(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
7 April 2011 (07.04.2011)

PCT

(10) International Publication Number
WO 2011/041402 A1

(51) International Patent Classification:
C12N 9/04 (2006.01) *C12N 9/10* (2006.01)
C12N 9/88 (2006.01) *C12N 15/52* (2006.01)
C12N 1/21 (2006.01) *C12P 7/16* (2006.01)
C12N 9/02 (2006.01) *C12N 15/74* (2006.01)

(21) International Application Number:
PCT/US2010/050705

(22) International Filing Date:
29 September 2010 (29.09.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
61/246,717 29 September 2009 (29.09.2009) US

(71) Applicant *(for all designated States except US)*: **BUTAMAX(TM) ADVANCED BIOFUELS LLC** [US/US]; 1007 Market Street, Wilmington, Delaware 19898 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **PAUL, Brian James** [US/US]; 115 Fairfax Boulevard, Wilmington, Delaware 19803 (US). **SUH, Wonchul** [KR/US]; 8 Salina Court, Piersons Ridge, Hockessin, Delaware 19707 (US).

(74) Agent: **LHULIER, Christine, M.**; E. I. du Pont de Nemours and Company, Legal Patent Records Center, 4417 Lancaster Pike, Wilmington, Delaware 19805 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

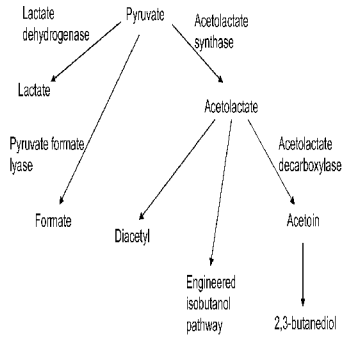
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2011/041402 A1

(54) **Title**: IMPROVED FLUX TO ACETOLACTATE-DERIVED PRODUCTS IN LACTIC ACID BACTERIA

Figure 1



(57) **Abstract**: An engineering method was developed to allow genetic modification and isolation of lactic acid bacteria cells that lack lactate dehydrogenase.and acetolactate decarboxylase activities. In cells with these modifications and an isobutanol biosynthetic pathway, improved production of isobutanol was observed.

IMPROVED FLUX TO ACETOLACTATE-DERIVED PRODUCTS IN LACTIC ACID BACTERIA

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of priority to U.S. Provisional Patent Application No. 61/246717, filed on September 29, 2009, the entirety of which is herein incorporated by reference.

FIELD OF THE INVENTION

The invention relates to the field of industrial microbiology and the metabolism of lactic acid bacteria. More specifically, engineered genetic modifications were made to reduce or eliminate enzyme activity of endogenously expressed acetolactate decarboxylase and lactate dehydrogenase genes to increase availability of acetolactate as a substrate for biosynthesis of desired products, including isobutanol.

BACKGROUND OF THE INVENTION

Metabolic flux in biosynthetic pathways endogenous to lactic acid bacteria has been altered for production of products that use pyruvate as a starting substrate. In lactic acid bacteria the major pyruvate metabolic pathway is conversion to lactate through activity of lactate dehydrogenase (LDH). Metabolic engineering to redirect pyruvate from lactate to other products in lactic acid bacteria has had unpredictable results. Production of alanine in LDH-deficient *Lactococcus lactis* expressing alanine dehydrogenase was shown by Hols et al. (Nature Biotech. 17:588-592 (1999)). However, production of ethanol in LDH-deficient *Lactobacillus plantarum* expressing pyruvate decarboxylase was very limited, with carbon flow not significantly improved toward ethanol and lactate still produced (Liu et la.(2006) J. Ind. Micro. Biotech. 33:1-7).

In lactic acid bacteria pyruvate is also converted in a pathway to acetolactate, which is then converted to acetoin by acetolactate decarboxylase, and then to 2,3-butanediol. Additional pathways convert

acetolactate to diacetyl, valine or leucine. Monnet et al. (Applied and Environmental Microbiology 66:5518-5520 (2000)) have through chemical mutagenesis eliminated acetolactate decarboxylase activity and reduced LDH activity to increase acetolactate, acetoin, and diacetyl production. Disclosed in US Patent Application Publication No. 20100112655 is engineering high flux from pyruvate to 2,3-butanediol in lactic acid bacteria by expressing heterologous butanediol dehydrogenase activity and substantially eliminating lactate dehydrogenase activity.

Disclosed in co-pending US Patent Application Publication No. 2010-0081183 is engineering lactic acid bacteria for high dihydroxy-acid dehydratase (DHAD) activity by expressing a heterologous DHAD and substantially eliminating lactate dehydrogenase activity. DHAD is one of the enzymes in a biosynthetic pathway for synthesis of isobutanol that is disclosed in co-pending US Patent Pub No. US20070092957 A1. Disclosed therein is engineering of recombinant microorganisms for production of isobutanol. Isobutanol is useful as a fuel additive, whose availability may reduce the demand for petrochemical fuels.

Disclosed in de Vos et al. (Int. Dairy J. 8:227-233 (1998)) is that it has appeared impossible to combine inactivation of *aldB*, encoding acetolactate decarboxylase, with inactivation of *ldh,* encoding lactate dehydrogenase, in rapidly growing cells of lactic acid bacteria.

There remains a need for altering metabolic flux in lactic acid bacteria away from lactate and away from the acetoin to 2,3-butanediol pathway, and into other biosynthetic pathways downstream of acetolactate, such as for production of isobutanol.

SUMMARY OF THE INVENTION

Disclosed herein are lactic acid bacteria cells that are genetically modified to eliminate lactate dehydrogenase activity and reduce or eliminate acetolactate decarboxylase activity as expressed endogenously by genes encoding lactate dehydrogenase *(ldh*) and acetolactate decarboxylase (*aldB*). The cells lack detectable dehydrogenase and acetolactate decarboxylase enzyme activity. These cells may be used to produce isobutanol and other products having acetolactate as an intermediate.

Accordingly, a recombinant lactic acid bacteria cell comprising at least one engineered genetic modification that reduces or eliminates enzyme activity of endogenously expressed acetolactate decarboxylase and at least one engineered genetic modification that eliminates enzyme activity of endogenously expressed lactate dehydrogenase is provided.

In another embodiment the recombinant lactic acid bacteria cell may further comprise at least one genetic modification that reduces pyruvate formate lyase activity. Further genetic modifications may also be included, such as additional biosynthetic pathways and/or additional modifications that provide for utilization of various substrates or production of other products.

In another embodiment, a method for producing a recombinant lactic acid bacteria cell is provided, said method comprises:

a) providing a lactic acid bacteria cell;

b) modifying by genetic engineering at least one endogenous gene encoding lactate dehydrogenase in the cell of (a) to eliminate enzyme activity of endogenously expressed lactate dehydrogenase;

c) expressing acetolactate decarboxylase activity from a plasmid in the cell of (b) to create a cell with non-chromosomally expressed acetolactate decarboxylase;

(d) modifying by genetic engineering an endogenous gene encoding acetolactate decarboxylase in the cell of (c) to eliminate enzyme activity of endogenously expressed acetolactate decarboxylase; and

(e) curing the plasmid expressing acetolactate decarboxylase activity from the cell of (d);

whereby a recombinant lactic acid bacteria cell lacking enzyme activity of endogenously expressed lactate dehydrogenase and acetolactate decarboxylase is produced.

In yet another embodiment the invention provides a method for producing isobutanol comprising:

(a) providing a lactic acid bacteria cell comprising:

i) at least one genetic modification that eliminates enzyme activity of endogenously expressed acetolactate decarboxylase

and at least one genetic modification that eliminates enzyme activity of endogenously expressed lactate dehydrogenase; and

ii) an isobutanol biosynthetic pathway; and

(b) culturing the cell of (a) under conditions wherein isobutanol is produced.

In yet another embodiment the invention provides an integration vector for lactic acid bacteria comprising:

a) a Tn-5 transposase coding region operably linked to a promoter that is active in LAB cells;

b) Tn5IE and TN5OE elements bounding a selection marker active in lactic acid bacteria cells and a DNA segment targeted for integration;

c) a selection marker active in *E. coli* cells;

d) an origin of replication for E. coli cells;

e) an origin of replication for lactic acid bacteria cells that is temperature sensitive;

wherein the Tn5IE and TN5OE elements direct random integration of the DNA segment of b).

In yet another embodiment the invention provides a method for randomly integrating a DNA segment into the LAB cell genome comprising:

a) providing a vector comprising:

(i) a Tn-5 transposase coding region operably linked to a promoter that is active in lactic acid bacteria cells;

(ii) Tn5IE and TN5OE elements bounding a selection marker that is active in *E. coli* and lactic acid bacteria cells;

(iii) a second selection marker active in lactic acid bacteria cells;

(iv) an origin of replication for *E. coli* cells;

(v) an origin of replication for lactic acid bacteria cells that is conditionally active;

b) placing a DNA segment for integration between the elements of step a (ii) creating an integration construction;

c) transforming the integration construction into a lactic acid bacteria cell whereby transformed cells are produced;

d) growing and selecting the transformed cells of step (c) in permissive conditions using the selection marker of step a (ii) to produce selected transformants; and

e) growing the selected transformants of step (d) in nonpermissive conditions;

wherein the vector is cured from the lactic acid bacteria cells and the DNA segment for integration is randomly integrated into the genome of said lactic acid bacteria cell.

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCES

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

Figure 1 shows a diagram of biosynthetic pathways initiating with pyruvate in lactic acid bacteria.

Figure 2 shows biosynthetic pathways for biosynthesis of isobutanol.

The following sequences conform with 37 C.F.R. 1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (2009) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

Table 1 SEQ ID NOs of lactate dehydrogenase coding regions and proteins

Organism and gene name	SEQ ID NO:	SEQ IDNO:

	nucleic acid	amino acid
Lactobacillus plantarum ldhD	1	2
Lactobacillus plantarum ldhL1	3	4
Lactobacillus plantarum ldhL2	5	6
Lactococcus lactis ldhL	7	8
Leuconostoc mesenteroides ldhD	9	10
Streptococcus thermophilus ldhL	11	12
Pediococcus pentosaceus ldhD	13	14
Pediococcus pentosaceus ldhL	15	16
Lactobacillus acidophilus ldhL1	17	18
Lactobacillus acidophilus ldhL2	19	20
Lactobacillus acidophilus ldhD	21	22

Table 2 SEQ ID NOs of acetolactate decarboxylase coding regions and proteins

Organism and gene name	SEQ ID NO: nucleic acid	SEQ IDNO: amino acid
aldB from *Lactobacillus plantarum*	23	24
aldB from *Lactobacillus rhamnosus*	25	26
aldB from *Pediococcus pentosaceus*	27	28
aldB from *Leuconostoc mesenteroides*	29	30
aldB from *Oenococcus oeni*	31	32
aldB from *Enterococcus faecalis*	33	34
aldB from *Streptococcus mutans*	35	36
aldB from *Lactococcus lactis*	37	38

Table 3. SEQ ID NOs of pyruvate formate lyase and pyruvate formate lyase activating enzyme coding regions and proteins

Organism and gene name	SEQ ID NO: nucleic acid	SEQ IDNO: amino acid
PflB1 from *Lactobacillus plantarum*	39	40
PflB2 from *Lactobacillus plantarum*	41	42
PflA1 from *Lactobacillus plantarum*	43	44

PflA2 from *Lactobacillus plantarum*	45	46
Pfl from *Lactococcus lactis*	47	48
PflA from *Lactococcus lactis*	49	50
Pfl from *Streptococcus thermophilus*	51	52
PflA from *Streptococcus thermophilus*	53	54

Table 4 SEQ ID NOs of expression coding regions and proteins

Description	SEQ ID NO: nucleic acid	SEQ ID NO: amino acid
ALS from *Bacillus subtilis*	55	56
ALS from *Bacillus subtilis* coding region optimized for *Lactobacillus plantarum*	57	56*
ALS from *Klebsiella pneumoniae (budB)*	58	59
ALS from *Lactococcus lactis*	60	61
ALS from *Staphylococcus aureus*	62	63
ALS from *Listeria monocytogenes*	64	65
ALS from *Streptococcus mutans*	66	67
ALS from *Streptococcus thermophilus*	68	69
ALS from *Vibrio angustum*	70	71
ALS from *Bacillus cereus*	72	73
KARI: *ilvC* gene of *Lactococcus lactis*	74	75
KARI from Vibrio cholerae	76	77
KARI from *Pseudomonas aeruginosa* I	78	79
KARI from *Pseudomonas fluorescens*	80	81
DHAD from *Lactococcus lactis ilvD*	82	83
DHAD from *Streptococcus mutans ilvD*	84	85
branched chain keto acid decarboxylase from *Lactococcus lactis kivD*	86	87
Lactococcus lactis kivD opt for *L. plantarum*	88	87*
secondary alcohol dehydrogenase from	91	92

Achromobacter xylosoxidans sadB		
A. xylosoxidans sadB opt for *L. plantarum*	157	92*
Tn5 transposase	93	94

* same protein sequence encoded by native and optimized sequence

SEQ ID NOs:95 and 96 are transposase recognition sites Tn5IE and Tn5OE.

SEQ ID NO:97 is the sequence of plasmid pFP996.

SEQ ID NOs:89, 90, 98-113, 117, 118, 120-122, 124-129, 131-136, 139-142, 144-147, 149-151, 153, 154, 156, 159-169, 171-175, 178-182, and 184-190 are PCR and sequencing primers.

SEQ ID NO:114 is a ribosome binding site (RBS).

SEQ ID NO:115 is the sequence of plasmid pDM20-ilvD(L. lactis).

SEQ ID NO:116 is the sequence of plasmid pDM1.

SEQ ID NO:119 is the sequence of a PCR fragment including a RBS and ilvD coding region from *Lactococcus lactis.*

SEQ ID NO:123 is a right homologous arm DNA fragment containing the 5' portion of the *suf* operon (*sufC* and part of *sufD*).

SEQ ID NO:130 is a left homologous arm DNA fragment containing the native *suf* promoter and sequences upstream into the *feoBA* operon.

SEQ ID NO:137 is the sequence of plasmid pTN6.

SEQ ID NO:138 is the sequence of a Tn5IE-loxP-cm-Pspac-loxP cassette.

SEQ ID NO:143 is the Pnpr promoter.

SEQ ID NO:148 is a Pnpr-tnp fusion DNA fragment.

SEQ ID NO:152 is a PgroE promoter sequence.

SEQ ID NO:155 is a PCR fragment containing the *kivD(o)* coding region together with a RBS.

SEQ ID NO:157 is the sadB coding region optimized for expression in *L. plantarum*.

SEQ ID NO:158 a DNA fragment containing an RBS and *sadB(o)* coding region.

SEQ ID NO:170 is a PrrnC1 promoter.

SEQ ID NO:176 is the sequence of plasmid pDM5.

SEQ ID NO:177 is a lacI-PgroE/lacO fragment.

SEQ ID NO:183 is the sequence of plasmid pDM5-PldhL1-ilvC(L. lactis).

DETAILED DESCRIPTION

The present invention relates to recombinant lactic acid bacteria (LAB) cells that are modified by genetic engineering to reduce or eliminate enzyme activity of an endogenously expressed enzymes encoded by genes encoding acetolactate decarboxylase (aldB) and lactate dehydrogenase (ldh). The cells have reduced or no acetolactate decarboxylase and no lactate dehydrogenase activity due to reduced or eliminated expression from these modified genes. The present invention also relates to the method of obtaining LAB cells which lack acetolactate decarboxylase and lactate dehydrogenase activities with engineered genetic modifications in aldB and ldh, which requires expressing one of the activities non-chromosomally while the chromosomal gene is modified. The non-chromosomal gene is then eliminated.

In these cells there is increased flux from pyruvate to acetolactate but away from acetoin. These cells may be used to produce isobutanol and other products having acetolactate as an intermediate. Isobutanol is useful as a fuel or fuel additive for replacing fossil fuels.

The following abbreviations and definitions will be used for the interpretation of the specification and the claims.

As used herein, the terms “comprises,” “comprising,” “includes,” “including,” “has,” “having,” “contains” or “containing,” or any other variation thereof, are intended to cover a non-exclusive inclusion. For example, a composition, a mixture, process, method, article, or apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such composition, mixture, process, method, article, or apparatus. Further, unless expressly stated to the contrary, “or” refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied

by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

Also, the indefinite articles “a” and “an” preceding an element or component of the invention are intended to be nonrestrictive regarding the number of instances (i.e. occurrences) of the element or component. Therefore “a” or “an” should be read to include one or at least one, and the singular word form of the element or component also includes the plural unless the number is obviously meant to be singular.

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

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

The term “isobutanol biosynthetic pathway” refers to an enzyme pathway to produce isobutanol from pyruvate.

The term “lactate dehydrogenase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of pyruvate to lactate. Lactate dehydrogenases are known as EC 1.1.1.27 (L-lactate dehydrogenase) or EC 1.1.1.28 (D-lactate dehydrogenase).

The term “acetolactate decarboxylase” refers to a polypeptide (or polypeptides) having an enzyme activity that catalyzes the conversion of

acetolactate to acetoin. Acetolactate decarboxylases are known as EC 4.1.1.5.

The term "pyruvate formate lyase", also called "formate C-acetyltransferase", refers to a polypeptide having enzyme activity that catalyzes the conversion of pyruvate to formate. Pyruvate formate lyases are known as EC 2.3.1.54.

The term "pyruvate formate lyase activating enzyme", also called "formate C-acetyltransferase activating enzyme", refers to a polypeptide that is required for activity of pyruvate formate lyase. Formate C-acetyltransferase activating enzymes are known as EC 1.97.1.4.

The term "a facultative anaerobe" refers to a microorganism that can grow in both aerobic and anaerobic environments.

The term "carbon substrate" or "fermentable carbon substrate" refers to a carbon source capable of being metabolized by host organisms of the present invention and particularly carbon sources selected from the group consisting of monosaccharides, oligosaccharides, polysaccharides, and one-carbon substrates or mixtures thereof.

The term "gene" refers to a nucleic acid fragment that is capable of being expressed as a specific protein, optionally including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. "Native gene" refers to a gene as found in nature with its own regulatory sequences. "Chimeric gene" refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. "Endogenous gene" refers to a native gene in its natural location in the genome of an organism. A "foreign gene" or "heterologous gene" refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. "Heterologous gene" includes a native coding region, or portion thereof, that is reintroduced into the source organism in a form that is different from the corresponding native gene. For example, a heterologous

gene may include a native coding region that is a portion of a chimeric gene including non-native regulatory regions that is reintroduced into the native host. Also a foreign gene can comprise native genes inserted into a non-native organism, or chimeric genes. A "transgene" is a gene that has been introduced into the genome by a transformation procedure.

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

The term "promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of effecting the expression of that coding sequence (i.e., that the coding sequence is under the

transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term “expression”, as used herein, refers to the transcription and stable accumulation of sense (mRNA). Expression may also refer to translation of mRNA into a polypeptide.

As used herein the term “transformation” refers to the transfer of a nucleic acid molecule into a host cell, which may be maintained as a plasmid or integrated into the genome. Host cells containing the transformed nucleic acid molecules are referred to as “transgenic” or “recombinant” or “transformed” cells.

The terms "plasmid" and "vector" as used herein, refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA molecules. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell.

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

The term “codon-optimized” as it refers to coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA.

As used herein, an "isolated nucleic acid fragment" or "isolated nucleic acid molecule" will be used interchangeably and will mean a polymer of RNA or DNA that is single- or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases. An isolated nucleic acid fragment in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA or synthetic DNA.

A nucleic acid fragment is "hybridizable" to another nucleic acid fragment, such as a cDNA, genomic DNA, or RNA molecule, when a single-stranded form of the nucleic acid fragment can anneal to the other nucleic acid fragment under the appropriate conditions of temperature and solution ionic strength. Hybridization and washing conditions are well known and exemplified in Sambrook, J., Fritsch, E. F. and Maniatis, T. Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularly Chapter 11 and Table 11.1 therein (entirely incorporated herein by reference). The conditions of temperature and ionic strength determine the "stringency" of the hybridization. Stringency conditions can be adjusted to screen for moderately similar fragments (such as homologous sequences from distantly related organisms), to highly similar fragments (such as genes that duplicate functional enzymes from closely related organisms). Post-hybridization washes determine stringency conditions. One set of preferred conditions uses a series of washes starting with 6X SSC, 0.5% SDS at room temperature for 15 min, then repeated with 2X SSC, 0.5% SDS at 45 °C for 30 min, and then repeated twice with 0.2X SSC, 0.5% SDS at 50 °C for 30 min. A more preferred set of stringent conditions uses higher temperatures in which the washes are identical to those above except for the temperature of the final two 30 min washes in 0.2X SSC, 0.5% SDS was increased to 60 °C. Another preferred set of highly stringent conditions uses two final washes in 0.1X SSC, 0.1% SDS at 65 °C. An additional set of stringent conditions include hybridization at 0.1X SSC, 0.1% SDS, 65 °C and washes with 2X SSC, 0.1% SDS followed by 0.1X SSC, 0.1% SDS, for example.

Hybridization requires that the two nucleic acids contain complementary sequences, although depending on the stringency of the

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

A "substantial portion" of an amino acid or nucleotide sequence is that portion comprising enough of the amino acid sequence of a polypeptide or the nucleotide sequence of a gene to putatively identify that polypeptide or gene, either by manual evaluation of the sequence by one skilled in the art, or by computer-automated sequence comparison and identification using algorithms such as BLAST (Altschul, S. F., et al., *J. Mol. Biol.*, 215:403-410 (1993)). In general, a sequence of ten or more contiguous amino acids or thirty or more nucleotides is necessary in order to putatively identify a polypeptide or nucleic acid sequence as homologous to a known protein or gene. Moreover, with respect to nucleotide sequences, gene specific oligonucleotide probes comprising 20-30 contiguous nucleotides may be used in sequence-dependent methods of gene identification (e.g., Southern hybridization) and isolation (e.g., *in situ* hybridization of bacterial colonies or bacteriophage plaques).

In addition, short oligonucleotides of 12-15 bases may be used as amplification primers in PCR in order to obtain a particular nucleic acid fragment comprising the primers. Accordingly, a "substantial portion" of a nucleotide sequence comprises enough of the sequence to specifically identify and/or isolate a nucleic acid fragment comprising the sequence. The instant specification teaches the complete amino acid and nucleotide sequence encoding particular proteins. The skilled artisan, having the benefit of the sequences as reported herein, may now use all or a substantial portion of the disclosed sequences for purposes known to those skilled in this art. Accordingly, the instant invention comprises the complete sequences as reported in the accompanying Sequence Listing, as well as substantial portions of those sequences as defined above.

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

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

Preferred methods to determine identity are designed to give the best match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs.

Sequence alignments and percent identity calculations may be performed using the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc., Madison, WI). Multiple alignment of the sequences is performed using the "Clustal method of alignment" which encompasses several varieties of the algorithm including the "Clustal V method of alignment" corresponding to the alignment method labeled Clustal V (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci.,* 8:189-191 (1992)) and found in the MegAlign™ program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). For multiple alignments, the default values correspond to GAP PENALTY=10 and GAP LENGTH PENALTY=10. Default parameters for pairwise alignments and calculation of percent identity of protein sequences using the Clustal method are KTUPLE=1, GAP PENALTY=3, WINDOW=5 and DIAGONALS SAVED=5. For nucleic acids these parameters are KTUPLE=2, GAP PENALTY=5, WINDOW=4 and DIAGONALS SAVED=4. After alignment of the sequences using the Clustal V program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program. Additionally the "Clustal W method of alignment" is available and corresponds to the alignment method labeled Clustal W (described by Higgins and Sharp, *CABIOS*. 5:151-153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci.* 8:189-191(1992), Thompson, J. D., Higgins, D. G., and Gibson T. J. (1994) *Nuc. Acid Res.* 22: 4673 4680) and found in the MegAlign™ v6.1 program of the LASERGENE bioinformatics computing suite (DNASTAR Inc.). Default parameters for multiple alignment (GAP PENALTY=10, GAP LENGTH PENALTY=0.2, Delay Divergen Seqs(%)=30, DNA Transition Weight=0.5, Protein Weight Matrix=Gonnet Series, DNA Weight Matrix=IUB). After alignment of the sequences using the Clustal W program, it is possible to obtain a "percent identity" by viewing the "sequence distances" table in the same program.

It is well understood by one skilled in the art that many levels of sequence identity are useful in identifying polypeptides, from other species, wherein such polypeptides have the same or similar function or

activity. Useful examples of percent identities include, but are not limited to: 24%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95%, or any integer percentage from 24% to 100% may be useful in describing the present invention, such as 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99%. Suitable nucleic acid fragments not only have the above homologies but typically encode a polypeptide having at least 50 amino acids, preferably at least 100 amino acids, more preferably at least 150 amino acids, still more preferably at least 200 amino acids, and most preferably at least 250 amino acids.

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

Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, J., Fritsch,

E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989) (hereinafter "Maniatis"); and by Silhavy, T. J., Bennan, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-Interscience (1987).

Improved production of isobutanol in lactic acid bacteria

The present invention provides for greatly improved isobutanol production in lactic acid bacteria (LAB) cells having genetic modifications, e.g., deletions, in certain genes, said modifications provide for the elimination of lactate dehydrogenase and reduction or elimination of acetolactate enzyme activity in these cells.

The primary flux of pyruvate in LAB cells, which is to lactic acid, is altered by decreased expression of lactate dehydrogenase (Ldh) activity. With reduced Ldh activity, there may be increased flux of pyruvate to production of acetolactate via acetolactate synthase, and from acetolactate to acetoin (see Figure 1). Acetolactate decarboxylase catalyzes conversion of acetolactate to acetoin. Decreased lactate dehydrogenase activity in an acetolactate decarboxylase null LAB cell has been found to result in increases in acetolactate and in acetoin after about 20 hours of growth (Monnet et al. Appl and Envrt. Microbiology 66:5518-5520 (2000). Thus efficient conversion of acetolactate to acetoin occured even in the absence of acetolactate decarboxylase activity. Modifications to the LAB cells made in Monnet et al (ibid.) were made by chemical mutagenesis followed by screening for reduced enzyme activities. Thus the nature of the alterations to the genome are unknown, in contrast to when engineered genetic modifications are made.

In the present invention a method was developed to engineer genetic modifications to eliminate the enzyme activity encoded by lactate dehydrogenase and acetolactate decarboxylase genes in LAB cells. Elimination of enzyme activity according to the invention means elimination of appreciable or detectable levels in functional activity. These modifications could not be obtained using standard engineering methods.

It was found, as described herein that in a LAB cell with these modifications in the presence of an isobutanol biosynthetic pathway, isobutanol production was increased 6-fold over isobutanol production in a cell with ldh gene deletions but no aldB deletion. Thus the isobutanol pathway was able to effectively divert flux from production of acetoin from acetolactate.

Engineered genetic modifications to eliminate enzyme activity resulting from modifications to genes encoding lactate dehydrogenase and acetolactate decarboxylase may be made as described below in any LAB, which may also be engineered for the presence of an isobutanol biosynthetic pathway. The LAB which may be host cells in the present disclosure include, but are not limited to, *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, and *Streptococcus.*

Eliminating lactate dehydrogenase enzyme activity

In the present invention genetic modifications are engineered in LAB to eliminate the enzyme activity from expression of endogenous lactate dehydrogenase genes that are naturally expressed under growth conditions used during fermentation for product production. LAB may have one or more genes, typically one, two or three genes, encoding lactate dehydrogenase. For example, *Lactobacillus plantarum* has three genes encoding lactate dehydrogenase which are named ldhL2 (protein SEQ ID NO:6, coding region SEQ ID NO:5), ldhD (protein SEQ ID NO:2, coding region SEQ ID NO:1), and ldhL1 (protein SEQ ID NO:4, coding region SEQ ID NO:3). *Lactococcus lactis* has one gene encoding lactate dehydrogenase which is named ldhL (protein SEQ ID NO:8, coding region SEQ ID NO:7), and *Pediococcus pentosaceus* has two genes named ldhD (protein SEQ ID NO:14, coding region SEQ ID NO:13) and ldhL (protein SEQ ID NO:16, coding region SEQ ID NO:15).

Genetic modification is made in at least one gene encoding lactate dehydrogenase to eliminate its activity. When more than one lactate dehydrogenase gene is expressed (is active) under the growth conditions to be used for production, a genetic modification may be made in each of these active genes to affect their expression such that enzyme activity is eliminated. For example, in *L. plantarum* ldhL1 and ldhD genes are

modified. It is not necessary to modify the third gene, ldhL2, for growth in typical conditions because this gene appears to be inactive in these conditions. Typically, expression of one or more genes encoding lactate dehydrogenase is disrupted to eliminate the expressed enzyme activity. Examples of LAB lactate dehydrogenase genes that may be targeted for disruption are represented by the coding regions of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21 listed in Table 1. Other target genes, such as those encoding lactate dehydrogenase proteins having at least about 80-85%, 85%- 90%, 90%-95%, or at least about 96%, 97%, 98%, or 99% sequence identity to a lactate dehydrogenase of SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, or 22 listed in Table 1 may be identified in the literature and using bioinformatics approaches, as is well known to the skilled person, since lactate dehydrogenases are well known. Typically BLAST (described above) searching of publicly available databases with known lactate dehydrogenase amino acid sequences, such as those provided herein, is used to identify lactate dehydrogenases, and their encoding sequences, that may be targets for disruption to eliminate expressed lactate dehydrogenase activity. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

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

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

In the present LAB cells at least one engineered genetic modification is made that affects expression of the target gene encoding lactate dehydrogenase such that enzyme activity is eliminated. Any genetic modification method known by one skilled in the art for eliminating expression of a gene may be used to eliminate expressed enzyme activity. Methods include, but are not limited to, deletion of the entire or a portion of the lactate dehydrogenase encoding gene, inserting a DNA fragment into the lactate dehydrogenase encoding gene (in either the promoter or coding region) so that the encoded protein cannot be expressed or expression does not occur to a level sufficient for the production of enzyme activity, introducing a mutation into the lactate dehydrogenase coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the lactate dehydrogenase coding region to alter amino acids so that a non-functional protein is expressed. In addition lactate dehydrogenase expression may be blocked by expression of an antisense RNA or an interfering RNA, and constructs may be introduced that result in cosuppression. All of these methods may be readily practiced by one skilled in the art making use of the known lactate dehydrogenase encoding sequences such as those of SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, and 21.

For some methods genomic DNA sequences that surround a lactate dehydrogenase encoding sequence are useful, such as for homologous recombination-based methods. These sequences may be available from genome sequencing projects such as for *Lactobacillus plantarum*, which is available through the National Center for Biotechnology Information (NCBI) database, with Genbank ™ identification gi|28376974|ref|NC_004567.1|[28376974]. Adjacent genomic DNA sequences may also be obtained by sequencing outward from a lactate dehydrogenase coding sequence using primers within the coding sequence, as well known to one skilled in the art.

A particularly suitable method for eliminating enzyme activity of a lactate dehydrogenase, as exemplified herein in Example 1, is using homologous recombination mediated by lactate dehydrogenase coding region flanking DNA sequences to delete the entire gene that encodes lactate dehydrogenase. The flanking sequences are cloned adjacent to each other so that a double crossover event using these flanking sequences deletes the lactate dehydrogenase coding region.

Eliminating acetolactate decarboxylase enzyme activity

In the present invention a genetic modification is engineered in LAB cells to reduce or eliminate enzyme activity of endogenously expressed acetolactate decarboxylase gene. Genes encoding acetolactate decarboxylase in LAB cells are typically called aldB. However alternative names of ald and aldC have sometimes been used. Thus ald and aldC are interchangeable with aldB herein as referring to a gene encoding acetolactate decarboxylase, as are any other names referring to the same gene.

Examples of acetolactate decarboxylase genes from LAB that may be targeted for modification are represented by the coding regions of SEQ ID NOs:23, 25, 27, 29, 31, 33, 35, and 37 listed in Table 2. Other target genes, such as those encoding an acetolactate decarboxylase protein having at least about 80-85%, 85%- 90%, 90%-95%, or at least about 96%, 97%, 98% or 99% *s*equence identity to an acetolactate decarboxylase of SEQ ID NO:24, 26, 28, 30, 32, 34, 36, or 38 listed in Table 2 may be identified in the literature and using bioinformatics approaches, as is well known to the skilled person, since acetolactate decarboxylases are well known. Typically BLAST (described above) searching of publicly available databases with known acetolactate decarboxylase amino acid sequences, such as those provided herein, is used to identify acetolactate decarboxylases, and their encoding sequences, that may be targets for modification to eliminate enzyme activity of acetolactate decarboxylase. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

Additionally, the acetolactate decarboxylase encoding sequences described herein or those recited in the art may be used to identify other homologs in nature as described above. In the present LAB cells at least one engineered genetic modification is made that affects expression of the target gene encoding acetolactate decarboxylase such that the enzyme activity of acetolactate decarboxylase is reduced or eliminated. Modifications are made as described for modifying the lactate dehydrogenase gene, using the method for combining ldh and aldB modifications as described below.

Transient expression allows ldh and ald gene knockouts

Similarly to what others had reported previously (de Vos et al. (1998) Int. Dairy J. 8:227-233), applicants were unable to recover a strain following genetic modification to eliminate aldB expression in LAB cells with genetic modifications engineered to eliminate expression of ldh genes as described in Example 4 herein. Both ldh genes that are active in typical growth conditions in *Lactobacillus plantarum*, ldhD and ldhL, had been modified to eliminate their expression.

In the present invention, acetolactate decarboxylase activity is expressed from a plasmid in a cell with ldh gene expression eliminated (as described above), during engineering of the chromosomal aldB gene. In the presence of the non-chromosomally expressed (from a plasmid) acetolactate decarboxylase activity, a genetic modification is engineered in the endogenous aldB gene to reduce or eliminate its expression. Then the plasmid is cured from the cell creating a cell with modifications that results in elimination of the enzyme activity resulting from expression of ldh and reduction or elimination of the enzyme activity resulting from expression of aldB genes. Through this method, cells with engineered modifications such that they are lacking lactate dehydrogenase activity and lack or have reduced acetolactate decarboxylase activity may be recovered.

Alternatively, lactate dehydrogenase activity may be expressed from a plasmid in a cell with aldB gene expression eliminated, during engineering of a chromosomal ldh gene. If more than one ldh gene is active, expression of one ldh gene may be eliminated prior to expressing lactate dehydrogenase activity from a plasmid. Then expression of the

second ldh gene is eliminated. Then the plasmid is cured from the cell creating a cell with modifications that affects expression of ldh and aldB genes such that enzyme activity is eliminated. Through this method, engineered cells lacking lactate dehydrogenase activity and acetolactate decarboxylase activity may be recovered.

Alternatively, lactate dehydrogenase activity may be expressed from a plasmid in a cell with ldh gene expression eliminated, during engineering of a chromosomal aldB gene. Then the plasmid is cured from the cell creating a cell with modifications that eliminate expression of ldh and reduce or eliminate expression of aldB genes. Through this method, engineered cells lacking lactate dehydrogenase activity and acetolactate decarboxylase activity may be recovered.

Acetolactate decarboxylase or lactate dehydrogenase activity may be expressed from a plasmid as is well known to one skilled in the art. Any of the sequences encoding acetolactate decarboxylase that are provided herein as SEQ ID NOs: 23, 25, 27, 29, 31, 33, 35, 37, or any acetolactate decarboxylase coding regions additionally identified through bioinformatics or experimental methods as described above, may be operably linked to a promoter for expression in LAB from a chimeric gene. Additionally, suitable acetolactate decarboxylase enzymes are classified as EC number 4.1.1.5. Alternatively, any of the sequences encoding lactate dehydrogenase that are provided herein as SEQ ID NOs:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, or any lactate dehydrogenase coding regions additionally identified through bioinformatics or experimental methods as described above, may be operably linked to a promoter for expression in LAB from a chimeric gene. Additionally, suitable lactate dehydrogenase enzymes are classified as EC number EC 1.1.1.27 (L-lactate dehydrogenase) or EC 1.1.1.28 (D-lactate dehydrogenase). A ribosome binding site and a termination control region may be included in the chimeric expression gene. The chimeric gene is typically constructed in an expression vector or plasmid containing a selectable marker and sequences allowing autonomous replication in LAB cells. In addition, a native ldh or aldB gene with a native promoter that is active in LAB cells may be used for expression from a plasmid.

Initiation control regions or promoters which are useful to drive expression of an acetolactate decarboxylase or lactate dehydrogenase coding region in LAB cells are familiar to those skilled in the art. Some examples include the *amy*, *apr*, *npr* and *rrnC1* promoters; *nisA* promoter (useful for expression Gram-positive bacteria (Eichenbaum et al. *Appl. Environ. Microbiol*. 64(8):2763-2769 (1998)); and the synthetic P11 promoter (useful for expression in *Lactobacillus plantarum*, Rud et al., *Microbiology* 152:1011-1019 (2006)). In addition, the *ldhL1,* and *fabZ1* promoters of *L plantarum* are useful for expression of chimeric genes in LAB. The *fabZ1* promoter directs transcription of an operon with the first gene, *fabZ1*, encoding (3R)-hydroxymyristoyl-[acyl carrier protein] dehydratase.

Termination control regions may also be derived from various genes, typically from genes native to the preferred hosts. Optionally, a termination site may be unnecessary, however, it is most preferred if included.

Vectors or plasmids useful in LAB cells include those having two origins of replication and two selectable markers which allow for replication and selection in both *Escherichia coli* and LAB. An example is pFP996, the sequence of which is provided as SEQ ID NO:97, which is useful in *L. plantarum*.and other LAB. Many plasmids and vectors used in the transformation of *Bacillus subtilis* and *Streptococcus* may be used generally for *LAB.* Non-limiting examples of suitable vectors include pAMβ1 and derivatives thereof (Renault et al., *Gene* 183:175-182 (1996); and O'Sullivan et al., *Gene* 137:227-231 (1993)); pMBB1 and pHW800, a derivative of pMBB1 (Wyckoff et al. *Appl. Environ. Microbiol*. 62:1481-1486 (1996)); pMG1, a conjugative plasmid (Tanimoto et al., *J. Bacteriol*. 184:5800-5804 (2002)); pNZ9520 (Kleerebezem et al., *Appl. Environ. Microbiol*. 63:4581-4584 (1997)); pAM401 (Fujimoto et al., *Appl. Environ. Microbiol*. 67:1262-1267 (2001)); and pAT392 (Arthur et al., *Antimicrob. Agents Chemother.* 38:1899-1903 (1994)). Several plasmids from *Lactobacillus plantarum* have also been reported (e.g., van Kranenburg R, Golic N, Bongers R, Leer RJ, de Vos WM, Siezen RJ, Kleerebezem M. *Appl. Environ. Microbiol*. 2005 Mar; 71(3): 1223-1230).

Vectors or plasmids may be introduced into a host cell using methods known in the art, such as electroporation (Cruz-Rodz et al. *Molecular Genetics and Genomics* 224:1252-154 (1990), Bringel, et al. *Appl. Microbiol. Biotechnol.* 33: 664-670 (1990), Alegre et al., *FEMS Microbiology letters* 241:73-77 (2004)), and conjugation (Shrago et al., *Appl. Environ. Microbiol.* 52:574-576 (1986)).

Following recovery of cells with ldh and aldB modifications, the cells are cured of the expression plasmid. Curing of the plasmid may be accomplished by any method known to one skilled in the art. Typically a temperature sensitive origin of replication is used, where growth of plasmid-harboring cells at the restrictive temperature causes the plasmid to be lost. Another method, for example, is to place a negative selection marker on the plasmid to be cured, where growth in the presence of the selective agent causes the plasmid to be lost.

Reducing pyruvate formate lyase activity

In addition to the modifications described above of ldh and aldB genes in the present cells, they may optionally have at least one modification that reduces endogenous pyruvate formate lyase activity. Pyruvate formate lyase activity converts pyruvate to formate (see Figure 1). Activity of pyruvate formate lyase in the cell may be reduced or eliminated. Preferably the activity is eliminated.

For expression of pyruvate formate lyase activity a gene encoding pyruvate formate lyase (pfl) and a gene encoding pyruvate formate lyase activating enzyme are required. To reduce pyruvate formate lyase activity a modification may be made in either or both of these genes. There may be one or more genes encoding each of pyruvate formate lyase and pyruvate formate lyase activating enzyme in a particular strain of LAB. For example, *Lactobacillus plantarum* WCFS1 contains two pfl genes (pflB1: coding region SEQ ID NO:39, protein SEQ ID NO:40; and pflB2: coding region SEQ ID NO:41, protein SEQ ID NO:42) and two pfl activating enzyme genes (pflA1: coding region SEQ ID NO:43, protein SEQ ID NO:44; and pflA2: coding region SEQ ID NO:45, protein SEQ ID NO:46), *Lactobacillus plantarum* PN0512 only contains one pfl gene (pflB2) and one pfl activating enzyme gene (pflA2). Expression is reduced for all pfl

encoding genes that are active in a production host cell under the desired production conditions and/or for all pfl activating enzyme encoding genes that are active in a production host cell under the desired production conditions.

Examples of pfl genes that may be modified to reduce pyruvate formate lyase activity are represented by the coding regions of SEQ ID NOs:39, 41, 47, and 51. Other target genes for modification include those encoding pyruvate formate lyase proteins having SEQ ID NOs:40, 42, 48, and 52 and those encoding a protein having at least about 80-85%, 85%-90%, 90%-95%, or at least about 96%, 97%, 98%, or 99% sequence identity to one of these proteins, which may be identified in the literature and using bioinformatics approaches, as is well known to the skilled person as described above for lactate dehydrogenase proteins. Additionally, the sequences described herein or those recited in the art may be used to identify other homologs in nature as described above.

Examples of pfl activating enzyme genes that may be modified to reduce pyruvate formate lyase activity are represented by the coding regions of SEQ ID NOs:43, 45, 49, and 53. Other target genes for modification include those encoding pyruvate formate lyase activating enzyme proteins having SEQ ID NOs:44, 46, 50, 54 and those encoding a protein having at least about 80-85%, 85%- 90%, 90%-95%, or at least about 96%, 97%, 98%, or 99% sequence identity to one of these proteins, which may be identified in the literature and using bioinformatics approaches, as is well known to the skilled person as described above for lactate dehydrogenase proteins. Additionally, the sequences described herein or those recited in the art may be used to identify other homologs in nature as described above.

Any genetic modification method known by one skilled in the art for reducing the expression of a protein may be used to alter pyruvate formate lyase activity. Methods to reduce or eliminate expression of the pyruvate formate lyase and/or pyruvate formate lyase activating enzyme genes include, but are not limited to, deletion of the entire or a portion of the gene, inserting a DNA fragment into the gene (in either the promoter or coding region) so that the encoded protein cannot be expressed or has

reduced expression, introducing a mutation into the coding region which adds a stop codon or frame shift such that a functional protein is not expressed, and introducing one or more mutations into the coding region to alter amino acids so that a non-functional or reduced-functional protein is expressed. In addition expression from the target gene may be partially or substantially blocked by expression of an antisense RNA or an interfering RNA, and constructs may be introduced that result in cosuppression.

Isobutanol production

In one embodiment of the present invention, a LAB cell with engineered modifications to the ldh and aldB genes as described above, and optionally reducing pyruvate formate lyase activity, produces isobutanol. Biosynthetic pathways for synthesis of isobutanol are disclosed in co-pending US Patent Pub No. US20070092957 A1, which is herein incorporated by reference. A diagram of the disclosed isobutanol biosynthetic pathways is provided in Figure 2. Production of isobutanol in a genetically engineered LAB cell disclosed herein is increased by eliminating the enzyme activity expressed by ldh and aldB genes, and increased further by eliminating expression of pfl and/or pflA genes.

In addition, an LAB host cell may be engineered for increased expression of Fe-S cluster forming proteins to improve the activity of the Fe-S cluster requiring dihydroxy-acid dehydratase enzyme of the isobutanol pathway as disclosed in co-pending US Patent Application Publication No. 20100081182, which is herein incorporated by reference. For example, expression of the endogenous suf operon encoding Fe-S cluster forming proteins may be increased as described in Example 2 herein.

As described in US Patent Pub No. US20070092957 A1, steps in an example isobutanol biosynthetic pathway include conversion of:

Pyruvate to acetolactate (Fig. 2 pathway step a) as catalyzed for example by acetolactate synthase (ALS) known by the EC number 2.2.1.6 9;

Acetolactate to 2,3-dihydroxyisovalerate (Fig. 2 pathway step b) as catalyzed for example by acetohydroxy acid isomeroreductase, also called ketol-acid reductoisomerase (KARI) known by the EC number 1.1.1.86;

2,3-dihydroxyisovalerate to α-ketoisovalerate (Fig. 2 pathway step c) as catalyzed for example by acetohydroxy acid dehydratase, also called dihydroxy-acid dehydratase (DHAD) known by the EC number 4.2.1.9;

α-ketoisovalerate to isobutyraldehyde (Fig. 2 pathway step d) as catalyzed for example by branched-chain α-keto acid decarboxylase known by the EC number 4.1.1.72; and

Isobutyraldehyde to isobutanol (Fig. 2 pathway step e) as catalyzed for example by branched-chain alcohol dehydrogenase known by the EC number 1.1.1.265, but may also be classified under other alcohol dehydrogenases (specifically, EC 1.1.1.1 or 1.1.1.2).

The substrate to product conversions, and enzymes involved in these reactions, for steps f, g, h, I, j, and k of alternative pathways are described in US Patent Pub No. US20070092957 A1.

Genes that may be used for expression of these enzymes, as well as those for two additional isobutanol pathways, are described in US Patent Pub No. US20070092957 A1, and additional genes that may be used can be identified in the literature and using bioinformatics approaches, as is well known to the skilled person as described above. Additionally, sequences provided therein may be used to isolate genes encoding homologous proteins using sequence-dependent protocols is well known in the art, as described above.

For example, some representative ALS enzymes that may be used include those encoded by *alsS* of *Bacillus* and *budB* of *Klebsiella* (Gollop et al., *J. Bacteriol.* 172(6):3444-3449 (1990); Holtzclaw et al., *J. Bacteriol.* 121(3):917-922 (1975)). ALS from *Bacillus subtilis* (DNA: SEQ ID NO:55; protein: SEQ ID NO:56), from *Klebsiella pneumoniae* (DNA: SEQ ID NO:58; protein:SEQ ID NO:59), and from *Lactococcus lactis* (DNA: SEQ ID NO:60; protein: SEQ ID NO:61) are provided herein. Additional Als coding regions and encoded proteins that may be used include those from *Staphylococcus aureus* (DNA: SEQ ID NO:62; protein:SEQ ID NO:63), *Listeria monocytogenes* (DNA: SEQ ID NO:64; protein:SEQ ID NO:65),

Streptococcus mutans (DNA: SEQ ID NO:66; protein:SEQ ID NO:67), *Streptococcus thermophilus* (DNA: SEQ ID NO:68; protein:SEQ ID NO:69), *Vibrio angustum* (DNA: SEQ ID NO:70; protein:SEQ ID NO:71), and *Bacillus cereus* (DNA: SEQ ID NO:72; protein:SEQ ID NO:73). Any Als gene that encodes an acetolactate synthase having at least about 80-85%, 85%- 90%, 90%- 95%, or at least about 96%, 97%, or 98% sequence identity to any one of those with SEQ ID NOs:56, 59, 61, 63, 65, 67, 69, 71, or 73 that converts pyruvate to acetolactate may be used. Identities are based on the Clustal W method of alignment using the default parameters of GAP PENALTY=10, GAP LENGTH PENALTY=0.1, and Gonnet 250 series of protein weight matrix.

Additionally, US Patent Application Publication No. 2009-0305363 provides a phylogenetic tree depicting acetolactate synthases that are the 100 closest neighbors of the *B. subtilis* AlsS sequence, any of which may be used. Additional Als sequences that may be used in the present strains may be identified in the literature and in bioinformatics databases as is well known to the skilled person. Identification of coding and/or protein sequences using bioinformatics is typically through BLAST (described above) searching of publicly available databases with known Als encoding sequences or encoded amino acid sequences, such as those provided herein. Identities are based on the Clustal W method of alignment as specified above. Additionally, the sequences listed herein or those recited in the art may be used to identify other homologs in nature as described above.

For example, KARI enzymes that may be used may be from the *ilvC* gene of *Lactococcus lactis (*DNA: SEQ ID NO:74; protein SEQ ID NO:75), *Vibrio cholerae (*DNA: SEQ ID NO:76; protein SEQ ID NO:77), *Pseudomonas aeruginosa* PAO1, *(*DNA: SEQ ID NO:78; protein SEQ ID NO:79), or *Pseudomonas fluorescens* PF5 (DNA: SEQ ID NO:80; protein SEQ ID NO:81). The later three are disclosed in US Patent Application Publication No, 20080261230, which is incorporated herein by reference. Additional KARI enzymes are described in US Application No. 61/246844,US Application Publication Nos. 2008026123, 2009016337, and 2010019751.

For example, DHAD enzymes that may be used may be from the *ilvD* gene of *Lactococcus lactis* (DNA: SEQ ID NO:82; protein SEQ ID NO:83) or *Streptococcus mutans* (DNA: SEQ ID NO:84; protein SEQ ID NO:85), and in addition sequences of DHAD coding regions and encoded proteins that may be used are provided in US Patent Application Publication No. 20100081183, which is incorporated herein by reference. This reference also includes descriptions for obtaining additional DHAD sequences that may be used.

For example, branched chain keto acid decarboxylase enzymes that may be used include one from the *kivD* gene of *Lactococcus lactis* (DNA: SEQ ID NO:86; protein SEQ ID NO:87) and others that may be identified by one skilled in the art using bioinformatics as described above.

For example, branched-chain alcohol dehydrogenases that may be used are known by the EC number 1.1.1.265, but may also be classified under other alcohol dehydrogenases (specifically, EC 1.1.1.1 or 1.1.1.2). These enzymes utilize NADH (reduced nicotinamide adenine dinucleotide) and/or NADPH as electron donors and sequences of branched-chain alcohol dehydrogenase enzymes and their coding regions that may be used are provided in US20070092957 A1.

In addition, useful for the last step of converting isobutyraldehyde to isobutanol is a new butanol dehydrogenase isolated from an environmental isolate of a bacterium identified as *Achromobacter xylosoxidans* (DNA: SEQ ID NO:91, protein SEQ ID NO:92) that is disclosed in US Patent Application Publication No. 20090269823, which is herein incorporated by reference

Improved activity of DHAD in LAB cells that are substantially free of lactate dehydrogenase activity was disclosed in US Patent Application Publication No. 20100081183, which is herein incorporated by reference. Additionally, increased expression of iron-sulfur cluster forming proteins to improve activity of DHAD was disclosed in US Patent Application Publication No. 20100081183, which is herein incorporated by reference.

Described in US Pub No. US20070092957 A1 is construction of chimeric genes and genetic engineering of LAB, exemplified by *Lactobacillus plantarum*, for isobutanol production using disclosed

biosynthetic pathways. Chimeric genes for pathway enzyme expression may be present in a cell on a replicating plasmid or integrated into the cell genome, as well known to one skilled in the art and described in Examples herein. A new method for integration developed herein is described below and used in Example 3.

Additional products

The present engineered LAB cells may be used for production of other products made from acetolactate that do not require acetolactate decarboxylase activity, to provide improved production. These products may include, but are not limited to valine, isoleucine, leucine, pantothenic acid (vitamin B5), 2-methyl-1-butanol, 3-methyl-1-butanol (isoamyl alcohol), and diacetyl. For production of these or other products the present LAB cells have in addition a biosynthetic pathway for the desired product, which may be endogenous, engineered, or a combination of both,

For example, a biosynthetic pathway for valine includes steps of acetolactate conversion to 2,3-dihydroxy-isovalerate by acetohydroxyacid reductoisomerase (ilvC), conversion of 2,3-dihydroxy-isovalerate to α-ketoisovalerate (also called 2-keto-isovalerate) by dihydroxy-acid dehydratase (ilvD), and conversion of α-ketoisovalerate to valine by branched-chain amino acid aminotransferase (ilvE). Biosynthesis of leucine includes the same steps to α-ketoisovalerate, followed by conversion of α-ketoisovalerate to leucine by enzymes encoded by leuA (2-isopropylmalate synthase), leuCD (isopropylmalate isomerase), leuB (3-isopropylmalate dehydrogenase), and tyrB/ ilvE (aromatic amino acid transaminase). Biosynthesis of pantothenate includes the same steps to α-ketoisovalerate, followed by conversion of α-ketoisovalerate to pantothenate by enzymes encoded by panB (3-methyl-2-oxobutanoate hydroxymethyltransferase), panE (2-dehydropantoate reductase), and panC (pantoate-beta-alanine ligase). Engineering expression of enzymes for enhanced production of pantothenic acid in microorganisms is described in US Patent No. 6177264.

2-methyl-1-butanol and 3-methyl-1-butanol may be produced by converting 2-ketoacids from amino acid biosynthetic pathways using 2-

ketoacid decarboxylases and alcohol dehyddrogenases (Atsumi and Liao (2008) Current Opinion in Biotechnology 19:414-419).

In combination with the elimination of ldh and aldB expression, increased expression of at least one gene in any of these pathways may be used to increase the production of the product of the pathway. Though some LAB naturally have the branched chain amino acid pathways for valine, isoleucine and leucine such as *Lactococcus lactis*, others such as *Lactobacillus plantarum* do not. LAB without an endogenous pathway producing the desired product, or precursor to a desired product, require engineering for expression of the missing pathway enzymes. One skilled in the art can readily assess which enzymes are present and missing for a desired pathway.

Diacetyl is produced from acetolactate spontaneously in the presence of oxygen, requiring no enzyme activity.

Tn5-mediated transposition in LAB

For long term maintenance and stability of foreign gene expression, such as for genes expressing enzymes of a desired biosynthetic pathway, it may be desired to integrate the expression gene into the cell genome. A vector was prepared herein to make use of the Tn5 transposition system in LAB cells, It was found that random integration into the genome of LAB cells was achieved using the Tn5 transposition vector developed herein. For integration, the vector includes a Tn5 transposase coding region (SEQ ID NO:93; encoded protein SEQ ID NO:94) operably linked to and expressed from a promoter that is active in LAB cells, examples of which are listed above, and transposase recognition sequences Tn5IE and Tn5OE (SEQ ID NOS:95 and 96). Any sequence that encodes a protein having at least about 90%, 95%, or 99% sequence identity with SEQ ID NO:94 and havingTn5 transposase activity may be used in the vector. Between Tn5IE and Tn5OE are a chloramphenicol resistance gene flanked by Cre recombinase sites, and a multiple cloning site (MCS). Any selection marker active in *E. coli* and LAB cells may substitute for the chloramphenicol resistance gene, examples of which are tetracycline resistance, spectinomycin resistance, and erythromycin resistance markers. The Cre recombinase sites are optional. In addition the vector

has a second marker gene, which is used for screening for transposition and loss of the Tn5 transposition vector. The second marker may be any marker active in LAB cells, including any of those listed above. The vector also has origins of replication for *E. coli* and LAB, the LAB origin being conditionally active, such as temperature sensitive. DNA segments placed between the Tn5IE and Tn5OE elements, typically in the MCS, may be randomly integrated into the genome of LAB cells using this vector. The described vector with a DNA segment between the Tn5IE and Tn5OE elements is an integration construction. For example, the vector has a termperature sensitive origin of replication for lactic acid bacteria cells and the chloramphenicol resistance marker is used to select transformants. The transformants are grown in permissive conditions (temperature typically of 30°C) for approximately 10 generations during which integration occurs. Transformants are then grown in nonpermissive conditions (temperature typically of 37°C) for approximately 20 generations to cure the plasmid, and chloramphenicol resistant colonies are screened for erythromycin sensitivity (loss of second marker) to confirm loss of the plasmid. The chloramphenicol resistance marker may be excised by expression of Cre recombinase in the cell, typically from a chimeric gene on a plasmid as is well known in the art.

Growth for production

Recombinant LAB cells disclosed herein may be used for fermentation production of isobutanol or other products as follows. The recombinant cells are grown in fermentation media which contains suitable carbon substrates. Suitable substrates may include but are not limited to monosaccharides such as glucose and fructose, oligosaccharides such as lactose or sucrose, polysaccharides such as starch or cellulose or mixtures thereof and unpurified mixtures from renewable feedstocks such as cheese whey permeate, cornsteep liquor, sugar beet molasses, and barley malt.

Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention, preferred carbon substrates are glucose, fructose, and sucrose. Sucrose

may be derived from renewable sugar sources such as sugar cane, sugar beets, cassava, sweet sorghum, and mixtures thereof. Glucose and dextrose may be derived from renewable grain sources through saccharification of starch based feedstocks including grains such as corn, wheat, rye, barley, oats, and mixtures thereof. In addition, fermentable sugars may be derived from renewable cellulosic or lignocellulosic biomass through processes of pretreatment and saccharification, as described, for example, in U.S. Patent Application Publication No. 2007/0031918A1, which is herein incorporated by reference. Biomass refers to any cellulosic or lignocellulosic material and includes materials comprising cellulose, and optionally further comprising hemicellulose, lignin, starch, oligosaccharides and/or monosaccharides. Biomass may also comprise additional components, such as protein and/or lipid. Biomass may be derived from a single source, or biomass can comprise a mixture derived from more than one source; for example, biomass may comprise a mixture of corn cobs and corn stover, or a mixture of grass and leaves. Biomass includes, but is not limited to, bioenergy crops, agricultural residues, municipal solid waste, industrial solid waste, sludge from paper manufacture, yard waste, wood and forestry waste. Examples of biomass include, but are not limited to, corn grain, corn cobs, crop residues such as corn husks, corn stover, grasses, wheat, wheat straw, barley, barley straw, hay, rice straw, switchgrass, waste paper, sugar cane bagasse, sorghum, soy, components obtained from milling of grains, trees, branches, roots, leaves, wood chips, sawdust, shrubs and bushes, vegetables, fruits, flowers, animal manure, and mixtures thereof.

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

Typically cells are grown at a temperature in the range of about 25 °C to about 40 °C in an appropriate medium. Suitable growth media are common commercially prepared media such as Bacto Lactobacilli MRS broth or Agar (Difco), Luria Bertani (LB) broth, Sabouraud Dextrose (SD)

broth or Yeast Medium (YM) broth. Other defined or synthetic growth media may also be used, and the appropriate medium for growth of the particular bacterial strain will be known by one skilled in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, e.g., cyclic adenosine 2':3'-monophosphate, may also be incorporated into the fermentation medium.

Suitable pH ranges for the fermentation are between pH 3.0 to pH 9.0, where pH 6.0 to pH 8.0 is preferred as the initial condition.

Fermentations may be performed under aerobic or anaerobic conditions, where anaerobic or microaerobic conditions are preferred.

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

Methods for Isobutanol Isolation from the Fermentation Medium

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

EXAMPLES

The meaning of abbreviations is as follows: "s" means second(s), "min" means minute(s), "h" means hour(s), "psi" means pounds per square inch, "nm" means nanometers, "d" means day(s), "µl" means microliter(s), "ml" means milliliter(s), "L" means liter(s), "mm" means millimeter(s), "nm" means nanometers, "mM" means millimolar, "M" means molar, "mmol"

means millimole(s), "µmol" means micromole(s)", "g" means gram(s), "µg" means microgram(s) and "ng" means nanogram(s), "PCR" means polymerase chain reaction, "OD" means optical density, "OD_{600}" means the optical density measured at a wavelength of 600 nm, "kDa" means kilodaltons, "*g*" means the gravitation constant, "bp" means base pair(s), "kbp" means kilobase pair(s), "% w/v" means weight/volume percent, % v/v" means volume/volume percent, "wt %" means percent by weight, "HPLC" means high performance liquid chromatography, and "GC" means gas chromatography. The term "molar selectivity" is the number of moles of product produced per mole of sugar substrate consumed and is reported as a percent. "SLPM" stands for Standard Liters per Minute (of air), "dO" is dissolved oxygen, q_p is "specific productivity" measured in grams isobutanol per gram of cells over time.

General Methods

Recombination plasmids were constructed using standard molecular biology methods known in the art. All restriction and modifying enzymes and Phusion High-Fidelity PCR Master Mix were purchased from New England Biolabs (Ipswich, MA). DNA fragments were purified with Qiaquick PCR Purification Kit (Qiagen Inc., Valencia, CA). Plasmid DNA was prepared with QIAprep Spin Miniprep Kit (Qiagen Inc., Valencia, CA). *L. plantarum* PN0512 genomic DNA was prepared with MasterPure DNA Purification Kit (Epicentre, Madison, WI). Oligonucleotides were synthesized by Sigma-Genosys (Woodlands, TX) or Invitrogen Corp (Carlsbad, CA).

Transformation

Lactobacillus plantarum PN0512 was transformed by the following procedure: 5 ml of Lactobacilli MRS medium (Accumedia, Neogen Corporation, Lansing, MI) containing 1% glycine (Sigma-Aldrich, St. Louis, MO) was inoculated with PN0512 cells and grown overnight at 30°C. 100 ml MRS medium with 1% glycine was inoculated with overnight culture to an OD600 of 0.1 and grown to an OD600 of 0.7 at 30°C. Cells were harvested at 3700xg for 8 min at 4 °C, washed with 100 ml cold 1 mM $MgCl_2$ (Sigma-Aldrich, St. Louis, MO), centrifuged at 3700xg for 8 min at 4

°C, washed with 100 ml cold 30% PEG-1000 (Sigma-Aldrich, St. Louis, MO), recentrifuged at 3700xg for 20 min at 4 °C, then resuspended in 1 ml cold 30% PEG-1000. 60 µl cells were mixed with ~100 ng plasmid DNA in a cold 1 mm gap electroporation cuvette and electroporated in a BioRad Gene Pulser (Hercules, CA) at 1.7 kV, 25 µF, and 400 Ω. Cells were resuspended in 1 ml MRS medium containing 500 mM sucrose (Sigma-Aldrich, St. Louis, MO) and 100 mM $MgCl_2$, incubated at 30°C for 2 hrs, plated on MRS medium plates containing 1 or 2 µg/ml of erythromycin (Sigma-Aldrich, St. Louis, MO), then placed in an anaerobic box containing a Pack-Anaero sachet (Mitsubishi Gas Chemical Co., Tokyo, Japan) and incubated at 30°C.

Example 1

Construction of the *ilvD* integration vector and PN0512ΔldhDΔldhL1::ilvDLl$^+$ integration strain

This example describes integration of the *Lactococcus lactis ilvD* gene into the chromosome of *L. plantarum* strain PN0512 ΔldhDΔldhL1 for expression of DHAD. The construction of *L. plantarum* PN0512 ΔldhDΔldhL1 was described in Example 1 of co-pending US Patent Appln No. 61/100786, which is herein incorporated by reference. This strain is deleted for the two genes that encode the major lactate dehydrogenases: *ldhD* and *ldhL1*. The double deletion was made in *Lactobacillus plantarum* PN0512 (ATCC strain # PTA-7727).

Gene knockouts were constructed using a process based on a two-step homologous recombination procedure to yield unmarked gene deletions (Ferain et al., 1994, *J. Bact.* 176:596). The procedure utilized a shuttle vector, pFP996 (SEQ ID NO:97). pFP996 is a shuttle vector for gram-positive bacteria. It can replicate in both *E. coli* and gram-postive bacteria. It contains the origins of replication from pBR322 (nucleotides #2628 to 5323) and pE194 (nucleotides #43 to 2627). pE194 is a small plasmid isolated originally from a gram positive bacterium, *Staphylococcus aureus* (Horinouchi and Weisblum J. Bacteriol. (1982) 150(2):804-814). In pFP996, the multiple cloning sites (nucleotides #1 to 50) contain restriction sites for EcoRI, BglII, XhoI, SmaI, ClaI, KpnI, and HindIII. There are two

antibiotic resistance markers; one is for resistance to ampicillin and the other for resistance to erythromycin. For selection purposes, ampicillin was used for transformation in *E. coli* and erythromycin was used for selection in *L. plantarum*.

Two segments of DNA, each containing 900 to 1200 bp of sequence either upstream or downstream of the intended deletion, were cloned into the plasmid to provide the regions of homology for the two genetic cross-overs. Cells were grown for an extended number of generations (30-50) to allow for the cross-over events to occur. The initial cross-over (single cross-over) integrated the plasmid into the chromosome by homologous recombination through one of the two homology regions on the plasmid. The second cross-over (double cross-over) event yielded either the wild type sequence or the intended gene deletion. A cross-over between the sequences that led to the initial integration event would yield the wild type sequence, while a cross-over between the other regions of homology would yield the desired deletion. The second cross-over event was screened for by antibiotic sensitivity. Single and double cross-over events were analyzed by PCR and DNA sequencing.

ΔldhD

The knockout cassette to delete the *ldhD* gene was created by amplifying from PN0512 genomic DNA an upstream flanking region with primers Top D F1 (SEQ ID NO:98) containing an EcoRI site and Top D R1 (SEQ ID NO:99). The downstream homology region including part of the coding sequence of *ldhD* was amplified with primers Bot D F2 (SEQ ID NO:100) and Bot D R2 (SEQ ID NO:101) containing an XhoI site. The two homology regions were joined by PCR SOE as follows. The 0.9 kbp upstream and downstream PCR products were gel-purified. The PCR products were mixed in equal amounts in a PCR reaction and re-amplified with primers Top D F1 and Bot D R2. The final 1.8 kbp PCR product was gel-purified and TOPO cloned into pCR4BluntII-TOPO (Invitrogen) to create vector pCRBluntII::ldhD. To create the integration vector carrying the internal deletion of the *ldhD* gene, pFP996 was digested with EcoRI and XhoI and the 5311-bp fragment gel-purified. Vector pCRBluntII::ldhD was digested with EcoRI and XhoI and the 1.8 kbp fragment gel- purified.

The *ldhD* knockout cassette and vector were ligated using T4 DNA ligase, resulting in vector pFP996::ldhD ko.

Electrocompetent *Lactobacillus plantarum* PN0512 cells were prepared, transformed with pFP996::ldhD ko, and plated on MRS containing 1 µg/ml of erythromycin. To obtain the single-crossover event (sco), transformants were passaged for approximately 50 generations in MRS medium at 37 °C. After growth, aliquots were plated for single colonies on MRS containing 1 µg/ml of erythromycin. The erythromycin-resistant colonies were screened by PCR amplification with primers ldhD Seq F1 (SEQ ID NO:102) and D check R (SEQ ID NO:103) to distinguish between wildtype and clones carrying the sco event. To obtain clones with a double crossover, the sco strains were passaged for approximately 30 generations in MRS medium with 20 mM D, L-lactate (Sigma, St. Louis, MO) at 37 °C and then plated for single colonies on MRS with lactate. Colonies were picked and patched onto MRS with lactate and MRS with lactate containing 1 µg/ml of erythromycin to find colonies sensitive to erythromycin. Sensitive colonies were screened by PCR amplification using primer D check R (SEQ ID NO:103) and D check F3 (SEQ ID NO:104). Wildtype colonies gave a 3.2 kbp product and deletion clones, called PN0512*ΔldhD,* gave a 2.3 kbp PCR product.

ΔldhDΔldhL1

A deletion of the *ldhL1* gene was made in the PN0512*ΔldhD* strain background in order to make a double *ΔldhL1ΔldhD* deletion strain. The knockout cassette to delete the *ldhL1* gene was amplified from PN0512 genomic DNA. The *ldhL1* left homologous arm was amplified using primers oBP31 (SEQ ID NO:105) containing a BglII restriction site and oBP32 (SEQ ID NO:106) containing an XhoI restriction site. The *ldhL1* right homologous arm was amplified using primers oBP33 (SEQ ID NO:107) containing an XhoI restriction site and oBP34 (SEQ ID NO:108) containing an XmaI restriction site. The *ldhL1* left homologous arm was cloned into the BglII/XhoI sites and the *ldhL1* right homologous arm was cloned into the XhoI/XmaI sites of pFP996pyrFΔerm, a derivative of pFP996. pFP996pyrFΔerm contains the *pyrF* sequence (SEQ ID NO:109) encoding orotidine-5'-phosphate decarboxylase from *Lactobacillus*

plantarum PN0512 in place of the erythromycin coding region in pFP996. The plasmid-borne *pyrF* gene, in conjunction with the chemical 5-fluoroorotic acid in a Δ*pyrF* strain, can be used as an effective counter-selection method in order to isolate the second homologous crossover. The XmaI fragment containing the *ldhL1* homologous arms was isolated following XmaI digestion and cloned into the XmaI restriction site of pFP996, yielding a 900 bp left homologous region and a 1200 bp right homologous region resulting in vector pFP996-ldhL1-arms.

PN0512Δ*ldhD* was transformed with pFP996-ldhL1-arms and grown at 30°C in Lactobacilli MRS medium with lactate (20 mM) and erythromycin (1µg/ml) for approximately 10 generations. Transformants were then grown under non-selective conditions at 37°C for about 50 generations by serial inoculations in MRS + lactate before cultures were plated on MRS containing lactate and erythromycin (1µg/ml). Isolates were screened by colony PCR for a single crossover using chromosomal specific primer oBP49 (SEQ ID NO:110) and plasmid specific primer oBP42 (SEQ ID NO:111). Single crossover integrants were grown at 37°C for approximately 40 generations by serial inoculations under non-selective conditions in MRS with lactate before cultures were plated on MRS medium with lactate. Isolates were patched to MRS with lactate plates, grown at 37°C, and then patched onto MRS plates with lactate and erythromycin (1µg/ml). Erythromycin sensitive isolates were screened by colony PCR for the presence of a wild-type or deletion second crossover using chromosomal specific primers oBP49 (SEQ ID NO:110) and oBP56 (SEQ ID NO:112). A wild-type sequence yielded a 3505 bp product and a deletion sequence yielded a 2545 bp product. The deletions were confirmed by sequencing the PCR product and absence of plasmid was tested by colony PCR with primers oBP42 (SEQ ID NO:111) and oBP57 (SEQ ID NO:113).

The *Lactobacillus plantarum* PN0512 double *ldhDldhL1* deletion strain was designated PNP0001. The Δ*ldhD* deletion included 83 bp upstream of where the *ldhD* start codon was through amino acid 279 of 332. The *ΔldhL1* deletion included the fMet through the final amino acid.

The chromosomal integration of a single copy of the *L. lactis ilvD* coding region expressed from the ldhL1 promoter was constructed by the same two-step homologous recombination procedure to yield an unmarked integration as described above using the pFP996 shuttle vector except that the second crossover event yielded the wild type sequence or the intended integration rather than the deletion. Two segments of DNA containing sequences upstream and downstream of the intended integration site were cloned into the plasmid to provide the regions of homology for two genetic crossovers.

Two DNA seqments (homologous arms) were designed to provide regions of homology for the two genetic cross-overs such that integration would place the *ilvD* coding region downstream of the *ldhL1* promoter in strain PN0512ΔldhDΔldhL1. The left and right homologous arms cloned into the plasmid were each approximately 1200 base pairs. The left homologous arm was amplified from *L. plantarum* PN0512 genomic DNA with primers oBP31 (SEQ ID NO:105), containing a BglII restriction site, and oBP32 (SEQ ID NO106), containing an XhoI restriction site using Phusion High-Fidelity PCR Master Mix. The right homologous arm was amplified from *L. plantarum* PN0512 genomic DNA with primers oBP33 (SEQ ID NO:107), containing an XhoI restriction site and oBP34 (SEQ ID NO:108), containing an XmaI restriction site using Phusion High-Fidelity PCR Master Mix. The left homologous arm was digested with BglII and XhoI and the right homologous arm was digested with XhoI and XmaI. The two homologous arms were ligated with T4 DNA Ligase into the corresponding restriction sites of pFP996, after digestion with the appropriate restriction enzymes, to generate the vector pFP996-ldhL1arms.

A DNA fragment containing the *ilvD* coding region from *Lactococcus lactis* (SEQ ID NO:82) and a ribosome binding sequence (RBS; SEQ ID NO:114) was amplified from pDM20-ilvD(L. lactis) (SEQ ID NO:115). Construction of pDM20-ilvD(L. lactis) was described in US Patent Appln No. 61/100809, which is herein incorporated by reference. This plasmid is pDM20 containing the ilvD coding region derived by PCR from *L. lactis subsp lactis* NCDO2118 (NCIMB 702118) (Godon et al., J.

Bacteriol. (1992) 174:6580-6589) and a ribosome binding sequence (SEQ ID NO:114) added in the 5" PCR primer. pDM20 is modified pDM1 (SEQ ID NO:116) which contains a minimal pLF1 replicon (~0.7 Kbp) and pemK-pemI toxin-antitoxin(TA) from *Lactobacillus plantarum* ATCC14917 plasmid pLF1, a P15A replicon from pACYC184, chloramphenicol resistance marker for selection in both *E. coli* and *L. plantarum*, and P30 synthetic promoter (Rud et al., *Microbiology* (2006) 152:1011-1019). Vector pDM1 was modified by deleting nucleotides 3281-3646 spanning the *lacZ* region which were replaced with a multi cloning site. Primers oBP120 (SEQ ID NO:117), containing an XhoI site, and oBP182 (SEQ ID NO:118), containing DrdI, PstI, HindIII, and BamHI sites, were used to amplify the P30 promoter from pDM1 with Phusion High-Fidelity PCR Master Mix. The resulting PCR product and pDM1 vector were digested with XhoI and DrdI, which drops out *lacZ* and P30. The PCR product and the large fragment of the pDM1 digestion were ligated to yield vector pDM20 in which the P30 promoter was reinserted, bounded by XhoI and DrdI restriction sites.

The DNA fragment containing the ilvD coding region and RBS (SEQ ID NO:119) was obtained by PCR using pDM20-ilvD(L. lactis) as the template with primers oBP246 (SEQ ID NO:120), containing an XhoI restriction site, and oBP237 (SEQ ID NO:121), containing an XhoI restriction site, using Phusion High-Fidelity PCR Master Mix. The resulting PCR product and pFP996-ldhL1arms were ligated with T4 DNA Ligase after digestion with XhoI. Clones were screened by PCR for the insert in the same orientation as the *ldhL1* promoter in the left homologous arm using vector specific primer oBP57 (SEQ ID NO:113) and *ilvD*-specific primer oBP237 (SEQ ID NO:121). A clone that had the correctly oriented insert was named pFP996-ldhL1arms-ilvDLl.

Integration of the *L. lactis ilvD* coding region was obtained by transforming *L. plantarum* PN0512ΔldhDΔldhL1 with pFP996-ldhL1arms-ilvDLl. 5 ml of Lactobacilli MRS medium (Accumedia, Neogen Corporation, Lansing, MI) containing 0.5% glycine (Sigma-Aldrich, St. Louis, MO) was inoculated with PN0512ΔldhDΔldhL1 and grown overnight at 30°C. 100 ml MRS medium with 0.5% glycine was inoculated with overnight culture to

an OD600 of 0.1 and grown to an OD600 of 0.7 at 30°C. Cells were harvested at 3700xg for 8 min at 4 °C, washed with 100 ml cold 1 mM $MgCl_2$ (Sigma-Aldrich, St. Louis, MO), centrifuged at 3700xg for 8 min at 4 °C, washed with 100 ml cold 30% PEG-1000 (Sigma-Aldrich, St. Louis, MO), recentrifuged at 3700xg for 20 min at 4 °C, then resuspended in 1 ml cold 30% PEG-1000. 60 µl of cells were mixed with ~100 ng of plasmid DNA in a cold 1 mm gap electroporation cuvette and electroporated in a BioRad Gene Pulser (Hercules, CA) at 1.7 kV, 25 µF, and 400 Ω. Cells were resuspended in 1 ml MRS medium containing 500 mM sucrose (Sigma-Aldrich, St. Louis, MO) and 100 mM $MgCl_2$, incubated at 30°C for 2 hrs, and then plated on MRS medium plates containing 2 µg/ml of erythromycin (Sigma-Aldrich, St. Louis, MO).

Transformants were screened by PCR using *ilvD* specific primers oBP237 (SEQ ID NO:121) and oBP246 (SEQ ID NO:120). Transformants were grown at 30°C in Lactobacilli MRS medium with erythromycin (1 µg/ml) for approximately 8 generations and then at 37°C for approximately 40 generations by serial inoculations in Lactobacilli MRS medium. The cultures were plated on Lactobacilli MRS medium with erythromycin (0.5 µg/ml).The isolates were screened by colony PCR for a single crossover with chromosomal specific primer oBP49 (SEQ ID NO:110) and plasmid specific primer oBP42 (SEQ ID NO:111).

Single crossover integrants were grown at 37°C for approximately 43 generations by serial inoculations in Lactobacilli MRS medium. The cultures were plated on MRS medium. Colonies were patched to MRS plates and grown at 37ºC. The isolates were then patched onto MRS medium with erythromycin (0.5 µg/ml). Erythromycin sensitive isolates were screened by (colony) PCR for the presence of a wild-type or integration second crossover using chromosomal specific primers oBP49 (SEQ ID NO:110) and oBP56 (SEQ ID NO:112). A wild-type sequence yielded a 2600 bp product and an integration sequence yielded a 4300 bp product. The integration was confirmed by sequencing the PCR product and an identified integration strain was designated PN0512ΔldhDΔldhL1::ilvDLI$^+$.

Example 2

Construction of a *suf* operon promoter integration vector and PN0512ΔldhDΔldhL1::ilvDLI$^+$suf::P5P4$^+$ integration strain

This Example describes integration of two promoters into the chromosome of *L. plantarum* PN0512ΔldhDΔldhL1::ilvDLI$^+$. The promoters were integrated upstream of the *suf* operon, whose gene products are responsible for Fe-S cluster assembly. The promoter integration results in a strain with increased expression of the endogenous Fe-S cluster machinery.

The *suf* operon chromosomal promoter integration was constructed by a two-step homologous recombination procedure to yield an unmarked integration using the shuttle vector pFP996 (SEQ ID NO:97) as described above.

The *suf* operon promoter integration vector was constructed in three steps. In the first step, a right homologous arm fragment containing the 5' portion of the *suf* operon (*sufC* and part of *sufD*) was cloned into pFP996. In the second step, the synthetic promoters P5 and P4 [Rud et al., *Microbiology* (2006) 152:1011] were cloned into the pFP996-right arm clone upstream of the right arm. In the final step, a left homologous arm fragment containing the native *suf* promoter and sequences upstream into the *feoBA* operon was cloned into the pFP996-P5P4-right arm clone upstream of the P5P4 promoters.

The right homologous arm DNA fragment (SEQ ID NO:123) was PCR amplified from *L. plantarum* PN0512 genomic DNA with primers AA199 (SEQ ID NO:124), containing an XmaI restriction site, and AA200 (SEQ ID NO:125), containing a KpnI restriction site, using Phusion High-Fidelity PCR Master Mix. The right homologous arm PCR fragment and pFP996 were ligated with T4 DNA Ligase after digestion with XmaI and KpnI to generate pFP996-sufCD. A DNA fragment containing promoters P5 and P4 was generated by performing PCR with two partially complementary primer sequences. Primer AA203 (SEQ ID NO:126), containing an XhoI site, the P5 promoter sequence, and part of the P4 promoter sequence, was combined with primer AA204 (SEQ ID NO:127), containing an XmaI site and the P4 promoter sequence, and PCR was

performed with Phusion High-Fidelity PCR Master Mix. The resulting PCR product was then amplified with primers AA206 (SEQ ID NO:128) and AA207 (SEQ ID NO:129) with Phusion High-Fidelity PCR Master Mix. The P5P4 PCR product and pFP996-sufCD were ligated after digestion with XhoI and XmaI to generate pFP996-P5P4-sufCD. The left homologous arm DNA fragment (SEQ ID NO:130) was amplified from *L. plantarum* PN0512 genomic DNA with primers AA201 (SEQ ID NO:131), containing an EcoRI restriction site, and AA202 (SEQ ID NO:132), containing an XhoI restriction site, using Phusion High-Fidelity PCR Master Mix. The left homologous arm and pFP996-P5P4-sufCD were ligated with T4 DNA Ligase after digestion with EcoRI and XhoI to generate pFP996-feoBA-P5P4-sufCD. The vector was confirmed by sequencing. The vector had a five base pair deletion (TTGTT), encompassing part of the -35 hexamer in the upstream P5 promoter.

Integration of the synthetic promoters (P5P4) upstream of the *suf* operon was obtained by transforming *L. plantarum* PN0512ΔldhDΔldhL1::ilvDLI$^+$ with pFP996-feoBA-P5P4-sufCD as described above. Transformants were grown at 30^{o}C in Lactobacilli MRS medium with erythromycin (2 µg/ml) for approximately 20 generations. The cultures were plated on Lactobacilli MRS medium with erythromycin (0.5 µg/ml). Isolates were screened by colony PCR for a single crossover with chromosomal specific primer AA209 (SEQ ID NO:133) and plasmid specific primer AA210 (SEQ ID NO:134). Single crossover integrants were grown at 37^{o}C for approximately 30 generations by serial inoculations in Lactobacilli MRS medium. The cultures were plated on MRS medium. Isolates were screened for erythromycin sensitivity. Isolates were screened by (colony) PCR for the presence of a wild-type or integration second crossover using P5 specific primer AA211 (SEQ ID NO:135) and chromosomal specific primer oBP126 (SEQ ID NO:136). An identified integration strain was designated PN0512ΔldhDΔldhL1::ilvDLI$^+$suf::P5P4$^+$.

Example 3

Construction of the Tn5-transposon vector (pTN6) and its use for integration of PgroE-kivD(o)-sadB(o) cassette

Tn5 is a bacterial transposon which has been well characterized in *E. coli* (Johnson & Reznikoff, *Nature* (1983) 304:280-282). A Tn5-mediated transposition system for lactic acid bacteria (LAB), however, has not been reported so far. In this Example, use of a Tn5-transposon vector as a delivery system for random gene integration into the chromosome of LAB was developed. The developed Tn5- transposon vector (pTN6) (SEQ ID NO:137) is an *E. coli-L. plantarum* shuttle vector. Plasmid pTN6 contains a transposase gene (*tnp*), transposase recognition nucleotide sequences Tn5IE (19 base pairs inside end) and Tn5OE (19 base pairs outside end), two antibiotic resistance markers; one for resistance to chloramphenicol and the other for resistance to erythromycin, P15A replication origin for *E. coli*, pE194 replication origin for *L. plantarum* which is temperature sensitive (Horinouchi and Weisblum *J. Bacteriol.* (1982) 150:804-814), and two loxP nucleotide sequences (34 base pairs). The chloramphenicol resistance gene is flanked by loxP sites for later excision by Cre recombinase. Multiple cloning sites (MSC) that contain restriction sites for BamHI, NotI, ScaI, and SpeI are located between the loxP and Tn5OE sites. The chloramphenicol resistance gene, two loxP sites, and MCS are flanked by Tn5IE and Tn5OE.

To construct the Tn5-transposon vector pTN6, first, the 1,048 bp Tn5IE-loxP-cm-loxP cassette containing Tn5IE, loxP, chloramphenicol resistant gene (cm), and loxP was synthesized by Genscript Corp (Piscataway, NJ) (SEQ ID NO:138). The Tn5IE-loxP-cm-Pspac-loxP cassette was cloned in the pUC57 vector (Genscript Corp, Piscataway, NJ), producing plasmid pUC57-Tn5IE-loxP-cm-loxP. The chloramphenicol resistance gene is expressed under the control of the spac promoter (Yansura & Henner, (1984) *Proc Natl Acad Sci USA*. 81:439–443) for selection in both *E. coli* and *L. plantarum*. Plasmid pUC57-Tn5IE-loxP-cm-loxP was digested with NsiI and SacI, and the1,044 bp Tn5IE-loxP-cm-loxP fragment was gel-purified. Plasmid pFP996 (SEQ ID NO:97) was digested with NsiI and SacI, and the 4,417bp pFP996 fragment containing the pBR322 and pE194 replication origins was gel-purified. The Tn5IE-loxP-cm-loxP fragment was ligated with the 4,417bp pFP996 fragment to generate pTnCm.

Second, the pBR322 replication origin on pTnCm was replaced by the P15A replication origin. Plasmid pTnCm was digested with AatII and SalI, and the 2,524 bp pTnCm fragment containing the pE194 replication origin and Tn5IE-loxP-cm-loxP cassette was gel-purified. The 913 bp p15A replication origin was PCR-amplified from pACYC184 [Chang and Cohen, *J. Bacteriol.* (1978)134:1141-1156] with primers T-P15A(SalITn5OE) (SEQ ID NO:139) that contains a SalI restriction site and 19 bp Tn5OE nucleotide sequence, and B-P15A(AatII) (SEQ ID NO:140) that contains an AatII restriction site by using Phusion High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, MA). The P15A fragment, after digestion with SalI and AatII restriction enzymes, was ligated with the 2,524 bp pTnCm fragment to generate pTN5.

Third, the erythromycin resistance gene (erm) was cloned into the HindIII site on pTN5. The 1,132bp erythromycin resistant gene (erm) DNA fragment was generated from vector pFP996 (SEQ ID NO:97) by PCR amplification with primers T-erm(HindIII) (SEQ ID NO:141) containing an NsiI restriction site and B-erm(HindIII) (SEQ ID NO:142) containing an NsiI restriction site by using Phusion High-Fidelity PCR Master Mix, and cloned into the HindIII restriction site on pTN5, producing pTN5-erm.

Finally, a *tnp* gene sequence encoding transposase was fused to the npr (neutral protease from *Bacillus amyloliquefaciens*) promoter [Nagarajan et al., *J. Bacteriol* (1984) 159:811-819] by SOE (splicing by overlap extension) PCR, and cloned into the NsiI site on pTN5-erm. A DNA fragment containing the Pnpr promoter (SEQ ID NO:143) was PCR-amplified from pBE83 [Nagarajan et al., *Appl Environ Microbiol* (1993) 59:3894-3898] with primer set T-Pnpr(NsiI) (SEQ ID NO:144) containing an NsiI restriction site and B-Pnpr(tnp) (SEQ ID NO:145) containing a 17 bp overlapping sequence by using Phusion High-Fidelity PCR Master Mix. A *tnp* coding region (SEQ ID NO:93) was PCR-amplified from pUTmTn5- (Sharpe et al., *Appl Environ Microbiol* (2007) 73:1721-1728) with primer set T-tnp(Pnpr) (SEQ ID NO:146) containing a 21 bp overlapping sequence and B-tnp(NsiI) (SEQ ID NO:147) containing an NsiI restriction site by using Phusion High-Fidelity PCR Master Mix. The PCR products of the two reactions were mixed and amplified using outer primers (T-

Pnpr(NsiI) and B-tnp(NsiI)), resulting in the production of a Pnpr-tnp fusion DNA fragment (SEQ ID NO:148). Plasmid pTN5-erm was digested with NsiI and treated with Calf Intestinal Phosphatase (New England Biolabs, MA) to prevent self-ligation. The digested pTN5-erm vector was ligated with the Pnpr-tnp fragment digested with NsiI. The ligation mixture was transformed into *E. coli* Top10 cells (Invitrogen Corp, Carlsbad, CA) by electroporation. Transformants were selected on LB plates containing 25 μg/mL chloramphenicol at 37 °C. Transformants then were screened by colony PCR with outer primers of the Pnpr-tnp cassette, and confirmed by DNA sequencing with primers pTnCm(711) (SEQ ID NO:149), pTnCm(1422) (SEQ ID NO: 150), and pTnCm(3025) (SEQ ID NO:151). The resulting plasmid was named pTN6.

This Tn5-transposon vector pTN6 was used as a random gene delivery system for integration of a PgroE-kivD(o)-sadB(o) cassette into the chromosome of the PN0512ΔldhDΔldhL1::ilvDLl$^+$ suf::P5P4$^+$ strain. A DNA fragment containing a PgroE promoter (Yuan and Wong, *J. Bacteriol* (1995) 177:5427-5433) (SEQ ID NO:152) was PCR-amplified from genomic DNA of *Bacillus subtilis* with primer set T-groE (SalIKpnI) (SEQ ID NO:153) containing SalI and KpnI restriction sites and B-groE (BamHI) (SEQ ID NO:154) containing a BamHI restriction site by using Phusion High-Fidelity PCR Master Mix. The resulting 154 bp PgroE promoter fragment, after digesting with SalI and BamHI restriction enzymes, was cloned into SalI and BamHI sites of plasmid pTN6, generating pTN6-PgroE. The coding region of the *kivD* gene encoding the branched-chain keto acid decarboxylase from *Lactococcus lactis* was codon optimized for expression in *L. plantarum*. The optimized coding region sequence called *kivD(o)* (SEQ ID NO:88) with a RBS was synthesized by Genscript Corp (Piscataway, NJ). The *kivD(o)* coding region together with a RBS (SEQ ID NO:155) was cloned in the pUC57 vector, producing plasmid pUC57-kivD(o). Plasmid pUC57-kivD(o) was digested with BamHI and NotI, and the 1,647bp RBS-*kivD(o)* fragment was gel-purified. The RBS-*kivD(o)* fragment was cloned into BamHI and NotI restriction sites on pTN6-PgroE, producing pTN6-PgroE-kivD(o). The correct clone was confirmed by

colony PCR with primers T-groE(SalIKpnI) and kivD(o)R (SEQ ID NO:153 and 156), producing a 1,822 bp fragment of the expected size. Then, the *sadB* gene coding region for branched-chain alcohol dehydrogenase from *Achromobacter xylosoxidans*, that was described in US Patent Application #12/430356, was cloned downstream of the *kivD*(o) coding region of pTN6-PgroE-kivD(o). The *A. xylosoxidans sadB* coding region was codon optimized for expression in *L. plantarum*. The new coding region called *sadB(o)* (SEQ ID NO:157) with a RBS was synthesized by Genscript Corp (Piscataway, NJ), and cloned in the pUC57 vector, producing plasmid pUC57-sadB(o). A 1,089bp DNA fragment (SEQ ID NO:158) containing the RBS and *sadB(o)* coding region was PCR-amplified from pUC57-sadB(o) with primer set T-sadB(o)(NotI) (SEQ ID NO:159) containing a NotI restriction site and B-sadB(o)(NotI) (SEQ ID NO:160) containing a NotI restriction site by using Phusion High-Fidelity PCR Master Mix. The RBS-*sadB(o)* gene fragment, after digesting with NotI, was cloned into NotI restriction site of pTN6-PgroE-kivD(o), producing pTN6-PgroE-kivD(o)-sadB(o). The correct clone was confirmed by DNA sequencing with kivD(o)1529 (SEQ ID NO:161) and B-spac(cm) (SEQ ID NO:162) primers. In this construction sadB(o) and kivD(o) coding regions are expressed in an operon from PgroE promoter.

The resulting plasmid pTN6-PgroE-kivD(o)-sadB(o) was transformed into PN0512ΔldhDΔldhL1::ilvDLI^{+}suf::P5P4^{+} by electroporation as described in General Methods. Transformants were selected on Lactobacilli MRS medium supplemented with 7.5 µg/ml chloramphenicol. The chloramphenicol resistant colonies were grown in Lactobacilli MRS medium with 7.5 µg/ml chloramphenicol at the permissive temperature of 30°C for approximately 10 generations. The culture was inoculated at 1/100 dilution in fresh MRS medium and grown at 37°C for approximately 20 generations by serial inoculation in Lactobacilli MRS medium. The cultures were plated on Lactobacilli MRS with 7.5 µg/ml chloramphenicol. The isolates were screened by re-streaking colonies on Lactobacilli MRS plates containing 1.5 µg/ml erythromycin for erythromycin sensitive colonies that were presumed to

contain a chromosomally integrated PgroE-kivD(o)-sadB(o) cassette along with the transposon. The transposon-mediated integrants were confirmed by colony PCR with the *kivD(o)* sequence specific primer KivD(o)1529 and *sadB(o)* sequence specific primer B-sadB(o)(NotI), to produce the expected sized PCR product (1,220 bp).

To excise the chloramphenicol resistance marker that is flanked by loxP sites from the chromosome, a helper plasmid pFP352 (SEQ ID NO:163) expressing a Cre recombinase was transformed into the transposon-mediated integrant, according to the protocol as described in General Methods, and grown on Lactobacillus MRS plate containing 1.5 µg/ml erythromycin at 30°C. The cre recombinase excises the chloramphenicol marker from the chromosome by a recombiniation event between the loxP sites. The erythromycin resistant transformants were inoculated in MRS medium and grown at 37°C for approximately 10 generations. The cultures were plated on Lactobacilli MRS without antibiotic and grown at 30°C. The isolates were screened for both erythromycin and chloramphenicol sensitive colonies by testing growth of colonies on Lactobacilli MRS plates containing 1.5 µg/ml erythromycin and Lactobacilli MRS plates containing chloramphenicol (7.5 µg/ml), separately, to verify loss of pFP352 and the chloramphenicol marker removal. Finally, the integrant was confirmed by genomic DNA sequencing with primer B-groE(BamHI). Genomic DNA was prepared using MasterPure DNA Purification® kit (Enpicentre, Inc., Madison, WI). The DNA sequencing result indicated that the PgroE-kivD(o)-sadB(o) cassette was Inserted within the coding region of the *glgB* gene encoding glycogen branching enzyme that catalyzes the transfer of a segment of a 1,4-alpha-D-glucan chain to a primary hydroxy group in a similar glucan chain. The resulting integrant was named PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o).

Example 4

Construction of an *aldB* deletion vector and initial deletion attempt

An attempt to delete the *aldB* gene, encoding acetolactate

decarboxylase, in strain PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o) is described.

A two-step homologous recombination procedure was used to try to create an unmarked deletion. The homologous recombination procedure utilized the shuttle vector, pFP996 (SEQ ID NO:97), described above. Two segments of DNA containing sequences upstream and downstream of the intended deletion were cloned into the plasmid to provide the regions of homology for two genetic crossovers. An initial single crossover integrates the plasmid into the chromosome. A second crossover event can then yield either the wild type sequence or the intended gene deletion.

The homologous DNA arms were designed such that the deletion would encompass the region encoding the first 23 amino acids of the AldB protein (nucleotides 1-68 of the *aldB* coding sequence). The left and right homologous arms cloned into the plasmid were 1186 and 700 base pairs, respectively. The homologous arms were separated by the sequence GGTACCT, which replaced the 68 nucleotide *aldB* sequence deletion. The left homologous arm was amplified from *L. plantarum* PN0512 genomic DNA with primers oBP23 (SEQ ID NO:122), containing an XhoI restriction site, and oBP24 (SEQ ID NO:164), containing a KpnI restriction site using Phusion High-Fidelity PCR Master Mix. The right homologous arm was amplified from *L. plantarum* PN0512 genomic DNA with primers oBP335 (SEQ ID NO:165), containing a KpnI restriction site and oBP336 (SEQ ID NO:166), containing a BsrGI restriction site using Phusion High-Fidelity PCR Master Mix. The left homologous arm DNA fragment was digested with XhoI and KpnI and the right homologous arm DNA fragment was digested with KpnI and BsrGI. The two homologous arms were ligated with T4 DNA Ligase into the corresponding restriction sites of pFP996, after digestion with the appropriate restriction enzymes, to generate the vector pFP996aldBdel23arms.

The single cross-over was obtained by transforming *Lactobacillus plantarum* PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o) with pFP996aldBdel23arms. 100 ml of Lactobacilli MRS medium (Accumedia, Neogen Corporation, Lansing, MI) containing 0.5% glycine (Sigma-Aldrich, St. Louis, MO) was inoculated with

PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o) and grown at 30°C to an OD600 of 0.7. Cells were harvested at 3700xg for 8 min at 4 °C, washed with 100 ml cold 1 mM $MgCl_2$ (Sigma-Aldrich, St. Louis, MO), centrifuged at 3700xg for 8 min at 4 °C, washed with 100 ml cold 30% PEG-1000 (Sigma-Aldrich, St. Louis, MO), recentrifuged at 3700xg for 20 min at 4 °C, then resuspended in 1 ml cold 30% PEG-1000. 60 µl of cells were mixed with ~100 ng of plasmid DNA in a cold 1 mm gap electroporation cuvette and electroporated in a BioRad Gene Pulser (Hercules, CA) at 1.7 kV, 25 µF, and 400 Ω. Cells were resuspended in 1 ml MRS medium containing 500 mM sucrose (Sigma-Aldrich, St. Louis, MO) and 100 mM $MgCl_2$, incubated at 30°C for 2 hrs, and then plated on MRS medium plates containing 1 µg/ml of erythromycin (Sigma-Aldrich, St. Louis, MO).

Transformants were screened by PCR using plasmid specific primers oBP42 (SEQ ID NO:111) and oBP57 (SEQ ID NO:113). Transformants were grown at 30°C in Lactobacilli MRS medium with erythromycin (1 µg/ml) for approximately 10 generations and then at 37°C for approximately 35 generations by serial inoculations in Lactobacilli MRS medium. The cultures were plated on Lactobacilli MRS medium with erythromycin (1 µg/ml). The isolates obtained were screened by colony PCR for a single crossover with chromosomal specific primer oBP47 (SEQ ID NO:167) and plasmid specific primer oBP42 (SEQ ID NO:111), and chromosomal specific primer oBP54 (SEQ ID NO:168) and plasmid specific primer oBP337 (SEQ ID NO:169).

Single crossover integrants were grown at 37°C for approximately 41 generations by serial inoculations in Lactobacilli MRS medium without glucose. The cultures were plated on MRS medium without glucose and grown at 37ºC. Colonies were patched to MRS plates without glucose and grown at 37ºC. 96 isolates were screened by (colony) PCR for the presence of a deletion second crossover using chromosomal specific primer oBP54 (SEQ ID NO:168) and deletion specific primer oBP337 (SEQ ID NO:169). None of the isolates tested contained the deletion.

Example 5

Construction of the pTN5-PrrnC1-aldB(L. lactis) vector

The purpose of this Example is to describe cloning of the *aldB* coding region (SEQ ID NO:37) for acetolactate decarboxylase from *Lactococcus lactis subsp lactis* NCDO2118 [Godon et al., J. Bacteriol. (1992) 174:6580-6589] into the *E. coli-L. plantarum* shuttle vector pTN5. The construction of the pTN5 vector was described in Example 3.

The DNA fragment containing a PrrnC1 promoter (SEQ ID NO:170) was PCR-amplified from genomic DNA of *Lactococcus lactis subsp lactis* NCDO2118 with primer set T-rrnC1(SalIKpnI) (SEQ ID NO:171) containing SalI and KpnI restriction sites and B-rrnC1(BamHI) (SEQ ID NO:172) containing a BamHI restriction site by using Phusion High-Fidelity PCR Master Mix. The resulting 149 bp PrrnC1 promoter fragment, after digesting with SalI and BamHI restriction enzymes, was cloned into SalI and BamHI sites of plasmid pTN5, generating pTN5-PrrnC1. A DNA fragment containing a RBS and *aldB* coding region was PCR-amplified from genomic DNA of *Lactococcus lactis subsp lactis* NCDO2118 with primer set T-aldBLI(BamHI) (SEQ ID NO:173) containing a BamHI restriction site and B-aldBLI(NotISpeI) (SEQ ID NO:174) containing NotI and SpeI restriction sites. The resulting 735 bp *aldB*(*L. lactis*) coding region and RBS fragment was digested with BamHI and NotI, and then cloned into BamHI and NotI sites on pTN5-PrrnC1, generating pTN5-PrrnC1-aldB(L. lactis). The correct clone was confirmed by restriction enzyme mapping with BamHI and NotI, showing expected size (3,569bp and 735bp) DNA fragments.

EXAMPLE 6

aldB deletion in the presence of plasmid-expressed acetolactate decarboxylase

In this Example, the second crossover to cause deletion of *aldB* was attempted in cells expressing an *aldB* gene on a plasmid.

A single cross-over integrant from Example 5 was transformed with the plasmid pTN5-PrrnC1-aldB(L. lactis) by electroporation. The electro-competent cells were prepared as described above in Example 4.

Transformants were selected following incubation at 30 °C for 5 days on Lactobacillus MRS agar plates containing chloramphenicol (7.5 µg/ml) and erythromycin (1 µg/ml). The chloramphenicol and erythromycin resistant transformants were grown at 30°C for approximately 20 generations by serial inoculations in Lactobacilli MRS medium with chloramphenicol (7.5 µg/ml), and then the cultures were plated on Lactobacillus MRS agar plates containing chloramphenicol (7.5 µg/ml). The resulting colonies were patched onto Lactobacillus MRS agar plates containing erythromycin (1 µg/ml) to test erythromycin sensitivity. 42 out of 130 colonies showed erythromycin sensitivity. Then, the 42 erythromycin sensitive colonies were screened for deletion of the region encoding the first 23 amino acids of the AldB protein (nucleotides 1-68 of the *aldB* coding sequence) by colony PCR analysis with the chromosomal specific primers OBP47 and OBP54 (expected size: ~3.3 kbp), and chromosomal specific primers OBP54 and OBP337 (expected size: ~1.9 kbp). The colony PCR analysis showed that 22 out of 42 erythromycin sensitive colonies had *Δ23aa aldB.*

To cure the plasmid pTN5-PrrnC1-aldB(L. lactis) the *Δ23aa aldB* deletion mutant strain was grown at 37°C for approximately 20 generations by serial inoculations in Lactobacilli MRS medium. The cultures were plated on Lactobacillus MRS agar plates. The plasmid removal of the *Δ23aa aldB* deletion mutant strain was confirmed by no growth of the strain on an MRS agar plate containing chloramphenicol (7.5 µg/ml). After removing the plasmid pTN5-PrrnC1-aldB(L. lactis), the deletion of the nucleotides 1-68 of the *aldB* coding sequence, corresponding to the first 23 amino acids of the AldB protein, was confirmed by DNA sequencing with AA213 primer (SEQ ID NO:175) showing that the endogenous *aldB* gene was successfully deleted in the presence of plasmid expression of AldB. The resulting *Δ23aa aldB* mutation strain was named PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB*.

Example 7

Construction of the pDM5-PldhL1-ilvC(L. lactis) vector

The purpose of this example is to describe cloning of the *ilvC* coding region (SEQ ID NO:74) for keto-acid reductoisomerase from *Lactococcus lactis subsp lactis* NCDO2118 (NCIMB 702118) [Godon et al., J. Bacteriol. (1992) 174:6580-6589] into the pDM5 vector.

Plasmid pDM5 (SEQ ID NO:176) was constructed by replacing the P30 promoter on pDM1 with the *B. subtilis* groE promoter (PgroE) fused to a lacO operator sequence and a *lacI* repressor gene. Plasmid pDM1 is described in Example 1. Plasmid pHTO1 (Mo Bi Tec, Goettingen, Germany) was digested with SacI, treated with Klenow fragment to make blunt ends, digested with BamHI, and then the 1,548 bp lacI-PgroE/lacO fragment (SEQ ID NO:177) was gel-purified. The lacI-PgroE/lacO fragment was cloned into KpnI (blunt end by Klenow fragment) and BamHI sites of pDM1 in place of the P30 promoter, generating pDM5.

A DNA fragment, PldhL1-ilvC(L. lactis), containing a ldhL1 (L-lactate dehydrogenase from *Lactobacillus plantarum* PN0512) promoter (PldhL1) and ilvC coding region from *Lactococcus lactis subsp lactis* NCDO2118 was generated by SOE (splicing by overlap extension) PCR. The DNA fragment containing a PldhL1 promoter was PCR-amplified from the genomic DNA of *Lactobacillus plantarum* PN0512 with primer set T-ldhL1(NotI) (SEQ ID NO:178) containing a NotI restriction site and B-ldhLI(CLI) (SEQ ID NO:179) containing a 19 bp overlapping sequence by using Phusion High-Fidelity PCR Master Mix. An *ilvC* coding region was PCR-amplified from the genomic DNA of *Lactococcus lactis subsp lactis* NCDO2118 with primer set T-CLI(ldh) (SEQ ID NO:180) containing a 17 bp overlapping sequence and B-CLI(PvuI) (SEQ ID NO:181) containing a PvuI restriction site by using Phusion High-Fidelity PCR Master Mix. The PCR products of the two fragments were mixed and amplified using outer primers T-ldhL1(NotI) and B-CLI(PvuI), resulting in the production of a PldhL1-ilvC(L. lactis) fusion DNA fragment. Plasmid pDM5 was digested with NotI and PvuI restriction enzymes, and ligated with the PldhL1-ilvC(L. lactis) cassette after digesting with NotI and PvuI restriction enzymes. The ligation mixture was transformed into *E. coli* Top10 cells (Invitrogen Corp, Carlsbad, CA) by electroporation. Transformants were selected on LB

plates containing 25 μg/mL chloramphenicol at 37 °C. Transformants then were screened by colony PCR with outer primers of the PldhL1-ilvC(L. lactis) cassette, and confirmed by DNA sequencing with T-ldhL1(NotI) (SEQ ID NO:178) and pDM(R)new (SEQ ID NO:182). The resulting plasmid was named pDM5-PldhL1-ilvC(L. lactis) (SEQ ID NO:183).

Example 8

Production of Isobutanol using PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB* containing vector pDM5-PldhL1-ilvC(L. lactis)

The purpose of this example is to demonstrate the increased production of isobutanol in the recombinant *Lactobacillus plantarum aldB-* strain background, compared to an *aldB*+ strain background.

To construct the recombinant *Lactobacillus plantarum* expressing the genes of the isobutanol biosynthetic pathway, competent cells of the two integrants PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) and PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB-* were prepared as described above, and transformed with plasmid pDM5- PldhL1-ilvC(L. lactis), yielding PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o)/pDM5-PldhL1-ilvC(L. lactis), named DWS2269, and PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB-*/pDM5-PldhL1-ilvC(L. lactis), named DWS2279. The first enzyme acetolactate synthase for the isobutanol pathway was provided by native expression from the endogenous gene.

The two strains DWS2269 and DWS2279 were inoculated in Lactobacilli MRS (100 mM 3-Morpholinopropanesulfonic acid (MOPS) pH7.0) medium containing 7.5 μg/ml chloramphenicol in 10 ml culture tubes and grown aerobically at 30°C overnight. Overnight cultures were inoculated with an initial OD600=0.4 into 40 ml MRS medium (100 mM MOPS pH7.0) containing 7.5 μg/ml chloramphenicol, 40 μM ferric citrate, 0.5 mM cysteine in 120 ml serum bottles, and grown with 100 rpm shaking anaerobically at 37°C for 96 hours. Samples of the cultures were centrifuged at 3700xg for

10 minutes at 4 °C and the supernatants filtered through a 0.2 µm filter (Pall Life Sciences, Ann Arbor, MI). The filtered supernatants were analyzed by HPLC (1200 Series, Agilent Technologies, Santa Clara, CA) with a SHODEX Sugar column, detected by UV210 and refractive index, mobile phase 10 mM H_2SO_4. Results in Table 5 show the production of isobutanol, acetoin, and ethanol for strains DWS2269 and DWS2279 grown in MRS medium (100 mM MOPS pH7.0) at 37°C anaerobically. The amount of isobutanol produced by DWS2279 that contains the *aldB*- mutation was 8 mM, which is approximately 6-fold higher than the isobutanol level (1.3 mM) produced by DWS2269 that contains wild type *aldB*+.

Table 5. Production of isobutanol, acetoin, and ethanol by DWS2269 (*aldB*+) and DWS2279 (*aldB*-) grown in grown in MRS medium (100 mM MOPS pH7.0) at 37°C anaerobically.

Strain	Ethanol (mM)	Acetoin (mM)	Isobutanol (mM)
DWS2269= PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o)/pDM5-PldhL1-ilvC(L. lactis)	25.9	25.4	1.3
DWS2279= PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB*-/pDM5-PldhL1-ilvC(L. lactis)	79.0	4.0	8.0

Example 9

Construction of the *Lactobacillus plantarum* PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o) strain

The purpose of this example is to describe the construction of a *Lactobacillus plantarum* strain in the PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB strain background that is deleted for the genes *pflB2*, encoding formate C-acetyltransferase (pyruvate formate lyase), and *pflA2*, encoding the formate C-acetyltransferase activating enzyme, and thus does not contain formate C-acetyltransferase activity. A gene (*alsS*), codon optimized for expression in *Lactobacillus plantarum*, encoding the *Bacillus subtilis* acetolactate synthase enzyme was integrated in place of the *pflB2A2*

genes of *Lactobacillus plantarum* PN0512.

The *pflB2A2* gene knockout and *alsS* gene integration were engineered using the two-step homologous recombination procedure described above. The knockout deleted the C-terminal 351 amino acids (nucleotides 1204 through 2256 of the coding sequence) of PflB2 and the entire coding sequence of *pflA2*. The deleted sequence was replaced with a stop codon, in frame with the truncated *pflB2*, followed by a ribosome binding sequence and the coding region of the *Bacillus subtilis alsS* gene codon optimized for expression in *Lactobacillus plantarum*.

The knockout/integration vector was constructed in plasmid pFP996 (SEQ ID NO:97) as follows. The homologous arms to delete the *pflB2A2* genes were amplified from *L. plantarum* PN0512 genomic DNA. The *pflB2A2* left homologous arm was amplified using primers oBP309 (SEQ ID NO:184) containing an XhoI restriction site and oBP310 (SEQ ID NO:185) containing a stop codon (complement of TAA) and XmaI restriction site. The *pflB2A2* right homologous arm was amplified using primers oBP271 (SEQ ID NO:186) containing a KpnI restriction site and oBP272 (SEQ ID NO:187) containing a BsrGI restriction site. The *pflB2A2* left homologous arm was cloned into the XhoI/XmaI sites and the *pflB2A2* right homologous arm was cloned into the KpnI/BsrGI sites of pFP996 to create pFP996-pflB2A2arms. The *Bacillus subtilis alsS* coding region codon optimized for expression in *Lactobacillus plantarum* (SEQ ID NO:57; synthesized by Genscript Corp, Piscataway, NJ) was amplified using primers oBP282 (SEQ ID NO:188) containing an XmaI restriction site and oBP283 (SEQ ID NO:189) containing a KpnI restriction site. The codon optimized *alsS* coding region was cloned into the XmaI/KpnI sites of pFP996-pflB2A2arms to create pFP996-pflB2A2arms-als(o).

PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4^{+} glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB (prepared in Example 6) was transformed with pFP996-pflB2A2arms-als(o) as above, except competent cells were prepared in the absence of glycine, and transformants were selected on MRS plates containing 1 µg/ml erythromycin. Transformants were streaked on MRS plates containing erythromycin (1µg/ml) and then re-streaked on MRS plates. Isolates were screened by colony PCR for a

single crossover using chromosomal specific primer oBP278 (SEQ ID NO:190) and als(o) specific primer oBP283 (SEQ ID NO:189). A single crossover integrant was grown at 37°C for approximately 25 generations by serial inoculations in MRS medium without glucose before cultures were plated on MRS medium without glucose. Erythromycin sensitive isolates were screened by colony PCR for the presence of a wild-type or deletion/integration second crossover using als(o) specific primer oBP282 (SEQ ID NO:188) and chromosomal specific primer oBP280 (SEQ ID NO:89). The resulting deletion/integration strain PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o) was confirmed by sequencing the PCR product amplified with chromosomal specific primers oBP279 (SEQ ID NO:90) and oBP280 (SEQ ID NO:89).

Example 10

Production of Isobutanol using PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o) containing a vector pDM5-PldhL1-ilvC(L. lactis)

The purpose of this example is to demonstrate the increased production of isobutanol in the *Lactobacillus plantarum* PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o) strain background, compared to the parental strain PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB strain background.

To construct a recombinant *Lactobacillus plantarum* expressing the genes of the isobutanol biosynthetic pathway, competent cells of the integrant PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o) were prepared as described in Example 1, and transformed with plasmid pDM5-PldhL1-ilvC(L. lactis) (construction described in Example 7), yielding PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) Δ23aa aldB ΔpflB2A2::alsS(o)/pDM5-PldhL1-ilvC(L. lactis), which was named DWS2307.

Production of isobutanol with strain DWS2307 was tested using the same medium, growth conditions, and sample preparation as described in Example 8. Strain DWS2279 (Example 8) was grown as the control. The filtered supernatants were analyzed by HPLC (1200 Series, Agilent Technologies, Santa Clara, CA) with a SHODEX Sugar column, detected by UV210 and refractive index, mobile phase 10 mM H_2SO_4. Results in Table 6 show the production of isobutanol, formate, acetoin, and ethanol for DWS2307, compared to DWS2279. The amount of isobutanol produced by DWS2307 that contains the ΔpflB2A2- mutation was 19.1 mM, which is approximately 2.4-fold higher than the isobutanol level (8 mM) produced by DWS2279 that contains wild type pflB2A2+. DWS2307 that is deleted for the genes *pflB2* and *pflA2* and thus does not contain formate C-acetyltransferase activity, showed no production of formate.

Table 6. Production of isobutanol, formate, acetoin, and ethanol by DWS2279 (pflB2A2+) and DWS2307 (ΔpflB2A2-) grown in grown in MRS medium (100 mM MOPS pH7.0) at 37°C anaerobically.

Strain	Ethanol (mM)	Formate (mM)	Acetoin (mM)	Isobutanol (mM)
DWS2279= PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB*-/pDM5-PldhL1-ilvC(L. lactis)	79.0	30.0	4.0	8.0
DWS2307= PN0512ΔldhDΔldhL1::ilvD(Ll) suf::P5P4$^+$ glgB::Tn5-PgroE-kivD(o)-sadB(o) *Δ23aa aldB*- ΔpflAB::alsS(B. subtilis)/pDM5-PldhL1-ilvC(L. lactis)	39.0	0.0	4.0	19.1

CLAIMS

What is claimed is:

1. A recombinant lactic acid bacteria cell comprising at least one engineered genetic modification that reduces or eliminates enzyme activity of endogenously expressed acetolactate decarboxylase and at least one engineered genetic modification that eliminates enzyme activity of endogenously expressed lactate dehydrogenase.

2. The bacteria cell of claim 1 comprising at least one engineered genetic modification that eliminates enzyme activity of endogenously expressed acetolactate decarboxylase and at least one engineered genetic modification that eliminates enzyme activity of endogenously expressed lactate dehydrogenase.

3. The bacteria cell of claim 1 wherein each of the engineered genetic modifications is a deletion of at least a portion of a gene encoding acetolactate decarboxylase or lactate dehydrogenase

4. The bacteria cell of claim 3 wherein said gene encoding acetolactate decarboxylase is selected from the group consisting of aldB, aldC, and ald.

5. The bacteria cell of claim 4 wherein the gene encoding acetolactate decarboxylase encodes a protein having an amino acid sequence that has at least about 95% identity to a sequence selected from the group consisting of SEQ ID NOs: 24, 26, 28, 30, 32, 34, 36, and 38.

6. The bacteria cell of claim 3 wherein said gene encodes lactate dehydrogenase and is selected from the group consisting of ldhL, ldhD, ldhL1, and ldhL2.

7. The bacteria cell of claim 6 wherein the gene encoding lactate dehydrogenase encodes a protein having an amino acid sequence that

has at least about 95% identity to a sequence selected from the group consisting of SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22.

8. The bacteria cell of claim 1 wherein the cell is a member of a genus selected from the group consisting of *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Oenococcus*, *Pediococcus*, and *Streptococcus*.

9. The bacteria cell of claim 1 further comprising at least one genetic modification that reduces pyruvate formate lyase activity.

10. The bacteria cell of claim 9 wherein the genetic modification is in a gene encoding pyruvate formate lyase, a gene encoding pyruvate formate lyase activating enzyme, or in both genes.

11. The bacteria cell of claim 10 wherein the said gene encoding pyruvate formate lyase is selected from the group consisting of pfl, pflB1 and pflB2 and said gene encoding formate C-acetyltransferase activating enzyme is selected from the group consisting of pflA, pflA1 and pflA2.

12. The bacteria cell of claim 1, 2, or 8 wherein the cell produces isobutanol.

13. The bacteria cell of claim 12 comprising an isobutanol biosynthetic pathway.

14. The bacteria cell of claim 13 wherein the isobutanol biosynthetic pathway comprises substrate to product conversions consisting of:

a) pyruvate to acetolactate;

b) acetolactate to 2,3-dihydroxyisovalerate;

c) 2,3-dihydroxyisovalerate to α-ketoisovalerate;

d) α-ketoisovalerate to isobutyraldehyde; and

e) isobutyraldehyde to isobutanol.

15. A method for producing a recombinant lactic acid bacteria cell comprising:

a) providing a lactic acid bacteria cell;

b) modifying by genetic engineering at least one endogenous gene encoding lactate dehydrogenase in the cell of (a) to eliminate enzyme activity of endogenously expressed lactate dehydrogenase;

c) expressing acetolactate decarboxylase activity from a plasmid in the cell of (b) to create a cell with non-chromosomally expressed acetolactate decarboxylase;

(d) modifying by genetic engineering an endogenous gene encoding acetolactate decarboxylase in the cell of (c) to eliminate enzyme activity of endogenously expressed acetolactate decarboxylase; and

(e) curing the plasmid expressing acetolactate decarboxylase activity from the cell of (d);

whereby a recombinant lactic acid bacteria cell lacking enzyme activity of endogenously expressed lactate dehydrogenase and acetolactate decarboxylase is produced.

16. A method for producing a recombinant lactic acid bacteria cell comprising:

a) providing a lactic acid bacteria cell;

b) modifying by genetic engineering an endogenous gene encoding acetolactate decarboxylase in the cell of (c) to eliminate enzyme activity of endogenously expressed acetolactate decarboxylase;

c) expressing lactate dehydrogenase activity from a plasmid in the cell of (b) to create a cell with non-chromosomally expressed lactate dehydrogenase;

(d) modifying by genetic engineering at least one endogenous gene encoding lactate dehydrogenase in the cell of (a) to eliminate enzyme activity of endogenously expressed lactate dehydrogenase; and

(e) curing the plasmid expressing lactate dehydrogenase activity from the cell of (d);

whereby a recombinant lactic acid bacteria cell lacking enzyme activity of endogenously expressed lactate dehydrogenase and acetolactate decarboxylase is produced.

17. The method of claim 15 wherein step (b) comprises a modification to a first gene encoding lactate dehydrogenase prior to (c), then a second gene encoding lactate dehydrogenase is modified by genetic engineering following step (c).

18. The method of claim 15, 16 or 17 further comprising modifying at least one endogenous gene to reduce pyruvate formate lyase activity.

19. A method for producing isobutanol comprising:

(a) providing a lactic acid bacteria cell comprising:

i) at least one genetic modification that eliminates enzyme activity of endogenously expressed acetolactate decarboxylase and at least one genetic modification that eliminates enzyme activity of endogenously expressed lactate dehydrogenase; and

ii) an isobutanol biosynthetic pathway; and

(b) culturing the cell of (a) under conditions wherein isobutanol is produced.

20. The method of claim 19 wherein the lactic acid bacteria cell of (a) further comprises at least one genetic modification that reduces pyruvate formate lyase activity.

21. An integration vector for lactic acid bacteria comprising:

a) a Tn-5 transposase coding region operably linked to a promoter that is active in lactic acid bacteria cells;

b) Tn5IE and TN5OE elements bounding a selection marker that is active in *E. coli* and lactic acid bacteria cells, and a DNA segment targeted for integration;

c) a second selection marker that is active in lactic acid bacteria cells;

d) an origin of replication for *E. coli* cells;

e) an origin of replication for lactic acid bacteria cells that is conditionally active;

wherein the Tn5IE and TN5OE elements direct random integration of the DNA segment of b) into the lactic acid bacteria cell genome.

22. A method for randomly integrating a DNA segment into a lactic acid bacteria cell genome comprising:

a) providing a vector comprising:

i) a Tn-5 transposase coding region operably linked to a promoter that is active in lactic acid bacteria cells;

ii) Tn5IE and TN5OE elements bounding a selection marker that is active in *E. coli* and lactic acid bacteria cells;

iii) a second selection marker active in lactic acid bacteria cells;

iv) an origin of replication for *E. coli* cells;

v) an origin of replication for lactic acid bacteria cells that is conditionally active;

b) placing a DNA segment for integration between the elements of step a (ii) creating an integration construction;

c) transforming the integration construction into a lactic acid bacteria cell whereby transformed cells are produced;

d) growing and selecting the transformed cells of step (c) in permissive conditions using the selection marker of step a (ii) to produce selected transformants; and

e) growing the selected transformants of step (d) in nonpermissive conditions;

wherein the vector is cured from the lactic acid bacteria cells and the DNA segment for integration is randomly integrated into the genome of said lactic acid bacteria cell.

Figure 1

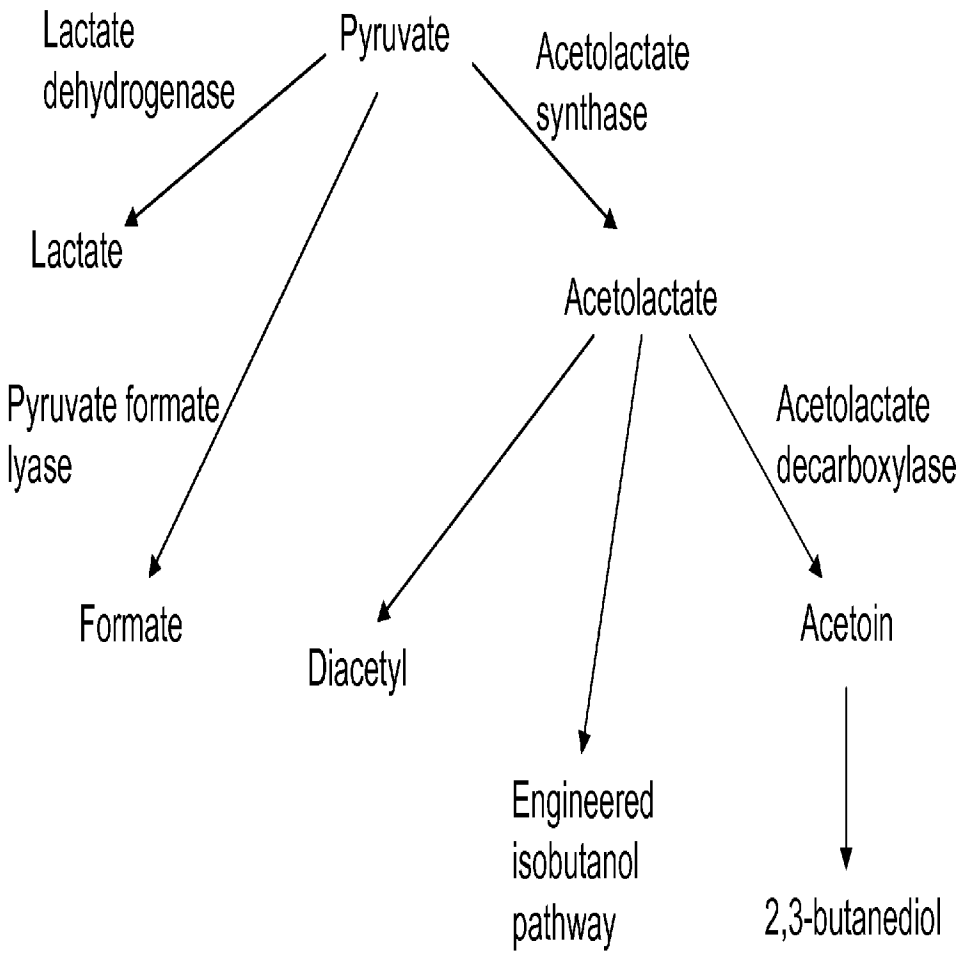
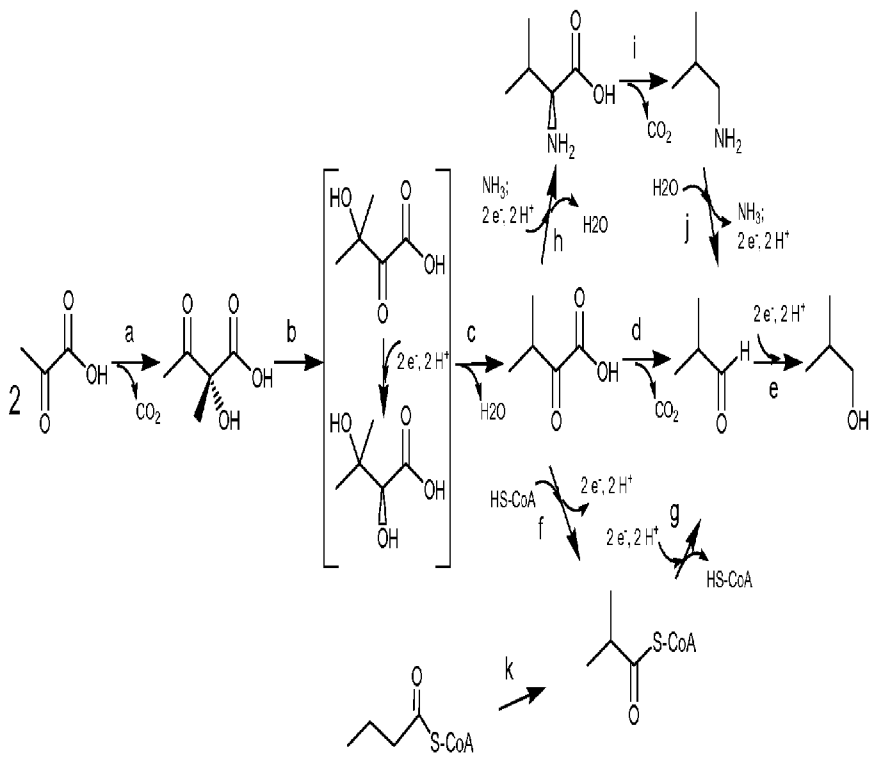

Lactate dehydrogenase
Pyruvate
Acetolactate synthase
Lactate
Acetolactate
Pyruvate formate lyase
Acetolactate decarboxylase
Formate
Diacetyl
Acetoin
Engineered isobutanol pathway
2,3-butanediol

Figure 2


2
OH
O
a
CO2
OH
OH
b
HO
OH
O
2 e-; 2 H+
HO
OH
OH
c
H2O
OH
O
d
CO2
H
O
2 e-; 2 H+
e
OH
NH3;
2 e-; 2 H+
h
H2O
H2O
j
NH3;
2 e-; 2 H+
i
OH
NH2
CO2
NH2
HS-CoA
f
2 e-; 2 H+
2 e-; 2 H+
g
HS-CoA
k
S-CoA
O
S-CoA

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2010/050705

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N9/04 C12N9/88 C12N1/21 C12N9/02 C12N9/10
C12N15/52 C12P7/16 C12N15/74

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C12N C12P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	DE VOS WILLEM M ET AL: "Making more of milk sugar by engineering lactic acid bacteria" INTERNATIONAL DAIRY JOURNAL, vol. 8, no. 3, March 1998 (1998-03), pages 227-233, XP002611985 ISSN: 0958-6946 cited in the application page 231; figure 1 ----- -/--	1-20

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 December 2010	17/02/2011

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	van Voorst, Frank

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MONNET C ET AL: "Diacetyl and alpha-acetolactate overproduction by Lactococcus lactis subsp. lactis biovar diacetylactis mutants that are deficient in alpha-acetolactate decarboxylase and have a low lactate dehydrogenase activity." APPLIED AND ENVIRONMENTAL MICROBIOLOGY DEC 2000 LNKD- PUBMED:11097941, vol. 66, no. 12, December 2000 (2000-12), pages 5518-5520, XP002611986 ISSN: 0099-2240 cited in the application abstract	1-20
X	FR 2 777 905 A1 (AGRONOMIQUE INST NAT RECH [FR]) 29 October 1999 (1999-10-29) abstract; example 3	1-20
X	OLIVEIRA ANA PAULA ET AL: "Modeling Lactococcus lactis using a genome-scale flux model" BMC MICROBIOLOGY, BIOMED CENTRAL, LONDON, GB, vol. 5, no. 1, 27 June 2005 (2005-06-27), page 39, XP021002638 ISSN: 1471-2180 DOI: DOI:10.1186/1471-2180-5-39 abstract	1-20
A	BISWAS I ET AL: "High-efficiency gene inactivation and replacement system for gram-positive bacteria." JOURNAL OF BACTERIOLOGY JUN 1993 LNKD- PUBMED:8501066, vol. 175, no. 11, June 1993 (1993-06), pages 3628-3635, XP002612131 ISSN: 0021-9193 abstract	1-20
Y	DE VOS W M: "Safe and sustainable systems for food-grade fermentations by genetically modified lactic acid bacteria" INTERNATIONAL DAIRY JOURNAL, ELSEVIER APPLIED SCIENCE, BARKING, GB, vol. 9, no. 1, 1 January 1999 (1999-01-01), pages 3-10, XP009087552 ISSN: 0958-6946 DOI: DOI:10.1016/S0958-6946(99)00038-2 page 7	1-20

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2010/050705

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-20

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-20

   Lactic acid bacteria comprising engineered genetic modifications of the acetolactate decarboxylase and the lactate dehydrogenase, methods for making said bacteria and methods for producing isobutanol using said bacteria.

   ---

2. claims: 21, 22

   An integration vector for lactic acid bacteria and a method for integrating said vector into a lactic acid bacteria cell genome.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2010/050705

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
FR 2777905 A1	29-10-1999	WO 9955836 A1	04-11-1999