa gctttagaag tatgaaaaaa tccttttttt tcattctttg 2220
caaccaaaat aagaagcttc ttttattcat tgaaatgatg aatataaacc taacaaaaga 2280
aaaagactcg aatatcaaac attaaaaaaa aataaaagag gttatctgtt ttcccattta 2340
gttggagttt gcattttcta atagatagaa ctctcaatta atgtggattt agtttctctg 2400
ttcgctgcag catacgatat atatacatgt gtatatatgt atacctatga atgtcagtaa 2460
gtatgtatac gaacagtatg atactgaaga tgacaaggta atgcatcatt ctatacgtgt 2520
cattctgaac gaggcgcgct ttccttttt cttttgcttt tttctttttt tttctcttga 2580
actcgacgga tctatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat 2640
caggaaattg taaacgttaa tattttgtta aaattcgcgt taaatttttg ttaaatcagc 2700
tcatttttta accaataggc cgaaatcggc aaaatccctt ataaatcaaa agaatagacc 2760
gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa gaacgtggac 2820
tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg tgaaccatca 2880
ccctaatcaa gtttttttggg gtcgaggtgc cgtaaagcac taaatcggaa ccctaaaggg 2940
agcccccgat ttagagcttg acggggaaag ccggcgaacg tggcgagaaa ggaagggaag 3000
aaagcgaaag gagcgggcgc tagggcgctg gcaagtgtag cggtcacgct gcgcgtaacc 3060
accacacccg ccgcgcttaa tgcgccgcta cagggcgcgt cgcgccattc gccattcagg 3120
ctgcgcaact gttgggaagg gcgatcggtg cgggcctctt cgctattacg ccagctggcg 3180
aaaggggat gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcacga 3240

cgttgtaaaa cgacggccag tgagcgcgcg taatacgact cactataggg cgaattgggt 3300
accgggcccc ccctcgaggt cgacgtgagt aaggaaagag tgaggaacta tcgcatacct 3360
gcatttaaag atgccgattt gggcgcgaat cctttatttt ggcttcaccc tcatactatt 3420
atcagggcca gaaaaaggaa gtgtttccct ccttcttgaa ttgatgttac cctcataaag 3480
cacgtggcct cttatcgaga aagaaattac cgtcgctcgt gatttgtttg caaaaagaac 3540
aaaactgaaa aaacccagac acgctcgact tcctgtcttc ctattgattg cagcttccaa 3600
tttcgtcaca caacaaggtc ctagcgacgg ctcacaggtt ttgtaacaag caatcgaagg 3660
ttctggaatg gcgggaaagg gtttagtacc acatgctatg atgcccactg tgatctccag 3720
agcaaagttc gttcgatcgt actgttactc tctctctttc aaacagaatt gtccgaatcg 3780
tgtgacaaca acagcctgtt ctcacacact cttttcttct aaccaagggg gtggtttagt 3840
ttagtagaac ctcgtgaaac ttacatttac atatatataa acttgcataa attggtcaat 3900
gcaagaaata catatttggt cttttctaat tcgtagtttt tcaagttctt agatgctttc 3960
tttttctctt ttttacagat catcaaggaa gtaattatct actttttaca acaaatataa 4020
aacaactagt atggtacgtc ctgtagaaac cccaacccgt gaaatcaaaa aactcgacgg 4080
cctgtgggca ttcagtctgg atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag 4140
cgcgttacaa gaaagccggg caattgctgt gccaggcagt tttaacgatc agttcgccga 4200
tgcagatatt cgtaattatg cgggcaacgt ctggtatcag cgcgaagtct ttataccgaa 4260
aggttgggca ggccagcgta tcgtgctgcg tttcgatgcg gtcactcatt acggcaaagt 4320
gtgggtcaat aatcaggaag tgatggagca tcagggcggc tatactccat ttgaagccga 4380
tgtcacgccg tatgttattg ccgggaaaag tgtacgtatc accgtttgtg tgaacaacga 4440
actgaactgg cagactatcc cgccgggaat ggtgattacc gacgaaaacg gcaagaaaaa 4500
gcagtcttac ttccatgatt tctttaacta tgccggaatc catcgcagcg taatgctcta 4560
caccacgccg aacacctggg tggacgatat caccgtggtg acgcatgtcg cgcaagactg 4620
taaccacgcg tctgttgact ggcaggtggt ggccaatggt gatgtcagcg ttgaactgcg 4680
tgatgcggat caacaggtgg ttgcaactgg acaaggcact agcgggactt tgcaagtggt 4740
gaatccgcac ctctggcaac cgggtgaagg ttatctctat gaactgtgcg tcacagccaa 4800
aagccagaca gagtgtgata tctacccgct tcgcgtcggc atccggtcag tggcagtgaa 4860
gggcgaacag ttcctgatta accacaaacc gttctacttt actggctttg gtcgtcatga 4920
agatgcggac ttgcgtggca aaggattcga taacgtgctg atggtgcacg accacgcatt 4980
aatggactgg attggggcca actcctaccg tacctcgcat tacccttacg ctgaagagat 5040
gctcgactgg gcagatgaac atggcatcgt ggtgattgat gaaactgctg ctgtcggctt 5100
taacctctct ttaggcattg gtttcgaagc gggcaacaag ccgaaagaac tgtacagcga 5160
agaggcagtc aacggggaaa ctcagcaagc gcacttacag gcgattaaag agctgatagc 5220
gcgtgacaaa aaccacccaa gcgtggtgat gtggagtatt gccaacgaac cggatacccg 5280
tccgcaaggt gcacgggaat atttcgcgcc actggcggaa gcaacgcgta aactcgaccc 5340
gacgcgtccg atcacctgcg tcaatgtaat gttctgcgac gctcacaccg ataccatcag 5400

cgatctcttt gatgtgctgt gcctgaaccg ttattacgga tggtatgtcc aaagcggcga 5460
tttggaaacg gcagagaagg tactggaaaa agaacttctg gcctggcagg agaaactgca 5520
tcagccgatt atcatcaccg aatacggcgt ggatacgtta gccgggctgc actcaatgta 5580
caccgacatg tggagtgaag agtatcagtg tgcatggctg gatatgtatc accgcgtctt 5640
tgatcgcgtc agcgccgtcg tcggtgaaca ggtatggaat ttcgccgatt ttgcgacctc 5700
gcaaggcata ttgcgcgttg gcggtaacaa gaaagggatc ttcactcgcg accgcaaacc 5760
gaagtcggcg gcttttctgc tgcaaaaacg ctggactggc atgaacttcg gtgaaaaacc 5820
gcagcaggga ggcaaacaat gattaattaa ctagagcggc cgcgttaatt caaattaatt 5880
gatatagttt tttaatgagt attgaatctg tttagaaata atggaatatt atttttattt 5940
atttatttat attattggtc ggctcttttc ttctgaaggt caatgacaaa atgatatgaa 6000
ggaaataatg atttctaaaa ttttacaacg taagatattt ttacaaaagc ctagctcatc 6060
ttttgtcatg cactatttta ctcacgcttg aaattaacgg ccagtccact gcggagtcat 6120
ttcaaagtca tcctaatcga tctatcgttt ttgatagctc attttggagt tcgcgattgt 6180
cttctgttat tcacaactgt tttaattttt atttcattct ggaactcttc gagttctttg 6240
taaagtcttt catagtagct tactttatcc tccaacatat ttaacttcat gtcaatttcg 6300
gctcttaaat tttccacatc atcaagttca acatcatctt ttaacttgaa tttattctct 6360
agctcttcca accaagcctc attgctcctt gatttactgg tgaaaagtga tacactttgc 6420
gcgcaatcca ggtcaaaact ttcctgcaaa gaattcacca atttctcgac atcatagtac 6480
aatttgtttt gttctcccat cacaatttaa tatacctgat ggattcttat gaagcgctgg 6540
gtaatggacg tgtcactcta cttcgccttt ttccctactc cttttagtac ggaagacaat 6600
gctaataaat aagagggtaa taataatatt attaatcggc aaaaaagatt aaacgccaag 6660
cgtttaatta tcagaaagca aacgtcgtac caatccttga atgcttccca attgtatatt 6720
aagagtcatc acagcaacat attcttgtta ttaaattaat tattattgat ttttgatatt 6780
gtataaaaaa accaaatatg tataaaaaaa gtgaataaaa aataccaagt atggagaaat 6840
atattagaag tctatacgtt aaaccaccgc ggtggagctc cagcttttgt tccctttagt 6900
gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct gtttcctgtg tgaaattgtt 6960
atccgctcac aattccacac aacataggag ccggaagcat aaagtgtaaa gcctggggtg 7020
cctaatgagt gaggtaactc acattaattg cgttgcgctc actgcccgct ttccagtcgg 7080
gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga ggcggtttgc 7140
gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc gttcggctgc 7200
ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa tcaggggata 7260
acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt aaaaaggccg 7320
cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa aatcgacgct 7380
caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttt ccccctggaa 7440
gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg tccgcctttc 7500
tcccttcggg aagcgtggcg ctttctcata gctcacgctg taggtatctc agttcggtgt 7560

aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc gaccgctgcg 7620
ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta tcgccactgg 7680
cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct acagagttct 7740
tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc tgcgctctgc 7800
tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa caaaccaccg 7860
ctggtagcgg tggtttttt gtttgcaagc agcagattac gcgcagaaaa aaaggatctc 7920
aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa aactcacgtt 7980
aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt ttaaattaaa 8040
aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac agttaccaat 8100
gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc atagttgcct 8160
gactccccgt cgtgtagata actacgatac gggagggctt accatctggc cccagtgctg 8220
caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata aaccagccag 8280
ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc cagtctatta 8340
attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc aacgttgttg 8400
ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca ttcagctccg 8460
gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgcaaaaaa gcggttagct 8520
ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca ctcatggtta 8580
tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt tctgtgactg 8640
gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt tgctcttgcc 8700
cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg ctcatcattg 8760
gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga tccagttcga 8820
tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc agcgtttctg 8880
ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg acacggaaat 8940
gttgaatact catactcttc ctttttcaat attattgaag catttatcag ggttattgtc 9000
tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg gttccgcgca 9060
catttccccg aaaagtgcca cctgaacgaa gcatctgtgc ttcattttgt agaacaaaaa 9120
tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt tacagaacag 9180
aaatgcaacg cgaaagcgct attttaccaa cgaagaatct gtgcttcatt tttgtaaaac 9240
aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag 9300
aacagaaatg caacgcgaga gcgctatttt accaacaaag aatctatact tctttttgt 9360
tctacaaaaa tgcatcccga gagcgctatt tttctaacaa agcatcttag attacttttt 9420
ttctcctttg tgcgctctat aatgcagtct cttgataact ttttgcactg taggtccgtt 9480
aaggttagaa gaaggctact ttggtgtcta ttttctcttc cataaaaaaa gcctgactcc 9540
acttcccgcg tttactgatt actagcgaag ctgcgggtgc atttttcaa gataaaggca 9600
tccccgatta tattctatac cgatgtggat tgcgcatact ttgtgaacag aaagtgatag 9660
cgttgatgat tcttcattgg tcagaaaatt atgaacggtt tcttctattt tgtctctata 9720

tactacgtat aggaaatgtt tacattttcg tattgttttc gattcactct atgaatagtt 9780

cttactacaa ttttttttgtc taaagagtaa tactagagat aaacataaaa aatgtagagg 9840

tcgagtttag atgcaagttc aaggagcgaa aggtggatgg gtaggttata tagggatata 9900

gcacagagat atatagcaaa gagatacttt tgagcaatgt ttgtggaagc ggtattcgca 9960

atattttagt agctcgttac agtccggtgc gtttttggtt ttttgaaagt gcgtcttcag 10020

agcgcttttg gttttcaaaa gcgctctgaa gttcctatac tttctagaga ataggaactt 10080

cggaatagga acttcaaagc gtttccgaaa acgagcgctt ccgaaaatgc aacgcgagct 10140

gcgcacatac agctcactgt tcacgtcgca cctatatctg cgtgttgcct gtatatatat 10200

atacatgaga agaacggcat agtgcgtgtt tatgcttaaa tgcgtactta tatgcgtcta 10260

tttatgtagg atgaaaggta gtctagtacc tcctgtgata ttatcccatt ccatgcgggg 10320

tatcgtatgc ttccttcagc actacccttt agctgttcta tatgctgcca ctcctcaatt 10380

ggattagtct catccttcaa tgctatcatt tcctttgata ttggatcatc taagaaacca 10440

ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt cgtc 10494

<210> 770
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> T-U/PGK1(KpnI) primer

<400> 770
actacagatg gtaccaatta ccgtcgctcg tgatttgttt gca 43

<210> 771
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> B-U/PGK1(SalI) primer

<400> 771
agcatccttg tcgacaggac cttgttgtgt gacgaaattg gaagc 45

<210> 772
<211> 7533
<212> DNA
<213> Artificial Sequence

<220>
<223> pYZ067DkivDDilvD plasmid

<400> 772
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240

gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tcttttttcta 300

ttactcttgg cctcctctag tacactctat atttttttat gcctcggtaa tgattttcat 360

```
ttttttttt ccacctagcg gatgactctt ttttttctt agcgattggc attatcacat      420
aatgaattat acattatata aagtaatgtg atttcttcga agaatatact aaaaaatgag      480
caggcaagat aaacgaaggc aaagatgaca gagcagaaag ccctagtaaa gcgtattaca      540
aatgaaacca agattcagat tgcgatctct ttaaagggtg gtcccctagc gatagagcac      600
tcgatcttcc cagaaaaaga ggcagaagca gtagcagaac aggccacaca atcgcaagtg      660
attaacgtcc acacaggtat agggtttctg gaccatatga tacatgctct ggccaagcat      720
tccggctggt cgctaatcgt tgagtgcatt ggtgacttac acatagacga ccatcacacc      780
actgaagact gcgggattgc tctcggtcaa gcttttaaag aggccctagg ggccgtgcgt      840
ggagtaaaaa ggtttggatc aggatttgcg cctttggatg aggcactttc cagagcggtg      900
gtagatcttt cgaacaggcc gtacgcagtt gtcgaacttg gtttgcaaag ggagaaagta      960
ggagatctct cttgcgagat gatcccgcat tttcttgaaa gctttgcaga ggctagcaga     1020
attaccctcc acgttgattg tctgcgaggc aagaatgatc atcaccgtag tgagagtgcg     1080
ttcaaggctc ttgcggttgc cataagagaa gccacctcgc ccaatggtac caacgatgtt     1140
ccctccacca aaggtgttct tatgtagtga caccgattat ttaaagctgc agcatacgat     1200
atatatacat gtgtatatat gtatacctat gaatgtcagt aagtatgtat acgaacagta     1260
tgatactgaa gatgacaagg taatgcatca ttctatacgt gtcattctga acgaggcgcg     1320
ctttcctttt ttctttttgc tttttctttt tttttctctt gaactcgacg gatctatgcg     1380
gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggaaat tgtaagcgtt     1440
aatattttgt taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag     1500
gccgaaatcg gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt     1560
gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga     1620
aaaaccgtct atcagggcga tggcccacta cgtggccggg atccactagt tctagagcgg     1680
ccgctctaga actagtacca caggtgttgt cctctgagga cataaaatac acaccgagat     1740
tcatcaactc attgctggag ttagcatatc tacaattggg tgaaatgggg agcgatttgc     1800
aggcatttgc tcggcatgcc ggtagaggtg tggtcaataa gagcgacctc atgctatacc     1860
tgagaaagca acctgaccta caggaaagag ttactcaaga ataagaattt tcgttttaaa     1920
acctaagagt cactttaaaa tttgtataca cttatttttt ttataactta tttaataata     1980
aaaatcataa atcataagaa attcgcttac tcttaattaa tcaaaaagtt aaaattgtac     2040
gaatagattc accacttctt aacaaatcaa acccttcatt gattttctcg aatggcaata     2100
catgtgtaat taaaggatca agagcaaact tcttcgccat aaagtcggca acaagttttg     2160
gaacactatc cttgctctta aaaccgccaa atatagctcc cttccatgta cgaccgctta     2220
gcaacagcat aggattcatc gacaaatttt gtgaatcagg aggaacacct acgatcacac     2280
tgactccata tgcctcttga cagcaggaca acgcagttac catagtatca agacggccta     2340
taacttcaaa agagaaatca actccaccgt ttgacatttc agtaaggact tcttgtattg     2400
gtttcttata atcttgaggg ttaacacatt cagtagcccc gacctcctta gcttttgcaa     2460
atttgtcctt attgatgtct acacctataa tcctcgctgc gcctgcagct ttacacccca     2520
```

taataacgct tagtcctact cctcctaaac cgaatactgc acaagtcgaa ccctgtgtaa 2580
cctttgcaac tttaactgcg gaaccgtaac cggtggaaaa tccgcaccct atcaagcaaa 2640
ctttttccag tggtgaagct gcatcgattt tagcgacaga tatctcgtcc accactgtgt 2700
attgggaaaa tgtagaagta ccaaggaaat ggtgtatagg tttccctctg catgtaaatc 2760
tgcttgtacc atcctgcata gtacctctag gcatagacaa atcattttta aggcagaaat 2820
taccctcagg atgtttgcag actctacact taccacattg aggagtgaac agtgggatca 2880
ctttatcacc aggacgaaca gtggtaacac cttcacctat ggattcaacg attccggcag 2940
cctcgtgtcc cgcgattact ggcaaaggag taactagagt gccactcacc acatggtcgt 3000
cggatctaca gattccggtg gcaaccatct tgattctaac ctcgtgtgct tttggtggcg 3060
ctacttctac ttcttctatg ctaaacggct ttttctcttc ccacaaaact gccgctttac 3120
acttaataac tttaccggct gttgacatcc tcagctagct attgtaatat gtgtgtttgt 3180
ttggattatt aagaagaata attacaaaaa aaattacaaa ggaaggtaat tacaacagaa 3240
ttaagaaagg acaagaagga ggaagagaat cagttcatta tttcttcttt gttatataac 3300
aaacccaagt agcgatttgg ccatacatta aaagttgaga accaccctcc ctggcaacag 3360
ccacaactcg ttaccattgt tcatcacgat catgaaactc gctgtcagct gaaatttcac 3420
ctcagtggat ctctcttttt attcttcatc gttccactaa ccttttcca tcagctggca 3480
gggaacggaa agtggaatcc catttagcga gcttcctctt ttcttcaaga aaagacgaag 3540
cttgtgtgtg ggtgcgcgcg ctagtatctt tccacattaa gaaatatacc ataaaggtta 3600
cttagacatc actatggcta tatatatata tatatatata tgtaacttag caccatcgcg 3660
cgtgcatcac tgcatgtgtt aaccgaaaag tttggcgaac acttcaccga cacggtcatt 3720
tagatctgtc gtctgcattg cacgtccctt agccttaaat cctaggcggg agcattctcg 3780
tgtaattgtg cagcctgcgt agcaactcaa catagcgtag tctacccagt ttttcaaggg 3840
tttatcgtta gaagattctc ccttttcttc ctgctcacaa atcttaaagt catacattgc 3900
acgactaaat gcaagcgcca gcttttgttc cctttagtga gggttaattg cgcgcttggc 3960
gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa 4020
catacgagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac 4080
attaattgcg ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca 4140
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct cttccgcttc 4200
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc 4260
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 4320
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 4380
gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 4440
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 4500
tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 4560
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 4620
ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 4680

tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat 4740
tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg 4800
ctacactaga agaacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 4860
aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt 4920
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 4980
tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 5040
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 5100
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 5160
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 5220
tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg 5280
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 5340
tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 5400
aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt 5460
gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt 5520
tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt 5580
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 5640
tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 5700
ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac 5760
cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa 5820
actctcaagg atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa 5880
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 5940
aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 6000
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 6060
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 6120
tgaacgaagc atctgtgctt cattttgtag aacaaaaatg caacgcgaga gcgctaattt 6180
ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat 6240
tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgagagcgct 6300
aatttttcaa acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgagagc 6360
gctattttac caacaaagaa tctatacttc ttttttgttc tacaaaaatg catcccgaga 6420
gcgctatttt tctaacaaag catcttagat tacttttttt ctcctttgtg cgctctataa 6480
tgcagtctct tgataacttt ttgcactgta ggtccgttaa ggttagaaga aggctacttt 6540
ggtgtctatt ttctcttcca taaaaaaagc ctgactccac ttcccgcgtt tactgattac 6600
tagcgaagct gcgggtgcat tttttcaaga taaaggcatc cccgattata ttctataccg 6660
atgtggattg cgcatacttt gtgaacagaa agtgatagcg ttgatgattc ttcattggtc 6720
agaaaattat gaacggtttc ttctattttg tctctatata ctacgtatag gaaatgttta 6780
cattttcgta ttgttttcga ttcactctat gaatagttct tactacaatt tttttgtcta 6840

aagagtaata ctagagataa acataaaaaa tgtagaggtc gagtttagat gcaagttcaa 6900
ggagcgaaag gtggatgggt aggttatata gggatatagc acagagatat atagcaaaga 6960
gatacttttg agcaatgttt gtggaagcgg tattcgcaat attttagtag ctcgttacag 7020
tccggtgcgt ttttggtttt ttgaaagtgc gtcttcagag cgcttttggt tttcaaaagc 7080
gctctgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttcaaagcgt 7140
ttccgaaaac gagcgcttcc gaaaatgcaa cgcgagctgc gcacatacag ctcactgttc 7200
acgtcgcacc tatatctgcg tgttgcctgt atatatatat acatgagaag aacggcatag 7260
tgcgtgttta tgcttaaatg cgtacttata tgcgtctatt tatgtaggat gaaaggtagt 7320
ctagtacctc ctgtgatatt atcccattcc atgcggggta tcgtatgctt ccttcagcac 7380
taccctttag ctgttctata tgctgccact cctcaattgg attagtctca tccttcaatg 7440
ctatcatttc ctttgatatt ggatcatact aagaaaccat tattatcatg acattaacct 7500
ataaaaatag gcgtatcacg aggccctttc gtc 7533

<210> 773
<211> 12518
<212> DNA
<213> Artificial Sequence

<220>
<223> pK9G9.OLE1p.ilvD plasmid

<400> 773
tcccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60
aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120
ttacttcacc acccttttatt tcaggctgat atcttagcct tgttactaga ttaatcatgt 180
aattagttat gtcacgctta cattcacgcc ctccccccac atccgctcta accgaaaagg 240
aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 300
taagaacgtt atttatattt caaatttttc ttttttttct gtacagacgc gtgtacgcat 360
gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt 420
gcgggcggcc gcacctggta aaacctctag tggagtagta gatgtaatca atgaagcgga 480
agccaaaaga ccagagtaga ggcctataga agaaactgcg atacctttg tgatggctaa 540
acaaacagac atctttttat atgtttttac ttctgtatat cgtgaagtag taagtgataa 600
gcgaatttgg ctaagaacgt tgtaagtgaa caagggacct cttttgcctt tcaaaaaagg 660
attaaatgga gttaatcatt gagatttagt tttcgttaga ttctgtatcc ctaaataact 720
cccttacccg acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga 780
ccaaataata ctagagttaa ctgatggtct taaacaggca ttacgtggtg aactccaaga 840
ccaatataca aaatatcgat aagttattct tgcccaccaa tttaaggagc ctacatcagg 900
acagtagtac cattcctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt 960
atataactta gcccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata 1020
cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgccccta tctgttcttc 1080
cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca 1140

ataaggctag ttcgaatgat gaacttgctt gctgtcaaac ttctgagttg ccgctgatgt 1200
gacactgtga caataaattc aaaccggtta tagcggtctc ctccggtacc ggttctgcca 1260
cctccaatag agctcagtag gagtcagaac ctctgcggtg gctgtcagtg actcatccgc 1320
gtttcgtaag ttgtgcgcgt gcacatttcg cccgttcccg ctcatcttgc agcaggcgga 1380
aattttcatc acgctgtagg acgcaaaaaa aaaataatta atcgtacaag aatcttggaa 1440
aaaaaattga aaaattttgt ataaaaggga tgacctaact tgactcaatg gcttttacac 1500
ccagtatttt ccctttcctt gtttgttaca attatagaag caagacaaaa acatatagac 1560
aacctattcc taggagttat attttttac cctaccagca atataagtaa aaaactgttt 1620
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 1680
cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 1740
cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgcgaagga 1800
gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 1860
ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 1920
catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca 1980
tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 2040
aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 2100
cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 2160
tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 2220
cggtgaacaa gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 2280
cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 2340
gttgatcgtt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 2400
cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 2460
ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt 2520
tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 2580
acacccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 2640
caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 2700
ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 2760
catttaagat ttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt 2820
ctgaagcagc ttcaaatata tatatttttt acatatttat tatgattcaa tgaacaatct 2880
aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt 2940
gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggctga gctactggtt 3000
gctgtgacgc agcggcatta gcgcggcgtt atgagctacc ctcgtggcct gaaagatggc 3060
gggaataaag cggaactaaa aattactgac tgagccatat tgaggtcaat ttgtcaactc 3120
gtcaagtcac gtttggtgga cggccccttt ccaacgaatc gtatatacta acatgcgcgc 3180
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca 3240
cctgtattta atttccttac tcgcgggttt ttctttttttc tcaattcttg gcttcctctt 3300

tctcgagcgg accggatcct cgcgaactcc aaaatgagct atcaaaaacg atagatcgat 3360
taggatgact ttgaaatgac tccgcagtgg actggccgtt aatttcaagc gtgagtaaaa 3420
tagtgcatga caaaagatga gctaggcttt tgtaaaaata tcttacgttg taaaatttta 3480
gaaatcatta tttccttcat atcattttgt cattgacctt cagaagaaaa gagccgacca 3540
ataatataaa taaataaata aaaataatat tccattattt ctaaacagat tcaatactca 3600
ttaaaaaact atatcaatta atttgaatta acgcggccgc ttaaccacag caaccaggac 3660
aacatttttt gccagtttct tcaggcttcc aaaagtctgt tacggctccc ctagaagcag 3720
acgaaacgat gtgagcatat ttaccaagga taccgcgtga atagagcggt ggcaattcaa 3780
tggtctcttg acgatgtttt aactcttcat cggagatatc aaagtgtaat tccttagtgt 3840
cttggtcaat agtgactatg tctcctgttt gcaggtaggc gattggaccg ccatcttgtg 3900
cttcaggagc gatatgaccc acgacaagac cataagtacc acctgagaag cggccatctg 3960
tcagaagggc aacttttca ccttgccctt taccaacaat cattgatgaa agggaaagca 4020
tttcaggcat accaggaccg ccctttggtc ctacaaaacg tacgacaaca acatcaccat 4080
caacaatatc atcattcaag acagcttcaa tggcttcttc ttcagaatta aagaccttag 4140
caggaccgac atgacgacgc acttttacac cagaaacttt ggcaacggca ccgtctggag 4200
ccaagttacc atggagaata atgagcggac catcttcacg tttaggattt tcaagcggca 4260
taataacctt ttgaccaggt gttaaatcat caaaagcctt caaattttca gcgactgttt 4320
tgccagtaca agtgatacgg tcaccatgaa ggaagccatt tttaaggaga tatttcataa 4380
ctgctggtac ccctccgacc ttgtaaaggt cttggaatac atattgacca gaaggtttca 4440
aatcagccaa atgaggaact ttttcttgga aagtattgaa atcatcaagt gtcaattcca 4500
cattagcagc atgggcaata gctaagaggt gaagggttga gttggttgaa cctcccagag 4560
ccatagttac agtaatagca tcttcaaaag cttcacgcgt taaaatgtca gaaggtttta 4620
agcccatttc gagcattttg acaacagcgc gaccagcttc ttcaatatct gctttctttt 4680
ctgcggattc agccgggtga gaagatgaac ccggaaggct aagtcccaaa acttcaatag 4740
ctgtcgccat tgtgttagca gtatacatac caccgcagcc tccaggaccg ggacaagcat 4800
tacattccaa agctttaact tcttctttgg tcatatcgcc gtggttccaa tggccgacac 4860
cttcaaagac agagactaaa tcgatatctt tgccgtctaa attaccaggt gcaattgttc 4920
cgccgtaagc aaaaatggct gggatatcca tgttagccat agcgataaca gaaccgggca 4980
tgtttttatc acaaccgcca atggctacaa aagcatccgc attatgacct cccatggctg 5040
cttcaataga atctgcaata atatcacgag atgtcaagga gaaacgcatt ccttgggttc 5100
ccatggcgat tccatcagaa accgtgattg ttccgaactg aactggccaa gcaccagctt 5160
ccttaacacc gactttggct agtttaccaa agtcatgtaa gtggatatta caaggtgtgt 5220
tttcagccca agttgaaatg acaccgacga taggtttttc aaagtcttca tcttgcatac 5280
cagttgcacg caacatagca cgattaggtg atttaaccat tgaatcgtaa acagaactac 5340
gatttcttaa gtctttaaga gttttttgt cagtcatact cacgtgcttt gttgtaatgt 5400
tttagtgctg tttataatat gatcaccaca actatctatt actatgatgt tctattctac 5460

gtaatacaaa atataaacgg aaacagaagt aggaaagatg gaaatagaac aataaatgaa 5520
tcaagatctg cccccatata tatatgtata tgctgatttg caagactcga tgagccagga 5580
gccgatgatt tgctgcatat attgttaact actattattt ccacctttgt gtgccatccc 5640
catagccgta acaataggga taggtgtgtc tgagtgagca agactcgtag aagcacacct 5700
ggttgggcac tagataaggt ttgttgagtg ttcaacgtcc gaaagaaagc tgccgactat 5760
gcgaagagaa ccttaagccg ttattacctc tgcctgtcac aggcgatgtg atgctaacga 5820
acagcaccag agccaagcca actggggcgg tctgcagaga aggctgggat acccgaaata 5880
gctcgctcaa cagcttttt tcttctacgg aagcccacca gataagcgcc tttgttgggc 5940
ccgctaaccc cgggacatgc ccgggctcgg agttagtttt tgcacggccg gcagatctat 6000
ttaaatggcg cgccgacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt 6060
tgtttatttt tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa 6120
atgcttcaat aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt 6180
attccctttt ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa 6240
gtaaaagatg ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac 6300
agcggtaaga tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt 6360
aaagttctgc tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt 6420
cgccgcatac actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat 6480
cttacggatg gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac 6540
actgcggcca acttacttct gacaacgatc ggaggaccga aggagctaac cgcttttttg 6600
cacaacatgg gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc 6660
ataccaaacg acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa 6720
ctattaactg gcgaactact tactctagct tcccggcaac aattaataga ctggatggag 6780
gcggataaag ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct 6840
gataaatctg gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat 6900
ggtaagccct cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa 6960
cgaaatagac agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac 7020
caagtttact catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc 7080
taggtgaaga tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc 7140
cactgagcgt cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg 7200
cgcgtaatct gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg 7260
gatcaagagc taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca 7320
aatactgttc ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg 7380
cctacatacc tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg 7440
tgtcttaccg ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga 7500
acgggggggtt cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac 7560
ctacagcgtg agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat 7620

ccggtaagcg gcagggtcgg aacaggagag cgcacgaggg agcttccagg gggaaacgcc 7680
tggtatcttt atagtcctgt cgggtttcgc cacctctgac ttgagcgtcg atttttgtga 7740
tgctcgtcag gggggcggag cctatggaaa aacgccagca acgcggcctt tttacggttc 7800
ctggcctttt gctggccttt tgctcacatg ttctttcctg cgttatcccc tgattctgtg 7860
gataaccgta ttaccgcctt tgagtgagct gataccgctc gccgcagccg aacgaccgag 7920
cgcagcgagt cagtgagcga ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc 7980
gcgcgttggc cgattcatta atgcagctgg cacgacaggt ttcccgactg gaaagcgggc 8040
agtgagcgca acgcaattaa tgtgagttag ctcactcatt aggcacccca ggctttacac 8100
tttatgcttc cggctcgtat gttgtgtgga attgtgagcg gataacaatt tcacacagga 8160
aacagctatg accatgatta cgccaagctt tttctttcca atttttttt tttcgtcatt 8220
ataaaaatca ttacgaccga gattcccggg taataactga tataattaaa ttgaagctct 8280
aatttgtgag tttagtatac atgcatttac ttataataca gttttttagt tttgctggcc 8340
gcatcttctc aaatatgctt cccagcctgc ttttctgtaa cgttcaccct ctaccttagc 8400
atcccttccc tttgcaaata gtcctcttcc aacaataata atgtcagatc ctgtagagac 8460
cacatcatcc acggttctat actgttgacc caatgcgtct cccttgtcat ctaaacccac 8520
accgggtgtc ataatcaacc aatcgtaacc ttcatctctt ccacccatgt ctctttgagc 8580
aataaagccg ataacaaaat ctttgtcgct cttcgcaatg tcaacagtac ccttagtata 8640
ttctccagta gatagggagc ccttgcatga caattctgct aacatcaaaa ggcctctagg 8700
ttcctttgtt acttcttctg ccgcctgctt caaaccgcta acaatacctg ggcccaccac 8760
accgtgtgca ttcgtaatgt ctgcccattc tgctattctg tatacacccg cagagtactg 8820
caatttgact gtattaccaa tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta 8880
cttggcggat aatgccttta gcggcttaac tgtgccctcc atggaaaaat cagtcaagat 8940
atccacatgt gtttttagta aacaaatttt gggacctaat gcttcaacta actccagtaa 9000
ttccttggtg gtacgaacat ccaatgaagc acacaagttt gtttgctttt cgtgcatgat 9060
attaaatagc ttggcagcaa caggactagg atgagtagca gcacgttcct tatatgtagc 9120
tttcgacatg atttatcttc gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa 9180
tactgggcaa tttcatgttt cttcaacact acatatgcgt atatatacca atctaagtct 9240
gtgctccttc cttcgttctt ccttctgttc ggagattacc gaatcaaaaa aatttcaagg 9300
aaaccgaaat caaaaaaaag aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg 9360
caggtcgact ctagtatact ccgtctactg tacgatacac ttccgctcag gtccttgtcc 9420
tttaacgagg ccttaccact cttttgttac tctattgatc cagctcagca aaggcagtgt 9480
gatctaagat tctatcttcg cgatgtagta aaactagcta gaccgagaaa gagactagaa 9540
atgcaaaagg cacttctaca atggctgcca tcattattat ccgatgtgac gctgcatttt 9600
tttttttttt tttttttttt tttttttttt tttttttttt tttttttgt acaaatatca 9660
taaaaaaaga gaatcttttt aagcaaggat tttcttaact tcttcggcga cagcatcacc 9720
gacttcggtg gtactgttgg aaccacctaa atcaccagtt ctgatacctg catccaaaac 9780

cttttaact gcatcttcaa tggctttacc ttcttcaggc aagttcaatg acaatttcaa 9840
catcattgca gcagacaaga tagtggcgat agggttgacc ttattctttg gcaaatctgg 9900
agcggaacca tggcatggtt cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc 9960
caaggacgca gatggcaaca aacccaagga gcctgggata acggaggctt catcggagat 10020
gatatcacca aacatgttgc tggtgattat aataccattt aggtgggttg ggttcttaac 10080
taggatcatg gcggcagaat caatcaattg atgttgaact ttcaatgtag ggaattcgtt 10140
cttgatggtt tcctccacag tttttctcca taatcttgaa gaggccaaaa cattagcttt 10200
atccaaggac caaataggca atggtggctc atgttgtagg gccatgaaag cggccattct 10260
tgtgattctt tgcacttctg gaacggtgta ttgttcacta tcccaagcga caccatcacc 10320
atcgtcttcc tttctcttac caaagtaaat acctcccact aattctctaa caacaacgaa 10380
gtcagtacct ttagcaaatt gtggcttgat tggagataag tctaaaagag agtcggatgc 10440
aaagttacat ggtcttaagt tggcgtacaa ttgaagttct ttacggattt ttagtaaacc 10500
ttgttcaggt ctaacactac cggtacccca tttaggacca cccacagcac ctaacaaaac 10560
ggcatcagcc ttcttggagg cttccagcgc ctcatctgga agtggaacac ctgtagcatc 10620
gatagcagca ccaccaatta aatgattttc gaaatcgaac ttgacattgg aacgaacatc 10680
agaaatagct ttaagaacct taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc 10740
tggcaaaacg acgatcttct taggggcaga cattacaatg gtatatcctt gaaatatata 10800
taaaaaaaaa aaaaaaaaaa aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc 10860
tttgaggaga tacagcctaa tatccgacaa actgttttac agatttacga tcgtacttgt 10920
tacccatcat tgaattttga acatccgaac ctgggagttt tccctgaaac agatagtata 10980
tttgaacctg tataataata tatagtctag cgctttacgg aagacaatgt atgtatttcg 11040
gttcctggag aaactattgc atctattgca taggtaatct tgcacgtcgc atccccggtt 11100
cattttctgc gtttccatct tgcacttcaa tagcatatct ttgttaacga agcatctgtg 11160
cttcattttg tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg 11220
agctgcattt ttacagaaca gaaatgcaac gcgaaagcgc tattttacca acgaagaatc 11280
tgtgcttcat ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaaga 11340
atctgagctg cattttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa 11400
gaatctatac ttctttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca 11460
aagcatctta gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac 11520
tttttgcact gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt 11580
ccataaaaaa agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg 11640
catttttca agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac 11700
tttgtgaaca gaaagtgata gcgttgatga ttcttcattg gtcagaaaat tatgaacggt 11760
ttcttctatt ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt 11820
cgattcactc tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga 11880
taaacataaa aaatgtagag gtcgagttta gatgcaagtt caaggagcga aaggtggatg 11940

ggtaggttat atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg 12000

tttgtggaag cggtattcgc aatattttag tagctcgtta cagtccggtg cgtttttggt 12060

tttttgaaag tgcgtcttca gagcgctttt ggttttcaaa agcgctctga agttcctata 12120

ctttctagag aataggaact tcggaatagg aacttcaaag cgtttccgaa aacgagcgct 12180

tccgaaaatg caacgcgagc tgcgcacata cagctcactg ttcacgtcgc acctatatct 12240

gcgtgttgcc tgtatatata tatacatgag aagaacggca tagtgcgtgt ttatgcttaa 12300

atgcgtactt atatgcgtct atttatgtag gatgaaaggt agtctagtac ctcctgtgat 12360

attatcccat tccatgcggg gtatcgtatg cttccttcag cactaccctt tagctgttct 12420

atatgctgcc actcctcaat tggattagtc tcatccttca atgctatcat ttcctttgat 12480

attggatcat atgcatagta ccgagaaact agaggatc 12518

<210> 774
<211> 1029
<212> DNA
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae KARI-AB1G9 variant

<400> 774
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 775
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer pOLE1-NgoMI

<400> 775
cgtggccggc cgtgcaaaaa ctaactccga g 31

<210> 776
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer pOLE1-PmlI

<400> 776
gtactcacgt gtttgttgta atgttttagt gctg 34

<210> 777
<211> 1029
<212> DNA
<213> Artificial Sequence

<220>
<223> K9G9 ilvC gene

<400> 777
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttatacg aaggtgcgaa ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatcaa cgatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaac 1029

<210> 778
<211> 9613
<212> DNA
<213> Artificial Sequence

<220>

<223> pLH701 plasmid containing K9G9 variant

<400> 778

ggccgcacct ggtaaaacct ctagtggagt agtagatgta atcaatgaag cggaagccaa 60

aagaccagag tagaggccta tagaagaaac tgcgatacct tttgtgatgg ctaaacaaac 120

agacatcttt ttatatgttt ttacttctgt atatcgtgaa gtagtaagtg ataagcgaat 180

ttggctaaga acgttgtaag tgaacaaggg acctcttttg cctttcaaaa aaggattaaa 240

tggagttaat cattgagatt tagttttcgt tagattctgt atccctaaat aactccctta 300

cccgacggga aggcacaaaa gacttgaata atagcaaacg gccagtagcc aagaccaaat 360

aatactagag ttaactgatg gtcttaaaca ggcattacgt ggtgaactcc aagaccaata 420

tacaaaatat cgataagtta ttcttgccca ccaatttaag gagcctacat caggacagta 480

gtaccattcc tcagagaaga ggtatacata acaagaaaat cgcgtgaaca ccttatataa 540

cttagcccgt tattgagcta aaaaaccttg caaaatttcc tatgaataag aatacttcag 600

acgtgataaa aatttacttt ctaactcttc tcacgctgcc cctatctgtt cttccgctct 660

accgtgagaa ataaagcatc gagtacggca gttcgctgtc actgaactaa aacaataagg 720

ctagttcgaa tgatgaactt gcttgctgtc aaacttctga gttgccgctg atgtgacact 780

gtgacaataa attcaaaccg gttatagcgg tctcctccgg taccggttct gccacctcca 840

atagagctca gtaggagtca gaacctctgc ggtggctgtc agtgactcat ccgcgtttcg 900

taagttgtgc gcgtgcacat ttcgcccgtt cccgctcatc ttgcagcagg cggaaatttt 960

catcacgctg taggacgcaa aaaaaaaata attaatcgta caagaatctt ggaaaaaaaa 1020

ttgaaaaatt ttgtataaaa gggatgacct aacttgactc aatggctttt acacccagta 1080

ttttcccttt ccttgtttgt tacaattata gaagcaagac aaaaacatat agacaaccta 1140

ttcctaggag ttatattttt ttaccctacc agcaatataa gtaaaaaact gtttaaacag 1200

tatggaagaa tgtaagatgg ctaagattta ctaccaagaa gactgtaact tgtccttgtt 1260

ggatggtaag actatcgccg ttatcggtta cggttctcaa ggtcacgctc atgccctgaa 1320

tgctaaggaa tccggttgta acgttatcat tggtttatac gaaggtgcga aggagtggaa 1380

aagagctgaa gaacaaggtt tcgaagtcta caccgctgct gaagctgcta agaaggctga 1440

catcattatg atcttgatca acgatgaaaa gcaggctacc atgtacaaaa acgacatcga 1500

accaaacttg gaagccggta acatgttgat gttcgctcac ggtttcaaca tccatttcgg 1560

ttgtattgtt ccaccaaagg acgttgatgt cactatgatc gctccaaagg gtccaggtca 1620

caccgttaga tccgaatacg aagaaggtaa aggtgtccca tgcttggttg ctgtcgaaca 1680

agacgctact ggcaaggctt tggatatggc tttggcctac gctttagcca tcggtggtgc 1740

tagagccggt gtcttggaaa ctaccttcag aaccgaaact gaaaccgact tgttcggtga 1800

acaagctgtt ttatgtggtg gtgtctgcgc tttgatgcag gccggttttg aaaccttggt 1860

tgaagccggt tacgacccaa gaaacgctta cttcgaatgt atccacgaaa tgaagttgat 1920

cgttgacttg atctaccaat ctggtttctc cggtatgcgt tactctatct ccaacactgc 1980

tgaatacggt gactacatta ccggtccaaa gatcattact gaagatacca agaaggctat 2040

gaagaagatt ttgtctgaca ttcaagatgg tacctttgcc aaggacttct tggttgacat 2100

gtctgatgct ggttcccagg tccacttcaa ggctatgaga aagttggcct ccgaacaccc 2160
agctgaagtt gtcggtgaag aaattagatc cttgtactcc tggtccgacg aagacaagtt 2220
gattaacaac tgaggccctg caggccagag gaaaataata tcaagtgctg gaaacttttt 2280
ctcttggaat ttttgcaaca tcaagtcata gtcaattgaa ttgacccaat ttcacattta 2340
agatttttt tttttcatcc gacatacatc tgtacactag gaagccctgt ttttctgaag 2400
cagcttcaaa tatatatatt ttttacatat ttattatgat tcaatgaaca atctaattaa 2460
atcgaaaaca agaaccgaaa cgcgaataaa taatttattt agatggtgac aagtgtataa 2520
gtcctcatcg ggacagctac gatttctctt tcggttttgg ctgagctact ggttgctgtg 2580
acgcagcggc attagcgcgg cgttatgagc taccctcgtg gcctgaaaga tggcgggaat 2640
aaagcggaac taaaaattac tgactgagcc atattgaggt caatttgtca actcgtcaag 2700
tcacgtttgg tggacggccc ctttccaacg aatcgtatat actaacatgc gcgcgcttcc 2760
tatatacaca tatacatata tatatatata tatatgtgtg cgtgtatgtg tacacctgta 2820
tttaatttcc ttactcgcgg gtttttcttt tttctcaatt cttggcttcc tctttctcga 2880
gcggaccgga tcctccgcgg tgccggcaga tctatttaaa tggcgcgccg acgtcaggtg 2940
gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa 3000
atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 3060
agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttttgcg gcattttgcc 3120
ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg 3180
gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc 3240
gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat 3300
tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg 3360
acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag 3420
aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa 3480
cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc 3540
gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca 3600
cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa ctacttactc 3660
tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc 3720
tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg 3780
ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta 3840
tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag 3900
gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga 3960
ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc 4020
tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 4080
agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 4140
aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttc 4200
cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgttcttcta gtgtagccgt 4260

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 4320
tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 4380
gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 4440
gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 4500
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 4560
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 4620
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggg cggagcctat 4680
ggaaaaacgc cagcaacgcg gccttttac ggttcctggc cttttgctgg ccttttgctc 4740
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt 4800
gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag 4860
cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca 4920
gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga 4980
gttagctcac tcattaggca ccccaggctt tacactttat gcttccggct cgtatgttgt 5040
gtggaattgt gagcggataa caatttcaca caggaaacag ctatgaccat gattacgcca 5100
agctttttct ttccaatttt ttttttttcg tcattataaa aatcattacg accgagattc 5160
ccgggtaata actgatataa ttaaattgaa gctctaattt gtgagtttag tatacatgca 5220
tttacttata atacagtttt ttagttttgc tggccgcatc ttctcaaata tgcttcccag 5280
cctgcttttc tgtaacgttc accctctacc ttagcatccc ttccctttgc aaatagtcct 5340
cttccaacaa taataatgtc agatcctgta gagaccacat catccacggt tctatactgt 5400
tgacccaatg cgtctccctt gtcatctaaa cccacaccgg gtgtcataat caaccaatcg 5460
taaccttcat ctcttccacc catgtctctt tgagcaataa agccgataac aaaatctttg 5520
tcgctcttcg caatgtcaac agtaccctta gtatattctc cagtagatag ggagcccttg 5580
catgacaatt ctgctaacat caaaaggcct ctaggttcct ttgttacttc ttctgccgcc 5640
tgcttcaaac cgctaacaat acctgggccc accacaccgt gtgcattcgt aatgtctgcc 5700
cattctgcta ttctgtatac acccgcagag tactgcaatt tgactgtatt accaatgtca 5760
gcaaattttc tgtcttcgaa gagtaaaaaa ttgtacttgg cggataatgc ctttagcggc 5820
ttaactgtgc cctccatgga aaaatcagtc aagatatcca catgtgtttt tagtaaacaa 5880
attttgggac ctaatgcttc aactaactcc agtaattcct tggtggtacg aacatccaat 5940
gaagcacaca agtttgtttg cttttcgtgc atgatattaa atagcttggc agcaacagga 6000
ctaggatgag tagcagcacg ttccttatat gtagctttcg acatgattta tcttcgtttc 6060
ctgcaggttt ttgttctgtg cagttgggtt aagaatactg ggcaatttca tgtttcttca 6120
acactacata tgcgtatata taccaatcta agtctgtgct ccttccttcg ttcttccttc 6180
tgttcggaga ttaccgaatc aaaaaaattt caaggaaacc gaaatcaaaa aaaagaataa 6240
aaaaaaaatg atgaattgaa aagcttgcat gcctgcaggt cgactctagt atactccgtc 6300
tactgtacga tacacttccg ctcaggtcct tgtcctttaa cgaggcctta ccactctttt 6360
gttactctat tgatccagct cagcaaaggc agtgtgatct aagattctat cttcgcgatg 6420

tagtaaaact agctagaccg agaaagagac tagaaatgca aaaggcactt ctacaatggc 6480
tgccatcatt attatccgat gtgacgctgc attttttttt ttttttttt ttttttttt 6540
ttttttttt ttttttttt tttgtacaaa tatcataaaa aaagagaatc ttttaagca 6600
aggattttct taacttcttc ggcgacagca tcaccgactt cggtggtact gttggaacca 6660
cctaaatcac cagttctgat acctgcatcc aaaacctttt taactgcatc ttcaatggct 6720
ttaccttctt caggcaagtt caatgacaat ttcaacatca ttgcagcaga caagatagtg 6780
gcgatagggt tgaccttatt ctttggcaaa tctggagcgg aaccatggca tggttcgtac 6840
aaaccaaatg cggtgttctt gtctggcaaa gaggccaagg acgcagatgg caacaaaccc 6900
aaggagcctg ggataacgga ggcttcatcg gagatgatat caccaaacat gttgctggtg 6960
attataatac catttaggtg ggttgggttc ttaactagga tcatggcggc agaatcaatc 7020
aattgatgtt gaactttcaa tgtagggaat tcgttcttga tggtttcctc cacagttttt 7080
ctccataatc ttgaagaggc caaaacatta gctttatcca aggaccaaat aggcaatggt 7140
ggctcatgtt gtagggccat gaaagcggcc attcttgtga ttctttgcac ttctggaacg 7200
gtgtattgtt cactatccca agcgacacca tcaccatcgt cttcctttct cttaccaaag 7260
taaatacctc ccactaattc tctaacaaca acgaagtcag tacctttagc aaattgtggc 7320
ttgattggag ataagtctaa aagagagtcg gatgcaaagt tacatggtct taagttggcg 7380
tacaattgaa gttctttacg gatttttagt aaaccttgtt caggtctaac actaccggta 7440
ccccatttag gaccacccac agcacctaac aaaacggcat cagccttctt ggaggcttcc 7500
agcgcctcat ctggaagtgg aacacctgta gcatcgatag cagcaccacc aattaaatga 7560
ttttcgaaat cgaacttgac attggaacga acatcagaaa tagctttaag aaccttaatg 7620
gcttcggctg tgatttcttg accaacgtgg tcacctggca aaacgacgat cttcttaggg 7680
gcagacatta caatggtata tccttgaaat atatataaaa aaaaaaaaaa aaaaaaaaaa 7740
aaaaaatgca gcttctcaat gatattcgaa tacgctttga ggagatacag cctaatatcc 7800
gacaaactgt tttacagatt tacgatcgta cttgttaccc atcattgaat tttgaacatc 7860
cgaacctggg agttttccct gaaacagata gtatatttga acctgtataa taatatatag 7920
tctagcgctt tacggaagac aatgtatgta tttcggttcc tggagaaact attgcatcta 7980
ttgcataggt aatcttgcac gtcgcatccc cggttcattt tctgcgtttc catcttgcac 8040
ttcaatagca tatctttgtt aacgaagcat ctgtgcttca ttttgtagaa caaaaatgca 8100
acgcgagagc gctaattttt caaacaaaga atctgagctg cattttaca gaacagaaat 8160
gcaacgcgaa agcgctattt taccaacgaa gaatctgtgc ttcatttttg taaaacaaaa 8220
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 8280
gaaatgcaac gcgagagcgc tattttacca acaaagaatc tatacttctt ttttgttcta 8340
caaaaatgca tcccgagagc gctatttttc taacaaagca tcttagatta ctttttttct 8400
cctttgtgcg ctctataatg cagtctcttg ataacttttt gcactgtagg tccgttaagg 8460
ttagaagaag gctactttgg tgtctatttt ctcttccata aaaaaagcct gactccactt 8520
cccgcgttta ctgattacta gcgaagctgc gggtgcattt tttcaagata aaggcatccc 8580

cgattatatt ctataccgat gtggattgcg catactttgt gaacagaaag tgatagcgtt 8640

gatgattctt cattggtcag aaaattatga acggtttctt ctattttgtc tctatatact 8700

acgtatagga aatgtttaca ttttcgtatt gttttcgatt cactctatga atagttctta 8760

ctacaatttt tttgtctaaa gagtaatact agagataaac ataaaaaatg tagaggtcga 8820

gtttagatgc aagttcaagg agcgaaaggt ggatgggtag gttatatagg gatatagcac 8880

agagatatat agcaaagaga tacttttgag caatgtttgt ggaagcggta ttcgcaatat 8940

tttagtagct cgttacagtc cggtgcgttt ttggtttttt gaaagtgcgt cttcagagcg 9000

cttttggttt tcaaaagcgc tctgaagttc ctatactttc tagagaatag gaacttcgga 9060

ataggaactt caaagcgttt ccgaaaacga gcgcttccga aaatgcaacg cgagctgcgc 9120

acatacagct cactgttcac gtcgcaccta tatctgcgtg ttgcctgtat atatatatac 9180

atgagaagaa cggcatagtg cgtgtttatg cttaaatgcg tacttatatg cgtctattta 9240

tgtaggatga aaggtagtct agtacctcct gtgatattat cccattccat gcggggtatc 9300

gtatgcttcc ttcagcacta ccctttagct gttctatatg ctgccactcc tcaattggat 9360

tagtctcatc cttcaatgct atcatttcct ttgatattgg atcatatgca tagtaccgag 9420

aaactagagg atctcccatt accgacattt gggcgctata cgtgcatatg ttcatgtatg 9480

tatctgtatt taaaacactt ttgtattatt tttcctcata tatgtgtata ggtttatacg 9540

gatgatttaa ttattacttc accacccttt atttcaggct gatatcttag ccttgttact 9600

agtcaccggt ggc 9613

<210> 779
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> N1090 primer

<400> 779
aatcccgatg tatgggtttg gttgccag 28

<210> 780
<211> 7589
<212> DNA
<213> Artificial Sequence

<220>
<223> Als integration plasmid

<400> 780
tatttgtatc gaggtgtcta gtcttctatt acactaatgc agtttcaggg ttttggaaac 60

cacactgttt aaacagtgtt ccttaatcaa ggatacctct ttttttttcc ttggttccac 120

taattcatcg gttttttttt tggaagacat cttttccaac gaaaagaata tacatatcgt 180

ttaagagaaa ttctccaaat ttgtaaagaa gcggacccag acttaaggcc gcccgcaaat 240

taaagccttc gagcgtccca aaaccttctc aagcaaggtt ttcagtataa tgttacatgc 300

gtacacgcgt ctgtacagaa aaaaaagaaa aatttgaaat ataaataacg ttcttaatac 360

taacataact ataaaaaaat aaatagggac ctagacttca ggttgtctaa ctccttcctt 420

ttcggttaga gcggatgtgg ggggagggcg tgaatgtaag cgtgacataa ctaattacat 480
gattaattaa ctagagagct ttcgttttca tgagttcccc gaattctttc ggaagcttgt 540
cacttgctaa attaatgtta tcactgtagt caaccgggac atcgatgatg acaggacctt 600
cagcgttcat gccttgacgc agaacatctg ccagctggtc tggtgattct acgcgcaagc 660
cagttgctcc gaagctttcc gcatatttca cgatatcgat atttccgaaa tcgaccgcag 720
atgtacggtt atatttttc aattgctgga atgcaaccat gtcatatgtg ctgtcgttcc 780
atacaatgtg tacaattggt gcttttagtc gaactgctgt ctctaattcc attgctgaga 840
ataagaaacc gccgtcacca gagacagaaa ccacttttc tcccggtttc accaatgaag 900
cgccgattgc ccaaggaagc gcaacgccga gtgtttgcat accgttactg atcattaatg 960
ttaacggctc gtagctgcgg aaataacgtg acatccaaat ggcgtgcgaa ccgatatcgc 1020
aagttactgt aacatgatca tcgactgcat tacgcaactc tttaacgatt tcaagagggt 1080
gcgctctgtc tgatttccaa tctgcaggca cctgctcacc ttcatgcata tattgtttta 1140
aatcagaaag gattttctgc tcacgctctg caaattccac tttcacagca tcgtgttcga 1200
tatgattgat cgtggacgga atgtcaccga tcaattcaag atcaggctgg taagcatgat 1260
caatgtcagc gataatctcg tctaaatgga taattgtccg gtctccattg atattccaga 1320
atttcggatc atattcaatc gggtcatagc cgatcgtcag aacaacatct gcctgctcta 1380
gcagtaaatc gccaggctgg ttgcggaaca aaccgatacg gccaaaatat tgatcctcta 1440
aatctctaga aagggtaccg gcagcttgat atgtttcaac aaatggaagc tgaacctttt 1500
tcaaaagctt gcgaaccgct ttaattgctt ccggtcttcc gcctttcatg ccgaccaaaa 1560
cgacaggaag ttttgctgtt tggatttttg ctatggccgc actgattgca tcatctgctg 1620
caggaccgag ttttggcgct gcaacagcac gcacgttttt cgtatttgtg acttcattca 1680
caacatcttg cggaaagctc acaaaagcgg ccccagcctg ccctgctgac gctatcctaa 1740
atgcatttgt aacagcttcc ggtatatttt ttacatcttg aacttctaca ctgtattttg 1800
taatcggctg gaatagcgcc gcattatcca aagattgatg tgtccgtttt aaacgatctg 1860
cacggatcac gtttccagca agcgcaacga cagggtctcc ttcagtgttc gctgtcagca 1920
ggcctgttgc caagttagag gcacccggtc ctgatgtgac taacacgact cccggttttc 1980
cagttaaacg gccgactgct tgggccatga atgctgcgtt ttgttcgtgc cgggcaacga 2040
taatttcagg tcctttatct tgtaaagcgt caaataccgc atcaattttt gcacctggaa 2100
tgccaaatac atgtgtgaca ccttgctcca ctaagcaatc aacaacaagc tccgcccctc 2160
tgtttttcac aagggatttt tgttcttttg ttgcttttgt caacatcctc agctctagat 2220
ttgaatatgt attacttggt tatggttata tatgacaaaa gaaaaagaag aacagaagaa 2280
taacgcaagg aagaacaata actgaaattg atagagaagt attatgtctt tgtcttttta 2340
taataaatca agtgcagaaa tccgttagac aacatgaggg ataaaattta acgtgggcga 2400
agaagaagga aaaaagtttt tgtgagggcg taattgaagc gatctgttga ttgtagattt 2460
ttttttttg aggagtcaaa gtcagaagag aacagacaaa tggtattaac catccaatac 2520
ttttttggag caacgctaag ctcatgcttt tccattggtt acgtgctcag ttgttagata 2580

tggaaagaga ggatgctcac ggcagcgtga ctccaattga gcccgaaaga gaggatgcca 2640
cgttttcccg acggctgcta gaatggaaaa aggaaaaata gaagaatccc attcctatca 2700
ttatttacgt aatgacccac acatttttga gattttcaac tattacgtat tacgataatc 2760
ctgctgtcat tatcattatt atctatatcg acgtatgcaa cgtatgtgaa gccaagtagg 2820
caattattta gtactgtcag tattgttatt catttcagat cttaagccta accaggccaa 2880
ttcaacagac tgtcggcaac ttcttgtctg gtctttccat ggtaagtgac agtgcagtaa 2940
taatatgaac caatttattt ttcgttacat aaaaatgctt ataaaacttt aactaataat 3000
tagagattaa atcgcaaacg gccgactcta gaggatcccc caccttggct aactcgttgt 3060
atcatcactg gataacttcg tatagcatac attatacgaa gttatctagg gattcataac 3120
cattttctca atcgaattac acagaacaca ccgtacaaac ctctctatca taactactta 3180
atagtcacac acgtactcgt ctaaatacac atcatcgtcc tacaagttca tcaaagtgtt 3240
ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc 3300
aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac 3360
agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa 3420
tttttctttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc 3480
tgagagttcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt 3540
ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt 3600
agtactggtt tatatgggtt tttctgtata acttcttttt attttagttt gtttaatctt 3660
attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa ctcgagatgg 3720
gtaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat gctgatttat 3780
atgggtataa atgggctcgc gataatgtcg ggcaatcagg tgcgacaatc tatcgattgt 3840
atgggaagcc cgatgcgcca gagttgtttc tgaaacatgg caaaggtagc gttgccaatg 3900
atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct cttccgacca 3960
tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg atccccggca 4020
aaacagcatt ccaggtatta gaagaatatc ctgattcagg tgaaaatatt gttgatgcgc 4080
tggcagtgtt cctgcgccgg ttgcattcga ttcctgtttg taattgtcct tttaacagcg 4140
atcgcgtatt tcgtctcgct caggcgcaat cacgaatgaa taacggtttg gttgatgcga 4200
gtgattttga tgacgagcgt aatggctggc ctgttgaaca agtctggaaa gaaatgcata 4260
agcttttgcc attctcaccg gattcagtcg tcactcatgg tgatttctca cttgataacc 4320
ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc ggaatcgcag 4380
accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct ccttcattac 4440
agaaacggct ttttcaaaaa tatggtattg ataatcctga tatgaataaa ttgcagtttc 4500
atttgatgct cgatgagttt ttctaagttt aacttgatac tactagattt tttctcttca 4560
tttataaaat ttttggttat aattgaagct ttagaagtat gaaaaaatcc ttttttttca 4620
ttctttgcaa ccaaaataag aagcttcttt tattcattga aatgatgaat ataaacctaa 4680
caaaagaaaa agactcgaat atcaaacatt aaaaaaaaat aaaagaggtt atctgttttc 4740

ccatttagtt ggagtttgca ttttctaata gatagaactc tcaattaatg tggatttagt 4800
ttctctgttc gataacttcg tatagcatac attatacgaa gttatctgaa cattagaata 4860
cgtaatccgc aatgcggggc cgcttaatta atctagagtc gacctgcagg catgcaagct 4920
tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg ttatccgctc acaattccac 4980
acaacatacg agccggaagc ataaagtgta aagcctgggg tgcctaatga gtgagctaac 5040
tcacattaat tgcgttgcgc tcactgcccg ctttccagtc gggaaacctg tcgtgccagc 5100
tgcattaatg aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg 5160
cttcctcgct cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc 5220
actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 5280
gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 5340
ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 5400
acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 5460
ctgttccgac cctgccgctt accggatacc tgtccgcctt tctcccttcg ggaagcgtgg 5520
cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 5580
tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 5640
gtcttgagtc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 5700
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 5760
acggctacac tagaaggaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 5820
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 5880
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 5940
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 6000
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 6060
tctaaagtat atatgagtaa acttggtctg acagttacca atgcttaatc agtgaggcac 6120
ctatctcagc gatctgtcta tttcgttcat ccatagttgc ctgactcccc gtcgtgtaga 6180
taactacgat acgggagggc ttaccatctg gccccagtgc tgcaatgata ccgcgagacc 6240
cacgctcacc ggctccagat ttatcagcaa taaaccagcc agccggaagg gccgagcgca 6300
gaagtggtcc tgcaacttta tccgcctcca tccagtctat taattgttgc cgggaagcta 6360
gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt tgccattgct acaggcatcg 6420
tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc cggttcccaa cgatcaaggc 6480
gagttacatg atcccccatg ttgtgcaaaa aagcggttag ctccttcggt cctccgatcg 6540
ttgtcagaag taagttggcc gcagtgttat cactcatggt tatggcagca ctgcataatt 6600
ctcttactgt catgccatcc gtaagatgct tttctgtgac tggtgagtac tcaaccaagt 6660
cattctgaga atagtgtatg cggcgaccga gttgctcttg cccggcgtca atacgggata 6720
ataccgcgcc acatagcaga actttaaaag tgctcatcat tggaaaacgt tcttcggggc 6780
gaaaactctc aaggatctta ccgctgttga gatccagttc gatgtaaccc actcgtgcac 6840
ccaactgatc ttcagcatct tttactttca ccagcgtttc tgggtgagca aaaacaggaa 6900

ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa atgttgaata ctcatactct 6960

tccttttca atattattga agcatttatc agggttattg tctcatgagc ggatacatat 7020

ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg cacatttccc cgaaaagtgc 7080

cacctgacgt ctaagaaacc attattatca tgacattaac ctataaaaat aggcgtatca 7140

cgaggccctt tcgtctcgcg cgtttcggtg atgacggtga aaacctctga cacatgcagc 7200

tcccggagac ggtcacagct tgtctgtaag cggatgccgg gagcagacaa gcccgtcagg 7260

gcgcgtcagc gggtgttggc gggtgtcggg gctggcttaa ctatgcggca tcagagcaga 7320

ttgtactgag agtgcaccat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 7380

accgcatcag gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc 7440

gggcctcttc gctattacgc cagctggcga aaggggatg tgctgcaagg cgattaagtt 7500

gggtaacgcc agggttttcc cagtcacgac gttgtaaaac gacggccagt gaattcgagc 7560

tcggtacccg gggatccggc gcgccgttt 7589

<210> 781
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> k9_Jarvis_SH

<400> 781

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Ile Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
130 135 140

2017200439 20 Jan 2017

```
Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 782
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Kelly

<400> 782

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                   10                  15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45
```

```
Glu Gly Gly Glu Glu Trp Asn Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335
```

Asn Asn

<210> 783
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Norman_SH

<400> 783

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                  10                  15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Cys Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220
```

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 784
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> N1237 primer

<400> 784
gatgatgcta tttggtgcag aggg 24

<210> 785
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> N1238 primer

<400> 785
ccttactgaa gtatttagtt atcatgg 27

<210> 786
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer N1240

<400> 786
ggatgtatgg gctaaatgta cggg 24

<210> 787
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> N1241 Primer

<400> 787
catccagtgt cgaaaacgag ctcg 24

<210> 788
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> Anaerostipes caccae KARI- K9D3 variant

<400> 788

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Asp Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 789
<211> 448
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae CUP1 promoter

<400> 789
cccattaccg acatttgggc gctatacgtg catatgttca tgtatgtatc tgtatttaaa 60

acacttttgt attatttttc ctcatatatg tgtataggtt tatacggatg atttaattat 120

tacttcacca ccctttattt caggctgata tcttagcctt gttactagtt agaaaaagac 180

atttttgctg tcagtcactg tcaagagatt ctttgctggg catttcttct agaagcaaaa 240

agagcgatgc gtcttttccg ctgaaccgtt ccagcaaaaa agactaccaa cgcaatatgg 300

attgtcagaa tcatataaaa gagaagcaaa taactccttg tcttgtatca attgcattat 360

aatatcttct tgttagtgca atatcatata gaagtcatcg aaatagatat taagaaaaac 420

aaactgtaca atcaatcaat caatcatc 448

<210> 790
<211> 1713
<212> DNA
<213> Artificial Sequence

<220>
<223> Bacillus subtilis AlsS

<400> 790
ttgacaaaag caacaaaaga acaaaaatcc cttgtgaaaa acagaggggc ggagcttgtt 60

gttgattgct tagtggagca aggtgtcaca catgtatttg gcattccagg tgcaaaaatt 120

gatgcggtat ttgacgcttt acaagataaa ggacctgaaa ttatcgttgc ccggcacgaa 180
caaaacgcag cattcatggc ccaagcagtc ggccgtttaa ctggaaaacc gggagtcgtg 240
ttagtcacat caggaccggg tgcctctaac ttggcaacag gcctgctgac agcgaacact 300
gaaggagacc ctgtcgttgc gcttgctgga aacgtgatcc gtgcagatcg tttaaaacgg 360
acacatcaat ctttggataa tgcggcgcta ttccagccga ttacaaaata cagtgtagaa 420
gttcaagatg taaaaaatat accggaagct gttacaaatg catttaggat agcgtcagca 480
gggcaggctg gggccgcttt tgtgagcttt ccgcaagatg ttgtgaatga agtcacaaat 540
acgaaaaacg tgcgtgctgt tgcagcgcca aaactcggtc ctgcagcaga tgatgcaatc 600
agtgcggcca tagcaaaaat ccaaacagca aaacttcctg tcgttttggt cggcatgaaa 660
ggcggaagac cggaagcaat taaagcggtt cgcaagcttt tgaaaaaggt tcagcttcca 720
tttgttgaaa catatcaagc tgccggtacc ctttctagag atttagagga tcaatatttt 780
ggccgtatcg gtttgttccg caaccagcct ggcgatttac tgctagagca ggcagatgtt 840
gttctgacga tcggctatga cccgattgaa tatgatccga aattctggaa tatcaatgga 900
gaccggacaa ttatccattt agacgagatt atcgctgaca ttgatcatgc ttaccagcct 960
gatcttgaat tgatcggtga cattccgtcc acgatcaatc atatcgaaca cgatgctgtg 1020
aaagtggaat ttgcagagcg tgagcagaaa atcctttctg atttaaaaca atatatgcat 1080
gaaggtgagc aggtgcctgc agattggaaa tcagacagag cgcaccctct tgaaatcgtt 1140
aaagagttgc gtaatgcagt cgatgatcat gttacagtaa cttgcgatat cggttcgcac 1200
gccatttgga tgtcacgtta tttccgcagc tacgagccgt taacattaat gatcagtaac 1260
ggtatgcaaa cactcggcgt tgcgcttcct tgggcaatcg gcgcttcatt ggtgaaaccg 1320
ggagaaaaag tggtttctgt ctctggtgac ggcggtttct tattctcagc aatggaatta 1380
gagacagcag ttcgactaaa agcaccaatt gtacacattg tatggaacga cagcacatat 1440
gacatggttg cattccagca attgaaaaaa tataaccgta catctgcggt cgatttcgga 1500
aatatcgata tcgtgaaata tgcggaaagc ttcggagcaa ctggcttgcg cgtagaatca 1560
ccagaccagc tggcagatgt tctgcgtcaa ggcatgaacg ctgaaggtcc tgtcatcatc 1620
gatgtcccgg ttgactacag tgataacatt aatttagcaa gtgacaagct tccgaaagaa 1680
ttcggggaac tcatgaaaac gaaagctctc tag 1713

<210> 791
<211> 571
<212> PRT
<213> Artificial Sequence

<220>
<223> Bacillus subtilis

<400> 791

Met Leu Thr Lys Ala Thr Lys Glu Gln Lys Ser Leu Val Lys Asn Arg
1 5 10 15

Gly Ala Glu Leu Val Val Asp Cys Leu Val Glu Gln Gly Val Thr His
20 25 30

```
Val Phe Gly Ile Pro Gly Ala Lys Ile Asp Ala Val Phe Asp Ala Leu
        35                  40                  45

Gln Asp Lys Gly Pro Glu Ile Ile Val Ala Arg His Glu Gln Asn Ala
    50                  55                  60

Ala Phe Met Ala Gln Ala Val Gly Arg Leu Thr Gly Lys Pro Gly Val
65                  70                  75                  80

Val Leu Val Thr Ser Gly Pro Gly Ala Ser Asn Leu Ala Thr Gly Leu
                85                  90                  95

Leu Thr Ala Asn Thr Glu Gly Asp Pro Val Val Ala Leu Ala Gly Asn
            100                 105                 110

Val Ile Arg Ala Asp Arg Leu Lys Arg Thr His Gln Ser Leu Asp Asn
        115                 120                 125

Ala Ala Leu Phe Gln Pro Ile Thr Lys Tyr Ser Val Glu Val Gln Asp
    130                 135                 140

Val Lys Asn Ile Pro Glu Ala Val Thr Asn Ala Phe Arg Ile Ala Ser
145                 150                 155                 160

Ala Gly Gln Ala Gly Ala Ala Phe Val Ser Phe Pro Gln Asp Val Val
                165                 170                 175

Asn Glu Val Thr Asn Thr Lys Asn Val Arg Ala Val Ala Ala Pro Lys
            180                 185                 190

Leu Gly Pro Ala Ala Asp Asp Ala Ile Ser Ala Ala Ile Ala Lys Ile
        195                 200                 205

Gln Thr Ala Lys Leu Pro Val Val Leu Val Gly Met Lys Gly Gly Arg
    210                 215                 220

Pro Glu Ala Ile Lys Ala Val Arg Lys Leu Leu Lys Lys Val Gln Leu
225                 230                 235                 240

Pro Phe Val Glu Thr Tyr Gln Ala Ala Gly Thr Leu Ser Arg Asp Leu
                245                 250                 255

Glu Asp Gln Tyr Phe Gly Arg Ile Gly Leu Phe Arg Asn Gln Pro Gly
            260                 265                 270

Asp Leu Leu Leu Glu Gln Ala Asp Val Val Leu Thr Ile Gly Tyr Asp
        275                 280                 285

Pro Ile Glu Tyr Asp Pro Lys Phe Trp Asn Ile Asn Gly Asp Arg Thr
    290                 295                 300

Ile Ile His Leu Asp Glu Ile Ile Ala Asp Ile Asp His Ala Tyr Gln
305                 310                 315                 320
```

```
Pro Asp Leu Glu Leu Ile Gly Asp Ile Pro Ser Thr Ile Asn His Ile
                325                 330                 335

Glu His Asp Ala Val Lys Val Glu Phe Ala Glu Arg Glu Gln Lys Ile
            340                 345                 350

Leu Ser Asp Leu Lys Gln Tyr Met His Glu Gly Glu Gln Val Pro Ala
        355                 360                 365

Asp Trp Lys Ser Asp Arg Ala His Pro Leu Glu Ile Val Lys Glu Leu
    370                 375                 380

Arg Asn Ala Val Asp Asp His Val Thr Val Thr Cys Asp Ile Gly Ser
385                 390                 395                 400

His Ala Ile Trp Met Ser Arg Tyr Phe Arg Ser Tyr Glu Pro Leu Thr
                405                 410                 415

Leu Met Ile Ser Asn Gly Met Gln Thr Leu Gly Val Ala Leu Pro Trp
            420                 425                 430

Ala Ile Gly Ala Ser Leu Val Lys Pro Gly Glu Lys Val Val Ser Val
        435                 440                 445

Ser Gly Asp Gly Gly Phe Leu Phe Ser Ala Met Glu Leu Glu Thr Ala
    450                 455                 460

Val Arg Leu Lys Ala Pro Ile Val His Ile Val Trp Asn Asp Ser Thr
465                 470                 475                 480

Tyr Asp Met Val Ala Phe Gln Gln Leu Lys Lys Tyr Asn Arg Thr Ser
                485                 490                 495

Ala Val Asp Phe Gly Asn Ile Asp Ile Val Lys Tyr Ala Glu Ser Phe
            500                 505                 510

Gly Ala Thr Gly Leu Arg Val Glu Ser Pro Asp Gln Leu Ala Asp Val
        515                 520                 525

Leu Arg Gln Gly Met Asn Ala Glu Gly Pro Val Ile Ile Asp Val Pro
    530                 535                 540

Val Asp Tyr Ser Asp Asn Ile Asn Leu Ala Ser Asp Lys Leu Pro Lys
545                 550                 555                 560

Glu Phe Gly Glu Leu Met Lys Thr Lys Ala Leu
                565                 570
```

<210> 792
<211> 250
<212> DNA
<213> Artificial Sequence

<220>

20 Jan 2017
2017200439

<223> Saccharomyces cerevisiae CYC1 terminator 2

<400> 792
ccgcaaatta aagccttcga gcgtcccaaa accttctcaa gcaaggtttt cagtataatg 60

ttacatgcgt acacgcgtct gtacagaaaa aaaagaaaaa tttgaaatat aaataacgtt 120

cttaatacta acataactat aaaaaaataa atagggacct agacttcagg ttgtctaact 180

ccttcctttt cggttagagc ggatgtgggg ggagggcgtg aatgtaagcg tgacataact 240

aattacatga 250

<210> 793
<211> 1181
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae ILV5 promoter

<400> 793
taaaacctct agtggagtag tagatgtaat caatgaagcg gaagccaaaa gaccagagta 60

gaggcctata gaagaaactg cgataccttt tgtgatggct aaacaaacag acatcttttt 120

atatgttttt acttctgtat atcgtgaagt agtaagtgat aagcgaattt ggctaagaac 180

gttgtaagtg aacaagggac ctcttttgcc tttcaaaaaa ggattaaatg gagttaatca 240

ttgagattta gttttcgtta gattctgtat ccctaaataa ctcccttacc cgacgggaag 300

gcacaaaaga cttgaataat agcaaacggc cagtagccaa gaccaaataa tactagagtt 360

aactgatggt cttaaacagg cattacgtgg tgaactccaa gaccaatata caaaatatcg 420

ataagttatt cttgcccacc aatttaagga gcctacatca ggacagtagt accattcctc 480

agagaagagg tatacataac aagaaaatcg cgtgaacacc ttatataact tagcccgtta 540

ttgagctaaa aaaccttgca aaatttccta tgaataagaa tacttcagac gtgataaaaa 600

tttactttct aactcttctc acgctgcccc tatctgttct tccgctctac cgtgagaaat 660

aaagcatcga gtacggcagt tcgctgtcac tgaactaaaa caataaggct agttcgaatg 720

atgaacttgc ttgctgtcaa acttctgagt tgccgctgat gtgacactgt gacaataaat 780

tcaaaccggt tatagcggtc tcctccggta ccggttctgc cacctccaat agagctcagt 840

aggagtcaga acctctgcgg tggctgtcag tgactcatcc gcgtttcgta agttgtgcgc 900

gtgcacattt cgcccgttcc cgctcatctt gcagcaggcg gaaattttca tcacgctgta 960

ggacgcaaaa aaaaaataat taatcgtaca agaatcttgg aaaaaaaatt gaaaaatttt 1020

gtataaaagg gatgacctaa cttgactcaa tggcttttac acccagtatt ttccctttcc 1080

ttgtttgtta caattataga agcaagacaa aaacatatag acaacctatt cctaggagtt 1140

atatttttt accctaccag caatataagt aaaaaactag t 1181

<210> 794
<211> 759
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae ILV5 terminator

<400> 794
ggccctgcag gcctatcaag tgctggaaac tttttctctt ggaatttttg caacatcaag 60

tcatagtcaa ttgaattgac ccaatttcac atttaagatt ttttttttt catccgacat 120

acatctgtac actaggaagc cctgtttttc tgaagcagct tcaaatatat atattttta 180

catatttatt atgattcaat gaacaatcta attaaatcga aaacaagaac cgaaacgcga 240

ataaataatt tatttagatg gtgacaagtg tataagtcct catcgggaca gctacgattt 300

ctctttcggt tttggctgag ctactggttg ctgtgacgca gcggcattag cgcggcgtta 360

tgagctaccc tcgtggcctg aaagatggcg ggaataaagc ggaactaaaa attactgact 420

gagccatatt gaggtcaatt tgtcaactcg tcaagtcacg tttggtggac ggccccttc 480

caacgaatcg tatatactaa catgcgcgcg cttcctatat acacatatac atatatatat 540

atatatatat gtgtgcgtgt atgtgtacac ctgtatttaa tttccttact cgcgggtttt 600

tctttttct caattcttgg cttcctcttt ctcgagtata taatttttca ggtaaaattt 660

agtacgatag taaaatactt ctcgaactcg tcacatatac gtgtacataa tgtctgaacc 720

agctcaaaag aaacaaaagg ttgctaacaa ctctctaga 759

<210> 795
<211> 643
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae FBA1 promoter

<400> 795
gaaatgaata acaatactga cagtactaaa taattgccta cttggcttca catacgttgc 60

atacgtcgat atagataata atgataatga cagcaggatt atcgtaatac gtaatagttg 120

aaaatctcaa aaatgtgtgg gtcattacgt aaataatgat aggaatggga ttcttctatt 180

tttccttttt ccattctagc agccgtcggg aaaacgtggc atcctctctt tcgggctcaa 240

ttggagtcac gctgccgtga gcatcctctc tttccatatc taacaactga gcacgtaacc 300

aatggaaaag catgagctta gcgttgctcc aaaaaagtat tggatggtta ataccatttg 360

tctgttctct tctgactttg actcctcaaa aaaaaaaaat ctacaatcaa cagatcgctt 420

caattacgcc ctcacaaaaa ctttttttcct tcttcttcgc ccacgttaaa ttttatccct 480

catgttgtct aacggatttc tgcacttgat ttattataaa aagacaaaga cataatactt 540

ctctatcaat ttcagttatt gttcttcctt gcgttattct tctgttcttc tttttctttt 600

gtcatatata accataacca agtaatacat attcaaatct aga 643

<210> 796
<211> 1188
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae KARI ILV5

<400> 796
atgttgagaa ctcaagccgc cagattgatc tgcaactccc gtgtcatcac tgctaagaga 60

acctttgctt tggccacccg tgctgctgct tacagcagac cagctgcccg tttcgttaag 120

ccaatgatca ctacccgtgg tttgaagcaa atcaacttcg gtggtactgt tgaaaccgtc 180

tacgaaagag ctgactggcc aagagaaaag ttgttggact acttcaagaa cgacactttt 240

gctttgatcg gttacggttc ccaaggttac ggtcaaggtt tgaacttgag agacaacggt 300

ttgaacgtta tcattggtgt ccgtaaagat ggtgcttctt ggaaggctgc catcgaagac 360

ggttgggttc caggcaagaa cttgttcact gttgaagatg ctatcaagag aggtagttac 420

gttatgaact tgttgtccga tgccgctcaa tcagaaacct ggcctgctat caagccattg 480

ttgaccaagg gtaagacttt gtacttctcc cacggtttct ccccagtctt caaggacttg 540

actcacgttg aaccaccaaa ggacttagat gttatcttgg ttgctccaaa gggttccggt 600

agaactgtca gatctttgtt caaggaaggt cgtggtatta actcttctta cgccgtctgg 660

aacgatgtca ccggtaaggc tcacgaaaag gcccaagctt tggccgttgc cattggttcc 720

ggttacgttt accaaaccac tttcgaaaga gaagtcaact ctgacttgta cggtgaaaga 780

ggttgtttaa tgggtggtat ccacggtatg ttcttggctc aatacgacgt cttgagagaa 840

aacggtcact ccccatctga agctttcaac gaaaccgtcg aagaagctac ccaatctcta 900

tacccattga tcggtaagta cggtatggat tacatgtacg atgcttgttc caccaccgcc 960

agaagaggtg ctttggactg gtacccaatc ttcaagaatg ctttgaagcc tgttttccaa 1020

gacttgtacg aatctaccaa gaacggtacc gaaaccaaga gatctttgga attcaactct 1080

caacctgact acagagaaaa gctagaaaag gaattagaca ccatcagaaa catggaaatc 1140

tggaaggttg gtaaggaagt cagaaagttg agaccagaaa accaataa 1188

<210> 797
<211> 395
<212> PRT
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae KARI coding region

<400> 797

Met Leu Arg Thr Gln Ala Ala Arg Leu Ile Cys Asn Ser Arg Val Ile
1 5 10 15

Thr Ala Lys Arg Thr Phe Ala Leu Ala Thr Arg Ala Ala Ala Tyr Ser
20 25 30

Arg Pro Ala Ala Arg Phe Val Lys Pro Met Ile Thr Thr Arg Gly Leu
35 40 45

Lys Gln Ile Asn Phe Gly Gly Thr Val Glu Thr Val Tyr Glu Arg Ala
50 55 60

Asp Trp Pro Arg Glu Lys Leu Leu Asp Tyr Phe Lys Asn Asp Thr Phe
65 70 75 80

Ala Leu Ile Gly Tyr Gly Ser Gln Gly Tyr Gly Gln Gly Leu Asn Leu
85 90 95

```
Arg Asp Asn Gly Leu Asn Val Ile Ile Gly Val Arg Lys Asp Gly Ala
            100                 105                 110

Ser Trp Lys Ala Ala Ile Glu Asp Gly Trp Val Pro Gly Lys Asn Leu
        115                 120                 125

Phe Thr Val Glu Asp Ala Ile Lys Arg Gly Ser Tyr Val Met Asn Leu
    130                 135                 140

Leu Ser Asp Ala Ala Gln Ser Glu Thr Trp Pro Ala Ile Lys Pro Leu
145                 150                 155                 160

Leu Thr Lys Gly Lys Thr Leu Tyr Phe Ser His Gly Phe Ser Pro Val
                165                 170                 175

Phe Lys Asp Leu Thr His Val Glu Pro Pro Lys Asp Leu Asp Val Ile
            180                 185                 190

Leu Val Ala Pro Lys Gly Ser Gly Arg Thr Val Arg Ser Leu Phe Lys
        195                 200                 205

Glu Gly Arg Gly Ile Asn Ser Ser Tyr Ala Val Trp Asn Asp Val Thr
    210                 215                 220

Gly Lys Ala His Glu Lys Ala Gln Ala Leu Ala Val Ala Ile Gly Ser
225                 230                 235                 240

Gly Tyr Val Tyr Gln Thr Thr Phe Glu Arg Glu Val Asn Ser Asp Leu
                245                 250                 255

Tyr Gly Glu Arg Gly Cys Leu Met Gly Gly Ile His Gly Met Phe Leu
            260                 265                 270

Ala Gln Tyr Asp Val Leu Arg Glu Asn Gly His Ser Pro Ser Glu Ala
        275                 280                 285

Phe Asn Glu Thr Val Glu Glu Ala Thr Gln Ser Leu Tyr Pro Leu Ile
    290                 295                 300

Gly Lys Tyr Gly Met Asp Tyr Met Tyr Asp Ala Cys Ser Thr Thr Ala
305                 310                 315                 320

Arg Arg Gly Ala Leu Asp Trp Tyr Pro Ile Phe Lys Asn Ala Leu Lys
                325                 330                 335

Pro Val Phe Gln Asp Leu Tyr Glu Ser Thr Lys Asn Gly Thr Glu Thr
            340                 345                 350

Lys Arg Ser Leu Glu Phe Asn Ser Gln Pro Asp Tyr Arg Glu Lys Leu
        355                 360                 365

Glu Lys Glu Leu Asp Thr Ile Arg Asn Met Glu Ile Trp Lys Val Gly
    370                 375                 380
```

Lys Glu Val Arg Lys Leu Arg Pro Glu Asn Gln
385 390 395

<210> 798
<211> 244
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae CYC1 terminator

<400> 798
attaaagcct tcgagcgtcc caaaaccttc tcaagcaagg ttttcagtat aatgttacat 60

gcgtacacgc gtctgtacag aaaaaaaaga aaaatttgaa atataaataa cgttcttaat 120

actaacataa ctataaaaaa ataaataggg acctagactt caggttgtct aactccttcc 180

ttttcggtta gagcggatgt gggggggaggg cgtgaatgta agcgtgacat aactaattac 240

atga 244

<210> 799
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> Pf5.IlvC-JEA1 variant

<400> 799
atgaaagttt tctacgataa agactgcgac ctgtcgatca tccaaggtaa gaaagttgcc 60

atcatcggct tcggttccca gggccacgct caagcactca acctgaagga ttccggcgta 120

gacgtgactg ttggcctgcc taaaggcttt gctgatgtag ccaaggctga agcccacggc 180

tttaaagtga ccgacgttgc tgcagccgtt gccggtgccg acttggtcat gatcctgatt 240

ccggacgagt tccagtccca gctgtacaag aacgaaatcg agccgaacat caagaagggc 300

gccactctgg ccttctccca cggcttcgcg atccactaca accaggttgt gcctcgtgcc 360

gacctcgacg tgatcatgat cgcgccgaag gctccaggcc acaccgtacg ttccgagttc 420

gtcaagggcg gaggtattcc tgacctgatc gcgatctacc aggacgtttc cggcaacgcc 480

aagaacgtcg ccctgtccta cgccgcaggc gtgggcggcg gccgtaccgg catcatcgaa 540

accaccttca aggacgagac tgaaaccgac ctgttcggtg agcaggctgt tctgtgtggc 600

ggtaccgtcg agctggtcaa agccggtttc gaaaccctgg ttgaagctgg ctacgctcca 660

gaaatggcct acttcgagtg cctgcacgaa ctgaagctga tcgttgacct catgtacgaa 720

ggcggtatcg ccaacatgaa ctactcgatc tccaacaacg ctgaatacgg cgagtacgtg 780

actggtccag aagtcatcaa cgccgaatcc cgtcaggcca tgcgcaatgc tctgaagcgc 840

atccaggacg gcgaatacgc gaagatgttc atcagcgaag gcgctaccgg ctacccatcg 900

atgaccgcca agcgtcgtaa caacgctgct cacggtatcg aaatcatcgg cgagcaactg 960

cgctcgatga tgccttggat cggtgccaac aaaatcgtcg acaaagccaa gaactaa 1017

<210> 800
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> Pf5.IlvC-JEA1 variant

<400> 800

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                  10                  15

Lys Lys Val Ala Ile Ile Gly Phe Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Leu Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Pro Lys
        35                  40                  45

Gly Phe Ala Asp Val Ala Lys Ala Glu Ala His Gly Phe Lys Val Thr
    50                  55                  60

Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Ile
65                  70                  75                  80

Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Val Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
            260                 265                 270
```

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
275 280 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
290 295 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305 310 315 320

Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
325 330 335

Lys Asn

<210> 801
<211> 1644
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimized sequence

<400> 801
atgtatacag taggtgacta tctgttggac agattacacg aattaggtat agaagaaata 60

ttcggagtac caggtgacta caatttgcaa tttctagatc aaattatttc acacaaagat 120

atgaaatggg tgggaaatgc taatgagtta aatgcctcct atatggccga cgggtacgca 180

agaacgaaaa aggctgcggc attcttgact acatttggtg ttggcgaatt atccgcagtt 240

aatggcttag cgggctccta tgctgagaac ctgcctgttg ttgagatcgt gggatctcct 300

acctcgaaag tgcagaacga aggtaagttt gttcaccata cgttggctga tggtgatttc 360

aagcacttta tgaagatgca cgaaccggtt actgctgcca ggactttatt gacagccgag 420

aatgcaactg ttgaaattga tagagtgttg tctgccttac taaaggaaag aaagccggtt 480

tacatcaatt tacctgtaga tgtagctgcc gctaaggctg aaaaaccatc cttgcctctt 540

aagaaggaaa attccacgtc gaatacatct gatcaagaga ttctgaacaa aatacaggaa 600

agtctgaaga atgccaagaa accaattgta atcacaggcc atgaaattat atcgttcggc 660

ctagagaaga ctgttactca gtttatttca aagactaagt tacctattac tactttgaac 720

tttggtaaat catctgttga tgaagcattg ccctcatttt tggggattta caacggtact 780

ctgtcagagc caaacttgaa ggaatttgtg gaatctgctg attttattct tatgttgggt 840

gtaaagctta ccgattctag tacgggtgca tttactcacc atcttaatga aaataaaatg 900

atttccttga atatcgatga aggtaaaatt ttcaacgaaa gaatccaaaa tttcgacttc 960

gaatccctga tatcatctct tcttgacttg tccgaaattg aatataaagg caagtacata 1020

gataaaaagc aagaagattt tgtaccttct aacgcgctgt tgtcacaaga tagactgtgg 1080

caagctgtcg aaaatttgac ccaaagtaat gagacgatcg tggctgaaca aggcacttct 1140

ttcttcggtg cctcatctat atttctgaaa tcgaaatcac attttattgg tcaacccttg 1200

tggggatcta taggatacac tttccccgca gctctaggca gccaaattgc agataaagaa 1260

tctagacatt tattgtttat cggagatgga tcattgcaac tgactgtcca agaattagga 1320

ctagccatta gagagaagat aaacccaatc tgctttatca ttaataacga tggttacacg 1380

gttgagaggg aaattcatgg tccgaaccag agttataatg acattcctat gtggaattac 1440

tcaaaactgc cagaaagttt cggggcaacg gaagacagag ttgtgtccaa aattgtgaga 1500

acagaaaatg aattcgtatc cgtgatgaaa gaagctcaag cagatccaaa taggatgtat 1560

tggatagaac ttattctagc aaaggagggt gcacctaaag ttttgaaaaa gatgggtaag 1620

ttatttgcag aacaaaacaa gagc 1644

<210> 802
<211> 1125
<212> DNA
<213> Artificial Sequence

<220>
<223> codon optimized sequence

<400> 802
atgtcaacag ccggtaaagt tattaagtgt aaagcggcag ttttgtggga agagaaaaag 60

ccgtttagca tagaagaagt agaagtagcg ccaccaaaag cacacgaggt tagaatcaag 120

atggttgcca ccggaatctg tagatccgac gaccatgtgg tgagtggcac tctagttact 180

cctttgccag taatcgcggg acacgaggct gccggaatcg ttgaatccat aggtgaaggt 240

gttaccactg ttcgtcctgg tgataaagtg atcccactgt tcactcctca atgtggtaag 300

tgtagagtct gcaaacatcc tgagggtaat ttctgcctta aaaatgattt gtctatgcct 360

agaggtacta tgcaggatgg tacaagcaga tttacatgca gagggaaacc tatacaccat 420

ttccttggta cttctacatt ttcccaatac acagtggtgg acgagatatc tgtcgctaaa 480

atcgatgcag cttcaccact ggaaaaagtt tgcttgatag ggtgcggatt ttccaccggt 540

tacggttccg cagttaaagt tgcaaaggtt acacagggtt cgacttgtgc agtattcggt 600

ttaggaggag taggactaag cgttattatg gggtgtaaag ctgcaggcgc agcgaggatt 660

ataggtgtag acatcaataa ggacaaattt gcaaaagcta aggaggtcgg ggctactgaa 720

tgtgttaacc ctcaagatta taagaaacca atacaagaag tccttactga aatgtcaaac 780

ggtggagttg atttctcttt tgaagttata ggccgtcttg atactatggt aactgcgttg 840

tcctgctgtc aagaggcata tggagtcagt gtgatcgtag gtgttcctcc tgattcacaa 900

aatttgtcga tgaatcctat gctgttgcta agcggtcgta catggaaggg agctatattt 960

ggcggtttta agagcaagga tagtgttcca aaacttgttg ccgactttat ggcgaagaag 1020

tttgctcttg atcctttaat tacacatgta ttgccattcg agaaaatcaa tgaagggttt 1080

gatttgttaa gaagtggtga atctattcgt acaattttaa ctttt 1125

<210> 803
<211> 375
<212> PRT
<213> Artificial Sequence

<220>
<223> Equus caballus

<400> 803

```
Met Ser Thr Ala Gly Lys Val Ile Lys Cys Lys Ala Ala Val Leu Trp
1               5                  10                  15

Glu Glu Lys Lys Pro Phe Ser Ile Glu Glu Val Glu Val Ala Pro Pro
            20                  25                  30

Lys Ala His Glu Val Arg Ile Lys Met Val Ala Thr Gly Ile Cys Arg
        35                  40                  45

Ser Asp Asp His Val Val Ser Gly Thr Leu Val Thr Pro Leu Pro Val
    50                  55                  60

Ile Ala Gly His Glu Ala Ala Gly Ile Val Glu Ser Ile Gly Glu Gly
65                  70                  75                  80

Val Thr Thr Val Arg Pro Gly Asp Lys Val Ile Pro Leu Phe Thr Pro
                85                  90                  95

Gln Cys Gly Lys Cys Arg Val Cys Lys His Pro Glu Gly Asn Phe Cys
            100                 105                 110

Leu Lys Asn Asp Leu Ser Met Pro Arg Gly Thr Met Gln Asp Gly Thr
        115                 120                 125

Ser Arg Phe Thr Cys Arg Gly Lys Pro Ile His His Phe Leu Gly Thr
    130                 135                 140

Ser Thr Phe Ser Gln Tyr Thr Val Val Asp Glu Ile Ser Val Ala Lys
145                 150                 155                 160

Ile Asp Ala Ala Ser Pro Leu Glu Lys Val Cys Leu Ile Gly Cys Gly
                165                 170                 175

Phe Ser Thr Gly Tyr Gly Ser Ala Val Lys Val Ala Lys Val Thr Gln
            180                 185                 190

Gly Ser Thr Cys Ala Val Phe Gly Leu Gly Gly Val Gly Leu Ser Val
        195                 200                 205

Ile Met Gly Cys Lys Ala Ala Gly Ala Ala Arg Ile Ile Gly Val Asp
    210                 215                 220

Ile Asn Lys Asp Lys Phe Ala Lys Ala Lys Glu Val Gly Ala Thr Glu
225                 230                 235                 240

Cys Val Asn Pro Gln Asp Tyr Lys Lys Pro Ile Gln Glu Val Leu Thr
                245                 250                 255

Glu Met Ser Asn Gly Gly Val Asp Phe Ser Phe Glu Val Ile Gly Arg
            260                 265                 270

Leu Asp Thr Met Val Thr Ala Leu Ser Cys Cys Gln Glu Ala Tyr Gly
        275                 280                 285
```

```
Val Ser Val Ile Val Gly Val Pro Pro Asp Ser Gln Asn Leu Ser Met
    290                 295                 300

Asn Pro Met Leu Leu Leu Ser Gly Arg Thr Trp Lys Gly Ala Ile Phe
305                 310                 315                 320

Gly Gly Phe Lys Ser Lys Asp Ser Val Pro Lys Leu Val Ala Asp Phe
                325                 330                 335

Met Ala Lys Lys Phe Ala Leu Asp Pro Leu Ile Thr His Val Leu Pro
            340                 345                 350

Phe Glu Lys Ile Asn Glu Gly Phe Asp Leu Leu Arg Ser Gly Glu Ser
        355                 360                 365

Ile Arg Thr Ile Leu Thr Phe
    370                 375


<210>  804
<211>  548
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Lactococcus lactis

<400>  804

Met Tyr Thr Val Gly Asp Tyr Leu Leu Asp Arg Leu His Glu Leu Gly
1               5                   10                  15

Ile Glu Glu Ile Phe Gly Val Pro Gly Asp Tyr Asn Leu Gln Phe Leu
            20                  25                  30

Asp Gln Ile Ile Ser His Lys Asp Met Lys Trp Val Gly Asn Ala Asn
        35                  40                  45

Glu Leu Asn Ala Ser Tyr Met Ala Asp Gly Tyr Ala Arg Thr Lys Lys
    50                  55                  60

Ala Ala Ala Phe Leu Thr Thr Phe Gly Val Gly Glu Leu Ser Ala Val
65                  70                  75                  80

Asn Gly Leu Ala Gly Ser Tyr Ala Glu Asn Leu Pro Val Val Glu Ile
                85                  90                  95

Val Gly Ser Pro Thr Ser Lys Val Gln Asn Glu Gly Lys Phe Val His
            100                 105                 110

His Thr Leu Ala Asp Gly Asp Phe Lys His Phe Met Lys Met His Glu
        115                 120                 125

Pro Val Thr Ala Ala Arg Thr Leu Leu Thr Ala Glu Asn Ala Thr Val
    130                 135                 140
```

```
Glu Ile Asp Arg Val Leu Ser Ala Leu Leu Lys Glu Arg Lys Pro Val
145                 150                 155                 160

Tyr Ile Asn Leu Pro Val Asp Val Ala Ala Ala Lys Ala Glu Lys Pro
                165                 170                 175

Ser Leu Pro Leu Lys Lys Glu Asn Ser Thr Ser Asn Thr Ser Asp Gln
            180                 185                 190

Glu Ile Leu Asn Lys Ile Gln Glu Ser Leu Lys Asn Ala Lys Lys Pro
        195                 200                 205

Ile Val Ile Thr Gly His Glu Ile Ile Ser Phe Gly Leu Glu Lys Thr
    210                 215                 220

Val Thr Gln Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
225                 230                 235                 240

Phe Gly Lys Ser Ser Val Asp Glu Ala Leu Pro Ser Phe Leu Gly Ile
                245                 250                 255

Tyr Asn Gly Thr Leu Ser Glu Pro Asn Leu Lys Glu Phe Val Glu Ser
            260                 265                 270

Ala Asp Phe Ile Leu Met Leu Gly Val Lys Leu Thr Asp Ser Ser Thr
        275                 280                 285

Gly Ala Phe Thr His His Leu Asn Glu Asn Lys Met Ile Ser Leu Asn
    290                 295                 300

Ile Asp Glu Gly Lys Ile Phe Asn Glu Arg Ile Gln Asn Phe Asp Phe
305                 310                 315                 320

Glu Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser Glu Ile Glu Tyr Lys
                325                 330                 335

Gly Lys Tyr Ile Asp Lys Lys Gln Glu Asp Phe Val Pro Ser Asn Ala
            340                 345                 350

Leu Leu Ser Gln Asp Arg Leu Trp Gln Ala Val Glu Asn Leu Thr Gln
        355                 360                 365

Ser Asn Glu Thr Ile Val Ala Glu Gln Gly Thr Ser Phe Phe Gly Ala
    370                 375                 380

Ser Ser Ile Phe Leu Lys Ser Lys Ser His Phe Ile Gly Gln Pro Leu
385                 390                 395                 400

Trp Gly Ser Ile Gly Tyr Thr Phe Pro Ala Ala Leu Gly Ser Gln Ile
                405                 410                 415

Ala Asp Lys Glu Ser Arg His Leu Leu Phe Ile Gly Asp Gly Ser Leu
            420                 425                 430
```

```
Gln Leu Thr Val Gln Glu Leu Gly Leu Ala Ile Arg Glu Lys Ile Asn
        435                 440                 445

Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
    450                 455                 460

Ile His Gly Pro Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
465                 470                 475                 480

Ser Lys Leu Pro Glu Ser Phe Gly Ala Thr Glu Asp Arg Val Val Ser
                485                 490                 495

Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
            500                 505                 510

Gln Ala Asp Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys
        515                 520                 525

Glu Gly Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
    530                 535                 540

Gln Asn Lys Ser
545
```

<210> 805
<211> 1023
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae TDH3 promoter

<400> 805

```
caccgcggtg gggcgcgccc tattttcgag gaccttgtca ccttgagccc aagagagcca       60

agatttaaat tttcctatga cttgatgcaa attcccaaag ctaataacat gcaagacacg      120

tacggtcaag aagacatatt tgacctctta acaggttcag acgcgactgc ctcatcagta      180

agacccgttg aaaagaactt acctgaaaaa aacgaatata tactagcgtt gaatgttagc      240

gtcaacaaca agaagtttaa tgacgcggag gccaaggcaa aaagattcct tgattacgta      300

agggagttag aatcattttg aataaaaaac acgctttttc agttcgagtt tatcattatc      360

aatactgcca tttcaaagaa tacgtaaata attaatagta gtgattttcc taactttatt      420

tagtcaaaaa attagccttt taattctgct gtaacccgta catgcccaaa atagggggcg      480

ggttacacag aatatataac atcgtaggtg tctgggtgaa cagtttattc ctggcatcca      540

ctaaatataa tggagcccgc tttttaagct ggcatccaga aaaaaaaaga atcccagcac      600

caaaatattg ttttcttcac caaccatcag ttcataggtc cattctctta gcgcaactac      660

agagaacagg ggcacaaaca ggcaaaaaac gggcacaacc tcaatggagt gatgcaacct      720

gcctggagta aatgatgaca caaggcaatt gacccacgca tgtatctatc tcattttctt      780

acaccttcta ttaccttctg ctctctctga tttggaaaaa gctgaaaaaa aaggttgaaa      840

ccagttccct gaaattattc ccctacttga ctaataagta tataaagacg gtaggtattg      900

attgtaattc tgtaaatcta tttcttaaac ttcttaaatt ctacttttat agttagtctt      960
```

tttttagtt ttaaaacacc aagaacttag tttcgaataa acacacataa actagtaaac 1020

aaa 1023

<210> 806
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer OT1068

<400> 806
caaaagctga gctccaccgc g 21

<210> 807
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer OT1067

<400> 807
gtttactagt ttatgtgtgt ttattcgaaa ctaagttctt ggtg 44

<210> 808
<211> 9333
<212> DNA
<213> Artificial Sequence

<220>
<223> pLH467

<400> 808
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataccac agcttttcaa ttcaattcat catttttttt ttattctttt ttttgatttc 240

ggtttctttg aaatttttt gattcggtaa tctccgaaca gaaggaagaa cgaaggaagg 300

agcacagact tagattggta tatatacgca tatgtagtgt tgaagaaaca tgaaattgcc 360

cagtattctt aacccaactg cacagaacaa aaacctgcag gaaacgaaga taaatcatgt 420

cgaaagctac atataaggaa cgtgctgcta ctcatcctag tcctgttgct gccaagctat 480

ttaatatcat gcacgaaaag caaacaaact tgtgtgcttc attggatgtt cgtaccacca 540

aggaattact ggagttagtt gaagcattag gtcccaaaat ttgtttacta aaaacacatg 600

tggatatctt gactgatttt tccatggagg gcacagttaa gccgctaaag gcattatccg 660

ccaagtacaa ttttttactc ttcgaagaca gaaaatttgc tgacattggt aatacagtca 720

aattgcagta ctctgcgggt gtatacagaa tagcagaatg ggcagacatt acgaatgcac 780

acggtgtggt gggcccaggt attgttagcg gtttgaagca ggcggcagaa gaagtaacaa 840

aggaacctag aggccttttg atgttagcag aattgtcatg caagggctcc ctatctactg 900

gagaatatac taagggtact gttgacattg cgaagagcga caaagatttt gttatcggct 960

ttattgctca aagagacatg ggtggaagag atgaaggtta cgattggttg attatgacac 1020

20 Jan 2017 2017200439

ccggtgtggg tttagatgac aagggagacg cattgggtca acagtataga accgtggatg 1080
atgtggtctc tacaggatct gacattatta ttgttggaag aggactattt gcaaagggaa 1140
gggatgctaa ggtagagggt gaacgttaca gaaaagcagg ctgggaagca tatttgagaa 1200
gatgcggcca gcaaaactaa aaaactgtat tataagtaaa tgcatgtata ctaaactcac 1260
aaattagagc ttcaatttaa ttatatcagt tattacccta tgcggtgtga aataccgcac 1320
agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa cgttaatatt ttgttaaaat 1380
tcgcgttaaa tttttgttaa atcagctcat tttttaacca ataggccgaa atcggcaaaa 1440
tcccttataa atcaaaagaa tagaccgaga tagggttgag tgttgttcca gtttggaaca 1500
agagtccact attaaagaac gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg 1560
gcgatggccc actacgtgaa ccatcaccct aatcaagttt tttggggtcg aggtgccgta 1620
aagcactaaa tcggaaccct aaagggagcc cccgatttag agcttgacgg ggaaagccgg 1680
cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa 1740
gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc gcttaatgcg ccgctacagg 1800
gcgcgtcgcg ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg 1860
cctcttcgct attacgccag ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg 1920
taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgag cgcgcgtaat 1980
acgactcact atagggcgaa ttgggtaccg ggcccccccct cgaggtcgac tggccattaa 2040
tctttcccat attagatttc gccaagccat gaaagttcaa gaaaggtctt tagacgaatt 2100
acccttcatt tctcaaactg gcgtcaaggg atcctggtat ggttttatcg ttttatttct 2160
ggttcttata gcatcgtttt ggacttctct gttcccatta ggcggttcag gagccagcgc 2220
agaatcattc tttgaaggat acttatcctt tccaattttg attgtctgtt acgttggaca 2280
taaactgtat actagaaatt ggactttgat ggtgaaacta gaagatatgg atcttgatac 2340
cggcagaaaa caagtagatt tgactcttcg tagggaagaa atgaggattg agcgagaaac 2400
attagcaaaa agatccttcg taacaagatt tttacatttc tggtgttgaa gggaaagata 2460
tgagctatac agcggaattt ccatatcact cagattttgt tatctaattt tttccttccc 2520
acgtccgcgg gaatctgtgt atattactgc atctagatat atgttatctt atcttggcgc 2580
gtacatttaa ttttcaacgt attctataag aaattgcggg agtttttttc atgtagatga 2640
tactgactgc acgcaaatat aggcatgatt tataggcatg atttgatggc tgtaccgata 2700
ggaacgctaa gagtaacttc agaatcgtta tcctggcgga aaaaattcat ttgtaaactt 2760
taaaaaaaaa agccaatatc cccaaaatta ttaagagcgc ctccattatt aactaaaatt 2820
tcactcagca tccacaatgt atcaggtatc tactacagat attacatgtg gcgaaaaaga 2880
caagaacaat gcaatagcgc atcaagaaaa aacacaaagc tttcaatcaa tgaatcgaaa 2940
atgtcattaa aatagtatat aaattgaaac taagtcataa agctataaaa agaaaattta 3000
tttaaatgca agatttaaag taaattcacg gccctgcagg cctcagctct tgttttgttc 3060
tgcaaataac ttacccatct ttttcaaaac tttaggtgca ccctcctttg ctagaataag 3120
ttctatccaa tacatcctat ttggatctgc ttgagcttct ttcatcacgg atacgaattc 3180

attttctgtt ctcacaattt tggacacaac tctgtcttcc gttgccccga aactttctgg 3240
cagttttgag taattccaca taggaatgtc attataactc tggttcggac catgaatttc 3300
cctctcaacc gtgtaaccat cgttattaat gataaagcag attgggttta tcttctctct 3360
aatggctagt cctaattctt ggacagtcag ttgcaatgat ccatctccga taaacaataa 3420
atgtctagat tctttatctg caatttggct gcctagagct gcggggaaag tgtatcctat 3480
agatccccac aagggttgac caataaaatg tgatttcgat ttcagaaata tagatgaggc 3540
accgaagaaa gaagtgcctt gttcagccac gatcgtctca ttactttggg tcaaattttc 3600
gacagcttgc cacagtctat cttgtgacaa cagcgcgtta gaaggtacaa aatcttcttg 3660
ctttttatct atgtacttgc ctttatattc aatttcggac aagtcaagaa gagatgatat 3720
cagggattcg aagtcgaaat tttggattct ttcgttgaaa attttacctt catcgatatt 3780
caaggaaatc attttatttt cattaagatg gtgagtaaat gcacccgtac tagaatcggt 3840
aagctttaca cccaacataa gaataaaatc agcagattcc acaaattcct tcaagtttgg 3900
ctctgacaga gtaccgttgt aaatccccaa aaatgagggc aatgcttcat caacagatga 3960
tttaccaaag ttcaaagtag taataggtaa cttagtcttt gaaataaact gagtaacagt 4020
cttctctagg ccgaacgata taatttcatg gcctgtgatt acaattggtt tcttggcatt 4080
cttcagactt tcctgtattt tgttcagaat ctcttgatca gatgtattcg acgtggaatt 4140
ttccttctta agaggcaagg atggtttttc agccttagcg gcagctacat ctacaggtaa 4200
attgatgtaa accggctttc tttcctttag taaggcagac aacactctat caatttcaac 4260
agttgcattc tcggctgtca ataaagtcct ggcagcagta accggttcgt gcatcttcat 4320
aaagtgcttg aaatcaccat cagccaacgt atggtgaaca aacttacctt cgttctgcac 4380
tttcgaggta ggagatccca cgatctcaac aacaggcagg ttctcagcat aggagcccgc 4440
taagccatta actgcggata attcgccaac accaaatgta gtcaagaatg ccgcagcctt 4500
tttcgttctt gcgtacccgt cggccatata ggaggcattt aactcattag catttcccac 4560
ccatttcata tctttgtgtg aaataatttg atctagaaat tgcaaattgt agtcacctgg 4620
tactccgaat atttcttcta tacctaattc gtgtaatctg tccaacagat agtcacctac 4680
tgtatacatt ttgtttacta gtttatgtgt gtttattcga aactaagttc ttggtgtttt 4740
aaaactaaaa aaaagactaa ctataaaagt agaatttaag aagtttaaga aatagattta 4800
cagaattaca atcaatacct accgtcttta tatacttatt agtcaagtag gggaataatt 4860
tcagggaact ggtttcaacc tttttttca gcttttcca aatcagagag agcagaaggt 4920
aatagaaggt gtaagaaaat gagatagata catgcgtggg tcaattgcct tgtgtcatca 4980
tttactccag gcaggttgca tcactccatt gaggttgtgc ccgtttttg cctgtttgtg 5040
cccctgttct ctgtagttgc gctaagagaa tggacctatg aactgatggt tggtgaagaa 5100
aacaatattt tggtgctggg attctttttt tttctggatg ccagcttaaa aagcgggctc 5160
cattatattt agtggatgcc aggaataaac tgttcaccca gacacctacg atgttatata 5220
ttctgtgtaa cccgcccccc attttgggca tgtacgggtt acagcagaat taaaaggcta 5280
atttttgac taaataaagt taggaaaatc actactatta attatttacg tattctttga 5340

aatggcagta ttgataatga taaactcgaa ctgaaaaagc gtgtttttta ttcaaaatga 5400
ttctaactcc cttacgtaat caaggaatct ttttgccttg gcctccgcgt cattaaactt 5460
cttgttgttg acgctaacat tcaacgctag tatatattcg ttttttttcag gtaagttctt 5520
ttcaacgggt cttactgatg aggcagtcgc gtctgaacct gttaagaggt caaatatgtc 5580
ttcttgaccg tacgtgtctt gcatgttatt agctttggga atttgcatca agtcatagga 5640
aaatttaaat cttggctctc ttgggctcaa ggtgacaagg tcctcgaaaa tagggcgcgc 5700
cccaccgcgg tggagctcca gcttttgttc cctttagtga gggttaattg cgcgcttggc 5760
gtaatcatgg tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa 5820
cataggagcc ggaagcataa agtgtaaagc ctggggtgcc taatgagtga ggtaactcac 5880
attaattgcg ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca 5940
ttaatgaatc ggccaacgcg cggggagagg cggtttgcgt attgggcgct cttccgcttc 6000
ctcgctcact gactcgctgc gctcggtcgt tcggctgcgg cgagcggtat cagctcactc 6060
aaaggcggta atacggttat ccacagaatc aggggataac gcaggaaaga acatgtgagc 6120
aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgctggcgt ttttccatag 6180
gctccgcccc cctgacgagc atcacaaaaa tcgacgctca agtcagaggt ggcgaaaccc 6240
gacaggacta taaagatacc aggcgtttcc ccctggaagc tccctcgtgc gctctcctgt 6300
tccgaccctg ccgcttaccg gatacctgtc cgcctttctc ccttcgggaa gcgtggcgct 6360
ttctcatagc tcacgctgta ggtatctcag ttcggtgtag gtcgttcgct ccaagctggg 6420
ctgtgtgcac gaacccccg ttcagcccga ccgctgcgcc ttatccggta actatcgtct 6480
tgagtccaac ccggtaagac acgacttatc gccactggca gcagccactg gtaacaggat 6540
tagcagagcg aggtatgtag gcggtgctac agagttcttg aagtggtggc ctaactacgg 6600
ctacactaga aggacagtat ttggtatctg cgctctgctg aagccagtta ccttcggaaa 6660
aagagttggt agctcttgat ccggcaaaca aaccaccgct ggtagcggtg gttttttgt 6720
ttgcaagcag cagattacgc gcagaaaaaa aggatctcaa gaagatcctt tgatcttttc 6780
tacggggtct gacgctcagt ggaacgaaaa ctcacgttaa gggattttgg tcatgagatt 6840
atcaaaaagg atcttcacct agatcctttt aaattaaaaa tgaagtttta aatcaatcta 6900
aagtatatat gagtaaactt ggtctgacag ttaccaatgc ttaatcagtg aggcacctat 6960
ctcagcgatc tgtctatttc gttcatccat agttgcctga ctccccgtcg tgtagataac 7020
tacgatacgg gagggcttac catctggccc cagtgctgca atgataccgc gagacccacg 7080
ctcaccggct ccagatttat cagcaataaa ccagccagcc ggaagggccg agcgcagaag 7140
tggtcctgca actttatccg cctccatcca gtctattaat tgttgccggg aagctagagt 7200
aagtagttcg ccagttaata gtttgcgcaa cgttgttgcc attgctacag gcatcgtggt 7260
gtcacgctcg tcgtttggta tggcttcatt cagctccggt tcccaacgat caaggcgagt 7320
tacatgatcc cccatgttgt gcaaaaaagc ggttagctcc ttcggtcctc cgatcgttgt 7380
cagaagtaag ttggccgcag tgttatcact catggttatg gcagcactgc ataattctct 7440
tactgtcatg ccatccgtaa gatgcttttc tgtgactggt gagtactcaa ccaagtcatt 7500

ctgagaatag tgtatgcggc gaccgagttg ctcttgcccg gcgtcaatac gggataatac 7560
cgcgccacat agcagaactt taaaagtgct catcattgga aaacgttctt cggggcgaaa 7620
actctcaagg atcttaccgc tgttgagatc cagttcgatg taacccactc gtgcacccaa 7680
ctgatcttca gcatctttta ctttcaccag cgtttctggg tgagcaaaaa caggaaggca 7740
aaatgccgca aaaaagggaa taagggcgac acggaaatgt tgaatactca tactcttcct 7800
ttttcaatat tattgaagca tttatcaggg ttattgtctc atgagcggat acatatttga 7860
atgtatttag aaaaataaac aaataggggt tccgcgcaca tttccccgaa aagtgccacc 7920
tgaacgaagc atctgtgctt cattttgtag aacaaaaatg caacgcgaga gcgctaattt 7980
ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat 8040
tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgagagcgct 8100
aatttttcaa acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgagagc 8160
gctattttac caacaaagaa tctatacttc ttttttgttc tacaaaaatg catcccgaga 8220
gcgctatttt tctaacaaag catcttagat tacttttttt ctcctttgtg cgctctataa 8280
tgcagtctct tgataacttt ttgcactgta ggtccgttaa ggttagaaga aggctacttt 8340
ggtgtctatt ttctcttcca taaaaaaagc ctgactccac ttcccgcgtt tactgattac 8400
tagcgaagct gcgggtgcat tttttcaaga taaaggcatc cccgattata ttctataccg 8460
atgtggattg cgcatacttt gtgaacagaa agtgatagcg ttgatgattc ttcattggtc 8520
agaaaattat gaacggtttc ttctattttg tctctatata ctacgtatag gaaatgttta 8580
cattttcgta ttgttttcga ttcactctat gaatagttct tactacaatt tttttgtcta 8640
aagagtaata ctagagataa acataaaaaa tgtagaggtc gagtttagat gcaagttcaa 8700
ggagcgaaag gtggatgggt aggttatata gggatatagc acagagatat atagcaaaga 8760
gatacttttg agcaatgttt gtggaagcgg tattcgcaat attttagtag ctcgttacag 8820
tccggtgcgt ttttggtttt ttgaaagtgc gtcttcagag cgcttttggt tttcaaaagc 8880
gctctgaagt tcctatactt tctagagaat aggaacttcg gaataggaac ttcaaagcgt 8940
ttccgaaaac gagcgcttcc gaaaatgcaa cgcgagctgc gcacatacag ctcactgttc 9000
acgtcgcacc tatatctgcg tgttgcctgt atatatatat acatgagaag aacggcatag 9060
tgcgtgttta tgcttaaatg cgtacttata tgcgtctatt tatgtaggat gaaaggtagt 9120
ctagtacctc ctgtgatatt atcccattcc atgcggggta tcgtatgctt ccttcagcac 9180
taccctttag ctgttctata tgctgccact cctcaattgg attagtctca tccttcaatg 9240
ctatcatttc ctttgatatt ggatcatact aagaaaccat tattatcatg acattaacct 9300
ataaaaatag gcgtatcacg aggccctttc gtc 9333

<210> 809
<211> 8994
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic construct pRS425::GPM-sadB

<400> 809
ctagttctag agcggccgcc accgcggtgg agctccagct tttgttccct ttagtgaggg 60
ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 120
ctcacaattc cacacaacat aggagccgga agcataaagt gtaaagcctg gggtgcctaa 180
tgagtgaggt aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac 240
ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 300
gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 360
gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 420
ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 480
ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt 540
cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 600
ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 660
tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 720
gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 780
tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 840
gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 900
tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 960
ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 1020
agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 1080
gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 1140
attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 1200
agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 1260
atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 1320
cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 1380
ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 1440
agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 1500
tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 1560
gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc 1620
caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc 1680
ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 1740
gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag 1800
tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 1860
tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa 1920
cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa 1980
cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga 2040
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 2100
atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 2160

agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 2220
ccccgaaaag tgccacctga acgaagcatc tgtgcttcat tttgtagaac aaaaatgcaa 2280
cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg 2340
caacgcgaaa gcgctatttt accaacgaag aatctgtgct tcatttttgt aaaacaaaaa 2400
tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt tacagaacag 2460
aaatgcaacg cgagagcgct attttaccaa caaagaatct atacttcttt tttgttctac 2520
aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac tttttttctc 2580
ctttgtgcgc tctataatgc agtctcttga taacttttg cactgtaggt ccgttaaggt 2640
tagaagaagg ctactttggt gtctattttc tcttccataa aaaaagcctg actccacttc 2700
ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa aggcatcccc 2760
gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt gatagcgttg 2820
atgattcttc attggtcaga aaattatgaa cggtttcttc tattttgtct ctatatacta 2880
cgtataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa tagttcttac 2940
tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt agaggtcgag 3000
tttagatgca agttcaagga gcgaaaggtg gatgggtagg ttatataggg atatagcaca 3060
gagatatata gcaaagagat acttttgagc aatgtttgtg gaagcggtat tcgcaatatt 3120
ttagtagctc gttacagtcc ggtgcgtttt tggtttttg aaagtgcgtc ttcagagcgc 3180
ttttggtttt caaaagcgct ctgaagttcc tatactttct agagaatagg aacttcggaa 3240
taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca 3300
catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata tatatataca 3360
tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc gtctatttat 3420
gtaggatgaa aggtagtcta gtacctcctg tgatattatc ccattccatg cggggtatcg 3480
tatgcttcct tcagcactac cctttagctg ttctatatgc tgccactcct caattggatt 3540
agtctcatcc ttcaatgcta tcatttcctt tgatattgga tcatactaag aaaccattat 3600
tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt 3660
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 3720
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 3780
tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatcga 3840
ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc accattatgg 3900
gaaatgcttc aagaaggtat tgacttaaac tccatcaaat ggtcaggtca ttgagtgttt 3960
tttatttgtt gtatttttt ttttttagag aaaatcctcc aatatcaaat taggaatcgt 4020
agtttcatga ttttctgtta cacctaactt tttgtgtggt gccctcctcc ttgtcaatat 4080
taatgttaaa gtgcaattct ttttccttat cacgttgagc cattagtatc aatttgctta 4140
cctgtattcc tttactatcc tccttttttct ccttcttgat aaatgtatgt agattgcgta 4200
tatagtttcg tctaccctat gaacatattc cattttgtaa tttcgtgtcg tttctattat 4260
gaatttcatt tataaagttt atgtacaaat atcataaaaa aagagaatct ttttaagcaa 4320

ggattttctt aacttcttcg gcgacagcat caccgacttc ggtggtactg ttggaaccac 4380
ctaaatcacc agttctgata cctgcatcca aaacctttt aactgcatct tcaatggcct 4440
taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac aagatagtgg 4500
cgatagggtc aaccttattc tttggcaaat ctggagcaga accgtggcat ggttcgtaca 4560
aaccaaatgc ggtgttcttg tctggcaaag aggccaagga cgcagatggc aacaaaccca 4620
aggaacctgg gataacggag gcttcatcgg agatgatatc accaaacatg ttgctggtga 4680
ttataatacc atttaggtgg gttgggttct taactaggat catggcggca gaatcaatca 4740
attgatgttg aaccttcaat gtagggaatt cgttcttgat ggtttcctcc acagtttttc 4800
tccataatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata ggcaatggtg 4860
gctcatgttg tagggccatg aaagcggcca ttcttgtgat tctttgcact tctggaacgg 4920
tgtattgttc actatcccaa gcgacaccat caccatcgtc ttcctttctc ttaccaaagt 4980
aaatacctcc cactaattct ctgacaacaa cgaagtcagt acctttagca aattgtggct 5040
tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatggtctt aagttggcgt 5100
acaattgaag ttctttacgg atttttagta aaccttgttc aggtctaaca ctaccggtac 5160
cccatttagg accagccaca gcacctaaca aaacggcatc aaccttcttg gaggcttcca 5220
gcgcctcatc tggaagtggg acacctgtag catcgatagc agcaccacca attaaatgat 5280
tttcgaaatc gaacttgaca ttggaacgaa catcagaaat agctttaaga accttaatgg 5340
cttcggctgt gatttcttga ccaacgtggt cacctggcaa aacgacgatc ttcttagggg 5400
cagacatagg ggcagacatt agaatggtat atccttgaaa tatatatata tattgctgaa 5460
atgtaaaagg taagaaaagt tagaaagtaa gacgattgct aaccacctat tggaaaaaac 5520
aataggtcct taaataatat tgtcaacttc aagtattgtg atgcaagcat ttagtcatga 5580
acgcttctct attctatatg aaaagccggt tccggcctct cacctttcct ttttctccca 5640
atttttcagt tgaaaaaggt atatgcgtca ggcgacctct gaaattaaca aaaaatttcc 5700
agtcatcgaa tttgattctg tgcgatagcg cccctgtgtg ttctcgttat gttgaggaaa 5760
aaaataatgg ttgctaagag attcgaactc ttgcatctta cgatacctga gtattcccac 5820
agttaactgc ggtcaagata tttcttgaat caggcgcctt agaccgctcg gccaaacaac 5880
caattacttg ttgagaaata gagtataatt atcctataaa tataacgttt ttgaacacac 5940
atgaacaagg aagtacagga caattgattt tgaagagaat gtggattttg atgtaattgt 6000
tgggattcca tttttaataa ggcaataata ttaggtatgt ggatatacta gaagttctcc 6060
tcgaccgtcg atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6120
aggaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt taaatcagct 6180
catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg 6240
agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag aacgtggact 6300
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac 6360
cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga 6420
gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga 6480

aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca 6540
ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc gcgccattcg ccattcaggc 6600
tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga 6660
aaggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac 6720
gttgtaaaac gacggccagt gagcgcgcgt aatacgactc actatagggc gaattgggta 6780
ccgggccccc cctcgaggtc gacggtatcg ataagcttga tatcgaattc ctgcagcccg 6840
ggggatccgc atgcttgcat ttagtcgtgc aatgtatgac tttaagattt gtgagcagga 6900
agaaaaggga gaatcttcta acgataaacc cttgaaaaac tgggtagact acgctatgtt 6960
gagttgctac gcaggctgca caattacacg agaatgctcc cgcctaggat ttaaggctaa 7020
gggacgtgca atgcagacga cagatctaaa tgaccgtgtc ggtgaagtgt tcgccaaact 7080
tttcggttaa cacatgcagt gatgcacgcg cgatggtgct aagttacata tatatatata 7140
tatatatata tagccatagt gatgtctaag taacctttat ggtatatttc ttaatgtgga 7200
aagatactag cgcgcgcacc cacacacaag cttcgtcttt tcttgaagaa aagaggaagc 7260
tcgctaaatg ggattccact ttccgttccc tgccagctga tggaaaaagg ttagtggaac 7320
gatgaagaat aaaaagagag atccactgag gtgaaatttc agctgacagc gagtttcatg 7380
atcgtgatga acaatggtaa cgagttgtgg ctgttgccag ggagggtggt tctcaacttt 7440
taatgtatgg ccaaatcgct acttgggttt gttatataac aaagaagaaa taatgaactg 7500
attctcttcc tccttcttgt cctttcttaa ttctgttgta attaccttcc tttgtaattt 7560
tttttgtaat tattcttctt aataatccaa acaaacacac atattacaat agctagctga 7620
ggatgaaggc attagtttat catggggatc acaaaatttc gttagaagac aaaccaaaac 7680
ccactctgca gaaaccaaca gacgttgtgg ttagggtgtt gaaaacaaca atttgcggta 7740
ctgacttggg aatatacaaa ggtaagaatc ctgaagtggc agatggcaga atcctgggtc 7800
atgagggcgt tggcgtcatt gaagaagtgg gcgaatccgt gacacaattc aaaaaggggg 7860
ataaagtttt aatctcctgc gttactagct gtggatcgtg tgattattgc aagaagcaac 7920
tgtattcaca ctgtagagac ggtggctgga ttttaggtta catgatcgac ggtgtccaag 7980
ccgaatacgt cagaatacca catgctgaca attcattgta taagatcccg caaactatcg 8040
atgatgaaat tgcagtacta ctgtccgata ttttacctac tggacatgaa attggtgttc 8100
aatatggtaa cgttcaacca ggcgatgctg tagcaattgt aggagcaggt cctgttggaa 8160
tgtcagtttt gttaactgct caattttact cgcctagtac cattattgtt atcgacatgg 8220
acgaaaaccg tttacaatta gcgaaggagc ttggggccac acacactatt aactccggta 8280
ctgaaaatgt tgtcgaagct gtgcatcgta tagcagccga aggagtggat gtagcaatag 8340
aagctgttgg tatacccgca acctgggaca tctgtcagga aattgtaaaa cccggcgctc 8400
atattgccaa cgtgggagtt catggtgtta aggtggactt tgaaattcaa aagttgtgga 8460
ttaagaatct aaccatcacc actggtttgg ttaacactaa tactacccca atgttgatga 8520
aggtagcctc tactgataaa ttgcctttaa agaaaatgat tactcacagg tttgagttag 8580
ctgaaatcga acacgcatat caggttttct tgaatggcgc taaagaaaaa gctatgaaga 8640

ttattctatc taatgcaggt gccgcctaat taattaagag taagcgaatt tcttatgatt 8700

tatgattttt attattaaat aagttataaa aaaaataagt gtatacaaat tttaaagtga 8760

ctcttaggtt ttaaaacgaa aattcttatt cttgagtaac tctttcctgt aggtcaggtt 8820

gctttctcag gtatagcatg aggtcgctct tattgaccac acctctaccg gcatgccgag 8880

caaatgcctg caaatcgctc cccatttcac ccaattgtag atatgctaac tccagcaatg 8940

agttgatgaa tctcggtgtg tattttatgt cctcagagga caacacctgt ggta 8994

<210> 810
<211> 753
<212> DNA
<213> Artificial Sequence

<220>
<223> Saccharomyces cerevisiae GPM1 promoter

<400> 810
gcatgcttgc atttagtcgt gcaatgtatg actttaagat ttgtgagcag gaagaaaagg 60

gagaatcttc taacgataaa cccttgaaaa actgggtaga ctacgctatg ttgagttgct 120

acgcaggctg cacaattaca cgagaatgct cccgcctagg atttaaggct aagggacgtg 180

caatgcagac gacagatcta aatgaccgtg tcggtgaagt gttcgccaaa cttttcggtt 240

aacacatgca gtgatgcacg cgcgatggtg ctaagttaca tatatatata tatagccata 300

gtgatgtcta agtaaccttt atggtatatt tcttaatgtg gaaagatact agcgcgcgca 360

cccacacaca agcttcgtct tttcttgaag aaaagaggaa gctcgctaaa tgggattcca 420

ctttccgttc cctgccagct gatggaaaaa ggttagtgga acgatgaaga ataaaaagag 480

agatccactg aggtgaaatt tcagctgaca gcgagtttca tgatcgtgat gaacaatggt 540

aacgagttgt ggctgttgcc agggagggtg gttctcaact tttaatgtat ggccaaatcg 600

ctacttgggt ttgttatata acaaagaaga aataatgaac tgattctctt cctccttctt 660

gtcctttctt aattctgttg taattacctt cctttgtaat ttttttgta attattcttc 720

ttaataatcc aaacaaacac acatattaca ata 753

<210> 811
<211> 1047
<212> DNA
<213> Artificial Sequence

<220>
<223> Achromobacter xylosoxidans sadB

<400> 811
atgaaagctc tggtttatca cggtgaccac aagatctcgc ttgaagacaa gcccaagccc 60

acccttcaaa agcccacgga tgtagtagta cgggttttga agaccacgat ctgcggcacg 120

gatctcggca tctacaaagg caagaatcca gaggtcgccg acgggcgcat cctgggccat 180

gaaggggtag gcgtcatcga ggaagtgggc gagagtgtca cgcagttcaa gaaaggcgac 240

aaggtcctga tttcctgcgt cacttcttgc ggctcgtgcg actactgcaa gaagcagctt 300

tactcccatt gccgcgacgg cgggtggatc ctgggttaca tgatcgatgg cgtgcaggcc 360

gaatacgtcc gcatcccgca tgccgacaac agcctctaca agatccccca gacaattgac 420

gacgaaatcg ccgtcctgct gagcgacatc ctgcccaccg gccacgaaat cggcgtccag 480

tatgggaatg tccagccggg cgatgcggtg gctattgtcg gcgcgggccc cgtcggcatg 540

tccgtactgt tgaccgccca gttctactcc ccctcgacca tcatcgtgat cgacatggac 600

gagaatcgcc tccagctcgc caaggagctc ggggcaacgc acaccatcaa ctccggcacg 660

gagaacgttg tcgaagccgt gcataggatt gcggcagagg gagtcgatgt tgcgatcgag 720

gcggtgggca taccggcgac ttgggacatc tgccaggaga tcgtcaagcc cggcgcgcac 780

atcgccaacg tcggcgtgca tggcgtcaag gttgacttcg agattcagaa gctctggatc 840

aagaacctga cgatcaccac gggactggtg aacacgaaca cgacgcccat gctgatgaag 900

gtcgcctcga ccgacaagct tccgttgaag aagatgatta cccatcgctt cgagctggcc 960

gagatcgagc acgcctatca ggtattcctc aatggcgcca aggagaaggc gatgaagatc 1020

atcctctcga acgcaggcgc tgcctga 1047

<210> 812
<211> 348
<212> PRT
<213> Artificial Sequence

<220>
<223> Achromobacter xylosoxidans sadB

<400> 812

Met Lys Ala Leu Val Tyr His Gly Asp His Lys Ile Ser Leu Glu Asp
1 5 10 15

Lys Pro Lys Pro Thr Leu Gln Lys Pro Thr Asp Val Val Val Arg Val
20 25 30

Leu Lys Thr Thr Ile Cys Gly Thr Asp Leu Gly Ile Tyr Lys Gly Lys
35 40 45

Asn Pro Glu Val Ala Asp Gly Arg Ile Leu Gly His Glu Gly Val Gly
50 55 60

Val Ile Glu Glu Val Gly Glu Ser Val Thr Gln Phe Lys Lys Gly Asp
65 70 75 80

Lys Val Leu Ile Ser Cys Val Thr Ser Cys Gly Ser Cys Asp Tyr Cys
85 90 95

Lys Lys Gln Leu Tyr Ser His Cys Arg Asp Gly Gly Trp Ile Leu Gly
100 105 110

Tyr Met Ile Asp Gly Val Gln Ala Glu Tyr Val Arg Ile Pro His Ala
115 120 125

Asp Asn Ser Leu Tyr Lys Ile Pro Gln Thr Ile Asp Asp Glu Ile Ala
130 135 140

Val Leu Leu Ser Asp Ile Leu Pro Thr Gly His Glu Ile Gly Val Gln
145 150 155 160

```
Tyr Gly Asn Val Gln Pro Gly Asp Ala Val Ala Ile Val Gly Ala Gly
                165                 170                 175

Pro Val Gly Met Ser Val Leu Leu Thr Ala Gln Phe Tyr Ser Pro Ser
            180                 185                 190

Thr Ile Ile Val Ile Asp Met Asp Glu Asn Arg Leu Gln Leu Ala Lys
        195                 200                 205

Glu Leu Gly Ala Thr His Thr Ile Asn Ser Gly Thr Glu Asn Val Val
    210                 215                 220

Glu Ala Val His Arg Ile Ala Ala Glu Gly Val Asp Val Ala Ile Glu
225                 230                 235                 240

Ala Val Gly Ile Pro Ala Thr Trp Asp Ile Cys Gln Glu Ile Val Lys
                245                 250                 255

Pro Gly Ala His Ile Ala Asn Val Gly Val His Gly Val Lys Val Asp
            260                 265                 270

Phe Glu Ile Gln Lys Leu Trp Ile Lys Asn Leu Thr Ile Thr Thr Gly
        275                 280                 285

Leu Val Asn Thr Asn Thr Thr Pro Met Leu Met Lys Val Ala Ser Thr
    290                 295                 300

Asp Lys Leu Pro Leu Lys Lys Met Ile Thr His Arg Phe Glu Leu Ala
305                 310                 315                 320

Glu Ile Glu His Ala Tyr Gln Val Phe Leu Asn Gly Ala Lys Glu Lys
                325                 330                 335

Ala Met Lys Ile Ile Leu Ser Asn Ala Gly Ala Ala
            340                 345
```

<210> 813
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer OT1074

<400> 813
cacacatatt acaatagcta gctgaggatg aaagctctg 39

<210> 814
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer OT1075

<400> 814
cagagctttc atcctcagct agctattgta atatgtgtg 39

<210> 815
<211> 9075
<212> DNA
<213> Artificial Sequence

<220>
<223> pLH435

<400> 815
ctagttctag agcggccgcc accgcggtgg agctccagct tttgttccct ttagtgaggg 60

ttaattgcgc gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg 120

ctcacaattc cacacaacat aggagccgga agcataaagt gtaaagcctg gggtgcctaa 180

tgagtgaggt aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac 240

ctgtcgtgcc agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt 300

gggcgctctt ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga 360

gcggtatcag ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca 420

ggaaagaaca tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg 480

ctggcgtttt tccataggct ccgcccccct gacgagcatc acaaaaatcg acgctcaagt 540

cagaggtggc gaaacccgac aggactataa agataccagg cgtttccccc tggaagctcc 600

ctcgtgcgct ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct 660

tcgggaagcg tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc 720

gttcgctcca agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta 780

tccggtaact atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca 840

gccactggta acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag 900

tggtggccta actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag 960

ccagttacct tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt 1020

agcggtggtt tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa 1080

gatcctttga tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg 1140

attttggtca tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga 1200

agttttaaat caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta 1260

atcagtgagg cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactc 1320

cccgtcgtgt agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg 1380

ataccgcgag acccacgctc accggctcca gatttatcag caataaacca gccagccgga 1440

agggccgagc gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt 1500

tgccgggaag ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt 1560

gctacaggca tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc 1620

caacgatcaa ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc 1680

ggtcctccga tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca 1740

gcactgcata attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag 1800

tactcaacca agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg 1860

tcaatacggg ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa 1920
cgttcttcgg ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa 1980
cccactcgtg cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga 2040
gcaaaaacag gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga 2100
atactcatac tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg 2160
agcggataca tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt 2220
ccccgaaaag tgccacctga acgaagcatc tgtgcttcat tttgtagaac aaaaatgcaa 2280
cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg 2340
caacgcgaaa gcgctatttt accaacgaag aatctgtgct tcatttttgt aaaacaaaaa 2400
tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt tacagaacag 2460
aaatgcaacg cgagagcgct attttaccaa caaagaatct atacttcttt tttgttctac 2520
aaaaatgcat cccgagagcg ctatttttct aacaaagcat cttagattac ttttttttctc 2580
ctttgtgcgc tctataatgc agtctcttga taacttttg cactgtaggt ccgttaaggt 2640
tagaagaagg ctactttggt gtctattttc tcttccataa aaaaagcctg actccacttc 2700
ccgcgtttac tgattactag cgaagctgcg ggtgcatttt ttcaagataa aggcatcccc 2760
gattatattc tataccgatg tggattgcgc atactttgtg aacagaaagt gatagcgttg 2820
atgattcttc attggtcaga aaattatgaa cggtttcttc tattttgtct ctatatacta 2880
cgtataggaa atgtttacat tttcgtattg ttttcgattc actctatgaa tagttcttac 2940
tacaattttt ttgtctaaag agtaatacta gagataaaca taaaaaatgt agaggtcgag 3000
tttagatgca agttcaagga gcgaaaggtg gatgggtagg ttatataggg atatagcaca 3060
gagatatata gcaaagagat actttgagc aatgtttgtg gaagcggtat tcgcaatatt 3120
ttagtagctc gttacagtcc ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc 3180
ttttggtttt caaaagcgct ctgaagttcc tatactttct agagaatagg aacttcggaa 3240
taggaacttc aaagcgtttc cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca 3300
catacagctc actgttcacg tcgcacctat atctgcgtgt tgcctgtata tatatataca 3360
tgagaagaac ggcatagtgc gtgtttatgc ttaaatgcgt acttatatgc gtctatttat 3420
gtaggatgaa aggtagtcta gtacctcctg tgatattatc ccattccatg cggggtatcg 3480
tatgcttcct tcagcactac cctttagctg ttctatatgc tgccactcct caattggatt 3540
agtctcatcc ttcaatgcta tcatttcctt tgatattgga tcatactaag aaaccattat 3600
tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt 3660
cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct 3720
gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg 3780
tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatcga 3840
ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc accattatgg 3900
gaaatgcttc aagaaggtat tgacttaaac tccatcaaat ggtcaggtca ttgagtgttt 3960
tttatttgtt gtatttttttt ttttttagag aaaatcctcc aatatcaaat taggaatcgt 4020

agtttcatga ttttctgtta cacctaactt tttgtgtggt gccctcctcc ttgtcaatat 4080
taatgttaaa gtgcaattct ttttccttat cacgttgagc cattagtatc aatttgctta 4140
cctgtattcc tttactatcc tccttttttct ccttcttgat aaatgtatgt agattgcgta 4200
tatagtttcg tctaccctat gaacatattc cattttgtaa tttcgtgtcg tttctattat 4260
gaatttcatt tataaagttt atgtacaaat atcataaaaa aagagaatct ttttaagcaa 4320
ggattttctt aacttcttcg gcgacagcat caccgacttc ggtggtactg ttggaaccac 4380
ctaaatcacc agttctgata cctgcatcca aaacctttttt aactgcatct tcaatggcct 4440
taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac aagatagtgg 4500
cgatagggtc aaccttattc tttggcaaat ctggagcaga accgtggcat ggttcgtaca 4560
aaccaaatgc ggtgttcttg tctggcaaag aggccaagga cgcagatggc aacaaaccca 4620
aggaacctgg gataacggag gcttcatcgg agatgatatc accaaacatg ttgctggtga 4680
ttataatacc atttaggtgg gttgggttct taactaggat catggcggca gaatcaatca 4740
attgatgttg aaccttcaat gtagggaatt cgttcttgat ggtttcctcc acagtttttc 4800
tccataatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata ggcaatggtg 4860
gctcatgttg tagggccatg aaagcggcca ttcttgtgat tctttgcact tctggaacgg 4920
tgtattgttc actatcccaa gcgacaccat caccatcgtc ttcctttctc ttaccaaagt 4980
aaatacctcc cactaattct ctgacaacaa cgaagtcagt acctttagca aattgtggct 5040
tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatggtctt aagttggcgt 5100
acaattgaag ttctttacgg atttttagta aaccttgttc aggtctaaca ctaccggtac 5160
cccatttagg accagccaca gcacctaaca aaacggcatc aaccttcttg gaggcttcca 5220
gcgcctcatc tggaagtggg acacctgtag catcgatagc agcaccacca attaaatgat 5280
tttcgaaatc gaacttgaca ttggaacgaa catcagaaat agctttaaga accttaatgg 5340
cttcggctgt gatttcttga ccaacgtggt cacctggcaa aacgacgatc ttcttagggg 5400
cagacatagg ggcagacatt agaatggtat atccttgaaa tatatatata tattgctgaa 5460
atgtaaaagg taagaaaagt tagaaagtaa gacgattgct aaccacctat tggaaaaaac 5520
aataggtcct taaataatat tgtcaacttc aagtattgtg atgcaagcat ttagtcatga 5580
acgcttctct attctatatg aaaagccggt tccggcctct cacctttcct ttttctccca 5640
atttttcagt tgaaaaaggt atatgcgtca ggcgacctct gaaattaaca aaaaatttcc 5700
agtcatcgaa tttgattctg tgcgatagcg cccctgtgtg ttctcgttat gttgaggaaa 5760
aaaataatgg ttgctaagag attcgaactc ttgcatctta cgatacctga gtattcccac 5820
agttaactgc ggtcaagata tttcttgaat caggcgcctt agaccgctcg gccaaacaac 5880
caattacttg ttgagaaata gagtataatt atcctataaa tataacgttt ttgaacacac 5940
atgaacaagg aagtacagga caattgattt tgaagagaat gtggattttg atgtaattgt 6000
tgggattcca tttttaataa ggcaataata ttaggtatgt ggatatacta gaagttctcc 6060
tcgaccgtcg atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 6120
aggaaattgt aaacgttaat attttgttaa aattcgcgtt aaatttttgt taaatcagct 6180

catttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa gaatagaccg 6240
agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag aacgtggact 6300
ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt gaaccatcac 6360
cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac cctaaaggga 6420
gcccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag gaagggaaga 6480
aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg cgcgtaacca 6540
ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc gcgccattcg ccattcaggc 6600
tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc gctattacgc cagctggcga 6660
aaggggatg tgctgcaagg cgattaagtt gggtaacgcc agggttttcc cagtcacgac 6720
gttgtaaaac gacggccagt gagcgcgcgt aatacgactc actatagggc gaattgggta 6780
ccgggcccccc cctcgaggtc gacggtatcg ataagcttga tatcgaattc ctgcagcccg 6840
ggggatccgc atgcttgcat ttagtcgtgc aatgtatgac tttaagattt gtgagcagga 6900
agaaaaggga gaatcttcta acgataaacc cttgaaaaac tgggtagact acgctatgtt 6960
gagttgctac gcaggctgca caattacacg agaatgctcc cgcctaggat ttaaggctaa 7020
gggacgtgca atgcagacga cagatctaaa tgaccgtgtc ggtgaagtgt tcgccaaact 7080
tttcggttaa cacatgcagt gatgcacgcg cgatggtgct aagttacata tatatatata 7140
tatatatata tagccatagt gatgtctaag taacctttat ggtatatttc ttaatgtgga 7200
aagatactag cgcgcgcacc cacacacaag cttcgtcttt tcttgaagaa aagaggaagc 7260
tcgctaaatg ggattccact ttccgttccc tgccagctga tggaaaaagg ttagtggaac 7320
gatgaagaat aaaaagagag atccactgag gtgaaatttc agctgacagc gagtttcatg 7380
atcgtgatga acaatggtaa cgagttgtgg ctgttgccag ggagggtggt tctcaacttt 7440
taatgtatgg ccaaatcgct acttgggttt gttatataac aaagaagaaa taatgaactg 7500
attctcttcc tccttcttgt cctttcttaa ttctgttgta attaccttcc tttgtaattt 7560
tttttgtaat tattcttctt aataatccaa acaaacacac atattacaat agctagctga 7620
ggatgtcaac agccggtaaa gttattaagt gtaaagcggc agttttgtgg gaagagaaaa 7680
agccgtttag catagaagaa gtagaagtag cgccaccaaa agcacacgag gttagaatca 7740
agatggttgc caccggaatc tgtagatccg acgaccatgt ggtgagtggc actctagtta 7800
ctcctttgcc agtaatcgcg ggacacgagg ctgccggaat cgttgaatcc ataggtgaag 7860
gtgttaccac tgttcgtcct ggtgataaag tgatcccact gttcactcct caatgtggta 7920
agtgtagagt ctgcaaacat cctgagggta atttctgcct taaaaatgat ttgtctatgc 7980
ctagaggtac tatgcaggat ggtacaagca gatttacatg cagagggaaa cctatacacc 8040
atttccttgg tacttctaca ttttcccaat acacagtggt ggacgagata tctgtcgcta 8100
aaatcgatgc agcttcacca ctggaaaaag tttgcttgat agggtgcgga ttttccaccg 8160
gttacggttc cgcagttaaa gttgcaaagg ttacacaggg ttcgacttgt gcagtattcg 8220
gtttaggagg agtaggacta agcgttatta tggggtgtaa agctgcaggc gcagcgagga 8280
ttataggtgt agacatcaat aaggacaaat ttgcaaaagc taaggaggtc ggggctactg 8340

aatgtgttaa ccctcaagat tataagaaac caatacaaga agtccttact gaaatgtcaa 8400

acggtggagt tgatttctct tttgaagtta taggccgtct tgatactatg gtaactgcgt 8460

tgtcctgctg tcaagaggca tatggagtca gtgtgatcgt aggtgttcct cctgattcac 8520

aaaatttgtc gatgaatcct atgctgttgc taagcggtcg tacatggaag ggagctatat 8580

ttggcggttt taagagcaag gatagtgttc caaaacttgt tgccgacttt atggcgaaga 8640

agtttgctct tgatccttta attacacatg tattgccatt cgagaaaatc aatgaagggt 8700

ttgatttgtt aagaagtggt gaatctattc gtacaatttt aacttttga ttaattaaga 8760

gtaagcgaat ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag 8820

tgtatacaaa ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa 8880

ctctttcctg taggtcaggt tgctttctca ggtatagcat gaggtcgctc ttattgacca 8940

cacctctacc ggcatgccga gcaaatgcct gcaaatcgct ccccatttca cccaattgta 9000

gatatgctaa ctccagcaat gagttgatga atctcggtgt gtattttatg tcctcagagg 9060

acaacacctg tggta 9075

<210> 816
<211> 11367
<212> DNA
<213> Artificial Sequence

<220>
<223> pLH441

<400> 816
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatagcca tcctcatgaa aactgtgtaa cataataacc gaagtgtcga aaaggtggca 240

ccttgtccaa ttgaacacgc tcgatgaaaa aaataagata tatataaggt taagtaaagc 300

gtctgttaga aaggaagttt ttccttttc ttgctctctt gtcttttcat ctactatttc 360

cttcgtgtaa tacagggtcg tcagatacat agatacaatt ctattacccc catccataca 420

atgccatctc atttcgatac tgttcaacta cacgccggcc aagagaaccc tggtgacaat 480

gctcacagat ccagagctgt accaatttac gccaccactt cttatgtttt cgaaaactct 540

aagcatggtt cgcaattgtt tggtctagaa gttccaggtt acgtctattc ccgtttccaa 600

aacccaacca gtaatgtttt ggaagaaaga attgctgctt tagaaggtgg tgctgctgct 660

ttggctgttt cctccggtca agccgctcaa acccttgcca tccaaggttt ggcacacact 720

ggtgacaaca tcgtttccac ttcttactta tacggtggta cttataacca gttcaaaatc 780

tcgttcaaaa gatttggtat cgaggctaga tttgttgaag gtgacaatcc agaagaattc 840

gaaaaggtct ttgatgaaag aaccaaggct gtttatttgg aaaccattgg taatccaaag 900

tacaatgttc cggattttga aaaaattgtt gcaattgctc acaaacacgg tattccagtt 960

gtcgttgaca acacatttgg tgccggtggt tacttctgtc agccaattaa atacggtgct 1020

gatattgtaa cacattctgc taccaaatgg attggtggtc atggtactac tatcggtggt 1080

attattgttg actctggtaa gttcccatgg aaggactacc cagaaaagtt ccctcaattc 1140
tctcaacctg ccgaaggata tcacggtact atctacaatg aagcctacgg taacttggca 1200
tacatcgttc atgttagaac tgaactatta agagatttgg gtccattgat gaacccattt 1260
gcctctttct tgctactaca aggtgttgaa acattatctt tgagagctga aagacacggt 1320
gaaaatgcat tgaagttagc caaatggtta gaacaatccc catacgtatc ttgggtttca 1380
taccctggtt tagcatctca ttctcatcat gaaaatgcta agaagtatct atctaacggt 1440
ttcggtggtg tcttatcttt cggtgtaaaa gacttaccaa atgccgacaa ggaaactgac 1500
ccattcaaac tttctggtgc tcaagttgtt gacaatttaa agcttgcctc taacttggcc 1560
aatgttggtg atgccaagac cttagtcatt gctccatact tcactaccca caaacaatta 1620
aatgacaaag aaaagttggc atctggtgtt accaaggact taattcgtgt ctctgttggt 1680
atcgaattta ttgatgacat tattgcagac ttccagcaat cttttgaaac tgttttcgct 1740
ggccaaaaac catgagtgtg cgtaatgagt tgtaaaatta tgtataaacc tactttctct 1800
cacaagttat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac cgcatcagga 1860
aattgtaaac gttaatattt tgttaaaatt cgcgttaaat ttttgttaaa tcagctcatt 1920
ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat agaccgagat 1980
agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg tggactccaa 2040
cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac catcacccta 2100
atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaacccta aagggagccc 2160
ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag ggaagaaagc 2220
gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg taaccaccac 2280
acccgccgcg cttaatgcgc cgctacaggg cgcgtcgcgc cattcgccat tcaggctgcg 2340
caactgttgg gaagggcgat cggtgcgggc ctcttcgcta ttacgccagc tggcgaaagg 2400
gggatgtgct gcaaggcgat taagttgggt aacgccaggg ttttcccagt cacgacgttg 2460
taaaacgacg gccagtgagc gcgcgtaata cgactcacta tagggcgaat tgggtaccgg 2520
gccccccctc gaggtcgacg gtatcgataa gcttgatatc gaattcctgc agcccggggg 2580
atccactagt tctagagcgg ccgctctaga actagtacca caggtgttgt cctctgagga 2640
cataaaatac acaccgagat tcatcaactc attgctggag ttagcatatc tacaattggg 2700
tgaaatgggg agcgatttgc aggcatttgc tcggcatgcc ggtagaggtg tggtcaataa 2760
gagcgacctc atgctatacc tgagaaagca acctgaccta caggaaagag ttactcaaga 2820
ataagaattt tcgttttaaa acctaagagt cactttaaaa tttgtataca cttatttttt 2880
ttataactta tttaataata aaaatcataa atcataagaa attcgcttac tcttaattaa 2940
tcaaaaagtt aaaattgtac gaatagattc accacttctt aacaaatcaa acccttcatt 3000
gattttctcg aatggcaata catgtgtaat taaaggatca agagcaaact tcttcgccat 3060
aaagtcggca acaagttttg gaacactatc cttgctctta aaaccgccaa atatagctcc 3120
cttccatgta cgaccgctta gcaacagcat aggattcatc gacaaatttt gtgaatcagg 3180
aggaacacct acgatcacac tgactccata tgcctcttga cagcaggaca acgcagttac 3240

catagtatca agacggccta taacttcaaa agagaaatca actccaccgt ttgacatttc 3300
agtaaggact tcttgtattg gtttcttata atcttgaggg ttaacacatt cagtagcccc 3360
gacctcctta gcttttgcaa atttgtcctt attgatgtct acacctataa tcctcgctgc 3420
gcctgcagct ttacacccca taataacgct tagtcctact cctcctaaac cgaatactgc 3480
acaagtcgaa ccctgtgtaa cctttgcaac tttaactgcg gaaccgtaac cggtggaaaa 3540
tccgcaccct atcaagcaaa ctttttccag tggtgaagct gcatcgattt tagcgacaga 3600
tatctcgtcc accactgtgt attgggaaaa tgtagaagta ccaaggaaat ggtgtatagg 3660
tttccctctg catgtaaatc tgcttgtacc atcctgcata gtacctctag gcatagacaa 3720
atcattttta aggcagaaat taccctcagg atgtttgcag actctacact taccacattg 3780
aggagtgaac agtgggatca ctttatcacc aggacgaaca gtggtaacac cttcacctat 3840
ggattcaacg attccggcag cctcgtgtcc cgcgattact ggcaaaggag taactagagt 3900
gccactcacc acatggtcgt cggatctaca gattccggtg gcaaccatct tgattctaac 3960
ctcgtgtgct tttggtggcg ctacttctac ttcttctatg ctaaacggct ttttctcttc 4020
ccacaaaact gccgctttac acttaataac tttaccggct gttgacatcc tcagctagct 4080
attgtaatat gtgtgtttgt ttggattatt aagaagaata attacaaaaa aaattacaaa 4140
ggaaggtaat tacaacagaa ttaagaaagg acaagaagga ggaagagaat cagttcatta 4200
tttcttcttt gttatataac aaacccaagt agcgatttgg ccatacatta aaagttgaga 4260
accaccctcc ctggcaacag ccacaactcg ttaccattgt tcatcacgat catgaaactc 4320
gctgtcagct gaaatttcac ctcagtggat ctctcttttt attcttcatc gttccactaa 4380
ccttttcca tcagctggca gggaacggaa agtggaatcc catttagcga gcttcctctt 4440
ttcttcaaga aaagacgaag cttgtgtgtg ggtgcgcgcg ctagtatctt tccacattaa 4500
gaaatatacc ataaaggtta cttagacatc actatggcta tatatatata tatatatata 4560
tatgtaactt agcaccatcg cgcgtgcatc actgcatgtg ttaaccgaaa agtttggcga 4620
acacttcacc gacacggtca tttagatctg tcgtctgcat tgcacgtccc ttagccttaa 4680
atcctaggcg ggagcattct cgtgtaattg tgcagcctgc gtagcaactc aacatagcgt 4740
agtctaccca gtttttcaag ggtttatcgt tagaagattc tccctttttct tcctgctcac 4800
aaatcttaaa gtcatacatt gcacgactaa atgcaagcat gcggatcccc cgggctgcag 4860
gaattcgata tcaagcttat cgataccgtc gactggccat taatctttcc catattagat 4920
ttcgccaagc catgaaagtt caagaaaggt ctttagacga attacccttc atttctcaaa 4980
ctggcgtcaa gggatcctgg tatggtttta tcgttttatt tctggttctt atagcatcgt 5040
tttggacttc tctgttccca ttaggcggtt caggagccag cgcagaatca ttctttgaag 5100
gatacttatc ctttccaatt ttgattgtct gttacgttgg acataaactg tatactagaa 5160
attggacttt gatggtgaaa ctagaagata tggatcttga taccggcaga aaacaagtag 5220
atttgactct tcgtagggaa gaaatgagga ttgagcgaga aacattagca aaaagatcct 5280
tcgtaacaag atttttacat ttctggtgtt gaagggaaag atatgagcta tacagcggaa 5340
tttccatatc actcagattt tgttatctaa ttttttcctt cccacgtccg cgggaatctg 5400

tgtatattac tgcatctaga tatatgttat cttatcttgg cgcgtacatt taattttcaa 5460
cgtattctat aagaaattgc gggagttttt ttcatgtaga tgatactgac tgcacgcaaa 5520
tataggcatg atttataggc atgatttgat ggctgtaccg ataggaacgc taagagtaac 5580
ttcagaatcg ttatcctggc ggaaaaaatt catttgtaaa ctttaaaaaa aaaagccaat 5640
atccccaaaa ttattaagag cgcctccatt attaactaaa atttcactca gcatccacaa 5700
tgtatcaggt atctactaca gatattacat gtggcgaaaa agacaagaac aatgcaatag 5760
cgcatcaaga aaaaacacaa agctttcaat caatgaatcg aaaatgtcat taaaatagta 5820
tataaattga aactaagtca taaagctata aaaagaaaat ttatttaaat gcaagattta 5880
aagtaaattc acggccctgc aggcctcagc tcttgttttg ttctgcaaat aacttaccca 5940
tctttttcaa aactttaggt gcaccctcct ttgctagaat aagttctatc caatacatcc 6000
tatttggatc tgcttgagct tctttcatca cggatacgaa ttcattttct gttctcacaa 6060
ttttggacac aactctgtct tccgttgccc cgaaactttc tggcagtttt gagtaattcc 6120
acataggaat gtcattataa ctctggttcg gaccatgaat ttccctctca accgtgtaac 6180
catcgttatt aatgataaag cagattgggt ttatcttctc tctaatggct agtcctaatt 6240
cttggacagt cagttgcaat gatccatctc cgataaacaa taaatgtcta gattctttat 6300
ctgcaatttg gctgcctaga gctgcgggga aagtgtatcc tatagatccc cacaagggtt 6360
gaccaataaa atgtgatttc gatttcagaa atatagatga ggcaccgaag aaagaagtgc 6420
cttgttcagc cacgatcgtc tcattacttt gggtcaaatt ttcgacagct tgccacagtc 6480
tatcttgtga caacagcgcg ttagaaggta caaaatcttc ttgcttttta tctatgtact 6540
tgcctttata ttcaatttcg gacaagtcaa gaagagatga tatcagggat tcgaagtcga 6600
aattttggat tctttcgttg aaaattttac cttcatcgat attcaaggaa atcattttat 6660
tttcattaag atggtgagta aatgcacccg tactagaatc ggtaagcttt acacccaaca 6720
taagaataaa atcagcagat tccacaaatt ccttcaagtt tggctctgac agagtaccgt 6780
tgtaaatccc caaaaatgag ggcaatgctt catcaacaga tgatttacca aagttcaaag 6840
tagtaatagg taacttagtc tttgaaataa actgagtaac agtcttctct aggccgaacg 6900
atataatttc atggcctgtg attacaattg gtttcttggc attcttcaga ctttcctgta 6960
ttttgttcag aatctcttga tcagatgtat tcgacgtgga attttccttc ttaagaggca 7020
aggatggttt ttcagcctta gcggcagcta catctacagg taaattgatg taaaccggct 7080
ttctttcctt tagtaaggca gacaacactc tatcaatttc aacagttgca ttctcggctg 7140
tcaataaagt cctggcagca gtaaccggtt cgtgcatctt cataaagtgc ttgaaatcac 7200
catcagccaa cgtatggtga acaaacttac cttcgttctg cactttcgag gtaggagatc 7260
ccacgatctc aacaacaggc aggttctcag cataggagcc cgctaagcca ttaactgcgg 7320
ataattcgcc aacaccaaat gtagtcaaga atgccgcagc ctttttcgtt cttgcgtacc 7380
cgtcggccat ataggaggca tttaactcat tagcatttcc cacccatttc atatctttgt 7440
gtgaaataat ttgatctaga aattgcaaat tgtagtcacc tggtactccg aatatttctt 7500
ctatacctaa ttcgtgtaat ctgtccaaca gatagtcacc tactgtatac attttgttta 7560

ctagtttatg tgtgtttatt cgaaactaag ttcttggtgt tttaaaacta aaaaaaagac 7620
taactataaa agtagaattt aagaagttta agaaatagat ttacagaatt acaatcaata 7680
cctaccgtct ttatatactt attagtcaag taggggaata atttcaggga actggtttca 7740
acctttttt tcagcttttt ccaaatcaga gagagcagaa ggtaatagaa ggtgtaagaa 7800
aatgagatag atacatgcgt gggtcaattg ccttgtgtca tcatttactc caggcaggtt 7860
gcatcactcc attgaggttg tgcccgtttt ttgcctgttt gtgcccctgt tctctgtagt 7920
tgcgctaaga gaatggacct atgaactgat ggttggtgaa gaaaacaata ttttggtgct 7980
gggattcttt ttttttctgg atgccagctt aaaaagcggg ctccattata tttagtggat 8040
gccaggaata aactgttcac ccagacacct acgatgttat atattctgtg taacccgccc 8100
cctattttgg gcatgtacgg gttacagcag aattaaaagg ctaatttttt gactaaataa 8160
agttaggaaa atcactacta ttaattattt acgtattctt tgaaatggca gtattgataa 8220
tgataaactc gaactgaaaa agcgtgtttt ttattcaaaa tgattctaac tcccttacgt 8280
aatcaaggaa tctttttgcc ttggcctccg cgtcattaaa cttcttgttg ttgacgctaa 8340
cattcaacgc tagtatatat tcgtttttt caggtaagtt cttttcaacg ggtcttactg 8400
atgaggcagt cgcgtctgaa cctgttaaga ggtcaaatat gtcttcttga ccgtacgtgt 8460
cttgcatgtt attagctttg ggaatttgca tcaagtcata ggaaaattta aatcttggct 8520
ctcttgggct caaggtgaca aggtcctcga aaatagggcg cgccccaccg cggtggagct 8580
cagcttttgt tccctttagt gagggttaat tgcgcgcttg gcgtaatcat ggtcatagct 8640
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 8700
aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgctc 8760
actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 8820
cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct 8880
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 8940
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9000
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 9060
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9120
ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 9180
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 9240
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 9300
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 9360
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 9420
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt 9480
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 9540
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 9600
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 9660
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 9720

ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 9780
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 9840
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 9900
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 9960
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 10020
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 10080
tagtttgcgc aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg 10140
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 10200
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 10260
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 10320
aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 10380
gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 10440
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 10500
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 10560
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 10620
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 10680
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 10740
acaaataggg gttccgcgca catttccccg aaaagtgcca cctgggtcct tttcatcacg 10800
tgctataaaa ataattataa tttaaatttt ttaatataaa tatataaatt aaaaatagaa 10860
agtaaaaaaa gaaattaaag aaaaaatagt ttttgttttc cgaagatgta aaagactcta 10920
gggggatcgc caacaaatac tacctttttat cttgctcttc ctgctctcag gtattaatgc 10980
cgaattgttt catcttgtct gtgtagaaga ccacacacga aaatcctgtg attttacatt 11040
ttacttatcg ttaatcgaat gtatatctat ttaatctgct tttcttgtct aataaatata 11100
tatgtaaagt acgctttttg ttgaaatttt ttaaaccttt gtttattttt ttttcttcat 11160
tccgtaactc ttctaccttc tttatttact ttctaaaatc caaatacaaa acataaaaat 11220
aaataaacac agagtaaatt cccaaattat tccatcatta aaagatacga ggcgcgtgta 11280
agttacaggc aagcgatccg tcctaagaaa ccattattat catgacatta acctataaaa 11340
ataggcgtat cacgaggccc tttcgtc 11367

<210> 817
<211> 9491
<212> DNA
<213> Artificial Sequence

<220>
<223> Synthetic construct pRS423 FBA ilvD(Strep)

<400> 817
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataaatt cccgttttaa gagcttggtg agcgctagga gtcactgcca ggtatcgttt 240
gaacacggca ttagtcaggg aagtcataac acagtccttt cccgcaattt tctttttcta 300
ttactcttgg cctcctctag tacactctat attttttttat gcctcggtaa tgattttcat 360
ttttttttttt cccctagcgg atgactcttt ttttttctta gcgattggca ttatcacata 420
atgaattata cattatataa agtaatgtga tttcttcgaa gaatatacta aaaaatgagc 480
aggcaagata aacgaaggca aagatgacag agcagaaagc cctagtaaag cgtattacaa 540
atgaaaccaa gattcagatt gcgatctctt taaagggtgg tcccctagcg atagagcact 600
cgatcttccc agaaaaagag gcagaagcag tagcagaaca ggccacacaa tcgcaagtga 660
ttaacgtcca cacaggtata gggtttctgg accatatgat acatgctctg gccaagcatt 720
ccggctggtc gctaatcgtt gagtgcattg gtgacttaca catagacgac catcacacca 780
ctgaagactg cgggattgct ctcggtcaag cttttaaaga ggccctactg gcgcgtggag 840
taaaaaggtt tggatcagga tttgcgcctt tggatgaggc actttccaga gcggtggtag 900
atctttcgaa caggccgtac gcagttgtcg aacttggttt gcaaagggag aaagtaggag 960
atctctcttg cgagatgatc ccgcattttc ttgaaagctt tgcagaggct agcagaatta 1020
ccctccacgt tgattgtctg cgaggcaaga atgatcatca ccgtagtgag agtgcgttca 1080
aggctcttgc ggttgccata agagaagcca cctcgcccaa tggtaccaac gatgttccct 1140
ccaccaaagg tgttcttatg tagtgacacc gattatttaa agctgcagca tacgatatat 1200
atacatgtgt atatatgtat acctatgaat gtcagtaagt atgtatacga acagtatgat 1260
actgaagatg acaaggtaat gcatcattct atacgtgtca ttctgaacga ggcgcgcttt 1320
ccttttttct ttttgctttt tctttttttt tctcttgaac tcgacggatc tatgcggtgt 1380
gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta aacgttaata 1440
ttttgttaaa attcgcgtta aatttttgtt aaatcagctc atttttaac caataggccg 1500
aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc 1560
cagtttggaa caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa 1620
ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt 1680
cgaggtgccg taaagcacta aatcggaacc ctaaagggag cccccgattt agagcttgac 1740
ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa agcgaaagga gcgggcgcta 1800
gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac cacacccgcc gcgcttaatg 1860
cgccgctaca gggcgcgtcg cgccattcgc cattcaggct gcgcaactgt tgggaagggc 1920
gatcggtgcg ggcctcttcg ctattacgcc agctggcgaa aggggatgt gctgcaaggc 1980
gattaagttg ggtaacgcca gggttttccc agtcacgacg ttgtaaaacg acggccagtg 2040
agcgcgcgta atacgactca ctatagggcg aattgggtac cgggcccccc ctcgaggtcg 2100
acggcgcgcc actggtagag agcgactttg tatgccccaa ttgcgaaacc cgcgatatcc 2160
ttctcgattc tttagtaccc gaccaggaca aggaaaagga ggtcgaaacg tttttgaaga 2220
aacaagagga actacacgga agctctaaag atggcaacca gccagaaact aagaaaatga 2280
agttgatgga tccaactggc accgctggct tgaacaacaa taccagcctt ccaacttctg 2340

taaataacgg cggtacgcca gtgccaccag taccgttacc tttcggtata cctcctttcc 2400
ccatgtttcc aatgcccttc atgcctccaa cggctactat cacaaatcct catcaagctg 2460
acgcaagccc taagaaatga ataacaatac tgacagtact aaataattgc ctacttggct 2520
tcacatacgt tgcatacgtc gatatagata ataatgataa tgacagcagg attatcgtaa 2580
tacgtaatag ttgaaaatct caaaaatgtg tgggtcatta cgtaaataat gataggaatg 2640
ggattcttct atttttcctt tttccattct agcagccgtc gggaaaacgt ggcatcctct 2700
ctttcgggct caattggagt cacgctgccg tgagcatcct ctctttccat atctaacaac 2760
tgagcacgta accaatggaa aagcatgagc ttagcgttgc tccaaaaaag tattggatgg 2820
ttaataccat ttgtctgttc tcttctgact ttgactcctc aaaaaaaaaa aatctacaat 2880
caacagatcg cttcaattac gccctcacaa aaacttttt ccttcttctt cgcccacgtt 2940
aaatttttatc cctcatgttg tctaacggat ttctgcactt gatttattat aaaaagacaa 3000
agacataata cttctctatc aatttcagtt attgttcttc cttgcgttat tcttctgttc 3060
ttctttttct tttgtcatat ataaccataa ccaagtaata catattcaaa ctagtatgac 3120
tgacaaaaaa actcttaaag acttaagaaa tcgtagttct gtttacgatt caatggttaa 3180
atcacctaat cgtgctatgt tgcgtgcaac tggtatgcaa gatgaagact ttgaaaaacc 3240
tatcgtcggt gtcatttcaa cttgggctga aaacacacct tgtaatatcc acttacatga 3300
ctttggtaaa ctagccaaag tcggtgttaa ggaagctggt gcttggccag ttcagttcgg 3360
aacaatcacg gtttctgatg gaatcgccat gggaacccaa ggaatgcgtt tctccttgac 3420
atctcgtgat attattgcag attctattga agcagccatg ggaggtcata atgcggatgc 3480
ttttgtagcc attggcggtt gtgataaaaa catgcccggt tctgttatcg ctatggctaa 3540
catggatatc ccagccattt ttgcttacgg cggaacaatt gcacctggta atttagacgg 3600
caaagatatc gatttagtct ctgtctttga aggtgtcggc cattggaacc acggcgatat 3660
gaccaaagaa gaagttaaag ctttggaatg taatgcttgt cccggtcctg gaggctgcgg 3720
tggtatgtat actgctaaca caatggcgac agctattgaa gttttgggac ttagccttcc 3780
gggttcatct tctcacccgg ctgaatccgc agaaaagaaa gcagatattg aagaagctgg 3840
tcgcgctgtt gtcaaaatgc tcgaaatggg cttaaaacct tctgacattt taacgcgtga 3900
agcttttgaa gatgctatta ctgtaactat ggctctggga ggttcaacca actcaaccct 3960
tcacctctta gctattgccc atgctgctaa tgtggaattg acacttgatg atttcaatac 4020
tttccaagaa aaagttcctc atttggctga tttgaaacct tctggtcaat atgtattcca 4080
agacctttac aaggtcggag gggtaccagc agttatgaaa tatctcctta aaaatggctt 4140
ccttcatggt gaccgtatca cttgtactgg caaaacagtc gctgaaaatt tgaaggcttt 4200
tgatgattta acacctggtc aaaaggttat tatgccgctt gaaaatccta aacgtgaaga 4260
tggtccgctc attattctcc atggtaactt ggctccagac ggtgccgttg ccaaagtttc 4320
tggtgtaaaa gtgcgtcgtc atgtcggtcc tgctaaggtc tttaattctg aagaagaagc 4380
cattgaagct gtcttgaatg atgatattgt tgatggtgat gttgttgtcg tacgttttgt 4440
aggaccaaag ggcggtcctg gtatgcctga aatgctttcc ctttcatcaa tgattgttgg 4500

taaagggcaa ggtgaaaaag ttgcccttct gacagatggc cgcttctcag gtggtactta 4560
tggtcttgtc gtgggtcata tcgctcctga agcacaagat ggcggtccaa tcgcctacct 4620
gcaaacagga gacatagtca ctattgacca agacactaag gaattacact ttgatatctc 4680
cgatgaagag ttaaaacatc gtcaagagac cattgaattg ccaccgctct attcacgcgg 4740
tatccttggt aaatatgctc acatcgtttc gtctgcttct aggggagccg taacagactt 4800
ttggaagcct gaagaaactg gcaaaaaatg ttgtcctggt tgctgtggtt aagcggccgc 4860
gttaattcaa attaattgat atagtttttt aatgagtatt gaatctgttt agaaataatg 4920
gaatattatt tttatttatt tatttatatt attggtcggc tcttttcttc tgaaggtcaa 4980
tgacaaaatg atatgaagga aataatgatt tctaaaattt tacaacgtaa gatattttta 5040
caaaagccta gctcatcttt tgtcatgcac tattttactc acgcttgaaa ttaacggcca 5100
gtccactgcg gagtcatttc aaagtcatcc taatcgatct atcgtttttg atagctcatt 5160
ttggagttcg cgattgtctt ctgttattca caactgtttt aatttttatt tcattctgga 5220
actcttcgag ttctttgtaa agtctttcat agtagcttac tttatcctcc aacatattta 5280
acttcatgtc aatttcggct cttaaatttt ccacatcatc aagttcaaca tcatctttta 5340
acttgaattt attctctagc tcttccaacc aagcctcatt gctccttgat ttactggtga 5400
aaagtgatac actttgcgcg caatccaggt caaaactttc ctgcaaagaa ttcaccaatt 5460
tctcgacatc atagtacaat ttgttttgtt ctcccatcac aatttaatat acctgatgga 5520
ttcttatgaa gcgctgggta atggacgtgt cactctactt cgccttttc cctactcctt 5580
ttagtacgga agacaatgct aataaataag agggtaataa taatattatt aatcggcaaa 5640
aaagattaaa cgccaagcgt ttaattatca gaaagcaaac gtcgtaccaa tccttgaatg 5700
cttcccaatt gtatattaag agtcatcaca gcaacatatt cttgttatta aattaattat 5760
tattgatttt tgatattgta taaaaaaacc aaatatgtat aaaaaaagtg aataaaaaat 5820
accaagtatg gagaaatata ttagaagtct atacgttaaa ccaccgcggt ggagctccag 5880
cttttgttcc ctttagtgag ggttaattgc gcgcttggcg taatcatggt catagctgtt 5940
tcctgtgtga aattgttatc cgctcacaat tccacacaac ataggagccg gaagcataaa 6000
gtgtaaagcc tggggtgcct aatgagtgag gtaactcaca ttaattgcgt tgcgctcact 6060
gcccgctttc agtcgggaa acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc 6120
ggggagaggc ggtttgcgta ttgggcgctc ttccgcttcc tcgctcactg actcgctgcg 6180
ctcggtcgtt cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc 6240
cacagaatca ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag 6300
gaaccgtaaa aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca 6360
tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca 6420
ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg 6480
atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag 6540
gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aacccccgt 6600
tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca 6660

cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg 6720
cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt 6780
tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc 6840
cggcaaacaa accaccgctg gtagcggtgg ttttttttgtt tgcaagcagc agattacgcg 6900
cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg 6960
gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta 7020
gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg 7080
gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg 7140
ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc 7200
atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc 7260
agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc 7320
ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag 7380
tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt cgtttggtat 7440
ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg 7500
caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt 7560
gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag 7620
atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg 7680
accgagttgc tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt 7740
aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct 7800
gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac 7860
tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat 7920
aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat 7980
ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca 8040
aatagggggtt ccgcgcacat ttccccgaaa agtgccacct gaacgaagca tctgtgcttc 8100
attttgtaga acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct 8160
gcatttttac agaacagaaa tgcaacgcga aagcgctatt ttaccaacga agaatctgtg 8220
cttcattttt gtaaaacaaa aatgcaacgc gagagcgcta atttttcaaa caaagaatct 8280
gagctgcatt tttacagaac agaaatgcaa cgcgagagcg ctattttacc aacaaagaat 8340
ctatacttct tttttgttct acaaaaatgc atcccgagag cgctattttt ctaacaaagc 8400
atcttagatt actttttttc tcctttgtgc gctctataat gcagtctctt gataactttt 8460
tgcactgtag gtccgttaag gttagaagaa ggctactttg gtgtctattt tctcttccat 8520
aaaaaaagcc tgactccact tcccgcgttt actgattact agcgaagctg cgggtgcatt 8580
ttttcaagat aaaggcatcc ccgattatat tctataccga tgtggattgc gcatactttg 8640
tgaacagaaa gtgatagcgt tgatgattct tcattggtca gaaaattatg aacggtttct 8700
tctattttgt ctctatatac tacgtatagg aaatgtttac attttcgtat tgttttcgat 8760
tcactctatg aatagttctt actacaattt ttttgtctaa agagtaatac tagagataaa 8820

cataaaaaat gtagaggtcg agtttagatg caagttcaag gagcgaaagg tggatgggta 8880

ggttatatag ggatatagca cagagatata tagcaaagag atacttttga gcaatgtttg 8940

tggaagcggt attcgcaata ttttagtagc tcgttacagt ccggtgcgtt tttggttttt 9000

tgaaagtgcg tcttcagagc gcttttggtt ttcaaaagcg ctctgaagtt cctatacttt 9060

ctagagaata ggaacttcgg aataggaact tcaaagcgtt tccgaaaacg agcgcttccg 9120

aaaatgcaac gcgagctgcg cacatacagc tcactgttca cgtcgcacct atatctgcgt 9180

gttgcctgta tatatatata catgagaaga acggcatagt gcgtgtttat gcttaaatgc 9240

gtacttatat gcgtctattt atgtaggatg aaaggtagtc tagtacctcc tgtgatatta 9300

tcccattcca tgcggggtat cgtatgcttc cttcagcact acccttttagc tgttctatat 9360

gctgccactc ctcaattgga ttagtctcat ccttcaatgc tatcatttcc tttgatattg 9420

gatcatctaa gaaaccatta ttatcatgac attaacctat aaaaataggc gtatcacgag 9480

gccctttcgt c 9491

<210> 818
<211> 1000
<212> DNA
<213> Artificial Sequence

<220>
<223> FBA1 terminator nt 4861 to 5860

<400> 818
gttaattcaa attaattgat atagtttttt aatgagtatt gaatctgttt agaaataatg 60

gaatattatt tttatttatt tatttatatt attggtcggc tcttttcttc tgaaggtcaa 120

tgacaaaatg atatgaagga aataatgatt tctaaaattt tacaacgtaa gatattttta 180

caaaagccta gctcatcttt tgtcatgcac tattttactc acgcttgaaa ttaacggcca 240

gtccactgcg gagtcatttc aaagtcatcc taatcgatct atcgtttttg atagctcatt 300

ttggagttcg cgattgtctt ctgttattca caactgtttt aatttttatt tcattctgga 360

actcttcgag ttctttgtaa agtctttcat agtagcttac tttatcctcc aacatattta 420

acttcatgtc aatttcggct cttaaatttt ccacatcatc aagttcaaca tcatctttta 480

acttgaattt attctctagc tcttccaacc aagcctcatt gctccttgat ttactggtga 540

aaagtgatac actttgcgcg caatccaggt caaaactttc ctgcaaagaa ttcaccaatt 600

tctcgacatc atagtacaat ttgttttgtt ctcccatcac aatttaatat acctgatgga 660

ttcttatgaa gcgctgggta atggacgtgt cactctactt cgccttttc cctactcctt 720

ttagtacgga agacaatgct aataaataag agggtaataa taatattatt aatcggcaaa 780

aaagattaaa cgccaagcgt ttaattatca gaaagcaaac gtcgtaccaa tccttgaatg 840

cttcccaatt gtatattaag agtcatcaca gcaacatatt cttgttatta aattaattat 900

tattgatttt tgatattgta taaaaaaacc aaatatgtat aaaaaaagtg aataaaaaat 960

accaagtatg gagaaatata ttagaagtct atacgttaaa 1000

<210> 819
<211> 1713

20172000439 20 Jan 2017

<212> DNA
<213> Artificial Sequence

<220>
<223> Streptococcus mutans ilvD coding region

<400> 819
atgactgaca aaaaaactct taaagactta agaaatcgta gttctgttta cgattcaatg 60

gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta tgcaagatga agactttgaa 120

aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca caccttgtaa tatccactta 180

catgactttg gtaaactagc caaagtcggt gttaaggaag ctggtgcttg gccagttcag 240

ttcggaacaa tcacggtttc tgatggaatc gccatgggaa cccaaggaat gcgtttctcc 300

ttgacatctc gtgatattat tgcagattct attgaagcag ccatgggagg tcataatgcg 360

gatgcttttg tagccattgg cggttgtgat aaaaacatgc ccggttctgt tatcgctatg 420

gctaacatgg atatcccagc catttttgct tacggcggaa caattgcacc tggtaattta 480

gacggcaaag atatcgattt agtctctgtc tttgaaggtg tcggccattg gaaccacggc 540

gatatgacca aagaagaagt taaagctttg gaatgtaatg cttgtcccgg tcctggaggc 600

tgcggtggta tgtatactgc taacacaatg gcgacagcta ttgaagtttt gggacttagc 660

cttccgggtt catcttctca cccggctgaa tccgcagaaa agaaagcaga tattgaagaa 720

gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa aaccttctga cattttaacg 780

cgtgaagctt ttgaagatgc tattactgta actatggctc tgggaggttc aaccaactca 840

acccttcacc tcttagctat tgcccatgct gctaatgtgg aattgacact tgatgatttc 900

aatactttcc aagaaaaagt tcctcatttg gctgatttga aaccttctgg tcaatatgta 960

ttccaagacc tttacaaggt cggaggggta ccagcagtta tgaaatatct ccttaaaaat 1020

ggcttccttc atggtgaccg tatcacttgt actggcaaaa cagtcgctga aaatttgaag 1080

gcttttgatg atttaacacc tggtcaaaag gttattatgc cgcttgaaaa tcctaaacgt 1140

gaagatggtc cgctcattat tctccatggt aacttggctc cagacggtgc cgttgccaaa 1200

gtttctggtg taaaagtgcg tcgtcatgtc ggtcctgcta aggtctttaa ttctgaagaa 1260

gaagccattg aagctgtctt gaatgatgat attgttgatg gtgatgttgt tgtcgtacgt 1320

tttgtaggac caaagggcgg tcctggtatg cctgaaatgc tttccctttc atcaatgatt 1380

gttggtaaag ggcaaggtga aaaagttgcc cttctgacag atggccgctt ctcaggtggt 1440

acttatggtc ttgtcgtggg tcatatcgct cctgaagcac aagatggcgg tccaatcgcc 1500

tacctgcaaa caggagacat agtcactatt gaccaagaca ctaaggaatt acactttgat 1560

atctccgatg aagagttaaa acatcgtcaa gagaccattg aattgccacc gctctattca 1620

cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg cttctagggg agccgtaaca 1680

gacttttgga agcctgaaga aactggcaaa aaa 1713

<210> 820
<211> 571
<212> PRT
<213> Artificial Sequence

<220>

20172000439 20 Jan 2017

<223> Streptococcus mutans ilvD

<400> 820

```
Met Thr Asp Lys Lys Thr Leu Lys Asp Leu Arg Asn Arg Ser Ser Val
1               5                  10                  15

Tyr Asp Ser Met Val Lys Ser Pro Asn Arg Ala Met Leu Arg Ala Thr
            20                  25                  30

Gly Met Gln Asp Glu Asp Phe Glu Lys Pro Ile Val Gly Val Ile Ser
        35                  40                  45

Thr Trp Ala Glu Asn Thr Pro Cys Asn Ile His Leu His Asp Phe Gly
    50                  55                  60

Lys Leu Ala Lys Val Gly Val Lys Glu Ala Gly Ala Trp Pro Val Gln
65                  70                  75                  80

Phe Gly Thr Ile Thr Val Ser Asp Gly Ile Ala Met Gly Thr Gln Gly
                85                  90                  95

Met Arg Phe Ser Leu Thr Ser Arg Asp Ile Ile Ala Asp Ser Ile Glu
            100                 105                 110

Ala Ala Met Gly Gly His Asn Ala Asp Ala Phe Val Ala Ile Gly Gly
        115                 120                 125

Cys Asp Lys Asn Met Pro Gly Ser Val Ile Ala Met Ala Asn Met Asp
    130                 135                 140

Ile Pro Ala Ile Phe Ala Tyr Gly Gly Thr Ile Ala Pro Gly Asn Leu
145                 150                 155                 160

Asp Gly Lys Asp Ile Asp Leu Val Ser Val Phe Glu Gly Val Gly His
                165                 170                 175

Trp Asn His Gly Asp Met Thr Lys Glu Glu Val Lys Ala Leu Glu Cys
            180                 185                 190

Asn Ala Cys Pro Gly Pro Gly Gly Cys Gly Gly Met Tyr Thr Ala Asn
        195                 200                 205

Thr Met Ala Thr Ala Ile Glu Val Leu Gly Leu Ser Leu Pro Gly Ser
    210                 215                 220

Ser Ser His Pro Ala Glu Ser Ala Glu Lys Lys Ala Asp Ile Glu Glu
225                 230                 235                 240

Ala Gly Arg Ala Val Val Lys Met Leu Glu Met Gly Leu Lys Pro Ser
                245                 250                 255

Asp Ile Leu Thr Arg Glu Ala Phe Glu Asp Ala Ile Thr Val Thr Met
            260                 265                 270
```

```
Ala Leu Gly Gly Ser Thr Asn Ser Thr Leu His Leu Leu Ala Ile Ala
        275                 280                 285

His Ala Ala Asn Val Glu Leu Thr Leu Asp Asp Phe Asn Thr Phe Gln
    290                 295                 300

Glu Lys Val Pro His Leu Ala Asp Leu Lys Pro Ser Gly Gln Tyr Val
305                 310                 315                 320

Phe Gln Asp Leu Tyr Lys Val Gly Gly Val Pro Ala Val Met Lys Tyr
                325                 330                 335

Leu Leu Lys Asn Gly Phe Leu His Gly Asp Arg Ile Thr Cys Thr Gly
            340                 345                 350

Lys Thr Val Ala Glu Asn Leu Lys Ala Phe Asp Asp Leu Thr Pro Gly
        355                 360                 365

Gln Lys Val Ile Met Pro Leu Glu Asn Pro Lys Arg Glu Asp Gly Pro
    370                 375                 380

Leu Ile Ile Leu His Gly Asn Leu Ala Pro Asp Gly Ala Val Ala Lys
385                 390                 395                 400

Val Ser Gly Val Lys Val Arg Arg His Val Gly Pro Ala Lys Val Phe
                405                 410                 415

Asn Ser Glu Glu Glu Ala Ile Glu Ala Val Leu Asn Asp Asp Ile Val
            420                 425                 430

Asp Gly Asp Val Val Val Val Arg Phe Val Gly Pro Lys Gly Gly Pro
        435                 440                 445

Gly Met Pro Glu Met Leu Ser Leu Ser Ser Met Ile Val Gly Lys Gly
    450                 455                 460

Gln Gly Glu Lys Val Ala Leu Leu Thr Asp Gly Arg Phe Ser Gly Gly
465                 470                 475                 480

Thr Tyr Gly Leu Val Val Gly His Ile Ala Pro Glu Ala Gln Asp Gly
                485                 490                 495

Gly Pro Ile Ala Tyr Leu Gln Thr Gly Asp Ile Val Thr Ile Asp Gln
            500                 505                 510

Asp Thr Lys Glu Leu His Phe Asp Ile Ser Asp Glu Glu Leu Lys His
        515                 520                 525

Arg Gln Glu Thr Ile Glu Leu Pro Pro Leu Tyr Ser Arg Gly Ile Leu
    530                 535                 540

Gly Lys Tyr Ala His Ile Val Ser Ser Ala Ser Arg Gly Ala Val Thr
545                 550                 555                 560
```

Asp Phe Trp Lys Pro Glu Glu Thr Gly Lys Lys
565 570

<210> 821
<211> 2145
<212> DNA
<213> Artificial Sequence

<220>
<223> GPM-sadB-ADHt segment

<400> 821
gcatgcttgc atttagtcgt gcaatgtatg actttaagat ttgtgagcag gaagaaaagg 60

gagaatcttc taacgataaa cccttgaaaa actgggtaga ctacgctatg ttgagttgct 120

acgcaggctg cacaattaca cgagaatgct cccgcctagg atttaaggct aagggacgtg 180

caatgcagac gacagatcta aatgaccgtg tcggtgaagt gttcgccaaa cttttcggtt 240

aacacatgca gtgatgcacg cgcgatggtg ctaagttaca tatatatata tatatatata 300

tatagccata gtgatgtcta agtaaccttt atggtatatt tcttaatgtg gaaagatact 360

agcgcgcgca cccacacaca agcttcgtct tttcttgaag aaaagaggaa gctcgctaaa 420

tgggattcca ctttccgttc cctgccagct gatggaaaaa ggttagtgga acgatgaaga 480

ataaaaagag agatccactg aggtgaaatt tcagctgaca gcgagtttca tgatcgtgat 540

gaacaatggt aacgagttgt ggctgttgcc agggagggtg gttctcaact tttaatgtat 600

ggccaaatcg ctacttgggt ttgttatata acaaagaaga aataatgaac tgattctctt 660

cctccttctt gtcctttctt aattctgttg taattacctt cctttgtaat ttttttgta 720

attattcttc ttaataatcc aaacaaacac acatattaca atagctagct gaggatgaag 780

gcattagttt atcatgggga tcacaaaatt tcgttagaag acaaaccaaa acccactctg 840

cagaaaccaa cagacgttgt ggttagggtg ttgaaaacaa caatttgcgg tactgacttg 900

ggaatataca aaggtaagaa tcctgaagtg gcagatggca gaatcctggg tcatgagggc 960

gttggcgtca ttgaagaagt gggcgaatcc gtgacacaat tcaaaaaggg ggataaagtt 1020

ttaatctcct gcgttactag ctgtggatcg tgtgattatt gcaagaagca actgtattca 1080

cactgtagag acggtggctg gattttaggt tacatgatcg acggtgtcca agccgaatac 1140

gtcagaatac cacatgctga caattcattg tataagatcc cgcaaactat cgatgatgaa 1200

attgcagtac tactgtccga tattttacct actggacatg aaattggtgt tcaatatggt 1260

aacgttcaac caggcgatgc tgtagcaatt gtaggagcag gtcctgttgg aatgtcagtt 1320

ttgttaactg ctcaatttta ctcgcctagt accattattg ttatcgacat ggacgaaaac 1380

cgtttacaat tagcgaagga gcttggggcc acacacacta ttaactccgg tactgaaaat 1440

gttgtcgaag ctgtgcatcg tatagcagcc gaaggagtgg atgtagcaat agaagctgtt 1500

ggtatacccg caacctggga catctgtcag gaaattgtaa aacccggcgc tcatattgcc 1560

aacgtgggag ttcatggtgt taaggtggac tttgaaattc aaaagttgtg gattaagaat 1620

ctaaccatca ccactggttt ggttaacact aatactaccc caatgttgat gaaggtagcc 1680

tctactgata aattgccttt aaagaaaatg attactcaca ggtttgagtt agctgaaatc 1740

gaacacgcat atcaggtttt cttgaatggc gctaaagaaa aagctatgaa gattattcta 1800

tctaatgcag gtgccgccta attaattaag agtaagcgaa tttcttatga tttatgattt 1860

ttattattaa ataagttata aaaaaaataa gtgtatacaa attttaaagt gactcttagg 1920

ttttaaaacg aaaattctta ttcttgagta actctttcct gtaggtcagg ttgctttctc 1980

aggtatagca tgaggtcgct cttattgacc acacctctac cggcatgccg agcaaatgcc 2040

tgcaaatcgc tccccatttc acccaattgt agatatgcta actccagcaa tgagttgatg 2100

aatctcggtg tgtattttat gtcctcagag gacaacacct gtggt 2145

<210> 822
<211> 4280
<212> DNA
<213> Artificial Sequence

<220>
<223> pUC19-URA3r

<400> 822
ggggatcctc tagagtcgac ctgcaggcat gcaagcttgg cgtaatcatg gtcatagctg 60

tttcctgtgt gaaattgtta tccgctcaca attccacaca acatacgagc cggaagcata 120

aagtgtaaag cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca 180

ctgcccgctt tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc 240

gcggggagag gcggtttgcg tattgggcgc tcttccgctt cctcgctcac tgactcgctg 300

cgctcggtcg ttcggctgcg gcgagcggta tcagctcact caaaggcggt aatacggtta 360

tccacagaat caggggataa cgcaggaaag aacatgtgag caaaaggcca gcaaaaggcc 420

aggaaccgta aaaaggccgc gttgctggcg tttttccata ggctccgccc ccctgacgag 480

catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cgacaggact ataaagatac 540

caggcgtttc cccctggaag ctccctcgtg cgctctcctg ttccgaccct gccgcttacc 600

ggatacctgt ccgcctttct cccttcggga agcgtggcgc tttctcatag ctcacgctgt 660

aggtatctca gttcggtgta ggtcgttcgc tccaagctgg gctgtgtgca cgaaccccc 720

gttcagcccg accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggtaaga 780

cacgacttat cgccactggc agcagccact ggtaacagga ttagcagagc gaggtatgta 840

ggcggtgcta cagagttctt gaagtggtgg cctaactacg gctacactag aaggacagta 900

tttggtatct gcgctctgct gaagccagtt accttcggaa aaagagttgg tagctcttga 960

tccggcaaac aaaccaccgc tggtagcggt ggtttttttg tttgcaagca gcagattacg 1020

cgcagaaaaa aaggatctca agaagatcct ttgatctttt ctacggggtc tgacgctcag 1080

tggaacgaaa actcacgtta agggattttg gtcatgagat tatcaaaaag gatcttcacc 1140

tagatccttt taaattaaaa atgaagtttt aaatcaatct aaagtatata tgagtaaact 1200

tggtctgaca gttaccaatg cttaatcagt gaggcaccta tctcagcgat ctgtctattt 1260

cgttcatcca tagttgcctg actccccgtc gtgtagataa ctacgatacg ggagggctta 1320

ccatctggcc ccagtgctgc aatgataccg cgagacccac gctcaccggc tccagattta 1380

tcagcaataa accagccagc cggaagggcc gagcgcagaa gtggtcctgc aactttatcc 1440

gcctccatcc agtctattaa ttgttgccgg gaagctagag taagtagttc gccagttaat 1500

agtttgcgca acgttgttgc cattgctaca ggcatcgtgg tgtcacgctc gtcgtttggt 1560
atggcttcat tcagctccgg ttcccaacga tcaaggcgag ttacatgatc ccccatgttg 1620
tgcaaaaaag cggttagctc cttcggtcct ccgatcgttg tcagaagtaa gttggccgca 1680
gtgttatcac tcatggttat ggcagcactg cataattctc ttactgtcat gccatccgta 1740
agatgctttt ctgtgactgg tgagtactca accaagtcat tctgagaata gtgtatgcgg 1800
cgaccgagtt gctcttgccc ggcgtcaata cgggataata ccgcgccaca tagcagaact 1860
ttaaaagtgc tcatcattgg aaaacgttct tcggggcgaa aactctcaag gatcttaccg 1920
ctgttgagat ccagttcgat gtaacccact cgtgcaccca actgatcttc agcatctttt 1980
actttcacca gcgtttctgg gtgagcaaaa acaggaaggc aaaatgccgc aaaaaaggga 2040
ataagggcga cacggaaatg ttgaatactc atactcttcc tttttcaata ttattgaagc 2100
atttatcagg gttattgtct catgagcgga tacatatttg aatgtattta gaaaaataaa 2160
caaatagggg ttccgcgcac atttccccga aaagtgccac ctgacgtcta agaaaccatt 2220
attatcatga cattaaccta taaaaatagg cgtatcacga ggccctttcg tctcgcgcgt 2280
ttcggtgatg acggtgaaaa cctctgacac atgcagctcc cggagacggt cacagcttgt 2340
ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg cgtcagcggg tgttggcggg 2400
tgtcggggct ggcttaacta tgcggcatca gagcagattg tactgagagt gcaccatatg 2460
cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc gcatcaggcg ccattcgcca 2520
ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct attacgccag 2580
ctggcgaaag ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttcccag 2640
tcacgacgtt gtaaaacgac ggccagtgaa ttcgagctcg gtacccccgg ctctgagaca 2700
gtagtaggtt agtcatcgct ctaccgacgc gcaggaaaag aaagaagcat tgcggattac 2760
gtattctaat gttcagcccg cggaacgcca gcaaatcacc acccatgcgc atgatactga 2820
gtcttgtaca cgctgggctt ccagtgtact gagagtgcac cataccacag cttttcaatt 2880
caattcatca ttttttttt attcttttt ttgatttcgg tttctttgaa atttttttga 2940
ttcggtaatc tccgaacaga aggaagaacg aaggaaggag cacagactta gattggtata 3000
tatacgcata tgtagtgttg aagaaacatg aaattgccca gtattcttaa cccaactgca 3060
cagaacaaaa acctgcagga aacgaagata aatcatgtcg aaagctacat ataaggaacg 3120
tgctgctact catcctagtc ctgttgctgc caagctattt aatatcatgc acgaaaagca 3180
aacaaacttg tgtgcttcat tggatgttcg taccaccaag gaattactgg agttagttga 3240
agcattaggt cccaaaattt gtttactaaa aacacatgtg gatatcttga ctgatttttc 3300
catggagggc acagttaagc cgctaaaggc attatccgcc aagtacaatt ttttactctt 3360
cgaagacaga aaatttgctg acattggtaa tacagtcaaa ttgcagtact ctgcgggtgt 3420
atacagaata gcagaatggg cagacattac gaatgcacac ggtgtggtgg gcccaggtat 3480
tgttagcggt ttgaagcagg cggcagaaga agtaacaaag gaacctagag gccttttgat 3540
gttagcagaa ttgtcatgca agggctccct atctactgga gaatatacta agggtactgt 3600
tgacattgcg aagagcgaca aagattttgt tatcggcttt attgctcaaa gagacatggg 3660

tggaagagat gaaggttacg attggttgat tatgacaccc ggtgtgggtt tagatgacaa 3720

gggagacgca ttgggtcaac agtatagaac cgtggatgat gtggtctcta caggatctga 3780

cattattatt gttggaagag gactatttgc aaagggaagg gatgctaagg tagagggtga 3840

acgttacaga aaagcaggct gggaagcata tttgagaaga tgcggccagc aaaactaaaa 3900

aactgtatta taagtaaatg catgtatact aaactcacaa attagagctt caatttaatt 3960

atatcagtta ttaccctatg cggtgtgaaa taccgcacag atgcgtaagg agaaaatacc 4020

gcatcaggaa attgtaaacg ttaatatttt gttaaaattc gcgttaaatt tttgttaaat 4080

cagctcattt tttaaccaat aggccgaaat cggcaaaatc ttcagcccgc ggaacgccag 4140

caaatcacca cccatgcgca tgatactgag tcttgtacac gctgggcttc cagtgatgat 4200

acaacgagtt agccaaggtg agcacggatg tctaaattag aattacgttt taatatcttt 4260

ttttccatat ctagggctag 4280

<210> 823
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-11A

<400> 823
gcatgcttgc atttagtcgt gcaatgtatg 30

<210> 824
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-11B

<400> 824
gaacattaga atacgtaatc cgcaatgcac tagtaccaca ggtgttgtcc tctg 54

<210> 825
<211> 54
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-11C

<400> 825
cagaggacaa cacctgtggt actagtgcat tgcggattac gtattctaat gttc 54

<210> 826
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-11D

<400> 826
caccttggct aactcgttgt atcatcac 28

20 Jan 2017

2017200439

<210> 827
<211> 100
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-13A

<400> 827
ttttaagccg aatgagtgac agaaaaagcc cacaacttat caagtgatat tgaacaaagg 60

gcgaaacttc gcatgcttgc atttagtcgt gcaatgtatg 100

<210> 828
<211> 98
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-13B

<400> 828
cccaattggt aaatattcaa caagagacgc gcagtacgta acatgcgaat tgcgtaattc 60

acggcgataa caccttggct aactcgttgt atcatcac 98

<210> 829
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 112590-34G

<400> 829
caaaagccca tgtcccacac caaaggatg 29

<210> 830
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 112590-34H

<400> 830
caccatcgcg cgtgcatcac tgcatg 26

<210> 831
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 112590-34F

<400> 831
tcggtttttg caatatgacc tgtgggcc 28

<210> 832
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 112590-49E

20 Jan 2017 2017200439

<400> 832
gagaagatgc ggccagcaaa ac 22

<210> 833
<211> 2745
<212> DNA
<213> Artificial Sequence

<220>
<223> pdc1:: PPDC1-ilvD-FBA1t-URA3r

<400> 833
atgactgaca aaaaaactct taaagactta agaaatcgta gttctgttta cgattcaatg 60

gttaaatcac ctaatcgtgc tatgttgcgt gcaactggta tgcaagatga agactttgaa 120

aaacctatcg tcggtgtcat ttcaacttgg gctgaaaaca caccttgtaa tatccactta 180

catgactttg gtaaactagc caaagtcggt gttaaggaag ctggtgcttg gccagttcag 240

ttcggaacaa tcacggtttc tgatggaatc gccatgggaa cccaaggaat gcgtttctcc 300

ttgacatctc gtgatattat tgcagattct attgaagcag ccatgggagg tcataatgcg 360

gatgcttttg tagccattgg cggttgtgat aaaaacatgc ccggttctgt tatcgctatg 420

gctaacatgg atatcccagc cattttgct tacggcggaa caattgcacc tggtaattta 480

gacggcaaag atatcgattt agtctctgtc tttgaaggtg tcggccattg gaaccacggc 540

gatatgacca aagaagaagt taaagctttg gaatgtaatg cttgtcccgg tcctggaggc 600

tgcggtggta tgtatactgc taacacaatg gcgacagcta ttgaagtttt gggacttagc 660

cttccgggtt catcttctca cccggctgaa tccgcagaaa agaaagcaga tattgaagaa 720

gctggtcgcg ctgttgtcaa aatgctcgaa atgggcttaa aaccttctga cattttaacg 780

cgtgaagctt ttgaagatgc tattactgta actatggctc tgggaggttc aaccaactca 840

acccttcacc tcttagctat tgcccatgct gctaatgtgg aattgacact tgatgatttc 900

aatactttcc aagaaaaagt tcctcatttg gctgatttga aaccttctgg tcaatatgta 960

ttccaagacc tttacaaggt cggaggggta ccagcagtta tgaaatatct ccttaaaaat 1020

ggcttccttc atggtgaccg tatcacttgt actggcaaaa cagtcgctga aaatttgaag 1080

gcttttgatg atttaacacc tggtcaaaag gttattatgc cgcttgaaaa tcctaaacgt 1140

gaagatggtc cgctcattat tctccatggt aacttggctc cagacggtgc cgttgccaaa 1200

gtttctggtg taaaagtgcg tcgtcatgtc ggtcctgcta aggtctttaa ttctgaagaa 1260

gaagccattg aagctgtctt gaatgatgat attgttgatg gtgatgttgt tgtcgtacgt 1320

tttgtaggac caaagggcgg tcctggtatg cctgaaatgc tttccctttc atcaatgatt 1380

gttggtaaag ggcaaggtga aaaagttgcc cttctgacag atggccgctt ctcaggtggt 1440

acttatggtc ttgtcgtggg tcatatcgct cctgaagcac aagatggcgg tccaatcgcc 1500

tacctgcaaa caggagacat agtcactatt gaccaagaca ctaaggaatt acactttgat 1560

atctccgatg aagagttaaa acatcgtcaa gagaccattg aattgccacc gctctattca 1620

cgcggtatcc ttggtaaata tgctcacatc gtttcgtctg cttctagggg agccgtaaca 1680

gacttttgga agcctgaaga aactggcaaa aaatgttgtc ctggttgctg tggttaagcg 1740

gccgcgttaa ttcaaattaa ttgatatagt tttttaatga gtattgaatc tgtttagaaa 1800

taatggaata ttatttttat ttatttattt atattattgg tcggctcttt tcttctgaag 1860

gtcaatgaca aaatgatatg aaggaaataa tgatttctaa aattttacaa cgtaagatat 1920

ttttacaaaa gcctagctca tcttttgtca tgcactattt tactcacgct tgaaattaac 1980

ggccagtcca ctgcggagtc atttcaaagt catcctaatc gatctatcgt ttttgatagc 2040

tcattttgga gttcgcgatt gtcttctgtt attcacaact gttttaattt ttatttcatt 2100

ctggaactct tcgagttctt tgtaaagtct ttcatagtag cttactttat cctccaacat 2160

atttaacttc atgtcaattt cggctcttaa attttccaca tcatcaagtt caacatcatc 2220

ttttaacttg aatttattct ctagctcttc caaccaagcc tcattgctcc ttgatttact 2280

ggtgaaaagt gatacacttt gcgcgcaatc caggtcaaaa ctttcctgca aagaattcac 2340

caatttctcg acatcatagt acaatttgtt ttgttctccc atcacaattt aatatacctg 2400

atggattctt atgaagcgct gggtaatgga cgtgtcactc tacttcgcct ttttccctac 2460

tccttttagt acggaagaca atgctaataa ataagagggt aataataata ttattaatcg 2520

gcaaaaaaga ttaaacgcca agcgtttaat tatcagaaag caaacgtcgt accaatcctt 2580

gaatgcttcc caattgtata ttaagagtca tcacagcaac atattcttgt tattaaatta 2640

attattattg atttttgata ttgtataaaa aaaccaaata tgtataaaaa aagtgaataa 2700

aaaataccaa gtatggagaa atatattaga agtctatacg ttaaa 2745

<210> 834
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-36D

<400> 834
gacttttgga agcctgaaga aactggc 27

<210> 835
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Ophelia_SH

<400> 835

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

```
Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Ser Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 836
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> 112590-30F

<400> 836
ccaggccaat tcaacagact gtcggc 26

<210> 837
<211> 2347
<212> DNA
<213> Artificial Sequence

<220>
<223> URA3r2 template DNA

<400> 837
gcattgcgga ttacgtattc taatgttcag gtgctggaag aagagctgct taaccgccgc 60

gcccagggtg aagatccacg ctactttacc ctgcgtcgtc tggatttcgg cggctgtcgt 120

ctttcgctgg caacgccggt tgatgaagcc tgggacggtc cgctctcctt aaacggtaaa 180

cgtatcgcca cctcttatcc tcacctgctc aagcgttatc tcgaccagaa aggcatctct 240

tttaaatcct gcttactgaa cggttctgtt gaagtcgccc cgcgtgccgg actggcggat 300

gcgatttgcg atctggtttc caccggtgcc acgctggaag ctaacggcct gcgcgaagtc 360

gaagttatct atcgctcgaa agcctgcctg attcaacgcg atggcgaaat ggaagaatcc 420

aaacagcaac tgatcgacaa actgctgacc cgtattcagg gtgtgatcca ggcgcgcgaa 480

tcaaaataca tcatgatgca cgcaccgacc gaacgtctgg atgaagtcat ggtacctact 540

gagagtgcac cataccacag cttttcaatt caattcatca ttttttttt attcttttt 600

ttgatttcgg tttctttgaa attttttga ttcggtaatc tccgaacaga aggaagaacg 660

aaggaaggag cacagactta gattggtata tatacgcata tgtagtgttg aagaaacatg 720

aaattgccca gtattcttaa cccaactgca cagaacaaaa acctgcagga aacgaagata 780

aatcatgtcg aaagctacat ataaggaacg tgctgctact catcctagtc ctgttgctgc 840

caagctattt aatatcatgc acgaaaagca aacaaacttg tgtgcttcat tggatgttcg 900

taccaccaag gaattactgg agttagttga agcattaggt cccaaaattt gtttactaaa 960

aacacatgtg gatatcttga ctgatttttc catggagggc acagttaagc cgctaaaggc 1020

attatccgcc aagtacaatt ttttactctt cgaagacaga aaatttgctg acattggtaa 1080

tacagtcaaa ttgcagtact ctgcgggtgt atacagaata gcagaatggg cagacattac 1140

gaatgcacac ggtgtggtgg gcccaggtat tgttagcggt ttgaagcagg cggcagaaga 1200

agtaacaaag gaacctagag gccttttgat gttagcagaa ttgtcatgca agggctccct 1260

atctactgga gaatatacta agggtactgt tgacattgcg aagagcgaca aagattttgt 1320

tatcggcttt attgctcaaa gagacatggg tggaagagat gaaggttacg attggttgat 1380

tatgacaccc ggtgtgggtt tagatgacaa gggagacgca ttgggtcaac agtatagaac 1440

cgtggatgat gtggtctcta caggatctga cattattatt gttggaagag gactatttgc 1500

aaagggaagg gatgctaagg tagagggtga acgttacaga aaagcaggct gggaagcata 1560

tttgagaaga tgcggccagc aaaactaaaa aactgtatta taagtaaatg catgtatact 1620

aaactcacaa attagagctt caatttaatt atatcagtta ttaccctatg cggtgtgaaa 1680

taccgcacag atgcgtaagg agaaaatacc gcatcaggaa attgtaaacg ttaatatttt 1740

gttaaaattc gcgttaaatt tttgttaaat cagctcattt tttaaccaat aggccgaaat 1800

cggcaaaatc tctagagtgc tggaagaaga gctgcttaac cgccgcgccc agggtgaaga 1860

tccacgctac tttaccctgc gtcgtctgga tttcggcggc tgtcgtcttt cgctggcaac 1920

gccggttgat gaagcctggg acggtccgct ctccttaaac ggtaaacgta tcgccacctc 1980

ttatcctcac ctgctcaagc gttatctcga ccagaaaggc atctctttta aatcctgctt 2040

actgaacggt tctgttgaag tcgccccgcg tgccggactg gcggatgcga tttgcgatct 2100

ggtttccacc ggtgccacgc tggaagctaa cggcctgcgc gaagtcgaag ttatctatcg 2160

ctcgaaagcc tgcctgattc aacgcgatgg cgaaatggaa gaatccaaac agcaactgat 2220

cgacaaactg ctgacccgta ttcagggtgt gatccaggcg cgcgaatcaa aatacatcat 2280

gatgcacgca ccgaccgaac gtctggatga agtcatccag tgatgataca acgagttagc 2340

caaggtg 2347

<210> 838
<211> 80
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-45A

<400> 838
cttcgaagaa tatactaaaa aatgagcagg caagataaac gaaggcaaag gcattgcgga 60

ttacgtattc taatgttcag 80

<210> 839
<211> 81
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 114117-45B

<400> 839
tatacacatg tatatatatc gtatgctgca gctttaaata atcggtgtca caccttggct 60

aactcgttgt atcatcactg g 81

<210> 840
<211> 90
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 384

<400> 840
atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcca 60

gcattgcgga ttacgtattc taatgttcag 90

<210> 841
<211> 91
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer 385

<400> 841
ttaagcaccg atgataccaa cggacttacc ttcagcaatt ctttttggg ccaaagcagc 60

caccttggct aactcgttgt atcatcactg g 91

<210> 842
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer N869

<400> 842
ctaggatgag tagcagcacg ttcc 24

<210> 843
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer N871

<400> 843
ccaattccgt gatgtctctt tgttgc 26

<210> 844
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer N946

<400> 844
gtgaacgagt tcacaaccgc 20

<210> 845
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer N947

<400> 845
gttcgttcca gaattatcac gc 22

<210> 846
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer PDC5:KanMXF

<400> 846
gacttgaata atgcagcggc gcttgc 26

<210> 847
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer PDC5::KanMXR

<400> 847
ccaccctctt caattagcta agatcatagc 30

<210> 848
<211> 90
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer T001

<400> 848
atgcctacct tgtatactga tatcgaaatc ccacaattga aaatctcttt aaagcaaccg 60

gcattgcgga ttacgtattc taatgttcag 90

<210> 849
<211> 97
<212> DNA
<213> Artificial Sequence

<220>
<223> T002 primer

<400> 849
ttagttgtcc aaagagagat ttatgtgaac tgcttttgtt tgaagatagg tatcaacacc 60

actttgcacc ttggctaact cgttgtatca tcactgg 97

<210> 850
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> T003 primer

<400> 850
gaacgttaca gaaaagcagg ctgggaagc 29

<210> 851
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> T004 primer

<400> 851
gacctagtaa caacaattgt tccgc 25

<210> 852
<211> 27
<212> DNA
<213> Artificial Sequence

<220>

<223> T005 primer

<400> 852
ccgtgctacc tcttaatgtg tcacaag 27

<210> 853
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Pat_SH

<400> 853

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Leu
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
130 135 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145 150 155 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
165 170 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
180 185 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
195 200 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
210 215 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225 230 235 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
245 250 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
260 265 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
275 280 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
290 295 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305 310 315 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
325 330 335

Asn Asn

<210> 854
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Quetnin_SH

<400> 854

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Thr Ile Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

```
Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 855
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Ralph_SH

<400> 855

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                   10                  15
```

20172000439 20 Jan 2017

```
Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Leu Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300
```

```
Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 856
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Ursala_SH

<400> 856

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                   10                  15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Val Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205
```

```
Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 857
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Watson_SH

<400> 857

```
Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1               5                   10                  15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
            20                  25                  30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
        35                  40                  45

Glu Gly Ala Glu Glu Cys Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
    50                  55                  60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65                  70                  75                  80

Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys Asn Asp Ile Glu Pro
                85                  90                  95
```

```
Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 858
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Watson_SH

<400> 858
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt gcaaaagagc tgaagaacaa 180

ggcttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 859
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Xavier_SH

<400> 859

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Lys Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

```
His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 860
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> k9_Zeke_SH

<400> 860
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaatgcaggc tgccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 861
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Zeke_SH

<400> 861

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Met Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

```
His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
        115                 120                 125

Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 862
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Annabel_SH

<400> 862
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaatgcaggc tatcatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 863
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> K9_Annabel_SH

<400> 863

Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn Leu Ser Leu Leu Asp
1 5 10 15

Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala His
20 25 30

Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val Ile Ile Gly Leu Phe
35 40 45

Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu Gln Gly Phe Glu Val
50 55 60

Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp Ile Ile Met Ile Leu
65 70 75 80

Ile Pro Asp Glu Met Gln Ala Ala Met Tyr Lys Asn Asp Ile Glu Pro
85 90 95

Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala His Gly Phe Asn Ile
100 105 110

His Phe Gly Cys Ile Val Pro Pro Lys Asp Val Asp Val Thr Met Ile
115 120 125

```
Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser Glu Tyr Glu Glu Gly
    130                 135                 140

Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln Asp Ala Thr Gly Lys
145                 150                 155                 160

Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala Ile Gly Gly Ala Arg
                165                 170                 175

Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu Thr Glu Thr Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val Cys Ala Leu Met Gln
        195                 200                 205

Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Asp Pro Arg Asn Ala
    210                 215                 220

Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile Val Asp Leu Ile Tyr
225                 230                 235                 240

Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile Thr Glu Asp Thr Lys
            260                 265                 270

Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln Asp Gly Thr Phe Ala
        275                 280                 285

Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly Ser Gln Val His Phe
    290                 295                 300

Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro Ala Glu Val Val Gly
305                 310                 315                 320

Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp Glu Asp Lys Leu Ile
                325                 330                 335

Asn Asn
```

<210> 864
<211> 340
<212> PRT
<213> Artificial Sequence

<220>
<223> Lactococcus lactis KARI encoded

<400> 864

```
Met Ala Val Thr Met Tyr Tyr Glu Asp Asp Val Glu Val Ser Ala Leu
1               5                   10                  15
```

20 Jan 2017 2017200439

```
Ala Gly Lys Gln Ile Ala Val Ile Gly Tyr Gly Ser Gln Gly His Ala
            20                  25                  30

His Ala Gln Asn Leu Arg Asp Ser Gly His Asn Val Ile Ile Gly Val
        35                  40                  45

Arg His Gly Lys Ser Phe Asp Lys Ala Lys Glu Asp Gly Phe Glu Thr
    50                  55                  60

Phe Glu Val Gly Glu Ala Val Ala Lys Ala Asp Val Ile Met Val Leu
65                  70                  75                  80

Ala Pro Asp Glu Leu Gln Gln Ser Ile Tyr Glu Glu Asp Ile Lys Pro
                85                  90                  95

Asn Leu Lys Ala Gly Ser Ala Leu Gly Phe Ala His Gly Phe Asn Ile
            100                 105                 110

His Phe Gly Tyr Ile Lys Val Pro Glu Asp Val Asp Val Phe Met Val
        115                 120                 125

Ala Pro Lys Ala Pro Gly His Leu Val Arg Arg Thr Tyr Thr Glu Gly
    130                 135                 140

Phe Gly Thr Pro Ala Leu Phe Val Ser His Gln Asn Ala Ser Gly His
145                 150                 155                 160

Ala Arg Glu Ile Ala Met Asp Trp Ala Lys Gly Ile Gly Cys Ala Arg
                165                 170                 175

Val Gly Ile Ile Glu Thr Thr Phe Lys Glu Glu Thr Glu Glu Asp Leu
            180                 185                 190

Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Leu Thr Ala Leu Val Glu
        195                 200                 205

Ala Gly Phe Glu Thr Leu Thr Glu Ala Gly Tyr Ala Gly Glu Leu Ala
    210                 215                 220

Tyr Phe Glu Val Leu His Glu Met Lys Leu Ile Val Asp Leu Met Tyr
225                 230                 235                 240

Glu Gly Gly Phe Thr Lys Met Arg Gln Ser Ile Ser Asn Thr Ala Glu
                245                 250                 255

Phe Gly Asp Tyr Val Thr Gly Pro Arg Ile Ile Thr Asp Glu Val Lys
            260                 265                 270

Lys Asn Met Lys Leu Val Leu Ala Asp Ile Gln Ser Gly Lys Phe Ala
        275                 280                 285

Gln Asp Phe Val Asp Asp Phe Lys Ala Gly Arg Pro Lys Leu Ile Ala
    290                 295                 300
```

```
Tyr Arg Glu Ala Ala Lys Asn Leu Glu Ile Glu Lys Ile Gly Ala Glu
305                 310                 315                 320

Leu Arg Gln Ala Met Pro Phe Thr Gln Ser Gly Asp Asp Asp Ala Phe
                325                 330                 335

Lys Ile Tyr Gln
            340
```

<210> 865
<211> 1074
<212> DNA
<213> Artificial sequence

<220>
<223> AOX1 from Pichia stipitis

<400> 865

```
atgctttctg tgcagactac aagagccgcc aagttgcaac taggtcaatt acctctgatt       60

gcctacaccg ccagaagtgg aagacttcac caccaattct actccaccgt tgctgaaaag      120

acagccaacc ctacgccaaa cacctcagat aaaactaata tttttgatat tagaaccaag      180

gtgtacgatg agactgatat aagaaaacat gacgacaatc agtttatcac ccatccttta      240

tttcctcatc ctacgttttc gcaggaagat tgtttgaaag ttggttacga acatcgtcct      300

cctcggactt ttggtgacaa aatggctttc agaggcattg aacttgtcag aggttctttc      360

gactttgtca ccggttacaa gaagccaaag acacaggctg atatcgattc aggttttaaa      420

ggtaccagat acgagatgac agaaggtaaa tggttgacca gatgtatatt cttagaaagt      480

attgctggag ttccaggtgc cgttgcttcc ttcatcagac acttgcattc tttacgtttg      540

ttgaagagag acaaagcctg gatcgaaacc ttacttgatg aagcattcaa cgaaagaatg      600

catttactta ccttcatcaa gattggcaaa cctagttggt tcaccagaac aatcatctac      660

gtcggccaag gtgtattctg caacttgttc ttcttgttct acttggccaa ccctaagtat      720

tgtcacaggt ttgttggcta ccttgaagaa gaggctgtta gtacctacac tcatttcgtt      780

catgaattac aatcaggtaa gcttcccaag tttgagaaca tcaagattcc aaccattgca      840

tggcaatact ggccagagtt gaccgagaat tcgtccatgt tggacttgat tttaagaatc      900

agagccgacg aagccaagca cagagaagtc aaccacacct tggccaattt agatcaaaga      960

aaggacagaa atccatttgg gttggcaatt ccggatctca aggaaccgca acccgaaagt     1020

ggcttaaagg ttacgaaacc ccacggctgg gagaaggaag aattgaaatt gtaa           1074
```

<210> 866
<211> 357
<212> PRT
<213> Artificial Sequence

<220>
<223> AOX1 from Pichia stipitis

<400> 866

```
Met Leu Ser Val Gln Thr Thr Arg Ala Ala Lys Leu Gln Leu Gly Gln
1               5                   10                  15
```

```
Leu Pro Ser Ile Ala Tyr Thr Ala Arg Ser Gly Arg Leu His His Gln
            20                  25                  30

Phe Tyr Ser Thr Val Ala Glu Lys Thr Ala Asn Pro Thr Pro Asn Thr
        35                  40                  45

Ser Asp Lys Thr Asn Ile Phe Asp Ile Arg Thr Lys Val Tyr Asp Glu
    50                  55                  60

Thr Asp Ile Arg Lys His Asp Asp Asn Gln Phe Ile Thr His Pro Leu
65                  70                  75                  80

Phe Pro His Pro Thr Phe Ser Gln Glu Asp Cys Leu Lys Val Gly Tyr
                85                  90                  95

Glu His Arg Pro Pro Arg Thr Phe Gly Asp Lys Met Ala Phe Arg Gly
            100                 105                 110

Ile Glu Leu Val Arg Gly Ser Phe Asp Phe Val Thr Gly Tyr Lys Lys
        115                 120                 125

Pro Lys Thr Gln Ala Asp Ile Asp Ser Gly Phe Lys Gly Thr Arg Tyr
    130                 135                 140

Glu Met Thr Glu Gly Lys Trp Leu Thr Arg Cys Ile Phe Leu Glu Ser
145                 150                 155                 160

Ile Ala Gly Val Pro Gly Ala Val Ala Ser Phe Ile Arg His Leu His
                165                 170                 175

Ser Leu Arg Leu Leu Lys Arg Asp Lys Ala Trp Ile Glu Thr Leu Leu
            180                 185                 190

Asp Glu Ala Phe Asn Glu Arg Met His Leu Leu Thr Phe Ile Lys Ile
        195                 200                 205

Gly Lys Pro Ser Trp Phe Thr Arg Thr Ile Ile Tyr Val Gly Gln Gly
    210                 215                 220

Val Phe Cys Asn Leu Phe Phe Leu Phe Tyr Leu Ala Asn Pro Lys Tyr
225                 230                 235                 240

Cys His Arg Phe Val Gly Tyr Leu Glu Glu Glu Ala Val Ser Thr Tyr
                245                 250                 255

Thr His Phe Val His Glu Leu Gln Ser Gly Lys Leu Pro Lys Phe Glu
            260                 265                 270

Asn Ile Lys Ile Pro Thr Ile Ala Trp Gln Tyr Trp Pro Glu Leu Thr
        275                 280                 285

Glu Asn Ser Ser Met Leu Asp Leu Ile Leu Arg Ile Arg Ala Asp Glu
    290                 295                 300
```

```
Ala Lys His Arg Glu Val Asn His Thr Leu Ala Asn Leu Asp Gln Arg
305                 310                 315                 320

Lys Asp Arg Asn Pro Phe Gly Leu Ala Ile Pro Asp Leu Lys Glu Pro
                325                 330                 335

Gln Pro Glu Ser Gly Leu Lys Val Thr Lys Pro His Gly Trp Glu Lys
            340                 345                 350

Glu Glu Leu Lys Leu
        355
```

<210> 867
<211> 1992
<212> DNA
<213> Artificial Sequence

<220>
<223> AOX2 from Pichia stipitis

<400> 867

```
atggccattc ctgaagaatt cgatattctt gtcctgggtg gtggatccag tggatcctgt       60
attgccggaa gattggccaa cttggaccac tccttgaaag ttggtcttat cgaggctggt      120
gagaacaatc ttaacaaccc atgggtctac cttccaggta tttacccaag aaacatgaag      180
ttggactcca aaactgcttc tttctacacc tccaaccctt ctcctcattt gaatggtaga      240
agagctattg tcccatgtgc caacatcttg ggtggtggtt cttcgatcaa cttcatgatg      300
tacaccagag gttccgcttc tgattacgat gactttgaag ctgagggatg gaagaccaag      360
gatttgcttc ctttgatgaa gaagactgag acttaccaaa gagcttgcaa caaccctgaa      420
attcacggtt ttgaaggtcc aatcaaggtt tctttcggta actacactta cccagtttgt      480
caagacttct tgagagcaac tgaatcccaa ggtattccat acgttgacga cttggaagac      540
ttggtgactg ctcatggtgc tgaacactgg ctgaaatgga tcaacagaga cactggtcgt      600
cgttccgact ctgctcatgc cttcgttcat tctacgatga gaaaccacga caatctgtac      660
ttgatctgca acaccaaagt tgacaagatt attgttgaag acggaagagc tgctgctgtc      720
agaaccgttc caagtaaacc tttgaacgca aagaagccaa ctcacaaggt ttatcgtgct      780
agaacgcaaa tcgttttgtc ttgtggtacc atctcttctc ctctggttct gcaaagatcc      840
ggttttggtg acccaatcaa attgagagcc gctggtgtta agcctttggt caacttgcca      900
ggtgttggaa gaaacttcca agaccactac tgcttcttct ctccttacag aattaagccc      960
caatacgagt ctttcgatga cttcgtacgt ggtgacgcta acattcaaaa gaaggtattc     1020
gaccaatggt acgctaacgg tactggtcca ttggccacca acggtattga agccggtgtc     1080
aagattagac caactccaga agaattatct cagatggacg agtccttcca agagggttac     1140
agagagtact tcgaagacaa accagacaag ccagttatgc actattccat cattgctggt     1200
ttcttcggtg accacaccaa gattccacct ggaaagtaca tgaccatgtt ccacttcttg     1260
gagtacccat tctccagagg ttctatccac atcacctctc cagacccata cgcaactcca     1320
gactttgacc caggtttcat gaacgatgaa agagacatgg ctcctatggt ctggtcttac     1380
aagaagtcca gagagactgc cagaaaaatg gaccactttg ctggtgaggt tacttcccac     1440
```

caccctctgt tcccatactc atccgaggcc agagcttacg agatggattt ggagacctcc 1500

aacgcctacg gtggaccact gaacttgact gctggtcttg ctcacggttc ttggactcag 1560

cctttgaaga agcctgctgg aagaaacgaa ggacatgtta cttccaacca agtcgagctt 1620

catccagaca ttgagtacga tgaggaggat gataaggcca ttgagaacta cattcgtgag 1680

cacactgaga ccacatggca ctgtctggga acctgttcca ttggtccaag agagggttcc 1740

aagatcgtca aatggggtgg tgttttggat cacagatcta acgtttacgg agtcaagggc 1800

ctgaaggttg gtgacttgtc cgtctgtcca gacaatgttg gttgtaacac ctacaccacc 1860

gctcttttga tcggtgaaaa gactgccacc ttggttggtg aagacttagg atacacaggt 1920

gaggccttag acatgactgt acctcagttc aagttgggca cttacgagaa gactggtctt 1980

gctagattct ag 1992

<210> 868
<211> 663
<212> PRT
<213> Artificial Sequence

<220>
<223> AOX2 from Pichia stipitis

<400> 868

Met Ala Ile Pro Glu Glu Phe Asp Ile Leu Val Leu Gly Gly Gly Ser
1 5 10 15

Ser Gly Ser Cys Ile Ala Gly Arg Leu Ala Asn Leu Asp His Ser Leu
20 25 30

Lys Val Gly Leu Ile Glu Ala Gly Glu Asn Asn Leu Asn Asn Pro Trp
35 40 45

Val Tyr Leu Pro Gly Ile Tyr Pro Arg Asn Met Lys Leu Asp Ser Lys
50 55 60

Thr Ala Ser Phe Tyr Thr Ser Asn Pro Ser Pro His Leu Asn Gly Arg
65 70 75 80

Arg Ala Ile Val Pro Cys Ala Asn Ile Leu Gly Gly Gly Ser Ser Ile
85 90 95

Asn Phe Met Met Tyr Thr Arg Gly Ser Ala Ser Asp Tyr Asp Asp Phe
100 105 110

Glu Ala Glu Gly Trp Lys Thr Lys Asp Leu Leu Pro Leu Met Lys Lys
115 120 125

Thr Glu Thr Tyr Gln Arg Ala Cys Asn Asn Pro Glu Ile His Gly Phe
130 135 140

Glu Gly Pro Ile Lys Val Ser Phe Gly Asn Tyr Thr Tyr Pro Val Cys
145 150 155 160

```
Gln Asp Phe Leu Arg Ala Thr Glu Ser Gln Gly Ile Pro Tyr Val Asp
                165                 170                 175

Asp Leu Glu Asp Leu Val Thr Ala His Gly Ala Glu His Trp Leu Lys
            180                 185                 190

Trp Ile Asn Arg Asp Thr Gly Arg Arg Ser Asp Ser Ala His Ala Phe
        195                 200                 205

Val His Ser Thr Met Arg Asn His Asp Asn Leu Tyr Leu Ile Cys Asn
    210                 215                 220

Thr Lys Val Asp Lys Ile Ile Val Glu Asp Gly Arg Ala Ala Ala Val
225                 230                 235                 240

Arg Thr Val Pro Ser Lys Pro Leu Asn Ala Lys Lys Pro Thr His Lys
                245                 250                 255

Val Tyr Arg Ala Arg Thr Gln Ile Val Leu Ser Cys Gly Thr Ile Ser
            260                 265                 270

Ser Pro Leu Val Leu Gln Arg Ser Gly Phe Gly Asp Pro Ile Lys Leu
        275                 280                 285

Arg Ala Ala Gly Val Lys Pro Leu Val Asn Leu Pro Gly Val Gly Arg
    290                 295                 300

Asn Phe Gln Asp His Tyr Cys Phe Phe Ser Pro Tyr Arg Ile Lys Pro
305                 310                 315                 320

Gln Tyr Glu Ser Phe Asp Asp Phe Val Arg Gly Asp Ala Asn Ile Gln
                325                 330                 335

Lys Lys Val Phe Asp Gln Trp Tyr Ala Asn Gly Thr Gly Pro Leu Ala
            340                 345                 350

Thr Asn Gly Ile Glu Ala Gly Val Lys Ile Arg Pro Thr Pro Glu Glu
        355                 360                 365

Leu Ser Gln Met Asp Glu Ser Phe Gln Glu Gly Tyr Arg Glu Tyr Phe
    370                 375                 380

Glu Asp Lys Pro Asp Lys Pro Val Met His Tyr Ser Ile Ile Ala Gly
385                 390                 395                 400

Phe Phe Gly Asp His Thr Lys Ile Pro Pro Gly Lys Tyr Met Thr Met
                405                 410                 415

Phe His Phe Leu Glu Tyr Pro Phe Ser Arg Gly Ser Ile His Ile Thr
            420                 425                 430

Ser Pro Asp Pro Tyr Ala Thr Pro Asp Phe Asp Pro Gly Phe Met Asn
        435                 440                 445
```

```
Asp Glu Arg Asp Met Ala Pro Met Val Trp Ser Tyr Lys Lys Ser Arg
    450                 455                 460

Glu Thr Ala Arg Lys Met Asp His Phe Ala Gly Glu Val Thr Ser His
465                 470                 475                 480

His Pro Leu Phe Pro Tyr Ser Ser Glu Ala Arg Ala Tyr Glu Met Asp
                485                 490                 495

Leu Glu Thr Ser Asn Ala Tyr Gly Gly Pro Leu Asn Leu Thr Ala Gly
            500                 505                 510

Leu Ala His Gly Ser Trp Thr Gln Pro Leu Lys Lys Pro Ala Gly Arg
        515                 520                 525

Asn Glu Gly His Val Thr Ser Asn Gln Val Glu Leu His Pro Asp Ile
    530                 535                 540

Glu Tyr Asp Glu Glu Asp Asp Lys Ala Ile Glu Asn Tyr Ile Arg Glu
545                 550                 555                 560

His Thr Glu Thr Thr Trp His Cys Leu Gly Thr Cys Ser Ile Gly Pro
                565                 570                 575

Arg Glu Gly Ser Lys Ile Val Lys Trp Gly Gly Val Leu Asp His Arg
            580                 585                 590

Ser Asn Val Tyr Gly Val Lys Gly Leu Lys Val Gly Asp Leu Ser Val
        595                 600                 605

Cys Pro Asp Asn Val Gly Cys Asn Thr Tyr Thr Thr Ala Leu Leu Ile
    610                 615                 620

Gly Glu Lys Thr Ala Thr Leu Val Gly Glu Asp Leu Gly Tyr Thr Gly
625                 630                 635                 640

Glu Ala Leu Asp Met Thr Val Pro Gln Phe Lys Leu Gly Thr Tyr Glu
                645                 650                 655

Lys Thr Gly Leu Ala Arg Phe
            660
```

<210> 869
<211> 1410
<212> DNA
<213> Artificial sequence

<220>
<223> YER081W Chr 5

<400> 869

```
atgacaagca ttgacattaa caacttacaa aataccttc aacaagctat gaatatgagc       60

ggctccccag gcgctgtttg tacttcacct acgcaatctt tcatgaatac cgttccacag      120

cgcttgaatg ctgtaaagca cccaaaaatt ttgaagcctt tctcaacggg tgatatgaag      180

attttactat tagaaaacgt taatcaaact gctattacaa tcttcgaaga gcaaggttac      240
```

caagtcgaat tctataaatc ttcattgccc gaggaagagt tgatcgaaaa gatcaaggac 300
gttcatgcta ttggtatcag atcaaagact agattaactt caaatgtctt acaacatgcg 360
aagaatctgg tttgtattgg ttgtttctgt atcggtacca accaagttga cttagactac 420
gctaccagca gaggtattgc tgttttcaac tcgcctttct ccaactcaag atcagtagca 480
gaattggtca tcgctgaaat cattagttta gcaagacaac taggtgatag atctatcgaa 540
ttacataccg gtacatggaa taaggttgct gctagatgtt gggaggtaag aggaaaaact 600
cttggtatta ttgggtacgg tcacattggt tcccaattat cagttcttgc agaagctatg 660
ggtttgcatg tgttgtacta cgatattgta actatcatgg ccttgggtac tgccagacaa 720
gtttctacat tagatgaatt gttgaataaa tctgattttg tgacactaca tgtaccagct 780
actcctgaaa ctgaaaaaat gttatctgcc ccacaatttg ctgctatgaa ggatggcgct 840
tatgttatta atgcttcaag aggtactgtc gtggacattc catctttgat ccaagccgtg 900
aaagccaaca aaattgcagg tgctgctttg gatgtttatc cacatgaacc agctaagaac 960
ggtgaaggtt catttaacga tgagctaaat agctggactt ctgaattagt ttcattacca 1020
aatatcatct tgacaccaca cattggtggc tctaccgaag aagcccaaag ctcaatcggt 1080
attgaagtgg ctaccgcatt gtccaaatac atcaatgaag gtaactctgt cggttcagtc 1140
aacttcccag aagtggcatt gaaatcattg tcttacgacc aagagaacac tgtgcgtgtg 1200
ttatacattc accaaaatgt accaggtgtt ttgaagaccg tcaatgatat tttatcgaac 1260
cataacatcg aaaagcaatt ttccgattca aatggtgaaa ttgcttattt aatggctgat 1320
atctcttctg ttgaccaaag cgatattaaa gatatttatg aacaactaaa tcaaacctct 1380
gctaagatct caattagatt gctatattaa 1410

<210> 870
<211> 469
<212> PRT
<213> Artificial Sequence

<220>
<223> YER081W Chr 5

<400> 870

Met Thr Ser Ile Asp Ile Asn Asn Leu Gln Asn Thr Phe Gln Gln Ala
1 5 10 15

Met Asn Met Ser Gly Ser Pro Gly Ala Val Cys Thr Ser Pro Thr Gln
20 25 30

Ser Phe Met Asn Thr Val Pro Gln Arg Leu Asn Ala Val Lys His Pro
35 40 45

Lys Ile Leu Lys Pro Phe Ser Thr Gly Asp Met Lys Ile Leu Leu Leu
50 55 60

Glu Asn Val Asn Gln Thr Ala Ile Thr Ile Phe Glu Glu Gln Gly Tyr
65 70 75 80

2017200439 20 Jan 2017

```
Gln Val Glu Phe Tyr Lys Ser Ser Leu Pro Glu Glu Glu Leu Ile Glu
                85                  90                  95

Lys Ile Lys Asp Val His Ala Ile Gly Ile Arg Ser Lys Thr Arg Leu
            100                 105                 110

Thr Ser Asn Val Leu Gln His Ala Lys Asn Leu Val Cys Ile Gly Cys
        115                 120                 125

Phe Cys Ile Gly Thr Asn Gln Val Asp Leu Asp Tyr Ala Thr Ser Arg
    130                 135                 140

Gly Ile Ala Val Phe Asn Ser Pro Phe Ser Asn Ser Arg Ser Val Ala
145                 150                 155                 160

Glu Leu Val Ile Ala Glu Ile Ile Ser Leu Ala Arg Gln Leu Gly Asp
                165                 170                 175

Arg Ser Ile Glu Leu His Thr Gly Thr Trp Asn Lys Val Ala Ala Arg
            180                 185                 190

Cys Trp Glu Val Arg Gly Lys Thr Leu Gly Ile Ile Gly Tyr Gly His
        195                 200                 205

Ile Gly Ser Gln Leu Ser Val Leu Ala Glu Ala Met Gly Leu His Val
    210                 215                 220

Leu Tyr Tyr Asp Ile Val Thr Ile Met Ala Leu Gly Thr Ala Arg Gln
225                 230                 235                 240

Val Ser Thr Leu Asp Glu Leu Leu Asn Lys Ser Asp Phe Val Thr Leu
                245                 250                 255

His Val Pro Ala Thr Pro Glu Thr Glu Lys Met Leu Ser Ala Pro Gln
            260                 265                 270

Phe Ala Ala Met Lys Asp Gly Ala Tyr Val Ile Asn Ala Ser Arg Gly
        275                 280                 285

Thr Val Val Asp Ile Pro Ser Leu Ile Gln Ala Val Lys Ala Asn Lys
    290                 295                 300

Ile Ala Gly Ala Ala Leu Asp Val Tyr Pro His Glu Pro Ala Lys Asn
305                 310                 315                 320

Gly Glu Gly Ser Phe Asn Asp Glu Leu Asn Ser Trp Thr Ser Glu Leu
                325                 330                 335

Val Ser Leu Pro Asn Ile Ile Leu Thr Pro His Ile Gly Gly Ser Thr
            340                 345                 350

Glu Glu Ala Gln Ser Ser Ile Gly Ile Glu Val Ala Thr Ala Leu Ser
        355                 360                 365
```

```
Lys Tyr Ile Asn Glu Gly Asn Ser Val Gly Ser Val Asn Phe Pro Glu
    370                 375                 380

Val Ala Leu Lys Ser Leu Ser Tyr Asp Gln Glu Asn Thr Val Arg Val
385                 390                 395                 400

Leu Tyr Ile His Gln Asn Val Pro Gly Val Leu Lys Thr Val Asn Asp
                405                 410                 415

Ile Leu Ser Asn His Asn Ile Glu Lys Gln Phe Ser Asp Ser Asn Gly
            420                 425                 430

Glu Ile Ala Tyr Leu Met Ala Asp Ile Ser Ser Val Asp Gln Ser Asp
        435                 440                 445

Ile Lys Asp Ile Tyr Glu Gln Leu Asn Gln Thr Ser Ala Lys Ile Ser
    450                 455                 460

Ile Arg Leu Leu Tyr
465
```

<210> 871
<211> 1365
<212> DNA
<213> Artificial Sequence

<220>
<223> YOR375C Chr 15

<400> 871

```
atgtcagagc cagaatttca acaagcttac gaagaagttg tctcctcttt ggaagactct       60

actcttttcg aacaacaccc agaatacaga aaggttttgc caattgtttc tgttccagaa      120

agaatcatac aattcagagt cacctgggaa aatgacaagg gtgaacaaga agttgctcaa      180

ggttacagag tgcaatataa ctccgccaag ggtccataca agggtggtct acgtttccat      240

ccttccgtga acttgtctat cttgaaattc ttgggtttcg aacaaatctt caagaactcc      300

ttgaccggcc tagacatggg tggtggtaaa ggtggtctat gtgtggactt gaagggaaga      360

tctaataacg aaatcagaag aatctgttat gctttcatga gagaattgag cagacacatt      420

ggtcaagaca ctgacgtgcc agctggtgat atcggtgttg gtggtcgtga aattggttac      480

ctgttcggtg cttacagatc atacaagaac tcctgggaag gtgtcttaac cggtaagggt      540

ttgaactggg gtggttcttt gatcagacca gaagccactg gttacggttt agtttactat      600

actcaagcta tgatcgacta tgccacaaac ggtaaggaat ctttcgaagg taagcgcgtc      660

accatctctg gtagtggtaa cgttgctcaa tacgctgcct tgaaggttat tgagctaggt      720

ggtactgtcg tttccctatc tgactccaag ggttgtatca tctctgaaac tggtatcacc      780

tccgaacaag tcgctgatat ttccagtgct aaggtcaact tcaagtcctt ggaacaaatc      840

gtcaacgaat actctacttt ctccgaaaac aaagtgcaat acattgctgg tgctcgtcca      900

tggacccacg tccaaaaggt cgacattgct ttgccatgtg ccacccaaaa tgaagtcagc      960

ggtgaagaag ccaaggcctt ggttgctcaa ggtgtcaagt ttattgccga aggttccaac     1020

atgggttcca ctccagaagc tattgccgtc tttgaaactg ctcgttccac cgccactgga     1080
```

ccaagcgaag ctgtttggta cggtccacca aaggctgcta acttgggtgg tgttgctgtt 1140

tctggtttag aaatggcaca aaactctcaa agaatcacat ggactagcga aagagttgac 1200

caagagttga agagaattat gatcaactgt ttcaatgaat gtatcgacta tgccaagaag 1260

tacactaagg acggtaaggt cttgccatct ttggtcaaag gtgctaatat cgcaagtttc 1320

atcaaggtct ctgatgctat gtttgaccaa ggtgatgtat tttaa 1365

<210> 872
<211> 454
<212> PRT
<213> Artificial Sequence

<220>
<223> YOR375C Chr 15 reverse complement

<400> 872

Met Ser Glu Pro Glu Phe Gln Gln Ala Tyr Glu Glu Val Val Ser Ser
1 5 10 15

Leu Glu Asp Ser Thr Leu Phe Glu Gln His Pro Glu Tyr Arg Lys Val
20 25 30

Leu Pro Ile Val Ser Val Pro Glu Arg Ile Ile Gln Phe Arg Val Thr
35 40 45

Trp Glu Asn Asp Lys Gly Glu Gln Glu Val Ala Gln Gly Tyr Arg Val
50 55 60

Gln Tyr Asn Ser Ala Lys Gly Pro Tyr Lys Gly Gly Leu Arg Phe His
65 70 75 80

Pro Ser Val Asn Leu Ser Ile Leu Lys Phe Leu Gly Phe Glu Gln Ile
85 90 95

Phe Lys Asn Ser Leu Thr Gly Leu Asp Met Gly Gly Gly Lys Gly Gly
100 105 110

Leu Cys Val Asp Leu Lys Gly Arg Ser Asn Asn Glu Ile Arg Arg Ile
115 120 125

Cys Tyr Ala Phe Met Arg Glu Leu Ser Arg His Ile Gly Gln Asp Thr
130 135 140

Asp Val Pro Ala Gly Asp Ile Gly Val Gly Gly Arg Glu Ile Gly Tyr
145 150 155 160

Leu Phe Gly Ala Tyr Arg Ser Tyr Lys Asn Ser Trp Glu Gly Val Leu
165 170 175

Thr Gly Lys Gly Leu Asn Trp Gly Gly Ser Leu Ile Arg Pro Glu Ala
180 185 190

Thr Gly Tyr Gly Leu Val Tyr Tyr Thr Gln Ala Met Ile Asp Tyr Ala
195 200 205

```
Thr Asn Gly Lys Glu Ser Phe Glu Gly Lys Arg Val Thr Ile Ser Gly
    210                 215                 220

Ser Gly Asn Val Ala Gln Tyr Ala Ala Leu Lys Val Ile Glu Leu Gly
225                 230                 235                 240

Gly Thr Val Val Ser Leu Ser Asp Ser Lys Gly Cys Ile Ile Ser Glu
                245                 250                 255

Thr Gly Ile Thr Ser Glu Gln Val Ala Asp Ile Ser Ser Ala Lys Val
            260                 265                 270

Asn Phe Lys Ser Leu Glu Gln Ile Val Asn Glu Tyr Ser Thr Phe Ser
        275                 280                 285

Glu Asn Lys Val Gln Tyr Ile Ala Gly Ala Arg Pro Trp Thr His Val
    290                 295                 300

Gln Lys Val Asp Ile Ala Leu Pro Cys Ala Thr Gln Asn Glu Val Ser
305                 310                 315                 320

Gly Glu Glu Ala Lys Ala Leu Val Ala Gln Gly Val Lys Phe Ile Ala
                325                 330                 335

Glu Gly Ser Asn Met Gly Ser Thr Pro Glu Ala Ile Ala Val Phe Glu
            340                 345                 350

Thr Ala Arg Ser Thr Ala Thr Gly Pro Ser Glu Ala Val Trp Tyr Gly
        355                 360                 365

Pro Pro Lys Ala Ala Asn Leu Gly Gly Val Ala Val Ser Gly Leu Glu
    370                 375                 380

Met Ala Gln Asn Ser Gln Arg Ile Thr Trp Thr Ser Glu Arg Val Asp
385                 390                 395                 400

Gln Glu Leu Lys Arg Ile Met Ile Asn Cys Phe Asn Glu Cys Ile Asp
                405                 410                 415

Tyr Ala Lys Lys Tyr Thr Lys Asp Gly Lys Val Leu Pro Ser Leu Val
            420                 425                 430

Lys Gly Ala Asn Ile Ala Ser Phe Ile Lys Val Ser Asp Ala Met Phe
        435                 440                 445

Asp Gln Gly Asp Val Phe
    450
```

<210> 873
<211> 343
<212> PRT
<213> Artificial Sequence

<220>

<223> 878 C1

<400> 873

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 874
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 879 A7

<400> 874

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Met Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

```
Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  875
<211>  343
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  879 C2

<400>  875

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala Gln Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 876
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 880 A11

<400> 876

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Thr Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225 230 235 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
245 250 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
260 265 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
275 280 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
290 295 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Gln
305 310 315 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
325 330 335

Glu Asp Lys Leu Ile Asn Asn
340

<210> 877
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 880 B4

<400> 877

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 878
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 880 D11

<400> 878

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45
```

```
Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Met Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Val Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335
```

Glu Asp Lys Leu Ile Asn Asn
340

<210> 879
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 881 A2

<400> 879

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

Ile Ile Met Ile Leu Ile Asn Asp Glu Met Gln Ala Thr Met Tyr Lys
85 90 95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
100 105 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
115 120 125

Asp Val Thr Met Val Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
130 135 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145 150 155 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
165 170 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
180 185 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
195 200 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
210 215 220

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Thr Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 880
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 881 G3

<400> 880

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Gly Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  881
<211>  343
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  881 G9

<400>  881

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15
```

20172000439 20 Jan 2017

```
Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Glu Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300
```

```
Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 882
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 882 B12

<400> 882

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile Arg Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205
```

```
Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340


<210>  883
<211>  343
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  882 C10

<400>  883

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Ser Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Ser Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95
```

```
Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Tyr Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asn Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Asp Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 884
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309I_111711r

20 Jan 2017 2017200439

<400> 884
ctttctcata gccttaatgt ggac 24

<210> 885
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309L_111711r

<400> 885
ctttctcata gcctttaagt ggac 24

<210> 886
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309Y_111711r

<400> 886
ctttctcata gccttatagt ggac 24

<210> 887
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309C_111711r

<400> 887
ctttctcata gccttacagt ggac 24

<210> 888
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309W_111711r

<400> 888
ctttctcata gccttccagt ggac 24

<210> 889
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309P_111711r

<400> 889
ctttctcata gcctttgggt ggac 24

<210> 890
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309H_111711r

20172000439 20 Jan 2017

<400> 890
ctttctcata gccttatggt ggac 24

<210> 891
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309Q_111711r

<400> 891
ctttctcata gcctttggt ggac 24

<210> 892
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309M_111711r

<400> 892
ctttctcata gccttcatgt ggac 24

<210> 893
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309N_111711r

<400> 893
ctttctcata gccttattgt ggac 24

<210> 894
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309V_111711r

<400> 894
ctttctcata gccttaacgt ggac 24

<210> 895
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309A_111711r

<400> 895
ctttctcata gccttagcgt ggac 24

<210> 896
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309D_111711r

20 Jan 2017 2017200439

<400> 896
ctttctcata gccttatcgt ggac 24

<210> 897
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309E_111711r

<400> 897
ctttctcata gcctttcgt ggac 24

<210> 898
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309G_111711r

<400> 898
ctttctcata gccttaccgt ggac 24

<210> 899
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309S_111711r

<400> 899
ctttctcata gccttagagt ggac 24

<210> 900
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309T_111711r

<400> 900
ctttctcata gccttagtgt ggac 24

<210> 901
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309R_111711r

<400> 901
ctttctcata gcctttctgt ggac 24

<210> 902
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_309K_111711r

<400> 902
ctttctcata gccttcttgt ggac 24

<210> 903
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Kshort1

<400> 903
aaaccggttt aaacagtatg gctaagattt actaccaaga agactg 46

<210> 904
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> YGrev

<400> 904
ttctgtttta tcagaccgct tc 22

<210> 905
<211> 4978
<212> DNA
<213> Artificial Sequence

<220>
<223> pBAD-ps-JEA1

<400> 905
aagaaaccaa ttgtccatat tgcatcagac attgccgtca ctgcgtcttt tactggctct 60

tctcgctaac caaaccggta accccgctta ttaaaagcat tctgtaacaa agcgggacca 120

aagccatgac aaaaacgcgt aacaaaagtg tctataatca cggcagaaaa gtccacattg 180

attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata agattagcgg 240

atcctacctg acgctttta tcgcaactct ctactgtttc tccatacccg ttttttgggc 300

taacaggagg aagtttaaac agtatgaaag ttttctacga taaagactgc gacctgtcga 360

tcatccaagg taagaaagtt gccatcatcg gcttcggttc ccagggccac gctcaagcac 420

tcaacctgaa ggattccggc gtagacgtga ctgttggcct gcctaaaggc tttgctgatg 480

tagccaaggc tgaagcccac ggctttaaag tgaccgacgt tgctgcagcc gttgccggtg 540

ccgacttggt catgatcctg attccggacg agttccagtc ccagctgtac aagaacgaaa 600

tcgagccgaa catcaagaag ggcgccactc tggccttctc ccacggcttc gcgatccact 660

acaaccaggt tgtgcctcgt gccgacctcg acgtgatcat gatcgcgccg aaggctccag 720

gccacaccgt acgttccgag ttcgtcaagg gcggaggtat tcctgacctg atcgcgatct 780

accaggacgt ttccggcaac gccaagaacg tcgccctgtc ctacgccgca ggcgtgggcg 840

gcggccgtac cggcatcatc gaaaccacct tcaaggacga gactgaaacc gacctgttcg 900

gtgagcaggc tgttctgtgt ggcggtaccg tcgagctggt caaagccggt ttcgaaaccc 960

tggttgaagc tggctacgct ccagaaatgg cctacttcga gtgcctgcac gaactgaagc 1020

tgatcgttga cctcatgtac gaaggcggta tcgccaacat gaactactcg atctccaaca 1080

acgctgaata cggcgagtac gtgactggtc cagaagtcat caacgccgaa tcccgtcagg 1140
ccatgcgcaa tgctctgaag cgcatccagg acggcgaata cgcgaagatg ttcatcagcg 1200
aaggcgctac cggctaccca tcgatgaccg ccaagcgtcg taacaacgct gctcacggta 1260
tcgaaatcat cggcgagcaa ctgcgctcga tgatgccttg gatcggtgcc aacaaaatcg 1320
tcgacaaagc caagaactaa ggccctgcag gccttggcgg atgagagaag attttcagcc 1380
tgatacagat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca 1440
gtagcgcggt ggtcccacct gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg 1500
atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaaacga 1560
aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggt gaacgctctc 1620
ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccggaggg 1680
tggcgggcag gacgcccgcc ataaactgcc aggcatcaaa ttaagcagaa ggccatcctg 1740
acggatggcc ttttgcgtt tctacaaact cttttgttta tttttctaaa tacattcaaa 1800
tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa 1860
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 1920
tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 1980
tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg agagttttcg 2040
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 2100
atcccgtgtt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 2160
cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga 2220
attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac ttctgacaac 2280
gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 2340
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 2400
gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 2460
agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 2520
gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 2580
gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat 2640
ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 2700
tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 2760
tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 2820
catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 2880
gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct tgcaaacaaa 2940
aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc 3000
gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 3060
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatcct 3120
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg 3180
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 3240

20172000439 20 Jan 2017

cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 3300
cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg 3360
agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt 3420
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 3480
gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca 3540
catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 3600
agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 3660
ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt cacaccgcat 3720
atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg 3780
ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg 3840
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg 3900
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gcagatcaat 3960
tcgcgcgcga aggcgaagcg gcatgcataa tgtgcctgtc aaatggacga agcagggatt 4020
ctgcaaaccc tatgctactc cgtcaagccg tcaattgtct gattcgttac caattatgac 4080
aacttgacgg ctacatcatt cacttttct tcacaaccgg cacggaactc gctcgggctg 4140
gccccggtgc atttttaaa tacccgcgag aaatagagtt gatcgtcaaa accaacattg 4200
cgaccgacgg tggcgatagg catccgggtg gtgctcaaaa gcagcttcgc ctggctgata 4260
cgttggtcct cgcgccagct taagacgcta atccctaact gctggcggaa aagatgtgac 4320
agacgcgacg gcgacaagca aacatgctgt gcgacgctgg cgatatcaaa attgctgtct 4380
gccaggtgat cgctgatgta ctgacaagcc tcgcgtaccc gattatccat cggtggatgg 4440
agcgactcgt taatcgcttc catgcgccgc agtaacaatt gctcaagcag atttatcgcc 4500
agcagctccg aatagcgccc ttccccttgc ccggcgttaa tgatttgccc aaacaggtcg 4560
ctgaaatgcg gctggtgcgc ttcatccggg cgaaagaacc ccgtattggc aaatattgac 4620
ggccagttaa gccattcatg ccagtaggcg cgcggacgaa agtaaaccca ctggtgatac 4680
cattcgcgag cctccggatg acgaccgtag tgatgaatct ctcctggcgg gaacagcaaa 4740
atatcacccg gtcggcaaac aaattctcgt ccctgatttt tcaccacccc ctgaccgcga 4800
atggtgagat tgagaatata acctttcatt cccagcggtc ggtcgataaa aaaatcgaga 4860
taaccgttgg cctcaatcgg cgttaaaccc gccaccagat gggcattaaa cgagtatccc 4920
ggcagcaggg gatcattttg cgcttcagcc atacttttca tactcccgcc attcagag 4978

<210> 906
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> KPSH1

<400> 906
aaacagtatg gctaagattt actaccaaga agactg 36

20 Jan 2017 2017200439

<210> 907
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Eliza_SH

<400> 907
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaagat gcggaggaat ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 908
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Frank_SH

<400> 908
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagctggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 909
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Grace_SH

<400> 909
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat catgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

ataaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg atatgtctga tgatggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 910
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ingrid_SH

<400> 910
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatttac aaaaacgaca tcgaaccaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
cagatccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960
gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 911
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Jarvis_SH

<400> 911
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60
gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtatcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 912
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Kelly_SH

<400> 912
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt ggggaggagt ggaatagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaaaaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 913
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> 9_Norman_SH

<400> 913
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaatgt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 914
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ophelia_SH

<400> 914
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgcttca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccat gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 915
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Pat_SH

<400> 915
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60
gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attagaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840
gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900
caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960
gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 916
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Quentin

<400> 916
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60
gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120
tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180
ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240
atcccagatg aaaagcaggc taccatttac aaaaacgaca tcgaaccaaa cttggaagcc 300
ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360
aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420
tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480
gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540
gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600
ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660
ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720
caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780
attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 917
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ralph_SH

<400> 917
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttagaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 918
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ursala

<400> 918
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gtggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc taccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgttt gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 919
<211> 1017
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Xavier_SH

<400> 919
atggctaaga tttactacca agaagactgt aacttgtcct tgttggatgg taagactatc 60

gccgttatcg gttacggttc tcaaggtcac gctcatgccc tgaatgctaa ggaatccggt 120

tgtaacgtta tcattggttt attcgaaggt gcggaggagt ggaaaagagc tgaagaacaa 180

ggtttcgaag tctacaccgc tgctgaagct gctaagaagg ctgacatcat tatgatcttg 240

atcccagatg aaaagcaggc tgccatgtac aaaaacgaca tcgaaccaaa cttggaagcc 300

ggtaacatgt tgatgttcgc tcacggtttc aacatccatt tcggttgtat tgttccacca 360

aaggacgttg atgtcactat gatcgctcca aagggtccag gtcacaccgt tagatccgaa 420

tacgaagaag gtaaaggtgt cccatgcttg gttgctgtcg aacaagacgc tactggcaag 480

gctttggata tggctttggc ctacgcttta gccatcggtg gtgctagagc cggtgtcttg 540

gaaactacct tcagaaccga aactgaaacc gacttgttcg gtgaacaagc tgttttatgt 600

ggtggtgtct gcgctttgat gcaggccggt tttgaaacct tggttgaagc cggttacgac 660

ccaagaaacg cttacttcga atgtatccac gaaatgaagt tgatcgttga cttgatctac 720

caatctggtt tctccggtat gcgttactct atctccaaca ctgctgaata cggtgactac 780

attaccggtc caaagatcat tactgaagat accaagaagg ctatgaagaa gattttgtct 840

gacattcaag atggtacctt tgccaaggac ttcttggttg acatgtctga tgctggttcc 900

caggtccact tcaaggctat gagaaagttg gcctccgaac acccagctga agttgtcggt 960

gaagaaatta gatccttgta ctcctggtcc gacgaagaca agttgattaa caactga 1017

<210> 920
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> oMT93A

<400> 920
gatcccagat gaaatgcagg ctgccatgta caaaaacgac atcg 44

<210> 921
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> oMT93Arev

<400> 921
cgatgtcgtt tttgtacatg gcagcctgca tttcatctgg gatc 44

<210> 922
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> oMT93I

<400> 922
gatcccagat gaaatgcagg ctatcatgta caaaaacgac atcg 44

<210> 923
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> oMT93Irev

<400> 923
cgatgtcgtt tttgtacatg atagcctgca tttcatctgg gatc 44

<210> 924
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> pHR81-F

<400> 924
acacccagta ttttccctttcc 22

<210> 925
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> pHR81- Rev

<400> 925
ctagtgtaca gatgtatgtc gg 22

<210> 926
<211> 570
<212> PRT
<213> Artificial Sequence

<220>
<223> Lactococcus lactis DHAD

<400> 926

```
Met Glu Phe Lys Tyr Asn Gly Lys Val Glu Ser Ile Glu Leu Asn Lys
1               5                  10                  15

Tyr Ser Lys Thr Leu Thr Gln Asp Pro Thr Gln Pro Ala Thr Gln Ala
            20                  25                  30

Met His Tyr Gly Ile Gly Phe Lys Asp Glu Asp Phe Lys Lys Ala Gln
        35                  40                  45

Val Gly Ile Val Ser Met Asp Trp Asp Gly Asn Pro Cys Asn Met His
    50                  55                  60

Leu Gly Thr Leu Gly Ser Lys Ile Lys Asn Ser Val Asn Gln Thr Asp
65                  70                  75                  80

Gly Leu Ile Gly Leu Gln Phe His Thr Ile Gly Val Ser Asp Gly Ile
                85                  90                  95

Ala Asn Gly Lys Leu Gly Met Arg Tyr Ser Leu Val Ser Arg Glu Val
            100                 105                 110

Ile Ala Asp Ser Ile Glu Thr Asn Ala Gly Ala Glu Tyr Tyr Asp Ala
        115                 120                 125

Ile Val Ala Val Pro Gly Cys Asp Lys Asn Met Pro Gly Ser Ile Ile
    130                 135                 140

Gly Met Ala Arg Leu Asn Arg Pro Ser Ile Met Val Tyr Gly Gly Thr
145                 150                 155                 160

Ile Glu His Gly Glu Tyr Lys Gly Glu Lys Leu Asn Ile Val Ser Ala
                165                 170                 175

Phe Glu Ala Leu Gly Gln Lys Ile Thr Gly Asn Ile Ser Glu Glu Asp
            180                 185                 190

Tyr His Gly Val Ile Cys Asn Ala Ile Pro Gly Gln Gly Ala Cys Gly
        195                 200                 205

Gly Met Tyr Thr Ala Asn Thr Leu Ala Ser Ala Ile Glu Thr Leu Gly
    210                 215                 220

Met Ser Leu Pro Tyr Ser Ala Ser Asn Pro Ala Val Ser Gln Glu Lys
225                 230                 235                 240

Glu Asp Glu Cys Asp Glu Ile Gly Leu Ala Ile Lys Asn Leu Leu Glu
                245                 250                 255

Lys Asp Ile Lys Pro Ser Asp Ile Met Thr Lys Glu Ala Phe Glu Asn
            260                 265                 270
```

```
Ala Ile Thr Ile Val Met Val Leu Gly Gly Ser Thr Asn Ala Val Leu
        275                 280                 285

His Ile Ile Ala Met Ala Asn Ala Ile Gly Val Glu Ile Thr Gln Asp
    290                 295                 300

Asp Phe Gln Arg Ile Ser Asp Val Thr Pro Val Leu Gly Asp Phe Lys
305                 310                 315                 320

Pro Ser Gly Lys Tyr Met Met Glu Asp Leu His Lys Ile Gly Gly Val
                325                 330                 335

Pro Ala Val Leu Lys Tyr Leu Leu Lys Glu Gly Lys Leu His Gly Asp
            340                 345                 350

Cys Leu Thr Val Thr Gly Lys Thr Leu Ala Glu Asn Val Glu Thr Ala
        355                 360                 365

Leu Asp Leu Asp Phe Asp Ser Gln Asp Ile Ile Arg Pro Leu Glu Asn
    370                 375                 380

Pro Ile Lys Ala Thr Gly His Leu Gln Ile Leu Tyr Gly Asn Leu Ala
385                 390                 395                 400

Glu Gly Gly Ser Val Ala Lys Ile Ser Gly Lys Glu Gly Glu Phe Phe
                405                 410                 415

Lys Gly Thr Ala Arg Val Phe Asp Gly Glu Gln His Phe Ile Asp Gly
            420                 425                 430

Ile Glu Ser Gly Arg Leu His Ala Gly Asp Val Ala Val Ile Arg Asn
        435                 440                 445

Ile Gly Pro Val Gly Gly Pro Gly Met Pro Glu Met Leu Lys Pro Thr
    450                 455                 460

Ser Ala Leu Ile Gly Ala Gly Leu Gly Lys Ser Cys Ala Leu Ile Thr
465                 470                 475                 480

Asp Gly Arg Phe Ser Gly Gly Thr His Gly Phe Val Val Gly His Ile
                485                 490                 495

Val Pro Glu Ala Val Glu Gly Gly Leu Ile Gly Leu Val Glu Asp Asp
            500                 505                 510

Asp Ile Ile Glu Ile Asp Ala Val Asn Asn Ser Ile Ser Leu Lys Val
        515                 520                 525

Ala Asp Asp Glu Ile Ala Arg Arg Arg Ala Asn Tyr Gln Lys Pro Ala
    530                 535                 540

Pro Lys Ala Thr Arg Gly Val Leu Ala Lys Phe Ala Lys Leu Thr Arg
545                 550                 555                 560
```

```
Pro Ala Ser Glu Gly Cys Val Thr Asp Leu
                565                 570


<210>  927
<211>  343
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  K9C4

<400>  927

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220
```

```
Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Ala Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 928
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> K9C8

<400> 928

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Phe Glu Gly Ala Glu Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Pro Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125
```

```
Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Thr Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 929
<211> 4978
<212> DNA
<213> Artificial Sequence

<220>
<223> K9SB2_SH in pBAD.KARI

<400> 929
aagaaaccaa ttgtccatat tgcatcagac attgccgtca ctgcgtcttt tactggctct 60

tctcgctaac caaaccggta accccgctta ttaaaagcat tctgtaacaa agcgggacca 120

aagccatgac aaaaacgcgt aacaaaagtg tctataatca cggcagaaaa gtccacattg 180

attatttgca cggcgtcaca ctttgctatg ccatagcatt tttatccata agattagcgg 240
atcctacctg acgcttttta tcgcaactct ctactgtttc tccatacccg ttttttgggc 300
taacaggagg aagtttaaac agtatggcta agatttacta ccaagaagac tgtaacttgt 360
ccttgttgga tggtaagact atcgccgtta tcggttacgg ttctcaaggt cacgctcatg 420
ccctgaatgc taaggaatcc ggttgtaacg ttatcattgg tttattcgaa ggtgcggagg 480
agtggaaaag agctgaagaa caaggtttcg aagtctacac cgctgctgaa gctgctaaga 540
aggctgacat cattatgatc ttgatcccag atgaaaagca ggctaccatg tacaaaaacg 600
acatcgaacc aaacttggaa gccggtaaca tgttgatgtt cgctcacggt ttcaacatcc 660
atttcggttg tattgttcca ccaaaggacg ttgatgtcac tatgatcgct ccaaagggtc 720
caggtcacac cgttagatcc gaatacgaag aaggtaaagg tgtcccatgc ttggttgctg 780
tcgaacaaga cgctactggc aaggctttgg atatggcttt ggcctacgct ttagccatcg 840
gtggtgctag agccggtgtc ttggaaacta ccttcagaac cgaaactgaa accgacttgt 900
tcggtgaaca agctgtttta tgtggtggtg tctgcgcttt gatgcaggcc ggttttgaaa 960
ccttggttga agccggttac gacccaagaa acgcttactt cgaatgtatc cacgaaatga 1020
agttgatcgt tgacttgatc taccaatctg gtttctccgg tatgcgttac tctatctcca 1080
acactgctga atacggtgac tacattaccg gtccaaagat cattactgaa gataccaaga 1140
aggctatgaa gaagattttg tctgacattc aagatggtac ctttgccaag gacttcttgg 1200
ttgacatgtc tgatgctggt tcccaggtcc acttcaaggc tatgagaaag ttggcctccg 1260
aacacccagc tgaagttgtc ggtgaagaaa ttagatcctt gtactcctgg tccgacgaag 1320
acaagttgat taacaactga ggccctgcag gccttggcgg atgagagaag attttcagcc 1380
tgatacagat taaatcagaa cgcagaagcg gtctgataaa acagaatttg cctggcggca 1440
gtagcgcggt ggtcccacct gaccccatgc cgaactcaga agtgaaacgc cgtagcgccg 1500
atggtagtgt ggggtctccc catgcgagag tagggaactg ccaggcatca aataaaacga 1560
aaggctcagt cgaaagactg ggcctttcgt tttatctgtt gtttgtcggt gaacgctctc 1620
ctgagtagga caaatccgcc gggagcggat ttgaacgttg cgaagcaacg gcccggaggg 1680
tggcgggcag gacgcccgcc ataaactgcc aggcatcaaa ttaagcagaa ggccatcctg 1740
acggatggcc ttttgcgtt tctacaaact cttttgttta tttttctaaa tacattcaaa 1800
tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa 1860
gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 1920
tcctgttttt gctcacccag aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 1980
tgcacgagtg ggttacatcg aactggatct caacagcggt aagatccttg agagttttcg 2040
ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 2100
atcccgtgtt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 2160
cttggttgag tactcaccag tcacagaaaa gcatcttacg gatggcatga cagtaagaga 2220
attatgcagt gctgccataa ccatgagtga taacactgcg gccaacttac ttctgacaac 2280
gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 2340

ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 2400
gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 2460
agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 2520
gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 2580
gtctcgcggt atcattgcag cactggggcc agatggtaag ccctcccgta tcgtagttat 2640
ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 2700
tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 2760
tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 2820
catgaccaaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 2880
gatcaaagga tcttcttgag atccttttt tctgcgcgta atctgctgct tgcaaacaaa 2940
aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa gagctaccaa ctctttttcc 3000
gaaggtaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 3060
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatcct 3120
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttgg actcaagacg 3180
atagttaccg gataaggcgc agcggtcggg ctgaacgggg ggttcgtgca cacagcccag 3240
cttggagcga acgacctaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 3300
cacgcttccc gaagggagaa aggcggacag gtatccggta agcggcaggg tcggaacagg 3360
agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt 3420
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 3480
gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc ttttgctggc cttttgctca 3540
catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 3600
agctgatacc gctcgccgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 3660
ggaagagcgc ctgatgcggt attttctcct tacgcatctg tgcggtattt cacaccgcat 3720
atggtgcact ctcagtacaa tctgctctga tgccgcatag ttaagccagt atacactccg 3780
ctatcgctac gtgactgggt catggctgcg ccccgacacc cgccaacacc cgctgacgcg 3840
ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac cgtctccggg 3900
agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgaggca gcagatcaat 3960
tcgcgcgcga aggcgaagcg gcatgcataa tgtgcctgtc aaatggacga agcagggatt 4020
ctgcaaaccc tatgctactc cgtcaagccg tcaattgtct gattcgttac caattatgac 4080
aacttgacgg ctacatcatt cacttttct tcacaaccgg cacggaactc gctcgggctg 4140
gccccggtgc atttttaaa tacccgcgag aaatagagtt gatcgtcaaa accaacattg 4200
cgaccgacgg tggcgatagg catccgggtg gtgctcaaaa gcagcttcgc ctggctgata 4260
cgttggtcct cgcgccagct taagacgcta atccctaact gctggcggaa aagatgtgac 4320
agacgcgacg gcgacaagca aacatgctgt gcgacgctgg cgatatcaaa attgctgtct 4380
gccaggtgat cgctgatgta ctgacaagcc tcgcgtaccc gattatccat cggtggatgg 4440
agcgactcgt taatcgcttc catgcgccgc agtaacaatt gctcaagcag atttatcgcc 4500

agcagctccg aatagcgccc ttcccttgc ccggcgttaa tgatttgccc aaacaggtcg 4560

ctgaaatgcg gctggtgcgc ttcatccggg cgaaagaacc ccgtattggc aaatattgac 4620

ggccagttaa gccattcatg ccagtaggcg cgcggacgaa agtaaaccca ctggtgatac 4680

cattcgcgag cctccggatg acgaccgtag tgatgaatct ctcctggcgg gaacagcaaa 4740

atatcacccg gtcggcaaac aaattctcgt ccctgatttt tcaccacccc ctgaccgcga 4800

atggtgagat tgagaatata acctttcatt cccagcggtc ggtcgataaa aaaatcgaga 4860

taaccgttgg cctcaatcgg cgttaaaccc gccaccagat gggcattaaa cgagtatccc 4920

ggcagcaggg gatcattttg cgcttcagcc atacttttca tactcccgcc attcagag 4978

<210> 930
<211> 12319
<212> DNA
<213> Artificial Sequence

<220>
<223> pLH744

<400> 930
aggccagagg aaaataatat caagtgctgg aaacttttc tcttggaatt tttgcaacat 60

caagtcatag tcaattgaat tgacccaatt tcacatttaa gatttttttt ttttcatccg 120

acatacatct gtacactagg aagccctgtt tttctgaagc agcttcaaat atatatattt 180

tttacatatt tattatgatt caatgaacaa tctaattaaa tcgaaaacaa gaaccgaaac 240

gcgaataaat aatttattta gatggtgaca agtgtataag tcctcatcgg gacagctacg 300

atttctcttt cggttttggc tgagctactg gttgctgtga cgcagcggca ttagcgcggc 360

gttatgagct accctcgtgg cctgaaagat ggcgggaata aagcggaact aaaaattact 420

gactgagcca tattgaggtc aatttgtcaa ctcgtcaagt cacgtttggt ggacggcccc 480

tttccaacga atcgtatata ctaacatgcg cgcgcttcct atatacacat atacatatat 540

atatatatat atatgtgtgc gtgtatgtgt acacctgtat ttaatttcct tactcgcggg 600

tttttctttt ttctcaattc ttggcttcct ctttctcgag cggaccggat cctcgcgaac 660

tccaaaatga gctatcaaaa acgatagatc gattaggatg actttgaaat gactccgcag 720

tggactggcc gttaatttca agcgtgagta aaatagtgca tgacaaaaga tgagctaggc 780

ttttgtaaaa atatcttacg ttgtaaaatt ttagaaatca ttatttcctt catatcattt 840

tgtcattgac cttcagaaga aaagagccga ccaataatat aaataaataa ataaaaataa 900

tattccatta tttctaaaca gattcaatac tcattaaaaa actatatcaa ttaatttgaa 960

ttaacgcggc cgcttaacca cagcaaccag gacaacattt tttgccagtt tcttcaggct 1020

tccaaaagtc tgttacggct cccctagaag cagacgaaac gatgtgagca tatttaccaa 1080

ggataccgcg tgaatagagc ggtggcaatt caatggtctc ttgacgatgt tttaactctt 1140

catcggagat atcaaagtgt aattccttag tgtcttggtc aatagtgact atgtctcctg 1200

tttgcaggta ggcgattgga ccgccatctt gtgcttcagg agcgatatga cccacgacaa 1260

gaccataagt accacctgag aagcggccat ctgtcagaag ggcaactttt tcaccttgcc 1320

ctttaccaac aatcattgat gaaagggaaa gcatttcagg cataccagga ccgccctttg 1380

gtcctacaaa acgtacgaca acaacatcac catcaacaat atcatcattc aagacagctt 1440
caatggcttc ttcttcagaa ttaaagacct tagcaggacc gacatgacga cgcactttta 1500
caccagaaac tttggcaacg gcaccgtctg gagccaagtt accatggaga ataatgagcg 1560
gaccatcttc acgtttagga ttttcaagcg gcataataac cttttgacca ggtgttaaat 1620
catcaaaagc cttcaaattt tcagcgactg ttttgccagt acaagtgata cggtcaccat 1680
gaaggaagcc atttttaagg agatatttca taactgctgg tacccctccg accttgtaaa 1740
ggtcttggaa tacatattga ccagaaggtt tcaaatcagc caaatgagga acttttttctt 1800
ggaaagtatt gaaatcatca agtgtcaatt ccacattagc agcatgggca atagctaaga 1860
ggtgaagggt tgagttggtt gaacctccca gagccatagt tacagtaata gcatcttcaa 1920
aagcttcacg cgttaaaatg tcagaaggtt ttaagcccat ttcgagcatt ttgacaacag 1980
cgcgaccagc ttcttcaata tctgctttct tttctgcgga ttcagccggg tgagaagatg 2040
aacccggaag gctaagtccc aaaacttcaa tagctgtcgc cattgtgtta gcagtataca 2100
taccaccgca gcctccagga ccgggacaag cattacattc caaagcttta acttcttctt 2160
tggtcatatc gccgtggttc caatggccga caccttcaaa gacagagact aaatcgatat 2220
ctttgccgtc taaattacca ggtgcaattg ttccgccgta agcaaaaatg gctgggatat 2280
ccatgttagc catagcgata acagaaccgg gcatgttttt atcacaaccg ccaatggcta 2340
caaaagcatc cgcattatga cctcccatgg ctgcttcaat agaatctgca ataatatcac 2400
gagatgtcaa ggagaaacgc attccttggg ttcccatggc gattccatca gaaaccgtga 2460
ttgttccgaa ctgaactggc caagcaccag cttccttaac accgactttg gctagtttac 2520
caaagtcatg taagtggata ttacaaggtg tgttttcagc ccaagttgaa atgacaccga 2580
cgataggttt ttcaaagtct tcatcttgca taccagttgc acgcaacata gcacgattag 2640
gtgatttaac cattgaatcg taaacagaac tacgatttct taagtcttta agagtttttt 2700
tgtcagtcat actcacgtgg aacttagatt agattgctat gctttctctc taacgagcaa 2760
gaagtaaaaa aagttgtaat agaacaagaa aaatgaaact gaagcttgag aaattgaaga 2820
ccgtttatta gcttaaatat caatgggagg tcatcgaaag agaaaaaaat caagaaagaa 2880
actctcaaga aaaagaaacg tgataaaaat ttttattgcc tctctcgacg aagagaaaga 2940
aacgaggcgg tccctttttt cttttccaaa cctttagtac gggtaattag cgacacccta 3000
gaggaagaaa gaggggaaat ttagtatgct gtgcttgggt gtcttgaagt ggtacggcga 3060
tgcgcggagt ccgagaaaat ctggaagagt aaaaaggggg tagaagcgtt ttgaagctat 3120
gccggcagat ctatttaaat ggcgcgccga cgtcaggtgg cactttcgg ggaaatgtgc 3180
gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac 3240
aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt 3300
tccgtgtcgc ccttattccc ttttttgcgg cattttgcct tcctgttttt gctcacccag 3360
aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg 3420
aactggatct caacagcggt aagatccttg agagttttcg ccccgaagaa cgttttccaa 3480
tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc 3540

aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag tactcaccag 3600
tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa 3660
ccatgagtga taacactgcg gccaacttac ttctgacaac gatcggagga ccgaaggagc 3720
taaccgcttt tttgcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg 3780
agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa 3840
caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa 3900
tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg 3960
gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag 4020
cactggggcc agatggtaag ccctcccgta tcgtagttat ctacacgacg gggagtcagg 4080
caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt 4140
ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt 4200
aatttaaaag gatctaggtg aagatccttt ttgataatct catgaccaaa atcccttaac 4260
gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 4320
atccttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 4380
tggtttgttt gccggatcaa gagctaccaa ctctttttcc gaaggtaact ggcttcagca 4440
gagcgcagat accaaatact gttcttctag tgtagccgta gttaggccac cacttcaaga 4500
actctgtagc accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca 4560
gtggcgataa gtcgtgtctt accgggttgg actcaagacg atagttaccg gataaggcgc 4620
agcggtcggg ctgaacgggg ggttcgtgca cacagcccag cttggagcga acgacctaca 4680
ccgaactgag atacctacag cgtgagctat gagaaagcgc cacgcttccc gaagggagaa 4740
aggcggacag gtatccggta agcggcaggg tcggaacagg agagcgcacg agggagcttc 4800
caggggaaa cgcctggtat ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc 4860
gtcgattttt gtgatgctcg tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg 4920
cctttttacg gttcctggcc ttttgctggc cttttgctca catgttcttt cctgcgttat 4980
cccctgattc tgtggataac cgtattaccg cctttgagtg agctgatacc gctcgccgca 5040
gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca 5100
aaccgcctct ccccgcgcgt tggccgattc attaatgcag ctggcacgac aggtttcccg 5160
actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag ttagctcact cattaggcac 5220
cccaggcttt acactttatg cttccggctc gtatgttgtg tggaattgtg agcggataac 5280
aatttcacac aggaaacagc tatgaccatg attacgccaa gctttttctt tccaattttt 5340
ttttttcgt cattataaaa atcattacga ccgagattcc cgggtaataa ctgatataat 5400
taaattgaag ctctaatttg tgagtttagt atacatgcat ttacttataa tacagttttt 5460
tagttttgct ggccgcatct tctcaaatat gcttcccagc ctgcttttct gtaacgttca 5520
ccctctacct tagcatccct tccctttgca aatagtcctc ttccaacaat aataatgtca 5580
gatcctgtag agaccacatc atccacggtt ctatactgtt gacccaatgc gtctcccttg 5640
tcatctaaac ccacaccggg tgtcataatc aaccaatcgt aaccttcatc tcttccaccc 5700

atgtctcttt gagcaataaa gccgataaca aaatctttgt cgctcttcgc aatgtcaaca 5760
gtacccttag tatattctcc agtagatagg gagcccttgc atgacaattc tgctaacatc 5820
aaaaggcctc taggttcctt tgttacttct tctgccgcct gcttcaaacc gctaacaata 5880
cctgggccca ccacaccgtg tgcattcgta atgtctgccc attctgctat tctgtataca 5940
cccgcagagt actgcaattt gactgtatta ccaatgtcag caaattttct gtcttcgaag 6000
agtaaaaaat tgtacttggc ggataatgcc tttagcggct taactgtgcc ctccatggaa 6060
aaatcagtca agatatccac atgtgttttt agtaaacaaa ttttgggacc taatgcttca 6120
actaactcca gtaattcctt ggtggtacga acatccaatg aagcacacaa gtttgtttgc 6180
ttttcgtgca tgatattaaa tagcttggca gcaacaggac taggatgagt agcagcacgt 6240
tccttatatg tagctttcga catgatttat cttcgtttcc tgcaggtttt tgttctgtgc 6300
agttgggtta agaatactgg gcaatttcat gtttcttcaa cactacatat gcgtatatat 6360
accaatctaa gtctgtgctc cttccttcgt tcttccttct gttcggagat taccgaatca 6420
aaaaaatttc aaggaaaccg aaatcaaaaa aaagaataaa aaaaaaatga tgaattgaaa 6480
agcttgcatg cctgcaggtc gactctagta tactccgtct actgtacgat acacttccgc 6540
tcaggtcctt gtcctttaac gaggccttac cactcttttg ttactctatt gatccagctc 6600
agcaaaggca gtgtgatcta agattctatc ttcgcgatgt agtaaaacta gctagaccga 6660
gaaagagact agaaatgcaa aaggcacttc tacaatggct gccatcatta ttatccgatg 6720
tgacgctgca ttttttttt ttttttttt ttttttttt ttttttttt ttttttttt 6780
ttgtacaaat atcataaaaa aagagaatct ttttaagcaa ggattttctt aacttcttcg 6840
gcgacagcat caccgacttc ggtggtactg ttggaaccac ctaaatcacc agttctgata 6900
cctgcatcca aaacctttt aactgcatct tcaatggctt taccttcttc aggcaagttc 6960
aatgacaatt tcaacatcat tgcagcagac aagatagtgg cgatagggtt gaccttattc 7020
tttggcaaat ctggagcgga accatggcat ggttcgtaca aaccaaatgc ggtgttcttg 7080
tctggcaaag aggccaagga cgcagatggc aacaaaccca aggagcctgg gataacggag 7140
gcttcatcgg agatgatatc accaaacatg ttgctggtga ttataatacc atttaggtgg 7200
gttgggttct taactaggat catggcggca gaatcaatca attgatgttg aactttcaat 7260
gtagggaatt cgttcttgat ggtttcctcc acagtttttc tccataatct tgaagaggcc 7320
aaaacattag ctttatccaa ggaccaaata ggcaatggtg gctcatgttg tagggccatg 7380
aaagcggcca ttcttgtgat tctttgcact tctggaacgg tgtattgttc actatcccaa 7440
gcgacaccat caccatcgtc ttcctttctc ttaccaaagt aaatacctcc cactaattct 7500
ctaacaacaa cgaagtcagt acctttagca aattgtggct tgattggaga taagtctaaa 7560
agagagtcgg atgcaaagtt acatggtctt aagttggcgt acaattgaag ttctttacgg 7620
atttttagta aaccttgttc aggtctaaca ctaccggtac cccatttagg accacccaca 7680
gcacctaaca aaacggcatc agccttcttg gaggcttcca gcgcctcatc tggaagtgga 7740
acacctgtag catcgatagc agcaccacca attaaatgat tttcgaaatc gaacttgaca 7800
ttggaacgaa catcagaaat agctttaaga accttaatgg cttcggctgt gatttcttga 7860

ccaacgtggt cacctggcaa aacgacgatc ttcttagggg cagacattac aatggtatat 7920
ccttgaaata tatataaaaa aaaaaaaaaa aaaaaaaaaa aaaaatgcag cttctcaatg 7980
atattcgaat acgctttgag gagatacagc ctaatatccg acaaactgtt ttacagattt 8040
acgatcgtac ttgttaccca tcattgaatt ttgaacatcc gaacctggga gttttccctg 8100
aaacagatag tatatttgaa cctgtataat aatatatagt ctagcgcttt acggaagaca 8160
atgtatgtat ttcggttcct ggagaaacta ttgcatctat tgcataggta atcttgcacg 8220
tcgcatcccc ggttcatttt ctgcgtttcc atcttgcact tcaatagcat atctttgtta 8280
acgaagcatc tgtgcttcat tttgtagaac aaaaatgcaa cgcgagagcg ctaatttttc 8340
aaacaaagaa tctgagctgc atttttacag aacagaaatg caacgcgaaa gcgctatttt 8400
accaacgaag aatctgtgct tcatttttgt aaaacaaaaa tgcaacgcga gagcgctaat 8460
ttttcaaaca aagaatctga gctgcatttt tacagaacag aaatgcaacg cgagagcgct 8520
attttaccaa caaagaatct atacttcttt tttgttctac aaaaatgcat cccgagagcg 8580
ctatttttct aacaaagcat cttagattac tttttttctc ctttgtgcgc tctataatgc 8640
agtctcttga taacttttg cactgtaggt ccgttaaggt tagaagaagg ctactttggt 8700
gtctattttc tcttccataa aaaaagcctg actccacttc ccgcgtttac tgattactag 8760
cgaagctgcg ggtgcatttt ttcaagataa aggcatcccc gattatattc tataccgatg 8820
tggattgcgc atactttgtg aacagaaagt gatagcgttg atgattcttc attggtcaga 8880
aaattatgaa cggtttcttc tattttgtct ctatatacta cgtataggaa atgtttacat 8940
tttcgtattg ttttcgattc actctatgaa tagttcttac tacaattttt ttgtctaaag 9000
agtaatacta gagataaaca taaaaaatgt agaggtcgag tttagatgca agttcaagga 9060
gcgaaaggtg gatgggtagg ttatataggg atatagcaca gagatatata gcaaagagat 9120
acttttgagc aatgtttgtg gaagcggtat tcgcaatatt ttagtagctc gttacagtcc 9180
ggtgcgtttt tggttttttg aaagtgcgtc ttcagagcgc ttttggtttt caaaagcgct 9240
ctgaagttcc tatactttct agagaatagg aacttcggaa taggaacttc aaagcgtttc 9300
cgaaaacgag cgcttccgaa aatgcaacgc gagctgcgca catacagctc actgttcacg 9360
tcgcacctat atctgcgtgt tgcctgtata tatatataca tgagaagaac ggcatagtgc 9420
gtgtttatgc ttaaatgcgt acttatatgc gtctatttat gtaggatgaa aggtagtcta 9480
gtacctcctg tgatattatc ccattccatg cggggtatcg tatgcttcct tcagcactac 9540
cctttagctg ttctatatgc tgccactcct caattggatt agtctcatcc ttcaatgcta 9600
tcatttcctt tgatattgga tcatatgcat agtaccgaga aactagagga tctcccatta 9660
ccgacatttg ggcgctatac gtgcatatgt tcatgtatgt atctgtattt aaaacacttt 9720
tgtattattt ttcctcatat atgtgtatag gtttatacgg atgatttaat tattacttca 9780
ccacccttta tttcaggctg atatcttagc cttgttacta gattaatcat gtaattagtt 9840
atgtcacgct tacattcacg ccctcccccc acatccgctc taaccgaaaa ggaaggagtt 9900
agacaacctg aagtctaggt ccctatttat ttttttatag ttatgttagt attaagaacg 9960
ttatttatat ttcaaatttt tctttttttt ctgtacagac gcgtgtacgc atgtaacatt 10020

atactgaaaa ccttgcttga gaaggttttg ggacgctcga aggctttaat ttgcgggcgg 10080
ccgcacctgg taaaacctct agtggagtag tagatgtaat caatgaagcg gaagccaaaa 10140
gaccagagta gaggcctata gaagaaactg cgataccttt tgtgatggct aaacaaacag 10200
acatcttttt atatgttttt acttctgtat atcgtgaagt agtaagtgat aagcgaattt 10260
ggctaagaac gttgtaagtg aacaagggac ctcttttgcc tttcaaaaaa ggattaaatg 10320
gagttaatca ttgagattta gttttcgtta gattctgtat ccctaaataa ctcccttacc 10380
cgacgggaag gcacaaaaga cttgaataat agcaaacggc cagtagccaa gaccaaataa 10440
tactagagtt aactgatggt cttaaacagg cattacgtgg tgaactccaa gaccaatata 10500
caaaatatcg ataagttatt cttgcccacc aatttaagga gcctacatca ggacagtagt 10560
accattcctc agagaagagg tatacataac aagaaaatcg cgtgaacacc ttatataact 10620
tagcccgtta ttgagctaaa aaaccttgca aaatttccta tgaataagaa tacttcagac 10680
gtgataaaaa tttactttct aactcttctc acgctgcccc tatctgttct tccgctctac 10740
cgtgagaaat aaagcatcga gtacggcagt tcgctgtcac tgaactaaaa caataaggct 10800
agttcgaatg atgaacttgc ttgctgtcaa acttctgagt tgccgctgat gtgacactgt 10860
gacaataaat tcaaaccggt tatagcggtc tcctccggta ccggttctgc cacctccaat 10920
agagctcagt aggagtcaga acctctgcgg tggctgtcag tgactcatcc gcgtttcgta 10980
agttgtgcgc gtgcacattt cgcccgttcc cgctcatctt gcagcaggcg gaaattttca 11040
tcacgctgta ggacgcaaaa aaaaaataat taatcgtaca agaatcttgg aaaaaaaatt 11100
gaaaaatttt gtataaaagg gatgacctaa cttgactcaa tggcttttac acccagtatt 11160
ttccctttcc ttgtttgtta caattataga agcaagacaa aaacatatag acaacctatt 11220
cctaggagtt atattttttt accctaccag caatataagt aaaaaactgt ttaaacagta 11280
tggaagaatg taagatggct aagatttact accaagaaga ctgtaacttg tccttgttgg 11340
atggtaagac tatcgccgtt atcggttacg gttctcaagg tcacgctcat gccctgaatg 11400
ctaaggaatc cggttgtaac gttatcattg gtttattcga aggtgcggag gagtggaaaa 11460
gagctgaaga acaaggtttc gaagtctaca ccgctgctga agctgctaag aaggctgaca 11520
tcattatgat cttgatccca gatgaaaagc aggctaccat gtacaaaaac gacatcgaac 11580
caaacttgga agccggtaac atgttgatgt tcgctcacgg tttcaacatc catttcggtt 11640
gtattgttcc accaaaggac gttgatgtca ctatgatcgc tccaaagggt ccaggtcaca 11700
ccgttagatc cgaatacgaa gaaggtaaag gtgtcccatg cttggttgct gtcgaacaag 11760
acgctactgg caaggctttg gatatggctt tggcctacgc tttagccatc ggtggtgcta 11820
gagccggtgt cttggaaact accttcagaa ccgaaactga aaccgacttg ttcggtgaac 11880
aagctgtttt atgtggtggt gtctgcgctt tgatgcaggc cggttttgaa accttggttg 11940
aagccggtta cgacccaaga aacgcttact tcgaatgtat ccacgaaatg aagttgatcg 12000
ttgacttgat ctaccaatct ggtttctccg gtatgcgtta ctctatctcc aacactgctg 12060
aatacggtga ctacattacc ggtccaaaga tcattactga agataccaag aaggctatga 12120
agaagatttt gtctgacatt caagatggta cctttgccaa ggacttcttg gttgacatgt 12180

ctgatgctgg ttcccaggtc cacttcaagg ctatgagaaa gttggcctcc gaacacccag 12240

ctgaagttgt cggtgaagaa attagatcct tgtactcctg gtccgacgaa gacaagttga 12300

ttaacaactg aggccctgc 12319

<210> 931
<211> 12304
<212> DNA
<213> Artificial Sequence

<220>
<223> K9SB2_SH_DHAD

<400> 931
tcccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60

aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120

ttacttcacc acccttttatt tcaggctgat atcttagcct tgttactaga ttaatcatgt 180

aattagttat gtcacgctta cattcacgcc ctccccccac atccgctcta accgaaaagg 240

aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 300

taagaacgtt atttatattt caaatttttc ttttttttct gtacagacgc gtgtacgcat 360

gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt 420

gcgggcggcc gcacctggta aaacctctag tggagtagta gatgtaatca atgaagcgga 480

agccaaaaga ccagagtaga ggcctataga agaaactgcg ataccttttg tgatggctaa 540

acaaacagac atctttttat atgttttttac ttctgtatat cgtgaagtag taagtgataa 600

gcgaatttgg ctaagaacgt tgtaagtgaa caagggacct cttttgcctt tcaaaaaagg 660

attaaatgga gttaatcatt gagatttagt tttcgttaga ttctgtatcc ctaaataact 720

cccttacccg acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga 780

ccaaataata ctagagttaa ctgatggtct taaacaggca ttacgtggtg aactccaaga 840

ccaatataca aaatatcgat aagttattct tgcccaccaa tttaaggagc ctacatcagg 900

acagtagtac cattcctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt 960

atataactta gcccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata 1020

cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgccccta tctgttcttc 1080

cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca 1140

ataaggctag ttcgaatgat gaacttgctt gctgtcaaac ttctgagttg ccgctgatgt 1200

gacactgtga caataaattc aaaccggtta tagcggtctc ctccggtacc ggttctgcca 1260

cctccaatag agctcagtag gagtcagaac ctctgcggtg gctgtcagtg actcatccgc 1320

gtttcgtaag ttgtgcgcgt gcacatttcg cccgttcccg ctcatcttgc agcaggcgga 1380

aattttcatc acgctgtagg acgcaaaaaa aaaataatta atcgtacaag aatcttggaa 1440

aaaaaattga aaaattttgt ataaaaggga tgacctaact tgactcaatg gcttttacac 1500

ccagtatttt ccctttcctt gtttgttaca attatagaag caagacaaaa acatatagac 1560

aacctattcc taggagttat atttttttac cctaccagca atataagtaa aaaactgttt 1620

aaacagtatg gctaagattt actaccaaga agactgtaac ttgtccttgt tggatggtaa 1680

gactatcgcc gttatcggtt acggttctca aggtcacgct catgccctga atgctaagga 1740
atccggttgt aacgttatca ttggtttatt cgaaggtgcg gaggagtgga aaagagctga 1800
agaacaaggt ttcgaagtct acaccgctgc tgaagctgct aagaaggctg acatcattat 1860
gatcttgatc ccagatgaaa agcaggctac catgtacaaa aacgacatcg aaccaaactt 1920
ggaagccggt aacatgttga tgttcgctca cggtttcaac atccatttcg gttgtattgt 1980
tccaccaaag gacgttgatg tcactatgat cgctccaaag ggtccaggtc acaccgttag 2040
atccgaatac gaagaaggta aaggtgtccc atgcttggtt gctgtcgaac aagacgctac 2100
tggcaaggct ttggatatgg ctttggccta cgctttagcc atcggtggtg ctagagccgg 2160
tgtcttggaa actaccttca gaaccgaaac tgaaaccgac ttgttcggtg aacaagctgt 2220
tttatgtggt ggtgtctgcg ctttgatgca ggccggtttt gaaaccttgg ttgaagccgg 2280
ttacgaccca agaaacgctt acttcgaatg tatccacgaa atgaagttga tcgttgactt 2340
gatctaccaa tctggtttct ccggtatgcg ttactctatc tccaacactg ctgaatacgg 2400
tgactacatt accggtccaa agatcattac tgaagatacc aagaaggcta tgaagaagat 2460
tttgtctgac attcaagatg gtacctttgc caaggacttc ttggttgaca tgtctgatgc 2520
tggttcccag gtccacttca aggctatgag aaagttggcc tccgaacacc cagctgaagt 2580
tgtcggtgaa gaaattagat ccttgtactc ctggtccgac gaagacaagt tgattaacaa 2640
ctgaggccct gcaggccaga ggaaaataat atcaagtgct ggaaactttt tctcttggaa 2700
tttttgcaac atcaagtcat agtcaattga attgacccaa tttcacattt aagatttttt 2760
ttttttcatc cgacatacat ctgtacacta ggaagccctg tttttctgaa gcagcttcaa 2820
atatatatat tttttacata tttattatga ttcaatgaac aatctaatta aatcgaaaac 2880
aagaaccgaa acgcgaataa ataatttatt tagatggtga caagtgtata agtcctcatc 2940
gggacagcta cgatttctct ttcggttttg gctgagctac tggttgctgt gacgcagcgg 3000
cattagcgcg gcgttatgag ctaccctcgt ggcctgaaag atggcgggaa taaagcggaa 3060
ctaaaaatta ctgactgagc catattgagg tcaatttgtc aactcgtcaa gtcacgtttg 3120
gtggacggcc cctttccaac gaatcgtata tactaacatg cgcgcgcttc ctatatacac 3180
atatacatat atatatatat atatatgtgt gcgtgtatgt gtacacctgt atttaatttc 3240
cttactcgcg ggtttttctt ttttctcaat tcttggcttc ctctttctcg agcggaccgg 3300
atcctcgcga actccaaaat gagctatcaa aaacgataga tcgattagga tgactttgaa 3360
atgactccgc agtggactgg ccgttaattt caagcgtgag taaaatagtg catgacaaaa 3420
gatgagctag gcttttgtaa aaatatctta cgttgtaaaa ttttagaaat cattatttcc 3480
ttcatatcat tttgtcattg accttcagaa gaaaagagcc gaccaataat ataaataaat 3540
aaataaaaat aatattccat tatttctaaa cagattcaat actcattaaa aaactatatc 3600
aattaatttg aattaacgcg gccgcttaac cacagcaacc aggacaacat tttttgccag 3660
tttcttcagg cttccaaaag tctgttacgg ctcccctaga agcagacgaa acgatgtgag 3720
catatttacc aaggataccg cgtgaataga gcggtggcaa ttcaatggtc tcttgacgat 3780
gttttaactc ttcatcggag atatcaaagt gtaattcctt agtgtcttgg tcaatagtga 3840

ctatgtctcc tgtttgcagg taggcgattg gaccgccatc ttgtgcttca ggagcgatat 3900
gacccacgac aagaccataa gtaccacctg agaagcggcc atctgtcaga agggcaactt 3960
tttcaccttg ccctttacca acaatcattg atgaaaggga aagcatttca ggcataccag 4020
gaccgccctt tggtcctaca aaacgtacga caacaacatc accatcaaca atatcatcat 4080
tcaagacagc ttcaatggct tcttcttcag aattaaagac cttagcagga ccgacatgac 4140
gacgcacttt tacaccagaa actttggcaa cggcaccgtc tggagccaag ttaccatgga 4200
gaataatgag cggaccatct tcacgtttag gattttcaag cggcataata acctttgac 4260
caggtgttaa atcatcaaaa gccttcaaat tttcagcgac tgttttgcca gtacaagtga 4320
tacggtcacc atgaaggaag ccatttttaa ggagatattt cataactgct ggtacccctc 4380
cgaccttgta aaggtcttgg aatacatatt gaccagaagg tttcaaatca gccaaatgag 4440
gaacttttc ttggaaagta ttgaaatcat caagtgtcaa ttccacatta gcagcatggg 4500
caatagctaa gaggtgaagg gttgagttgg ttgaacctcc cagagccata gttacagtaa 4560
tagcatcttc aaaagcttca cgcgttaaaa tgtcagaagg ttttaagccc atttcgagca 4620
ttttgacaac agcgcgacca gcttcttcaa tatctgcttt cttttctgcg gattcagccg 4680
ggtgagaaga tgaacccgga aggctaagtc ccaaaacttc aatagctgtc gccattgtgt 4740
tagcagtata cataccaccg cagcctccag gaccgggaca agcattacat tccaaagctt 4800
taacttcttc tttggtcata tcgccgtggt tccaatggcc gacaccttca aagacagaga 4860
ctaaatcgat atctttgccg tctaaattac caggtgcaat tgttccgccg taagcaaaaa 4920
tggctgggat atccatgtta gccatagcga taacagaacc gggcatgttt ttatcacaac 4980
cgccaatggc tacaaaagca tccgcattat gacctcccat ggctgcttca atagaatctg 5040
caataatatc acgagatgtc aaggagaaac gcattccttg ggttcccatg gcgattccat 5100
cagaaaccgt gattgttccg aactgaactg gccaagcacc agcttcctta acaccgactt 5160
tggctagttt accaaagtca tgtaagtgga tattacaagg tgtgttttca gcccaagttg 5220
aaatgacacc gacgataggt ttttcaaagt cttcatcttg cataccagtt gcacgcaaca 5280
tagcacgatt aggtgattta accattgaat cgtaaacaga actacgattt cttaagtctt 5340
taagagtttt tttgtcagtc atactcacgt ggaacttaga ttagattgct atgctttctc 5400
tctaacgagc aagaagtaaa aaaagttgta atagaacaag aaaaatgaaa ctgaagcttg 5460
agaaattgaa gaccgtttat tagcttaaat atcaatggga ggtcatcgaa agagaaaaaa 5520
atcaagaaag aaactctcaa gaaaaagaaa cgtgataaaa atttttattg cctctctcga 5580
cgaagagaaa gaaacgaggc ggtccctttt ttcttttcca aacctttagt acgggtaatt 5640
agcgacaccc tagaggaaga aagaggggaa atttagtatg ctgtgcttgg gtgtcttgaa 5700
gtggtacggc gatgcgcgga gtccgagaaa atctggaaga gtaaaaaggg ggtagaagcg 5760
ttttgaagct atgccggcag atctatttaa atggcgcgcc gacgtcaggt ggcacttttc 5820
ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca aatatgtatc 5880
cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg aagagtatga 5940
gtattcaaca tttccgtgtc gcccttattc cctttttgc ggcattttgc cttcctgttt 6000

ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag 6060
tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgccccgaag 6120
aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta ttatcccgta 6180
ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat gacttggttg 6240
agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga gaattatgca 6300
gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca acgatcggag 6360
gaccgaagga gctaaccgct tttttgcaca acatggggga tcatgtaact cgccttgatc 6420
gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc acgatgcctg 6480
tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact ctagcttccc 6540
ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt ctgcgctcgg 6600
cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt gggtctcgcg 6660
gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt atctacacga 6720
cggggagtca ggcaactatg gatgaacgaa atagacagat cgctgagata ggtgcctcac 6780
tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag attgatttaa 6840
aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat ctcatgacca 6900
aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa aagatcaaag 6960
gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca aaaaaaccac 7020
cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt ccgaaggtaa 7080
ctggcttcag cagagcgcag ataccaaata ctgttcttct agtgtagccg tagttaggcc 7140
accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc ctgttaccag 7200
tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga cgatagttac 7260
cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc agcttggagc 7320
gaacgaccta caccgaactg agatacctac agcgtgagct atgagaaagc gccacgcttc 7380
ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca ggagagcgca 7440
cgagggagct tccaggggga aacgcctggt atctttatag tcctgtcggg tttcgccacc 7500
tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gcggagccta tggaaaaacg 7560
ccagcaacgc ggcctttta cggttcctgg ccttttgctg gccttttgct cacatgttct 7620
ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag tgagctgata 7680
ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa gcggaagagc 7740
gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc agctggcacg 7800
acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg agttagctca 7860
ctcattaggc accccaggct ttacacttta tgcttccggc tcgtatgttg tgtggaattg 7920
tgagcggata acaatttcac acaggaaaca gctatgacca tgattacgcc aagcttttc 7980
tttccaattt ttttttttc gtcattataa aaatcattac gaccgagatt cccgggtaat 8040
aactgatata attaaattga agctctaatt tgtgagttta gtatacatgc atttacttat 8100
aatacagttt tttagttttg ctggccgcat cttctcaaat atgcttccca gcctgctttt 8160

20 Jan 2017 2017200439

ctgtaacgtt caccctctac cttagcatcc cttccctttg caaatagtcc tcttccaaca 8220
ataataatgt cagatcctgt agagaccaca tcatccacgg ttctatactg ttgacccaat 8280
gcgtctccct tgtcatctaa acccacaccg ggtgtcataa tcaaccaatc gtaaccttca 8340
tctcttccac ccatgtctct ttgagcaata aagccgataa caaaatcttt gtcgctcttc 8400
gcaatgtcaa cagtaccctt agtatattct ccagtagata gggagccctt gcatgacaat 8460
tctgctaaca tcaaaaggcc tctaggttcc tttgttactt cttctgccgc ctgcttcaaa 8520
ccgctaacaa tacctgggcc caccacaccg tgtgcattcg taatgtctgc ccattctgct 8580
attctgtata cacccgcaga gtactgcaat ttgactgtat taccaatgtc agcaaatttt 8640
ctgtcttcga agagtaaaaa attgtacttg gcggataatg cctttagcgg cttaactgtg 8700
ccctccatgg aaaaatcagt caagatatcc acatgtgttt ttagtaaaca aattttggga 8760
cctaatgctt caactaactc cagtaattcc ttggtggtac gaacatccaa tgaagcacac 8820
aagtttgttt gcttttcgtg catgatatta aatagcttgg cagcaacagg actaggatga 8880
gtagcagcac gttccttata tgtagctttc gacatgattt atcttcgttt cctgcaggtt 8940
tttgttctgt gcagttgggt taagaatact gggcaatttc atgtttcttc aacactacat 9000
atgcgtatat ataccaatct aagtctgtgc tccttccttc gttcttcctt ctgttcggag 9060
attaccgaat caaaaaaatt tcaaggaaac cgaaatcaaa aaaaagaata aaaaaaaaat 9120
gatgaattga aaagcttgca tgcctgcagg tcgactctag tatactccgt ctactgtacg 9180
atacacttcc gctcaggtcc ttgtccttta acgaggcctt accactcttt tgttactcta 9240
ttgatccagc tcagcaaagg cagtgtgatc taagattcta tcttcgcgat gtagtaaaac 9300
tagctagacc gagaaagaga ctagaaatgc aaaaggcact tctacaatgg ctgccatcat 9360
tattatccga tgtgacgctg catttttttt ttttttttt ttttttttt ttttttttt 9420
ttttttttt ttttgtacaa atatcataaa aaaagagaat ctttttaagc aaggattttc 9480
ttaacttctt cggcgacagc atcaccgact tcggtggtac tgttggaacc acctaaatca 9540
ccagttctga tacctgcatc caaaaccttt ttaactgcat cttcaatggc tttaccttct 9600
tcaggcaagt tcaatgacaa tttcaacatc attgcagcag acaagatagt ggcgataggg 9660
ttgaccttat tctttggcaa atctggagcg gaaccatggc atggttcgta caaaccaaat 9720
gcggtgttct tgtctggcaa agaggccaag gacgcagatg gcaacaaacc caaggagcct 9780
gggataacgg aggcttcatc ggagatgata tcaccaaaca tgttgctggt gattataata 9840
ccatttaggt gggttgggtt cttaactagg atcatggcgg cagaatcaat caattgatgt 9900
tgaactttca atgtagggaa ttcgttcttg atggtttcct ccacagtttt tctccataat 9960
cttgaagagg ccaaaacatt agctttatcc aaggaccaaa taggcaatgg tggctcatgt 10020
tgtagggcca tgaaagcggc cattcttgtg attctttgca cttctggaac ggtgtattgt 10080
tcactatccc aagcgacacc atcaccatcg tcttcctttc tcttaccaaa gtaaatacct 10140
cccactaatt ctctaacaac aacgaagtca gtacctttag caaattgtgg cttgattgga 10200
gataagtcta aaagagagtc ggatgcaaag ttacatggtc ttaagttggc gtacaattga 10260
agttctttac ggatttttag taaaccttgt tcaggtctaa cactaccggt accccattta 10320

ggaccaccca cagcacctaa caaaacggca tcagccttct tggaggcttc cagcgcctca 10380
tctggaagtg gaacacctgt agcatcgata gcagcaccac caattaaatg attttcgaaa 10440
tcgaacttga cattggaacg aacatcagaa atagctttaa gaaccttaat ggcttcggct 10500
gtgatttctt gaccaacgtg gtcacctggc aaaacgacga tcttcttagg ggcagacatt 10560
acaatggtat atccttgaaa tatatataaa aaaaaaaaaa aaaaaaaaaa aaaaaaatgc 10620
agcttctcaa tgatattcga atacgctttg aggagataca gcctaatatc cgacaaactg 10680
ttttacagat ttacgatcgt acttgttacc catcattgaa ttttgaacat ccgaacctgg 10740
gagttttccc tgaaacagat agtatatttg aacctgtata ataatatata gtctagcgct 10800
ttacggaaga caatgtatgt atttcggttc ctggagaaac tattgcatct attgcatagg 10860
taatcttgca cgtcgcatcc ccggttcatt ttctgcgttt ccatcttgca cttcaatagc 10920
atatctttgt taacgaagca tctgtgcttc attttgtaga acaaaaatgc aacgcgagag 10980
cgctaatttt tcaaacaaag aatctgagct gcatttttac agaacagaaa tgcaacgcga 11040
aagcgctatt ttaccaacga agaatctgtg cttcattttt gtaaaacaaa aatgcaacgc 11100
gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa 11160
cgcgagagcg ctattttacc aacaaagaat ctatacttct tttttgttct acaaaaatgc 11220
atcccgagag cgctattttt ctaacaaagc atcttagatt actttttttc tcctttgtgc 11280
gctctataat gcagtctctt gataactttt tgcactgtag gtccgttaag gttagaagaa 11340
ggctactttg gtgtctattt tctcttccat aaaaaaagcc tgactccact tcccgcgttt 11400
actgattact agcgaagctg cgggtgcatt ttttcaagat aaaggcatcc ccgattatat 11460
tctataccga tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct 11520
tcattggtca gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg 11580
aaatgtttac attttcgtat tgttttcgat tcactctatg aatagttctt actacaattt 11640
ttttgtctaa agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg 11700
caagttcaag gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata 11760
tagcaaagag atacttttga gcaatgtttg tggaagcggt attcgcaata ttttagtagc 11820
tcgttacagt ccggtgcgtt tttggttttt tgaaagtgcg tcttcagagc gcttttggtt 11880
ttcaaaagcg ctctgaagtt cctatacttt ctagagaata ggaacttcgg aataggaact 11940
tcaaagcgtt tccgaaaacg agcgcttccg aaaatgcaac gcgagctgcg cacatacagc 12000
tcactgttca cgtcgcacct atatctgcgt gttgcctgta tatatatata catgagaaga 12060
acggcatagt gcgtgtttat gcttaaatgc gtacttatat gcgtctattt atgtaggatg 12120
aaaggtagtc tagtacctcc tgtgatatta tcccattcca tgcggggtat cgtatgcttc 12180
cttcagcact acccttttagc tgttctatat gctgccactc ctcaattgga ttagtctcat 12240
ccttcaatgc tatcatttcc tttgatattg gatcatatgc atagtaccga gaaactagag 12300
gatc 12304

<210> 932
<211> 9612

<212> DNA
<213> Artificial Sequence

<220>
<223> K9SB2_SH_81

<400> 932
catcatcaca gtttaaacag tatggctaag atttactacc aagaagactg taacttgtcc 60

ttgttggatg gtaagactat cgccgttatc ggttacggtt ctcaaggtca cgctcatgcc 120

ctgaatgcta aggaatccgg ttgtaacgtt atcattggtt tattcgaagg tgcggaggag 180

tggaaaagag ctgaagaaca aggtttcgaa gtctacaccg ctgctgaagc tgctaagaag 240

gctgacatca ttatgatctt gatcccagat gaaaagcagg ctaccatgta caaaaacgac 300

atcgaaccaa acttggaagc cggtaacatg ttgatgttcg ctcacggttt caacatccat 360

ttcggttgta ttgttccacc aaaggacgtt gatgtcacta tgatcgctcc aaagggtcca 420

ggtcacaccg ttagatccga atacgaagaa ggtaaaggtg tcccatgctt ggttgctgtc 480

gaacaagacg ctactggcaa ggctttggat atggctttgg cctacgcttt agccatcggt 540

ggtgctagag ccggtgtctt ggaaactacc ttcagaaccg aaactgaaac cgacttgttc 600

ggtgaacaag ctgttttatg tggtggtgtc tgcgctttga tgcaggccgg ttttgaaacc 660

ttggttgaag ccggttacga cccaagaaac gcttacttcg aatgtatcca cgaaatgaag 720

ttgatcgttg acttgatcta ccaatctggt ttctccggta tgcgttactc tatctccaac 780

actgctgaat acggtgacta cattaccggt ccaaagatca ttactgaaga taccaagaag 840

gctatgaaga agattttgtc tgacattcaa gatggtacct ttgccaagga cttcttggtt 900

gacatgtctg atgctggttc ccaggtccac ttcaaggcta tgagaaagtt ggcctccgaa 960

cacccagctg aagttgtcgg tgaagaaatt agatccttgt actcctggtc cgacgaagac 1020

aagttgatta acaactgagg ccctgcaggc cagaggaaaa taatatcaag tgctggaaac 1080

tttttctctt ggaatttttg caacatcaag tcatagtcaa ttgaattgac ccaatttcac 1140

atttaagatt ttttttttt catccgacat acatctgtac actaggaagc cctgtttttc 1200

tgaagcagct tcaaatatat atatttttta catatttatt atgattcaat gaacaatcta 1260

attaaatcga aaacaagaac cgaaacgcga ataaataatt tatttagatg gtgacaagtg 1320

tataagtcct catcgggaca gctacgattt ctctttcggt tttggctgag ctactggttg 1380

ctgtgacgca gcggcattag cgcggcgtta tgagctaccc tcgtggcctg aaagatggcg 1440

ggaataaagc ggaactaaaa attactgact gagccatatt gaggtcaatt tgtcaactcg 1500

tcaagtcacg tttggtggac ggccccttc caacgaatcg tatatactaa catgcgcgcg 1560

cttcctatat acacatatac atatatatat atatatatat gtgtgcgtgt atgtgtacac 1620

ctgtatttaa tttccttact cgcgggtttt tctttttttct caattcttgg cttcctcttt 1680

ctcgagcgga ccggatcctc cgcggtgccg gcagatctat ttaaatggcg cgccgacgtc 1740

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 1800

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 1860

aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 1920

ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 1980

gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 2040
ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 2100
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 2160
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 2220
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 2280
gacaacgatc ggaggaccga aggagctaac cgcttttttg cacaacatgg gggatcatgt 2340
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 2400
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 2460
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 2520
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga 2580
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 2640
agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac agatcgctga 2700
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 2760
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga 2820
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 2880
agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 2940
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 3000
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgttc ttctagtgta 3060
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 3120
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 3180
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 3240
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga 3300
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 3360
aacaggagag cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt 3420
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag 3480
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 3540
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 3600
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga 3660
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 3720
atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 3780
tgtgagttag ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat 3840
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 3900
cgccaagctt tttctttcca attttttttt tttcgtcatt ataaaaatca ttacgaccga 3960
gattcccggg taataactga tataattaaa ttgaagctct aatttgtgag tttagtatac 4020
atgcatttac ttataataca gtttttttagt tttgctggcc gcatcttctc aaatatgctt 4080
cccagcctgc ttttctgtaa cgttcaccct ctaccttagc atcccttccc tttgcaaata 4140

gtcctcttcc aacaataata atgtcagatc ctgtagagac cacatcatcc acggttctat 4200
actgttgacc caatgcgtct cccttgtcat ctaaacccac accgggtgtc ataatcaacc 4260
aatcgtaacc ttcatctctt ccacccatgt ctctttgagc aataaagccg ataacaaaat 4320
ctttgtcgct cttcgcaatg tcaacagtac ccttagtata ttctccagta gatagggagc 4380
ccttgcatga caattctgct aacatcaaaa ggcctctagg ttcctttgtt acttcttctg 4440
ccgcctgctt caaaccgcta acaatacctg ggcccaccac accgtgtgca ttcgtaatgt 4500
ctgcccattc tgctattctg tatacacccg cagagtactg caatttgact gtattaccaa 4560
tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta cttggcggat aatgccttta 4620
gcggcttaac tgtgccctcc atggaaaaat cagtcaagat atccacatgt gtttttagta 4680
aacaaatttt gggacctaat gcttcaacta actccagtaa ttccttggtg gtacgaacat 4740
ccaatgaagc acacaagttt gtttgctttt cgtgcatgat attaaatagc ttggcagcaa 4800
caggactagg atgagtagca gcacgttcct tatatgtagc tttcgacatg atttatcttc 4860
gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa tactgggcaa tttcatgttt 4920
cttcaacact acatatgcgt atatatacca atctaagtct gtgctccttc cttcgttctt 4980
ccttctgttc ggagattacc gaatcaaaaa aatttcaagg aaaccgaaat caaaaaaaag 5040
aataaaaaaa aaatgatgaa ttgaaaagct tgcatgcctg caggtcgact ctagtatact 5100
ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact 5160
cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatcttcg 5220
cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 5280
atggctgcca tcattattat ccgatgtgac gctgcatttt tttttttttt tttttttttt 5340
tttttttttt tttttttttt tttttttgt acaaatatca taaaaaaaga gaatcttttt 5400
aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 5460
aaccacctaa atcaccagtt ctgatacctg catccaaaac ctttttaact gcatcttcaa 5520
tggctttacc ttcttcaggc aagttcaatg acaatttcaa catcattgca gcagacaaga 5580
tagtggcgat agggttgacc ttattctttg gcaaatctgg agcggaacca tggcatggtt 5640
cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 5700
aacccaagga gcctgggata acggaggctt catcggagat gatatcacca aacatgttgc 5760
tggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat 5820
caatcaattg atgttgaact ttcaatgtag ggaattcgtt cttgatggtt tcctccacag 5880
tttttctcca taatcttgaa gaggccaaaa cattagcttt atccaaggac caaataggca 5940
atggtggctc atgttgtagg gccatgaaag cggccattct tgtgattctt tgcacttctg 6000
gaacggtgta ttgttcacta tcccaagcga caccatcacc atcgtcttcc tttctcttac 6060
caaagtaaat acctcccact aattctctaa caacaacgaa gtcagtacct ttagcaaatt 6120
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 6180
tggcgtacaa ttgaagttct ttacggattt ttagtaaacc ttgttcaggt ctaacactac 6240
cggtacccca tttaggacca cccacagcac ctaacaaaac ggcatcagcc ttcttggagg 6300

cttccagcgc ctcatctgga agtggaacac ctgtagcatc gatagcagca ccaccaatta 6360
aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 6420
taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 6480
taggggcaga cattacaatg gtatatcctt gaaatatata taaaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa atgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa 6600
tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat tgaattttga 6660
acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata 6720
tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc 6780
atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct 6840
tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa 6900
atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 6960
gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa 7020
caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg cattttttaca 7080
gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac ttctttttg 7140
ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta gattactttt 7200
tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact gtaggtccgt 7260
taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa agcctgactc 7320
cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttca agataaaggc 7380
atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 7440
gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt ttgtctctat 7500
atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc tatgaatagt 7560
tcttactaca attttttttgt ctaaagagta atactagaga taaacataaa aaatgtagag 7620
gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat atagggatat 7680
agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaag cggtattcgc 7740
aatattttag tagctcgtta cagtccggtg cgtttttggt tttttgaaag tgcgtcttca 7800
gagcgctttt ggttttcaaa agcgctctga agttcctata ctttctagag aataggaact 7860
tcggaatagg aacttcaaag cgtttccgaa aacgagcgct tccgaaaatg caacgcgagc 7920
tgcgcacata cagctcactg ttcacgtcgc acctatatct gcgtgttgcc tgtatatata 7980
tatacatgag aagaacggca tagtgcgtgt ttatgcttaa atgcgtactt atatgcgtct 8040
atttatgtag gatgaaaggt agtctagtac ctcctgtgat attatcccat tccatgcggg 8100
gtatcgtatg cttccttcag cactaccctt tagctgttct atatgctgcc actcctcaat 8160
tggattagtc tcatccttca atgctatcat ttcctttgat attggatcat atgcatagta 8220
ccgagaaact agaggatctc ccattaccga catttgggcg ctatacgtgc atatgttcat 8280
gtatgtatct gtatttaaaa cacttttgta ttatttttcc tcatatatgt gtataggttt 8340
atacggatga tttaattatt acttcaccac cctttatttc aggctgatat cttagccttg 8400
ttactagtca ccggtggcgg ccgcacctgg taaaacctct agtggagtag tagatgtaat 8460

caatgaagcg gaagccaaaa gaccagagta gaggcctata gaagaaactg cgataccttt 8520
tgtgatggct aaacaaacag acatcttttt atatgttttt acttctgtat atcgtgaagt 8580
agtaagtgat aagcgaattt ggctaagaac gttgtaagtg aacaagggac ctcttttgcc 8640
tttcaaaaaa ggattaaatg gagttaatca ttgagattta gttttcgtta gattctgtat 8700
ccctaaataa ctcccttacc cgacgggaag gcacaaaaga cttgaataat agcaaacggc 8760
cagtagccaa gaccaaataa tactagagtt aactgatggt cttaaacagg cattacgtgg 8820
tgaactccaa gaccaatata caaaatatcg ataagttatt cttgcccacc aatttaagga 8880
gcctacatca ggacagtagt accattcctc agagaagagg tatacataac aagaaaatcg 8940
cgtgaacacc ttatataact tagcccgtta ttgagctaaa aaaccttgca aaatttccta 9000
tgaataagaa tacttcagac gtgataaaaa tttactttct aactcttctc acgctgcccc 9060
tatctgttct tccgctctac cgtgagaaat aaagcatcga gtacggcagt tcgctgtcac 9120
tgaactaaaa caataaggct agttcgaatg atgaacttgc ttgctgtcaa acttctgagt 9180
tgccgctgat gtgacactgt gacaataaat tcaaaccggt tatagcggtc tcctccggta 9240
ccggttctgc cacctccaat agagctcagt aggagtcaga acctctgcgg tggctgtcag 9300
tgactcatcc gcgtttcgta agttgtgcgc gtgcacattt cgcccgttcc cgctcatctt 9360
gcagcaggcg gaaattttca tcacgctgta ggacgcaaaa aaaaaataat taatcgtaca 9420
agaatcttgg aaaaaaaatt gaaaaatttt gtataaaagg gatgacctaa cttgactcaa 9480
tggcttttac acccagtatt ttccctttcc ttgtttgtta caattataga agcaagacaa 9540
aaacatatag acaacctatt cctaggagtt atatttttt accctaccag caatataagt 9600
aaaaaactgt tt 9612

<210> 933
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 882 C7

<400> 933

Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1 5 10 15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
20 25 30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
35 40 45

Ile Ile Gly Leu Tyr Gly Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
50 55 60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65 70 75 80

```
Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 934
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 882 F9

<400> 934

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Glu Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu His
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270
```

```
Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 935
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 882 G6

<400> 935

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Ala Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160
```

```
Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 936
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 882 G12

<400> 936

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60
```

```
Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Leu Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Tyr Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 937
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 883 C4

<400> 937

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                  10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140

Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Val Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Leu Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240
```

```
Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Lys Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340
```

<210> 938
<211> 343
<212> PRT
<213> Artificial Sequence

<220>
<223> 883 G9

<400> 938

```
Met Glu Glu Cys Lys Met Ala Lys Ile Tyr Tyr Gln Glu Asp Cys Asn
1               5                   10                  15

Leu Ser Leu Leu Asp Gly Lys Thr Ile Ala Val Ile Gly Tyr Gly Ser
            20                  25                  30

Gln Gly His Ala His Ala Leu Asn Ala Lys Glu Ser Gly Cys Asn Val
        35                  40                  45

Ile Ile Gly Leu Tyr Glu Gly Ala Lys Glu Trp Lys Arg Ala Glu Glu
    50                  55                  60

Gln Gly Phe Glu Val Tyr Thr Ala Ala Glu Ala Ala Lys Lys Ala Asp
65                  70                  75                  80

Ile Ile Met Ile Leu Ile Asn Asp Glu Lys Gln Ala Thr Met Tyr Lys
                85                  90                  95

Asn Asp Ile Glu Pro Asn Leu Glu Ala Gly Asn Met Leu Met Phe Ala
            100                 105                 110

His Gly Phe Asn Ile His Phe Gly Cys Ile Val Pro Pro Lys Asp Val
        115                 120                 125

Asp Val Thr Met Ile Ala Pro Lys Gly Pro Gly His Thr Val Arg Ser
    130                 135                 140
```

```
Glu Tyr Glu Glu Gly Lys Gly Val Pro Cys Leu Val Ala Val Glu Gln
145                 150                 155                 160

Asp Ala Thr Gly Lys Ala Leu Asp Met Ala Leu Ala Tyr Ala Leu Ala
                165                 170                 175

Ile Gly Gly Ala Arg Ala Gly Val Leu Glu Thr Thr Phe Arg Thr Glu
            180                 185                 190

Thr Glu Thr Asp Met Phe Gly Glu Gln Ala Val Leu Cys Gly Gly Val
        195                 200                 205

Cys Ala Leu Met Gln Ala Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr
    210                 215                 220

Asp Pro Arg Asn Ala Tyr Phe Glu Cys Ile His Glu Met Lys Leu Ile
225                 230                 235                 240

Val Asp Leu Ile Tyr Gln Ser Gly Phe Ser Gly Met Arg Tyr Ser Ile
                245                 250                 255

Ser Asn Thr Ala Glu Tyr Gly Asp Tyr Ile Thr Gly Pro Lys Ile Ile
            260                 265                 270

Thr Glu Asp Thr Lys Lys Ala Met Lys Lys Ile Leu Ser Asp Ile Gln
        275                 280                 285

Asp Gly Thr Phe Ala Lys Asp Phe Leu Val Asp Met Ser Asp Ala Gly
    290                 295                 300

Ser Gln Val His Phe Met Ala Met Arg Lys Leu Ala Ser Glu His Pro
305                 310                 315                 320

Ala Glu Val Val Gly Glu Glu Ile Arg Ser Leu Tyr Ser Trp Ser Asp
                325                 330                 335

Glu Asp Lys Leu Ile Asn Asn
            340

<210>  939
<211>  571
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Streptococcus mutans DHAD

<400>  939

Met Thr Asp Lys Lys Thr Leu Lys Asp Leu Arg Asn Arg Ser Ser Val
1               5                   10                  15

Tyr Asp Ser Met Val Lys Ser Pro Asn Arg Ala Met Leu Arg Ala Thr
            20                  25                  30
```

```
Gly Met Gln Asp Glu Asp Phe Glu Lys Pro Ile Val Gly Val Ile Ser
        35                  40                  45

Thr Trp Ala Glu Asn Thr Pro Cys Asn Ile His Leu His Asp Phe Gly
    50                  55                  60

Lys Leu Ala Lys Val Gly Val Lys Glu Ala Gly Ala Trp Pro Val Gln
65                  70                  75                  80

Phe Gly Thr Ile Thr Val Ser Asp Gly Ile Ala Met Gly Thr Gln Gly
                85                  90                  95

Met Arg Phe Ser Leu Thr Ser Arg Asp Ile Ile Ala Asp Ser Ile Glu
            100                 105                 110

Ala Ala Met Gly Gly His Asn Ala Asp Ala Phe Val Ala Ile Gly Gly
        115                 120                 125

Cys Asp Lys Asn Met Pro Gly Ser Val Ile Ala Met Ala Asn Met Asp
    130                 135                 140

Ile Pro Ala Ile Phe Ala Tyr Gly Gly Thr Ile Ala Pro Gly Asn Leu
145                 150                 155                 160

Asp Gly Lys Asp Ile Asp Leu Val Ser Val Phe Glu Gly Val Gly His
                165                 170                 175

Trp Asn His Gly Asp Met Thr Lys Glu Glu Val Lys Ala Leu Glu Cys
            180                 185                 190

Asn Ala Cys Pro Gly Pro Gly Gly Cys Gly Gly Met Tyr Thr Ala Asn
        195                 200                 205

Thr Met Ala Thr Ala Ile Glu Val Leu Gly Leu Ser Leu Pro Gly Ser
    210                 215                 220

Ser Ser His Pro Ala Glu Ser Ala Glu Lys Lys Ala Asp Ile Glu Glu
225                 230                 235                 240

Ala Gly Arg Ala Val Val Lys Met Leu Glu Met Gly Leu Lys Pro Ser
                245                 250                 255

Asp Ile Leu Thr Arg Glu Ala Phe Glu Asp Ala Ile Thr Val Thr Met
            260                 265                 270

Ala Leu Gly Gly Ser Thr Asn Ser Thr Leu His Leu Leu Ala Ile Ala
        275                 280                 285

His Ala Ala Asn Val Glu Leu Thr Leu Asp Asp Phe Asn Thr Phe Gln
    290                 295                 300

Glu Lys Val Pro His Leu Ala Asp Leu Lys Pro Ser Gly Gln Tyr Val
305                 310                 315                 320
```

Phe Gln Asp Leu Tyr Lys Val Gly Gly Val Pro Ala Val Met Lys Tyr
325 330 335

Leu Leu Lys Asn Gly Phe Leu His Gly Asp Arg Ile Thr Cys Thr Gly
340 345 350

Lys Thr Val Ala Glu Asn Leu Lys Ala Phe Asp Asp Leu Thr Pro Gly
355 360 365

Gln Lys Val Ile Met Pro Leu Glu Asn Pro Lys Arg Glu Asp Gly Pro
370 375 380

Leu Ile Ile Leu His Gly Asn Leu Ala Pro Asp Gly Ala Val Ala Lys
385 390 395 400

Val Ser Gly Val Lys Val Arg Arg His Val Gly Pro Ala Lys Val Phe
405 410 415

Asn Ser Glu Glu Glu Ala Ile Glu Ala Val Leu Asn Asp Asp Ile Val
420 425 430

Asp Gly Asp Val Val Val Val Arg Phe Val Gly Pro Lys Gly Gly Pro
435 440 445

Gly Met Pro Glu Met Leu Ser Leu Ser Ser Met Ile Val Gly Lys Gly
450 455 460

Gln Gly Glu Lys Val Ala Leu Leu Thr Asp Gly Arg Phe Ser Gly Gly
465 470 475 480

Thr Tyr Gly Leu Val Val Gly His Ile Ala Pro Glu Ala Gln Asp Gly
485 490 495

Gly Pro Ile Ala Tyr Leu Gln Thr Gly Asp Ile Val Thr Ile Asp Gln
500 505 510

Asp Thr Lys Glu Leu His Phe Asp Ile Ser Asp Glu Glu Leu Lys His
515 520 525

Arg Gln Glu Thr Ile Glu Leu Pro Pro Leu Tyr Ser Arg Gly Ile Leu
530 535 540

Gly Lys Tyr Ala His Ile Val Ser Ser Ala Ser Arg Gly Ala Val Thr
545 550 555 560

Asp Phe Trp Lys Pro Glu Glu Thr Gly Lys Lys
565 570

<210> 940
<211> 1641
<212> DNA
<213> Artificial Sequence

<220>
<223> M. caseolyticus KAD

<400> 940
atgaaacaac gtatcgggca atacttgatc gatgccctac acgttaatgg tgtcgataag 60

atctttggag tcccaggtga tttcacttta gccttttttgg acgatatcat aagacatgac 120

aacgtggaat gggtgggaaa tactaatgag ttgaacgccg cttacgccgc tgatggttac 180

gctagagtta atggattagc cgctgtatct accacttttg gggttggcga gttatctgct 240

gtgaatggta ttgctggaag ttacgcagag cgtgttcctg taatcaaaat ctcaggcggt 300

ccttcatcag ttgctcaaca agagggtaga tatgtccacc attcattggg tgaaggaatc 360

tttgattcat attcaaagat gtacgctcac ataaccgcaa caactacaat cttatccgtt 420

gacaacgcag tcgacgaaat tgatagagtt attcattgtg ctttgaagga aaagaggcca 480

gtgcatattc atttgcctat tgacgtagcc ttaactgaga ttgaaatccc tcatgcacca 540

aaagtttaca cacacgaatc ccagaacgtc gatgcttaca ttcaagctgt tgagaaaaag 600

ttaatgtctg caaaacaacc agtaatcata gcaggtcatg aaatcaattc attcaagttg 660

cacgaacaac tggaacagtt tgtcaatcag acaaacatcc ctgttgcaca actttccttg 720

ggtaagtctg ctttcaatga agagaatgaa cattaccttg gtatctacga tggcaaaatc 780

gcaaaggaaa atgtgagaga gtacgtcgac aatgctgatg tcatattgaa cataggtgcc 840

aaactgactg attctgctac agctggattt tcctacaagt tcgatacaaa caacataatc 900

tacattaacc ataatgactt caaagctgaa gatgtgattt ctgataatgt ttcactgatt 960

gatcttgtga atggcctgaa ttctattgac tatagaaatg aaacacacta cccatcttat 1020

caaagatctg atatgaaata cgaattgaat gacgcaccac ttacacaatc taactatttc 1080

aaaatgatga acgcttttct agaaaaagat gacatcctac tagctgaaca aggtacatcc 1140

tttttcggcg catatgactt atccctatac aagggaaatc agtttatcgg tcagccttta 1200

tgggggtcaa tagggtatac ttttccatct ttactaggaa gtcaactagc agacatgcat 1260

aggagaaaca ttttgcttat aggcgatggt agtttacaac ttactgttca agccctaagt 1320

acaatgatta gaaaggatat caaaccaatc attttcgtta tcaataacga cggttacacc 1380

gtcgaaagac ttatccacgg catggaagag ccatacaatg atatccaaat gtggaactac 1440

aagcaattgc cagaagtatt tggtggaaaa gatactgtaa aagttcatga tgctaaaacc 1500

tccaacgaac tgaaaactgt aatggattct gttaaagcag acaaagatca catgcatttc 1560

attgaagtgc atatggcagt agaggacgcc ccaaagaagt tgattgatat agctaaagcc 1620

tttagtgatg ctaacaagta a 1641

<210> 941
<211> 1647
<212> DNA
<213> Artificial Sequence

<220>
<223> L. grayi KAD

<400> 941
atgtacaccg tcggccaata cttagtagac cgcttagaag agatcggcat cgataaggtt 60

tttggtgtcc cgggtgacta caacctgacc tttttggact acatccagaa ccacgaaggt 120

ctgagctggc aaggtaatac gaatgaactg aatgccgcgt acgcagctga tggctatgct 180

cgtgaacgcg gtgttagcgc tttggtcacg accttcggcg ttggtgagct gtccgcaatc 240

aatggcaccg caggtagctt cgcggagcaa gttccggtga ttcatatcgt gggcagcccg 300

accatgaatg ttcagagcaa caagaaactg gttcatcaca gcctgggtat gggcaacttt 360

cacaacttca gcgagatggc gaaagaagtc accgccgcaa ccacgatgct gacggaagag 420

aatgcggcgt cggagattga tcgtgttctg gaaaccgccc tgctggagaa acgcccagtg 480

tacatcaatc tgccgatcga cattgctcac aaggcgatcg tcaagccggc gaaagccctg 540

caaaccgaga agagctctgg cgagcgtgag gcacaactgg cggagatcat tctgagccat 600

ctggagaagg ctgcacagcc gattgtgatt gcgggtcacg agatcgcgcg cttccagatc 660

cgtgagcgtt tcgagaattg gattaatcaa acgaaactgc cggtgaccaa tctggcctac 720

ggcaagggta gcttcaacga agaaaacgag catttcattg gtacctatta tcctgcattt 780

agcgataaga acgtgctgga ctacgtggat aactccgact ttgtcctgca ctttggtggt 840

aaaatcattg ataacagcac ctccagcttc tcccaaggct tcaaaaccga gaacaccctg 900

actgcggcga acgatatcat tatgctgccg gacggtagca cgtattctgg tattagcctg 960

aatggcctgc tggccgagct ggaaaaactg aatttcacgt ttgccgacac cgcagcaaag 1020

caggcggagt tggcggtgtt tgagccgcag gctgaaaccc cgttgaaaca ggaccgtttt 1080

caccaggcgg tgatgaattt tctgcaagct gacgatgtcc tggttacgga acagggcacc 1140

tcttcttttg gcttgatgct ggcgcctctg aaaaagggta tgaacttgat ctcgcaaacg 1200

ctgtggggta gcattggtta cacgttgccg gcgatgattg gtagccaaat tgcggcaccg 1260

gagcgtcgtc atatcctgag cattggtgat ggtagctttc agctgactgc gcaggaaatg 1320

agcaccattt tccgtgagaa actgacccca gtcatcttca tcattaacaa tgatggctat 1380

accgttgagc gtgcgatcca tggcgaagat gaaagctata acgacattcc gacgtggaac 1440

ttgcaactgg tggcggaaac cttcggtggt gacgccgaaa ccgtcgacac tcacaatgtg 1500

ttcacggaga ctgatttcgc caacaccctg gcggcaattg acgcgacgcc gcagaaagca 1560

cacgttgtgg aagttcacat ggaacaaatg gatatgccgg agagcctgcg ccagatcggt 1620

ctggcactgt ccaagcagaa tagctaa 1647

<210> 942
<211> 491
<212> PRT
<213> Artificial Sequence

<220>
<223> E. coli KARI

<400> 942

Met Ala Asn Tyr Phe Asn Thr Leu Asn Leu Arg Gln Gln Leu Ala Gln
1 5 10 15

Leu Gly Lys Cys Arg Phe Met Gly Arg Asp Glu Phe Ala Asp Gly Ala
20 25 30

Ser Tyr Leu Gln Gly Lys Lys Val Val Ile Val Gly Cys Gly Ala Gln
35 40 45

```
Gly Leu Asn Gln Gly Leu Asn Met Arg Asp Ser Gly Leu Asp Ile Ser
    50                  55                  60

Tyr Ala Leu Arg Lys Glu Ala Ile Ala Glu Lys Arg Ala Ser Trp Arg
65                  70                  75                  80

Lys Ala Thr Glu Asn Gly Phe Lys Val Gly Thr Tyr Glu Glu Leu Ile
                85                  90                  95

Pro Gln Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gln His Ser
            100                 105                 110

Asp Val Val Arg Thr Val Gln Pro Leu Met Lys Asp Gly Ala Ala Leu
        115                 120                 125

Gly Tyr Ser His Gly Phe Asn Ile Val Glu Val Gly Glu Gln Ile Arg
    130                 135                 140

Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145                 150                 155                 160

Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
                165                 170                 175

Val His Pro Glu Asn Asp Pro Lys Gly Glu Gly Met Ala Ile Ala Lys
            180                 185                 190

Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Glu Ser
        195                 200                 205

Ser Phe Val Ala Glu Val Lys Ser Asp Leu Met Gly Glu Gln Thr Ile
    210                 215                 220

Leu Cys Gly Met Leu Gln Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
225                 230                 235                 240

Val Glu Glu Gly Thr Asp Pro Ala Tyr Ala Glu Lys Leu Ile Gln Phe
                245                 250                 255

Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Gln Gly Gly Ile Thr Leu
            260                 265                 270

Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
        275                 280                 285

Ser Glu Gln Leu Lys Glu Ile Met Ala Pro Leu Phe Gln Lys His Met
    290                 295                 300

Asp Asp Ile Ile Ser Gly Glu Phe Ser Ser Gly Met Met Ala Asp Trp
305                 310                 315                 320

Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Glu Glu Thr Gly Lys
                325                 330                 335
```

```
Thr Ala Phe Glu Thr Ala Pro Gln Tyr Glu Gly Lys Ile Gly Glu Gln
            340                 345                 350

Glu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly
        355                 360                 365

Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
    370                 375                 380

Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
385                 390                 395                 400

Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
                405                 410                 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
            420                 425                 430

Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
        435                 440                 445

Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
    450                 455                 460

Ile Arg Ser His Ala Ile Glu Gln Val Gly Lys Lys Leu Arg Gly Tyr
465                 470                 475                 480

Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly
                485                 490
```

<210> 943
<211> 338
<212> PRT
<213> Artificial Sequence

<220>
<223> Pseudomonas fluorescens KARI

<400> 943

```
Met Lys Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Gln Gly
1               5                   10                  15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Gln Gly His Ala Gln Ala
            20                  25                  30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
        35                  40                  45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
    50                  55                  60

Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
65                  70                  75                  80
```

```
Pro Asp Glu Phe Gln Ser Gln Leu Tyr Lys Asn Glu Ile Glu Pro Asn
                85                  90                  95

Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
            100                 105                 110

Tyr Asn Gln Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
        115                 120                 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
    130                 135                 140

Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gln Asp Ala Ser Gly Asn Ala
145                 150                 155                 160

Lys Asn Val Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Thr
                165                 170                 175

Gly Ile Ile Glu Thr Thr Phe Lys Asp Glu Thr Glu Thr Asp Leu Phe
            180                 185                 190

Gly Glu Gln Ala Val Leu Cys Gly Gly Thr Val Glu Leu Val Lys Ala
        195                 200                 205

Gly Phe Glu Thr Leu Val Glu Ala Gly Tyr Ala Pro Glu Met Ala Tyr
    210                 215                 220

Phe Glu Cys Leu His Glu Leu Lys Leu Ile Val Asp Leu Met Tyr Glu
225                 230                 235                 240

Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Glu Tyr
                245                 250                 255

Gly Glu Tyr Val Thr Gly Pro Glu Val Ile Asn Ala Glu Ser Arg Gln
            260                 265                 270

Ala Met Arg Asn Ala Leu Lys Arg Ile Gln Asp Gly Glu Tyr Ala Lys
        275                 280                 285

Met Phe Ile Ser Glu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
    290                 295                 300

Arg Arg Asn Asn Ala Ala His Gly Ile Glu Ile Ile Gly Glu Gln Leu
305                 310                 315                 320

Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
                325                 330                 335

Lys Asn
```

<210> 944
<211> 348
<212> PRT
<213> Artificial Sequence

<220>
<223> Achromobacter xylosoxidans ADH

<400> 944

```
Met Lys Ala Leu Val Tyr His Gly Asp His Lys Ile Ser Leu Glu Asp
1               5                  10                  15

Lys Pro Lys Pro Thr Leu Gln Lys Pro Thr Asp Val Val Val Arg Val
            20                  25                  30

Leu Lys Thr Thr Ile Cys Gly Thr Asp Leu Gly Ile Tyr Lys Gly Lys
        35                  40                  45

Asn Pro Glu Val Ala Asp Gly Arg Ile Leu Gly His Glu Gly Val Gly
    50                  55                  60

Val Ile Glu Glu Val Gly Glu Ser Val Thr Gln Phe Lys Lys Gly Asp
65                  70                  75                  80

Lys Val Leu Ile Ser Cys Val Thr Ser Cys Gly Ser Cys Asp Tyr Cys
                85                  90                  95

Lys Lys Gln Leu Tyr Ser His Cys Arg Asp Gly Gly Trp Ile Leu Gly
            100                 105                 110

Tyr Met Ile Asp Gly Val Gln Ala Glu Tyr Val Arg Ile Pro His Ala
        115                 120                 125

Asp Asn Ser Leu Tyr Lys Ile Pro Gln Thr Ile Asp Asp Glu Ile Ala
    130                 135                 140

Val Leu Leu Ser Asp Ile Leu Pro Thr Gly His Glu Ile Gly Val Gln
145                 150                 155                 160

Tyr Gly Asn Val Gln Pro Gly Asp Ala Val Ala Ile Val Gly Ala Gly
                165                 170                 175

Pro Val Gly Met Ser Val Leu Leu Thr Ala Gln Phe Tyr Ser Pro Ser
            180                 185                 190

Thr Ile Ile Val Ile Asp Met Asp Glu Asn Arg Leu Gln Leu Ala Lys
        195                 200                 205

Glu Leu Gly Ala Thr His Thr Ile Asn Ser Gly Thr Glu Asn Val Val
    210                 215                 220

Glu Ala Val His Arg Ile Ala Ala Glu Gly Val Asp Val Ala Ile Glu
225                 230                 235                 240

Ala Val Gly Ile Pro Ala Thr Trp Asp Ile Cys Gln Glu Ile Val Lys
                245                 250                 255

Pro Gly Ala His Ile Ala Asn Val Gly Val His Gly Val Lys Val Asp
            260                 265                 270
```

Phe Glu Ile Gln Lys Leu Trp Ile Lys Asn Leu Thr Ile Thr Thr Gly
275 280 285

Leu Val Asn Thr Asn Thr Thr Pro Met Leu Met Lys Val Ala Ser Thr
290 295 300

Asp Lys Leu Pro Leu Lys Lys Met Ile Thr His Arg Phe Glu Leu Ala
305 310 315 320

Glu Ile Glu His Ala Tyr Gln Val Phe Leu Asn Gly Ala Lys Glu Lys
325 330 335

Ala Met Lys Ile Ile Leu Ser Asn Ala Gly Ala Ala
340 345

<210> 945
<211> 347
<212> PRT
<213> Artificial Sequence

<220>
<223> B. indica ADH

<400> 945

Met Lys Ala Leu Val Tyr Arg Gly Pro Gly Gln Lys Leu Val Glu Glu
1 5 10 15

Arg Gln Lys Pro Glu Leu Lys Glu Pro Gly Asp Ala Ile Val Lys Val
20 25 30

Thr Lys Thr Thr Ile Cys Gly Thr Asp Leu His Ile Leu Lys Gly Asp
35 40 45

Val Ala Thr Cys Lys Pro Gly Arg Val Leu Gly His Glu Gly Val Gly
50 55 60

Val Ile Glu Ser Val Gly Ser Gly Val Thr Ala Phe Gln Pro Gly Asp
65 70 75 80

Arg Val Leu Ile Ser Cys Ile Ser Ser Cys Gly Lys Cys Ser Phe Cys
85 90 95

Arg Arg Gly Met Phe Ser His Cys Thr Thr Gly Gly Trp Ile Leu Gly
100 105 110

Asn Glu Ile Asp Gly Thr Gln Ala Glu Tyr Val Arg Val Pro His Ala
115 120 125

Asp Thr Ser Leu Tyr Arg Ile Pro Ala Gly Ala Asp Glu Glu Ala Leu
130 135 140

Val Met Leu Ser Asp Ile Leu Pro Thr Gly Phe Glu Cys Gly Val Leu
145 150 155 160

```
Asn Gly Lys Val Ala Pro Gly Ser Ser Val Ala Ile Val Gly Ala Gly
                165                 170                 175

Pro Val Gly Leu Ala Ala Leu Leu Thr Ala Gln Phe Tyr Ser Pro Ala
            180                 185                 190

Glu Ile Ile Met Ile Asp Leu Asp Asp Asn Arg Leu Gly Leu Ala Lys
        195                 200                 205

Gln Phe Gly Ala Thr Arg Thr Val Asn Ser Thr Gly Gly Asn Ala Ala
    210                 215                 220

Ala Glu Val Lys Ala Leu Thr Glu Gly Leu Gly Val Asp Thr Ala Ile
225                 230                 235                 240

Glu Ala Val Gly Ile Pro Ala Thr Phe Glu Leu Cys Gln Asn Ile Val
                245                 250                 255

Ala Pro Gly Gly Thr Ile Ala Asn Val Gly Val His Gly Ser Lys Val
            260                 265                 270

Asp Leu His Leu Glu Ser Leu Trp Ser His Asn Val Thr Ile Thr Thr
        275                 280                 285

Arg Leu Val Asp Thr Ala Thr Thr Pro Met Leu Leu Lys Thr Val Gln
    290                 295                 300

Ser His Lys Leu Asp Pro Ser Arg Leu Ile Thr His Arg Phe Ser Leu
305                 310                 315                 320

Asp Gln Ile Leu Asp Ala Tyr Glu Thr Phe Gly Gln Ala Ala Ser Thr
                325                 330                 335

Gln Ala Leu Lys Val Ile Ile Ser Met Glu Ala
            340                 345
```

<210> 946
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Eliza

<400> 946

```
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg       60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat      120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aagatgcgga ggaatggaaa      180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac      240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa      300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt      360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac      420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa      480
```

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 947
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Frank

<400> 947
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag ctggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 948
<211> 912
<212> DNA
<213> Artificial Sequence

20 Jan 2017 2017200439

<220>
<223> K9_Grace

<400> 948
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 60

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 120

atcatcatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 180

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 240

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 300

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 360

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 420

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 480

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 540

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 600

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 660

gaatacggtg actacataac cggtccaaag atcattactg aagataccaa gaaggctatg 720

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgatatg 780

tctgatgatg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 840

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 900

attaacaact ga 912

<210> 949
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ingrid

<400> 949
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tttacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccagat ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 950
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Jarvis

<400> 950
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggtat cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 951
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Kelly

<400> 951
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgggga ggagtggaat 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa aaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 952
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Norman

<400> 952
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aatgtgcgga ggagtggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 953
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ophelia

<400> 953
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg cttcaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccatgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 954
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Pat

<400> 954
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattag aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

<210> 955
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Quentin

<400> 955
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60
gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120
gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180
agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240
atcattatga tcttgatccc agatgaaaag caggctacca tttacaaaaa cgacatcgaa 300
ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360
tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420
accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480
gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540
agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600
caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660
gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720
gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780
gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840
aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900
tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960
gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020
attaacaact ga 1032

20 Jan 2017 2017200439

<210> 956
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Ralph

<400> 956
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt agaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 957
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Sophia

<400> 957
atggaagaat gtaagatggc taatatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaatg caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

atcgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 958
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Tiberius

<400> 958
atggaagaat gtaagatggc taagatttac taccaagtag actgtaactt gtccttgttg 60

gatggtaaga ctattgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagccgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tttacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

atcgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 959
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>

20 Jan 2017
2017200439

<223> K9_Ursala

<400> 959
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgtgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtttgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 960
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Victor

<400> 960
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtggaaa 180

agagctgaag aacaaggttt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatgc tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gcttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 961
<211> 1032
<212> DNA
<213> Artificial Sequence

<220>
<223> K9_Watson

<400> 961
atggaagaat gtaagatggc taagatttac taccaagaag actgtaactt gtccttgttg 60

gatggtaaga ctatcgccgt tatcggttac ggttctcaag gtcacgctca tgccctgaat 120

gctaaggaat ccggttgtaa cgttatcatt ggtttattcg aaggtgcgga ggagtgcaaa 180

agagctgaag aacaaggctt cgaagtctac accgctgctg aagctgctaa gaaggctgac 240

atcattatga tcttgatccc agatgaaaag caggctacca tgtacaaaaa cgacatcgaa 300

ccaaacttgg aagccggtaa catgttgatg ttcgctcacg gtttcaacat ccatttcggt 360

tgtattgttc caccaaagga cgttgatgtc actatgatcg ctccaaaggg tccaggtcac 420

accgttagat ccgaatacga agaaggtaaa ggtgtcccat gcttggttgc tgtcgaacaa 480

gacgctactg gcaaggcttt ggatatggct ttggcctacg ctttagccat cggtggtgct 540

agagccggtg tcttggaaac taccttcaga accgaaactg aaaccgactt gttcggtgaa 600

caagctgttt tatgtggtgg tgtctgcgct ttgatgcagg ccggttttga aaccttggtt 660

gaagccggtt acgacccaag aaacgcttac ttcgaatgta tccacgaaat gaagttgatc 720

gttgacttga tctaccaatc tggtttctcc ggtatgcgtt actctatctc caacactgct 780

gaatacggtg actacattac cggtccaaag atcattactg aagataccaa gaaggctatg 840

aagaagattt tgtctgacat tcaagatggt acctttgcca aggacttctt ggttgacatg 900

tctgatgctg gttcccaggt ccacttcaag gctatgagaa agttggcctc cgaacaccca 960

gctgaagttg tcggtgaaga aattagatcc ttgtactcct ggtccgacga agacaagttg 1020

attaacaact ga 1032

<210> 962
<211> 978
<212> DNA
<213> Lactobacillus plantarum

<400> 962
atggatttat ttgagtcatt agcacaaaaa attactggta aagatcaaac aattgttttc 60

cctgaaggaa ctgaaccccg aattgtcggt gcggcagcgc gattagctgc agacggcttg 120

gttaagccga ttgttttagg tgcaacggac aaagttcagg ctgtggctaa cgatttgaat 180

gcggatttaa caggcgttca agtccttgat cctgcgacat acccggctga agataagcaa 240
gcaatgcttg atgccctcgt tgaacggcgg aaaggtaaga atacgccaga acaagcggct 300
aaaatgctgg aagatgaaaa ctactttggc acgatgctcg tttatatggg caaagcggat 360
gggatggttt caggtgcaat ccatccaact ggtgatacgg tacggccagc gttacaaatt 420
attaagacca agcccggttc acaccgaatc tcgggtgcat ttatcatgca aaagggtgag 480
gaacgctacg tctttgctga ctgtgccatc aatattgatc ccgatgccga tacgttagcg 540
gaaattgcca ctcagagtgc ggctactgct aaggtcttcg atattgaccc gaaagttgcg 600
atgctcagct tctcaactaa gggttcggct aagggtgaaa tggtcactaa agtgcaagaa 660
gcaacggcca aggcgcaagc tgctgaaccg gaattggcta tcgatggtga acttcaattt 720
gacgcggcct tcgttgaaaa agttggtttg caaaaggctc ctggttccaa agtagctggt 780
catgccaatg tctttgtatt tccagagctt cagtctggta atattggcta taagattgcg 840
caacgatttg gtcattttga agcggtgggt cctgtcttgc aaggcctgaa caagccggtc 900
tccgacttgt cacgtggatg cagtgaagaa gacgtttata aggttgcgat tattacagca 960
gcccaaggat tagcttaa 978

<210> 963
<211> 2367
<212> DNA
<213> Lactobacillus plantarum

<400> 963
atgacaacag attactcatc accagcatat ttgcaaaaag ttgataagta ctggcgtgct 60
gccaactact tatcagttgg tcaactttat ttaaaagata atccactatt acaacggcca 120
ttgaaggcca gtgacgttaa ggttcatcca attggtcact gggggacgat tgccggtcaa 180
aactttatct atgctcatct taaccgggtc atcaacaagt acggtttgaa gatgttctac 240
gttgaaggtc caggtcatgg tggtcaagtg atggtttcaa actcttacct tgacggtact 300
tacaccgata tttatccaga aattacgcag gatgttgaag ggatgcaaaa gctcttcaag 360
caattctcat tcccaggtgg ggttgcttcc catgcggcac ctgaaacacc cggttcaatc 420
cacgaaggtg gcgaacttgg ttactcaatt tcacacgggg ttggggcaat tcttgacaat 480
cctgacgaaa tcgccgcggt tgttgttggt gatggggaat ccgaaacggg tccattagca 540
acttcatggc aatcaacgaa gttcattaac ccaatcaacg acggggctgt tttaccaatc 600
ttgaacttaa atggttttaa gatttctaat ccaacgattt ttggtcggac ttctgatgct 660
aagattaagg aatacttcga aagcatgaat tgggaaccaa tcttcgttga aggtgacgat 720
cctgaaaagg ttcacccagc cttagctaag gccatggatg aagccgttga aaagatcaag 780
gcaatccaga agcatgctcg cgaaaataac gatgcaacat tgccagtatg gccaatgatc 840
gtcttccgcg cacctaaggg ctggactggt ccgaagtcat gggacggtga taagatcgaa 900
ggttcattcc gtgctcatca aattccgatt cctgttgatc aaaatgacat ggaacatgcg 960
gatgctttag ttgattggct cgaatcatat caaccaaaag aactcttcaa tgaagatggc 1020
tctttgaagg atgatattaa agaaattatt cctactgggg acagtcggat ggctgctaac 1080
ccaatcacca atggtggggt cgatccgaaa gccttgaact taccaaactt ccgtgattat 1140

gcggtcgata cgtccaaaga aggcgcgaat gttaagcaag atatgatcgt ttggtcagac 1200

tatttgcggg atgtcatcaa gaaaaatcct gataacttcc ggttgttcgg acctgatgaa 1260

accatgtcta accgtttata tggtgtcttc gaaaccacta atcgtcaatg gatggaagac 1320

attcatccag atagtgacca atatgaagca ccagctggcc gggtcttaga tgctcagtta 1380

tctgaacacc aagctgaagg ttggttagaa ggttacgtct taactggacg tcatgggtta 1440

tttgccagtt atgaagcctt cctacgcgtt gtggactcaa tgttgacgca acacttcaag 1500

tggttacgta aagccaatga acttgattgg cgtaaaaagt acccatcact taacattatc 1560

gcggcttcaa ctgtattcca acaagaccat aatggttata cccaccaaga tccaggtgca 1620

ttaactcatt tggccgaaaa gaaaccagaa tacattcgtg aatatttacc agccgatgcc 1680

aacacgttat tagctgtcgg tgacgtcatt ttccggagcc aagaaaagat caactacgtg 1740

gttacgtcaa aacacccacg tcaacaatgg ttcagcattg aagaagctaa gcaattagtt 1800

gacaatggtc ttggtatcat tgattgggca agtacggacc aaggtagcga accagacatt 1860

gtctttgcag ctgctgggac ggaaccaacg cttgaaacgt tggctgccat ccaattacta 1920

cacgacagtt tcccagagat gaagattcgt ttcgtgaacg tggtcgacat cttgaagtta 1980

cgtagtcctg aaaaggatcc gcggggcttg tcagatgctg agtttgacca ttactttact 2040

aaggacaaac cagtggtctt tgctttccac ggttacgaag acttagttcg tgacatcttc 2100

tttgatcgtc acaaccataa cttatacgtc cacggttacc gtgaaaatgg tgatattacc 2160

acaccattcg acgtacgggt catgaaccag atggaccgct tcgacttagc taagtcggca 2220

attgcggcgc aaccagcaat ggaaaacact ggtgcggcct tcgttcaatc catggataat 2280

atgcttgcta aacacaatgc ctatatccgg gatgccggaa ctgacttgcc agaagttaat 2340

gattggcaat ggaagggttt aaaataa 2367

<210> 964
<211> 9613
<212> DNA
<213> Artificial Sequence

<220>
<223> pHR81-ILV5p-K9SB

<400> 964
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 60

cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 120

cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttattcgaag gtgcggagga 180

gtggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 240

ggctgacatc attatgatct tgatcccaga tgaaaagcag gctaccatgt acaaaaacga 300

catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca 360

tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 420

aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 480

cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 540

tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 600

cggtgaacaa gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 660
cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 720
gttgatcgtt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 780
cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 840
ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt 900
tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 960
acacccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 1020
caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 1080
ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 1140
catttaagat ttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt 1200
ctgaagcagc ttcaaatata tatatttttt acatatttat tatgattcaa tgaacaatct 1260
aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt 1320
gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggctga gctactggtt 1380
gctgtgacgc agcggcatta gcgcggcgtt atgagctacc ctcgtggcct gaaagatggc 1440
gggaataaag cggaactaaa aattactgac tgagccatat tgaggtcaat ttgtcaactc 1500
gtcaagtcac gtttggtgga cggccccttt ccaacgaatc gtatatacta acatgcgcgc 1560
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca 1620
cctgtattta atttccttac tcgcgggttt ttctttttc tcaattcttg gcttcctctt 1680
tctcgagcgg accggatcct ccgcggtgcc ggcagatcta tttaaatggc gcgccgacgt 1740
caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac 1800
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 1860
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 1920
tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc 1980
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 2040
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 2100
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 2160
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 2220
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc 2280
tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg 2340
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 2400
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 2460
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 2520
cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 2580
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 2640
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 2700
agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 2760

tttagattga tttaaaactt cattttaat ttaaaaggat ctaggtgaag atccttttg 2820
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 2880
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 2940
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc 3000
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt 3060
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 3120
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 3180
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac 3240
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 3300
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg 3360
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 3420
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca gggggggcgga 3480
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 3540
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct 3600
ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 3660
aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt 3720
aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 3780
atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 3840
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 3900
acgccaagct ttttctttcc aatttttttt ttttcgtcat tataaaaatc attacgaccg 3960
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata 4020
catgcattta cttataatac agtttttag ttttgctggc cgcatcttct caaatatgct 4080
tcccagcctg cttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat 4140
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta 4200
tactgttgac ccaatgcgtc tcccttgtca tctaaaccca caccgggtgt cataatcaac 4260
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa 4320
tctttgtcgc tcttcgcaat gtcaacagta cccttagtat attctccagt agatagggag 4380
cccttgcatg acaattctgc taacatcaaa aggcctctag gttcctttgt tacttcttct 4440
gccgcctgct tcaaaccgct aacaatacct gggcccacca caccgtgtgc attcgtaatg 4500
tctgcccatt ctgctattct gtatacaccc gcagagtact gcaatttgac tgtattacca 4560
atgtcagcaa attttctgtc ttcgaagagt aaaaaattgt acttggcgga taatgccttt 4620
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt 4680
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca 4740
tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag cttggcagca 4800
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt 4860
cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt 4920

tcttcaacac tacatatgcg tatatatacc aatctaagtc tgtgctcctt ccttcgttct 4980
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaa 5040
gaataaaaaa aaaatgatga attgaaaagc ttgcatgcct gcaggtcgac tctagtatac 5100
tccgtctact gtacgataca cttccgctca ggtccttgtc ctttaacgag gccttaccac 5160
tcttttgtta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc 5220
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac 5280
aatggctgcc atcattatta tccgatgtga cgctgcattt ttttttttt ttttttttt 5340
ttttttttt ttttttttt ttttttttg tacaaatatc ataaaaaaag agaatctttt 5400
taagcaagga ttttcttaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg 5460
gaaccaccta aatcaccagt tctgatacct gcatccaaaa ccttttaac tgcatcttca 5520
atggctttac cttcttcagg caagttcaat gacaatttca acatcattgc agcagacaag 5580
atagtggcga tagggttgac cttattcttt ggcaaatctg gagcggaacc atggcatggt 5640
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac 5700
aaacccaagg agcctgggat aacggaggct tcatcggaga tgatatcacc aaacatgttg 5760
ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggcggcagaa 5820
tcaatcaatt gatgttgaac tttcaatgta gggaattcgt tcttgatggt ttcctccaca 5880
gtttttctcc ataatcttga agaggccaaa acattagctt tatccaagga ccaaataggc 5940
aatggtggct catgttgtag ggccatgaaa gcggccattc ttgtgattct ttgcacttct 6000
ggaacggtgt attgttcact atcccaagcg acaccatcac catcgtcttc ctttctctta 6060
ccaaagtaaa tacctcccac taattctcta acaacaacga agtcagtacc tttagcaaat 6120
tgtggcttga ttggagataa gtctaaaaga gagtcggatg caaagttaca tggtcttaag 6180
ttggcgtaca attgaagttc tttacggatt tttagtaaac cttgttcagg tctaacacta 6240
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatcagc cttcttggag 6300
gcttccagcg cctcatctgg aagtggaaca cctgtagcat cgatagcagc accaccaatt 6360
aaatgatttt cgaaatcgaa cttgacattg gaacgaacat cagaaatagc tttaagaacc 6420
ttaatggctt cggctgtgat ttcttgacca acgtggtcac ctggcaaaac gacgatcttc 6480
ttaggggcag acattacaat ggtatatcct tgaaatatat ataaaaaaaa aaaaaaaaaa 6540
aaaaaaaaaa aatgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta 6600
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg 6660
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat 6720
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg 6780
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc 6840
ttgcacttca atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa 6900
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac 6960
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa 7020
acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac 7080

agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaatctata cttctttttt 7140
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt 7200
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtccg 7260
ttaaggttag aagaaggcta ctttggtgtc tattttctct tccataaaaa aagcctgact 7320
ccacttcccg cgtttactga ttactagcga agctgcgggt gcatttttc aagataaagg 7380
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat 7440
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta 7500
tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag 7560
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga 7620
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata 7680
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg 7740
caatatttta gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc 7800
agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac 7860
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag 7920
ctgcgcacat acagctcact gttcacgtcg cacctatatc tgcgtgttgc ctgtatatat 7980
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc 8040
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg 8100
ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa 8160
ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt 8220
accgagaaac tagaggatct cccattaccg acatttgggc gctatacgtg catatgttca 8280
tgtatgtatc tgtatttaaa acacttttgt attatttttc ctcatatatg tgtataggtt 8340
tatacggatg atttaattat tacttcacca ccctttattt caggctgata tcttagcctt 8400
gttactagtc accggtggcg gccgcacctg gtaaaacctc tagtggagta gtagatgtaa 8460
tcaatgaagc ggaagccaaa agaccagagt agaggcctat agaagaaact gcgatacctt 8520
ttgtgatggc taaacaaaca gacatctttt tatatgtttt tacttctgta tatcgtgaag 8580
tagtaagtga taagcgaatt tggctaagaa cgttgtaagt gaacaaggga cctcttttgc 8640
ctttcaaaaa aggattaaat ggagttaatc attgagattt agttttcgtt agattctgta 8700
tccctaaata actcccttac ccgacgggaa ggcacaaaag acttgaataa tagcaaacgg 8760
ccagtagcca agaccaaata atactagagt taactgatgg tcttaaacag gcattacgtg 8820
gtgaactcca agaccaatat acaaaatatc gataagttat tcttgcccac caatttaagg 8880
agcctacatc aggacagtag taccattcct cagagaagag gtatacataa caagaaaatc 8940
gcgtgaacac cttatataac ttagcccgtt attgagctaa aaaaccttgc aaaatttcct 9000
atgaataaga atacttcaga cgtgataaaa atttactttc taactcttct cacgctgccc 9060
ctatctgttc ttccgctcta ccgtgagaaa taaagcatcg agtacggcag ttcgctgtca 9120
ctgaactaaa acaataaggc tagttcgaat gatgaacttg cttgctgtca aacttctgag 9180
ttgccgctga tgtgacactg tgacaataaa ttcaaaccgg ttatagcggt ctcctccggt 9240

accggttctg ccacctccaa tagagctcag taggagtcag aacctctgcg gtggctgtca 9300

gtgactcatc cgcgtttcgt aagttgtgcg cgtgcacatt tcgcccgttc ccgctcatct 9360

tgcagcaggc ggaaattttc atcacgctgt aggacgcaaa aaaaaaataa ttaatcgtac 9420

aagaatcttg gaaaaaaaat tgaaaaattt tgtataaaag ggatgaccta acttgactca 9480

atggctttta cacccagtat tttccctttc cttgtttgtt acaattatag aagcaagaca 9540

aaaacatata gacaacctat tcctaggagt tatatttttt taccctacca gcaatataag 9600

taaaaaactg ttt 9613

<210> 965
<211> 12319
<212> DNA
<213> Artificial Sequence

<220>
<223> pBP2090

<400> 965
tcccattacc gacatttggg cgctatacgt gcatatgttc atgtatgtat ctgtatttaa 60

aacacttttg tattattttt cctcatatat gtgtataggt ttatacggat gatttaatta 120

ttacttcacc acccttttatt tcaggctgat atcttagcct tgttactaga ttaatcatgt 180

aattagttat gtcacgctta cattcacgcc ctccccccac atccgctcta accgaaaagg 240

aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 300

taagaacgtt atttatattt caaatttttc ttttttttct gtacagacgc gtgtacgcat 360

gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag gctttaattt 420

gcgggcggcc gcacctggta aaacctctag tggagtagta gatgtaatca atgaagcgga 480

agccaaaaga ccagagtaga ggcctataga agaaactgcg atacctttg tgatggctaa 540

acaaacagac atctttttat atgtttttac ttctgtatat cgtgaagtag taagtgataa 600

gcgaatttgg ctaagaacgt tgtaagtgaa caagggacct cttttgcctt tcaaaaaagg 660

attaaatgga gttaatcatt gagatttagt tttcgttaga ttctgtatcc ctaaataact 720

cccttacccg acgggaaggc acaaaagact tgaataatag caaacggcca gtagccaaga 780

ccaaataata ctagagttaa ctgatggtct taaacaggca ttacgtggtg aactccaaga 840

ccaatataca aaatatcgat aagttattct tgcccaccaa tttaaggagc ctacatcagg 900

acagtagtac cattcctcag agaagaggta tacataacaa gaaaatcgcg tgaacacctt 960

atataactta gcccgttatt gagctaaaaa accttgcaaa atttcctatg aataagaata 1020

cttcagacgt gataaaaatt tactttctaa ctcttctcac gctgccccta tctgttcttc 1080

cgctctaccg tgagaaataa agcatcgagt acggcagttc gctgtcactg aactaaaaca 1140

ataaggctag ttcgaatgat gaacttgctt gctgtcaaac ttctgagttg ccgctgatgt 1200

gacactgtga caataaattc aaaccggtta tagcggtctc ctccggtacc ggttctgcca 1260

cctccaatag agctcagtag gagtcagaac ctctgcggtg gctgtcagtg actcatccgc 1320

gtttcgtaag ttgtgcgcgt gcacatttcg cccgttcccg ctcatcttgc agcaggcgga 1380

aattttcatc acgctgtagg acgcaaaaaa aaaataatta atcgtacaag aatcttggaa 1440

aaaaaattga aaaattttgt ataaaaggga tgacctaact tgactcaatg gcttttacac 1500
ccagtatttt ccctttcctt gtttgttaca attatagaag caagacaaaa acatatagac 1560
aacctattcc taggagttat attttttac cctaccagca atataagtaa aaaactgttt 1620
aaacagtatg gaagaatgta agatggctaa gatttactac caagaagact gtaacttgtc 1680
cttgttggat ggtaagacta tcgccgttat cggttacggt tctcaaggtc acgctcatgc 1740
cctgaatgct aaggaatccg gttgtaacgt tatcattggt ttatacgaag gtgctaagga 1800
ttggaaaaga gctgaagaac aaggtttcga agtctacacc gctgctgaag ctgctaagaa 1860
ggctgacatc attatgatct tgatcaacga tgaaaagcag gctaccatgt acaaaaacga 1920
catcgaacca aacttggaag ccggtaacat gttgatgttc gctcacggtt tcaacatcca 1980
tttcggttgt attgttccac caaaggacgt tgatgtcact atgatcgctc caaagggtcc 2040
aggtcacacc gttagatccg aatacgaaga aggtaaaggt gtcccatgct tggttgctgt 2100
cgaacaagac gctactggca aggctttgga tatggctttg gcctacgctt tagccatcgg 2160
tggtgctaga gccggtgtct tggaaactac cttcagaacc gaaactgaaa ccgacttgtt 2220
cggtgaacaa gctgttttat gtggtggtgt ctgcgctttg atgcaggccg gttttgaaac 2280
cttggttgaa gccggttacg acccaagaaa cgcttacttc gaatgtatcc acgaaatgaa 2340
gttgatcgtt gacttgatct accaatctgg tttctccggt atgcgttact ctatctccaa 2400
cactgctgaa tacggtgact acattaccgg tccaaagatc attactgaag ataccaagaa 2460
ggctatgaag aagattttgt ctgacattca agatggtacc tttgccaagg acttcttggt 2520
tgacatgtct gatgctggtt cccaggtcca cttcaaggct atgagaaagt tggcctccga 2580
acacccagct gaagttgtcg gtgaagaaat tagatccttg tactcctggt ccgacgaaga 2640
caagttgatt aacaactgag gccctgcagg ccagaggaaa ataatatcaa gtgctggaaa 2700
ctttttctct tggaattttt gcaacatcaa gtcatagtca attgaattga cccaatttca 2760
catttaagat ttttttttt tcatccgaca tacatctgta cactaggaag ccctgttttt 2820
ctgaagcagc ttcaaatata tatatttttt acatatttat tatgattcaa tgaacaatct 2880
aattaaatcg aaaacaagaa ccgaaacgcg aataaataat ttatttagat ggtgacaagt 2940
gtataagtcc tcatcgggac agctacgatt tctctttcgg ttttggctga gctactggtt 3000
gctgtgacgc agcggcatta gcgcggcgtt atgagctacc ctcgtggcct gaaagatggc 3060
gggaataaag cggaactaaa aattactgac tgagccatat tgaggtcaat ttgtcaactc 3120
gtcaagtcac gtttggtgga cggccccttt ccaacgaatc gtatatacta acatgcgcgc 3180
gcttcctata tacacatata catatatata tatatatata tgtgtgcgtg tatgtgtaca 3240
cctgtattta atttccttac tcgcgggttt ttctttttttc tcaattcttg gcttcctctt 3300
tctcgagcgg accggatcct cgcgaactcc aaaatgagct atcaaaaacg atagatcgat 3360
taggatgact ttgaaatgac tccgcagtgg actggccgtt aatttcaagc gtgagtaaaa 3420
tagtgcatga caaaagatga gctaggcttt tgtaaaaata tcttacgttg taaaatttta 3480
gaaatcatta tttccttcat atcattttgt cattgacctt cagaagaaaa gagccgacca 3540
ataatataaa taaataaata aaaataatat tccattattt ctaaacagat tcaatactca 3600

2017200439 20 Jan 2017

ttaaaaaact atatcaatta atttgaatta acgcggccgc ttaaccacag caaccaggac 3660
aacatttttt gccagtttct tcaggcttcc aaaagtctgt tacggctccc ctagaagcag 3720
acgaaacgat gtgagcatat ttaccaagga taccgcgtga atagagcggt ggcaattcaa 3780
tggtctcttg acgatgtttt aactcttcat cggagatatc aaagtgtaat tccttagtgt 3840
cttggtcaat agtgactatg tctcctgttt gcaggtaggc gattggaccg ccatcttgtg 3900
cttcaggagc gatatgaccc acgacaagac cataagtacc acctgagaag cggccatctg 3960
tcagaagggc aacttttca ccttgccctt taccaacaat cattgatgaa agggaaagca 4020
tttcaggcat accaggaccg ccctttggtc ctacaaaacg tacgacaaca acatcaccat 4080
caacaatatc atcattcaag acagcttcaa tggcttcttc ttcagaatta aagaccttag 4140
caggaccgac atgacgacgc acttttacac cagaaacttt ggcaacggca ccgtctggag 4200
ccaagttacc atggagaata atgagcggac catcttcacg tttaggattt tcaagcggca 4260
taataacctt ttgaccaggt gttaaatcat caaaagcctt caaattttca gcgactgttt 4320
tgccagtaca agtgatacgg tcaccatgaa ggaagccatt tttaaggaga tatttcataa 4380
ctgctggtac ccctccgacc ttgtaaaggt cttggaatac atattgacca gaaggtttca 4440
aatcagccaa atgaggaact ttttcttgga aagtattgaa atcatcaagt gtcaattcca 4500
cattagcagc atgggcaata gctaagaggt gaagggttga gttggttgaa cctcccagag 4560
ccatagttac agtaatagca tcttcaaaag cttcacgcgt taaaatgtca gaaggtttta 4620
agcccatttc gagcattttg acaacagcgc gaccagcttc ttcaatatct gctttctttt 4680
ctgcggattc agccgggtga gaagatgaac ccggaaggct aagtcccaaa acttcaatag 4740
ctgtcgccat tgtgttagca gtatacatac caccgcagcc tccaggaccg ggacaagcat 4800
tacattccaa agctttaact tcttctttgg tcatatcgcc gtggttccaa tggccgacac 4860
cttcaaagac agagactaaa tcgatatctt tgccgtctaa attaccaggt gcaattgttc 4920
cgccgtaagc aaaaatggct gggatatcca tgttagccat agcgataaca gaaccgggca 4980
tgtttttatc acaaccgcca atggctacaa aagcatccgc attatgacct cccatggctg 5040
cttcaataga atctgcaata atatcacgag atgtcaagga gaaacgcatt ccttgggttc 5100
ccatggcgat tccatcagaa accgtgattg ttccgaactg aactggccaa gcaccagctt 5160
ccttaacacc gactttggct agtttaccaa agtcatgtaa gtggatatta caaggtgtgt 5220
tttcagccca agttgaaatg acaccgacga taggtttttc aaagtcttca tcttgcatac 5280
cagttgcacg caacatagca cgattaggtg atttaaccat tgaatcgtaa acagaactac 5340
gatttcttaa gtctttaaga gttttttgt cagtcatact cacgtggaac ttagattaga 5400
ttgctatgct ttctctctaa cgagcaagaa gtaaaaaaag ttgtaataga acaagaaaaa 5460
tgaaactgaa gcttgagaaa ttgaagaccg tttattagct taaatatcaa tgggaggtca 5520
tcgaaagaga aaaaaatcaa gaaagaaact ctcaagaaaa agaaacgtga taaaaatttt 5580
tattgcctct ctcgacgaag agaaagaaac gaggcggtcc ctttttctt ttccaaacct 5640
ttagtacggg taattagcga caccctagag gaagaaagag gggaaattta gtatgctgtg 5700
cttgggtgtc ttgaagtggt acggcgatgc gcggagtccg agaaaatctg gaagagtaaa 5760

aaggggtag aagcgttttg aagctatgcc ggcagatcta tttaaatggc gcgccgacgt 5820
caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt ttctaaatac 5880
attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa taatattgaa 5940
aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt tttgcggcat 6000
tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat gctgaagatc 6060
agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag atccttgaga 6120
gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg ctatgtggcg 6180
cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata cactattctc 6240
agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat ggcatgacag 6300
taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc aacttacttc 6360
tgacaacgat cggaggaccg aaggagctaa ccgcttttt gcacaacatg ggggatcatg 6420
taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac gacgagcgtg 6480
acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact ggcgaactac 6540
ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa gttgcaggac 6600
cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct ggagccggtg 6660
agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc tcccgtatcg 6720
tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga cagatcgctg 6780
agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac tcatatatac 6840
tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag atccttttg 6900
ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg tcagaccccg 6960
tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc 7020
aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc 7080
tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtt cttctagtgt 7140
agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac ctcgctctgc 7200
taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc gggttggact 7260
caagacgata gttaccggat aaggcgcagc ggtcgggctg aacgggggt tcgtgcacac 7320
agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt gagctatgag 7380
aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc ggcagggtcg 7440
gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt tatagtcctg 7500
tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca gggggggcga 7560
gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt tgctggcctt 7620
ttgctcacat gttctttcct gcgttatccc ctgattctgt ggataaccgt attaccgcct 7680
ttgagtgagc tgataccgct cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg 7740
aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt 7800
aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 7860
atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 7920

tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgatt 7980
acgccaagct ttttctttcc aatttttttt ttttcgtcat tataaaaatc attacgaccg 8040
agattcccgg gtaataactg atataattaa attgaagctc taatttgtga gtttagtata 8100
catgcattta cttataatac agtttttag ttttgctggc cgcatcttct caaatatgct 8160
tcccagcctg cttttctgta acgttcaccc tctaccttag catcccttcc ctttgcaaat 8220
agtcctcttc caacaataat aatgtcagat cctgtagaga ccacatcatc cacggttcta 8280
tactgttgac ccaatgcgtc tcccttgtca tctaaaccca caccgggtgt cataatcaac 8340
caatcgtaac cttcatctct tccacccatg tctctttgag caataaagcc gataacaaaa 8400
tctttgtcgc tcttcgcaat gtcaacagta cccttagtat attctccagt agatagggag 8460
cccttgcatg acaattctgc taacatcaaa aggcctctag gttcctttgt tacttcttct 8520
gccgcctgct tcaaaccgct aacaatacct gggcccacca caccgtgtgc attcgtaatg 8580
tctgcccatt ctgctattct gtatacaccc gcagagtact gcaatttgac tgtattacca 8640
atgtcagcaa attttctgtc ttcgaagagt aaaaaattgt acttggcgga taatgccttt 8700
agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga tatccacatg tgtttttagt 8760
aaacaaattt tgggacctaa tgcttcaact aactccagta attccttggt ggtacgaaca 8820
tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga tattaaatag cttggcagca 8880
acaggactag gatgagtagc agcacgttcc ttatatgtag ctttcgacat gatttatctt 8940
cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga atactgggca atttcatgtt 9000
tcttcaacac tacatatgcg tatatatacc aatctaagtc tgtgctcctt ccttcgttct 9060
tccttctgtt cggagattac cgaatcaaaa aaatttcaag gaaaccgaaa tcaaaaaaaa 9120
gaataaaaaa aaaatgatga attgaaaagc ttgcatgcct gcaggtcgac tctagtatac 9180
tccgtctact gtacgataca cttccgctca ggtccttgtc ctttaacgag gccttaccac 9240
tcttttgtta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc 9300
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac 9360
aatggctgcc atcattatta tccgatgtga cgctgcattt ttttttttt ttttttttt 9420
ttttttttt ttttttttt ttttttttg tacaaatatc ataaaaaaag agaatctttt 9480
taagcaagga ttttcttaac ttcttcggcg acagcatcac cgacttcggt ggtactgttg 9540
gaaccaccta aatcaccagt tctgatacct gcatccaaaa ccttttaac tgcatcttca 9600
atggctttac cttcttcagg caagttcaat gacaatttca acatcattgc agcagacaag 9660
atagtggcga tagggttgac cttattcttt ggcaaatctg gagcggaacc atggcatggt 9720
tcgtacaaac caaatgcggt gttcttgtct ggcaaagagg ccaaggacgc agatggcaac 9780
aaacccaagg agcctgggat aacggaggct tcatcggaga tgatatcacc aaacatgttg 9840
ctggtgatta taataccatt taggtgggtt gggttcttaa ctaggatcat ggcggcagaa 9900
tcaatcaatt gatgttgaac tttcaatgta gggaattcgt tcttgatggt ttcctccaca 9960
gtttttctcc ataatcttga agaggccaaa acattagctt tatccaagga ccaaataggc 10020
aatggtggct catgttgtag ggccatgaaa gcggccattc ttgtgattct ttgcacttct 10080

ggaacggtgt attgttcact atcccaagcg acaccatcac catcgtcttc ctttctctta 10140
ccaaagtaaa tacctcccac taattctcta acaacaacga agtcagtacc tttagcaaat 10200
tgtggcttga ttggagataa gtctaaaaga gagtcggatg caaagttaca tggtcttaag 10260
ttggcgtaca attgaagttc tttacggatt tttagtaaac cttgttcagg tctaacacta 10320
ccggtacccc atttaggacc acccacagca cctaacaaaa cggcatcagc cttcttggag 10380
gcttccagcg cctcatctgg aagtggaaca cctgtagcat cgatagcagc accaccaatt 10440
aaatgatttt cgaaatcgaa cttgacattg gaacgaacat cagaaatagc tttaagaacc 10500
ttaatggctt cggctgtgat ttcttgacca acgtggtcac ctggcaaaac gacgatcttc 10560
ttaggggcag acattacaat ggtatatcct tgaaatatat ataaaaaaaa aaaaaaaaaa 10620
aaaaaaaaaa aatgcagctt ctcaatgata ttcgaatacg ctttgaggag atacagccta 10680
atatccgaca aactgtttta cagatttacg atcgtacttg ttacccatca ttgaattttg 10740
aacatccgaa cctgggagtt ttccctgaaa cagatagtat atttgaacct gtataataat 10800
atatagtcta gcgctttacg gaagacaatg tatgtatttc ggttcctgga gaaactattg 10860
catctattgc ataggtaatc ttgcacgtcg catccccggt tcattttctg cgtttccatc 10920
ttgcacttca atagcatatc tttgttaacg aagcatctgt gcttcatttt gtagaacaaa 10980
aatgcaacgc gagagcgcta atttttcaaa caaagaatct gagctgcatt tttacagaac 11040
agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat ctgtgcttca tttttgtaaa 11100
acaaaaatgc aacgcgagag cgctaatttt tcaaacaaag aatctgagct gcatttttac 11160
agaacagaaa tgcaacgcga gagcgctatt ttaccaacaa agaatctata cttctttttt 11220
gttctacaaa aatgcatccc gagagcgcta tttttctaac aaagcatctt agattacttt 11280
ttttctcctt tgtgcgctct ataatgcagt ctcttgataa ctttttgcac tgtaggtccg 11340
ttaaggttag aagaaggcta ctttggtgtc tattttctct tccataaaaa aagcctgact 11400
ccacttcccg cgtttactga ttactagcga agctgcgggt gcatttttc aagataaagg 11460
catccccgat tatattctat accgatgtgg attgcgcata ctttgtgaac agaaagtgat 11520
agcgttgatg attcttcatt ggtcagaaaa ttatgaacgg tttcttctat tttgtctcta 11580
tatactacgt ataggaaatg tttacatttt cgtattgttt tcgattcact ctatgaatag 11640
ttcttactac aatttttttg tctaaagagt aatactagag ataaacataa aaaatgtaga 11700
ggtcgagttt agatgcaagt tcaaggagcg aaaggtggat gggtaggtta tatagggata 11760
tagcacagag atatatagca aagagatact tttgagcaat gtttgtggaa gcggtattcg 11820
caatatttta gtagctcgtt acagtccggt gcgtttttgg ttttttgaaa gtgcgtcttc 11880
agagcgcttt tggttttcaa aagcgctctg aagttcctat actttctaga gaataggaac 11940
ttcggaatag gaacttcaaa gcgtttccga aaacgagcgc ttccgaaaat gcaacgcgag 12000
ctgcgcacat acagctcact gttcacgtcg cacctatatc tgcgtgttgc ctgtatatat 12060
atatacatga gaagaacggc atagtgcgtg tttatgctta aatgcgtact tatatgcgtc 12120
tatttatgta ggatgaaagg tagtctagta cctcctgtga tattatccca ttccatgcgg 12180
ggtatcgtat gcttccttca gcactaccct ttagctgttc tatatgctgc cactcctcaa 12240

20 Jan 2017 2017200439

ttggattagt ctcatccttc aatgctatca tttcctttga tattggatca tatgcatagt 12300

accgagaaac tagaggatc 12319