



- (51) **International Patent Classification:**
C07K 14/395 (2006.01) *C12R 1/865* (2006.01)
C12N 1/18 (2006.01)
- (21) **International Application Number:**
PCT/EP2015/066610
- (22) **International Filing Date:**
21 July 2015 (21.07.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
14178396.9 24 July 2014 (24.07.2014) EP
14187968.4 7 October 2014 (07.10.2014) EP
- (71) **Applicant:** DSM IP ASSETS B.V. [NL/NL]; Het Overloon 1, NL-6411 TE Heerlen (NL).
- (72) **Inventors:** DE WAAL, Paulus Petrus; P.O. Box 4, NL-6100 AA Echt (NL). DE JONG, René Marcel; P.O. Box 4, NL-6100 AA Echt (NL). KLAASSEN, Paul; P.O. Box 4, NL-6100 AA Echt (NL).
- (74) **Agent:** KLEIBORN, Paul Erik; DSM Intellectual Property, P.O. Box 4, 6100 AA Echt (NL).

- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

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

(54) **Title:** YEAST CELL WITH IMPROVED PENTOSE TRANSPORT

(57) **Abstract:** The invention relates to a yeast cell comprising at least one functional pentose conversion pathway from pentose to fermentation product, wherein the yeast cell comprises one or more transporter construct, wherein the transporter construct comprises DNA sequences: a) an transporter promoter b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and; c) a terminator, wherein the transporter promoter is regulated by cellular regulation induced by SNF3-and/or RGT2-signaling.



YEAST CELL WITH IMPROVED PENTOSE TRANSPORT

5

Field of the invention

The invention is directed to fermentative production of alcohols and chemicals, including ethanol, and processes for improving alcohol fermentation employing yeasts with improved pentose conversion properties.

10

Background of the invention

There is much interest in the use of lignocellulosic biomass as an alternative substrate for the production of bio-based chemicals and energy production (such as e.g. biofuels). Numerous initiatives are employed for the development and design of technologies enabling the conversion of lignocellulosic feedstocks into valuable products, both by academia and industry.

Due to the complex structure of lignocellulosic materials, pretreatment of the feedstock and hydrolysis of hemicellulose and cellulose into five- and six-carbon sugars has to be carried out. Hydrolysis takes place prior to, or concurrently with, the fermentation. During fermentation, the sugars mixtures are converted into the valuable products, such as e.g. bio-ethanol.

Thus, hydrolysates of lignocellulosic biomass are a valuable feedstock for the production of biofuels and chemicals that provide both 5- and 6-carbon sugars. However, these hydrolysates can contain compounds that are inhibitory to the growth and metabolism of microorganisms.

Saccharomyces cerevisiae has a long history in the ethanol production industry due to its proven ability to produce high ethanol titers rapidly. The ethanol industry is adapted to the use of *Saccharomyces cerevisiae* as an ethanologen, basically due to its relatively high level resistance to fermentation inhibitors and other stresses. However, *Saccharomyces cerevisiae* cannot metabolize pentoses (like xylose and arabinose) by nature. Thus, the amount of ethanol that can be produced from lignocellulosic hydrolysates is limited because the five-carbon sugars are not readily usable without certain genetic modifications and without some processing to improve inhibitor activity

35

Methods of making microorganisms that express enzymes allowing the conversion of pentoses are known in the art. For example, International Publication No. WO 2009/109630, which is hereby incorporated by reference in its entirety, illustrates the production of pentose sugar-fermenting cells that express xylose isomerase which enables the cell to ferment xylose into ethanol. Several other groups have introduced heterologous metabolic pathways which allow pentose catabolism in *Saccharomyces cerevisiae*. The mere introduction of these pathways is usually in itself not sufficient to allow pentose conversion. Subsequently, a process called “adaptive evolution” or “evolutionary engineering” is needed to enable pentose conversion and/or to speed up the rate of pentose metabolism.

During evolutionary engineering, spontaneous genomic changes that positively affect the growth rate on pentose(s) are selected for. The working hypothesis is that cells that grow faster on pentoses are also able to anaerobically ferment pentoses in sugar mixtures at a higher rate. Cell(s) which have picked up a beneficial genomic modification, allowing faster growth on the respective pentose, will eventually outgrow the other cells in the population.

As stated above, a major issue in the conversion of saccharified cellulosic biomass is the utilization of D-xylose, which is the most abundant pentose sugar present in such feedstocks. Although solutions in bioengineering have been successful in (re-)programming *Saccharomyces cerevisiae* into a xylose-fermenting organism, in typical fermentations first glucose is converted before xylose is metabolized. Since this process occurs late during exponential growth, it often results in incomplete conversion of the xylose. For a more robust conversion process, it is desirable that the xylose and glucose are co-consumed. In addition, fast and complete consumption of all biomass sugars will have a positive impact on the product yield.

Several factors may contribute to the incomplete pentose sugar utilization, among which inefficient transport of xylose across the plasma membrane, which takes place via the endogenous hexose transporters of *Saccharomyces cerevisiae*. Xylose transport is strongly affected by glucose. One strategy to overcome this problem has been to introduce heterologous xylose transporters (see e.g. Leandro et al, 2008; Runquist et al, 2010; Young et al, 2011) but this approach has only met little success because of low transport rates, inefficient plasma membrane targeting and protein instability as such heterologous transporters are not well integrated in the cellular regulatory network of carbon metabolism of *Saccharomyces cerevisiae*. In addition, all known xylose

transporters that have been successfully expressed in *Saccharomyces cerevisiae* are not selective for xylose. As the affinity for glucose is also higher than for xylose, competitive inhibition of xylose transport by glucose still takes place.

Through the use of *in vivo* engineering and directed mutagenesis, however, endogenous hexose transporters were converted into a specific xylose transporters that are no longer inhibited by glucose (PCT/EP2014/061635; Farwick et al, 2014). In a xylose-fermenting yeast, these transporters allow for efficient uptake of xylose even in the presence of high glucose and as a result both sugars are co-metabolized resulting in complete sugar utilization in a shorter time, resulting in improved bioethanol yield and productivity. Since the transporters involved are endogenous to the yeast cell, they are well embedded in the cellular metabolic network and properly targeted to the plasma membrane.

The latter also implicates, however, that the expressed endogenous pentose transporters are subject to cellular regulation regimes, such as endocytosis. The yeast cell possesses multiple endogenous hexose transporters with different affinities for the substrate, which allows the yeast cells to grow well over a wide range of substrate concentrations in the extracellular environment. The extracellular hexose concentration is constantly sensed by two glucose sensors, Rgt2 and Snf3. Rgt2 has a low affinity for glucose, while Snf3 has a high affinity for glucose (Özcan et al, 1998; Roy and Kim, 2014). These two cell membrane glucose sensors in *Saccharomyces cerevisiae* are thought to be evolutionary derived from genuine glucose transporters, but appeared to have lost their ability to actually transport sugars across the plasma membrane into the cytoplasm of the yeast cell. The amount of glucose available in the extracellular environment of *Saccharomyces cerevisiae* dictates the levels of the glucose sensors at the cell surface, and ultimately, through a complex intracellular regulatory network (reviewed by Horák, 2013), the composition of the hexose transporter landscape at the plasma membrane. Yeast cells express only the sugar transporters that are most appropriate for the amounts of glucose available in the extracellular environment. Both sensors (Rgt2 and Snf3) and the hexose transporters are internalized (endocytosis) and degraded in the vacuole if the external glucose concentration no longer matches their affinity for glucose.

Rewiring sugar transport preference and kinetics, as described above, turning hexose transporters into pentose-specific transporters, are still subject to regulation at the protein level, dictated by the external glucose concentration. Thus, although the

substrate preference of these transporters have dramatically increased, the timely presence of these transporters at the cell surface needs to be adjusted.

Summary of the invention

An object of the invention is to provide solutions which allow for expression of pentose-specific sugar transporters which are expressed at the right time during the fermentation of lignocellulosic hydrolysates, resulting in improved hexose and pentose co-consumption characteristics of the yeast cell. Another object is to provide timely presence of pentose transporters at the cell surface. Another object is to embed the pentose transporters in the yeast regulatory network.

One or more of these objects are attained according to the invention. According to the present invention, there is provided a yeast cell comprising at least one functional pentose conversion pathway from pentose to fermentation product, wherein the yeast cell comprises one or more transporter construct, wherein the transporter construct comprises DNA sequences:

- a) a transporter promoter
 - b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 - c) a terminator,
- wherein the transporter promoter is regulated by cellular regulation initiated by *SNF3*- and/or *RGT2*-signaling.

This will be described in more detail below and will be illustrated by examples below.

Brief description of the drawings

FIG. 1 shows a schematic representation of the integration of different fragments, containing different elements, into the genome of *Saccharomyces cerevisiae*

FIG. 2 shows the sugar consumption and ethanol production profiles of the anaerobic growth tests for (A) DS65307-1T, (B) DS65307-2T, (C) DS65307-3T, (D), DS65307-4T, (E) DS65307-7T, and (F) the ratio of the consumption rate of xylose over glucose (Q_x/Q_g) was plotted against the residual glucose concentration in the medium for each tested strain.

Brief description of the sequence listing

SEQ ID NO: 1: sequence of the *HXT11*-gene (synthetic; *HXT11* (N366N) wt);
 SEQ ID NO:2: sequence of the variant *HXT11* (N366M) (synthetic; *HXT11* (N366M));
 SEQ ID NO:3: sequence of the variant *HXT11* (N366T) (synthetic; *HXT11* (N366T));
 5 SEQ ID NO: 4: *S. cerevisiae* strain DS65307 *HXT1* promoter;
 SEQ ID NO: 5: *S. cerevisiae* strain DS65307 *HXT2* promoter;
 SEQ ID NO: 6: *S. cerevisiae* strain DS65307 *HXT3* promoter;
 SEQ ID NO: 7: *S. cerevisiae* strain DS65307 *HXT4* promoter;
 SEQ ID NO: 8: *S. cerevisiae* strain DS65307 *HXT5* promoter;
 10 SEQ ID NO: 9: *S. cerevisiae* strain DS65307 *HXT6* promoter;
 SEQ ID NO: 10: *S. cerevisiae* strain DS65307 *HXT7* promoter;
 SEQ ID NO: 11: *HXT1* Promoter (+Bsal) Forward primer;
 SEQ ID NO: 12: *HXT1* Promoter (+Bsal) Reverse primer;
 SEQ ID NO: 13: *HXT2* Promoter (+Bsal) Forward primer;
 15 SEQ ID NO: 14: *HXT2* Promoter (+Bsal) Reverse primer;
 SEQ ID NO: 15: *HXT3* Promoter (+Bsal) Forward primer;
 SEQ ID NO: 16: *HXT3* Promoter (+Bsal) Reverse primer;
 SEQ ID NO: 17: *HXT4* Promoter (+Bsal) Forward primer;
 SEQ ID NO: 18: *HXT4* Promoter (+Bsal) Reverse primer;
 20 SEQ ID NO: 19: *HXT5* Promoter (+Bsal) Forward primer;
 SEQ ID NO: 20: *HXT5* Promoter (+Bsal) Reverse primer;
 SEQ ID NO: 21: *HXT6* Promoter (+Bsal) Forward primer;
 SEQ ID NO: 22: *HXT67* Promoter (+Bsal) Reverse primer;
 SEQ ID NO: 23: *HXT7* Promoter (+Bsal) Forward primer;
 25 SEQ ID NO: 24: *HXT11*-PNTORF Forward primer;
 SEQ ID NO: 25: *HXT11*-PNTORF Reverse primer;
 SEQ ID NO: 26: *S. cerevisiae* DS65307 strain *HXT1* terminator;
 SEQ ID NO: 27: *S. cerevisiae* DS65307 strain *HXT2* terminator;
 SEQ ID NO: 28: *S. cerevisiae* DS65307 strain *HXT3* terminator;
 30 SEQ ID NO: 29: *S. cerevisiae* DS65307 strain *HXT4* terminator;
 SEQ ID NO: 30: *S. cerevisiae* DS65307 strain *HXT5* terminator;
 SEQ ID NO: 31: *S. cerevisiae* DS65307 strain *HXT6* terminator;
 SEQ ID NO: 32: *S. cerevisiae* DS65307 strain *HXT7* terminator;
 SEQ ID NO: 33: *HXT1* Terminator (+Bsal) Forward primer;
 35 SEQ ID NO: 34: *HXT1* Terminator (+Bsal) Reverse primer;

- SEQ ID NO: 35: *HXT2* Terminator (+Bsal) Forward primer;
SEQ ID NO: 36: *HXT2* Terminator (+Bsal) Reverse primer;
SEQ ID NO: 37: *HXT3* Terminator (+Bsal) Forward primer;
SEQ ID NO: 38: *HXT3* Terminator (+Bsal) Reverse primer;
5 SEQ ID NO: 39: *HXT4* Terminator (+Bsal) Forward primer;
SEQ ID NO: 40: *HXT4* Terminator (+Bsal) Reverse primer;
SEQ ID NO: 41: *HXT5* Terminator (+Bsal) Forward primer;
SEQ ID NO: 42: *HXT5* Terminator (+Bsal) Reverse primer;
SEQ ID NO: 43: *HXT6* Terminator (+Bsal) Forward primer;
10 SEQ ID NO: 44: *HXT6* Terminator (+Bsal) Reverse primer;
SEQ ID NO: 45: *HXT7* Terminator (+Bsal) Forward primer;
SEQ ID NO: 46: *HXT7* Terminator (+Bsal) Reverse primer;
SEQ ID NO: 47: *HXT1* promoter Forward primer;
SEQ ID NO: 48: *HXT2* promoter Forward primer;
15 SEQ ID NO: 49: *HXT3* promoter Forward primer;
SEQ ID NO: 50: *HXT4* promoter Forward primer;
SEQ ID NO: 51: *HXT5* promoter Forward primer;
SEQ ID NO: 52: *HXT6* promoter Forward primer;
SEQ ID NO: 53: *HXT7* promoter Forward primer;
20 SEQ ID NO: 54: primer sequence conA Reverse primer;
SEQ ID NO: 55: kanMX-marker conferring resistance to G418 flanked by connector
sequences a (5'-end) and b (3'-end);
SEQ ID NO: 56: hphMX-marker conferring resistance to hygromycin B flanked by
connector sequences a (5'-end) and b (3'-end);
25 SEQ ID NO: 57: natMX-marker conferring resistance to nourseotrycin flanked by connector
sequences a (5'-end) and b (3'-end);
SEQ ID NO: 58: zeoMX-marker conferring resistance to phleomycin flanked by connector
sequences a (5'-end) and b (3'-end);
SEQ ID NO: 59: amdSYM-marker conferring growth on acetamide as sole nitrogen source
30 flanked by connector sequences a (5'-end) and b (3'-end);
SEQ ID NO: 60: connector a-marker forward primer;
SEQ ID NO: 61: marker-connector b reverse primer;
SEQ ID NO: 62: connector a-amdSYM marker forward primer;
SEQ ID NO: 63: amdSYM-connector b reverse primer;

SEQ ID NO: 64: *S. cerevisiae TDH3* promoter synthetic DNA sequence flanked by *BsaI* sites);

SEQ ID NO: 65: Dasher GFP (+*BsaI*) Forward primer;

SEQ ID NO: 66: Dasher GFP (+*BsaI*) Reverse primer;

5 SEQ ID NO: 67: *S. cerevisiae ADH1* terminator synthetic DNA sequence;

SEQ ID NO: 68: connector b Forward primer;

SEQ ID NO: 69: connector 3 Reverse primer;

SEQ ID NO: 70: connector 3-terminator *HXT1* forward primer to generate CAS29;

SEQ ID NO: 71: terminator *HXT1* reverse primer to generate CAS29;

10 SEQ ID NO: 72: connector 3-terminator *HXT2* forward primer to generate CAS30;

SEQ ID NO: 73: terminator *HXT2* reverse primer to generate CAS30;

SEQ ID NO: 74: connector 3-terminator *HXT3* forward primer to generate CAS31;

SEQ ID NO: 75: terminator *HXT3* reverse primer to generate CAS31;

SEQ ID NO: 76: connector 3-terminator *HXT4* forward primer to generate CAS32;

15 SEQ ID NO: 77: terminator *HXT4* reverse primer to generate CAS32;

SEQ ID NO: 78: connector 3-terminator *HXT5* forward primer to generate CAS33;

SEQ ID NO: 79: terminator *HXT5* reverse primer to generate CAS33;

SEQ ID NO: 80: connector 3-terminator *HXT6* forward primer to generate CAS34;

SEQ ID NO: 81: terminator *HXT6* reverse primer to generate CAS34;

20 SEQ ID NO: 82: connector 3-terminator *HXT7* forward primer to generate CAS35;

SEQ ID NO: 83: terminator *HXT7* reverse primer to generate CAS35;

SEQ ID NO: 84: Forward primer RT-Q-PCR *HXT11*, *HXT11* (N366M), and *HXT11*(N366T);

SEQ ID NO: 85: Reverse primer RT-Q-PCR *HXT11*;

SEQ ID NO: 86: Reverse primer RT-Q-PCR *HXT11* (N366M);

25 SEQ ID NO: 87: Reverse primer RT-Q-PCR *HXT11*(N366T).

Detailed description of the invention

Throughout the present specification and the accompanying claims, the words
 30 "comprise" and "include" and variations such as "comprises", "comprising", "includes"
 and "including" are to be interpreted inclusively. That is, these words are intended to
 convey the possible inclusion of other elements or integers not specifically recited, where
 the context allows. The articles "a" and "an" are used herein to refer to one or to more
 than one (i.e. to one or at least one) of the grammatical object of the article. By way of
 35 example, "an element" may mean one element or more than one element.

The invention thus relates to a yeast cell comprising at least one functional pentose conversion pathway from pentose to fermentation product, wherein the yeast cell comprises one or more transporter construct, wherein the transporter construct comprises DNA sequences:

- a) a transporter promoter
 - b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 - c) a terminator,
- wherein the transporter promoter is regulated by cellular regulation induced by *SNF3*- and/or *RGT2*-signaling.

The transporter promoter

The first and any following transporter promoter is regulated by cellular regulation induced by *SNF3*- and/or *RGT2*-signaling.

As used herein, the term "promoter" refers to a nucleic acid fragment that functions to control the transcription of one or more genes, located upstream with respect to the direction of transcription of the transcription initiation site of the gene, and is structurally identified by the presence of a binding site for DNA-dependent RNA polymerase, transcription initiation sites and any other DNA sequences known to one of skilled in the art. A "constitutive" promoter is a promoter that is active under most environmental and developmental conditions. An "inducible" promoter is a promoter that is active under environmental or developmental regulation. According to an embodiment of the invention the first transporter is a inducible promoter, either native or prepared synthetically, that are regulated by hexose concentration. Examples of such promoters are the promoters of *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT6*, *HXT7* and *GAL2*, or other promoters which show similar expression profiles and which are regulated by cellular regulation mechanisms after signaling is induced by glucose sensors *Snf3/Rgt2*, *cAMP/PKA*, or *Mig1/Snf1/Hxk2* (for review on glucose sensing and downstream signaling, see e.g. Rodkaer and Faergeman, 2014).

Those skilled in the art can perform experiments to identify and characterize previously unidentified promoters eliciting similar expression profiles as those regulated by the above mentioned cellular regulation mechanisms (again see Rodkaer and Faergeman, 2014). By performing either batch fermentations, or chemostat cultivations on hexoses sensed by the above mentioned glucose sensors, samples can be taken at either different hexose concentrations, or different growth rates of the micro-organism, respectively. From the cells in the samples RNA can be extracted. The RNA can be

prepared in a suitable manner to perform experiments to determine the differential expression of genes by using techniques such as hybridization to microarray chips (e.g. Affymetrix format) or more state of the art methodologies such as RNA sequencing. Promoters of the differentially expressed genes could then be suitable promoters. Such promoters herein are included as transport promoter.

Those skilled in the art could also prepare a hexose-inducible promoter, more preferably glucose-inducible promoter, by preparing such a promoter synthetically that it contains in the appropriate part of the DNA sequence of the promoter binding sites for transcription factors regulated by the signaling induced by hexoses sensed by glucose sensors Snf3/Rgt2, e.g. Rgt1. Binding sites in the DNA sequence for specific transcription factors are retrieved by those skilled in the art from publicly accessible databases such as Yeastract (<http://www.yeastract.com/>). Such promoters herein are included as transport promoter.

The pentose transporter

A pentose transporter is herein a transporter that has a ratio K_m pentose (i.e. xylose or arabinose) to K_m glucose that is 1.5 or less, preferably 1.0 or less. The last indicates that the pentose transporter has a higher affinity to pentose (xylose and/or arabinose) than to hexose.

Those skilled in the art know how to determine the K_m (Michaelis constant), or rather a measure of the affinity of an enzyme for a ligand, or even more specifically of a transporter for a sugar, by conducting an experiment measuring enzyme kinetics with increasing amounts of radio-labeled ligand (even more specifically ^{14}C -xylose or ^{14}C -glucose) in time to determine reaction velocities (or more specifically sugar uptake velocities). Also by applying the Michaelis-Menten kinetics model the maximal velocity ($V_{\max}$) may be calculated.

In an embodiment, the transporter is capable of transporting pentose across the cell membrane. In an embodiment, b) is a ORF of an exogenous pentose transporter. In another embodiment b) is an ORF of a mutant of an endogenous hexose transporter. In an embodiment b) is an ORF of a mutant of any hexose transporter. It may e.g. be a HXT mutant having reduced affinity for glucose or higher affinity for pentose (xylose and/or arabinose), chosen from the group consisting of *GAL2*, *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT6*, and *HXT7*.

In an embodiment the ORF (Open Reading Frame) of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell

is a mutant *HXT11*, which has a reduced affinity for glucose compared to native *HXT11* (SEQ ID NO: 1). In an embodiment thereof, has mutation N336M (DNA SEQ ID NO: 2) or N366T (DNA SEQ ID NO: 3). In an embodiment the transporter construct is integrated into the yeast cell genome.

Suitable pentose transporter are pentose transporters with Km ratio ($K_m\text{-xylose}/K_m\text{-glucose}$) of 1.5 or lower, preferably 1.0 or lower. Such pentose transporters are disclosed in PCT/EP2014/061635. Examples are mutant *GAL2*, mutant *HXT11*, mutant *HXT36*, mutant *HXT1* (N370x), mutant *HXT2* (N361x). Examples of some transporters with Km values and Km ratio are given in table 1, the suitable pentose transporters are those with with Km ratio of 1.5 or lower.

In an embodiment, the pentose transporter sequences can contain one or more of the following motifs:

- a) G-R-x(3)-G-x(3)-G-x(11)-E-x(5)-[LIVM]-R-G-x(12)-[GA];
- b) R-x(14)-G-x(2)-Y-x(2)-[YF]-[YF]-[GSAL] and/or
- c) V-x(15)-[GNR]-[RH]-R-x(2)-[LM]-x(2)-[GA]

Motif (a) is corresponds to residues 179-221 in *Gal2*; motif (b) is corresponds to residues 330-353 in *Gal2*; motif (c) is corresponds to residues 375-399 in *Gal2*.

The terminator

The terminator may be any terminator that performs its function. In a transformed host cell, the 3'-end of the nucleotide acid sequence encoding the pentose transporter is preferably is operably linked to a transcription terminator sequence. Preferably the terminator sequence is operable in a host cell of choice, such as e.g. the yeast species of choice. In any case the choice of the terminator is not critical; it may e.g. be from any yeast gene, although terminators may sometimes work if from a non-yeast, eukaryotic, gene. Preferably, such terminators are combined with mutations that prevent nonsense mediated mRNA decay in the host transformed host cell (see for example: Shirley et al., 2002, Genetics 161:1465-1482).

In an embodiment the native terminator sequences of HXT genes are used. The advantage is that the pentose transporter ORF is expressed in a natural yeast context.

The transporter construct

The transporter construct comprises the DNA sequences:

- a) a transporter promoter

- b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 c) a terminator,
 wherein the transporter promoter is regulated by cellular regulation induced by *SNF3*- and/or *RGT2*-signaling.

In an embodiment, the yeast cell comprises the DNA sequences:

- a) an endogenous *HXT_n*-promoter
 b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 c) a *HXT_n*-terminator

wherein n is 1,2,3,4,6 or 7 and

wherein a), b), c) is designated as P_{HXT_n}-PNT-T_{HXT_n}.

In an embodiment the the transporter construct comprises:

- a) P_{HXT1}-PNT_a-T_{HXT1} or P_{HXT3}-PNT_a-T_{HXT3}; b) a construct P_{HXT2}-PNT_b-T_{HXT2} or P_{HXT4}-PNT_b-T_{HXT4} and c) a construct P_{HXT6}-PNT_c-T_{HXT6} or P_{HXT7}-PNT_c-T_{HXT7}. Herein P_{HXT1} means a promoter that is connected to the *HXT1* gene in native yeast. PNT_a means any pentose transporter. PNT_a, PNT_b and PNT_c may be the same or different pentose transporters.

Other suffixes have meaning analogous to the foregoing. In an embodiment the the transporter construct comprises:

- a) a construct P_{HXT1}-PNT_a-T_{HXT1};
 b) a construct P_{HXT2}-PNT_b-T_{HXT2} and
 c) a construct P_{HXT6}-PNT_c-T_{HXT6}

This construct with promoters (HXT1,2,6) is advantageous in yeast since it allows expression of suitable yeast transporters PNT_a to PNT_c at all common concentrations of glucose, which results in improved co-consumption of pentose and glucose as defined hereafter. The other promoter combinations below (HXT1,2,7) , (HXT 1,4,6), (HXT 1,4,7), (HXT 3,2,6), (HXT 3,2,7), (HXT 3,4,6) and (HXT 3,4,7) give similar advantages.

In an embodiment the the transporter construct comprises:

- a) a construct P_{HXT1}-PNT_a-T_{HXT1};
 b) a construct P_{HXT2}-PNT_b-T_{HXT2} and
 c) a construct P_{HXT7}-PNT_c-T_{HXT7}

In an embodiment the the transporter construct comprises:

- 5
- a) a construct P_{HXT1} -PNT_a-T_{HXT1};
 - b) a construct P_{HXT4} -PNT_b-T_{HXT4} and
 - c) a construct P_{HXT6} -PNT_c-T_{HXT6}

In an embodiment the the transporter construct comprises:

- 10
- a) a construct P_{HXT1} -PNT_a-T_{HXT1} or P_{HXT3} -PNT_a-T_{HXT3};
 - b) a construct P_{HXT4} -PNT_b-T_{HXT4} and
 - c) a construct P_{HXT7} -PNT_c-T_{HXT7}

In an embodiment the the transporter construct comprises:

- 15
- a) a construct P_{HXT3} -PNT_a-T_{HXT3};
 - b) a construct P_{HXT2} -PNT_b-T_{HXT2} and
 - c) a construct P_{HXT6} -PNT_c-T_{HXT6}

In an embodiment the the transporter construct comprises:

- 20
- a) a construct P_{HXT3} -PNT_a-T_{HXT3};
 - b) a construct P_{HXT2} -PNT_b-T_{HXT2} and
 - c) a construct P_{HXT7} -PNT_c-T_{HXT7}

In an embodiment the the transporter construct comprises:

- 25
- a) a construct P_{HXT3} -PNT_a-T_{HXT3};
 - b) a construct P_{HXT4} -PNT_b-T_{HXT4} and
 - c) a construct P_{HXT6} -PNT_c-T_{HXT6}

30 In an embodiment the the transporter construct comprises:

- a) a construct P_{HXT3} -PNT_a-T_{HXT3};
- b) a construct P_{HXT4} -PNT_b-T_{HXT4} and
- c) a construct P_{HXT7} -PNT_c-T_{HXT7}.

In an embodiment the ORF (Open Reading Frame) of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell is a mutant *HXT11*, which has a reduced affinity for glucose or an increased affinity for xylose compared to native *HXT11*. In an embodiment thereof, has mutation N336M or N366T of *HXT11*. In an embodiment the transporter construct is integrated into the yeast cell genome.

Table 1: Affinities (K_m) towards xylose and glucose and calculated K_m ratio (K_m -xylose/ K_m -glucose) of several sugar transporters found in (patent) literature

Sugar transporter	Donor organism	K_m -xylose (mM)	K_m -glucose (mM)	K_m -xylose/ K_m -glucose	Reference
<i>GXS1</i>	<i>Candida intermedia</i>	0.4	0.012	33.3	Leandro <i>et al</i> , 2006
<i>GFX1</i>	<i>Candida intermedia</i>	48.7	2	24.34	Leandro <i>et al</i> , 2006
<i>GAL2</i>	<i>Saccharomyces cerevisiae</i>	225.6	1.5	150.4	Farwick <i>et al</i> , 2014
<i>GAL2 (N376F)*</i>	<i>Saccharomyces cerevisiae</i>	91.4	n.d.	No affinity to glucose*	Farwick <i>et al</i> , 2014
<i>GAL2 (N376V)</i>	<i>Saccharomyces cerevisiae</i>	168.3	22.1	7.6	Farwick <i>et al</i> , 2014
<i>HXT7</i>	<i>Saccharomyces cerevisiae</i>	200.3	0.5	400.6	Farwick <i>et al</i> , 2014
<i>HXT7 (N370S)</i>	<i>Saccharomyces cerevisiae</i>	169.9	10.8	15.7	Farwick <i>et al</i> , 2014
<i>SUT1</i>	<i>Scheffersomyces stipitis</i>	145	1.5	96.7	Weierstall <i>et al</i> , 1999
<i>SUT2</i>	<i>Scheffersomyces stipitis</i>	49	1.1	44.5	Weierstall <i>et al</i> , 1999
<i>SUT3</i>	<i>Scheffersomyces stipitis</i>	103	0.8	128.8	Weierstall <i>et al</i> , 1999
<i>HXT36</i>	<i>Saccharomyces cerevisiae</i>	107.9	6.1	17.7	PCT/EP2014/061635
<i>HXT36 (N376I)*</i>	<i>Saccharomyces cerevisiae</i>	39.8	n.d.	No affinity to glucose*	PCT/EP2014/061635
<i>HXT36 (N376A)*</i>	<i>Saccharomyces cerevisiae</i>	24.9	170.7	0.15*	PCT/EP2014/061635
<i>HXT11</i>	<i>Saccharomyces cerevisiae</i>	84.2	33.4	2.52	PCT/EP2014/061635
<i>HXT11(N366D)</i>	<i>Saccharomyces cerevisiae</i>	106.7	87	1.23	PCT/EP2014/061635
<i>HXT11(N366T)*</i>	<i>Saccharomyces cerevisiae</i>	46.7	194.4	0.24*	PCT/EP2014/061635
<i>HXT11(N366M)*</i>	<i>Saccharomyces cerevisiae</i>	50.1	144.9	0.35*	PCT/EP2014/061635
<i>mstA</i>	<i>Aspergillus niger</i>	0.3	0.025	12	Vankuyk <i>et al</i> , 2004

n.d.: not detectable in assay; * ratio K_m -xylose/ K_m -glucose < 1 or no affinity detected for glucose suggests that respective sugar transporter is a pentose transporter, more specifically a xylose transporter

Table 2: Sugar transporters expressed under control of *HXT* promoters (besides the *HXT1-17* genes as present in the genome of *Saccharomyces cerevisiae*) screened in the search for suitable pentose transporters as found in prior published literature.

Sugar transporter	Donor organism	Promoter	<i>Saccharomyces cerevisiae</i> host strain	Reference
GXS1	<i>Candida intermedia</i>	<i>S. cerevisiae</i> HXT7(-392)	MY1	Leandro <i>et al</i> , 2006
GFX1	<i>Candida intermedia</i>	<i>S. cerevisiae</i> HXT7(-392)	MY2	Leandro <i>et al</i> , 2006
GAL2	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (N376F)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (N376V)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
HXT7	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
HXT7 (N370S)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (T219S)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (T219N)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (N376Y)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I218A)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I218G)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I218D)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I218K)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I218F)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222G)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222A)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222D)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222K)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222F)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (I222Q)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (Y446W)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350A)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350D)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350G)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350K)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350L)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014

GAL2 (F350M)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350W)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
GAL2 (F350Y)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	AFY10x	Farwick <i>et al</i> , 2014
HXT1	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT3	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT4	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT5	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT7	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT8	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT9	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT10	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT11	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT13	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT14	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
HXT15	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
<i>xylE</i>	<i>Escherichia coli</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
STP2	<i>Arabidopsis thaliana</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
STP3	<i>Arabidopsis thaliana</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
Hug1	<i>Chlorella kessleri</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
GAL2	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	TMB3201	Hamacher <i>et al</i> , 2002
<i>xtrD</i>	<i>Aspergillus nidulans</i> <i>et al</i> , 2010	<i>S. cerevisiae</i> HXT7(-392)	EBY.VW4000+ XR+XDH	Colabardini <i>et al</i> , 2014
HXT36	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT36 (N376I)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT36 (N376A)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
<i>araT</i>	<i>Scheffersomyces stipitis</i>	<i>S. cerevisiae</i> HXT7(-392)	MKY06	Subtil and Boles, 2011
STP2	<i>Arabidopsis thaliana</i>	<i>S. cerevisiae</i> HXT7(-392)	MKY06	Subtil and Boles, 2011
HXT9	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	MKY06	Subtil and Boles, 2011
HXT10	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	MKY06	Subtil and Boles, 2011

HXT11	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT11(N366D)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT11(N366T)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT11(N366M)	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
HXT2	<i>Saccharomyces cerevisiae</i>	<i>S. cerevisiae</i> HXT7(-392)	DS68625	PCT/EP2014/061635
Xyp29	<i>Scheffersomyces stipitis</i>	<i>S. cerevisiae</i> HXT7(-392)	EBY.VW4000	Du et al, 2010
An25	<i>Neurospora crassa</i>	<i>S. cerevisiae</i> HXT7(-392)	EBY.VW4000	Du et al, 2010

Polynucleotide sequence

With the pentose transporter polypeptides and its amino acid sequence as disclosed herein, the skilled person may determine suitable polynucleotides that encode the pentose transporter polypeptide.

The polynucleotides of encoding the pentose transporter and its promoter according to the invention may be isolated or synthesized. The pentose transporter polypeptides and pentose transporter polypeptide polynucleotides herein may be synthetic polypeptides, respectively polynucleotides. The synthetic polynucleotides may be optimized in codon use, preferably according to the methods described in WO2006/077258 and/or PCT/EP2007/055943, which are herein incorporated by reference. PCT/EP2007/055943 addresses codon-pair optimization.

The term refers to a polynucleotide molecule, which is a ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) molecule, either single stranded or double stranded. A polynucleotide may either be present in isolated form, or be comprised in recombinant nucleic acid molecules or vectors, or be comprised in a host cell.

The word "polypeptide" is used herein for chains containing more than seven amino acid residues. All oligopeptide and polypeptide formulas or sequences herein are written from left to right and in the direction from amino terminus to carboxy terminus. The one-letter code of amino acids used herein is commonly known in the art.

By "isolated" polypeptide or protein is intended a polypeptide or protein removed from its native environment. For example, recombinantly produced polypeptides and proteins expressed in host cells are considered isolated for the purpose of the invention as are native or recombinant polypeptides which have been substantially purified by any suitable technique such as, for example, the single-step purification method disclosed in Smith and Johnson, Gene 67:31-40 (1988).

The host cell

The host cell may be any host cell suitable for production of a useful product. A
5 host cell may be any suitable cell, such as a prokaryotic cell, such as a bacterium, or a
eukaryotic cell. Typically, the cell will be a eukaryotic cell, for example a yeast or a
filamentous fungus.

Yeasts are herein defined as eukaryotic microorganisms and include all species
of the subdivision Eumycotina (Alexopoulos, C. J. (1962_, In : Introductory
10 Mycology, John Wiley & Sons, Inc. , New York) that predominantly grow in unicellular
form.

Yeasts may either grow by budding of a unicellular thallus or may grow by fission
of the organism. A preferred yeast as a transformed host cell may belong to the genera
Saccharomyces, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*,
15 *Kloeckera*, *Schwanniomyces* or *Yarrowia*. Preferably the yeast is one capable of
anaerobic fermentation, more preferably one capable of anaerobic alcoholic
fermentation.

In one embodiment the host cell may be yeast.

Preferably the host is an industrial host, more preferably an industrial yeast. An
20 industrial host and industrial yeast cell may be defined as follows. The living
environments of yeast cells in industrial processes are significantly different from that in
the laboratory. Industrial yeast cells must be able to perform well under multiple
environmental conditions that may vary during the process. Such variations include
change in nutrient sources, pH, ethanol concentration, temperature, oxygen
25 concentration, etc., which together have potential impact on the cellular growth and
ethanol production of *Saccharomyces cerevisiae*. Under adverse industrial conditions,
the environmental tolerant strains should allow robust growth and production. Industrial
yeast strains are generally more robust towards these changes in environmental
conditions which may occur in the applications they are used, such as in the baking
30 industry, brewing industry, wine making and the ethanol industry. Examples of industrial
yeast (*S. cerevisiae*) are Ethanol Red® (Fermentis) Fermiol® (DSM) and Thermosacc®
(Lallemand).

In an embodiment the host is inhibitor tolerant. Inhibitor tolerant host cells may be
selected by screening strains for growth on inhibitors containing materials, such as

illustrated in Kadar *et al.*, Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, wherein an inhibitor tolerant *S. cerevisiae* strain ATCC 26602 was selected.

Inhibitor tolerance is resistance to inhibiting compounds. The presence and level of inhibitory compounds in lignocellulose may vary widely with variation of feedstock, pretreatment method hydrolysis process. Examples of categories of inhibitors are carboxylic acids, furans and/or phenolic compounds. Examples of carboxylic acids are lactic acid, acetic acid or formic acid. Examples of furans are furfural and hydroxymethylfurfural. Examples of phenolic compounds are vanillin, syringic acid, ferulic acid and coumaric acid. The typical amounts of inhibitors are for carboxylic acids: several grams per liter, up to 20 grams per liter or more, depending on the feedstock, the pretreatment and the hydrolysis conditions.

For phenolics: several tens of milligrams per liter, up to a gram per liter, depending on the feedstock, the pretreatment and the hydrolysis conditions.

Transformation

The polynucleotides and constructs of the present invention, such as a polynucleotide encoding the pentose transporter polypeptide can be isolated or synthesized using standard molecular biology techniques and the sequence information provided herein.

The polynucleotide encoding the pentose transporter polypeptide of the invention can be amplified using cDNA, mRNA or alternatively, genomic DNA, as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques.

The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis.

The polynucleotides and constructs according to the invention may be expressed in a suitable yeast. The invention thus relates to a transformed host cell. In an embodiment, the host cell may be transformed with a nucleic acid construct that comprises a polynucleotide that encodes the polypeptide according to the invention defined before. Therefore standard transformation techniques may be used.

For most yeast, the vector or expression construct is preferably integrated in the genome of the host cell in order to obtain stable transformants. However, for certain yeasts also suitable episomal vectors are available into which the expression construct can be incorporated for stable and high level expression, examples thereof include

vectors derived from the 2 μ and pKD1 plasmids of *Saccharomyces* and *Kluyveromyces*, respectively, or vectors containing an AMA sequence (e.g. AMA1 from *Aspergillus*). In case the expression constructs are integrated in the host cells genome, the constructs are either integrated at random loci in the genome, or at predetermined target loci using homologous recombination, in which case the target loci preferably comprise a highly expressed gene.

Accordingly, expression vectors useful in the present invention include chromosomal-, episomal- and virus-derived vectors e.g., vectors derived from bacterial plasmids, bacteriophage, yeast episome, yeast chromosomal elements, viruses such as baculoviruses, papova viruses, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids.

The vector may further include sequences flanking the polynucleotide giving rise to RNA which comprise sequences homologous to eukaryotic genomic sequences or viral genomic sequences. This will allow the introduction of the polynucleotides of the invention into the genome of a host cell.

An integrative cloning vector may integrate at random or at a predetermined target locus in the chromosome(s) of the host cell into which it is to be integrated.

The vector system may be a single vector, such as a single plasmid, or two or more vectors, such as two or more plasmids, which together contain the total DNA to be introduced into the genome of the host cell.

The vector may contain a polynucleotide of the invention oriented in an antisense direction to provide for the production of antisense RNA.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, transduction, infection, lipofection, cationic lipidmediated transfection or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*Molecular Cloning: A Laboratory Manual*, 2nd, ed. Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989), Davis et al., *Basic Methods in Molecular Biology* (1986) and other laboratory manuals.

Provided also are host cells, comprising a polynucleotide or vector of the invention. The polynucleotide may be heterologous to the genome of the host cell. The term "heterologous", usually with respect to the host cell, means that the polynucleotide does not naturally occur in the genome of the host cell or that the polypeptide is not naturally produced by that cell.

In another embodiment, the invention features cells, e.g., transformed host cells or recombinant host cells that contain a nucleic acid encompassed by the invention. A "transformed cell" or "recombinant cell" is a cell into which (or into an ancestor of which) has been introduced, by means of recombinant DNA techniques, a nucleic acid according to the invention. All yeast cells are included as host, e.g. *Saccharomyces*, for example *Saccharomyces cerevisiae*.

A host cell can be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in a specific, desired fashion. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may facilitate optimal functioning of the protein. Polynucleotides of the invention may be incorporated into a recombinant replicable vector, e. g. an expression vector. The vector may be used to replicate the nucleic acid in a compatible host cell. Thus in a further embodiment, the invention provides a method of making a polynucleotide of the invention by introducing a polynucleotide of the invention into a replicable vector, introducing the vector into a compatible host cell, and growing the host cell under conditions which bring about the replication of the vector. The vector may be recovered from the host cell.

The vectors may be transformed or transfected into a suitable host cell as described above to provide for expression of a polypeptide of the invention. This process may comprise culturing a host cell transformed with an expression vector as described above.

Herein standard isolation, hybridization, transformation and cloning techniques are used (e. g., as described in Sambrook, J., Fritsh, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

Optionally, a selectable marker may be present in a nucleic acid construct suitable for use in the invention. As used herein, the term "marker" refers to a gene encoding a trait or a phenotype which permits the selection of, or the screening for, a host cell containing the marker. The marker gene may be an antibiotic resistance gene whereby the appropriate antibiotic can be used to select for transformed cells from

among cells that are not transformed. Examples of suitable antibiotic resistance markers include e.g. dihydrofolate reductase, hygromycin-B-phosphotransferase, 3'-O-phosphotransferase II (kanamycin, neomycin and G418 resistance). Antibiotic resistance markers may be most convenient for the transformation of polyploid host cells. Also non-antibiotic resistance markers may be used, such as auxotrophic markers (*URA3*, *TRP1*, *LEU2*) or the *S. pombe* *TPI* gene (described by Russell P R, 1985, Gene 40: 125-130). In a preferred embodiment the host cells transformed with the nucleic acid constructs are marker gene free. Methods for constructing recombinant marker gene free microbial host cells are disclosed in EP-A-O 635 574 and are based on the use of bidirectional markers such as the *A. nidulans amdS* (acetamidase) gene or the yeast *URA3* and *LYS2* genes. Alternatively, a screenable marker such as Green Fluorescent Protein, lacL, luciferase, chloramphenicol acetyltransferase, beta-glucuronidase may be incorporated into the nucleic acid constructs of the invention allowing to screen for transformed cells.

Optional further elements that may be present in the nucleic acid constructs suitable for use in the invention include, but are not limited to, one or more leader sequences, enhancers, integration factors, and/or reporter genes, intron sequences, centromeres, telomeres and/or matrix attachment (MAR) sequences. The nucleic acid constructs of the invention may further comprise a sequence for autonomous replication, such as an ARS sequence.

The heterologous DNA is usually introduced into the organism in the form of extra-chromosomal plasmids (YE_p, YC_p and YR_p). Unfortunately, it has been found with both bacteria and yeasts that the new characteristics may not be retained, especially if the selection pressure is not applied continuously. This is due to the segregational instability of the hybrid plasmid when recombinant cells grow for a long period of time. This leads to population heterogeneity and clonal variability, and eventually to a cell population in which the majority of the cells has lost the properties that were introduced by transformation. If vectors with auxotrophic markers are being used, cultivation in rich media often leads to rapid loss of the vector, since the vector is only retained in minimal media. The alternative, the use of dominant antibiotic resistance markers, is often not compatible with production processes. The use of antibiotics may not be desired from a registration point of view (the possibility that trace amounts of the antibiotic end up in the end product) or for economic reasons (costs of the use of antibiotics at industrial scale).

Genetic engineering, i.e. transformation of yeast cells with recombinant DNA, became feasible for the first time in 1978 [Beggs, 1978; Hinnen et al., 1978]. Recombinant DNA technology in yeast has established itself since then. A multitude of

different vector constructs are available. Generally, these plasmid vectors, called shuttle vectors, contain genetic material derived from *E.coli* vectors consisting of an origin of replication and a selectable marker (often the β lactamase gene, *ampR*), which enable them to be propagated in *E.coli* prior to transformation into yeast cells. Additionally, the shuttle vectors contain a selectable marker for selection in yeast. Markers can be genes encoding enzymes for the synthesis of a particular amino acid or nucleotide, so that cells carrying the corresponding genomic deletion (or mutation) are complemented for auxotrophy or autotrophy. Alternatively, these vectors contain heterologous dominant resistance markers, which provides recombinant yeast cells (i.e. the cells that have taken up the DNA and express the marker gene) resistance towards certain antibiotics, like g418 (Geneticin), hygromycinB or phleomycin. In addition, these vectors may contain a sequence of (combined) restriction sites (multiple cloning site or MCS) which will allow to clone foreign DNA into these sites, although alternative methods exist as well.

Traditionally, four types of shuttle vectors can be distinguished by the absence or presence of additional genetic elements:

- Integrative plasmids (YIp) which by homologous recombination are integrated into the host genome at the locus of the marker or another gene, when this is opened by restriction and the linearized DNA is used for transformation of the yeast cells. This generally results in the presence of one copy of the foreign DNA inserted at this particular site in the genome.
- Episomal plasmids (YEp) which carry part of the 2 μ plasmid DNA sequence necessary for autonomous replication in yeast cells. Multiple copies of the transformed plasmid are propagated in the yeast cell and maintained as episomes.
- Autonomously replicating plasmids (YRp) which carry a yeast origin of replication (ARS, autonomously replicated sequence) that allows the transformed plasmids to be propagated several hundred-fold.
- CEN plasmids (YCp) which carry in addition to an ARS sequence a centromeric sequence (derived from one of the nuclear chromosomes) which normally guarantees stable mitotic segregation and usually reduces the copy number of self-replicated plasmid to just one.

These plasmids are being introduced into the yeast cells by transformation. Transformation of yeast cells may be achieved by several different techniques, such as permeabilization of cells with lithium acetate (Ito et al, 1983) and electroporation methods.

In commercial application of recombinant microorganisms, plasmid instability is the most important problem. Instability is the tendency of the transformed cells to lose their engineered properties because of changes to, or loss of, plasmids. This issue is discussed in detail by Zhang et al (Plasmid stability in recombinant *Saccharomyces cerevisiae*. Biotechnology Advances, Vol. 14, No. 4, pp. 401-435, 1996). Strains transformed with integrative plasmids are extremely stable, even in the absence of selective pressure (Sherman, F. http://dbb.urmc.rochester.edu/bs/sherman_f/yeast/9.html and references therein).

Loss of vectors leads to problems in large scale production situations. Alternative methods for introduction of DNA do exist for yeasts, such as the use of integrating plasmids (YIp). The DNA is integrated into the host genome by recombination, resulting in high stability. (Caunt, P. (1988) Stability of recombinant plasmids in yeast. Journal of Biotechnology 9: 173 – 192). We have found that an integration method using the host transposons are a good alternative. In an embodiment genes may be integrated into the transformed host cell genome. Initial introduction (i.e. before adaptive evolution) of multiple copies be executed in any way known in the art that leads to introduction of the genes. In an embodiment, this may be accomplished using a vector with parts homologous to repeated sequences (transposons), of the host cell. When the host cell is a yeast cell, suitable repeated sequences are the long terminal repeats (LTR) of the Ty element, known as delta sequence. Ty elements fall into two rather similar subfamilies called Ty1 and Ty2. These elements are about 6 kilobases (kb) in length and are bounded by long terminal repeats (LTR), sequences of about 335 base pairs (Boeke JD et al, (1988) The *Saccharomyces cerevisiae* Genome Contains Functional and Nonfunctional Copies of Transposon Ty1. Molecular and Cellular Biology 8: 1432-1442). In the fully sequenced *S. cerevisiae* strain, S288c, the most abundant transposons are Ty1 (31 copies) and Ty2 (13 copies) (Gabriel A, Daprich J, Kunkel M, Gresham D, Pratt SC, et al. (2006) Global mapping of transposon location. PLoS Genet 2: e212). These transposons consist of two overlapping open reading frames (ORFs), each of which encode several proteins. The coding regions are flanked by the aforementioned, nearly identical LTRs. Other, but less abundant and more distinct Ty elements in *S. cerevisiae* comprise Ty3, Ty4 and Ty5. For each family of full-length Ty elements there are an order of magnitude more solo LTR elements dispersed through the genome. These are thought to arise by LTR–LTR recombination of full-length elements, with looping out of the internal protein encoding regions.

The retrotransposition mechanism of the Ty retrotransposon has been exploited to integrate multiple copies throughout the genome (Boeke et al., 1988; Jacobs et al., 1988). The long terminal repeats (LTR) of the Ty element, known as delta sequences, are also good targets for integration by homologous recombination as they exist in about 150-200 copies that are either Ty associated or solo sites (Boeke, 1989; Kingsman and Kingsman, 1988). (Parekh R.N. (1996). An Integrating Vector for Tunable, High Copy, Stable Integration into the Dispersed Ty DELTA Sites of *Saccharomyces cerevisiae*. Biotechnol. Prog. 12:16-21). By adaptive evolution, see later herein, the number of copies may change.

The recombination process may thus be executed with known recombination techniques. Various means are known to those skilled in the art for expression and overexpression of enzymes in a transformed host cell. In particular, an enzyme may be overexpressed by increasing the copy number of the gene coding for the enzyme in the host cell, e.g. by integrating additional copies of the gene in the host 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.

To get to the claimed yeast cell and constructs in particular the following transformations may be advantageous:

- a) Cloning method according to WO2013144257. This allows expression of multiple gene constructs
- b) CrispR/CAS-9 technology. This allows to convert any specific HXT transporter native in the yeast cell to be mutated into a pentose transporter. A summary of this technology is given in articles Sheridon 2014, Nature Biotechn. 32 (7) 599-601 and DiCarlo (2013) NAR doi:10.1093/nar/gkt135.

Co-consumption

In an embodiment the transformed host is capable of co-consumption of glucose and at least one pentose. This pentose may be arabinose or xylose, in an embodiment it is xylose. Co-consumption (or co-fermentation) of two substrates is defined herein as a simultaneous uptake and intracellular conversion of two different carbon sources (e.g. xylose and glucose), at an appreciable level. Said carbon sources are simultaneously converted into products, such as e.g. biomass, ethanol, glycerol, and the like. This can be determined in a fermentation experiments wherein sugar consumption is monitored.

Co-consumption of a cell is may be quantified and expressed as co-consumption index. The co-consumption index is herein the co-consumption index for glucose and xylose and is calculated as the sum over the time interval of the experiment (measured at multiple instants) of the absolute difference of the glucose uptake rate (Qg) and the specific xylose uptake rate (Qx), expressed as grams of sugar consumed per time unit, in an anaerobic batch culture fermentation at 1 g/l dry yeast pitch, approximately 30 degrees C temperature and wherein the fermentation medium contains approximately 70 grams of glucose per liter and approximately 40 grams xylose per liter, at the start of the fermentation. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

Homology & Identity

Amino acid or nucleotide sequences are said to be homologous when exhibiting a certain level of similarity. Two sequences being homologous indicate a common evolutionary origin. Whether two homologous sequences are closely related or more distantly related is indicated by “percent identity” or “percent similarity”, which is high or low respectively. Although disputed, to indicate “percent identity” or “percent similarity”, “level of homology” or “percent homology” are frequently used interchangeably.

A comparison of sequences and determination of percent identity between two sequences can be accomplished using a mathematical algorithm. The skilled person will be aware of the fact that several different computer programs are available to align two sequences and determine the homology between two sequences (Kruskal, J. B. (1983) An overview of sequence comparison In D. Sankoff and J. B. Kruskal, (ed.), Time warps, string edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). The percent identity between two amino acid sequences can be determined using the Needleman and Wunsch algorithm for the alignment of two sequences. (Needleman, S. B. and Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). The algorithm aligns amino acid sequences as well as nucleotide sequences. The Needleman-Wunsch algorithm has been implemented in the computer program NEEDLE. For the purpose of this invention the NEEDLE program from the EMBOSS package was used (version 2.8.0 or higher, *EMBOSS: The European Molecular Biology Open Software Suite* (2000) Rice, P. Longden, I. and Bleasby, A. Trends in Genetics 16, (6) pp276—277, <http://emboss.bioinformatics.nl/>). For protein sequences, EBLOSUM62 is used for the substitution matrix. For nucleotide sequences, EDNAFULL is used. Other matrices can be specified. The optional parameters used for alignment of amino acid

sequences are a gap-open penalty of 10 and a gap extension penalty of 0.5. The skilled person will appreciate that all these different parameters will yield slightly different results but that the overall percentage identity of two sequences is not significantly altered when using different algorithms.

5

Global Homology Definition

The homology or identity is the percentage of identical matches between the two full sequences over the total aligned region including any gaps or extensions. The homology or identity between the two aligned sequences is calculated as follows:

10 Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment including the gaps. The identity defined as herein can be obtained from NEEDLE and is labelled in the output of the program as "IDENTITY".

15 ***Longest Identity Definition***

The homology or identity between the two aligned sequences is calculated as follows: Number of corresponding positions in the alignment showing an identical amino acid in both sequences divided by the total length of the alignment after subtraction of the total number of gaps in the alignment. The identity defined as herein can be obtained

20 from NEEDLE by using the NOBRIEF option and is labelled in the output of the program as "longest-identity".

The various embodiments of the invention described herein may be cross-combined.

25

The sugar composition

The sugar composition according to the invention comprises glucose, arabinose and xylose. Any sugar composition may be used in the invention that suffices those criteria. Optional sugars in the sugar composition are galactose and mannose. In a preferred embodiment, the sugar composition is a hydrolysate of one or more

30 lignocellulosic material. Lignocellulose herein includes hemicellulose and hemicellulose parts of biomass. Also lignocellulose includes lignocellulosic fractions of biomass. Suitable lignocellulosic materials may be found in the following list: orchard primings, chaparral, mill waste, urban wood waste, municipal waste, logging waste, forest

35 thinnings, short-rotation woody crops, industrial waste, wheat straw, oat straw, rice

straw, barley straw, rye straw, flax straw, soy hulls, rice hulls, rice straw, corn gluten feed, oat hulls, sugar cane, corn stover, corn stalks, corn cobs, corn husks, switch grass, miscanthus, sweet sorghum, canola stems, soybean stems, prairie grass, gamagrass, foxtail; sugar beet pulp, citrus fruit pulp, seed hulls, cellulosic animal wastes, lawn clippings, cotton, seaweed, trees, softwood, hardwood, poplar, pine, shrubs, grasses, wheat, wheat straw, sugar cane bagasse, corn, corn husks, corn hobs, corn kernel, fiber from kernels, products and by-products from wet or dry milling of grains, municipal solid waste, waste paper, yard waste, herbaceous material, agricultural residues, forestry residues, municipal solid waste, waste paper, pulp, paper mill residues, branches, bushes, canes, corn, corn husks, an energy crop, forest, a fruit, a flower, a grain, a grass, a herbaceous crop, a leaf, bark, a needle, a log, a root, a sapling, a shrub, switch grass, a tree, a vegetable, fruit peel, a vine, sugar beet pulp, wheat midlings, oat hulls, hard or soft wood, organic waste material generated from an agricultural process, forestry wood waste, or a combination of any two or more thereof.

An overview of some suitable sugar compositions derived from lignocellulose and the sugar composition of their hydrolysates is given in table 3. The listed lignocelluloses include: corn cobs, corn fiber, rice hulls, melon shells, sugar beet pulp, wheat straw, sugar cane bagasse, wood, grass and olive pressings.

Table 3: Overview of sugar compositions from lignocellulosic materials. Gal=galactose, Xyl=xylose, Ara=arabinose, Man=mannose, Glu=glucose, Rham=rhamnose.

Lignocellulosic material	Gal	Xyl	Ara	Man	Glu	Rham	Sum
Corn cob a	10	286	36		227	11	570
Corn cob b	131	228	160		144		663
Rice hulls a	9	122	24	18	234	10	417
Rice hulls b	8	120	28		209	12	378
Melon Shells	6	120	11		208	16	361
Sugar beet pulp	51	17	209	11	211	24	523
Wheat straw Idaho	15	249	36		396		696
Corn fiber	36	176	113		372		697
Cane Bagasse	14	180	24	5	391		614
Corn stover	19	209	29		370		626
Athel (wood)	5	118	7	3	493		625

Lignocellulosic material	Gal	Xyl	Ara	Man	Glu	Rham	Sum
Eucalyptus (wood)	22	105	8	3	445		583
CWR (grass)	8	165	33		340		546
JTW (grass)	7	169	28		311		515
MSW	4	24	5	20	440		493
Reed Canary Grass Veg	16	117	30	6	209	1	379
Reed Canary Grass Seed	13	163	28	6	265	1	476
Olive pressing residue	15	111	24	8	329		487

It is clear from table 3 that in these lignocelluloses a high amount of sugar is presence in de form of glucose, xylose, arabinose and galactose. The conversion of glucose, xylose, arabinose and galactose to fermentation product is thus of great economic importance. Also mannose is present in some lignocellulose materials be it usually in lower amounts than the previously mentioned sugars. Advantageously therefore also mannose is converted by the transformed host cell.

The transformed host cell

In an embodiment, the transformed host cell may comprise one or more copies of xylose isomerase gene and/or one or more copies of xylose reductase and/or xylitol dehydrogenase, and two to ten copies of *araA*, *araB* and *araD*, genes, wherein these genes are integrated into the cell genome.

In one embodiment, the transformed host cell comprises genes, for example the above xylose isomerase gene and/or one or more copies of xylose reductase and/or xylitol dehydrogenase, and two to ten copies of *araA*, *araB* and *araD*, genes, are integrated into the transformed host cell genome.

The number of copies may be determined by the skilled person by any known method.

In an embodiment, the transformed host cell is able to ferment glucose, arabinose, xylose and galactose.

In an embodiment, the cell is capable of converting 90% or more glucose, xylose arabinose, galactose and mannose available, into a fermentation product. In an

embodiment, cell is capable of converting 91% or more, 92% or more, 94% or more, 95% or more, 96% or more, 97% or more, 98% or more or 100% of all glucose, xylose arabinose, galactose and mannose available, into a fermentation product.

In one embodiment of the invention the transformed host cell is able to ferment one or more additional sugar, preferably C5 and/or C6 sugar e.g. mannose. In an embodiment of the invention the transformed host cell comprises one or more of: a *xyIA*-gene, *XYL1* gene and *XYL2* gene and/or *XKS1*-gene, to allow the transformed host cell to ferment xylose; deletion of the aldose reductase (*GRE3*) gene; overexpression of PPP-genes *TAL1*, *TKL1*, *RPE1* and *RKI1* to allow the increase of the flux through the pentose phosphate pathway in the cell.

The transformed host cells according to the invention may be inhibitor tolerant, i.e. they can withstand common inhibitors at the level that they typically have with common pretreatment and hydrolysis conditions, so that the transformed host cells can find broad application, i.e. it has high applicability for different feedstock, different pretreatment methods and different hydrolysis conditions.

In one embodiment, the industrial transformed host cell is constructed on the basis of an inhibitor tolerant host cell, wherein the construction is conducted as described hereinafter. Inhibitor tolerant host cells may be selected by screening strains for growth on inhibitors containing materials, such as illustrated in Kadar et al, Appl. Biochem. Biotechnol. (2007) Vol. 136-140, 847-858, wherein an inhibitor tolerant *S. cerevisiae* strain ATCC 26602 was selected.

In an embodiment, the transformed host cell is marker-free. As used herein, the term "marker" refers to a gene encoding a trait or a phenotype which permits the selection of, or the screening for, a host cell containing the marker. Marker-free means that markers are essentially absent in the transformed host cell. Being marker-free is particularly advantageous when antibiotic markers have been used in construction of the transformed host cell and are removed thereafter. Removal of markers may be done using any suitable prior art technique, e.g. intramolecular recombination.

A transformed host cell may be able to convert plant biomass, celluloses, hemicelluloses, pectins, starch, starch derivatives,, for example into fermentable sugars. Accordingly, a transformed host cell may express one or more enzymes such as a cellulase (an endocellulase or an exocellulase), a hemicellulase (an endo- or exo-xylanase or arabinase) necessary for the conversion of cellulose into glucose monomers and hemicellulose into xylose and arabinose monomers, a pectinase able to convert

pectins into glucuronic acid and galacturonic acid or an amylase to convert starch into glucose monomers.

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

In an embodiment, the transformed host cell is a cell that is naturally capable of alcoholic fermentation, preferably, anaerobic alcoholic fermentation. A transformed host cell preferably has a high tolerance to ethanol, a high tolerance to low pH (i.e. capable of growth at a pH lower than about 5, about 4, about 3, or about 2.5) and towards organic and/or a high tolerance to elevated temperatures.

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

Additional modifications in a pentose fermenting cell

According to an embodiment, the genes may be introduced in the host cell by introduction into a host cell:

- a) a cluster consisting of the genes *araA*, *araB* and *araD* under control of a strong constitutive promoter
- b) a cluster consisting of PPP-genes *TAL1*, *TKL1*, *RPE1* and *RKI1*, optionally under control of strong constitutive promoter; and deletion of an aldose reductase gene;
- c) a cluster consisting of a *xylA*-gene or *XYL1* and *XYL2* and a *XKS1*-gene under control of strong constitutive promoter;
- d) a construct comprising a *xylA* or *XYL1* and *XYL2* gene under control of a strong constitutive promoter, which has the ability to integrate into the genome on multiple loci;

and adaptive evolution to produce the transformed host cell. The above cell may be constructed using recombinant expression techniques.

Adaptation

The yeast cells according to the invention may be subjected to adaptive evolution. Adaptation is the evolutionary process whereby a population becomes better

suited (adapted) to its habitat or habitats. This process takes place over several to many generations, and is one of the basic phenomena of biology.

The term adaptation may also refer to a feature that is especially important for an organism's survival. Such adaptations are produced in a variable population by the better-suited forms reproducing more successfully, by natural selection.

Changes in environmental conditions alter the outcome of natural selection, affecting the selective benefits of subsequent adaptations that improve an organism's fitness under the new conditions. In the case of an extreme environmental change, the appearance and fixation of beneficial adaptations can be essential for survival. A large number of different factors, such as e.g. nutrient availability, temperature, the availability of oxygen, etcetera, can drive adaptive evolution.

Fitness

There is a clear relationship between adaptedness (the degree to which an organism is able to live and reproduce in a given set of habitats) and fitness. Fitness is an estimate and a predictor of the rate of natural selection. By the application of natural selection, the relative frequencies of alternative phenotypes will vary in time, if they are heritable.

Genetic changes

When natural selection acts on the genetic variability of the population, genetic changes are the underlying mechanism. By this means, the population adapts genetically to its circumstances. Genetic changes may result in visible structures, or may adjust the physiological activity of the organism in a way that suits the changed habitat.

It may occur that habitats frequently change. Therefore, it follows that the process of adaptation is never finally complete. In time, it may happen that the environment changes gradually, and the species comes to fit its surroundings better and better. On the other hand, it may happen that changes in the environment occur relatively rapidly, and then the species becomes less and less well adapted. Adaptation is a genetic process, which goes on all the time to some extent, also when the population does not change the habitat or environment.

The adaptive evolution

The transformed host cells may in their preparation be subjected to adaptive evolution (also called evolutionary engineering). A transformed host cell may be adapted

to sugar utilisation by selection of mutants, either spontaneous or induced (e.g. by radiation or chemicals), for growth on the desired sugar, preferably as sole carbon source, and more preferably under anaerobic conditions. Selection of mutants may be performed by techniques including serial transfer of cultures as e.g. described by Kuyper *et al.* (2004, FEMS Yeast Res. 4: 655-664) or by cultivation under selective pressure in a chemostat culture. E.g. in a preferred host cell at least one of the genetic modifications described above, including modifications obtained by selection of mutants, confer to the host cell the ability to grow on the xylose as carbon source, preferably as sole carbon source, and preferably under anaerobic conditions. When XI is used as gene to convert xylose, preferably the cell produce essentially no xylitol, e.g. the xylitol produced is below the detection limit or e.g. less than about 5, about 2, about 1, about 0.5, or about 0.3 % of the carbon consumed on a molar basis.

Adaptive evolution is also described e.g. in Wisselink H.W. *et al* (2007) Appl. Environ. Microbiol. 73:4881–4891.

In one embodiment of adaptive evolution a regimen consisting of repeated batch cultivation with repeated cycles of consecutive growth in different media is applied, e.g. three media with different compositions (glucose, xylose, and arabinose; xylose and arabinose. See Wisselink *et al.* (2009) Applied and Environ. Microbiol. 75: 907–914.

araA, *araB* and *araD* genes

A transformed host cell is capable of using arabinose. A transformed host cell is therefore, be capable of converting L-arabinose into L-ribulose and/or xylulose 5-phosphate and/or into a desired fermentation product, for example one of those mentioned herein.

Organisms, for example *S. cerevisiae* strains, able to produce ethanol from L-arabinose may be produced by modifying a cell introducing the *araA* (L-arabinose isomerase), *araB* (L-ribulokinase) and *araD* (L-ribulose-5-P4-epimerase) genes from a suitable source. Such genes may be introduced into a transformed host cell in order that it is capable of using arabinose. Such an approach is given is described in WO2003/095627. *araA*, *araB* and *araD* genes from *Lactobacillus plantarum* may be used and are disclosed in WO2008/041840. The *araA* gene from *Bacillus subtilis* and the *araB* and *araD* genes from *Escherichia coli* may be used and are disclosed in EP1499708. In another embodiment, *araA*, *araB* and *araD* genes may be derived from at least one of the genus *Clavibacter*, *Arthrobacter* and/or *Gramella*, in particular one of

Clavibacter michiganensis, *Arthrobacter aurescens*, and/or *Gramella forsetii*, as disclosed in WO 2009011591.

PPP-genes

5 A transformed host cell may comprise one or more genetic modifications that increases the flux of the pentose phosphate pathway. In particular, the genetic modification(s) may lead to an increased flux through the non-oxidative part of the 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
10 modification that increases the flux by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 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 by growing the modified host on xylose as sole carbon source, determining the specific
15 xylose consumption rate and subtracting the specific xylitol production rate from the specific xylose consumption rate, if any xylitol is produced. However, the flux of the non-oxidative part of the pentose phosphate pathway is proportional with the growth rate on xylose as sole carbon source, preferably with the anaerobic growth rate on xylose as sole carbon source. There is a linear relation between the growth rate on xylose as sole
20 carbon source ($\mu_{\max}$) and the flux of the non-oxidative part of the pentose phosphate pathway. The specific xylose consumption rate (Q_s) is equal to the growth rate (μ) divided by the yield of biomass on sugar (Y_{xs}) because the yield of biomass on sugar is constant (under a given set of conditions: anaerobic, growth medium, pH, genetic background of the strain, etc.; i.e. $Q_s = \mu / Y_{xs}$). Therefore the increased flux of the non-oxidative part of the pentose phosphate pathway may be deduced from the increase in
25 maximum growth rate under these conditions unless transport (uptake is limiting).

One or more genetic modifications that increase the flux of the pentose phosphate pathway may be introduced in the host cell in various ways. These including e.g. achieving higher steady state activity levels of xylulose kinase and/or one or more of
30 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
35 genes.

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

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

ribulose-5-P 3-epimerase; D-xylulose-5-phosphate 3-epimerase; pentose-5-phosphate 3-epimerase; or D-ribulose-5-phosphate 3-epimerase. A ribulose 5-phosphate epimerase may be further defined by its amino acid sequence. Likewise a ribulose 5-phosphate epimerase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a ribulose 5-phosphate epimerase. The nucleotide sequence encoding for ribulose 5-phosphate epimerase is herein designated *RPE1*.

The enzyme "ribulose 5-phosphate isomerase" (EC 5.3.1.6) is herein defined as an enzyme that catalyses direct isomerisation of D-ribose 5-phosphate into D-ribulose 5-phosphate and vice versa. The enzyme is also known as phosphopentosisomerase; phosphoriboisomerase; ribose phosphate isomerase; 5-phosphoribose isomerase; D-ribose 5-phosphate isomerase; D-ribose-5-phosphate ketol-isomerase; or D-ribose-5-phosphate aldose-ketose-isomerase. A ribulose 5-phosphate isomerase may be further defined by its amino acid sequence. Likewise a ribulose 5-phosphate isomerase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a ribulose 5-phosphate isomerase. The nucleotide sequence encoding for ribulose 5-phosphate isomerase is herein designated *RKI1*.

The enzyme "transketolase" (EC 2.2.1.1) is herein defined as an enzyme that catalyses the reaction: D-ribose 5-phosphate + D-xylulose 5-phosphate $\leftrightarrow$ sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate and vice versa. The enzyme is also known as glycolaldehydetransferase or sedoheptulose-7-phosphate:D-glyceraldehyde-3-phosphate glycolaldehydetransferase. A transketolase may be further defined by its amino acid. Likewise a transketolase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a transketolase. The nucleotide sequence encoding for transketolase is herein designated *TKL1*.

The enzyme "transaldolase" (EC 2.2.1.2) is herein defined as an enzyme that catalyses the reaction: sedoheptulose 7-phosphate + D-glyceraldehyde 3-phosphate $\leftrightarrow$ D-erythrose 4-phosphate + D-fructose 6-phosphate and vice versa. The enzyme is also known as dihydroxyacetone transferase; dihydroxyacetone synthase; formaldehyde transketolase; or sedoheptulose-7-phosphate :D-glyceraldehyde-3-phosphate glyceronetransferase. A transaldolase may be further defined by its amino acid sequence. Likewise a transaldolase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide

sequence encoding a transaldolase. The nucleotide sequence encoding for transketolase from is herein designated *TAL1*.

Xylose Isomerase or xylose reductase genes

5 According to the invention, one or more copies of one or more xylose isomerase gene and/or one or more xylose reductase and xylitol dehydrogenase are introduced into the genome of the host cell. The presence of these genetic elements confers on the cell the ability to convert xylose by isomerisation or reduction.

10 In one embodiment, the one or more copies of one or more xylose isomerase gene are introduced into the genome of the host cell.

A "xylose isomerase" (EC 5.3.1.5) is herein defined as an enzyme that catalyses the direct isomerisation of D-xylose into D-xylulose and/or *vice versa*. The enzyme is also known as a D-xylose ketoisomerase. A xylose isomerase herein may also be capable of catalysing the conversion between D-glucose and D-fructose (and
15 accordingly may therefore be referred to as a glucose isomerase). A xylose isomerase herein may require a bivalent cation, such as magnesium, manganese or cobalt as a cofactor.

20 Accordingly, such a transformed host cell is capable of isomerising xylose to xylulose. The ability of isomerising xylose to xylulose is conferred on the host cell by transformation of the host cell with a nucleic acid construct comprising a nucleotide sequence encoding a defined xylose isomerase. A transformed host cell isomerises xylose into xylulose by the direct isomerisation of xylose to xylulose.

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

The Xylose isomerase gene may have various origin, such as for example *Piromyces* sp. as disclosed in WO2006/009434. Other suitable origins are e.g. *Bacteroides*, in particular *Bacteroides uniformis* as described in PCT/EP2009/52623, *Bacteroides thetaiotamicron*, *Clostridium phytofermentans* (WO2010000464),
30 *Lactococcus lactis* ssp. *Lactis* (WO2010070549), *XI*'s disclosed in WO2010074577, including *Ciona*, *Ruminococcus flavefaciens* (WO2011006126), *Reticulitermis speratus*, *Mastotermes darwiniensis* (WO2011078262), *Abiothrophia defectiva* (WO2011150131).

35 In another embodiment, one or more copies of one or more xylose reductase and xylitol dehydrogenase genes are introduced into the genome of the host cell. In this embodiment the conversion of xylose is conducted in a two step conversion of xylose

into xylulose via a xylitol intermediate as catalysed by xylose reductase and xylitol dehydrogenase, respectively. In an embodiment thereof xylose reductase (XR), xylitol dehydrogenase (XDH), and xylokinase (XK) may be overexpressed, and optionally one or more of genes encoding NADPH producing enzymes are up-regulated and one or more of the genes encoding NADH consuming enzymes are up-regulated, as disclosed in WO 2004085627.

XKS1 gene

A transformed host cell may comprise one or more genetic modifications that increase the specific xylulose kinase activity. Preferably the genetic modification or modifications 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 host cell or may be a xylulose kinase that is heterologous to the host cell. A nucleotide sequence used for overexpression of xylulose kinase in the host cell is a nucleotide sequence encoding a polypeptide with xylulose kinase activity.

The enzyme "xylulose kinase" (EC 2.7.1.17) is herein defined as an enzyme that catalyses the reaction $\text{ATP} + \text{D-xylulose} = \text{ADP} + \text{D-xylulose 5-phosphate}$. The enzyme is also known as a phosphorylating xylulokinase, D-xylulokinase or ATP :D- xylulose 5-phosphotransferase. A xylulose kinase of the invention may be further defined by its amino acid sequence. Likewise a xylulose kinase may be defined by a nucleotide sequence encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding a xylulose kinase.

In a transformed host cell, a genetic modification or modifications that increase(s) the specific xylulose kinase activity may be combined with any of the modifications increasing the flux of the pentose phosphate pathway as described above. This is not, however, essential.

Thus, a host cell may comprise only a genetic modification or modifications that increase the specific xylulose kinase activity. The various means available in the art for achieving and analysing overexpression of a xylulose kinase in the host cells of the invention are the same as described above for enzymes of the pentose phosphate pathway. Preferably in the host cells of the invention, a xylulose kinase to be overexpressed is overexpressed by at least a factor of about 1.1, about 1.2, about 1.5, about 2, about 5, about 10 or about 20 as compared to a strain which is genetically identical except for the genetic modification(s) 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.

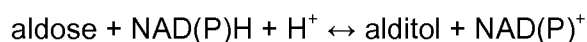
Aldose reductase (*GRE3*) gene deletion

5 In the embodiment, where XI is used as gene to convert xylose, it may be advantageous to reduce aldose reductase activity. A transformed host cell may therefore comprise one or more genetic modifications that reduce unspecific aldose reductase activity in the host cell. Preferably, unspecific aldose reductase activity is reduced in the host cell by one or more genetic modifications that reduce the expression of or inactivates a gene encoding an unspecific aldose reductase. Preferably, the genetic modification(s) reduce or inactivate the expression of each endogenous copy of a gene encoding an unspecific aldose reductase in the host cell (herein called *GRE3* deletion). Transformed host cells may comprise multiple copies of genes encoding unspecific aldose reductases as a result of di-, poly- or aneu-ploidy, and/or the host cell may contain several different (iso)enzymes with aldose reductase activity that differ in amino acid sequence and that are each encoded by a different gene. Also in such instances preferably the expression of each gene that encodes an unspecific aldose reductase is reduced or inactivated. Preferably, the gene is inactivated by deletion of at least part of the gene or by disruption of the gene, whereby in this context the term gene also includes any non-coding sequence up- or down-stream of the coding sequence, the (partial) deletion or inactivation of which results in a reduction of expression of unspecific aldose reductase activity in the host cell.

A nucleotide sequence encoding an aldose reductase whose activity is to be reduced in the host cell is a nucleotide sequence encoding a polypeptide with aldose reductase activity.

Thus, a host cell comprising only a genetic modification or modifications that reduce(s) unspecific aldose reductase activity in the host cell is specifically included in the invention.

The enzyme "aldose reductase" (EC 1.1.1.21) is herein defined as any enzyme that is capable of reducing xylose or xylulose to xylitol. In the context of the present invention an aldose reductase may be any unspecific aldose reductase that is native (endogenous) to a host cell of the invention and that is capable of reducing xylose or xylulose to xylitol. Unspecific aldose reductases catalyse the reaction:



The enzyme has a wide specificity and is also known as aldose reductase; polyol dehydrogenase (NADP⁺); alditol:NADP oxidoreductase; alditol:NADP⁺ 1-oxidoreductase; NADPH-aldopentose reductase; or NADPH-aldose reductase.

A particular example of such an unspecific aldose reductase that is endogenous to *S. cerevisiae* and that is encoded by the GRE3 gene (Traff *et al.*, 2001, Appl. Environ. Microbiol. 67: 5668-74). Thus, an aldose reductase of the invention may be further defined by its amino acid sequence. Likewise an aldose reductase may be defined by the nucleotide sequences encoding the enzyme as well as by a nucleotide sequence hybridising to a reference nucleotide sequence encoding an aldose reductase.

Bioproducts production

Over the years suggestions have been made for the introduction of various organisms for the production of bio-ethanol from crop sugars. In practice, however, all major bio-ethanol production processes have continued to use the yeasts of the genus *Saccharomyces* as ethanol producer. This is due to the many attractive features of *Saccharomyces* species for industrial processes, i. e. , a high acid-, ethanol-and osmo-tolerance, capability of anaerobic growth, and of course its high alcoholic fermentative capacity. Preferred yeast species as host cells include *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* or *K fragilis*.

A transformed host cell may be a cell suitable for the production of ethanol. A transformed host cell may, however, be suitable for the production of fermentation products other than ethanol

Such non-ethanolic fermentation products include in principle any bulk or fine chemical that is producible by a eukaryotic microorganism such as a yeast or a filamentous fungus.

A transformed host cell that may be used for production of non-ethanolic fermentation products is a host cell that contains a genetic modification that results in decreased alcohol dehydrogenase activity.

In an embodiment the transformed host cell may be used in a process wherein sugars originating from lignocellulose are converted into ethanol.

Lignocellulose

Lignocellulose, which may be considered as a potential renewable feedstock, generally comprises the polysaccharides cellulose (glucans) and hemicelluloses (xylans, heteroxylans and xyloglucans). In addition, some hemicellulose may be present as

glucomannans, for example in wood-derived feedstocks. The enzymatic hydrolysis of these polysaccharides to soluble sugars, including both monomers and multimers, for example glucose, cellobiose, xylose, arabinose, galactose, fructose, mannose, rhamnose, ribose, galacturonic acid, glucuronic acid and other hexoses and pentoses occurs under the action of different enzymes acting in concert.

In addition, pectins and other pectic substances such as arabinans may make up considerably proportion of the dry mass of typically cell walls from non-woody plant tissues (about a quarter to half of dry mass may be pectins).

Pretreatment

Before enzymatic treatment, the lignocellulosic material may be pretreated. The pretreatment may comprise exposing the lignocellulosic material to an acid, a base, a solvent, heat, a peroxide, ozone, mechanical shredding, grinding, milling or rapid depressurization, or a combination of any two or more thereof. This chemical pretreatment is often combined with heat-pretreatment, e.g. between 150-220 °C for 1 to 30 minutes.

Enzymatic hydrolysis

The pretreated material is commonly subjected to enzymatic hydrolysis to release sugars that may be fermented according to the invention. This may be executed with conventional methods, e.g. contacting with cellulases, for instance cellobiohydrolase(s), endoglucanase(s), beta-glucosidase(s) and optionally other enzymes. The conversion with the cellulases may be executed at ambient temperatures or at higher temperatures, at a reaction time to release sufficient amounts of sugar(s). The result of the enzymatic hydrolysis is hydrolysis product comprising C5/C6 sugars, herein designated as the sugar composition.

Fermentation

The fermentation process may be an aerobic or an anaerobic fermentation process. An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, more preferably 0 mmol/L/h is consumed (i.e. oxygen consumption is not detectable), and wherein organic molecules serve as both electron donor and electron acceptors. In the absence of oxygen, NADH produced in glycolysis and biomass formation, cannot be oxidised by oxidative

phosphorylation. To solve this problem many microorganisms use pyruvate or one of its derivatives as an electron and hydrogen acceptor thereby regenerating NAD⁺.

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

The fermentation process is preferably run at a temperature that is optimal for the cell. Thus, for most yeasts or fungal host cells, the fermentation process is performed at a temperature which is less than about 42°C, preferably less than about 38°C. For yeast or filamentous fungal host cells, the fermentation process is preferably performed at a temperature which is about 35, about 34, about 33, about 32, about 31 or about 30°C and at a temperature which is higher than about 20, about 22, or about 25°C or about 28 °C.

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

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

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

An anaerobic fermentation process is herein defined as a fermentation process run in the absence of oxygen or in which substantially no oxygen is consumed, preferably less than about 5, about 2.5 or about 1 mmol/L/h, and wherein organic molecules serve as both electron donor and electron acceptors.

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

In a preferred process the cell ferments both the xylose and glucose, preferably simultaneously in which case preferably a cell is used which is insensitive to glucose repression to prevent diauxic growth. In addition to a source of xylose (and glucose) as

carbon source, the fermentation medium will further comprise the appropriate ingredient required for growth of the cell. Compositions of fermentation media for growth of microorganisms such as yeasts are well known in the art

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

SSF mode

For Simultaneous Saccharification and Fermentation (SSF) mode, the reaction time for liquefaction/hydrolysis or presaccharification step is dependent on the time to realize a desired yield, i.e. cellulose to glucose conversion yield. Such yield is preferably as high as possible, preferably 60% or more, 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more, 96% or more, 97% or more, 98% or more, 99% or more, even 99.5% or more or 99.9% or more.

According to the invention very high sugar concentrations in SHF mode and very high product concentrations (e.g. ethanol) in SSF mode are realized. In SHF operation the glucose concentration is 25g/L or more, 30 g/L or more, 35g/L or more, 40 g/L or more, 45 g/L or more, 50 g/L or more, 55 g/L or more, 60 g/L or more, 65 g/L or more, 70 g/L or more, 75 g/L or more, 80 g/L or more, 85 g/L or more, 90 g/L or more, 95 g/L or more, 100 g/L or more, 110 g/L or more, 120g/L or more or may e.g. be 25g/L-250 g/L, 30g/L-200g/L, 40g/L-200 g/L, 50g/L-200g/L, 60g/L-200g/L, 70g/L-200g/L, 80g/L-200g/L, 90 g/L, 80g/L-200g/L.

Product concentration in SSF mode

In SSF operation, the product concentration (g/L) is dependent on the amount of glucose produced, but this is not visible since sugars are converted to product in the SSF, and product concentrations can be related to underlying glucose concentration by multiplication with the theoretical maximum yield (Y_{ps} max in gr product per gram glucose)

The theoretical maximum yield (Y_{ps} max in gr product per gram glucose) of a fermentation product can be derived from textbook biochemistry. For ethanol, 1 mole of glucose (180 gr) yields according to normal glycolysis fermentation pathway in yeast 2

moles of ethanol ($=2 \times 46 = 92$ gr ethanol. The theoretical maximum yield of ethanol on glucose is therefore $92/180 = 0.511$ gr ethanol/gr glucose.

For Butanol (MW 74 gr/mole) or iso butanol, the theoretical maximum yield is 1 mole of butanol per mole of glucose. So Yps max for (iso-)butanol = $74/180 = 0.411$ gr (iso-)butanol/gr glucose.

For lactic acid the fermentation yield for homolactic fermentation is 2 moles of lactic acid (MW = 90 gr/mole) per mole of glucose. According to this stoichiometry, the Yps max = 1 gr lactic acid/gr glucose.

For other fermentation products a similar calculation may be made.

SSF mode

In SSF operation the product concentration is 25g * Yps g/L /L or more, 30 * Yps g/L or more, 35g * Yps /L or more, 40 * Yps g/L or more, 45 * Yps g/L or more, 50 * Yps g/L or more, 55 * Yps g/L or more, 60 * Yps g/L or more, 65 * Yps g/L or more, 70 * Yps g/L or more, 75 * Yps g/L or more, 80 * Yps g/L or more, 85 * Yps g/L or more, 90 * Yps g/L or more, 95 * Yps g/L or more, 100 * Yps g/L or more, 110 * Yps g/L or more, 120g/L * Yps or more or may e.g. be 25 * Yps g/L-250 * Yps g/L, 30 * Yps g/L-200 * Yps g/L, 40 * Yps g/L-200 * Yps g/L, 50 * Yps g/L-200 * Yps g/L, 60 * Yps g/L-200 * Yps g/L, 70 * Yps g/L-200 * Yps g/L, 80 * Yps g/L-200 * Yps g/L, 90 * Yps g/L, 80 * Yps g/L-200 * Yps g/L

Accordingly, the invention provides a method for the preparation of a fermentation product, which method comprises:

- a. degrading lignocellulose using a method as described herein; and
- b. fermenting the resulting material, thereby to prepare a fermentation product.

Fermentation product

The fermentation product of the invention may be any useful product. In one embodiment, it is a product selected from the group consisting of ethanol, n-butanol, isobutanol, lactic acid, di-terpene, glycosylated di-terpene, 3-hydroxy-propionic acid, acrylic acid, acetic acid, succinic acid, fumaric acid, malic acid, itaconic acid, maleic acid,

citric acid, adipic acid, an amino acid, such as lysine, methionine, tryptophan, threonine, and aspartic acid, 1,3-propane-diol, ethylene, glycerol, a β -lactam antibiotic and a cephalosporin, vitamins, pharmaceuticals, animal feed supplements, specialty chemicals, chemical feedstocks, plastics, solvents, fuels, including biofuels and biogas or organic polymers, and an industrial enzyme, such as a protease, a cellulase, an amylase, a glucanase, a lactase, a lipase, a lyase, an oxidoreductases, a transferase or a xylanase. For example the fermentation products may be produced by cells according to the invention, following prior art cell preparation methods and fermentation processes, which examples however should herein not be construed as limiting. n-butanol may be produced by cells as described in WO2008121701 or WO2008086124; lactic acid as described in US2011053231 or US2010137551; 3-hydroxy-propionic acid as described in WO2010010291; acrylic acid as described in WO2009153047.

Recovery of the fermentation product

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

The following examples illustrate the invention:

EXAMPLES

METHODS

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

Media

10 The media used in the experiments was either YEP-medium (10 g/l yeast extract, 20 g/l peptone) or solid YNB-medium (6.7 g/l yeast nitrogen base, 15 g/l agar), supplemented with sugars as indicated in the examples. For solid YEP medium, 15 g/l agar was added to the liquid medium prior to sterilization.

15 In the anaerobic growth experiments, Mineral Medium was used. The composition of Mineral Medium has been described by Verduyn et al. (Yeast (1992), Volume 8, 501-517). The use of ammoniumsulphate was however omitted; instead, as a nitrogen source, 2.3 g/l urea was used. In addition, ergosterol (0.01 g/L), Tween80 (0.42 g/L) and sugars (as indicated) were added.

Strains

20 Strain CEN.PK113-7D has been described by Entian (2007). The genotype of strain CEN.PK113-7D is MATa MAL2-8c SUC2, and was obtained from Dr P. Kötter, Frankfurt, Germany.

25 Strain DS65307 has the following genotype: MATa, MAL2-8c, sit2::TDH3p-araA; ENO1p-araB; PGI1p-araD} , gre3::TPI1p-TAL1; ADH1p-TKL1; PGI1p-RPE1; ENO1p-RKI1}, sit4::TPI1-xylA; TDH1-XKS1}, TY1::TPI1p-xylA. The DS65307 strain construction has been described in WO2012143513.

PCR reactions

30 PCR reactions were carried out using Phusion High-Fidelity DNA Polymerase (Thermo-Scientific Landsmeer, the Netherlands) according to the instructions of the supplier.

Expression cassette construction

The open reading frames (ORFs), promoter sequences and terminators were either synthesized at DNA 2.0 (Menlo Park, CA 94025, USA) or generated by PCR as indicated in the Examples, using standard molecular biology techniques. The promoter, ORF and terminator sequences were assembled by using the Golden Gate technology, as described by Engler et al (2011) and references therein.

Strain construction

The strain construction approach is described in patent application WO2013144257. It describes the techniques enabling the construction of expression cassettes from various genes of interest in such a way, that these cassettes are combined into a pathway and integrated in a specific locus of the yeast genome upon transformation of this yeast. A schematic representation is depicted in figure 1. Firstly, an integration site in the yeast genome is chosen. In this case, the integration site is a *HXT* gene (e.g. *HXT1* as an example in figure 1). The integration of the DNA construct in these examples has two aims: the introduction of a certain coding sequence (i.e. of pentose transporter) and the deletion of a gene (i.e. endogenous *HXT*). A DNA fragment of approximately 600 bp of the up- and downstream part of the integration locus is amplified using PCR, flanked by a connector. These connectors are 50 bp sequences that allow for correct *in vivo* recombination of the expression cassettes upon transformation in yeast (*Saccharomyces cerevisiae* e.g.). The genes of interest, as well as a selectable resistance marker (e.g. *kanMX*), are generated by PCR, incorporating a different connector at each flank, as is displayed in figure 1. Upon transformation of yeast cells with the DNA fragments, *in vivo* recombination and integration into the genome takes place at the desired genomic locus. This technique allows for pathway tuning in the broadest sense, as one or more genes from the pathway can be replaced with (an)other gene(s) or genetic element(s) (patent application WO2013144257).

Transformation of yeast cells

Yeast transformation was done according to the method described by Schiestl and Gietz (Current Genetics (1989), Volume 16, 339-346).

Aerobic growth experiments

Growth experiments are performed in flat bottom NUNC microplates (Thermo Scientific, Landsmeer, The Netherlands; MTPs; 275 µl of medium) or shake flasks (50 ml

of medium). The composition of the medium is as described above. Sugar concentrations and pH-values are indicated in the experimental details of the examples.

Alternatively, strains are grown in shake flasks of an appropriate volume, filled with Mineral (Verduyn) medium, as indicated above.

5

Anaerobic growth experiments

Growth experiments are performed in flat bottom NUNC microplates (MTPs; 275 μ l of medium) or shake flasks sealed with water locks (25 ml of medium). The composition of the medium is as described above. Sugar concentrations and pH-values are indicated in the experimental details of the examples.

10

Alternatively, strains are grown in shake flasks of an appropriate volume, filled with Mineral (Verduyn) medium, as indicated above. In order to establish anaerobic conditions, the shake flasks are closed with water locks.

15

NMR analysis

For the quantification of glucose, xylose, glycerol, acetic acid and ethanol in the sample, 150 μ l sample is transferred accurately into a suitable vial. Subsequently 100 μ l internal standard solution, containing maleic acid (20 g/l), EDTA (40 g/l) and trace amounts of DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) in D_2O , and 450 μ L D_2O is added.

20

1D ^1H NMR spectra are recorded on a Bruker Avance III 700 MHz, equipped with a cryo-probe, using a pulse program with water suppression (power corresponding to 3 Hz) at a temperature of 27°C.

25

The analyte concentrations are calculated based on the following signals (δ relative to DSS):

30

- α -glucose peak at 5.22 ppm (d, 0.38 H, $J = 4$ Hz)
- α -xylose peak at 5.18 ppm (d, 0.37 H, $J = 4$ Hz)
- glycerol peak at 3.55 ppm (dd, 2 H, $J_{1,2} = 6$ Hz and $J_{1a,1b} = 12$ Hz)
- acetic acid peak at 1.91 ppm (s, 3 H)
- ethanol peak at 1.17 ppm (t, 3 H, $J = 7$ Hz)

The signal user for the standard:

- Maleic acid peak at 6.05 ppm (s, 2H)

HPLC analysis

At least, one mL sample was filtrated to separate medium from yeast. The filtrate was inserted into the appropriate vials for HPLC analysis. The concentrations of glucose, xylose, glycerol, acetic acid and ethanol in the medium were determined using a Shimadzu 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) and an Aminex HPX-87H column (300 x 7.8 mm; Bio-Rad). Elution 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).

EXAMPLE 1

Cloning of P_{HXT}-PNT-P_{HXT} integration cassettes.

In PCT/EP2014/061635 the engineered pentose-transporters *HXT11* (N366T) and *HXT11* (N366M) have been described. These transporters display a diminished affinity to glucose whereas their affinity to xylose remained intact. Their K_m-xylose/K_m-glucose is lower than 1 (see Table 1) indicating these transporters are pentose transporters in the fact that these transporters have a higher affinity to xylose than to glucose. When expressed in a xylose-fermenting yeast strain these transporters facilitate simultaneous co-fermentation of glucose and xylose. These *HXT11* variants are expressed in industrially relevant backgrounds having wild type hexose transporter landscape. Unfortunately, a limited benefit is expected of the expression of a single pentose transporter (under control of a constitutive promoter) in a strain with intact hexose transporter landscape with complete glucose transport capacity. Therefore, in this example *HXT11* variants, (N366N) (wild type), (N366T) and (N366M) were integrated seamlessly behind the native *HXT* promoters (*HXT1-7*). In this way, the native gene was replaced by the *HXT11* variants, and the *HXT11* variants were expressed under control of the regulation exerted on the respective HXT promoter.

To this aim, the DNA sequences of *HXT11* (N366N) (wild type; SEQ ID NO: 1), *HXT11* (N366M) (SEQ ID NO: 2), and *HXT11* (N366T) (SEQ ID NO:3) have been fused using Golden Gate cloning techniques (Engler et al, 2011) to the promoters and terminators of *HXT1-7* using type II restriction enzymes (*Bsa*I). *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT5*, *HXT6*, and *HXT7* promoter sequences (SEQ ID NO. 4-10 respectively; P1-P7, see Table 4; approximately 600 bp upstream of ORF) have been PCR-amplified with DS65307 genomic DNA as a template using primer pairs SEQ ID NO: 11-12, 13-14, 15-

16, 17-18, 19-20, 21-22, 23-22, respectively. PCR amplification of *HXT6* and *HXT7* promoter made use of the same reverse primer (SEQ ID NO: 22). The coding sequences for the *HXT11* variants were PCR-amplified using synthetically prepared DNA constructs (O1-O3, Table 4; DNA2.0, Menlo Park, CA 94025, USA) using primers of SEQ ID NO: 24 and SEQ ID NO: 25. *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT5*, *HXT6*, and *HXT7* terminator sequences (SEQ ID NO: 26-32 respectively; T1-T7 Table 4; ca. 300 bp downstream of ORF) were PCR-amplified using DS65307 genomic DNA as a template using primer pairs with SEQ ID NO: 33-34, 35-36, 37-38, 39-40, 41-42, 43-44, 45-46, respectively.

PCR fragments were purified from gel and were used as templates for Golden Gate cloning reactions resulting in the expression cassettes typed as brick P_{HXT}-PNT-T_{HXT} (CAS1-21, respectively; see Table 5). These expression cassettes were sub-cloned in a suitable vector. Besides expression cassettes, the P_{HXT}-PNT-T_{HXT} bricks served as 5' integration flank (because of the HXT promoter). The final P_{HXT}-PNT-T_{HXT} bricks were PCR-amplified using *HXT* promoter primers (without *Bsa*I sites) for *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT5*, *HXT6*, and *HXT7* (SEQ ID NO: 47, 48, 49, 50, 51, 52, 53, respectively) as forward and SEQ ID NO: 54 reverse primer hybridizing to the connector a in the sub-cloning vector.

Table 4: Different bricks needed for the construction glucose-regulated pentose transporter expression cassettes

Brick	Description	Template	DNA Sequence (SEQ ID NO)	Primers (SEQ ID NO)
P1	Promoter <i>HXT1</i>	gDNA DS65307	4	11-12
P2	Promoter <i>HXT2</i>	gDNA DS65307	5	13-14
P3	Promoter <i>HXT3</i>	gDNA DS65307	6	15-16
P4	Promoter <i>HXT4</i>	gDNA DS65307	7	17-18
P5	Promoter <i>HXT5</i>	gDNA DS65307	8	19-20
P6	Promoter <i>HXT6</i>	gDNA DS65307	9	21-22
P7	Promoter <i>HXT7</i>	gDNA DS65307	10	23-22
P8	Promoter <i>TDH3</i>	Synthetic DNA	64	n.a.
O1	<i>HXT11</i> (N366N)	Synthetic DNA	1	17-18
O2	<i>HXT11</i> (N366M)	Synthetic DNA	2	17-18
	<i>HXT11</i> (N366T)	Synthetic DNA	3	17-18

Brick	Description	Template	DNA Sequence (SEQ ID NO)	Primers (SEQ ID NO)
O3				
O4	Dasher GFP (co)	Synthetic DNA (DNA2.0)	n.a.	65-66
O5	Donner Magenta	Synthetic DNA (DNA2.0)	n.a.	n.a.
T1	Terminator <i>HXT1</i>	gDNA DS65307	26	33-34
T2	Terminator <i>HXT2</i>	gDNA DS65307	27	35-36
T3	Terminator <i>HXT3</i>	gDNA DS65307	28	37-38
T4	Terminator <i>HXT4</i>	gDNA DS65307	29	39-40
T5	Terminator <i>HXT5</i>	gDNA DS65307	30	41-42
T6	Terminator <i>HXT6</i>	gDNA DS65307	31	43-44
T7	Terminator <i>HXT7</i>	gDNA DS65307	32	45-46
T8	Terminator <i>ADH1</i>	Synthetic DNA (part of CAS27)	67	n.a.

n.a.: not applicable

Table 5: Combinations of different promoters (P), ORFs (O) and terminators (T) resulting in glucose-regulated pentose transporter expression cassettes

Cassette	Promoter	ORF	Terminator	Description	Connector (5'end)	Connector (3'end)
CAS1	P1	O1	T1	$P_{HXT1}-HXT11(N366N)-T_{HXT1}$	5	a
CAS2	P1	O2	T1	$P_{HXT1}-HXT11(N366M)-T_{HXT1}$	5	a
CAS3	P1	O3	T1	$P_{HXT1}-HXT11(N366T)-T_{HXT1}$	5	a
CAS4	P2	O1	T2	$P_{HXT2}-HXT11-T_{HXT2}$	5	a
CAS5	P2	O2	T2	$P_{HXT2}-HXT11(N366M)-T_{HXT2}$	5	a
CAS6	P2	O3	T2	$P_{HXT2}-HXT11(N366T)-T_{HXT2}$	5	a
CAS7	P3	O1	T3	$P_{HXT3}-HXT11-T_{HXT3}$	5	a
CAS8	P3	O2	T3	$P_{HXT3}-HXT11(N366M)-T_{HXT3}$	5	a
CAS9	P3	O3	T3	$P_{HXT3}-HXT11(N366T)-T_{HXT3}$	5	a
CAS10	P4	O1	T4	$P_{HXT4}-HXT11-T_{HXT4}$	5	a
CAS11	P4	O2	T4	$P_{HXT4}-HXT11(N366M)-T_{HXT4}$	5	a
CAS12	P4	O3	T4	$P_{HXT4}-HXT11(N366T)-T_{HXT4}$	5	a
CAS13	P5	O1	T5	$P_{HXT5}-HXT11-T_{HXT5}$	5	a

Cassette	Promoter	ORF	Terminator	Description	Connector (5'end)	Connector (3'end)
CAS14	P5	O2	T5	P_{HXT5} - $HXT11(N366M)$ - T_{HXT5}	5	a
CAS15	P5	O3	T5	P_{HXT5} - $HXT11(N366T)$ - T_{HXT5}	5	a
CAS16	P6	O1	T6	P_{HXT6} - $HXT11-T_{HXT6}$	5	a
CAS17	P6	O2	T6	P_{HXT6} - $HXT11(N366M)$ - T_{HXT6}	5	a
CAS18	P6	O3	T6	P_{HXT6} - $HXT11(N366T)$ - T_{HXT6}	5	a
CAS19	P7	O1	T7	P_{HXT7} - $HXT11-T_{HXT7}$	5	a
CAS20	P7	O2	T7	P_{HXT4} - $HXT11(N366M)$ - T_{HXT4}	5	a
CAS21	P7	O3	T7	P_{HXT4} - $HXT11(N366T)$ - T_{HXT4}	5	a
CAS27	P8	O4	T8	P_{TDH3} -GFP- T_{ADH1}	b	3
CAS28	P8	O5	T8	P_{TDH3} -GFP- T_{ADH1}	b	3

To allow for selection of transformants in yeast dominant resistance marker cassettes were integrated in the genome of yeast (see Table 6). The following markers were used:

- 1) A DNA construct, which upon integration into the genome of yeast is conferring resistance to the antibiotic G418, is enclosed herein as SEQ ID NO: 55.
- 2) A DNA construct, which upon integration into the genome of yeast is conferring resistance to the antibiotic hygromycin B, is enclosed herein as SEQ ID NO: 56.
- 3) A DNA construct, which upon integration into the genome of yeast is conferring resistance to the antibiotic nourseotrycin, is enclosed herein as SEQ ID NO: 57.
- 4) A DNA construct, which upon integration into the genome of yeast is conferring resistance to the antibiotic phleomycin, is enclosed herein as SEQ ID NO: 58.
- 5) A DNA construct, which upon integration into the genome of yeast allows growth in the presence of acetamide as sole nitrogen source, is enclosed herein as SEQ ID NO: 59.

The marker cassettes were PCR-amplified from the suitable vectors containing connector sequences a and b using primers with SEQ ID NO: 60 and SEQ ID NO: 61 with kanMX, hphMX or natMX as template. For the amplification of amdSYM marker cassette SEQ ID NO: 62 and SEQ ID NO: 63 were used. The flanking connector sequences allow for correct integration with other introduced integration cassettes to allow for mini-pathway integration sharing the same connector sequences (see Figure 1).

Table 6: Cassettes for different dominant resistance markers

Cassette	Description	Compound resistance	Template	Connector (5'end)	Connector (3'end)
CAS22	kanMX	G418	SEQ ID NO: 55	a	b
CAS23	hphMX	Hygromycin B	SEQ ID NO: 56	a	b
CAS24	natMX	Nourseotrycin	SEQ ID NO: 57	a	b
CAS25	phleoMX	phleomycin	SEQ ID NO: 58	a	b
CAS26	amdSYM	acetamide	SEQ ID NO: 59	a	b

5

To allow monitoring of the presence of a marker in the transformants a reporter expression cassette (CAS27 or CAS28; Table 4) was introduced as well. For the promoter part of these cassettes a synthetically prepared *S. cerevisiae* *TDH3* promoter (DNA2.0, Menlo Park, CA 94025, USA, SEQ ID NO 64; P8, Table 4) was ordered. The sequences codon-optimized for *S. cerevisiae* for the reporter genes, either Dasher GFP (green fluorescent protein; Table 4; O4) or Donner Magenta (red/pink chromogenic protein giving red/pink coloration in yeast; Table 4; O5) were ordered from DNA2.0 (Menlo Park, CA 94025, USA). The GFP ORF (Table 4; O4) was PCR-amplified using primers with SEQ ID NO: 65 and SEQ ID NO: 66. The Donner Magenta ORF (Table 4; O5) was digested from DNA2.0-provided plasmid with *BsaI* restriction. For the terminator part of these cassettes a synthetically prepared *S. cerevisiae* *ADH1* terminator (DNA2.0, Menlo Park, CA 94025, USA, SEQ ID NO 67; T8) was ordered. The promoter (Table 4; P8), either of the two ORFs (Table 4; O4 or O5), and the terminator (Table 4, T8) were fused by using Golden Gate cloning techniques resulting in the expression cassette CAS27 or CAS28 (see Table 5) and cloned in the appropriate host plasmid with appropriate connectors. The expression cassette was flanked by connector b on the 5' end and connector 3 on the 3' end. Reporter cassettes CAS27 and CAS28 were PCR-amplified using primers SEQ ID NO: 68 and SEQ ID NO: 69.

10

15

20

To have the opportunity to induce marker removal by homologous recombination a cassette was generated with the exact same *HXT* terminator sequence as the ones in CAS1-21. Using genomic DNA from strain DS65307 the *HXT* terminators were PCR-amplified with primers SEQ ID NOs displayed in Table 7.

5

Table 7: cassettes with HXT terminator sequences to allow for homologous recombination for the removal of marker and GFP-reporter cassettes

Cassette	Terminator <i>HXT</i>	Connector (5'end)	Connector (3 'end)	Primers (SEQ ID NO)
CAS29	<i>HXT1</i>	3	None	70-71
CAS30	<i>HXT2</i>	3	None	72-73
CAS31	<i>HXT3</i>	3	None	74-75
CAS32	<i>HXT4</i>	3	None	76-77
CAS33	<i>HXT5</i>	3	None	78-79
CAS34	<i>HXT6</i>	3	None	80-81
CAS35	<i>HXT7</i>	3	None	82-83

10

EXAMPLE 2

Integration of *HXT11* sequences under control of HXT regulation in pentose-fermenting *Saccharomyces cerevisiae* strains

15

The PCR-amplified cassettes flanked by overlapping connectors generated in Example 1 were integrated in the genome of the glucose-, xylose- and arabinose-fermenting strain DS65307 in the following manner (also displayed schematically in Figure 1):

20

First round:

25

Yeast strain DS65307 was transformed with the following DNA fragments displayed in Table 8:

Table 8: scheme of DNA sequences transformed to yeast resulting in the PCR-verified strains

30

Transformation	Resulting Strain	Strain code	I	II	III	IV
TRAFO1	DS65307-HXT1-HXT11 (N366N)	DS65307-1N	CAS1	CAS24	CAS27	CAS29
TRAFO2	DS65307-HXT1-HXT11 (N366M)	DS65307-1M	CAS2	CAS24	CAS27	CAS29
TRAFO3	DS65307-HXT1-HXT11 (N366T)	DS65307-1T	CAS3	CAS24	CAS27	CAS29
TRAFO4	DS65307-HXT2-HXT11 (N366N)	DS65307-2N	CAS4	CAS24	CAS27	CAS30
TRAFO5	DS65307-HXT2-HXT11 (N366M)	DS65307-2M	CAS5	CAS24	CAS27	CAS30
TRAFO6	DS65307-HXT2-HXT11 (N366T)	DS65307-2T	CAS6	CAS24	CAS27	CAS30
TRAFO7	DS65307-HXT3-HXT11 (N366N)	DS65307-3N	CAS7	CAS24	CAS27	CAS31
TRAFO8	DS65307-HXT3-HXT11 (N366M)	DS65307-3M	CAS8	CAS24	CAS27	CAS31
TRAFO9	DS65307-HXT3-HXT11 (N366T)	DS65307-3T	CAS9	CAS24	CAS27	CAS31
TRAFO10	DS65307-HXT4-HXT11 (N366N)	DS65307-4N	CAS10	CAS24	CAS27	CAS32
TRAFO11	DS65307-HXT4-HXT11 (N366M)	DS65307-4M	CAS11	CAS24	CAS27	CAS32
TRAFO12	DS65307-HXT4-HXT11 (N366T)	DS65307-4T	CAS12	CAS24	CAS27	CAS32
TRAFO13	DS65307-HXT5-HXT11 (N366N)	DS65307-5N	CAS13	CAS24	CAS27	CAS33
TRAFO14	DS65307-HXT5-HXT11 (N366M)	DS65307-5M	CAS14	CAS24	CAS27	CAS33
TRAFO15	DS65307-HXT5-HXT11 (N366T)	DS65307-5T	CAS15	CAS24	CAS27	CAS33
TRAFO16	DS65307-HXT6-HXT11 (N366N)	DS65307-6N	CAS16	CAS24	CAS27	CAS34
TRAFO17	DS65307-HXT6-HXT11 (N366M)	DS65307-6M	CAS17	CAS24	CAS27	CAS34
TRAFO18	DS65307-HXT6-HXT11 (N366T)	DS65307-6T	CAS18	CAS24	CAS27	CAS34
TRAFO19	DS65307-HXT7-HXT11 (N366N)	DS65307-7N	CAS19	CAS24	CAS27	CAS35
TRAFO20	DS65307-HXT7-HXT11 (N366M)	DS65307-7M	CAS20	CAS24	CAS27	CAS35
TRAFO21	DS65307-HXT7-HXT11 (N366T)	DS65307-7T	CAS21	CAS24	CAS27	CAS35

1 µg of the largest DNA fragment was used, and the amounts of the other fragments were adjusted in an equimolar manner. After heatshock, cells were plated on YEPA-agar medium containing 200 µg nourseotrycin/ml. The agar plates were incubated at 30°C. Colonies were visible after two days of incubation.

Colonies were checked by GFP fluorescence as first selection of colonies presenting putatively the correct integrations. Furthermore, GFP fluorescence-emitting

colonies were checked by PCR in order to verify the correct integration of the DNA fragments at the different *HXT* loci. A colony showing the right integration pattern of the transforming DNA, as schematically depicted in figure 1, was designated with the name of the resulting strain as depicted in Table 8.

5

EXAMPLE 3

Expression of *HXT11* sequences under control of HXT regulation in pentose-fermenting *Saccharomyces cerevisiae* strains

10 A qualitative reverse transcriptase (RT)-PCR (RT-PCR) experiment is carried out in order to show that *HXT11* (N366N), *HXT11* (N366M), *HXT11* (N366T) are expressed (or at a higher level) in DS65307-1N, DS65307-1M, DS65307-1T, DS65307-2N, DS65307-2M, DS65307-2T, DS65307-3N, DS65307-3M, DS65307-3T, DS65307-4N, DS65307-4M, DS65307-4T, DS65307-5N, DS65307-5M, DS65307-5T, DS65307-6N, DS65307-6M, DS65307-6T, DS65307-7N, DS65307-7M, and DS65307-7T (see Table 15 8), but not in DS65307 (or at a lower level). To this end, both strains are cultured in Mineral Medium supplemented with 5% glucose; cells are harvested in the early exponential phase, mid-exponential phase and in stationary phase. As a control strain, DS65307 is taken along. RNA is isolated from these cells using the Nucleospin RNA kit 20 from Macherey-Nagel (supplied by BioKé, Leiden, the Netherlands) using the manufacturers manual. cDNA is synthesized using the RevertAid Kit from Thermo Scientific using the manufacturer's instructions. RT-PCR is conducted using the DyNAzyme Colorflash SYBR Green Master Mix from Thermo Scientific, using the supplier's manual. Primers are designated as SEQ ID NO: 84 and 85 (*HXT11*), SEQ ID 25 NO: 84 and 86 (*HXT11* (N366M)), and SEQ ID NO: 84 and 87 (*HXT11* (N366T)).

The results are summarized in the tables below:

Table 9: RTPCR results of *HXT11* expression in constructed DS65307-derived strains with primers SEQ ID NO: 84-85 designed on *HXT11* (N366N)

Strain	Primers (SEQ ID NO:)	Early exponential phase	Mid- exponential phase	Stationary phase
DS65307	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-1N	84 and 85	Band of ± 169 bp	Band of ± 169 bps	No band(s)

DS65307-1M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-1T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-2N	84 and 85	No band(s)	Band of $\pm$ 169 bps	No band(s)
DS65307-2M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-2T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-3N	84 and 85	Band of $\pm$ 169 bps	Band of $\pm$ 169 bps	No band(s)
DS65307-3M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-3T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-4N	84 and 85	No band(s)	Band of $\pm$ 169 bps	No band(s)
DS65307-4M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-4T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-5N	84 and 85	No band(s)	No band(s)	Band of $\pm$ 169 bps
DS65307-5M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-5T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-6N	84 and 85	No band(s)	No band(s)	Band of $\pm$ 169 bps
DS65307-6M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-6T	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-7N	84 and 85	No band(s)	No band(s)	Band of $\pm$ 169 bps
DS65307-7M	84 and 85	No band(s)	No band(s)	No band(s)
DS65307-7T	84 and 85	No band(s)	No band(s)	No band(s)

Table 10: RTPCR results of HXT11 expression in constructed DS65307-derived strains with primers SEQ ID NO: 84-86 designed on HXT11 (N366M)

Strain	Primers (SEQ ID NO:)	Early exponential phase	Mid-exponential phase	Stationary phase
DS65307	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-1N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-1M	84 and 86	Band of $\pm$ 169 bps	Band of $\pm$ 169 bps	No band(s)
DS65307-1T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-2N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-2M	84 and 86	No band(s)	Band of $\pm$ 169 bps	No band(s)
DS65307-2T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-3N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-3M	84 and 86	Band of $\pm$ 169 bps	Band of $\pm$ 169 bps	No band(s)

DS65307-3T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-4N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-4M	84 and 86	No band(s)	Band of ± 169 bps	No band(s)
DS65307-4T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-5N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-5M	84 and 86	No band(s)	No band(s)	Band of ± 169 bps
DS65307-5T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-6N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-6M	84 and 86	No band(s)	No band(s)	Band of ± 169 bps
DS65307-6T	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-7N	84 and 86	No band(s)	No band(s)	No band(s)
DS65307-7M	84 and 86	No band(s)	No band(s)	Band of ± 169 bps
DS65307-7T	84 and 86	No band(s)	No band(s)	No band(s)

Table 11: RTPCR results of HXT11 expression in constructed DS65307-derived strains with primers SEQ ID NO: 84-87 designed on HXT11 (N366T)

5

Strain	Primers (SEQ ID NO:)	Early exponential phase	Mid-exponential phase	Stationary phase
DS65307	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-1N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-1M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-1T	84 and 87	Band of ± 169 bps	Band of ± 169 bps	No band(s)
DS65307-2N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-2M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-2T	84 and 87	No band(s)	Band of ± 169 bps	No band(s)
DS65307-3N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-3M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-3T	84 and 87	Band of ± 169 bps	Band of ± 169 bps	No band(s)
DS65307-4N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-4M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-4T	84 and 87	No band(s)	Band of ± 169 bps	No band(s)
DS65307-5N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-5M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-5T	84 and 87	No band(s)	No band(s)	Band of ± 169 bps
DS65307-6N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-6M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-6T	84 and 87	No band(s)	No band(s)	Band of ± 169 bps
DS65307-7N	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-7M	84 and 87	No band(s)	No band(s)	No band(s)
DS65307-7T	84 and 87	No band(s)	No band(s)	Band of ± 169 bps

The results are showing (as displayed in Table 9, 10 and 11):

1) no expression of *HXT11*, *HXT11* (N366M) or *HXT11* (N366T) in DS65307 in all growth phases;

2) no expression of *HXT11* (N366N) in DS65307-1N and DS65307-3N in stationary phase; but expression of *HXT11* (N366N) in DS65307-1N and DS65307-3N in the early and mid-exponential phase

3) no expression of *HXT11* (N366N) in DS65307-2N and DS65307-4N in the early-exponential phase stationary phase; but expression of *HXT11* (N366N) in DS65307-2N and DS65307-4N in the mid-exponential phase

4) no expression of *HXT11* (N366N) in DS65307-5N in the early-exponential phase; but expression of *HXT11* (N366N) in DS65307-5N in the stationary phase

5) no expression of *HXT11* (N366N) in DS65307-6N and DS65307-7N, in early- or mid-exponential phase; but expression of *HXT11* (N366N) in DS65307-6N and DS65307-7N in the stationary phase

6) no expression of *HXT11* (N366M) in DS65307-1M and DS65307-3M in stationary phase; but expression of *HXT11* (N366M) in DS65307-1M and DS65307-3M in the early and mid-exponential phase

7) no expression of *HXT11* (N366M) in DS65307-2M and DS65307-4M in the early-exponential phase stationary phase; but expression of *HXT11* (N366N) in DS65307-2M and DS65307-4M in the mid-exponential phase

8) no expression of *HXT11* (N366M) in DS65307-5M, in the early- or mid-exponential phase; but expression of *HXT11* (N366M) in DS65307-5M in the stationary phase

9) no expression of *HXT11* (N366M) in DS65307-6M and DS65307-7M, in early- or mid-exponential phase; but expression of *HXT11* (N366M) in DS65307-6M and DS65307-7M in the stationary phase

10) no expression of *HXT11* (N366T) in DS65307-1T and DS65307-3T in stationary phase; but expression of *HXT11* (N366T) in DS65307-1T, and DS65307-3T in the early and mid-exponential phase

11) no expression of *HXT11* (N366M) in DS65307-2T and DS65307-4T in the early-exponential phase stationary phase; but expression of *HXT11* (N366N) in DS65307-2T and DS65307-4T in the mid-exponential phase

12) no expression of *HXT11* (N366T) in DS65307-5T in the early- or mid-exponential phase; but expression of *HXT11* (N366T) in DS65307-5T in the stationary phase

13) no expression of *HXT11* (N366T) in DS65307-6T and DS65307-7T, in early- or mid-exponential phase; but expression of *HXT11* (N366M) in DS65307-6T and DS65307-7T in the stationary phase

5 In other words, the native *HXT11* (N366N) gene is never expressed in reference strain DS65307. In strain DS65307-1N, and DS65307-3N, DS65307-1M and DS65307-3M, or DS65307-1T and DS65307-3T, *HXT11*, *HXT11* (N366M) or *HXT11* (N366T), respectively, are only expressed in the early and mid-exponential phases, *i.e.* when the glucose concentrations are sufficiently high, because they are expressed under control
10 of the *HXT1* or *HXT3* promoter.

In strain DS65307-2N, and DS65307-4N, DS65307-2M and DS65307-4M, or DS65307-2T and DS65307-4T, *HXT11*, *HXT11*-N366M, or *HXT11* (N366T), respectively, are only expressed in the early and mid-exponential phases, *i.e.* when the glucose concentrations are under a certain threshold level (ca. $<15 \text{ g L}^{-1}$), because they
15 are expressed under control of the *HXT2* or *HXT4* promoter.

In strain DS65307-5N, DS65307-5M, or DS65307-5T, *HXT11*, *HXT11* (N366M), or *HXT11* (N366T), respectively, are only expressed in the stationary phase, *i.e.* when yeast's growth rates are low, because they are expressed under control of the *HXT5* promoter.

20 In strain DS65307-6N and DS65307-7N, DS65307-6M and DS65307-7M, or DS65307-6T and DS65307-7T, *HXT11*, *HXT11* (N366M), or *HXT11* (N366T), respectively, are only expressed in the stationary phase, *i.e.* glucose concentrations are scarce or depleted, because they are expressed under control of the *HXT6* or *HXT7* promoter.

25 EXAMPLE 4

Expression of *HXT11* (N366T) by replacing specifically the *HXT1* or *HXT3* gene in a yeast strain enhances the xylose fermentation at higher glucose concentrations of the fermentation

30 In Example 2, transformants of *S. cerevisiae* DS65307 were described, in which *HXT11* (N366N), *HXT11* (N366M) or *HXT11* (N366T) were stably integrated into the genome having two consequences: a) the deletion of an HXT gene, b) expression of *HXT11* (N366N), *HXT11* (N366M) or *HXT11* (N366T) under control of the promoter of the
35 deleted HXT gene. In other words, the native ORF has been replaced by one of the

three HXT11-variants designated above. Strain codes as displayed in Table 8 are maintained throughout this text.

Strains DS65307, DS65307-1T, DS65307-2T, DS65307-3T, DS65307-4T, and DS65307-7T were pre-cultured in a YEP-medium containing 2% maltose. After overnight cultivation at 30°C and 280 rpm in a rotary shaker, cells were harvested, and inoculated at a starting cell density of 0.5 g L⁻¹ for an anaerobic growth test at 32°C using an Alcohol Fermentation Monitor (Applikon, Delft, The Netherlands) with stirring speed 250 rpm. As medium, 400 ml mineral medium supplemented with 7.5 % glucose, 3.8 % xylose, 0.5% arabinose, 0.25 % galactose, 0.1 % mannose was used. During the experiment, biomass, sugar levels, glycerol, acetic acid and ethanol were monitored every four hours.

All strains finished the glucose and xylose within 48 hours (see figure 2). After 8 hours, the sugar consumption profile became distinctly different for DS65307-1T and DS65307-3T which express *HXT11*(N366T) under control of the *HXT1* or *HXT3* promoter, respectively (Figure 2A, 2B), compared with the reference DS65307 (transparent lines in each graph of Figure 2) and the other tested strains (Figure 2C, 2D, 2E). For DS65307-1T and DS65307-3T xylose levels were already decreased after 8 hours into the fermentation at glucose concentrations of 50 g L⁻¹, whereas for the other strains, DS65307-2T, DS65307-4T, DS65307-7T and DS65307 (reference strain), xylose concentrations started to decrease at lower glucose concentrations (10 g L⁻¹ at 16 hours of fermentation).

This indicates that co-fermentation of glucose and xylose was clearly facilitated in DS65307-1T and DS65307-3T at higher glucose concentrations than for the other strains. When expressing the level of co-fermentation by calculating the ratio of the xylose fermentation rate (Q_x) over the glucose fermentation rate (Q_g), clearly higher Q_x/Q_g ratio's were observed at higher glucose concentrations for DS65307-1T and DS65307-3T compared to the other strains (Figure 2F). These results indicate that replacing one of the hexose transporters expressed at high glucose concentrations, *i.e.* *HXT1* or *HXT3*, for a pentose transporter, *i.e.* *HXT11*(N366T) and expressing the pentose transporter under control of the promoters of these hexose transporters resulted in a co-fermentation phenotype at higher glucose concentrations than for the reference strain.

In summary, DS65307-1T and DS65307-3T exhibited a higher degree of glucose and xylose co-fermentation as compared to DS65307 at high glucose concentrations (higher than > 15 g L⁻¹).

EXAMPLE 5**Expression of several modified *HXT* genes by replacing multiple *HXT* genes in a yeast strain enhances xylose fermentation during all phases of the fermentation**

In this Example, transformants of *S. cerevisiae* DS65307 are described, in which *HXT11* (N366T) is stably integrated into the genome at multiple *HXT* loci having two consequences: a) the deletion of multiple native *HXT* genes, b) expression of *HXT11* (N366T) under control of the promoter of the deleted *HXT* gene it has replaced. Strain codes as displayed in Table 12 are maintained throughout this text.

In the strains generated in Example 2, additional constructs are integrated which were described in Example 1 (Table 5, 6 and 7).

Table 12: scheme of DNA sequences transformed (in one or two rounds) to yeast strains generated in Example 2 resulting in the PCR-verified strains

Transformant	Reference strain	Resulting Strain	Strain code	Transformation round	I	II	III	IV
TRAFO 22	DS65307-1T	DS65307- <i>P_{HXT1}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT1}</i> , - <i>P_{HXT2}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT2}</i> , - <i>P_{HXT6}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT6}</i>	DS65307-126T	1T	CAS3	CAS2 4	CAS2 7	CAS 29
				12T	CAS6	CAS2 2	CAS2 7	CAS 30
				126T	CAS1 8	CAS2 3	CAS2 8	CAS 34
TRAFO 23	DS65307-3T	DS65307- <i>P_{HXT3}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT3}</i> , - <i>P_{HXT4}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT4}</i> , - <i>P_{HXT7}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT7}</i>	DS65307-347T	3T	CAS9	CAS2 4	CAS2 7	CAS 31
				34T	CAS1 2	CAS2 2	CAS2 7	CAS 32
				347T	CAS2 1	CAS2 3	CAS2 8	CAS 35
TRAFO 24	DS65307-1T	DS65307- <i>P_{HXT1}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT1}</i> , - <i>P_{HXT2}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT2}</i> , - <i>P_{HXT7}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT7}</i>	DS65307-127T	1T	CAS3	CAS2 4	CAS2 7	CAS 29
				12T	CAS6	CAS2 2	CAS2 7	CAS 30
				127T	CAS2 1	CAS2 3	CAS2 8	CAS 35
TRAFO 25	DS65307-3T	DS65307- <i>P_{HXT3}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT3}</i> , - <i>P_{HXT4}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT4}</i> , - <i>P_{HXT6}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT6}</i>	DS65307-346T	3T	CAS9	CAS2 4	CAS2 7	CAS 31
				34T	CAS1 2	CAS2 2	CAS2 7	CAS 32
				346T	CAS1 8	CAS2 3	CAS2 8	CAS 34
TRAFO 26	DS65307-1T	DS65307- <i>P_{HXT1}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT1}</i> , - <i>P_{HXT4}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT4}</i> , - <i>P_{HXT7}</i> - <i>HXT11</i> (N366T)- <i>T_{HXT7}</i>	DS65307-147T	1T	CAS3	CAS2 4	CAS2 7	CAS 29
				14T	CAS1 2	CAS2 2	CAS2 7	CAS 32
				147T	CAS2 1	CAS2 3	CAS2 8	CAS 35

TRAFO 27	DS65307-3T	DS65307- P_{HXT3^-} - $HXT11$ (N366T)- T_{HXT3} , - P_{HXT2^-} - $HXT11$ (N366T)- T_{HXT2} , - P_{HXT7^-} - $HXT11$ (N366T)- T_{HXT7}	DS65307-326T	3T	CAS9	CAS2 4	CAS2 7	CAS 31
				32T	CAS6	CAS2 2	CAS2 7	CAS 30
				326T	CAS1 8	CAS2 3	CAS2 8	CAS 34
TRAFO 28	DS65307-1T	DS65307- P_{HXT1^-} - $HXT11$ (N366T)- T_{HXT1} , - P_{HXT4^-} - $HXT11$ (N366T)- T_{HXT4} , - P_{HXT6^-} - $HXT11$ (N366T)- T_{HXT6}	DS65307-146T	1T	CAS3	CAS2 4	CAS2 7	AS2 9
				14T	CAS1 2	CAS2 2	CAS2 7	CAS 32
				146T	CAS1 8	CAS2 3	CAS2 8	CAS 34
TRAFO 29	DS65307-3T	DS65307- P_{HXT3^-} - $HXT11$ (N366T)- T_{HXT3} , - P_{HXT2^-} - $HXT11$ (N366T)- T_{HXT2} , - P_{HXT7^-} - $HXT11$ (N366T)- T_{HXT7}	DS65307-327T	3T	CAS9	CAS2 4	CAS2 7	CAS 31
				32T	CAS6	CAS2 2	CAS2 7	CAS 30
				327T	CAS2 1	CAS2 3	CAS2 8	CAS 35

100 ng of each DNA fragment is used, except for the marker, where 20 ng was used. After heatshock, cells are plated on YEPD-agar medium containing either 200 μ g G418/ml, 200 μ g hygromycin B/ml, 200 μ g nourseothricin/ml or all antibiotics. The agar plates are incubated at 30°C. Colonies are visible after two days of incubation.

Colonies are checked by GFP fluorescence or red/pink coloration (due to Donner Magenta expression) as first selection of colonies presenting putatively the correct integrations. Furthermore, GFP fluorescence-emitting or pink/red-colored colonies are checked by PCR in order to verify the correct integration of the DNA fragments at the different *HXT* loci. A colony showing the right integration pattern of the transforming DNA, as schematically depicted in figure 1, is designated with the name of the resulting strain code as depicted in Table 12.

This results in the end in a seamless integration of *HXT11* (N366N), *HXT11* (N366M), or *HXT11* (N366T) behind the targeted HXT promoter. To do so, previously verified correct transformants which emit GFP-fluorescence, are serially transferred several instances on YePD without antibiotic selection pressure after which colonies are plated on YePD agar medium. The agar plates are incubated at 30°C. Colonies are visible after two days of incubation. Colonies that have lost GFP fluorescence or pink/red-coloration are verified for marker removal by colony PCR with the appropriate primers. Furthermore, seamless integration is validated by standard sequencing techniques.

Strains DS65307, DS65307-1T, DS65307-3T, DS65307-126T, DS65307-347T, DS65307-127T, DS65307-146T, DS65307-147T, DS65307-346T, DS65307-326T and DS65307-327T are pre-cultured in a YEP-medium containing 2% maltose. After

overnight cultivation at approximately 30°C and 280 rpm in a rotary shaker, cells are harvested and employed in an anaerobic growth test, in a mineral medium containing glucose (4% or more) and xylose (2% or more).

- 5 DS65307-126T, DS65307-347T, DS65307-127T, DS65307-146T, DS65307-147T, DS65307-346T, DS65307-326T and DS65307-327T are exhibiting a larger degree of glucose and xylose co-fermentation and faster xylose fermentation as compared to DS65307 throughout the fermentation resulting in a higher ethanol productivity and yield.

REFERENCES

Colabardini AC, Ries LN, Brown NA, Dos Reis TF, Savoldi M, Goldman MH, Menino JF, Rodrigues F, Goldman GH 2014 Functional characterization of a xylose transporter in *Aspergillus nidulans*. *Biotechnol Biofuels* 7 p.46. doi: 10.1186/1754-6834-7-46;

Du J, Li S, Zhao H 2010 Discovery and characterization of novel d-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol Biosyst* 6 pp.2150-2156;

Farwick, A., Bruder, S., Schadeweg, V., Oreb, M. and Boles, E. (2014) Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. DOI10.1073/pnas.1323464111; PubMed ID24706835, *Proc Natl Acad Sci USA*. 111 pp.5159-5164;

Hamacher T, Becker J, Gárdonyi M, Hahn-Hägerdal B, Boles E 2002 Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 148 pp.2783-2788;

Horák, J. (2013) Regulations of sugar transporters: insights from yeast. *Curr Genet* DOI 10.1007/s00294-013-0388-8;

Leandro MJ, Gonçalves P, Spencer-Martins I 2006 Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose-H⁺ symporter. *Biochem J*. 395 pp. 543-549;

Leandro, M.J., Spencer-Martins, I. and Gonçalves, P. (2008) The expression in *Saccharomyces cerevisiae* of a glucose/xylose symporter from *Candida intermedia* is affected by the presence of a glucose/xylose facilitator. *Microbiology* vol. 154 no. 6 1646-1655;

Ozcan, S., Dover, J., and Johnston, M. (1998) Glucose sensing and signaling by two glucose receptors in the yeast *Saccharomyces cerevisiae*. *EMBO J*. 17, 2566–2573

Rødkaer SV, Faergeman NJ, 2014, Glucose- and nitrogen sensing and regulatory mechanisms in *Saccharomyces cerevisiae*. *FEMS Yeast Res* Apr 17. doi: 10.1111/1567-1364.12157.

Roy, A. and Kim, J-H. (2014) Endocytosis and Vacuolar Degradation of the Yeast Cell Surface Glucose Sensors Rgt2 and Snf3. *The Journal of Biological Chemistry* vol. 289 no. 10 pp7247-7256;

Runquist, D., Hahn-Hägerdal, B. and Rådström, P. (2010) Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*. *Biotechnology for Biofuels* 2010, 3:5 (doi:10.1186/1754-6834-3-5);

Subtil T, Boles E 2011 Improving L-arabinose utilization of pentose fermenting *Saccharomyces cerevisiae* cells by heterologous expression of L-arabinose transporting sugar transporters. *Biotechnol Biofuels* 4 p.38. doi: 10.1186/1754-6834-4-38;

Vankuyk PA, Diderich JA, MacCabe AP, Hererro O, Ruijter GJ, Visser J 2004 *Aspergillus niger* mstA encodes a high-affinity sugar/H⁺ symporter which is regulated in response to extracellular pH. *Biochem J* 379 pp.375-383;

De Waal PP, Klaassen P, De Jong RM, Nijland JG, Shin HY, Driessen AJ 2013 POLYPEPTIDES WITH PERMEASE ACTIVITY. Patent Application (PCT/EP2014/061635)

Weierstall T, Hollenberg CP, Boles E 1999 Cloning and characterization of three genes (*SUT1-3*) encoding glucose transporters of the yeast *Pichia stipitis*. *Mol Microbiol* 31 pp.871-883;

Young, E., Poucher, E., Comer, A., Bailey, A. and Alper, A. (2011) Functional Survey for Heterologous Sugar Transport Proteins, Using *Saccharomyces cerevisiae* as a Host. *Appl. Environ. Microbiol.* vol. 77 no. 10 3311-3319.

CLAIMS

1. Yeast cell comprising at least one functional pentose conversion pathway from pentose to fermentation product, wherein the yeast cell comprises one or more transporter construct, wherein the transporter construct comprises DNA sequences:
 - a) a transporter promoter
 - b) an ORF of a pentose transporter that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 - c) a terminator,
 wherein the transporter promoter is regulated by cellular regulation by *SNF3*- and/or *RGT2*-signaling.
2. Yeast cell according to claim 1, wherein the transporter is capable of transporting pentose across the cell membrane.
3. Yeast cell according to any of claim 1 or 2, wherein b) is an ORF of a heterologous pentose transporter.
4. Yeast cell according to claim 1 or 2, wherein the b) is an ORF of a variant of an endogenous hexose transporter.
5. Yeast cell according to any of claim 1 to 4, wherein b) is an ORF of a variant of a hexose transporter chosen from the group consisting of *HXT1*, *HXT2*, *HXT3*, *HXT36*, *HXT4*, *HXT5*, *HXT6*, and *HXT7*, *GAL2* and *HXT11*.
6. Yeast cell according to claim 5 comprising DNA sequences:
 - a) an endogenous *HXT_n*-promoter
 - b) an ORF of a pentose transporter (PNT) that replaces a hexose transporter connected to the transporter promoter in native yeast cell and;
 - c) a *HXT_n*-terminator
 wherein n is 1,2,3,4,6 or 7 and
 wherein a), b), c) is designated as P_{HXTn} -PNT-T_{HXTn}.
7. Yeast cell according to claim 6, comprising:

- a) a construct P_{HXT1} -PNT_a-T_{HXT1} or P_{HXT3} -PNT_a-T_{HXT3};
- b) a construct P_{HXT2} -PNT_b-T_{HXT2} or P_{HXT4} -PNT_b-T_{HXT4} and
- c) a construct P_{HXT6} -PNT_c-T_{HXT6} or P_{HXT7} -PNT_c-T_{HXT7}.

5 8. Yeast cell according to claim 7, comprising

- d) a construct P_{HXT1} -PNT_a-T_{HXT1};
- e) a construct P_{HXT2} -PNT_b-T_{HXT2} and
- f) a construct P_{HXT6} -PNT_c-T_{HXT6}

10 or

- g) a construct P_{HXT1} -PNT_a-T_{HXT1};
- h) a construct P_{HXT2} -PNT_b-T_{HXT2} and
- i) a construct P_{HXT7} -PNT_c-T_{HXT7}

15 or

- j) a construct P_{HXT1} -PNT_a-T_{HXT1};
- k) a construct P_{HXT4} -PNT_b-T_{HXT4} and
- l) a construct P_{HXT6} -PNT_c-T_{HXT6}

20 or

- m) a construct P_{HXT1} -PNT_a-T_{HXT1}
- n) a construct P_{HXT4} -PNT_b-T_{HXT4} and
- o) a construct P_{HXT7} -PNT_c-T_{HXT7}

25 or

- p) a construct P_{HXT3} -PNT_a-T_{HXT3};
- q) a construct P_{HXT2} -PNT_b-T_{HXT2} and
- r) a construct P_{HXT6} -PNT_c-T_{HXT6}

30 or

35

- s) a construct P_{HXT3} -PNT_a-T_{HXT3};
- t) a construct P_{HXT2} -PNT_b-T_{HXT2} and
- u) a construct P_{HXT7} -PNT_c-T_{HXT7}

5 or

- v) a construct P_{HXT3} -PNT_a-T_{HXT3};
- w) a construct P_{HXT4} -PNT_b-T_{HXT4} and
- x) a construct P_{HXT6} -PNT_c-T_{HXT6}

10 or

- y) a construct P_{HXT3} -PNT_a-T_{HXT3};
- z) a construct P_{HXT4} -PNT_b-T_{HXT4} and
- 15 aa) a construct P_{HXT7} -PNT_c-T_{HXT7}.

9. Yeast cell according to any of claim 1 to 8, wherein b) in claim 1 is a mutant *HXT11*, which has a reduced affinity for glucose compared to native *HXT11*

20 10. Yeast cell according to claim 9, having mutation N366M or N366T of SEQ ID NO: 1

11. Yeast cell according to any of claims 1-10, wherein the transporter construct is integrated into the yeast cell genome.

25

Fig. 1

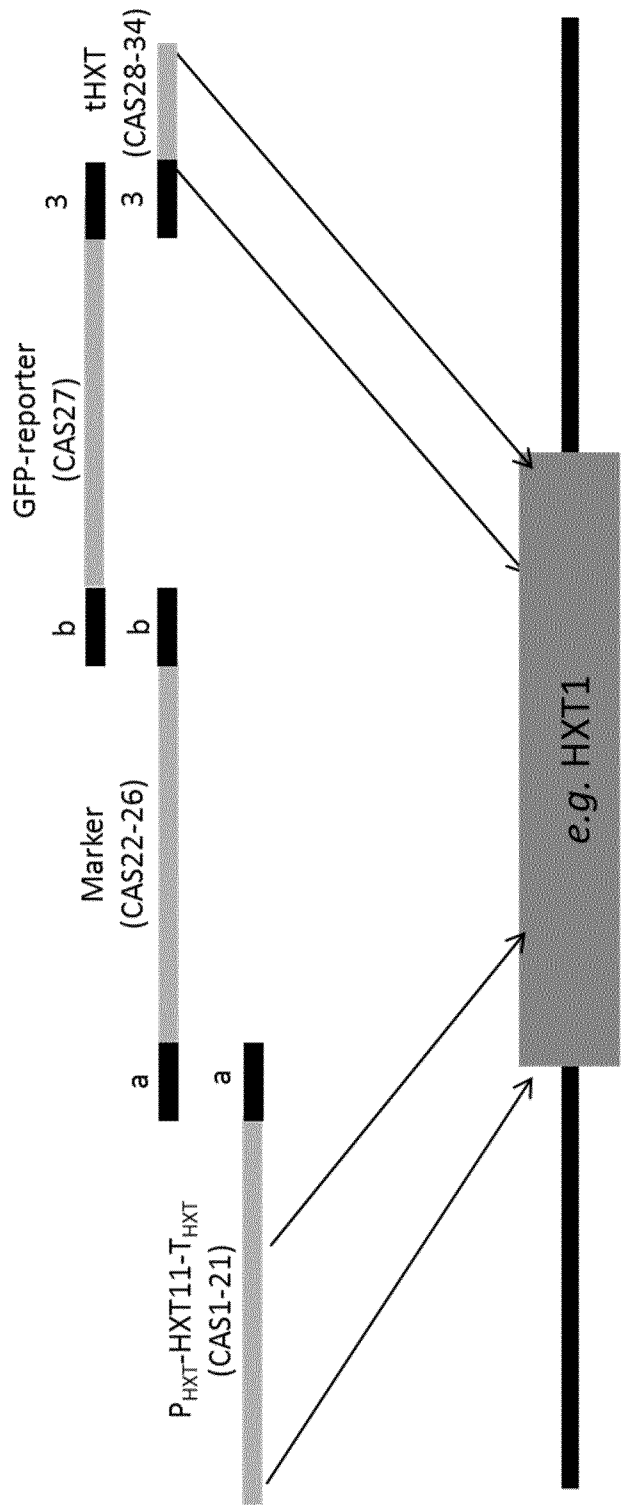


Fig. 2A

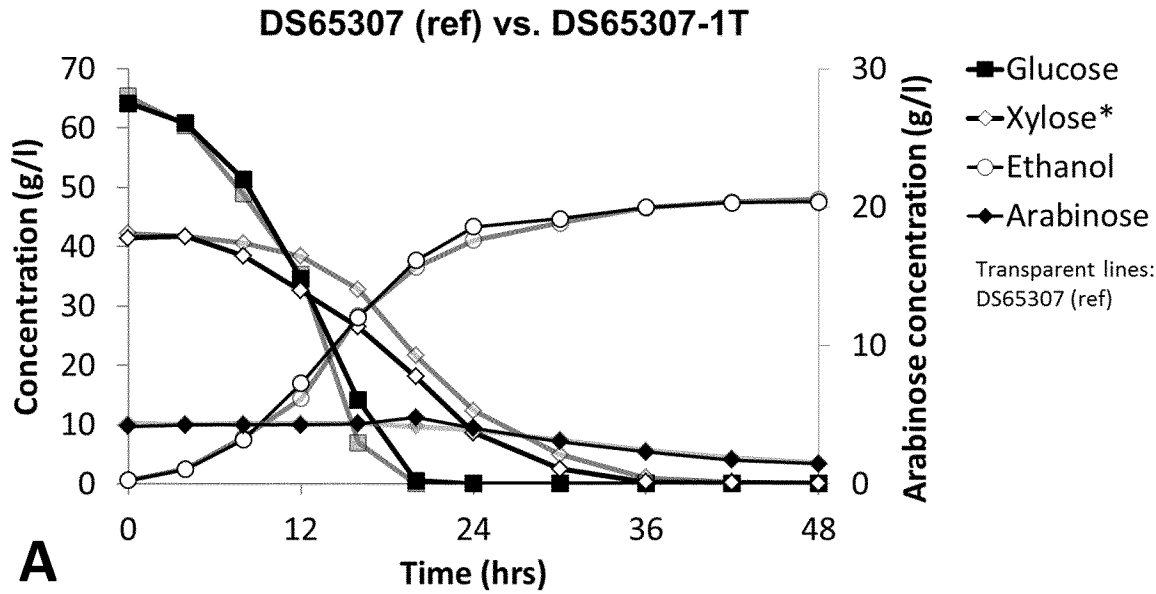


Fig. 2B

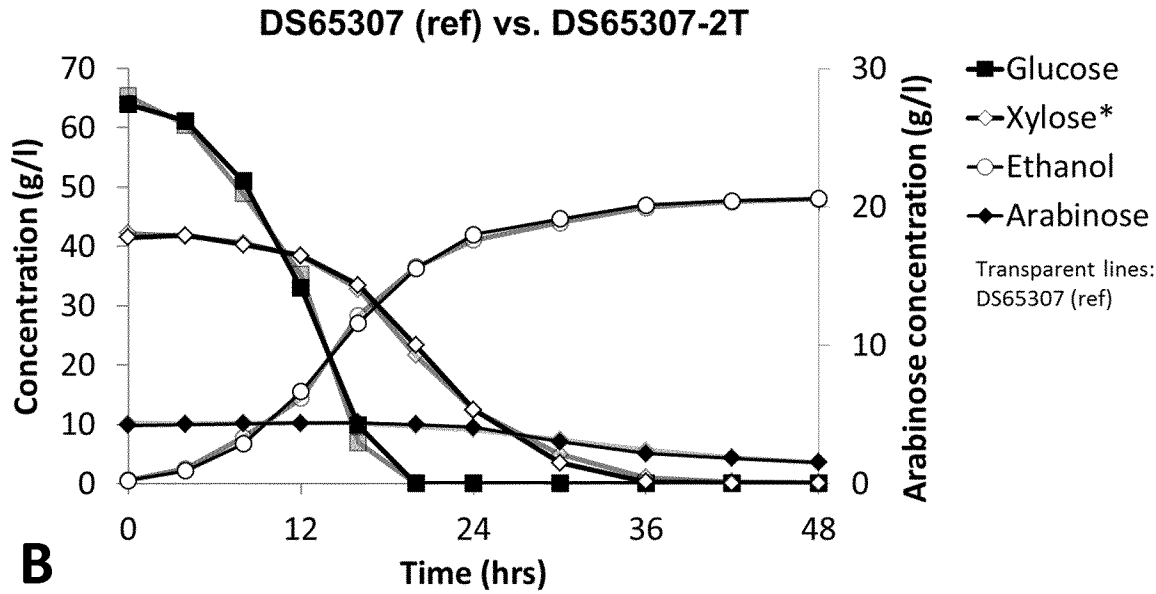


Fig. 2C

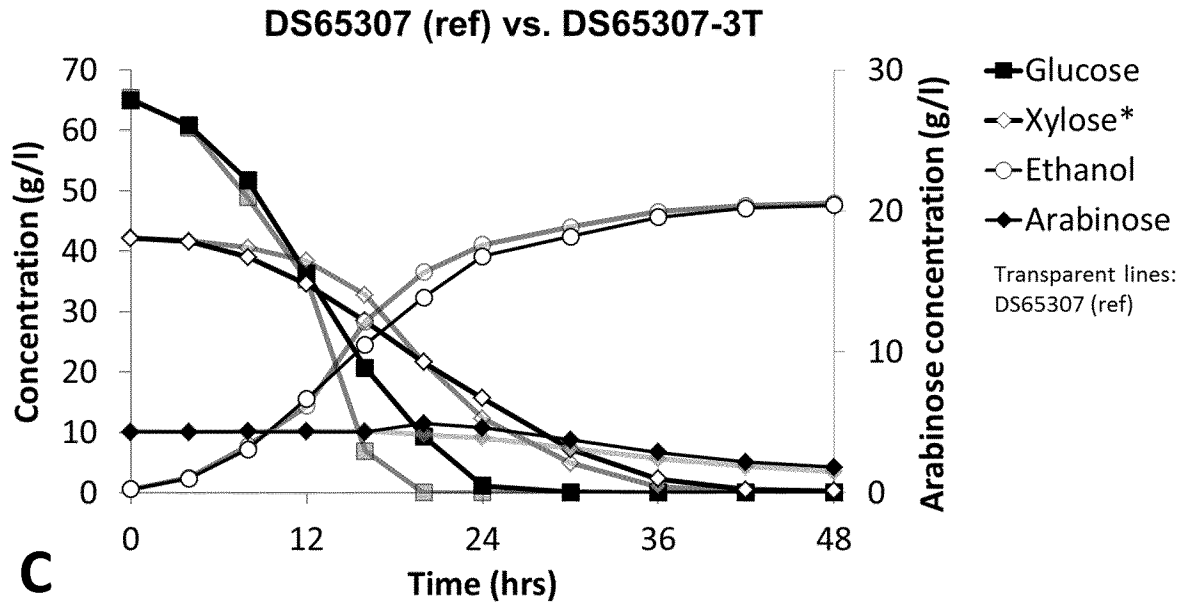


Fig. 2D

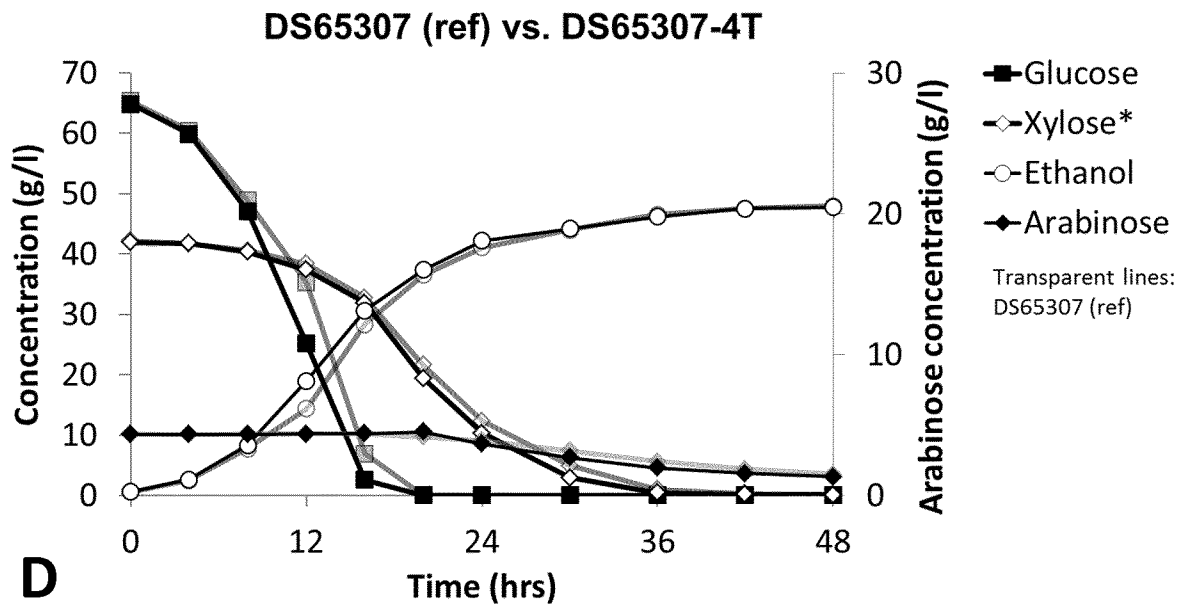


FIG. 2CD

Fig. 2E

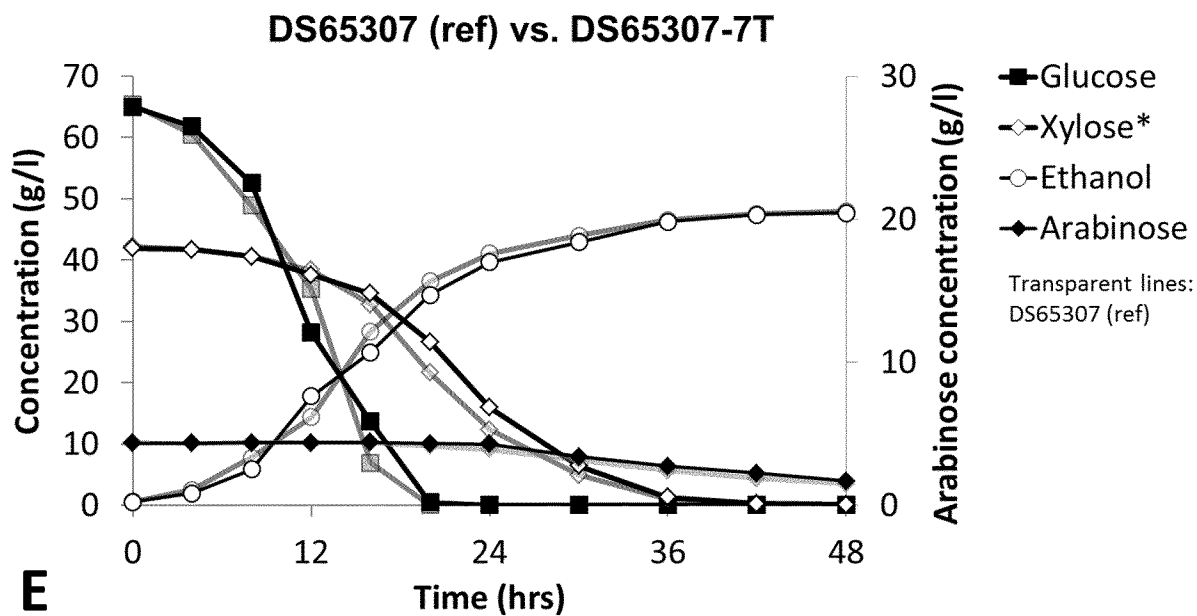
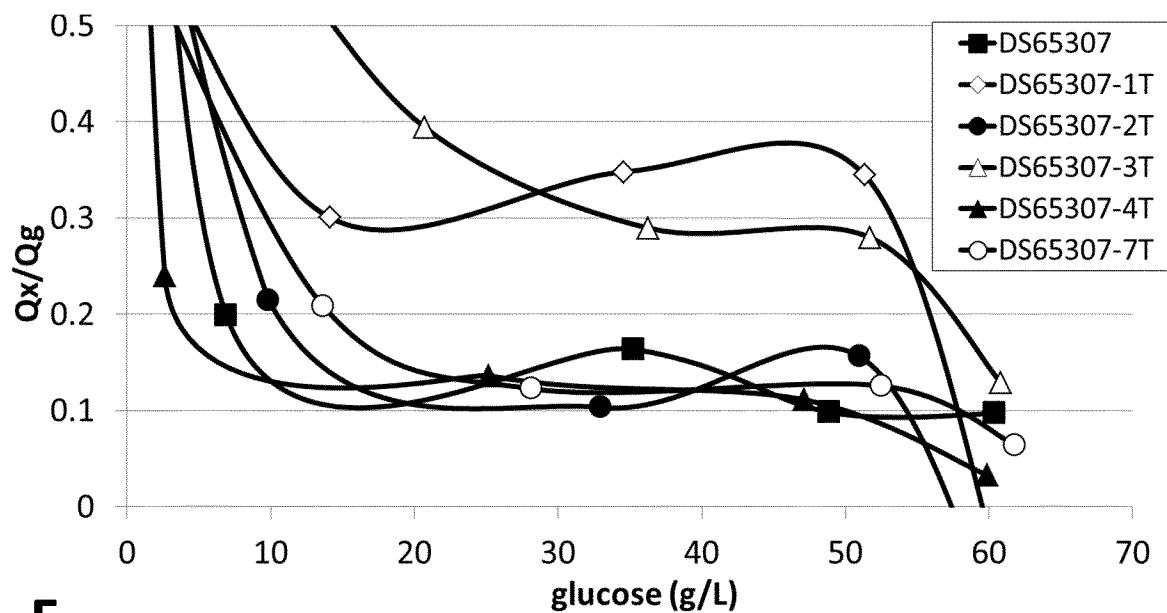


Fig. 2F



INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2015/066610A. CLASSIFICATION OF SUBJECT MATTER
INV. C07K14/395 C12N1/18
ADD. C12R1/865

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)
C07K

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

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

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	A. FARWICK ET AL: "Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 111, no. 14, 24 March 2014 (2014-03-24), pages 5159-5164, XP055161224, ISSN: 0027-8424, DOI: 10.1073/pnas.1323464111 the whole document -/--	1-11



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

16 September 2015

Date of mailing of the international search report

28/09/2015

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

Sonnerat, Isabelle

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/066610

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	-& A. FARWICK ET AL: "Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose", PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES, vol. 111, no. 14, 24 March 2014 (2014-03-24), XP055161297, ISSN: 0027-8424, DOI: 10.1073/pnas.1323464111 supporting information, table S3 -----	
X	Jeroen G Nijland ET AL: "Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in <i>Saccharomyces cerevisiae</i>", Biotechnology for biofuels, 1 January 2014 (2014-01-01), pages 168-168, XP055160601, England DOI: 10.1186/s13068-014-0168-9 Retrieved from the Internet: URL: http://www.ncbi.nlm.nih.gov/pubmed/25505932 [retrieved on 2015-01-12] the whole document -----	1-11
X	WO 2011/011796 A2 (UNIV CALIFORNIA; BOARD OF TRUSTEES OF THE UNIV ILLINOS; BP CORP [US]) 27 January 2011 (2011-01-27) example 11 -----	1-3,11
X	JING DU ET AL: "Discovery and characterization of novel d-xylose-specific transporters from <i>Neurospora crassa</i> and <i>Pichia stipitis</i>", MOLECULAR BIOSYSTEMS, vol. 6, no. 11, 1 January 2010 (2010-01-01), page 2150, XP055160928, ISSN: 1742-206X, DOI: 10.1039/c0mb00007h abstract page 2151, left-hand column, paragraph 4 - right-hand column, paragraph 3 page 2154, right-hand column, paragraph 2 - paragraph 3 -----	1-3,11
X	WO 2007/018442 A2 (FUNDACAO DA FACULDADE DE CIENC [PT]; TRAVASSOS LEANDRO MARIA JOSE [PT]) 15 February 2007 (2007-02-15) abstract paragraph [0033] - paragraph [0035] ----- -/--	1-3,11

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2015/066610

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 2 495 306 A1 (BUTALCO GMBH [CH]) 5 September 2012 (2012-09-05) abstract paragraphs [0037], [0053], [0054], [0128]	1-3,11
A	----- SOO RIN KIM ET AL: "Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol", TRENDS IN BIOTECHNOLOGY, vol. 30, no. 5, 1 May 2012 (2012-05-01), pages 274-282, XP028415841, ISSN: 0167-7799, DOI: 10.1016/J.TIBTECH.2012.01.005 [retrieved on 2012-01-27] page 278, left-hand column, paragraph 4 - page 279, right-hand column, paragraph 1 table 1 -----	1-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2015/066610

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2011011796 A2	27-01-2011	AU 2010275388 A1 CA 2765313 A1 CN 102625844 A EA 201290068 A1 EP 2456878 A2 JP 2013500019 A KR 20120089631 A NZ 597001 A US 2011020910 A1 US 2013316437 A1 US 2015010985 A1 WO 2011011796 A2	12-01-2012 27-01-2011 01-08-2012 30-05-2013 30-05-2012 07-01-2013 13-08-2012 30-05-2014 27-01-2011 28-11-2013 08-01-2015 27-01-2011
WO 2007018442 A2	15-02-2007	CA 2618273 A1 EP 1960424 A2 JP 2009502191 A PT 103331 A US 2009053784 A1 WO 2007018442 A2	15-02-2007 27-08-2008 29-01-2009 28-02-2007 26-02-2009 15-02-2007
EP 2495306 A1	05-09-2012	EP 2495306 A1 US 2012225464 A1	05-09-2012 06-09-2012