



(86) Date de dépôt PCT/PCT Filing Date: 2008/07/21
(87) Date publication PCT/PCT Publication Date: 2009/01/22
(85) Entrée phase nationale/National Entry: 2010/01/19
(86) N° demande PCT/PCT Application No.: NL 2008/050500
(87) N° publication PCT/PCT Publication No.: 2009/011591
(30) Priorités/Priorities: 2007/07/19 (US60/929,951);
2007/07/19 (EP07112791.4)

(51) Cl.Int./Int.Cl. *C12N 1/14* (2006.01),
C12N 1/18 (2006.01), *C12N 15/01* (2006.01),
C12N 9/92 (2006.01), *C12P 7/00* (2006.01)
(71) Demandeur/Applicant:
ROYAL NEDALCO B.V., NL
(72) Inventeur/Inventor:
DE BONT, JOHANNES ADRIANUS MARIA, NL
(74) Agent: BORDEN LADNER GERVAIS LLP

(54) Titre : NOUVELLES CELLULES EUCARYOTES FERMENTANT L'ARABINOSE
(54) Title: NOVEL ARABINOSE-FERMENTING EUKARYOTIC CELLS

(57) **Abrégé/Abstract:**

The present invention relates to eukaryotic cells which have the ability to convert L-arabinose into D-xylulose 5-phosphate. The cells have acquired this ability by transformation with nucleotide sequences coding for an arabinose isomerase, a ribulokinase, and a ribulose-5-P4-epimerase from a bacterium that belongs to a *Clavibacter*, *Arthrobacter* or *Gramella* genus. The cell preferably is a yeast or a filamentous fungus, more preferably a yeast is capable of anaerobic alcoholic fermentation. The may further comprise one or more genetic modifications that increase the flux of the pentose phosphate pathway, reduce unspecific aldose reductase activity, confer to the cell the ability to directly isomerise xylose into xylulose, increase the specific xylulose kinase activity, increase transport of at least one of xylose and arabinose into the host cell, decrease sensitivity to catabolite repression, increase tolerance to ethanol, osmolarity or organic acids; and/or reduce production of by- products. The cell preferably is a cell that has the ability to produce a fermentation product such as ethanol, lactic acid, 3- hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, -lactam antibiotics and cephalosporins. The invention further relates to processes for producing these fermentation products wherein a cell of the invention is used to ferment arabinose into the fermentation products.



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

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
22 January 2009 (22.01.2009)

PCT

(10) International Publication Number
WO 2009/011591 A3

(51) International Patent Classification:

C12N 1/14 (2006.01) *C12N 15/01* (2006.01)
C12N 1/18 (2006.01) *C12P 7/00* (2006.01)
C12N 9/92 (2006.01)

(21) International Application Number:

PCT/NL2008/050500

(22) International Filing Date: 21 July 2008 (21.07.2008)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

07112791.4 19 July 2007 (19.07.2007) EP
 60/929,951 19 July 2007 (19.07.2007) US

(71) Applicant (for all designated States except US): Royal
Nedalco B.V. [NL/NL]; Van Konijnenburgweg 100,
NL-4600 AA Bergen Op Zoom (NL).

(72) Inventor; and

(75) Inventor/Applicant (for US only): DE BONT, Johannes
Adrianus Maria [NL/NL]; Lawickse Allee 68, NL-6707
AK Wageningen (NL).(74) Agent: VAN WESTENBRUGGE, Andries; Nederland-
sch Octrooibureau, Postbus 29720, NL-2502 LS Den Haag
(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, BR, BW, BY, BZ, CA,
CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE,
EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID,
IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK,
LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW,
MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT,
RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM,
ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI,
FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL,
NO, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG,
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

- with international search report
- before the expiration of the time limit for amending the
claims and to be republished in the event of receipt of
amendments
- with sequence listing part of description published sepa-
rately in electronic form and available upon request from
the International Bureau

(88) Date of publication of the international search report:
30 April 2009

(54) Title: NOVEL ARABINOSE-FERMENTING EUKARYOTIC CELLS

(57) Abstract: The present invention relates to eukaryotic cells which have the ability to convert L-arabinose into D-xylulose 5-phosphate. The cells have acquired this ability by transformation with nucleotide sequences coding for an arabinose isomerase, a ribulokinase, and a ribulose-5-P-4-epimerase from a bacterium that belongs to a *Clavibacter*, *Arthrobacter* or *Gramella* genus. The cell preferably is a yeast or a filamentous fungus, more preferably a yeast is capable of anaerobic alcoholic fermentation. The may further comprise one or more genetic modifications that increase the flux of the pentose phosphate pathway, reduce unspecific aldose reductase activity, confer to the cell the ability to directly isomerise xylose into xylulose, increase the specific xylulose kinase activity, increase transport of at least one of xylose and arabinose into the host cell, decrease sensitivity to catabolite repression, increase tolerance to ethanol, osmolarity or organic acids; and/or reduce production of by-products. The cell preferably is a cell that has the ability to produce a fermentation product such as ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, -lactam antibiotics and cephalosporins. The invention further relates to processes for producing these fermentation products wherein a cell of the invention is used to ferment arabinose into the fermentation products.

WO 2009/011591 A3

Novel Arabinose-Fermenting Eukaryotic Cells

Field of the invention

The invention relates to the fields of fermentation technology, molecular biology and biofuel production. In particular the invention relates to an eukaryotic cell having the ability to convert L-arabinose into a fermentation product and to a process for producing a fermentation product wherein this cell is used.

Background of the invention

Economically viable ethanol production from the hemicellulose fraction of plant biomass requires the simultaneous conversion of both pentoses and hexoses at comparable rates and with high yields. Yeasts, in particular *Saccharomyces spp.*, are the most appropriate candidates for this process since they can grow fast on hexoses, both aerobically and anaerobically. Furthermore they are much more resistant to the toxic environment of lignocellulose hydrolysates than (genetically modified) bacteria. Although wild-type *S. cerevisiae* strains rapidly ferment hexoses with high efficiency, they cannot grow on nor use pentoses such as D-xylose and L-arabinose. This inspired various studies to expand the substrate range of *S. cerevisiae*.

EP 1 499 708 discloses the construction of a L-arabinose-fermenting *S. cerevisiae* strain by overexpression of the bacterial L-arabinose pathway. In the bacterial pathway, the enzymes L-arabinose isomerase (*araA*), L-ribulokinase (*araB*), and L-ribulose-5-phosphate 4-epimerase (*araD*) are involved converting L-arabinose to L-ribulose, L-ribulose-5-P, and D-xylulose-5-P, respectively. Using the *Bacillus subtilis* *araA* gene and the *Escherichia coli* *araB*, and *araD* genes, combined with evolutionary engineering, a *S. cerevisiae* strain capable of aerobic growth on L-arabinose was obtained. The evolved strain was reported to have acquired a mutation in the L-ribulokinase gene (*araB*), that resulted in a reduced activity of this enzyme. Enhanced transaldolase (TAL1) activity was also reported to be required for L-arabinose fermentation. Moreover, EP 1 499 708 discloses that overexpression of the gene encoding the *S. cerevisiae* galactose permease (GAL2) - also known to transport arabinose - improved growth on arabinose. However, although the evolved *S. cerevisiae* strain produced ethanol from arabinose at a low

specific production rate of $60 - 80 \text{ mg h}^{-1} (\text{g dry weight})^{-1}$ under oxygen-limited conditions, no anaerobic fermentation of arabinose was observed.

Wisselink et al. (2007, AEM Accepts, published online ahead of print on 1 June 2007; Appl. Environ. Microbiol. doi:10.1128/AEM.00177-07) disclose a *S. cerevisiae* strain obtained by expression of the L-arabinose isomerase (*araA*), L-ribulokinase (*araB*), and L-ribulose-5-phosphate 4-epimerase (*araD*) of the L-arabinose utilization pathway of *Lactobacillus plantarum*, overexpression of *S. cerevisiae* genes encoding the enzymes of the non-oxidative pentose-phosphate pathway, and extensive evolutionary engineering. The resulting *S. cerevisiae* strain exhibits a rate of arabinose consumption of $0.70 \text{ g h}^{-1} (\text{g dry weight})^{-1}$ and a rate of ethanol production of $0.29 \text{ g h}^{-1} (\text{g dry weight})^{-1}$ with an ethanol yield of 0.43 g g^{-1} during anaerobic growth on L-arabinose as sole carbon source.

WO 03/062430 and WO 06/009434 disclose yeast strains able to convert xylose into ethanol. These yeast strains are able to directly isomerise xylose into xylulose. WO 06/096130 discloses yeast strains able to convert xylose and arabinose simultaneously into ethanol.

Description of the invention

Definitions

Arabinose isomerase

The enzyme "arabinose isomerase" (EC 5.3.1.4) is herein defined as an enzyme that catalyses the direct isomerisation of L-arabinose into L-ribulose and vice versa. The enzyme is also known as a L-arabinose ketol-isomerase. Arabinose isomerases of the invention may be further defined by their amino acid sequence as herein described below. Likewise arabinose isomerases may be defined by the nucleotide sequences encoding the enzyme as well as by nucleotide sequences hybridising to a reference (*araA*) nucleotide sequence encoding a arabinose isomerase as herein described below.

L-ribulokinase

The enzyme "L-ribulokinase" (EC 2.7.1.16) is herein defined as an enzyme that catalyses the reaction $\text{ATP} + \text{L-ribulose} = \text{ADP} + \text{L-ribulose 5-phosphate}$. A ribulose kinase of the invention may be further defined by its amino acid sequence as herein described below. Likewise a ribulose kinase may be defined by the nucleotide sequences encoding the enzyme as well as by

nucleotide sequences hybridising to a reference nucleotide sequence (*araB*) encoding a xylulose kinase as herein described below.

L-ribulose-5-phosphate 4-epimerase

The enzyme "L-ribulose-5-phosphate 4-epimerase" (5.1.3.4) is herein defined as an enzyme that catalyses the epimerisation of L-ribulose 5-phosphate into D-xylulose 5-phosphate and vice versa. The enzyme is also known as L-ribulose phosphate 4-epimerase or ribulose phosphate 4-epimerase. A ribulose 5-phosphate 4-epimerase of the invention may be further defined by its amino acid sequence as herein described below. Likewise a ribulose 5-phosphate 4-epimerase may be defined by the nucleotide sequences encoding the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide sequence (*araD*) encoding a ribulose 5-phosphate 4-epimerase as herein described below.

D-ribulose 5-phosphate 3-epimerase

The enzyme "D-ribulose 5-phosphate 3-epimerase" (5.1.3.1) is herein defined as an enzyme that catalyses the epimerisation of D-xylulose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphoribulose epimerase; erythrose-4-phosphate isomerase; phosphoketopentose 3-epimerase; xylulose phosphate 3-epimerase; phosphoketopentose epimerase; ribulose 5-phosphate 3-epimerase; D-ribulose phosphate-3-epimerase; D-ribulose 5-phosphate epimerase; D-ribulose-5-*P* 3-epimerase; D-xylulose-5-phosphate 3-epimerase; pentose-5-phosphate 3-epimerase; or D-ribulose-5-phosphate 3-epimerase.

Ribulose 5-phosphate isomerase

The enzyme "ribulose 5-phosphate isomerase" (EC 5.3.1.6) is herein defined as an enzyme that catalyses direct isomerisation of D-ribose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphopentosisomerase; phosphoriboisomerase; ribose phosphate isomerase; 5-phosphoribose isomerase; D-ribose 5-phosphate isomerase; D-ribose-5-phosphate ketol-isomerase; or D-ribose-5-phosphate aldose-ketose-isomerase.

Transketolase

The enzyme "transketolase" (EC 2.2.1.1) is herein defined as an enzyme that catalyses the reaction: D-ribose 5-phosphate + D-xylulose 5-phosphate into sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate and vice versa. The enzyme is also known as

glycolaldehydetransferase or sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate
glycolaldehydetransferase.

Transaldolase

The enzyme "transaldolase" (EC 2.2.1.2) is herein defined as an enzyme that catalyses the
reaction: sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate into D-erythrose 4-
phosphate + D-fructose 6-phosphate and vice versa. The enzyme is also known as
dihydroxyacetone transferase; dihydroxyacetone synthase; formaldehyde transketolase; or
sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycerone- transferase. A
transaldolase of the invention may be further defined by its amino acid sequence as herein
described below.

Aldose reductase

The enzyme "aldose reductase" (EC 1.1.1.21) is herein defined as any enzyme that is
capable of reducing an aldose to the corresponding alditol and vice versa. In the context of the
present invention an aldose reductase may be any unspecific aldose reductase that is native
(endogenous) to a host cell of the invention and that is capable of reducing aldopentoses such as
arabinose, xylose or xylulose to arabinitol or xylitol, respectively. Unspecific aldose reductases
catalyse the reaction: aldose + NAD(P)H + H⁺ ↔ alditol + NAD(P)⁺. The enzyme has a wide
specificity and is also known as aldose reductase; polyol dehydrogenase (NADP⁺); alditol:NADP
oxidoreductase; alditol:NADP⁺ 1-oxidoreductase; NADPH-aldopentose reductase; or NADPH-
aldose reductase. A particular example of such an unspecific aldose reductase that is endogenous
to *S. cerevisiae* and that is encoded by the *GRE3* gene (Träff et al., 2001, Appl. Environ.
Microbiol. 67: 5668-74).

Xylose isomerase

The enzyme "xylose isomerase" (EC 5.3.1.5) is herein defined as an enzyme that catalyses
the direct isomerisation of D-xylose into D-xylulose and vice versa. The enzyme is also known
as a D-xylose ketoisomerase. Some xylose isomerases are also capable of catalysing the
conversion between D-glucose and D-fructose and are therefore sometimes referred to as glucose
isomerase. Xylose isomerases require bivalent cations like magnesium or manganese as cofactor.
Xylose isomerases of the invention may be further defined by their amino acid sequence as
herein described below. Likewise xylose isomerases may be defined by the nucleotide sequences
encoding the enzyme as well as by nucleotide sequences hybridising to a reference nucleotide

sequence encoding a xylose isomerase as herein described below. A unit (U) of xylose isomerase activity is herein defined as the amount of enzyme producing 1 nmol of xylulose per minute, under conditions as described by Kuyper et al. (2003, FEMS Yeast Res. 4: 69-78).

Xylulose kinase

The enzyme "xylulose kinase" (EC 2.7.1.17) is herein defined as an enzyme that catalyses the reaction $\text{ATP} + \text{D-xylulose} = \text{ADP} + \text{D-xylulose 5-phosphate}$. The enzyme is also known as a phosphorylating xylulokinase, D-xylulokinase or ATP:D-xylulose 5-phosphotransferase.

Sequence identity and similarity

Sequence identity is herein defined as a relationship between two or more amino acid (polypeptide or protein) sequences or two or more nucleic acid (polynucleotide) sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between amino acid or nucleic acid sequences, as the case may be, as determined by the match between strings of such sequences. "Similarity" between two amino acid sequences is determined by comparing the amino acid sequence and its conserved amino acid substitutes of one polypeptide to the sequence of a second polypeptide. "Identity" and "similarity" can be readily calculated by known methods. The terms "substantially identical", "substantial identity" or "essentially similar" or "essential similarity" means that two peptide or two nucleotide sequences, when optimally aligned, such as by the programs GAP or BESTFIT using default parameters, share at least a certain percentage of sequence identity as defined elsewhere herein. GAP uses the Needleman and Wunsch global alignment algorithm to align two sequences over their entire length, maximizing the number of matches and minimizes the number of gaps. Generally, the GAP default parameters are used, with a gap creation penalty = 50 (nucleotides) / 8 (proteins) and gap extension penalty = 3 (nucleotides) / 2 (proteins). For nucleotides the default scoring matrix used is nwsgapdna and for proteins the default scoring matrix is Blosum62 (Henikoff & Henikoff, 1992, PNAS 89, 915-919). It is clear that when RNA sequences are said to be essentially similar or have a certain degree of sequence identity with DNA sequences, thymine (T) in the DNA sequence is considered equal to uracil (U) in the RNA sequence. Sequence alignments and scores for percentage sequence identity may be determined using computer programs, such as the GCG Wisconsin Package, Version 10.3, available from Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 USA or the open-source software Emboss for Windows (current version 2.7.1-07). Alternatively

1 percent similarity or identity may be determined by searching against databases such as FASTA,
2 BLAST, etc.

3 Optionally, in determining the degree of amino acid similarity, the skilled person may also
4 take into account so-called "conservative" amino acid substitutions, as will be clear to the skilled
5 person. Conservative amino acid substitutions refer to the interchangeability of residues having
6 similar side chains. For example, a group of amino acids having aliphatic side chains is glycine,
7 alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side
8 chains is serine and threonine; a group of amino acids having amide-containing side chains is
9 asparagine and glutamine; a group of amino acids having aromatic side chains is phenylalanine,
10 tyrosine, and tryptophan; a group of amino acids having basic side chains is lysine, arginine, and
11 histidine; and a group of amino acids having sulphur-containing side chains is cysteine and
12 methionine. Preferred conservative amino acids substitution groups are: valine-leucine-
13 isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.
14 Substitutional variants of the amino acid sequence disclosed herein are those in which at least
15 one residue in the disclosed sequences has been removed and a different residue inserted in its
16 place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for
17 each of the naturally occurring amino acids are as follows: Ala to ser; Arg to lys; Asn to gln or
18 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
19 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;
20 Thr to ser; Trp to tyr; Tyr to trp or phe; and, Val to ile or leu.

21 Hybridising nucleic acid sequences

22 Nucleotide sequences encoding the enzymes of the invention may also be defined by their
23 capability to hybridise with the nucleotide sequences of SEQ ID NO.'s 10 – 18, respectively,
24 under moderate, or preferably under stringent hybridisation conditions. Stringent hybridisation
25 conditions are herein defined as conditions that allow a nucleic acid sequence of at least about
26 25, preferably about 50 nucleotides, 75 or 100 and most preferably of about 200 or more
27 nucleotides, to hybridise at a temperature of about 65°C in a solution comprising about 1 M salt,
28 preferably 6 x SSC or any other solution having a comparable ionic strength, and washing at
29 65°C in a solution comprising about 0.1 M salt, or less, preferably 0.2 x SSC or any other
30 solution having a comparable ionic strength. Preferably, the hybridisation is performed
31 overnight, i.e. at least for 10 hours and preferably washing is performed for at least one hour with

1 at least two changes of the washing solution. These conditions will usually allow the specific
2 hybridisation of sequences having about 90% or more sequence identity.

3 Moderate conditions are herein defined as conditions that allow a nucleic acid sequences of
4 at least 50 nucleotides, preferably of about 200 or more nucleotides, to hybridise at a temperature
5 of about 45°C in a solution comprising about 1 M salt, preferably 6 x SSC or any other solution
6 having a comparable ionic strength, and washing at room temperature in a solution comprising
7 about 1 M salt, preferably 6 x SSC or any other solution having a comparable ionic strength.
8 Preferably, the hybridisation is performed overnight, i.e. at least for 10 hours, and preferably
9 washing is performed for at least one hour with at least two changes of the washing solution.
10 These conditions will usually allow the specific hybridisation of sequences having up to 50%
11 sequence identity. The person skilled in the art will be able to modify these hybridisation
12 conditions in order to specifically identify sequences varying in identity between 50% and 90%.

13 Operably linked

14 As used herein, the term "operably linked" refers to a linkage of polynucleotide elements in
15 a functional relationship. A nucleic acid is "operably linked" when it is placed into a functional
16 relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably
17 linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked
18 means that the DNA sequences being linked are typically contiguous and, where necessary to
19 join two protein coding regions, contiguous and in reading frame.

20 Promoter

21 As used herein, the term "promoter" refers to a nucleic acid fragment that functions to
22 control the transcription of one or more genes, located upstream with respect to the direction of
23 transcription of the transcription initiation site of the gene, and is structurally identified by the
24 presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites and
25 any other DNA sequences, including, but not limited to transcription factor binding sites,
26 repressor and activator protein binding sites, and any other sequences of nucleotides known to
27 one of skill in the art to act directly or indirectly to regulate the amount of transcription from the
28 promoter. A "constitutive" promoter is a promoter that is active under most environmental and
29 developmental conditions. An "inducible" promoter is a promoter that is active under
30 environmental or developmental regulation.

31 Protein

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

4 Homologous

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

21 Heterologous

22 The term "heterologous" when used with respect to a nucleic acid (DNA or RNA) or
23 protein refers to a nucleic acid or protein that does not occur naturally as part of the organism,
24 cell, genome or DNA or RNA sequence in which it is present, or that is found in a cell or
25 location or locations in the genome or DNA or RNA sequence that differ from that in which it is
26 found in nature. Heterologous nucleic acids or proteins are not endogenous to the cell into which
27 it is introduced, but has been obtained from another cell or synthetically or recombinantly
28 produced. Generally, though not necessarily, such nucleic acids encode proteins that are not
29 normally produced by the cell in which the DNA is transcribed or expressed. Similarly
30 exogenous RNA encodes for proteins not normally expressed in the cell in which the exogenous
31 RNA is present. Heterologous nucleic acids and proteins may also be referred to as foreign

1 nucleic acids or proteins. Any nucleic acid or protein that one of skill in the art would recognize
2 as heterologous or foreign to the cell in which it is expressed is herein encompassed by the term
3 heterologous nucleic acid or protein. The term heterologous also applies to non-natural
4 combinations of nucleic acid or amino acid sequences, i.e. combinations where at least two of the
5 combined sequences are foreign with respect to each other.

6 7 Detailed description of the invention

8 In a first aspect the present invention relates to a eukaryotic cell comprising nucleotide
9 sequences as defined in (a), (b) and (c), whereby the expression of the nucleotide sequences
10 confers to the cell the ability to convert L-arabinose into D-xylulose 5-phosphate. Expressly
11 included in the invention are eukaryotic cells that may already have the ability to convert L-
12 arabinose into D-xylulose 5-phosphate (at a low level) and wherein expression of the nucleotide
13 sequences as defined in (a), (b) and (c) increases the cell's ability to convert L-arabinose into D-
14 xylulose 5-phosphate. Preferably, in the cells of the invention, the ability to convert L-arabinose
15 into D-xylulose 5-phosphate is the ability to convert L-arabinose into D-xylulose 5-phosphate
16 through the subsequent reactions of 1) isomerisation of arabinose into ribulose; 2)
17 phosphorylation of ribulose to ribulose 5-phosphate; and, 3) epimerisation of ribulose 5-
18 phosphate into D-xylulose 5-phosphate. Preferably expression of the nucleotide sequences
19 confers to, or increases in the cell the ability to grow on arabinose as sole carbon and/or energy
20 source, more preferably expression of the nucleotide sequences confers to the cell, or increases in
21 the ability to grow on arabinose as sole carbon and/or energy source through conversion of
22 arabinose into D-xylulose 5-phosphate (and further metabolism of D-xylulose 5-phosphate).

23 The nucleotide sequence (a) preferably is a nucleotide sequence encoding an arabinose
24 isomerase, preferably a L-arabinose isomerase as herein defined above. The nucleotide sequence
25 encoding the arabinose isomerase preferably is selected from the group consisting of:

- 26 (i) a nucleotide sequence encoding an arabinose isomerase comprising an amino acid
27 sequence that has at least 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one
28 of the amino acid sequences of SEQ ID NO's: 1, 2 and 3;
29 (ii) a nucleotide sequence comprising a nucleotide sequence that has at least 70, 80, 90, 95, 98,
30 99 or 100% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 10,
31 11 and 12;

(iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,

(iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code.

The nucleotide sequence (b) preferably is a nucleotide sequence encoding a ribulokinase, preferably a L-ribulokinase as herein defined above. The nucleotide sequence encoding the ribulokinase preferably is selected from the group consisting of:

(i) a nucleotide sequence encoding a ribulokinase comprising an amino acid sequence that has at least 55, 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one of the amino acid sequences of SEQ ID NO's: 4, 5 and 6;

(ii) a nucleotide sequence comprising a nucleotide sequence that has at least 65, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 13, 14 and 15;

(iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,

(iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code.

The nucleotide sequence (c) preferably is a nucleotide sequence encoding a ribulose-5-P-4-epimerase, preferably a L-ribulose-5-P-4-epimerase as herein defined above. The nucleotide sequence encoding the ribulose-5-P-4-epimerase preferably is selected from the group consisting of:

(i) a nucleotide sequence encoding a ribulose-5-P-4-epimerase comprising an amino acid sequence that has at least 55, 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one of the amino acid sequences of SEQ ID NO's: 7, 8 and 9;

(ii) a nucleotide sequence comprising a nucleotide sequence that has at least 65, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 16, 17 and 18;

(iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,

(iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code.

1 A nucleotide sequence encoding an arabinose isomerase comprising an amino acid
2 sequence that has at least 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one
3 of the amino acid sequences of SEQ ID NO's: 1, 2 and 3, preferably encodes an amino acid
4 sequence wherein active site residues, and/or residues involved in metal ion- and/or substrate-
5 binding are conserved. Such residues may be derived by comparison of the amino acid sequences
6 of SEQ ID NO's: 1, 2 and 3 with the crystal structure of the E. coli L-arabinose isomerase
7 (Manjasetty and Chance, 2006, J Mol Biol. 360(2):297-309). In addition more than 166 amino
8 acid sequences of arabinose isomerases are known in the art. Sequence alignments of SEQ ID
9 NO's: 1, 2 and 3 with these known arabinose isomerase amino acid sequences will indicate
10 conserved regions and amino acid positions, the conservation of which are important for
11 structure and enzymatic activity. These regions and positions will tolerate no or only
12 conservative amino acid substitutions. Amino acid substitutions outside of these regions and
13 positions are unlikely to greatly affect arabinose isomerase activity.

14 A nucleotide sequence encoding an L-ribulokinase comprising an amino acid sequence that
15 has at least 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one of the amino
16 acid sequences of SEQ ID NO's: 4, 5 and 6, preferably encodes an amino acid sequence wherein
17 active site residues, and/or residues involved in substrate-binding are conserved. Such residues
18 may be derived by comparison of the amino acid sequences of SEQ ID NO's: 4, 5 and 6 with the
19 crystal structure of the E. coli L-ribulokinase (Lee and Bendet, 1967, Biol Chem. 242(9):2043-
20 50; Lee et al., 1970, J Biol Chem. 245(6):1357-61). In addition more than 5000 amino acid
21 sequences of ribulokinases are known in the art. Sequence alignments of SEQ ID NO's: 4, 5 and
22 6 with these known ribulokinase amino acid sequences will indicate conserved regions and
23 amino acid positions, the conservation of which are important for structure and enzymatic
24 activity. These regions and positions will tolerate no or only conservative amino acid
25 substitutions. Amino acid substitutions outside of these regions and positions are unlikely to
26 greatly affect ribulokinase activity.

27 A nucleotide sequence encoding a ribulose-5-P-4-epimerase comprising an amino acid
28 sequence that has at least 60, 70, 80, 90, 95, 98, 99 or 100% sequence identity with at least one
29 of the amino acid sequences of SEQ ID NO's: 7, 8 and 9, preferably encodes an amino acid
30 sequence wherein active site residues, residues involved in metal ion- and substrate-binding
31 and/or residues involved in intersubunit interface are conserved. Such residues may be derived

1 by comparison of the amino acid sequences of SEQ ID NO's: 7, 8 and 9 with the crystal
2 structure of the E. coli ribulose-5-P-4-epimerase (Luo et al., 2001, Biochemistry. 40(49):14763-
3 71) and comparisons with the structurally related aldolases (Kroemer and Schulz, 2002, Acta
4 Crystallogr D Biol Crystallogr. 58(Pt 5):824-32; Joerger et al., 2000, Biochemistry. 39(20):6033-
5 41). In addition more than 600 amino acid sequences of ribulose-5-P-4-epimerases and related
6 aldolases are known in the art. Sequence alignments of SEQ ID NO's: 7, 8 and 9 with these
7 known epimerase/aldolase amino acid sequences will indicate conserved regions and amino acid
8 positions, the conservation of which are important for structure and enzymatic activity. These
9 regions and positions will tolerate no or only conservative amino acid substitutions. Amino acid
10 substitutions outside of these regions and positions are unlikely to greatly affect ribulose-5-P-4-
11 epimerase activity.

12 In accordance with the invention the eukaryotic host cell may comprise any possible
13 combination of at least one nucleotide sequence as defined in (a), at least one nucleotide
14 sequence as defined in (b) and at least one nucleotide sequence as defined in (c). Herein a
15 nucleotide sequence as defined in (a) can be a nucleotide sequence with a percentage of
16 sequence identity as indicated with an amino acid sequences of an arabinose isomerase (araA) of
17 at least one of *Clavibacter michiganensis* (C), *Arthrobacter aurescens* (A) and *Gramella forsetii*
18 (G); a nucleotide sequence as defined in (b) can be a nucleotide sequence with a percentage of
19 sequence identity as indicated with an amino acid sequences of a L-ribulose kinase (araB) of at
20 least one of *Clavibacter michiganensis* (C), *Arthrobacter aurescens* (A) and *Gramella forsetii*
21 (G); and, a nucleotide sequence as defined in (c) can be a nucleotide sequence with a percentage
22 of sequence identity as indicated with an amino acid sequences of an ribulose-5-P-4-epimerase
23 (araD) of at least one of *Clavibacter michiganensis* (C), *Arthrobacter aurescens* (A) and
24 *Gramella forsetii* (G). In particular the following combinations are included in the invention:
25 AAA; AAC; AAG; ACA; ACC; ACG; AGA; AGC; AGG; CAA; CAC; CAG; CCA; CCC;
26 CCG; CGA; CGC; CGG; GAA; GAC; GAG; GCA; GCC; GCG; GGA; GGC; GGG. Herein the
27 first position in each triplet indicates the type of the araA sequence, the second position indicates
28 the type of araB sequence, and the third position indicates the type of araD sequence, whereby
29 the letters "C", "A" and "G" indicate amino acid sequences with a percentage amino acid identity
30 as indicated to the corresponding enzymes of *Clavibacter michiganensis* (C), *Arthrobacter*
31 *aurescens* (A) and *Gramella forsetii* (G), respectively.

1 In a preferred embodiment of the invention, at least one of the nucleotide sequences as
2 defined in (a), (b) and (c) of claim 1 encodes an amino acid sequences that originate from a
3 bacterial genus selected from the group consisting of *Clavibacter*, *Arthrobacter* and *Gramella*,
4 i.e. the amino acid sequence is identical to an amino acid sequence as it naturally occurs in one
5 of these genera. More preferably, at least one of the nucleotide sequences as defined in (a), (b)
6 and (c) of claim 1 encodes an amino acid sequences that originate from a bacterial species
7 selected from the group consisting of *Clavibacter michiganensis*, *Arthrobacter aurescens* and
8 *Gramella forsetii*, i.e. the amino acid sequence is identical to an amino acid sequence as it
9 naturally occurs in one of these species.

10 To increase the likelihood that the arabinose isomerase, the ribulokinase and the ribulose-
11 5-P-4-epimerase are expressed at sufficient levels and in active form in the cells of the invention,
12 the nucleotide sequence encoding these enzymes, as well as other enzymes of the invention (see
13 below), are preferably adapted to optimise their codon usage to that of the host cell in question.
14 The adaptiveness of a nucleotide sequence encoding an enzyme to the codon usage of a host cell
15 may be expressed as codon adaptation index (CAI). The codon adaptation index is herein defined
16 as a measurement of the relative adaptiveness of the codon usage of a gene towards the codon
17 usage of highly expressed genes in a particular host cell or organism. The relative adaptiveness
18 (w) of each codon is the ratio of the usage of each codon, to that of the most abundant codon for
19 the same amino acid. The CAI index is defined as the geometric mean of these relative
20 adaptiveness values. Non-synonymous codons and termination codons (dependent on genetic
21 code) are excluded. CAI values range from 0 to 1, with higher values indicating a higher
22 proportion of the most abundant codons (see Sharp and Li , 1987, Nucleic Acids Research 15:
23 1281-1295; also see: Jansen et al., 2003, Nucleic Acids Res. 31(8):2242-51). An adapted
24 nucleotide sequence preferably has a CAI of at least 0.2, 0.3, 0.4, 0.5, 0.6 or 0.7. Most preferred
25 are the sequences as listed in SEQ ID NO's: 10 - 18, which have been codon optimised for
26 expression in *S. cerevisiae* cells.

27 The cell of the invention, preferably is a cell capable of active or passive pentose
28 (arabinose and xylose) transport into the cell. The cell preferably contains active glycolysis. The
29 cell further preferably contains an endogenous pentose phosphate pathway. The cell further
30 preferably contains enzymes for conversion of arabinose (and xylose), optionally through
31 pyruvate, to a desired fermentation product such as ethanol, lactic acid, 3-hydroxy-propionic

1 acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene,
2 glycerol, β -lactam antibiotics and cephalosporins. A particularly preferred cell is a cell that is
3 naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. The
4 cell further preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of
5 growth at a pH lower than 5, 4, or 3) and towards organic acids like lactic acid, acetic acid or
6 formic acid and sugar degradation products such as furfural and hydroxy-methylfurfural, and a
7 high tolerance to elevated temperatures. Any of these characteristics or activities of the cell may
8 be naturally present in the cell or may be introduced or modified by genetic modification,
9 preferably by self cloning or by the methods of the invention described below. A suitable cell is a
10 cultured cell, a cell that may be cultured in fermentation process e.g. in submerged or solid state
11 fermentation. Particularly suitable cells are eukaryotic microorganism like e.g. fungi, however,
12 most suitable for use in the present inventions are yeasts or filamentous fungi.

13 Yeasts are herein defined as eukaryotic microorganisms and include all species of the
14 subdivision Eumycotina (Alexopoulos, C. J., 1962, In: Introductory Mycology, John Wiley &
15 Sons, Inc., New York) that predominantly grow in unicellular form. Yeasts may either grow by
16 budding of a unicellular thallus or may grow by fission of the organism. Preferred yeasts as host
17 cells belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*,
18 *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, and *Yarrowia*. Preferably the
19 yeast is capable of anaerobic fermentation, more preferably anaerobic alcoholic fermentation.
20 Over the years suggestions have been made for the introduction of various organisms for the
21 production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol
22 production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol
23 producer. This is due to the many attractive features of *Saccharomyces* species for industrial
24 processes, i.e., a high acid-, ethanol- and osmo-tolerance, capability of anaerobic growth, and of
25 course its high alcoholic fermentative capacity. Preferred yeast species as fungal host cells
26 include *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* and *Schizosaccharomyces*
27 *pombe*.

28 Filamentous fungi are herein defined as eukaryotic microorganisms that include all
29 filamentous forms of the subdivision Eumycotina. These fungi are characterized by a vegetative
30 mycelium composed of chitin, cellulose, and other complex polysaccharides. The filamentous
31 fungi of the present invention are morphologically, physiologically, and genetically distinct from

1 yeasts. Vegetative growth by filamentous fungi is by hyphal elongation and carbon catabolism of
2 most filamentous fungi is obligately aerobic. Preferred filamentous fungi as host cells belong to
3 the genera *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium*, and *Penicillium*.

4 In a cell of the invention, the nucleotide sequence as defined in (a), (b) and (c) are
5 preferably operably linked to a promoter that causes sufficient expression of the nucleotide
6 sequences in the cell to confer to the cell the ability to convert L-arabinose into D-xylulose 5-
7 phosphate. Preferably, each of the nucleotide sequence as defined in (a), (b) and (c) is operably
8 linked to a promoter that causes sufficient expression of the nucleotide sequences in the cell to
9 confer to the cell the ability to convert L-arabinose into D-xylulose 5-phosphate. More preferably
10 the promoter(s) cause sufficient expression of the nucleotide sequences confers to the cell the
11 ability to grow on arabinose as sole carbon and/or energy source, most preferably the
12 promoter(s) cause sufficient expression of the nucleotide sequences confers to the cell the ability
13 to grow on arabinose as sole carbon and/or energy source through conversion of arabinose into
14 D-xylulose 5-phosphate (and further metabolism of D-xylulose 5-phosphate). Suitable promoters
15 for expression of the nucleotide sequence as defined in (a), (b) and (c) include promoters that are
16 insensitive to catabolite (glucose) repression and/or that do require xylose for induction.
17 Promoters having these characteristics are widely available and known to the skilled person.
18 Suitable examples of such promoters include e.g. promoters from glycolytic genes such as the
19 phosphofructokinase (*PPK*), triose phosphate isomerase (*TPI*), glyceraldehyde-3-phosphate
20 dehydrogenase (*GPD*, *TDH3* or *GAPDH*), pyruvate kinase (*PYK*), phosphoglycerate kinase
21 (*PGK*), glucose-6-phosphate isomerase promoter (*PGII*) promoters from yeasts or filamentous
22 fungi; more details about such promoters from yeast may be found in (WO 93/03159). Other
23 useful promoters are ribosomal protein encoding gene promoters, the lactase gene promoter
24 (*LAC4*), alcohol dehydrogenase promoters (*ADH1*, *ADH4*, and the like), the enolase promoter
25 (*ENO*), the hexose(glucose) transporter promoter (*HXT7*), and the cytochrome c1 promoter
26 (*CYC1*). Other promoters, both constitutive and inducible, and enhancers or upstream activating
27 sequences will be known to those of skill in the art. Preferably the promoter that is operably
28 linked to nucleotide sequence as defined in (a), (b) and (c) is homologous to the host cell. It is
29 preferred that for expression of each of the nucleotide sequence as defined in (a), (b) and (c) a
30 different promoter is used. This will improved stability of the expression construct by avoiding

1 homologous recombination between repeated promoter sequences and it avoids competition
2 different copies of the promoter for limiting trans-acting factors.

3 A cell of the invention further preferably comprises a genetic modification that increases
4 the flux of the pentose phosphate pathway as described in WO 06/009434. In particular, the
5 genetic modification causes an increased flux of the non-oxidative part pentose phosphate
6 pathway. A genetic modification that causes an increased flux of the non-oxidative part of the
7 pentose phosphate pathway is herein understood to mean a modification that increases the flux
8 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
9 genetically identical except for the genetic modification causing the increased flux. The flux of
10 the non-oxidative part of the pentose phosphate pathway may be measured as described in WO
11 06/009434.

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

20 In a preferred cell of the invention, the genetic modification comprises overexpression of at
21 least one enzyme of the (non-oxidative part) pentose phosphate pathway. Preferably the enzyme
22 is selected from the group consisting of the enzymes encoding for ribulose-5-phosphate
23 isomerase, ribulose-5-phosphate 3-epimerase, transketolase and transaldolase. Various
24 combinations of enzymes of the (non-oxidative part) pentose phosphate pathway may be
25 overexpressed. E.g. the enzymes that are overexpressed may be at least the enzymes ribulose-5-
26 phosphate isomerase and ribulose-5-phosphate 3-epimerase; or at least the enzymes ribulose-5-
27 phosphate isomerase and transketolase; or at least the enzymes ribulose-5-phosphate isomerase
28 and transaldolase; or at least the enzymes ribulose-5-phosphate 3-epimerase and transketolase; or
29 at least the enzymes ribulose-5-phosphate 3-epimerase and transaldolase; or at least the enzymes
30 transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate 3-epimerase,
31 transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase,

1 transketolase and transaldolase; or at least the enzymes ribulose-5-phosphate isomerase,
2 ribulose-5-phosphate 3-epimerase, and transaldolase; or at least the enzymes ribulose-5-
3 phosphate isomerase, ribulose-5-phosphate 3-epimerase, and transketolase. In one embodiment
4 of the invention each of the enzymes ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-
5 epimerase, transketolase and transaldolase are overexpressed in the cell of the invention.
6 Preferred is a cell in which the genetic modification comprises at least overexpression of the
7 enzyme transaldolase. More preferred is a cell in which the genetic modification comprises at
8 least overexpression of both the enzymes transketolase and transaldolase as such a host cell is
9 already capable of anaerobic growth on arabinose. In fact, under some conditions we have found
10 that cells overexpressing only the transketolase and the transaldolase already have the same
11 anaerobic growth rate on arabinose as do cells that overexpress all four of the enzymes, i.e. the
12 ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, transketolase and
13 transaldolase. Moreover, cells of the invention overexpressing both of the enzymes ribulose-5-
14 phosphate isomerase and ribulose-5-phosphate 3-epimerase are preferred over cells
15 overexpressing only the isomerase or only the 3-epimerase as overexpression of only one of
16 these enzymes may produce metabolic imbalances.

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

25 Alternatively overexpression of enzymes in the cells of the invention may be achieved by
26 using a promoter that is not native to the sequence coding for the enzyme to be overexpressed,
27 i.e. a promoter that is heterologous to the coding sequence to which it is operably linked.
28 Although the promoter preferably is heterologous to the coding sequence to which it is operably
29 linked, it is also preferred that the promoter is homologous, i.e. endogenous to the cell of the
30 invention. Preferably the heterologous promoter is capable of producing a higher steady state
31 level of the transcript comprising the coding sequence (or is capable of producing more transcript

1 molecules, i.e. mRNA molecules, per unit of time) than is the promoter that is native to the
2 coding sequence, preferably under conditions where arabinose or arabinose and glucose are
3 available as carbon sources, more preferably as major carbon sources (i.e. more than 50% of the
4 available carbon source consists of arabinose or arabinose and glucose), most preferably as sole
5 carbon sources. Suitable promoters in this context include promoters as described above for
6 expression of the nucleotide sequences as defined in (a), (b) and (c).

7 A further preferred cell of the invention comprises a genetic modification that reduces
8 unspecific aldose reductase activity in the cell. Preferably, unspecific aldose reductase activity is
9 reduced in the host cell by one or more genetic modifications that reduce the expression of or
10 inactivates a gene encoding an unspecific aldose reductase. Preferably, the genetic modifications
11 reduce or inactivate the expression of each endogenous copy of a gene encoding an unspecific
12 aldose reductase that is capable of reducing an aldopentose, including arabinose, xylose and
13 xylulose, in the cell's genome. A given cell may comprise multiple copies of genes encoding
14 unspecific aldose reductases as a result of di-, poly- or aneu-ploidy, and/or a cell may contain
15 several different (iso)enzymes with aldose reductase activity that differ in amino acid sequence
16 and that are each encoded by a different gene. Also in such instances preferably the expression of
17 each gene that encodes an unspecific aldose reductase is reduced or inactivated. Preferably, the
18 gene is inactivated by deletion of at least part of the gene or by disruption of the gene, whereby
19 in this context the term gene also includes any non-coding sequence up- or down-stream of the
20 coding sequence, the (partial) deletion or inactivation of which results in a reduction of
21 expression of unspecific aldose reductase activity in the host cell. A nucleotide sequence
22 encoding an aldose reductase whose activity is to be reduced in the cell of the invention and
23 amino acid sequences of such aldose reductases are described in WO 06/009434 and include e.g.
24 the (unspecific) aldose reductase genes of *S. cerevisiae* *GRE3* gene (Träff et al., 2001, Appl.
25 Environm. Microbiol. 67: 5668-5674) and orthologues thereof in other species.

26 In a further preferred embodiment, the cell of the invention that has the ability to convert
27 L-arabinose into D-xylulose 5-phosphate expressing in addition has the ability of isomerising
28 xylose to xylulose as e.g. described in WO 03/0624430 and in WO 06/009434. The ability of
29 isomerising xylose to xylulose is preferably conferred to the cell by transformation with a nucleic
30 acid construct comprising a nucleotide sequence encoding a xylose isomerase. Preferably the cell
31 thus acquires the ability to directly isomerise xylose into xylulose. More preferably the cell thus

1 acquires the ability to grow aerobically and/or anaerobically on xylose as sole energy and/or
2 carbon source though direct isomerisation of xylose into xylulose (and further metabolism of
3 xylulose). It is herein understood that the direct isomerisation of xylose into xylulose occurs in a
4 single reaction catalysed by a xylose isomerase, as opposed to the two step conversion of xylose
5 into xylulose via a xylitol intermediate as catalysed by xylose reductase and xylitol
6 dehydrogenase, respectively.

7 Several xylose isomerases (and their amino acid and coding nucleotide sequences) that
8 may be successfully used to confer to the cell of the invention the ability to directly isomerise
9 xylose into xylulose have been described in the art. These include the xylose isomerase of
10 *Piromyces* sp. and of other anaerobic fungi that belongs to the families *Neocallimastix*,
11 *Caecomyces*, *Piromyces*, *Orpinomyces*, or *Ruminomyces* (WO 03/0624430), the xylose
12 isomerase of the bacterial genus *Bacteroides*, including e.g. *B. thetaiotaomicron* (WO
13 06/009434) and *B. fragilis*, and the xylose isomerase of the anaerobic fungus *Cyllamyces*
14 *aberensis* (US 20060234364). Preferably, a xylose isomerase that may be used to confer to the
15 cell of the invention the ability to directly isomerise xylose into xylulose is a xylose isomerase
16 comprising an amino acid sequence that has at least 70, 75, 80, 83% amino acid identity with the
17 amino acid sequence of SEQ ID NO. 19 or 20.

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

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

9 The cells according to the invention may comprises further genetic modifications that
10 result in one or more of the characteristics selected from the group consisting of (a) increased
11 transport of arabinose and/or xylose into the cell; (b) decreased sensitivity to catabolite
12 repression; (c) increased tolerance to ethanol, osmolarity or organic acids; and, (e) reduced
13 production of by-products. By-products are understood to mean carbon-containing molecules
14 other than the desired fermentation product and include e.g. arabinitol, xylitol, glycerol and/or
15 acetic acid. Any genetic modification described herein may be introduced by classical
16 mutagenesis and screening and/or selection for the desired mutant, or simply by screening and/or
17 selection for the spontaneous mutants with the desired characteristics. Alternatively, the genetic
18 modifications may consist of overexpression of endogenous genes and/or the inactivation of
19 endogenous genes.

20 Genes the overexpression of which is desired for increased transport of arabinose and/or
21 xylose into the cell are preferably chosen from genes encoding a hexose or pentose transporter.
22 In *S. cerevisiae* these genes include *HXT1*, *HXT2*, *HXT4*, *HXT5*, *HXT7* and *GAL2*, of which
23 *HXT7*, *HXT5* and *GAL2* are most preferred (see Sedlack and Ho, Yeast 2004; 21: 671–684).
24 Similarly orthologues of these genes in other species may be overexpressed.

25 Other genes that may be overexpressed in the cells of the invention include genes coding
26 for glycolytic enzymes and/or ethanologenic enzymes such as alcohol dehydrogenases.

27 Preferred endogenous genes for inactivation include hexose kinase genes e.g. the *S.*
28 *cerevisiae* *HXK2* gene (see Diderich et al., 2001, Appl. Environ. Microbiol. 67: 1587-1593); the
29 *S. cerevisiae* *MIG1* or *MIG2* genes; genes coding for enzymes involved in glycerol metabolism
30 such as the *S. cerevisiae* glycerol-phosphate dehydrogenase 1 and/or 2 genes; or (hybridising)
31 orthologues of these genes in other species.

1 Other preferred further modifications of host cells for xylose fermentation are described in
2 van Maris et al. (2006, Antonie van Leeuwenhoek 90:391–418), WO2006/009434,
3 WO2005/023998, WO2005/111214, and WO2005/091733.

4 Any of the genetic modifications of the cells of the invention as described herein are, in as
5 far as possible, preferably introduced or modified by self cloning genetic modification.

6 A preferred cell of the invention with one or more of the genetic modifications described
7 above, including modifications obtained by selection of (spontaneous) mutants, has the ability to
8 grow on L-arabinose and optionally xylose as carbon/energy source, preferably as sole carbon
9 source, and preferably under anaerobic conditions. Preferably the cell produces essentially no
10 arabinitol, e.g. the arabinitol produced is below the detection limit or e.g. less than 5, 2, 1, 0.5, or
11 0.3 % of the carbon consumed on a molar basis. Preferably, in case carbon/energy source also
12 includes xylose, the cell produces essentially no xylitol, e.g. the xylitol produced is below the
13 detection limit or e.g. less than 5, 2, 1, 0.5, or 0.3 % of the carbon consumed on a molar basis.

14 A cell of the invention preferably has the ability to grow on L-arabinose as sole
15 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
16 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
17 or 0.2 h⁻¹ under anaerobic conditions. A cell of the invention preferably has the ability to grow
18 on a mixture of glucose and L-arabinose (in a 1:1 weight ratio) as sole carbon/energy source at a
19 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
20 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
21 anaerobic conditions.

22 Preferably, a cell of the invention has a specific L-arabinose consumption rate of at least
23 346, 400, 600, 700, 800, 900 or 1000 mg h⁻¹ (g dry weight)⁻¹. Preferably, a cell of the invention
24 has a yield of fermentation product (such as ethanol) on L-arabinose that is at least 20, 40, 50,
25 60, 80, 90, 95 or 98% of the cell's yield of fermentation product (such as ethanol) on glucose.
26 More preferably, the modified host cell's yield of fermentation product (such as ethanol) on L-
27 arabinose is equal to the host cell's yield of fermentation product (such as ethanol) on glucose.
28 Likewise, the modified host cell's biomass yield on L-arabinose is preferably at least 55, 60, 70,
29 80, 85, 90, 95 or 98% of the host cell's biomass yield on glucose. More preferably, the modified
30 host cell's biomass yield on L-arabinose is equal to the host cell's biomass yield on glucose. It is

1 understood that in the comparison of yields on glucose and L-arabinose both yields are compared
2 under aerobic conditions or both under anaerobic conditions.

3 In another aspect the invention relates to a eukaryotic cell comprising nucleotide sequences
4 as encoding (a') an arabinose isomerase, (b') a xylulose kinase, and (c') a ribulose-5-P-4-
5 epimerase, whereby the expression of the nucleotide sequences confers to the cell the ability to
6 convert L-arabinose into D-xylulose 5-phosphate. In this embodiment the broad substrate
7 specificity of xylulose kinases, in particular eukaryotic xylulose kinases, is exploited to
8 phosphorylate ribulose (and optionally xylulose). Expressly included in also this embodiment of
9 the invention are eukaryotic cells that may already have the ability to convert L-arabinose into
10 D-xylulose 5-phosphate (at a low level) and wherein expression of the nucleotide sequences as
11 defined in (a'), (b') and (c') increases the cell's ability to convert L-arabinose into D-xylulose 5-
12 phosphate. Preferably, in the cells of the invention, the ability to convert L-arabinose into D-
13 xylulose 5-phosphate is the ability to convert L-arabinose into D-xylulose 5-phosphate through
14 the subsequent reactions of 1) isomerisation of arabinose into ribulose; 2) phosphorylation of
15 ribulose to ribulose 5-phosphate; and, 3) epimerisation of ribulose 5-phosphate into D-xylulose
16 5-phosphate. Preferably expression of the nucleotide sequences confers to, or increases in the
17 cell the ability to grow on arabinose as sole carbon and/or energy source, more preferably
18 expression of the nucleotide sequences confers to the cell, or increases in the ability to grow on
19 arabinose as sole carbon and/or energy source through conversion of arabinose into D-xylulose
20 5-phosphate (and further metabolism of D-xylulose 5-phosphate).

21 The nucleotide sequence (a') encoding the arabinose isomerase may be a nucleotide
22 sequence (a) as defined above, however the nucleotide sequence may also encode any other,
23 preferably bacterial, arabinose isomerase, e.g. those from *E. coli*, *Bacillus* and *Lactobacillus* as
24 described in e.g. EP 1499708 and Wisselink et al. (2007, *supra*). Preferably, the nucleotide
25 sequence encoding the arabinose isomerase comprises an amino acid sequence that has at least
26 30, 35, 40, 45, or 50% sequence identity with at least one of the amino acid sequences of SEQ ID
27 NO's: 1, 2 and 3.

28 The nucleotide sequence (b') encoding a polypeptide with xylulose kinase activity
29 preferably comprises an amino acid sequence having at least 50, 60, 70, 80, 90 or 95% identity
30 with SEQ ID NO. 21.

1 The nucleotide sequence (c') encoding the ribulose-5-P-4-epimerase may be a nucleotide
2 sequence (c) as defined above, however the nucleotide sequence may also encode any other,
3 preferably bacterial, ribulose-5-P-4-epimerase, e.g. those from *E. coli*, *Bacillus* and *Lactobacillus*
4 as described in e.g. EP 1499708 and Wisselink et al. (2007, *supra*). Preferably, the nucleotide
5 sequence encoding the ribulose-5-P-4-epimerase comprises an amino acid sequence that has at
6 least 30, 35, 40, 45, or 50% sequence identity with at least one of the amino acid sequences of
7 SEQ ID NO's: 7, 8 and 9.

8 The eukaryotic cell comprising the nucleotide sequence encoding an eukaryotic xylulose
9 kinase, in stead of a bacterial ribulose kinase, may the same as the above described cells
10 comprising the nucleotide sequence encoding a bacterial ribulose kinase sequences in all aspects
11 except for the more broadly defined nucleotide sequences (a') and (c') and the different
12 nucleotide sequence (b').

13 In another aspect the invention relates to a process for producing a fermentation product
14 selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid,
15 acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam
16 antibiotics and cephalosporins. The process preferably comprises the steps of: a) fermenting a
17 medium containing a source of arabinose, and optionally xylose, with a cell as defined
18 hereinabove, whereby the cell ferments arabinose, and optionally xylose, to the fermentation
19 product, and optionally, b) recovery of the fermentation product.

20 In addition to a source of arabinose the carbon source in the fermentation medium may also
21 comprise a source of glucose. The skilled person will further appreciate that the fermentation
22 medium may further also comprise other types of carbohydrates such as e.g. in particular a
23 source of xylose. The sources of arabinose, glucose and xylose may be arabinose, glucose and
24 xylose as such (i.e. as monomeric sugars) or they may be in the form of any carbohydrate oligo-
25 or polymer comprising arabinose, glucose and/or xylose units, such as e.g. lignocellulose,
26 arabinans, xylans, cellulose, starch and the like. For release of arabinose, glucose and/or xylose
27 units from such carbohydrates, appropriate carbohydrases (such as arabinases, xylanases,
28 glucanases, amylases, cellulases, glucanases and the like) may be added to the fermentation
29 medium or may be produced by the modified host cell. In the latter case the modified host cell
30 may be genetically engineered to produce and excrete such carbohydrases. An additional
31 advantage of using oligo- or polymeric sources of glucose is that it enables to maintain a low(er)

1 concentration of free glucose during the fermentation, e.g. by using rate-limiting amounts of the
2 carbohydrases preferably during the fermentation. This, in turn, will prevent repression of
3 systems required for metabolism and transport of non-glucose sugars such as arabinose and
4 xylose. In a preferred process the modified host cell ferments both the arabinose and glucose,
5 and optionally xylose, preferably simultaneously in which case preferably a modified host cell is
6 used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source
7 of arabinose (and glucose) as carbon source, the fermentation medium will further comprise the
8 appropriate ingredient required for growth of the modified host cell. Compositions of
9 fermentation media for growth of eukaryotic microorganisms such as yeasts and filamentous
10 fungi are well known in the art.

11 The fermentation process may be an aerobic or an anaerobic fermentation process. An
12 anaerobic fermentation process is herein defined as a fermentation process run in the absence of
13 oxygen or in which substantially no oxygen is consumed, preferably less than 5, 2.5 or 1
14 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable),
15 and wherein organic molecules serve as both electron donor and electron acceptors. In the
16 absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidised by
17 oxidative phosphorylation. To solve this problem many microorganisms use pyruvate or one of
18 its derivatives as an electron and hydrogen acceptor thereby regenerating NAD^+ . Thus, in a
19 preferred anaerobic fermentation process pyruvate is used as an electron (and hydrogen acceptor)
20 and is reduced to fermentation products such as ethanol, lactic acid, 3-hydroxy-propionic acid,
21 acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene,
22 glycerol, β -lactam antibiotics and cephalosporins. Anaerobic processes of the invention are
23 preferred over aerobic processes because anaerobic processes do not require investments and
24 energy for aeration and in addition, anaerobic processes produce higher product yields than
25 aerobic processes. Alternatively, the fermentation process of the invention may be run under
26 aerobic oxygen-limited conditions. Preferably, in an aerobic process under oxygen-limited
27 conditions, the rate of oxygen consumption is at least 5.5, more preferably at least 6 and even
28 more preferably at least 7 mmol/L/h.

29 The fermentation process is preferably run at a temperature that is optimal for the modified
30 cells of the invention. Thus, for most yeasts or fungal cells, the fermentation process is
31 performed at a temperature which is less than 42°C, preferably less than 38°C. For yeast or

1 filamentous fungal cells, the fermentation process is preferably performed at a temperature
2 which is lower than 35, 33, 30 or 28°C and at a temperature which is higher than 20, 22, or 25°C.

3 Preferably in the fermentation processes of the invention, the cells stably maintain the
4 nucleic acid constructs that confer to the cell the ability of converting arabinose into D-xylulose
5 5-phosphate, and optionally isomerising xylose to xylulose. Preferably in the process at least 10,
6 20, 50 or 75% of the cells retain the abilities to convert arabinose into D-xylulose 5-phosphate,
7 and optionally isomerise xylose to xylulose after 50 generations of growth, preferably under
8 industrial fermentation conditions.

9 A preferred fermentation process according to the invention is a process for the production
10 of ethanol, whereby the process comprises the steps of: a) fermenting a medium containing a
11 source of arabinose, and optionally xylose, with a cell as defined hereinabove, whereby the cell
12 ferments arabinose, and optionally xylose, to ethanol, and optionally, b) recovery of the ethanol.
13 The fermentation medium may further be performed as described above. In the process the
14 volumetric ethanol productivity is preferably at least 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 5.0 or 10.0 g
15 ethanol per litre per hour. The ethanol yield on arabinose and/or glucose and/or xylose in the
16 process preferably is at least 50, 60, 70, 80, 90, 95 or 98%. The ethanol yield is herein defined as
17 a percentage of the theoretical maximum yield, which, for arabinose, glucose and xylose is 0.51
18 g. ethanol per g. arabinose, glucose or xylose.

19 A further preferred fermentation process according to the invention is a process which
20 comprises fermenting a medium containing a source of arabinose and a source of xylose wherein
21 however two separate strains of cells are used, a first strain of cells as defined hereinabove
22 except that cells of the first strain do not have the ability to (directly) isomerise xylose into
23 xylulose, which cells of the first strain ferment arabinose to the fermentation product; and a
24 second strain of cells as defined hereinabove except that cells of the second strain do not have
25 the ability to convert arabinose to xylulose 5-phosphate, which cells of the second strain ferment
26 xylose to the fermentation product. The process optionally comprises the step of recovery of the
27 fermentation product. The cells of the first and second are further as otherwise described
28 hereinabove.

29 In this document and in its claims, the verb "to comprise" and its conjugations is used in its
30 non-limiting sense to mean that items following the word are included, but items not specifically
31 mentioned are not excluded. In addition, reference to an element by the indefinite article "a" or

1 "an" does not exclude the possibility that more than one of the element is present, unless the
2 context clearly requires that there be one and only one of the elements. The indefinite article "a"
3 or "an" thus usually means "at least one".

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

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

8

Description of the figuresFigure 1

Physical map of plasmid pRS316 GGA showing the three ara genes. The most important restriction-enzyme recognition sites used for cloning are indicated.

Figure 2

Colony PCR on RN1002 and as a negative control on the host strain RN1000. The Fermentas 1kb ladder is used to control the length of the amplified fragments. On the left side RN1002 and on the right side RN1000 results are shown. All fragment sizes are as expected. Used primers are indicated in Table 1.

Examples

1. Example 1

1.1. Plasmids

1.1.1 *araA*

For high level of expression of the bacterial *araA* and *araD* genes the corresponding expression cassettes are inserted into the 2 μ plasmid pAKX002 that already comprises the *Piromyces xylA* gene linked the *S. cerevisiae TPI* promoter. The *araA* expression cassettes is constructed by amplifying the *S. cerevisiae TDH3* promoter (P_{TDH3}) with oligo's that allow to link the *TDH3* promoter to the 5' end of the synthetic *araA* coding sequences of *Arthrobacter aurescens* (SEQ ID NO. 10), *Clavibacter michiganensis* (SEQ ID NO. 11) and *Gramella forsetii* (SEQ ID NO. 12), and amplifying the *S. cerevisiae ADHI* terminator with oligo's that allow to link the 3' end of the synthetic *araA* coding sequences to the *ADHI* terminator (T_{ADHI}). The two fragments are extracted from gel and mixed in roughly equimolar amounts with the fragments of the synthetic *araA* coding sequences. On this mixture a PCR is performed using the 5' P_{TDH3} oligo and the 3' T_{ADHI} oligo. The resulting P_{TDH3} -*araA*- T_{ADHI} cassette is gel purified, cut at the 5' and 3' restriction sites and then ligated into pAKX002, resulting in plasmids pRN-AAaraA, pRN-CMaraA and pRN-GFaraA, respectively.

1.1.2 *araD*

The three *araD* constructs are made by first amplifying a truncated version of the *S. cerevisiae HXT7* promoter (P_{HXT7}) with oligo's that allow to link the *HXT7* promoter to the 5' end of the synthetic *araD* coding sequences of *Arthrobacter aurescens* (SEQ ID NO. 16), *Clavibacter michiganensis* (SEQ ID NO. 17) and *Gramella forsetii* (SEQ ID NO. 18), and amplifying the *PGII* terminator with oligo's that allow to link the 3' end of the synthetic *araD* coding sequences to the *PGII* terminator region (T_{PGI}). The resulting fragments were extracted from gel and mixed in roughly equimolar amounts with the synthetic *araD* coding sequences, after which a PCR was performed using the 5' P_{HXT7} oligo and the 3' T_{PGI} oligo. The resulting P_{HXT7} -*araD*- T_{PGI} cassettes are gel purified, cut at the 5' and 3' restriction sites and ligated into pRN-AAaraA, pRN-CMaraA and pRN-GFaraA, respectively, resulting in plasmids pRN-AAaraAD, pRN-CMaraAD and pRN-GFaraAD, respectively.

1 1.1.3 *araB*

2 For the expression of the three bacterial *araB* genes, the integrational plasmid pRS305 is used
 3 (Gietz and Sugino, 1988, Gene 74:527-534). Aside from the bacterial *AraB* genes, the
 4 *S.cerevisiae XKS1* gene was also included on this vector. For this, the P_{ADHI}-XKS1-T_{CYCI}
 5 containing *PvuI* fragment from p415ADHXKS was ligated into the *PvuI* digested vector
 6 backbone from the integration plasmid pRS305, resulting in pRN-XKS1. For expression of the
 7 bacterial *araB* genes, three cassettes containing the synthetic *araB* coding sequences of
 8 *Arthrobacter aureus* (SEQ ID NO. 13), *Clavibacter michiganensis* (SEQ ID NO. 14) and
 9 *Gramella forsetii* (SEQ ID NO. 15) genes between the *PGII* promoter (P_{PGI}) and *ADHI*
 10 terminator (T_{ADHI}) is constructed by PCR amplification. The *AraB* expression cassettes are made
 11 by amplifying the *PGII* promoter with oligonucleotides that allow to link the *PGII* promoter to
 12 the 5' end of the synthetic *araB* coding sequences, and amplifying the *ADHI* terminator with
 13 oligo's that allow to link the 3' end of the synthetic *araB* coding sequences to the *ADHI*
 14 terminator (T_{ADHI}). The resulting P_{PGI}-*araB*-T_{ADHI} cassettes are gel purified, digested at the
 15 5' and 3' restriction sites and are then ligated into pRN-XKS1, to yield plasmids pRN-XKS1-
 16 AAaraB, pRN-XKS1-CMaraB and pRN-XKS1-GFaraB, respectively.

17 1.2 Strains

18 Media for cultivations of *Saccharomyces cerevisiae* strains, shake flask and fermenter
 19 cultivations as well as sequential batch fermentation under aerobic, oxygen-limited and
 20 anaerobic conditions were performed as described in Wisselink et al. (2007, AEM Accepts,
 21 published online ahead of print on 1 June 2007; Appl. Environ. Microbiol.
 22 doi:10.1128/AEM.00177-07).

23 1.2.1 Derivation of host strain RN679 from RWB218

24 The *S. cerevisiae* strains in this work are derived from the xylose-fermenting strain RWB217
 25 (Kuyper et al., 2005a, FEMS Yeast Res. 5:399-409): RWB217 has the following genotype:
 26 MATA ura3-52 leu2-112 loxP-PTPI::(-266,-1)TAL1 gre3::hphMX pUGPTPI-TKL1 pUGPTPI-
 27 RPE1 KanloxP-PTPI::(-?,-1)RKI1 {p415ADHXKS, pAKX002}. Strain RWB218 is obtained by
 28 selection of RWB217 for improved growth on D-xylose (Kuyper et al., 2005b, FEMS Yeast Res.
 29 5:925-934) by plating and restreaking on MYD plates. RWB218 is grown non-selectively on
 30 YPD in order to facilitate the loss of plasmids pAKX002 and p415ADHXKS1 (Kuyper et al.,
 31 2005a, *supra*), harbouring the URA3 and LEU2 selective markers, respectively. RWB218 is

plated on YPD, single colonies are screened for plasmid loss by testing for uracil and leucine auxotrophy. In order to remove a KanMX cassette - still present after integrating the RKI1 overexpression construct (Kuyper et al., 2005a, *supra*) - a strain from which both plasmids are lost is transformed with pSH47, containing the cre recombinase (Güldener et al., 1996, Nucleic Acids Res., 24:2519-252410). Transformants containing pSH47 are resuspended in YP with 1% D-galactose and incubated for 1 hour at 30°C. Cells are plated on YPD and colonies are screened for loss of the KanMX marker (G418 resistance) and pSH47 (URA3). A strain that has lost both the KanMX marker and the pSH47 plasmid is designated as RN679. The genotype of RN679 is: MATA ura3-52 leu2-112 loxP-PTPI::(-266,-1)TAL1 gre3::hphMX pUGPTPI-TKL1 pUGPTPI-RPE1 KanloxP-PTPI::(-?, -1)RKI1.

1.2.2 Transformations of RN679

RN679 is transformed with:

- 1) pRN-AAaraAD and pRN-XKS1-AAaraB, resulting in strain RN680;
- 2) pRN-CMaraAD and pRN-XKS1-CMaraB, resulting in strain RN681; and
- 3) pRN-GFaraAD and pRN-XKS1-GFaraB, resulting in strain RN681.

1.2.3 Selection of strains RN680, RN681 and RN682 for aerobic growth on L-arabinose

Strains RN680, RN681 and RN682 do not grow on solid synthetic medium supplemented with 2% (w/v) L-arabinose (MYA). Therefore, evolutionary engineering is applied for the selection of cells of the strains RN680, RN681 and RN682 with an improved specific growth rate on arabinose. Prior to the selection in synthetic medium supplemented with 2% of arabinose, cells are pre-grown in synthetic medium with galactose, as it is known that galactose-induced *S. cerevisiae* cells can transport L-arabinose via the galactose permease GAL2p (Kou et al., 1970, J. Bacteriol. 103:671-67817). Galactose-grown cells of strains RN680, RN681, RN682 and control strain RWB218 are transferred to shake flasks containing MY supplemented with 0.1% D-galactose and 2% L-arabinose. After approximately several weeks of cultivation in the single initial shake flask, the cultures of strains RN680, RN681, RN682 IMS0001 show very slow growth after depletion of the galactose, in contrast to the reference strain RWB218 which does not grow after depletion of galactose. Cells of the cultures are next transferred to fresh synthetic medium supplied with 2% of L-arabinose (MYA). After again 1-3 weeks of cultivation in MYA descendants of strains RN680, RN681, RN682 grow with an improved doubling time, whereas strain RWB219 still does not grow. Next cells are sequentially transferred each time an OD660

of 2-3 is reached to fresh MYA with a start OD660 of approximately 0.05 and gradually the specific growth rate of the sequentially transferred cultures increases.

1.2.4 Selection of strains RN680, RN681 and RN682 for anaerobic growth on L-arabinose

To allow for a more gradual transfer to anaerobic conditions, the aerobically evolved strains, as obtained in Example 2.3 above, are first grown under oxygen-limited conditions. As soon as growth is observed under oxygen-limited conditions, the culture is switched to anaerobic conditions in the next batch cycle. Upon arabinose depletion, as indicated by the CO₂ percentage dropping below 0.05% after the CO₂ production peak, a new cycle is initiated by either manual or automated replacement of approximately 90% of the culture with fresh synthetic medium containing 20 g l⁻¹ L-arabinose. In 10-15 cycles, the anaerobic specific growth rate increases as estimated from the CO₂ profile. After 20-25 cycles no significant further increase of the growth rate is noticed. Single colonies are isolated on solid MYA for anaerobically evolved descendants of each of RN680, RN681 and RN682.

Example 2

2.1 Donor organisms and genes

As described in Example 1, three donor organisms were selected:

- *Arthrobacter aurescens* (A)
- *Clavibacter michiganensis* (C)
- *Gramella forsetii* (G)

The arabinose genes selected were:

- araA: arabinose isomerase EC 3.5.1.4
- araB: ribulokinase EC 2.7.1.16
- araD: L-ribulose-5-phosphate 4-epimerase EC 5.1.3.4

The 9 genes were synthesized by EXONBIO based on sequences that were optimized for codon usage in yeast by Nextgen Sciences. See sequence listings.

To express the araA gene in *Saccharomyces cerevisiae* the HXT7 promoter (410bp) and the PGI1 terminator (329bp) sequences were used.

To express the araB gene in *Saccharomyces cerevisiae* the TPI1 promoter (899bp) and the ADH1 terminator (351bp) sequences were used.

1 To express the *araD* gene in *Saccharomyces cerevisiae* the TDH3 promoter (686bp) and the
2 CYC1 terminator (288bp) sequences were used

3 The first three nucleotides in front of the ATG were modified into AAA in order to optimize
4 expression.

5 2.2 Host organism

6 The yeast host strain was RN1000. This strain is a derivative of strain RWB 218 (Kuyper et al.,
7 FEMS Yeast Research 5, 2005, 399-409). The plasmid pAKX002 encoding the *Piromyces XylA*
8 is lost in RN1000. The genotype of the host strain is: *MatA*, *ura3-52*, *leu2-112*, *gre3::hphMX*,
9 *loxP-Ptpi::TAL1*, *KanloxP-Ptpi::RKII*, *pUGPtpi-TKL1*, *pUGPtpi-RPE1*, {*p415*
10 *Padh1XKS1Tcyc1-LEU2*}

11 2.3 Molecular techniques employed in plasmid construction

12 The synthetic genes were amplified using the 'polymerase chain reaction (PCR)' technique
13 facilitating cloning. For each reaction two short synthetic oligomers 'primers' were used. The
14 one in the 'forward' and the other in the 'reverse' mode. Constitutive promoter sequences and
15 terminator sequences from *Saccharomyces cerevisiae* were also amplified using PCR. In Table 1
16 an overview of all primers used in this study is given. To minimize PCR-induced sequence
17 mistakes, the Finnzymes proofreading enzyme Phusion was used.

18 The plasmid used to express the *ara* genes into yeast is pRS316 (Sikorski R.S., Hieter P., "A
19 system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in
20 *Saccharomyces cerevisiae*" Genetics 122:19-27(1989), accession U03442, ATCC77145). This
21 plasmid is a centromeric plasmid (low copynumber in yeast) that has the *URA3* gene for
22 selection.

23 The construction of the pRS316 GGA plasmid is given below. The primers used contained
24 specific restriction-enzyme recognition sites. Construction involved standard molecular
25 biological techniques.

26 GaraA: promoter cut with *NotI* and *PstI*; ORF cut with *PstI* and *XhoI*; terminator cut with *XhoI*
27 and *BsiWI*.

28 GaraB: promoter cut with *AgeI* and *XbaI*; ORF cut with *XbaI* and *BssHII*; terminator cut with
29 *BssHII* and *BsiWI*.

30 AaraD: promoter cut with *AgeI* and *HindIII*; ORF cut with *HindIII* and *BamHI*; terminator cut
31 with *BamHI* and *XhoI*.

- 1
- 2 Table 1 Overview of the primers used in this study. Explanation code: e.g. DPF = araD promoter
- 3 Forward, BTR = araB terminator Reverse and CMDR = *Clavibacter michiganensis* araD Reverse.

1

DPF	AAGAGCTCACCGGTTTATCATTATCAATACTGCC
DPR	AAGAATTCAAGCTTTATGTGTGTTTATTCGAAACTAAGTTCTTG
DTF	AAGAATTCGGATCCCCTTTTCCTTTGTCTGA
DTR	AACTCGAGCCTAGGAAGCCTTCGAGCGTC
AADF	AAAAGCTTAAGAAAATGAGTTCACCTTCTGGAGTC
AADR	TTGGATCCGACGTCACCTACCGTAAACGTTTTGG
CMDF	AAAAGCTTAAGAAAATGTCCACGTATGCCCC
CMDR	TTGGATCCGACGTCATTTTAAACGCACCTTGCG
GFDF	AAAAGCTTAAGAAAATGTCGAGCCAATACAAAGA
GFDR	TTGGATCCGACGTCAGTTCTGTCCATAATATGCG
BPF	AACCGGTTTCTTCTTCAGATTCCCTC
BPR	TTAGATCTCTAGATTTATGTATGTGTTTTTTGTAGT
BTF	AAAGATCTGCGCGCGAATTTCTTATGATTTATG
BTR	TTAAGCTTCGTACGTGTGGAAGAACGAT
AABF	AATCTAGATTAATAAAAATGAATACGTCCGAAAACATACCC
AABR	TTGCGCGCGACGTCACGCGGACGCCCC
CMBF	AATCTAGATTAATAAAAATGCCTTCGGCTCCCG
CMBR	TTGCGCGCGACGTCAGGCCCTGGCTTCCCTTTTC
GFBF	AATCTAGATTAATAAAAATGTCGAATTATGTCATCGGG
GFBR	TTGCGCGCGACGTCAAACAGCGAATTCGTTC
APF	AAGCGGCCGCGGCTACTTCTCGTAGGAAC
APR	TTAGATCTGCAGAATTAAAAAACTTTTTGTTTTTTGTG
ATF	AAAGATCTCGAGACAAATCGCTCTTAAATATATACC
ATR	TTAAGCTTCGTACGTTTTTAAACAGTTGATGAGAACC
AAAF	AACTGCAGATATCAAAATGCCATCAGCTACCAGC
AAAR	TTCTCGAGAGCGCTAAAGACCACCAGCTAGTTTG
CMAF	AACTGCAGATATCAAAATGAGCAGAATCACCAC
CMAR	TTCTCGAGAGCGTCATAAACCTTGAGCTAACCTATGG
GFAF	AACTGCAGATATCAAAATGACAAATTTTGAGAATAAAGAAGTC
GFAR	TTCTCGAGAGCGCTACATTCCGTGCTGAAACAAG

2

The expression constructs were first assembled per gene and then ligated together into the plasmid pRS316 cut with *NotI* and *XhoI*. A and B in opposite direction (adjacent terminator sequences), B and D in opposite direction (adjacent promoter sequences). A physical map of the final plasmid p RS316 GGA is shown in Figure 1 and its sequence is depicted in SEQ ID NO: 22. Other combinations of AraA, AraB and AraD including the respective promoters were obtained as well and corresponding plasmids were constructed.

2.4 Transformation of the host organism and selection of transformants

RN1000 was transformed with plasmids using the 'Gietz method' (Gietz et al., 1992, Nucleic Acids Res. 1992 Mar 25;20(6):1425.). Primary selection of transformants was done on mineral medium (YNB+2% glucose) via uracil complementation. Further selection for transformants containing plasmid pRS316 GGA was done on YNB + 2% L-arabinose. Colonies emerging on plates of the latter medium grew slowly. However, via Colony PCR it was demonstrated that all three ara genes are present in the transformants (Figure 2). The yeast transformant thus obtained was designated Royal Nedalco collection number RN1002 and harbours a plasmid with an expression construct for the expression of araA, araB and araD genes.

2.5 Oxidic growth of the engineered *Saccharomyces cerevisiae* strain RN1002 at the expense of L-arabinose

The purpose of the experiment reported here was to demonstrate that strain RN1002 has the ability to grow at the expense of L-arabinose under oxic (aerobic) conditions.

2.5.1 Media

Yeast nitrogen base (YNB, Difco) buffered with 0.17M KH_2PO_4 and 0.72M K_2HPO_4 at pH 5,5 was used for assessing oxic growth at the expense of arabinose. Incubation were performed in the presence of galactose in order to stimulate cell biomass production. After heat sterilization of the medium for 20 min at 120 °C, the sugars galactose (0.05%) and/or L-arabinose (1%) were added after filter sterilization.

2.5.2 Oxic cultivation

25 ml YNB with 0,5 g/l galactose with or without 10 g/l L-arabinose was inoculated with material derived from a single colony grown on solid medium (YNB agar with 1% L-arabinose and 0,05 % galactose). A culture without any sugar added served as an additional blank. The OD of this culture was below detection level. Cultures were incubated while shaking at 30°C with oxygen from the air allowed to enter into the liquid medium. The concentrations of L-arabinose

1 and galactose were determined at various times. Cell growth was monitored by measuring the
2 OD.

3 2.5.3 Measurement of the optical densities

4 Optical densities were analyzed by an (Perkin Elmer lambda 2S) spectrophotometer at 700 nm.

5 2.5.4 Determination of monomeric sugars

6 Sugar concentrations in filtered supernatants were determined by high-performance anion-
7 exchange. It was performed on a Dionex system equipped with a CarboPac PA-1 column (4 mm
8 ID x 250 mm) in combination with a CarboPac PA guard column (4 mm x 50 mm). For the
9 analysis of both L-arabinose and galactose, an isocratic elution (1 ml/min) of 25 minutes was
10 carried out with water. Each elution was followed by a washing and equilibration step. Detection
11 of the compounds was accomplished by the post-column addition of NaOH to the column eluent
12 to raise the pH (>12) before it entered the PAD (Electrochemical detector ED40, Dionex).

13 2.5.5 Results

14 The results obtained are summarized in Table 2, which demonstrates that strain RN1002 has the
15 ability to metabolize L-arabinose as witnessed by the consumption of L-arabinose and to grow at
16 its expense as demonstrated by the increase in time of OD values of the L-arabinose-containing
17 culture.

18 2.6 Anoxic production of ethanol at the expense of L-arabinose by the engineered 19 *Saccharomyces cerevisiae* strain RN1002

20 The purpose of the experiment reported here was to demonstrate that strain RN1002 has the
21 ability to produce ethanol from L-arabinose under anoxic (anaerobic) conditions.

22 2.6.1 Media

23 For assessing anoxic ethanol production from L-arabinose, a medium containing yeast extract
24 (1% w/w) and peptone (2% w/w) was used. After heat sterilization of the medium for 20 min at
25 120 °C, the sugars galactose(0.5%) and/or arabinose(2%) were added separately after heat
26 sterilization at 110 °C.

27 2.6.2 Anoxic cultivation

28 To prepare a preculture, strain RN1002 was grown at 32 °C and pH5 in a shake flask culture on
29 100 ml medium containing yeast extract with peptone and with addition of the sugars galactose
30 (0.5%) and arabinose (2%). After 70 h incubation, this culture was centrifuged twice and cells
31 were resuspended to an OD of 112. This suspension was used to inoculate four anoxic operated

1 stirred fermenters (BAM fermenters purchased from Halotec) with 1 ml each. The subsequent
2 batch fermentations were performed at 32 °C and the working volumes of the four fermentations
3 used in this study were 150 ml each.

4 2.6.3 Gas analysis

5 The exhaust gas was cooled by a condenser connected to a cryostat set at 4°C. The exhaust gas
6 flow rate was measured with a Brooks Smart mass flow meter, which is calibrated for CO₂ flow.
7 This mass flow meter was located in a valve box interface (Halotec). The valve box contains all
8 the mechanical parts of the system and its purpose is to control the gas flow of each flask and to
9 house the sensors.

10 2.6.4 Measurement of the optical densities

11 Optical densities were analyzed by an (Perkin Elmer lambda 2S) spectrophotometer at 700 nm.

12 2.6.5 Determination of ethanol concentration

13 Ethanol concentrations in filtered supernatants were determined by HPLC analysis with a Bio-
14 rad Aminex HPX-87H column at 65°C. The column was eluted with 0.25 M sulfuric acid at a
15 flow rate of 0.55 ml min⁻¹.

16 2.6.6 Determination of monomeric sugars

17 Sugar concentrations in filtered supernatants were determined by high-performance anion-
18 exchange. It was performed on a Dionex system equipped with a CarboPac PA-1 column (4 mm
19 ID x 250 mm) in combination with a CarboPac PA guard column (4 mm x 50 mm). For the
20 analysis of both L-arabinose and galactose, an isocratic elution (1 ml/min) of 25 minutes was
21 carried out with water. Each elution was followed by a washing and equilibration step. Detection
22 of the compounds was accomplished by the post-column addition of NaOH to the column eluent
23 to raise the pH (>12) before it entered the PAD (Electrochemical detector ED40, Dionex).

24 2.6.7 Results

25 The results obtained are summarized in Table 3 and demonstrate that strain RN1002 has the
26 ability to convert L-arabinose into ethanol.

Table 2 Time course of the optical density (A700) and cumulative L-arabinose and galactose consumption of strain RN1002 during oxic incubations.

Additions to YNB medium (g/l)	Time of incubation(h)	OD (A700)	Arabinose consumed g/l	Galactose consumed g/l
No addition	0	0,00		
	48	0,00		
	144	0,00		
	192	0,00		
	240	0,00		
	312	0,00		
	384	0,00		
Galactose (0.5)	0	0,00		0,00
	48	0,98		
	144	1,24		
	192	1,02		0,50
Galactose (0.5) + Arabinose 10)	0	0,01	0,00	0,00
	48	1,42		
	144	1,51		
	192	1,44	1,14	0,50
	240	1,75		
	312	2,38	3,32	
	384	4,08	5,26	

- 1 Table 3 Time course of the optical density (A700) and cumulative L-arabinose and galactose
 2 consumption of strain RN1002 during anoxic incubations as well as the production of ethanol.
 3

Additions to medium (g/l)	Time of incubation	OD (A700)	Arabinose consumed	Galactose consumed	Ethanol produced
	(h)		g/l	g/l	(g/l)
No addition	0	0,2			0,00
	18	1,5			0,00
	42	1,5			0,00
Arabinose (20)	0	0,2	0,00		0,00
	18	2,0	0,38		0,25
	42	2,3	0,73		0,55
	66	2,3	2,20		0,82
Galactose (5)	0	0,2			0,00
	18	4,2		5,00	2,20
	42	4,0			2,16
Arabinose (20) + Galactose (5)	0	0,2	0,00		0,00
	18	4,4	1,61	4,94	2,48
	42	4,4	2,59	5,01	3,01
	66	4,5	3,95		3,39

Claims

1. A eukaryotic cell comprising nucleotide sequences as defined in (a), (b) and (c), whereby the expression of the nucleotide sequences confers to, or increases in the cell the ability to convert L-arabinose into D-xylulose 5-phosphate, wherein the nucleotide sequences (a), (b) and (c) are:
 - (a) a nucleotide sequence encoding an arabinose isomerase, wherein the nucleotide sequence selected from the group consisting of:
 - (i) a nucleotide sequence encoding an arabinose isomerase comprising an amino acid sequence that has at least 60% sequence identity with at least one of the amino acid sequences of SEQ ID NO's: 1, 2 and 3;
 - (ii) a nucleotide sequence comprising a nucleotide sequence that has at least 70% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 10, 11 and 12;
 - (iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,
 - (iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code,
 - (b) a nucleotide sequence encoding a ribulokinase, wherein the nucleotide sequence is selected from the group consisting of:
 - (i) a nucleotide sequence encoding a ribulokinase comprising an amino acid sequence that has at least 55% sequence identity with at least one of the amino acid sequences of SEQ ID NO's: 4, 5 and 6;
 - (ii) a nucleotide sequence comprising a nucleotide sequence that has at least 65% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 13, 14 and 15;
 - (iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,
 - (iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code; and,
 - (c) a nucleotide sequence encoding a ribulose-5-P-4-epimerase, wherein the nucleotide sequence is selected from the group consisting of:

- (i) a nucleotide sequence encoding a ribulose-5-P-4-epimerase comprising an amino acid sequence that has at least 55% sequence identity with at least one of the amino acid sequences of SEQ ID NO's: 7, 8 and 9;
 - (ii) a nucleotide sequence comprising a nucleotide sequence that has at least 65% sequence identity with at least one of the nucleotide sequences of SEQ ID NO's: 16, 17 and 18;
 - (iii) a nucleotide sequence the complementary strand of which hybridises to a nucleotide sequence of (i) or (ii); and,
 - (iv) a nucleotide sequence the sequences of which differs from the sequence of a nucleotide sequence of (iii) due to the degeneracy of the genetic code.
2. A cell according to claim 1, wherein at least one of the nucleotide sequences as defined in (a), (b) and (c) of claim 1 encodes an amino acid sequence that originate from a bacterial genus selected from the group consisting of *Clavibacter*, *Arthrobacter* and *Gramella*.
3. A cell according to claim 1 or 2, wherein the nucleotide sequence as defined in (a), (b) and (c) of claim 1 encode amino acid sequences that originate from a bacterial genus selected from the group consisting of *Clavibacter michiganensis*, *Arthrobacter aurescens* and *Gramella forsetii*.
4. A cell according to any one of the preceding claims, wherein the cell is a yeast or a filamentous fungus of a genus selected from the group consisting of *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, *Yarrowia*, *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium*, and *Penicillium*.
5. A cell according to claim 4, wherein the cell is a yeast is capable of anaerobic alcoholic fermentation.
6. A cell according to claim 5, wherein the yeast belongs to a species selected from the group consisting of *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* and *Schizosaccharomyces pombe*.

7. A cell according to any one of the preceding claims, wherein the nucleotide sequence as defined in (a), (b) and (c) of claim 1 are operably linked to a promoter that causes sufficient expression of the nucleotide sequences in the cell to confer to the cell the ability to convert L-arabinose into D-xylulose 5-phosphate.
8. A cell according to any one the preceding claims, wherein the cell comprises a genetic modification that increases the flux of the pentose phosphate pathway.
9. A cell according to claim 8, wherein by the genetic modification comprises overexpression of at least one gene of the non-oxidative part pentose phosphate pathway.
10. A cell according to claim 9, wherein the genetic modification comprises overexpression of at least the gene coding for a transaldolase.
11. A cell according to claim 10, wherein the genetic modification comprises overexpression of at least the genes coding for a transketolase and a transaldolase.
12. A cell according to claim 11, wherein the genetic modification comprises overexpression of all four genes encoding a D-ribulose 5-phosphate 3-epimerase, a ribulose 5-phosphate isomerase, a transketolase and a transaldolase.
13. The cell according to any one the preceding claims, whereby the host cell comprises a genetic modification that reduces unspecific aldose reductase activity in the cell.
14. The cell according to claim 13, whereby the genetic modification reduces the expression of, or inactivates a gene encoding an unspecific aldose reductase.
15. The cell according to claim 14, whereby the gene is inactivated by deletion of at least part of the gene or by disruption of the gene.

16. The cell according to any one of claims 13 - 15, whereby the expression of each gene in the cell that encodes an unspecific aldose reductase that is capable of reducing an aldopentose is reduced or inactivated.
17. A cell according to any one the preceding claims, wherein the cell exhibits the ability to directly isomerise xylose into xylulose.
18. A cell according to claim 17, whereby the cell further comprises a genetic modification that increases the specific xylulose kinase activity.
19. A cell according to claim 18, whereby the genetic modification comprises overexpression of a gene encoding a xylulose kinase.
20. A cell according to claim 18 or 19, whereby the xylulose kinase gene that is overexpressed is endogenous to the cell.
21. A cell according to any one of claims 1 - 7, wherein the cell comprises at least one further genetic modification that result in a characteristic selected from the group consisting of:
- (a) increase transport of at least one of xylose and arabinose into the host cell;
 - (b) decreased sensitivity to catabolite repression;
 - (c) increased tolerance to ethanol, osmolarity or organic acids; and,
 - (d) reduced production of by-products.
22. A cell according to any one of the preceding claims, whereby the cell expresses one or more enzymes that confer to the host cell the ability to produce at least one fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins.

23. A eukaryotic cell comprising nucleotide sequences as defined in (a'), (b') and (c'), whereby the expression of the nucleotide sequences confers to, or increases in the cell the ability to convert L-arabinose into D-xylulose 5-phosphate, wherein the nucleotide sequences (a'), (b') and (c') are:

- (a') a nucleotide sequence encoding an arabinose isomerase;
- (b') a nucleotide sequence encoding a xylulose kinases; and,
- (c') a nucleotide sequence encoding a ribulose-5-P-4-epimerase.

24. A process for producing a fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins, whereby the process comprises the steps of:

- (a) fermenting a medium containing a source of arabinose with cell as defined in any one of claims 1 - 23, whereby the host cell ferments arabinose to the fermentation product, and optionally,
- (b) recovery of the fermentation product.

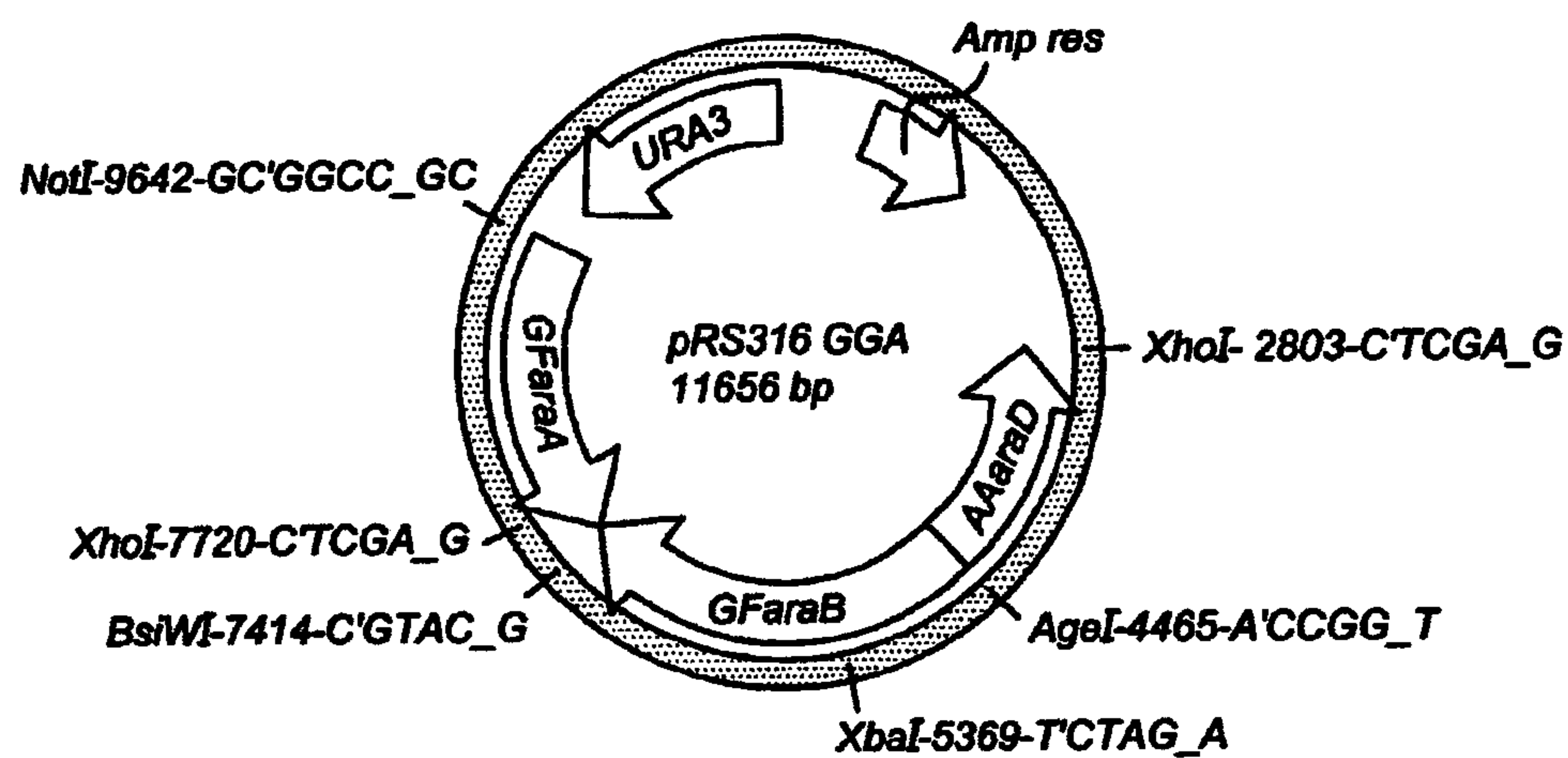
25. A process for producing a fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, citric acid, amino acids, 1,3-propane-diol, ethylene, glycerol, β -lactam antibiotics and cephalosporins, whereby the process comprises the steps of:

- (a) fermenting a medium containing at least one of a source of xylose and a source of arabinose with a cell as defined in any one of claims 17 - 23, whereby the cell ferments at least one of xylose and arabinose to the fermentation product, and optionally,
- (b) recovery of the fermentation product.

26. A process according to claim 24 or 25, whereby the medium also contains a source of glucose.

27. The process according to any one of claim 24 - 26, wherein the fermentation product is ethanol.

28. The process according to claim 27, whereby the volumetric ethanol productivity is at least 0.5 g ethanol per litre per hour.
29. The process according to claim 27 or 28, whereby the ethanol yield is at least 50 %.
30. The process according to any one of claims 24 to 29, wherein the process is anaerobic.

Fig 1**Fig 2**