

(19) **DANMARK**



Patent- og
Varemærkestyrelsen

(12)

Oversættelse af europæisk patentskrift

(10) **DK/EP 3559234 T3**

-
- (51) Int.Cl.: **C 12 N 15/52 (2006.01)** **C 12 N 1/18 (2006.01)** **C 12 N 1/22 (2006.01)**
C 12 N 9/04 (2006.01) **C 12 N 9/12 (2006.01)** **C 12 P 7/10 (2006.01)**
C 12 R 1/865 (2006.01)
- (45) Oversættelsen bekendtgjort den: **2021-11-15**
- (80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2021-10-27**
- (86) Europæisk ansøgning nr.: **17832062.8**
- (86) Europæisk indleveringsdag: **2017-12-19**
- (87) Den europæiske ansøgnings publiceringsdag: **2019-10-30**
- (86) International ansøgning nr.: **EP2017083586**
- (87) Internationalt publikationsnr.: **WO2018114973**
- (30) Prioritet: **2016-12-20 US 201662436648 P** **2016-12-20 US 201662436723 P**
2016-12-20 US 201662436706 P **2016-12-20 US 201662436687 P**
- (84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**
- (73) Patenthaver: **Novozymes A/S, Krogshøjvej 36, 2880 Bagsværd, Danmark**
- (72) Opfinder: **SENDELIUS, Malin, Mercuriusvägen 22, 24533 Staffanstorps, Sverige**
KARHUMAA, Kaisa, Blåmesvägen 17, 24735 Södra Sandby, Sverige
SÁNCHEZ I NÓGUE, Violeta, Tejevägen19, 22478 Lund, Sverige
- (54) Benævnelse: **Rekombinante gærstammer til pentosefermentering**
- (56) Fremdragne publikationer:
WO-A2-2010/039692
WO-A2-2012/045088
JP-A- 2015 177 760
US-A1- 2012 309 093
C. GUO ET AL: "Cloning and molecular characterization of a gene coding d-xylulokinase (CmXYL3) from Candida maltosa", JOURNAL OF APPLIED MICROBIOLOGY., vol. 101, no. 1, 1 July 2006 (2006-07-01), pages 139-150, XP055459154, GB ISSN: 1364-5072, DOI: 10.1111/j.1365-2672.2006.02915.x
D. J. WOHLBACH ET AL: "Comparative genomics of xylose-fermenting fungi for enhanced biofuel production", PROCEEDINGS NATIONAL ACADEMY OF SCIENCES PNAS, vol. 108, no. 32, 25 July 2011 (2011-07-25), pages 13212-13217, XP055459033, US ISSN: 0027-8424, DOI: 10.1073/pnas.1103039108 -& WOHLBACH ET AL.: "Supplementary Information Appendix for: Comparative genomics of xylose-fermenting fungi for enhanced biofuel production", PNAS, vol. 108, no. 32, 9 August 2011 (2011-08-09), pages 1-29, XP002779114,
KUYPER M ET AL: "Metabolic engineering of a xylose-isomerase-expressing Saccharomyces cerevisiae strain for rapid anaerobic xylose fermentation", FEMS YEAST RESEARCH, WILEY-BLACKWELL PUBLISHING LTD, GB, NL, vol. 5, no. 4-5, 26 November 2004 (2004-11-26), pages 399-409, XP027680444, ISSN: 1567-1356 [retrieved on 2005-02-01]

Fortsættes ...

DESCRIPTION

Field of the Invention

[0001] Described herein are genetically modified recombinant yeast cells and strains capable of fermenting pentoses, as well as to the preparation and use of such cells and strains.

Reference to a Sequence Listing

[0002] This application contains a Sequence Listing in computer readable form.

Background

[0003] Bioethanol production from renewable feedstock by baker's yeast *Saccharomyces cerevisiae* has become an attractive alternative to fossil fuels, but the use of lignocellulosic feedstocks for such purposes poses challenges. For example, a substantial fraction of lignocellulosic material consists of pentoses such as xylose and arabinose. Native *Saccharomyces* species cannot ferment these pentoses, but genetic engineering techniques to provide *Saccharomyces* with this ability are now well-established (Kim *et al.*, 2013). These include heterologous expression of xylose reductase (XR) and xylitol dehydrogenase (XDH) from naturally xylose fermenting yeasts such as *Scheffersomyces (Pichia) stipitis* and various *Candida* species, as well as the overexpression of xylulokinase (XK) and the four enzymes in the non-oxidative pentose phosphate pathway (PPP), namely transketolase (TKL), transaldolase (TAL), ribose-5-phosphate ketol-isomerase (RKI) and D-ribulose-5-phosphate 3-epimerase (RPE). Modifying the co-factor preference of *S. stipitis* XR towards NADH in such systems has been found to provide metabolic advantages as well as improving anaerobic growth, and replacing the XR/XDH with heterologous XI has been reported to reduce unwanted xylitol by-product. These and other modifications have been described in, e.g., WO2009/017441, WO2010/059095, WO2012/135110, Karhumaa *et al.*, 2005; Karhumaa *et al.*, 2007; Kuyper *et al.*, 2005).

[0004] Some degree of overexpression of XK is generally considered necessary for ethanol formation, directing the xylose metabolism towards central metabolism (Chang and Ho, 1988; Eliasson *et al.*, 2001), and XK from various species, e.g., *S. cerevisiae*, *E. coli*, *Pichia stipitis* and *P. tannophilus*, have been used or proposed (WO 95/13362; US 2009/0246857). It has also been reported that XK reduces the production of unwanted xylitol and acetate byproducts (Johansson *et al.*, 2001; Parachin *et al.*, 2011; Matsushika *et al.*, 2011). Several investigators have concluded, however, that it is necessary to limit the levels of XK since XK overexpression inhibited growth of *S. cerevisiae* on xylose (Jin *et al.*, 2003; Matsushika *et al.*, 2011), reduced xylose consumption (Johansson *et al.*, 2001), or drained ATP (Eliasson *et al.*, 2001), suggesting that moderate or low XK levels are needed for optimal xylose fermentation. A lower XK activity could, however, limit the metabolic flux.

[0005] Some yeast species are naturally capable of fermenting xylose, e.g., *Pichia stipites*, *Spathaspora passalidarum*, *Candida jeffriesii* and *Candida tenuis* (Nguyen *et al.*, 2006; US 2012/270289 A1; Wohlbach *et al.*, 2011). Wohlbach and co-workers applied a comparative genomic approach to identify genes involved in xylose metabolism in *Spathaspora passalidarum* and *Candida tenuis*, and found that a *Cten* aldo/keto reductase, CtAKR, significantly improved xylose consumption in engineered *S. cerevisiae* strains during both aerobic and anaerobic growth, although this did not result in improved ethanol production. Hou (2012) then found that the ability of *Spathaspora passalidarum* to utilize xylose under anaerobic conditions was possibly due to the balance of cofactors in the XR-XDH pathway.

[0006] Lignocellulosic hydrolysates contain complex mixtures of other compounds, many of which are inhibitory to microbial fermentation, growth or viability. When the lignocellulosic material is heated during pretreatment, some of the sugars are dehydrated to furans such as furfural (from pentoses) and HMF (from hexoses), which are toxic to most micro-organisms. Further, when the hemicellulose is hydrolysed to release the monomeric sugars, acetic acid is formed by the deacetylation of this fraction. More acids are formed if the lignocellulosic hydrolysate containing furfural and HMF is further heated, since these compounds can degrade into formic and levulinic acids, which are even more potent inhibitors of microorganisms than acetic acid. The toxicity and acidity of the pretreated and hydrolysed lignocellulosic material presents a strong limitation on the fermenting micro-organism.

[0007] WO 2009/017441 describes a mutant of alcohol dehydrogenase (ADH1) from *S. cerevisiae*, which was capable of reducing HMF.

[0008] Mollapour and Piper (Molecular Microbiology (2001) 42(4):919-930) describes that the YME2p gene from the food spoilage yeast *Zygosaccharomyces bailii*, heterologously expressed in *S. cerevisiae* cells, could enable growth of the latter on benzoate, sorbate and phenylalanine.

[0009] For the purpose of further examining the factors improving xylose utilization, Karhumaa *et al.* (2009) compared the proteome of mutant *S. cerevisiae* strain TMB 3400, which has good xylose fermentation properties, with that of its parental strain. Although the level of acetaldehyde dehydrogenase (Ald6) and some other proteins were found to be increased, the most significant changes detected by proteome analysis were 6-10-fold increased levels of XR, XDH and TKL1 in the mutant, which was in accordance with previous knowledge from rational engineering of xylose metabolism in yeast.

[0010] WO 2010/059095 describes that increased levels of phosphoglucomutase obtained by constitutive overexpression of the PGM2 gene with a constitutive promoter improved ethanol production from xylose.

[0011] Tiwari *et al.* (2008) describes that PGM1, PGM2 encode minor and major isozymes of phosphoglucomutase, and further notes that the protein product of YMR278w exhibits phosphoglucomutase activity, designating YMR278w as PGM3. PGM3 has now, however, also been named PRM15 (Xu *et al.*, 2013, Walther *et al.*, 2012), since the enzyme, apart from phosphoglucomutase activity, also has a significant phosphoribomutase activity (Xu *et al.*, 2013).

[0012] Despite these and other advances in the art, there is still a need for improved yeast strains providing for cost-effective production of ethanol and other fermentation products from pentoses such as xylose.

Summary

[0013] Described herein are improved recombinant yeast cells for production of ethanol and other fermentation products. This is based, at least in part, on the discovery of xylulokinases (XKs) which, even if present or expressed at high XK activities, have one or more improved properties such as high aerobic or anaerobic growth rates, increased xylose consumption and/or improved ethanol production.

[0014] Accordingly, in one aspect is a recombinant *Saccharomyces* cell which is capable of fermenting xylose and which comprises a heterologous gene encoding a xylulokinase (XK), wherein the XK:

- provides an enzymatic activity for converting D-xylulose to xylulose 5-phosphate at least twice that provided by *S. cerevisiae* XK disclosed as SEQ ID NO: 32, and
- provides for an anaerobic growth rate of the recombinant cell on xylose which is higher than that provided by a *S. cerevisiae* XK of SEQ ID NO: 32, and
wherein the XK has a sequence identity of at least 80% with the amino acid sequence of SEQ ID NO: 6 or a sequence identity of at least 80% with the amino acid sequence of SEQ ID NO: 22. One aspect is a process for producing a recombinant *Saccharomyces* cell comprising transforming a *Saccharomyces* cell with one or more vectors comprising genes encoding
- an XK comprising the amino acid sequence of SEQ ID NO: 6 or with a sequence identity of at least 80% with the amino acid sequence of SEQ ID NO: 6, such as at least 90%, 95%, 97%, 98%, or 99% with the amino acid sequence of SEQ ID NO: 6, the amino acid sequence of SEQ ID NO: 22, or with a sequence identity of at least 80% with SEQ ID NO: 22, such as at least 90%, 95%, 97%, 98%, or 99% with the amino acid sequence of SEQ ID NO: 22,
- an XR, and
- an XDH, and optionally, regulatory sequences for expressing the genes in the host cell.

[0015] In one aspect is a method for producing a fermentation product, comprising contacting the recombinant cell of of the invention with a medium comprising a carbon source comprising xylose or arabinose under anaerobic conditions, and
(b) isolating the fermentation product from the medium.

[0016] These and other aspects and aspect of the disclosures are described in more detail below.

Brief Description of the Figures

[0017]

Fig. 1: Curves for the aerobic growth on xylose for C5LTe1035 (diamonds) and C5LTe1042 (triangles).

Fig. 2: Curves over anaerobic growth of strain C5LTe1042 ("+") compared with control strain C5LTe1035 ("-").

Fig. 3: Curves over anaerobic growth of strain C5LTe1043 ("+") compared with control strain C5LTe1035 ("-").

Fig. 4: Curves over anaerobic growth of strain C5LTe1040 ("+") compared with control strain C5LTe1035 ("-").

Fig. 5: Graph showing xylose fermentation by strains C5LTe1035 (dashed line) and C5LTe1042 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line.

Fig. 6: Graph showing xylose fermentation by strains C5LTe1036 (dashed line) and C5LTe1048 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line.

Fig. 7: Graph showing xylose fermentation by 5 strains C5LTe1204 (dashed line) and C5LTe1208 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line.

Fig. 8: Graph showing a summary of xylose consumption and ethanol production in anaerobic xylose fermentation of 50 g/l xylose and 20 g/l glucose in mineral medium within 72 hours of fermentation.

Fig. 9: Graph showing change in OD (620nm) in microplate experiments of cells growing on xylose in the presence of various concentrations of acetic (A) and formic (B) acids (g/g, vertical axis).

Fig. 10: Graph showing fermentation results of fermentation of xylose in the presence of acetic acid with strains C5LTe1202 (A) and C5LTe1212 (B).

Fig. 11: Graph showing fermentation profiles of strains C5LTe1202 and C5LTe1212 in mineral medium in the presence of acetic acid.

Fig. 12: Graph showing overlay of xylose consumption (A) and ethanol production (B) profiles of strains C5LTe1202 and C5LTe1212.

Fig. 13: Graph showing fermentation profiles of strains C5LTe1202 (A) and C5LTe1212 (B) in mineral medium in the presence of formic acid.

Fig. 14: Graph showing fermentation results of fermentation of xylose in the presence of formic acid with strains C5LTe1202 and C5LTe1212.

Fig. 15: Graph showing overlay of xylose consumption (A) and ethanol production (B) profiles of strains C5LTe1202 and C5LTe1212.

Fig. 16: Graphic representation of the growth characteristics (measured as change in OD at 620 nm) of strains C5LTe1202 (upper curve) and C5LTe1212 (lower curve).

Fig. 17: Graph showing aerobic growth on xylose of clones expressing various glycolytic genes. Normalized OD (620 nm) for clones with the expressed genes is shown.

Fig. 18: Graph showing anaerobic growth on xylose of clones expressing various glycolytic genes. Normalized OD (620 nm) for clones with the expressed genes is shown, the black bars and the striped bars representing results from two independent experiments.

Fig. 19: Graph showing anaerobic growth of C5LTe1051 ("x") compared with control strain C5LTe1035 ("-").

Fig. 20: Graph showing anaerobic growth of C5LTe1052 compared with control strain C5LTe1035.

Fig. 21: Graph showing anaerobic growth of C5LTe1054 compared with control strain C5LTe1035.

Fig. 22: Graph showing anaerobic growth of C5LTe1055 compared with control strain C5LTe1035.

Fig. 23: Fermentation graphs of xylose fermentation by strains MC2 (A), MC3 (B), MC4 (C), MC11 (D), MC14 (E) and MC22 (F) compared with control strain (G). Symbols: diamond - xylose, square - xylitol, triangle - glycerol, star - ethanol.

Fig. 24: Graphic representation of anaerobic growth on xylose of control strain and strains carrying PGM1 and PGM3.

Fig. 25: Graphic representation of consumed xylose after fermentation in mineral medium for 120h with control strain and strains carrying PGM1 and PGM3.

Fig. 26: Graphic representation of produced ethanol after fermentation in mineral medium for 120h with control strain and strains carrying PGM1 and PGM3.

Definitions

[0018] The term "pathway", "biometabolic pathway" and the like herein refers to an enzymatic pathway present in a cell, typically a yeast cell, which converts or processes an initial substrate to an intermediate or a final product in a series of enzyme-catalyzed reactions.

[0019] The term "gene" refers to a nucleic acid sequence that is capable of being expressed as a specific protein, such as an enzyme, optionally including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence.

[0020] The term "transformation" refers to the transfer of a nucleic acid sequence such as, e.g., a gene, into a host cell, typically a yeast host cell, resulting in genetically stable inheritance.

[0021] As used herein, "recombinant" refers to a host cell into which a nucleic acid sequence, such as a gene, has been transferred, typically by transformation.

[0022] A "yeast" is any of various small, single-celled eukaryotic fungi of the phylum *Ascomycota* that reproduce by fission or budding, and that are capable of fermenting carbohydrates into alcohol and carbon dioxide. Preferably, a yeast cell as used herein refers to a cell of a genus selected from the group consisting of *Saccharomyces*, *Rhodotorula*, *Schizosaccharomyces*, *Kluyveromyces*, *Pichia*, *Hansenula*, *Rhodospiridium*, *Candida*, *Yarrowia*, *Lipomyces*, *Cryptococcus*, and *Dekkera* sp.

[0023] A metabolic pathway, protein, polypeptide, enzyme, nucleic acid sequence or gene may be "heterologous" or "foreign" to a host cell, meaning that the pathway, enzyme, nucleic acid sequence or gene is not normally found in the host cell, typically a yeast host cell of a specific taxonomic classification. "Endogenous" refers to a pathway, protein, polypeptide, enzyme, nucleic acid or gene normally present in the host cell, typically a yeast host cell of a specific taxonomic classification.

[0024] The term "heterologous gene" is defined herein as a gene that is not native to the host cell; a native gene in which structural modifications have been made to the coding region; a native gene whose expression is quantitatively altered (e.g., "overexpressed") as a result of a manipulation of the DNA by recombinant DNA techniques, e.g., a different (foreign) promoter; or a gene in a host cell having one or more extra copies of the coding sequence to quantitatively alter expression. For avoidance of doubt, as used herein, a described *Saccharomyces* gene shall be considered a "heterologous gene" when expressed in a *Saccharomyces* host so long as the gene is not in its native form and is altered by any means as described above (e.g., transformed into the host).

[0025] As used herein, an "overexpressed" gene encoding an enzyme means that the enzyme is produced at a higher level of specific enzymatic activity as compared to the unmodified host cell under identical conditions. Usually this means that the enzymatically active protein (or proteins in case of multi-subunit enzymes) is produced in greater amounts, or rather at a higher steady state level as compared to the unmodified host cell under identical conditions. Typically, this is the result of the mRNA coding for the enzymatically active protein being produced in greater amounts, or again rather at a higher steady state level as compared to the unmodified host cell under identical conditions. Overexpression of an enzyme is thus preferably determined by measuring the level of the enzyme specific activity in the host cell using appropriate enzyme assays as described herein. Alternatively, overexpression of the enzyme may be determined indirectly by quantifying the specific steady state level of enzyme protein, e.g. using antibodies specific for the enzyme, or by quantifying the specific steady level of the mRNA coding for the enzyme. The latter may particularly be suitable for enzymes of the pentose phosphate pathway for which enzymatic assays are not easily feasible as substrates for the enzymes are not commercially available. Preferably, in the host cells described herein, a heterologous gene is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to a host cell which is genetically identical except for the genetic modification causing the overexpression. It is to be understood that these levels of overexpression may apply to the steady state level of the enzyme's activity, the steady state level

of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

[0026] The term "isolated" means a substance in a form or environment that does not occur in nature. Non-limiting examples of isolated substances include (1) any non-naturally occurring substance, (2) any substance including, but not limited to, any host cell, enzyme, variant, nucleic acid, protein, peptide or cofactor, that is at least partially removed from one or more or all of the naturally occurring constituents with which it is associated in nature; (3) any substance modified by the hand of man relative to that substance found in nature; or (4) any substance modified by increasing the amount of the substance relative to other components with which it is naturally associated.

[0027] The term "operably linked" means a configuration in which a control sequence is placed at an appropriate position relative to the coding sequence of a polynucleotide such that the control sequence directs expression of the coding sequence.

[0028] A "variant" of a reference enzyme as used herein is similar in its amino acid sequence to the reference enzyme, such as a "parent" or wild-type enzyme, having an amino acid sequence identity of at least 30%, such as at least 40%, such as at least 50%, such as at least 60%, such as at least 70%, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% to the amino acid sequence of the reference. Enzyme variants can be made by a wide variety of different mutagenesis techniques well known to those skilled in the art. In addition, mutagenesis kits are also available from many commercial molecular biology suppliers. Methods are available to make specific substitutions at defined amino acids (site-directed), specific or random mutations in a localized region of the gene (regio-specific) or random mutagenesis over the entire gene (e.g., saturation mutagenesis). Numerous suitable methods are known to those in the art to generate enzyme variants, including but not limited to site-directed mutagenesis of single-stranded DNA or double-stranded DNA using PCR, cassette mutagenesis, gene synthesis, error-prone PCR, shuffling, and chemical saturation.

[0029] As used herein, a "fragment" of a reference enzyme such as a parent or wild-type enzyme comprises a segment of the reference enzyme amino acid sequence. The amino acid sequence of the fragment typically comprises at least 50%, such as at least 60%, such as at least 70%, such as at least 80%, such as at least 90%, such as at least 95% of the amino acid sequence of the reference enzyme, lacking either an N-terminal portion, a C-terminal portion, or both N-terminal and C-terminal portions of the reference. Typically, a fragment is a catalytically active, at least retaining the enzyme activity of the reference enzyme, though fragments having improved enzyme activity, improved thermostability, altered co-factor dependency, or the like, are also encompassed. The fragment can be designed and expressed using recombinant methods, simply omitting the coding sequences for the relevant N-terminal and/or C-terminal portions.

[0030] An "ortholog" of a wild-type reference enzyme from a particular organism can readily be identified as being similar in its amino acid sequence to the reference enzyme though being encoded by a gene from another organism. Preferred orthologs are yeast orthologs, which can be readily identified by, e.g., searching public genomic sequence databases or screening EST libraries for nucleic acid sequences which hybridize to the wild-type nucleic acid sequence

encoding the reference enzyme under moderate or stringent hybridization conditions. Typically, the ortholog has an amino acid sequence identity of at least 30%, such as at least 40%, such as at least 50%, such as at least 60%, such as at least 70%, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% to the amino acid sequence of the reference enzyme, and is catalytically active such that it at least retains the enzyme activity of the reference enzyme, though orthologs having improved enzyme activity, improved thermostability, altered co-factor dependency, or the like, are also encompassed.

[0031] The relatedness between two amino acid sequences or between two nucleotide sequences is described by the parameter "sequence identity".

[0032] For purposes described herein, the degree of sequence identity between two amino acid sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970) as implemented in the Needle program of the EMBOSS package (Rice et al., 2000), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EBLOSUM62 (EMBOSS version of BLOSUM62) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

$$\frac{(\text{Identical Residues} \times 100)}{(\text{Length of the Referenced Sequence} - \text{Total Number of Gaps in Alignment})}$$

[0033] For purposes described herein, the degree of sequence identity between two deoxyribonucleotide sequences is determined using the Needleman-Wunsch algorithm (Needleman and Wunsch, 1970) as implemented in the Needle program of the EMBOSS package (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferably version 3.0.0 or later. The optional parameters used are gap open penalty of 10, gap extension penalty of 0.5, and the EDNAFULL (EMBOSS version of NCBI NUC4.4) substitution matrix. The output of Needle labeled "longest identity" (obtained using the -nobrief option) is used as the percent identity and is calculated as follows:

$$\frac{(\text{Identical Deoxyribonucleotides} \times 100)}{(\text{Length of Referenced Sequence} - \text{Total Number of Gaps in Alignment})}$$

[0034] The variants, fragments, and orthologs of reference enzymes as described herein are typically "catalytically active", meaning that at least retain the enzyme activity of the reference enzyme, though variants, fragments and orthologs having improved enzyme activity, improved thermostability, altered co-factor dependency, improved affinity, improved catalytic rate, or the like, are also encompassed. For example, a catalytically active variant, fragment or ortholog of *Sp. passalidarum* XK (SEQ ID NO: 6) or *K. marxianus* XK (SEQ ID NO: 22) thus has, e.g., at least twice the enzymatic activity of *S. cerevisiae* XK (SEQ ID NO: 32) for converting D-xylulose to xylulose 5-phosphate, preferably when measured according to Example 7, and provides for an anaerobic growth rate of a recombinant *S. cerevisiae* or other yeast cell on xylose which is higher than that provided by a *S. cerevisiae* XK, preferably when measured according to Example 8. In another example, a functional variant, fragment or ortholog of YME2p from *Zygosaccharomyces bailii* (SEQ ID NO:50) thus provides for one or more, preferably all, of an increased anaerobic

growth, an increased xylose consumption rate, and an increased ethanol production rate of a recombinant yeast cell in the presence of formic acid and/or acetic acid, preferably when measured according to Example 14 or 13, respectively. In another example, a catalytically active variant, fragment or yeast ortholog of *S. cerevisiae* ENO1, PFK2, PFK26, PGI1, GPM1 or TPI1 thus provides for one or more, preferably all, of an increased aerobic growth on xylose, an increased anaerobic growth rate on xylose, an increased xylose consumption rate, and an increased ethanol production rate from xylose, preferably when measured according to Example 17, 18, 23 and/or 24. In another example, a catalytically active variant, fragment or ortholog of *S. cerevisiae* PGM3/PRM15 thus provides for one or more, preferably all, of an increased anaerobic growth on xylose, an increased xylose consumption rate, and an increased ethanol production rate from xylose, preferably when measured according to Example 26 and/or 27 and, typically, substantially retained or improved phosphoglucomutase activity, phosphoribomutase activity, or both. Phosphoglucomutase activity and phosphoribomutase activity can be determined as known in the art (e.g., See Xu *et al.*, 2013).

[0035] Reference to "about" a value or parameter herein includes aspects that are directed to that value or parameter *per se*. For example, description referring to "about X" includes the aspect "X". When used in combination with measured values, "about" includes a range that encompasses at least the uncertainty associated with the method of measuring the particular value, and can include a range of plus or minus two standard deviations around the stated value.

[0036] As used herein and in the appended claims, the singular forms "a," "or," and "the" include plural referents unless the context clearly dictates otherwise. It is understood that the aspects described herein include "consisting" and/or "consisting essentially of" aspects.

[0037] Unless defined otherwise or clearly indicated by context, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art.

Detailed Description

Recombinant cells

[0038] Table 1 below summarizes some key results from the Examples. When tested in cell extracts of recombinant *S. cerevisiae* cells capable of fermenting xylose, the extracts from those cells overexpressing *Sp. passalidarum* XK (SEQ ID NO: 6) had a consistently higher XK activity than the control strains overexpressing *S. cerevisiae* XK (SEQ ID NO: 32). Surprisingly, the anaerobic growth rate on xylose of the cells overexpressing XK from either *Sp. passalidarum* or *K. marxianus* (SEQ ID NO: 22) was nonetheless higher than those of the control strains overexpressing *S. cerevisiae* XK or *E. coli* XK (SEQ ID NO: 18). In addition, the cell overexpressing *Sp. passalidarum* XK also provided for higher xylose consumption and ethanol production in fermentation on xylose than the corresponding *S. cerevisiae* XK control strains.

Table 1. Overview of strains constructed and tested as described in the Examples.

Strain	Heterologous genes	XK activity (U/mg)	Aerobic growth on xylose (h ⁻¹)	Anaerobic growth on xylose (h ⁻¹)	Anaerobic fermentation on xylose 72h	
					Consumed xylose (g/L)	Produced ethanol (g/L)
C5LTe1000	None	0.11 ± 0.00				
C5LTe1001	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1035	<i>P. stipitis</i> mXR(N272D)	1.65 ± 0.27	0.16		42	21
	<i>P. stipitis</i> XDH					
	<i>S. cerevisiae</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1036	<i>P. stipitis</i> mXR(N272D)	0.84 ± 0.20			21	15
	<i>P. stipitis</i> XDH					
	<i>S. cerevisiae</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1040	<i>P. stipitis</i> mXR(N272D)			Slower than e1035		
	<i>P. stipitis</i> XDH					
	<i>E. coli</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i>					

Strain	Heterologous genes	XK activity (U/mg)	Aerobic growth on xylose (h ⁻¹)	Anaerobic growth on xylose (h ⁻¹)	Anaerobic fermentation on xylose 72h	
					Consumed xylose (g/L)	Produced ethanol (g/L)
	TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1042	<i>P. stipitis</i> mXR(N272D)	5.21 ± 0.37	0.19	Faster than e1035	46	23
	<i>P. stipitis</i> XDH					
	<i>S. passalidarum</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1043	<i>P. stipitis</i> mXR(N272D)			Faster than e1035		
	<i>P. stipitis</i> XDH					
	<i>K. marxianus</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1048	<i>P. stipitis</i> mXR(N272D)	4.31 ± 0.06			35	19
	<i>P. stipitis</i> XDH					
	<i>S. passalidarum</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i>					

Strain	Heterologous genes	XK activity (U/mg)	Aerobic growth on xylose (h ⁻¹)	Anaerobic growth on xylose (h ⁻¹)	Anaerobic fermentation on xylose 72h	
					Consumed xylose (g/L)	Produced ethanol (g/L)
	RKI1					
C5LTe1204	<i>P. stipitis</i> mXR(N272D)	0.92 ± 0.27			25	15
	<i>P. stipitis</i> XDH					
	<i>S. cerevisiae</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe1208	<i>P. stipitis</i> mXR(N272D)	3.67 ± 0.19			37	19
	<i>P. stipitis</i> XDH					
	<i>S. passalidarum</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					

[0039] Accordingly, in one aspect of the disclosure, is a recombinant yeast *Saccharomyces* cell, optionally a *S. cerevisiae* cell, which is capable of fermenting xylose and which comprises a heterologous gene encoding a polypeptide having enzymatic xylulokinase (XK) activity.

[0040] In one aspect of the disclosure, is a recombinant yeast *Saccharomyces* cell, optionally a *S. cerevisiae* cell, which is capable of fermenting xylose and which comprises a heterologous gene encoding a xylulokinase (XK) providing an enzymatic activity for converting D-xylulose to xylulose 5-phosphate which is higher than that provided by *S. cerevisiae* XK (SEQ ID NO: 32) and yet provides for an anaerobic growth rate which is higher than that provided by *S. cerevisiae* XK, *E. coli* XK, or both.

[0041] In one aspect of the disclosure, is a recombinant *Saccharomyces* cell which is capable of

fermenting xylose and which comprises a heterologous gene encoding a xylulokinase (XK), wherein the XK has an enzymatic XK activity, which is higher than that of *S. cerevisiae* XK, and provides for an anaerobic growth rate of the recombinant cell on xylose which is higher than that provided by a *S. cerevisiae* XK.

[0042] In one aspect of the disclosure, is a recombinant yeast *Saccharomyces* cell, optionally a *S. cerevisiae* cell, capable of fermenting xylose and which comprises a heterologous gene encoding a xylulokinase (XK) comprising SEQ ID NO: 6 or a catalytically active variant thereof as defined in the claims.

[0043] In separate and specific aspects of the disclosure of any aforementioned aspect, the XK provides for an enzymatic XK activity which is at least 1.1, e.g., at least 1.2, at least 1.5, at least 1.7, at least twice (2), at least 2.5 or at least 3 times that provided by *S. cerevisiae* XK. In one aspect of the disclosure, the XK has an enzymatic activity for converting D-xylulose to xylulose 5-phosphate at least twice that of *S. cerevisiae* XK. In one aspect of the disclosure, the XK further provides for an aerobic growth rate of the cell on xylose which is higher than that provided by *S. cerevisiae* XK, *E. coli* XK, or both. For example, the XK may provide for an aerobic growth rate which is 1.1, 1.2, 1.3, 1.4 or 1.5 times that provided by *S. cerevisiae* XK. Additionally, the XK may provide for a higher xylose consumption, a higher ethanol production, or both, of the recombinant cell than that provided by *S. cerevisiae* XK, *E. coli* XK or both.

[0044] Advantageously, the XK activity, anaerobic and aerobic growth rates, and xylose consumption and ethanol production, can be tested in the assays and strain constructs according to the Examples. For example, in one aspect of the disclosure, the enzymatic activity for converting D-xylulose to xylulose 5-phosphate is measured according to Example 7. In an additional or alternative aspect of the disclosure, the anaerobic growth rate can be measured according to Example 9. Likewise, the aerobic growth rate can be measured according to Example 8, and/or the xylose consumption and ethanol production can be measured according to Example 10. In these tests, the recombinant strains can, for example, be prepared from C5LTe1000 or an equivalent or similar laboratory or commercially available *S. cerevisiae* strain, for example CEN.PK or s288c, and then tested in the form of live cells or cell extracts as described. Typically, for conducting such test assays, *S. cerevisiae* cells are transformed with genes encoding *P. stipitis* mXR with an N272D mutation, *P. stipitis* XDH, the XK to be examined or control *S. cerevisiae* or *E. coli* XK, and *S. cerevisiae* TAL1, TKL1, and RKI1, so that each gene is chromosomally integrated and operably linked to a constitutive promoter.

[0045] In one aspect of the disclosure, the XK comprises the amino acid sequence of SEQ ID NO: 6, the amino acid sequence of SEQ ID NO: 22 or a catalytically active variant or fragment of any thereof as defined in the claims.

[0046] In one aspect of the disclosure, the XK is a catalytically active variant of SEQ ID NO: 6, comprising an amino acid sequence which is at least 80%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99% identical to SEQ ID NO: 6. In one aspect of the disclosure, the XK is a variant of SEQ ID NO: 22, comprising an amino acid sequence which is at least 80%, such as at least 90%, such as at least 95%, such as at least 98%, such as at least 99%

identical to SEQ ID NO: 6. In one aspect of the disclosure, the XK is a fragment (e.g., a catalytically active fragment) of SEQ ID NO: 6 corresponding to at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the full-length enzyme amino acid sequence. In one aspect of the disclosure, the XK is a fragment (e.g., a catalytically active fragment) of SEQ ID NO: 22 corresponding to at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the full-length enzyme amino acid sequence.

[0047] Table 2 below additionally summarizes additional key results from the Examples. Surprisingly, expressing the heterologous YME2 gene (*i.e.*, overexpressing the YME2 gene) in recombinant *S. cerevisiae* cells capable of fermenting xylose resulted in a higher anaerobic growth rate, more efficient xylose consumption and increased ethanol production in fermentation on xylose.

Table 2. Overview of strains constructed and tested in the presence of formic acid as described in the Examples. The growth and fermentation experiments reported here were conducted on xylose and under anaerobic conditions.

Strain	Heterologous genes	Growth in the presence of formic acid (h ⁻¹)	Fermentation in the presence of acetic acid, 144h		Fermentation in the presence of formic acid, 144h	
			Consumed xylose (g/L)	Consumed ethanol (g/L)	Consumed xylose (g/L)	Consumed ethanol (g/L)
C5LTe 1202	<i>P. stipitis</i> mXR(N272D)	0.03	37.6	25.4	42.7	27.1
	<i>P. stipitis</i> ^a XDH					
	<i>S. cerevisiae</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					
	<i>S. cerevisiae</i> RKI1					
C5LTe 1212	<i>P. stipitis</i> mXR(N272D)	0.05	42.4	28.3	45.8	30.3
	<i>P. stipitis</i> ^a XDH					
	<i>S. cerevisiae</i> XK					
	<i>S. cerevisiae</i> TAL1					
	<i>S. cerevisiae</i> TKL1					

Strain	Heterologous genes	Growth in the presence of formic acid (h ⁻¹)	Fermentation in the presence of acetic acid, 144h		Fermentation in the presence of formic acid, 144h	
			Consumed xylose (g/L)	Consumed ethanol (g/L)	Consumed xylose (g/L)	Consumed ethanol (g/L)
	<i>S. cerevisiae</i> RKI1					
	<i>Z. bailii</i> YME2					

[0048] Accordingly, in one aspect of the disclosure is a recombinant yeast *Saccharomyces* cell, optionally a *S. cerevisiae* cell, which is capable of fermenting xylose and which comprises a heterologous gene encoding an Yme2p polypeptide comprising the amino acid sequence of SEQ ID NO: 50 or a catalytically active variant, fragment or yeast ortholog thereof, wherein the Yme2p polypeptide provides for an increased tolerance of the recombinant cell to formic acid, acetic acid, or both. In one aspect of the disclosure, the Yme2p polypeptide further provides for (a) an increased anaerobic growth, (b) an increased xylose consumption rate, (c) an increased ethanol production rate, or (d) a combination of two or all of (a) to (c), of the recombinant cell in the presence of formic acid. In one aspect of the disclosure, the Yme2p polypeptide further provides for (a) an increased xylose consumption rate, (b) an increased ethanol production rate, or (c) a combination of (a) and (b), of the recombinant cell in the presence of acetic acid.

[0049] Advantageously, the anaerobic growth, xylose consumption, ethanol production rate, or combination can be measured according to Examples 13 or 14. Notably, these Example show that xylose fermentation in presence of formic or acetic acid was improved in yeast overexpressing YME2. Specifically, xylose consumption and ethanol production increased by 13% and 12%, respectively, in the presence of acetic acid, and by 7% and 12%, respectively, in the presence of formic acid. Formic acid is common in lignocellulosic hydrolysates, and strongly contributes to the toxicity of such hydrolysates.

[0050] In separate aspect of the disclosure, the Yme2p polypeptide comprising the amino acid sequence of SEQ ID NO: 50 or a functional variant, fragment or yeast ortholog thereof, provides for a xylose consumption which is at least 5%, e.g., at least 10% higher, an ethanol production which is at least 5%, e.g., at least 10% higher, or both, of the recombinant cell when tested according to Example 14. In these tests, the recombinant strains can, for example, be prepared from TMB 3000 or an equivalent or similar laboratory or commercially available *S. cerevisiae* strain, and then tested in as described. Typically, for conducting such test assays, *S. cerevisiae* cells are transformed with genes encoding *P. stipitis* mXR with an N272D mutation, *P. stipitis* XDH, the Yme2p polypeptide to be examined, and *S. cerevisiae* XK, TAL1, TKL1, and RKI1, so that each gene is chromosomally integrated and operably linked to a constitutive promoter.

[0051] In one aspect of the disclosure, the Yme2p polypeptide comprises the amino acid sequence of SEQ ID NO: 50 or a functional variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the Yme2p polypeptide is a functional variant or yeast ortholog of the

amino acid sequence of SEQ ID NO: 50, comprising an amino acid sequence which is at least 50%, such as at least 60%, such as at least 70%, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 97%, such as at least 98%, such as at least 99% identical to the amino acid sequence of SEQ ID NO: 50. In one aspect of the disclosure, the Yme2p polypeptide is a functional fragment of the amino acid sequence of SEQ ID NO: 50 corresponding to at least 50%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the full-length enzyme amino acid sequence.

[0052] Tables 3 and 8 below additionally summarize some key results from the Examples, showing that overexpression of genes encoding ENO1, PFK2, PFK26, PGI1, GPM1 or TPI1 in recombinant *S. cerevisiae* cells capable of fermenting xylose resulted in a higher anaerobic growth rate, more efficient xylose consumption and/or increased ethanol production in fermentation on xylose.

Table 3. Overview of strains constructed and tested as described in the Examples. See also Table 8, for results on xylose consumption and ethanol production.

Strain	Heterologous genes	Aerobic growth on xylose	Anaerobic growth on xylose (h ⁻¹)	Xylose consumed in 140h, (g/L)	Ethanol produced in 140h (g/L)
C5LTe1035	<i>P. stipitis</i> mXR(N272D)		0.058		
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
C5LTe1051	<i>P. stipitis</i> mXR(N272D)		0.066		
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i>				

Strain	Heterologous genes	Aerobic growth on xylose	Anaerobic growth on xylose (h ⁻¹)	Xylose consumed in 140h, (g/L)	Ethanol produced in 140h (g/L)
	ENO1				
C5LTe1052	<i>P. stipitis</i> mXR(N272D)		0.063		
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> PFK2				
C5LTe1054	<i>P. stipitis</i> mXR(N272D)		0.071		
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> PGI1				
C5LTe1055	<i>P. stipitis</i> mXR(N272D)		0.075		
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				

Strain	Heterologous genes	Aerobic growth on xylose	Anaerobic growth on xylose (h ⁻¹)	Xylose consumed in 140h, (g/L)	Ethanol produced in 140h (g/L)
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> PFK26				
MC2	<i>P. stipitis</i> mXR(N272D)	Faster than average	Faster than average	51	18
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> PGI1				
MC3	<i>P. stipitis</i> mXR(N272D)	Slower than average	Faster than average	53	21
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> PFK26				
MC4	<i>P. stipitis</i> mXR(N272D)	Faster than average	Faster than average	50	22
	<i>P. stipitis</i> ^a				

Strain	Heterologous genes	Aerobic growth on xylose	Anaerobic growth on xylose (h ⁻¹)	Xylose consumed in 140h, (g/L)	Ethanol produced in 140h (g/L)
	XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> PFK2				
MC11	<i>P. stipitis</i> mXR(N272D)	Slower than average	Faster than average	50	18
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> PFK2				
MC14	<i>P. stipitis</i> mXR(N272D)	Faster than average	Faster than average	54	20
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				

Strain	Heterologous genes	Aerobic growth on xylose	Anaerobic growth on xylose (h ⁻¹)	Xylose consumed in 140h, (g/L)	Ethanol produced in 140h (g/L)
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> ENO1				
MC22	<i>P. stipitis</i> mXR(N272D)	Slower than average	Faster than average	51	18
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	<i>S. cerevisiae</i> ENO1				
Control	<i>P. stipitis</i> mXR(N272D)	--	--	36	13
	<i>P. stipitis</i> ^a XDH				
	<i>S. cerevisiae</i> XK				
	<i>S. cerevisiae</i> TAL1				
	<i>S. cerevisiae</i> TKL1				
	<i>S. cerevisiae</i> RKI1				
	<i>S. cerevisiae</i> RPE1				
	URA3+				

[0053] Accordingly, in one aspect is a recombinant yeast cell such as a *Saccharomyces* cell, optionally a *S. cerevisiae* cell, which is capable of fermenting xylose and which comprises a heterologous gene encoding an enolase comprising the amino acid sequence of SEQ ID NO: 132 (ENO1), a phosphofructokinase beta subunit polypeptide comprising the amino acid sequence of SEQ ID NO: 134 (PFK2), a 6-phosphofructo-2-kinase comprising the amino acid sequence of SEQ ID NO: 136 (PFK26), a glucose-6-phosphate isomerase comprising the amino acid sequence of SEQ ID NO: 138 (PGI1), a phosphoglycerate mutase comprising the amino acid sequence of SEQ ID NO: 140 (GPM1), a triose-phosphate isomerase comprising the amino acid sequence of SEQ ID NO: 142 (TPI1), or a catalytically active variant, fragment or yeast ortholog of any one of ENO1, PFK2, PFK26, PGI1, GPM1 and TPI1, or a combination of any two or more of ENO1, PFK2, PFK26, PGI1, GPM1 and TPI1 thereof, such as three, four, five or all of ENO1, PFK2, PFK26, PGI1, GPM1 and TPI1. In one aspect of the disclosure, the heterologous gene or combination provides for an increased anaerobic growth rate on xylose, an increased aerobic growth rate on xylose, an increased ethanol production from xylose, or a combination of any two or all thereof. Advantageously, the anaerobic growth rate can be measured according to Example 23, the aerobic growth can be measured according to Example 5, and/or the ethanol production can be measured according to Example 24.

[0054] In separate aspect of the disclosure, the ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 or catalytically active variant, fragment or yeast ortholog of any thereof, provides for an ethanol yield from xylose which is at least 5%, e.g., at least 10% higher, at least 15% higher, or at least 20% higher, of the recombinant cell when tested according to Example 12. In these tests, the recombinant strains can, for example, be prepared from CEN.PK or an equivalent or similar laboratory or commercially available *S. cerevisiae* strain, and then tested in as described. Typically, for conducting such test assays, *S. cerevisiae* cells are transformed with genes encoding *P. stipitis* mXR with an N272D mutation, *P. stipitis* XDH, the ENO1, PFK2, PFK26, PGI1, GPM1, or TPI1 to be examined, and *S. cerevisiae* XK, TAL1, TKL1, and RKI1, so that each gene is chromosomally integrated and operably linked to a constitutive promoter.

[0055] In one aspect of the disclosure, the cell comprises a heterologous gene encoding an enolase comprising the amino acid sequence of SEQ ID NO: 132 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 132. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 132.

[0056] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a phosphofructokinase beta subunit polypeptide comprising the amino acid sequence of SEQ ID NO: 134 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 134. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 134.

[0057] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a 6-

phosphofructo-2-kinase comprising the amino acid sequence of SEQ ID NO: 136 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 136. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 136.

[0058] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a glucose-6-phosphate isomerase comprising the amino acid sequence of SEQ ID NO: 138 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 138. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 138.

[0059] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a phosphoglycerate mutase comprising the amino acid sequence of SEQ ID NO: 140 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 140. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 140.

[0060] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a triose-phosphate isomerase comprising the amino acid sequence of SEQ ID NO: 142 or a catalytically active variant, fragment or yeast ortholog thereof. In one aspect of the disclosure, the catalytically active variant or yeast ortholog has a sequence identity of at least 80%, e.g., at least 90%, at least 95%, at least 97%, at least 98%, or at least 99% to the amino acid sequence of SEQ ID NO: 142. In one aspect of the disclosure, the heterologous gene encodes the amino acid sequence of SEQ ID NO: 142.

[0061] In one aspect of the disclosure, the heterologous gene encodes a fragment of the amino acid sequence of SEQ ID NO: 132, 134, 136, 138, 140 or 142 corresponding to at least 50%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the full-length amino acid sequence.

[0062] Table 4 below additionally summarizes some key results from the Examples, showing that overexpression of PGM3 in recombinant *S. cerevisiae* cells capable of fermenting xylose resulted in a higher anaerobic growth rate, and more efficient xylose consumption and increased ethanol production in anaerobic fermentation of xylose.

Table 4. *Overview of strains constructed and tested as described in the Examples.*

Strain	Heterologous genes	Anaerobic growth on xylose (h ⁻¹)	Anaerobic growth on xylose 120h, consumed xylose (g/L)	Anaerobic growth on xylose 120h, produced ethanol (g/L)
Parent	<i>P. stipitis</i> mXR(N272D)	0.010	20	6.6
strain	<i>P. stipitis</i> XDH			
	<i>S. cerevisiae</i> XK			
	<i>S. cerevisiae</i> TAL1			
	<i>S. cerevisiae</i> TKL1			
	<i>S. cerevisiae</i> RKI1			
	<i>S. cerevisiae</i> RPE1			
PGM1	<i>P. stipitis</i> mXR(N272D)	0.009	24	7.3
	<i>P. stipitis</i> XDH			
	<i>S. cerevisiae</i> XK			
	<i>S. cerevisiae</i> TAL1			
	<i>S. cerevisiae</i> TKL1			
	<i>S. cerevisiae</i> RKI1			
	<i>S. cerevisiae</i> RPE1			
	<i>S. cerevisiae</i> PGM1			
PGM3	<i>P. stipitis</i> mXR(N272D)	0.019	40	14.3
	<i>P. stipitis</i> XDH			
	<i>S. cerevisiae</i> XK			
	<i>S. cerevisiae</i> TAL1			
	<i>S. cerevisiae</i> TKL1			
	<i>S. cerevisiae</i> RKI1			

Strain	Heterologous genes	Anaerobic growth on xylose (h ⁻¹)	Anaerobic growth on xylose 120h, consumed xylose (g/L)	Anaerobic growth on xylose 120h, produced ethanol (g/L)
	<i>S. cerevisiae</i> RPE1			
	<i>S. cerevisiae</i> PGM3			

[0063] Accordingly, in one aspect is a recombinant yeast *Saccharomyces* cell optionally a *S. cerevisiae* cell, which is capable of fermenting xylose and which comprises an heterologous gene encoding a phosphoglucumutase and/or phosphoribomutase comprising the amino acid sequence of SEQ ID NO: 150 (PGM3) or a catalytically active variant or fragment thereof. In one aspect of the disclosure, the heterologous gene or combination provides for an increased anaerobic growth rate on xylose, an increased xylose consumption, an increased ethanol production from xylose, or a combination of any two or all thereof.

[0064] In separate aspect of the disclosure, the PGM3 or catalytically active fragment or variant provides for an anaerobic growth rate of the recombinant cell which is at least 10%, e.g., at least 20%, at least 50%, at least 80%, or at least 100% higher than a relevant control not comprising the heterologous PGM3 gene (in which the PGM3 is not overexpressed), (e.g., when tested according to Example 26). In one aspect of the disclosure, the PGM3 or catalytically active fragment or variant also or alternatively provides for a xylose consumption rate in anaerobic fermentation which is at least 10%, e.g., at least 20%, at least 50%, at least 80%, or at least 100% higher than a relevant control not comprising the heterologous PGM3 gene (in which the PGM3 is not overexpressed), (e.g., when tested according to Example 27). In aspect of the disclosure, the PGM3 or catalytically active fragment or variant also or alternatively provides for an ethanol yield in anaerobic fermentation which is at least 10%, e.g., at least 20%, at least 50%, at least 80%, or at least 100% higher than a relevant control not comprising the heterologous PGM3 gene (in which the PGM3 is not overexpressed), (e.g., when tested according to Example 27). In these tests, the recombinant strains can, for example, be prepared from CEN.PK or an equivalent or similar laboratory or commercially available *S. cerevisiae* strain, and then tested in as described. Typically, for conducting such test assays, *S. cerevisiae* cells are transformed with genes encoding *P. stipitis* mXR with an N272D mutation, *P. stipitis* XDH, the PGM3 or catalytically active fragment or variant to be examined, and *S. cerevisiae* XK, TAL1, TKL1, and RKI1, so that each gene is chromosomally integrated and operably linked to a constitutive promoter.

[0065] In one aspect of the disclosure, the cell comprises a heterologous gene encoding a PGM3 comprising the amino acid sequence of SEQ ID NO: 150 or a catalytically active variant or fragment thereof. In one aspect of the disclosure, the catalytically active variant has a sequence identity of at least 30%, such as at least 50%, such as at least 80%, such as at least 90%, such as at least 95%, such as at least 97%, such as at least 98%, such as at least 99% to the amino acid sequence of SEQ ID NO: 150. In one aspect of the disclosure, the heterologous gene encodes a PGM3 comprising the amino acid sequence of SEQ ID NO: 150.

[0066] In one aspect of the disclosure, the heterologous gene encoding PGM3 or catalytically active fragment or variant encodes a fragment of SEQ ID NO: 150 corresponding to at least 50%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% of the full-length amino acid sequence.

[0067] The skilled artisan is well aware that variants and fragments of an enzyme sequence can be modified by replacing, inserting, or deleting amino acids using standard recombinant techniques while still retaining, or even improving, the enzyme activity of interest. Although such variants include those having amino acid sequences with one or more conservative or nonconservative substitutions relative to the amino acid sequence of SEQ ID NO: 6, 22, 50, 132, 134, 136, 138, 140, 142 and/or 150, conservative substitutions are typically of most interest. As used herein, the term "conservative substitution" refers to the substitution of a residue for another residue that does not generally alter the specific activity of the encoded polypeptide. An exemplary conservative substitution is a substitution that is within the same group of basic amino acids (arginine, lysine and histidine), acidic amino acids (glutamic acid and aspartic acid), polar amino acids (glutamine and asparagine), hydrophobic amino acids (leucine, isoleucine and valine), aromatic amino acids (phenylalanine, tryptophan and tyrosine), and small amino acids (glycine, alanine, serine, threonine, proline, cysteine and methionine). Amino acid substitutions that do not generally alter the specific activity are well-known in the art. The most commonly occurring exchanges are Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, and Asp/Gly, as well as these in reverse. In some aspect of the disclosures, the substitutions are of a low percentage, typically less than about 10%, more typically less than 5%, and often less than about 2% of the amino acids of the polypeptide sequence, with a conservatively selected amino acid of the same conservative substitution group. In one preferred aspect of the disclosure, the XK comprises SEQ ID NO: 6. In another preferred aspect of the disclosure, the XK comprises SEQ ID NO: 22. In another preferred aspect of the disclosure, the Yme2p comprises the amino acid sequence of SEQ ID NO: 50. In another preferred aspect of the disclosure, the enzyme comprises the amino acid sequence of SEQ ID NO: 134, 136, 138, 140, 142 or 150.

[0068] Essential amino acids of enzymes can be identified according to procedures known in the art, such as site-directed mutagenesis or alanine-scanning mutagenesis (Cunningham and Wells, 1989, *Science* 244: 1081-1085). In the latter technique, single alanine mutations are introduced at every residue in the molecule, and the resultant mutant molecules are tested for enzymatic activity to identify amino acid residues that are critical to the activity of the molecule. See also, Hilton et al., 1996, *J. Biol. Chem.* 271: 4699-4708. The active site of the enzyme or other biological interaction can also be determined by physical analysis of structure, as determined by such techniques as nuclear magnetic resonance, crystallography, electron diffraction, or photoaffinity labeling, in conjunction with mutation of putative contact site amino acids. See, for example, de Vos et al., 1992, *Science* 255: 306-312; Smith et al., 1992, *J. Mol. Biol.* 224: 899-904; Wlodaver et al., 1992, *FEBS Lett.* 309: 59-64. The identities of essential amino acids can also be inferred from analysis of identities with other enzymes that are related to the referenced enzymes.

[0069] Additional guidance on the structure-activity relationship of the enzymes herein can be

determined using multiple sequence alignment (MSA) techniques well-known in the art. Based on the teachings herein, the skilled artisan could make similar alignments with any number of enzymes described herein or known in the art. Such alignments aid the skilled artisan to determine potentially relevant domains (e.g., binding domains or catalytic domains), as well as which amino acid residues are conserved and not conserved among the different enzyme sequences. It is appreciated in the art that changing an amino acid that is conserved at a particular position between the disclosed enzyme will more likely result in a change in biological activity (Bowie et al., 1990, Science 247: 1306-1310: "Residues that are directly involved in protein functions such as binding or catalysis will certainly be among the most conserved"). In contrast, substituting an amino acid that is not highly conserved among the enzymes will not likely or significantly alter the biological activity.

[0070] Single or multiple amino acid substitutions, deletions, and/or insertions can be made and tested using known methods of mutagenesis, recombination, and/or shuffling, followed by a relevant screening procedure, such as those disclosed by Reidhaar-Olson and Sauer, 1988, Science 241: 53-57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; or WO 95/22625. Other methods that can be used include error-prone PCR, phage display (e.g., Lowman et al., 1991, Biochemistry 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), and region-directed mutagenesis (Derbyshire et al., 1986, Gene 46: 145; Ner et al., 1988, DNA 7: 127).

[0071] Mutagenesis/shuffling methods can be combined with high-throughput, automated screening methods to detect activity of cloned, mutagenized polypeptides expressed by host cells (Ness et al., 1999, Nature Biotechnology 17: 893-896). Mutagenized DNA molecules that encode active enzymes (e.g., xylulokinases) can be recovered from the host cells and rapidly sequenced using standard methods in the art. These methods allow the rapid determination of the importance of individual amino acid residues in a polypeptide.

[0072] The host cells for preparing the recombinant cells of the invention are *Saccharomyces* host cells such as *Saccharomyces cerevisiae*, *bayanus* or *carlsbergensis* cells. Preferably, the yeast cell is a *Saccharomyces cerevisiae* cell. Suitable cells can, for example, be derived from commercially available strains and polyploid or aneuploid industrial strains, including but not limited to those from Superstart™, THERMOSACC®, etc. (Lallemand); RED STAR and ETHANOL RED® (Fermentis/Lesaffre); FALI (AB Mauri); Baker's Best Yeast, Baker's Compressed Yeast, etc. (Fleishmann's Yeast); BIOFERM AFT, XP, CF, and XR (North American Bioproducts Corp.); Turbo Yeast (Gert Strand AB); and FERMIOL® (DSM Specialties). Other useful yeast strains are available from biological depositories such as the American Type Culture Collection (ATCC) or the Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), such as, e.g., BY4741 (e.g., ATCC 201388); Y108-1 (ATCC PTA.10567) and NRRL YB-1952 (ARS Culture Collection). Still other *S. cerevisiae* strains suitable as host cells DBY746, [Alpha][Eta]22, S150-2B, GPY55-15Ba, CEN.PK, USM21, TMB3500, TMB3400, VTT-A-63015, VTT-A-85068, VTT-c-79093 and their derivatives as well as *Saccharomyces* sp. 1400, 424A (LNH-ST), 259A (LNH-ST) and derivatives thereof.

[0073] As previously mentioned, some wild-type yeast cells, e.g., *Saccharomyces* cells, cannot

naturally ferment xylose. However, it is now well within the level of skill in the art to apply genetic engineering technology to prepare recombinant *Saccharomyces* cells which are capable of fermenting xylose. For example, XR and XDH enzymes from naturally xylose fermenting yeasts such as *Scheffersomyces (Pichia) stipitis* and various *Candida* species can be expressed in *Saccharomyces* cells to provide this ability and/or an XI enzyme suitable for expression in a yeast host cell. Additionally, it is contemplated to use catalytically active variants of XR and/or XDH. For example, variants of *P. stipitis* XR exist which change the cofactor preference of the XR from NADPH to NADH. These variants are referred to herein as "NADH-preferring" XR variants, and include, but are not limited to, N272D, K270R and P275Q. Still other variants of XR and XK are described in WO 2012/135110. As for XI, XI from the fungus *Piromyces sp.* (Kuyper *et al.*, 2005) or other sources (Madhavan *et al.*, 2009) have been expressed in *S. cerevisiae* host cells. Still other XIs suitable for expression in yeast have been described in US 2012/0184020 (an XI from *Ruminococcus flavefaciens*), WO 2011/078262 (several XIs from *Reticulitermes speratus* and *Mastotermes darwiniensis*) and WO 2012/009272 (constructs and fungal cells containing an XI from *Abiotrophia defectiva*). Optionally, additional improvements in xylose fermentation can also be achieved by strain adaptation to selective conditions, according to techniques known in the art. Additionally, the xylose fermentative capability of a yeast cell can also be increased by increasing the flux of the pentose-phosphate pathway (PPP) by overexpressing one or more genes encoding enzymes of the non-oxidative pathway, which includes TAL (EC 2.2.1.2), TKL (EC 2.2.1.1), RKI (EC 5.3.1.6) and RPE (EC 5.1.3.4) (Karhumaa *et al.*, 2005). Preferably, in the yeast cells of the invention, the genes encoding TAL, TKL and RKI are overexpressed. Typically, although not necessarily, the endogenous genes of the PPP are overexpressed. An increased flux in the PPP can be measured by metabolic flux analysis with ¹³C-labeled glucose as described in, e.g., van Winden *et al.* (2005, FEMS Yeast Research 5:559-568).

[0074] Accordingly, in one aspect of the disclosure, the recombinant cell comprises a heterologous gene encoding a transaldolase (TAL), a heterologous gene encoding a transketolase (TKL), and/or a heterologous gene encoding a ribose 5-phosphate ketol-isomerase (RKI). In some aspect of the disclosures, the recombinant cell comprises a heterologous gene encoding a transaldolase (TAL), a heterologous gene encoding a transketolase (TKL), and a heterologous gene encoding a ribose 5-phosphate ketol-isomerase (RKI).

[0075] In one aspect of the disclosure, the recombinant cell comprises a heterologous gene encoding a xylose reductase (XR), a heterologous gene encoding a xylitol dehydrogenase (XDH), and/or a heterologous gene encoding a xylose isomerase (XI). Preferred XRs are *Pichia stipitis* XR and NADH-preferring variants thereof, such as *Pichia stipites* XR comprising one or more amino acid substitutions selected from N272D, K270R and P275Q. Most preferred is *Pichia stipitis* XR(N272D). Preferred XDHs are *Pichia stipitis* XDH and catalytically active variants thereof. Preferred TALs, TKLs and RKIs are those that are endogenous to the cell.

[0076] The specific coding or amino acid sequences for the various *S. cerevisiae* or other yeast or fungal enzymes referred to above can be identified in the literature and in bioinformatics databases well known to the skilled person, such as the BRENDA comprehensive enzyme information system, available at www.brenda-enzymes.org, KEGG (www.genome.jp/kegg/kegg2.html/), UniProt (<http://ca.expasy.org/sprot/>), Metacyc

(www.metacyc.com/), and the *Saccharomyces* genome database (www.yeastgenome.org). Particular XR, XDH, XK, TAL, TKL and RKL nucleic acid and amino acid sequences can also be prepared as described in the Examples, or are provided in the accompanying sequence listing.

[0077] In another aspect is a method for increasing the tolerance of a yeast cell such as a *Saccharomyces* cell to formic acid, comprising transforming the cell with a gene encoding a Yme2p polypeptide comprising the amino acid sequence of SEQ ID NO: 50 or a catalytically active variant, fragment or yeast ortholog thereof, and expressing the gene.

[0078] In one aspect of the disclosure, each of the heterologous gene or genes in the recombinant cell according to the invention is operably linked to an inducible, a regulated or a constitutive promoter. Optionally, one or more, optionally all, of the genes are integrated into the genome of the cell. In a specific aspect of the disclosure, the gene encoding an XK is operably linked to a constitutive promoter endogenous to the cell. Related recombinant techniques are described in more detail below.

[0079] Strains or clones of the recombinant cells of any of the preceding aspects of the disclosures are also provided by the disclosure. A "clone" in this context refers to a number of cells which all are derived from the same parent cell by cell division.

Recombinant methods

[0080] The disclosure also relates to vectors comprising genes encoding an XK as described in the preceding aspects which can be used for transforming a yeast *Saccharomyces* cell, such as an *S. cerevisiae* cell. The yeast host cell can either be naturally capable of fermenting xylose, or may be transformed with genes providing this capability, as described above.

[0081] Accordingly, in one aspect is a process for producing a recombinant cell described herein (a *Saccharomyces* cell), comprising transforming the cell with a heterologous gene (e.g., a vector) that encodes a XK comprising the amino acid sequence of SEQ ID NO: 6, the amino acid sequence of SEQ ID NO: 22 or a catalytically variant or fragment of any thereof.

[0082] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a XK comprising the amino acid sequence of SEQ ID NO: 6, the amino acid sequence of SEQ ID NO: 22 or a catalytically variant or fragment of any thereof; a heterologous gene (e.g., a vector) that encodes a XR, and a heterologous gene (e.g., a vector) that encodes a XDH. The heterologous genes may be in the form of one or more vectors.

[0083] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a XK comprising the amino acid sequence of SEQ ID NO: 6, the amino acid sequence of SEQ ID NO: 22 or a catalytically variant or fragment of any thereof; and a heterologous gene (e.g., a vector) that encodes an XI. The heterologous genes may be in the form of one or more vectors.

[0084] In one aspect is a vector comprising a gene encoding a xylulokinase (XK) comprising the amino acid sequence of SEQ ID NO: 6, SEQ ID NO: 22 or a catalytically active variant or fragment of any thereof, as defined in the claims.

[0085] In some aspects of the disclosure, the vector further comprises a gene encoding a xylose reductase (XR), a gene encoding a xylitol dehydrogenase (XDH), and/or a gene encoding a xylose isomerase (XI). The vector may further comprise regulatory sequences for expressing the genes in a yeast *Saccharomyces* or *Saccaromyces cerevisiae* host cell.

[0086] The invention also relates to vectors comprising genes encoding an Yme2p as described in the preceding aspects of the disclosures which can be used for transforming a yeast host *Saccharomyces* cell, such as an *S. cerevisiae* cell. The yeast host cell can either be naturally capable of fermenting xylose, or may be transformed with genes providing this capability, as described above.

[0087] Accordingly, in one aspect is a process for producing a recombinant cell described herein (a *Saccharomyces* cell), comprising transforming the cell with a heterologous gene (e.g., a vector) that encodes a Yme2p comprising the amino acid sequence of SEQ ID NO: 50, or a catalytically variant or fragment of any thereof.

[0088] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a Yme2p comprising the amino acid sequence of SEQ ID NO: 50 or a catalytically variant or fragment of any thereof; a heterologous gene (e.g., a vector) that encodes a XR, and a heterologous gene (e.g., a vector) that encodes a XDH. The heterologous genes may be in the form of one or more vectors.

[0089] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a Yme2p comprising the amino acid sequence of SEQ ID NO: 50 or a catalytically variant or fragment of any thereof; and a heterologous gene (e.g., a vector) that encodes an XI. The heterologous genes may be in the form of one or more vectors.

[0090] In one aspect is a vector comprising a gene encoding a Yme2p comprising the amino acid sequence of SEQ ID NO: 50 or a catalytically active variant or fragment of any thereof. In some aspects of the disclosure, the vector further comprises a gene encoding a xylose reductase (XR), a gene encoding a xylitol dehydrogenase (XDH), and/or a gene encoding a xylose isomerase (XI). The vector may further comprise regulatory sequences for expressing the genes in a yeast *Saccharomyces* or *Saccaromyces cerevisiae* host cell

[0091] The invention also relates to vectors comprising genes encoding an ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 polypeptide or catalytically active variant, fragment or yeast ortholog thereof as described in the preceding aspects of the disclosures which can be used for transforming a yeast *Saccharomyces* cell, such as an *S. cerevisiae* cell. The yeast host cell can either be naturally capable of fermenting xylose, or may be transformed with genes providing this capability, as described above.

[0092] Accordingly, in one aspect is a process for producing a recombinant cell described herein (a *Saccharomyces* cell), comprising transforming the cell with one or more heterologous genes (e.g., vectors) that encode ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 or a catalytically active variant, fragment or yeast ortholog of any thereof.

[0093] In one aspect of the disclosure, the process comprises transforming the cell with one or more heterologous genes (e.g., vectors) that encodes ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 or a catalytically active variant, fragment or yeast ortholog of any thereof; a heterologous gene (e.g., a vector) that encodes a XR, and a heterologous gene (e.g., a vector) that encodes a XDH. In some aspect of the disclosures, the process further comprises transforming the cell with a heterologous gene encoding a xylulose kinase (XK). The heterologous genes may be in the form of one or more vectors.

[0094] In one aspect of the disclosure, the process comprises transforming the cell with one or more heterologous genes (e.g., vectors) that encodes ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 or a catalytically active variant, fragment or yeast ortholog of any thereof; and a heterologous gene (e.g., a vector) that encodes an XI. In some aspect of the disclosures, the process further comprises transforming the cell with a heterologous gene encoding a xylulose kinase (XK). The heterologous genes may be in the form of one or more vectors.

[0095] In another aspect is a vector comprising one or more heterologous genes encoding ENO1, PFK2, PFK26, PGI1, GPM1, TPI1 or a functional variant, fragment or yeast ortholog of any thereof. In some aspect of the disclosures, the vector further comprises a gene encoding a xylose reductase (XR), a gene encoding a xylitol dehydrogenase (XDH). In some aspect of the disclosures, the process further comprises transforming the cell with a heterologous gene encoding a xylulose kinase (XK), and/or a gene encoding a xylose isomerase (XI). The vector may further comprise regulatory sequences for expressing the genes in a yeast *Saccharomyces* or *Saccaromyces cerevisiae* host cell.

[0096] The invention also relates to vectors comprising genes encoding an PGM3 polypeptide or a catalytically active variant or fragment thereof as described in the preceding aspects and aspect of the disclosures which can be used for transforming a yeast host *Saccharomyces* cell, such as an *S. cerevisiae* cell. The yeast host cell can either be naturally capable of fermenting xylose, or may be transformed with genes providing this capability, as described above.

[0097] Accordingly, in one aspect is a process for producing a recombinant cell described herein (a *Saccharomyces* cell), comprising transforming the cell with a heterologous gene (e.g., a vector) that encodes the PGM3 comprising the amino acid sequence of SEQ ID NO: 150 or a catalytically active variant or fragment thereof.

[0098] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a PGM3 comprising the amino acid sequence of SEQ ID NO: 150 or a catalytically variant or fragment of any thereof; a heterologous gene (e.g., a vector) that encodes a XR, and a heterologous gene (e.g., a vector) that encodes a XDH. The

heterologous genes may be in the form of one or more vectors.

[0099] In one aspect of the disclosure, the process comprises transforming the cell with a heterologous gene (e.g., a vector) that encodes a PGM3 comprising the amino acid sequence of SEQ ID NO: 150 or a catalytically variant or fragment of any thereof; and a heterologous gene (e.g., a vector) that encodes an XI. The heterologous genes may be in the form of one or more vectors.

[0100] In another aspect is a vector comprising a gene encoding the PGM3 comprising the amino acid sequence of SEQ ID NO: 150 or a functional variant or fragment thereof In some aspect of the disclosures, the vector further comprises a gene encoding a xylose reductase (XR), a gene encoding a xylitol dehydrogenase (XDH), and/or a gene encoding a xylose isomerase (XI). The vector may further comprise regulatory sequences for expressing the genes in a yeast *Saccharomyces* or *Saccaromyces cerevisiae* host cell

[0101] Many methods for genetic modification, including transformation of yeast host cells are known to one skilled in the art and may be used to create the present recombinant cells, some of which are exemplified below. Standard recombinant DNA and molecular cloning techniques useful for transforming microbial cells with a desired nucleic acid sequence or gene, or otherwise manipulate the microbial cell, are described, e.g., in Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); and by Silhavy, T. J., Bennis, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984); Ausubel et al., *Short Protocols in Molecular Biology*, 3rd ed., Wiley & Sons, 1995; and by Ausubel, F. M. et al., *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, Calif.).

[0102] Overexpressing a gene can be achieved by increasing the copy number of the gene coding for the enzyme in the cell, e.g. by integrating additional copies of the gene in the cell genome, expressing the gene from an episomal multicopy expression vector or introducing an episomal expression vector that comprises multiple copies of the gene; upregulating the endogenous gene, and the like. In a preferred aspect of the disclosure, the gene is introduced into the microbial cell via, e.g., transformation with one or more expression vectors. For example, for a yeast host cell, the level of a recombinantly expressed enzyme in the cell can be increased by clone one or more recombinant genes in a multicopy plasmid in the manner described by Mumberg *et al.* (1995).

[0103] The gene can either be synthesized or cloned from a host organism in which the corresponding DNA sequence is endogenous (see, e.g., Examples 1-5). Standard cloning procedures used in genetic engineering can be used to relocate a nucleic acid sequence from its natural location to a different site where it will be reproduced. The cloning procedures may involve excision and isolation of a desired DNA fragment comprising the DNA sequence encoding the polypeptide of interest, insertion of the fragment into a vector molecule, and incorporation of the

recombinant vector into a host cell where multiple copies or clones of the DNA sequence will be replicated. An isolated DNA sequence may be manipulated in a variety of ways to provide for expression of the polypeptide of interest. Manipulation of the DNA sequence prior to its insertion into a vector may be desirable or necessary depending on the expression vector and the host cell.

[0104] For example, to increase the likelihood that a, e.g., a bacterial enzyme gene is expressed in a yeast cell, the nucleotide sequence encoding the heterologous sequence may be adapted to optimize its codon usage to that of the yeast cell. The adaptiveness of a nucleotide sequence encoding enzyme to the codon usage of the host cell may be expressed as codon adaptation index (CAI). The CAI is herein defined as a measurement of the relative adaptiveness of the codon usage of a gene towards the codon usage of highly expressed genes. The relative adaptiveness (w) of each codon is the ratio of the usage of each codon, to that of the most abundant codon for the same amino acid. The CAI is defined as the geometric mean of these relative adaptiveness values. Non-synonymous codons and termination codons (dependent on genetic code) are excluded. CAI values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons (see Sharp and Li, 1987, *Nucleic Acids Research* 15: 1281-1295; also see: Jansen et al., 2003, *Nucleic Acids Res.* 31(8):2242-51). An adapted or "optimized" nucleotide sequence preferably has a CAI of at least 0.2, 0.3, 0.4, 0.5, 0.6 or 0.7.

[0105] The nucleotide sequence to be introduced into the DNA of the host cell may be integrated in vectors comprising the nucleotide sequence operably linked to one or more control sequences that direct the expression of the coding sequence in a suitable host cell under conditions compatible with the control sequences. The control sequence may be an appropriate promoter sequence, a nucleotide sequence which is recognized by a host cell for expression of the nucleotide sequence. The promoter sequence contains transcriptional control sequences, which mediate the expression of the coding sequence. The promoter may be any nucleotide sequence which shows transcriptional activity in the host cell of choice including native, mutant, truncated, and hybrid promoters, and may be obtained from genes encoding extracellular or intracellular polypeptides either homologous or heterologous to the host cell. The promoter may be a weak or a strong promoter that is constitutive or regulated in the host to be used. Strong, constitutive promoters are generally preferred for overexpression of the genes. Examples of suitable promoters for directing the transcription of the genes and vector constructs of the present invention in a yeast host cell are promoters obtained for example from the genes for *Saccharomyces cerevisiae* enolase (ENO1), *S. cerevisiae* galactokinase (GAL1), *S. cerevisiae* alcohol dehydrogenase 2 (ADH2), *S. cerevisiae* glyceraldehyde-3-phosphate dehydrogenase (TDH1), *S. cerevisiae* glyceraldehyde-3-phosphate dehydrogenase (TDH3), *S. cerevisiae* alcohol dehydrogenase 1 (ADH1), *S. cerevisiae* 3-phosphoglycerate kinase (PGK1) or *S. cerevisiae* cytochrome C (CYC1) (Karhumaa et al, 2005), translation elongation factor 1alpha (TEF1/TEF2) (Mumberg et al., 1995), PDC1 pyruvate decarboxylase (PDC1), pyruvate kinase (PYK1), and the constitutive truncated HXT7 promoter (Hauf et al. *Enzym Microb Technol* (2000) 26:688-698.) Other suitable vectors and promoters for use in yeast expression are further described in EP A-73,657 to Hitzeman, which is hereby incorporated by reference. Preferred promoters for overexpressing a gene encoding an XK in a recombinant yeast cell according to the invention include, but are not limited to, PGK1, TPI1, HXT7, TDH3, ADH2 and TEF2 promoters. Preferably, the promoter for overexpressing a gene encoding an XK according to the invention is TPI1.

Promoters for overexpressing the genes encoding an XR, an XDH, an XI, a TAL, a TKL and an RPI in a recombinant yeast cell according to the invention are preferably separately selected from the following: PGK1, TDH3, TEF2, PDC1, HXT truncated, TPI1 and PYK1.

[0106] The above disclosed vectors may comprise a gene encoding the enzyme polypeptide, a promoter, and transcriptional and translational stop signals as well as other regulatory or structural DNA sequences known to a person of skill in the art. The vector may be any vector or nucleic acid (e.g., a plasmid, virus, integration vector or integration fragment), which can be conveniently subjected to recombinant procedures and can bring about the expression of the gene in the yeast host cell. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids, and may contain any means for assuring self-replication.

[0107] Alternatively, the vector may be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. The vector system may be a single vector or plasmid or two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The vector may also be an integration vector comprising solely the gene or part of the gene to be integrated. For integration, the vector may rely on the DNA sequence encoding the polypeptide of interest or any other element of the vector for stable integration of the vector into the genome by homologous or non homologous recombination. More than one copy of a DNA sequence encoding a polypeptide of interest may be inserted into the host cell to amplify expression of the DNA sequence.

[0108] Optionally, the vectors of the present invention may contain one or more selectable markers, which permit easy selection of transformed cells. A selectable marker is a gene the product of which provides for biocide or viral resistance, resistance to heavy metals, prototrophy to auxotrophs and the like. Useful expression vectors for yeast cells include, for example, the 2 [mu] (micron) plasmid and derivatives thereof, the Yip, YE_p and YC_p vectors described by Gietz and Sugino (1988. "New yeast vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sites", *Gene* 74:527-534), the vectors described in Mumberg *et al.* (1995, *supra*), YEplac-HXT vector (Karhumaa *et al.*, 2005), the POT1 vector (U.S. Pat. No. 4,931,373), the pJS037 vector described in Okkels, *Ann. New York Acad. Sci.* 782, 202-207, 1996, the pPICZ A, B or C vectors (Invitrogen).

[0109] To achieve overexpression of an endogenous gene, promoter replacement methods may also be used to exchange the endogenous transcriptional control elements of the gene for another promoter (see, e.g., Mnaimneh *et al.* (2004) *Cell* 118(1):31-44). Deletions of DNA control elements preventing a high expression of an endogenous target gene may be made using mitotic recombination as described in Wach *et al.* ((1994) *Yeast* 10:1793-1808). This method involves preparing a DNA fragment that contains a selectable marker between genomic regions that may be as short as 20 bp, and which bound a target DNA sequence. 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. The linear DNA fragment can be efficiently transformed into a

cell and recombined into the genome resulting in gene replacement including with deletion of the target DNA sequence (as described in Methods in Enzymology, v194, pp 281-301 (1991)).

[0110] For yeasts such as for *Saccharomyces cerevisiae*, DNA sequences surrounding a target gene coding sequence can be identified, e.g., in the complete genome sequence coordinated by Genome Project ID9518 of Genome Projects coordinated by NCBI (National Center for Biotechnology Information) with identifying GOPIID #13838. Additional examples of yeast genomic sequences include that of *Yarrowia lipolytica*, GOPIC #13837, and of *Candida albicans*, which is included in GPID #10771, #10701 and #16373. Additional genomes have been completely sequenced and annotated and are publicly available for the following yeast strains *Candida glabrata* CBS 138, *Kluyveromyces lactis* NRRL Y-1140, *Pichia stipitis* CBS 6054, and *Schizosaccharomyces pombe* 972h-, available at <http://biocyc.org/>.

Fermentation methods

[0111] In one aspect a method for producing a fermentation product, comprising contacting the recombinant cell of any one of the preceding aspects and aspect of the disclosures with a medium comprising a carbon source comprising xylose and/or arabinose under anaerobic conditions, and recovering or isolating the fermentation product from the medium.

[0112] The fermentation product may be or comprise, for example, at least one of ethanol, butanol, isobutanol, isopentanol, lactate, isoamylacetate, glycerol, sorbitol, mannitol, xylitol, arabinitol; 3-hydroxybutyrolactone; hydrogen gas; L-ascorbic acid, 2,5-furan dicarboxylic acid, 3-hydroxy propionic acid, aspartic acid, glutaric acid, glutamic acid, itaconic acid, levulinic acid; succinic acid, fumaric acid, malic acid or other 1,4-diacid; a fatty acid, a fatty acid derived molecule; an isoprenoid, an isoprenoid-derived molecule; and an alkane.

[0113] However, it is contemplated that other fermentation products can also be produced using the methods of the present invention. Preferably, the carbon source comprises xylose and the fermentation product comprises ethanol.

[0114] In some aspect of the disclosures, the medium, *i.e.*, the fermentation medium, is feedstock from a cellulosic saccharification process and/or feedstock from a hemicellulose pretreatment process. Such feedstocks include, but are not limited to carbohydrates (e.g., lignocellulose, xylans, cellulose, starch, etc.), other sugars (e.g., glucose, xylose, arabinose, etc.), and other compositions. Compositions of fermentation media suitable for the growth of yeast are well known in the art and there are various reference texts that provide recipes for these media.

[0115] Typically, the fermentation takes place under conditions known to be suitable for generating the fermentation product. Fermentation conditions suitable for generating desired fermentation products are well known in the art and any suitable method finds use in the present invention. In some aspect of the disclosures, the fermentation process is carried out under aerobic or microaerophilic (*i.e.*, where the concentration of oxygen is less than that in air), or anaerobic conditions. In some aspect of the disclosures, fermentation is conducted under

anaerobic conditions (*i.e.*, no detectable oxygen), or less than about 5, about 2.5, or about 1 mmol/L/h oxygen. In the absence of oxygen, the NADH produced in glycolysis cannot be oxidized by oxidative phosphorylation. Under anaerobic conditions, pyruvate or a derivative thereof may be utilized by the host cell as an electron and hydrogen acceptor in order to generate NAD⁺.

[0116] The fermentation process is typically run at a temperature that is optimal for the recombinant fungal cell. For example, in some aspect of the disclosures, the fermentation process is performed at a temperature in the range of from about 25°C to about 42°C. Typically the process is carried out a temperature that is less than about 38°C, less than about 35°C, less than about 33°C, or less than about 38°C, but at least about 20°C, 22°C, or 25°C. Example 10 describes an exemplary assay for evaluating xylose consumption and/or ethanol production during fermentation of a xylose-containing fermentation medium.

[0117] In some aspect of the disclosures, the recombinant cells of the present invention are grown under batch or continuous fermentation conditions. Classical batch fermentation is a closed system, wherein the composition of the medium is set at the beginning of the fermentation and is not subject to artificial alterations during the fermentation. A variation of the batch system is a fed-batch fermentation, which also finds use in the present invention. In this variation, the substrate is added in increments as the fermentation progresses. Fed-batch systems are useful when catabolite repression is likely to inhibit the metabolism of the cells and/or where it is desirable to have limited amounts of substrate in the medium. Batch and fed-batch fermentations are common and well known in the art. Continuous fermentation is an open system where a defined fermentation generally maintains the culture at a constant high density where cells are primarily in log phase growth. Continuous fermentation systems strive to maintain steady state growth conditions. Methods for modulating nutrients and growth factors for continuous fermentation processes, as well as techniques for 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.

Examples

Example 1: Construction of a genetically modified *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway (TAL1 (transaldolase), TKL1 (transketolase) and RKI1 (ribose 5-phosphate ketol isomerase))

Strains, media and genetic techniques

[0118] *Escherichia coli* strain DH5α (Life Technologies, Rockville, MD, USA) was used for subcloning. *E. coli* was grown in LB-medium (Ausubel *et al.*, 1995). Yeast cells from freshly streaked YPD plates (Ausubel *et al.*, 1995) were used for inoculation. Plasmid DNA was prepared with the GeneJET™ Plasmid Miniprep Kit (Fermentas UAB, Vilnius, Lithuania). Agarose gel DNA

extraction was made with QIAquick® Gel Extraction Kit (Qiagen GmbH, Hilden, Germany). Primers from MWG-Biotech AG (Ebersberg, Germany) and Pfu DNA Polymerase and dNTP from Fermentas (Vilnius, Lithuania) were used for polymerase chain reactions (PCR). PCR product purification was made with the E.Z.N.A.® Cycle Sequencing Kit (Omega Bio-tek Inc., Doraville, GA, USA). Sequencing was performed by MWG-Biotech AG (Ebersberg, Germany). Restriction endonucleases, FastAP Thermosensitive Alkaline Phosphatase and T4 DNA Ligase from Fermentas (Vilnius, Lithuania) were used for DNA manipulation. Competent *E. coli* DH5α cells were transformed as described elsewhere (Inoue *et al.*, 1990) and transformed *E. coli* cells were selected on LB plates (Ausubel *et al.*, 1995) containing 100 mg/l ampicillin (IBI Shelton Scientific Inc., Shelton, CT, USA). *E. coli* strains were grown in LB medium containing 100 mg/l ampicillin for plasmid amplifications. Yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on YPD plates with a pH set to 7.5 containing 50 µg/ml zeocin (Invitrogen, Groningen, The Netherlands).

Construction of YlpTAL containing the *S. cerevisiae* transaldolase (TAL1) gene

[0119] Plasmid pB3 PGK TAL1 (Johansson and Hahn-Hägerdal, 2002) containing the *S. cerevisiae* transaldolase (TAL1) gene under control of the PGK1 promoter and the GCY1 terminator from *S. cerevisiae* was digested with restriction enzymes XcmI and EheI, and ends of the resulting fragment were made blunt by the use of T4 DNA polymerase (Boehringer Mannheim, Indianapolis, IN) and the vector was re-ligated resulting in YlpTAL. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing. Construction of YlpTALTKL containing the *S. cerevisiae* transaldolase (TAL1) gene and transketolase (TKL1) gene

[0120] The DNA cassette PGKp-TKL1-GCYt containing the *S. cerevisiae* transketolase (TKL1) gene under control of the PGK1 promoter and the GCY1 terminator from *S. cerevisiae* was PCR amplified having as template plasmid pB3 PGK TKL1 (Johansson and Hahn-Hägerdal, 2002) and using primers FwdTKL and RevTKL identified by GGTACCGAGCTCTAACTGATCTATCCAAAAGT (SEQ ID NO: 1) and GGTACCGATCAGCATGCGATCGCTCGACATTTGATATAC (SEQ ID NO: 2), which were including the restriction site KpnI at the ends of the amplified DNA cassette. The PCR product PGKp-TKL1-GCYt was then digested with KpnI restriction enzyme. The resulting purified DNA fragment was inserted into the plasmid YlpTAL, which had also been cleaved with the restriction enzyme KpnI and which had been dephosphorylated. The resulting plasmid was named YlpTALTKL. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of YlpTTR containing the *S. cerevisiae* transaldolase (TAL1) gene, transketolase (TKL1) gene and ribose 5-phosphate ketol isomerase (RKI1) gene

[0121] The DNA cassette PGKp-RKI1-GCYt containing the *S. cerevisiae* ribose 5-phosphate ketol isomerase (RKI1) gene under control of the PGK1 promoter and the GCY1 terminator from *S.*

cerevisiae was PCR amplified having as template plasmid pB3 PGK RKI1 (Johansson and Hahn-Hägerdal, 2002) and using primers FwdRKI and RevRKI identified by 5'-CCGCGGGAGCTCTAACTGATCTATCCAAACTG-3' (SEQ ID NO: 3) and 5'-CCGCGGGATCAGCATGCGATCGCTCGACATTTGATATAC-3' (SEQ ID NO: 4) which were including the restriction site SacII at the ends of the amplified DNA cassette. The PCR product PGKp-RKI1-GCYt was then digested with SacII restriction enzyme. The resulting purified DNA fragment was inserted into the plasmid YlpTALTKL, which also had been cleaved with the restriction enzyme KpnI and which had been dephosphorylated. The resulting plasmid was named YlpTTR. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1001 containing YlpTTR

[0122] YlpTTR was cleaved with SpeI within the RKI1 gene and transformed into strain C5LTe1000. This resulted in strain C5LTe1001. Strain C5LTe1000 was deposited in accordance with the terms of the Budapest Treaty on 4 November 2014 with DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Inhoffenstr. 7B, 38124 Braunschweig, Germany - under DSMZ accession number DSM 29597.

Example 2: Strains, Media and Genetic Techniques

[0123] *Escherichia coli* strain NEB 5-alpha (New England BioLabs, Ipswich, MA, USA) was used for subcloning. *E. coli* was grown in LB-medium (Ausubel *et al.*, 1995). Yeast cells from freshly streaked YPD plates (Ausubel *et al.*, 1995) were used for inoculation. Plasmid DNA was prepared with the GeneJET™ Plasmid Miniprep Kit (Fermentas UAB, Vilnius, Lithuania). Agarose gel DNA extraction was made with QIAquick® Gel Extraction Kit (Qiagen GmbH, Hilden, Germany). Primers from Eurofins MWG Operon (Ebersberg, Germany) and Phusion Hot Start II High-Fidelity DNA Polymerase and dNTP from Fermentas (Vilnius, Lithuania) were used for polymerase chain reactions (PCR). PCR product purification was made with the E.Z.N.A.® Cycle Sequencing Kit (Omega Bio-tek Inc., Doraville, GA, USA). Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany). Restriction endonucleases from Fermentas (Vilnius, Lithuania) were used for DNA manipulation. Competent *E. coli* cells were transformed as described elsewhere (Inoue *et al.*, 1990) and transformed *E. coli* cells were selected on LB plates (Ausubel *et al.*, 1995) containing 100 mg/l ampicillin (IBI Shelton Scientific Inc., Shelton, CT, USA). *E. coli* strains were grown in LB medium containing 100 mg/l ampicillin for plasmid amplifications. Yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on Yeast Nitrogen Base plates (YNB) (6.7 g/l Difco Yeast Nitrogen Base without amino acids; Becton Dickinson and Company, Sparks, MD, USA) supplemented with 40 g/l xylose and buffered at pH 5.5 with 10.21 g/l potassium hydrogen phthalate.

Example 3: Construction of an *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway and expressing a *Spathaspora passalidarum* xylulose kinase gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated

Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene*Strains, media and genetic techniques**

[0124] The strains, media and genetic techniques described in Example 2 were used.

Construction of synthetic xylulose kinase gene encoding *Spathaspora passalidarum* xylulose kinase based on NCBI Accession Code XP 007373112 under control of the TPI1 Promoter and the PGK1 terminator from *S. cerevisiae*

[0125] The entire *Spathaspora passalidarum* xylulose kinase gene (XKsp) was synthesized and assembled by Eurofins MWG Operon (Ebersberg, Germany). Codon usage in the sequence was optimised based on the yeast codon usage table from the Kazusa codon usage database. The TPI1 promoter from *S. cerevisiae* and the PGK1 terminator from *S. cerevisiae* were also included in the synthetic construct; the TPI1 promoter before the ATG-start codon and the PGK1 terminator after the stop-codon. The nucleotide sequence of TPI1 promoter, XKsp and PGK1 terminator is identified as SEQ ID NO: 5, showing the synthesized DNA sequence. The corresponding amino acid sequence of the coding region of XKsp is identified in SEQ ID NO: 6. The harboring plasmid was named pC5e0022.

Construction of pC5e0049 containing a *Spathaspora passalidarum* xylulose kinase gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0126] The DNA cassette TPI1p-XKsp-PGK1t was PCR amplified having as template plasmid pC5e0022 and using the following primers.

TPI1p_fwd (SEQ ID NO: 7):

5'-TCTTC CACAC CTGCA GTATA TCTAG GAACC CATCA G-3'

reverse-PGK1t (SEQ ID NO: 8):

5'-ATCAG TTAGA CTGCA GGAAC ATAGA AATAT CGAAT GGGAA-3'

[0127] The resulting purified DNA fragment was inserted into plasmid YlpDR7 (Runquist *et al.* 2010), containing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene under control of the PGK1 promoter and the PGK1 terminator from *S. cerevisiae* and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene under control of the TDH3 promoter and the ADH1 terminator from *S. cerevisiae*. YlpDR7 which had been cleaved with restriction enzyme PstI and the DNA fragment was inserted by In-Fusion cloning (Clontech, California, CA, USA). The

resulting plasmid was named pC5e0049. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1042 containing YlpTTR and pC5e0049

[0128] The Plasmid pC5e0049 was digested with restriction enzyme EcoRV within the URA3 gene and it was thereafter transformed into strain C5LTe1001. This resulted in strain C5LTe1042.

Example 4: Construction of an *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

Strains, media and genetic techniques

[0129] The strains, media and genetic techniques described in Example 2 were used.

Construction of pC5e0024 containing a *S. cerevisiae* xylulose kinase gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0130] The TPI1 promoter from *S. cerevisiae* (TPI1p) was PCR amplified having as template plasmid pC5e0022 and using the following primers.

TPI1p_fwd (SEQ ID NO: 9):

5'-TCTTC CACAC CTGCA GTATA TCTAG GAACC CATCA G-3'

R_TPI1p (SEQ ID NO: 10):

5'-CTGTC TCTGA ATTAC TGAAC ACAAC ATTTT TAGTT TATGT ATGTG TTT-3'

[0131] The PGK1 terminator from *S. cerevisiae* (PGK1t) was PCR amplified having as template plasmid pC5e0022 and using the following primers.

fwdS_PGK1t (SEQ ID NO: 11):

5'-GCGAACTGGAAAAGACTCTCATCTAAAGATCTCCCATGTCTCTACTGG-3'

reverse_PGK1t (SEQ ID NO: 12):

5'-ATCAGTTAGACTGCAGGAACATAGAAATATCGAATGGGAA-3'

[0132] The *S. cerevisiae* xylulose kinase gene (Xkcere; SEQ ID NO: 31), encoding *S. cerevisiae* XK (SEQ ID NO: 32) was PCR amplified having as template plasmid YlpXK (Lönn *et al.*, 2003) and using the following primers

XKcere_fwd (SEQ ID NO: 13):

5'-AAACACATACATAAACTAAAAATGTTGTGTTTCAGTAATTCAGAGACAG-3'

XKcere_rev (SEQ ID NO: 14):

5'-CCAGTAGAGACATGGGAGATCTTTAGATGAGAGTCTTTTCCAGTTCGC-3'

[0133] The DNA cassette TPI1p-XKcere-PGK1t was PCR amplified by overlap extension PCR having as template the three purified DNA fragments TPI1p, PGK1t and XKcere using the following primers

TPI1p_fwd (SEQ ID NO: 15):

5'-TCTTCCACACCTGCAGTATATCTAGGAACCCATCAG-3'

reverse_PGK1t (SEQ ID NO: 16):

5'-ATCAGTTAGACTGCAGGAACATAGAAATATCGAATGGGAA-3'

[0134] The resulting purified DNA fragment was inserted into the plasmid YlpDR7 (Runquist *et al.* 2010), which had been cleaved with restriction enzyme PstI. The DNA fragment was inserted by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0024. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of of C5LTe1035 containing YlpTTR and pC5e0024

[0135] Plasmid pC5e0024 was cleaved with restriction enzyme EcoRV within the URA3 gene and it was thereafter transformed into strain C5LTe1001. This resulted in strain C5LTe1035.

Example 5: Construction of an *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway and expressing an *Escherichia coli* xylulose kinase (XK) gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

Strains, media and genetic techniques

[0136] The strains, media and genetic techniques described in Example 2 were used.

Construction of synthetic xylulose kinase gene encoding *Escherichia coli* xylulose kinase based on NCBI Accession Code YP 001460359 under control of the TPI1 Promoter and the PGK1 terminator from *S. cerevisiae*

[0137] The entire *E. coli* xylulose kinase gene (XKcoli) was synthesized and assembled by Eurofins MWG Operon (Ebersberg, Germany). Codon usage in the sequence was optimised based on the yeast codon usage table from the Kazusa codon usage database. The TPI1 promoter from *S. cerevisiae* and the PGK1 terminator from *S. cerevisiae* were also included in the synthetic construct; the TPI1 promoter before the ATG-start codon and the PGK1 terminator after the stop-codon. The nucleotide sequence of TPI1 promoter, XKcoli and PGK1 terminator is identified as SEQ ID NO: 17, showing the synthesized DNA sequence. The corresponding amino acid sequence of the coding region of XKcoli is identified in SEQ ID NO: 18. The harboring plasmid was named pC5e0012.

Construction of pC5e0046 containing a *E. coli* xylulose kinase gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0138] The DNA cassette TPI1p-XKcoli-PGK1t was PCR amplified having as template plasmid pC5e0012 and using the following primers.

TPI1p_fwd (SEQ ID NO: 19):

5'-TCTTCCACACCTGCAGTATATCTAGGAACCCATCAG-3'

reverse-PGK1t (SEQ ID NO: 20):

5'-ATCAGTTAGACTGCAGGAACATAGAAATATCGAATGGGAA-3'.

[0139] The resulting purified DNA fragment was inserted into plasmid YlpDR7 (Runquist *et al.* 2010), containing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene under control of the PGK1 promoter and the PGK1 terminator from *S. cerevisiae* and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene under control of the TDH3 promoter and the ADH1 terminator from *S. cerevisiae*. YlpDR7 which had been cleaved with restriction enzyme PstI and the DNA fragment was inserted by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0046. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1040 containing YlpTTR and pC5e0046

[0140] Plasmid pC5e0046 was digested with restriction enzyme EcoRV within the URA3 gene and

it was thereafter transformed into strain C5LTe1001. This resulted in strain C5LTe1040.

Example 6: Construction of an *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway and expressing a *Kluyveromyces marxianus* xylulose kinase (XK) gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

Strains, media and genetic techniques

[0141] The strains, media and genetic techniques described in Example 2 were used.

Construction of synthetic xylulose kinase gene encoding *Kluyveromyces marxianus* xylulose kinase based on NCBI Accession Code ADW23548

[0142] The entire *Kluyveromyces marxianus* xylulose kinase gene (XKmarx) was synthesized and assembled by Eurofins MWG Operon (Ebersberg, Germany). Codon usage in the sequence was optimised based on the yeast codon usage table from the Kazusa codon usage database. The nucleotide sequence of XKmarx is identified as SEQ ID NO:21, showing the synthesized DNA sequence. The corresponding amino acid sequence of the coding region of XKmarx is identified in SEQ ID NO: 22. The harboring plasmid was named pC5e0043.

Construction of pC5e0051 containing a *Kluyveromyces marxianus* xylulose kinase gene, a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0143] The TPI1 promoter from *S. cerevisiae* (TPI1p) was PCR amplified having as template plasmid pC5e0022 and using the following primers.

TPI1p_fwd (SEQ ID NO: 23):

5'-TCTTCCACACCTGCAGTATATCTAGGAACCCATCAG-3'

Rev_TPI1p (SEQ ID NO: 24):

5'-GCCTAAGTAATATGGAGTCGACATTTTTAGTTTATGTATGTGTTT-3'.

[0144] The PGK1 terminator from *S. cerevisiae* (PGK1t) was PCR amplified having as template plasmid pC5e0022 and using the following primers.

PGK1t_forS (SEQ ID NO: 25):

5'-CTTTAGCACAATCTCAGGGTCAATAAAGATCTCCCATGTCTCTACTGG-3'

reverse_PGK1t (SEQ ID NO: 26):

5'-ATCAGTTAGACTGCAGGAACATAGAAATATCGAATGGGAA-3'

[0145] The *Kluyveromyces marxianus* kinase gene (XKmarx) was PCR amplified having as template plasmid pC5e0043 and using the following primers

XKmarx_fwd (SEQ ID NO: 27):

5'-AAACACATACATAAACTAAAAATGTCGACTCCATATTACTTAGGC-3'

XKmarx_rev (SEQ ID NO: 28):

5'-CCAGTAGAGACATGGGAGATCTTTATTGACCCTGAGATTGTGCTAAAG-3'

[0146] The DNA cassette TPI1p-XKmarx-PGK1t was PCR amplified by overlap extension PCR having as template the three purified DNA fragments TPI1p, PGK1t and XKmarx using the following primers:

TPI1p_fwd (SEQ ID NO: 29):

5'-TCTTCCACACCTGCAGTATATCTAGGAACCCATCAG-3'

reverse_PGK1t (SEQ ID NO: 30):

5'-ATCAGTTAGACTGCAGGAACATAGAAATATCGAATGGGAA-3'

[0147] The resulting purified DNA fragment was inserted into the plasmid YlpDR7 (Runquist *et al.* 2010), which had been cleaved with restriction enzyme PstI. The DNA fragment was inserted by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0051. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1043 containing YlpTTR and pC5e0051

[0148] Plasmid pC5e0051 was digested with restriction enzyme EcoRV within the URA3 gene and it was thereafter transformed into strain C5LTe1001. This resulted in strain C5LTe1043.

Example 7: Enzyme activity assays

[0149] Cell extracts for activity assays were prepared from exponentially growing aerobic batch cultures. Cells were collected by centrifugation, washed with sterile water, resuspended in an appropriate amount of Y-PER reagent (Pierce; Rockford, Ill., USA), and processed according to the instructions. Protein concentrations were determined with the Bradford Protein Assay (Pierce,

Rockford, IL, USA) against a bovine serum albumin standard. Xylulose kinase (XK) and xylitol dehydrogenase (XDH) were measured as described by Shamanna and Sanderson (1979). XK activity was determined in two steps. First, the XDH activity was determined in the absence of ATP, and then the sum of the XK and XDH activities in the presence of ATP was determined, the XK activity being the difference. All enzyme activity measurements were performed at 30°C. Specific activities are expressed as units per milligram of protein. One unit of enzyme activity is defined as 1 μ mol of substrate converted per minute.

[0150] As shown in Table 5, strains C5LTe1042, C5LTe1048 and C5LT1208 had higher XK activity than their corresponding control strains. All strains with a heterologous XK gene had higher XK activity when compared to the wild-type strain C5LTe1000.

Table 5. *XK activities in constructed strains.*

Strain	XK activity
C5LTe1000	0.11 $\pm$ 0.00
C5LTe1035	1.65 $\pm$ 0.28
C5LTe1042	5.21 $\pm$ 0.37
C5LTe1036	0.84 $\pm$ 0.20
C5LTe1048	4.31 $\pm$ 0.06
C5LTe1204	0.92 $\pm$ 0.27
C5LTe1208	3.67 $\pm$ 0.19

Example 8: Aerobic growth on xylose

[0151] Cells were grown on YPD medium (Yeast extract, Peptone and 20 g/l glucose) overnight, and inoculated into mineral medium with xylose as the sole carbon source (13.4 g/l Yeast Nitrogen Base and 50 g/l xylose, buffered with 10.2 g/l potassium hydrogen phthalate to pH 5.5). Starting OD at 620 nm was around 0.1. Growth was measured by increase in OD 620nm. Maximum specific growth rate was calculated from time points 8 to 30 hours, and they were 0.16 h⁻¹ for C5LTe1035 and 0.19 h⁻¹ for C5LTe1042. See Fig. 1, which shows curves for the aerobic growth on xylose for C5LTe1035 and C5LTe1042.

[0152] In conclusion, strains C5LTe1042, C5LTe1048 and C5LT1208 grew more than their corresponding control strains.

Example 9: Anaerobic growth on xylose

[0153] Yeast strains were analysed for anaerobic growth on xylose in 96-well microplate yeast cultures. Prior to experiments, yeast was grown in semi-aerobic precultures in YNB medium with 50 g/l xylose, in microplates, overnight. For measurement of anaerobic growth, cells were inoculated in 200 μ l of the same medium where solution of ergosterol and Tween80 had been

added at final concentration of 0.03 and 1.2 g/l, respectively. 50 µl mineral oil was added on top of each well to keep culture anaerobic. 20 µl of pre-cultured cells were added. The growth was followed in Multiskan FC (Thermo Scientific) at 30°C and growth was measured as increase of OD (620nm).

[0154] In conclusion, strains with *K. marxianus* XK or *S. passalidarum* XK grow faster and more under anaerobic conditions than the control strain with *S. cerevisiae* XK. See Fig. 2, which shows curves over anaerobic growth of C5LTe1042 compared with control strain C5LTe1035, Fig. 3, which shows curves over anaerobic growth of C5LTe1043 compared with control strain C5LTe1035, and Fig. 4, which shows curves over anaerobic growth of C5LTe1040 compared with control strain C5LTe1035.

Example 10: Anaerobic fermentation on xylose

[0155] Cells were pre-grown on mineral medium with glucose as a carbon source (YNB 6.7 g/l and 20 g/l glucose). Cells were then inoculated in fermenters (Applikon) at starting biomass concentration of about 0.15 g/l cell dry weight. Mixing was set at 200 rpm and the gas outlet was closed by a waterlock. Fermentation medium consisted of YNB (6.7 g/l) with 50 g/l xylose and 20 g/l glucose, supplemented with Tween80 (1.2 g/l) and ergosterol (0.03 g/l). The pH was controlled at 5.5 with KOH.

[0156] Concentrations of glucose, xylose, ethanol, glycerol, and xylitol were determined by high performance liquid chromatography (Waters, Milford, MA, USA). The compounds were separated with a Shodex SUGAR SP0810 Pb2+ copolymer-based column (Showa Denko America, NY, USA) preceded by a Micro-Guard Carbo-C guard column (Bio-Rad, Hercules, CA, USA). Separation was performed at 80°C, with H₂O at a flow rate of 0.6 ml min⁻¹ as mobile phase. Compounds were quantified by refractive index detection (Waters). A seven-point calibration curve was made for each compound to calculate concentrations. See Fig. 5, which shows xylose fermentation by strains C5LTe1035 (dashed line) and C5LTe1042 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line. See also Fig. 6, which shows xylose fermentation by strains C5LTe1036 (dashed line) and C5LTe1048 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line. And, see also Fig. 7, which shows xylose fermentation by strains C5LTe1204 (dashed line) and C5LTe1208 (solid line). Glucose: circles, Xylose: squares, Xylitol: stars, Biomass: line.

Table 6. Summary of xylose consumption and ethanol production in anaerobic xylose fermentation of 50 g/l xylose and 20 g/l glucose in mineral medium within 72 hours of fermentation. See also a graphic depiction of this data in Fig. 8.

Strain	Consumed xylose (g/L)	Produced ethanol (g/L)	Xylitol yield (g/g xylose consumed)	Biomass yield (g/g total sugar consumed)
C5LTe1036	21	15	0.30	0.08
C5LTe1048	35	19	0.25	0.09
C5LTe1035	42	21	0.24	0.07

Strain	Consumed xylose (g/L)	Produced ethanol (g/L)	Xylitol yield (g/g xylose consumed)	Biomass yield (g/g total sugar consumed)
C5LTe1042	46	23	0.14	0.06
C5LTe1204	25	15	0.21	0.10
C5LTe1208	37	19	0.25	0.07

[0157] In conclusion, these data show that strains C5LTe1042, C5LTe1048 and C5LT1208 carrying *S. passalidarum* XK consumed more xylose and produced more ethanol than their corresponding control strains C5LTe1035, C5LTe1036 and C5LT1204, respectively.

Example 11: Construction of an *S. cerevisiae* strain overexpressing three of the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

Strains, media and genetic techniques

[0158] Yeast cells from freshly streaked YPD plates (Ausubel et al., 1995) were used for inoculation. Plasmid DNA was prepared with the GeneJET™ Plasmid Miniprep Kit (Fermentas UAB, Vilnius, Lithuania). Agarose gel DNA extraction was made with QIAquick® Gel Extraction Kit (Qiagen GmbH, Hilden, Germany). Primers from Eurofins MWG Operon (Ebersberg, Germany) and Phusion Hot Start II High-Fidelity DNA Polymerase and dNTP from Fermentas (Vilnius, Lithuania) were used for polymerase chain reactions (PCR). PCR product purification was made with the E.Z.N.A.® Cycle Sequencing Kit (Omega Bio-tek Inc., Doraville, GA, USA). Sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany). Restriction endonucleases from Fermentas (Vilnius, Lithuania) were used for DNA manipulation. Competent *E. coli* cells were transformed as described elsewhere (Inoue et al., 1990) and transformed *E. coli* cells were selected on LB plates (Ausubel et al., 1995) containing 100 mg/l ampicillin (IBI Shelton Scientific Inc., Shelton, CT, USA). *E. coli* strains were grown in LB medium containing 100 mg/l ampicillin for plasmid amplifications. Yeast strains were transformed with the lithium acetate method (Gietz et al., 2007) and transformed yeast strains were selected on Yeast Nitrogen Base plates (YNB) (6.7 g/l Difco Yeast Nitrogen Base without amino acids; Becton Dickinson and Company, Sparks, MD, USA) supplemented with 40 g/l xylose and buffered at pH 5.5 with 10.21 g/l potassium hydrogen phthalate.

Construction of C5LTe1201 containing YlpTTR

[0159] YlpTTR was cleaved with *SpeI* within the RKI1 gene and transformed into strain TMB 3000, a robust strain of *Saccharomyces cerevisiae* (Lindén et al., 1992). This resulted in strain\

C5LTe1201.

Construction of C5LTe1202 containing YlpTTR and pC5e0024

[0160] Plasmid pC5e0024 was cleaved with restriction enzyme EcoRV within the URA3 gene and it was thereafter transformed into strain C5LTe1201. This resulted in strain C5LTe1202.

Example 12: Preparation of plasmid pC50042 and transformant C5LTe1212

Generation of fragments

[0161] YME2-PGK1t was synthesized and assembled by Eurofins MWG Operon (Ebersberg, Germany) and that the harboring plasmid was named pEX-A1-YME2. The fragment YME2-PGK1t was generated by PCR using the plasmid pEX-A1-YME2 as template and the following primers (Eurofins MWG Operon):

5'-GATCCCCGGGCTGCAATGTTGCCCATTTCTGGACCTT-3' (SEQ ID NO: 45)

5'-CGCTGCAGGTCGACGTGTTACATGCGTACACGCGTCT-3' (SEQ ID NO: 46)

[0162] The vector pUG6-HXT7'p was generated by PCR using the plasmid pUG6-HXT-PGM2 (WO2010/059095 A1) as template and the following primers (Eurofins MWG Operon):

5'-CGTCGACCTGCAGCGTAC -3' (SEQ ID NO: 47)

5'-TGCAGCCCGGGGATCCTTTTT-3' (SEQ ID NO: 48)

[0163] The three PCR products were purified using QIAquick Gel Extraction Kit (Qiagen, Venlo, Netherlands) and DNA concentration was determined.

[0164] SEQ ID NO: 49 shows the nucleic acid sequence for codon-optimized YME2 tolerance gene, whereas SEQ ID NO: 50 shows the encoded amino acid sequence.

In-Fusion cloning and *E.coli* transformation

[0165] Fragment YME2-PGK1t and vector pUG6-HXT7'p were introduced into NEB5α *E. coli* competent cells (New England Biolabs, Ipswich MA, USA) using In-Fusion HD Cloning Kit (Clontech, Mountain View CA, USA) following the manufacturer instructions procedure.

Transformants were selected on LB agar plates containing 100 µg/mL ampicillin and incubated overnight.

[0166] Four colonies were randomly selected and growth overnight on LB containing 100 µg/mL ampicillin. Plasmid DNA was prepared using the GeneJET Plasmid MiniPrep Kit (Thermo Scientific, Waltham MA, USA) by following the manufacturer instructions and concentration was determined.

[0167] Plasmid DNA was digested with restriction enzymes XhoI and SmaI (Thermo Scientific) by following the manufacturer instructions. All the evaluated clones containing the YME2-plasmid displayed the expected size fragments (1904 and 3105 bp) and one of them was chosen to be sequenced (Eurofins MWG Operon) using the following primers:

5'-CCTGCGTGTTCTTCTGAGGTTC-3' (SEQ ID NO: 21)

5'-ATATTGTCGTTAGAACGCGG-3' (SEQ ID NO: 22)

[0168] The obtained sequence was as predicted when using *in silico* cloning tools and the plasmid was named pC5e0042.

Genomic integration of linearized plasmid pC50042 into C5LTe1202

[0169] 5 µg of plasmid was linearized with Bmrl (New England Biolabs) following the manufacturer instructions procedure and kept on ice until used in the transformation system.

[0170] Strain C5LTe1202 was transformed using a lithium acetate-based method (Gietz and Schiestl, 2007) and 20 µL of linearized pC50042 were used in the transformation system. YPD agar plates supplemented with 20g/L glucose and 500 mg/L geneticin were incubated aerobically at 30°C and transformant C5LTe1212 was selected after 2 days.

Example 13: Toxicity of formic acid on yeast growing on xylose

[0171] The inhibitory effect of acetic and formic acids on cells growing on xylose was initially demonstrated with strain C5LTe1202.

[0172] Mineral medium supplemented with 110 g/L of xylose, 50 mM potassium phthalate buffer and containing the concentrations (in g/L) of single inhibitors indicated in Fig. 9. Anaerobic growth was followed in Multiskan FC (Thermo Scientific) at 30°C and growth was measured as increase of OD (620nm). The increment of OD after 60 hours of cultivation was chosen to estimate the tolerance to single lignocellulosic inhibitors, since no exponential growth can be observed when using initial high concentration of xylose.

Example 14: Fermentation in the presence of acetic acid

[0173] Cells aerobically grown overnight on YNB medium supplemented with 20 g/L glucose were used to inoculate a serum flask containing 30 mL of YNB medium supplemented with glucose (20 g/L) and xylose (50 g/L) as carbon source and acetic acid (8 g/L) at initial concentration of 1 g cell dry weight/L.

[0174] Concentrations of glucose, xylose, ethanol, glycerol, xylitol and acetic acid were determined by high performance liquid chromatography (Waters, Milford, MA, USA). The compounds were separated with a Shodex SUGAR SP0810 Pb²⁺ copolymer-based column (Showa Denko America, NY, USA), or a Rezex H+ column, preceded by a Micro-Guard Carbo-C guard column (Bio-Rad, Hercules, CA, USA). Separation was performed at 80°C, with H₂O at a flow rate of 0.6 ml min⁻¹ as mobile phase. Compounds were quantified by refractive index detection (Waters). A seven-point calibration curve was made for each compound to calculate concentrations.

[0175] Results are presented graphically in Fig. 10, Fig. 11, and Fig. 12. In conclusion, more xylose was consumed, and more ethanol was produced, in strain C5LTe1212 expressing the YME2 gene.

Example 15: Fermentation and growth characteristics in the presence of formic acid

[0176] Cells aerobically grown overnight on YNB medium supplemented with 20 g/L glucose were used to inoculate a serum flask containing 30 mL of YNB medium supplemented with glucose (20 g/L) and xylose (50 g/L) as carbon source and formic acid (4.5 g/L) at initial concentration of 1 g CDW/L. Glucose, xylose, ethanol, xylitol, glycerol and acetate were analysed using HPLC (Waters) using the same procedure as in Example 7.

[0177] Results are shown graphically in Fig. 13, Fig. 14, and Fig. 15. In conclusion, the more xylose was consumed, and more ethanol was produced, in strain C5LTe1212 expressing the YME2 gene.

[0178] In another tolerance experiment, cells grown overnight on YNB medium supplemented with 50 g/L xylose were used to inoculate a microtiter plate containing YNB medium supplemented with 110 g/L xylose as carbon source and 4.5 g/L formic acid. Fig. 16 shows a graphic representation of the anaerobic growth characteristics (measured as change in OD at 620 nm) of strains C5LTe1202 (upper curve) and C5LTe1212 (lower curve).

[0179] To conclude, overexpression of the YME2 gene clearly improves tolerance towards both formic acid and acetic acid. Xylose fermentation in presence of formic or acetic acid was improved in yeast overexpressing YME2. Specifically, xylose consumption and ethanol production increased by 13% and 12%, respectively, in the presence of acetic acid, and by 7% and 12%, respectively, in

the presence of formic acid. Formic acid is common in lignocellulosic hydrolysates, and strongly contributes to the toxicity of such hydrolysates.

Example 16: Construction and selection of MC strains

[0180] For this Example, certain *S. cerevisiae* genes coding for enzymes in the main pathways of the central metabolism were tested.

[0181] Yeast strain C5LTe1101 was constructed by transforming yeast strain TMB 3043 (Karhumaa *et al.* 2005) in its *ura3* locus with a DNA fragment containing URA3 gene, TDH3p-XYL1(N272D)-ADH1t, and PGK1p-XYL2-PGK1t. Selection was made on YNB (Yeast nitrogen base 6,7 g/l and 20 g/l glucose, 20 g/l agar) plates supplemented with 200 mg/L leucine. To C5LTe1101 were transformed DNA fragments obtained with PCR with primers shown in Table 7 together with the plasmid p245GPD (Mumberg *et al.*, 1995) linearized with *Sma*I. A control strain was constructed by transforming with unlinearized plasmid p245GPD. Selection was made on YNB agar plates with 20 g/l glucose as a carbon source. Selected colonies were tested by colony PCR and clones with PCR-products indicating right size of plasmid insert were chosen.

Table 7. Genes Included:

Strain	Gene	Forward primer (5' → 3') (SEQ ID NO:)	Reverse primer (5' → 3') (SEQ ID NO:)
MC1	GLK1	<u>CTAGAACTAGTGGATCCCCCATGTCA</u> TTCGACGACTTACACAAAG (63)	<u>ATATCGAATTCCTGCAGCCCTCATGCTA</u> CAAGCGCACACAA (64)
MC2	PGI1	<u>CTAGAACTAGTGGATCCCCCATGTCC</u> AATAACTCATTCACTAACTTCA (65)	<u>ATATCGAATTCCTGCAGCCCTCACATC</u> CATTCTTGAATTG (66)
MC3	PFK2 6	<u>CTAGAACTAGTGGATCCCCCATGTTT</u> AAACCAGTAGACTTCTCTGA (67)	<u>ATATCGAATTCCTGCAGCCCTTAAACGT</u> GACTTTGGCTGC (68)
MC4	PFK2	<u>CTAGAACTAGTGGATCCCCCATGACT</u> GTTACTACTCCTTTTGTGAATG (69)	<u>ATATCGAATTCCTGCAGCCCTTAATCAA</u> CTCTCTTCTTCCAACC (70)
MC5	PFK2 7	<u>CTAGAACTAGTGGATCCCCCATGGGT</u> GGTTCTCCGATTCA (71)	<u>ATATCGAATTCCTGCAGCCCTCAAGCA</u> AATCCGTTGCTTTC (72)
MC6	FBA1	<u>CTAGAACTAGTGGATCCCCCATGGGT</u> GTTGAACAAATCTTAAAGAG (73)	<u>ATATCGAATTCCTGCAGCCCTTAATCAA</u> CTCTCTTCTTCCAACC (74)
MC7	TDH1	<u>CTAGAACTAGTGGATCCCCCATGATC</u> AGAATTGCTATTAACGGTTTC (75)	<u>ATATCGAATTCCTGCAGCCCTTAAGCCT</u> TGGCAACATATTCG (76)

Strain	Gene	Forward primer (5' → 3') (SEQ ID NO:)	Reverse primer (5' → 3') (SEQ ID NO:)
MC8	TDH2 /3	<u>CTAGAACTAGTGGATCCCCCATGGTT</u> AGAGTTGCTATTAACGGTTTC (77)	<u>ATATCGAATTCCTGCAGCCCTTAAGCCT</u> TGGCAACGTGTT (78)
MC11	GPM 1	<u>CTAGAACTAGTGGATCCCCCATGCCA</u> AAGTTAGTTTTAGTTAGACACG (79)	<u>ATATCGAATTCCTGCAGCCCTTATTTCT</u> TACCTTGGTTGGCAAC (80)
MC12	GPM 2	<u>CTAGAACTAGTGGATCCCCCATGACT</u> GCAAGCACACCATCCAAT (81)	<u>ATATCGAATTCCTGCAGCCCTTAAGGAT</u> TTTTATGAAACCCTCA (82)
MC13	GPM 3	<u>CTAGAACTAGTGGATCCCCCATGACT</u> GTTACTGACACTTTTAACTG (83)	<u>ATATCGAATTCCTGCAGCCCTCATGGA</u> TTCTTTTCGAAACCC (84)
MC14	ENO1	<u>CTAGAACTAGTGGATCCCCCATGGCT</u> GTCTCTAAAGTTTACGC (85)	<u>ATATCGAATTCCTGCAGCCCTTATAATT</u> TGTCACCGTGGTGG (86)
MC15	PYK1	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> AGATTAGAAAGATTGACCTCA (87)	<u>ATATCGAATTCCTGCAGCCCTTAAACG</u> GTAGAGACTTGCAAAGTG (88)
MC16	PYK2	<u>CTAGAACTAGTGGATCCCCCATGCCA</u> GAGTCCAGATTGCA (89)	<u>ATATCGAATTCCTGCAGCCCTTAGAATT</u> CTTGACCAACAGTAGAAATG (90)
MC17	PDC1	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> GAAATTACTTTGGGTAAA (91)	<u>ATATCGAATTCCTGCAGCCCTTATTGCT</u> TAGCGTTGGTAGCAG (92)
MC18	PDC6	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> GAAATTACTCTTGGAATACT (93)	<u>ATATCGAATTCCTGCAGCCCTTATTGTT</u> TGGCATTGTAGCGG (94)
MC19	ALD6	<u>CTAGAACTAGTGGATCCCCCATGACT</u> AAGCTACACTTTGACACTGC (95)	<u>ATATCGAATTCCTGCAGCCCTTACAAC</u> TAATTCTGACAGCTTTTAC (96)
MC20	ADH5	<u>CTAGAACTAGTGGATCCCCCATGCCT</u> TCGCAAGTCATTCT (97)	<u>ATATCGAATTCCTGCAGCCCTCATTTAG</u> AAGTCTCAACAACATATC (98)
MC21	ADH6	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> TATCCTGAGAAATTTGAAGGT (99)	<u>ATATCGAATTCCTGCAGCCCTAGTCT</u> GAAAATTCCTTGTCTAGCC (100)

Strain	Gene	Forward primer (5' → 3') (SEQ ID NO:)	Reverse primer (5' → 3') (SEQ ID NO:)
MC22	TPI1	<u>CTAGAACTAGTGGATCCCCCATGGCT</u> AGAACTTTCTTTGTCGG (101)	<u>ATATCGAATTCCTGCAGCCCTTAGTTTC</u> TAGAGTTGATGATATCAACA (102)
MC23	GPD1	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> GCTGCTGCTGATAGATT (103)	<u>ATATCGAATTCCTGCAGCCCCTAATCTT</u> CATGTAGATCTAATTCTTCA (104)
MC24	GPD2	<u>CTAGAACTAGTGGATCCCCCATGCTT</u> GCTGTCAGAAGATTAACA (105)	<u>ATATCGAATTCCTGCAGCCCCTATTTCGT</u> CATCGATGCTAGCTCT (106)
MC25	HOR 2	<u>CTAGAACTAGTGGATCCCCCATGGGA</u> TTGACTACTAAACCTCTATCT (107)	<u>ATATCGAATTCCTGCAGCCCTTACCATT</u> TCAACAGATCGTCC (108)
MC26	SNF3	<u>CTAGAACTAGTGGATCCCCCATGGAT</u> CCTAATAGTAACAGTTCTAGCG (109)	<u>ATATCGAATTCCTGCAGCCCCTATTTC</u> AATCATTATTTTCATTACAGGTTG (110)
MC27	RGT2	<u>CTAGAACTAGTGGATCCCCCATGAAC</u> GATAGCCAAAACCTGC (111)	<u>ATATCGAATTCCTGCAGCCCTTATTGG</u> GGGAAGTGTATTG (112)
MC28	MIG1	<u>CTAGAACTAGTGGATCCCCCATGCAA</u> AGCCCATATCCAATG (113)	<u>ATATCGAATTCCTGCAGCCCCTCAGTCC</u> ATGTGTGGGAAGG (114)
MC29	STD1	<u>CTAGAACTAGTGGATCCCCCATGTTT</u> GTTTCACCACCTCCA (115)	<u>ATATCGAATTCCTGCAGCCCCTAGGAC</u> ATTCCATCAGGCTT (116)
MC30	ADH1	<u>CTAGAACTAGTGGATCCCCCATGTCT</u> ATCCCGAAGAACTCAAAAAGG (117)	<u>ATATCGAATTCCTGCAGCCCTTATTAG</u> AAGTGTCAACCGTATCTACC (118)
MC31	PGM 1	<u>CTAGAACTAGTGGATCCCCCATGAAC</u> GATAGCCAAAACCTGC (119)	<u>ATATCGAATTCCTGCAGCCCTTATTGG</u> GGGAAGTGTATTG (120)
MC32	PGM 3	<u>CTAGAACTAGTGGATCCCCCATGTTG</u> CAAGGAATTTTAGAAACCG (121)	<u>ATATCGAATTCCTGCAGCCCCTCAAAATT</u> TTGTAATATATTCATTTTCATCTG (122)
MC33	GAL3	<u>CTAGAACTAGTGGATCCCCCATGAAT</u> ACAAACGTTCCAATATTCAG (123)	<u>ATATCGAATTCCTGCAGCCCTTATTGTT</u> CGTACAAACAAGTACCC (124)
MC34	GAL1	CTAGAACTAGTGGATCCCCCATGACT	ATATCGAATTCCTGCAGCCCTTATAATT

Strain	Gene	Forward primer (5' → 3') (SEQ ID NO:)	Reverse primer (5' → 3') (SEQ ID NO:)
		AAATCTCATTGAGAAGAAGTGA (125)	CATATAGACAGCTGCCCA (126)
MC35	GAL4	CTAGAACTAGTGGATCCCCCATGAAG CTACTGTCTTCTATCGAACAAG (127)	ATATCGAATTCCTGCAGCCCTTACTCTT TTTTGGGTTTGGTGG (128)
MC36	GAL1	CTAGAACTAGTGGATCCCCCATGACT GCTGAAGAATTTGATTTTTC (129)	ATATCGAATTCCTGCAGCCCTTACAGTC TTGTAGATAATGAATCTGACC (130)

[0182] The nucleic acid and amino acid sequences of ENO1 are shown in SEQ ID NOs: 131 and 132, respectively. The nucleic acid and amino acid sequences of PFK2 are shown in SEQ ID NOs: 133 and 134, respectively. The nucleic acid and amino acid sequences of PFK26 are shown in SEQ ID NOs: 135 and 136, respectively. The nucleic acid and amino acid sequences of PGI1 are shown in SEQ ID NOs: 137 and 138, respectively. The nucleic acid and amino acid sequences of GMP1 are shown in SEQ ID NOS: 139 and 140, respectively. The nucleic acid and amino acid sequences of TPI1 are shown in SEQ ID NOs: 141 and 142, respectively.

Example 17: Aerobic growth on xylose

[0183] Aerobic growth on xylose was measured in mineral medium (Yeast nitrogen base, 13.4 g/l, xylose 50 g/l) in 5 ml cultures in 50 ml falcon tubes. Samples were taken and OD at 620 nm was measured. See Fig. 17: For clarity result is presented as normalized OD, against average of all results, for OD at time point of 48h. Values exceeding 1 indicate that the strain is better than the average of the strains in the experiment.

Example 18: Anaerobic growth on xylose

[0184] Yeast strains were analysed for anaerobic growth on xylose in 96-well microplate yeast cultures. Prior to experiments, yeast was grown in semi-aerobic precultures in YNB medium with 50 g/l xylose, in microplates, overnight. For measurement of anaerobic growth, cells were inoculated in 200 µl of the same medium where solution of ergosterol and Tween80 had been added at final concentration of 0.03 and 1.2 g/l, respectively. 50 µl mineral oil was added on top of each well to keep culture anaerobic. 20 µl of precultured cells were added. The growth was followed in Multiskan FC (Thermo Scientific) at 30°C and growth was measured as increase of OD (620nm). See Fig. 18: For clarity result is presented as normalized OD, against average of all results, for OD at time point of 80h. Values exceeding 1 indicate that the strain is better than the average of the strains in the experiment.

Example 19: Construction of an *S. cerevisiae* strain overexpressing an enolase, three of

the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0185] In this example, the strains, media and genetic techniques described in above were used, except that yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on YPD plates containing 500 µg mL⁻¹ geneticin (Gibco Invitrogen, Paisley, UK).

Construction of pC5e0057 containing a *S. cerevisiae* enolase (ENO1) gene

[0186] Plasmid pUG6 HXT-PGM2 (WO2010 059095 (A1)) contains the truncated HXT7' promoter and the PGK1 terminator from *S. cerevisiae* in pUG6. The DNA cassette HXT7' p-pUG6-PGK1t was PCR amplified having as template pUG6 HXT-PGM2 and using the following primers

2_fwd:

5'-TGCAGCCCGGGGATCCTTTTT-3' (SEQ ID NO: 93)

2_rev:

5'-AGGAATTCTAGATCTCCCATGTCTCT-3' (SEQ ID NO: 94).

[0187] The entire *S. cerevisiae* enolase (*ENO1*) gene was PCR amplified having as template genomic DNA from CEN.PK and using the following primers.

EN1_fwd

5'-GATCCCCGGGCTGCAATGGCTGTCTCTAAAGTTTACGC-3' (SEQ ID NO: 95)

EN1_rev

5'-AGATCTAGAATTCCTTTATAATTTGTCACCGTGGTGG-3' (SEQ ID NO: 96)

[0188] The two DNA fragments were fused together by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0057. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1051 overexpressing the gene coding for ENO1 and capable of growing on solely xylose

[0189] Plasmid pC5e0057 was digested with restriction enzyme KpnI within the ENO1 gene and it was thereafter transformed into strain C5LTe1035. This resulted in strain C5LTe1051.

Example 20: Construction of an *S. cerevisiae* strain overexpressing a 6-phosphofructokinase, three of the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0190] In this example, the strains, media and genetic techniques described above were used, except that yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on YPD plates containing 500 µg mL⁻¹ geneticin (Gibco Invitrogen, Paisley, UK).

Construction of pC5e0058 containing a *S. cerevisiae* 6-phosphofructokinase subunit beta (PFK2) gene

[0191] Plasmid pUG6 HXT-PGM2 (WO 2010/059095 (A1)) contains the truncated HXT7' promoter and the PGK1 terminator from *S. cerevisiae* in pUG6. The DNA cassette HXT7'p-pUG6-PGK1t was PCR amplified having as template pUG6 HXT-PGM2 and using the following primers.

2_fwd:

5'-TGCAGCCCGGGGATCCTTTTT-3' (SEQ ID NO: 97)

2_rev:

5'-AGGAATTCTAGATCTCCCATGTCTCT-3' (SEQ ID NO: 98)

[0192] The entire *S. cerevisiae* 6-phosphofructokinase subunit beta (PFK2) gene was PCR amplified having as template genomic DNA from CEN.PK and using the following primers.

PF2_fwd

5'-GATCCCCGGGCTGCAATGACTGTTACTACTCCTTTTGTGAATG-3' (SEQ ID NO: 99)

PF2_rev:

5'-AGATCTAGAATTCTTTAATCAACTCTTTCTTCCAACC-3' (SEQ ID NO: 100)

[0193] The two DNA fragments were fused together by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0058. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1052 overexpressing the gene coding for PFK2 and capable of growing on solely xylose

[0194] Plasmid pC5e0058 was digested with restriction enzyme KpnI within the PFK2 gene and it was thereafter transformed into strain C5LTe1035, resulting in strain C5LTe1052.

Example 21: Construction of an *S. cerevisiae* strain overexpressing a glucose-6-phosphate isomerase, three of the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0195] In this example, the strains, media and genetic techniques described above were used, except that yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on YPD plates containing 500 µg mL⁻¹ geneticin (Gibco Invitrogen, Paisley, UK).

Construction of pC5e0060 containing a *S. cerevisiae* glucose-6-phosphate isomerase (PGI1) gene

[0196] Plasmid pUG6 HXT-PGM2 (WO2010/059095 (A1)) contains the truncated HXT7' promoter and the PGK1 terminator from *S. cerevisiae* in pUG6. The DNA cassette HXT7'p-pUG6-PGK1t was PCR amplified having as template pUG6 HXT-PGM2 and using the following primers.

2_fwd:

5'-TGCAGCCCGGGGATCCTTTTT-3' (SEQ ID NO: 101)

2_rev:

5'-AGGAATTCTAGATCTCCCATGTCTCT-3' (SEQ ID NO: 102)

[0197] The entire *S. cerevisiae* glucose-6-phosphate isomerase (PGI1) gene was PCR amplified having as template genomic DNA from CEN PK and using the following primers.

PG1_fwd:

5'-GATCCCCGGGCTGCAATGTCCAATAACTCATCTCACTTCA-3' (SEQ ID NO: 103)

PG1_rev:

5'-AGATCTAGAATTCCTTCACATCCATTCCTTGAATTG-3' (SEQ ID NO: 104)

[0198] The two DNA fragments were fused together by In-Fusion cloning (Clontech, California, CA, USA). The resulting plasmid was named pC5e0060. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1054 overexpressing the gene coding for PGI1 and capable of growing on solely xylose

[0199] Plasmid pC5e0060 was digested with restriction enzyme KpnI within the PGI1 gene and it was thereafter transformed into strain C5LTe1035. This resulted in strain C5LTe1054.

Example 22: Construction of an *S. cerevisiae* strain overexpressing a 6-phosphofructo-2-kinase three of the genes in the pentose phosphate pathway and a xylulose kinase (XK) gene and expressing a *Scheffersomyces stipitis* xylitol dehydrogenase (XDH) gene and a mutated *Scheffersomyces stipitis* xylose reductase (XR(N272D)) gene

[0200] In this example, the strains, media and genetic techniques described above were used, except that yeast strains were transformed with the lithium acetate method (Gietz and Schiestl, 2007) and transformed yeast strains were selected on YPD plates containing 500 µg mL⁻¹ geneticin (Gibco Invitrogen, Paisley, UK).

Construction of pC5e0060 containing a *S. cerevisiae* 6-phosphofructo-2-kinase (PFK26) gene

[0201] Plasmid pUG6 HXT-PGM2 (WO2010/059095 A1) contains the truncated HXT7' promoter and the PGK1 terminator from *S. cerevisiae* in pUG6. The DNA cassette HXT7'-pUG6-PGK1t was PCR amplified having as template pUG6 HXT-PGM2 and using the following primers.

2_fwd:

5'-TGCAGCCCGGGGATCCTTTTT-3' (SEQ ID NO: 105)

2_rev:

5'-AGGAATTCTAGATCTCCCATGTCTCT-3' (SEQ ID NO: 106)

[0202] The entire *S. cerevisiae* 6-phosphofructo-2-kinase (PFK26) gene was PCR amplified having as template genomic DNA from CEN PK and using the following primers.

PF26_fwd:

5'-GATCCCCGGGCTGCAATGTTCAAACAGTAGACTTCTCTGA-3' (SEQ ID NO: 107)

PF26_rev:

5'-AGATCTAGAATTCCTTTAAACGTGACTTTGGCTGC-3' (SEQ ID NO: 108)

[0203] The two DNA fragments were fused together by In-Fusion cloning (Clontech, California,

CA, USA). The resulting plasmid was named pC5e0061. Correct orientations and sequences of the inserts were verified by restriction analysis and sequencing.

Construction of C5LTe1055 overexpressing the gene coding for PFK26 and capable of growing on solely xylose

[0204] Plasmid pC5e0061 was digested with restriction enzyme KpnI within the PFK26 gene and it was thereafter transformed into strain C5LTe1035. This resulted in strain C5LTe1055.

Example 23: Anaerobic growth on xylose

[0205] Yeast strains were analysed for anaerobic growth on xylose in 96-well microplate yeast cultures. Prior to experiments, yeast was grown in semi-aerobic pre-cultures in YNB medium with 50 g/l xylose, in microplates, overnight. For measurement of anaerobic growth, cells were inoculated in 250 µl of the same medium where solution of ergosterol and Tween80® had been added at final concentration of 0.03 and 1.2 g/l, respectively. 50 µl mineral oil was added on top of each well to keep culture anaerobic. 10 µl of pre-cultured cells were added. The growth was followed in Multiskan FC (Thermo Scientific) at 30°C and growth was measured as increase of OD (620nm). Data is shown graphically in Fig. 19, Fig. 20, Fig. 21, and Fig. 22.

[0206] In conclusion, strains with the gene coding for ENO1 or the gene coding for PFK2 or the gene coding for PGI1 or the gene coding for PFK26 grow faster and more under anaerobic conditions than the control strain without overexpression of any of these four genes.

Example 24: Serum flask fermentation

Fermentation procedure:

[0207] Cells aerobically grown overnight on YNB medium supplemented with 20 g/L glucose were used to inoculate a shake flask containing 50 mL of YNB medium supplemented with xylose (70 g/L) as carbon source. When the cells were growing exponentially they were used to inoculate a serum flask containing 30 mL of YNB medium supplemented with xylose (55 g/L) and a solution of ergosterol and Tween80 with final concentration of 0.03 and 1.2 g/l, respectively. 7 ml mineral oil was added on top of each serum flask to keep culture anaerobic. Concentrations of xylose, ethanol, glycerol, and xylitol were determined by high performance liquid chromatography (Waters, Milford, MA, USA).

[0208] The results are shown in Table 8 and Fig. 23. Notably, all chosen strains carrying overexpressed genes consumed more xylose and produced more ethanol than the control strain without overexpressed genes.

Table 8.

Strain	Xylose consumed in 140h (g/L)	Ethanol produced in 140h (g/L)	Ethanol yield (g EtOH/ g consumed xylose)
MC4	50	22	0.43
MC22	51	18	0.35
MC2	51	18	0.36
MC14	54	20	0.37
MC3	53	21	0.40
MC11	50	18	0.36
Control	36	13	0.35

Example 25: Preparation of strains carrying PGM1 and PGM3

[0209] Yeast strain C5LTe1101 was constructed by transforming yeast strain TMB 3043 (Karhumaa *et al.* 2005) in its *ura3* locus with a DNA fragment containing *URA3* gene, *TDH3p-XYL1(N272D)-ADH1t*, and *PGK1p-XYL2-PGK1t*. Selection was made on YNB (Yeast nitrogen base 6,7 g/l and 20 g/l glucose, 20 g/l agar) plates supplemented with 200 mg/L leucine.

[0210] C5LTe1101 was transformed with DNA fragments obtained with the following PGM1-specific PCR primers

5'-CTAGAACTAGTGGATCCCCCATGAACGATAGCCAAACTGC-3' (SEQ ID NO: 143)

5'-ATATCGAATTCCTGCAGCCCTTATTGGGGGGAAGTGTATTG-3' (SEQ ID NO: 144)

and with DNA fragments obtained with the following PGM3-specific PCR primers

5'-CTAGAACTAGTGGATCCCCCATGTTGCAAGGAATTTAGAAACCG-3' (SEQ ID NO: 145)

5'-ATATCGAATTCCTGCAGCCCTCAAATTTTGTAATATATTCATTTTCATCTG-3' (SEQ ID NO: 146)

together with the plasmid p245GPD (Mumberg et al) linearized with *Sma*I. Selection was made on YNB agar plates 5 with 20 g/l glucose as a carbon source. Selected colonies were tested by colony PCR and clones with PCR-products indicating right size of plasmid insert were chosen.

[0211] The nucleic acid and amino acid sequences of PGM1 are shown as SEQ ID NOs: 147 and 148, respectively. The nucleic acid and amino acid sequences of PGM3 are shown as SEQ ID NOs: 149 and 150, respectively.

Example 26: Anaerobic growth of strains carrying PGM1 and PGM3

[0212] Yeast strains obtained in Example 25 were analysed for anaerobic growth on xylose in 96-well microplate yeast cultures. Prior to experiments, yeast was grown in semi-aerobic precultures in YNB medium with 50 g/l xylose, in microplates, overnight. For measurement of anaerobic growth, cells were inoculated in 250 µl of the same medium where a solution of ergosterol and Tween80 had been added at final concentration of 0.03 and 1.2 g/l, respectively. 50 µl mineral oil was added on top of each well to keep the cultures anaerobic. 10 µl of precultured cells were added. The growth was followed in Multiskan FC (Thermo Scientific) at 30°C and growth was measured as increase of OD (620nm). The growth curves are presented in Fig. 24.

Example 27: Small scale fermentation of strains carrying PGM1 and PGM3

[0213] Cells aerobically grown overnight on YNB medium supplemented with 20 g/L glucose were used to inoculate a serum flask containing 30 mL of YNB medium supplemented with glucose (20 g/L) and xylose (50 g/L) as carbon source and formic acid (4.5 g/L) at initial concentration of 1 g CDW/L.

[0214] Concentrations of glucose, xylose, ethanol, glycerol, and xylitol were determined by high performance liquid chromatography (Waters, Milford, MA, USA). The compounds were separated with a Shodex SUGAR SP0810 Pb2+ copolymer-based column (Showa Denko America, NY, USA) preceded by a Micro-Guard Carbo-C guard column (Bio-Rad, Hercules, CA, USA). Separation was performed at 80°C, with H₂O at a flow rate of 0.6 ml min⁻¹ as mobile phase. Compounds were quantified by refractive index detection (Waters). A seven-point calibration curve was made for each compound to calculate concentrations. Results are presented graphically in Fig. 25 (xylose consumption) and Fig. 26 (ethanol production) as well as in Table 4.

[0215] In conclusion, these data show that strains carrying PGM1 or PGM3 consumed more xylose and produced more ethanol than the control strain.

REFERENCES

[0216]

Ausubel et al., Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons, 1995.

Chang and Ho, Appl Biochem Biotechnol. 1988, 17, 313-318.

Eliasson et al., Enzyme and Microbial Technology (2001) 29:288-297.

Gietz and Schiestl, Nature protocols (2) (2007) 31-34.

Hou, Appl Microbiol Biotechnol (2012) 94:205-214.

Inoue et al., Gene 96 (1) (1990) 23-28.

Jin et al., App Environmental Microbiol (2003) 69(1):495-503.

- Johansson and Hahn-Hägerdal, *Yeast* 19 (3) (2002) 225-231.
- Johansson et al., *Appl Environmental Microbiol* (2001) 67(9):4249-4255.
- Karhumaa et al., *Microb Cell Fact.* (2007) 5;6:5.
- Karhumaa et al., *Yeast* (2005) 22(5):359-68.
- Karhumaa et al., *Yeast* 2009; 26(7):371-82.
- Kim et al., *Biotechnol Adv.* (2013) 31(6):851-61.
- Kuyper et al. *FEMS Yeast Res.* (2005) 5(4-5):399-409.
- Linden T, Peetre J, Hahn-Hägerdal B. (1992) *Appl Environ Microbiol.*58(5):1661-9.
- Lönn et al., *Enzyme Microb. Technol.* 32 (5) (2003) 567-573.
- Madhavan et al., *Appl Microbiol Biotechnol.* 2009, 82(6), 1067-1078.
- Matsushika et al, *Enzyme and Microbial Technology* (2011) 48:466-471.
- Mollapour M, Piper PW. *Mol Microbiol.* (2001) 42(4):919-30.
- Mumberg D, Müller R, Funk M. *Gene.* 1995 Apr 14;156(1):119-22.
- Needleman and Wunsch, J. *Mol. Biol.* (1970), 48:443-453.
- Nguyen et al., *Mycological Res* (2006) 110:1232-1241.
- Parachin et al., *Metabolic Engineering* (2011) 13:508-517.
- Rice et al., *Trends Genet* 2000, 16:276-277.
- Runquist et al., *Appl. Environ. Microbiol.* 76 (23) (2010) 7796-7802.
- Shamanna and Sanderson, J. *Bacteriol.* 139 (1979) 64-70.
- Tiwari et al., *Biochemical and Biophysical Research Communications* 366 (2008) 340-345.
- Walther et al. *FEBS Lett.* 2012 Nov 30;586(23):4114-8.
- Wohlbach et al., *Proc Natl Acad Sci U S A* (2011) 108(32):13212-7).
- Xu et al., *Mol Syst Biol* 9 (2013) 665.
- US 2009/0246857
- US 2012/270289 A1
- WO 2009/017441
- WO 2010/059095

TATATCTAGGAACCCATCAGGTTGGTGGAAGATTACCGTTCCTAAGACTTTTCAGCTTCCTCTATTGA
TGTTACACCTGGACACCCCTTTTCTGGCATCCAGTTTTTAATCTTCAGTGGCATGTGAGATTCTCCGA
AATTAATTAAGCAATCACACAATTCTCTCGGATACCACCTCGGTTGAAACTGACAGGTGGTTTGTTA
CGCATGCTAATGCAAAGGAGCCTATATACCTTTGGCTCGGCTGCTGTAACAGGGAATATAAAGGGCAG
CATAATTTAGGAGTTTAGTGAAC TTGCAACATTTACTATTTTCCCCTCTTACGTAAATATTTTCTTT
TTAATTTCTAAATCAATCTTTTTCATTTTTTGTTTGTATTCTTTTCTTGCTTAAATCTATAACTACA
AAAACACATACATAAACTAAAAATGACAGTAGAACTACCCGCTTCAGAACCTTTGTTTCTTGGGTTTG
ATCTTAGCACTCAACAGTTGAAAATCATAGTGACAAACCAGAAATTAGCTGCACTAAAATCTTACAAC
GTTGAATTCGATGTGGCATTTAAGGAGAAATATGGGATAACCAAGGGAGTCTTAACGAACAAAGAGGA
CGGAGAAGTGGTTTCTCCAGTTGGTATGTGGTTAGATTCCATAAACCATGTATTTCGACCAAATGAAAC
AAGATGACTTTCGGTTCAATCAAGTTGCAGGCATTTCAGGCTCTTGTCAACAACATGGTTCTGTGTTT
TGGTCACATGAAGCTGAGAAGCTTTTATCAGGTTTACAGAAGGATCAAGATCTGTGCACTCAACTAAA
GGACGCTTTATCTTGGGACAAAAGTCCCAATTGGCAAGATCATTCGACTTTAGAGGAAAGTAAGGCTT
TCGTAGATGCTGTAGGGAGGGAAGAGTTAGCCGATATTACTGGTAGTAGAGATCACTTAAGATTCACT
GGATTGCAAATTAGGAAGTTTGCCACTAGATCACATCCCGATAAGTATGCGAATACTAGTAGAATCTC
ACTGGTTAGCTCCTTCATAACAAGTGTTCTTCTGGGTGAGATTACCGAATTGGAAGAATCTGATGCTT
GTGGCATGAAC TTGTATGACATCAAAGCCGGTGATTTCAATGAAGAATTGTTGGCTCTAGCAGCCGGT
GTTTCATCTAAAGTTGACAACATAACGAAAGATGATCCGAAATATAAAGCCGGAATTGAGGACATCAA
AGCGAACTTGGAAGATCTCCCAATTACATATAAAAGCTCCGGATCCATTGCTTCATATTACGTTG
AAAAGTACGGTTTGAATCCTAAGTGCCAGATTTACAGCTTTACCGGTGACAATTTGGCTACAATCTTG
AGTTTGCCATTACAGCCTAACGATTGCTTGATTTCGTTAGGTACTTCGACTACGGTCTTGCTAATCAC
TAAGAATTACCAACCTTCTTCTCAATATCACTTGTTTAAGCATCCAACCATACCAGATGGATATATGG
GCATGATCTGCTATTGCAATGGCTCTTTGGCCAGAGAGAAGATAAGAGATGAAGTTAATGAATACTAT
AAGGTGGAAGATAAGAAGAGTTGGGATAAAATTTAGTGAAATTC TGATAAGTCGACCAAATTCGATAA
TAAGCTGGGTATTTTCTTTCCGTTAGGTGAAATCGTTCCACAGGCAAGGCACAACTGTCAGAGCAG
TATTGGAGAATGACAAAGTCATAGAAGTAGGTTTGATACACACGGATTTGATATTGATCAGCAGCA
AGAGCTATTGTGAAAGCCAAGCCTTATCTTG TAGACTTAGAGCTGGCCCTATGTTATCCAAATCATC
ACGTGCTTCCGTCACATCACCAACGGAGTTAAAAGGCGTATACCATGACATAGTGGCCAAATATGGTG
ACCTGTACACAGATGGTAAACAACAAACCTATGAGTCACTTACATCTAGGCCAAATCGTTGTTTCTAT
GTTGGCGCGGGGAGCAATAACATTTCCATCATTAGTAAAATGGGTTCTATTCTAGGTCTGTTACGG
TAAGTCTAACTGCATATTCGAAAGCCGCTGCTCTAGCTGCAGGATGCAAGCATGCGCTCTAT

TAACCTTAAGTCGATATCCAAACGGTGTCTCTAGGTGGAGCATACAAAGCATCCCTGGCTTATG
 AATGTGAACAGAAAGGTGAATGGATTAAATTATGACCAATACATAAAATCAGTTATTGAAAGAATTGAAG
 TCATTTAACGTGGAGGACAAATGGTTAGAATACTTTGATGGAGTTGCGCTTTTGGCTAAGATGGAAGA
 AACCCCTGTTGAAATAAAGATCTCCCATGTCTCTACTGGTGGTGGTGCTTCTTTGGAATTATTGGAAGG
 TAAGGAATTGCCAGGTGTTGCTTTCTTATCCGAAAAGAAATAAATTGAATTGAATTGAAATCGATAGA
 TCAATTTTTTTCTTTTCTCTTTCCCATCCTTTACGCTAAAATAATAGTTTATTTTATTTTGAATA
 TTTTTTATTTATATACGTATATATAGACTATTATTTATCTTTTAATGATTATTAAGATTTTTATTAAA
 AAAAAATTCGCTCCTCTTTTAATGCCTTTATGCAGTTTTTTTTTCCCATTGATATTCTATGTTT

SEQ ID NO: 6 - *Spathaspora passalidarum* XK

MTVELPASEPLFLGFDLSTQQLKIIIVTNQKLAALKSYNVEFDVAFKEYGITKGVLTNKEDGEVVS
 GMWLD SINHVFDQMKQDDFPFNQVAGISGSCQQHGSVFWSEAEKLLSGLQKDQDLSTQLKDALS
 SWDKSPNWQDHSTLEESKAFVDAVGREELADITGSRDHLRFTGLQIRKFATRSHPKYANTSRISLV
 SSFITSVLLGEITELEESDACGMNLYDIKAGDFNEELLALAAGVHPKVDNITKDDPKYKAGIEDI
 KAKLGKIS PITYKSSGSIASYVEKYGLNPKCQIYSFTGDNLATILSLPLQPNDC LISLGTSTTV
 LLITKNYQPSS QYHLFKHPTIPDGYMGMICYCNGSLAREKIRDEVNEYKVEDKKSWDKFSEIL
 DKSTKFDNKLGIFFP LGEIVPQAKAQTVRAVLNDKVIEVGLDTHGFDIDHDARAIVESQALSC
 RL RAGPMLS KSSRASVTSP TELKGVYHDIVAKYGDLYTDGKQQTYESLTSRPNRCFYVGG
 SNNISII SKMGSILGPVHGNFKVDIP NACSLGGAYKASWSYECEQKGEWINYDQYINQLL
 KELKSFNVEDKWLEYFDGVALLAKMEETLLK

SEQ ID NO: 17 - Artificial *Escherichia coli* xylulose kinase gene

Artificial *Escherichia coli* xylulose kinase gene with TPI1 promoter before the ATG-start codon and the PGK1 terminator after the stop-codon, and coding region based on the amino acid sequence from NCBI Accession Code YP_001460359:

TATATCTAGGAACCCATCAGGTTGGTGGAGATTACCGTTCTAAGACTTTTCAGCTTCCTCTATTGA
 TGTTACACCTGGACACCCCTTTTCTGGCATCCAGTTTTTAATCTTCAGTGGCATGTGAGATTCTCCGA
 AATTAATTAAAGCAATCACACAATTCTCTCGGATACCACCTCGGTTGAAACTGACAGGTGGTTTGTTA
 CGCATGCTAATGCAAAGGAGCCTATATACCTTTGGCTCGGCTGCTGTAACAGGGAATATAAAGGGCAG
 CATAATTTAGGAGTTTAGTGAACCTTGCAACATTTACTATTTTCCCTTCTTACGTAAATATTTTCTTT
 TTAATTCTAAATCAATCTTTTTCAATTTTTTGTGTGATTCTTTTCTTGCTTAAATCTATAACTACAA
 AAAACACATACATAAACTAAAAATGTATATCGGCATTGATTTGGGTACTTCTGGCGTAAAGGTTATCC
 TGCTGAATGAACAGGTTGAAGTGGTTGCCTCACAAACGGAAAAGTTGACTGTATCTAGGCCACATCCT
 TTGTGGAGCGAACAAGATCCAGAACAGTGGTGGCAAGCTACAGATAGAGCAATGAAAGCGTTAGGTGA
 CCAGCATTCCTTACAGGACGTTAAAGCCTTGGGGATTGCTGGCCAAATGCATGGTGGGACACTGCTTG
 ATGCCAACAAAGGGTCTTAAGGCCTGCAATACTGTGGAATGATGGACGTTGTGCTCAGGAGTGTAAC
 TTATTGGAAGCAAGAGTGCCTCAATCCAGGGTGATAACCGGTAACCTTGATGATGCCTGGATTACAGC
 CCCAAAATTGTTATGGGTTCAAAGACACGAACCAGAGATCTTCCGTCAAATCGACAAGGTCTTATTAC
 CGAAGGACTACTTGAGACTACGTATGACTGGTGAATTCGCTTCAGACATGAGTGACGCAGCAGGAACC
 ATGTGGTTGGATGTCGCGAAAAGAGATTGGAGTGACGTTATGTTACAAGCTTGGATCTATCTAGAGA
 TCAAATGCCAGCTCTGTATGAGGGCTCAGAAATTACCGGTGCATTATTACCTGAAGTCGCTAAAGCAT
 GGGGTATGGCTACTGTCCCAGTTGTTGCCGGTGGTGGTGACAATGCCGCAGGAGCTGTTGGAGTTGGT
 ATGGTGGATGCAAATCAAGCGATGTTGTCTCTTGGCACATCAGCGCTCTATTTGCCGTATCGGAAGG
 GTTTCTGTGCGAAACCAGAATCAGCCGTACATTCTTTTGTACGCTCTTCCACAAAGATGGCATCTAA
 TGAGCGTGATGCTTTCTGCAGCATCATGCTTGGATTGGGCCGCTAAATTGACGGGTTTGAGTAATGTT
 CCGGCACCTTATAGCAGCTGCACAACAAGCAGATGAAAGTGCTGAACCCGTTTGGTTCTTGCCCTATCT
 TTCCGGAGAGAGAACACCACACAACAATCCTCAAGCCAAAGGTGTGTTCTTTGGGTAACTCACCAAC

ATGGTCCAAACGAATTGGCGAGAGCAGTATTGGAAGGAGTAGGGTATGCTCTTGCTGATGGTATGGAT
 GTTGTCCATGCATGTGGCATAAAGCCGCAATCTGTTACGCTTATTGGAGGTGGTGCCAGAAGCGAATA
 CTGGAGACAAATGTTAGCCGATATTTCCGGTCAACAACCTAGACTACAGAACAGGAGGCGATGTAGGGC
 CAGCTTTGGGTGCTGCTAGATTGGCTCAGATTGCTGCTAACCTGAGAAGTCGTTGATTGAGCTACTA
 CCTCAGTTACCCTTAGAACAGTCTCATCTACCAGATGCCCAGAGATATGCTGCGTACCAACCTAGAAG
 AGAGACTTTTCGTAGGTTATACCAGCAATTACTACCCTTGATGGCGTAAAGATCTCCCATGTCTCTAC
 TGGTGGTGGTGCTTCTTTGGAATTATTGGAAGGTAAGGAATTGCCAGGTGTTGCTTTCTTATCCGAAA
 AGAAATAAATTGAATTGAATTGAAATCGATAGATCAATTTTTTTCTTTTCTTTTCCCCATCCTTTAC
 GCTAAAATAATAGTTTATTTTATTTTTTGAATATTTTTTATTATATACGTATATATAGACTATTATT
 TATCTTTTAATGATTATTAAGATTTTTTATAAAAAAAATTCGCTCCTCTTTTAATGCCTTTATGCAG
 TTTTTTTTTCCCATTCGATATTTCTATGTTC

SEQ ID NO: 18 - *Escherichia coli* xylulose kinase

MYIGIDLGTSGVKVILLNEQGEVVASQTEKLTVSRPHPLWSEQDPEQWWQATDRAMKALGDQHSLODV
 KALGIAGQMHGATLLDAQQRVLRPAILLWNDGRCAQECTLLEARVPQSRVITGNLMPGFTAPKLLWVQ
 RHEPEIFRQIDKVLLEPKDYLRLRMTGEFASDMSDAAGTMWLDVAKRDWSDVMLQACDLSDQMPALYE
 GSEITGALLPEVAKAWGMATVPVVAGGGDNAAGAVGVGMVDANQAMLSLGTSGVYFAVSEGFLSKPES
 AVHSFCHALPQRWHLSVMLSAASCLDWAAKLTGLSNVPALIAAAQQADESAEPVWFLPYLSGERTPH
 NNPQAKGVFFGLTHQHGPNELARAVLEGVGYALADGMDVVHACGIKPQSVTLIGGGARSEYWRQMLAD
 ISGQQLDYRTGGDVGPALGAARLAQIAANPEKSLIELLPQLPLEQSHLPDAQRYAAYQPRRETFRRLY
 QQLLPLMA

SEQ ID NO: 21 - *Kluyveromyces marxianus* xylulose kinase gene

Artificial *Kluyveromyces marxianus* xylulose kinase gene, with coding region based on the amino acid sequence from NCBI Accession Code ADW23548:

ATGTCGACTCCATATTACTTAGGCTTTGATTTGTCAACTCAGCAGTTGAAATGTCTAGCAATAGATGA
 TCAGTTAAACATCGTGACTTCAGTTAGCATTGAATTTGACCGTAATTTCCCAGCTTACAACACAAAGA
 AAGGGGTATACATCAAGAATGGTGGTGTGATAGACGCACCAAGTTGCTATGTGGTTAGAAGCTGTTGAT
 TTATGCTTTAGTCAGCTTGCTGAAAGAATCGACTTGAAGAGAGTTCAATCAATGTGAGGTTCTTGCCA
 ACAACATGGCACCCTCTACTGGAAGTGTGAGCATCTACCAAGTAACCTTGATCCTGCTTCAACCTTGA
 GAGAGCAACTTCAAGGCAGTTTATCAAGACCAGTTGCACCCAATTGGCAAGATCATTCCACCAAGAAA
 CAATGTGATGAATTGGCAGAATCGGTAGGAGGACCTGAAGAACTAGCAAGGATTACAGGTTCTGGAGC
 AACTATAGATTTTCCGGTTCCCAAATTGCCAAAATCCATGAACTGAACCTGAAGTCTATGAGGCTA
 CTAAACGTATTTTCGTGGTAAGTAGCTTTCTAGCGTCTGTTTTAGTAGGTGACATTGTGCCCTTGGA
 GAAGCGGATGCTTGTGGCATGAACCTTATACGATCTATCCAAACACGACTTTGACGAAACATTACTGGC
 TGTGCTTGACCATGATACAGCGAGATTGAGAAGGAACTATCAGATCCACCCGTTGGAGCTCCTACCA
 GAGAAAGTCCTCTGACCTCCTTGGGTAAAGTCTCTAAGTACTTTCAGGACAAATATGGGGTTAACTGT
 GAATGTGAGATCTTCCCGTTTACTGGCGATAACCTGGCAACGATTTGTTCCCTACCTTTGCAAAAGAA
 TGATGTCTTGATTAGTCTAGGTACTTCGACCACGATTTTGTGGTAACTGACCAATATCACTCTTCTC
 CCAATTATCACTTGTTTATACATCCGACAGTGCCAGGTTATTACATGGGTATGATTTGCTATTGCAAT
 GGGTCTTTGGCTCGTGAGAGAGTAAGAGATGATCTGGCTGGACCACAAGCCTCTCAAGCTCCTGGGGA
 GCAAGTTCCATGGACTCAATTCAATGACGCATTACTGGATGACTCATTGAGCAATGACAATGAGATAG
 GCCTTTACTTCCCTCTTGGTGAGATTGTCCCAAATGTTGATGCCGTCACCAAAGATGGACATTCGAA
 AGAAAAGAGAACCATTGCAACAAAAGTATCGTTCTTCACGAGTTGGATCAATTCACGCCAAAGAGGAA
 AGATGCAAGAACATAGTGGAAGTCAGGCCTTAAGCTGTAGAGTGCGTATCTCTCCATTGCTGTCTG
 ATGAAACAGATGCCTTAAGCGAACTCAAGTGTGTCAAAGAAGGAGAATACCCAAGTTACGTTTGAC
 TACGATGCATTTCCGTTGTGGACGTATGCCAAAAGACCGAATAGAGCGTTCTTTGTTGGTGGTGGCTC

CAAGAATGATGCCATAGTCAGGACAAATGGCAAATGTAATAGGTGCTAGGAATGGAAATTATAGACTTG
 AAACCTCCCAATTCCTGTGCTTTAGGAGGCTGTTATAAAGCGATGTGGTCATGGTTAAAGGTACATGAA
 CCTACAACCTACACCATCTTTTCGATGTTTGGTTAAACGCAAGCTTTAACTGGCAGAGAGATTGCGAATT
 CGTGTGCCAGTCTGACGCCGCTAAGTGGGAACAATCTAATGGTAAAATTCAAGCTTTATCAGAAGCTG
 AAGCCTATGTTAAAGCTTTAGCACAAATCTCAGGGTCAA

SEQ ID NO: 22 - *Kluyveromyces marxianus* xylulose kinase

MSTPYYLGFDLSTQQLKCLAIDDLNIVTSVSIIEFDRNFPAYNTKKGVIYIKNGGVIDAPVAMWLEAVD
 LCFSQLAERIDLKRVQSMGSCQQHGTIVWNCEHLPSNLDPASTLREQLQGSLSRPVAPNWQDHSTKK
 QCDELAESVGGPEELARITGSGAHYRFGSGQIAKIHETEPEVEYATKRISLVSSFLASVLVGDIVPLE
 EADACGMNLYDLSKHDFDETLAVVDHDTARLRRKLSDPVVGAPTRESPLTSLGKVSKEYQDKYGVNC
 ECEIFPFTGDNLATICSPLQKNDVLISLGTSTTILLVTDQYHSSPNYHLFIHPTVPGYYMGMICYCN
 GSLARERVRDDLAGEPQASQAPGEQVPWTQFNDALLDDSLSDNEIGLYFPLGEIVPNVDAVTKRWTFE
 RKENHSNKSIVLHELDQFTPKRKDAKNIVESQALSCRVRISPLLSDETDLSETQVLSKKENTQVTFD
 YDAFPLWTYAKRENRAFFVGGASKNDIVRTMANVIGARNGNYRLETPNSCALGGCYKAMWSWLKVHE
 PTTTPSFDVWLNASEFNWQRDCEFVCQSDAAKWEQSNKIQALSEAEAYVKALAQSQGQ

SEQ ID NO: 31 - *Saccharomyces cerevisiae* xylulokinase gene (XKS1)

ATGTTGTGTTTCAGTAATTCAGAGACAGACAAGAGAGGTTTCCAACACAATGTCTTTAGACTCATACTA
 TCTTGGGTTTGATCTTTTCGACCCAACAACCTGAAATGTCTCGCCATTAAACCAGGACCTAAAAATTGTCC
 ATTCAGAAACAGTGGAAATTTGAAAAGGATCTTCCGCATTATCACACAAAGAAGGGTGTCTATATACAC
 GCGGACACTATCGAATGTCCCGTAGCCATGTGGTTAGAGGCTCTAGATCTGGTTCTCTCGAAATATCG
 CGAGGCTAAATTTCCATTGAACAAAGTTATGGCCGTCTCAGGGTCCTGCCAGCAGCACGGGTCTGTCT
 ACTGGTCTCTCCCAAGCCGAATCTCTGTTAGAGCAATTGAATAAGAAACCGGAAAAAGATTTATTGCAC
 TACGTGAGCTCTGTAGCATTTGCAAGGCAAACCGCCCCCAATTGGCAAGACCACAGTACTGCAAAGCA
 ATGTCAAGAGTTTGAAAGAGTGATAGGTGGGCCTGAAAAAATGGCTCAATTAACAGGGTCCAGAGCCC
 ATTTTAGATTTACTGGTCTCTCAAATCTGAAAATTGCACAATTAGAACCAGAAGCTTACGAAAAAACA
 AAGACCATTCTTTTAGTGTCTAATTTTTTGACTTCTATCTTAGTGGGCCATCTTGTGAATTAGAGGA
 GGCAGATGCCTGTGGTATGAACCTTTATGATATACGTGAAAGAAAATTCAAGTATGAGCTACTACATC
 TAATTGATAGTTCTTCTAAGGATAAACTATCAGACAAAAATTAATGAGAGCACCCATGAAAAATTTG
 ATAGCGGGTACCATCTGTAAATATTTTATTGAGAAGTACGGTTTCAATACAACTGCAAGGTCTCTCC
 CATGACTGGGGATAATTTAGCCACTATATGTTCTTTACCCCTGCGGAAGAATGACGTTCTCGTTTCCC
 TAGGAACAAGTACTACAGTTCTTCTGGTCACCGATAAGTATCACCCCTCTCCGAACCTATCATCTTTTC
 ATTCATCCAACCTCTGCCAAACCATTATATGGGTATGATTTGTTATTGTAATGGTTCTTTGGCAAGGGA
 GAGGATAAGAGACGAGTTAAACAAAGAACGGGAAAATAATTATGAGAAGACTAACGATTGGACTCTTT
 TTAATCAAGCTGTGCTAGATGACTCAGAAAAGTAGTGAAAATGAATTAGGTGTATATTTTCTCTGGGG
 GAGATCGTTCTTAGCGTAAAAGCCATAAAACAAAAGGGTTATCTCAATCCAAAACGGGTATGATTGA
 AAGAGAGGTGGCCAAGTTCAAAGACAAGAGGCACGATGCCAAAAATATTGTAGAATCACAGGCTTTAA
 GTTGCAAGGTGAAGAATATCTCCCTGCTTTTCGGATTCAAACGCAAGCTCACAAACAGAGACTGAACGAA
 GATACAATCGTGAAGTTTGATTACGATGAATCTCCGCTGCGGGACTACCTAAATAAAAGGCCAGAAAG
 GACTTTTTTTGTAGGTGGGGCTTCTAAAAACGATGCTATTGTGAAGAAGTTGCTCAAGTCATTGGTG
 CTACAAAGGGTAATTTTAGGCTAGAAACACCAAACCTCATGTGCCCTTGGTGGTTGTTATAAGGCCATG
 TGGTCATTGTTATATGACTCTAATAAAATTGCAGTTCCTTTTGATAAATTTCTGAATGACAATTTTCC
 ATGGCATGTAATGAAAAGCATATCCGATGTGGATAATGAAAATTGGGATCGCTATAATTCCAAGATTG
 TCCCCTTAAGCGAACTGGAAAAGACTCTCATCTAA

SEQ ID NO: 32 - *Saccharomyces cerevisiae* xylulokinase

MLCSVIQRQTREVSNTMSLDSYYLGFDLSTQQLKCLAINQDLKIVHSETVEFEKDLPHYHTKKGVIYIH
 GDTIECPVAMWLEALDLVLSKYREAKFPLNKVMAVSGSCQQHGSVYWSSQAESLLEQLNKKPEKDLLH
 YVSSVAFARQTAPNWQDHSTAKQCQEFEECIGGPEKMAQLTGSRAHFRFTGPQILKIAQLEPEAYEKT
 KTISLVSNFLTSLVGHLEEEADACGMNLYDIRERKFSDELLHLIDSSSKDKTIRQKLMRAPMKNL
 IAGTICKYFIEKYGFNTNCKVSPMTGDNLATICSLPLRKNDVLVSLGTSTTVLLVTDKYHPSPNYHLF
 IHPTLPNHYMGMICYCNGSLARERIRDELNKERENNYEKTNDWTLEFNQAVLDDSESENELGVYFPLG
 EIVPSVKAINKRVI FNPKTGMIEREVAKFKDKRHDAKNIVESQALSCRVRISPLSDSNASSQORLNE
 DTIVKFDYDESPLRDYLNKRPERTEFFVGGASKNDAIVKKFAQVIGATKGNFRLETNPSCALGGCYKAM
 WSLLYDSNKIAVPFDKFLNDNFPWHVMESISDVDNENWDRYNSKIVPLSELEKTLI *

SEQ ID NO: 49 - Codon optimized *Zygosaccharomyces bailii* YME2 gene

ATGTTGCCCATTTCTGGACCTTCCAACATGCTGCATGGCCTCGTTTCAGCCCGTTGTGCAGGGGGTTG
 GAGGCCACTTATCTCGCATTTGCGTAGGGGAGTTTTTCCTAAGATGCTTACCATGACAGGTATTGGGG
 CCAAGAGATTTGTCTCCAGCGAAATACAGGAGAAAGACGAACAAGCTGGTGAGTCTACTACTGCTACA
 GATACTGGTATCATTATAAAACGGAGCAGGAGACCTAGTATATTTGACAACGTCTATCCACGGAC
 CGCATCTCTATGGAGCCCTGCGCAATGGTACAATCTACTTCTAACTAATCAATCGAGGGAGGCTGTTA
 GGCAAAAGATCAGCGGTTGCGATCCCGCTAGAGACCATTTTTTGGCTTCATTGA

SEQ ID NO: 50- *Zygosaccharomyces bailii* YME2

MLPISGPSNMLHGLVSARCAGGWRPLISHLRRGVFPKMLTMTGIGAKRFVSSEIQEKDEQAGESTTATDT
 GIIHKTEQETLVYFDNVYPRASLWSPAQWYNLLLTNQSREAVRQKISGSIPLETIFWLH

SEQ ID NO: 131 - *Saccharomyces cerevisiae* ENO1 DNA

ATGGCTGTCTCTAAAGTTTACGCTAGATCCGCTCTACGACTCCCGTGGTAACCCAACCGTCGAAGTCGAAT
 TAACCACCGAAAAGGGTGTTCAGATCCATTGTCCCATCTGGTGCTTCTACCGGTGTCCACGAAGCTTT
 GGAAATGAGAGATGGTGACAAATCCAAGTGGATGGGTAAGGGTGTTCACGCTGTTAAGAACGTCAAC
 GATGTCATTGCTCCAGCTTTTGGTAAAGGCTAACATTGATGTTAAGGACCAAAAGGCCGTCGATGACTTCT
 TGATTTCTTTGGACGGTACTGCCAACAAATCCAAGTGGGTGCTAACGCTATCTGGGTGTTTCTTTGGC
 TGCTTCCAGAGCTGCCGCTGCTGAAAAGAAATGTCCCATTTATACAAGCACTTGGCTGACTTGCTTAAGTCC
 AAGACCTCTCCATACGTTTTTGGCAGTTCCATTCTTGAACGTTTTGAACGGTGGTCCCACGCTGGTGGTG
 CTTTGGCTTTGCAAGAATTTATGATTGCTCCAAGTGGTGCTAAGACCTTCGCTGAAGCTTTGAGAATTGG
 TTCCGAAGTTTACCACAACCTTGAAGTCTTTGACCAAGAAGAGATACGGTGCTTCTGCCGGTAACGTCGGT
 GACGAAGGTGGTGTGCTCCAAACATTCAAAGTCTGAAGAAGCTTTGGACTTGATTGTTGACGCTATCA
 AGGCTGCTGGTCACGACGGTAAGATCAAGATCGGTTTGGACTGTGCTTCCTCTGAATTCTTCAAGGACGG
 TAAGTACGACTTGGACTTCAAGAATCCAAACTCTGACAAATCCAAGTGGTTGACTGGTCTCAATTGGCT
 GACTTGTACCACTCCTTGATGAAGAGATACCAATTGTCTCCATCGAAGATCCATTTGCTGAAGATGACT
 GGAAGCTTGGTCTCACTTCTTCAAGACCGCTGGTATTCAAATTGTTGCTGATGACTTGACTGTCACCAA
 CCCAAAGAGAATTGCTACCGCTATCGAAAAGAAGGCTGCCGACGCTTTGTTGTTGAAGGTCAACCAAATC
 GGTACCTTGTCTGAATCCATCAAGGCTGCTCAAGACTCTTCGCTGCCGGTTGGGGTGTTATGGTTTCCC
 ACAGATCTGGTGAAGTGAAGACACTTTCATTGCTGACTTGGTCGTCGGTTTGAGAACTGGTCAAATCAA
 GACTGGTGCTCCAGCTAGATCCGAAAGATTGGCTAAATTGAACCAATTGTTGAGAATCGAAGAAGAATTA
 GGTGACAACGCTGTTTTGCTGGTGAAAACCTCCACCACGGTGACAAATTATAA

SEQ ID NO: 132 - *Saccharomyces cerevisiae* ENO1

MAVSKVYARSVYDSRGNPTVEVELTTEKGVFRSIVPSGASTGVHEALEMRDGDKSKWMGKGV LHAVKNVN
 DVIAPAFVKANIDVKDQKAVDDFLISLDGTANKSKLGANAILGVSLAASRAAAAEKNVPLYKHLADLSKS

KTSPLYVLPVPFLNLVNGGSHAGGALALQEFMIAPTGAKTFAEALRIGSEVYHNLKSLTKKRYGASAGNVG
DEGGVAPNIQTAEELDLIVDAIKAAGHDGKIKIGLDCASSEFFKDGKYDLDFKNPNSDKSKWLTGPQLA
DLYHSLMKRYPIVSIEDPFAEDDWEAWSHFFKTAGIQIVADDLTVTNPKRIATAIEKKAADALLLKVNQI
GTLSESIKAAQDSFAAGWGMVSHRSGETEDTFIADLVVGLRTGQIKTGAPARSERLAKLNQLLRIEEEEL
GDNVAFAGENFHHGDKL*

SEQ ID NO: 133 - *Saccharomyces cerevisiae* PFK2 DNA

ATGACTGTTACTACTCCTTTTGTGAATGGTACTTCTTATTGTACCGTCACTGCATATTCCGTTCAATCTT
ATAAAGCTGCCATAGATTTTACACCAAGTTTTTGTCAATTAGAAAACCGCTCTTCTCCAGATGAAAACCTC
CACTTTATTGTCTAACGATTCCATCTCTTTGAAGATCCTTCTACGTCCTGATGAAAAAATCAATAAAAAAT
GTTGAGGCTCATTGAAGGAATTGAACAGTATTACCAAGACTCAAGACTGGAGATCACATGCCACCCCAAT
CCTTGGTATTTAACACTTCCGACATCTTGGCAGTCAAGGACACTCTAAATGCTATGAACGCTCCTCTTCA
AGGCTACCCAACAGAAGTATTTCCAATGCAGTTGTACACTTTGGACCCATTAGGTAAAGCTTGGTGGTGT
ACTTCTACTAAGAACGCAGTTTCAACCAAGCCAACTCCACCACCAGCACCAGAAGCTTCTGCTGAGTCTG
GTCTTTCTCTAAAGTTCACTCTTACACTGATTTGGCTTACCGTATGAAAACCACCGACACCTATCCATC
TCTGCCAAAGCCATTGAACAGGCCCTCAAAGGCAATTGCCGTCATGACTTCCGGTGGTGATGCTCCAGGT
ATGAAGTCTAACGTTAGAGCCATCGTGCGTTCCGCTATCTTCAAAGGTTGTCGTGCCTTTGTTGTCATGG
AAGGTTATGAAGGTTTGGTTTCGTGGTGGTCCAGAATACATCAAGGAATTCCACTGGGAAGACGTCCGTGG
TTGGTCTGCTGAAGGTGGTACCAACATTTGGTACTGCCCGTTGTATGGAATTCAGAAGCGCGAAGGTAGA
TTATTGGGTGCCCCAATTTGATTGAGGCCGGTGTGATGCTTTGATCGTTTGTGGTGGTGACGGTTCTT
TGACTGGTGTGATCTGTTTAGATCAGAATGGCCTTCTTTGATCGAGGAATTGTTGAAAACAAACAGAAT
TTCCAACGAACAATACGAAAGAATGAAGCATTTGAATATTTGCCGTTACTGTCGGTTCTATTGATAACGAT
ATGTCCACCACGGATGCTACTATTGGTGCTTACTCTGCCTTGGACAGAATCTGTAAGGCCATCGATTACG
TTGAAGCCACTGCCAACTCTCACTCAAGAGCTTTCTGTTGTTGAAGTTATGGGTAGAACTGTGGTTGGTT
AGCTTTATTAGCTGGTATCGCCACTTCCGCTGACTATATCTTTATTCCAGAGAAGCCAGCCACTTCCAGC
GAATGGCAAGATCAAATGTGTGACATTGTCTCCAAGCACAGATCAAGGGGTAAGAGAACACCATTGTTG
TTGTTGCAGAAGGTGCTATCGCTGCTGACTTGACCCCAATTTCTCCAAGCGACGTCCACAAAGTTCTAGT
TGACAGATTAGGTTTGGATACAAGAATTACTACCTTAGGTACAGTTCAAAGAGGTGGTACTGCTGTTGCT
TACGACCGTATCTTGGCTACTTTACAAGGCTTTGAGGCCGTTAATGCCGTTTGGGAATCCACTCCAGACA
CCCCATCACCATTGATTGCTGTTAACGAAAACAAATTTGTTGTAACCATTAATGGAATCCGTCAAGTT
GACCAAAGCAGTTGCAGAAGCCATTCAAGCTAAGGATTTCAAGAGAGCTATGCTTTAAGAGACACTGAG
TTCATTGAACATTTAAACAATTTCAATGGCTATCAACTCTGCTGACCACAACGAACCAAAGCTACCAAAGG
ACAAGAGACTGAAGATTGCCATTTGTTAATGTGCGTGCTCCAGCTGGTGGTATCAACTCTGCCGTCTACTC
GATGGCTACTTACTGTATGTCCCAAGGTCACAGACCATACGCTATCTACAATGGTTGGTCTGGTTTGGCA
AGACATGAAAGIGTTTCGTTCTTTGAACTGGAAGGATATGTTGGGTTGGCAATCCCGTGGTGGTTCTGAAA
TCGGTACTAACAGAGTCACTCCAGAAGAAGCAGATCTAGGTATGATTGCTTACTATTTCCAAAAGTACGA
ATTTGATGGTTTGATCATCGTTGGTGGTTTTCGAAGCTTTTGAATCTTTACATCAATTAGAGAGAGCAAGA
GAAAGTTATCCAGCTTTTCAAGATCCCAATGGTCTTGATACCAGCTACTTTGTCTAACAATGTTCCAGGTA
CTGAATACTCTTTGGGTTCTGATACCGCTTTGAATGCTCTAATGGAATACTGTGATGTTGTTAAACAATC
CGCTTCTTCAACCAGAGGTAGAGCCTTCGTTGTGATGTCAGGTTGGTAACTCAGGCTATTTGGCCACT
TACGCTTCTTTGGCTGTTGGTGTCTCAAGTCTCTTATGTCCTCAGAGAAGGATTTCTTTGGAGCAATTGT
CCGAGGATATTGAATACTTAGCTCAATCTTTTGAAGGCGAGAAGGTAGAGGTAGATTGGTAAATTTGAT
TTTGAAGAGTACAAACGCTTCTAAGGCTTTATCAGCCACTAAATTTGGCTGAAGTTATTACTGCTGAAGCC
GATGGCAGATTTGACGCTAAGCCAGCTTATCCAGGTCATGTACAACAAGGTGGTTTGGCATCTCCAAATTG
ATAGAACAAGAGCCACTAGAATGGCCATTAAAGCTGTGCGCTTCATCAAAGACAACCAAGCTGCCATTGC

TGAAGCTCGTGCTGCCGAAGAAAACCTTCAACGCTGATGACAAGACCATTTCTGACACTGCTGCTGTCCTT
GGTGTAAAGGGTTACATGTCGTTTACAACCTCATTAGACAATTGTATGACTATGAACTGAAGTTTCCA
TGAGAATGCCAAAGGTCATTCACCTGGCAAGCTACCAGACTCATTGCTGACCATTGGTTGGAAGAAAGAG
AGTTGATTAA

SEQ ID NO: 134 - *Saccharomyces cerevisiae* PFK2

MTVTTTFFVNGTSYCTVTAYSVQSYKAAIDFYTKFLSLENRSPDENSTLLSNDISLKILLRPDEKINKN
VEAHLKELNSITKTQDWRSHATQSLVFENTSDILAVKDTLNMNAPLQGYPTLFPMLQYTLDPNGVVG
TSTKNAVSTKPTPPPAFEASAESGLSSKVHSYTDLAYRMKTTDTYPSLPKPLNRPQKAIAMVMTSGGDAPG
MNSNVRAIVRSALFKGCRAFFVMEGYEGLVRGGPEYIKEFHWDVVRGWSAEGGTNIGTARCMFEKKREGR
LLGAQHLIEAGVDALIVCGDGLTGADLFRSEWPSLIEELLKTNRISNEQYERMKHLNICGTVGSIDND
MSTTDATIGAYSALDRICKAIDYVEATANSHSRAFFVVEVMGRNCGWLALLAGIATSADYIFIPEKPATSS
EWQDQMCDIVSKHRSRGKRTTIVVVAEGAIADLTPIPSDVKHVLVDRLGLDTRITTLGHVQRGGTAVA
YDRILATLQGLEAVNAVLESTPDTPSPLIAVNENKIVRKPLMESVKLTKAVAEAIQAKDFKRAMSLRDE
FIEHLNNFMAINSADHNEPKLPKDKRLKIAIVNVGAPAGGINSVYSMATYCMSQGHRPYAIYNGWSGLA
RHESVRSNLNWKMLGWQSRGGSEIGTNRVTPEEADLGMIAYYFQKYEFDGLIIVGGFEAFESLHQLERAR
ESYPAFRIPMVLIPATLSNNVPGTEYSLGSDTALNALMEYCDVVKQSASSTRGRAFFVDCQGGNSGYLAT
YASLAVGAQVSYVPEEGISLEQLSEIDIEYLAQSFKAEGRGFRGKLILKSTNASKALSATKLAEVITAEA
DGRFDAKPAYPGHVQGGGLPSPIDRTRATRMAIKAVGFIKDNQAAIAEARAEEENFNADDKTI SDTAAVV
GVKGS HVVYNSIRQLDYETEVS MRMPKVIHWQATRLIADHLVGRKRVD*

SEQ ID NO: 135 - *Saccharomyces cerevisiae* PFK26 DNA

ATGTTCAAACCATGACTTCTCTGAAACATCTCCTGTGCCGCTGATATTGATCTTGCTCCTACACAAT
CTCCACACCATGTGGCACCTAGTCAAGACTCCAGTTATGATCTTTTATCCCGGAGTTCCGATGATAAAAT
TGATGCTGAAAAGGGTCCGCATGATGAATTATCTAAGCACTTACCCTTTTTTTCAGAAAAGACCTTTGAGC
GATACTCCTATATCGAGCAATTGGAACCTCTCCTGGAATCACTGAAGAAAATACACCTTCTGACTCTCCTG
AAAAATAGCGCTACTAATTTGAAATCGCTACATCGATTGCATATTAACGACGAAACGCAACTAAAAATGC
TAAAATTTCCACAAAACGATACTACTGACTACATGCCTCCTTCAGATGGAGCAAATGAGGTAACTCGGATT
GATTTGAAAGACATTAAATCACCTACGAGACACCATAAAAGAAGACCTACCACCATCGATGTTCTCGGTT
TAACAAAGTCTAAAACATCTCCAGATGGTCTCATATCAAAGGAAGATAGTGGATCAAAGTTAGTGATTGT
CATGGTCGGACTGCCAGCTACGGGAAAGTCATTTATTACAAATAAATTATCCAGATTTTAAATTATTTCT
TTATACTATTGTAAAGTGTTTAATGTCGGTAACACTAGAAGGAAGTTTGCTAAGGAGCATGGCCTAAAGG
ACCAGGATTCAAAGTTTTTCGAGCCGAAAAACGCCGACTCTACTAGGTTGAGAGACAAATGGGCCATGGA
TACTCTGGATGAATTGCTAGATTATTTATTAGAAGGTTCCAGGATCTGTGGGAATTTTGTATGCTACAAAT
ACCTCTCGTGAAGAAGAAAAACGTTCTGGCTAGAATCAGAAAGAGAAGTCCTCATTTGAAGGTTTTAT
TTTTAGAATCTGTTTGTTCGGATCATGCACTGGTACAGAAAAATATTAGACTCAAATTATTTGGTCCAGA
TTACAAAGGTAAAGATCCTGAAAGCTCTTTAAAGATTTTAAAAGTCGCCTGGCAAACCTACTTGAAAGCC
TATGAACCAATTGAGGATGACGAAAATTTGCAGTACATCAAATGATAGATGTGGGAAAGAAAGTCATCG
CATACAATATTCAAGGGTTTTTAGCTTCGCAGACGGTATATTATTTGTTAAATTTCAATTTGGCTGACAG
ACAAATTTGCATAACGAGAAGTGGCGAGAGCGAAGATAATGTTAGTGGCCGTATAGCGGAAATTTCCAT
TTGACTCCTCGTGGTCTAAGATTTGCTAAAAGTCTACCAAATTCATTGCCAGACAGAGAGAAATATTTT
ATCAAATCTCATGCAACAAAAAAGAATAATGAAAATACAGATGGGAACATTTATAATGACTTTTTTCGT
TTGGACCAGCATGCGTGCTAGGACTATAGGGACTGCTCAATATTTCAACGAAGATGATTATCCTATCAAA
CAAATGAAAATGTTAGATGAGTTAAGTGCAGGTGATTATGATGGTATGACATATCCAGAAATTA AAAACA
ACTTTCTGAGAATTCGAAAAAAGACAGAAAGATAAGTTGAGATACAGATACCCTGGTATTGGCGGTGA
ATCGTATATGGACCTTATTAATAGACTCAGACCTGTTATCAGACAAGTAGAAAGAAATCGAGGATAACCTT

CTTATTATTACACACCGGGTGGTGGCAAGAGCCTTATTGGGTATTATTTATGAACTTGAGTATGGGTATTA
 TTGCCAATTTGGATGTCCCATTACATTGTGTATATTGCCTAGAACCAAAACCATATGGAATCACTTGGTC
 ATTATGGGAGTATGATGAAGCATCGGATTCAATTTTCTAAGGTCCCACAAACGGACTTGAATACCACCAGA
 GTAAAGGAGGTTGGCCTTGTATTATAATGAAAGAAGATATTCTGTTATACCAACAGCTCCGCCAAGTGCAA
 GAAGCAGCTTTGCAAGTGACTTTTTGTCAAGAAAAAGATCTAATCCTACTTCTGCATCTTCATCCCAGAG
 TGAATTATCAGAACAACCCAAGAATAGCGTTAGTGCTCAAACCTGGCAGCAATAATACCACTCTCATTGGG
 AGCAACTTTAACATCAAGAATGAAAATGGTGATTGAGAATACCATTATCTGCACCACTTATGGCCACTA
 ATACTTCTAATAACATCTTAGATGGTGGAGGTACCTCAATTTGATACATCGTCCCAGGGTTGTTCCAAA
 TCAAAACAACGTGAATCCTCTTTTGGCTAACAACAATAAAGCGGCTTCTAATGTACCTAATGTAAAGAAG
 TCAGCGGCTACACCAAGGCAAATTTTTGAAATAGATAAAGTGGACGAAAAGTTATCCATGTTGAAAAATA
 AAAGTTTTCTATTACATGGAAAGGATTATCCTAATAATGCTGATAATAATGACAACGAAGATATAAGGGC
 AAAAACCATGAATCGCAGCCAAAGTCACGTTTAA

SEQ ID NO: 136 - *Saccharomyces cerevisiae* PFK26

MFKPVDFSETSPVPPDIDLAPTQSPHHVAPSQDSSYDLLSRS SDDKIDAEKGPHDELSKHLPLFQKRPLS
 DTPISSNWNSPGITEENTPSDSPENSATNLKSLHRLHINDETQLKNAKIPTNDTDDYMPSPDGANEVTRI
 DLKDIKSPTRHHKRRPTTIDVPLTKSKTSPDGLISKEDSGSKLVIVMVGLPATGKSFITNKLRSFLNYS
 LYYCKVFNVGNTRRKFAKEHGLKDQDSKFEPKNADSTRLRDKWAMDTLDEL DYLLEGSGSVGIFDATN
 TSRRERRKNVLRIRKRSPHLKVLFLSVCS DHALVQKNIRLKLFGPDYKGKDPESSLKDFKSRLANYLKA
 YEP IEDDENLQYIKMIDVGKKVIAYNIQGFLASQT VYYLLNFNLADRQIWITRSGESEDNVSGRIGGN SH
 LTPRGLRFAKSLPKFIARQREIFYQNLMMQKKNNENTDGNINYNDFVWTSMRARTIGTAQYFNEDDYP I K
 QMKMLDELSAGDYDGMTYPEIKNNFPEEF EK RQDKLRYRYPGIGGESYMDVINRLRPVITELERIEDNV
 LIITHRVVARALLGYFMNLSMGI IANLDVPLHCVYCLEPKPYGITWSLWEYDEASDSFSKVPQTDLNTTR
 VKEVGLVYNERRYSVIPTAPPSARSSFSASDFLSRKRSNPTSASSSQSELSEQPKNSVSAQTGSNNTTLIG
 SNFNIKNENGDSRIPLSAPLMATNTSNNILDGGGTSIS IHRPRVVPNQNNVNPLLANNNKAASNVPNVKK
 SAATPRQIFEIDKVDKLSMLKNKSFLHLHGKDY PNNADNNDNEDIRAKTMNRSQSHV*

SEQ ID NO: 137 - *Saccharomyces cerevisiae* PGI1 DNA

ATGTCCAATAACTCATTCACTAACTTCAAACCTGGCCACTGAATTGCCAGCCTGGTCTAAGTTGCAAAAAA
 TTTATGAATCTCAAGGTAAGACTTTGTCTGTCAAGCAAGAATTCCAAAAGATGCCAAGCGTTTTGAAA
 ATTGAACAAGACTTTCACTCACTATGATGGTTCCAAAATCTTGTTGCGACTACTCAAAGAAGTTGGTCAAC
 GATGAAATCATTGCTGCATTGATTGAACTGGCCAAGGAGGCTAACGTCACCGGTTTGAGAGATGCTATGT
 TCAAAGGTGAACACATCAACTCCACTGAAGATCGTGCTGTCTACCACGTCGCATTGAGAAACAGAGCTAA
 CAAGCCAATGTACGTTGATGGTGTCAACGTTGCTCCAGAAGTCGACTCTGTCTTGAAGCACATGAAGGAG
 TTCTCTGAACAAGTTCGTTCTGGTGAATGGAAGGGTTATACCGGTAAGAAGATCACCGATGTTGTTAACA
 TCGGTATTGGTGGTTCGATTTGGGTCCAGTCATGGTCACTGAGGCTTTGAAGCACTACGCTGGTGTCTT
 GGATGTCCACTTCGTTTCCAACATTGACGGTACTCACATTGCTGAAACCTTGAAGGTTGTTGACCAGAA
 ACTACTTTGTTTTTGATTGCTTCCAAGACTTTCACTACCCTGAAACTATCACTAACGCTAACACTGCCA
 AGAACTGGTTCCTGTGCGAAGACAGGTAATGATCCATCTCACATTGCTAAGCATTTCGCTGCTTTGTCCAC
 TAACGAAACCGAAGTTGCCAAGTTCCGTATTGACACCAAAAACATGTTTGGTTTCGAAAGTTGGGTCCGT
 GGTGCTTACTCTGTCTGGTTCGGCTATTGGTTTGTCTGTTGCCTTGACATTGGCTATGACAACCTTGAGG
 CTTTCTTGAAGGGTGCTGAAGCCGTCGACAACCACTTCACCCAAACCCCATTTGAAGACAACATTCCATT
 GTTGGGTGGTTTTGTTGTCTGTCTGGTACAACAACCTTCTTTGGTGCTCAAACCCATTTGGTTGCTCCATT
 GACCAATACTTGACAGATTCCCAGCCTACTTGCAACAATTGTCAATGGAATCTAACGGTAAGTCTGTTA
 CCAGAGGTAACGTGTTTACTGACTACTCTACTGGTTCTATCTTGTGTTGGTGAACCAGCTACCAACGCTCA

ACACTCTTTCTTCCAATTGGTTCACCAAGGTACCAAGTTGATTCCATCTGATTTTCATCTTAGCTGCTCAA
TCTCATAACCCAATTGAGAACAAATTACATCAAAAGATGTTGGCTTCAAACCTCTTTGCTCAAGCTGAAG
CTTTAATGGTTGGTAAGGATGAAGAACAAAGTTAAGGCTGAAGGTGCCACTGGTGGTTTGGTCCCACACAA
GGTCTTCTCAGGTAACAGACCAACTACCTCTATCTTGGCTCAAAAGATTACTCCAGCTACTTTGGGTGCT
TTGATTGCCTACTACGAACATGTTACTTTCACTGAAGGTGCCATTTGGAATATCAACTCTTTGACCAAT
GGGGTGTTGAATTGGGTAAAGTCTTGGCTAAAGTCATCGGCAAGGAATTGGACAACTCCTCCACCATTTC
TACCCACGATGCTTCTACCAACGGTTTAATCAATCAATTCAAGGAATGGATGTGA

SEQ ID NO: 138 - *Saccharomyces cerevisiae* PGI1

MSNNSFTNFKLATELPAWSKLQKIYESQGKTL SVKQEFQKDAKRFEKLNKFTNYDGSKILFDYSKNLVN

DEITAAALIELAKEANVTGLRDAMFKGEHINSTEDRAVYHVALRNRANKPMYVDGVNVAPEVDSVLKHMKE
FSEQVRSGEWKGYTGKKITDVVNIGIGGSDLGPVMVTEALKHYAGVLDVHFVSNIDGTHIAETLKVVDPE
TTLFLIASKTFTTAETITNANTAKNWFLSKTGNDPSHIAKHFAALSTNETEVAKFGIDTKNMFGEFSWVG
GRYSVWSAIGLSVALYIGYDNFEAFLKGAEAVDNHFTQTPLEDNIPLLGLLSVWYNNFFGAQTHLVAF
DQYLHRFPAYLQQLSMESNGKSVTRGNVFTDYSTGSILFGEPATNAQHSFFQLVHQGTKLIPSDFILAAQ
SHNPIENKLHQKMLASNFFAQAEALMVGKDEEQVKAEGATGGLVPHKVFSGNRPTTSILAQKITPATLGA
LIAYYEHVFTFEGAIWNINSFDQWGVELGKVLAKVIGKELDNSTISTHDASTNGLINQFKEM*

SEQ ID NO: 139 - *Saccharomyces cerevisiae* GPM1 DNA

ATGCCAAAGTTAGTTTTAGTTAGACACGGTCAATCCGAATGGAACGAAAAGAAGTTATTACCCGGTTGGG
TTGATGTTAAATTGTCTGCCAAGGGTCAACAAGAAGCCGCTAGAGCCGGTGAATTGTTGAAGGAAAAGAA
GGTCTACCCAGACGTCTTGACACTTCCAAGTTGTCCAGAGCTATCCAAACTGCTAACATTGCTTTGGAA
AAGGCTGACAGATTATGGATTCCAGTCAACAGATCCTGGAGATTGAACGAAAGACATTACGGTGACTTAC
AAGGTAAAGACAAGGCTGAAACTTTGAAGAAGTTCGGTGAAGAAAAATTCAACACCTACAGAAGATCCTT
CGATGTTCCACCTCCCCCAATCGACGCTTCTTCTCCATTCTCTCAAAAGGGTGATGAAAGATACAAGTAC
GTTGACCCAAATGCTTGCCAGAACTGAATCTTTGGCTTTGGTCATTGACAGATTGTTGCCATACTGGC
AAGATGTCATTGCCAAGGACTTGTTGAGTGGTAAGACCGTCATGATCGCCGCTCACGGTAACTCCTTGAG
AGGTTTGGTTAAGCACTTGGAAGGTATCTCTGATGCTGACATTGCTAAGTTGAACATCCCAACTGGTATT
CCATTGGTCTTCGAATTGGACGAAAAGTTGAAGCCATCTAAGCCATCTTACTACTTGGACCCAGAAGCTG
CCGCTGCTGGTGCCGCTGCTGTTGCCAACCAAGGTAAGAAATAA

SEQ ID NO: 140 - *Saccharomyces cerevisiae* GPM1

MPKLLVLRHGQSEWNEKNLFTGWVDVKLAKGQQEFAARAGELLKEKKVYPDVLVTSKLSRAIQTANIALE
KADRLWIPVNRSWRLNERHYGDLQGKDKAETLKKFGEEKFNTRYRSFDPVPPPIDASSPFSQKGDERYKY
VDPNVLPETESLALVIDRLLPYWQDVIKDLLSGKTVMIAAHGNSLRGLVKHLEGISDADIAKLNIPTGI
PLVFELDENLKPSKPSYYLDPEAAAAAGAAAVANQGKK

SEQ ID NO: 141 - *Saccharomyces cerevisiae* TPI1 DNA

[illegible]

TCGCGTAACGGTAGCAACGCCGTTACCTTCAAGGACAAGGCTGATGTGATGGTTCTTTGGTCGGTGGTG
CTTCTTTGAAGCCAGAATTTGTTGATATCATCAACTCTAGAACTAA

SEQ ID NO: 142 - *Saccharomyces cerevisiae* TPI1

MARTFFVGGNFKLNGSKQSIKEIVERLNTASIPENVEVVICPPATYLDYSVSLVKKPQVTVGAQNAYLKA
SGAFTGENSV DQIKDVGA K WVILGHSERRSYFHEDDKFIADKTKFALGQGVGVILCIGETLEEKKAGKTL
DVVERQLNAVLEEVKDWNTNVVVA YEPVWAIGTGLAATPEDAQDIHASIRKFLASKLGDKAASELRILYGG
SANGSNAVT FKDKADVDGFLVGGASLKPEFVDI INSRN

SEQ ID NO: 147 - *Saccharomyces cerevisiae* PGM1 DNA

ATGTCACTTCTAATAGATTCTGTACCAACAGTTGCTTATAAGGACCAAAAACCGGGTACTTCAGGTTTAC
GTAAGAAGACCAAGGTTTTTCATGGATGAGCCTCATTATACTGAGAACTTCATTCAAGCAACAATGCAATC
TATCCCTAATGGCTCAGAGGGAACCACTTTAGTTGTTGGAGGAGATGGTCGTTTCTACAACGATGTTATC
ATGAACAAGATTGCCGCAGTAGGTGCTGCAAACGGTGTGAGAAAGTTAGTCATTGGTCAAGGCGGTTTAC
TTTCAACACCAGCTGCTTCTCATATAATTAGAACATACGAGGAAAAGTGACCGGTGGTGGTATCATATT
AACTGCCTCACACAACCCAGGCGGTCCAGAGAATGATTTAGGTATCAAGTATAATTTACCTAATGGTGGG
CCAGCTCCAGAGAGTGTCACTAACGCTATCTGGGAAGCGTCTAAAAAATTAACCTCACTATAAAATTATAA
AGAACTTCCCCAAGTTGAATTTGAACAAGCTTGGTAAAAACCAAAAATATGGCCCATTTGTTAGTGGACAT
AATTGATCCTGCCAAAGCATA CGTTCAATTTCTGAAGGAAATTTTGTATTTGACTTAATTAAGGCTTC
TTAGCGAAACAGCGCAAAGACAAAGGGTGGAAGTTGTTGTTGACTCCTTAAATGGTATTACAGGACCAT
ATGGTAAGGCTATATTTGTTGATGAATTTGGTTTACCGGCAGAGGAAGTTCTTCAAAATTTGGCACCCTTT
ACCTGATTTTCGGCGGTTTACATCCCGATCCGAATCTAACCTATGCACGAACCTCTTGTGACAGGGTTGAC
CGCGAAAAAATTGCCTTTGGAGCAGCCTCCGATGGTGATGGTGATAGGAATATGATTTACGGTTATGGCC
CTGCTTTTCGTTTCGCCAGGTGATTCTGTTGCCATTATTGCCGAATATGCACCCGAAATTCATACCTTCGC
CAACAAGGTATTATGGCTTGGCACGTTTCAATTCCTACATCCTCAGCCATTGATCGTGTGACAGCAAAA
AAGGGATTAAGATGTTACGAAGTTCCAACCGGCTGGAAATTTCTTCTGTGCCTTATTTGATGCTAAAAAGC
TATCAATCTGTGGTGAAGAATCCTTCGGTACAGGTTCCAATCATATCAGAGAAAAGGACGGTCTATGGGC
CATTATTGCTTGGTTAAATATCTTGGCTATCTACCATAGGCGTAACCCTGAAAAGGAAGCTTCGATCAAA
ACTATTCAGGACGAATTTTGAACGAGTATGGCCGTACTTTCTTCACAAGATACGATTACGAACATATCG
AATGCGAGCAGGCCGAAAAAGTTGTAGCTCTTTTGAAGTGAATTTGTATCAAGGCCAAACGTTTGTGGCTC
CCACTTCCCAGCTGATGAGTCTTTAACCGTTATCGATTGTGGTGATTTTTCGTATAGAGATCTAGATGGC
TCCATCTCTGAAAATCAAGGCCTTTTCGTAAAGTTTTCGAATGGGACTAAATTTGTTTTGAGGTTATCCG
GCACAGGCAGTTCTGGTGCAACAATAAGATTATACGTAGAAAAGTATACTGATAAAAAGGAGAACTATCG
CCAAACAGCTGACGTCTTCTTGAACCCGTCATCAACTCCATTGTAAAATTTCTTAAGATTTAAAGAAATT
TTAGGAACAGACGAACCAACAGTCCGCACATAG

SEQ ID NO: 148 - *Saccharomyces cerevisiae* PGM1

MSLLIDSVPTVAYKDQKPGTSGLRKKT KVFMDEPHYTENFIQATMQSIPNGSEGTTLVVGGDGRFYNDVI
MNKIAAVGAANGVRKLVIGQGGLLSTPAASHIIRT YEEKCTGGGIILTASHNPGGPENDLGIKYNLPNGG
PAPESVTNAIWEASKKLTHYKIIKNFPKLNLNKLGNQKYGPLLVDIIDPAKAYVQFLKEIFDFDLIKSF
LAKQRKDKGWKLLFDSLNGITGPYKAI FVDEFGLPAEEVLQNWHLPLPDFGGLHPDENLTYARTLVDRVD
REKIAFGAASDGDGRNMIYGYGPAFVSPGDSVAIIAEYAPEIPYFAKQGIYGLARSFPTSSAIDRVAAK
KGLRCYEVPTGWKFFCALFDAKKLSICGEESFGTGSNHIREKDGLWAI IAWLNILAIYHRRNPEKEASIK
TIQDEFWNEYGRFFTRYDYEHEICEQA EKVVALLSEFVSRPNVCGSHFPADESLTVIDCGDFS YRDL DG
SISENQGLFVKFSNGTKFVLRLSGTGSSGATIRLYVEKYTDKKENYGQTADVFLKPVINSIVKFLRFKEI
LGTDEPTVRT

SEQ ID NO: 149 - *Saccharomyces cerevisiae* PGM3 DNA

ATGTTGCAAGGAATTTTAGAAACCGTACCATCTGACTTGAAAGATCCGATATCATTATGGTTTAAGCAAG
 ACCGCAACCCAAAACTATAGAAGAGGTCACCGCTCTCTGCAAAAAATCCGACTGGAATGAGTTACACAA
 AAGATTTGATTCTAGAATTCAGTTTGGCACTGCTGGTTTAAGATCGCAAATGCAAGCTGGCTTTAGCAGG
 ATGAATACTTTAGTAGTCATACAAGCGTCTCAGGGATTGGCAACTTATGTAAGACAACAGTTTCCAGACA
 ATTTGGTAGCTGTTGTGGGACACGATCATAGATTCCATTCTAAGGAGTTCGCTAGAGCTACTGCTGCTGC
 ATTTCTTTTAAAAGGATTTAAGGTACATTATTTGAATCCTGACCACGAATTTGTTTCATACCCCTTTAGTT
 CCCTTTGCACTGGATAAGCTAAAGGCTCCGTTGGCGTAATGATAACAGCAAGTCACAACCCAAAAATGG
 ATAATGGATATAAAGTATACTATTCCAATGGATGCCAAATCATTCCACCTCAGCATCATGCCATCTCTGA
 TTCCATTGACGCAAATTTAGAACCATGGGCCAATGTGTGGGATTTGACGATGTTCTAAATAAGGCTCTC
 AAACAAGGGAAATTGATGTATTCCAGAGAAGAAATGCTGAAGTTATATTTAGAGGAGGTTTCTAAAAATC
 TGGTAGAAATCAACCCATTAAAGCTTGAAGTAAAAGCCAAACCTTGGTTCGTTTACACTCCAATGCATGG
 GGTGGATTGACATTTTCAGCACCATCGTAAAAAAAACACTGTGCCTGGTAGAAGGTAAGGATTACCTA
 TGTGTTCCCTGAACAACAAAATCCAGATCCTTCTTTCCCAACTGTTGGATTTCCTAACCCCTGAAGAAAAAG
 GTGCTTTAGACATTGGTATAAACTTGGCTGAAAAACATGACATTGACTTACTTGTGTTGCCAACGACCTGA
 CGCTGATAGATTCTCTGTTGCTGTTAAAGATATGCAGTCAGGCGAATGGCGACAACCTAACAGGTAACGAA
 ATCGGTTTTCTTTTGCATTTTATGAATATCAGAAATATAAAAGTATGGACAAAGAATTTTCAGCACGTTT
 ATCCGTTGGCTATGTTAAATTCACAGTGTCTTCACAAATGATAAAAAAATGGCAGAAATAGAAGGGTT
 CCATTATGAGGATACATTAACAGGATTTAAGTGGATCGGAAATCGTGCCATACTCTTGGAAAAGAAAGGC
 TATTACGTTCCCTTTTGGATTGAGGAAGCAATAGGCTACATGTTTCCAGCAATGGAGCATGATAAGGATG
 GTATCAGTGCATCCATTGTCTTCTTGCAGGCTACTGTAAGTGGAAAATAGACCACAATTTGGACCCGCT
 AAATGTCTTAGAAAAATGGCTTCAAAAAATATGGCGTGTTCAAAGAGTACAATGGCTATTATGTCGTTCCA
 AATCCAACCTGTTACAAAAGATATATTTGACTACATCAGGAATGTCTACACTCCTGAGGGCGCGTCATATC
 CTTTCATCTATTGGTGAAGAAATCGAAGTACTTTACTATCGAGATTTAACCCTGGTTACCAATCGGATAC
 CATAAATCATAAACCTACTCTACCCGTCGATCCTACATCACAATGATAACAGTATCTGCTAGACCAAGT
 AACGGTAGTGAGAATGAGCATATCCGCTTCACTATTCGCGGGTCCGGAACAGAACCAAACTTAAAGTAT
 ATATTGAAGCTTGCGCAAATGAAGAACAAAGAGCCTCTTTCTTGGCGAAATTGACTTGGAATGTGCTGAG
 ACGTGAATGGTTTAGACCAGATGAAATGAATATAGTTACAAAATTTTGA

SEQ ID NO: 150 - *Saccharomyces cerevisiae* PGM3

MLQGILETVPSDLKDPISLWFKQDRNPKTIEEVTALCKSDWNELHKRFDRIQFGTAGLRSQMAGFSR
 MNTLVVIQASQGLATYVRQQFPDNLVAVVGHDRHFSKEFARATAAFLKGFVHYLNPDEHFVHTPLV
 PFAVDKLKASVGVMITASHNPKMDNGYKVYYSNQCIIIPPHDHAISDSIDANLEPWANVWDFDDVLNKL
 KQGKLMYSREMLKLYLEEVSKNLVEINPLKLEVKAKPWVYTPMHGVGFDIFSTIVKKTLCLEVGKDYL
 CVPEQQNPDPSPFTVGFNPPEEKALDIGINLAEKHDIDLIVANDPDADRFSAVKMDQSGEWRQLTGNE
 IGFLFAFYEQKYKSMDKEFQHVHPLAMLNSTVSSQMIKKMAEIEGFHYEDTLTGFKWIGNRAILLEKKG
 YYVPFGFEEAIGYMFAMEHDKDGISASIVFLQAYCKWKIDHNLDPNLVLENGFKKYGVFKEYNGYYVVP
 NPTVTKDIFDYIRNVYTPEGASYPSSIGEEIEVLYYRDLTTGYQSDTINHKTLPVDPTSQMITVSARPS
 NGSENEHIRFTIRSGTEPKLVYIEACANEEQRASFLAKLTWNVLRREWFRPDEMNIIVTKF

SEQUENCE LISTING

<110> Novozymes A/S

Karhumaa, Kaisa

Sendelius, Malin

Sanchez i Nogue, Violeta

<120> Recombinant Yeast Strains for Pentose Fermentation

<130> 14257-WO-PCT

<150> US 62/436,648

<151> 2016-12-20

<150> US 62/436,687

<151> 2016-12-20

<150> US 62/436,706

<151> 2016-12-20

<150> US 62/436,723

<151> 2016-12-20

<160> 150

<170> PatentIn version 3.5

<210> 1

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 1

ggtagcgagc tctaactgat ctatcaaaa ctg 33

<210> 2

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 2

ggtagcgatc agcatgcat cgctcgacat ttgatatac 39

<210> 3

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 3

ccgcgggagc tctaactgat ctatccaaaa ctg 33

<210> 4

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 4

ccgcgggatc agcatgcat cgctcgacat ttgatatac 39

<210> 5

<211> 2582

<212> DNA

<213> *Spathaspora passalidarum*

<400> 5

tatatctagg aacccatcag gttggtggaa gattaccogt tctaagactt ttcagcttcc	60
tctattgatg ttacacctgg acaccccttt tctggcatcc agtttttaat cttcagtggc	120
atgtgagatt ctccgaaatt aattaaagca atcacacaat tctctcggat accacotcgg	180
ttgaaactga caggtggttt gttacgcatg ctaatgcaaa ggagcctata tacctttggc	240
tcggctgctg taacagggaa tataaagggc agcataattt aggagttag tgaacttgca	300
acatttacta ttttcccttc ttacgtaaatt atttttcttt ttaattctaa atcaatcttt	360
ttcaattttt tgtttgtatt cttttcttgc ttaaattctat aactacaaaa aacacataca	420
taaaactaaa atgacagtag aactaccgcg ttcagaacct ttgtttcttg ggtttgatct	480
tagcactcaa cagttgaaaa tcatagtgac aaaccagaaa ttagctgcac taaaatctta	540
caacggtgaa ttcgatgtgg catttaagga gaaatatggg ataaccaagg gagtccctaac	600
gaacaaagag gacggagaag tggtttctcc agttggtatg tggtttagatt ccataaacca	660
tgtattcgac caaatgaaac aagatgactt tccgttcaat caagttgcag gcatttcagg	720
ctcttgtaaa caacatgggt ctgtgttttg gtcacatgaa gctgagaagc ttttatcagg	780
tttacagaag gatcaagatc tgtcgactca actaaaggac gctttatctt gggacaaaag	840
tcccaattgg caagatcatt cgactttaga ggaaagtaag gctttcgtag atgctgtagg	900
gaggggaagag ttagccgata ttactggtag tagagatcac ttaagattca ctggattgca	960
aattaggaag tttgccacta gatcacatcc cgataagtat gcgaatacta gtagaatctc	1020
actgggttagc tccttcataa caagtgttct tctgggtgag attaccgaat tggaagaatc	1080
tgatgcttgt ggcatagaact tgtatgacat caaagccggt gatttcaatg aagaattggt	1140
ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct ctctctctct	1200

```

ggctctagca gccggtgttc atcctaaagt tgacaacata acgaaagatg atccgaaata 1200
taaagccgga attgaggaca tcaaagcgaa acttggaag atctcccaa ttacatataa 1260
aagctccgga tccattgctt catattacgt tgaaaagtac ggtttgaatc ctaagtgcc 1320
gatttacagc ttaccgggtg acaatttggc tacaatcttg agtttgccat tacagcctaa 1380
cgattgcttg atttcgttag gtacttcgac tacggctcttg ctaatcacta agaattacca 1440
accttcttct caatatcact tgtttaagca tccaaccata ccagatggat atatgggcat 1500
gatctgctat tgcaatggct ctttggccag agagaagata agagatgaag ttaatgaata 1560
ctataaggtg gaagataaga agagttggga taaatttagt gaaattcttg ataagtcgac 1620
caaattcgat aataagctgg gtattttctt tccgttaggt gaaatcgtc cacaggcaaa 1680
ggcacaaact gtcagagcag tattggagaa tgacaaagtc atagaagtag gtttggatac 1740
acacggattt gatattgatc acgacgcaag agctattgtc gaaagccaag ccttatcttg 1800
tagacttaga gctggcccta tgttatccaa atcatcacgt gcttccgtca catcaccaac 1860
ggagttaaaa ggcgtatacc atgacatagt ggccaaatat ggtgacctgt acacagatgg 1920
taaacaacaa acctatgagt cacttacatc taggccaaat cgttgtttct atgttggcgg 1980
cgggagcaat aacatttcca tcattagtaa aatgggttct attctaggtc ctgttcacgg 2040
taactttaaa gtcgatattc caaacgcgtg ttctctaggt ggagcataca aagcatcctg 2100
gtcttatgaa tgtgaacaga aaggtgaatg gattaattat gaccaatata taaatcagtt 2160
attgaaagaa ttgaagtcac ttaacgtgga ggacaaatgg ttagaatact ttgatggagt 2220
tgcgcttttg gctaagatgg aagaaaccct gttgaaataa agatctcca tgtctctact 2280
ggtggtggtg cttcttttga attattggaa ggtaaggaat tgccaggtgt tgctttctta 2340
tccgaaaaga aataaattga attgaattga aatcgataga tcaatttttt tcttttctct 2400
ttcccatcc tttacgctaa aataatagtt tattttattt tttgaatatt ttttatttat 2460
atcgtatat atagactatt atttatcttt taatgattat taagattttt attaaaaaaa 2520
aattcgctcc tcttttaatg cctttatgca gttttttttt cccattcgat atttctatgt 2580
tc 2582

```

<210> 6

<211> 609

<212> PRT

<213> Spathaspora passalidarum

<400> 6

```

Met Thr Val Glu Leu Pro Ala Ser Glu Pro Leu Phe Leu Gly Phe Asp
1          5          10          15

Leu Ser Thr Gln Gln Leu Lys Ile Ile Val Thr Asn Gln Lys Leu Ala
20          25          30

Ala Leu Lys Ser Tyr Asn Val Glu Phe Asp Val Ala Phe Lys Glu Lys
35          40          45

Tyr Gly Ile Thr Lys Gly Val Leu Thr Asn Lys Glu Asp Gly Glu Val
50          55          60

Val Ser Pro Val Glu Met Trp Leu Asn Ser Ile Asn His Val Phe Asn

```

Val Ser Leu Val Gly Asp Asp Phe Pro Phe Asn Gln Val Ala Gly Ile Ser
 65 70 75 80
 Gln Met Lys Gln Asp Asp Phe Pro Phe Asn Gln Val Ala Gly Ile Ser
 85 90 95
 Gly Ser Cys Gln Gln His Gly Ser Val Phe Trp Ser His Glu Ala Glu
 100 105 110
 Lys Leu Leu Ser Gly Leu Gln Lys Asp Gln Asp Leu Ser Thr Gln Leu
 115 120 125
 Lys Asp Ala Leu Ser Trp Asp Lys Ser Pro Asn Trp Gln Asp His Ser
 130 135 140
 Thr Leu Glu Glu Ser Lys Ala Phe Val Asp Ala Val Gly Arg Glu Glu
 145 150 155 160
 Leu Ala Asp Ile Thr Gly Ser Arg Asp His Leu Arg Phe Thr Gly Leu
 165 170 175
 Gln Ile Arg Lys Phe Ala Thr Arg Ser His Pro Asp Lys Tyr Ala Asn
 180 185 190
 Thr Ser Arg Ile Ser Leu Val Ser Ser Phe Ile Thr Ser Val Leu Leu
 195 200 205
 Gly Glu Ile Thr Glu Leu Glu Glu Ser Asp Ala Cys Gly Met Asn Leu
 210 215 220
 Tyr Asp Ile Lys Ala Gly Asp Phe Asn Glu Glu Leu Leu Ala Leu Ala
 225 230 235 240
 Ala Gly Val His Pro Lys Val Asp Asn Ile Thr Lys Asp Asp Pro Lys
 245 250 255
 Tyr Lys Ala Gly Ile Glu Asp Ile Lys Ala Lys Leu Gly Lys Ile Ser
 260 265 270
 Pro Ile Thr Tyr Lys Ser Ser Gly Ser Ile Ala Ser Tyr Tyr Val Glu
 275 280 285
 Lys Tyr Gly Leu Asn Pro Lys Cys Gln Ile Tyr Ser Phe Thr Gly Asp
 290 295 300
 Asn Leu Ala Thr Ile Leu Ser Leu Pro Leu Gln Pro Asn Asp Cys Leu
 305 310 315 320
 Ile Ser Leu Gly Thr Ser Thr Thr Val Leu Leu Ile Thr Lys Asn Tyr
 325 330 335
 Gln Pro Ser Ser Gln Tyr His Leu Phe Lys His Pro Thr Ile Pro Asp
 340 345 350
 Gly Tyr Met Gly Met Ile Cys Tyr Cys Asn Gly Ser Leu Ala Arg Glu
 355 360 365
 Lys Ile Arg Asp Glu Val Asn Glu Tyr Tyr Lys Val Glu Asp Lys Lys
 370 375 380

Ser Trp Asp Lys Phe Ser Glu Ile Leu Asp Lys Ser Thr Lys Phe Asp
385 390 395 400

Asn Lys Leu Gly Ile Phe Phe Pro Leu Gly Glu Ile Val Pro Gln Ala
405 410 415

Lys Ala Gln Thr Val Arg Ala Val Leu Glu Asn Asp Lys Val Ile Glu
420 425 430

Val Gly Leu Asp Thr His Gly Phe Asp Ile Asp His Asp Ala Arg Ala
435 440 445

Ile Val Glu Ser Gln Ala Leu Ser Cys Arg Leu Arg Ala Gly Pro Met
450 455 460

Leu Ser Lys Ser Ser Arg Ala Ser Val Thr Ser Pro Thr Glu Leu Lys
465 470 475 480

Gly Val Tyr His Asp Ile Val Ala Lys Tyr Gly Asp Leu Tyr Thr Asp
485 490 495

Gly Lys Gln Gln Thr Tyr Glu Ser Leu Thr Ser Arg Pro Asn Arg Cys
500 505 510

Phe Tyr Val Gly Gly Gly Ser Asn Asn Ile Ser Ile Ile Ser Lys Met
515 520 525

Gly Ser Ile Leu Gly Pro Val His Gly Asn Phe Lys Val Asp Ile Pro
530 535 540

Asn Ala Cys Ser Leu Gly Gly Ala Tyr Lys Ala Ser Trp Ser Tyr Glu
545 550 555 560

Cys Glu Gln Lys Gly Glu Trp Ile Asn Tyr Asp Gln Tyr Ile Asn Gln
565 570 575

Leu Leu Lys Glu Leu Lys Ser Phe Asn Val Glu Asp Lys Trp Leu Glu
580 585 590

Tyr Phe Asp Gly Val Ala Leu Leu Ala Lys Met Glu Glu Thr Leu Leu
595 600 605

Lys

<210> 7

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 7

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 8

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 8

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 9

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 9

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 10

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 10

ctgtctctga attactgaac acaacatttt tagtttatgt atgtgttt 48

<210> 11

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 11

gcgaactgga aaagactctc atctaaagat ctcccatgtc tctactgg 48

<210> 12

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 12

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 13

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 13

aaacacatac ataaactaaa aatgttgtgt tcagtaattc agagacag 48

<210> 14

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 14

ccagtagaga catgggagat cttagatga gagtcttttc cagttcgc 48

<210> 15

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 15

tctccacac ctgcagtata tctaggaacc catcag 36

<210> 16

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 16

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 17

<211> 2207

<212> DNA

<213> Escherichia coli

<400> 17

tatatctagg aacccatcag gttggtggaa gattaccogt tctaagactt ttcagcttcc	60
tctattgatg ttacacctgg acaccccttt tctggcatcc agtttttaat cttcagtggc	120
atgtgagatt ctccgaaatt aattaaagca atcacacaat tctctcggat accacctcgg	180
ttgaaactga caggtggttt gttacgcatg ctaatgcaaa ggagcctata tacctttggc	240
tcggctgctg taacagggaa tataaagggc agcataattt aggagttag tgaacttgca	300
acatttacta ttttccttcc ttacgtaaat atttttcttt ttaattctaa atcaatcttt	360
ttcaattttt tgtttgtatt cttttcttgc ttaaattctat aactacaaaa aacacataca	420
taaactaaaa atgtatatcg gcattgattt gggtaacttct ggcgtaaagg ttatcctgct	480
gaatgaacag ggtgaagtgg ttgcctcaca aacggaaaag ttgactgtat ctaggccaca	540
tcctttgtgg agcgaacaag atccagaaca gtggtggcaa gctacagata gagcaatgaa	600
agcgttaggt gaccagcatt ccttacagga cgttaaagcc ttggggattg ctggccaaat	660
gcatggtgcg aactgcttg atgccaaca aagggtctta aggcctgcaa tactgtggaa	720
tgatggacgt tgtgctcagg agtgtacctt attggaagca agagtgcctc aatccagggt	780
gataaccggt aacttgatga tgcttgatt tacagcccca aaattgttat gggttcaaag	840
acacgaacca gagatcttcc gtcaaataca caagggtctta ttaccgaagg actacttgag	900
actacgtatg actggtgaat tcgcttcaga catgagtgaac gcagcaggaa ccatgtggtt	960
ggatgtcgcg aaaagagatt ggagtgaagt tatgttaca gcttgcgac tatctagaga	1020
tcaaatacca gctctgtatg agggctcaga aattaccggt gcattattac ctgaagtcgc	1080
taaagcatgg ggtatggcta ctgtccaggt tgttgccggt ggtggtgaca atgccgcagg	1140
agctgttgga gttggtatgg tggatgcaaa tcaagcgatg ttgtctcttg gcacatcagg	1200
cgtctatttt gccgtatcgg aagggtttct gtcgaaacca gaatcagccg tacattcctt	1260
ttgtcacgct cttccacaaa gatggcatct aatgagcgtg atgctttctg cagcatcatg	1320
cttggtattg gccgctaaat tgacgggttt gagtaatgtt ccggcactta tagcagctgc	1380
acaacaagca gatgaaagtg ctgaacccgt ttggttcttg ccctatcttt ccggagagag	1440
aacaccacac aacaatcttc aagccaaagg tgtgttcttt gggtaactc accaacaatg	1500
tccaaacgaa ttggcgagag cagtattgga aggagtaggg tatgctcttg ctgatggtat	1560
ggatgttgct catgcatgtg gcataaagcc gcaatctggt acgcttattg gagtggtgc	1620
cagaagcgaa tactggagac aaatgttagc cgatatttcc ggtcaacaac tagactacag	1680
aacaggaggc gatgtagggc cagctttggg tgctgctaga ttggctcaga ttgctgctaa	1740
ccctgagaag tcgttgattg agctactacc tcagttaccc ttagaacagt ctcatctacc	1800
agatgcccag agatatgtcg cgtaccaacc tagaagagag acttttcgta ggttatacca	1860
gcaattacta cccttgatgg cgtaaagatc tcccatgtct ctactggtgg tgggtgcttct	1920
ttggaattat tgggaagtaa ggaattgcc aagtgtgctt tcttatccga aaagaaataa	1980
attgaattga attgaaatcg atagatcaat ttttttcttt tctctttccc catcctttac	2040
gctaaaaata tagtttattt tattttttga atatttttta tttatatagc tatatataga	2100

ctattatttta tcttttaaatg attattaaga tttttatttaa aaaaaaatc gctcctcttt 2160
 taatgccttt atgcagtttt tttttcccat tcgatatttc tatgttc 2207

<210> 18

<211> 484

<212> PRT

<213> Escherichia coli

<400> 18

Met Tyr Ile Gly Ile Asp Leu Gly Thr Ser Gly Val Lys Val Ile Leu
 1 5 10 15

Leu Asn Glu Gln Gly Glu Val Val Ala Ser Gln Thr Glu Lys Leu Thr
 20 25 30

Val Ser Arg Pro His Pro Leu Trp Ser Glu Gln Asp Pro Glu Gln Trp
 35 40 45

Trp Gln Ala Thr Asp Arg Ala Met Lys Ala Leu Gly Asp Gln His Ser
 50 55 60

Leu Gln Asp Val Lys Ala Leu Gly Ile Ala Gly Gln Met His Gly Ala
 65 70 75 80

Thr Leu Leu Asp Ala Gln Gln Arg Val Leu Arg Pro Ala Ile Leu Trp
 85 90 95

Asn Asp Gly Arg Cys Ala Gln Glu Cys Thr Leu Leu Glu Ala Arg Val
 100 105 110

Pro Gln Ser Arg Val Ile Thr Gly Asn Leu Met Met Pro Gly Phe Thr
 115 120 125

Ala Pro Lys Leu Leu Trp Val Gln Arg His Glu Pro Glu Ile Phe Arg
 130 135 140

Gln Ile Asp Lys Val Leu Leu Pro Lys Asp Tyr Leu Arg Leu Arg Met
 145 150 155 160

Thr Gly Glu Phe Ala Ser Asp Met Ser Asp Ala Ala Gly Thr Met Trp
 165 170 175

Leu Asp Val Ala Lys Arg Asp Trp Ser Asp Val Met Leu Gln Ala Cys
 180 185 190

Asp Leu Ser Arg Asp Gln Met Pro Ala Leu Tyr Glu Gly Ser Glu Ile
 195 200 205

Thr Gly Ala Leu Leu Pro Glu Val Ala Lys Ala Trp Gly Met Ala Thr
 210 215 220

Val Pro Val Val Ala Gly Gly Gly Asp Asn Ala Ala Gly Ala Val Gly
 225 230 235 240

Val Gly Met Val Asp Ala Asn Gln Ala Met Leu Ser Leu Gly Thr Ser
 245 250 255

Gly Val Tyr Phe Ala Val Ser Glu Gly Phe Leu Ser Lys Pro Glu Ser
260 265 270

Ala Val His Ser Phe Cys His Ala Leu Pro Gln Arg Trp His Leu Met
275 280 285

Ser Val Met Leu Ser Ala Ala Ser Cys Leu Asp Trp Ala Ala Lys Leu
290 295 300

Thr Gly Leu Ser Asn Val Pro Ala Leu Ile Ala Ala Ala Gln Gln Ala
305 310 315 320

Asp Glu Ser Ala Glu Pro Val Trp Phe Leu Pro Tyr Leu Ser Gly Glu
325 330 335

Arg Thr Pro His Asn Asn Pro Gln Ala Lys Gly Val Phe Phe Gly Leu
340 345 350

Thr His Gln His Gly Pro Asn Glu Leu Ala Arg Ala Val Leu Glu Gly
355 360 365

Val Gly Tyr Ala Leu Ala Asp Gly Met Asp Val Val His Ala Cys Gly
370 375 380

Ile Lys Pro Gln Ser Val Thr Leu Ile Gly Gly Gly Ala Arg Ser Glu
385 390 395 400

Tyr Trp Arg Gln Met Leu Ala Asp Ile Ser Gly Gln Gln Leu Asp Tyr
405 410 415

Arg Thr Gly Gly Asp Val Gly Pro Ala Leu Gly Ala Ala Arg Leu Ala
420 425 430

Gln Ile Ala Ala Asn Pro Glu Lys Ser Leu Ile Glu Leu Leu Pro Gln
435 440 445

Leu Pro Leu Glu Gln Ser His Leu Pro Asp Ala Gln Arg Tyr Ala Ala
450 455 460

Tyr Gln Pro Arg Arg Glu Thr Phe Arg Arg Leu Tyr Gln Gln Leu Leu
465 470 475 480

Pro Leu Met Ala

<210> 19

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 19

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 20

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 20

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 21

<211> 1806

<212> DNA

<213> Kluyveromyces marxianus

<400> 21

atgtcgaactc catattactt aggctttgat ttgtcaactc agcagttgaa atgtctagca	60
atagatgatac agttaaacat cgtgacttca gttagcattg aatttgaccg taatttccca	120
gcttacaca caaagaaagg ggtatacatc aagaatgggt gtgtgataga cgcaccagtt	180
gctatgtggt tagaagctgt tgattttatgc tttagtcagc ttgctgaaag aatcgacttg	240
aagagagttc aatcaatgtc aggtttcttc caacaacatg gcaccgtcta ctggaactgt	300
gagcatctac caagtaacct tgatcctgct tcaacctga gagagcaact tcaaggcagt	360
ttatcaagac cagttgcacc caattggcaa gatcattcca ccaagaaaca atgtgatgaa	420
ttggcagaat cggtaggagg acctgaagaa ctagcaagga ttacaggttc tggagcacac	480
tatagatttt ccggttccca aattgccaaa atccatgaaa ctgaacctga agtctatgag	540
gctactaaac gtatttcgtt ggtaagtagc tttctagcgt ctgttttagt aggtgacatt	600
gtgcccttgg aagaagcggg tgcttgtggc atgaacttat acgatctatc caaacacgac	660
tttgacgaaa cattactggc tgtcgttgac catgatacag cgagattgag aaggaaacta	720
tcagatccac ccgttggagc tcctaccaga gaaagtcctc tgacctcctt gggtaaagtc	780
tctaagtact ttcaggacaa atatggggtt aactgtgaat gtgagatctt cccgtttact	840
ggcgataacc tggcaacgat ttgttcctca cttttgcaaa agaagatgtt cttgattagt	900
ctaggtactt cgaccacgat tttgttggtg actgaccaat atcactcttc tcccaattat	960
cacttgttta tacatccgac agtgccaggt tattacatgg gtatgatttg ctattgcaat	1020
gggtccttgg ctcgtagag agtaagagat gatctggctg gaccacaagc ctctcaagct	1080
cctggggagc aagttccatg gactcaattc aatgacgcat tactggatga ctcatgagc	1140
aatgacaatg agataggcct ttacttccct cttggtgaga ttgtcccaaa tgttgatgcc	1200
gtcaccaaaa gatggacatt cgaaagaaaa gagaaccatt cgaacaaaag tatcgttctt	1260
cacgagttgg atcaattcac gccaaagagg aaagatgcaa agaacatagt ggaaagtcag	1320
gccttaagct gtagagtgcg tatctctcca ttgctgtctg atgaaacaga tgccttaagc	1380
gaaactcaag tgttgtcaaa gaaggagaat acccaagtta cgtttgacta cgatgcattt	1440
ccgttgtgga cgtatgcaa aagaccgaat agagcgttct ttgttggtgg tgcctccaag	1500

```

aatgatgcc a tagtcaggac aatggcaaat gtaataggtg ctaggaatgg aaattataga 1560
cttgaaactc ccaattcctg tgcttttagga ggctgttata aagcgatgtg gtcatggtta 1620
aaggtagcatg aacctacaac tacaccatct ttcatgtttt ggtaaaccgc aagctttaac 1680
tggcagagag attgcgaatt cgtgtgccag tctgacgccg ctaagtggga acaatctaata 1740
ggtaaaattc aagctttatc agaagctgaa gcctatgtta aagcttttagc acaatctcag 1800
gggtcaa 1806

```

<210> 22

<211> 602

<212> PRT

<213> Kluyveromyces marxianus

<400> 22

```

Met Ser Thr Pro Tyr Tyr Leu Gly Phe Asp Leu Ser Thr Gln Gln Leu
1          5          10          15

Lys Cys Leu Ala Ile Asp Asp Gln Leu Asn Ile Val Thr Ser Val Ser
20          25          30

Ile Glu Phe Asp Arg Asn Phe Pro Ala Tyr Asn Thr Lys Lys Gly Val
35          40          45

Tyr Ile Lys Asn Gly Gly Val Ile Asp Ala Pro Val Ala Met Trp Leu
50          55          60

Glu Ala Val Asp Leu Cys Phe Ser Gln Leu Ala Glu Arg Ile Asp Leu
65          70          75          80

Lys Arg Val Gln Ser Met Ser Gly Ser Cys Gln Gln His Gly Thr Val
85          90          95

Tyr Trp Asn Cys Glu His Leu Pro Ser Asn Leu Asp Pro Ala Ser Thr
100         105         110

Leu Arg Glu Gln Leu Gln Gly Ser Leu Ser Arg Pro Val Ala Pro Asn
115         120         125

Trp Gln Asp His Ser Thr Lys Lys Gln Cys Asp Glu Leu Ala Glu Ser
130         135         140

Val Gly Gly Pro Glu Glu Leu Ala Arg Ile Thr Gly Ser Gly Ala His
145         150         155         160

Tyr Arg Phe Ser Gly Ser Gln Ile Ala Lys Ile His Glu Thr Glu Pro
165         170         175

Glu Val Tyr Glu Ala Thr Lys Arg Ile Ser Leu Val Ser Ser Phe Leu
180         185         190

Ala Ser Val Leu Val Gly Asp Ile Val Pro Leu Glu Glu Ala Asp Ala
195         200         205

Cys Gly Met Asn Leu Tyr Asp Leu Ser Lys His Asp Phe Asp Glu Thr
210         215         220

```

Leu Leu Ala Val Val Asp His Asp Thr Ala Arg Leu Arg Arg Lys Leu
 225 230 235 240
 Ser Asp Pro Pro Val Gly Ala Pro Thr Arg Glu Ser Pro Leu Thr Ser
 245 250 255
 Leu Gly Lys Val Ser Lys Tyr Phe Gln Asp Lys Tyr Gly Val Asn Cys
 260 265 270
 Glu Cys Glu Ile Phe Pro Phe Thr Gly Asp Asn Leu Ala Thr Ile Cys
 275 280 285
 Ser Leu Pro Leu Gln Lys Asn Asp Val Leu Ile Ser Leu Gly Thr Ser
 290 295 300
 Thr Thr Ile Leu Leu Val Thr Asp Gln Tyr His Ser Ser Pro Asn Tyr
 305 310 315 320
 His Leu Phe Ile His Pro Thr Val Pro Gly Tyr Tyr Met Gly Met Ile
 325 330 335
 Cys Tyr Cys Asn Gly Ser Leu Ala Arg Glu Arg Val Arg Asp Asp Leu
 340 345 350
 Ala Gly Pro Gln Ala Ser Gln Ala Pro Gly Glu Gln Val Pro Trp Thr
 355 360 365
 Gln Phe Asn Asp Ala Leu Leu Asp Asp Ser Leu Ser Asn Asp Asn Glu
 370 375 380
 Ile Gly Leu Tyr Phe Pro Leu Gly Glu Ile Val Pro Asn Val Asp Ala
 385 390 395 400
 Val Thr Lys Arg Trp Thr Phe Glu Arg Lys Glu Asn His Ser Asn Lys
 405 410 415
 Ser Ile Val Leu His Glu Leu Asp Gln Phe Thr Pro Lys Arg Lys Asp
 420 425 430
 Ala Lys Asn Ile Val Glu Ser Gln Ala Leu Ser Cys Arg Val Arg Ile
 435 440 445
 Ser Pro Leu Leu Ser Asp Glu Thr Asp Ala Leu Ser Glu Thr Gln Val
 450 455 460
 Leu Ser Lys Lys Glu Asn Thr Gln Val Thr Phe Asp Tyr Asp Ala Phe
 465 470 475 480
 Pro Leu Trp Thr Tyr Ala Lys Arg Pro Asn Arg Ala Phe Phe Val Gly
 485 490 495
 Gly Ala Ser Lys Asn Asp Ala Ile Val Arg Thr Met Ala Asn Val Ile
 500 505 510
 Gly Ala Arg Asn Gly Asn Tyr Arg Leu Glu Thr Pro Asn Ser Cys Ala
 515 520 525

Leu Gly Gly Cys Tyr Lys Ala Met Trp Ser Trp Leu Lys Val His Glu
530 535 540

Pro Thr Thr Thr Pro Ser Phe Asp Val Trp Leu Asn Ala Ser Phe Asn
545 550 555 560

Trp Gln Arg Asp Cys Glu Phe Val Cys Gln Ser Asp Ala Ala Lys Trp
565 570 575

Glu Gln Ser Asn Gly Lys Ile Gln Ala Leu Ser Glu Ala Glu Ala Tyr
580 585 590

Val Lys Ala Leu Ala Gln Ser Gln Gly Gln
595 600

<210> 23

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 23

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 24

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 24

gcctaagtaa tatggagtcg acattttag ttatgtatg tgttt 45

<210> 25

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 25

ctttagcaca atctcagggt caataaagat ctcccatgtc tctactgg 48

<210> 26

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 26

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 27

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 27

aaacacatac ataaactaaa aatgtcgact ccatattact taggc 45

<210> 28

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 28

ccagtagaga catgggagat ctttattgac cctgagattg tgctaaag 48

<210> 29

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 29

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 30

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 30

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 31

<211> 1803

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 31

atgttgtgtt cagtaattca gagacagaca agagaggttt ccaacacaat gtcttttagac	60
tcatactatc ttgggtttga tctttcgacc caacaactga aatgtctcgc cattaaccag	120
gacctaaaaa ttgtccattc agaaacagtg gaatttgaaa aggatcttcc gcattatcac	180
acaaagaagg gtgtctatat acacggcgac actatcgaat gtcccgtagc catgtggtta	240
gaggctctag atctggttct ctcgaaatat cgcgaggcta aatttccatt gaacaaagtt	300
atggccgtct cagggtcctg ccagcagcac gggctctgtct actggtcctc ccaagccgaa	360
tctctgttag agcaattgaa taagaaaccg gaaaaagatt tattgacta cgtgagctct	420
gtagcatttg caaggcaaac cgccccaat tggcaagacc acagtactgc aaagcaatgt	480
caagagtttg aagagtgcac aggtgggcct gaaaaaatgg ctcaattaac agggtcaga	540
gccattttta gatttactgg tcctcaaatt ctgaaaattg cacaattaga accagaagct	600
tacgaaaaaa caaagaccat ttcttttagtg tctaattttt tgacttctat cttagtgggc	660
catcttgttg aattagagga ggcagatgcc tgtggtatga acctttatga tatacgtgaa	720
agaaaattca gtgatgagct actacatcta attgatagtt cttctaagga taaaactatc	780
agacaaaaat taatgagagc acccatgaaa aatttgatag cgggtaccat ctgtaaatat	840
tttattgaga agtacggttt caatacaaac tgcaaggtct ctcccatgac tggggataat	900
ttagccacta tatgttcttt acccctgcgg aagaatgacg ttctcgtttc cctaggaaca	960
agtactacag ttcttctggt caccgataag taccacctt ctccgaacta tcatcttttc	1020
attcatccaa ctctgccaaa ccatttatatg ggtatgattt gttattgtaa tggttctttg	1080
gcaagggaga ggataagaga cgagttaaac aaagaacggg aaaataatta tgagaagact	1140
aacgattgga ctctttttta tcaagctgtg ctagatgact cagaaagtag tgaaaatgaa	1200
ttaggtgtat attttctctt gggggagatc gttcctagcg taaaagccat aaacaaaagg	1260
gttatcttca atccaaaaac gggatgattt gaaagagagg tggccaagtt caaagacaag	1320
aggcacgatg ccaaaaatat tgtagaatca caggctttaa gttgcagggt aagaatatct	1380
cccctgcttt cggattcaaa cgcaagctca caacagagac tgaacgaaga tacaatcgtg	1440
aagtttgatt acgatgaatc tccgctgcgg gactacctaa ataaaaggcc agaaaggact	1500
ttttttgtag gtggggcttc taaaaacgat gctattgtga agaagtttgc tcaagtcatt	1560
ggtgctacaa agggtaattt taggctagaa acaccaaact catgtgccct tgggtggtgt	1620
tataaggcca tgtggtcatt gttatatgac tctaataaaa ttgcagttcc ttttgataaa	1680
tttctgaatg acaatttttc atggcatgta atggaaagca tatccgatgt ggataatgaa	1740
aattgggacg gctataatc caagattgtc cccttaagcg aactggaaaa gactctcatc	1800
taa	1803

<210> 32

<211> 600

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 32

Met Leu Cys Ser Val Ile Gln Arg Gln Thr Arg Glu Val Ser Asn Thr
 1 5 10 15

Met Ser Leu Asp Ser Tyr Tyr Leu Gly Phe Asp Leu Ser Thr Gln Gln
 20 25 30

Leu Lys Cys Leu Ala Ile Asn Gln Asp Leu Lys Ile Val His Ser Glu
 35 40 45

Thr Val Glu Phe Glu Lys Asp Leu Pro His Tyr His Thr Lys Lys Gly
 50 55 60

Val Tyr Ile His Gly Asp Thr Ile Glu Cys Pro Val Ala Met Trp Leu
 65 70 75 80

Glu Ala Leu Asp Leu Val Leu Ser Lys Tyr Arg Glu Ala Lys Phe Pro
 85 90 95

Leu Asn Lys Val Met Ala Val Ser Gly Ser Cys Gln Gln His Gly Ser
 100 105 110

Val Tyr Trp Ser Ser Gln Ala Glu Ser Leu Leu Glu Gln Leu Asn Lys
 115 120 125

Lys Pro Glu Lys Asp Leu Leu His Tyr Val Ser Ser Val Ala Phe Ala
 130 135 140

Arg Gln Thr Ala Pro Asn Trp Gln Asp His Ser Thr Ala Lys Gln Cys
 145 150 155 160

Gln Glu Phe Glu Glu Cys Ile Gly Gly Pro Glu Lys Met Ala Gln Leu
 165 170 175

Thr Gly Ser Arg Ala His Phe Arg Phe Thr Gly Pro Gln Ile Leu Lys

180

185

190

Ile Ala Gln Leu Glu Pro Glu Ala Tyr Glu Lys Thr Lys Thr Ile Ser
 195 200 205

Leu Val Ser Asn Phe Leu Thr Ser Ile Leu Val Gly His Leu Val Glu
 210 215 220

Leu Glu Glu Ala Asp Ala Cys Gly Met Asn Leu Tyr Asp Ile Arg Glu
 225 230 235 240

Arg Lys Phe Ser Asp Glu Leu Leu His Leu Ile Asp Ser Ser Ser Lys
 245 250 255

Asp Lys Thr Ile Arg Gln Lys Leu Met Arg Ala Pro Met Lys Asn Leu
 260 265 270

260	265	270
Ile Ala Gly Thr Ile Cys Lys Tyr Phe Ile Glu Lys Tyr Gly Phe Asn 275 280 285		
Thr Asn Cys Lys Val Ser Pro Met Thr Gly Asp Asn Leu Ala Thr Ile 290 295 300		
Cys Ser Leu Pro Leu Arg Lys Asn Asp Val Leu Val Ser Leu Gly Thr 305 310 315 320		
Ser Thr Thr Val Leu Leu Val Thr Asp Lys Tyr His Pro Ser Pro Asn 325 330 335		
Tyr His Leu Phe Ile His Pro Thr Leu Pro Asn His Tyr Met Gly Met 340 345 350		
Ile Cys Tyr Cys Asn Gly Ser Leu Ala Arg Glu Arg Ile Arg Asp Glu 355 360 365		
Leu Asn Lys Glu Arg Glu Asn Asn Tyr Glu Lys Thr Asn Asp Trp Thr 370 375 380		
Leu Phe Asn Gln Ala Val Leu Asp Asp Ser Glu Ser Ser Glu Asn Glu 385 390 395 400		
Leu Gly Val Tyr Phe Pro Leu Gly Glu Ile Val Pro Ser Val Lys Ala 405 410 415		
Ile Asn Lys Arg Val Ile Phe Asn Pro Lys Thr Gly Met Ile Glu Arg 420 425 430		
Glu Val Ala Lys Phe Lys Asp Lys Arg His Asp Ala Lys Asn Ile Val 435 440 445		
Glu Ser Gln Ala Leu Ser Cys Arg Val Arg Ile Ser Pro Leu Leu Ser 450 455 460		
Asp Ser Asn Ala Ser Ser Gln Gln Arg Leu Asn Glu Asp Thr Ile Val 465 470 475 480		
Lys Phe Asp Tyr Asp Glu Ser Pro Leu Arg Asp Tyr Leu Asn Lys Arg 485 490 495		
Pro Glu Arg Thr Phe Phe Val Gly Gly Ala Ser Lys Asn Asp Ala Ile 500 505 510		
Val Lys Lys Phe Ala Gln Val Ile Gly Ala Thr Lys Gly Asn Phe Arg 515 520 525		
Leu Glu Thr Pro Asn Ser Cys Ala Leu Gly Gly Cys Tyr Lys Ala Met 530 535 540		
Trp Ser Leu Leu Tyr Asp Ser Asn Lys Ile Ala Val Pro Phe Asp Lys 545 550 555 560		
Phe Leu Asn Asp Asn Phe Pro Trp His Val Met Glu Ser Ile Ser Asp 565 570 575		

Val	Asp	Asn	Glu	Asn	Trp	Asp	Arg	Tyr	Asn	Ser	Lys	Ile	Val	Pro	Leu
			580					585					590		

Ser Glu Leu Glu Lys Thr Leu Ile
595 600

<210> 33

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 33

ggtaccgagc tctaactgat ctatccaaaa ctg 33

<210> 34

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 34

ggtaaccgatc agcatgcat cgctcgacat ttgatatac 39

<210> 35

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 35

ccgcgggagc tctaactgat ctatccaaaa ctg 33

<210> 36

<211> 39

<212> DNA

<213> Artificial Sequence

$\langle 220 \rangle$

<223> Artificial DNA Primer

<400> 36

ccgcgggatac agcatgcgat cgctcgacat ttgatatac 39

<210> 37

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 37

tctccacac ctgcagtata tctaggaacc catcag 36

<210> 38

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 38

ctgtctctga attactgaac acaacatttt tagtttatgt atgtgttt 48

<210> 39

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 39

gcgaactgga aaagactctc atctaaagat ctcccatgtc tctactgg 48

<210> 40

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 40

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 41

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 41

aaacacatac ataaactaaa aatgttgtgt tcagtaattc agagacag 48

<210> 42

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 42

ccagtagaga catgggagat ctttagatga gagtcttttc cagttcgc 48

<210> 43

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 43

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 44

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 44

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 45

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 45

gatccccggg ctgcaatgtt gccatttct ggacctt 37

<210> 46
 <211> 37
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial DNA Primer

<400> 46
 cgctgcaggt cgacgtgtta catgcgtaca cgctct 37

<210> 47
 <211> 18
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial DNA Primer

<400> 47
 cgtcgacctg cagcgtac 18

<210> 48
 <211> 21
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Artificial DNA Primer

<400> 48
 tgcagcccgg ggatccttt t 21

<210> 49
 <211> 393
 <212> DNA
 <213> Zygosaccharomyces bailii

<400> 49
 atgttgccca tttctggacc ttccaacatg ctgcatggcc tcgtttcagc ccgttgtgca 60
 ggggggttga ggccacttat ctcgcatctg cgtaggggag tttttcctaa gatgcttacc 120
 atgacaggta ttggggccaa gagatttgtc tccagcgaaa tacaggagaa agacgaacaa 180
 gctggtgagt ctactactgc tacagatact ggtatcatc ataaaacgga gcaggagacc 240
 ctagtatatt tcgacaacgt ctatccacgg accgcatctc tatggagccc tgcgcaatgg 300
 tacaatctac ttctaactaa tcaatcgagg gaggtgtta ggcaaaagat cagcggttcg 360
 atcccgctag agaccatttt ttggcttcat tga 393

<210> 50

<211> 130

<212> PRT

<213> Zygosaccharomyces bailii

<400> 50

Met Leu Pro Ile Ser Gly Pro Ser Asn Met Leu His Gly Leu Val Ser
 1 5 10 15

Ala Arg Cys Ala Gly Gly Trp Arg Pro Leu Ile Ser His Leu Arg Arg
 20 25 30

Gly Val Phe Pro Lys Met Leu Thr Met Thr Gly Ile Gly Ala Lys Arg
 35 40 45

Phe Val Ser Ser Glu Ile Gln Glu Lys Asp Glu Gln Ala Gly Glu Ser
 50 55 60

Thr Thr Ala Thr Asp Thr Gly Ile Ile His Lys Thr Glu Gln Glu Thr
 65 70 75 80

Leu Val Tyr Phe Asp Asn Val Tyr Pro Arg Thr Ala Ser Leu Trp Ser
 85 90 95

Pro Ala Gln Trp Tyr Asn Leu Leu Leu Thr Asn Gln Ser Arg Glu Ala
 100 105 110

Val Arg Gln Lys Ile Ser Gly Ser Ile Pro Leu Glu Thr Ile Phe Trp
 115 120 125

Leu His
 130

<210> 51

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 51

ggtaccgagc tctaactgat ctatccaaaa ctg 33

<210> 52

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 52

ggtaccgatc agcatgcat cgctcgacat ttgatatac 39

<210> 53

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 53

ccgcgggagc tctaactgat ctatccaaaa ctg 33

<210> 54

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 54

ccgcgggatc agcatgcat cgctcgacat ttgatatac 39

<210> 55

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 55

tctccacac ctgcagtata tctaggaacc catcag 36

<210> 56

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 56

ctgtctctga attactgaac acaacatttt tagtttatgt atgtgttt 48

<210> 57

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 57

gcgaactgga aaagactctc atctaaagat ctcccatgtc tctactgg 48

<210> 58

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 58

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 59

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 59

aaacacatac ataaactaaa aatgttgtgt tcagtaattc agagacag 48

<210> 60

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 60

ccagtagaga catgggagat cttagatga gagtcttttc cagttcgc 48

<210> 61

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 61

tcttcacac ctgcagtata tctaggaacc catcag 36

<210> 62

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 62

atcagttaga ctgcaggaac atagaaatat cgaatgggaa 40

<210> 63

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 63

ctagaactag tggatccccc atgtcattcg acgacttaca caaag 45

<210> 64

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 64

atatcgaatt cctgcagccc tcatgctaca agcgcacaca a 41

<210> 65

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 65

ctagaactag tggatccccc atgtccaata actcattcac taacttca 48

<210> 66

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 66

atatcgaatt cctgcagccc tcacatccat tccttgaatt g 41

<210> 67

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 67

ctagaactag tggatccccc atgtcaaac cagtagactt ctctga 46

<210> 68

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 68

atatcgaatt cctgcagccc ttaaactga ctttggtgc 40

<210> 69

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 69

ctagaactag tggatccccc atgactgtta ctactcctt tgtgaatg 48

<210> 70

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 70

atatcgaatt cctgcagccc ttaatcaact ctcttcttc caacc 45

<210> 71

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 71

ctagaactag tggatccccc atgggtgggtt ctccgattc a 41

<210> 72

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 72

atatcgaatt cctgcagccc tcaagcaaat ccgttgcttt c 41

<210> 73

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 73

ctagaactag tggatccccc atgggtgttg aacaaatctt aaagag 46

<210> 74

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 74

atatcgaatt cctgcagccc ttaatcaact ctctttcttc caacc 45

<210> 75

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 75

ctagaactag tggatccccc atgacagaa ttgctattaa cggtttc 47

<210> 76

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 76

atatcgaatt cctgcagccc ttaagccttg gcaacatatt cg 42

<210> 77

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 77

ctagaactag tggatccccc atggtagag ttgctattaa cggtttc 47

<210> 78

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 78

atatcgaatt cctgcagccc ttaagccttg gcaacgtgtt 40

<210> 79

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 79

ctagaactag tggatccccc atgccaaagt tagtttagt tagacacg 48

<210> 80

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 80

atatcgaatt cctgcagccc ttatttcta ccttggttg caac 44

<210> 81

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 81

ctagaactag tggatccccc atgactgcaa gcacaccatc caat 44

<210> 82

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 82

atatcgaatt cctgcagccc ttaaggattt ttatgaaac cctca 45

<210> 83

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 83

ctagaactag tggatccccc atgactgtta ctgacacttt taaactg 47

<210> 84

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 84

atatcgaatt cctgcagccc tcatggattc ttttcgaaac cc 42

<210> 85

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 85

ctagaactag tggatccccc atggctgtct ctaaagtta cgc 43

<210> 86

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 86

atatcgaatt cctgcagccc ttataatttg tcaccgtggt gg 42

<210> 87

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 87

ctagaactag tggatccccc atgtctagat tagaaagatt gacctca 47

<210> 88

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 88

atatcgaatt cctgcagccc ttaaacggta gagacttgca aagtg 45

<210> 89

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 89

ctagaactag tggatcccc atgccagagt ccagattgca 40

<210> 90

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 90

atatcgaatt cctgcagccc ctagaattct tgaccaacag tagaaatg 48

<210> 91

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 91

ctagaactag tggatcccc atgtctgaaa ttactttggg taaa 44

<210> 92

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 92

atatcgaatt cctgcagccc ttattgctta gcgttggtag cag 43

<210> 93

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 93

ctagaactag tggatcccc atgtctgaaa ttactcttg aaaatact 48

<210> 94

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 94

atatcgaatt cctgcagccc ttattgttg gcattttag cgg 43

<210> 95

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 95

ctagaactag tggatcccc atgactaagc tacacttga cactgc 46

<210> 96

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 96

atatcgaatt cctgcagccc ttacaacta attctgacag cttttac 47

<210> 97

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 97

ctagaactag tggatcccc atgccttcgc aagtcattcc t 41

<210> 98

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 98

atatcgaatt cctgcagccc tcatttagaa gtctcaacaa catatc 46

<210> 99

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 99

ctagaactag tggatccccc atgtcttctc ctgagaaatt tgaaggt 47

<210> 100

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 100

atatcgaatt cctgcagccc ctagtctgaa aattctttgt cgtagcc 47

<210> 101

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 101

ctagaactag tggatccccc atggctagaa ctttctttgt cgg 43

<210> 102

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 102

atatcgaatt cctgcagccc ttagtttcta gagttgatga tatcaaca 48

<210> 103

<211> 43

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 103

ctagaactag tggatccccc atgtctgctg ctgctgatag att 43

<210> 104

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 104

atatcgaatt cctgcagccc ctaatcttca ttagatctca attcttca 48

<210> 105

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 105

ctagaactag tggatccccc atgcttgctg tcagaagatt aaca 44

<210> 106

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 106

atatcgaatt cctgcagccc ctattcgtca tcgatgtcta gctct 45

<210> 107

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 107

ctagaactag tggatccccc atgggattga ctactaaacc tctatct 47

<210> 108

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 108

atatcgaatt cctgcagccc ttaccatttc aacagatcgt cc 42

<210> 109

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 109

ctagaactag tggatccccc atggatccta atagtaacag ttctagcg 48

<210> 110

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 110

atatcgaatt cctgcagccc ttattcaaa tcattatatt catttacagg ttg 53

<210> 111

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 111

ctagaactag tggatcccc atgaacgata gccaaaactg c 41

<210> 112

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 112

atatcgaatt cctgcagccc ttattggggg gaagtgtatt g 41

<210> 113

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 113

ctagaactag tggatcccc atgcaaagcc catatccaat g 41

<210> 114

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 114

atatcgaatt cctgcagccc tcagtccatg tgtgggaagg 40

<210> 115

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 115

ctagaactag tggatcccc atgtttgttt caccacctcc a 41

<210> 116

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 116

atatcgaatt cctgcagccc ctaggacatt ccatcaggct t 41

<210> 117

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 117

ctagaactag tggatccccc atgtctatcc cagaaactca aaaagg 46

<210> 118

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 118

atatcgaatt cctgcagccc ttatttagaa gtgtcacaac gtatctacc 49

<210> 119

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 119

ctagaactag tggatccccc atgaacgata gccaaaactg c 41

<210> 120

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 120

atatcgaatt cctgcagccc ttattggggg gaagtgtatt g 41

<210> 121

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 121

ctagaactag tggatcccc atgttgcaag gaattttaga aaccg 45

<210> 122

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 122

atatcgaatt cctgcagccc tcaaaatatt gtaactatat tcatttcac tg 52

<210> 123

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 123

ctagaactag tggatcccc atgaatacaa acgttccaat attcag 46

<210> 124

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 124

atatcgaatt cctgcagccc ttattgttcg taaaacaag taccc 45

<210> 125

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 125

ctagaactag tggatccccc atgactaaat ctcattcaga agaagtga 48

<210> 126

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 126

atatcgaatt cctgcagccc ttataattca tatagacagc tgccca 46

<210> 127

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 127

ctagaactag tggatccccc atgaagctac tgtcttctat cgaacaag 48

<210> 128

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 128

atatcgaatt cctgcagccc ttactctttt ttggggttg gtgg 44

<210> 129

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 129

ctagaactag tggatcccc atgactgctg aagaattga ttttc 46

<210> 130

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 130

atatcgaatt cctgcagccc ttacagtctt ttagataat gaatctgacc 50

<210> 131

<211> 1314

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 131

atggctgtct ctaaagttta cgctagatcc gtctacgact cccgtggtaa cccaaccgtc	60
gaagtgaat taaccaccga aaagggtgtt ttcagatcca ttgtcccatc tgggtgcttct	120
accggtgtcc acgaagcttt ggaaatgaga gatggtgaca aatccaagtg gatgggtaag	180
ggtgttttgc acgctgttaa gaacgtcaac gatgtcattg ctccagcttt cgttaaggct	240
aacattgatg ttaaggacca aaaggccgtc gatgacttct tgatttcttt ggacggtact	300
gccaacaaat ccaagttggg tgctaacgct atcttggtg tttctttggc tgcttccaga	360
gctgccgctg ctgaaaagaa tgtcccatta tacaagcact tggctgactt gtctaagtcc	420
aagacctctc catacgtttt gccagttcca ttcttgaacg ttttgaacgg tggttcccac	480
gctggtggtg ctttggcttt gcaagaattt atgattgctc caactggtgc taagaccttc	540
gctgaagctt tgagaattgg ttccgaagtt taccacaact tgaagtcttt gaccaagaag	600
agatacggtg cttctgccgg taacgtcggg gacgaagggt gtgttgctcc aaacattcaa	660
actgtgaag aagcttttga cttgattgtt gacgctatca aggctgctgg tcacgacggt	720
aagatcaaga tcggttttga ctgtgcttcc tctgaattct tcaaggacgg taagtacgac	780
ttggacttca agaatccaaa ctctgacaaa tccaagtggt tgactggtcc tcaattggct	840
gacttgtagc actccttgat gaagagatac ccaattgtct ccatcgaaga tccatttgct	900
gaagatgact gggaagcttg gtctcacttc ttcaagaccg ctggtattca aattggtgct	960
gatgacttga ctgtcaccaa cccaagaga attgctaccg ctatcgaaaa gaaggctgcc	1020
gacgctttgt tgttgaaggt caaccaaatac ggtacottgt ctgaatccat caaggctgct	1080
caagactctt tcgctgccgg ttggggtgtt atggtttccc acagatctgg tgaaactgaa	1140
gacactttca ttgctgactt ggtcgtcggg ttgagaactg gtcaaataca gactggtgct	1200
ccagctagat ccgaaagatt ggctaaattg aaccaattgt tgagaatcga agaagaatta	1260
ggtgacaacg ctgttttcgc tgggtgaaaac ttccaccacg gtgacaaatt ataa	1314

<210> 132

<211> 437

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 132

Met Ala Val Ser Lys Val Tyr Ala Arg Ser Val Tyr Asp Ser Arg Gly
 1 5 10 15

Asn Pro Thr Val Glu Val Glu Leu Thr Thr Glu Lys Gly Val Phe Arg
 20 25 30

Ser Ile Val Pro Ser Gly Ala Ser Thr Gly Val His Glu Ala Leu Glu
 35 40 45

Met Arg Asp Gly Asp Lys Ser Lys Trp Met Gly Lys Gly Val Leu His
 50 55 60

Ala Val Lys Asn Val Asn Asp Val Ile Ala Pro Ala Phe Val Lys Ala
 65 70 75 80

Asn Ile Asp Val Lys Asp Gln Lys Ala Val Asp Asp Phe Leu Ile Ser
 85 90 95

Leu Asp Gly Thr Ala Asn Lys Ser Lys Leu Gly Ala Asn Ala Ile Leu
 100 105 110

Gly Val Ser Leu Ala Ala Ser Arg Ala Ala Ala Ala Glu Lys Asn Val
 115 120 125

Pro Leu Tyr Lys His Leu Ala Asp Leu Ser Lys Ser Lys Thr Ser Pro
 130 135 140

Tyr Val Leu Pro Val Pro Phe Leu Asn Val Leu Asn Gly Gly Ser His
 145 150 155 160

Ala Gly Gly Ala Leu Ala Leu Gln Glu Phe Met Ile Ala Pro Thr Gly
 165 170 175

Ala Lys Thr Phe Ala Glu Ala Leu Arg Ile Gly Ser Glu Val Tyr His
 180 185 190

Asn Leu Lys Ser Leu Thr Lys Lys Arg Tyr Gly Ala Ser Ala Gly Asn
 195 200 205

Val Gly Asp Glu Gly Gly Val Ala Pro Asn Ile Gln Thr Ala Glu Glu
 210 215 220

Ala Leu Asp Leu Ile Val Asp Ala Ile Lys Ala Ala Gly His Asp Gly
 225 230 235 240

Lys Ile Lys Ile Gly Leu Asp Cys Ala Ser Ser Glu Phe Phe Lys Asp
 245 250 255

Gly Lys Tyr Asp Leu Asp Phe Lys Asn Pro Asn Ser Asp Lys Ser Lys
260 265 270

Trp Leu Thr Gly Pro Gln Leu Ala Asp Leu Tyr His Ser Leu Met Lys
275 280 285

Arg Tyr Pro Ile Val Ser Ile Glu Asp Pro Phe Ala Glu Asp Asp Trp
290 295 300

Glu Ala Trp Ser His Phe Phe Lys Thr Ala Gly Ile Gln Ile Val Ala
305 310 315 320

Asp Asp Leu Thr Val Thr Asn Pro Lys Arg Ile Ala Thr Ala Ile Glu
325 330 335

Lys Lys Ala Ala Asp Ala Leu Leu Leu Lys Val Asn Gln Ile Gly Thr
340 345 350

Leu Ser Glu Ser Ile Lys Ala Ala Gln Asp Ser Phe Ala Ala Gly Trp
355 360 365

Gly Val Met Val Ser His Arg Ser Gly Glu Thr Glu Asp Thr Phe Ile
370 375 380

Ala Asp Leu Val Val Gly Leu Arg Thr Gly Gln Ile Lys Thr Gly Ala
385 390 395 400

Pro Ala Arg Ser Glu Arg Leu Ala Lys Leu Asn Gln Leu Leu Arg Ile
405 410 415

Glu Glu Glu Leu Gly Asp Asn Ala Val Phe Ala Gly Glu Asn Phe His
420 425 430

His Gly Asp Lys Leu
435

<210> 133

<211> 2880

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 133

atgactgtta ctactccttt tgtgaatggt acttcttatt gtaccgtcac tgcattattcc	60
gttcaatcct ataaagctgc catagatttt tacaccaagt ttttgtcatt agaaaaccgc	120
tcttctccag atgaaaactc cactttattg tctaacgatt ccatctcttt gaagatcctt	180
ctacgtcctg atgaaaaaat caataaaaat gttgaggctc atttgaagga attgaacagt	240
attaccaaga ctcaagactg gagatcacat gccacccaat ccttggtatt taacacttcc	300
gacatcttgg cagtcaagga cactctaaat gctatgaacg ctccctcttca aggetaccca	360
acagaactat ttccaatgca gttgtacact ttggacccat taggtaacgt tgttggtgtt	420
acttctacta agaacgcagt ttcaaccaag ccaactccac caccagcacc agaagcttct	480
gctgagctct gtcttttctc taaagttcac tcttactctg atttggctta ccgtatgaaa	540
accaccgaca cctatccatc tctgccaaag ccattgaaca ggccatcaaaa ggcaattgcc	600

gtcatgactt cgggtggtga tgctccaggt atgaactcta acgttagagc catcgtgcgt	660
tccgctatct tcaaaggttg tcgtgccttt gttgtcatgg aaggttatga aggtttggtt	720
cgtggtggtc cagaatacat caaggaattc cactgggaag acgtccgtgg ttggtctgct	780
gaaggtggta ccaacattgg tactgccgtg tgtatggaat tcaagaagcg cgaaggtaga	840
ttattgggtg cccaacattht gattgaggcc ggtgtcgatg ctttgatcgt ttgtgggtgt	900
gacggttctt tgactggtgc tgatctgttt agatcagaat ggccttcttt gatcgaggaa	960
ttgttgaaaa caaacagaat ttccaacgaa caatacgaaa gaatgaagca tttgaatatt	1020
tgcggtactg tcggttctat tgataacgat atgtccacca cggatgctac tattggtgct	1080
tactctgcct tggacagaat ctgtaaggcc atcgattacg ttgaagccac tgccaactct	1140
cactcaagag ctttcgttgt tgaagttatg ggtagaaaact gtggttggtt agctttatta	1200
gctggtatcg ccacttcgcg tgactatata tttattccag agaagccagc cacttccagc	1260
gaatggcaag atcaaatgtg tgacattgtc tccaagcaca gatcaagggg taagagaacc	1320
accattgttg ttgttgacaga aggtgctatc gctgctgact tgacccaat ttctccaagc	1380
gacgtccaca aagttctagt tgacagatta ggtttggata caagaattac taccttaggt	1440
cacgttcaaa gaggtggtac tgctgttgct tacgaccgta tcttggctac tttacaaggt	1500
cttgaggccg ttaatgcctg tttggaatcc actccagaca ccccatcacc attgattgct	1560
gttaacgaaa acaaaattgt tcgtaaacca ttaatggaat ccgtcaagtt gaccaaagca	1620
gttgacagaag ccattcaagc taaggatttc aagagagcta tgtctttaag agacactgag	1680
ttcattgaac atttaaacaa tticattggt atcaactctg ctgaccacaa cgaaccaaag	1740
ctaccaaaag acaagagact gaagattgcc attgttaatg tcggtgctcc agctgggtgt	1800
atcaactctg ccgtctactc gatggctact tactgtatgt cccaaggtca cagaccatac	1860
gctatctaca atggttggtc tgggttggtc agacatgaaa gtgttcgttc tttgaactgg	1920
aaggatatgt tgggttggtc atcccggtgt ggttctgaaa tcggtactaa cagagtcact	1980
ccagaagaag cagatctagg tatgattgct tactatttcc aaaagtacga atttgatggt	2040
ttgatcatcg ttggtggttt cgaagctttt gaatctttac atcaattaga gagagcaaga	2100
gaaagttatc cagcttctag aatcccaatg gtcttgatac cagctacttt gtctaacaat	2160
gttccaggta ctgaatactc tttgggttct gataccgctt tgaatgctct aatggaatac	2220
tgtgatgttg ttaaacatc cgcttcttca accagaggta gagccttcgt tgcgattgt	2280
caaggtggta actcaggcta tttggccact tacgcttctt tggctgttg tgctcaagtc	2340
tcttatgtcc cagaagaagg tatttctttg gagcaattgt ccgaggatat tgaatactta	2400
gctcaatctt ttgaaaaggc agaaggtaga ggtagatttg gtaaattgat tttgaagagt	2460
acaaacgctt ctaaggcttt atcagccact aaattggctg aagttattac tgctgaagcc	2520
gatggcagat ttgacgctaa gccagcttat ccaggtcatg tacaacaagg ttggttgcca	2580
tctccaattg atagaacaag agccactaga atggccatta aagctgtcgg cttcatcaaa	2640
gacaaccaag ctgccattgc tgaagctcgt gctgccgaag aaaacttcaa cgctgatgac	2700
aagaccattt ctgacactgc tgctgtcgtt ggtgttaagg gttcacatgt cgtttacaac	2760
tccattagac aattgtatga ctatgaaact gaagtttcca tgagaatgcc aaaggtcatt	2820
cactggcaag ctaccagact cattgctgac catttggttg gaagaaagag agttgattaa	2880

<210> 134

<211> 959

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 134

Met Thr Val Thr Thr Pro Phe Val Asn Gly Thr Ser Tyr Cys Thr Val
 1 5 10 15

Thr Ala Tyr Ser Val Gln Ser Tyr Lys Ala Ala Ile Asp Phe Tyr Thr
 20 25 30

Lys Phe Leu Ser Leu Glu Asn Arg Ser Ser Pro Asp Glu Asn Ser Thr
 35 40 45

Leu Leu Ser Asn Asp Ser Ile Ser Leu Lys Ile Leu Leu Arg Pro Asp
 50 55 60

Glu Lys Ile Asn Lys Asn Val Glu Ala His Leu Lys Glu Leu Asn Ser
 65 70 75 80

Ile Thr Lys Thr Gln Asp Trp Arg Ser His Ala Thr Gln Ser Leu Val
 85 90 95

Phe Asn Thr Ser Asp Ile Leu Ala Val Lys Asp Thr Leu Asn Ala Met
 100 105 110

Asn Ala Pro Leu Gln Gly Tyr Pro Thr Glu Leu Phe Pro Met Gln Leu
 115 120 125

Tyr Thr Leu Asp Pro Leu Gly Asn Val Val Gly Val Thr Ser Thr Lys
 130 135 140

Asn Ala Val Ser Thr Lys Pro Thr Pro Pro Ala Pro Glu Ala Ser
 145 150 155 160

Ala Glu Ser Gly Leu Ser Ser Lys Val His Ser Tyr Thr Asp Leu Ala
 165 170 175

Tyr Arg Met Lys Thr Thr Asp Thr Tyr Pro Ser Leu Pro Lys Pro Leu
 180 185 190

Asn Arg Pro Gln Lys Ala Ile Ala Val Met Thr Ser Gly Gly Asp Ala
 195 200 205

Pro Gly Met Asn Ser Asn Val Arg Ala Ile Val Arg Ser Ala Ile Phe
 210 215 220

Lys Gly Cys Arg Ala Phe Val Val Met Glu Gly Tyr Glu Gly Leu Val
 225 230 235 240

Arg Gly Gly Pro Glu Tyr Ile Lys Glu Phe His Trp Glu Asp Val Arg
 245 250 255

Gly Trp Ser Ala Glu Gly Gly Thr Asn Ile Gly Thr Ala Arg Cys Met
 260 265 270

Glu Phe Lys Lys Arg Glu Gly Arg Leu Leu Gly Ala Gln His Leu Ile
 275 280 285

Glu Ala Gly Val Asp Ala Leu Ile Val Cys Gly Gly Asp Gly Ser Leu
 290 295 300

Thr Gly Ala Asp Leu Phe Arg Ser Glu Trp Pro Ser Leu Ile Glu Glu
 305 310 315 320

Leu Leu Lys Thr Asn Arg Ile Ser Asn Glu Gln Tyr Glu Arg Met Lys
 325 330 335

His Leu Asn Ile Cys Gly Thr Val Gly Ser Ile Asp Asn Asp Met Ser
 340 345 350

Thr Thr Asp Ala Thr Ile Gly Ala Tyr Ser Ala Leu Asp Arg Ile Cys
 355 360 365

Lys Ala Ile Asp Tyr Val Glu Ala Thr Ala Asn Ser His Ser Arg Ala
 370 375 380

Phe Val Val Glu Val Met Gly Arg Asn Cys Gly Trp Leu Ala Leu Leu
 385 390 395 400

Ala Gly Ile Ala Thr Ser Ala Asp Tyr Ile Phe Ile Pro Glu Lys Pro
 405 410 415

Ala Thr Ser Ser Glu Trp Gln Asp Gln Met Cys Asp Ile Val Ser Lys
 420 425 430

His Arg Ser Arg Gly Lys Arg Thr Thr Ile Val Val Val Ala Glu Gly
 435 440 445

Ala Ile Ala Ala Asp Leu Thr Pro Ile Ser Pro Ser Asp Val His Lys
 450 455 460

Val Leu Val Asp Arg Leu Gly Leu Asp Thr Arg Ile Thr Thr Leu Gly
 465 470 475 480

His Val Gln Arg Gly Gly Thr Ala Val Ala Tyr Asp Arg Ile Leu Ala
 485 490 495

Thr Leu Gln Gly Leu Glu Ala Val Asn Ala Val Leu Glu Ser Thr Pro
 500 505 510

Asp Thr Pro Ser Pro Leu Ile Ala Val Asn Glu Asn Lys Ile Val Arg
 515 520 525

Lys Pro Leu Met Glu Ser Val Lys Leu Thr Lys Ala Val Ala Glu Ala
 530 535 540

Ile Gln Ala Lys Asp Phe Lys Arg Ala Met Ser Leu Arg Asp Thr Glu
 545 550 555 560

Phe Ile Glu His Leu Asn Asn Phe Met Ala Ile Asn Ser Ala Asp His
 565 570 575

Asn Glu Pro Lys Leu Pro Lys Asp Lys Arg Leu Lys Ile Ala Ile Val
 580 585 590

```

    595      600      605
Asn Val Gly Ala Pro Ala Gly Gly Ile Asn Ser Ala Val Tyr Ser Met
    610      615      620
Ala Thr Tyr Cys Met Ser Gln Gly His Arg Pro Tyr Ala Ile Tyr Asn

Gly Trp Ser Gly Leu Ala Arg His Glu Ser Val Arg Ser Leu Asn Trp
625      630      635      640

Lys Asp Met Leu Gly Trp Gln Ser Arg Gly Gly Ser Glu Ile Gly Thr
    645      650      655

Asn Arg Val Thr Pro Glu Glu Ala Asp Leu Gly Met Ile Ala Tyr Tyr
    660      665      670

Phe Gln Lys Tyr Glu Phe Asp Gly Leu Ile Ile Val Gly Gly Phe Glu
    675      680      685

Ala Phe Glu Ser Leu His Gln Leu Glu Arg Ala Arg Glu Ser Tyr Pro
    690      695      700

Ala Phe Arg Ile Pro Met Val Leu Ile Pro Ala Thr Leu Ser Asn Asn
705      710      715      720

Val Pro Gly Thr Glu Tyr Ser Leu Gly Ser Asp Thr Ala Leu Asn Ala
    725      730      735

Leu Met Glu Tyr Cys Asp Val Val Lys Gln Ser Ala Ser Ser Thr Arg
    740      745      750

Gly Arg Ala Phe Val Val Asp Cys Gln Gly Gly Asn Ser Gly Tyr Leu
    755      760      765

Ala Thr Tyr Ala Ser Leu Ala Val Gly Ala Gln Val Ser Tyr Val Pro
    770      775      780

Glu Glu Gly Ile Ser Leu Glu Gln Leu Ser Glu Asp Ile Glu Tyr Leu
785      790      795      800

Ala Gln Ser Phe Glu Lys Ala Glu Gly Arg Gly Arg Phe Gly Lys Leu
    805      810      815

Ile Leu Lys Ser Thr Asn Ala Ser Lys Ala Leu Ser Ala Thr Lys Leu
    820      825      830

Ala Glu Val Ile Thr Ala Glu Ala Asp Gly Arg Phe Asp Ala Lys Pro
    835      840      845

Ala Tyr Pro Gly His Val Gln Gln Gly Gly Leu Pro Ser Pro Ile Asp
    850      855      860

Arg Thr Arg Ala Thr Arg Met Ala Ile Lys Ala Val Gly Phe Ile Lys
865      870      875      880

Asp Asn Gln Ala Ala Ile Ala Glu Ala Arg Ala Ala Glu Glu Asn Phe

```

885	890	895
Asn Ala Asp Asp Lys Thr Ile Ser Asp Thr Ala Ala Val Val Gly Val		
900	905	910
Lys Gly Ser His Val Val Tyr Asn Ser Ile Arg Gln Leu Tyr Asp Tyr		
915	920	925
Glu Thr Glu Val Ser Met Arg Met Pro Lys Val Ile His Trp Gln Ala		
930	935	940
Thr Arg Leu Ile Ala Asp His Leu Val Gly Arg Lys Arg Val Asp		
945	950	955

<210> 135

<211> 2484

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 135

atgttcaaac cagtagactt ctctgaaaca tctcctgtgc cgctgatata tgatcttgct	60
cctacacaat ctccacacca tgtggcacct agtcaagact ccagttatga tcttttatcc	120
cggagttccg atgataaaat tgatgctgaa aagggtccgc atgatgaatt atctaagcac	180
ttaccacttt ttcagaaaaag acctttgagc gatactccta tatcgagcaa ttggaactct	240
cctggaatca ctgaagaaaa tacaccttct gactctcctg aaaatagcgc tactaatttg	300
aaatcgctac atcgattgca tattaacgac gaaacgcaac taaaaaatgc taaaattccc	360
acaaacgata ctactgacta catgcctcct tcagatggag caaatgaggt aactcggatt	420
gatttgaaag acattaaatc acctacgaga caccataaaa gaagacctac caccatcgat	480
gttctcggtt taacaaagtc taaaacatct ccagatggtc tcatatcaaa ggaagatagt	540
ggatcaaagt tagtgattgt catggtcgga ctgccagcta cgggaaagtc atttattaca	600
aataaattat ccagattttt aaattattct ttatactatt gtaaagtgtt taatgtcgggt	660
aacactagaa ggaagtttgc taaggagcat ggcctaaagg accaggattc aaagtttttc	720
gagccgaaaa acgccgactc tactagggtg agagacaaat gggccatgga tactctggat	780
gaattgctag attatttatt agaaggttca ggatctgtgg gaatttttga tgctacaaat	840
acctctcgtg aaagaagaaa aaacgttctg gctagaatca gaaagagaag tcctcatttg	900
aaggttttat ttttagaatc tgtttgttcg gatcatgcac tggtagagaa aaatattaga	960
ctcaaattat ttggtccaga ttacaaaggt aaagatcctg aaagctcttt aaaagatttt	1020
aaaagtcgcc tggcaaaacta cttgaaagcc tatgaaccaa ttgaggatga cgaaaatttg	1080
cagtacatca aaatgataga tgtgggaaag aaagtcacgc catacaatat tcaagggttt	1140
ttagcttcgc agacgggtata ttatttggtt aatttcaatt tggctgacag acaaatttgg	1200
ataacgagaa gtggcgagag cgaagataat gttagtggcc gtataggcgg aaattcccat	1260
ttgactcctc tgggtctaag atttgctaaa agtctaccaa aattcattgc cagacagaga	1320
gaaatatttt atcaaaatct catgcaacaa aaaaagaata atgaaaatac agatgggaac	1380
atttataatg actttttcgt ttggaccagc atgctgtgcta ggactatagg gactgctcaa	1440

tatttcaacg aagatgatta tcctatcaaa caaatgaaaa tgtagatga gttaagtga 1500
 ggtgattatg atggtatgac atatccagaa attaaaaaca actttcctga agaattcgaa 1560
 aaaagacaga aagataagtt gagatacaga taccctggta ttggcggta atcgtatatg 1620
 gacgttatta atagactcag acctgttatc acagaactag aaagaatcga ggataacgtt 1680
 cttattatta cacaccgggt ggtggcaaga gccttattgg gttattttat gaacttgagt 1740
 atgggtatta ttgccaatth ggatgtccca ttacattgtg tatattgcct agaaccacaaa 1800
 ccatatggaa tcacttggtc attatgggag tatgatgaag catcggattc attttctaag 1860
 gtcccacaaa cggacttgaa taccaccaga gtaaaggagg ttggccttgt ttataatgaa 1920
 agaagatatt ctgttatacc aacagctccg ccaagtgcga gaagcagctt tgcaagtgcac 1980
 tttttgtcaa gaaaaagatc taatcctact tctgcatctt catcccagag tgaattatca 2040
 gaacaaccca agaatacgct tagtgcctcaa actggcagca ataataccac tctcattggg 2100
 agcaacttta acatcaagaa tgaaaatggt gattcgagaa taccattatc tgcaccactt 2160
 atggccacta atacttctaa taacatctta gatggtggag gtacctcaat ttcgatacat 2220
 cgtcccaggg ttgttccaaa tcaaaacaac gtgaatcctc ttttggctaa caacaataaa 2280
 gcggcttcta atgtacctaa tgtaaagaag tcagcggcta caccaaggca aatttttgaa 2340
 atagataaag tggacgaaaa gttatccatg ttgaaaaata aaagttttct attacatgga 2400
 aaggattatc ctaataatgc tgataataat gacaacgaag atataagggc aaaaaccatg 2460
 aatcgagcc aaagtcacgt ttaa 2484

<210> 136

<211> 827

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 136

Met Phe Lys Pro Val Asp Phe Ser Glu Thr Ser Pro Val Pro Pro Asp
 1 5 10 15

Ile Asp Leu Ala Pro Thr Gln Ser Pro His His Val Ala Pro Ser Gln
 20 25 30

Asp Ser Ser Tyr Asp Leu Leu Ser Arg Ser Ser Asp Asp Lys Ile Asp
 35 40 45

Ala Glu Lys Gly Pro His Asp Glu Leu Ser Lys His Leu Pro Leu Phe
 50 55 60

Gln Lys Arg Pro Leu Ser Asp Thr Pro Ile Ser Ser Asn Trp Asn Ser
 65 70 75 80

Pro Gly Ile Thr Glu Glu Asn Thr Pro Ser Asp Ser Pro Glu Asn Ser
 85 90 95

Ala Thr Asn Leu Lys Ser Leu His Arg Leu His Ile Asn Asp Glu Thr
 100 105 110

Gln Leu Lys Asn Ala Lys Ile Pro Thr Asn Asp Thr Thr Asp Tyr Met
 115 120 125

Pro Pro Ser Asp Gly Ala Asn Glu Val Thr Arg Ile Asp Leu Lys Asp
 130 135 140
 Ile Lys Ser Pro Thr Arg His His Lys Arg Arg Pro Thr Thr Ile Asp
 145 150 155 160
 Val Pro Gly Leu Thr Lys Ser Lys Thr Ser Pro Asp Gly Leu Ile Ser
 165 170 175
 Lys Glu Asp Ser Gly Ser Lys Leu Val Ile Val Met Val Gly Leu Pro
 180 185 190
 Ala Thr Gly Lys Ser Phe Ile Thr Asn Lys Leu Ser Arg Phe Leu Asn
 195 200 205
 Tyr Ser Leu Tyr Tyr Cys Lys Val Phe Asn Val Gly Asn Thr Arg Arg
 210 215 220
 Lys Phe Ala Lys Glu His Gly Leu Lys Asp Gln Asp Ser Lys Phe Phe
 225 230 235 240
 Glu Pro Lys Asn Ala Asp Ser Thr Arg Leu Arg Asp Lys Trp Ala Met
 245 250 255
 Asp Thr Leu Asp Glu Leu Leu Asp Tyr Leu Leu Glu Gly Ser Gly Ser
 260 265 270
 Val Gly Ile Phe Asp Ala Thr Asn Thr Ser Arg Glu Arg Arg Lys Asn
 275 280 285
 Val Leu Ala Arg Ile Arg Lys Arg Ser Pro His Leu Lys Val Leu Phe
 290 295 300
 Leu Glu Ser Val Cys Ser Asp His Ala Leu Val Gln Lys Asn Ile Arg
 305 310 315 320
 Leu Lys Leu Phe Gly Pro Asp Tyr Lys Gly Lys Asp Pro Glu Ser Ser
 325 330 335
 Leu Lys Asp Phe Lys Ser Arg Leu Ala Asn Tyr Leu Lys Ala Tyr Glu
 340 345 350
 Pro Ile Glu Asp Asp Glu Asn Leu Gln Tyr Ile Lys Met Ile Asp Val
 355 360 365
 Gly Lys Lys Val Ile Ala Tyr Asn Ile Gln Gly Phe Leu Ala Ser Gln
 370 375 380
 Thr Val Tyr Tyr Leu Leu Asn Phe Asn Leu Ala Asp Arg Gln Ile Trp
 385 390 395 400
 Ile Thr Arg Ser Gly Glu Ser Glu Asp Asn Val Ser Gly Arg Ile Gly
 405 410 415
 Gly Asn Ser His Leu Thr Pro Arg Gly Leu Arg Phe Ala Lys Ser Leu
 420 425 430
 Pro Lys Phe Ile Ala Arg Gln Arg Glu Ile Phe Tyr Gln Asn Leu Met
 435 440 445

Gln Gln Lys Lys Asn Asn Glu Asn Thr Asp Gly Asn Ile Tyr Asn Asp
 450 455 460
 Phe Phe Val Trp Thr Ser Met Arg Ala Arg Thr Ile Gly Thr Ala Gln
 465 470 475 480
 Tyr Phe Asn Glu Asp Asp Tyr Pro Ile Lys Gln Met Lys Met Leu Asp
 485 490 495
 Glu Leu Ser Ala Gly Asp Tyr Asp Gly Met Thr Tyr Pro Glu Ile Lys
 500 505 510
 Asn Asn Phe Pro Glu Glu Phe Glu Lys Arg Gln Lys Asp Lys Leu Arg
 515 520 525
 Tyr Arg Tyr Pro Gly Ile Gly Gly Glu Ser Tyr Met Asp Val Ile Asn
 530 535 540
 Arg Leu Arg Pro Val Ile Thr Glu Leu Glu Arg Ile Glu Asp Asn Val
 545 550 555 560
 Leu Ile Ile Thr His Arg Val Val Ala Arg Ala Leu Leu Gly Tyr Phe
 565 570 575
 Met Asn Leu Ser Met Gly Ile Ile Ala Asn Leu Asp Val Pro Leu His
 580 585 590
 Cys Val Tyr Cys Leu Glu Pro Lys Pro Tyr Gly Ile Thr Trp Ser Leu
 595 600 605
 Trp Glu Tyr Asp Glu Ala Ser Asp Ser Phe Ser Lys Val Pro Gln Thr
 610 615 620
 Asp Leu Asn Thr Thr Arg Val Lys Glu Val Gly Leu Val Tyr Asn Glu
 625 630 635 640
 Arg Arg Tyr Ser Val Ile Pro Thr Ala Pro Pro Ser Ala Arg Ser Ser
 645 650 655
 Phe Ala Ser Asp Phe Leu Ser Arg Lys Arg Ser Asn Pro Thr Ser Ala
 660 665 670
 Ser Ser Ser Gln Ser Glu Leu Ser Glu Gln Pro Lys Asn Ser Val Ser
 675 680 685
 Ala Gln Thr Gly Ser Asn Asn Thr Thr Leu Ile Gly Ser Asn Phe Asn
 690 695 700
 Ile Lys Asn Glu Asn Gly Asp Ser Arg Ile Pro Leu Ser Ala Pro Leu
 705 710 715 720
 Met Ala Thr Asn Thr Ser Asn Asn Ile Leu Asp Gly Gly Gly Thr Ser
 725 730 735
 Ile Ser Ile His Arg Pro Arg Val Val Pro Asn Gln Asn Asn Val Asn
 740 745 750

Pro Leu Leu Ala Asn Asn Asn Lys Ala Ala Ser Asn Val Pro Asn Val
755 760 765

Lys Lys Ser Ala Ala Thr Pro Arg Gln Ile Phe Glu Ile Asp Lys Val
770 775 780

Asp Glu Lys Leu Ser Met Leu Lys Asn Lys Ser Phe Leu Leu His Gly
785 790 795 800

Lys Asp Tyr Pro Asn Asn Ala Asp Asn Asn Asp Asn Glu Asp Ile Arg
805 810 815

Ala Lys Thr Met Asn Arg Ser Gln Ser His Val
820 825

<210> 137

<211> 1665

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 137

atgtccaata actcattcac taacttcaaa ctggccactg aattgccagc ctggtctaag	60
ttgcaaaaaa tttatgaatc tcaaggtaag actttgtctg tcaagcaaga attccaaaaa	120
gatgccaaagc gttttgaaaa attgaacaag actttcacca actatgatgg ttccaaaatc	180
ttgttcgact actcaaagaa cttggtcaac gatgaaatca ttgctgcatt gattgaactg	240
gccaaggagg ctaacgtcac cggtttgaga gatgctatgt tcaaaggatga acacatcaac	300
tccactgaag atcgtgctgt ctaccacgtc gcattgagaa acagagctaa caagccaatg	360
tacgttgatg gtgtcaacgt tgctccagaa gtcgactctg tcttgaagca catgaaggag	420
ttctctgaac aagttcgttc tggatgaatgg aagggttata ccggtgaagaa gatcaccgat	480
gttggttaaca tcggtattgg tgggtccgat ttgggtccag tcatgggtcac tgaggctttg	540
aagcactacg ctggtgtcctt ggatgtccac ttcgtttcca acattgacgg tactcacatt	600
gctgaaacct tgaaggttgt tgaccagaa actactttgt ttttgattgc ttccaagact	660
ttcactaccg ctgaaactat cactaacgtc aacactgcc aagaactggt cttgtcgaag	720
acaggtaatg atccatctca cattgctaag ctttcgctg ctttgcacac taacgaaacc	780
gaagttgcc aagtcggtat tgacacaaaa aacatgtttg gtttcgaaag ttgggtcggg	840
ggtcgttact ctgtctggtc ggctattggg ttgtctgttg ccttgatcat tggctatgac	900
aactttgagg ctttcttgaa ggggtctgaa gccgtcgaca accacttcac ccaaacccca	960
ttggaagaca acattccatt gttgggtggg ttgtgtctg tctggtacaa caacttcttt	1020
ggtgctcaaa cccatttggt tgctccattc gaccaatact tgcacagatt cccagcctac	1080
ttgcaacaat tgtcaatgga atctaacggg aagtcgttta ccagaggtaa cgtgtttact	1140
gactactcta ctggttctat cttgtttggg gaaccagcta ccaacgctca aactctttc	1200
ttccaattgg ttcaccaagg taccaagttg attccatctg atttcatctt agctgctcaa	1260
tctcataacc caattgagaa caaattacat caaaagatgt tggcttcaaa cttctttgct	1320
caagctgaag ctttaattggg tggttaaggat gaagaacaag ttaaggctga aggtgccact	1380

```

ggtaggtttgg tcccacacaa ggtcttctca ggtaacagac caactacctc tatcttggtc 1440
caaaagatta ctccagctac ttgggtgct ttgattgcct actacgaaca tgttactttc 1500

actgaaggtg ccatttgga tatcaactct ttcgaccaat ggggtgttga attgggtaaa 1560
gtcttggtta aagtcacgga caaggaattg gacaactcct ccaccatttc taccacgat 1620
gcttctacca acggtttaat caatcaattc aaggaatgga tgtga 1665

```

<210> 138

<211> 554

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 138

```

Met Ser Asn Asn Ser Phe Thr Asn Phe Lys Leu Ala Thr Glu Leu Pro
1          5          10          15

Ala Trp Ser Lys Leu Gln Lys Ile Tyr Glu Ser Gln Gly Lys Thr Leu
20          25          30

Ser Val Lys Gln Glu Phe Gln Lys Asp Ala Lys Arg Phe Glu Lys Leu
35          40          45

Asn Lys Thr Phe Thr Asn Tyr Asp Gly Ser Lys Ile Leu Phe Asp Tyr
50          55          60

Ser Lys Asn Leu Val Asn Asp Glu Ile Ile Ala Ala Leu Ile Glu Leu
65          70          75          80

Ala Lys Glu Ala Asn Val Thr Gly Leu Arg Asp Ala Met Phe Lys Gly
85          90          95

Glu His Ile Asn Ser Thr Glu Asp Arg Ala Val Tyr His Val Ala Leu
100         105         110

Arg Asn Arg Ala Asn Lys Pro Met Tyr Val Asp Gly Val Asn Val Ala
115         120         125

Pro Glu Val Asp Ser Val Leu Lys His Met Lys Glu Phe Ser Glu Gln
130         135         140

Val Arg Ser Gly Glu Trp Lys Gly Tyr Thr Gly Lys Lys Ile Thr Asp
145         150         155         160

Val Val Asn Ile Gly Ile Gly Gly Ser Asp Leu Gly Pro Val Met Val
165         170         175

Thr Glu Ala Leu Lys His Tyr Ala Gly Val Leu Asp Val His Phe Val
180         185         190

Ser Asn Ile Asp Gly Thr His Ile Ala Glu Thr Leu Lys Val Val Asp
195         200         205

Pro Glu Thr Thr Leu Phe Leu Ile Ala Ser Lys Thr Phe Thr Thr Ala
210         215         220

```

```

---
Glu Thr Ile Thr Asn Ala Asn Thr Ala Lys Asn Trp Phe Leu Ser Lys
225                230                235                240

Thr Gly Asn Asp Pro Ser His Ile Ala Lys His Phe Ala Ala Leu Ser
                245                250                255

Thr Asn Glu Thr Glu Val Ala Lys Phe Gly Ile Asp Thr Lys Asn Met
                260                265                270

Phe Gly Phe Glu Ser Trp Val Gly Gly Arg Tyr Ser Val Trp Ser Ala
                275                280                285

Ile Gly Leu Ser Val Ala Leu Tyr Ile Gly Tyr Asp Asn Phe Glu Ala
                290                295                300

Phe Leu Lys Gly Ala Glu Ala Val Asp Asn His Phe Thr Gln Thr Pro
305                310                315                320

Leu Glu Asp Asn Ile Pro Leu Leu Gly Gly Leu Leu Ser Val Trp Tyr
                325                330                335

Asn Asn Phe Phe Gly Ala Gln Thr His Leu Val Ala Pro Phe Asp Gln
                340                345                350

Tyr Leu His Arg Phe Pro Ala Tyr Leu Gln Gln Leu Ser Met Glu Ser
                355                360                365

Asn Gly Lys Ser Val Thr Arg Gly Asn Val Phe Thr Asp Tyr Ser Thr
                370                375                380

Gly Ser Ile Leu Phe Gly Glu Pro Ala Thr Asn Ala Gln His Ser Phe
385                390                395                400

Phe Gln Leu Val His Gln Gly Thr Lys Leu Ile Pro Ser Asp Phe Ile
                405                410                415

Leu Ala Ala Gln Ser His Asn Pro Ile Glu Asn Lys Leu His Gln Lys
                420                425                430

Met Leu Ala Ser Asn Phe Phe Ala Gln Ala Glu Ala Leu Met Val Gly
                435                440                445

Lys Asp Glu Glu Gln Val Lys Ala Glu Gly Ala Thr Gly Gly Leu Val
                450                455                460

Pro His Lys Val Phe Ser Gly Asn Arg Pro Thr Thr Ser Ile Leu Ala
465                470                475                480

Gln Lys Ile Thr Pro Ala Thr Leu Gly Ala Leu Ile Ala Tyr Tyr Glu
                485                490                495

His Val Thr Phe Thr Glu Gly Ala Ile Trp Asn Ile Asn Ser Phe Asp
                500                505                510

Gln Trp Gly Val Glu Leu Gly Lys Val Leu Ala Lys Val Ile Gly Lys
                515                520                525

```

Glu Leu Asp Asn Ser Ser Thr Ile Ser Thr His Asp Ala Ser Thr Asn
530 535 540

Gly Leu Ile Asn Gln Phe Lys Glu Trp Met
545 550

<210> 139

<211> 744

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 139

```

atgccaaagt tagttttagt tagacacggt caatccgaat ggaacgaaaa gaacttattc      60
accggttggg ttgatgttaa attgtctgcc aagggtcaac aagaagccgc tagagccggt      120
gaattgttga aggaaaagaa ggtctacca gacgtcttgt acacttccaa gttgtccaga      180
gctatccaaa ctgctaacat tgctttggaa aaggctgaca gattatggat tccagtcaac      240
agatcctgga gattgaacga aagacattac ggtgacttac aaggaagga caaggctgaa      300
actttgaaga agttcgggtga agaaaaattc aacacctaca gaagatcctt cgatgttcca      360
cctcccccaa tcgacgcttc ttctccattc tctcaaaagg gtgatgaaag atacaagtac      420
gttgacccaa atgtcttgcc agaaactgaa tctttggcct tggtcattga cagattgttg      480
ccatactggc aagatgtcat tgccaaggac ttgttgagtg gtaagaccgt catgatcgcc      540
gctcacggta actccttgag aggtttggtt aagcacttgg aaggatatct tgatgctgac      600
attgctaagt tgaacatccc aactgggtatt ccattggctc tcgaattgga cgaaaacttg      660
aagccatcta agccatctta ctacttgac ccagaagctg ccgctgctgg tgccgctgct      720
gttgccaacc aaggtaagaa ataa                                     744

```

<210> 140

<211> 247

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 140

```

Met Pro Lys Leu Val Leu Val Arg His Gly Gln Ser Glu Trp Asn Glu
1          5          10          15

Lys Asn Leu Phe Thr Gly Trp Val Asp Val Lys Leu Ser Ala Lys Gly
20          25          30

Gln Gln Glu Ala Ala Arg Ala Gly Glu Leu Leu Lys Glu Lys Lys Val
35          40          45

Tyr Pro Asp Val Leu Tyr Thr Ser Lys Leu Ser Arg Ala Ile Gln Thr
50          55          60

Ala Asn Ile Ala Leu Glu Lys Ala Asp Arg Leu Trp Ile Pro Val Asn
65          70          75          80

Arg Ser Trp Arg Leu Asn Glu Arg His Tyr Gly Asp Leu Gln Gly Lys
85          90          95

```

Asp Lys Ala Glu Thr Leu Lys Lys Phe Gly Glu Glu Lys Phe Asn Thr
 100 105 110
 Tyr Arg Arg Ser Phe Asp Val Pro Pro Pro Pro Ile Asp Ala Ser Ser
 115 120 125
 Pro Phe Ser Gln Lys Gly Asp Glu Arg Tyr Lys Tyr Val Asp Pro Asn
 130 135 140
 Val Leu Pro Glu Thr Glu Ser Leu Ala Leu Val Ile Asp Arg Leu Leu
 145 150 155 160
 Pro Tyr Trp Gln Asp Val Ile Ala Lys Asp Leu Leu Ser Gly Lys Thr
 165 170 175
 Val Met Ile Ala Ala His Gly Asn Ser Leu Arg Gly Leu Val Lys His
 180 185 190
 Leu Glu Gly Ile Ser Asp Ala Asp Ile Ala Lys Leu Asn Ile Pro Thr
 195 200 205
 Gly Ile Pro Leu Val Phe Glu Leu Asp Glu Asn Leu Lys Pro Ser Lys
 210 215 220
 Pro Ser Tyr Tyr Leu Asp Pro Glu Ala Ala Ala Ala Gly Ala Ala Ala
 225 230 235 240
 Val Ala Asn Gln Gly Lys Lys
 245

<210> 141

<211> 747

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 141

atggctagaa ctttctttgt cgggtgtaac tttaaattaa acggttccaa acaatccatt	60
aaggaaattg ttgaaagatt gaacactgct tctatcccag aaaatgtcga agttgttatt	120
tgtcctccag ctacctactt agactactct gtctctttgg ttaagaagcc acaagtcact	180
gtcggtgctc aaaacgccta cttgaaggct tctggtgctt tcaccggtga aaactccgtt	240
gaccaaataca aggatgttgg tgctaagtgg gttattttgg gtcactccga aagaagatct	300
tacttccacg aagatgacaa gttcattgct gacaagacca agttcgcttt aggtcaagg	360
gtcggtgtca tcttggtgat cggtgaaact ttggaagaaa agaaggccgg taagactttg	420
gatgttggtg aaagacaatt gaacgctgtc ttggaagaag ttaaggactg gactaacgtc	480
gttgtcgctt acgaaccagt ctgggccatt ggtaccgggt tggctgctac tccagaagat	540
gctcaagata ttcacgcttc catcagaaag ttcttggctt ccaagttggg tgacaaggct	600
gccagcgaat tgagaatctt atacggtggg tccgctaacg gtagcaacgc cgttaccttc	660
aaggacaagg ctgatgtcga tggtttcttg gtcggtggtg cttctttgaa gccagaattt	720
gttgatatca tcaactctag aaactaa	747

<210> 142

<211> 248

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 142

Met Ala Arg Thr Phe Phe Val Gly Gly Asn Phe Lys Leu Asn Gly Ser
 1 5 10 15

Lys Gln Ser Ile Lys Glu Ile Val Glu Arg Leu Asn Thr Ala Ser Ile
 20 25 30

Pro Glu Asn Val Glu Val Val Ile Cys Pro Pro Ala Thr Tyr Leu Asp
 35 40 45

Tyr Ser Val Ser Leu Val Lys Lys Pro Gln Val Thr Val Gly Ala Gln
 50 55 60

Asn Ala Tyr Leu Lys Ala Ser Gly Ala Phe Thr Gly Glu Asn Ser Val

65 70 75 80

Asp Gln Ile Lys Asp Val Gly Ala Lys Trp Val Ile Leu Gly His Ser
 85 90 95

Glu Arg Arg Ser Tyr Phe His Glu Asp Asp Lys Phe Ile Ala Asp Lys
 100 105 110

Thr Lys Phe Ala Leu Gly Gln Gly Val Gly Val Ile Leu Cys Ile Gly
 115 120 125

Glu Thr Leu Glu Glu Lys Lys Ala Gly Lys Thr Leu Asp Val Val Glu
 130 135 140

Arg Gln Leu Asn Ala Val Leu Glu Glu Val Lys Asp Trp Thr Asn Val
 145 150 155 160

Val Val Ala Tyr Glu Pro Val Trp Ala Ile Gly Thr Gly Leu Ala Ala
 165 170 175

Thr Pro Glu Asp Ala Gln Asp Ile His Ala Ser Ile Arg Lys Phe Leu
 180 185 190

Ala Ser Lys Leu Gly Asp Lys Ala Ala Ser Glu Leu Arg Ile Leu Tyr
 195 200 205

Gly Gly Ser Ala Asn Gly Ser Asn Ala Val Thr Phe Lys Asp Lys Ala
 210 215 220

Asp Val Asp Gly Phe Leu Val Gly Gly Ala Ser Leu Lys Pro Glu Phe
 225 230 235 240

Val Asp Ile Ile Asn Ser Arg Asn
 245

<210> 143

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 143

ctagaactag tggatcccc atgaacgata gccaaaactg c 41

<210> 144

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 144

atatcgaatt cctgcagccc ttattggggg gaagtgtatt g 41

<210> 145

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 145

ctagaactag tggatcccc atgttgcaag gaattttaga aaccg 45

<210> 146

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Artificial DNA Primer

<400> 146

atatcgaatt cctgcagccc tcaaaatgtt gtaactatat tcatttcac tg 52

<210> 147

<211> 3426

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 147

atgtcacttc taatagattc tgtaccaaca gttgcttata aggacccaaa accgggtact	60
tcagggtttac gtaagaagac caagggttttc atggatgagc ctcattatac tgagaacttc	120
attcaagcaa caatgcaatc tatccctaata ggctcagagg gaaccacttt agttgttgga	180
ggagatggtc gtttctacaa cgatgttatc atgaacaaga ttgccgcagt agtgctgca	240
aacgggtgca gaaagttagt cattgggtcaa ggcgggttac tttcaacacc agctgcttct	300
catataatta gaacatacga ggaaaagtgt accgggtggg gtatcatatt aactgcctca	360
cacaacccag gcgggtccaga gaatgattta ggtatcaagt ataatttacc taatgggtggg	420
ccagctccag agagtgtcac taacgctatc tgggaagcgt ctaaaaaatt aactcactat	480
aaaattataa agaacttccc caagttgaat ttgaacaagc ttggtaaaaa ccaaaaatat	540
ggcccattgt tagtggacat aattgatcct gccaaagcat acgttcaatt tctgaaggaa	600
atttttgatt ttgacttaat taaaagcttc ttagcgaaac agcgcaaaga caaagggtgg	660
aagttgttgt ttgactcctt aaatgggtatt acaggaccat atggtaaggc tatatttggt	720
gatgaatttg gtttaccggc agaggaagtt cttcaaaatt ggcacccttt acctgatttc	780
ggcgggttac atcccgatcc gaatctaacc tatgcacgaa ctcttggtga cagggttgac	840
cgcgaaaaaa ttgccttttg agcagcctcc gatggtgatg gtgataggaa tatgatttac	900
ggttatggcc ctgctttogt ttogccaggt gattctgttg ccattattgc cgaatatgca	960
cccgaatttc catacttgcg caaacaaggt atttatggct tggcacgttc atttctaca	1020
tcctcagcca ttgatcgtgt tgcagcaaaa aagggattaa gatgttacga agttccaacc	1080
ggctggaaat tcttctgtgc cttatttgat gctaaaaagc tatcaatctg tgggaagaa	1140
tccttcggta caggttccaa tcatatcaga gaaaaggacg gtctatgggc cattattgct	1200
tgggttaaata tcttggctat ctaccatagg cgtaaccctg aaaaggaagc ttcgatcaaa	1260
actattcagg acgaattttg gaacgagtat ggcggtactt tcttcacaag atacgattac	1320
gaacatatcg aatgcgagca ggccgaaaaa gttgtagctc ttttgagtga atttgtatca	1380
aggccaaacg tttgtggctc ccacttccca gctgatgagt ctttaaccgt tatcgattgt	1440
ggtgattttt cgtatagaga tctagatggc tccatctctg aaaatcaagg ccttttcgta	1500
aagttttcga atgggactaa atttgttttg aggttatccg gcacaggcag ttctgggtgca	1560
acaataagat tatacgtaga aaagtatact gataaaaagg agaactatgg ccaaacagct	1620
gacgtcttct tgaaaccogt catcaactcc attgtaaaat tcttaagatt taaagaaatt	1680
ttaggaacag acgaaccaac agtccgcaca tagatgtcac ttctaataga ttctgtacca	1740
acagttgctt ataaggacca aaaaccgggt acttcagggt tacgtaagaa gaccaagggt	1800
ttcatggatg agcctcatta tactgagaac ttcatccaag caacaatgca atctatccct	1860
aatggctcag agggaaccac tttagttggt ggaggagatg gtcgtttcta caacgatgtt	1920
atcatgaaca agattgcgcg agtaggtgct gcaaaccgtg tcagaaagt agtcattggt	1980
caaggcgggt tactttcaac accagctgct tctcatataa ttagaacata cgaggaaaag	2040
tgtaccgggt gtggtatcat attaactgcc tcacacaacc caggcgggtcc agagaatgat	2100
ttaggtatca agtataattt acctaatggt gggccagctc cagagagtgt cactaacgct	2160
atctgggaag cgtctaaaaa attaactcac tataaaatta taaagaactt cccaagtgtg	2220
aatttgaaca agcttggtta aaacccaaaa tatggcccat tgtagtgga cataattgat	2280

cctgccaaag catacgttca atttctgaag gaaatTTTTg attttgactt aattaaaagc 2340
 ttcttagcga aacagcgcaa agacaaaagg tggaagttgt tgtttgactc cttaaattgt 2400
 attacaggac catatggtta ggctatatTT gtgatgaat ttggtttacc ggcagaggaa 2460
 gttcttcaaa attggcacc tttacctgat ttggcggtt tacatccga tccgaatcta 2520
 acctatgcac gaactcttgt tgacagggtt gaccgcgaaa aaattgcctt tggagcagcc 2580
 tccgatggtg atggtgatag gaatatgatt tacggttatg gccctgcttt cgtttcgcca 2640

 ggtgattctg ttgccattat tgccgaatat gcaccgaaa ttccatactt cgccaaacaa 2700
 ggtatttatg gcttggcacg ttcatttcct acatcctcag ccattgatcg tgttgcagca 2760
 aaaaagggat taagatgta cgaagtcca accggctgga aattcttctg tgccttattt 2820
 gatgctaaaa agctatcaat ctgtggtgaa gaatccttcg gtacaggttc caatcatatc 2880
 agagaaaagg acggtctatg ggccattatt gcttgggtta atatcttggc tatctaccat 2940
 aggcgtaacc ctgaaaagga agcttcgatc aaaactattc aggacgaatt ttggaacgag 3000
 tatggccgta ctttcttcac aagatacgat tacgaacata tcgaatgcga gcaggccgaa 3060
 aaagtgttag ctcttttgag tgaatttgta tcaaggccaa acgtttgtgg ctcccacttc 3120
 ccagctgatg agtccttaac cgttatcgat tgtggtgatt tttcgtatag agatctagat 3180
 ggctccatct ctgaaaatca aggccttttc gtaaagtgtt cgaatgggac taaatttggt 3240
 ttgaggttat ccggcacagg cagttctggt gcaacaataa gattatacgt agaaaagtat 3300
 actgataaaa aggagaacta tggccaaaca gctgacgtct tcttgaaacc cgtcatcaac 3360
 tccattgtaa aattcttaag atttaaagaa attttaggaa cagacgaacc aacagtccgc 3420
 acatag 3426

<210> 148

<211> 570

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 148

Met Ser Leu Leu Ile Asp Ser Val Pro Thr Val Ala Tyr Lys Asp Gln
 1 5 10 15

 Lys Pro Gly Thr Ser Gly Leu Arg Lys Lys Thr Lys Val Phe Met Asp
 20 25 30

 Glu Pro His Tyr Thr Glu Asn Phe Ile Gln Ala Thr Met Gln Ser Ile
 35 40 45

 Pro Asn Gly Ser Glu Gly Thr Thr Leu Val Val Gly Gly Asp Gly Arg
 50 55 60

 Phe Tyr Asn Asp Val Ile Met Asn Lys Ile Ala Ala Val Gly Ala Ala
 65 70 75 80

 Asn Gly Val Arg Lys Leu Val Ile Gly Gln Gly Gly Leu Leu Ser Thr
 85 90 95

 Pro Ala Ala Ser His Ile Ile Arg Thr Tyr Glu Glu Lys Cys Thr Gly
 100 105 110

Gly Gly Ile Ile Leu Thr Ala Ser His Asn Pro Gly Gly Pro Glu Asn
 115 120 125

Asp Leu Gly Ile Lys Tyr Asn Leu Pro Asn Gly Gly Pro Ala Pro Glu
 130 135 140

Ser Val Thr Asn Ala Ile Trp Glu Ala Ser Lys Lys Leu Thr His Tyr
 145 150 155 160

Lys Ile Ile Lys Asn Phe Pro Lys Leu Asn Leu Asn Lys Leu Gly Lys
 165 170 175

Asn Gln Lys Tyr Gly Pro Leu Leu Val Asp Ile Ile Asp Pro Ala Lys
 180 185 190

Ala Tyr Val Gln Phe Leu Lys Glu Ile Phe Asp Phe Asp Leu Ile Lys
 195 200 205

Ser Phe Leu Ala Lys Gln Arg Lys Asp Lys Gly Trp Lys Leu Leu Phe
 210 215 220

Asp Ser Leu Asn Gly Ile Thr Gly Pro Tyr Gly Lys Ala Ile Phe Val
 225 230 235 240

Asp Glu Phe Gly Leu Pro Ala Glu Glu Val Leu Gln Asn Trp His Pro
 245 250 255

Leu Pro Asp Phe Gly Gly Leu His Pro Asp Pro Asn Leu Thr Tyr Ala
 260 265 270

Arg Thr Leu Val Asp Arg Val Asp Arg Glu Lys Ile Ala Phe Gly Ala
 275 280 285

Ala Ser Asp Gly Asp Gly Asp Arg Asn Met Ile Tyr Gly Tyr Gly Pro
 290 295 300

Ala Phe Val Ser Pro Gly Asp Ser Val Ala Ile Ile Ala Glu Tyr Ala
 305 310 315 320

Pro Glu Ile Pro Tyr Phe Ala Lys Gln Gly Ile Tyr Gly Leu Ala Arg
 325 330 335

Ser Phe Pro Thr Ser Ser Ala Ile Asp Arg Val Ala Ala Lys Lys Gly
 340 345 350

Leu Arg Cys Tyr Glu Val Pro Thr Gly Trp Lys Phe Phe Cys Ala Leu
 355 360 365

Phe Asp Ala Lys Lys Leu Ser Ile Cys Gly Glu Glu Ser Phe Gly Thr
 370 375 380

Gly Ser Asn His Ile Arg Glu Lys Asp Gly Leu Trp Ala Ile Ile Ala
 385 390 395 400

Trp Leu Asn Ile Leu Ala Ile Tyr His Arg Arg Asn Pro Glu Lys Glu
 405 410 415

Ala Ser Ile Lys Thr Ile Gln Asn Glu Phe Trp Asn Glu Tyr Glu Arg

```

-----
      420              425              430
Thr Phe Phe Thr Arg Tyr Asp Tyr Glu His Ile Glu Cys Glu Gln Ala
  435              440              445

Glu Lys Val Val Ala Leu Leu Ser Glu Phe Val Ser Arg Pro Asn Val
  450              455              460

Cys Gly Ser His Phe Pro Ala Asp Glu Ser Leu Thr Val Ile Asp Cys
  465              470              475              480

Gly Asp Phe Ser Tyr Arg Asp Leu Asp Gly Ser Ile Ser Glu Asn Gln
              485              490              495

Gly Leu Phe Val Lys Phe Ser Asn Gly Thr Lys Phe Val Leu Arg Leu
              500              505              510

Ser Gly Thr Gly Ser Ser Gly Ala Thr Ile Arg Leu Tyr Val Glu Lys
              515              520              525

Tyr Thr Asp Lys Lys Glu Asn Tyr Gly Gln Thr Ala Asp Val Phe Leu
  530              535              540

Lys Pro Val Ile Asn Ser Ile Val Lys Phe Leu Arg Phe Lys Glu Ile
  545              550              555              560

Leu Gly Thr Asp Glu Pro Thr Val Arg Thr
              565              570

```

<210> 149

<211> 1869

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 149

```

atgttgcaag gaattttaga aaccgtacca tctgacttga aagatccgat atcattatgg      60
ttaagcaag accgcaaccc aaaaactata gaagagggtca ccgctctctg caaaaaatcc      120

gactggaatg agttacacaa aagatttgat tctagaattc agtttggcac tgctggttta      180
agatcgcaaa tgcaagctgg cttagcagg atgaatactt tagtagtcat acaagcgtct      240
cagggattgg caacttatgt aagacaacag tttccagaca atttggtagc tgttgtggga      300
cacgatcata gattccattc taaggagttc gctagagcta ctgctgctgc atttctttta      360
aaaggattta aggtacatta tttgaatcct gaccacgaat ttgttcatac ccctttagtt      420
ccctttgcag tggataagct aaaggcctcc gttggcgtaa tgataacagc aagtcacaac      480
ccaaaaatgg ataatggata taaagtatac tattccaatg gatgccaaat cattccacct      540
cacgatcatg ccatctctga ttccattgac gcaaatttag aaccatgggc caatgtgtgg      600
gatttcgacg atgttctaaa taaggctctc aaacaaggga aattgatgta ttccagagaa      660
gaaatgctga agttatatatt agaggagggt tctaaaaatc tggtagaaat caaccatta      720
aagcttgaag taaaagccaa accttgggtc gtttactc caatgcatgg ggttggattt      780
gacattttca gcaccatcgt aaaaaaaca ctgtgcctgg tagaaggtaa ggattaccta      840

```

```

tgtgttcctg aacaacaaaa tccagatcct tctttcccaa ctgttgatt tcctaaccct      900
gaagaaaaag gtgctttaga cattgggtata aacttggtg aaaaacatga cattgactta      960
cttggtgcca acgaccctga cgctgataga ttctctgttg ctgttaaaga tatgcagtca     1020
ggcgaatggc gacaactaac aggtaacgaa atcggttttc tttttgcatt ttatgaatat     1080
cagaaatata aaagtatgga caaagaattt cagcacgttc atccgttggc tatgttaaata     1140
tcaacagtgt cttcacaaat gataaaaaaa atggcagaaa tagaagggtt ccattatgag     1200
gatacattaa caggatttaa gtggatcgga aatcgtgcca tactcttgga aaagaaaggc     1260
tattacgttc cttttggatt cgaggaagca ataggctaca tgtttccagc aatggagcat     1320
gataaggatg gtatcagtgc atccattgtc ttcttgcaag cctactgtaa gtggaaaata     1380
gaccacaatt tggaccgct aaatgtctta gaaaatggct tcaaaaaata tggcgtgttc     1440
aaagagtaca atggctatta tgcgttcca aatccaactg ttacaaaaga tatatttgac     1500
tacatcagga atgtctacac tcctgagggc ggcgtcatatc cttcatctat tggagaagaa     1560
atcgaagtac tttactatcg agatttaacc actggttacc aatcggatac cataaatcat     1620
aaacctactc taccgctcga tcctacatca caaatgataa cagtatctgc tagaccaagt     1680
aacggtagtg agaataagca tatccgcttc actattcgcg ggtccggaac agaaccacaaa     1740
cttaaagtat atattgaagc ttgcgcaaat gaagaacaaa gagcctcttt cttggcgaaa     1800
ttgacttgga atgtgctgag acgtgaatgg tttagaccag atgaaatgaa tatagttaca     1860
aaattttga                                     1869

```

<210> 150

<211> 622

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 150

```

Met Leu Gln Gly Ile Leu Glu Thr Val Pro Ser Asp Leu Lys Asp Pro
1           5           10          15

Ile Ser Leu Trp Phe Lys Gln Asp Arg Asn Pro Lys Thr Ile Glu Glu
20          25          30

Val Thr Ala Leu Cys Lys Lys Ser Asp Trp Asn Glu Leu His Lys Arg
35          40          45

Phe Asp Ser Arg Ile Gln Phe Gly Thr Ala Gly Leu Arg Ser Gln Met
50          55          60

Gln Ala Gly Phe Ser Arg Met Asn Thr Leu Val Val Ile Gln Ala Ser
65          70          75          80

Gln Gly Leu Ala Thr Tyr Val Arg Gln Gln Phe Pro Asp Asn Leu Val
85          90          95

Ala Val Val Gly His Asp His Arg Phe His Ser Lys Glu Phe Ala Arg
100         105         110

Ala Thr Ala Ala Ala Phe Leu Leu Lys Gly Phe Lys Val His Tyr Leu

```

```

-----
115                               120                               125

Asn Pro Asp His Glu Phe Val His Thr Pro Leu Val Pro Phe Ala Val
130                               135                               140

Asp Lys Leu Lys Ala Ser Val Gly Val Met Ile Thr Ala Ser His Asn
145                               150                               155                               160

Pro Lys Met Asp Asn Gly Tyr Lys Val Tyr Tyr Ser Asn Gly Cys Gln
165                               170                               175

Ile Ile Pro Pro His Asp His Ala Ile Ser Asp Ser Ile Asp Ala Asn
180                               185                               190

Leu Glu Pro Trp Ala Asn Val Trp Asp Phe Asp Asp Val Leu Asn Lys
195                               200                               205

Ala Leu Lys Gln Gly Lys Leu Met Tyr Ser Arg Glu Glu Met Leu Lys
210                               215                               220

Leu Tyr Leu Glu Glu Val Ser Lys Asn Leu Val Glu Ile Asn Pro Leu
225                               230                               235                               240

Lys Leu Glu Val Lys Ala Lys Pro Trp Phe Val Tyr Thr Pro Met His
245                               250                               255

Gly Val Gly Phe Asp Ile Phe Ser Thr Ile Val Lys Lys Thr Leu Cys
260                               265                               270

Leu Val Glu Gly Lys Asp Tyr Leu Cys Val Pro Glu Gln Gln Asn Pro
275                               280                               285

Asp Pro Ser Phe Pro Thr Val Gly Phe Pro Asn Pro Glu Glu Lys Gly
290                               295                               300

Ala Leu Asp Ile Gly Ile Asn Leu Ala Glu Lys His Asp Ile Asp Leu
305                               310                               315                               320

Leu Val Ala Asn Asp Pro Asp Ala Asp Arg Phe Ser Val Ala Val Lys
325                               330                               335

Asp Met Gln Ser Gly Glu Trp Arg Gln Leu Thr Gly Asn Glu Ile Gly
340                               345                               350

Phe Leu Phe Ala Phe Tyr Glu Tyr Gln Lys Tyr Lys Ser Met Asp Lys
355                               360                               365

Glu Phe Gln His Val His Pro Leu Ala Met Leu Asn Ser Thr Val Ser
370                               375                               380

Ser Gln Met Ile Lys Lys Met Ala Glu Ile Glu Gly Phe His Tyr Glu
385                               390                               395                               400

Asp Thr Leu Thr Gly Phe Lys Trp Ile Gly Asn Arg Ala Ile Leu Leu
405                               410                               415

```

Glu Lys Lys Gly Tyr Tyr Val Pro Phe Gly Phe Glu Glu Ala Ile Gly
 420 425 430

Tyr Met Phe Pro Ala Met Glu His Asp Lys Asp Gly Ile Ser Ala Ser
 435 440 445

Ile Val Phe Leu Gln Ala Tyr Cys Lys Trp Lys Ile Asp His Asn Leu
 450 455 460

Asp Pro Leu Asn Val Leu Glu Asn Gly Phe Lys Lys Tyr Gly Val Phe
 465 470 475 480

Lys Glu Tyr Asn Gly Tyr Tyr Val Val Pro Asn Pro Thr Val Thr Lys
 485 490 495

Asp Ile Phe Asp Tyr Ile Arg Asn Val Tyr Thr Pro Glu Gly Ala Ser
 500 505 510

Tyr Pro Ser Ser Ile Gly Glu Glu Ile Glu Val Leu Tyr Tyr Arg Asp
 515 520 525

Leu Thr Thr Gly Tyr Gln Ser Asp Thr Ile Asn His Lys Pro Thr Leu
 530 535 540

Pro Val Asp Pro Thr Ser Gln Met Ile Thr Val Ser Ala Arg Pro Ser
 545 550 555 560

Asn Gly Ser Glu Asn Glu His Ile Arg Phe Thr Ile Arg Gly Ser Gly
 565 570 575

Thr Glu Pro Lys Leu Lys Val Tyr Ile Glu Ala Cys Ala Asn Glu Glu
 580 585 590

Gln Arg Ala Ser Phe Leu Ala Lys Leu Thr Trp Asn Val Leu Arg Arg
 595 600 605

Glu Trp Phe Arg Pro Asp Glu Met Asn Ile Val Thr Lys Phe
 610 615 620

REFERENCES CITED IN THE DESCRIPTION

Cited references

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- [WO2009017441A](#) [0003] [0007] [0216]
- [WO2010059095A](#) [0003] [0010] [0216]
- [WO2012135110A](#) [0003] [0073] [0216]
- [WO9513362A](#) [0004] [0216]
- [US20090246857A](#) [0004] [0216]
- [US2012270289A1](#) [0005] [0216]
- [WO9517413A](#) [0070]
- [WO9522625A](#) [0070]
- [US5223409A](#) [0070]
- [WO9206204A](#) [0070]
- [US20120184020A](#) [0073]
- [WO2011078262A](#) [0073]
- [WO2012009272A](#) [0073]
- [EP73657A](#) [0105]
- [US4931373A](#) [0108]
- [WO2010059095A1](#) [0162] [0186] [0191] [0196] [0201]
- [US62436648B](#) [0218]
- [US62436687B](#) [0218]
- [US62436706B](#) [0218]
- [US62436723B](#) [0218]

Non-patent literature cited in the description

- **MOLLAPOURPIPER** Molecular Microbiology, 2001, vol. 42, 4919-930 [0008]
- **RICE et al.** EMBOSS: The European Molecular Biology Open Software Suite, 2000, [0033]
- **CUNNINGHAMWELL** Science, 1989, vol. 244, 1081-1085 [0068]
- **HILTONJ.** Biol. Chem., 1996, vol. 271, 4699-4708 [0068]
- **DE VO** Science, 1992, vol. 255, 306-312 [0068]
- **SMITHJ.** Mol. Biol., 1992, vol. 224, 899-904 [0068]
- **WLODAVER et al.** FEBS Lett., 1992, vol. 309, 59-64 [0068]
- **BOWIE et al.** Residues that are directly involved in protein functions such as binding or catalysis will certainly be among the most conserved Science, 1990, vol. 247, 1306-1310 [0069]
- **REIDHAAR-OLSONSAUER** Science, 1988, vol. 241, 53-57 [0070]
- **BOWIESAUER** Proc. Natl. Acad. Sci. USA, 1989, vol. 86, 2152-2156 [0070]
- **LOWMAN et al.** Biochemistry, 1991, vol. 30, 10832-10837 [0070]
- **DERBYSHIRE** Gene, 1986, vol. 46, 145- [0070]

- **NER et al.**DNA, 1988, vol. 7, 127- [\[0070\]](#)
- **NESS et al.**Nature Biotechnology, 1999, vol. 17, 893-896 [\[0071\]](#)
- **VAN WINDEN et al.**FEMS Yeast Research, 2005, vol. 5, 559-568 [\[0073\]](#)
- **SAMBROOK, J.FRITSCH, E. F.MANIATIS, T.**Molecular Cloning: A Laboratory ManualCold Spring Harbor Laboratory Press19890000 [\[0101\]](#)
- **SILHAVY, T. J.BENNAN, M. L.ENQUIST, L. W.**Experiments with Gene FusionsCold Spring Harbor Laboratory Press19840000 [\[0101\]](#)
- **AUSUBEL et al.**Short Protocols in Molecular BiologyWiley & Sons19950000 [\[0101\]](#) [\[0216\]](#)
- **AUSUBEL, F. M. et al.**Current Protocols in Molecular BiologyGreene Publishing Assoc. and Wiley-Interscience19870000 [\[0101\]](#)
- **Methods in Enzymology**Guide to Yeast Genetics and Molecular and Cell BiologyElsevier Academic Press20040000vol. 194, [\[0101\]](#)
- **SHARPLIN**Nucleic Acids Research, 1987, vol. 15, 1281-1295 [\[0104\]](#)
- **JANSEN et al.**Nucleic Acids Res., 2003, vol. 31, 82242-51 [\[0104\]](#)
- **HAUF et al.**Enzym Microb Technol, 2000, vol. 26, 688-698 [\[0105\]](#)
- **GIETZSUGINO**New yeast vectors constructed with in vitro mutagenized yeast genes lacking six-base pair restriction sitesGene, 1988, vol. 74, 527-534 [\[0108\]](#)
- **OKKELS**Ann. New York Acad. Sci., 1996, vol. 782, 202-207 [\[0108\]](#)
- **MNAIMNEH et al.**Cell, 2004, vol. 118, 131-44 [\[0109\]](#)
- **WACH et al.**Yeast, 1994, vol. 10, 1793-1808 [\[0109\]](#)
- **Methods in Enzymology**, 1991, vol. 194, 281-301 [\[0109\]](#)
- **CHANGHO**Appl Biochem Biotechnol., 1988, vol. 17, 313-318 [\[0216\]](#)
- **ELIASSON et al.**Enzyme and Microbial Technology, 2001, vol. 29, 288-297 [\[0216\]](#)
- **GIETZSCHIEL**Nature protocols, 2007, vol. 2, 31-34 [\[0216\]](#)
- **HOU**Appl Microbiol Biotechnol, 2012, vol. 94, 205-214 [\[0216\]](#)
- **INOUE et al.**Gene, 1990, vol. 96, 123-28 [\[0216\]](#)
- **JIN et al.**App Environmental Microbiol, 2003, vol. 69, 1495-503 [\[0216\]](#)
- **JOHANSSONHAHN-HÄGERDAL**Yeast, 2002, vol. 19, 3225-231 [\[0216\]](#)
- **JOHANSSON et al.**Appl Environmental Microbiol, 2001, vol. 67, 94249-4255 [\[0216\]](#)
- **KARHUMAA et al.**Microb Cell Fact., 2007, vol. 5, 65- [\[0216\]](#)
- **KARHUMAA et al.**Yeast, 2005, vol. 22, 5359-68 [\[0216\]](#)
- **KARHUMAA et al.**Yeast, 2009, vol. 26, 7371-82 [\[0216\]](#)
- **KIM et al.**Biotechnol Adv., 2013, vol. 31, 6851-61 [\[0216\]](#)
- **KUYPER et al.**FEMS Yeast Res., 2005, vol. 5, 4-5399-409 [\[0216\]](#)
- **LINDEN TPEETRE JHAHN-HÄGERDAL B.**Appl Environ Microbiol., 1992, vol. 58, 51661-9 [\[0216\]](#)
- **LÖNN et al.**Enzyme Microb. Technol., 2003, vol. 32, 5567-573 [\[0216\]](#)
- **MADHAVAN et al.**Appl Microbiol Biotechnol., 2009, vol. 82, 61067-1078 [\[0216\]](#)
- **MATSUSHIKA et al.**Enzyme and Microbial Technology, 2011, vol. 48, 466-471 [\[0216\]](#)
- **MOLLAPOUR MPIPER PW.**Mol Microbiol., 2001, vol. 42, 4919-30 [\[0216\]](#)
- **MUMBERG DMÜLLER RFUNK M.**Gene, 1995, vol. 156, 1119-22 [\[0216\]](#)
- **NEEDLEMANWUNSCHJ.** Mol. Biol., 1970, vol. 48, 443-453 [\[0216\]](#)
- **NGUYEN et al.**Mycological Res, 2006, vol. 110, 1232-1241 [\[0216\]](#)
- **PARACHIN et al.**Metabolic Engineering, 2011, vol. 13, 508-517 [\[0216\]](#)
- **RICE et al.**Trends Genet, 2000, vol. 16, 276-277 [\[0216\]](#)

- **RUNQUIST et al.**Appl. Environ. Microbiol., 2010, vol. 76, 237796-7802 [\[0216\]](#)
- **SHAMANNASANDERSONJ.** Bacteriol., 1979, vol. 139, 64-70 [\[0216\]](#)
- **TIWARI et al.**Biochemical and Biophysical Research Communications, 2008, vol. 366, 340-345 [\[0216\]](#)
- **WALTHER et al.**FEBS Lett., 2012, vol. 586, 234114-8 [\[0216\]](#)
- **WOHLBACH et al.**Proc Natl Acad Sci U S A, 2011, vol. 108, 3213212-7 [\[0216\]](#)
- **XU et al.**Mol Syst Biol, 2013, vol. 9, 665- [\[0216\]](#)

KRAV

1. Rekombinant *Saccharomyces*-celle, som er i stand til at fermentere xylose, og som omfatter et heterologt gen, der koder for en xylulokinase (XK), hvor XK:

- tilvejebringer en enzymatisk aktivitet til omdannelse af D-xylulose til xylulose-5-phosphat på mindst to gange den tilvejebragt af *S. cerevisiae* XK ifølge SEQ ID NO: 32, og
 - tilvejebringer en anærob vækstrate for den rekombinante celle på xylose, som er højere end den tilvejebragt af en *S. cerevisiae* XK ifølge SEQ ID NO: 32, og
- hvor XK har en sekvensidentitet på mindst 80% med aminosyresekvensen ifølge SEQ ID NO: 6 eller en sekvensidentitet på mindst 80% med aminosyresekvensen ifølge SEQ ID NO: 22.

2. Rekombinant celle ifølge krav 1, hvor XK har en sekvensidentitet på mindst 90% med aminosyresekvensen ifølge SEQ ID NO: 6 eller en sekvensidentitet på mindst 90% med aminosyresekvensen ifølge SEQ ID NO: 22.

3. Rekombinant celle ifølge krav 1, hvor XK har en sekvensidentitet på mindst 95% med aminosyresekvensen ifølge SEQ ID NO: 6 eller en sekvensidentitet på mindst 95% med aminosyresekvensen ifølge SEQ ID NO: 22.

4. Fremgangsmåde til fremstilling af en rekombinant *Saccharomyces*-celle, hvilken fremgangsmåde omfatter transformering af en *Saccharomyces*-celle med en eller flere vektorer, der omfatter gener, der koder for

- en XK, der omfatter aminosyresekvensen ifølge SEQ ID NO: 6 eller en sekvens med en sekvensidentitet på mindst 80% med aminosyresekvensen ifølge SEQ ID NO: 6, såsom mindst 90%, 95%, 97%, 98% eller 99% med aminosyresekvensen ifølge SEQ ID NO: 6, aminosyresekvensen ifølge SEQ ID NO: 22 eller en sekvens med sekvensidentitet på mindst 80% med SEQ ID NO: 22, såsom mindst 90%, 95%, 97%, 98% eller 99% med aminosyresekvensen ifølge SEQ ID NO: 22,

- en XR, og
- en XDH, og valgfrit, regulatoriske sekvenser til ekspresion af generne i værtscellen.

5. Rekombinant celle ifølge krav 1, hvilken celle stammer fra en *Saccharomyces cerevisiae*-celle.

6. Rekombinant celle ifølge ethvert af kravene 1-5, hvilken celle omfatter et heterologt gen, der koder for en transaldolase (TAL), et heterologt gen, der koder for en transketolase (TKL) og et

heterologt gen, der koder for en ribose-5-phosphate-ketolisomerase (RKI), og

(a) et heterologt gen, der koder for en xylosereduktase (XR) og et heterologt gen, der koder for en xylitoldehydrogenase (XDH), og/eller

(b) et heterologt gen, der koder for en xyloseisomerase (XI).

5

7. Fremgangsmåde til fremstilling af et fermenteringsprodukt, hvilken fremgangsmåde omfatter

(a) at bringe den rekombinante celle ifølge ethvert af kravene 1-5 i kontakt med et medium, der omfatter en carbonkilde, der omfatter xylose eller arabinose, under anærobe betingelser, og

10 (b) at isolere fermenteringsproduktet fra mediet.

DRAWINGS

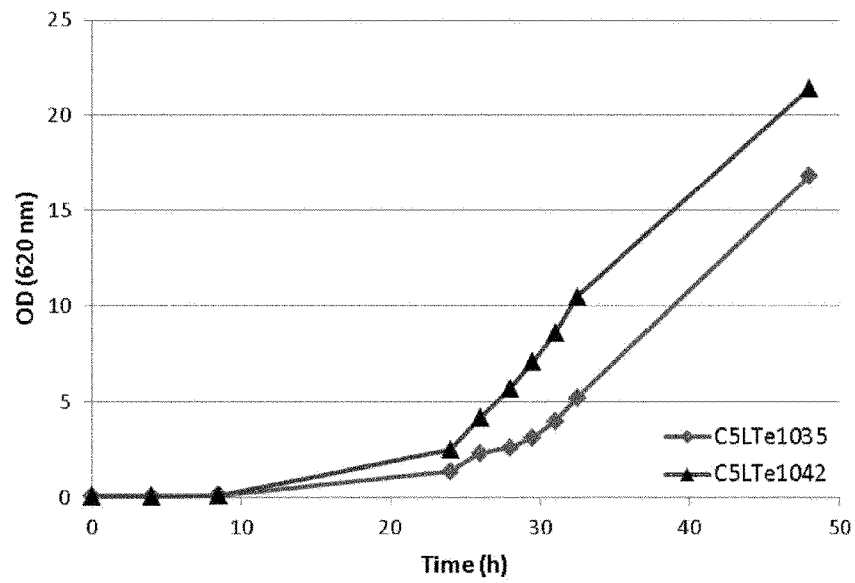
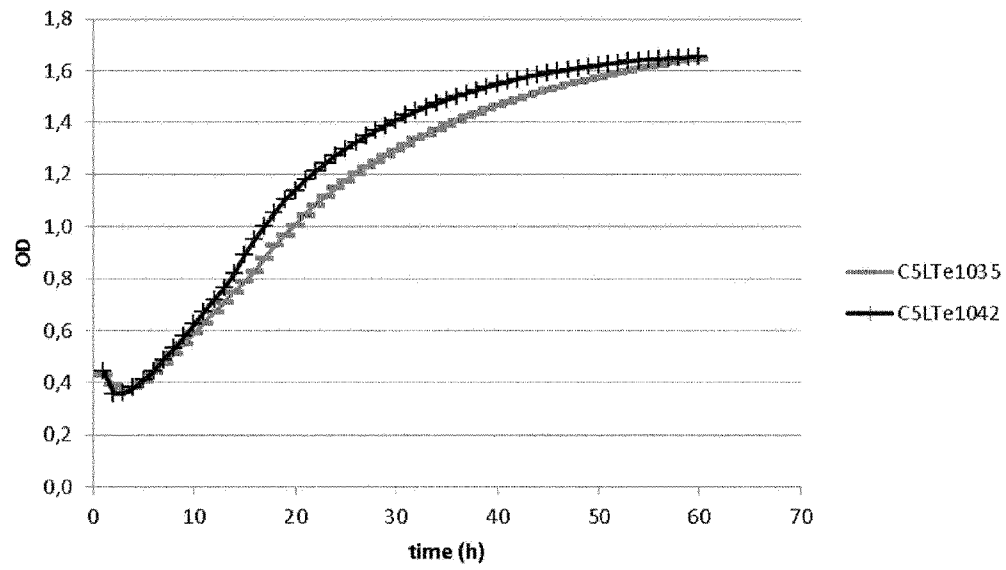
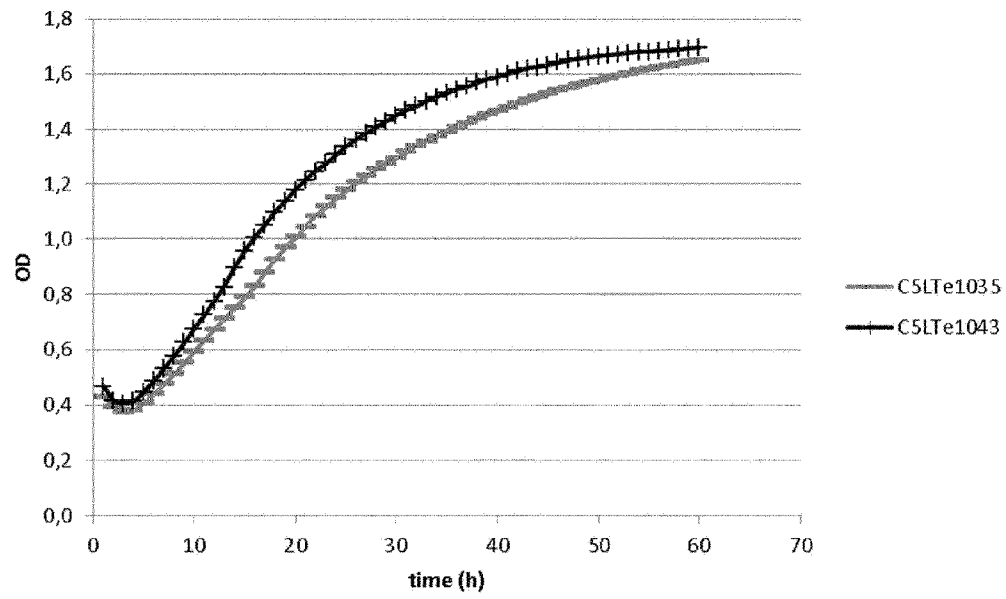
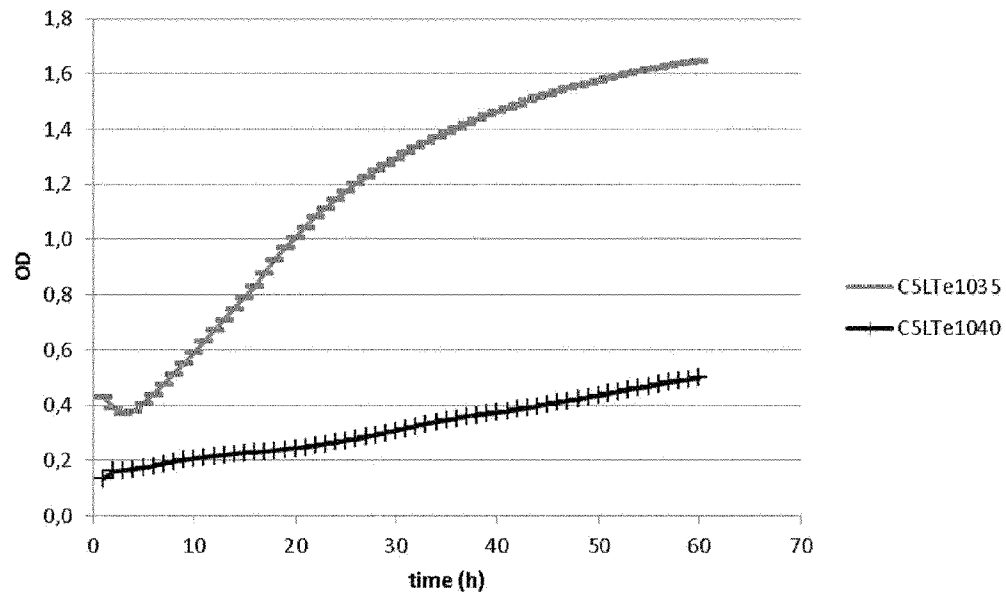
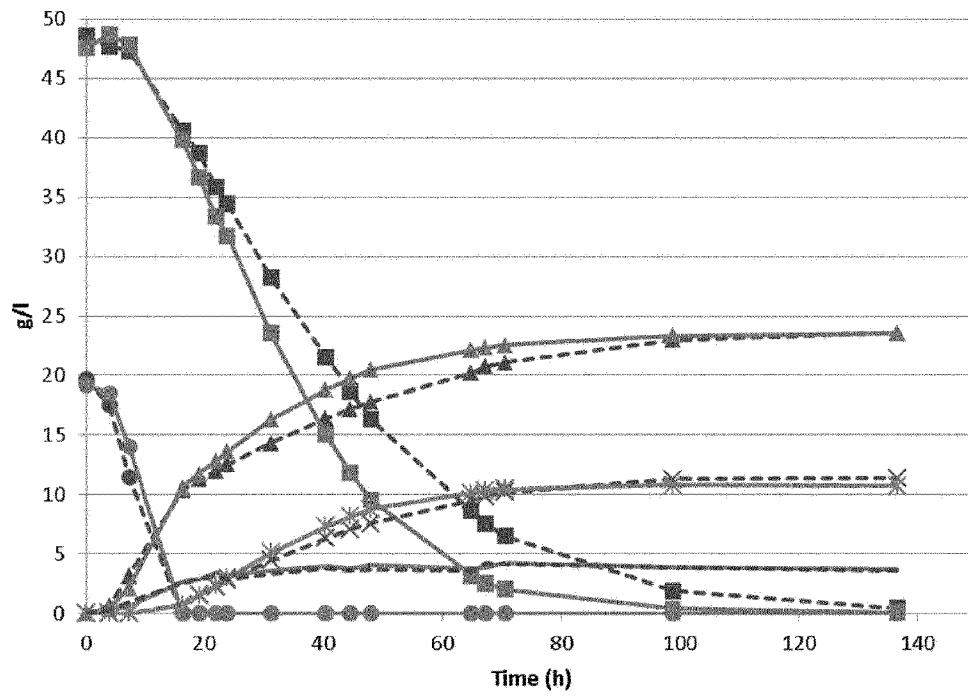


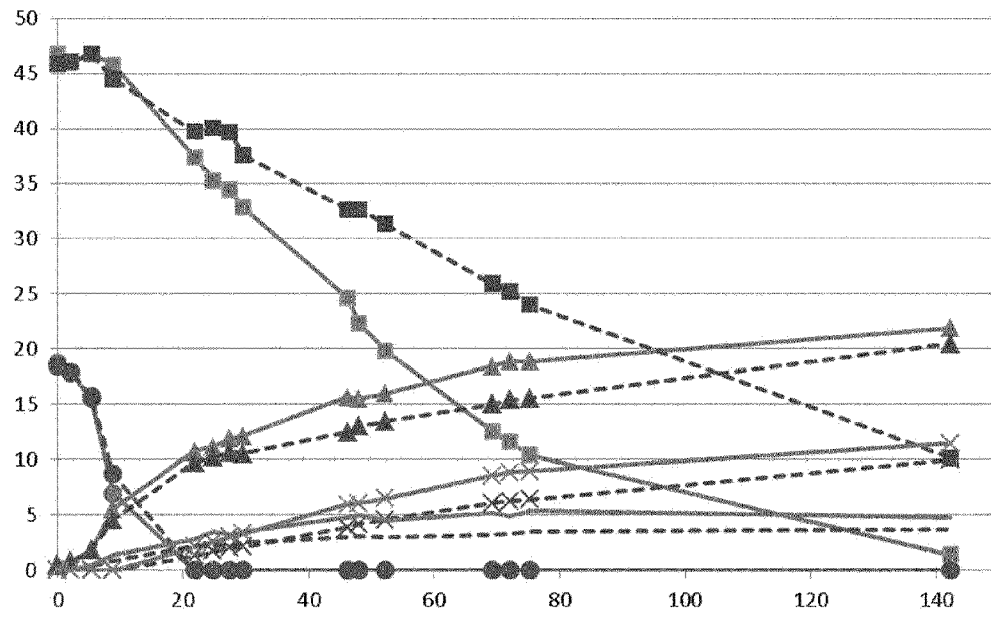
FIG. 1

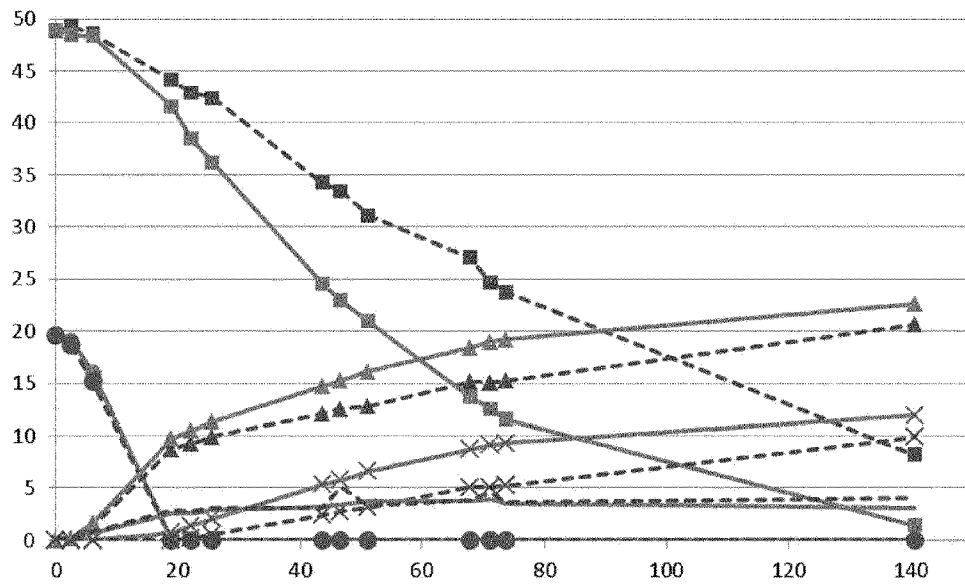
**FIG. 2**

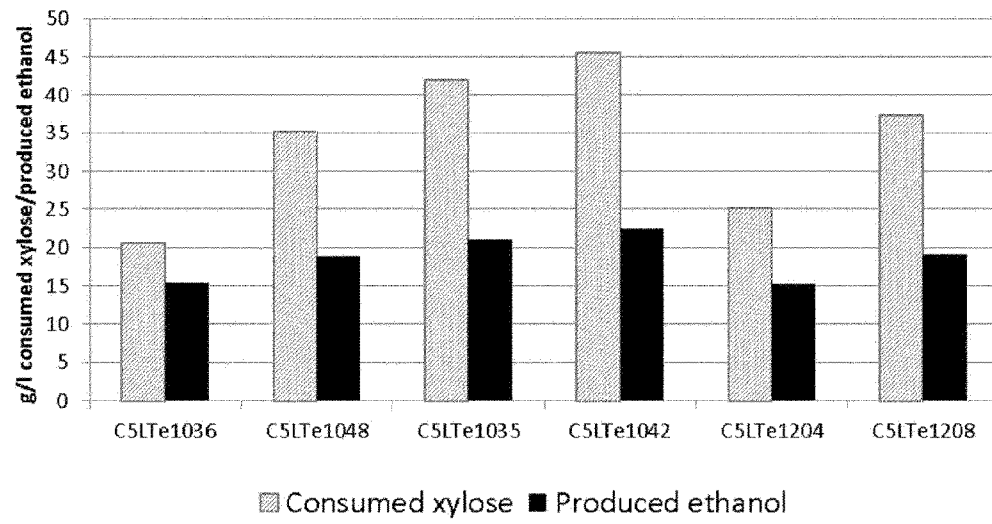
**FIG. 3**

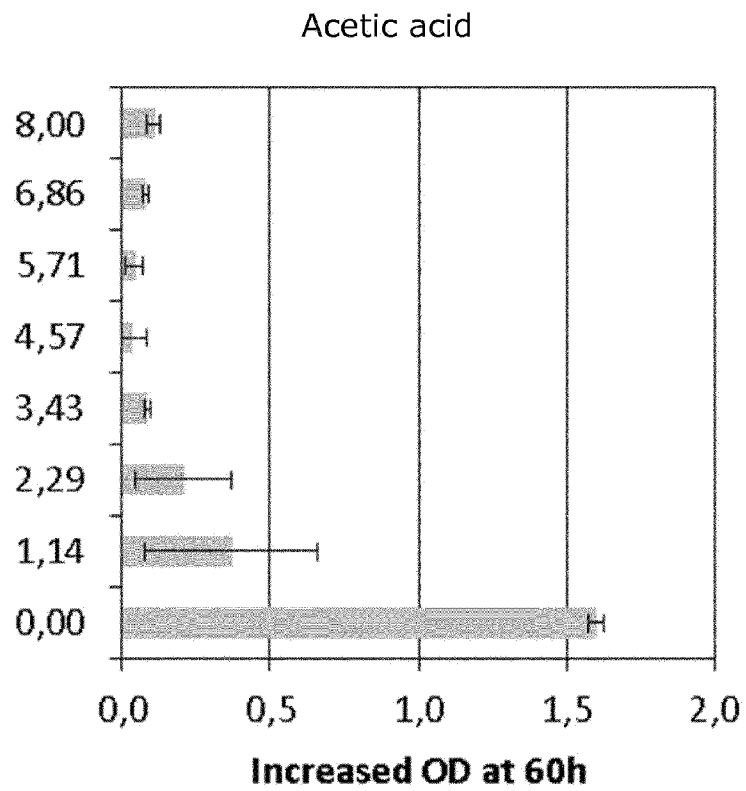
**FIG. 4**

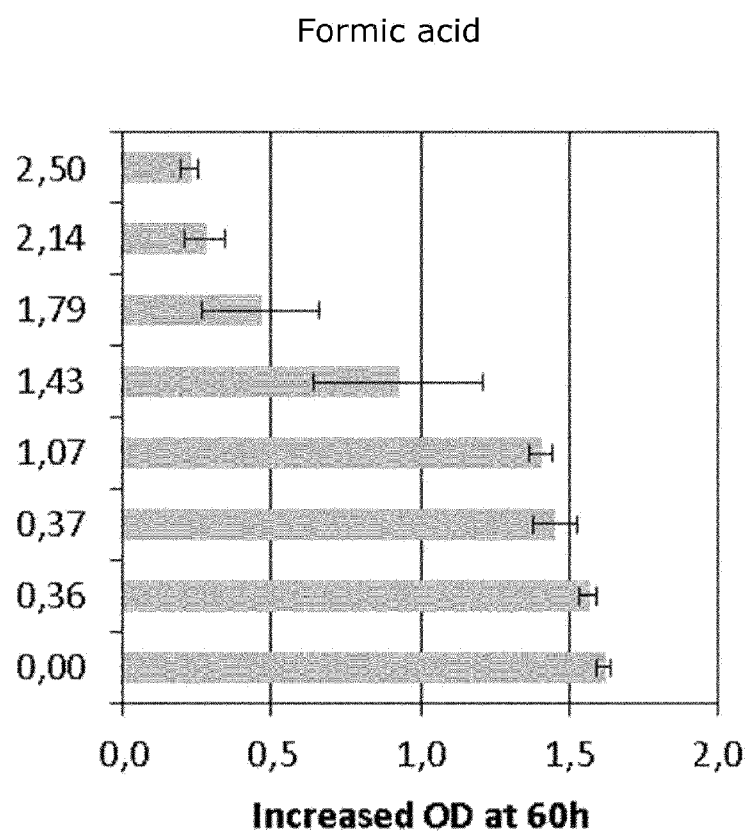
**FIG. 5**

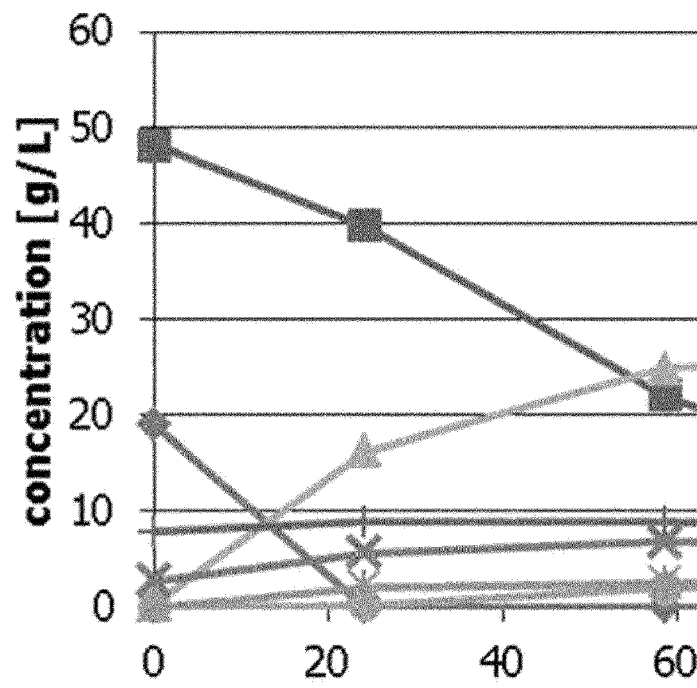
**FIG. 6**

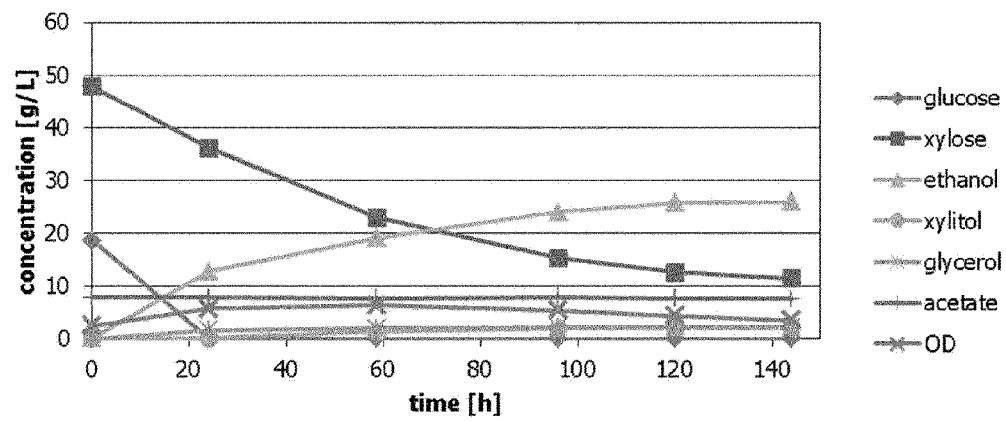
**FIG. 7**

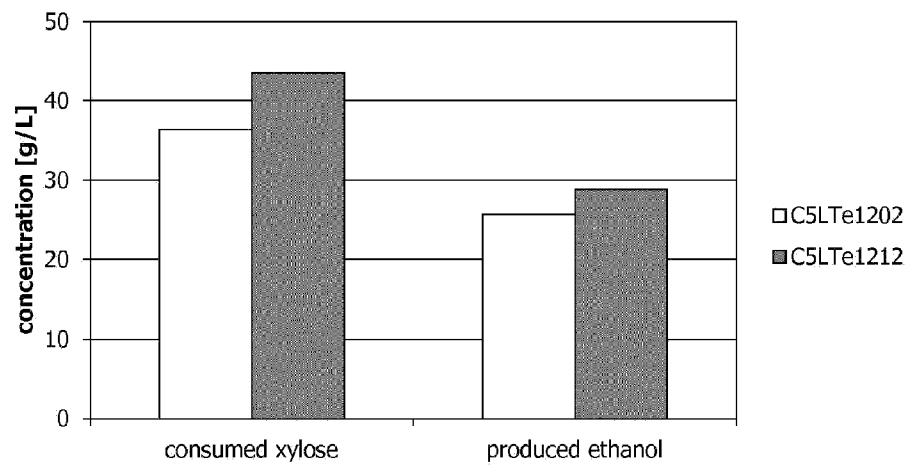
**FIG. 8**

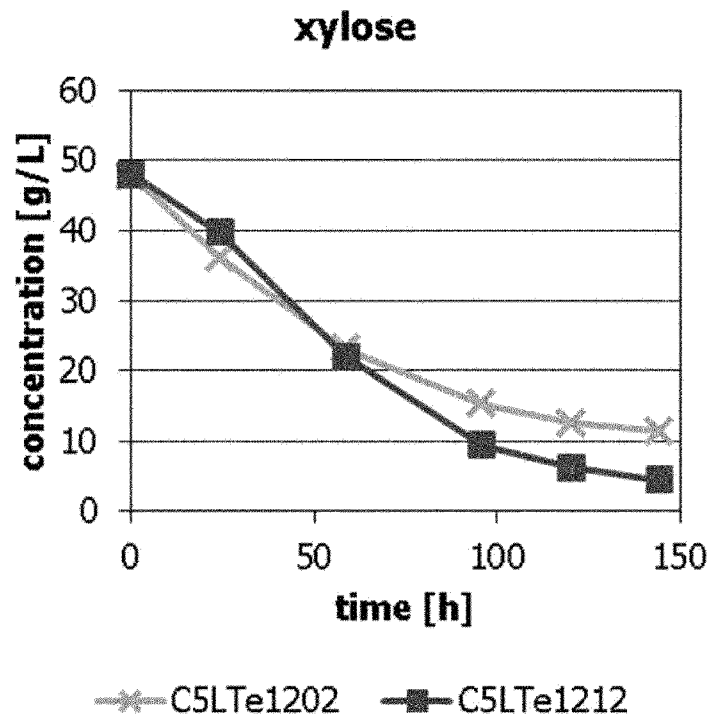
**FIG. 9A**

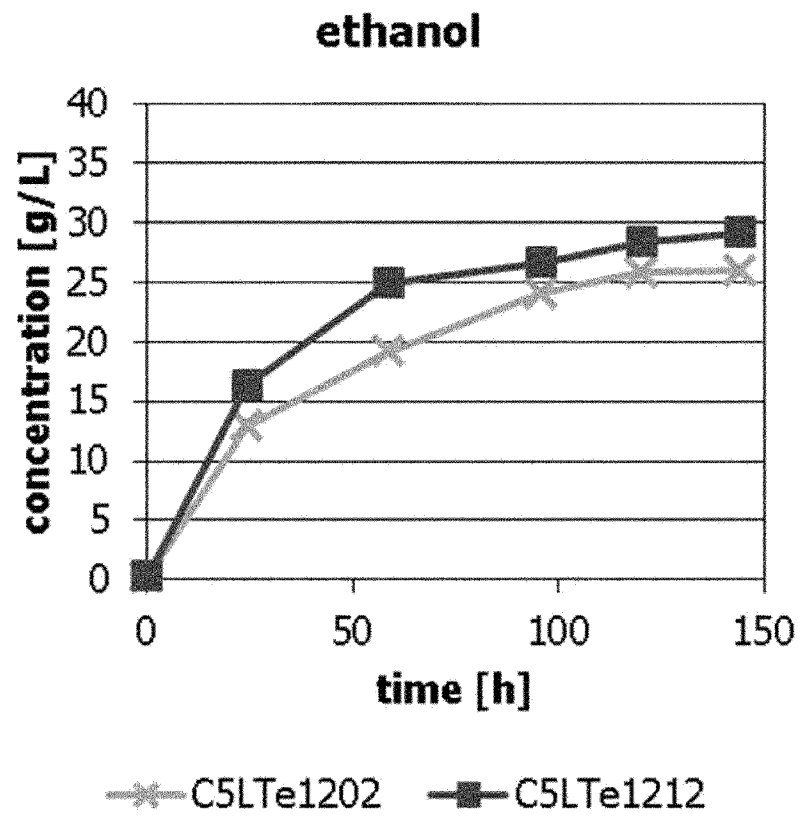
**FIG. 9B**

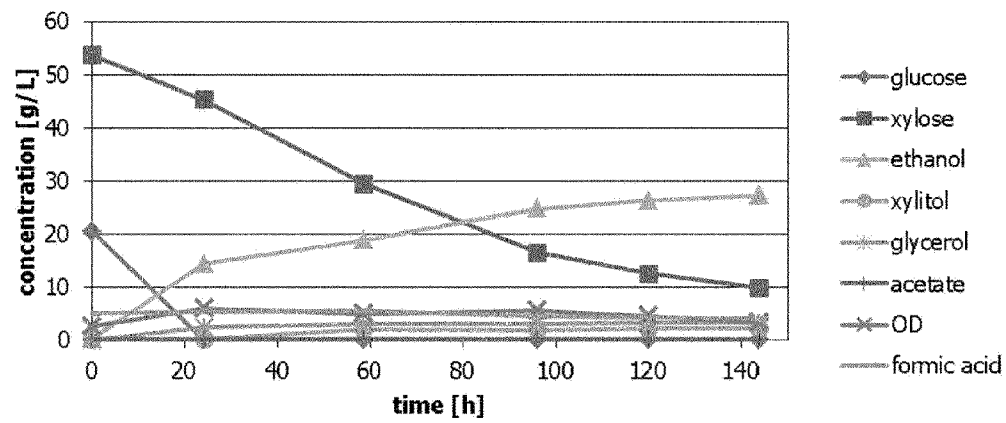
**FIG. 10A**

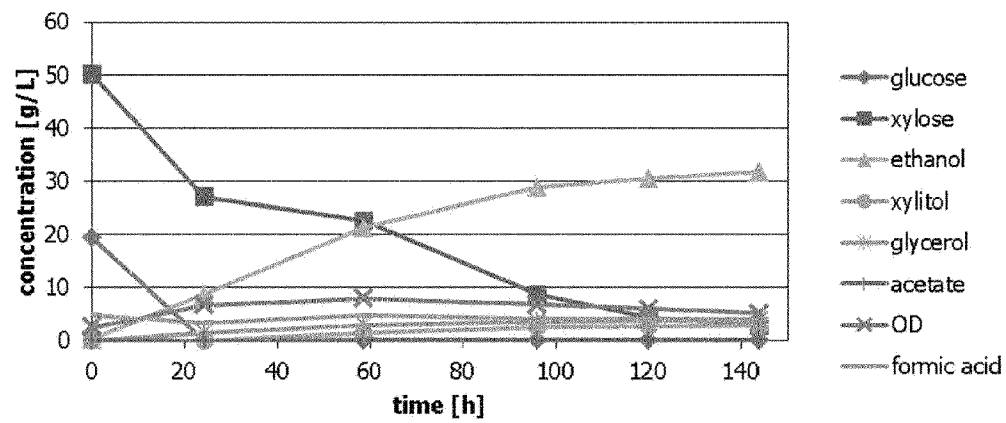
C5LTe1202**FIG. 10B**

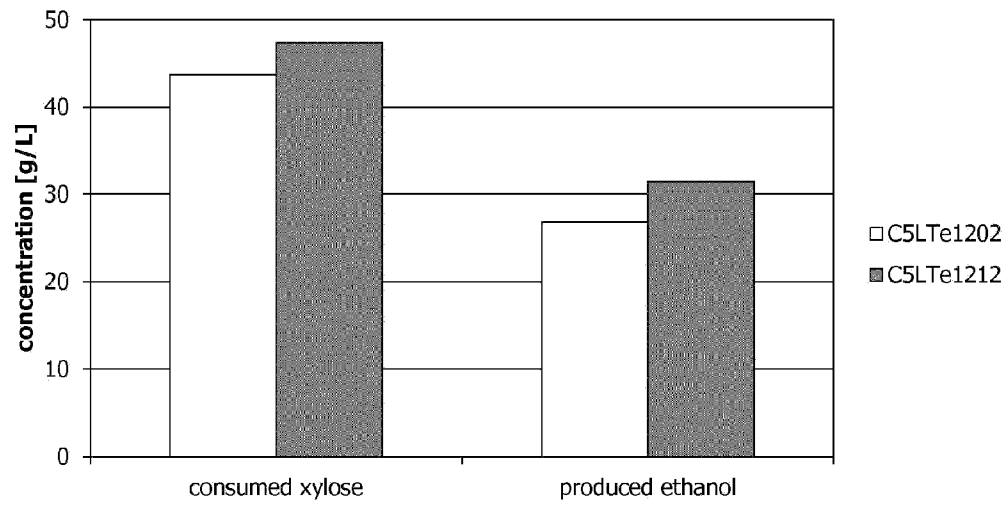
**FIG. 11**

**FIG. 12A**

**FIG. 12B**

1202**FIG. 13A**

1212**FIG. 13B**

**FIG. 14**

xylose

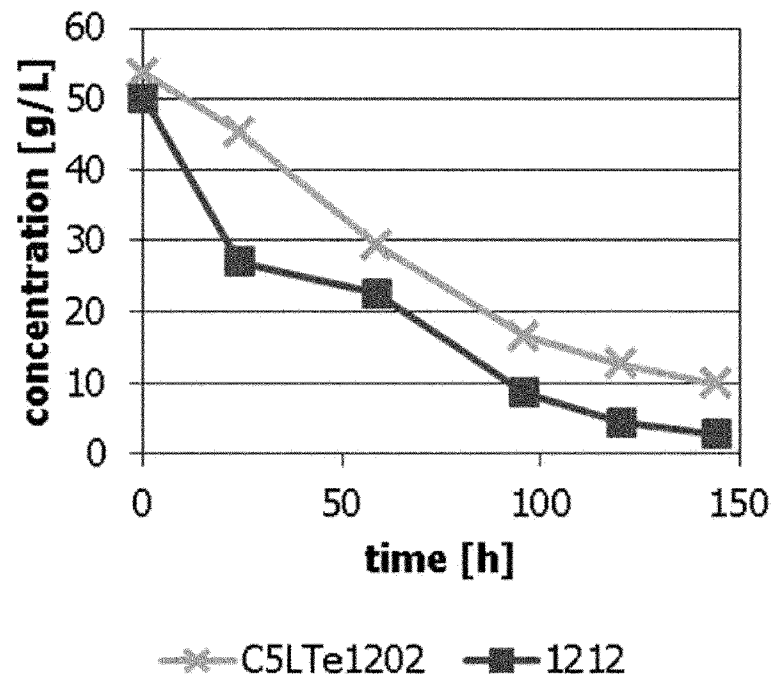


FIG. 15A

ethanol

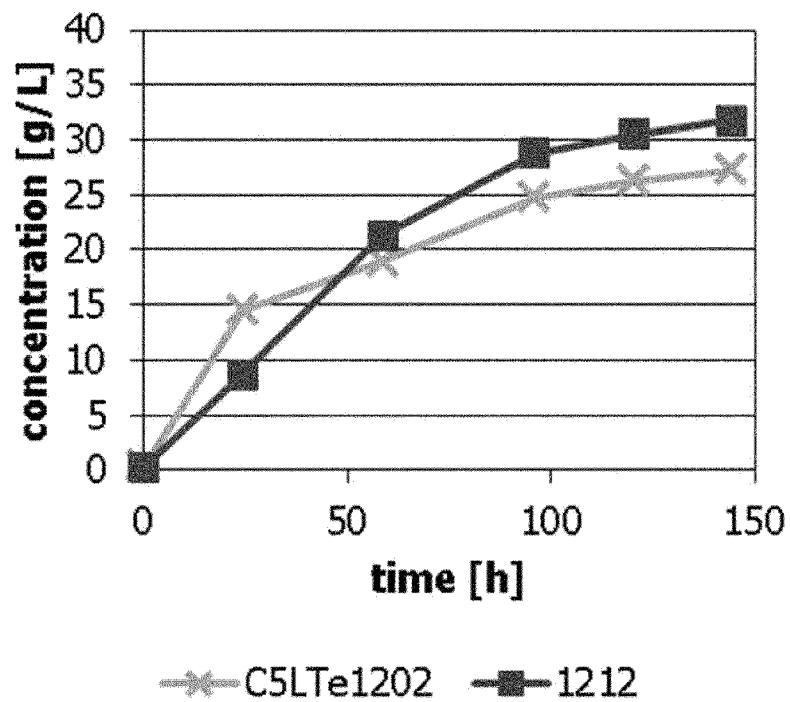
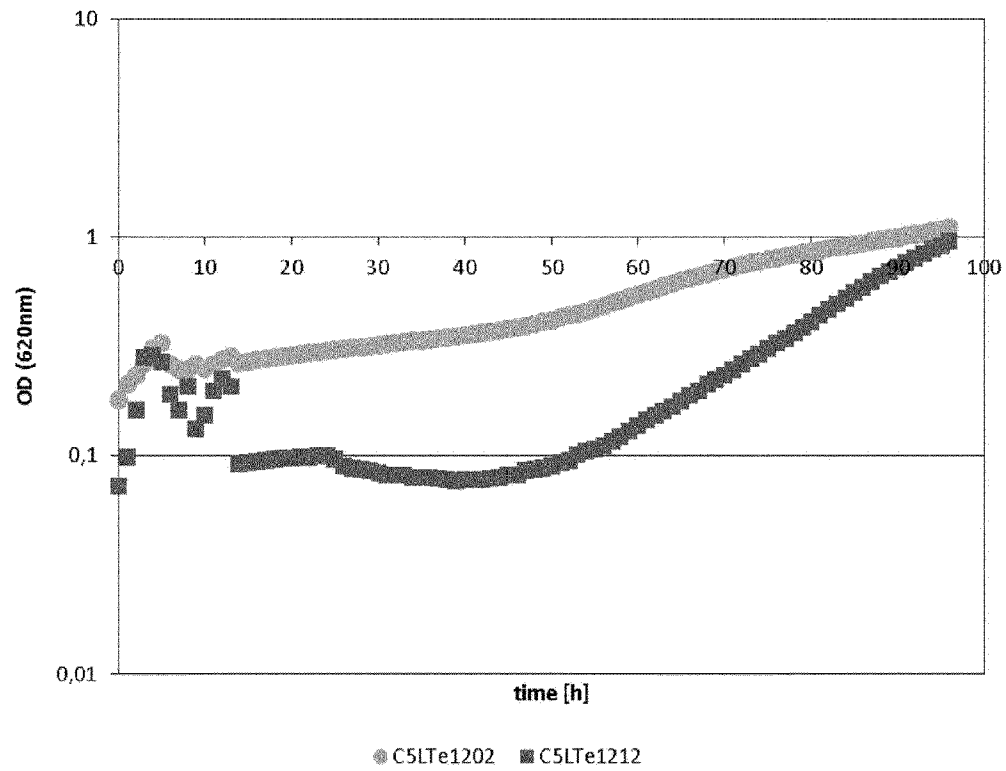
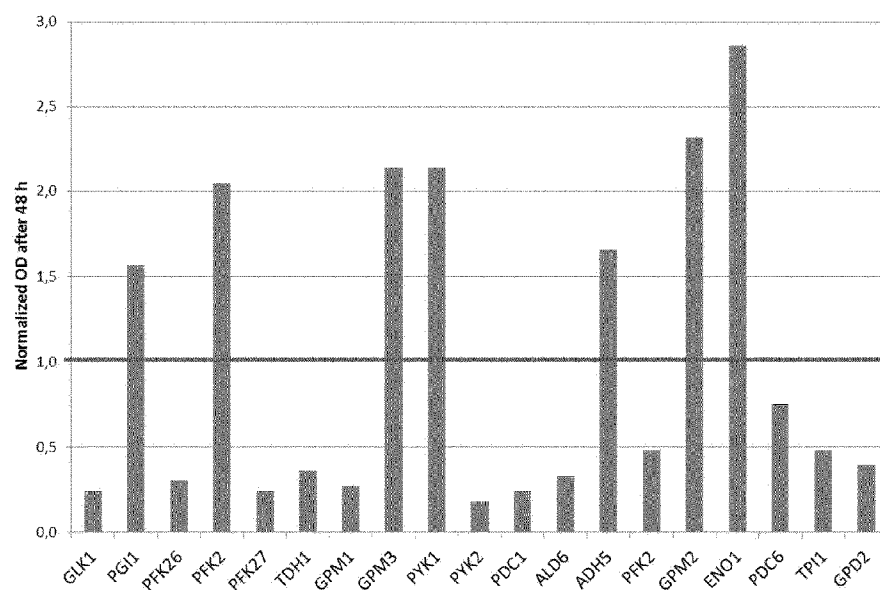
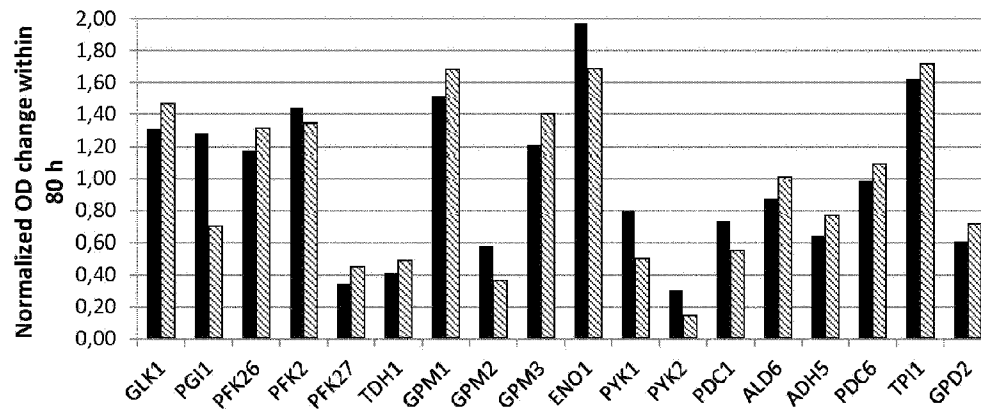
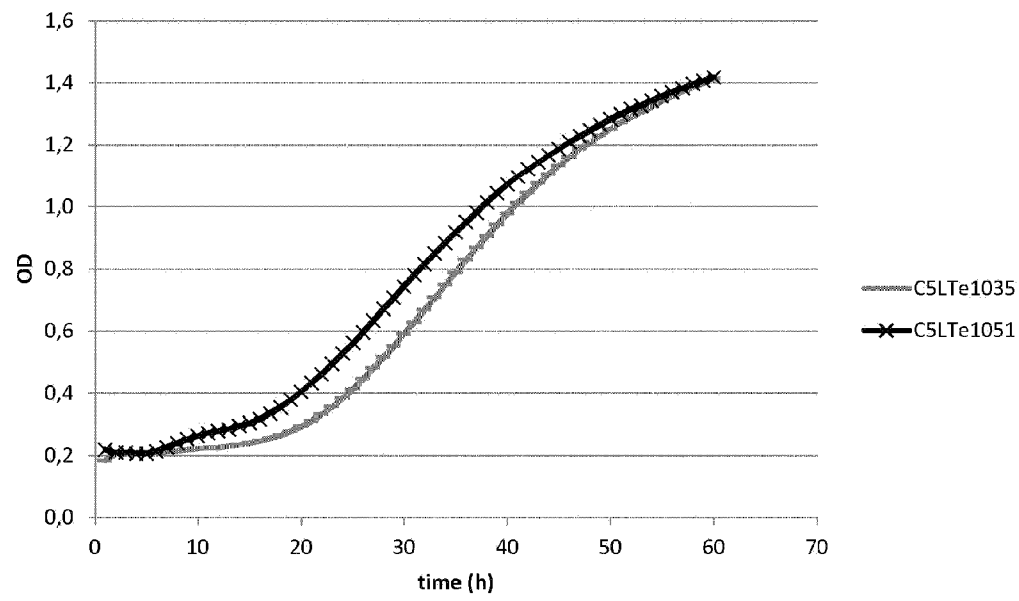


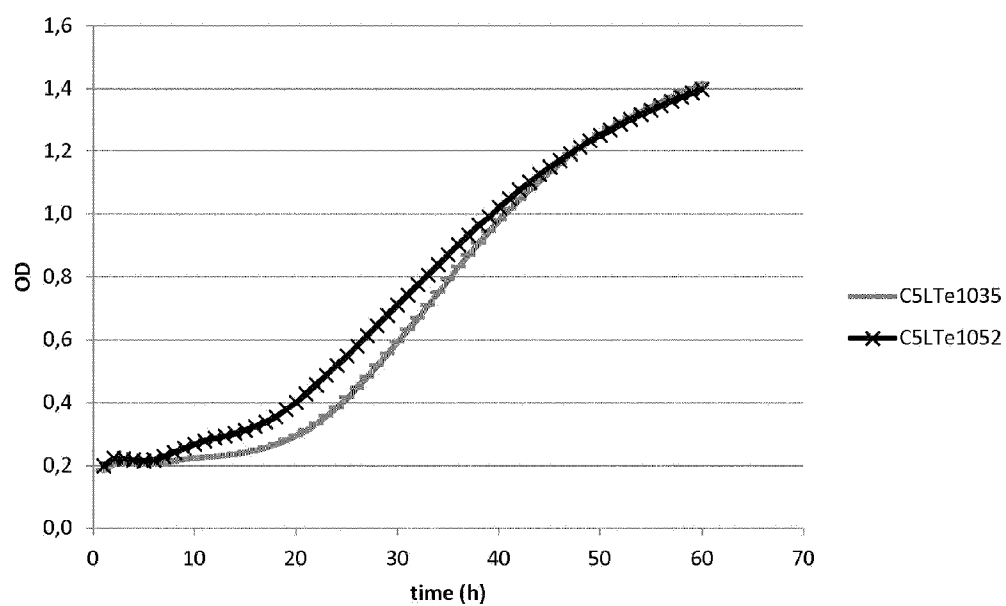
FIG. 15B

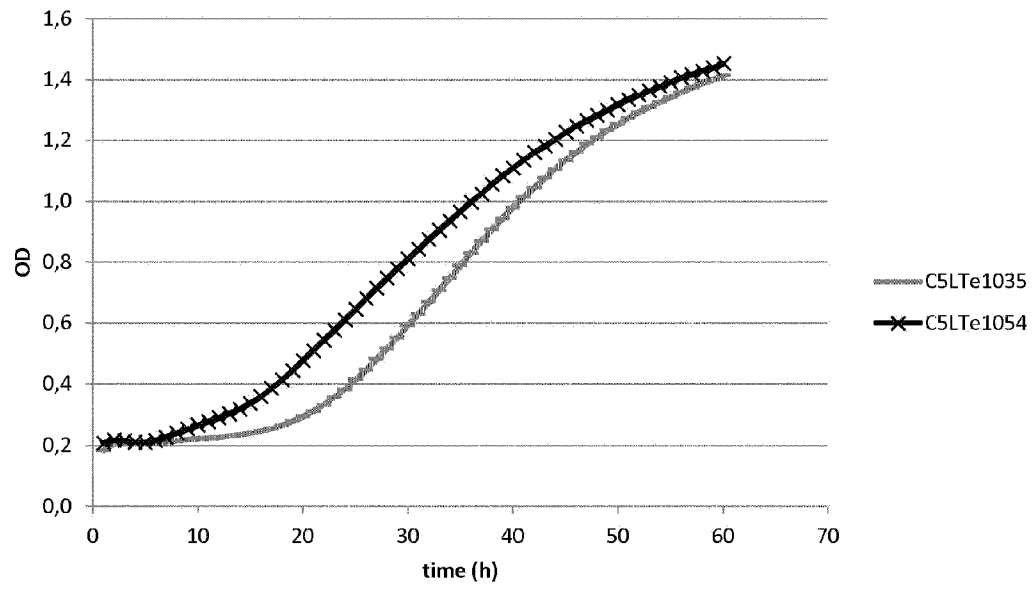
**FIG. 16**

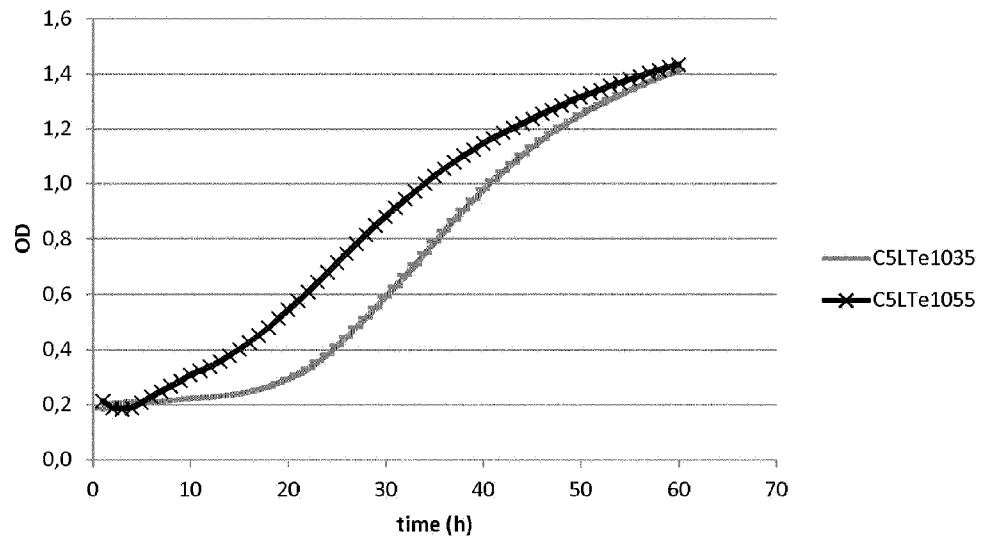
**FIG. 17**

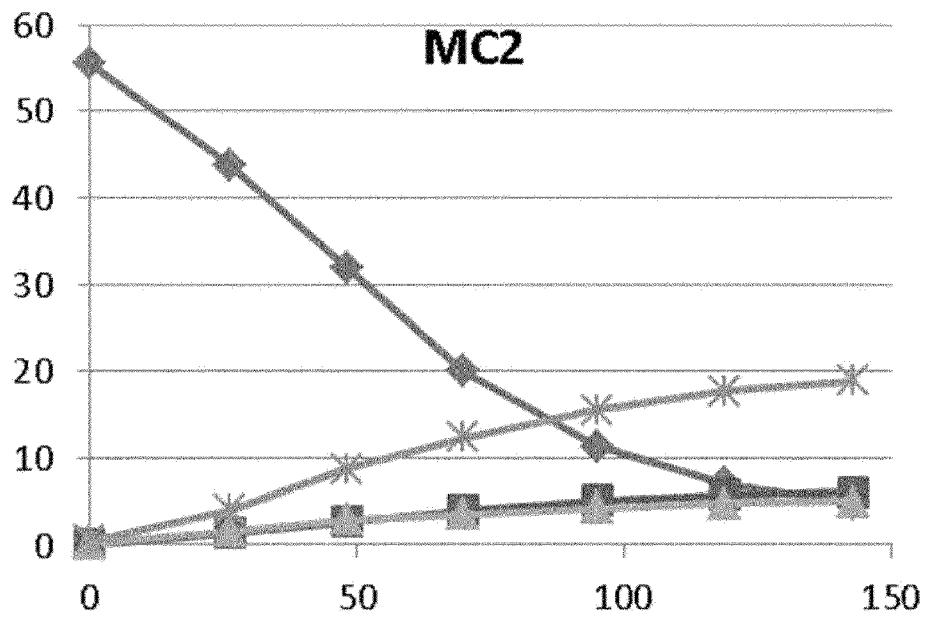
**FIG. 18**

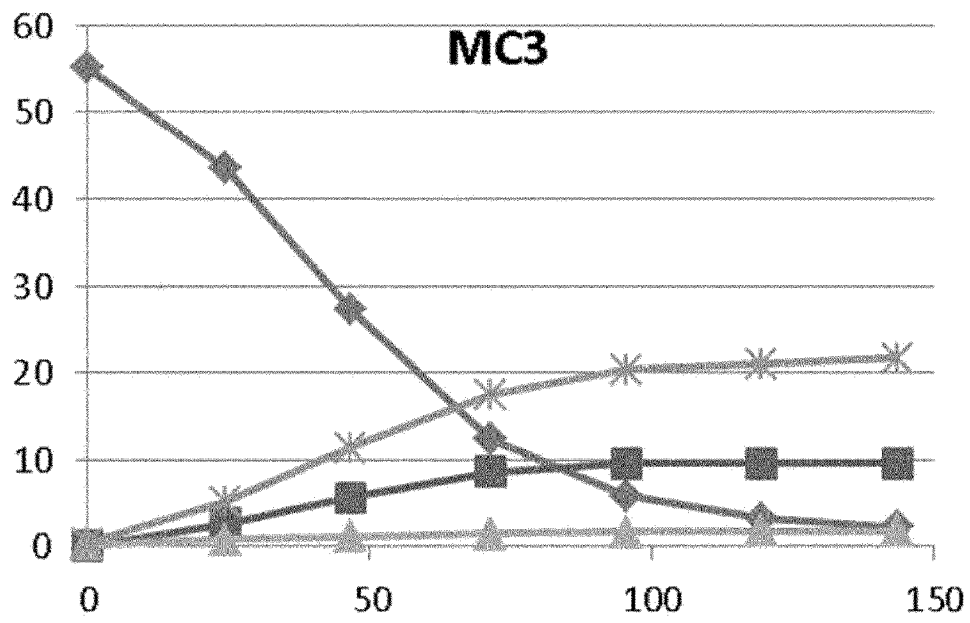
**FIG. 19**

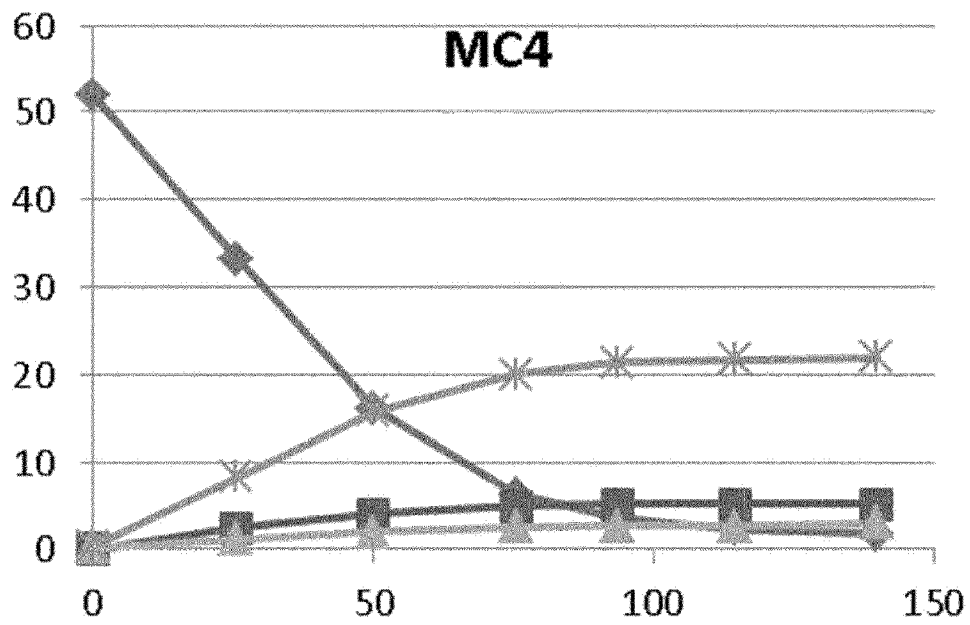
**FIG. 20**

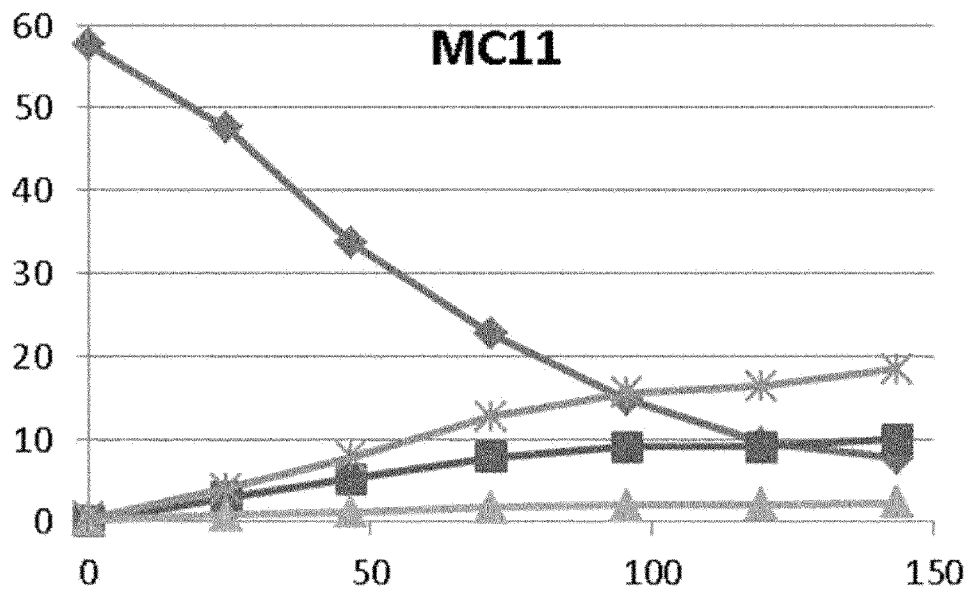
**FIG. 21**

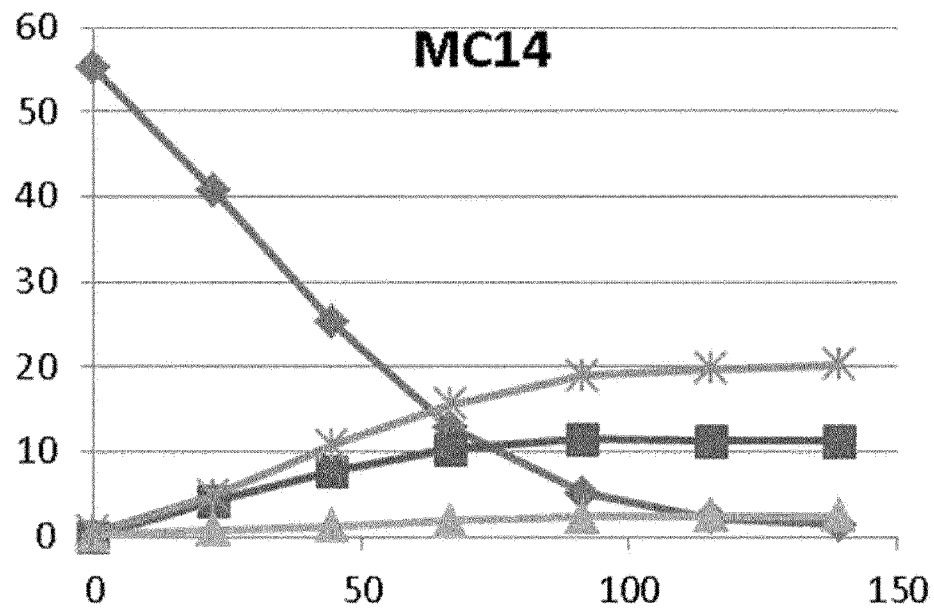
**FIG. 22**

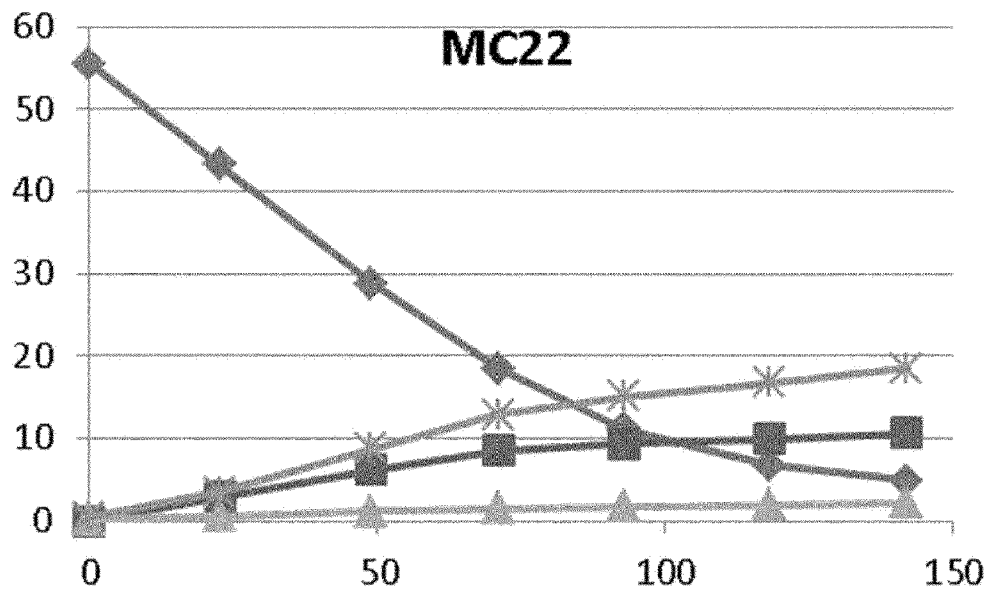
**FIG. 23A**

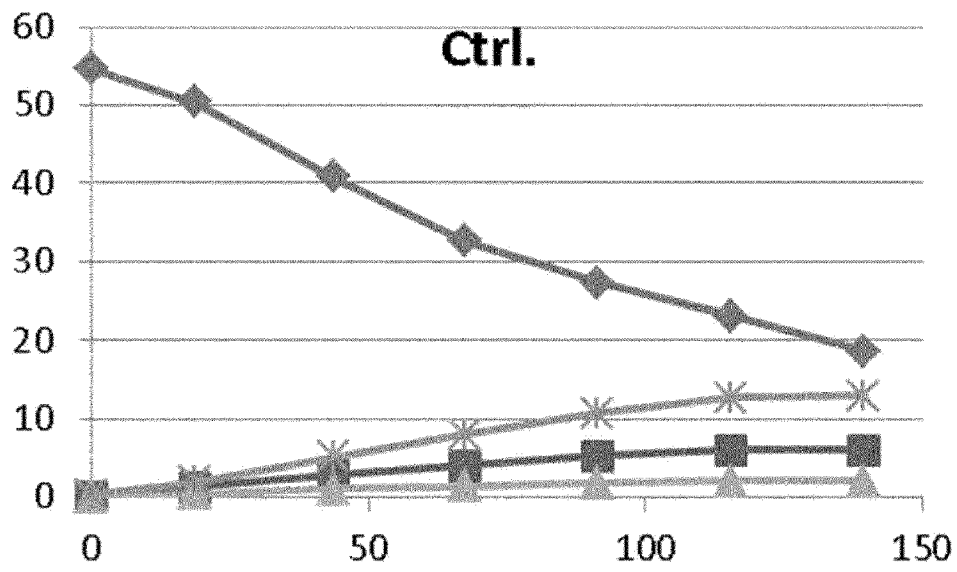
**FIG. 23B**

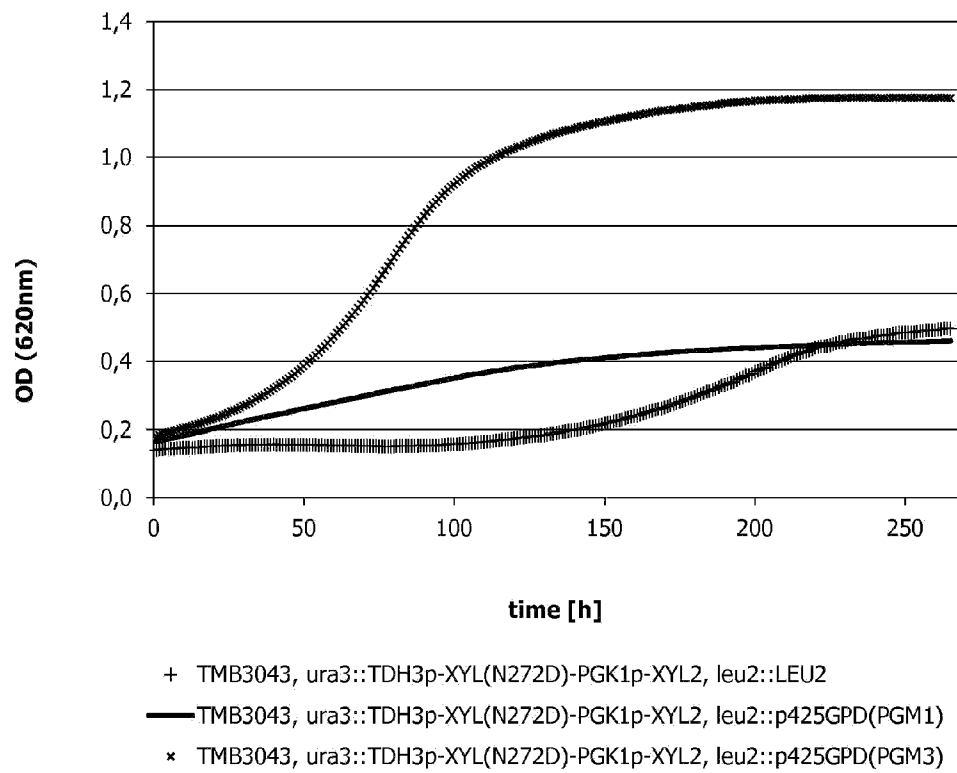
**FIG. 23C**

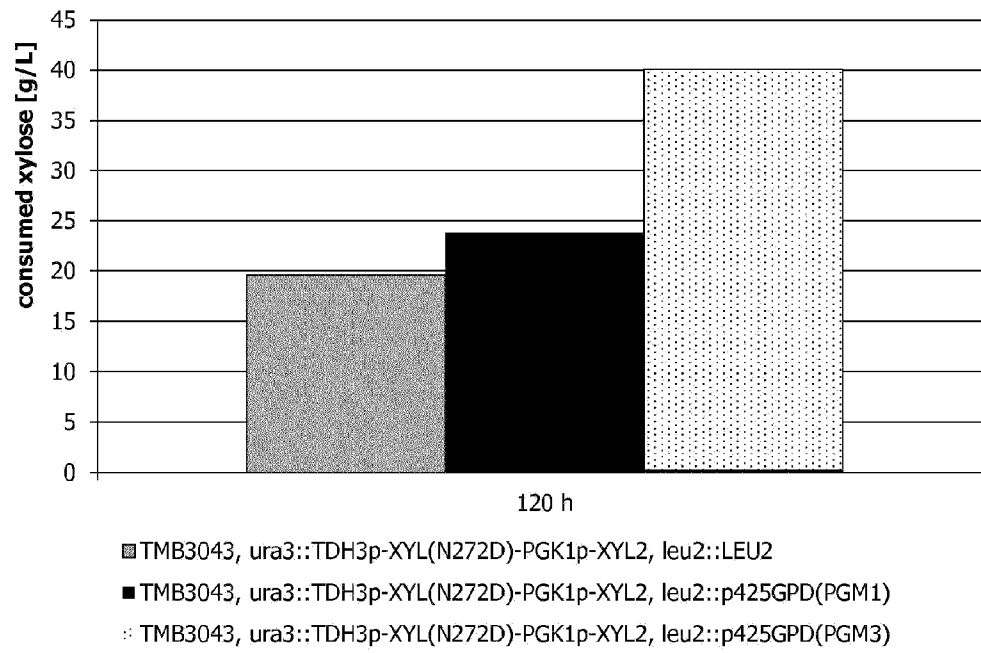
**FIG. 23D**

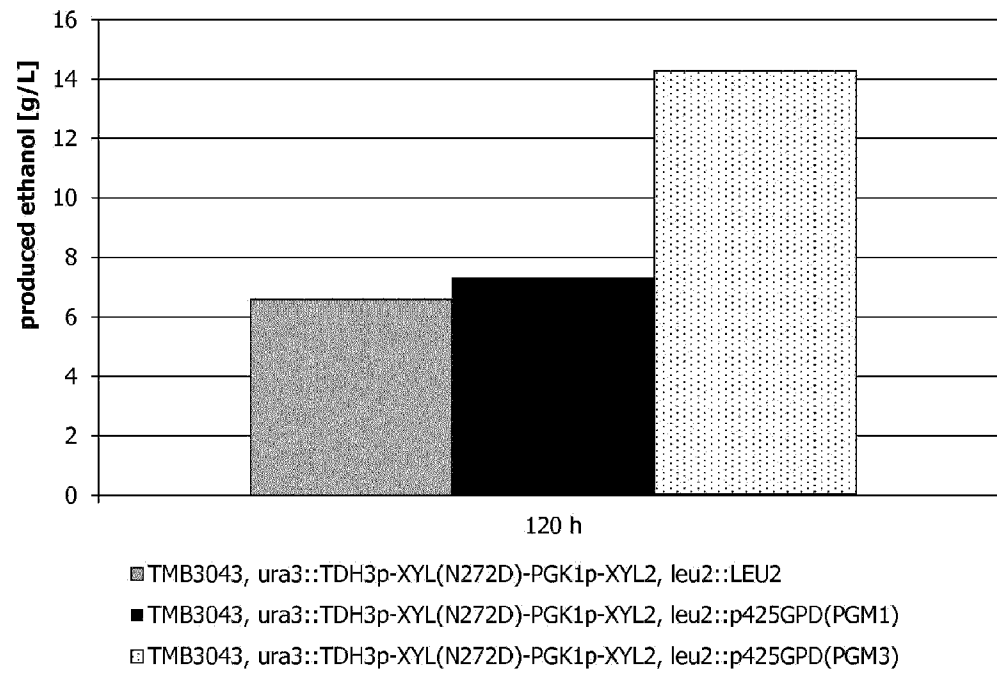
**FIG. 23E**

**FIG. 23F**

**FIG. 23G**

**FIG. 24**

**FIG. 25**

**FIG. 26**