



(43) International Publication Date
20 September 2012 (20.09.2012)

WIPO | PCT

(10) International Publication Number
WO 2012/125027 A1

(51) International Patent Classification:

C12N 9/00 (2006.01) C12N 9/90 (2006.01)
C12N 9/12 (2006.01) C12P 7/06 (2006.01)
C12N 9/88 (2006.01)

(21) International Application Number:

PCT/NL2012/050151

(22) International Filing Date:

13 March 2012 (13.03.2012)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

61/452,366 14 March 2011 (14.03.2011) US

(71) Applicant (for all designated States except US): **DSM IP Assets B.V.** [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DE BONT, Johannes Adrianus Maria** [NL/NL]; Lawickse Allee 68, NL-6707 AK Wageningen (NL). **TEUNISSEN, Aloysius Wilhelmus Rudolphus Hubertus** [NL/NL]; Tjaskerlaan 28, NL-3052 HP Rotterdam (NL).

(74) Agent: **PALLARD, Caroline**; J.W. Frisolaan 13, NL-2517 JS The Hague (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, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

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

Published:

- with international search report (Art. 21(3))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: YEAST STRAINS THAT FERMENT URONIC ACIDS

(57) Abstract: The present invention relates to yeast strains that have been engineered to produce fermentation products from uronic acids such as galacturonic acid or glucuronic acid. For this purpose the yeast strain have been modified by introduction of a) genes coding for enzymes of a plant salvage pathway with the ability to convert a D uronic acid into a pentose- 1 -phosphate and carbon dioxide; b) a gene coding for enzyme with the ability to convert a pentose- 5 -phosphate into a pentose; and, c) genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate. The invention further relates to the use of such engineered yeast strains for the production of fermentation products and to process for the production of a fermentation product. The process comprises the steps of a) fermenting a medium with a yeast cell of the invention, whereby the medium contains or is fed with a source of a uronic acid and whereby the yeast cell ferments the uronic acid to the fermentation product; and optionally, b) recovery of the fermentation product. A preferred fermentation product is ethanol.



WO 2012/125027 A1

Yeast strains that ferment uronic acids

Field of the invention

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

Background of the invention

There is a need for bioproduction of fuels or chemicals from second generation feedstocks. Second generation feedstocks may contain, apart from cellulose, also hemicelluloses and/or pectin. Hemicellulose contains mainly the pentose sugars xylose and arabinose, but also glucuronate. Pectin contains apart from various sugars, vast quantities of galacturonate. Pectin-rich residues become available as by-product when sugar is extracted from e.g. sugar beet or when juices are produced from citrus fruits. The uronic acids glucuronate and galacturonate are thus available as important constituents in second generation feedstocks, however, they cannot be metabolized by industrially important yeasts including *Saccharomyces cerevisiae* e.g. for the production of fuel ethanol.

No suggestions have been reported for engineering *S. cerevisiae* to introduce the ability to ferment glucuronate into ethanol. For galacturonate, however, three routes have been suggested that potentially may result in the formation of ethanol from galacturonate in *S. cerevisiae* (Richard and Hilditch, 2009, Appl Microbiol Biotechnol 82:597–604; Hilditch, Thesis 2010, VTT publications 739, ISBN 978-951-38-7398-1): two bacterial routes and one fungal route. However, for none of these three routes, it has been demonstrated that ethanol is actually formed by *S. cerevisiae* that is transformed with the respective sets of genes. Furthermore, even if the expression of these suggested pathways were successful, then they would result in biological systems

that are not compatible with large-scale ethanol production under anoxic conditions because in all three cases input of reducing equivalents would be required in the form of NAD(P)H in order to counterbalance the relatively oxidized status of galacturonate as compared to sugars such as glucose or xylose.

5 There is therefore still a need in the art for yeasts engineered to produce fermentation products such as ethanol from uronic acids such as galacturonic acid or glucuronic acid, whereby preferably the yeast uses a redox neutral pathway. It is an object of the invention to provide such yeasts, as well as processes wherein they are used to produce fermentation products such as ethanol from these uronic acids.

10

Summary of the invention

In a first embodiment the invention relates to a yeast cell comprising: a) genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide; b) a gene coding for enzyme with the ability to convert a pentose-1P into a pentose; and, c) genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate.

In a second embodiment the invention relates to a yeast cell according to the first embodiment, wherein: a) the genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide are genes coding for enzymes with the ability to convert at least one of galacturonic acid and glucuronic acid into carbon dioxide and at least one of xylose-1-phosphate and arabinose-1-phosphate; b) the gene coding for enzyme with the ability to convert a pentose-1P into a pentose is a gene coding for an enzyme with the ability to convert at least one of xylose-1-phosphate into xylose and arabinose-1-phosphate into arabinose; and, c) the genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate are selected from the group consisting of genes coding for enzymes with the ability to i) directly isomerise xylose into xylulose, and, ii) convert L-arabinose into D-xylulose 5-phosphate.

30 In a third embodiment the invention relates to a yeast cell according to the first or second embodiment, wherein the yeast cell has the ability of metabolizing glucuronic acid, and wherein the yeast cell comprises genes coding for: a) a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a

broader substrate specificity; b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64); c) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and d) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and, ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).

In a fourth embodiment the invention relates to a yeast cell according to the first or second embodiment, wherein the yeast cell has the ability of metabolizing galacturonic acid, and wherein the yeast cell comprises genes coding for: a) a galacturonokinase (EC 2.7.1.44); b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64); c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6); d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and e) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and, ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).

In a fifth embodiment the invention relates to a yeast cell according to the first or second embodiment, wherein the yeast cell has the ability of metabolizing galacturonic acid, and wherein the yeast cell comprises genes coding for: a) a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a broader substrate specificity; b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64); c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6); d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and e) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and, ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).

In a sixth embodiment the invention relates to a yeast cell according to any one of the third to fifth embodiments, wherein at least one of: a) the gene coding for an enzyme with galacturokinase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 1; b) the gene coding for an enzyme with glucuronokinase activity comprises a

nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 2; c) the gene coding for an enzyme with UTP-monosaccharide-1-phosphate uridylyltransferase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with at least one of SEQ ID NO: 3 and 4; d) the gene coding for an enzyme with UDP-D-glucuronate 4-epimerase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 5; e) the gene coding for an enzyme with UDP-glucuronic acid decarboxylase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 6; f) the gene coding for an enzyme with UDP-D-xylose 4-epimerase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 7; g) the gene coding for an enzyme with arabinose kinase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 8; and, h) the gene coding for an enzyme with sugar phosphatase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 9.

In a seventh embodiment the invention relates to a yeast cell according to any one of the preceding embodiments, wherein the cell is a yeast cell selected from the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, *Brettanomyces*, and *Yarrowia*. In particular in this embodiment, the cell is a yeast cell selected from the species *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* and *Schizosaccharomyces pombe*.

In a eighth embodiment the invention relates to the use of a yeast cell according to any of the preceding embodiments for the preparation of a fermentation product. Preferably in this embodiment, the fermentation product is selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propanediol, butanols and isoprenoid-derived products.

In a ninth embodiment the invention relates to a process for producing a fermentation product, whereby the process comprises the steps of: a) fermenting a medium with a yeast cell as defined in any one of claim 1 - 8, whereby the medium contains or is fed with a source of a uronic acid and whereby the yeast cell ferments the

uronic acid to the fermentation product; and optionally, b) recovery of the fermentation product. Preferably in this embodiment, the fermentation product is selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, butanols and isoprenoid-derived products.

5 In a tenth embodiment the invention relates to a process according to the ninth embodiment, wherein the uronic acid is at least one of galacturonic acid and glucuronic acid.

In a eleventh embodiment the invention relates to a process according to the ninth or tenth embodiment, wherein the medium further contains or is fed with a source
10 of at least one of a hexose, a pentose, acetic acid and glycerol.

In a twelfth embodiment the invention relates to a process according to any one of the ninth to eleventh embodiments, wherein the medium contains or is fed with at least one of a hydrolyzed pectin-rich residue and hydrolyzed lignocellulosic biomass.

In a thirteenth embodiment the invention relates to a process according to any
15 one of the ninth to twelfth embodiments, wherein the yeast cell ferments under anaerobic conditions.

In a fourteenth embodiment the invention relates to a process according to any one of the ninth to thirteenth embodiments, wherein the fermentation product is ethanol.

20

Description of the invention

Definitions

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)
25 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
30 sequence of a second polypeptide. "Identity" and "similarity" can be readily calculated by known methods. The terms "sequence identity" or "sequence similarity" means that two (poly)peptide or two nucleotide sequences, when optimally aligned, preferably over the entire length (of at least the shortest sequence in the comparison) and

maximizing the number of matches and minimizes the number of gaps such as by the programs ClustalW (1.83), 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). A preferred multiple alignment program for aligning protein sequences of the invention is ClustalW (1.83) using a Blosum matrix and default settings (Gap opening penalty: 10; Gap extension penalty: 0.05). 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 2.10.0-0.8). Alternatively percent similarity or identity may be determined by searching against databases such as FASTA, BLAST, etc.

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

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

arginine, and histidine; and a group of amino acids having sulphur-containing side chains is cysteine and methionine. Preferred conservative amino acids substitution groups are: valine-leucine-isoleucine, phenylalanine-tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine. Substitutional variants of the amino acid sequence disclosed herein are those in which at least one residue in the disclosed sequences has been removed and a different residue inserted in its place. Preferably, the amino acid change is conservative. Preferred conservative substitutions for each of the naturally occurring amino acids are as follows: Ala to Ser; Arg to Lys; Asn to Gln or His; Asp to Glu; Cys to Ser or Ala; Gln to Asn; Glu to Asp; Gly to Pro; His to Asn or Gln; Ile to Leu or Val; Leu to Ile or Val; Lys to Arg; Gln to Glu or Asn; Met to Leu or Ile; Phe to Met, Leu or Tyr; Ser to Thr; Thr to Ser; Trp to Tyr; Tyr to Trp or Phe; and, Val to Ile or Leu.

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

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

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

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

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

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

used interchangeably with marker, although it is mainly used to refer to visible markers, such as green fluorescent protein (GFP). Selectable markers may be dominant or recessive or bidirectional.

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

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

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

The term “gene” means a DNA fragment comprising a region (transcribed region), which is transcribed into an RNA molecule (e.g. an mRNA) in a cell, operably linked to suitable regulatory regions (e.g. a promoter). A gene will usually comprise several operably linked fragments, such as a promoter, a 5' leader sequence, a coding region and a 3' nontranslated sequence (3' end) comprising a polyadenylation site. The coding region of the gene may still comprise intron but usually will comprise an uninterrupted open reading frame, such as e.g. a cDNA. “Expression of a gene” refers to the process wherein a DNA region which is operably linked to appropriate regulatory regions, particularly a promoter, is transcribed into an RNA, which is biologically active, i.e. which is capable of being translated into a biologically active protein or peptide.

The term “homologous” when used to indicate the relation between a given (recombinant) nucleic acid or polypeptide molecule and a given host organism or host cell, is understood to mean that in nature the nucleic acid or polypeptide molecule is produced by a host cell or organisms of the same species, preferably of the same

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

The terms "heterologous" and "exogenous" when used with respect to a nucleic acid (DNA or RNA) or protein refers to a nucleic acid or protein that does not occur naturally as part of the organism, cell, genome or DNA or RNA sequence in which it is present, or that is found in a cell or location or locations in the genome or DNA or RNA sequence that differ from that in which it is found in nature. Heterologous and exogenous nucleic acids or proteins are not endogenous to the cell into which it is introduced, but have been obtained from another cell or synthetically or recombinantly produced. Generally, though not necessarily, such nucleic acids encode proteins, i.e. exogenous proteins, that are not normally produced by the cell in which the DNA is transcribed or expressed. Similarly exogenous RNA encodes for proteins not normally expressed in the cell in which the exogenous RNA is present. Heterologous or exogenous nucleic acids and proteins may also be referred to as foreign nucleic acids or proteins. Any nucleic acid or protein that one of skill in the art would recognize as foreign to the cell in which it is expressed is herein encompassed by the term heterologous or exogenous nucleic acid or protein. The terms heterologous and exogenous also apply to non-natural combinations of nucleic acid or amino acid sequences, i.e. combinations where at least two of the combined sequences are foreign with respect to each other.

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

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

When an enzyme is mentioned with reference to an enzyme class (EC), the enzyme class is a class wherein the enzyme is classified or may be classified, on the basis of the Enzyme Nomenclature provided by the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), which nomenclature may be found at <http://www.chem.qmul.ac.uk/iubmb/enzyme/>. Other suitable enzymes that have not (yet) been classified in a specified class but may be classified as such, are meant to be included.

If referred herein to a protein or a nucleic acid sequence, such as a gene, by reference to a accession number, this number in particular is used to refer to a protein or nucleic acid sequence (gene) having a sequence as can be found in the GenBank database (found at World Wide Web URL via www.ncbi.nlm.nih.gov/) and refers to the version of that sequence as available on 14 March 2011, unless specified otherwise.

25

Detailed description of the invention

In a first aspect the invention relates to a yeast cell that has the ability of metabolizing uronic acids. A uronic acid is herein understood to refer to a sugar acid with both a carbonyl and a carboxylic acid function. Preferred uronic acids in the context of the invention are glucuronic acid and galacturonic acid. It is further understood herein that the terms “uronic acid”, “glucuronic acid” and “galacturonic acid” as used herein are interchangeable with the term “uronate”, “glucuronate” and “galacturonate”, respectively. A yeast cell that has the ability of metabolizing uronic

acids preferably is a yeast that can grow at the expense of the uronic acid, more preferably the yeast cell has the ability to grow on the uronic acid as sole carbon source. The yeast cell preferably has the ability to grow at the expense of the uronic acid or to grow on the uronic acid as sole carbon source under anaerobic conditions.

- 5 More preferably the yeast cell has the ability to ferment the uronic acid under anaerobic conditions to fermentation products such as ethanol or lactic acid.

A yeast cell of the invention that has the ability of metabolizing a uronic acid preferably is a yeast that has been engineered to have this metabolic ability. Preferably the yeast cell has been engineered by introducing one or more expression constructs for
10 the expression of one or more nucleotide sequences (genes) coding for enzymes that are required for metabolizing a uronic acid, i.e. the cell has been transformed with the corresponding expression constructs.

Thus, in one embodiment, a yeast cell of the invention comprises: a) genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic
15 acid into a pentose-1-phosphate and carbon dioxide; b) a gene coding for enzyme with the ability to convert a pentose-1P into a pentose; and, c) genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate. The genes in a) , b) and c) will usually be genes that are exogenous to the yeast cell but they can also be endogenous to the yeast cell.

20 Genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide are herein understood as genes coding for enzymes that generate nucleotide sugars from uronic acids such as glucuronic acid and galacturonic acid and convert these nucleotide sugars into carbon dioxide and pentose-1-phosphates such as xylose-1-phosphate and arabinose-1-
25 phosphate as e.g. described in Kotake et al. (2010, Plant Biotechnol. 27: 231–236). The relevant parts of the plant salvage pathway are also depicted in Figure 1 wherein the relevant enzyme activities are indicated by the corresponding EC numbers. It is understood herein that an enzyme of a plant salvage pathway for use in the present invention does not necessarily has to be a plant enzyme as the same enzymatic activity
30 may also be present in and obtained from organisms other than plants, including e.g. bacteria, fungi and animals. Preferably, in the yeast cell of the invention the genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide are genes coding for enzymes with

the ability to convert at least one of galacturonic acid and glucuronic acid into carbon dioxide and at least one of xylose-1-phosphate and arabinose-1-phosphate. Genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide are further specified herein below
5 under “Enzyme activities introduced into the yeast cell of the invention”, in paragraphs A) to E).

The gene coding for an enzyme with the ability to convert a pentose-1-phosphate into a pentose preferably is a gene coding for an enzyme with the ability to convert at least one of xylose-1-phosphate into xylose and arabinose-1-phosphate into arabinose.
10 Genes coding for enzymes with the ability to convert at least one of xylose-1-phosphate into xylose and arabinose-1-phosphate into arabinose are further specified herein below under “Enzyme activities introduced into the yeast cell of the invention”, in paragraph F) and include e.g. genes encoding arabinose kinases (EC 2.7.1.46) and genes encoding sugar phosphatases (EC 3.1.3.23) or sugar-1-phosphatases (EC 3.1.3.10).

15 The genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate are selected from the group consisting of genes coding for enzymes with the ability to i) directly isomerise xylose into xylulose, and, ii) convert L-arabinose into D-xylulose 5-phosphate.

In one embodiment the yeast cell is a cell that has the ability of metabolizing
20 glucuronic acid. Such a yeast cell will at least comprise genes coding for a) a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a broader substrate specificity as described in A) herein below; b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64) as described in B) herein below; c) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35) as described in D)
25 herein below; and d) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) as described in F) herein below and a xylose isomerase (EC 5.3.1.5) as described in G) herein below; and, ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5) as described in E), an arabinose kinase (EC 2.7.1.46) as described in F) herein below, and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a
30 L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) as described in G) herein below. The expression of these genes in the yeast cell confers to the yeast cell the ability of metabolizing glucuronic acid.

In another embodiment the yeast cell is a cell that has the ability of metabolizing galacturonic acid. Such a yeast cell will at least comprise genes coding for a) a galacturonokinase (EC 2.7.1.44) as described in A) herein below; b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64) as described in B) herein below; c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6) as described in C) herein below; d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35) as described in D) herein below; and, e) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) as described in F) herein below and a xylose isomerase (EC 5.3.1.5) as described in G) herein below; and ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5) as described in E), an arabinose kinase (EC 2.7.1.46) as described in F) herein below, and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) as described in G) herein below. The expression of these genes in the yeast cell confers to the yeast cell the ability of metabolizing galacturonic acid.

In preferred embodiment the yeast cell is a cell that has the ability of metabolizing both galacturonic acid and glucuronic acid. Such a yeast cell will at least comprise genes coding for a) a galacturonokinase (EC 2.7.1.44) and a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a broader substrate specificity as described in A) herein below; b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64) as described in B) herein below; c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6) as described in C) herein below; d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35) as described in D) herein below; and, e) at least one of: i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) as described in F) herein below and a xylose isomerase (EC 5.3.1.5) as described in G) herein below; and ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5) as described in E), an arabinose kinase (EC 2.7.1.46) as described in F) herein below, and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4) as described in G) herein below. The expression of these genes in the yeast cell confers to the yeast cell the ability of metabolizing both galacturonic acid and glucuronic acid.

Enzyme activities introduced into the yeast cell of the invention

A) Uronokinases (2.7.1.44 and 2.7.1.43)

A yeast cell of the invention preferably comprises a gene coding for a enzyme with the ability to convert a uronic acid into a 1-P-uronic acid. Preferably the enzyme is a galacturonokinase (EC 2.7.1.44) or a glucuronokinase (EC 2.7.1.43). A galacturonokinase is herein understood as an enzyme that catalyses the ATP-dependent conversion of α -D-galacturonic acid to α -D-galacturonic acid-1-phosphate. Likewise, a glucuronokinase is herein understood as an enzyme that catalyzes the ATP-dependent conversion of α -D-glucuronic acid to α -D-glucuronic acid-1-phosphate. A preferred gene coding for a uronokinase is an exogenous gene, preferably a gene encoding an enzyme of plant origin or a variant thereof.

10 An example of a suitable gene coding for a uronokinase is e.g. the GalAK gene (cDNA) encoding the galacturokinase of *Arabidopsis thaliana* as described by Yang et al. (2009, J. Biol Chem. 284: 21526–21535; accession no.: At3g10700). A gene coding for an enzyme with galacturokinase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 15 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 1. A gene coding for an enzyme with galacturokinase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 1. Preferably the amino acid sequence has no more than 210, 200, 150, 100, 75, 50, 40, 30, 20, 10 or 20 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 1. Other suitable and available galacturokinases are from *Oryza sativa*, *Populus trichocarpa*, *Vitis vinifera*, *Sorghum bicolor* and *Selaginella moellendorffii* (accession no.'s: *Oryza sativa* GalAK: Os4g51880; *Populus trichocarpa* GalAK: JGI:427630; *Vitis vinifera* GalAK: GSVIVT00007137001; *Sorghum bicolor* GalAK: Sb06g027910; 25 *Selaginella moellendorffii* GalAK: SmGalAK, JGI:82393). The amino acid sequences of these enzymes are available in public databases and can be used by the skilled person to design codon-optimized nucleotide sequences coding for the corresponding enzyme. A preferred codon-optimized nucleotide sequences coding for a galacturokinase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence 30 identity with the open reading frame in SEQ ID NO: 10 (see Table 1).

A particularly preferred uronokinase is a uronokinase having substrate specificity for both glucuronic acid and galacturonic acid. An example thereof is a galacturokinase

engineered for broader substrate specificity, such as e.g. the GalAK^{Y250F} variant described by Yang et al. (2009, *supra*), which also phosphorylates α -D-glucuronic acid.

An example of a suitable gene coding for a uronokinase is e.g. the GlcAK gene (cDNA) encoding the glucurokinase of *Arabidopsis thaliana* as described by Pieslinger et al. (2010, J. Biol Chem. 285: 2902–2910; accession no.'s: At3g01640; At5g14470).
 5 A gene coding for an enzyme with glucurokinase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 2. A gene coding for an enzyme with glucurokinase activity may also comprises a nucleotide
 10 sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 2. Preferably the amino acid sequence has no more than 180, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 2. Other suitable and available glucurokinases are from *Zea mays*, *Oryza sativa*,
 15 *Populus trichocarpa*, *Vitis vinifera*, *Sorghum bicolor* and *Physcomitrella patens* (accession no.'s: Os11g11060 and Os11g0217300 from rice; *Vitis vinifera*, CAO14683; *Sorghum bicolor*, EES15489; *Zea mays*, ACG36196; *Populus trichocarpa*, XM_002298517; *Physcomitrella patens* subsp. *patens*, EDQ81053). The amino acid sequences of these enzymes are available in public databases and can be used by the
 20 skilled person to design codon-optimized nucleotide sequences coding for the corresponding enzyme. A preferred codon-optimized nucleotide sequences coding for a glucurokinase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in SEQ ID NO: 11 (see Table 1).

A yeast cell of the invention with the ability to grow on galacturonic acid
 25 contains at least a gene coding for a galacturonokinase. A yeast cell of the invention with the ability to grow glucuronic acid contains at least a gene coding for a glucuronokinase. A yeast cell of the invention with the ability to grow on both galacturonic acid and glucuronic acid, contains either a gene coding for a galacturokinase engineered for broader substrate specificity, such as e.g. the
 30 GalAK^{Y250F} variant described above, or it contains a gene coding for a galacturokinase and a gene coding for a glucuronokinase.

B) Uridylyltransferase/UDP-sugar pyrophosphorylase (EC 2.7.7.64)

A yeast cell of the invention further preferably comprises a gene coding for a enzyme with the ability to convert a D-uronate-1P into UDP-D-uronate. The enzyme preferably has to ability to convert D-galacturonate-1P into UDP-D-galacturonate and/or to convert D-glucuronate-1P into UDP-D-glucuronate. A suitable enzyme for catalyzing this reaction is an UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64), also referred to as a UDP-sugar pyrophosphorylase. A preferred gene coding for a uridylyltransferase is an exogenous gene, preferably a gene encoding an enzyme of plant origin or a variant thereof. Plant UDP-sugar pyrophosphorylase/uridylyltransferase (EC 2.7.7.64) are a novel class of nucleotide sugar pyrophosphorylases with broad substrate specificity toward monosaccharide 1-Ps, which at least catalyze the formation of UDP-D-galacturonate and UDP-D-glucuronate from the respective 1-P uronic acids.

Examples of a suitable genes coding for a uridylyltransferase are e.g. the PsUSP gene (cDNA) of *Pisum sativum* as described by Kotake et al. (2004, J. Biol. Chem. 279: 45728–45736; accession no.: AB178642) or the AtUSP gene (cDNA) encoding the UDP-sugar pyrophosphorylase of *Arabidopsis thaliana* with accession no.: At5g52560 as described by Litterer et al. (2006, Plant Physiol. Biochem. 44: 171–180) and by Kotake et al. (2007, Biosci. Biotechnol. Biochem. 71: 761–771). A gene coding for an enzyme with uridylyltransferase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with at least one of SEQ ID NO: 3 and 4. A gene coding for an enzyme with uridylyltransferase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of at least one of SEQ ID NO: 3 and 4. Preferably the amino acid sequence has no more than 300, 250, 200, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to at least one of SEQ ID NO: 3 and 4. Another suitable and available uridylyltransferase is from *Oryza sativa* (accession no. rice: AK064009). The amino acid sequences of this enzyme is available in public databases and can be used by the skilled person to design codon-optimized nucleotide sequences coding for the corresponding enzyme. A preferred codon-optimized nucleotide sequences coding for a uridylyltransferase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99%

sequence identity with the open reading frames in at least one of SEQ ID NO's: 12 and 13 (see Table 1).

C) UDP-glucuronate-4-epimerase (EC 5.1.3.6)

A yeast cell of the invention further preferably comprises a gene coding for a
5 enzyme with the ability to epimerize UDP-D-glucuronate into UDP-D-galacturonate
and vice versa. A suitable enzyme for catalyzing this reaction is a UDP-D-glucuronate
4-epimerase (EC 5.1.3.6). A preferred gene coding for a UDP-D-glucuronate 4-
epimerase is an exogenous gene, more preferably a gene encoding an enzyme of plant
origin or a variant thereof. However, other preferred genes coding for the UDP-D-
10 glucuronate 4-epimerase are genes encoding enzymes of bacterial origin or variants
thereof, such as e.g. the Cap1J gene from *S. pneumoniae* as described by Muñoz et al.
(1999, Mol. Microbiol. 31: 703–713).

An example of a suitable gene coding for a UDP-D-glucuronate 4-epimerase is
e.g. the GAE1 gene (cDNA) encoding the UDP-D-glucuronate 4-epimerase of
15 *Arabidopsis thaliana* as described by Mølhøj et al. (2004, Plant Physiol. 135: 1221–
1230; accession no.: At4g30440). A gene coding for an enzyme with UDP-D-
glucuronate 4-epimerase activity preferably comprises a nucleotide sequence coding for
an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99%
amino acid sequence identity with SEQ ID NO: 5. A gene coding for an enzyme with
20 UDP-D-glucuronate 4-epimerase activity may also comprises a nucleotide sequence
coding for an amino acid sequence having one or several substitutions, insertions
and/or deletions as compared to the amino acid sequence of SEQ ID NO: 5. Preferably
the amino acid sequence has no more than 200, 150, 100, 75, 50, 40, 30, 20, 10 or 5
amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 5.

25 A preferred codon-optimized nucleotide sequences coding for a UDP-D-
glucuronate 4-epimerase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99%
sequence identity with the open reading frame in SEQ ID NO: 14 (see Table 1).

D) UDP-uronic acid decarboxylase (EC 4.1.1.35)

A yeast cell of the invention further preferably comprises a gene coding for a
30 UDP-uronic acid decarboxylase, e.g. an enzyme with the ability to convert UDP-D-
glucuronate into UDP-D-xylose and CO₂. A suitable enzyme for catalyzing this
reaction is an UDP-glucuronic acid decarboxylase (EC 4.1.1.35), also referred to as a
UDP glucuronate carboxy-lyase or a UDP- α -D-xylopyranose synthase. A preferred

gene coding for a UDP-glucuronic acid decarboxylase is an exogenous gene, more preferably a gene encoding an enzyme of plant origin or a variant thereof.

Examples of a suitable genes coding for a UDP-glucuronic acid decarboxylase are e.g. the UXS gene (cDNA) of *Oryza sativa* as described by Suzuki et al. (2003, J. Exp. Botany, 54: 1997-1999; accession no.: AB079064), a UDP-glucuronic acid decarboxylase from pea with accession no: BAB40967, or the AtUXS3 gene (cDNA) encoding a UDP-glucuronic acid decarboxylase of *Arabidopsis thaliana* with Genbank accession no.: AF387789 or NP_200737 as described by Oka and Jigami (2006, FEBS J. 273: 2645–2657). A gene coding for an enzyme with UDP-glucuronic acid decarboxylase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 6. A gene coding for an enzyme with UDP-glucuronic acid decarboxylase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 6. Preferably the amino acid sequence has no more than 170, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 6.

Further suitable and available genes coding for a UDP-glucuronic acid decarboxylase include e.g. is the genes listed in the GenBank database under the following accession numbers: BAB84334 (rice), CAC14890 (reed), CAB61752, (chick-pea), AAK59981 (Filobasidiella), AAM45939 (rat), AAK85410 (mouse), CAC48840 (rhizobium), and BAA18111 (cyanobacterium).f from *Oryza sativa* (accession no. rice: AK064009). The amino acid sequences of these enzymes are available in public databases and can be used by the skilled person to design codon-optimized nucleotide sequences coding for the corresponding enzyme. A preferred codon-optimized nucleotide sequences coding for a UDP-glucuronic acid decarboxylase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in SEQ ID NO: 15 (see Table 1).

Another suitable UDP-uronic acid decarboxylase for expression in a yeast cell of the invention is an enzyme with the ability to convert UDP-D-galacturonate into UDP-D-arabinose and CO₂. A suitable enzyme for catalyzing this reaction is an UDP-galacturonate decarboxylase (EC 4.1.1.67), also referred to as a UDP galacturonate carboxy-lyase. A preferred gene coding for a UDP-galacturonic acid decarboxylase is

an exogenous gene, more preferably a gene encoding an enzyme of plant origin or a variant thereof. Expression of a gene coding for a UDP-galacturonic acid decarboxylase in a yeast cell of the invention is straightforward once the sequence of a UDP-galacturonic acid decarboxylase becomes available in the sequence databases.

5 E) UDP-D-xylose 4-epimerase (EC 5.1.3.5)

A yeast cell of the invention further preferably comprises a gene coding for a enzyme with the ability to convert UDP-D-xylose into UDP-L-arabinose and *vice versa*. A suitable enzyme for catalyzing this reaction is a UDP-D-xylose 4-epimerase (EC 5.1.3.5), also referred to as a UDP-arabinose 4-epimerase. A preferred gene coding
10 for a UDP-D-xylose 4-epimerase is an exogenous gene, more preferably a gene encoding an enzyme of plant origin or a variant thereof.

Examples of a suitable genes coding for a UDP-D-xylose 4-epimerase are e.g. the barley (*Hordeum vulgare*) UXE genes (cDNAs), designated HvUXE1, HvUXE2, and HvUXE3 as described by Zhang et al. (2010, Plant Physiol. 153: 555–568; accession
15 no.'s in the GenBank/EMBL databases are: DQ336893, DQ336894, and DQ336895 for HvUXE1, HvUXE2, and HvUXE3, respectively). A gene coding for an enzyme with UDP-D-xylose 4-epimerase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 7. A gene coding for an enzyme with
20 UDP-D-xylose 4-epimerase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 7. Preferably the amino acid sequence has no more than 200, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 7.

25 Further suitable and available genes coding for a UDP-D-xylose 4-epimerase include e.g. genes (cDNAs) coding for the *Arabidopsis* UXEs, having GenBank accession no.'s as follows: AtUXE1, Q9SA77; AtUXE2, O64749; AtUXE3, Q9SUN3; and AtUXE4, Q9FI17. The amino acid sequences of these enzymes are available in public databases and can be used by the skilled person to design codon-optimized
30 nucleotide sequences coding for the corresponding enzyme. A preferred codon-optimized nucleotide sequences coding for a UDP-D-xylose 4-epimerase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in SEQ ID NO: 16 (see Table 1).

F) Enzymes for converting a pentose-1P into a pentose

A yeast cell of the invention further preferably comprises a gene coding for an enzyme with the ability to convert a pentose-1P into a pentose. Preferably the enzyme has the ability to convert at least one of xylose-1P and arabinose-1P into at least one of xylose and arabinose, respectively. At least two different types of enzymes can be used to catalyze this conversion: 1) an arabinose kinase and 2) a sugar phosphatase.

Thus, in a preferred embodiment the yeast cell of the invention comprises a gene coding for an arabinose kinase (EC 2.7.1.46). Arabinose kinases catalyze the reaction $\text{ATP} + \text{L-arabinose} \leftrightarrow \text{ADP} + \beta\text{-L-arabinose-1-phosphate}$. Even though the reaction is favored in the direction from arabinose to arabinose 1-phosphate, the reaction is sufficiently reversible in a dynamic metabolic environment within a yeast cell for forming arabinose from arabinose-1P. A preferred gene coding for an arabinose kinase is an exogenous gene, more preferably the exogenous gene is a gene encoding an enzyme of plant origin or a variant thereof.

An example of a suitable gene coding for an arabinose kinase is e.g. the *Arabidopsis* ARA1 gene (cDNAs) as described by Sherson et al. (1999, Plant Mol. Biol. 39: 1003–1012; accession no.: NP_193348). A gene coding for an enzyme with arabinose kinase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 8. A gene coding for an enzyme with arabinose kinase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 8. Preferably the amino acid sequence has no more than 500, 400, 300, 250, 200, 150, 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 8. A preferred codon-optimized nucleotide sequences coding for a arabinose kinase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in SEQ ID NO: 17 (see Table 1).

In one embodiment, a N-terminally truncated version of the enzyme with arabinose kinase activity is used. An N-terminally truncated version of the enzyme with arabinose kinase preferably at least comprises amino acids 343 - 1039, 486 - 1039, or 542 - 1039 of SEQ ID NO: 8, or an amino acids sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with amino acids 343 -

1039, 486 - 1039, or 542 - 1039 of SEQ ID NO: 8. In a preferred N-terminally truncated version of the enzyme with arabinose kinase the amino terminus of the enzyme starts at position 2, 10, 20, 50, 100, 200, 343, 486, 542, 543, 544, 545, 546, 547, 548, 549, or 550. One example of an N-terminally truncated enzyme with
5 arabinose kinase is an enzyme having the amino acid sequence of SEQ ID NO: 47.

Thus, in a preferred embodiment the yeast cell of the invention comprises a gene coding for an xylose kinase. Xylose kinases catalyze the reaction $\text{ATP} + \text{xylose} \leftrightarrow \text{ADP} + \text{xylose-1-phosphate}$. Even though the reaction is favored in the direction from xylose to xylose-1-phosphate, the reaction is sufficiently reversible in a dynamic
10 metabolic environment within a yeast cell for forming xylose from xylose-1P. A preferred gene coding for an xylose kinase is an exogenous gene, more preferably the exogenous gene is a gene encoding an enzyme of plant or bacterial origin or a variant thereof. Expression of a gene coding for a xylose kinase in a yeast cell of the invention is straightforward once the sequence of a xylose kinase becomes available in the
15 sequence databases.

In an alternatively preferred embodiment, the yeast cell of the invention comprises a gene coding for a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10), collectively referred to herein as sugar phosphatase. The sugar phosphatase at least catalyzes the reaction $\text{xylose-1P} \rightarrow \text{xylose} + \text{phosphate}$ and/or $\text{arabinose-P} \rightarrow \text{arabinose}$. The gene coding for a sugar phosphatase can be an
20 endogenous or exogenous gene. An exogenous gene coding for a sugar phosphatase preferably is a gene encoding an enzyme of bacterial, fungal or plant origin or a variant thereof. An example of a suitable gene coding for a sugar phosphatase is e.g. the *E. coli* YihX (HAD4) gene as described by Kuznetsova et al. (2006, J. Biol. Chem. 281: 36149–36161; accession no.: NP_193348). An gene coding for an enzyme with sugar
25 phosphatase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with SEQ ID NO: 9. A gene coding for an enzyme with sugar phosphatase activity may also comprises a nucleotide sequence coding for an amino
30 acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of SEQ ID NO: 9. Preferably the amino acid sequence has no more than 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to SEQ ID NO: 9. A preferred codon-optimized

nucleotide sequences coding for a sugar phosphatase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in SEQ ID NO: 18 (see Table 1).

Other examples of suitable genes coding for sugar phosphatases are e.g. the *E. coli* EC 3.1.3.10 agp gene, the *Prevotella ruminicola* EC 3.1.3.10 agp gene (accession no.: YP_003576035) or the *Neisseria meningitidis* gene coding for the EC 3.1.3.23 sugar-1P phosphatase (accession no.: YP_003083641). A gene coding for an enzyme with sugar phosphatase and/or sugar-1P phosphatase activity preferably comprises a nucleotide sequence coding for an amino acid sequence with at least 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% amino acid sequence identity with at least one of SEQ ID NO: 41, 43 and 45. A gene coding for an enzyme with sugar phosphatase and/or sugar-1P phosphatase activity may also comprises a nucleotide sequence coding for an amino acid sequence having one or several substitutions, insertions and/or deletions as compared to the amino acid sequence of at least one of SEQ ID NO: 41, 43 and 45. Preferably the amino acid sequence has no more than 100, 75, 50, 40, 30, 20, 10 or 5 amino acid substitutions, insertions and/or deletions as compared to at least one of SEQ ID NO: 41, 43 and 45. A preferred codon-optimized nucleotide sequences coding for a sugar phosphatase is a nucleotide sequence with at least 80, 85, 90, 95, 98, 99% sequence identity with the open reading frame in at least one of SEQ ID NO: 42, 44 and 46.

Most sugar phosphatases and sugar-1P phosphatases have a broad substrate specificity. Expression of such phosphatases in the yeast cell of the invention may therefore interfere with other, e.g. anabolic reactions in the cell. In a preferred embodiment the sugar phosphatase or sugar-1P phosphatase is therefore expressed under control of an inducible promoter such as e.g. an anoxic promoter. This will allow for a process wherein the yeast cells of the invention are first grown under conditions under which the promoter controlling the phosphatase is not active (e.g. aerobic conditions), to efficiently accumulate sufficient biomass without interference of the phosphatase. Then in a second stage of the process, conditions are used which induce the promoter (e.g. anaerobic conditions) so that phosphatase activity is expressed in the cell, allowing the accumulated yeast biomass to ferment the uronic acid(s) into a fermentation product. In an alternative embodiment the sugar phosphatase or sugar-1P

phosphatase is expressed under control of an relatively weak promoter such as e.g. an INO4 or a CYC1 promoter.

G) Enzymes for fermenting pentoses

In a further preferred embodiment, the yeast cell of the invention has at least one of: a) the ability of isomerising xylose to xylulose (a xylose isomerase; EC 5.3.1.5); and, b) the ability to convert L-arabinose into D-xylulose 5-phosphate. For a) the cell preferably has a functional xylose isomerase gene, which gene confers to the cell the ability to isomerise xylose into xylulose. The xylose isomerase preferably is an exogenous gene. For b) the cell preferably has functional exogenous genes coding for a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4), which genes together confers to the cell the ability to isomerise convert L-arabinose into D-xylulose 5-phosphate. The genes conferring to the cell the ability to isomerise convert L-arabinose into D-xylulose 5-phosphate are preferably exogenous genes.

Yeast cells having the ability of isomerising xylose to xylulose as e.g. described in WO 03/0624430, WO 06/009434, and WO 10/074577. The ability of isomerising xylose to xylulose is preferably conferred to the cell by transformation with a nucleic acid construct comprising a nucleotide sequence encoding a xylose isomerase. Preferably the cell thus acquires the ability to directly isomerise xylose into xylulose. More preferably the cell thus acquires the ability to grow aerobically and/or anaerobically on xylose as sole energy and/or carbon source though direct isomerisation of xylose into xylulose (and further metabolism of xylulose). It is herein understood that the direct isomerisation of xylose into xylulose occurs in a single reaction catalyzed by a xylose isomerase, as opposed to the two step conversion of xylose into xylulose via a xylitol intermediate as catalyzed by xylose reductase and xylitol dehydrogenase, respectively.

Several xylose isomerases (and their amino acid and coding nucleotide sequences) that may be successfully used to confer to the cell of the invention the ability to directly isomerise xylose into xylulose have been described in the art. These include the xylose isomerases of *Piromyces* sp. and of other anaerobic fungi that belongs to the families *Neocallimastix*, *Caecomyces*, *Piromyces* or *Ruminomyces* (WO 03/0624430), *Cyllamyces aberensis* (US 20060234364), *Orpinomyces* (Madhavan et al., 2008, DOI 10.1007/s00253-008-1794-6), the xylose isomerase of the bacterial

genus *Bacteroides*, including e.g. *B. thetaiotaomicron* (WO 06/009434), *B. fragilis*, and *B. uniformis* (WO 09/109633), the xylose isomerase of the anaerobic bacterium *Clostridium phytofermentans* (Brat et al., 2009, Appl. Environ. Microbiol. 75: 2304–2311), *Lactococcal* xylose isomerases (WO 10/070549), the xylose isomerases of
5 *Clostridium difficile*, *Ciona intestinales* and *Fusobacterium mortiferum* (WO 10/074577) and *Ruminococcal* xylose isomerases and chimeric isomerases therewith (WO 2011/006136).

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

Construction of yeast cells of the invention

In a second aspect the invention relates to methods for preparing or constructing the yeast cells of the invention. For this purpose standard genetic and molecular biology techniques are used that are generally known in the art and have e.g. been
30 described by Sambrook and Russell (2001, "Molecular cloning: a laboratory manual" (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press) and Ausubel et al. (1987, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York). Furthermore, the construction of

mutated host yeast strains is carried out by genetic crosses, sporulation of the resulting diploids, tetrad dissection of the haploid spores containing the desired auxotrophic markers, and colony purification of such haploid host yeasts in the appropriate selection medium. All of these methods are standard yeast genetic methods known to those in the art. See, for example, Sherman et al., *Methods Yeast Genetics*, Cold Spring Harbor Laboratory, NY (1978) and Guthrie et al. (Eds.) *Guide To Yeast Genetics and Molecular Biology* Vol. 194, Academic Press, San Diego (1991).

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

Suitable promoters for expression of the nucleotide sequences coding for the enzymes to be introduced into the yeast cell of the invention include promoters that are preferably insensitive to catabolite (glucose) repression, that are active under anaerobic conditions and/or that preferably do not require xylose or arabinose for induction. Promoters having these characteristics are widely available and known to the skilled person. Suitable examples of such promoters include e.g. promoters from glycolytic genes such as the phosphofructokinase (*PPK*), triose phosphate isomerase (*TPI*), glyceraldehyde-3-phosphate dehydrogenase (*GPD*, *TDH3* or *GAPDH*), pyruvate kinase (*PYK*), phosphoglycerate kinase (*PGK*), glucose-6-phosphate isomerase promoter (*PGII*) promoters from yeasts. More details about such promoters from yeast may be found in (WO 93/03159). Other useful promoters are ribosomal protein encoding gene promoters (*TEF1*), the lactase gene promoter (*LAC4*), alcohol dehydrogenase

promoters (*ADHI*, *ADH4*, and the like), the enolase promoter (*ENO*) and the hexose(glucose) transporter promoter (*HXT7*). Alternatively, a nucleotide sequences encoding an enzyme to be introduced into the yeast cell of the invention can be expressed under anaerobic conditions by using an anoxic promoter such as e.g. the *S. cerevisiae ANBI* promoter (SEQ ID NO: 24). Other promoters, both constitutive and inducible, and enhancers or upstream activating sequences will be known to those of skill in the art. Preferably the promoter that is operably linked to nucleotide sequence as defined above is homologous to the host cell. Suitable terminator sequences are e.g. obtainable from the cytochrome c1 (*CYCI*) gene or an alcohol dehydrogenase gene (e.g. *ADHI*).

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

The yeast cell transformed with the nucleic acid construct(s) comprising the nucleotide sequence encoding the enzymes to be introduced into the yeast cell of the invention preferably is a yeast cell capable of passive or active uronic acid transport into the cell. The yeast cell of the invention, further preferably is a cell capable of active or passive pentose (xylose and preferably also arabinose) transport into the cell.

The yeast cell preferably contains active glycolysis. The yeast cell may further preferably contain an endogenous pentose phosphate pathway and may contain endogenous xylulose kinase activity so that xylulose isomerised from xylose may be metabolized to pyruvate. The yeast further preferably contains enzymes for conversion of a pentose (preferably through pyruvate) to a desired fermentation product such as ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, butanols (1-butanol, 2-butanol, isobutanol) isoprenoid-derived products. A particularly preferred yeast cell is a yeast cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. The yeast cell further preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than 5, 4, or 3) and towards organic acids like lactic acid, acetic acid or formic acid and sugar degradation products such as furfural and hydroxy-methylfurfural, and a high tolerance to elevated temperatures. Any of these characteristics or activities of the yeast cell may be naturally present in the yeast cell or may be introduced or modified by genetic modification, preferably by self cloning or by the methods of the invention described below. A suitable cell is a cultured cell, a cell that may be cultured in fermentation process e.g. in submerged or solid state fermentation.

Yeasts are herein defined as eukaryotic microorganisms and include all species of the subdivision Eumycotina (Yeasts: characteristics and identification, J.A. Barnett, R.W. Payne, D. Yarrow, 2000, 3rd ed., Cambridge University Press, Cambridge UK; and, The yeasts, a taxonomic study, C.P. Kurtzman and J.W. Fell (eds) 1998, 4th ed., Elsevier Science Publ. B.V., Amsterdam, The Netherlands) that predominantly grow in unicellular form. Yeasts may either grow by budding of a unicellular thallus or may grow by fission of the organism. Preferred yeasts cells for use in the present invention belong to the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, *Brettanomyces*, and *Yarrowia*. Preferably the yeast is capable of anaerobic fermentation, more preferably anaerobic alcoholic fermentation. Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i.e., a high tolerance for acidity, ethanol and high osmolarity, capability of anaerobic growth, and

of course its high alcoholic fermentative capacity. Preferred yeast species for use in the present invention include *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* and *Schizosaccharomyces pombe*.

A yeast cell of the invention may be a haploid cell, however, preferably the yeast
5 cell is a non-haploid cell such as e.g. a diploid, aneuploid or polyploid cell. Thus, the yeast cell preferably has a greater number of chromosomes than the haploid number (n) of chromosomes in a gamete, i.e. yeast spore. Non-haploid yeast strains are more robust and more stable compared to haploid strains with the same genotype and are therefore preferably used for industrial fermentation processes, including the production of
10 ethanol.

A yeast cell of the invention further preferably comprises xylulose kinase activity so that xylulose isomerised from xylose may be metabolized to pyruvate. Preferably, the cell contains endogenous xylulose kinase activity. More preferably, a cell of the invention comprises a genetic modification that increases the specific xylulose kinase
15 activity. Preferably the genetic modification causes overexpression of a xylulose kinase, e.g. by overexpression of a nucleotide sequence encoding a xylulose kinase. The gene encoding the xylulose kinase may be endogenous to the cell or may be a xylulose kinase that is heterologous to the cell. A nucleotide sequence that may be used for overexpression of xylulose kinase in the cells of the invention is e.g. the xylulose
20 kinase gene from *S. cerevisiae* (*XKSI*) as described by Deng and Ho (1990, Appl. Biochem. Biotechnol. 24-25: 193-199). Another preferred xylulose kinase is a xylose kinase that is related to the xylulose kinase from *Piromyces* (*xylB*; see WO 03/0624430). In the cells of the invention, a xylulose kinase to be overexpressed is overexpressed by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to a strain
25 which is genetically identical except for the genetic modification causing the overexpression. It is to be understood that these levels of overexpression may apply to the steady state level of the enzyme's activity, the steady state level of the enzyme's protein as well as to the steady state level of the transcript coding for the enzyme.

A yeast cell of the invention further preferably comprises a genetic modification
30 that increases the flux of the pentose phosphate pathway as described in WO 06/009434. In particular, the genetic modification causes an increased flux of the non-oxidative part pentose phosphate pathway. A genetic modification that causes an increased flux of the non-oxidative part of the pentose phosphate pathway is herein

understood to mean a modification that increases the flux by at least a factor 1.1, 1.2, 1.5, 2, 5, 10 or 20 as compared to the flux in a strain which is genetically identical except for the genetic modification causing the increased flux. The flux of the non-oxidative part of the pentose phosphate pathway may be measured as described in WO 06/009434. Genetic modifications that increase the flux of the pentose phosphate pathway may be introduced in the cells of the invention in various ways. These including e.g. achieving higher steady state activity levels of xylulose kinase and/or one or more of the enzymes of the non-oxidative part pentose phosphate pathway and/or a reduced steady state level of unspecific aldose reductase activity. These changes in steady state activity levels may be effected by selection of mutants (spontaneous or induced by chemicals or radiation) and/or by recombinant DNA technology e.g. by overexpression or inactivation, respectively, of genes encoding the enzymes or factors regulating these genes. In a preferred cell of the invention, the genetic modification comprises overexpression of at least one enzyme of the (non-oxidative part) pentose phosphate pathway. Preferably the enzyme is selected from the group consisting of the enzymes encoding for ribulose-5-phosphate isomerase, ribulose-5-phosphate 3-epimerase, transketolase and transaldolase.

There are various means available in the art for overexpression of enzymes in the yeast cells of the invention. In particular, an enzyme may be overexpressed by increasing the copy number of the gene coding for the enzyme in the cell, e.g. by integrating additional copies of the gene in the cell's genome, by expressing the gene from an episomal multicopy expression vector or by introducing a episomal expression vector that comprises multiple copies of the gene. The coding sequence used for overexpression of the enzymes preferably is homologous to the yeast cell of the invention. However, coding sequences that are heterologous to the yeast cell of the invention may likewise be applied. Alternatively overexpression of enzymes in the cells of the invention may be achieved by using a promoter that is not native to the sequence coding for the enzyme to be overexpressed, i.e. a promoter that is heterologous to the coding sequence to which it is operably linked. Although the promoter preferably is heterologous to the coding sequence to which it is operably linked, it is also preferred that the promoter is homologous, i.e. endogenous to the cell of the invention. Preferably the heterologous promoter is capable of producing a higher steady state level of the transcript comprising the coding sequence (or is capable of

producing more transcript molecules, i.e. mRNA molecules, per unit of time) than is the promoter that is native to the coding sequence, preferably under conditions where one or more of xylose, arabinose and glucose are available as carbon sources, more preferably as major carbon sources (i.e. more than 50% of the available carbon source consists of one or more of xylose, arabinose and glucose), most preferably as sole carbon sources.

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

A further preferred transformed yeast cell according to the invention may comprises further genetic modifications that result in one or more of the characteristics selected from the group consisting of (a) increased transport of xylose and/or arabinose into the cell; (b) decreased sensitivity to catabolite repression; (c) increased tolerance to ethanol, osmolarity or organic acids; and, (d) reduced production of by-products. By-

products are understood to mean carbon-containing molecules other than the desired fermentation product and include e.g. xylitol, arabinitol, glycerol and/or acetic acid. Expression of an NAD^+ -dependent acetylating acetaldehyde dehydrogenase (EC 1.2.1.10) presents an alternative to glycerol synthesis as a means of recycling excess NADH in anaerobically grown yeasts as suggested by Waks and Silver (2009, Appl. Environ. Microbiol. 75: 1867–1875). In yeast cells expressing an NAD^+ -dependent acetylating acetaldehyde dehydrogenase one or more genes involved in glycerol formation, such as the *S. cerevisiae* glycerol-phosphate dehydrogenases (GPD1 and/or GPD2) genes and/or glycerol-phosphate phosphatase (GPP1 and/or GPP2) genes, are therefore preferably inactivated or reduced in expression (WO 11/01092).

Any genetic modification described herein can be introduced by classical mutagenesis and screening and/or selection for the desired mutant, or simply by screening and/or selection for the spontaneous mutants with the desired characteristics.

Alternatively, the genetic modifications may consist of overexpression of endogenous genes and/or the inactivation of endogenous genes. Genes the overexpression of which is desired for increased transport of arabinose and/or xylose into the cell are preferably chosen from genes encoding a hexose or pentose transporter. In *S. cerevisiae* and other yeasts these genes include *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT5*, *HXT7* and *GAL2*, of which *HXT7*, *HXT5* and *GAL2* are most preferred (see Sedlack and Ho, Yeast 2004; 21: 671–684). Another preferred transporter for expression in yeast is the glucose transporter encoded by the *P. stipitis* *SUT1* gene (Katahira et al., 2008, Enzyme Microb. Technol. 43: 115–119). Similarly orthologues of these transporter genes in other species may be overexpressed. Other genes that may be overexpressed in the cells of the invention include genes coding for glycolytic enzymes and/or ethanologenic enzymes such as alcohol dehydrogenases. Preferred endogenous genes for inactivation include hexose kinase genes e.g. the *S. cerevisiae* *HXK2* gene (see Diderich et al., 2001, Appl. Environ. Microbiol. 67: 1587–1593); the *S. cerevisiae* *MIG1* or *MIG2* genes; or (hybridizing) orthologues of these genes in other species. Other preferred further modifications of yeast cells for xylose fermentation are described in van Maris et al. (2006, Antonie van Leeuwenhoek 90:391–418), WO2006/009434, WO2005/023998, WO2005/111214, and WO2005/091733. Any of the genetic modifications of the cells of the invention as described herein are, in as far as possible, preferably introduced or modified by self cloning genetic modification.

In a preferred embodiment of a method for preparing or constructing the yeast cells of the invention, a yeast of the invention is further improved by evolutionary engineering. Thus, a yeast of the invention that has the ability of metabolizing uronic acids as described herein above, can be further subjected to a method for identifying a yeast cell/strain that shows improved consumption of uronic acids in comparison to the starting or reference yeast cell/strain to which the method is applied. The method may thus be used to improve the performance of an existing strain of yeast cells of the invention with respect to its ability to consume a uronic acid such as galacturonic acid or glucuronic acid, e.g. by selecting a strain of the organism which shows faster consumption of the uronic acid(s). Preferably, the method is used to select a strain which has improved consumption on a carbon source comprising a uronic acid so that it shows improved fermentation characteristics. Thus, a strain which has been selected according to the invention may show improved performance in terms of increased productivity, e.g. on a volumetric basis, of the fermentation product in question. Also, or alternatively, a strain selected using this method may show an increase in yield of the fermentation product (in comparison to the starting strain from which it was selected). In the method of the invention, a population of the yeast cell of the invention is grown, that is to say selected, in the presence a uronic acid, preferably at least one of galacturonic acid and glucuronic acid. Preferably the yeast cell of the invention is grown/ selected on the uronic acid(s) as sole carbon source. Growth of the population of the yeast cell on the indicated carbon sources exerts selection pressure on the population. Thus, mutants in the population may be selected for with an increased maximum specific growth rate (μ_{max}) on the uronic acid(s) carbon source. If the selection pressure is maintained, e.g. by sequentially transferring batch-wise grown cultures to new batches, eventually (mutant) cells with a higher specific growth rate will overgrow all other cells with a lower specific growth rate. The process of growing the yeast cells may e.g. be operated in batch culture, as a fed batch fermentation with constant feed or as a continuous fermentation, e.g. in a chemostat. A method for selecting a strain the yeast cell of the invention capable of improved consumption of a uronic acid carbon sources as compared to a reference strain of the yeast cell, which method comprises: a) growing a population of the reference strain of the yeast cell in the presence of a uronic acid carbon source and b) selecting from the population a resulting strain of the yeast cell capable of improved consumption of a uronic acid

carbon source as compared to the reference strain of the organism. Preferably the method is a method, wherein the growth of the population of the yeast cell is carried out by cultivation in sequential batch reactors (SBR). The method can be carried out under anaerobic conditions or the method can be carried out under aerobic conditions, preferably under oxygen limited conditions aerobic conditions.

A preferred yeast cell according to the invention has the ability to grow on at least one of glucuronic acid and galacturonic acid as carbon/energy source, preferably as sole carbon/energy source, and preferably under anaerobic conditions, i.e. conditions as defined herein below for anaerobic fermentation process.

A preferred yeast cell of the invention has the ability to grow on at least one of a hexose, a pentose, a uronic acid and combinations thereof at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0.25 or 0.3 h⁻¹ under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h⁻¹ under anaerobic conditions. Therefore, preferably the yeast cell has the ability to grow on at least one of glucuronic acid and galacturonic acid as sole carbon/energy source at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0.25 or 0.3 h⁻¹ under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h⁻¹ under anaerobic conditions. More preferably, the yeast cell has the ability to grow on a mixture of glucuronic acid and galacturonic acid and at least on one of xylose and arabinose and of a hexose (e.g. at least on one of sucrose, glucose, fructose or galactose) as carbon/energy source at a rate of at least 0.01, 0.02, 0.05, 0.1, 0.2, 0.25 or 0.3 h⁻¹ under aerobic conditions, or, more preferably, at a rate of at least 0.005, 0.01, 0.02, 0.05, 0.08, 0.1, 0.12, 0.15 or 0.2 h⁻¹ under anaerobic conditions.

Further aspects of the invention

In a third aspect, the invention relates to the use of a yeast cell according to the invention for the preparation of a fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propanediol, butanols and isoprenoid-derived products. Preferably the yeast cell is used for the preparation of a fermentation product from a uronic acid as carbon source, more preferably at least one of galacturonic acid and glucuronic acid as carbon source.

In another aspect the invention relates to a process for producing a fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic

acid, acrylic acid, 1,3-propane-diol, butanols (1-butanol, 2-butanol, isobutanol) and isoprenoid-derived products. The process preferably comprises the step of: a) fermenting a medium with a yeast cell, whereby the medium contains or is fed with: a) a source of a uronic acid and whereby the yeast cell ferments the uronic acid to the fermentation product. The source of the uronic acid preferably is a source of at least one of galacturonic acid and glucuronic acid. The yeast cell preferably is a yeast cell as herein defined above. The process preferably comprises one or more further steps wherein the fermentation product is recovered. The process may be a batch process, a fed-batch process or a continuous process as are well known in the art.

In a preferred process the medium further contains or is fed with a source of at least one of a hexose, a pentose and glycerol. A source of at least one of a hexose, a pentose and glycerol comprises or consist of: hexose and pentose; hexose and glycerol; pentose and glycerol; hexose, pentose and glycerol. In a further preferred process, the medium further contains or is fed with a source of acetic acid.

Preferably, the medium fermented by the cells of the invention comprises or is fed with (fractions of) hydrolyzed biomass comprising the uronic acid, preferably at least one of galacturonic acid and glucuronic acid. Suitable sources of uronic acids are pectin-rich residues which accumulate when sugar is extracted from sugar beet, i.e. sugar beet pulp, and when juices are produced from fruits such as citrus and apple, i.e. apple or citrus fruit pulp, peel, and/or rag. Another suitable source of uronic acids are pectins in hemicellulose, e.g. as present in lignocellulosic biomass.

Pectins are complex and heterogeneous polymers that primarily act as hydrating and cementing agents for the cellulosic matrix of plant cell walls. The principal unit in pectin chains is α -(1-4) linked galacturonic acid. The galacturonic acid residues can be esterified with methyl and acetyl groups. Additionally, pectin contains the branched polysaccharides rhamnogalacturonan I, rhamnogalacturonan II and xylogalacturonan (see Richard and Hilditch, 2009, *supra*). The various pectins or pectic substances can be hydrolyzed by simply boiling in 2.5% sulphuric acid. However, preferably they are hydrolyzed by pectinolytic enzymes as are known in the art. Hydrolysis of pectin can e.g. be accomplished by a mixture of enzymes, comprising one or more of a pectin methyl esterase, a pectin acetyl esterase, an endo-polygalacturonase, an endo-pectin lyase, a rhamnogalacturonan hydrolyase, a rhamnogalacturonan lyase, a rhamnogalacturonan acetyl esterase, an arabinofuranosidase, an endo-arabinose, an

endo-galactanase and β -galactosidase see e.g. Kashyap et al. (2001, Bioresour Technol. 77:215–227), Hoondal et al. (2002, Appl Microbiol Biotechnol 59:409–418), Baciú and Jördening (2004, Enzyme Microb. Technol. 34:505–512), Jayani et al. (2005, Process Biochem 40:2931–2944) and references therein. In addition to galacturonic acid the
5 (fractions of) hydrolyzed pectic substances acid can also comprise glucuronic acid, galactose, xylose, arabinose, acetic acid (or a salt thereof), all of which can be consumed by the yeast cells of the invention.

In the process of the invention, the sources of uronic acid may be galacturonic acid and glucuronic acid as such (i.e. as monomeric uronic acids) or they may be in a
10 polymeric form such as e.g. above described pectic substances. For release of the various monomeric units from the polymeric pectic substances, appropriate pectinolytic enzymes may be added to the fermentation medium or may be produced by a yeast cell of the invention. In the latter case the yeast cell may be genetically engineered to produce and excrete such pectinolytic enzymes.

15 In a preferred process the medium fermented by the cells of the invention comprises or is fed with, in addition to the above-described sources of uronic acid, sources of hexoses and/or pentoses. The source of hexose comprises or consists of at least one of glucose, fructose, sucrose, maltose, galactose and lactose. Preferably the source of pentose comprises or consists of at least one of xylose and arabinose, of
20 which xylose is preferred. Preferably, the medium fermented by the cells of the invention comprises or is fed with (fractions of) hydrolyzed biomass comprising at least one at least one of a hexose and a pentose such as glucose, xylose and/or arabinose. The (fractions of) hydrolyzed biomass comprising the hexoses and pentose will usually also comprise acetic acid (or a salt thereof). An example of hydrolyzed
25 biomass to be fermented in the processes of the invention is e.g. hydrolyzed lignocellulosic biomass. Examples of lignocellulosic biomass to be hydrolyzed for use in the present invention include agricultural residues (including e.g. empty fruit bunches (EFB) of oil palm, corn stover and sugarcane bagasse), wood residues (including sawmill and paper mill discards and (municipal) paper waste. Methods for
30 hydrolysis of biomass such as lignocelluloses are known in the art per se and include e.g. acids, such as sulphuric acid and enzymes such as cellulases and hemicellulases. In the process of the invention, the sources of xylose, glucose and arabinose may be xylose, glucose and arabinose as such (i.e. as monomeric sugars) or they may be in the

form of any carbohydrate oligo- or polymer comprising xylose, glucose and/or arabinose units, such as e.g. lignocellulose, arabinans, xylans, cellulose, starch and the like. For release of xylose, glucose and/or arabinose units from such carbohydrates, appropriate carbohydrases (such as arabinases, xylanases, glucanases, amylases, cellulases, glucanases and the like) may be added to the fermentation medium or may be produced by the modified host cell. In the latter case the modified host cell may be genetically engineered to produce and excrete such carbohydrases. An additional advantage of using oligo- or polymeric sources of glucose is that it enables to maintain a low(er) concentration of free glucose during the fermentation, e.g. by using rate-limiting amounts of the carbohydrases preferably during the fermentation. This, in turn, will prevent repression of systems required for metabolism and transport of non-glucose sugars such as xylose and arabinose. In a preferred process the modified host cell ferments both the glucose and at least one of xylose and arabinose, preferably simultaneously in which case preferably a modified host cell is used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source of at least one of a hexose, a pentose and glycerol, as carbon source, the fermentation medium will further comprise the appropriate ingredients required for growth of the modified host cell. Compositions of fermentation media for growth of eukaryotic microorganisms such as yeasts are well known in the art.

The fermentation process may be an aerobic or an anaerobic fermentation process. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than 5, 2.5 or 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors. In the absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidized by oxidative phosphorylation. To solve this problem many microorganisms use pyruvate or one of its derivatives as an electron and hydrogen acceptor thereby regenerating NAD^+ . Thus, in a preferred anaerobic fermentation process pyruvate is used as an electron (and hydrogen acceptor) and is reduced to fermentation products such as ethanol, as well as non-ethanol fermentation products such as lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, butanols (1-butanol, 2-butanol, isobutanol) and isoprenoid-derived products, preferably under concomitant production of formate. Anaerobic processes of

the invention are preferred over aerobic processes because anaerobic processes do not require investments and energy for aeration and in addition, anaerobic processes produce higher product yields than aerobic processes. Alternatively, the fermentation process of the invention may be run under aerobic oxygen-limited conditions.

- 5 Preferably, in an aerobic process under oxygen-limited conditions, the rate of oxygen consumption is at least 5.5, more preferably at least 6 and even more preferably at least 7 mmol/L/h.

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

- The fermentation process is preferably run at a pH that is acceptable to the yeast cell that is used in the process, based on common general knowledge or as can be routinely determined. Usually fermentation process is run at a neutral or acidic pH, preferably at a pH less than pH, 7.0, 6.0, 5.5, 5.0 or 4.5 and at pH higher than pH 2.0, 3.0, 3.5 or 4.0.

- A preferred fermentation process according to the invention is a process for the production of ethanol, whereby the process comprises the step of fermenting a medium with a yeast cell, whereby the medium contains or is fed with a source of a uronic acid and optionally a source of at least one of a hexose, a pentose and glycerol and whereby the yeast cell ferments the uronic acid and optionally the at least one of a hexose, pentose and glycerol to ethanol. Optionally the process comprises the step of recovery of at least one of ethanol. The fermentation may further be performed as described above. In the process the 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 ethanol per liter per hour. The ethanol yield on the uronic acid, and optionally the hexose and/or pentose and/or glycerol (and/or acetate) in the process preferably is at least 50, 60, 70, 80, 90, 95 or 98%. The ethanol yield is herein defined as a percentage of the theoretical maximum yield, which, for xylose, glucose and arabinose is 0.51 g. ethanol per g. hexose or pentose. For glycerol the theoretical maximum yield is 0.50 g. ethanol per g. glycerol and for acetic acid the theoretical maximum yield is 0.77 g. ethanol per g. acetic acid. For the uronic acids

galacturonic acid and glucuronic acid the theoretical maximum yield is 0.39 g. ethanol per g. uronic acid.

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

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

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

Description of the figures

Figure 1. Plant salvage pathways for converting uronic acids into a pentose-1-phosphates. The relevant enzymes of the pathway are indicated by the corresponding EC numbers.

Examples

20

1. Construction of yeast cells of the invention

All promoter and terminator sequences used in the construction of expression cassettes are publicly available and their sequences are indicated in the sequence listing. Synthetic DNA fragments were prepared comprising codon optimized (for expression in *S. cerevisiae*) open reading frames coding for the enzymes to be expressed in the yeast cells of the invention. In Table 1 the Codon Adaptation Index (CAI) is given for the various coding sequences including references to the sequence listing.

Table 1: Overview of the codon optimized synthetic DNA fragments comprising the open reading frames coding for the various enzymes identified by their respective EC numbers. SEQ ID NO.'s for the corresponding amino acid sequences (aa) and nucleotide sequences (nt) are also given.

Enzyme	CAI	L_aa	SEQ ID NO: aa	SEQ ID NO: nt	ORF
2.7.1.44_	0.956	424	1	10	10 - 1284
2.7.1.43	0.968	362	2	11	10 - 1098
2.7.7.64_(Ps)	0.956	600	3	12	10 - 1812
2.7.7.64_(At)	0.996	614	4	13	10 - 1854
5.1.3.6	1.000	429	5	14	11 - 1300
4.1.1.35	0.968	342	6	15	10 - 1038
5.1.3.5	0.994	419	7	16	10 - 1269
2.7.1.46_	0.965	1039	8	17	10 - 3129
YihX	0.931	199	9	18	1 - 609
3.1.3.10 Ec	0.960	413	41	42	824-2065
3.1.3.10 Pr	0.976	469	43	44	824-2233
3.1.3.23	1.000	262	45	46	824-1612

1.1 Construction of expression vector plasmids pRN757, pRN758, pRN759 and pRN760

First, expression cassettes for the expression of the genes encoding the enzymes
 5 EC 2.7.7.64 (based on *Pisum sativum* USP - AB178642), EC 4.1.1.35 (based on
Arabidopsis thaliana UXS3 - AT5G59290) EC2.7.1.46 (based on *Arabidopsis thaliana*
 ARA1 - AT4G16130) and a sugar phosphatase YihX (based on *Escherichia coli* YihX
 (HAD4) - EG11850) are prepared as follows.

The plasmids are based on pRN599, a shuttle vector plasmid with a 2 μ origin of
 10 replication and a kanMX selection marker encoding G418 resistance (SEQ ID NO: 19).

The EC 2.7.7.64 Ps expression cassette (SEQ ID NO: 20) is prepared by ligating
 the TEF1 promoter (cut with the restriction enzymes *Age*I and *Hind*III), the synthetic
 ORF (cut with *Hind*III and *Bss*HII) and the ADH1 terminator (cut with *Bss*HII and
*Sal*I) together in pCRII (InVitrogen) to yield plasmid pPS27764.

15 The EC 4.1.1.35 At expression cassette (SEQ ID NO: 21) is prepared by ligating
 the PGK1 promoter (cut with the restriction enzymes *Spe*I and *Pst*I), the synthetic ORF
 (cut with *Nsi*I and *Sal*I) and the PGI1 terminator (cut with *Xho*I and *Bsi*WI) together in
 pCRII (invitrogen) to yield plasmid pAT41135.

20 The EC 2.7.1.46 At expression cassette (SEQ ID NO: 22) is prepared by ligating
 the ADHm1 promoter (cut with the restriction enzymes *Sac*I and *Pst*I), the synthetic

ORF (cut with *Pst*I and *Bam*HI) and the CYC1 terminator (cut with *Bam*HI and *Xho*I) together in pCRII (invitrogen) to yield plasmid pAT27146.

The YihX Ec expression cassette (SEQ ID NO: 23) is prepared by ligating the anaerobic ANB1 promoter (cut with the restriction enzymes *Sac*I and *Pst*I), the
5 synthetic ORF (cut with *Nsi*I and *Bgl*II) and the CYC1 terminator (cut with *Bam*HI and *Sal*I) together in pCRII (invitrogen) to yield plasmid pECYihX.

Next, expression vector plasmids pRN757, pRN758, pRN759 en pRN760 are constructed as follows.

For the final construction of pRN757, the shuttle vector pRN599 is cut with *Xma*I
10 and *Acc*65I, this vector is combined with the inserts from pPS27764 cut with *Sal*I and *Age*I (2577bp), pAT41135 cut with *Bsi*WI and *Sac*I (2127bp) and pECYihX cut with *Sac*I and *Sal*I (1701bp) to produce pRN757 (SEQ ID NO: 24).

For the final construction of pRN758, the shuttle vector pRN599 is cut with *Xma*I
15 and *Acc*65I, this vector is combined with the inserts from pPS27764 cut with *Sal*I and *Age*I (2577bp), pAT41135 cut with *Bsi*WI and *Sac*I (2127bp) and pAT27146 cut with *Sac*I and *Xho*I (4146bp) to produce pRN758 (SEQ ID NO: 25).

For the final construction of pRN759, the shuttle vector pRN599 is cut with *Xho*I
20 and *Acc*65I, this vector is combined with the inserts from pAT41135 cut with *Bsi*WI and *Sac*I (2127bp) and pECYihX cut with *Sac*I and *Sal*I (1701bp) to produce pRN759 (SEQ ID NO: 26).

For the final construction of pRN760, the shuttle vector pRN599 is cut with *Xho*I
and *Acc*65I, this vector is combined with the inserts from pAT41135 cut with *Bsi*WI and *Sac*I (2127bp) and pAT27146 cut with *Sac*I and *Xho*I (4146bp) to produce pRN760 (SEQ ID NO: 27).

25

1.2 Construction of expression vector plasmid pRN761

First, expression cassettes for the expression of the genes encoding the enzymes EC 2.7.1.44 Y250F mutant (Yang *et al.*, 2009, *supra*; based on *Arabidopsis thaliana* GalAK- FJ439676) and EC 5.1.3.6 (based on *Arabidopsis thaliana* GAE1 -
30 AT4G30440) are prepared as follows.

The plasmid is based on pRN600, a shuttle vector plasmid with a CEN IV origin of replication and a hphMX selection marker encoding hygromycine resistance (SEQ ID NO: 28).

The EC 2.7.1.44 At Y250F expression cassette (SEQ ID NO: 29) is prepared by ligating the TDH3 promoter (cut with the restriction enzymes *Acc65I* and *EcoRI*), the synthetic ORF (cut with *EcoRI* and *BamHI*) and the CYC1 terminator (cut with *BamHI* and *XhoI*) together in pCRII (invitrogen) to yield plasmid pAT27144.

- 5 The EC 5.1.3.6 At expression cassette (SEQ ID NO: 30) is prepared by ligating the ACT1 promoter (cut with the restriction enzymes *SpeI* and *PstI*), the synthetic ORF (cut with *PstI* and *BssHII*) and the ADH1 terminator (cut with *BssHII* and *BsiWI*) together in pCRII (invitrogen) to yield plasmid pAT5136.

10 For the final construction of pRN761, the shuttle vector pRN600 is cut with *XhoI* and *SpeI*, this vector is combined with the inserts from pAT5136 cut with *SpeI* and *BsiWI* (2242bp) and pAT27144 cut with *Acc65I* and *XhoI* (2189bp) to produce pRN761 (SEQ ID NO: 31).

1.3 Construction of expression vector plasmid pRN762, pRN763 and pRN764

- 15 First, expression cassettes for the expression of the genes encoding the enzymes EC 2.7.1.43 (based on Arabidopsis thaliana GlcAK - Q93ZC9) and or EC 5.1.3.5 (based on Arabidopsis thaliana MUR4 - AY195742) are prepared as follows.

20 The plasmids are based on pRN615, a shuttle vector plasmid with a 2 μ origin of replication and a zeoMX selection marker encoding G418 phleomycin (SEQ ID NO: 32).

The EC 2.7.1.43 At expression cassette (SEQ ID NO: 33) is prepared by ligating the TDH3 promoter (cut with the restriction enzymes *Acc65I* and *EcoRI*), the synthetic ORF (cut with *EcoRI* and *BamHI*) and the CYC1 terminator (cut with *BamHI* and *XhoI*) together in pCRII (invitrogen) to yield plasmid pAT27143.

- 25 The EC 5.1.3.5 At expression cassette (SEQ ID NO: 34) is prepared by ligating the TPI1 promoter (cut with the restriction enzymes *AgeI* and *XbaI*), the synthetic ORF (cut with *SpeI* and *SalI*) and the PGI1 terminator (cut with *XhoI* and *HindIII*) together in pCRII (invitrogen) to yield plasmid pAT5135.

30 For the final construction of pRN762, the shuttle vector pRN615 is cut with *Acc65I* and *XhoI*, this vector is combined with the insert from pAT27143 cut with *Acc65I* and *XhoI* (2056bp) to produce pRN762 (SEQ ID NO: 35).

For the final construction of pRN763, pRN762 is cut with *AgeI* and *HindIII*, this vector is combined with the insert from pAT5135 cut with *AgeI* and *HindIII* (2493bp) to produce pRN763 (SEQ ID NO: 36).

For the final construction of pRN764, the shuttle vector pRN615 is cut with
5 *Acc65I* and *XhoI*, this vector is combined with the inserts from pAT27143 cut with *AgeI* and *XhoI* (2056bp) and pAT5135 cut with *AgeI* and *BsiWI* (2493bp) to produce pRN764 (SEQ ID NO: 37).

1.4 Construction of expression vector plasmid pRN765

10 First, an expression cassette for the expression of the gene encoding the enzyme EC 2.7.7.64 (based on *Arabidopsis thaliana* USP1- AT5G52560) is prepared as follows.

The plasmid is based on pRN656, a shuttle vector plasmid with a CEN IV origin of replication and a natMX selection marker encoding nourseotricine resistance (SEQ
15 ID NO: 38).

The EC 27764 At expression cassette (SEQ ID NO: 39) is prepared by ligating the TPI1 promoter (cut with the restriction enzymes *AflIII* and *XbaI*), the synthetic ORF (cut with *NheI* and *BssHII*) and the ADH1 terminator (cut with *BssHII* and *XhoI*) together in pCRII (invitrogen) to yield plasmid pAT27764.

20 For the final construction of pRN765, the shuttle vector pRN656 is cut with *AflIII* and *XhoI*, this vector is combined with the insert from pAT27764 cut with *AflIII* and *XhoI* (3113bp) to produce pRN765 (SEQ ID NO: 40).

1.5 Construction of expression vector plasmids pRN766, pRN767 and pRN768

25 First, an expression cassette for the expression of the gene encoding the *E.coli* EC 3.1.3.10 agp enzyme (SEQ ID NO: 41) is prepared by ligating the anaerobic ANB1 promoter (cut with the restriction enzymes *SacI* and *PstI*), the synthetic ORF (cut with *PstI* and *BamHI*) and the CYC1 terminator (cut with *BamHI* and *SalI*) together in pCRII (invitrogen) to yield plasmid pEC31310 (SEQ ID NO: 42).

30 For the final construction of pRN766, pRN656 is cut with *SacI* and *XhoI*, this vector is combined with the insert from pEC31310 cut with *SacI* and *SalI* (2343bp) to produce pRN766.

First, an expression cassette for the expression of the gene encoding the *Prevotella ruminicola* EC 3.1.3.10 agp enzyme (YP_003576035; SEQ ID NO: 43) is prepared by ligating the anaerobic ANB1 promoter (cut with the restriction enzymes *SacI* and *PstI*), the synthetic ORF (cut with *PstI* and *BamHI*) and the CYC1 terminator (cut with *BamHI* and *SalI*) together in pCRII (invitrogen) to yield plasmid pPr31310 (SEQ ID NO: 44).

For the final construction of pRN767, pRN656 is cut with *SacI* and *XhoI*, this vector is combined with the insert from pPr31310 cut with *SacI* and *SalI* (2511bp) to produce pRN767.

First, an expression cassette for the expression of the gene encoding the *Neisseria meningitidis* EC 3.1.3.23 enzyme (YP_003083641; SEQ ID NO: 45) is prepared by ligating the anaerobic ANB1 promoter (cut with the restriction enzymes *SacI* and *PstI*), the synthetic ORF (cut with *PstI* and *BamHI*) and the CYC1 terminator (cut with *BamHI* and *SalI*) together in pCRII (invitrogen) to yield plasmid pNm31323 (SEQ ID NO: 46).

For the final construction of pRN768, pRN656 is cut with *SacI* and *XhoI*, this vector is combined with the insert from pNm31323 cut with *SacI* and *SalI* (1890bp) to produce pRN768.

1.6 Construction of host of yeast cells and transformation with expression constructs

RN1070 is constructed from RN1001. RN1001 is derived from RN1000 (genotype described in WO 2010/074577), has a CEN.PK102-3A background and further has the genotype: *MAT a*, *ura3-52*, *leu2-112*, *gre3::loxP*, *loxP-Ptpi::TALI*, *loxP-Ptpi::RKI1*, *loxP-Ptpi-TKL1*, *loxP-Ptpi-RPE1*, *delta::Padh1XKS1Tcyc1-LEU2*, *delta::URA3-Ptpi-xylA-Tcyc1*.

RN1070 was obtained as an autodiploid of RN1001 by ectopic expression of the HO gene using plasmid pFL39 KanMX-GAL1HO (comprising a G418 resistance marker and the GAL1 promoter fused to HO as described in Teunissen et al., 2002, Appl. Environ. Microbiol. 68: 4780–4787). The forced mating type switch followed by mating with the original cells results in diploid MATa/MAT α cells. The diploid nature was confirmed by specific PCR on the MAT locus. The pFL39 plasmid is cured from the cells by cultivation (>10 generations) in non selective medium (xylose as carbon source) and testing for G418 sensitivity. The genotype of RN1070 is: *MATa /MAT α* ,

ura3-52/ ura3-52, leu2-112/leu2-112, gre3::loxP/ gre3::loxP, loxP-Ptpi::TAL1/ loxP-Ptpi::TAL1, loxP-Ptpi::RKI1/ loxP-Ptpi::RKI1, loxP-Ptpi-TKL1/ loxP-Ptpi-TKL1, loxP-Ptpi-RPE1/ loxP-Ptpi-RPE1, delta::Padh1XKSI Tcyc1-LEU2/ delta::Padh1XKSI Tcyc1-LEU2, delta::URA3-Ptpi-xylA-Tcyc1 delta::URA3-Ptpi-xylA-Tcyc1

RN1030 (n) and RN1100 (2n) were constructed from CEN.PK113-7A. In the haploid laboratory strain CEN.PK113-7A the HIS3 gene is deleted by insertion of the loxP sequence (one step gene deletion method: the loxP-KanMX –loxP construct is flanked by his3 up and downstream sequences). This fragment is integrated into the genome (selected for G418 resistance) CRE mediated recombination resulted in marker removal leaving a footprint of one loxP site). The same method is used to delete in this strain also the LYS2 gene. Deletions are PCR verified. The resulting strain requires both histidine and lysine for growth.

Plasmid pRN792 (SEQ ID NO: 49) contains the N-terminal part of the GAL2 gene fused to the TPI1 promoter sequence. This plasmid contains the kanMX dominant marker. The plasmid lacks 2μ, ARS or CEN sequences and therefore not able to replicate in yeast. By restriction enzyme digestion with *Bst*BI (unique in pRN792) the plasmid is linearized. Transformation of yeast with this fragment results in forced integration in the GAL2 locus by homologous recombination. Cre mediated recombination results in removal of both marker and plasmid sequences. The resulting yeast displays a TPI1 promoter driven GAL2 (over-)expression (qPCR verified).

Next, two Ty constructs were used to transform the GAL2 overexpressing yeast strain. The first construct (the insert of plasmid pRN693, linearized by restriction digestion with *Bam*HI; see SEQ ID NO 50) contains the overexpression constructs for the oxidative part of the pentose phosphate pathway (Ptpi1-TAL1-Tadh1, Ptefl1-TKL1-Tpgi1, Ptdh3- RPE1-Tpgi1 and Ptdh3-RKI1-Tcyc1), in addition to the HIS3 gene. This construct is integrated by homologous recombination at one of the endogenous TY1 elements. Transformants are selected by complementation of the histidine auxotrophy. Similarly the second construct, the insert of pRN755 (linearized by restriction digestion with *Nsi*I; see SEQ ID NO: 51) is integrated by selection for complementation of the lysine auxotrophy. This insert of pRN755 contains, in addition to the LYS2 gene, expression cassettes for the *araA*, -B and -D genes of *Arthrobacter aureescens* as are described in WO 2009/011591 (Phxt7-*araA*-Tpgi1; Ptpi1-*araB*-Tadh1; Ptdh3-*araD*-

Tcyc1). Transformants were selected on mineral medium with arabinose as single carbon source. The best growing transformant RN1030 was selected for further modification. The genotype of RN1030 is: *MATa*, *his3::loxP*, *lys2::loxP*, *loxP-Ptpi-GAL2*, *delta::araA-araB-araD-LYS2*, *delta:: Ptpi1-TAL1-Tadh1+Ptef1-TKL1-Tpgi1+Ptdh3- RPE1-Tpgi1+Ptdh3-RKI1-Tcyc1+HIS3*.

An autodiploid of RN1030 is obtained by ectopic expression of the HO gene using pFL39 KanMX-GAL1 HO as described above (Teunissen et al., 2002, *supra*). The forced mating type switch followed by mating with the original cells results in diploid MATa/MAT α cells. The diploid nature was confirmed by specific PCR on the MAT locus. The pFL39 plasmid is cured from the cells by cultivation (>10 generations) in non selective medium (arabinose as carbon source) and testing for G418 sensitivity. A thus obtained diploid (MATa/MAT α) strain is named RN1100 (genotype is PCR verified). The genotype of RN1100 is: *MATa/MAT α* , *his3::loxP /his3::loxP*, *lys2::loxP/lys2::loxP*, *loxP-Ptpi-GAL2/ loxP-Ptpi-GAL2*, *delta::araA-araB-araD-LYS/ delta::araA-araB-araD-LYS2*, *delta:: Ptpi1-TAL1-Tadh1+Ptef1- TKL1-Tpgi1+Ptdh3-RPE1-Tpgi1+Ptdh3-RKI1-Tcyc1+HIS3/ delta:: Ptpi1-TAL1-Tadh1+Ptef1- TKL1-Tpgi1+Ptdh3- RPE1-Tpgi1+Ptdh3-RKI1-Tcyc1+HIS3*.

RN1070 and RN1100 are transformed with plasmids using the 'Gietz method' (Gietz et al., 1992, Nucleic Acids Res. 20:1425) as indicated in Table 3. Primary selection of transformants is done on mineral medium (YNB + 2% glucose) using antibiotic selection as indicated in Table 4. Table 2 summarizes the strains thus obtained and the enzymes expressed therein.

2. Fermentations with the constructed of yeast cells of the invention

2.1 Experimental set-up

Precultures of strains were prepared by inoculating a frozen glycerol stock culture of the yeast in an YP (Yeast extract Peptone) medium with addition of the sugar glucose (2.5% w/v) at 32 °C and pH 5.5. After 24 h incubation, this culture was used to inoculate the fermenter cultures. Cells were harvested by centrifugation and washed with cold distilled water (dH₂O). Yeast inoculation used was 5 gram dry matter yeast per liter of fermentation medium.

Oxic growth experiments (oxygen present) were performed in a defined mineral medium. Its composition was based on the description by Verduyn et al. (1992, Yeast

8:501-517) but the ammonium in the medium was replaced by 2.3 g/l of urea. To compensate for the reduced sulfate content of this media, 6.6 g of K_2SO_4 per liter was added (Luttik et al., 2000, J Bacteriol. 182:7007-7013). Cell growth was assessed in Erlenmeyer flasks (100 ml with 10 ml medium) in the absence or presence of a carbon
5 source. Flasks were incubated in a rotary shaker (200 rpm) at 32 °C. Cell growth was assessed in time by optical density (700nm) readings of the media. From the OD measurements the dry weight was calculated. For this a linear calibration curve between the optical density and dry weight yeast was used. The dry weight of yeast culture samples were determined using 0.45- μ m-pore-size nitrocellulose filters
10 (Millipore) and a microwave oven (Postma et al, 1989, Appl Environ Microbiol. 55:3214-3220).

Anoxic fermentations (no oxygen present) were performed by using a fermentation setup which can run six fermentations simultaneously. Anoxic batch fermentations were performed in YP medium to which either glucose, xylose, arabinose, galacturonate or glucuronate were added. The pH of the medium was set at
15 pH = 7.0 by adding K_2HPO_4 and KH_2PO_4 resulting in 100 mM phosphate in the YP medium. The temperature during the fermentations was set at 32 °C. The working volumes of the fermentations used in this study were 200 ml.

Samples for analysis of glucose, xylose, arabinose, galacturonate, glucuronate and ethanol were taken after 72 hours of incubation. Ethanol concentrations were monitored by HPLC analysis. Sugars and uronates were determined by HPAEC (Dionex) analysis.

2.2 Results of anoxic fermentations

25 Table 2 summarizes the strains used in the fermentations and the enzymes expressed therein.

Fermentation of strain RN1101 (Table 2) on either glucose (25 mM) or galacturonate (25 mM) or xylose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose or
30 galacturonate or xylose by this strain within 72 hours. The concentrations of ethanol in the glucose, galacturonate or xylose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either galacturonate or xylose.

Fermentation of strain RN1104 (Table 2) on either glucose (25 mM) or glucuronate (25 mM) or xylose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose, glucuronate or xylose by this strain within 72 hours. The concentrations of ethanol in the glucose, glucuronate and xylose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either glucuronate or xylose.

Fermentation of strain RN1105 (Table 2) on either glucose (25 mM) or galacturonate or arabinose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose, galacturonate or arabinose by this strain within 72 hours. The concentrations of ethanol in the glucose, galacturonate and arabinose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either galacturonate or arabinose.

Fermentation of strain RN1108 (Table 2) on either glucose (25 mM) or galacturonate or arabinose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose, galacturonate or arabinose by this strain within 72 hours. The concentrations of ethanol in the glucose, galacturonate and arabinose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either galacturonate or arabinose.

Fermentation of strain RN1109 (Table 2) on either glucose (25 mM) or glucuronate or arabinose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose, glucuronate or arabinose by this strain within 72 hours. The concentrations of ethanol in the glucose, glucuronate and arabinose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either glucuronate or arabinose.

Fermentation of strain RN1111 (Table 2) on either glucose (25 mM) or glucuronate or arabinose (25 mM) added to the YP medium, under the anoxic conditions described above, leads to a full consumption of either glucose, glucuronate or arabinose by this strain within 72 hours. The concentrations of ethanol in the glucose, glucuronate and arabinose media, respectively, at the end of the fermentations are 40 mM (for glucose) and 33 mM for either glucuronate or arabinose.

Fermentations with strains RN1102, RN1103, RN1104, RN1106, RN1107, RN1110, RN1112, RN1113, RN1114, RN1115 en RN1116 gives similar results.

3. Additional fermentation experiments

5 The following experiments were conducted to provide further proof of principle that xylose- or arabinose-fermenting strains, transformed with the plasmids expressing the enzymes necessary to convert uronic acids via the salvage pathway into xylose or arabinose, consume galacturonic or glucuronic acid under aerobic culture conditions. These strains were primary transformants and were not optimized for consumption of
10 uronic acids by evolutionary engineering.

3.1 Material and methods

Medium compositions

Saccharomyces cerevisiae strains were grown on YP with either 20 g/L glucose (YPD), 5 g/L galacturonic acid (YP-galA) or 5 g/L glucuronic acid (YP-glucA). To the
15 transformants antibiotics (G418, hygrogold [HG], phleomycin [phleo], nourseotrycin [nour]) were added to maintain the introduced plasmids according to the scheme in Table 5.

Uronic acid consumption experiments

Saccharomyces cerevisiae strains were pre-cultured on YPD. Fifty µl of the fully
20 grown overnight culture was used to inoculate 4 ml YP-galA or YP-glucA. The scheme of which strain was inoculated to a certain medium is displayed in Table 5. Cultures were incubated at 30° in a rotary shaker (200 rpm).

Cultures were sampled after 144 hours. Samples were filter-sterilized to separate medium from yeast. The filtrate was inserted into the appropriate vials for HPLC
25 analysis. The concentrations of galacturonic acid, glucuronic acid and ethanol in the medium were determined using a Shimadzu ('s-Hertogenbosch, The Netherlands) HPLC system. The system is equipped with column oven CTO-10A-vp and Autoinjector SIL-10AD-vp with a guard column (Bio-Rad H cartridge, Bio-Rad, Hercules, USA) and an Aminex HPX-87H column (300 x 7.8 mm; Bio-Rad). Elution
30 took place at 80 °C with 5 mM H₂SO₄ at 0.6 mL/min. The eluate was monitored using a Refractive Index detector RID-10A (Shimadzu).

Table 5: Inoculation scheme

Strain	YP-galA	YP-glucA	Antibiotics
RN1070 (reference xylose strain)	X	X	-
RN1100 (reference arabinose strain)	X	X	-
RN1102	X		G418, HG, nour
RN1103	X		G418, HG, nour
RN1104		X	G418, nour, phleo
RN1113		X	G418, nour, phleo
RN1114	X		G418, HG, nour

3.2 Results

Galacturonic acid consumption

- 5 After 144 hours of aerobic culture on YP-galA strains, RN1102 and RN1103 and RN1114, displayed a lower amount of galacturonic acid in the medium compared to the medium and both reference strains, RN1070 and RN1100. Translated in consumption RN1102, RN1103 and RN1114 consumed at least 0.2 g/l galacturonic acid after 144 hours of culture. Galacturonic acid concentrations in the YP-galA medium and consumption (compared to the starting medium concentration) are shown in Table 6.

Table 6: Galacturonic acid consumption

strain	144 hours	
	[galacturonic acid] (g/L)	galacturonic acid consumption (g/L)
YP-glucA medium	5.30	-
RN1070 (reference xylose strain)	5.25	0.05
RN1100 (reference arabinose strain)	5.18	0.12
RN1102	5.04	0.26
RN1103	5.09	0.21
RN1114	5.07	0.23

Glucuronic acid consumption

After 144 hours of aerobic culture on YP-glucA strains, RN1104 and RN1113, displayed a lower amount of glucuronic acid in the medium compared to the medium and both reference strains, RN1070 and RN1100. Translated in consumption rate, RN1104 and RN1113 consumed at least 0.2 g/l glucuronic acid after 144 hours of culture. For RN1104 an amount of ethanol was detected in the medium. Glucuronic acid and ethanol concentrations in the YP-glucA medium and consumption (compared to the starting medium concentration) are shown in Table 7.

10 Table 7: Glucuronic acid consumption

strain	144 hours		
	[glucuronic acid] (g/L)	glucuronic acid consumption (g/L)	[ethanol] (g/L)
YP-glucA medium	4.93	-	-
RN1070 (reference xylose strain)	4.95	-0.02	-
RN1100 (reference arabinose strain)	4.87	0.06	-
RN1104	4.67	0.25	0.18
RN1113	4.7	0.23	-

Table 2

Enzyme	2.7.7.64(Ps)	4.1.1.35	2.7.1.46	YihX	2.7.1.44	5.1.3.6	2.7.1.43	5.1.3.5	2.7.7.64(At)	3.1.3.10	3.1.3.23	host	xylose fermenting strain
RN1070	-	-	-	-	-	-	-	-	-	-	-		xylose fermenting strain
RN1100	-	-	-	-	-	-	-	-	-	-	-		arabinose fermenting strain
RN1101	x	x	-	x	x	x	-	-	-	-	-	RN1070	GalUA
RN1102	x	x	-	x	x	x	-	-	x	-	-	RN1070	GalUA
RN1103	-	x	-	x	x	x	-	-	x	-	-	RN1070	GalUA
RN1104	x	x	-	x	-	-	x	-	x	-	-	RN1070	GlcUA
RN1105	x	x	x	-	x	x	-	x	-	-	-	RN1100	GalUA
RN1106	x	x	x	-	x	x	-	x	x	-	-	RN1100	GalUA
RN1107	-	x	x	-	x	x	-	x	x	-	-	RN1100	GalUA
RN1108	x	x	-	x	x	x	-	x	-	-	-	RN1100	GalUA
RN1109	x	x	x	-	-	-	x	x	x	-	-	RN1100	GlcUA
RN1110	x	x	x	-	-	-	x	x	-	-	-	RN1100	GlcUA
RN1111	x	x	-	x	-	-	x	x	x	-	-	RN1100	GlcUA
RN1112	x	x	-	x	-	-	x	x	-	-	-	RN1100	GlcUA
RN1113	-	x	-	x	-	-	x	x	x	-	-	RN1100	GlcUA
RN1114	x	x	-	x	x	x	-	-	-	x	-	RN1070	GalUA
RN1115	x	x	-	x	x	x	-	-	-	x	-	RN1070	GalUA
RN1116	x	x	-	x	x	x	-	-	-	-	x	RN1070	GalUA

Claims

1. A yeast cell comprising:
 - a) genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide;
 - b) a gene coding for enzyme with the ability to convert a pentose-1P into a pentose; and,
 - c) genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate.
2. A yeast cell according to claim 1, wherein:
 - a) the genes coding for enzymes of a plant salvage pathway with the ability to convert a D-uronic acid into a pentose-1-phosphate and carbon dioxide are genes coding for enzymes with the ability to convert at least one of galacturonic acid and glucuronic acid into carbon dioxide and at least one of xylose-1-phosphate and arabinose-1-phosphate;
 - b) the gene coding for enzyme with the ability to convert a pentose-1P into a pentose is a gene coding for an enzyme with the ability to convert at least one of xylose-1-phosphate into xylose and arabinose-1-phosphate into arabinose; and,
 - c) the genes coding for enzymes with the ability to convert a pentose into at least one of xylulose and xylulose 5-phosphate are selected from the group consisting of genes coding for enzymes with the ability to i) directly isomerise xylose into xylulose, and, ii) convert L-arabinose into D-xylulose 5-phosphate.
3. A yeast cell according to claim 1 or 2, wherein the yeast cell has the ability of metabolizing glucuronic acid, and wherein the yeast cell comprises genes coding for:
 - a) a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a broader substrate specificity;
 - b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64);
 - c) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and
 - d) at least one of:
 - i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and,

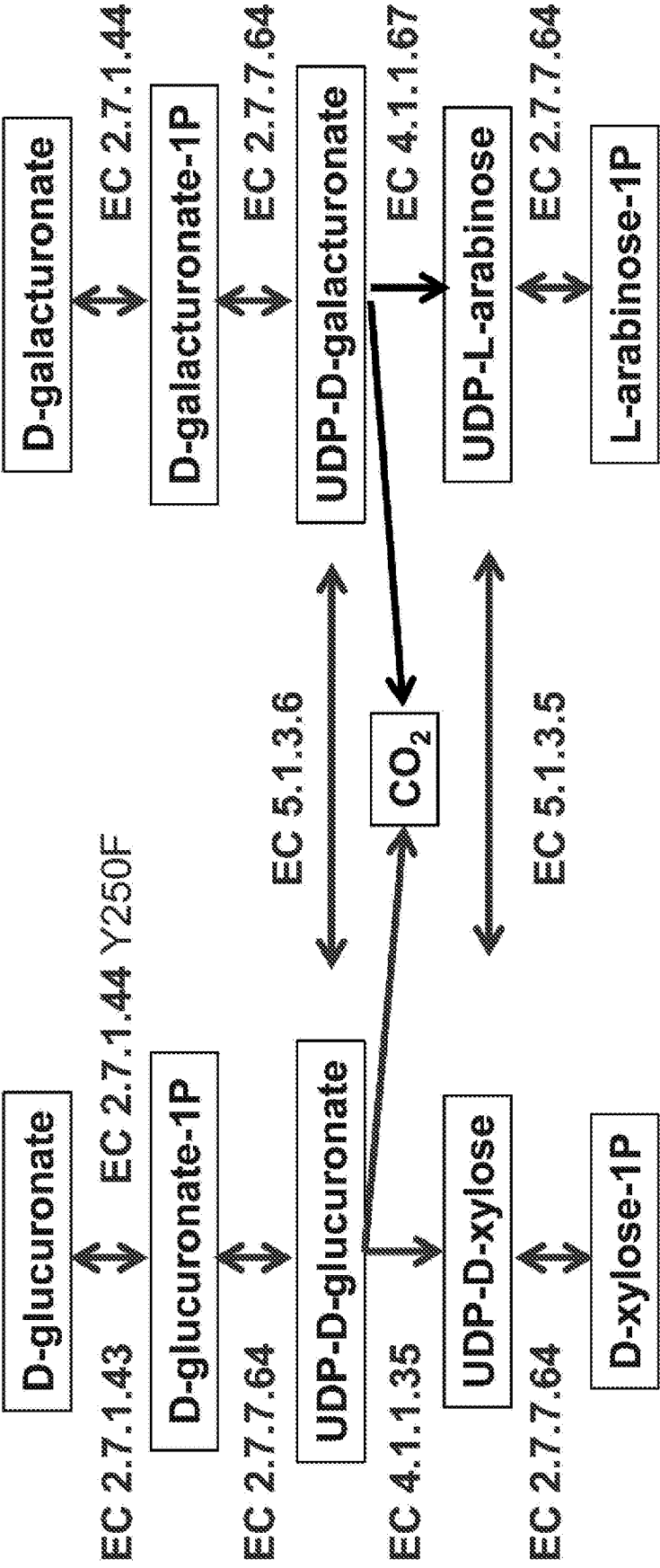
ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).

- 5 4. A yeast cell according to claim 1 or 2, wherein the yeast cell has the ability of metabolizing galacturonic acid, and wherein the yeast cell comprises genes coding for:
- a) a galacturonokinase (EC 2.7.1.44);
 - b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64);
 - c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6);
 - 10 d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and
 - e) at least one of:
 - i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and,
 - 15 ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).
5. A yeast cell according to claim 1 or 2, wherein the yeast cell has the ability of metabolizing galacturonic acid, and wherein the yeast cell comprises genes coding for:
- 20 a) a glucuronokinase (EC 2.7.1.43) or a modified galacturonokinase (EC 2.7.1.44) having a broader substrate specificity;
 - b) a UTP-monosaccharide-1-phosphate uridylyltransferase (EC 2.7.7.64);
 - c) a UDP-glucuronate-4-epimerase (EC 5.1.3.6);
 - d) a UDP-glucuronic acid decarboxylase (EC 4.1.1.35); and
 - 25 e) at least one of:
 - i) a sugar phosphatase (EC 3.1.3.23) or a sugar-1-phosphatase (EC 3.1.3.10) and a xylose isomerase (EC 5.3.1.5); and,
 - ii) a UDP-D-xylose 4-epimerase (EC 5.1.3.5), an arabinose kinase (EC 2.7.1.46), and a L-arabinose isomerase (EC 5.3.1.3), a L-ribulokinase (EC 2.7.1.16) and a
 - 30 L-ribulose-5-phosphate 4-epimerase (EC 5.1.3.4).
6. A yeast cell according to any one of claims 3 - 5, wherein at least one of:

- a) the gene coding for an enzyme with galacturokinase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 1;
- b) the gene coding for an enzyme with glucuronokinase activity comprises a nucleotide
5 sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 2;
- c) the gene coding for an enzyme with UTP-monosaccharide-1-phosphate uridylyltransferase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with at least one of SEQ ID
10 NO: 3 and 4;
- d) the gene coding for an enzyme with UDP-D-glucuronate 4-epimerase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 5;
- e) the gene coding for an enzyme with UDP-glucuronic acid decarboxylase activity
15 comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 6;
- f) the gene coding for an enzyme with UDP-D-xylose 4-epimerase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 7;
- 20 g) the gene coding for an enzyme with arabinose kinase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid sequence identity with SEQ ID NO: 8; and,
- h) the gene coding for an enzyme with sugar phosphatase activity comprises a nucleotide sequence coding for an amino acid sequence with at least 60% amino acid
25 sequence identity with SEQ ID NO: 9.

7. A yeast cell according to any of the preceding claims, wherein the cell is a yeast cell selected from the genera *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces*, *Brettanomyces*, and
30 *Yarrowia*, in particular the cell is a yeast cell selected from the species *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* and *Schizosaccharomyces pombe*.

8. Use of a yeast cell according to any of the preceding claims for the preparation of a fermentation product, wherein preferably the fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, butanols and isoprenoid-derived products.
- 5
9. A process for producing a fermentation product, whereby the process comprises the steps of:
- a) fermenting a medium with a yeast cell as defined in any one of claim 1 - 8, whereby the medium contains or is fed with a source of a uronic acid and whereby the yeast cell
- 10 ferments the uronic acid to the fermentation product; and optionally,
- b) recovery of the fermentation product,
- wherein preferably the fermentation product selected from the group consisting of ethanol, lactic acid, 3-hydroxy-propionic acid, acrylic acid, 1,3-propane-diol, a butanol and an isoprenoid-derived product.
- 15
10. A process according to claim 9, wherein the uronic acid is at least one of galacturonic acid and glucuronic acid
11. A process according to claim 9 or 10, wherein the medium further contains or is
- 20 fed with a source of at least one of a hexose, a pentose, acetic acid and glycerol.
12. A process according to any one of the claims 9 - 11, wherein the medium contains or is fed with at least one of a hydrolyzed pectin-rich residue and hydrolyzed lignocellulosic biomass.
- 25
13. A process according to any one of the claims 9 - 12, wherein the yeast cell ferments under anaerobic conditions.
14. A process according to any one of the claims 9 - 13, wherein the fermentation
- 30 product is ethanol.



INTERNATIONAL SEARCH REPORT

International application No
PCT/NL2012/050151

A. CLASSIFICATION OF SUBJECT MATTER

INV. C12N9/00 C12N9/12 C12N9/88 C12N9/90 C12P7/06
ADD.

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

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C12N C12P

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

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EP0-Internal, BIOSIS, EMBASE, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KOTAKE TOSHIHISA ET AL: "Generation of nucleotide sugars for biomass formation in plants", PLANT BIOTECHNOLOGY, vol. 27, no. 3, 2010, pages 231-236 URL, XP002675164, ISSN: 1342-4580 the whole document ----- -/--	1-14



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents :

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

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

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

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

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

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

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

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

"&" document member of the same patent family

Date of the actual completion of the international search

3 May 2012

Date of mailing of the international search report

25/05/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Roscoe, Richard

INTERNATIONAL SEARCH REPORT

International application No

PCT/NL2012/050151

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	A. M. PIESLINGER ET AL: "Cloning of Glucuronokinase from Arabidopsis thaliana, the Last Missing Enzyme of the myo-Inositol Oxygenase Pathway to Nucleotide Sugars", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 285, no. 5, 29 January 2010 (2010-01-29), pages 2902-2910, XP55014095, ISSN: 0021-9258, DOI: 10.1074/jbc.M109.069369 the whole document	1-14
A	----- WO 2010/074577 A1 (ROYAL NEDALCO B V [NL]; TEUNISSEN ALOYSIUS WILHELMUS RUDOLPHUS HUBERTU) 1 July 2010 (2010-07-01) the whole document	1-14
A	----- PETER RICHARD ET AL: "d-Galacturonic acid catabolism in microorganisms and its biotechnological relevance", APPLIED MICROBIOLOGY AND BIOTECHNOLOGY, SPRINGER, BERLIN, DE, vol. 82, no. 4, 22 January 2009 (2009-01-22), pages 597-604, XP019705514, ISSN: 1432-0614 the whole document	1-14
A	----- VAN MARIS A J A ET AL: "Alcoholic fermentation of carbon sources in biomass hydrolysates by Saccharomyces cerevisiae: current status", ANTONIE VAN LEEUWENHOEK, SPRINGER NETHERLANDS, DORDRECHT, vol. 90, no. 4, 1 November 2006 (2006-11-01), pages 391-418, XP002500454, ISSN: 0003-6072, DOI: 10.1007/S10482-006-9085-7 [retrieved on 2006-10-11] the whole document	1-14
A	----- WO 2010/072902 A1 (VALTION TEKNILLINEN [FI]; BOER HANNU [FI]; HILDICH SATU [FI]; RICHARD) 1 July 2010 (2010-07-01) the whole document	1-14

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/NL2012/050151

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010074577 A1	01-07-2010	CA 2746818 A1 CN 102307987 A EP 2367928 A1 US 2011318790 A1 WO 2010074577 A1	01-07-2010 04-01-2012 28-09-2011 29-12-2011 01-07-2010
WO 2010072902 A1	01-07-2010	EP 2376620 A1 US 2012045804 A1 WO 2010072902 A1	19-10-2011 23-02-2012 01-07-2010