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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
6 March 2014 (06.03.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/033018 A1

(51) **International Patent Classification**:
C12N 1/19 (2006.01) *C12N 9/04* (2006.01)
C12N 9/02 (2006.01) *C12P 7/10* (2006.01)

(21) **International Application Number**:
PCT/EP2013/067341

(22) **International Filing Date**:
20 August 2013 (20.08.2013)

(25) **Filing Language**: English

(26) **Publication Language**: English

(30) **Priority Data**:
12181938.7 28 August 2012 (28.08.2012) EP

(71) **Applicant**: **DSM IP ASSETS B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).

(72) **Inventors**: **KLAASSEN, Paul**; P.O.Box 130, NL-6100 AC Echt (NL). **KOLEN, Catharina Petronella Antonia Maria**; P.O.Box 130, NL-6100 AC Echt (NL). **VAN MARIS, Antonius Jeroen Adriaan**; Technische Universiteit Delft, Sectie Industriële Microbiologie Delft, Afdeling Biotechnologie, P.O.Box 5, NL-2600 AA Delft (NL). **PRONK, Jacobus Thomas**; Technische Universiteit Delft, Sectie Industriële Microbiologie Delft, Afdeling Biotechnologie, P.O.Box 5, NL-2600 AA Delft (NL).

(74) **Agent**: **KLEIBORN, Paul Erik**; DSM Intellectual Property, P.O. Box 130, NL-6100 AC Echt (NL).

(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, BN, 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, 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, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, 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, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, 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, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

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

WO 2014/033018 A1

(54) **Title**: YEAST STRAINS ENGINEERED TO PRODUCE ETHANOL FROM ACETATE

(57) **Abstract**: The present invention relates to a yeast cell comprising: a) a disruption of deletion of one or more gene native in the yeast cell coding for an enzyme having glycerolphosphate dehydrogenase (GPD) activity; b) an exogenous polynucleotide coding for an enzyme having aldehyde oxidoreductase activity, wherein the polynucleotide comprises a nucleotide sequence coding for an amino acid sequence with at least 50% amino acid sequence identity with SEQ ID NO: 12.

YEAST STRAINS ENGINEERED TO PRODUCE ETHANOL FROM ACETATE

Field of the invention

The present invention relates to metabolic engineering in microorganisms such as yeast. In particular the invention relates to yeast strains that have been engineered to produce ethanol from acetate and have reduced glycerol production. These strains have retained their natural ability to produce ethanol from hexoses (glucose, fructose, galactose, etc) and comprise an engineered ability to produce ethanol from pentoses like xylose and arabinose. The invention further relates to the processes wherein the engineered strains of the invention produce ethanol from acetic acid, either as main fermentation feedstock, or concomitantly with one or more of hexoses and pentoses.

Background of the invention

Large-scale consumption of traditional, fossil fuels (petroleum-based fuels) in recent decades has contributed to high levels of pollution. This, along with the realisation that the world stock of fossil fuels is not limited and a growing environmental awareness, has stimulated new initiatives to investigate the feasibility of alternative fuels such as ethanol, which is a particulate-free burning fuel source that releases less CO_2 than unleaded gasoline on a per litre basis.

Although biomass-derived ethanol may be produced by the fermentation of hexose sugars obtained from many different sources, the substrates typically used for commercial scale production of fuel alcohol, such as cane sugar and corn starch, are expensive. Increases in the production of fuel ethanol will therefore require the use of lower-cost feedstocks.

Currently, only lignocellulosic feedstock derived from plant biomass is available in sufficient quantities to substitute the crops currently used for ethanol production. In most lignocellulosic material, the second-most-common sugar, after glucose, is xylose. Thus, for an economically feasible fuel production process, both hexose and pentose sugars must be fermented to form ethanol. The yeast *Saccharomyces cerevisiae* is robust and well adapted for ethanol production, but it

is unable to produce ethanol using xylose as a carbon source. Also, no naturally-occurring organisms are known which can ferment xylose to ethanol with both a high ethanol yield and high ethanol productivity. Also it is desirable to convert the available sugars to the desired fermentation product, e.g. ethanol meaning that the formation of byproducts, e.g. in the case of ethanol, glycerol is avoided as much as possible.

There is therefore a need for an organism possessing these properties so as to enable the commercially-viable production of ethanol from lignocellulosic feedstocks.

Guadalupe Medina et al. (2009, Appl. Environ. Microbiol., 76: 190–195) disclose a *S. cerevisiae* strain wherein production of the by-product glycerol is eliminated by disruption of the endogenous NAD-dependent glycerol 3-phosphate dehydrogenase genes (*GPD1* and *GPD2*). Expression of the *E. coli mhpF* gene, encoding the acetylating NAD-dependent aldehyde oxidoreductase restored the ability of the GPD-disrupted strain to grow anaerobically. The GPD-disrupted strain grew anaerobic in a medium supplemented with acetic acid. However, the ethanol production and/or acetic acid consumption of these strains is not particularly high.

Summary of the invention

It is therefore an object of the present invention to provide for yeasts that are capable of producing ethanol from acetic acid or acetate while retaining their abilities of fermenting hexoses (glucose, fructose, galactose, etc) as well as pentoses like xylose, as well as processes wherein these strains are used for the production of ethanol and/or other fermentation products. One or more of the objects are attained according to the invention that provides a yeast cell comprising:

a) a disruption or deletion of one or more gene native in the yeast cell coding for an enzyme having glycerolphosphate dehydrogenase activity;

b) an exogenous polynucleotide coding for an enzyme having aldehyde oxidoreductase activity, wherein the polynucleotide comprises a nucleotide sequence coding for an amino acid sequence with at least 50 % amino acid sequence identity with SEQ ID NO: 12.

Brief description of the drawings

Figure 1 sets out plasmid pSUC227;

Figure 2 sets out plasmid pSUC225;

Figure 3 sets out: schematically the mechanism of gene replacement in example 1;

Figure 4 sets out a scheme of genomic recombination between the loxP-sites is catalyzed by the cre-recombinase, resulting in the excision of the gene encoding cre-recombinase and the KanMX marker conferring resistance to G418.

Figure 5 sets out a physical map of plasmid pBOL033

Figure 6 sets out a scheme of the introduction of the KanMX-marker in yeast expression plasmids, wherein pBOL033-Acdh should be read as pBOL033-eutE, and Acdh > as eutE >.

Figure 7: CO_2 production of strains BIE301-pPWT240 and BIE301-pPWT252 in CFMM1M medium supplemented with 2 grams HAc per liter

Brief description of the sequence listing

SEQ ID NO: 1 sets out forward primer *GPD1*- in *pSUC227*;

SEQ ID NO: 2 sets out reverse primer in *pSUC227*;

SEQ ID NO: 3 sets out forward primer in *pSUC225*;

SEQ ID NO: 4 sets out reverse primer *GPD1*- in pSUC225;

SEQ ID NO: 5 sets out forward primer for *GPD1*;

SEQ ID NO: 6 sets out reverse primer for *GPD1*;

SEQ ID NO: 7 sets out forward primer *GPD2*-in *pSUC227*;

SEQ ID NO: 8 sets out reverse primer *GPD2*-in *pSUC225*;

SEQ ID NO: 9 sets out forward primer for *GPD2*;

SEQ ID NO: 10 sets out reverse primer for *GPD2*;

SEQ ID NO: 11 sets out nucleotide sequence of aldehyde oxidoreductase (*Listeria innocula*);

SEQ ID NO: 12 sets out amino acid sequence of aldehyde oxidoreductase (*Listeria innocula*);

SEQ ID NO: 13 sets out forward primer *KanMX TRP1*;

SEQ ID NO: 14 sets out reverse primer *KanMX TRP1*;

SEQ ID NO: 15 sets out forward primer *KanMX URA3*;

SEQ ID NO: 16 sets out reverse primer *KanMX URA3*

SEQ ID NO: 17 sets out amino acid sequence of GPD1 of *S. cerevisiae*;
SEQ ID NO: 18 sets out amino acid sequence of GPD2 of *S. cerevisiae*
SEQ ID NO: 19 sets out the amino acid sequence of ACS1 of *S. cerevisiae*
SEQ ID NO: 20 sets out the amino acid sequence of ACS2 of *S. cerevisiae*
SEQ ID NO: 21 sets out the nucleotide sequence of GPD1 of *S. cerevisiae*;
SEQ ID NO: 22 sets out the nucleotide sequence of GPD2 of *S. cerevisiae*
SEQ ID NO: 23 sets out the nucleotide sequence of ACS1 of *S. cerevisiae*
SEQ ID NO: 24 sets out the nucleotide sequence of ACS2 of *S. cerevisiae*

Detailed description of the invention

Definitions

Amino acid or nucleotide sequences are said to be homologous when exhibiting a certain level of similarity. Two sequences being homologous indicate a common evolutionary origin. Whether two homologous sequences are closely related or more distantly related is indicated by “percent identity” or “percent similarity”, which is high or low respectively. Although disputed, to indicate “percent identity” or “percent similarity”, “level of homology” or “percent homology” are frequently used interchangeably. A comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. The skilled person will be aware of the fact that several different computer programs are available to align two sequences and determine the homology between two sequences (Kruskal, J. B. (1983) An overview of sequence comparison In D. Sankoff and J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). The percent identity between two amino acid sequences can be determined using the Needleman and Wunsch algorithm for the alignment of two sequences. (Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). The algorithm aligns amino acid sequences as well as nucleotide sequences. The Needleman-Wunsch algorithm has been implemented in the computer program NEEDLE. For the purpose of this invention the NEEDLE program from the EMBOSS package was used (version 2.8.0 or higher, EMBOSS: The European Molecular Biology Open Software Suite (2000) Rice,P. Longden,I. and Bleasby,A. Trends in Genetics 16, (6) pp276—277,

http://emboss.bioinformatics.nl/). For protein sequences, EBLOSUM62 is used for the substitution matrix. For nucleotide sequences, EDNAFULL is used. Other matrices can be specified. The optional parameters used for alignment of amino acid sequences are a gap-open penalty of 10 and a gap extension penalty of 0.5. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

Global Homology Definition

The homology or identity is the percentage of identical matches between the two full sequences over the total aligned region including any gaps or extensions. The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment including the gaps. The identity defined as herein can be obtained from NEEDLE and is labelled in the output of the program as “IDENTITY”.

Longest Identity Definition

The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment after subtraction of the total number of gaps in the alignment. The identity defined as herein can be obtained from NEEDLE by using the NOBRIEF option and is labeled in the output of the program as “longest-identity”.

A variant of a nucleotide or amino acid sequence disclosed herein may also be defined as a nucleotide or amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the nucleotide or amino acid sequence specifically disclosed herein (e.g. in de the sequence listing).

Optionally, in determining the degree of amino acid similarity, the skilled person may also take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled person. Conservative amino acid substitutions refer to the interchangeability of residues having similar side chains. For example, a group of amino acids having aliphatic side chains is glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains is serine and threonine; a group of amino acids having amide-containing side chains is asparagine and glutamine; a group of amino acids having

aromatic side chains is phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or his; Asp to glu; Cys to ser or ala; Gln to asn; Glu to asp; Gly to pro; His to asn or gln; Ile to leu or val; Leu to ile or val; Lys to arg; gln or glu; Met to leu or ile; Phe to met, leu or tyr; Ser to thr; Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

Nucleotide sequences of the invention may also be defined by their capability to hybridise with parts of specific nucleotide sequences disclosed herein, respectively, under moderate, or preferably under stringent hybridisation conditions. Stringent hybridisation conditions are herein defined as conditions that allow a nucleic acid sequence of at least about 25, preferably about 50 nucleotides, 75 or 100 and most preferably of about 200 or more nucleotides, to hybridise at a temperature of about 65˚C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at 65˚C in a solution comprising about 0.1 M salt, or less, preferably 0.2 x SSC or any other solution having a comparable ionic strength. Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours and preferably washing is performed for at least one hour with at least two changes of the washing solution. These conditions will usually allow the specific hybridisation of sequences having about 90% or more sequence identity.

Moderate conditions are herein defined as conditions that allow a nucleic acid sequences of at least 50 nucleotides, preferably of about 200 or more nucleotides, to hybridise at a temperature of about 45˚C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at room temperature in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength. Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours, and preferably washing is performed for at least one hour with at least two changes of

the washing solution. These conditions will usually allow the specific hybridisation of sequences having up to 50% sequence identity. The person skilled in the art will be able to modify these hybridisation conditions in order to specifically identify sequences varying in identity between 50% and 90%.

A “nucleic acid construct” or “nucleic acid vector” is herein understood to mean a nucleic acid molecule designed by man resulting from the use of recombinant DNA technology. The term “nucleic acid construct” therefore does not include naturally occurring nucleic acid molecules although a nucleic acid construct may comprise (parts of) naturally occurring nucleic acid molecules. The terms “expression vector” or expression construct" refer to nucleotide sequences that are capable of affecting expression of a gene in host cells or host organisms compatible with such sequences. These expression vectors typically include at least suitable transcription regulatory sequences and optionally, 3’ transcription termination signals. Additional factors necessary or helpful in effecting expression may also be present, such as expression enhancer elements. The expression vector will be introduced into a suitable host cell and be able to effect expression of the coding sequence in an in vitro cell culture of the host cell. The expression vector will be suitable for replication in the host cell or organism of the invention.

As used herein, the term “promoter” or “transcription regulatory sequence” refers to a nucleic acid fragment that functions to control the transcription of one or more coding sequences, and is located upstream with respect to the direction of transcription of the transcription initiation site of the coding sequence, and is structurally identified by the presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites and any other DNA sequences, including, but not limited to transcription factor binding sites, repressor and activator protein binding sites, and any other sequences of nucleotides known to one of skill in the art to act directly or indirectly to regulate the amount of transcription from the promoter. A “constitutive” promoter is a promoter that is active in most tissues under most physiological and developmental conditions. An “inducible” promoter is a promoter that is physiologically or developmentally regulated, e.g. by the application of a chemical inducer.

The term “selectable marker” is a term familiar to one of ordinary skill in the art and is used herein to describe any genetic entity which, when expressed, can be used to select for a cell or cells containing the selectable marker. The term “reporter” may be used interchangeably with marker, although it is mainly used to

refer to visible markers, such as green fluorescent protein (GFP). Selectable markers may be dominant or recessive or bidirectional.

As used herein, the term “operably linked” refers to a linkage of polynucleotide elements in a functional relationship. A nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For instance, a transcription regulatory sequence is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous and, where necessary to join two protein encoding regions, contiguous and in reading frame.

The terms “protein” or “polypeptide” are used interchangeably and refer to molecules consisting of a chain of amino acids, without reference to a specific mode of action, size, 3-dimensional structure or origin.

Yeasts are herein defined as eukaryotic microorganisms and include all species of the subdivision Eumycotina (Yeasts: characteristics and identification, J.A. Barnett, R.W. Payne, D. Yarrow, 2000, 3rd ed., Cambridge University Press, Cambridge UK; and, The yeasts, a taxonomic study, C.P. Kurtzman and J.W. Fell (eds) 1998, 4th ed., Elsevier Science Publ. B.V., Amsterdam, The Netherlands) that predominantly grow in unicellular form. Yeasts may either grow by budding of a unicellular thallus or may grow by fission of the organism. Preferred yeasts cells for use in the present invention belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia, Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, and *Yarrowia*. Preferably the yeast is capable of anaerobic fermentation, more preferably anaerobic alcoholic fermentation.

"Fungi" (singular: fungus) are herein understood as heterotrophic eukaryotic microorganism that digest their food externally, absorbing nutrient molecules into their cells. Fungi are a separate kingdom of eukaryotic organisms and include yeasts, molds, and mushrooms. The terms fungi, fungus and fungal as used herein thus expressly includes yeasts as well as filamentous fungi.

The term “gene” means a DNA fragment comprising a region (transcribed region), which is transcribed into an RNA molecule (e.g. an mRNA) in a cell, operably linked to suitable regulatory regions (e.g. a promoter). A gene will usually comprise several operably linked fragments, such as a promoter, a 5’ leader sequence, a coding region and a 3’nontranslated sequence (3’end) comprising a polyadenylation site. “Expression of a gene” refers to the process wherein a DNA

region which is operably linked to appropriate regulatory regions, particularly a promoter, is transcribed into an RNA, which is biologically active, i.e. which is capable of being translated into a biologically active protein or peptide.

The term “homologous” when used to indicate the relation between a given (recombinant) nucleic acid or polypeptide molecule and a given host organism or host cell, is understood to mean that in nature the nucleic acid or polypeptide molecule is produced by a host cell or organisms of the same species, preferably of the same variety or strain. If homologous to a host cell, a nucleic acid sequence encoding a polypeptide will typically (but not necessarily) be operably linked to another (heterologous) promoter sequence and, if applicable, another (heterologous) secretory signal sequence and/or terminator sequence than in its natural environment. It is understood that the regulatory sequences, signal sequences, terminator sequences, etc. may also be homologous to the host cell. In this context, the use of only “homologous” sequence elements allows the construction of “self-cloned” genetically modified organisms (GMO’s) (self-cloning is defined herein as in European Directive 98/81/EC Annex II). When used to indicate the relatedness of two nucleic acid sequences the term “homologous” means that one single-stranded nucleic acid sequence may hybridize to a complementary single-stranded nucleic acid sequence. The degree of hybridization may depend on a number of factors including the amount of identity between the sequences and the hybridization conditions such as temperature and salt concentration as discussed later.

The terms "heterologous" and “exogenous” when used with respect to a nucleic acid (DNA or RNA) or protein refers to a nucleic acid or protein that does not occur naturally as part of the organism, cell, genome or DNA or RNA sequence in which it is present, or that is found in a cell or location or locations in the genome or DNA or RNA sequence that differ from that in which it is found in nature. Heterologous and exogenous nucleic acids or proteins are not endogenous to the cell into which it is introduced, but have been obtained from another cell or synthetically or recombinantly produced. Generally, though not necessarily, such nucleic acids encode proteins, i.e. exogenous proteins, that are not normally produced by the cell in which the DNA is transcribed or expressed. Similarly exogenous RNA encodes for proteins not normally expressed in the cell in which the exogenous RNA is present. Heterologous/exogenous nucleic acids and proteins may also be referred to as foreign nucleic acids or proteins. Any nucleic

acid or protein that one of skill in the art would recognize as foreign to the cell in which it is expressed is herein encompassed by the term heterologous or exogenous nucleic acid or protein. The terms heterologous and exogenous also apply to non-natural combinations of nucleic acid or amino acid sequences, i.e. combinations where at least two of the combined sequences are foreign with respect to each other.

The "specific activity" of an enzyme is herein understood to mean the amount of activity of a particular enzyme per amount of total host cell protein, usually expressed in units of enzyme activity per mg total host cell protein. In the context of the present invention, the specific activity of a particular enzyme may be increased or decreased as compared to the specific activity of that enzyme in an (otherwise identical) wild type host cell.

"Anaerobic conditions" or an anaerobic fermentation process is herein defined as conditions or a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than 5, 2.5 or 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors.

Detailed description of the invention

In a first aspect the invention relates to a yeast cell comprising:

a) a disruption or deletion of one or more gene(s) native to the yeast cell coding for an enzyme having glycerolphosphate dehydrogenase (GPD) activity;

b) an exogenous polynucleotide coding for an enzyme having aldehyde oxidoreductase activity, wherein the polynucleotide comprises a nucleotide sequence coding for an amino acid sequence with at least 50% amino acid sequence identity with SEQ ID NO: 12.

The enzyme b) is herein understood as an enzyme having ethanol amine activity and is referred to as aldehyde oxidoreductase and/or aldehyde dehydrogenase (eutE, EC 1.2.1.10).

Therefore the host cell of the invention comprises an exogenous gene coding for an enzyme with the ability to hydrogenate acetaldehyde and/or is herein understood as an enzyme which catalyzes the reaction (eutE):

$$\text{acetyl-Coenzyme A} + \text{NADH} + H^+ \leftrightarrow \text{acetaldehyde} + NAD^+ + \text{Coenzyme A}$$

Thus, the enzyme catalyzes the conversion of acetaldehyde into Acetyl-CoA (and vice versa) and is also referred to as a ethanolamine utilization enzyme and/or as CoA-dependent propionaldehyde dehydrogenase. The enzyme may be a bifunctional enzyme which further catalyzes the conversion of acetaldehyde into ethanol (and vice versa; see below). For convenience we shall refer herein to an enzyme having at least the ability to reduce acetylCoA into either acetaldehyde or ethanol as an "aldehyde oxidoreductase". It is further understood, that in a preferred embodiment, that the host cell has endogenous alcohol dehydrogenase activities which allow the cell, being provided with aldehyde oxidoreductase activity, to complete the conversion of acetyl-CoA into ethanol. It is further also preferred that the host cell has endogenous acetyl-CoA synthetase which allow the cell, being provided with aldehyde oxidoreductase activity, to complete the conversion of acetic acid (via acetyl-CoA) into ethanol.

The exogenous gene may encode for a monofunctional enzyme having only aldehyde oxidoreductase activity (i.e. an enzyme only having the ability to reduce acetylCoA into acetaldehyde) such as e.g. the aldehyde oxidoreductase encoded by the *L. innocula* gene. The amino acid sequences of these monofunctional enzymes are available in public databases and can be used by the skilled person to design codon-optimised nucleotide sequences coding for the corresponding monofunctional enzyme (see e.g. SEQ ID NO: 11). The exogenous gene coding for the monofunctional enzyme having only aldehyde oxidoreductase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to SEQ ID NO: 11.

A suitable exogenous gene coding for an enzyme with aldehyde oxidoreductase activity is e.g. a prokaryotic aldehyde oxidoreductase, such as the aldehyde oxidoreductase from *Listeria innocula* (SEQ ID NO: 12).

The exogenous gene coding for an enzyme with aldehyde oxidoreductase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 12. Suitable examples of organisms comprising an enzyme with aldehyde oxidoreductase activity include eutE's from genus *Listeria, Veillonella, Yersinia, Shewanella, Klebsiella, Citrobacter, Salmonella, Escherichia, Enterobacter*. Examples of suitable species are given in

Table 1. The amino acid sequences of these enzymes are available in public databases and can be used by the skilled person to design codon-optimised nucleotide sequences coding for the corresponding enzyme with aldehyde oxidoreductase activity (see e.g. SEQ ID NO: 11). The exogenous gene coding for an enzyme with aldehyde oxidoreductase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 11. Preferably the amino acid sequence has no more than 420, 380, 300, 250, 200, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 11.

Table 1: Enzymes with aldehyde oxidoreductase (eutE) activity

Organism	**Amino acid identity (%)**
Listeria innocula	100%
Listeria welshimeri serovar	96%
Listeria marthii	95%
Listeria monocytogenes	95%
Listeria seeligeri	94%
Listeria ivanovii	93%
Veillonella sp. oral taxon	61%
Veillonella sp. 6	60%
Yersinia enterocolitica subsp. palearctica	59%
Yersinia intermedia	58%
Shewanella putrefaciens	57%
Salmonella arizonae	58%
Enterobacter cloacae	58%
Citrobacter rodentium	58%
Salmonella enterica subsp. enterica serovar Kentucky	57%
Escherichia fergusonii	56%
Klebsiella pneumoniae	56%

To increase the likelihood that the enzyme having aldehyde oxidoreductase activity is expressed at sufficient levels and in active form in the transformed host cells of the invention, the nucleotide sequence encoding these enzymes, as well as the aldehyde oxidoreductase activating enzyme and other enzymes of the invention (see below), are preferably adapted to optimise their codon usage to that of the host cell in question. The adaptiveness of a nucleotide sequence encoding an enzyme to the codon usage of a host cell may be expressed as codon adaptation index (CAI). The codon adaptation index is herein defined as a measurement of the relative adaptiveness of the codon usage of a gene towards

the codon usage of highly expressed genes in a particular host cell or organism. The relative adaptiveness (w) of each codon is the ratio of the usage of each codon, to that of the most abundant codon for the same amino acid. The CAI index is defined as the geometric mean of these relative adaptiveness values. Non-synonymous codons and termination codons (dependent on genetic code) are excluded. CAI values range from 0 to 1, with higher values indicating a higher proportion of the most abundant codons (see Sharp and Li , 1987, Nucleic Acids Research 15: 1281-1295; also see: Jansen et al., 2003, Nucleic Acids Res. 31(8):2242-51). An adapted nucleotide sequence preferably has a CAI of at least 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8 or 0.9. Most preferred are the sequences which have been codon optimised for expression in the fungal host cell in question such as e.g. *S. cerevisiae* cells.

According to an aspect of the invention, the host cell of the invention further comprises a genetic modification that reduces or deletes specific glycerolphosphate dehydrogenase activity in the cell. Glycerolphosphate dehydrogenase, also called glycerol-3-phosphate dehydrogenase (NAD+) (EC1.1.1.8) is an enzyme that catalyzes the chemical reaction:

This enzyme belongs to the family of oxidoreductases, specifically those acting on the CH-OH group of donor with NAD+ or NADP+ as acceptor. The systematic name of this enzyme class is sn-glycerol-3-phosphate:NAD+ 2-oxidoreductase. Other names in common use include alpha-glycerol phosphate dehydrogenase (NAD+), alpha-glycerophosphate dehydrogenase (NAD+), glycerol 1-phosphate dehydrogenase, glycerol phosphate dehydrogenase (NAD+), glycerophosphate dehydrogenase (NAD+), hydroglycerophosphate dehydrogenase, L-alpha-glycerol phosphate dehydrogenase, L-alpha-glycerophosphate dehydrogenase, L-glycerol phosphate dehydrogenase, L-glycerophosphate dehydrogenase, NAD+-alpha-glycerophosphate dehydrogenase, NAD+-dependent glycerol phosphate dehydrogenase, NAD+-dependent glycerol-3-phosphate dehydrogenase, NAD+-L-glycerol-3-phosphate dehydrogenase, NAD+-linked glycerol 3-phosphate dehydrogenase, NADH-dihydroxyacetone phosphate reductase, and glycerol-3-phosphate dehydrogenase (NAD+). The enzyme participates in glycerophospholipid metabolism.

In the cells of the invention, the specific glycerolphosphate dehydrogenase activity is preferably reduced by at least a factor 0.8, 0.5, 0.3, 0.1, 0.05 or 0.01 as compared to a strain which is genetically identical except for the genetic modification causing the reduction in specific activity, preferably under anaerobic conditions. Glycerolphosphate dehydrogenase activity may be determined as described by Overkamp et al. (2002, Yeast 19:509-520).

Preferably, glycerolphosphate dehydrogenase activity is reduced in the host cell by one or more genetic modifications that reduce the expression of or inactivates a gene encoding an glycerolphosphate dehydrogenase. Preferably, the genetic modifications reduce or inactivate the expression of each endogenous copy of the gene encoding a specific glycerolphosphate dehydrogenase in the cell's genome. A given cell may comprise multiple copies of the gene encoding a specific glycerolphosphate dehydrogenase with one and the same amino acid sequence as a result of di-, poly- or aneuploidy. In such instances preferably the expression of each copy of the specific gene that encodes the glycerolphosphate dehydrogenase is reduced or inactivated. Alternatively, a cell may contain several different (iso)enzymes with glycerolphosphate dehydrogenase activity that differ in amino acid sequence and that are each encoded by a different gene. In such instances, in some embodiments of the invention it may be preferred that only certain types of the isoenzymes are reduced or inactivated while other types remain unaffected. Preferably, however, expression of all copies of genes encoding (iso)enzymes with glycerolphosphate dehydrogenase activity is reduced or inactivated.

Preferably, a gene encoding glycerolphosphate dehydrogenase activity is inactivated by deletion of at least part of the gene or by disruption of the gene, whereby in this context the term gene also includes any non-coding sequence up- or down-stream of the coding sequence, the (partial) deletion or inactivation of which results in a reduction of expression of glycerolphosphate dehydrogenase activity in the host cell.

A preferred gene encoding a glycerolphosphate dehydrogenase whose activity is to be reduced or inactivated in the cell of the invention is the *S. cerevisiae GPD1* as described by van den Berg and Steensma (1997, Yeast 13:551-559), encoding the amino acid sequence of SEQ ID NO: 17 and orthologues thereof in other species. Therefore a gene encoding a

glycerolphosphate dehydrogenase whose activity is to be reduced or inactivated in the cell of the invention preferably is a gene encoding a glycerolphosphate dehydrogenase having an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98 or 99% sequence identity to SEQ ID NO: 17 or a gene encoding a glycerolphosphate dehydrogenase having an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO:17.

Suitable examples of organisms (hosts) comprising an enzyme with glycerolphosphate dehydrogenase activity belonging to the genus *Saccharomyces, Naumovozyna, Candida Vanderwaltozyma and Zygosaccharomyces* are provided in Table 4.

Table 4: Enzymes with glycerolphosphate dehydrogenase (GPD1) activity

Organism	**Amino acid identity (%)**
S. cerevisiae	100%
Naumovozyma dairenensis	79%
Naumovozyma castellii	80%
Candida glabrata	77%
Vanderwaltozyma polyspora	77%
Zygosaccharomyces rouxii	74%
Saccharomycopsis fibuligera	61%

However, in some strains e.g. of *Saccharomyces, Candida* and *Zygosaccharomyces* a second gene encoding a glycerolphosphate dehydrogenase is active, i.e. the *GPD2*, see e.g. Overkamp et al. (2002, *supra*). Another preferred gene encoding a glycerolphosphate dehydrogenase whose activity is to be reduced or inactivated in the cell of the invention therefore is an *S. cerevisiae GPD2* as described by Overkamp et al. (2002, *supra*), encoding the amino acid sequence of SEQ ID NO: 18 and orthologues thereof in other species. Therefore a gene encoding a glycerolphosphate dehydrogenase*GPD2* whose activity is to be reduced or inactivated in the cell of the invention preferably is a gene encoding a glycerolphosphate dehydrogenasehaving an amino acid

sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98 or 99% sequence identity to SEQ ID NO: 18 or a gene encoding a glycerolphosphate dehydrogenasehaving an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 18.

Suitable examples of organisms (hosts) comprising an enzyme with glycerolphosphate dehydrogenase activity belonging to the genus (Zygo)*Saccharomyces* and *Candida* are provided in Table 5.

Table 5: Enzymes with glycerolphosphate dehydrogenase(GPD2) activity

Organism	**Amino acid identity (%)**
S. cerevisiae	100%
Candida glabrata	75%
Zygosaccharomyces rouxii	73%
Spathaspora passalidarum	62%
Scheffersomyces stipitis	61%

In one aspect the invention relates to methods for preparing or constructing the yeast cells of the invention. For this purpose standard genetic and molecular biology techniques are used that are generally known in the art and have e.g. been described by Sambrook and Russell (2001, "Molecular cloning: a laboratory manual" (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press) and Ausubel et al. (1987,eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York). Furthermore, the construction of mutated host yeast strains is carried out by genetic crosses, sporulation of the resulting diploids, tetrad dissection of the haploid spores containing the desired auxotrophic markers, and colony purification of such haploid host yeasts in the appropriate selection medium. All of these methods are standard yeast genetic methods known to those in the art. See, for example, Sherman et al., Methods Yeast Genetics, Cold Spring Harbor Laboratory, NY (1978) and Guthrie et al. (Eds.) Guide To Yeast Genetics and Molecular Biology Vol. 194, Academic Press, San Diego (1991).

The exogenous genes coding for the enzyme having aldehyde oxidoreductase activity preferably are expression constructs comprising the nucleotide sequence coding for the enzymes operably linked to suitable expression regulatory regions/sequences to ensure expression of the enzymes upon transformation of the expression constructs into the host cell of the invention. Thus, the gene or expression construct will at least comprise a promoter that is functional in the host cell operably linked to the coding sequence. The gene or construct may further comprise a 5' leader sequence upstream of the coding region and a 3'-nontranslated sequence (3'end) comprising a polyadenylation site and a transcription termination site downstream of the coding sequence. It is understood that the nucleotide sequences coding for the enzyme having aldehyde oxidoreductase activity may be present together on a single expression construct, or each enzyme may be present on a separate expression construct.

Suitable promoters for expression of the nucleotide sequences coding for the enzyme having aldehyde oxidoreductase activity and the (as well as other enzymes of the invention; see below) include promoters that are preferably insensitive to catabolite (glucose) repression, that are active under anaerobic conditions and/or that preferably do not require xylose or arabinose for induction. Promoters having these characteristics are widely available and known to the skilled person. Suitable examples of such promoters include e.g. promoters from glycolytic genes such as the phosphofructokinase (*PFK*), triose phosphate isomerase (*TPI*), glyceraldehyde-3-phosphate dehydrogenase (*GPD, TDH3* or *GAPDH*), pyruvate kinase (*PYK*), phosphoglycerate kinase (*PGK*), glucose-6-phosphate isomerase promoter (*PGI1*) promoters from yeasts. More details about such promoters from yeast may be found in (WO 93/03159). Other useful promoters are ribosomal protein encoding gene promoters (*TEF1*), the lactase gene promoter (*LAC4*), alcohol dehydrogenase promoters (*ADH1, ADH4*, and the like), the enolase promoter (*ENO*) and the hexose (glucose) transporter promoter (*HXT7*). Alternatively, the nucleotide sequences encoding the enzyme having aldehyde oxidoreductase activity are expressed under anaerobic conditions by using an anoxic promoter such as e.g. the *S. cerevisiae ANB1* promoter. Other promoters, both constitutive and inducible, and enhancers or upstream activating sequences will be known to those of skill in the art. Preferably the promoter that is operably linked to nucleotide sequence as defined above is homologous to the

host cell. Suitable terminator sequences are e.g. obtainable from the cytochrome c1 (*CYC1*) gene or an alcohol dehydrogenase gene (e.g. *ADH1*).

The fungal host cell to be transformed with a nucleic acid construct comprising a nucleotide sequence encoding an enzyme with aldehyde oxidoreductase activity preferably is a yeast host cell. Preferably the host cell is a cultured cell. The host cell of the invention preferably is a host capable of active or passive pentose (xylose and preferably also arabinose) transport into the cell. The host cell preferably contains active glycolysis. The host cell may further preferably contain an endogenous pentose phosphate pathway and may contain endogenous xylulose kinase activity so that xylulose isomerised from xylose may be metabolised into pyruvate. The host further preferably contains enzymes for conversion of a pentose (preferably through pyruvate) to a desired fermentation product such as ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, butanols (1-butanol, 2-butanol, isobutanol) and isoprenoid-derived products. A particularly preferred host cell is a yeast cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. The yeast host cell further preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than 5, 4, or 3) and towards organic acids like lactic acid, acetic acid or formic acid and sugar degradation products such as furfural and hydroxy-methylfurfural, and a high tolerance to elevated temperatures. Any of these characteristics or activities of the host cell may be naturally present in the host cell or may be introduced or modified by genetic modification, preferably by self cloning or by the methods of the invention described below. A suitable cell is a cultured cell, a cell that may be cultured in fermentation process e.g. in submerged or solid state fermentation. Particularly suitable host cells are eukaryotic microorganism like e.g. fungi, however, most suitable for use in the present inventions are yeasts.

Yeasts are herein defined as eukaryotic microorganisms and include all species of the subdivision Eumycotina (Yeasts: characteristics and identification, J.A. Barnett, R.W. Payne, D. Yarrow, 2000, 3rd ed., Cambridge University Press, Cambridge UK; and, The yeasts, a taxonomic study, C.P. Kurtzman and J.W. Fell (eds) 1998, 4th ed., Elsevier Science Publ. B.V., Amsterdam, The Netherlands) that predominantly grow in unicellular form. Yeasts may either grow by budding of a unicellular thallus or may grow by fission of the organism. Preferred yeasts cells for use in the present invention belong to the genera *Saccharomyces*, *Kluyveromyces*,

Candida, *Pichia, Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, and *Yarrowia*. Preferably the yeast is capable of anaerobic fermentation, more preferably anaerobic alcoholic fermentation. Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i.e., a high acid-, ethanol- and osmotolerance, capability of anaerobic growth, and of course its high alcoholic fermentative capacity. Preferred yeast species as host cells include *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis, K. marxianus* and *Schizosaccharomyces pombe*.

In a further embodiment, the host cell of the invention further comprises a genetic modification that increases the specific acetyl-CoA synthetase activity in the cell, preferably under anaerobic conditions as this activity is rate-limiting under these conditions. Acetyl-CoA synthetase or acetate-CoA ligase (EC 6.2.1.1) is herein understood as an enzyme that catalyzes the formation of a new chemical bond between acetate and coenzyme A (CoA). Preferably the genetic modification causes overexpression of a acetyl-CoA synthetase, e.g. by overexpression of a nucleotide sequence encoding a acetyl-CoA synthetase. The nucleotide sequence encoding the acetyl-CoA synthetase may be endogenous to the cell or may be a acetyl-CoA synthetase that is heterologous to the cell. Nucleotide sequences that may be used for overexpression of acetyl-CoA synthetase in the cells of the invention are e.g. the acetyl-CoA synthetase genes from *S. cerevisiae* (*ACS1* and *ACS2*) as e.g. described by de Jong-Gubbels et al. (1998, FEMS Microbiol Lett. 165: 15-20). Preferably, the nucleotide sequence encoding the acetyl-CoA synthetase comprises an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with at least one of SEQ ID NO's: 19 and 20, or a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to at least one of SEQ ID NO's: 19 and 20.

In one embodiment, the nucleotide sequence that is overexpressed encodes an acetyl-CoA synthetase with a high affinity for acetate. Use of an acetyl-CoA synthetase with a high affinity for acetate is preferred for conditions under which there is a relatively low concentration of acetic acid in the culture medium, e.g. no more than 2 g acetic acid/L culture medium. An acetyl-CoA synthetase with a high

affinity for acetate is herein defined as an acetyl-CoA synthetase with a higher affinity for acetate than the acetyl-CoA synthetase encoded by the *S. cerevisiae ACS2* (SEQ ID NO: 20). Preferably, an acetyl-CoA synthetase with a high affinity for acetate has a Km for acetate of no more than 10, 5, 2, 1, 0.5, 0.2 or 0.1 mM, such e.g. the acetyl-CoA synthetase encoded by the *S. cerevisiae ACS1* gene. More preferably a codon-optimised (see above) nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 19 is overexpressed.

In another embodiment, the nucleotide sequence that is overexpressed encodes an acetyl-CoA synthetase with a high maximum rate (V_{max}). Use of an acetyl-CoA synthetase with a high maximum rate is preferred for condition under which there is a relatively high concentration of acetic acid in the culture medium, e.g. at least 2, 3, 4 or 5 g acetic acid/L culture medium. An acetyl-CoA synthetase with a high maximum rate is herein defined as an acetyl-CoA synthetase with a higher maximum rate than the acetyl-CoA synthetase encoded by the *S. cerevisiae ACS2*. Preferably, the acetyl-CoA synthetase with a high maximum rate is the acetyl-CoA synthetase encoded by the *S. cerevisiae ACS1* gene. More preferably a codon-optimised (see above) nucleotide sequence encoding the amino acid sequence of SEQ ID NO: 19 is overexpressed.

For overexpression of the nucleotide sequence encoding the acetyl-CoA synthetase (to be overexpressed) can be placed in an expression construct wherein it is operably linked to suitable expression regulatory regions/sequences to ensure overexpression of the acetyl-CoA synthetase enzyme upon transformation of the expression construct into the host cell of the invention (see above). Suitable promoters for (over)expression of the nucleotide sequence coding for the enzyme having acetyl-CoA synthetase activity include promoters that are preferably insensitive to catabolite (glucose) repression, that are active under anaerobic conditions and/or that preferably do not require xylose or arabinose for induction. Examples of such promoters are given above. In the cells of the invention, an acetyl-CoA synthetase to be overexpressed is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to a strain which is genetically identical except for the genetic modification causing the overexpression. Preferably, the acetyl-CoA synthetase is overexpressed under anaerobic conditions by at least a factor 2, 5, 10, 20, 50, or 100 as compared to a strain which is genetically identical except for the genetic modification causing the overexpression. It is to be understood that these levels of overexpression may

apply to the steady state level of the enzyme's activity (specific activity), the steady state level of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

In a further preferred embodiment, the host cell of the invention has at least one of: a) the ability of isomerising xylose to xylulose; and, b) the ability to convert L-arabinose into D-xylulose 5-phosphate. For a) the cell preferably has a functional exogenous xylose isomerase gene, which gene confers to the cell the ability to isomerise xylose into xylulose. For b) the cell preferably has functional exogenous genes coding for a L-arabinose isomerase, a L-ribulokinase and a L-ribulose-5-phosphate 4-epimerase, which genes together confers to the cell the ability to isomerise convert L-arabinose into D-xylulose 5-phosphate.

Fungal host cells having the ability of isomerising xylose to xylulose as e.g. described in WO 03/0624430 and in WO 06/009434. The ability of isomerising xylose to xylulose is preferably conferred to the cell by transformation with a nucleic acid construct comprising a nucleotide sequence encoding a xylose isomerase. Preferably the cell thus acquires the ability to directly isomerise xylose into xylulose. More preferably the cell thus acquires the ability to grow aerobically and/or anaerobically on xylose as sole energy and/or carbon source though direct isomerisation of xylose into xylulose (and further metabolism of xylulose). It is herein understood that the direct isomerisation of xylose into xylulose occurs in a single reaction catalysed by a xylose isomerase, as opposed to the two step conversion of xylose into xylulose via a xylitol intermediate as catalysed by xylose reductase and xylitol dehydrogenase, respectively.

Several xylose isomerases (and their amino acid and coding nucleotide sequences) that may be successfully used to confer to the cell of the invention the ability to directly isomerise xylose into xylulose have been described in the art. These include the xylose isomerases of *Piromyces* sp. and of other anaerobic fungi that belongs to the families *Neocallimastix*, *Caecomyces*, *Piromyces* or *Ruminomyces* (WO 03/0624430), *Cyllamyces aberensis* (US 20060234364), *Orpinomyces* (Madhavan et al., 2008, DOI 10.1007/s00253-008-1794-6), the xylose isomerase of the bacterial genus *Bacteroides*, including e.g. *B.thetaiotaomicron* (WO 06/009434), *B. fragilis*, and *B. uniformis* (WO 09/109633), the xylose isomerase of the anaerobic bacterium *Clostridium phytofermentans* (Brat et al., 2009, Appl. Environ. Microbiol. 75:2304–2311), and the xylose

isomerases of *Clostridium difficile*, *Ciona intestinales* and *Fusobacterium mortiferum* (WO 10/074577).

Fungal host cells having the ability to convert L-arabinose into D-xylulose 5-phosphate as e.g. described in Wisselink et al. (2007, Appl. Environ. Microbiol. doi:10.1128/AEM.00177-07) and in EP 1 499 708. The ability of to converting L-arabinose into D-xylulose 5-phosphate is preferably conferred to the cell by transformation with a nucleic acid construct(s) comprising nucleotide sequences encoding a) an arabinose isomerase; b) a ribulokinase, preferably a L-ribulokinase a xylose isomerase; and c) a ribulose-5-P-4-epimerase, preferably a L-ribulose-5-P-4-epimerase. Preferably, in the cells of the invention, the ability to convert L-arabinose into D-xylulose 5-phosphate is the ability to convert L-arabinose into D-xylulose 5-phosphate through the subsequent reactions of 1) isomerisation of arabinose into ribulose; 2) phosphorylation of ribulose to ribulose 5-phosphate; and, 3) epimerisation of ribulose 5-phosphate into D-xylulose 5-phosphate. Suitable nucleotide sequences encoding arabinose isomerases, a ribulokinases and ribulose-5-P-4-epimerases may be obtained from *Bacillus subtilis*, *Listeria innocula* (see e.g. EP 1 499 708), *Lactobacilli*, e.g. *Listeria innocula* (see e.g. Wisselink et al. *supra*; WO2008/041840), or species of *Clavibacter*, *Arthrobacter* and *Gramella*, of which preferably *Clavibacter michiganensis, Arthrobacter aurescens* and *Gramella forsetii* (see WO2009/011591).

The transformed host cell of the invention further preferably comprises xylulose kinase activity so that xylulose isomerised from xylose may be metabolised to pyruvate. Preferably, the cell contains endogenous xylulose kinase activity. More preferably, a cell of the invention comprises a genetic modification that increases the specific xylulose kinase activity. Preferably the genetic modification causes overexpression of a xylulose kinase, e.g. by overexpression of a nucleotide sequence encoding a xylulose kinase. The gene encoding the xylulose kinase may be endogenous to the cell or may be a xylulose kinase that is heterologous to the cell. A nucleotide sequence that may be used for overexpression of xylulose kinase in the cells of the invention is e.g. the xylulose kinase gene from *S. cerevisiae* (*XKS1*) as described by Deng and Ho (1990, Appl. Biochem. Biotechnol. 24-25: 193-199). Another preferred xylulose kinase is a xylose kinase that is related to the xylulose kinase from *Piromyces* (*xylB*; see WO 03/0624430). This *Piromyces* xylulose kinase is actually more related to prokaryotic kinase than to all of the known eukaryotic kinases such as the yeast

kinase. The eukaryotic xylulose kinases have been indicated as non-specific sugar kinases, which have a broad substrate range that includes xylulose. In contrast, the prokaryotic xylulose kinases, to which the *Piromyces* kinase is most closely related, have been indicated to be more specific kinases for xylulose, i.e. having a narrower substrate range. In the cells of the invention, a xylulose kinase to be overexpressed is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to a strain which is genetically identical except for the genetic modification causing the overexpression. It is to be understood that these levels of overexpression may apply to the steady state level of the enzyme's activity, the steady state level of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

A cell of the invention further preferably comprises a genetic modification that increases the flux of the pentose phosphate pathway as described in WO 06/009434. In particular, the genetic modification causes an increased flux of the non-oxidative part pentose phosphate pathway. A genetic modification that causes an increased flux of the non-oxidative part of the pentose phosphate pathway is herein understood to mean a modification that increases the flux by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to the flux in a strain which is genetically identical except for the genetic modification causing the increased flux. The flux of the non-oxidative part of the pentose phosphate pathway may be measured as described in WO 06/009434.

Genetic modifications that increase the flux of the pentose phosphate pathway may be introduced in the cells of the invention in various ways. These including e.g. achieving higher steady state activity levels of xylulose kinase and/or one or more of the enzymes of the non-oxidative part pentose phosphate pathway and/or a reduced steady state level of unspecific aldose reductase activity. These changes in steady state activity levels may be effected by selection of mutants (spontaneous or induced by chemicals or radiation) and/or by recombinant DNA technology e.g. by overexpression or inactivation, respectively, of genes encoding the enzymes or factors regulating these genes.

In a preferred cell of the invention, the genetic modification comprises overexpression of at least one enzyme of the (non-oxidative part) pentose phosphate pathway. Preferably the enzyme is selected from the group consisting of the enzymes encoding for ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, transketolase and transaldolase. Various combinations of

enzymes of the (non-oxidative part of the) pentose phosphate pathway may be overexpressed. E.g. the enzymes that are overexpressed may be at least the enzymes ribulose-5-phosphate isomerase and ribulose-5-phosphate 3-epimerase; or at least the enzymes ribulose-5-phosphate isomerase and transketolase; or at least the enzymes ribulose-5-phosphate isomerase and transaldolase; or at least the enzymes ribulose-5-phosphate 3-epimerase and transketolase; or at least the enzymes ribulose-5-phosphate 3-epimerase and transaldolase; or at least the enzymes transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate 3-epimerase, transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, and transketolase. In one embodiment of the invention each of the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, transketolase and transaldolase are overexpressed in the cell of the invention. Preferred is a cell in which the genetic modification comprises at least overexpression of the enzyme transaldolase. More preferred is a cell in which the genetic modification comprises at least overexpression of both the enzymes transketolase and transaldolase. In fact, under some conditions we have found that cells overexpressing only the transketolase and the transaldolase already have the same anaerobic growth rate on xylose as do cells that overexpress all four of the enzymes, i.e. the ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, transketolase and transaldolase. Moreover, cells of the invention overexpressing both of the enzymes ribulose-5-phosphate isomerase and ribulose-5-phosphate 3-epimerase are preferred over cells overexpressing only the isomerase or only the 3-epimerase as overexpression of only one of these enzymes may produce metabolic imbalances.

There are various means available in the art for overexpression of enzymes in the host cells of the invention. In particular, an enzyme may be overexpressed by increasing the copy number of the gene coding for the enzyme in the cell, e.g. by integrating additional copies of the gene in the cell's genome, by expressing the gene from an episomal multicopy expression vector or by introducing a episomal expression vector that comprises multiple copies of the gene. The coding sequence used for overexpression of the enzymes preferably is homologous to the host cell of the invention. However, coding sequences that are heterologous to the

host cell of the invention may likewise be applied. Alternatively overexpression of enzymes in the cells of the invention may be achieved by using a promoter that is not native to the sequence coding for the enzyme to be overexpressed, i.e. a promoter that is heterologous to the coding sequence to which it is operably linked. Although the promoter preferably is heterologous to the coding sequence to which it is operably linked, it is also preferred that the promoter is homologous, i.e. endogenous to the cell of the invention. Preferably the heterologous promoter is capable of producing a higher steady state level of the transcript comprising the coding sequence (or is capable of producing more transcript molecules, i.e. mRNA molecules, per unit of time) than is the promoter that is native to the coding sequence, preferably under conditions where one or more of xylose, arabinose and glucose are available as carbon sources, more preferably as major carbon sources (i.e. more than 50% of the available carbon source consists of one or more of xylose, arabinose and glucose), most preferably as sole carbon sources.

A further preferred cell of the invention comprises a genetic modification that reduces unspecific aldose reductase activity in the cell. Preferably, unspecific aldose reductase activity is reduced in the host cell by one or more genetic modifications that reduce the expression of or inactivates a gene encoding an unspecific aldose reductase. Preferably, the genetic modifications reduce or inactivate the expression of each endogenous copy of a gene encoding an unspecific aldose reductase that is capable of reducing an aldopentose, including, xylose, xylulose and arabinose, in the cell's genome. A given cell may comprise multiple copies of genes encoding unspecific aldose reductases as a result of di-, poly- or aneuploidy, and/or a cell may contain several different (iso)enzymes with aldose reductase activity that differ in amino acid sequence and that are each encoded by a different gene. Also in such instances preferably the expression of each gene that encodes an unspecific aldose reductase is reduced or inactivated. Preferably, the gene is inactivated by deletion of at least part of the gene or by disruption of the gene, whereby in this context the term gene also includes any non-coding sequence up- or down-stream of the coding sequence, the (partial) deletion or inactivation of which results in a reduction of expression of unspecific aldose reductase activity in the host cell. A nucleotide sequence encoding an aldose reductase whose activity is to be reduced in the cell of the invention and amino acid sequences of such aldose reductases are described in WO 06/009434 and include e.g. the (unspecific) aldose reductase genes of *S. cerevisiae GRE3*

gene (Träff et al., 2001, Appl. Environm. Microbiol. 67: 5668-5674) and orthologues thereof in other species.

A further preferred transformed host cell according to the invention may comprises further genetic modifications that result in one or more of the characteristics selected from the group consisting of (a) increased transport of xylose and/or arabinose into the cell; (b) decreased sensitivity to catabolite repression; (c) increased tolerance to ethanol, osmolarity or organic acids; and, (d) reduced production of by-products. By-products are understood to mean carbon-containing molecules other than the desired fermentation product and include e.g. xylitol, arabinitol, glycerol and/or acetic acid. Any genetic modification described herein may be introduced by classical mutagenesis and screening and/or selection for the desired mutant, or simply by screening and/or selection for the spontaneous mutants with the desired characteristics. Alternatively, the genetic modifications may consist of overexpression of endogenous genes and/or the inactivation of endogenous genes. Genes the overexpression of which is desired for increased transport of arabinose and/or xylose into the cell are preferably chosen form genes encoding a hexose or pentose transporter. In *S. cerevisiae* and other yeasts these genes include *HXT1, HXT2, HXT3, HXT4, HXT5, HXT7* and *GAL2*, of which *HXT7, HXT5* and *GAL2* are most preferred (see Sedlack and Ho, Yeast 2004; 21: 671–684). Another preferred transporter for expression in yeast is the glucose transporter encoded by the *P. stipitis SUT1* gene (Katahira et al., 2008, Enzyme Microb. Technol. 43: 115-119). Similarly orthologues of these transporter genes in other species may be overexpressed. Other genes that may be overexpressed in the cells of the invention include genes coding for glycolytic enzymes and/or ethanologenic enzymes such as alcohol dehydrogenases. Preferred endogenous genes for inactivation include hexose kinase genes e.g. the *S. cerevisiae HXK2* gene (see Diderich et al., 2001, Appl. Environ. Microbiol. 67: 1587-1593); the *S. cerevisiae MIG1* or *MIG2* genes; genes coding for enzymes involved in glycerol metabolism such as the *S. cerevisiae* glycerol-phosphate dehydrogenase 1 and/or 2 genes; or (hybridising) orthologues of these genes in other species. Other preferred further modifications of host cells for xylose fermentation are described in van Maris et al. (2006, Antonie van Leeuwenhoek 90:391–418), WO2006/009434, WO2005/023998, WO2005/111214, and WO2005/091733. Any of the genetic modifications of the cells of the invention as described herein are, in as far as possible, preferably introduced or modified by self cloning genetic modification.

A preferred host cell according to the invention has the ability to grow on at least one of xylose and arabinose as carbon/energy source, preferably as sole carbon/energy source, and preferably under anaerobic conditions, i.e. conditions as defined herein below for anaerobic fermentation process. Preferably, when grown on xylose as carbon/energy source the host cell produces essentially no xylitol, e.g. the xylitol produced is below the detection limit or e.g. less than 5, 2, 1, 0.5, or 0.3% of the carbon consumed on a molar basis. Preferably, when grown on arabinose as carbon/energy source, the cell produces essentially no arabinitol, e.g. the arabinitol produced is below the detection limit or e.g. less than 5, 2, 1, 0.5, or 0.3 % of the carbon consumed on a molar basis.

A preferred host cell of the invention has the ability to grow on at least one of a hexose, a pentose, glycerol, acetic acid and combinations thereof at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0,25 or 0,3 h^{-1} under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h^{-1} under anaerobic conditions. Therefore, preferably the host cell has the ability to grow on at least one of xylose and arabinose as sole carbon/energy source at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0,25 or 0,3 h^{-1} under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h^{-1} under anaerobic conditions. More preferably, the host cell has the ability to grow on a mixture of a hexose (e.g. glucose) and at least one of xylose and arabinose (in a 1:1 weight ratio) as sole carbon/energy source at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0,25 or 0,3 h^{-1} under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h^{-1} under anaerobic conditions. Most preferably, the host cell has the ability to grow on a mixture of a hexose (e.g. glucose), at least one of xylose and arabinose and glycerol (in a 1:1:1 weight ratio) as sole carbon/energy source at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0,25 or 0,3 h^{-1} under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h^{-1} under anaerobic conditions.

Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i. e. , a high acid-, ethanol-and osmo- tolerance, capability of anaerobic growth, and of course its high

alcoholic fermentative capacity. Preferred yeast species as host cells include *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* or *K fragilis*.

A cell of the invention may be able to convert plant biomass, celluloses, hemicelluloses, pectins,rhamnose, galactose, frucose, maltose, maltodextrines, ribose, ribulose, or starch, starch derivatives, sucrose, lactose and glycerol, for example into fermentable sugars. Accordingly, a cell of the invention may express one or more enzymes such as a cellulase (an endocellulase or an exocellulase), a hemicellulase (an endo- or exo-xylanase or arabinase) necessary for the conversion of cellulose into glucose monomers and hemicellulose into xylose and arabinose monomers, a pectinase able to convert pectins into glucuronic acid and galacturonic acid or an amylase to convert starch into glucose monomers.

The cell further preferably comprises those enzymatic activities required for conversion of pyruvate to a desired fermentation product, such as ethanol, butanol, lactic acid, 3 -hydroxy- propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, fumaric acid, malic acid, itaconic acid, an amino acid, 1,3- propane-diol, ethylene, glycerol, a ß-lactam antibiotic or a cephalosporin.

A preferred cell of the invention is a cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. A cell of the invention preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than about 5, about 4, about 3, or about 2.5) and towards organic acids like lactic acid, acetic acid or formic acid and/or sugar degradation products such as furfural and hydroxy- methylfurfural and/or a high tolerance to elevated temperatures.

Any of the above characteristics or activities of a cell of the invention may be naturally present in the cell or may be introduced or modified by genetic modification.

A cell of the invention may be a cell suitable for the production of ethanol. A cell of the invention may, however, be suitable for the production of fermentation products other than ethanol. Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible by a eukaryotic microorganism such as a yeast or a filamentous fungus.

Such fermentation products may be, for example, butanol, lactic acid, 3 -hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, adipic acid, citric acid, malic acid, fumaric acid, itaconic acid, an amino acid, 1,3-propane-diol,

ethylene, glycerol, a ß-lactam antibiotic or a cephalosporin. A preferred cell of the invention for production of non-ethanolic fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

In a further aspect the invention relates to fermentation processes in which the cells of the invention are used for the fermentation of a carbon source comprising a source of pentose, such as xylose. In addition to a source of xylose the carbon source in the fermentation medium may also comprise a source of glucose. The source of xylose or glucose may be xylose or glucose as such or may be any carbohydrate oligo- or polymer comprising xylose or glucose units, such as e.g. lignocellulose, xylans, cellulose, starch and the like. For release of xylose or glucose units from such carbohydrates, appropriate carbohydrases (such as xylanases, glucanases, amylases and the like) may be added to the fermentation medium or may be produced by the cell. In the latter case the cell may be genetically engineered to produce and excrete such carbohydrases. An additional advantage of using oligo- or polymeric sources of glucose is that it enables to maintain a low(er) concentration of free glucose during the fermentation, e.g. by using rate- limiting amounts of the carbohydrases. This, in turn, will prevent repression of systems required for metabolism and transport of non-glucose sugars such as xylose.

In a preferred process the cell ferments both the xylose and glucose, preferably simultaneously in which case preferably a cell is used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source of xylose (and glucose) as carbon source, the fermentation medium will further comprise the appropriate ingredient required for growth of the cell. Compositions of fermentation media for growth of microorganisms such as yeasts are well known in the art. The fermentation process is a process for the production of a fermentation product such as e.g. ethanol, butanol, lactic acid, 3 -hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, adipic acid, citric acid, malic acid, fumaric acid, itaconic acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β-lactam antibiotic, such as Penicillin G or Penicillin V and fermentative derivatives thereof, and a cephalosporin.

The fermentation process may be an aerobic or an anaerobic fermentation process. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is

consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors. In the absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidised by oxidative phosphorylation. To solve this problem many microorganisms use pyruvate or one of its derivatives as an electron and hydrogen acceptor thereby regenerating NAD^+.

Thus, in a preferred anaerobic fermentation process pyruvate is used as an electron (and hydrogen acceptor) and is reduced to fermentation products such as ethanol, butanol, lactic acid, 3 -hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, adipic acid, citric acid, malic acid, fumaric acid, an amino acid, 1,3-propane-diol, ethylene, glycerol, a β-lactam antibiotic and a cephalosporin.

The fermentation process is preferably run at a temperature that is optimal for the cell. Thus, for most yeasts or fungal host cells, the fermentation process is performed at a temperature which is less than about 42°C, preferably less than about 38°C. For yeast or filamentous fungal host cells, the fermentation process is preferably performed at a temperature which is lower than about 35, about 33, about 30 or about 28°C and at a temperature which is higher than about 20, about 22, or about 25°C.

The ethanol yield on xylose and/or glucose in the process preferably is at least about 50, about 60, about 70, about 80, about 90, about 95 or about 98%. The ethanol yield is herein defined as a percentage of the theoretical maximum yield.

The invention also relates to a process for producing a fermentation product.,

The fermentation processes may be carried out in batch, fed-batch or continuous mode. A separate hydrolysis and fermentation (SHF) process or a simultaneous saccharification and fermentation (SSF) process may also be applied. A combination of these fermentation process modes may also be possible for optimal productivity.

The fermentation process according to the present invention may be run under aerobic and anaerobic conditions. Preferably, the process is carried out under micro-aerophilic or oxygen limited conditions.

An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is

consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, and wherein organic molecules serve as both electron donor and electron acceptors.

An oxygen-limited fermentation process is a process in which the oxygen consumption is limited by the oxygen transfer from the gas to the liquid. The degree of oxygen limitation is determined by the amount and composition of the ingoing gasflow as well as the actual mixing/mass transfer properties of the fermentation equipment used. Preferably, in a process under oxygen-limited conditions, the rate of oxygen consumption is at least about 5.5, more preferably at least about 6, such as at least 7 mmol/L/h. A process of the invention comprises recovery of the fermentation product.

The fermentation product of the invention may be any useful product. In one embodiment, it is a product selected from the group consisting of ethanol, n-butanol, isobutanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, adipic acid, fumaric acid, malic acid, itaconic acid, maleic acid, citric acid, adipic acid, an amino acid, such as lysine, methionine, tryptophan, threonine, and aspartic acid, 1,3-propane-diol, ethylene, glycerol, a β-lactam antibiotic and a cephalosporin, vitamins, pharmaceuticals, animal feed supplements, specialty chemicals, chemical feedstocks, plastics, solvents, fuels, including biofuels and biogas or organic polymers, and an industrial enzyme, such as a protease, a cellulase, an amylase, a glucanase, a lactase, a lipase, a lyase, an oxidoreductases, a transferase or a xylanase.

For the recovery of the fermentation product existing technologies are used. For different fermentation products different recovery processes are appropriate. Existing methods of recovering ethanol from aqueous mixtures commonly use fractionation and adsorption techniques. For example, a beer still can be used to process a fermented product, which contains ethanol in an aqueous mixture, to produce an enriched ethanol-containing mixture that is then subjected to fractionation (e.g., fractional distillation or other like techniques). Next, the fractions containing the highest concentrations of ethanol can be passed through an adsorber to remove most, if not all, of the remaining water from the ethanol.

In this document and in its claims, the verb "to comprise" and its conjugations is used in its non-limiting sense to mean that items following the word are included, but items not specifically mentioned are not excluded. In addition, reference to an element by the indefinite article "a" or "an" does not exclude the possibility that more than one of the element is present, unless the context clearly

requires that there be one and only one of the elements. The indefinite article "a" or "an" thus usually means "at least one".

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

EXAMPLES

General molecular biology techniques

Unless indicated otherwise, the methods used are standard biochemical techniques. Examples of suitable general methodology textbooks include Sambrook *et al.*, Molecular Cloning, a Laboratory Manual (1989) and Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

Growth experiments were performed using either Verduyn medium (Verduyn, 1992) or YEPh-medium (10 g/l yeast extract, 20 g/l phytone), supplemented with sugars as indicated in the examples. For solid YEPh medium, 20 g/l agar was added to the liquid medium prior to sterilization.

Transformation of yeast cells

Yeast transformation was done according to the method described by Chen et al (Current Genetics (1992), Volume 21, Number 1, 83-84).

Transformation of yeast cells with linear DNA fragments for integration

In cases that a high efficiency of transformation was required, like transformation with linear DNA fragments for integration into the yeast genome, the method described by Gietz et al (Methods in Enzymology 350: 87-96) was performed.

Example 1

Deletion of *gpd1* and/or *gpd2* in BIE272

In order to delete the *GDP1* gene in the yeast strain BIE272 (European Patent Application EP11163579.3), PCR fragments were generated with which the deletion of the *GPD1* gene could be established. To this end, PCR reaction 1 was performed using primers of SEQ ID NO 1 and SEQ ID NO 2 on plasmid pSUC227 as a template in order to generate the first gene deletion fragment. In PCR reaction 2, primers of SEQ ID NO 3 and SEQ ID NO 4 were used, with plasmid pSUC225 as a template, generating the second deletion fragment.

The physical maps of plasmids pSUC227 and pSUC225 are given in figures 1 and 2.

Plasmid pSUC227 contains the following relevant elements: a lox66 mutant loxP-site (Zhang et al, 2002), followed by a kanMX marker (consisting of a *TEF1*-promoter, a KanMX resistance gene and the *TEF1* terminator from pUG7 (Qian et al (2009))), and a partial cre-recombinase gene (consisting of the *GAL1* promoter and a first 533 nucleotides of the *cre*-gene from plasmid pSH47, Güldener et al (1996)).

Plasmid pSUC225 contains the following relevant elements: a partial cre-recombinase gene (the final 597 nucleotides of the *cre*-gene from plasmid pSH47 followed by the *CYC1*-terminator, Güldener et al (1996)) and a lox71 recombination site (a mutant loxP site, Zhang et al (2002)).

The sequence of the primer of SEQ ID NO 1 contains a tail of 60 nucleotides on the 5'-end, identical to a region of 60 nucleotides in front of the ATG (startcodon) of the *GPD1*-gene, followed by 20 nucleotides corresponding to the lox66 site in pSUC227.

The sequence of the primer of SEQ ID NO 2 is identical to the last part of the *cre* gene fragment in pSUC227.

The sequence of the primer of SEQ ID NO 3 is identical to the first part of the *cre* gene fragment in pSUC225.

The sequence of the primer of SEQ ID NO 4 contains a tail of 60 nucleotides on the 5'-end, identical to a region at the 3'-end of the *GPD1*-gene and *GPD1*-terminator, followed by 21 nucleotides corresponding to the lox71 site in pSUC225.

The mechanism of gene replacement is schematically shown in figure 3. PCR fragments 1 and 2 were used in a transformation experiment. Transformants were first selected on

YEPhD plates containing G418 for selection of the correct integration of both fragments into the genome of strain BIE272, resulting in the replacement of the coding region of the *GPD1*-gene by the KanMX/*cre* cassette.

Subsequently, transformants were transferred to YEPh plates containing 2% galactose in order to induce the expression of the *cre* gene. As a consequence, genomic recombination between the loxP-sites is catalyzed by the cre-recombinase, resulting in the excision of the gene encoding cre-recombinase and the KanMX marker conferring resistance to G418. This is schematically shown in figure 4. Transformants which have lost the KanMX marker are sensitive to G418, which was tested by replica plating.

The *GPD1* deletion was verified by PCR using the primers Gpd-5'-F and Gpd-3'-R (for *GPD1* the sequence of these primers are SEQ ID NO 5 and SEQ ID 6).

BIE272 containing the *GPD1* deletion was called BIE299.

Primer SEQ ID NO 5 is identical to a sequence upstream of the ATG of the *GPD1* gene and also in front of the sequence used for integration of the deletion fragment.

Primer SEQ ID NO 6 is identical to a sequence downstream of the *GPD1* stopcodon and the sequence used for integration of the deletion fragment.

The same strategy was used for making of the *GPD2* deletion in BIE272. Two comparable PCR fragments were generated. PCR reaction 3 was performed using primers of SEQ ID NO 7 and SEQ ID NO 2 in a PCR with pSUC227 as template. PCR reaction 4 was performed using the primers of SEQ ID NO 3 and SEQ ID NO 8 in a PCR with template pSUC225.

The sequence of the primer of SEQ ID NO 7 contains a tail of 60 nucleotides on the 5'-end, identical to a region of 60 nucleotides in front of the ATG (startcodon) of the *GPD2*-gene, followed by 20 nucleotides corresponding to the lox66 site in pSUC227.

The sequence of the primer of SEQ ID NO 8 contains a tail of 60 nucleotides on the 5'-end, identical to a region at the 3'-end of the *GPD2*-gene and *GPD2*-terminator, followed by 21 nucleotides corresponding to the lox71 site in pSUC225.

Both *GPD2* replacement fragments, the result of PCR reactions 3 and 4, were used to transform to BIE272 and transformants were selected on YEPhD plates containing G418 for selection of the correct recombinants. Subsequently these transformants were transferred to YEPh plates containing 2% galactose in order to induce the expression of

the *cre* gene. As a consequence, genomic recombination between the loxP-sites was catalyzed by the cre-recombinase, resulting in the excision of the gene encoding cre-recombinase and the KanMX marker conferring resistance to G418. This is schematically shown in figure 4. Transformants which have lost the KanMX marker are sensitive to G418.

Correctness of the *GPD2* gene deletion was verified by PCR with primers SEQ ID NO 9 and SEQ ID NO 10.

BIE272 containing the *GPD2* deletion was called BIE300.

Primer SEQ ID NO 9 is identical to a sequence upstream of the ATG of the *GPD2* gene and also in front of the sequence used for integration of the deletion fragment.

Primer SEQ ID NO 10 is identical to a sequence downstream of the *GPD2* stopcodon and the sequence used for integration of the deletion fragment.

For construction of the *gpd1 gpd2* double deletion mutant of BIE272 the same procedure was followed, starting with a single deletion mutant of BIE272 (BIE299 or BIE300) and using the deletion fragments for the other gene. Both lines of experiments were followed. Replacement and subsequent recombination was immediately successful when the *GPD1* gene was deleted in strain BIE300. This double deletion strain is named BIE301.

Example 2

Construction of an aldehyde oxidoreductase overexpression plasmid

A synthetic codon-pair optimized *eutE* gene from *Listeria innocula* was designed based on the primary amino acid sequence and synthesized by GeneArt (Regensburg, Germany). Codon-pair optimization was performed as previously disclosed in WO2006/077258 and/or WO2008/000632. WO2008/000632 addresses codon-pair optimization. The nucleotide sequence is included in here as SEQ ID NO 11, the corresponding amino acid sequence as SEQ ID NO 12.

The open reading frame was cloned into the yeast shuttle vector YEplac112 (Gietz et al, 1988) between the *TDH3* promoter and *ADH1* terminator (pBOL033; figure 5) by GeneArt. The plasmid was called pBOL066.

pBOL066 does not contain dominant antibiotic resistance marker genes. Therefore, for transformation of the prototrophic strains BIE272, BIE299, BIE300 and BIE301, the plasmid pBOL066 was adapted. A KanMX selection marker gene was inserted into the PmlI site of the vector backbone by use of the yeast own gap repair mechanism (Raymond et al, 1999), as schematically shown in figure 6. A KanMX PCR product was generated using primers SEQ ID NO 13 and SEQ ID NO 14 on plasmid p427-TEF (DualSystems Biotech, Schlieren, Switzerland). The linearized plasmid fragment and the KanMX PCR fragment were used to transform CEN.PK113-7D (Dr. P. Kötter, Frankfurt, Germany) in a 1:1 molar ratio. Transformants were selected on YEPhD plates containing G418 as selection. Plasmids were isolated from the transformants by using the Zymoprep™ Yeast Plasmid Miniprep I (Zymo Research, Irvine, CA, USA). 1 µl of the plasmid isolate was used to transform electrocompetent DH10b *E.*coli cells (Invitrogen, Carlsbad, CA, USA). Miniprep DNA isolations were performed on several colonies. The obtained plasmids were checked by restriction enzyme analyses. The correct KanMX containing pBOL066 was called pPWT251.

The empty plasmid pBOL033 was adapted the same way and the KanMX containing version was called pPWT242.

Primer SEQ ID NO 13 contains 50 nucleotides identical to the upstream sequence of the PmlI site of pBOL033 followed by 21 nucleotides identical to the KanMX promoter region of p427-TEF.

Primer SEQ ID NO 14 consists of 50 nucleotides identical to the downstream sequence of the PmlI site of pBOL033 and 22 nucleotides identical to the KanMX terminator region of p427-TEF.

Example 3

Construction of an *E. coli mhpF* containing plasmid

Construct pUDE43 contains the *E. coli mhpF* gene, encoding the acetylating NAD-dependent acetaldehyde dehydrogenase, has been described by Medina et al (2010). For the transformation of the prototrophic strains BIE272, BIE299, BIE300 and BIE301 this plasmid needs to be adapted. Therefore the *URA3* gene of the plasmid was replaced by the KanMX selection marker by using the yeast own gap repair mechanism (Raymond et al, 1999), as shown in figure 6.

A KanMX marker fragment was generated by a PCR with primers SEQ ID NO 15 and SEQ ID NO 16 and p427-TEF as template. pUDE43 was linearized with StuI. Both

fragments were used to transform yeast strain CEN.PK113-7D in a 1:1 molar ratio. Transformants were selected on YEPhD plates containing G418 as selection. Plasmids were isolated from the transformants by using the Zymoprep™ Yeast Plasmid Miniprep I (Zymo Research, Irvine, CA, USA). 1 µl of the plasmid isolate was used to transform electrocompetent DH10b *E.*coli cells (Invitrogen, Carlsbad, CA, USA). Miniprep DNA isolations were performed on several colonies. The obtained plasmids were checked by restriction enzyme analyses. The KanMX marker containing pUDE43 was called pPWT240.

The empty vector p426GPD (Mumberg et al (1995)) was adapted the same way and the KanMX containing version was called pPWT239.

Primer SEQ ID NO15 exists of 50 nucleotides identical to the promoter region of the URA3 gene in p426GPD, followed by 21 nucleotides identical to the KanMX promoter region of p427-TEF.

Primer SEQ ID NO16 exists of 50 nucleotides identical to the terminator region of the URA3 gene in p426GPD, followed by 22 nucleotides identical to the KanMX terminator region of p427-TEF.

Example 4

Transformation of yeast strains with overexpression plasmids

Plasmids pPWT239, pPWT240, pPWT242 and pPWT251 were used to transform yeast strains BIE272, BIE299, BIE300 and BIE301 with the method described by Chen et al (1992). Transformants were selected on YEPhD plates containing G418 as selection. Two independent colonies of each transformation were randomly picked for growth tests.

Example 5

Growth experiments

Two independent clones of each transformation (Example 4) were randomly picked. The colonies were named as follows: parent strain – plasmid – number; e.g. BIE272-pPWT240-1 was the first selected clone of the transformed BIE272 with plasmid pPWT240.

All selected clones were pregrown in Verduyn medium supplemented with 2% glucose. Subsequently, the growth was assessed in the Biolector (DasGip, 52428 Jülich, Germany). To this end, 1 ml of CFMM2M medium (Verduyn medium, containing 50 g/l of

glucose, 50 g/l xylose, 35 g/l arabinose, 10 g/l galactose, 5 g/l mannose, 3 g/l acetic acid, 0.03 g/l coumaric acid, 0.2 g/l ferulic acid, 0.1 g/l furfural, 0.1 g/l HMF and 0.1 g/l formic acid) was inoculated with 50 µl of the preculture. The parental strains (BIE272, BIE299, BIE300 and BIE301) were also included, in duplicate. To all cultures, except the parental strains, G418 was added.

The growth of the culture was recorded for 72 hours. At the end of the experiment, the Biolector cultures were collected and spun down in an Eppendorf centrifuge. The supernatant was analyzed by NMR for residual sugars, acetic acid, ethanol and glycerol.

The *gpd1* and/or *gpd2* strains (i.e. BIE299-pPWT251, BIE300-pPWT251 and BIE301-pPWT251), expressing the *eutE* gene from *L. innocula* exhibited a higher acetate consumption and a lower glycerol production than the control strain BIE272-pPWT251. Hence, the ethanol yield on sugar is improved if the genes *GPD1* and/or *GPD2* have been deleted.

The strains expressing the *eutE* gene from *L. innocula*, exhibit a higher acetate consumption and a lower glycerol production than the respective reference strains expressing the *mhpF*-gene from *E. coli*. Hence, the ethanol yield on sugar is improved if the *eutE* gene from *L. innocula* is used as compared to the *mhpF*-gene from *E. coli.*

Example 6

AFM experiments

The AFM (Alcohol Fermentation Monitor, Halotec, Veenendaal) is a laboratory device for monitoring metabolic activity of six alcoholic fermentations simultaneously. For each individual channel, the amount of carbon dioxide that is produced is accurately measured. Temperature and stirrer speed can be set for each channel individually.

BIE301 cells, transformed with the plasmids pPWT240 (expressing the *mhpF*-gene from *E. coli*) and pPWT252 (expressing the *Lin1129*-gene from *Listeria innocua*) were inoculated from fresh YEPhD agarplates containing 300 µg G418 per ml, into 100 ml Verduyn containing 2% glucose and 50 µg G418 per ml. The cultures were incubated overnight at 30°C and 280 rpm. The optical density of the culture at 600 nm was determined. The amount of cells needed to inoculate 200 ml culture at an optical density

of 5.0 was calculated. The necessary amount of cells was collected by centrifugation, washed with Physiological Salt (0.9% NaCl) and resuspended in 200 ml of CFMM1M medium (Verduyn medium, containing Tween80 and ergosterol to final concentrations of 420 mg/l and 10 mg/l respectively, 50 g/l of glucose, 50 g/l xylose, 35 g/l arabinose, 10 g/l galactose and 5 g/l mannose), supplemented with 2 grams acetic acid (HAc) per ml. The AFM was started according to the manual of the supplier. The temperature was 33°C and the rotation speed 250 rpm. CO_2 emission was followed in time.

As is clear from figure 7, the CO_2 emission rate was higher and faster in time in case of BIE301 transformed with pPWT252 compared to the reference strain, BIE301 transformed with pPWT240. Under these conditions, BIE301-pPWT240 starts to produce CO_2 after approximately 80 hours, while the strain BIE301-pPWT252 is capable of doing so directly after inoculation. This indicates that the protein which is encoded by the *Lin1129* gene from *Listeria innocua*, in concert with the yeast enzymes acetylCoA synthase and alcoholdehydrogenase, is better capable of catalyzing the conversion of acetate (acetic acid) into ethanol than the protein encoded by the *mhpF* gene from *E.coli*. Since the expression of the endogenous acetylCoA synthase and alcoholdehydrogenase were not modified in these strains, the difference in performance in this experiment can be assigned to the performance of the corresponding acetylating acetaldehyde dehydrogenase (i.e. mhpF versus Lin1129).

In addition, it indicates that NAD^+ regeneration, necessary for growth, since this is not possible via the production of glycerol, as the genes *GPD1* and *GPD2* were deleted, is more efficient in case of BIE301-pPWT252 compared to BIE301-pPWT240.

The study by Guadelupe Medina (2010) provided a proof of principle for the potential of the metabolic engineering strategy to improve ethanol yields, eliminate glycerol production, and partially convert acetate (a well-known inhibitor of yeast performance in lignocellulosic hydrolysates) into ethanol, by overexpressing the *mhpF* gene from *Escherichia coli* in a *Saccharomyces cerevisiae* strain in which the two genes encoding NAD-dependent glycerol-3-phosphate dehydrogenase (*GPD1* and *GPD2*) were deleted. In the paper it was mentioned that certain properties like the kinetic aspects of acetate reduction needed to be addressed in follow-up research. We now have found new enzymes that allow a more efficient complementation of the inability of *gpd1 gpd2* cells to reoxidize NADH, leading to a faster growth phenotype under relevant conditions.

References

Gietz et al (1988) Gene, 14, 527-534

Güldener et al (1996) Nucleic Acids Research, Vol. 24, no.13, 2519–2524

Medina et al (2010) Applied and Environmental Microbiology, Vol. 76, no. 1, 190-195

Mumberg et al (1995) Gene,156, 119-122

Qian et al (2009) Journal of Microbiological Methods 79 (2009) 253–259

Raymond et al (1999) BioTechniques 26:134-141

Zhang et al (2002) Nucl. Acids Res. (2002) 30 (17): e90. doi: 10.1093/nar/gnf089

CLAIMS

1. Yeast cell comprising:

a) a disruption of deletion or one or more gene native to the yeast cell coding for an enzyme having glycerolphosphate dehydrogenase (GPD) activity;

b) an exogenous polynucleotide coding for an enzyme having aldehyde oxidoreductase activity, wherein the polynucleotide comprises a nucleotide sequence coding for an amino acid sequence with at least 50% amino acid sequence identity with SEQ ID NO: 12.

2. Yeast cell according to claim 1, wherein the polynucleotide comprises a nucleotide sequence coding for an amino acid sequence with at least 90% amino acid sequence identity with SEQ ID NO: 12.

3. Yeast cell according to claims 1 or 2, wherein the yeast cell has a disruption or deletion of one or more GPD1 gene and/or one ore more GPD2 gene native in yeast.

4. Yeast cell according to claim 3, wherein the yeast cell has a disruption or deletion of one or more GPD1 gene and one or more GPD2 gene native in yeast.

5. Yeast cell according to any of claims 1 to 4, wherein the cell comprises at least one of:

i) an exogenous xylose isomerase gene, which gene confers to the cell the ability to isomerise xylose into xylulose; and,

ii) an exogenous genes coding for a L-arabinose isomerase, a L-ribulokinase and a L-ribulose-5-phosphate 4-epimerase, which genes together confer to the cell the ability to convert L-arabinose into D-xylulose 5-phosphate.

6. Yeast cell according to any one of claims 1 to 5, wherein at least one of:

a) the genetic modification that reduces specific glycerolphosphate dehydrogenase activity in the cell comprises the inactivation of at least one endogenous copy of a gene encoding a glycerolphosphate dehydrogenase in the cell's genome;

b) the genetic modification that increases the specific activity of the NAD^+-linked glycerol dehydrogenase is expression of a heterologous gene encoding an NAD^+-linked glycerol dehydrogenase;

c) the genetic modification that increases transport of glycerol into the cell is overexpression of a nucleotide sequence encoding at least one of a glycerol uptake protein and a glycerol channel; and,

d) the genetic modification that reduces specific NAD^+-dependent glycerol 3-phosphate dehydrogenase activity in the cell comprises the inactivation of at least one endogenous copy of a gene encoding a glycerol 3-phosphate dehydrogenase in the cell's genome and,

e) the genetic modification that increases the specific acetyl-CoA synthetase activity is overexpression of a nucleotide sequence encoding an acetyl-CoA synthetase; and,

f) the genetic modification that increases the specific activity of dihydroxyacetone kinase is overexpression of a nucleotide sequence encoding a dihydroxyacetone kinase.

7. Yeast cell according to any one of claims 1 to 6, wherein the yeast cell comprises at least one further genetic modification that results in a characteristic selected from the group consisting of:

a) increased xylulose kinase specific activity;

b) increased flux of the pentose phosphate pathway

c) reduced unspecific aldose reductase specific activity

d) increased transport of at least one of xylose and arabinose into the host cell;

e) decreased sensitivity to catabolite repression;

f) increased tolerance to ethanol, osmolarity or organic acids; and,

g) reduced production of by-products.

8. Yeast cell according to any one of claims 1 to 7, wherein the yeast cell is of a genus selected from the group consisting of *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia, Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces,* and *Yarrowia*.

9. Yeast cell according to claim 8, wherein the yeast cell belongs to a species selected from the group consisting of *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis, K. marxianus* and *Schizosaccharomyces pombe*.

10. A process for producing a fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, a butanol and an isoprenoid-derived product, whereby the process comprises the steps of:
a) fermenting a medium with a yeast cell as defined in any one of claim 1 – 9.

11. A process according to claim 11, wherein the fermentation product is ethanol.

12. A process according to claim 10 or 11, wherein the medium contains or is fed with a lignocellulosic hydrolysate.

13. A process according to any one of the claims 10 - 12, wherein the yeast cell ferments under anaerobic conditions.

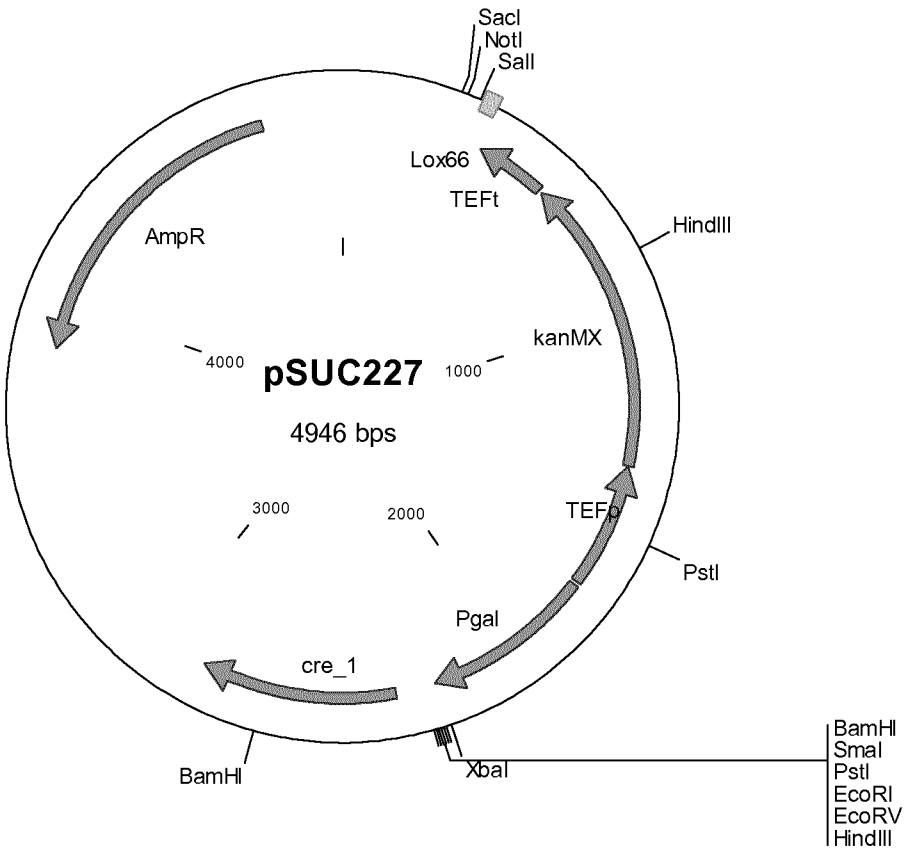
SacI
NotI
SalI
Lox66
TEFt
HindIII
AmpR
kanMX
4000
pSUC227
1000
4946 bps
3000
2000
TEFp
PstI
Pgal
cre_1
BamHI
XbaI
BamHI
SmaI
PstI
EcoRI
EcoRV
HindIII

Fig. 1

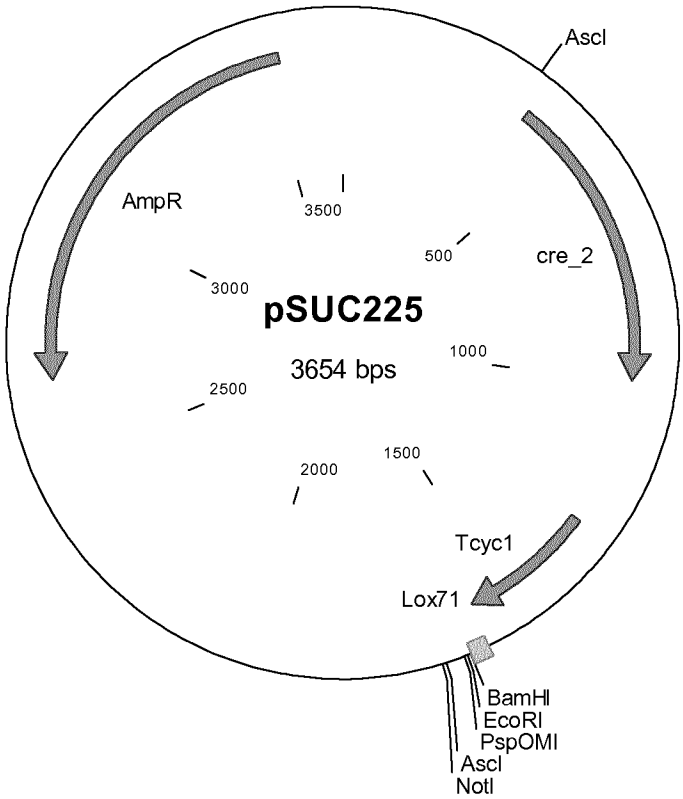
AscI
AmpR
3500
500
cre_2
3000
pSUC225
3654 bps
1000
2500
1500
2000
Tcyc1
Lox71
BamHI
EcoRI
PspOMI
AscI
NotI

Fig. 2

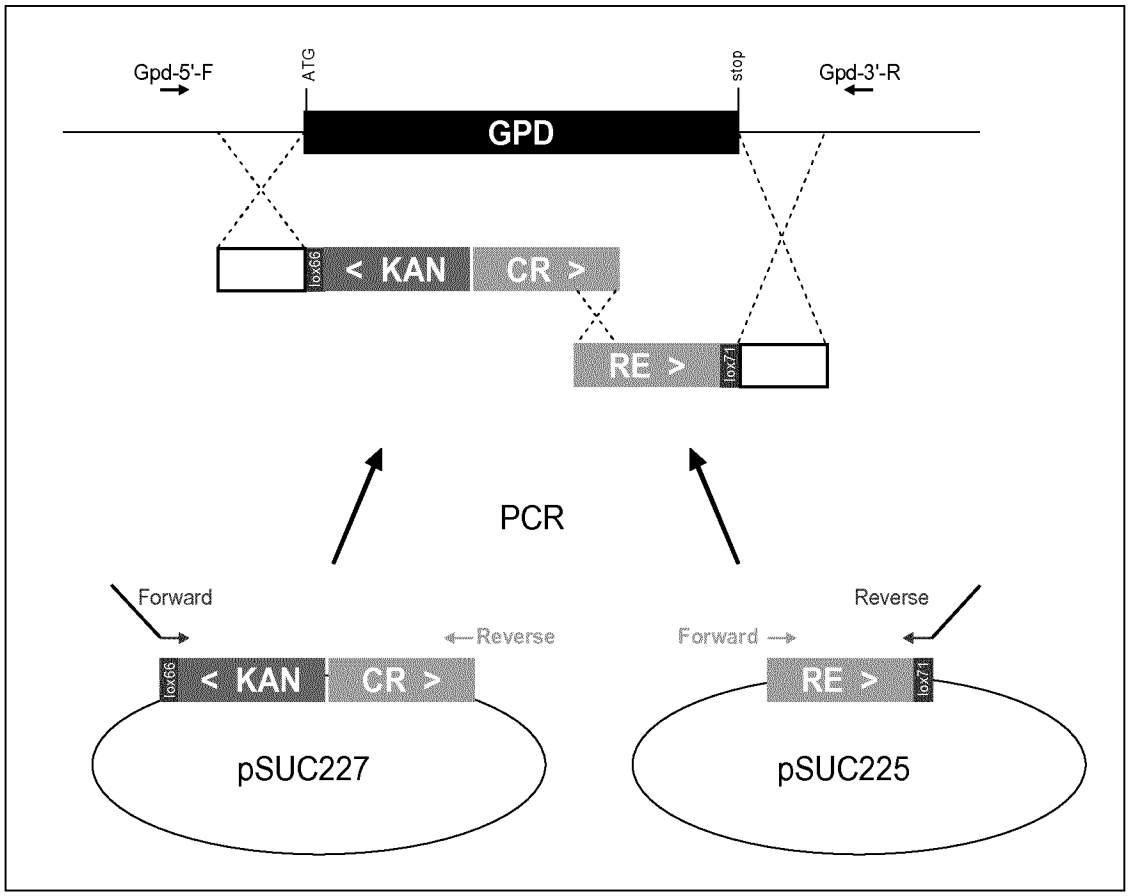
Gpd-5'-F
ATG
stop
Gpd-3'-R
GPD
lox66
< KAN
CR >
RE >
lox71
PCR
Forward
Reverse
Forward
Reverse
lox66
< KAN
CR >
RE >
lox71
pSUC227
pSUC225

Fig. 3

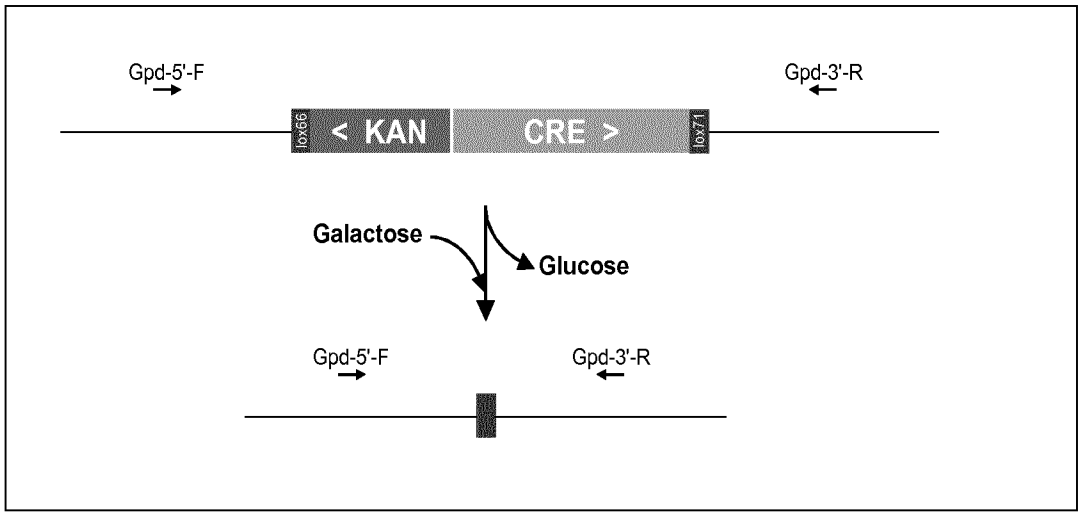
Gpd-5'-F
Gpd-3'-R
lox66
< KAN
CRE >
lox71
Galactose
Glucose
Gpd-5'-F
Gpd-3'-R

Fig. 4

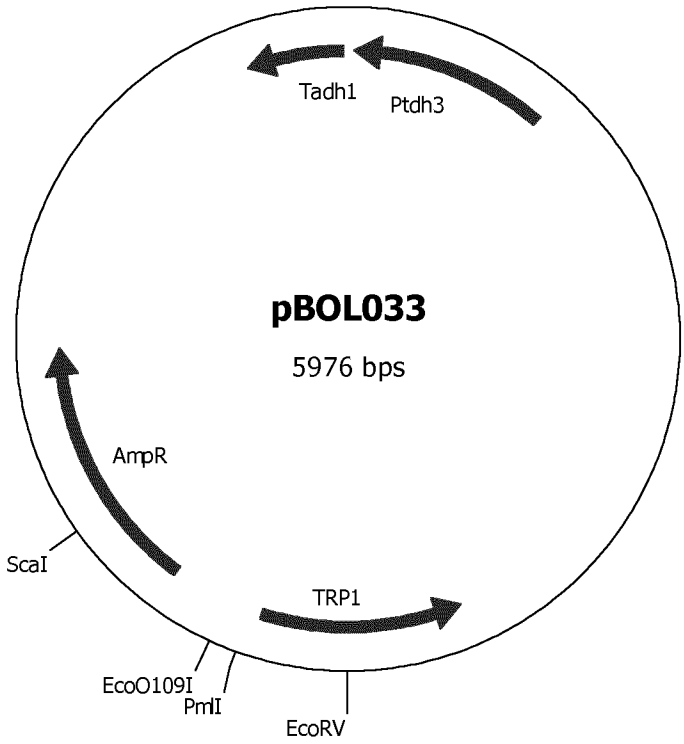
Tadh1
Ptdh3
pBOL033
5976 bps
AmpR
ScaI
TRP1
EcoO109I
PmlI
EcoRV

Fig. 5

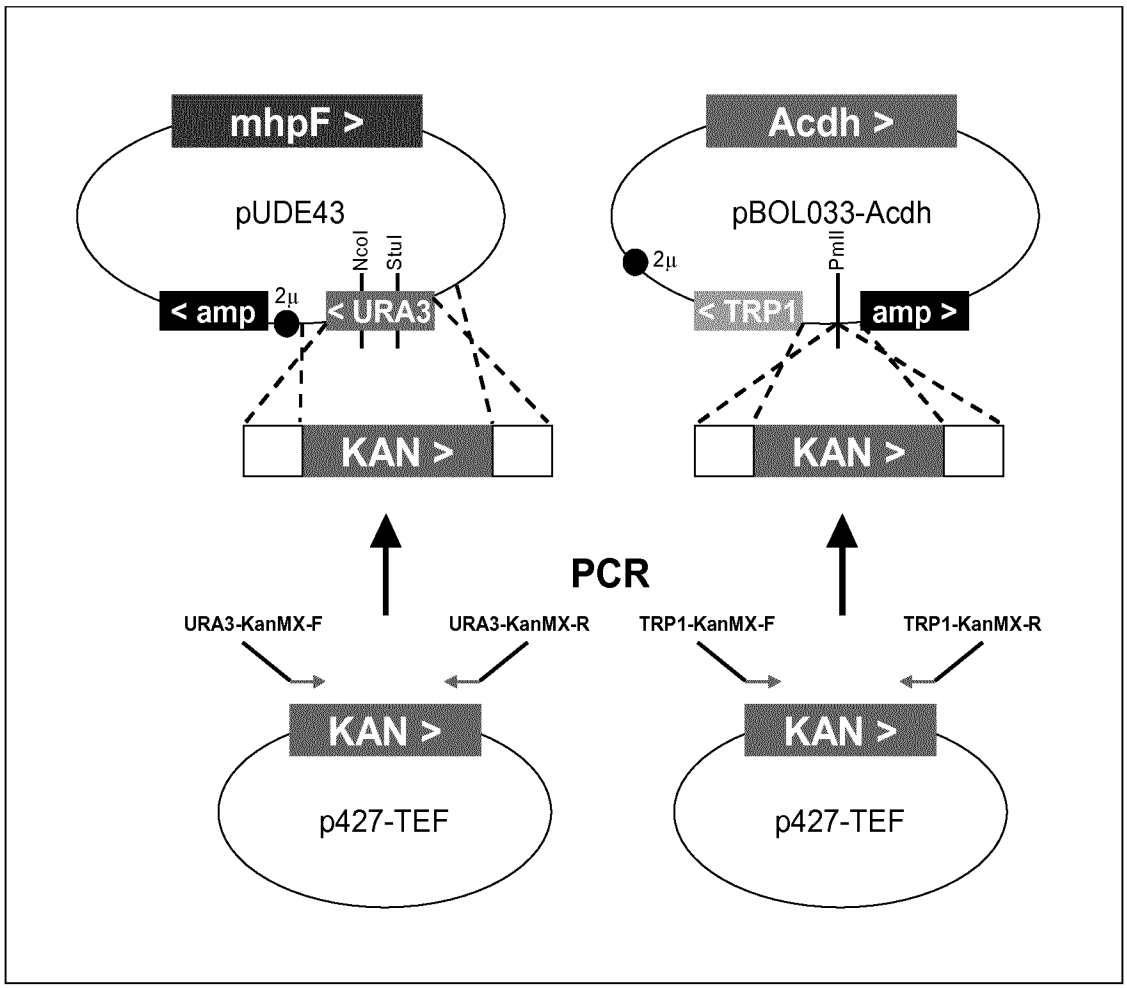
mhpF >
pUDE43
NcoI
StuI
< amp
2µ
< URA3
KAN >
Acdh >
pBOL033-Acdh
2µ
PmlI
< TRP1
amp >
KAN >
PCR
URA3-KanMX-F
URA3-KanMX-R
TRP1-KanMX-F
TRP1-KanMX-R
KAN >
p427-TEF
KAN >
p427-TEF

Fig.6

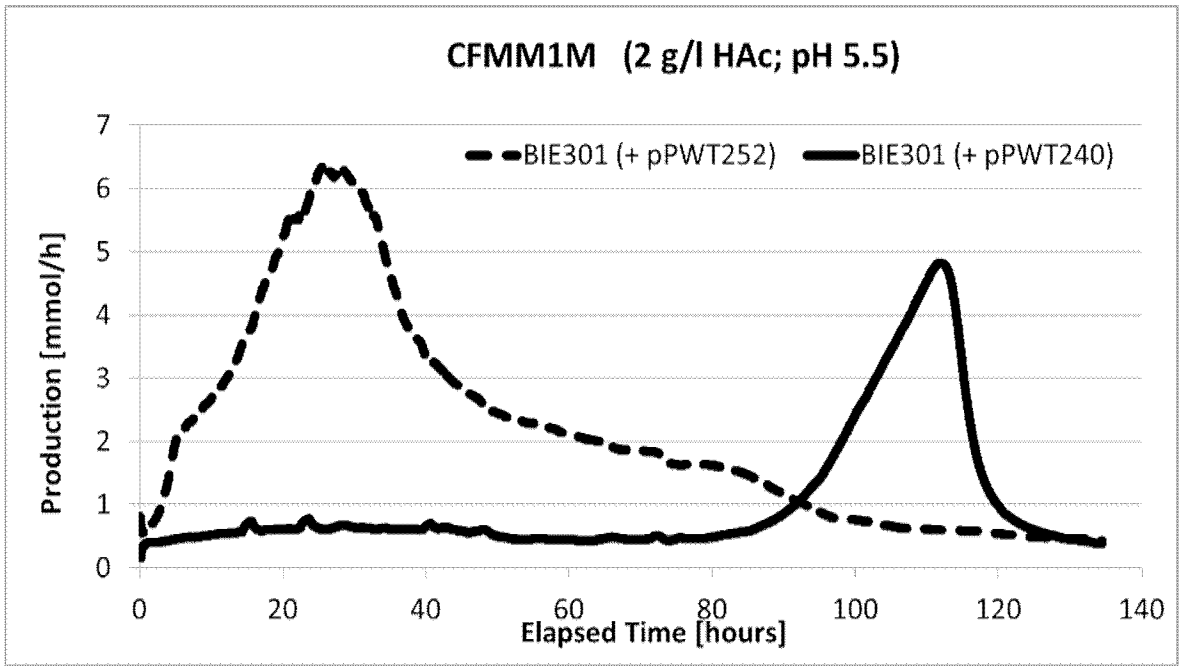
CFMM1M (2 g/l HAc; pH 5.5)
BIE301 (+ pPWT252)
BIE301 (+ pPWT240)
Production [mmol/h]
Elapsed Time [hours]
0
1
2
3
4
5
6
7
0
20
40
60
80
100
120
140

Fig. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2013/067341

A. CLASSIFICATION OF SUBJECT MATTER
INV. C12N1/19 C12N9/02 C12N9/04 C12P7/10
ADD.

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 C12Y 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 practicable, search terms used)
EPO-Internal, Sequence Search, BIOSIS, COMPENDEX, EMBASE, FSTA, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2011/010923 A1 (UNIV DELFT TECH [NL]; PRONK JACOBUS THOMAS [NL]; VAN MARIS ANTONIUS JE) 27 January 2011 (2011-01-27)	1-4,6, 8-13
Y	claims 1-15 page 16, lines 11-20	5,7

Y	WO 2006/009434 A1 (UNIV DELFT TECH [NL]; WINKLER AARON ADRIAAN [NL]; KUYPER SIPKO MAARTEN) 26 January 2006 (2006-01-26) cited in the application the whole document	5,7

A	WO 2009/013159 A2 (DSM IP ASSETS BV [NL]; MUELLER ULRIKE MARIA [DE]; WU LIANG [NL]; RAAMS) 29 January 2009 (2009-01-29) table 1	1-13

	-/--	

[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 application or patent 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
5 November 2013	12/11/2013

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	Lejeune, Robert

1

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MEDINA VICTOR GUADALUPE ET AL: "Elimination of Glycerol Production in Anaerobic Cultures of a Saccharomyces cerevisiae Strain Engineered To Use Acetic Acid as an Electron Acceptor", APPLIED AND ENVIRONMENTAL MICROBIOLOGY, AMERICAN SOCIETY FOR MICROBIOLOGY, US, vol. 76, no. 1, 1 January 2010 (2010-01-01), pages 190-195, XP002603125, ISSN: 0099-2240, DOI: 10.1128/AEM.01772-09 [retrieved on 2009-11-13] cited in the application the whole document -----	1-13
T	WO 2012/143513 A2 (DSM IP ASSETS BV [NL]; KLAASSEN PAUL [NL]; GIELESEN BIANCA ELISABETH M) 26 October 2012 (2012-10-26) * L: source of strain BIE272 used in the examples (software prevents citation as L-document) * -----	1-13

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2013/067341

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011010923 A1	27-01-2011	AU 2010275085 A1	23-02-2012
		CA 2768897 A1	27-01-2011
		CN 102712895 A	03-10-2012
		EA 201270187 A1	28-09-2012
		EP 2277989 A1	26-01-2011
		EP 2456851 A1	30-05-2012
		JP 2013500006 A	07-01-2013
		US 2011275130 A1	10-11-2011
		WO 2011010923 A1	27-01-2011
WO 2006009434 A1	26-01-2006	BR PI0513437 A	06-05-2008
		CA 2573483 A1	26-01-2006
		JP 2008506383 A	06-03-2008
		JP 2013055948 A	28-03-2013
		US 2008261287 A1	23-10-2008
		US 2012225451 A1	06-09-2012
		WO 2006009434 A1	26-01-2006
WO 2009013159 A2	29-01-2009	CN 101939433 A	05-01-2011
		EP 2173881 A2	14-04-2010
		US 2010248233 A1	30-09-2010
		WO 2009013157 A1	29-01-2009
		WO 2009013159 A2	29-01-2009
WO 2012143513 A2	26-10-2012	AU 2012244691 A1	17-10-2013
		WO 2012143513 A2	26-10-2012