

PRODUCTION OF XYLITOL FROM GLUCOSE BY A RECOMBINANT STRAIN

Field of the Invention

5 The present invention relates to a method of using genetically modified microorganisms for the manufacture of xylitol, and a method of preparing a genetically modified microorganism that is capable of converting in one step readily available carbon sources, such as D-glucose, into xylitol.

Background of the Invention

10 Xylitol is a polyalcohol or sugar alcohol (alditol) of formula $(\text{CHOH})_3(\text{CH}_2\text{OH})_2$, that has applications in hygiene and nutraceutical formulations and products.

Xylitol is used as a diabetic sweetener which is roughly as sweet as sucrose with 33% fewer calories. Unlike other natural or synthetic sweeteners, xylitol is actively beneficial for dental health by reducing caries to a third in regular use and helpful to remineralization.

15 Xylitol is naturally found in low concentrations in the fibers of many fruits and vegetables, and can be extracted from various berries, oats, and mushrooms, as well as fibrous material such as corn husks and sugar cane bagasse, and birch.

However, industrial production starts from xylan (a hemicellulose) extracted from hardwoods or corncobs, which is hydrolyzed into xylose and catalytically hydrogenated into
20 xylitol.

Purification of xylose and also xylitol presents therefore a significant problem. A number of processes of this type are known. U.S. patents 4,075,406 and 4,008,285 can be mentioned as examples.

25 The reduction of D-xylose into xylitol can also be achieved in a microbiological process using either yeast strains isolated from nature (wild type strains) or genetically engineered strains.

However, obtaining the substrate, D-xylose, in a form suitable for yeast fermentation is a problem because inexpensive xylose sources such as sulphite liquor from pulp and paper processes contain impurities which inhibit yeast growth.

30 An attractive alternative method for the manufacture of xylitol is obtaining it by fermentation of a cheap and readily available substrate, such as D-glucose.

In the state of the art, there are some recombinant microorganisms described able to produce xylitol in certain amounts during a one-step fermentation of any common carbon sources other than D-xylose and D-xylulose.

35 These recombinant microorganisms, especially osmophilic yeasts, are for example *Zygosaccharomyces rouxii*, *Candida polymorpha*, and *Torulopsis candida*, initially known as

producers of significant amounts of a xylitol closely related pentitol, which is D-arabitol, from D-glucose (*Lewis D.H. & Smith D.C., 1967, New Phytol. 66:143-184*).

Thus, the international patent application WO 94/10325 provides methods for constructing such recombinant hosts being capable of producing xylitol when grown on carbon sources other than D-xylulose or D-xylose, and other than polymers or oligomers or mixtures thereof.

In the current patent application, this goal is achieved through modification of the metabolism of the desired microorganism, preferably a naturally occurring yeast microorganism, by introducing and expressing desired heterologous genes.

This goal is also achieved by further modification of the metabolism of such desired microorganism, so as to overexpress and/or inactivate the activity or expression of certain genes homologous to such microorganism in its native state.

The method provided in this patent application for the production of xylitol utilized an altered D-arabitol biosynthesis pathway, and such pathway being notably altered by extending the preexisting D-arabitol pathway by the introduction and overexpression of the genes coding for D-xylulose-forming D-arabitol dehydrogenase (EC 1.1.1.11) and xylitol dehydrogenase (EC 1.1.1.9) into an D-arabitol-producing microorganism.

However, the yield of xylitol in the trials described in WO 94/10325 was only approximately 7.7 g/l after 48 hours of cultivation in a medium with yeast extract.

To try to optimize this first result, it was further proposed in WO 94/10325 to inactivate, using mutagenesis or gene disruption, the genes coding for transketolase (EC 2.2.1.1) and/or the gene coding for D-xylulokinase (EC 2.7.1.17), and also to overexpress the genes coding for the enzymes of the oxidative branch of the pentose-phosphate pathway, and specifically D-glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and/or 6-phospho-D-gluconate dehydrogenase (EC 1.1.1.44) and/or D-ribulose-5-phosphate epimerase gene (EC 5.1.3.1) in such microorganisms.

But, whatever the genetic combination employed, the xylitol titer was never more than 9 g/l.

There is therefore still an unsatisfied need for a better genetic manipulation of xylitol producing strains in order to optimize its production, and thus make it commercially profitable.

Summary of the Invention

The present invention relates to a recombinant host cell capable of producing xylitol, wherein said host cell comprises:

an heterologous nucleic acid sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) using D-arabitol as substrate and producing D-xylulose as product; and,

5 an heterologous nucleic acid sequence encoding a NADPH-specific xylitol dehydrogenase using D-xylulose as substrate and producing xylitol as product.

Preferably, the host cell does not consume D-arabitol as a sole carbon source. More preferably, the host cell is selected from bacteria, fungi and yeast. In a preferred embodiment, the host cell is an osmophilic or osmotolerant yeast, in particular *Pichia ohmeri*.

10 Preferably, the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) is from *E. coli* or *Ralstonia solanacearum*. More preferably, the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) comprises or consists in the sequence of SEQ ID No 2 or 43 or a sequence with 1-3 additions, substitutions or deletions of amino acids. In a preferred embodiment, the sequence encoding the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) comprises or consists in the sequence of SEQ ID No 3 or 42.

15 Preferably, the NADPH-specific xylitol dehydrogenase is a xylitol dehydrogenase from *Pichia stipitis* or *Gluconobacter oxydans* mutated for changing the cofactor specificity from NADH to NADPH. More preferably, the NADPH-specific xylitol dehydrogenase comprises or consists in the sequence of SEQ ID No 5 or 8 or a sequence with 1-3 additions, substitutions or deletions of amino acids. In a preferred embodiment, the sequence
20 encoding the NADPH-specific xylitol dehydrogenase comprises or consists in the sequence of SEQ ID No 6 or 9.

Preferably, the host cell is capable of producing a xylitol titer of at least 15 g/l in the supernatant after a 48 h culture.

25 Preferably, the host cell is a strain selected from strains I-4982, I-4960 and I-4981 deposited at the CNCM.

Preferably, the host cell comprises several copies of a sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase and/or several copies of a sequence encoding the NADPH-specific xylitol dehydrogenase.

30 The present invention also relates to a method for producing xylitol comprising culturing a recombinant host cell according to any one of claims 1-13, and recovering xylitol.

It additionally relates to a nucleic acid comprising or consisting in a nucleic acid sequence selected from the group consisting of SEQ ID No 1, 3, 7 and 9, an expression cassette or vector comprising said nucleic acid.

35 Finally, the present invention relates to the use of a recombinant host cells according to the present invention for producing xylitol.

Detailed Description of the Invention

Definitions

5 As used herein, by a "carbon source other than D-xylose and D-xylulose" is meant a carbon substrate for xylitol production other than D-xylose and D-xylulose or polymers or oligomers or mixtures thereof (such as xylan and hemicellulose). The carbon source preferably includes D-glucose, and various D-glucose-containing syrups and mixtures of D-glucose with other sugars.

10 As used herein, by "gene" is meant a nucleic acid sequence that may code for a protein, in particular a DNA sequence.

 As used herein, by "vector" is meant a plasmid or any other DNA sequence which is able to carry genetic information, specifically DNA, into a host cell. The vector can further contain a marker or reporter suitable for use in the identification of cells transformed with the vector, and origins of replication that allow for the maintenance and replication of the vector
15 in one or more prokaryotic or eukaryotic hosts. A "plasmid" is a vector, generally circular DNA that is maintained and replicates autonomously in at least one host cell.

 As used herein, by "expression vector" is meant a vector similar to a vector but which supports expression of a gene or encoding nucleic acid that has been cloned into it,
20 after transformation into a host. The cloned gene or encoding nucleic acid is usually placed under the control of (i.e., operably linked to) certain control sequences such as promoter sequences, that can be provided by the vector or by the recombinant construction of the cloned gene. Expression control sequences will vary depending on whether the vector is designed to express the operably linked gene in a prokaryotic or eukaryotic host and can
25 additionally contain transcriptional elements such as enhancer elements (upstream activation sequences) and termination sequences, and/or translational initiation and termination sites.

 As used herein, by "host" is meant a cell, prokaryotic or eukaryotic, that is utilized as the recipient and carrier of recombinant material.

30 As used herein, by "Oxidative Branch of the Pentose-Phosphate Pathway" is meant to include the part of the pentose-phosphate shunt that catalyzes oxidative reactions, such as reactions catalyzed by D-glucose-6-phosphate dehydrogenase (EC 1.1.1.49) gluconolactonase (EC 3.1.1.17), and 6-phospho-D-gluconate dehydrogenase (EC 1.1.1.44), and that utilizes hexose substrates to form pentose phosphates. The "non-oxidative" part of
35 the pentose-phosphate pathway (which also catalyzes the net formation of ribose from D-glucose) is characterized by non-oxidative isomerizations such as the reactions catalyzed by transketolase (EC 2.2.1.1), ribose-5-phosphate isomerase (EC 5.3.1.6), D-ribulose-5-

phosphate-3-epimerase (EC 5.1.3.1) and transaldolase (EC 2.2.1.2). See Biological Chemistry, H.R. Mahler & E.H.Cordes, Harper & Row, publishers, New York, 1966, pp. 448-454.

As used herein, by "encoding nucleic acid" is meant a nucleic acid molecule (preferably DNA). Encoding nucleic acid is capable of encoding a protein and can be prepared from a variety of sources. These sources include genomic DNA, cDNA, synthetic DNA, and combinations thereof.

"Heterologous", as used herein, is understood to mean that a gene or encoding sequence has been introduced into the cell by genetic engineering. It can be present in episomal or chromosomal form. The gene or encoding sequence can originate from a source different from the host cell in which it is introduced. However, it can also come from the same species as the host cell in which it is introduced but it is considered heterologous due to its environment which is not natural. For example, the gene or encoding sequence is referred to as heterologous because it is under the control of a promoter which is not its natural promoter, it is introduced at a location which differs from its natural location. The host cell may contain an endogenous copy of the gene prior to introduction of the heterologous gene or it may not contain an endogenous copy.

Object of the invention

According to the invention, the native metabolic pathways of a specific microbial host are manipulated so as to decrease or eliminate the utilization of carbon into purposes other than xylitol production.

Such a genetically modified host strain is thus able to produce xylitol in one fermentation step with a high yield. For instance, the xylitol titer after 48 h of culture in the supernatant is more than 15 g/l, preferably more than 25 g/l, still more preferably more than 50, 60, 70, 80, 90 or 100 g/l.

In the practical realization of the invention, the genetically modified host of the invention is also characterized by its ability to synthesize xylitol from structurally unrelated carbon sources such as D-glucose, and not just from D-xylose and/or D-xylulose.

Preferably, the genetically modified host of the invention is also capable of secreting the synthesized xylitol into the medium.

Specifically, in the exemplified and preferred embodiments, the genetically modified host of the invention is characterized by a pathway in which arabitol is an intermediate in xylitol formation.

Accordingly, the recombinant host strain of the invention is characterized by the following genetic alterations:

(1) a heterologous nucleic acid encoding a protein possessing NAD⁺-specific D-arabitol 4-oxidoreductase (D-xylulose-forming) activity has been introduced into the host cell - thus providing for the conversion of D-arabitol to D-xylulose; and

5 (2) a heterologous nucleic acid encoding a protein possessing NADPH-specific xylitol dehydrogenase activity has been introduced into the host cell- thus providing for the conversion of D-xylulose to xylitol.

The choice of the microorganism

10 The microorganisms or host strains suitable for the present invention are capable of producing D-arabitol from glucose. More particularly, they are capable of producing significant amounts of D-arabitol from glucose under high osmotic pressure medium.

By "high osmotic pressure medium" is intended here to refer to medium containing 10-60% D-glucose, preferably about 25% D-glucose.

15 By "significant amounts of D-arabitol" is intended at least 100 g/L of D-arabitol. In particular, a microorganism or host strain is considered as producing significant amounts of D-arabitol when the microorganism or host strain produces 100g/L D-arabitol in a medium containing 25 % D-glucose in batch conditions.

20 Examples of host strains capable of producing significant amounts of D-arabitol from glucose include the osmophilic or osmotolerant yeasts, in particular those belonging to the species *Pichia*, *Kodamaea*, *Candida*, *Zygoaccharomyces*, *Debaromyces*, *Metschnikowia* and *Hansenula*; or the D-arabitol producing fungi, in particular those belonging to the species *Dendryphiella* and *Schizophyllum*, in particular *Dendryphiella salina* and *Schizophyllum commune*.

25 Examples of the microorganisms of the genus *Pichia* include *Pichia ohmeri*, *Pichia stipitis*, *Pichia farinosa*, *Pichia haplophila*. Examples of the microorganisms of the genus *Candida* include *Candida polymorpha* and *Candida tropicalis*. Examples of the microorganisms of the genus *Zygoaccharomyces* include *Zygoaccharomyces rouxii*. Other examples include *Torulopsis candida* and *Torulaspora hansenii*. Examples of the microorganisms of the genus *Metschnikowia* include *Metschnikowia pulcherrima*,
30 *Metschnikowia reukaufii*, *Metschnikowia bicuspidata*, *Metschnikowia lunata* and *Metschnikowia zobellii*. As specific strains, *Metschnikowia pulcherrima* ATCC 18406, *Metschnikowia reukaufii* ATCC 18407, *Metschnikowia bicuspidata* ATCC 24179, *Metschnikowia lunata* ATCC 22033, *Metschnikowia zobellii* ATCC 22302 and *Metschnikowia pulcherrima* FERM BP-7161 can be mentioned. These strains can be
35 obtained from American Type Culture Collection, Address: 12301 Parklawn Drive, Rockville, Maryland 20852, United States of America. *Metschnikowia pulcherrima* FERN BP-7161 was originally deposited at the National Institute of Bioscience and Human-Technology, Agency

of Industrial Science and Technology, Ministry of International Trade and Industry (postal code: 305-8566, 1-3 Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken, Japan) on January 16, 1998, under the deposition number of FERM P-16592 and transferred from the original deposit to international deposit based on Budapest Treaty on May 15, 2000, and has been deposited as deposition number of FERM BP-7161. In a specific aspect, the microorganism has the accession number FERM BP-7161. For more information, refer to EP1065276.

The microorganism can be genetically engineered in order to improve its capacity of producing D-arabitol and/or reducing its capacity to use D-arabitol for a goal distinct from xylitol production.

For the invention, the host strain is advantageously chosen by its specific metabolic attributes:

- it may be a producer of significant amounts of D-arabitol from glucose as detailed above, in particular under high osmotic pressure medium, for example medium containing 10-60% D-glucose, and preferably 25% D-glucose ("Normal" medium usually contains only 2-3% glucose.)
- it may not consume D-arabitol as a sole carbon source;
- its redox balance permits the generation of the cofactors needed for the corresponding ketopentose / pentose alcohol conversion.

In one embodiment of the invention, the osmophilic yeast *Pichia ohmeri* (and their mutagenized derivatives) has been employed as a model and as a preferred host. *Pichia ohmeri* has initially been isolated from cucumber brine and commonly used in food industry for fermentation in pickles, rinds, and fruits.

It is known by the one skilled in the art that yeasts species such as *Pichia*, *Zygosaccharomyces*, *Debaromyces* and *Hansenula* are able to grow in low water activity environments, on the opposite of *Saccharomyces cerevisiae*. These osmotolerant or osmophilic yeasts accumulate compatible solute like glycerol, D-arabitol, erythritol and mannitol which protect and stabilize enzymes, thereby enabling the cellular functions in osmotic conditions of growth. The polyols produced also play a role in redox balancing.

In a preferred aspect, the microorganism is *Pichia ohmeri*. Indeed, the main characteristic of the host strain *Pichia ohmeri* is to produce only D-arabitol as compatible solute, in contrast to *Zygosaccharomyces rouxii* producing glycerol and D-arabitol. In addition, the metabolic pathway from glucose to D-arabitol is well known in *Pichia ohmeri*.

As described in *Zygosaccharomyces rouxii* (J.M. INGRAM and W.A. WOOD, 1965, Journal of Bacteriology, Vol.89, N°5, 1186-1194), the carbon flux in *Pichia ohmeri* goes through the oxidative part of the Pentose-Phosphate Pathway (PPP) to convert D-Glucose into D-ribulose-5-P with the concomitant production of two molecules of NADPH. D-ribulose-

5-P is dephosphorylated to D-ribulose and then reduced to D-arabitol. In *Pichia ohmeri* host strain, the Pentose Phosphate Pathway (PPP) is very active and has been determined to be higher than 50%.

In a preferred embodiment, the host cell is a mutant *Pichia ohmeri* deposited on March 7, 2012, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 PARIS Cedex 15, under number I-4605.

Redox reactions and enzymes

The cofactors NADH and NADPH are essential to a multitude of biological functions, acting in so-called redox reactions as carriers of electrons from one reaction to another. Cells need to maintain the metabolic equilibrium of the two redox couples NADH/NAD⁺ and NADPH/NADP⁺ knowing that the NADPH/NADP⁺ couple is maintained in a more reduced state than the NADH/NAD⁺ couple to provide a thermodynamic driving force. NADH, which is mostly found in the oxidized form NAD⁺, is the main co-factor in catabolic reactions where it is involved in the oxidative release of energy from nutrients. In contrast to NADH, NADPH is re-oxidized exclusively in anabolic reactions or during times of oxidative stress.

Any metabolic engineering strategy that involves redox reactions has to function under these cellular constraints. It has been done in the genetically modified strain that is the object of the invention.

As found by the inventor, notably described in the PhD Thesis entitled "*Contribution à l'étude du métabolisme des pentitols chez Pichia ohmeri*" (Sophie Huchette, University of Sciences and Technics of Lille, 1992), it has been demonstrated that the reactions involved into the oxidoreduction of ketopentoses are catalyzed by two different enzymes.

Thus, the host strain has an enzyme defined as a NADPH-specific D-ketopentose-oxidoreductase, forming D-arabitol from D-ribulose and forming xylitol from D-xylulose. The host strain also possesses a NADH-specific D-ketopentose-oxidoreductase, forming ribitol and xylitol respectively from D-ribulose and D-xylulose. This enzyme is closed to the well-known NAD⁺-specific xylitol dehydrogenase E.C 1.1.1.9 from *Pichia stipitis* (XYL2). As only intracellular D-ribulose is available in contrast to D-xylulose, the host strain balances the NADPH/NADP⁺ redox couple directly with the re-oxidization of NADPH through cytosolic formation of D-arabitol from D-ribulose. Then, D-arabitol is secreted into the broth via a passive diffusion.

The inventors found that the lack in intracellular D-xylulose would be the main reason for the non-production of xylitol by the host strain even if *Pichia ohmeri* possesses all

the enzymatic tools to produce this polyol via NADH- or NADPH-specific-D-ketopentose-oxidoreductase.

Indeed, it was chosen to clone into the wild type host strain *Pichia ohmeri* a gene encoding a protein possessing NAD⁺-specific D-arabitol 4-oxidoreductase (D-xylulose-forming) activity (E.C.1.1.1.11) allowing the cytosolic D-arabitol to be converted to D-xylulose and NADH.

So, intracellular D-xylulose becomes available into the genetically modified strain and could be reduced by the intrinsic NADH- and NADPH-specific-D-ketopentose-oxidoreductase. However, the strain is devoid of the endogenous enzymes able to efficiently transform D-xylulose into xylitol. Therefore, it is necessary to genetically engineer the strain in order to introduce a heterologous xylitol dehydrogenase.

In the patent WO 94/10325, it was chosen to clone the NAD⁺-specific Xylitol dehydrogenase (E.C 1.1.1.9) from *Pichia stipitis* (XYL2) allowing the production of xylitol and balancing the NADH/NAD⁺ redox couple with the oxidation of NADH produced by the previous metabolic step. But as mentioned before, the results are not really convincing.

The inventors found that, by cloning a gene encoding a mutated protein possessing NADPH-specific xylitol dehydrogenase, the D-xylulose is converted to xylitol to balance the NADPH/NADP⁺ redox couple such as done by the intrinsic production of D-arabitol from D-ribulose.

Due to its low affinity of the NADPH-specific D-ketopentose-oxidoreductase for D-arabitol, the *Pichia ohmeri* wild type host strain does not consume the extracellular D-arabitol.

Because of the introduction of a NAD⁺-specific D-arabitol 4-oxidoreductase (D-xylulose-forming) activity into the genetically modified strain, the D-arabitol produced into the broth could be well consumed by the modified strain the same way as the cytosolic D-arabitol.

Consequently, xylitol is produced at the same time from intracellular and extracellular D-arabitol.

Its production could be improved by enhancing the efficiency of the xylitol pathway extension to totally avoid the exportation of the intermediary D-arabitol.

Thus only xylitol would be produced from D-glucose with the same physiological effect as D-arabitol. This improvement could be the result of the genetic modifications but also of the adaptation of the culture conditions.

The choice of the two enzymatic activities to be cloned in the host strain

The choice of these two enzymatic activities is supported by their cofactor specificity, as described above.

The first enzyme oxidizes D-arabitol into D-xylulose.

Two types of D-arabitol dehydrogenases are known: D-xylulose-forming (EC 1.1.1.11) (D-arabinitol NAD⁺ 4-oxidoreductase) and D-ribulose-forming (EC 1.1.1.250). Unless otherwise stated, it is the D-xylulose-forming arabitol dehydrogenase that is intended
5 herein and referred to herein as arabitol dehydrogenase. D-ribulose-forming dehydrogenases are found in wild-type yeasts and fungi.

D-xylulose-forming arabitol dehydrogenases are mainly known in bacteria. For instance, they have been identified in *Enterobacteriaceae*, in particular *E. coli*, *Klebsiella aerogenes*, and *Aerobacter aerogenes* strain PRL-R3, in *Gluconobacter oxydans*, and
10 additionally also in *Pichia stipitis*. In particular, several enzymes are referenced in UniprotKB database, such as, *Klebsiella pneumoniae* (# O52720), *Ralstonia solanacearum* (# P58708), *Yersinia pestis* (# P58709), *Aerobacter aerogenes* (#L8BEF0), *E. coli* (# K3EX35, I2ZSJ5, W1BYD6, W1H8N7, E7U4R7).

For the purposes of the present invention, *Escherichia coli* is the preferred source
15 of the NAD⁺-specific D-arabitol 4-oxidoreductase (D-xylulose-forming) gene. More specifically, its amino acid sequence is disclosed in SEQ ID No 2. In particular, SEQ ID Nos 1 and 3 disclose nucleic acids encoding NAD⁺-specific D-arabitol 4-oxidoreductase of *Escherichia coli*. The encoding sequence has been optimized for *Pichia ohmeri* by taking into account its codon specificity.

In addition, *Ralstonia solanacearum* is also a preferred source of the NAD⁺-specific
20 D-arabitol 4-oxidoreductase (D-xylulose-forming) gene. More specifically, its amino acid sequence is disclosed in SEQ ID No 43. In particular, SEQ ID No 42 disclose nucleic acids encoding NAD⁺-specific D-arabitol 4-oxidoreductase of *Ralstonia solanacearum*. The encoding sequence has been optimized for *Pichia ohmeri* by taking into account its codon
25 specificity.

The second enzyme converts D-xylulose into xylitol.

Although the majority of yeasts and fungi possess an endogenous xylitol dehydrogenase (EC 1.1.1.9) gene, the change of their cofactor specificity from NADH to
30 NADPH is necessary for the implementation of the present invention. Indeed, a key aspect of the present invention is to use a NADPH-specific xylitol dehydrogenase. In addition, this enzyme is preferably overexpressed in the host.

Numerous xylitol dehydrogenases are known and several scientific articles teach how to change the cofactor specificity from NADH to NADPH. Watanabe et al (*J. Biol. Chem.*, 2005, 280, 10340-10345) discloses mutated xylitol dehydrogenase of *Pichia stipitis*
35 with a modified cofactor specificity, especially the triple mutant (D207A/I208R/F209S) and the quadruple mutant (D207A/I208R/F209S/N211R). The amino acid sequence of the

quadruple mutant is disclosed in SEQ ID No 5. A double mutant of xylitol dehydrogenase of *Gluconobacter oxydans* (D38S/M39R) with a NADPH cofactor specificity is disclosed in Ehrensberger et al (2006, *Structure*, 14, 567-575). The amino acid sequence of the double mutant is disclosed in SEQ ID No 8.

5 The mutation and cloning of the *Pichia stipitis* XYL2 nucleic acid sequence encoding the NADPH-specific xylitol dehydrogenase have been prepared by the inventors. In particular, SEQ ID Nos 4 and 6 disclose nucleic acids encoding specific NADPH xylitol dehydrogenase of *Pichia stipitis*.

10 Alternatively, the inventors have also performed the mutation and cloning of the *Gluconobacter oxydans* nucleic acid sequence encoding the NADPH-specific xylitol dehydrogenase. In particular, SEQ ID Nos 7 and 9 disclose nucleic acids encoding NADPH specific xylitol dehydrogenase of *Gluconobacter oxydans*. The encoding sequence has been optimized for *Pichia ohmeri* by taking into account its codon specificity.

15 Expression cassette, vector and recombinant host cell

 In a particular aspect, the present invention relates to a nucleic acid comprising an encoding sequence optimized for *Pichia ohmeri* selected from the group consisting of SEQ ID No 1, 3, 7, 9 and 42.

20 It further relates to an expression cassette comprising a nucleic acid comprising an encoding sequence optimized for *Pichia ohmeri* selected from the group consisting of SEQ ID No 1, 3, 7, 9 and 42.

 It also relates to the nucleic acid construct of SEQ ID No 4 and a nucleic acid comprising said nucleic acid construct.

25 In addition, it relates to a recombinant vector, in particular an expression vector, comprising said nucleic acid or expression cassette. Generally, an expression cassette comprises all the elements required for gene transcription and translation into a protein. In particular, it comprises a promoter, optionally an enhancer, a transcription terminator and the elements for translation. More particularly, the promoter used to control the expression of the NADPH-specific xylitol dehydrogenase is selected in order to drive a strong
30 expression. Indeed, this enzyme is preferably overexpressed in the host cell. Such promoters are well-known in the art. For instance, the promoter could be the *P. ohmeri* ribulose reductase promoter (poRR) or the *P. ohmeri* phosphoglycerate kinase (poPGK1).

 It relates to a recombinant vector, in particular an expression vector, comprising a nucleic acid encoding a NAD⁺-specific D-arabitol 4-oxidoreductase and a nucleic acid
35 encoding NADPH-specific xylitol dehydrogenase. It also relates to a kit comprising a recombinant vector, in particular an expression vector, comprising a nucleic acid encoding a NAD⁺-specific D-arabitol 4-oxidoreductase, and a recombinant vector, in particular an

expression vector, comprising a nucleic acid encoding NADPH-specific xylitol dehydrogenase.

Preferably, said NAD⁺-specific D-arabitol 4-oxidoreductase and NADPH-specific xylitol dehydrogenase are selected among the enzymes disclosed above. In particular, said NAD⁺-specific D-arabitol 4-oxidoreductase comprises or consists of an amino acid sequence of SEQ ID No 2 or 42 or a sequence with 1-3 additions, substitutions or deletions of amino acids. In particular, said NADPH-specific xylitol dehydrogenase comprises or consists of an amino acid sequence of SEQ ID No 5 or 8 or a sequence with 1-3 additions, substitutions or deletions of amino acids.

A preferred vector is a plasmid. Suitable plasmids are well-known by the person skilled in the art and can be for instance selected among those specifically disclosed in Examples.

Genetically modified host of the invention are first produced by cloning the genes coding for NAD⁺-specific D-arabitol 4-oxidoreductase and for NADPH-specific xylitol dehydrogenase under control of suitable promoters into a recombinant vector and introduced into the host cells of the D-arabitol producing organism by transformation.

The present invention relates to a recombinant or genetically engineering host cell comprising an heterologous nucleic acid sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) and an heterologous nucleic acid sequence encoding a NADPH-specific xylitol dehydrogenase. The NAD⁺-specific D-arabitol 4-oxidoreductase uses D-arabitol as substrate and produces D-xylulose as product. The NADPH-specific xylitol dehydrogenase uses D-xylulose as substrate and produces xylitol. The sequence encoding NADPH-specific xylitol dehydrogenase and NAD⁺-specific D-arabitol 4-oxidoreductase can be episomal or be integrated into the chromosome of the host cell. Indeed, genetically stable transformants are preferably constructed through transformation systems using a vector, whereby a desired DNA is integrated into the host chromosome. Such integration occurs *de novo* within the cell or can be assisted by transformation with a vector which functionally inserts itself into the host chromosome, with DNA elements which promote integration of DNA sequences in chromosomes.

The recombinant or genetically engineering host cell can comprise several copies of a sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase and/or several copies of a sequence encoding the NADPH-specific xylitol dehydrogenase, preferably integrated into the host cell chromosome. In particular, the recombinant or genetically engineering host cell can comprise two, three or four sequences encoding a NAD⁺-specific D-arabitol 4-oxidoreductase and/or two, three or four sequences encoding the NADPH-specific xylitol dehydrogenase. For instance, the host cell may comprise two or three NAD⁺-specific D-arabitol 4-oxidoreductases from *E. coli* and/or one or two NAD⁺-specific D-arabitol 4-

oxidoreductases from *R. solanacearum*, more specifically two or three NAD⁺-specific D-arabitol 4-oxidoreductases from *E. coli* and/or one NAD⁺-specific D-arabitol 4-oxidoreductase from *R. solanacearum*. The NAD⁺-specific D-arabitol 4-oxidoreductases can be from the same organism or from different organisms. The NADPH-specific xylitol dehydrogenases can be from the same organism or from different organisms. For instance, the host cell may comprise one, two or three NADPH-specific xylitol dehydrogenases from *P. stipitis* and/or one, two or three NADPH-specific xylitol dehydrogenases from *G. oxydans*, more specifically one NADPH-specific xylitol dehydrogenase from *P. stipitis* and/or three NADPH-specific xylitol dehydrogenases from *G. oxydans*.

In a particular aspect of the invention, the recombinant or genetically engineering host cell is a *Pichia ohmeri* strain comprising:

- two NAD⁺-specific D-arabitol 4-oxidoreductases and two NADPH-specific xylitol dehydrogenases; or
- two NAD⁺-specific D-arabitol 4-oxidoreductases from *E. coli* and two NADPH-specific xylitol dehydrogenases, one from *P. stipitis* and the other from *G. oxydans*; or
- two NAD⁺-specific D-arabitol 4-oxidoreductases and three NADPH-specific xylitol dehydrogenases; or
- two NAD⁺-specific D-arabitol 4-oxidoreductases from *E. coli* and three NADPH-specific xylitol dehydrogenases, one from *P. stipitis* and two from *G. oxydans*; or
- three NAD⁺-specific D-arabitol 4-oxidoreductases and three NADPH-specific xylitol dehydrogenases; or
- three NAD⁺-specific D-arabitol 4-oxidoreductases, two from *E. coli* and one from *R. solanacearum*, and three NADPH-specific xylitol dehydrogenases, one from *P. stipitis* and two from *G. oxydans*; or
- four NAD⁺-specific D-arabitol 4-oxidoreductases and four NADPH-specific xylitol dehydrogenases; or
- four NAD⁺-specific D-arabitol 4-oxidoreductases, three from *E. coli* and one from *R. solanacearum*, and four NADPH-specific xylitol dehydrogenases, one from *P. stipitis* and three from *G. oxydans*.

The host cell is selected among the microorganisms detailed above. In a preferred embodiment, the host cell is *Pichia ohmeri*. The starting host cell is preferably the mutant *Pichia ohmeri* deposited at the CNCM under number I-4605.

In a particular aspect of the invention, the host cell is a strain selected from strains I-4982, I-4960 and I-4981 deposited at the CNCM.

The present invention relates to a method for producing xylitol comprising culturing the recombinant or genetically engineering host cell in a culture medium and recovering the produced xylitol. Preferably, the culture medium provides the microorganism with the convenient carbon source. The carbon source preferably includes D-glucose, and various D-glucose-containing syrups and mixtures of D-glucose with other sugars. The method may further comprises a step of purifying xylitol.

The present invention relates to the use of a recombinant or genetically engineering host cell as disclosed herein for producing xylitol.

Xylitol produced by such genetically modified strains can be purified from the medium of the hosts of the invention according to any technique known in the art. For example, US 5,081,026, incorporated herein by reference, described the chromatographic separation of xylitol from yeast cultures. Thus, from the fermentation step, xylitol can be purified from the culture medium using chromatographic steps as described in US 5,081,026, followed by crystallization.

Other characteristic features and advantages of the invention will be apparent on reading the following Examples. However, they are given here only as an illustration and are not limiting.

Figures and sequences

Figure 1: 12 ABYWMP: Restriction map of the synthesized NAD⁺-specific D-arabitol 4-oxidoreductase from *E. coli* flanked by *Ascl* and *SphI* restriction sites.

Figure 2a: lig7.78: Restriction map of the NADH-specific xylitol dehydrogenase from *Pichia stipitis*.

Figure 2b: 12AALQTP: Restriction map of the synthesized NADPH-specific xylitol dehydrogenase from *Pichia stipitis* flanked by *HindIII* and *SacII* restriction sites.

Figure 3: 13AAYSYP: Restriction map of the synthesized NADPH-specific xylitol dehydrogenase from *Gluconobacter oxydans* flanked by *Ascl* and *SphI* restriction sites.

Figure 4: Construction of an expression cassette consisting of an open reading frame flanked by a poRR promoter and terminator using overlap PCR.

Figure 5: 12 AAMCJP: Restriction map of the synthesized tagatose-3-epimerase of *Pseudomonas cichorii* flanked by *HindIII* and *SacII* restriction sites.

Figure 6: Construction of *P. ohmeri* shuttle vectors with poLEU2 and poURA3 selection markers.

Figure 7: pEVE2523: Restriction map of the *P. ohmeri* poURA3 expression vector pEVE2523, with a cloned expression cassette containing the open reading frame of tagatose-3-epimerase of *Pseudomonas cichorii* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

Figure 8: pEVE2560: Restriction map of the *P. ohmeri* poLEU2 expression vector pEVE2560, with a cloned expression cassette containing the open reading frame of tagatose-3-epimerase of *Pseudomonas cichorii* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

5 **Figure 9:** Construction of a *P. ohmeri* vector for overexpression of *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase.

10 **Figure 10: pEVE3284:** Restriction map of the *P. ohmeri* pEVE3284 expression vector, with a cloned expression cassette containing the NADPH-specific xylitol dehydrogenase of *Gluconobacter oxydans* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

Figure 11: Construction of a *P. ohmeri* vectors for overexpression of *Pichia stipitis* NADPH-specific xylitol dehydrogenase.

15 **Figure 12: pEVE2562/pEVE2564:** Restriction map of the *P. ohmeri* pEVE2562/pEVE2564 expression vectors, with a cloned expression cassette containing the NADPH-specific xylitol dehydrogenase of *Pichia stipitis* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator with either a poURA3 or poLEU2 selection marker, respectively.

Figure 13: Construction of a *P. ohmeri* vector for overexpression of *Pichia stipitis* NADH-specific xylitol dehydrogenase.

20 **Figure 14: pEVE2563:** Restriction map of the *P. ohmeri* pEVE2563 expression vector, with a cloned expression cassette containing the NADH-specific xylitol dehydrogenase of *Pichia stipitis* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

25 **Figure 15:** Construction of a *P. ohmeri* vector for overexpression of *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase under the control of the *P. ohmeri* ribulose reductase (poRR) promoter and terminator using a poURA3 selection marker.

30 **Figure 16: pEVE2839:** Restriction map of the *P. ohmeri* pEVE2839 expression vector, with a cloned expression cassette containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

Figure 17: Construction of a *P. ohmeri* vector for overexpression of *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator using a poURA3 selection marker.

35 **Figure 18: pEVE3102:** Restriction map of the *P. ohmeri* pEVE3102 expression vector, with a cloned expression cassette containing the NAD⁺-specific D-arabitol 4-

oxidoreductase of *E. coli* flanked by a *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and ribulose reductase (poRR) terminator.

Figure 19: pEVE3123: Restriction map of the *P. ohmeri* pEVE3123 expression vector, with a cloned expression cassette containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* flanked by a *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and a transketolase (poTKL) terminator and a poURA3 selection marker.

Figure 20: Construction of a *P. ohmeri* vector for overexpression of *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator using a poLEU2 selection marker.

Figure 21: pEVE3157: Restriction map of the *P. ohmeri* pEVE3157 expression vector, with a cloned expression cassette containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* flanked by a *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and a transketolase (poTKL) terminator and a poLEU2 selection marker.

Figure 22: Construction of a *P. ohmeri* loxP vector with a poLEU2 selection marker

Figure 23: pEVE2787: Restriction map of the *P. ohmeri* pEVE2787 integration vector, with a cloned *P. ohmeri* LEU2 selection marker under the control of the endogenous promoter and terminator, flanked by two loxP sites.

Figure 24: 12ABTV4P: Restriction map of the synthesized *nat1* gene from *Streptomyces noursei* flanked by *Ascl* and *SphI* restriction sites.

Figure 25: pEVE2798: Restriction map of the *P. ohmeri* pEVE2798 expression vector, with a cloned *nat1* marker under the control of a ribulose reductase (poRR) promoter and an orotidine-5'-phosphate decarboxylase (poURA3) terminator.

Figure 26: Construction of a *P. ohmeri* loxP vector with a *nat1* selection marker.

Figure 27: pEVE2852: Restriction map of the *P. ohmeri* pEVE2852 integration vector, with a cloned with a cloned *nat1* marker under the control of a ribulose reductase (poRR) promoter and an orotidine-5'-phosphate decarboxylase (poURA3) terminator, flanked by two loxP sites.

Figure 28: pEVE2855: Restriction map of the *P. ohmeri* pEVE2855 integration vector, with a cloned fragment homologous to the 5' region upstream of the LEU2 open reading frame and a *nat1* selection marker flanked by two loxP sites.

Figure 29: Construction of a *P. ohmeri* loxP vector for the deletion of the LEU2 open reading frame.

Figure 30: pEVE2864: Restriction map of the *P. ohmeri* pEVE2864 integration vector, with a cloned fragment homologous to the 5' region upstream of the LEU2 open reading frame and fragment homologous to the 3' region downstream of the LEU2 open reading frame, and a *nat1* selection marker flanked by two loxP sites.

Figure 31: Construction of a double expression plasmids comprising the NADPH-specific xylitol dehydrogenase of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*.

5 **Figure 32: pEVE3318:** Restriction map of the *P. ohmeri* pEVE3318 expression vector, containing the double expression construct of the NADPH-specific xylitol dehydrogenase of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*.

Figure 33: pEVE2862: Restriction map of the *P. ohmeri* pEVE2862 expression vector, containing the *P. ohmeri* *LEU2* marker flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and an orotidine-5'-phosphate decarboxylase (poURA3) terminator.

10 **Figure 34:** Construction of an integrative vector for the genomic expression of the *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *P. stipitis* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*.

Figure 35: pEVE2865: Restriction map of the *P. ohmeri* pEVE2865 integration vector, containing the *P. ohmeri* *LEU2* marker flanked by two loxP sites.

15 **Figure 36: pEVE3387:** Restriction map of the *P. ohmeri* pEVE3387 integration vector, containing the double expression construct of the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* with a *P. ohmeri* *LEU2* selection marker flanked by two loxP sites.

20 **Figure 37:** Construction of double/triple expression plasmids comprising the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*.

25 **Figure 38: pEVE3322/pEVE3324:** Restriction map of the *P. ohmeri* pEVE3322/pEVE3324 expression vectors, containing either the double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* or the triple expression construct of two NADPH-specific xylitol dehydrogenase genes of *G. oxydans* and one NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*.

30 **Figure 39:** Construction of an integrative vector for the genomic expression of the *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *G. oxydans* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*.

35 **Figure 40: pEVE3390/pEVE3392:** Restriction map of the *P. ohmeri* pEVE3390/pEVE3392 integration vectors, containing either the double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* or the triple expression construct of two NADPH-specific xylitol dehydrogenase genes of *G. oxydans* and one NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* with a *P. ohmeri* *LEU2* selection marker flanked by two loxP sites.

Figure 41: Construction of an integrative vector for the genomic expression of the *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *G. oxydans* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*.

Figure 42: pEVE4390: Restriction map of the *P. ohmeri* pEVE4390 expression vector, containing the double expression construct of the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* and the NADPH-specific xylitol dehydrogenase gene of *G. oxydans* with a *P. ohmeri* LEU2 selection marker flanked by two loxP sites.

Figure 43: 13AB2EGF: Restriction map of the synthesized NAD⁺-specific D-arabitol 4-oxidoreductase from *R. solanacearum* flanked by *Ascl* and *SphI* restriction sites.

Figure 44: Construction of an integrative vector for the genomic expression of the *R. solanacearum* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *G. oxydans* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*

Figure 45: pEVE3898: Restriction map of the *P. ohmeri* pEVE3898 expression vector, with a cloned expression cassette containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *Ralstonia solanacearum* flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and terminator.

Figure 46: pEVE4077: Restriction map of the *P. ohmeri* pEVE4077 expression vector, with a double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum*.

Figure 47: pEVE4377: Restriction map of the *P. ohmeri* pEVE4377 integration vector, with a double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum* and the poLEU2 selection marker flanked by two loxP sites.

25 SEQUENCE LISTING

SEQ ID No	Description
1	Sequence encoding NAD ⁺ -specific D-arabitol 4-oxidoreductase from <i>E. coli</i> flanked by <i>Ascl</i> and <i>SphI</i> restriction sites
2	Amino acid sequence of NAD ⁺ -specific D-arabitol 4-oxidoreductase from <i>E. coli</i>
3	Sequence encoding NAD ⁺ -specific D-arabitol 4-oxidoreductase from <i>E. coli</i>
4	Sequence encoding NADPH-specific xylitol dehydrogenase from <i>Pichia stipitis</i> flanked by <i>HindIII</i> and <i>SacII</i> restriction sites
5	Amino acid sequence of NADPH-specific xylitol dehydrogenase from <i>P. stipitis</i>
6	Sequence encoding NADPH-specific xylitol dehydrogenase from <i>Pichia stipitis</i>
7	Sequence encoding NADPH-specific xylitol dehydrogenase from <i>Gluconobacter oxydans</i> flanked by <i>Ascl</i> and <i>SphI</i> restriction sites

8	Amino acid sequence of NADPH-specific xylitol dehydrogenase from <i>Gluconobacter oxydans</i>
9	Sequence encoding NADPH-specific xylitol dehydrogenase from <i>Gluconobacter oxydans</i>
10	Sequence encoding tagatose-3-epimerase of <i>Pseudomonas cichorii</i> ST24
11	Amino acid sequence of tagatose-3-epimerase of <i>Pseudomonas cichorii</i> ST24
28	Sequence encoding the <i>nat1</i> gene of <i>Streptomyces noursei</i> flanked by <i>Ascl</i> and <i>SphI</i> restriction sites
42	Sequence encoding the NAD ⁺ -specific D-arabitol 4-oxidoreductase from <i>R. solanacearum</i> flanked by <i>Ascl</i> and <i>SphI</i> restriction sites
43	Amino acid sequence of NAD ⁺ -specific D-arabitol 4-oxidoreductase from <i>R. solanacearum</i>

EXAMPLES

Example 1. Choice of a *Pichia ohmeri* strain as preferred host for genetic engineering

5 As host strain of choice, *Pichia ohmeri*:

- is a producer of significant amounts of arabitol from glucose, under high osmotic pressure medium, for example medium containing 10-60% D-glucose, and preferably 25% D-glucose ("Normal" medium usually contains only 2-3% glucose.)

10 - has a redox balance that permits the generation of the cofactors needed.

As an illustration of its performances, the following tables indicate the enzyme activities involved in the arabitol metabolic pathway of *Pichia ohmeri* (Sophie HUCHETTE Thesis, 1992)

15 *The Hexose Monophosphate Pathway: from Glucose-6-P to D-Ribulose-5-P and D-Xylulose-5-P*

The oxidative part of the PPP, also named the Hexose Monophosphate Pathway (HMP), is a NADPH-producing pathway. The two NADP⁺-dependent enzymes which are Glucose-6-P dehydrogenase (E.C.1.1.1.49) and 6-P-Gluconate dehydrogenase (E.C.1.1.1.44) participate to the oxidation of 1 mole of glucose-6-P in 1 mole of D-ribulose-5-P and generate 2 moles of NADPH.

Table 1Hexose Monophosphate Pathway in *P. ohmeri* ATCC 20209

<i>Enzymes</i>	<i>Specific activity U/mg</i>
NADP ⁺ G6P dehydrogenase	1.5
NADP ⁺ 6PG dehydrogenase	0.55

One unit of enzyme activity was defined as the consumption of 1 μ mole of NAD(P)H or NAD(P)⁺ per minute per mL of crude extract. One unit of specific activity was defined as one unit of enzyme activity per mg of proteins in crude extract.

5

The kinetic parameters of the following enzymes were determined: D-ribulose-5-P 3-epimerase (E.C. 5.1.3.1), D-ribose-5-P keto-isomerase (E.C.5.3.1.6), transketolase (E.C.2.2.1.1) and acidic phosphatases (E.C. 3.1.3.2).

10 **Table 2**Kinetic parameters of enzymes using D-Ribulose-5-P as substrate
in *P. ohmeri* ATCC 20209

<i>Enzymes</i>	<i>K_M mM</i>	<i>V_M U/mg</i>
D-Ribulose-5-P 3-epimerase	6.3	3
D-Ribose-5-P keto-isomerase	0.35	1.8
Acid phosphatase	4.3	0.65

One unit of enzyme activity was defined as the consumption of 1 μ mole of NAD(P)H or NAD(P)⁺ per minute per mL of crude extract. One unit of specific activity was defined as one unit of enzyme activity per mg of proteins in crude extract.

15 **Table 3**Kinetic parameters of enzymes using D-Xylulose-5-P as substrate
in *P. ohmeri* ATCC 20209

<i>Enzymes</i>	<i>K_M mM</i>	<i>V_M U/mg</i>
D-Ribulose-5-P 3-epimerase	6.6	0.7
Transketolase (D-ribose-5-P)	0.2	0.9
Transketolase (Erythrose-4-P)	0.6	1.45
Acid phosphatase	16	0.11

One unit of enzyme activity was defined as the consumption of 1 μ mole of NAD(P)H or NAD(P)⁺ per minute per mL of crude extract. One unit of specific activity was defined as one unit of enzyme activity per mg of proteins in crude extract.

In vivo, D-xylulose-5-P, synthesized from the epimerization of D-ribulose-5-P, enters efficiently into the non-oxidative part of the PPP via the transketolization. Consequently, D-xylulose-5-P is not available for its dephosphorylation into D-xylulose.

5 NADH and NADPH specific D-Ketopentose-oxidoreductases

D-Ribulose and D-Xylulose are produced by dephosphorylization of D-Ribulose-5-P and D-Xylulose-5-P.

The Michaelis-Menten constants highlight the affinities of the NADH and NADPH-D-ketopentose-oxidoreductases for each substrate and the corresponding maximum velocities.

Table 4

NADH-specific D-ketopentose-oxidoreductase kinetic parameters of <i>P. ohmeri</i> ATCC 20209		
Substrate	K_M mM	V_M U/mg
D-Ribulose	90	1
Ribitol	16	0.16
D-Xylulose	5	0.6
Xylitol	7	0.2

One unit of enzyme activity was defined as the consumption of 1 μ mole of NAD(P)H or NAD(P)⁺ per minute per mL of crude extract. One unit of specific activity was defined as one unit of enzyme activity per mg of proteins in crude extract.

NADH-specific D-ketopentose-oxidoreductase, forming ribitol and xylitol respectively from D-ribulose and D-xylulose shows a higher affinity for D-xylulose than D-ribulose. The reverse reaction shows a good affinity for xylitol and ribitol explaining the good growth of the host strain on these two polyols.

Table 5

NADPH-specific D-ketopentose-oxidoreductase kinetic parameters of <i>P. ohmeri</i> ATCC 20209		
Substrate	K_M mM	V_M U/mg
D-Ribulose	72	3.4
D-Arabitol	1300	0.8
D-Xylulose	262	1.5
Xylitol	200	0.15

One unit of enzyme activity was defined as the consumption of 1 μ mole of NAD(P)H or NAD(P)⁺ per minute per mL of crude extract. One unit of specific activity was defined as one unit of enzyme activity per mg of proteins in crude extract.

5 NADPH-specific D-ketopentose-oxidoreductase, forming D-arabitol from D-ribulose and forming xylitol from D-xylulose shows a higher affinity for D-ribulose than D-xylulose. The reverse reaction shows a very low affinity for D-arabitol explaining the non-growth of the host strain on this polyol.

10 The two ketopentose-oxidoreductases from the host strain were characterized as different from the previous enzymes described in *Saccharomyces rouxii* by Ingram and Wood, 1965 (Journal of Bacteriology, vol.89, n°5, 1186-1194). Indeed, in *Saccharomyces rouxii*, no forward reaction was detected on D-ribulose and NADH and a backward reaction was detected on D-arabitol with NADPH.

The Haldane relationship predicts *in vivo* enzyme kinetic behaviors.

15 **Table 6**

Haldane constants determination: NADH-specific D-Ketopentose-oxidoreductase

<i>Substrat/Product</i>	<i>K_{eq} mM⁻¹</i>
D-Ribulose/Ribitol	78
D-Xylulose/Xylitol	104

Haldane constants determination: NADPH-specific D-Ketopentose-oxidoreductase

<i>Substrat/Product</i>	<i>K_{eq} mM⁻¹</i>
D-Ribulose/D-Arabitol	104
D-Xylulose/Xylitol	24

20 The two enzymes favor the forward reaction (D-ketopentose oxidation) over the backward reaction (pentitol reduction).

The PPP in the host strain is extremely efficient and 2 moles of NADPH are generated from 1 mole of glucose consumed. Consequently, NADPH would be available in excess for both anabolic reactions and maintenance reactions. The host strain must produce D-arabitol from D-ribulose or xylitol from D-xylulose to balance the NADPH/NADP⁺ redox couple.

The inhibitory effect of NADP⁺ on NADPH-specific D-ketopentose-oxidoreductase has been determined *in vitro*. The activity is 80% less when NADP⁺ is added in excess. Even if this concentration is not compatible with the intracellular NADP⁺ concentration, this

result gives some overview of the role of the NADPH-specific D-ketopentose oxidoreductase into the balance of the NADPH/NADP⁺ redox couple.

The host strain produces only D-arabitol from D-ribulose as D-xylulose is not available because of the entrance of D-xylulose-5-P into the non-oxidative part of the PPP.

5 The link between the production of D-arabitol and the NADPH/NADP⁺ redox balance has been demonstrated in the host strain by evaluating the impact of the overexpression of Glucose-6-P dehydrogenase onto the D-arabitol production. So, the obtained strain harbors a G6PDH activity 1.5 times higher and produces 10% more of D-arabitol compared to the host strain (FR2772788).

10

Example 2. *Pichia ohmeri* codon usage

The codon usage of *P. ohmeri* was determined from the available DNA and corresponding amino acid sequence of five *P. ohmeri* genes: transketolase, glucose-6-phosphate dehydrogenase (FR 2772788), ribulose reductase, beta-isopropylmalate dehydrogenase - *LEU2* (Piredda and Gaillardin, Yeast, vol.10:1601-1612 (1994) and orotidine-5'-phosphate decarboxylase - *URA3* (Piredda and Gaillardin, 1994, *supra*).

Every individual gene was divided in nucleotide triplets encoding for a single amino acid. The five genes consisted of a total of 2091 codons.

For each amino acid, the number of every codon present in the five genes was counted, divided by 2091 and multiplied by 1000. This way, the frequency of a specific codon in 1000 codons was estimated.

The preliminary codon usage of *P. ohmeri* is depicted in **Table 7**.

All heterologous genes expressed in *P. ohmeri*, except the xylitol dehydrogenase from *P. stipitis*, were codon optimized using this table and the Optimizer program obtained from <http://genomes.urv.es/OPTIMIZER/>.

The obtained sequence was sent for gene synthesis after manual addition of recognition sites for restriction enzymes at the respective 5' and 3' ends of the sequence encoding the enzyme.

30 **Table 7.** Codon usage table of *P. ohmeri* derived from 5 coding sequences (CDS)

***Pichia ohmeri* [gbpln]: 5 CDS's (2091 codons)**

fields: [triplet] [frequency: **per thousand**] ([number])

TTT 10.5 (22)	TCT 30.6 (64)	TAT 7.7 (16)	TGT 5.7 (12)
TTC 30.1 (63)	TCC 23.4 (49)	TAC 27.3 (57)	TGC 1.4 (3)
TTA 5.3 (11)	TCA 4.8 (10)	TAA 1.4 (3)	TGA 0.0 (0)
TTG 64.1 (134)	TCG 9.6 (20)	TAG 1.0 (2)	TGG 12.9 (27)
CTT 10.5 (22)	CCT 12.0 (25)	CAT 3.3 (7)	CGT 5.7 (12)

24

CTC 12.0 (25)	CCC 0.0 (0)	CAC 15.8 (33)	CGC 1.0 (2)
CTA 0.0 (0)	CCA 34.0 (71)	CAA 12.4 (26)	CGA 0.0 (0)
CTG 2.9 (6)	CCG 0.5 (1)	CAG 17.7 (37)	CGG 0.5 (1)
ATT 27.7 (58)	ACT 22.0 (46)	AAT 7.7 (16)	AGT 1.9 (4)
ATC 30.6 (64)	ACC 24.4 (51)	AAC 29.2 (61)	AGC 2.4 (5)
ATA 2.4 (5)	ACA 3.3 (7)	AAA 11.0 (23)	AGA 26.3 (55)
ATG 14.3 (30)	ACG 1.4 (3)	AAG 64.1 (134)	AGG 0.0 (0)
GTT 27.3 (57)	GCT 46.9 (98)	GAT 23.0 (48)	GGT 60.7 (127)
GTC 19.1 (40)	GCC 27.7 (58)	GAC 35.9 (75)	GGC 10.5 (22)
GTA 1.9 (4)	GCA 11.0 (23)	GAA 18.7 (39)	GGA 12.0 (25)
GTG 21.5 (45)	GCG 3.3 (7)	GAG 46.9 (98)	GGG 1.0 (2)

Example 3. Cloning of the *E. coli* bacterial NAD⁺-specific D-arabitol 4-oxidoreductase (D-xylulose-forming) gene

A DNA fragment encoding the NAD⁺-specific D-arabitol 4-oxidoreductase altD from *E. coli* was chemically synthesized (GeneArt® Gene Synthesis, Life Technologies, Regensburg, Germany), according to the submitted sequence of SEQ ID NO: 1.

Nucleotides 1441 to 2808 of sequence AF378082.1 (obtained from <http://www.ncbi.nlm.nih.gov/nuccore/AF378082>) coding for the altD gene were used as template and subjected to codon optimization for use in *P. ohmeri* ATCC 20209 according to **Table 7** of **example 2**, using the Optimizer program obtained from <http://genomes.urv.es/OPTIMIZER/>.

At the 5' and 3' ends of the resulting sequence, nucleotides encoding for the recognition sites of the restriction enzymes *Ascl* (GGCGCGCC) and *SphI* (GCATGC) respectively, were added in order to facilitate further cloning.

Additionally, an adenosine triplet was included in front of the start ATG to account for an adenosine at the -3 position in the Kozak-like sequence of yeasts.

The final sequence (SEQ ID NO: 1) was then submitted for synthesis (GeneArt, Regensburg, Germany).

The synthesized DNA fragment encoding the NAD⁺-specific D-arabitol 4-oxidoreductase from *E. coli* was delivered as 5 µg lyophilized plasmid DNA in a pMK-RQ derived vector (12ABYWMP, **Figure 1**).

For further sub-cloning the gene was released by restriction cutting with *Ascl* and *SphI* enzymes (New England Biolabs, Ipswich, Massachusetts).

Example 4. Mutagenesis and cloning of the *Pichia stipitis* NADH and NADPH-specific xylitol dehydrogenase

Cloning of the Pichia stipitis NADH-specific xylitol dehydrogenase gene

The known nucleotide sequence of the yeast (*Pichia stipitis*) gene XYL2, encoding xylitol dehydrogenase (Kötter et al., Curr. Genet. 18:493-500 (1990)) was cloned in the plasmidic vector lig 7.78 following the teaching of FR 2 765 589 (see **example 4** and **Figure 7** of this patent). The restriction map of the vector is presented in **Figure 2a**.

Mutagenesis and cloning of the Pichia stipitis NADPH-specific xylitol dehydrogenase gene

A DNA fragment encoding the NADPH-specific xylitol dehydrogenase XYL2 from *Pichia stipitis* was chemically synthesized (GeneArt® Gene Synthesis, Life Technologies, Regensburg, Germany) according to the sequence of SEQ ID NO: 4.

Nucleotides 319 to 1410 of sequence X55392.1 (obtained from <http://www.ncbi.nlm.nih.gov/nuccore/X55392.1>) coding for the XYL2 gene were used as template.

According to the paper from Watanabe et al. (*J. Biol. Chem.*, 2005, 280, 10340-10345), the cofactor preference of the xylitol dehydrogenase could be changed from NADH to NADPH by introducing four published amino acid mutations: D207A/I208R/F209S/N211R (numbering based on P22144 protein sequence obtained from <http://www.uniprot.org/uniprot/P22144>).

Accordingly, the codons encoding for D207, I208, F209 and N211 were manually replaced by GCT, AGA, TCA and AGA in the corresponding sequence, respectively.

Additionally, nucleotides coding for the recognition sites of the restriction enzymes *Hind*III (AAGCTT) and *Sac*II (CCGCGG) were manually included at the respective 5' and 3' ends, in order to facilitate further cloning.

Furthermore, an adenosine triplet was included in front of the start ATG to account for an adenosine at the -3 position in the Kozak-like sequence of yeasts. The final sequence (SEQ ID NO: 4) was submitted for synthesis (GeneArt, Regensburg, Germany).

The synthesized DNA fragment encoding the NADPH-specific xylitol dehydrogenase from *P. stipitis* was delivered as 5 µg lyophilized plasmid DNA in a pMA-T derived vector (12AALQTP, **Figure 2b**).

Example 5. Mutagenesis and cloning of the *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase gene

A DNA fragment encoding the NADPH-specific xylitol dehydrogenase *Xdh* from *Gluconobacter oxydans* was chemically synthesized (GeneArt® Gene Synthesis, Life

Technologies, Regensburg, Germany), according to the submitted sequence of SEQ ID NO: 7.

Nucleotides 1063 to 1851 of sequence AB091690.1 (obtained from <http://www.ncbi.nlm.nih.gov/nuccore/AB091690.1>) coding for the Xdh gene were used as template and subjected to codon optimization for use in *P. ohmeri* ATCC 20209 according to **Table 7 (Example 2)** using the Optimizer program obtained from <http://genomes.urv.es/OPTIMIZER/>.

Based on the publication by Ehrensberger et al. (*Structure*, 2006, 14, 567-575), the cofactor specificity of the enzyme could be changed from NADH to NADPH by introducing two published amino acid mutations: D38S/M39R (numbering based on Q8GR61 protein sequence obtained from <http://www.uniprot.org/uniprot/Q8GR61>).

Thus, the codons encoding for D38 and M39 were manually replaced by TCT and AGA in the corresponding sequence, respectively. Additionally, nucleotides encoding for the recognition sites of the restriction enzymes *AscI* (GGCGCGCC) and *SphI* (GCATGC) were manually included at the respective 5' and 3' ends, in order to enable further cloning.

Furthermore, an adenosine triplet was included in front of the start ATG to account for an adenosine at the -3 position in the Kozak-like sequence of yeasts. The final sequence (SEQ ID NO: 7), was submitted for synthesis (GeneArt, Regensburg, Germany).

The synthesized DNA fragment encoding the NADPH-specific xylitol dehydrogenase from *Gluconobacter oxydans* was delivered as 5 µg lyophilized plasmid DNA in a pMA-T derived vector (13AAYSYP, **Figure 3**). For further subcloning, the gene was released by restriction cutting with *AscI* and *SphI* enzymes (New England Biolabs, Ipswich, Massachusetts).

Example 6. Construction of a *P. ohmeri* vector for heterologous gene expression using the poURA3 selection marker

The cloning of a vector with replaceable:

- promoter,
- open reading frame, and
- terminator elements

was performed by two successive overlap PCRs of three individual fragments (**Figure 4**).

The vector was originally planned as an expression model, to test the cloning and the overexpression of the tagatose 3-epimerase gene in the recombinant *Pichia ohmeri* strain.

As it will be described below, the tagatose 3-epimerase gene has been cloned into specific *Ascl* – *SphI* restriction sites cassette, allowing the cloning of any gene of interest by using these same sites of insertion.

The cloning was conceived by the following way.

5 In a first PCR (PCR1), a 490 bp long ribulose reductase promoter fragment of *P. ohmeri* flanked by *SpeI* and *Ascl* sites (underlined in primer sequence) was amplified using:

- primer EV2960:

GAACTAGTGGATCCGTAGAAATCTTG (SEQ ID No 12)

and

10 - primer EV2961:

CTTTGTTCAATTTTGGCGCGCCTTTAGTTTAATAAGGGTCCGTG (SEQ ID No 13)

Additionally, at the 5' end of the reverse primer EV2961, a 13 nucleotide long fragment representing the 5' end of the tagatose-3-epimerase gene was added.

15 This fragment together with the 8 nucleotides of the *Ascl* site and the 10 following nucleotides of 3' end of the ribulose reductase promoter were needed as overlap for fusing the fragment of PCR1 with the fragment of PCR2 described below. Genomic DNA of *P. ohmeri* ATCC 20209 was used as template.

For this purpose, a freshly streaked out *P. ohmeri* colony was resuspended in 30 µl of 0.2% SDS and heated for 4 min at 95°C. After full speed centrifugation, 0.5 µl of the
20 supernatant was used for PCR.

The template was amplified in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

25 The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 25 cycles with 10 sec at 98°C / 20 sec at 50°C / 15 sec at 72°C, and a final extension step of 10 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

30 In a second PCR (PCR2), a 911 bp long fragment of the tagatose-3-epimerase of *Pseudomonas cichorii* ST24 flanked by *Ascl* and *SphI* sites (underlined in primer sequence) was amplified using:

- primer EV2962:

AACTAAAAGGGCGCGCCAAATGAACAAAGTTGGCATG (SEQ ID No 14)

and

35 - primer EV2963:

TTCTCTTCGAGAGGCATGCTCAGGCCAGCTTGTCACG (SEQ ID No 15).

The 5' end of primer EV2962 contains a 9 nucleotide long fragment representing the 3' of the ribulose reductase promoter.

This fragment together with the 8 nucleotides of the *Ascl* site and the following 12 nucleotides of the tagatose-3-epimerase open reading frame, is used for the overlap PCR to fuse the PCR2 product to the previously described PCR1 product.

Additionally, the 5' end of reverse primer EV2963 contains a 12 nucleotide long fragment representing the 5' end of the ribulose reductase terminator of *P. ohmeri*.

This fragment, together with the 6 nucleotides of the *SphI* site and the following 12 nucleotides of the 3' end of the tagatose-3-epimerase open reading frame, is needed as overlap for fusing PCR2 with the PCR fragment of PCR3 described below.

As template 25 ng of vector 12AAMCJP (**Figure 5**) (GeneArt, Regensburg, Germany) containing a synthesized copy of the tagatose-3-epimerase gene of *Pseudomonas cichorii* ST24 was used (nucleotide 719 to 1591 of AB000361.1, from <http://www.ncbi.nlm.nih.gov/nuccore/AB000361>) – SEQ ID No: 11.

The template was amplified in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 25 cycles with 10 sec at 98°C / 20 sec at 48°C / 30 sec at 72°C, and a final extension step of 10 minutes at 72°C.

In a third PCR (PCR3), a 380 bp long fragment of the ribulose reductase terminator of *P. ohmeri* flanked by *SphI* and *SacII* sites (underlined in primer sequence) was amplified using:

- primer EV2964
AAGCTGGCCTGAGCATGCTCTCGAAGAGAATCTAG (SEQ ID No 16)
and
- primer EV2965
GTTCCGCGGAGAATGACACGGCCGAC (SEQ ID No 17)

The 5' end of primer EV2964 contains a 12 nucleotide long fragment of the 3' end of the tagatose-3-epimerase open reading frame that, together with the 6 nucleotides of the *SphI* site and the following 12 nucleotides of the ribulose reductase terminator of *P. ohmeri* is used for the fusion of PCR3 to the previously described PCR2.

Genomic DNA of *P. ohmeri* ATCC 20209 was used as template. After full speed centrifugation, 0.5 µl of the supernatant was used in PCR. For this purpose, a freshly streaked out *P. ohmeri* colony was resuspended in 30 µl of 0.2% SDS and heated for 4 min at 95°C.

The template was amplified in a reaction mix consisting of 200 μ M of each dNTP, 0.5 μ M of each primer and 0.02 U/ μ l of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

5 The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 25 cycles with 10 sec at 98°C / 20 sec at 50°C / 15 sec at 72°C, and a final extension step of 10 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research
10 Corporation, Irvine, California).

Fusion of the three individual PCR fragments was performed as follows: 50 ng of each gel purified product of PCR1 and PCR2 was used as template in a PCR reaction with EV2960 and EV2963.

15 A 30 nucleotide long homologous segment in the two fragments, resulting from the primer design described above, was used as overlap in the fusion reaction.

This way, a 1.4 kb long fragment, consisting of a ribulose reductase promoter of *P. ohmeri* flanked by *SpeI* and *Ascl* sites was fused to the open reading frame of the tagatose-3-epimerase of *Pseudomonas cichorii* ST24.

20 The templates were amplified in a reaction mix consisting of 200 μ M of each dNTP, 0.5 μ M of each primer and 0.02 U/ μ l of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

25 The PCR was performed with an initial denaturation step of 30 sec at 98°C, followed by 30 cycles with 10 sec at 98°C / 20 sec at 62°C / 45 sec at 72°C, and a final extension step of 10 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

The purified fragment was fused in a second overlap PCR to the product of PCR3. 40 ng of each fragment was used as template and amplified with EV2960 and EV2965.

30 A 30 nucleotide long homologous segment in the two fragments, resulting from the primer design described above, was used as overlap in the fusion.

This way, a 1.8 kb long fragment, consisting of a ribulose reductase promoter of *P. ohmeri* flanked by *SpeI* and *Ascl* and the open reading frame of the tagatose-3-epimerase of *Pseudomonas cichorii* ST24 flanked by *Ascl* and *SphI* sites was fused to the ribulose reductase terminator of *P. ohmeri*.

35 The templates were amplified in a reaction mix consisting of 200 μ M of each dNTP, 0.5 μ M of each primer and 0.02 U/ μ l of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 20 sec at 65°C / 55 sec at 72°C, and a final extension step of 10 minutes at 72°C. The PCR product was separated on an agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

The final PCR product consisting of a 1.7 kb long fragment of the tagatose-3-epimerase of *Pseudomonas cichorii* ST24 flanked by a ribulose reductase promoter and terminator was digested with restriction enzymes *SpeI* and *SacII* (New England Biolabs, Ipswich, Massachusetts), gel purified and ligated overnight at 16°C with a 9.8 kb long isolated *SpeI/SacII* fragment of a lig7.78 vector backbone using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 6**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California). The purified plasmid DNA was used for further characterization by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The newly cloned expression plasmid pEVE2523 (**Figure 7**) is a shuttle *E. coli* - *P. ohmeri* vector consisting of a bacterial (*E. coli*) origin of replication and an ampicillin resistance gene, the yeast (*P. ohmeri*) autonomous replication sequence, and the *poURA3* (*P. ohmeri*) gene for selection in yeast.

Moreover, it contains an exchangeable *P. ohmeri* ribulose reductase promoter element (via *SpeI* and *Ascl* restriction) and terminator element (via *SphI* and *SacII*) flanking an open reading frame of the tagatose-3-epimerase of *Pseudomonas cichorii* (exchangeable via *Ascl* and *SphI* restriction).

Example 7. Construction of a *P. ohmeri* vector for heterologous gene expression using the *poLEU2* selection marker

For the construction of a second *P. ohmeri* expression vector, the expression cassette of plasmid pEVE2523 (**Figure 7**) described previously in **Example 6** was cloned into a vector containing the *P. ohmeri poLEU2* selection marker (**Figure 6**).

A blunted 1.7 kb fragment of vector pEVE2523 (**Figure 7**) cut with *SpeI* and *SacII* (New England Biolabs, Ipswich, Massachusetts) was used as insert. Blunting was performed with the Blunting Enzyme Mix (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C.

The vector backbone was obtained from a *poARS* vector (plig3 – FR 2772788) linearized with *SalI* (New England Biolabs, Ipswich, Massachusetts), blunted and

dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). Gel purified insert and vector backbone using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) were ligated for 1 h at RT using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts).

5 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zypzy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and used for further characterization by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

10 The new cloned expression plasmids pEVE2560 (**Figure 8**) is a shuttle *E. coli* - *P. ohmeri* vector containing a bacterial (*E. coli*) origin of replication and an ampicillin resistance gene, the yeast (*P. ohmeri*) autonomous replication sequence, and the poLEU2 (*P. ohmeri*) gene for selection in yeast.

Moreover, the open reading frame of the tagatose-3-epimerase of *Pseudomonas cichorii* flanked by a *P. ohmeri* ribulose reductase promoter and terminator is exchangeable via *Ascl* and *SphI* restriction.

Example 8. Construction of a *P. ohmeri* vector for overexpression of *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase

20 A *P. ohmeri* vector for overexpression of *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase was constructed.

For cloning into the expression vector, the DNA fragment encoding the *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase was released from vector 13AAYSYP (**Figure 3**) by cutting with *Ascl* and *SphI* restriction enzymes (New England Biolabs, Ipswich, Massachusetts).

25 The 803 bp fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and ligated for 2 h at room temperature to the 9.8 kb *Ascl/SphI*-digested and gel-purified vector backbone of pEVE2523 (**Figure 7**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 9**).

30 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zypzy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

35 The resulting plasmid pEVE3284 (**Figure 10**) contains the codon-optimized NADPH-specific xylitol dehydrogenase of *Gluconobacter oxydans* flanked by a ribulose reductase promoter and terminator of *P. ohmeri* and the poURA3 selection marker.

Example 9. Construction of a *P. ohmeri* vector for overexpression of *Pichia stipitis* NADPH-specific xylitol dehydrogenase

For sub-cloning into the expression vector, the DNA fragment encoding the NADPH-specific xylitol dehydrogenase from *Pichia stipitis* had to be flanked with *Ascl* and *SphI* restriction sites.

For this purpose :

- EV3101 primer

AAGGCGCGCCAAA ATGACTGCTAACCCTTCC (SEQ ID No 18) containing an *Ascl* site (underlined) and

- EV3102 primer

GAGCATGCTTACTCAGGGCCGTCAATG (SEQ ID No 19) containing a *SphI* (underlined)

were used in a PCR reaction with 30 ng of vector 12AALQTP (**Figure 2b**) as template.

The template was amplified in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 25 cycles with 10 sec at 98°C / 20 sec at 55°C / 30 sec at 72°C, and a final extension step of 10 minutes at 72°C.

The 1.1 kb PCR product was separated on a 1% agarose gel, extracted, purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and restriction digested with *Ascl* and *SphI* (New England Biolabs, Ipswich, Massachusetts). After column purification with the DNA Clean & Concentrator™-5 Kit (Zymo Research Corporation, Irvine, California), it was ligated for 2 h at room temperature to the 10.6 kb *Ascl/SphI*-digested and gel-purified vector backbone of pEVE2523 (**Figure 7**) and the 11.8 kb *Ascl/SphI*-digested and gel-purified vector backbone of pEVE2560 (**Figure 8**) respectively, using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 11**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmids pEVE2562 and pEVE2564 (**Figure 12**) contain the codon optimised NADPH-specific xylitol dehydrogenase of *Pichia stipitis* flanked by a ribulose reductase promoter and terminator of *P. ohmeri* and either the *poURA3* or *poLEU2* selection marker, respectively.

Example 10. Construction of a *P. ohmeri* vector for overexpression of *Pichia stipitis* NADH-specific xylitol dehydrogenase

For sub-cloning into the expression vector, the DNA fragment encoding the NADH-specific xylitol dehydrogenase from *Pichia stipitis* had to be flanked with *Ascl* and *SphI* restriction sites.

For this purpose:

- EV3101 (AAGGCGCGCCAAA ATGACTGCTAACCCTTCC) (SEQ ID No 18) containing an *Ascl* site (underlined) and
- EV3102 (GAGCATGCTTACTCAGGGCCGTCAATG) (SEQ ID No 19) containing a *SphI* (underlined)

were used in a PCR reaction with 30 ng of vector lig7.78 (**Figure 2a**) as template.

The template was amplified in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 25 cycles with 10 sec at 98°C / 20 sec at 55°C / 30 sec at 72°C, and a final extension step of 10 minutes at 72°C.

The 1.1 kb PCR product was separated on a 1% agarose gel, extracted, purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and restriction digested with *Ascl* and *SphI* (New England Biolabs, Ipswich, Massachusetts).

After column purification with the DNA Clean & Concentrator™-5 Kit (Zymo Research Corporation, Irvine, California), it was ligated for 2 h at room temperature to the 10.5 kb *Ascl/SphI*-digested and gel-purified vector backbone of pEVE2560 (**Figure 8**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 13**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2563 (**Figure 14**) contains the codon optimised NADH-specific xylitol dehydrogenase of *Pichia stipitis* flanked by a ribulose reductase promoter and terminator of *P. ohmeri* and the poLEU2 selection marker.

Example 11. Construction of *P. ohmeri* vectors for overexpression of *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase

A *P. ohmeri* vector for overexpression of *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase was constructed.

For cloning into the expression vector, the DNA fragment encoding the codon-optimised *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase was released from vector 12ABYWMP (**Figure 1**) by cutting with *Ascl* and *SphI* restriction enzymes (New England Biolabs, Ipswich, Massachusetts).

5 The 1.4 kb fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and ligated for 2 h at room temperature to the 9.8 kb *Ascl/SphI*-digested and gel-purified vector backbone of pEVE2523 (**Figure 7**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 15**).

10 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2839 (**Figure 16**) contains the codon-optimised *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase flanked by a ribulose reductase promoter and terminator of *P. ohmeri* and the *poURA3* selection marker.

In addition to the *P. ohmeri* ribulose reductase promoter, the NAD⁺-specific D-arabitol 4-oxidoreductase from *E. coli* was also cloned under the control of the *P. ohmeri* phosphoglycerate kinase (*poPGK1*) promoter and transketolase (*poTKL*) terminator.

20 Cloning was performed in two consecutive steps, by first replacing the ribulose reductase promoter by the *poPGK1* promoter, followed by an exchange of the ribulose reductase terminator for the *poTKL* terminator.

A 611 bp long fragment of the *P. ohmeri poPGK1* promoter was amplified from genomic DNA of *P. ohmeri* using:

- 25 - primer EV3177 (GAAGACTAGTTCACGTGATCTC) (SEQ ID No 20) containing a *SpeI* site (underlined) and
- primer EV3178 (CACTGCGCGCCTTTTGTGTGGTGGTGTCC) (SEQ ID No 21), containing an *Ascl* site (underlined).

30 The genomic DNA template was prepared by resuspending a freshly streaked out *P. ohmeri* colony in 30 µl of 0.2% SDS and heating for 4 min at 95°C. After full speed centrifugation, 0.5 µl of the supernatant was used for PCR.

Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

35 The PCR was accomplished with an initial denaturation step of 2 min at 96°C followed by 25 cycles with 10 sec at 96°C / 10 sec at 58°C / 30 sec at 72°C, and a final extension step of 2 minutes at 72°C.

The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

The amplified 610 bp long poPGK1 promoter fragment was restriction digested with *SpeI* and *Ascl* (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 11.5 kb *SpeI/Ascl*-digested and gel-purified vector backbone of pEVE2839 (**Figure 16**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 17**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE3102 (**Figure 18**) contains the codon-optimised *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase flanked by a phosphoglycerate kinase (poPGK1) promoter and ribulose reductase terminator of *P. ohmeri* and the po*URA3* selection marker.

In the next step the ribulose reductase terminator of pEVE3102 was exchanged for the tranketolase (poTKL) terminator of *P. ohmeri*.

A 213 bp long fragment of the *P. ohmeri* poTKL terminator was amplified from genomic DNA of *P. ohmeri* using:

- primer EV3817 (TAGCAGCATGCATAGGTTAGTGAATGAGGTATG) (SEQ ID No 22) containing a *SphI* site (underlined) and
- primer EV3818 (TAGGTCCGCGGGAGCTTCGTTAAAGGGC) (SEQ ID No 23) containing a *SacII* site (underlined).

The genomic DNA template was prepared as described above.

Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer.

The PCR was accomplished with an initial denaturation step of 2 min at 96°C followed by 25 cycles with 10 sec at 96°C / 10 sec at 57°C / 30 sec at 72°C, and a final extension step of 2 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

The amplified 213 bp long poTKL terminator fragment was restriction digested with *SphI* and *SacII* (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 11.5 kb *SphI/SacII*-digested and gel-purified vector backbone of pEVE3102 (**Figure 18**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 17**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

5 The resulting plasmid pEVE3123 (**Figure 19**) contains the codon-optimised *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase flanked by a phosphoglycerate kinase (poPGK1) promoter and a transketolase (poTKL) terminator of *P. ohmeri* and the poURA3 selection marker.

10 In order to be able to express the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* from a plasmid using another selection, the poURA3 marker of pEVE3123 was exchanged for the poLEU2 marker.

For this purpose the poURA3 marker was released from vector pEVE3123 (**Figure 19**) by restriction digestion with *PsiI* and *AfeI* (New England Biolabs, Ipswich, Massachusetts).

15 The 9.1 kb vector backbone was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California), blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C and dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts).

20 As insert, a 3 kb blunted and gel-purified fragment of the poLEU2 marker released from vector pEVE2560 (**Figure 8**) by *Asel* and *AfeI* restriction digestion was used. Ligation of the fragments was performed for 2 h at room temperature using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 20**).

25 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

30 The resulting plasmid pEVE3157 (**Figure 21**) contains the codon-optimised *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase flanked by a phosphoglycerate kinase (poPGK1) promoter and a transketolase (poTKL) terminator of *P. ohmeri* and the poLEU2 selection marker.

Example 12. Expression of the plasmidic *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase and of the plasmidic *Pichia stipitis* NADPH-specific xylitol dehydrogenase gene in *Pichia ohmeri* strain ATCC 20209

For the biosynthetic conversion of arabitol into xylitol, the simultaneous expression
 5 of the NAD⁺-specific *E. coli* D-arabitol 4-oxidoreductase and the NADP-specific xylitol dehydrogenase of *P. stipitis* is necessary.

The first enzyme leads to the formation of xylulose and the second ones convert xylulose into xylitol.

P. ohmeri strain SRLU (*MAT^h leu2 ura3*) derived from ATCC 20209 and
 10 auxotrophic for leucine and uracil (Piredda and Gaillardin, 1994, *supra*) was used as host for the construction of a yeast strains secreting xylitol by transformation with plasmids:

- pEVE2839 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*) and
- pEVE2564 (NADPH-specific xylitol dehydrogenase of *P. stipitis*)

leading to strain EYS2755

15 Additionally, as a control (following the teaching of WO 94/10325) a strain expressing the NADH-specific wild type xylitol dehydrogenase of *P. stipitis* was also constructed by transformation with plasmids:

- pEVE2839 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*) and
- pEVE2563 (NADH-specific xylitol dehydrogenase of *P. stipitis*)

20 into the SRLU host, leading to strain EYS2962.

As control, strains transformed with the single plasmids:

- pEVE2839 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*),
- pEVE2563 (NADH-specific xylitol dehydrogenase of *P. stipitis*), and
- pEVE2564 (NADPH-specific xylitol dehydrogenase of *P. stipitis*)

25 leading to EYS2943, EYS2696 and EYS2697 respectively, were also generated.

Yeast transformation was carried out in essential by the spheroplasting method of Green et al. (Green E.D., Hieter, P., and Spencer F. A., chapter 5 in Genome Analysis: A Laboratory Manual, Vol. 3, Cloning Systems, Birren et al. (eds.), Cold Spring Harbor Press, New York, 1999) with the following modifications: Instead of Lyticase, Zymolyase 100T was
 30 used for generation of spheroplasts and the incubation with the enzyme was performed at 37°C until the OD of the cell suspension reached 20-30% of the original OD before Zymolyase treatment.

Briefly, *P. ohmeri* cells were grown overnight at 30°C in YPD medium (Yeast extract 1% (w/v), Peptone 2% (w/v), Dextrose 2% (w/v)) to a final OD₆₀₀ of 3-5.

35 200 OD₆₀₀ units were harvested by centrifugation, washed once with water and 1M sorbitol, and resuspended in SCE buffer (1 M sorbitol, 100 mM citric acid trisodium salt dihydrate, 10 mM EDTA) to a final concentration of 70 ODs/ml.

DTT and Zymolase (LuBio Science, Luzern, Switzerland) were added to a final concentration of 10 mM and 0.5 U/OD, respectively and the mixture incubated at 37°C with slow shaking.

5 The cell wall digestion was followed by measuring the optical density of the solution diluted in water. When this value dropped to 80% of the original, the digestion was terminated by careful centrifugation and washing with 1 M sorbitol and STC buffer (0.98 M sorbitol, 10 mM Tris pH 7.5, 10 mM CaCl₂).

10 Spheroplasts were carefully resuspended in STC buffer containing 50 µg/ml calf-thymus DNA (Calbiochem/VWR, Dietikon, Switzerland) to a final concentration of 200 OD/ml. Aliquots of 100 µl were mixed with 100-200 ng of plasmid DNA and incubated for 10 min at room temperature.

15 1 ml PEG solution (19.6% PEG 8000 w/v, 10 mM Tris pH 7.5, 10 mM CaCl₂) was added to the suspension, incubated for 10 minutes and pelleted. Spheroplasts were regenerated at 30°C for 1-2 h in 1 ml of a 1 M sorbitol solution containing 25% YPD and 7 mM CaCl₂.

20 To the regenerated cells 7 ml of 50°C warm top agar (0.67% yeast nitrogen base w/o amino acids, 0.13% drop-out powder without leucine/uracil/histidine/tryptophan/methionine, 0.086% of required missing amino acid, 2% glucose, 1 M sorbitol, pH5.8 and 2.5% Noble agar) was added and the mixture was poured evenly onto pre-warmed, sorbitol containing selective plates (0.67% yeast nitrogen base w/o amino acids, 0.13% drop-out powder without leucine/uracil/histidine/tryptophan/methionine, 0.086% of required missing amino acid, 2% glucose, 1 M sorbitol, pH5.8).

Plates were incubated for 3-5 days at 30°C. Transformants were reselected on the appropriate selective plates.

25 Each generated strain was tested in triplicates for arabitol, xylitol and ribitol production.

30 For this purpose clones were first grown at 30°C overnight in seed media (0.67% yeast nitrogen base without amino acids; 0.13% drop-out powder without leucine/uracil/histidine/tryptophan/methionine; 0.086% of required missing amino acid; 5% glucose; pH5.7).

Out of this overnight culture a main culture in production media (0.67% yeast nitrogen base without amino acids; 0.13% drop-out powder without leucine/uracil/histidine/tryptophan/methionine; 0.086% of required missing amino acid; 15% glucose; pH5.7) at a starting OD₆₀₀ of 0.2 was inoculated.

35 This culture was grown at 37°C for 48 hours and the arabitol, xylitol and ribitol concentrations of the supernatants were determined by HPLC/MS using a Aminex® HPX-87 column (Bio-Rad, Hercules, California) and a Waters® TQ-Detector (Acquity® UPLC linked

to a triple quadrupol detector, Waters, Milford, Massachusetts) and isocratic conditions with 100% water as mobile phase.

Polyol titers of all tested strains are depicted in **Table 8**.

5 **Table 8**

Polyol production of *P. ohmeri* SRLU strains transformed with NADH- and NADPH-specific xylitol dehydrogenase of *P. stipitis* and/or with NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* (average of triplicates)

Strain	Arabitol (g/L)	Xylitol (g/L)	Ribitol (g/L)
SRLU	32.9±2.4	nd	nd
EYS2943 [pEVE2839]	26.4±2.8	2.3±0.1	0.7±1.2
EYS2696 [pEVE2563]	36.0±2.7	nd.	0.8±0.1
EYS2697 [pEVE2564]	31.1±1.6	nd	6.3±0.1
EYS2962 [pEVE2839/pEVE2563]	29.6±0.8	7.0±0.3	2.3±0.1
EYS2755 [pEVE2839/pEVE2564]	16.4±2.2	19.9±0.8	10.9±0.4

nd – not detected

Use of the NADPH-specific xylitol dehydrogenase of *P. stipitis* leads to a significant increase in xylitol titers, as compared to the wild type NADH-specific enzyme.

10

Example 13. Expression of the plasmidic *Gluconobacter oxydans* NADPH-specific xylitol dehydrogenase gene in *Pichia ohmeri*

In addition to a xylitol producing strain using the NADP-specific xylitol dehydrogenase of *P. stipitis* a second strain expressing the NADP-specific xylitol dehydrogenase of *G. oxydans* was engineered.

15

P. ohmeri strain SRLU (*MATh⁻ leu2 ura3*) derived from ATCC 20209 and auxotrophic for leucine and uracil (Piredda and Gaillardin, 1994, *supra*) was used as host for the construction of a yeast strains secreting xylitol by transformation with plasmids pEVE3157 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*) and pEVE3284 (NADPH-specific xylitol dehydrogenase of *G. oxydans*) leading to strain EYS3324.

20

As control, strains transformed with the single plasmids:

- pEVE3157 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*) and
- pEVE3284 (NADH-specific xylitol dehydrogenase of *G. oxydans*), leading to EYS3067 and EYS3323 respectively, were also generated.

25

The *E. coli* D-arabitol 4-oxidoreductase used for the construction of the above strains is controlled by poPGK1 promoter in contrast to the poRR promoter used in strains expressing the xylitol dehydrogenase of *P. stipitis*.

However, to exclude a promoter influence and therefore, to be able to compare polyol levels in strains expressing the xylitol dehydrogenase from *G. oxydans* with those expressing the corresponding enzyme from *P. stipitis*, an additional strain has been generated.

- 5 This strain EYS2963 was obtained by transforming the SRLU host with
- pEVE3123 (NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*) and
 - pEVE2564 (NADPH-specific xylitol dehydrogenase of *P. stipitis*).

Yeast transformation was carried out as described in **Example 12**. Each generated strain was tested in triplicates for arabitol, xylitol and ribitol production as described in

10 **Example 12.**

Polyol titers of all tested strains are depicted in **Table 9**.

Table 9

Polyol production of <i>P. ohmeri</i> SRLU strains transformed with NADPH-specific xylitol dehydrogenase of <i>G. oxydans</i> and/or with NAD ⁺ -specific D-arabitol 4-oxidoreductase of <i>E. coli</i> (average of triplicates)			
<i>Strain</i>	<i>Arabitol (g/L)</i>	<i>Xylitol (g/L)</i>	<i>Ribitol (g/L)</i>
SRLU	32.9±2.4	nd	nd
EYS3067 [pEVE3157]	29.0±3.8	1.5±0.3	1.8±0.4
EYS3323 [pEVE3284]	32.8±0.6	nd	nd
EYS3324 [pEVE3157/pEVE3284]	26.3±1.7	21.1±1.1	1.2±0.1
EYS2963 [pEVE3123/pEVE2564]	27.3±2.5	17.7±1.7	13.9±0.7

nd – not detected

15

Xylitol titers in strains expressing the NADPH-specific xylitol dehydrogenase from *G. oxydans* (EYS3324) are similar to those of strains expressing the corresponding enzyme from *P. stipitis* (EYS2963). However, the *G. oxydans* enzyme leads to much lower ribitol titers, thus showing a higher substrate specificity towards xylulose.

20

Example 14. Generation of a mutant *P. ohmeri* strain with increased arabitol secretion

A higher arabitol producer mutant has been selected from an UV irradiated suspension of *P. ohmeri* ATCC 20209.

25

The UV-irradiation system (Vilber Lourmat, France), was equipped with a microprocessor-controlled RMX-3 W radiometer. *P.ohmeri* was grown on YPD agar (Dextrose 20g/L) at 37°C overnight.

5 A suspension was prepared to reach 10^6 cfu/mL ($OD_{620}=0.4$) and 5 mL were put into a sterile Petri dish. The suspension was irradiated after removing the cover from the dish. The UV wavelength was 254 nm and the irradiation energy was $1.8 \cdot 10^{-2}$ J/cm². 90% of mortality of the yeast cells was obtained. After stopping the irradiation and replacing the lid on the dish, the suspension was transferred into a sterile tube located into an iced bath.

10 20 mL of YPD liquid medium was inoculated with the mutated suspension and was incubated for 12 hours at 37°C, 250 rpm.

After incubation the mutated culture was diluted with sterile 40 % glycerol (V/V). Aliquots were distributed into 5 mL vials and frozen at -80°C.

The screening was based on the osmophilic property of *Pichia ohmeri* which is able to grow on very high concentrations of Dextrose (up to 600 g/L).

15 Our goal was to select mutants able to grow faster than the mother strain on YPD agar containing Dextrose 600 g/L or 700 g/L.

Defrosted aliquots were spread on YPD₆₀₀ and YPD₇₀₀ and the first appearing colonies were selected and tested for the production of arabitol in shake flasks.

20 The subculture and production medium were made of glucose 50g/L or 100 g/L respectively, yeast extract 3g/L, MgSO₄ 1 g/L and KH₂PO₄ 2g/L, pH 5.7. The subculture (10 mL in a 100 mL flask) was incubated for 24 h at 37°C, 250 rpm. The production (40 mL in a 500mL flask) was inoculated by 5 mL of subculture and incubated for 64 hours at 37°C, 250rpm.

	Glucose g/L 64h	Arabitol g/L 64h
<i>P. ohmeri</i> ATCC 20209	6.0	52.7
<i>P. ohmeri</i> CNCM I-4605	0	58.6

25

The mutant *P. ohmeri* strain was selected for its faster consumption of glucose and its higher production of arabitol and was deposited in France on March 7, 2012, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 PARIS Cedex 15, 30 under number I-4605.

Example 15. Construction of a *LEU2* deletion plasmid

In order to be able to use the newly generated CNCM I-4605 strain for plasmid selection and gene integrations, a plasmid for the deletion of the *LEU2* open reading frame was constructed.

5 In a first step, a general integration vector that can be used in *P. ohmeri* was adapted from the *S. cerevisiae* CRE/loxP system. The vector backbone was isolated from pUG73 (Gueldener et al., 2002, *Nucleic Acid Res*, **30**, e23) by restriction cutting with *Pst*I and *EcoRV* enzymes (New England Biolabs, Ipswich, Massachusetts).

10 As insert served a PCR fragment containing a *LEU2* selection marker of *P. ohmeri* flanked by loxP sites, generated with primer pair:

- EV3043

(CACTGGCGCGCCCACTGCATGCGTCGACAACCCTTAATATAACTTCGTA
TAATGTATGCTATACGAAGTTATTAGGTCTAGACACATCGTGGATCCAAG
CTATCAACGAGAGAGTC) (SEQ ID No 24) and

15 - EV3044

(AGTGGCTAGCAGTGCCATGGCCTAATAACTTCGTATAGCATACATTATAC
GAAGTTATATTAAGGGTTCTCGAGACGCGTCATCTAGCATCTCATCTACCA
ACTC) (SEQ ID No 25) and

- poARS (plig3 - FR 2772788 - see **Figure 6**) as template.

20 The forward primer EV3043 contains an *Ascl* (underlined) site preceding a *Sph*I site (underlined), followed by a 48 bp long loxP fragment (bold) and a *Dra*III site (underlined). The 3' end of EV3043 contains an additional a 25 bp long fragment for amplification of the *P. ohmeri LEU2* gene. The reversed primer EV3044 on the other hand, contains a *Nhe*I (underlined) site preceding a *Nco*I site (underlined), followed by a 48 bp long loxP fragment
25 (bold) and a *Mlu*I site (underlined). The 3' end of EV3044 contains an additional a 25 bp long fragment for amplification of the *P. ohmeri LEU2* gene. The template was amplified in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was performed with an initial denaturation step of 30 sec at 98°C followed by 30 cycles
30 with 10 sec at 98°C / 10 sec at 65°C / 50 sec at 72°C, and a final extension step of 7 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

35 The amplified fragment was flanked by a *Pst*I and *EcoRV* site in a second PCR reaction for further subcloning. Amplification was performed with :

- primer EV3056 (CACTCTGCAGCACTGGCGCGCCCACTGCAT) (SEQ ID No 26) containing the *Pst*I site (underlined) and

- primer EV3057 (CACTGATATCAGTGGCTAGCAGTGCCATGG) (SEQ ID No 27) containing the *EcoRV* site (underlined)

in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 45 sec at 72°C, and a final extension step of 7 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

The amplified 2.5 kb *LEU2* marker was restriction digested with *Pst*I and *EcoRV* enzymes (New England Biolabs, Ipswich, Massachusetts), gel-purified with the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and ligated for 2 h at room temperature to the 2.4 kb *Pst*I/*EcoRV* (New England Biolabs, Ipswich, Massachusetts), gel-purified (Zymoclean™ Gel DNA Recovery Kit - Zymo Research Corporation, Irvine, California) backbone of vector pUG73 (Gueldener et al., 2002 *Nucleic Acid Res*, **30**, e23) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) – (Figure 22). After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2787 (Figure 23) contains the *P. ohmeri LEU2* selection marker under the control of the endogenous promoter and terminator, flanked by two loxP sites. Additionally, a *Ascl* and *Sph*I site have been introduced upstream of the first loxP and a *Nhe*I and *Nco*I site downstream of the second loxP, in order to help in cloning of regions homologous to the integration sites in the genome.

The *LEU2* marker of the integration vector was then replaced by the *nat1* resistance gene of *Streptomyces noursei*, in a second cloning step, since a deletion of the endogenous *LEU2* open reading frame was aimed.

A DNA fragment encoding the *nat1* gene of *Streptomyces noursei* was chemically synthesized by GeneArt® Gene Synthesis (Life Technologies, Regensburg, Germany) according to the submitted sequence of SEQ ID No 28.

Nucleotides 204 to 776 of sequence S60706.1 (obtained from <http://www.ncbi.nlm.nih.gov/nuccore/S60706.1>) coding for the *nat1* gene were used as template and subjected to codon optimization for use in *P. ohmeri* ATCC 20209 according to Table 7 (above), using the Optimizer program obtained from <http://genomes.urv.es/OPTIMIZER/>.

At the 5' and 3' ends of the resulting sequence, nucleotides encoding for the recognition sites of the restriction enzymes *Ascl* (GGCGCGCC) and *SphI* (GCATGC) respectively, were manually added in the text file, in order to facilitate further cloning. Additionally, an adenosine triplet was included in front of the start ATG to account for an adenosine at the -3 position in the Kozak-like sequence of yeasts.

The final sequence (SEQ ID No 28) was then submitted for synthesis to GeneArt (Regensburg, Germany). The synthesized DNA fragment encoding the *nat1* gene was delivered as 5 µg lyophilized plasmid DNA in a pMA-T derived vector (12ABTV4P, **Figure 24**).

For the cloning of the *nat1* gene a vector containing a ribulose reductase (poRR) promoter and terminator was used. The terminator was exchanged by an orotidine-5'-phosphate decarboxylase (poURA3) terminator and the *nat1* gene was introduced between the promoter and terminator sequences.

For this purpose, the orotidine-5'-phosphate decarboxylase (poURA3) terminator was generated by PCR with :

- primer EV3393 (CAAGCATGCGGGGAATGATAAGAGACTTTG) (SEQ ID No 29) containing a *SphI* site (underlined) and
- primer EV3394 (GGACCGCGGAAAGGTGAGGAAGTATATGAAC) (SEQ ID No 30) containing a *SacII* site (underlined) and
- pEVE2523 (**Figure 7**) as template.

Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 59°C / 10 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The 239 bp poURA3 terminator was restriction digested with *SphI* and *SacII* enzymes (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 11 kb vector backbone of pEVE2681 linearized with *SphI* and *SacII* restriction enzymes (New England Biolabs, Ipswich, Massachusetts) and gel-purified with Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts). After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

In a second cloning step, the *nat1* gene was released from 12ABTV4P (**Figure 24**) by restriction cutting with *SphI* and *Ascl* enzymes (New England Biolabs, Ipswich, Massachusetts). Additionally, a blunting of the *SphI* site with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C was performed in between the *SphI* and *Ascl* digestion. The 587 bp gel-purified fragment (Zymoclean™ Gel DNA Recovery Kit - Zymo Research Corporation, Irvine, California) was then ligated to the gel-purified 10.5 kb vector backbone of the vector described above cut with *SphI* and *Ascl* restriction enzymes (New England Biolabs, Ipswich, Massachusetts).

Also the *SphI* site of the vector was blunted for 15 min at room temperature with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts), followed by a heat inactivation step of 10 min at 70°C before the digestion with *Ascl* was performed. Additionally, the vector was dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). The ligation was performed for 2 h at room temperature using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2798 (**Figure 25**) contains the *nat1* drug resistance marker flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and an orotidine-5'-phosphate decarboxylase (poURA3) terminator.

The *nat1* expression cassette was used to replace the *P. ohmeri* *LEU2* selection marker in the integrative vector. In order to facilitate further cloning the *nat1* cassette had to be flanked with *XbaI* (underlined in primer EV3643) and *MluI* (underlined in primer EV3644) sites by PCR with:

- primer EV3643
(CACTTCTAGACACTATCGATGGATCCGTAGAAATCTTG) (SEQ ID No 31)
and

- primer EV3644 (CACTACGCGTAAAGGTGAGGAAGTATATG) (SEQ ID No 32).

Primer EV3643 contains an additional *ClaI* site (dotted line) following the *XbaI* site. pEVE2798 served as template (**Figure 25**).

Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 54°C / 25 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product

was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The 1.3 kb *nat1* expression cassette was restriction digested with *MluI* and *XbaI* enzymes (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 2.6 kb vector backbone of pEVE2787 (**Figure 23**) linearized with *MluI* and *XbaI* enzymes (New England Biolabs, Ipswich, Massachusetts) and gel-purified with the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 26**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2852 (**Figure 27**) contains the *nat1* selection marker under the control of the ribulose reductase (poRR) promoter and orotidine-5'-phosphate decarboxylase (poURA3) terminator and flanked by two loxP sites.

The integration plasmid does not contain any *P. ohmeri* homologous fragments needed for site specific integration into the genome, so far. This sites were attached in the next steps.

The 5' homologous region upstream of the *LEU2* open reading frame was amplified from 50 ng poARS vector (**Figure 6**) with :

- primer EV3548 (CACTCTGCAGGATCCAAGCTATCAACGAGA) (SEQ ID No 33) containing a *PstI* site (underlined) and
- primer EV3549 (CACTGCATGCGTTGCGGAAAAACAGCC) (SEQ ID No 34) containing a *SphI* site (underlined).

The PCR was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The amplification was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 61°C / 15 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The 567 bp fragment was restriction digested with *PstI* and *SphI* enzymes (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 3.9 kb vector backbone of pEVE2852 (**Figure 27**) linearized with *PstI* and *SphI* restriction enzymes (New England Biolabs, Ipswich, Massachusetts) and gel-purified with Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 29**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

5 The resulting plasmid pEVE2855 (**Figure 28**) contains a fragment homologous to the 5' region upstream of the *LEU2* open reading frame and a *nat1* marker flanked by two loxP sites.

The 3' homologous region downstream of the *LEU2* open reading frame was amplified from 50 ng poARS vector (**Figure 6**) with:

- 10 - primer EV3550 (CACT CCATGG AGTAGGTATATAAAAATATAAGAG) (SEQ ID No 35) containing a *NcoI* site (underlined) and
- primer EV3551 (CACTGCTAGCGTGACAAACAGCAACTAG) (SEQ ID No 36) containing a *NheI* site (underlined).

15 The PCR was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The amplification was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 51°C / 25 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel

20 DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The 1.3 kb fragment was restriction digested with *NcoI* and *NheI* enzymes (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 4.4 kb vector backbone of pEVE2855 (**Figure 28**) linearized with *NcoI* and *NheI* restriction enzymes (New England Biolabs, Ipswich, Massachusetts) and gel-purified with Zymoclean™ Gel DNA Recovery Kit

25 (Zymo Research Corporation, Irvine, California) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 29**).

30 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting final *LEU2* deletion plasmid pEVE2864 (**Figure 30**) contains a fragment homologous to the 5' region upstream and a fragment homologous to the 3' region downstream of the *LEU2* open reading frame and a *nat1* marker flanked by two loxP sites.

Example 16. Generation of a mutant *P. ohmeri* strain auxotrophic for leucine

Since the generated *P. ohmeri* CNCM I-4605 strain did not display any auxotrophy so far, a *LEU2* open reading frame deletion was performed, so as to be able to use the *LEU2* selection marker for gene integrations.

5 For this purpose plasmid pEVE2864 (**Figure 30**) was restriction digested with *EcoRV* and *PstI* enzymes (New England Biolabs, Ipswich, Massachusetts) for 2.5 h at 37°C and the mixture used to transform the *Mut165* strain according to the procedure described in Example 12.

10 To the regenerated cells, 7 ml of 50°C warm top agar (1% yeast extract, 2% peptone, 2% glucose, 1 M sorbitol, pH 5.8 and 2.5% Noble agar) with 25 µg/ml natamycin was added and the mixture was poured evenly onto pre-warmed, sorbitol containing selection plates (1% yeast extract, 2% peptone, 2% glucose, 1 M sorbitol, pH 5.8 and 2% agar) with 25 µg/ml natamycin. Plates were incubated for 4 days at 30°C. Deletion of the *LEU2* open reading frame was verified by no growth on selective plates without leucine and
15 confirmed by colony PCR using:

- primer EV3393 (CAAGCATGCGGGAATGATAAGAGACTTTG) (SEQ ID No 29) and
- primer EV3795 (CAAGTCGTGGAGATTCTGC) (SEQ ID No 37).

20 The 1.6 kb fragment was amplified with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 51°C / 25 sec at 72°C, and a final extension step of 5 minutes at 72°C.

The resulting strain contains the full open reading frame deletion of the *LEU2* gene in a CNCM I-4605 background and was deposited in France on February 5, 2015, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism
25 Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 PARIS cedex 15, under number I- 4955.

Example 17. Construction of a double expression plasmids comprising the NADPH-specific xylitol dehydrogenase of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*

30 In order to be able to express the NADPH-specific xylitol dehydrogenase of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* in the mutant *P. ohmeri* strain only auxotrophic for leucine, construction of a double expression plasmid was required.

35 The expression cassette containing the NADPH-specific xylitol dehydrogenase of *P. stipitis* was released from pEVE2562 (**Figure 12**) by restriction cutting with *SpeI* and *SacII* enzymes (New England Biolabs, Ipswich, Massachusetts). The 1.9 kb fragment was gel-

purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C. The insert was then ligated for 2 h at room temperature to the 12.1 kb *SpeI*-linearized, blunted, dephosphorylated (1 h at 37°C using Antarctic phosphatase - New England Biolabs, Ipswich, Massachusetts) and gel-purified pEVE3157 backbone (**Figure 21**) containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 31**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE3318 (**Figure 32**) contains the double expression construct of the NADPH-specific xylitol dehydrogenase of *P. stipitis* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and ribulose reductase (poRR) terminator and the poLEU2 selection marker.

Example 18. Construction of integrative vectors for the expression of the *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *P. stipitis* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*

The NAD⁺-specific D-arabitol 4-oxidoreductase gene of *E. coli* and the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* should ultimately become an integral part of the *P. ohmeri* genome. Therefore, an integrative vector with a *LEU2* selection marker had to be constructed, by replacing the *nat1* selection marker of pEVE2852 and incorporating the double expression construct of arabitol oxidoreductase and xylitol dehydrogenase.

For this purpose, the *P. ohmeri* *LEU2* open reading frame, flanked by an *Ascl* and *SphI* sites, was generated by PCR with:

- primer EV3645 (CAAGGCGCGCCAAAATGTCTACCAAAACCATTAC) (SEQ ID No 38) and

- primer EV3646 (GGAGCATGCCTACTTTCCCTCAGCCAAG) (SEQ ID No 39).

Amplification was performed with 50 ng of poARS (**Figure 6**) template in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 57°C / 20 sec at 72°C, and a final extension step of 5

minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The amplified *LEU2* open reading frame was subsequently restriction digested with *Ascl* and *SphI* enzymes (New England Biolabs, Ipswich, Massachusetts).

5 Additionally, a blunting of the *SphI* site with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C was performed in between the *SphI* and *Ascl* digestion. The 1.1 kb gel-purified fragment was then ligated to the gel-purified 11 kb vector backbone of pEVE2811 cut with *SphI* and *Ascl* restriction enzymes (New England Biolabs, Ipswich, Massachusetts). Also the *SphI* site of the vector was blunted for 15 min at room temperature with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts), followed by a heat inactivation step of 10 min at 70°C before the digestion with *Ascl* was performed. Additionally, the vector was dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). The ligation of the *LEU2* open reading frame and the vector backbone was performed for 2 h at room temperature using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts).

20 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2862 (**Figure 33**) contains the *P. ohmeri LEU2* marker flanked by a *P. ohmeri* ribulose reductase (poRR) promoter and an orotidine-5'-phosphate decarboxylase (poURA3) terminator.

Subsequently, the *LEU2* marker was amplified by PCR using:

- 25 - primer EV3643 (CACTATCGATGGATCCGTAGAAATCTTG) (SEQ ID No 31) containing a *ClaI* site and
- primer EV3644 (CACTACGCGTAAAGGTGAGGAAGTATATG) (SEQ ID No 32) containing a *MluI* site (underline) and pEVE2862 (**Figure 33**) as template.

30 Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 54°C / 30 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The amplified 1.8 kb long *LEU2* fragment was restriction digested with *ClaI* and *MluI* enzymes (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 2.6 kb *ClaI* and *MluI*

35

(New England Biolabs, Ipswich, Massachusetts) restriction digested and gel-purified vector backbone of pEVE2852 (**Figure 27**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 34**).

5 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppey™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE2865 (**Figure 35**) contains the *P. ohmeri* LEU2 marker flanked by two loxP sites.

10 For cloning of the integration vector, pEVE2865 was restriction digested with *Sal*I enzyme (New England Biolabs, Ipswich, Massachusetts), blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C and dephosphorylated dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, 15 Ipswich, Massachusetts).

The 4.5 kb gel-purified fragment of the vector backbone was used for ligation. As insert served a double expression construct of the NADPH-specific xylitol dehydrogenase genes of *P.stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* released from pEVE3318 (**Figure 32**) by restriction cutting with *Nde*I and *Sac*II enzymes (New 20 England Biolabs, Ipswich, Massachusetts).

The 4.4 kb fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C, followed by an additional 25 gel purification. The vector backbone of pEVE2865 and the insert of pEVE3318 were ligated for 2 h at room temperature using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 34**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the 30 Zyppey™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE3387 (**Figure 36**) contains the double expression construct of the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator. As selection marker serves a *P. ohmeri* LEU2 gene flanked by two loxP sites.

35

Example 19. Construction of a first generation integrative *P. ohmeri* strain secreting xylitol into the media

The previously described vector was used to randomly integrate the NAD⁺-specific D-arabitol 4-oxidoreductase gene of *E. coli* and the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* into the genome of *P. ohmeri*.

For this purpose strain CNCM I-4955 (**Example 16**) auxotrophic for leucine was transformed with pEVE3387 (**Figure 36**) restriction digested with *NotI* (New England Biolabs, Ipswich, Massachusetts) for 3 h at 37°C according to the procedure described in **Example 12**. Transformants were selected on sorbitol plates without any leucine.

The resulting strain contains the NAD⁺-specific D-arabitol 4-oxidoreductase gene of *E. coli* and the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* randomly integrated into the *P. ohmeri* genome and was deposited in France on May 20, 2015, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 Cedex 15, under number I-4982.

Example 20. Construction of a double/triple expression plasmid comprising the NADPH- specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*

In order to be able to express the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* in the mutant *P. ohmeri* strain only auxotrophic for leucine, construction of a double expression plasmid was required.

The expression cassette containing the NADPH-specific xylitol dehydrogenase of *G. oxydans* was released from pEVE3284 (**Figure 10**) by restriction cutting with *SpeI* and *SacI* enzymes (New England Biolabs, Ipswich, Massachusetts). The 1.6 kb fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C. The vector backbone used consisted of the 12.1 kb *SpeI*-linearized (New England Biolabs, Ipswich, Massachusetts) and gel-purified (Zymoclean™ Gel DNA Recovery Kit - Zymo Research Corporation, Irvine, California) pEVE3157 backbone (**Figure 21**) containing the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*.

The backbone has additionally been blunted for 15 min at room temperature with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts), followed by heat inactivation of the enzymes for 10 min at 70°C and dephosphorylated for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). Ligation was

performed for 2 h at room temperature using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 37**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmids pEVE3322 and pEVE3324 (**Figure 38**) contain either the double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator or the triple expression construct of two NADPH-specific xylitol dehydrogenase genes of *G. oxydans* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator and the poLEU2 selection marker.

Example 21. Construction of integrative vectors for the expression of the *E. coli* NAD⁺-specific D-arabitol 4-oxidoreductase gene and the *G. oxydans* NADPH-specific xylitol dehydrogenase gene in *P. ohmeri*

Besides the integrative vector containing the NADPH-specific xylitol dehydrogenase of *P. stipitis* and the NAD⁺-specific D-arabitol 4-oxidoreductase gene of *E. coli* also plasmids containing the NADPH-specific xylitol dehydrogenase of *G. oxydans* were generated.

For this purpose, the double and triple expression cassettes containing either one or two NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* were released from pEVE3322 and pEVE3324 (**Figure 38**) respectively, by restriction cutting with *Nde*I and *Sac*II enzymes (New England Biolabs, Ipswich, Massachusetts).

The 4.1 kb and 5.7 kb fragments were gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C. As vector served the gel-purified (Zymoclean™ Gel DNA Recovery Kit - Zymo Research Corporation, Irvine, California), 5.7 kb *Sac*II-linearized pEVE2865 (**Figure 35**).

The vector backbone has additionally been blunted for 15 min at room temperature with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts), followed

by heat inactivation of the enzymes for 10 min at 70°C and dephosphorylation for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). Ligation of vector and insert was performed for 2 h at room temperature to using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 39**).

5 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

10 The resulting plasmids pEVE3390 and pEVE3392 (**Figure 40**) contain the double or triple expression constructs of either one or two NADPH-specific xylitol dehydrogenase genes of *G. oxydans* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator. As selection marker serves a *P. ohmeri* *LEU2* gene flanked by two loxP sites.

15

Example 22. Construction of second generation integrative strains capable of secreting more than 100 g/L xylitol

First generation strain CNCM I-4982 containing a randomly integrated copy of the NAD⁺-specific D-arabitol 4-oxidoreductase gene of *E. coli* and the NADPH-specific xylitol dehydrogenase gene of *P. stipitis* was used to further integrate additional copies of the two heterologous enzymes.

20 However, in order to be able to integrate above constructs the *LEU2* selection marker had to be removed. For this purpose first generation strain CNCM I-4982 was transformed with vector pEVE3163 according to the procedure described in **Example 12**.
25 The vector pEVE3163 contains the CRE recombinase of bacteriophage P1 (codon optimized according to Table 7) flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR). Removal of the *LEU2* selection marker was confirmed by no-growth of clones on plates without leucine.

30 The resulting strain EYS3842 was transformed with pEVE3390 or pEVE3392 (**Figure 40**) restriction digested with *NotI* (New England Biolabs, Ipswich, Massachusetts) for 3 h at 37°C according to the procedure described in **Example 12**. Transformants were selected on sorbitol plates without any leucine.

35 Resulting second generation strain EYS3929 contains two NAD⁺-specific D-arabitol 4-oxidoreductase genes of *E. coli* and two NADPH-specific xylitol dehydrogenase genes, one from *G. oxydans* and a second one from *P. stipitis* randomly integrated into the genome. Strain EYS3930, on the other hand, contains an additional NADPH-specific xylitol dehydrogenase gene of *G. oxydans*.

Example 23. Construction of a further vector used for the integration of additional gene copies of the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* and the NADPH-specific xylitol dehydrogenase of *G. oxydans*

5 In order to construct a further integration vector, a double expression cassette of the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* and the NADPH-specific xylitol dehydrogenase of *G. oxydans* was amplified by PCR using:

- primer EV4904 (ATATCCCGGGCACCGTCATCACCGAAACGC) (SEQ ID No 40) containing a *Sma*I site and

10 - primer EV4905 (ATATCCCGGGCACGACCACGCTGATGAGC) (SEQ ID No 41) containing a *Sma*I site (underline) and pEVE3321 as template.

Amplification was performed in a reaction mix consisting of 200 µM of each dNTP and 0.5 µM of each primer with 0.02 U/µl of iProof™ polymerase (BIO-RAD, Hercules, California) in the appropriate 1X buffer. The PCR was accomplished with an initial denaturation step of 30 sec at 98°C followed by 30 cycles with 10 sec at 98°C / 10 sec at 68°C / 75 sec at 72°C, and a final extension step of 5 minutes at 72°C. The PCR product was separated on a 1% agarose gel, extracted and purified using the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California).

20 The amplified 3.9 kb long fragment was restriction digested with *Sma*I (New England Biolabs, Ipswich, Massachusetts) and ligated for 2 h at room temperature to the 4.4 kb *Pvu*II (New England Biolabs, Ipswich, Massachusetts) linearized, Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts) dephosphorylated and gel-purified vector backbone of pEVE2865 (**Figure 35**) using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 41**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

30 The resulting plasmids pEVE4390 (**Figure 42**) contains the double expression construct of the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli* under the control of the *P. ohmeri* phosphoglycerate kinase (poPGK1) promoter and transketolase (poTKL) terminator and the NADPH-specific xylitol dehydrogenase gene of *G. oxydans* flanked by a *P. ohmeri* ribulose reductase promoter and terminator (poRR). As selection marker serves a *P. ohmeri* *LEU2* gene flanked by two loxP sites.

Example 24. Construction of a vector used for the integration of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum*

5 An additional integrative vector for the expression of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and of the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum* was constructed as follows: In a first step a double expression vector containing the two above genes was generated. This double expression cassette was the cloned into an integrative loxP vector.

10 A DNA fragment encoding the NAD⁺-specific D-arabitol 4-oxidoreductase gene of *Ralstonia solanacearum* was chemically synthesized by GeneArt® Gene Synthesis (Life Technologies, Regensburg, Germany) according to the submitted sequence of sequence SEQ ID No 42.

15 Nucleotides 2310548 to 2309151 of sequence AL646052.1 (obtained from <http://www.ncbi.nlm.nih.gov/nuccore/AL646052>) coding for the *dalD* gene were used as template and subjected to codon optimization for use in *P. ohmeri* ATCC 20209 according to Table 7 (above), using the Optimizer program obtained from <http://genomes.urv.es/OPTIMIZER/>. At the 5' and 3' ends of the resulting sequence, nucleotides encoding for the recognition sites of the restriction enzymes *Ascl* (GGCGCGCC) and *SphI* (GCATGC) respectively, were manually added in the text file, in order to facilitate further cloning. Additionally, an adenosine triplet was included in front of the start ATG to account for an adenosine at the -3 position in the Kozak-like sequence of yeasts.

20 The final sequence (SEQ ID No 42) was then submitted for synthesis to GeneArt (Regensburg, Germany). The synthesized DNA fragment encoding the *dalD* gene was delivered as 5 µg lyophilized plasmid DNA in a pMA-RQ derived vector (13AB2EGP, **Figure 25 43**).

The 1.4 kb fragment of the D-arabitol 4-oxidoreductase from *R. solanacearum* was released from vector 13AB2EGP (**Figure 43**) by restriction digested with *Ascl* and *SphI* (New England Biolabs, Ipswich, Massachusetts) and gel-purified with the Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California). The insert was then 30 ligated with the 11.8 kb backbone of pEVE2560 (**Figure 8**) linearized with *Ascl* and *SphI* (New England Biolabs, Ipswich, Massachusetts) and gel purified using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 44**).

35 After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zippy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE3898 (**Figure 45**) contains the codon-optimised *R. solanacearum* NAD⁺-specific D-arabitol 4-oxidoreductase flanked by a ribulose reductase promoter and terminator of *P. ohmeri* and the poLEU2 selection marker.

In a next step the expression cassette containing the NADPH-specific xylitol dehydrogenase of *G. oxydans* flanked by a phosphoglycerate kinase promoter (poPGK) and ribulose reductase terminator (poRR) was released from pEVE3960 by restricon digest with *SpeI* and *SacII* (New England Biolabs, Ipswich, Massachusetts). The 1.8 kb fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C. As vector served the gel-purified (Zymoclean™ Gel DNA Recovery Kit - Zymo Research Corporation, Irvine, California), 13.2 kb *SaII*-linearized pEVE3898. The vector backbone has additionally been blunted for 15 min at room temperature with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts), followed by heat inactivation of the enzymes for 10 min at 70°C and dephosphorylation for 1 h at 37°C using Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). Ligation of vector and insert was performed for 2 h at room temperature to using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 44**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE4077 (**Figure 46**) contains the double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* flanked by a *P. ohmeri* phosphoglycerate kinase promoter (poPGK) and a ribulose reductase terminator (poRR) and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum* under the control of the *P. ohmeri* ribulose reductase promoter and (poRR) terminator and the poLEU2 selection marker.

Finally, the double expression cassette of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum* was released from pEVE4077 (**Figure 46**) by restriction cutting with *SapI* (New England Biolabs, Ipswich, Massachusetts). The 5.9 kb fragment was gel-purified using Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Irvine, California) and blunted with the Blunting Enzyme Mix kit (New England Biolabs, Ipswich, Massachusetts) for 15 min at room temperature, followed by heat inactivation of the enzymes for 10 min at 70°C. As vector served the gel-purified (Zymoclean™ Gel DNA Recovery Kit - Zymo

Research Corporation, Irvine, California), 4.4 kb *EcoRV*-linearized pEVE2865 (**Figure 35**), dephosphorylated for 1 h at 37°C with Antarctic phosphatase (New England Biolabs, Ipswich, Massachusetts). Ligation of vector and insert was performed for 2 h at room temperature to using T4 DNA ligase (New England Biolabs, Ipswich, Massachusetts) (**Figure 44**).

After transformation of XL10 Gold ultracompetent cells (Agilent Technologies, Santa Clara, California) with the ligation mixture, plasmid DNA was isolated using the Zyppy™ Plasmid Miniprep Kit (Zymo Research Corporation, Irvine, California) and further characterized by restriction digestion and sequencing (Microsynth, Balgach, Switzerland).

The resulting plasmid pEVE4377 (**Figure 47**) contains the double expression construct of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and the NAD⁺-specific D-arabitol 4-oxidoreductase of *R. solanacearum* and the poLEU2 selection marker flanked by two loxP sites.

Example 25. Construction of third generation integrative strains with increased productivity of xylitol

The *LEU2* marker of second generation strains EYS3929 and EYS3930 (**Example 22**) was loxed out as described in **Example 18** using vector pEVE3163. The resulting strains EYS4118 and EYS4119 were transformed with pEVE4377 (**Figure 47**) and pEVE4390 (**Figure 42**), respectively. The vectors were restriction digested with *NotI* (New England Biolabs, Ipswich, Massachusetts) for 3 h at 37°C according to the procedure described in **Example 12**. Transformants were selected on sorbitol plates without any leucine.

Resulting third generation strain EYS4353 contains three NAD⁺-specific D-arabitol 4-oxidoreductase genes, two from *E. coli* and one from *R. solanacearum* and three NADPH-specific xylitol dehydrogenase genes, two from *G. oxydans* and one from *P. stipitis* randomly integrated into the genome.

The second third generation strain, on the other hand, contains three copies of the NAD⁺-specific D-arabitol 4-oxidoreductase of *E. coli*, three copies of the NADPH-specific xylitol dehydrogenase of *G. oxydans* and one copy from *P. stipitis*, background and was deposited in France on March 5, 2015, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 PARIS Cedex 15, under number I-4960.

Example 26. Construction of fourth generation integrative strains

The *LEU2* marker of third generation strains CNCM I-4960 (**Example 25**) was loxed out as described in **Example 18** using vector pEVE3163. The resulting strain EYS4955 was

transformed with pEVE4377 (**Figure 47**) restriction digested with *NotI* (New England Biolabs, Ipswich, Massachusetts) for 3 h at 37°C according to the procedure described in **Example 12**. Transformants were selected on sorbitol plates without any leucine.

Resulting fourth generation strain contains four NAD⁺-specific D-arabitol 4-oxidoreductase genes, three from *E. coli* and one from *R. solanacearum* and four NADPH-specific xylitol dehydrogenase genes, three from *G. oxydans* and one from *P. stipitis* randomly integrated into the genome and was deposited in France on May 20, 2015, with the Collection Nationale de Cultures de Microorganismes [National Collection of Microorganism Cultures] of the Institut Pasteur (CNCM), 25 rue du Docteur Roux, 75724 PARIS Cedex 15, under number I-4981.

Example 27. Polyol production with *Pichia ohmeri* strains (synthetic medium)

The yeast strains CNCM I-4605, CNCM I-4982, CNCM I-4960 & CNCM I-4981 constructed as described above, were fermented according to the following protocol.

The fermentation process is run under Nitrogen-limitation and can be separated into a growth phase and a production phase. During the growth phase the ammonia in the medium is completely consumed to produce biomass, once the biomass formation stops the production phase starts and Polyol levels increase. The platform used for the described fermentation process was a Multifors 2 from INFORS HT, using vessels with a working volume of 1 L. The fermenters were equipped with two Rushton six-blade disc turbines. Air was used for sparging the fermenters.

Temperature, pH, agitation, and aeration rate were controlled throughout the cultivation. The temperature was maintained at 36 °C. The pH was kept at 3 by automatic addition of 5M KOH.

The aeration rate was kept at 1.0 vvm and the initial stirrer speed was set to 300 rpm. In order to prevent the Dissolved Oxygen (DO) to drop below 20% an automatic stirring cascade was employed. The operating conditions used in the fermentation process are summarized in Table 10.

Table 10: Operating conditions for the Polyol production fermentations

Parameter	Set-point
Volume of liquid [L]	1
Temperature [° C]	36
pH	3
Agitation speed [rpm]	Initially 300, then DO setpoint (20%) controlled stirrer cascade
Air flow rate [vvm]	1

For inoculation of the fermenters a 1-stage propagation culture was used. The composition of the used propagation culture medium is described in table 11. Propagation cultures were prepared by inoculating 100 ml of medium in a 500-ml shake flask with 4 baffles (indent). The shake flasks were incubated on a shaking table at 30°C and 150rpm.

- 5 The cells were grown for ~24 hrs into mid-exponential phase.

Table 11: Propagation culture medium composition.

Raw material		Concentration [g/L]
Glucose monohydrate	$C_6H_{12}O_6 \cdot H_2O$	46
Antifoam Erol 18		1 drop
Potassium dihydrogenphosphate	KH_2PO_4	6
Magnesium sulfate heptahydrate	$MgSO_4 \cdot 7H_2O$	2.4
Ammonium sulfate	$(NH_4)_2SO_4$	0.16
Iron(II) ammonium sulfate hexahydrate	$Fe(SO_4)_2(NH_4)_2 \cdot 6H_2O$	0.012
Manganese (II) sulfate monohydrate	$MnSO_4 \cdot H_2O$	0.0007
Zinc sulfate heptahydrate	$ZnSO_4 \cdot 7H_2O$	0.00007
Biotine	$C_{10}H_{16}N_2O_3S$	0.0004
Sodium phosphate	Na_2HPO_4	0.292
Citric acid monohydrate	$C_6H_8O_7 \cdot H_2O$	0.835

- 10 Prior to inoculation, an amount of the medium in the fermenter equivalent to the amount of inoculum was removed and an aliquot of the propagation culture was used for inoculation of the fermenter to a final volume of 1 L and an OD_{600} -at-start of ca. 0.2 (CDW ca. 0.03 g/L). The composition of the medium used in the fermenter is described in table 12.

Table 12: Fermentation medium composition.

Raw material		Concentration [g/L]
Glucose monohydrate	$C_6H_{12}O_6 \cdot H_2O$	250
Antifoam Erol 18		0.67
Potassium dihydrogenphosphate	KH_2PO_4	6
Magnesium sulfate heptahydrate	$MgSO_4 \cdot 7H_2O$	2.4
Ammonium sulfate	$(NH_4)_2SO_4$	4
Iron(II) ammonium sulfate hexahydrate	$Fe(SO_4)_2(NH_4)_2 \cdot 6H_2O$	0.012
Manganese (II) sulfate monohydrate	$MnSO_4 \cdot H_2O$	0.0007
Zinc sulfate heptahydrate	$ZnSO_4 \cdot 7H_2O$	0.00007
Biotine	$C_{10}H_{16}N_2O_3S$	0.0004

Samples were withdrawn in regular intervals and the total fermentation broth was analyzed for Glucose consumption and extracellular Polyol (Xylitol, Arabitol and Ribitol) formation. Furthermore common fermentation metabolites (Glycerol, Acetate, Ethanol, Pyruvate, Malate, Fumarate & Succinate) were determined. The increase in biomass was on one hand followed by OD₆₀₀ and on the other hand by cell dry weight (CDW) determination. The above mentioned measurements were used to determine Polyol production, Arabitol or Xylitol yield and productivity; the results are shown in table 13.

Table 13: Polyol production with *Pichia ohmeri* strains (synthetic medium).

	CNCM I-4605	CNCM I-4982	CNCM I- 4960	CNCM I-4981	
Elapsed Fermentation Time (EFT) [h]	67	79	146	64	66
Glucose [g/L]	0	0	0	0	0
Arabitol [g/L]	118	74	0	0	0
Ribitol [g/L]	0	6	2	7	5
Xylitol [g/L]	0	28	60	110	120
Yield Arabitol [%]	52	-	-	-	-
Yield Xylitol [%]	-	12	26	44	48
Productivity [g/L/h]	1.76	0.35	0.41	1.71	1.81

Pichia ohmeri CNCM I-4605 produces arabitol only.

Pichia ohmeri CNCM I-4982 produces arabitol, xylitol and ribitol. In this strain one copy of NAD⁺- D-arabitol 4-oxidoreductase gene and one copy of NADPH-specific xylitol dehydrogenase gene have been integrated. The modified strain is now able to consume arabitol. Consequently, after total consumption of glucose, arabitol and ribitol are re-consumed by CNCM I-4982 to produce more xylitol.

Pichia ohmeri CNCM I-4960 (third generation) and CNCM I-4981 (fourth generation) produce xylitol and ribitol but no more arabitol. The intracellular conversion of arabitol in xylulose and xylitol is efficient enough to avoid the excretion of arabitol into the broth.

The more copies of the genes encoding for the NAD⁺-specific D-arabitol oxidoreductase and the NADPH-specific xylitol dehydrogenase have been introduced into *P. ohmeri*, the higher are the titer, yield and productivity of xylitol.

WE CLAIM:

- 1- A recombinant host cell, wherein said host cell comprises:
a heterologous nucleic acid sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) using D-arabitol as substrate and producing D-xylulose as product; and,
a heterologous nucleic acid sequence encoding a NADPH-specific xylitol dehydrogenase using D-xylulose as substrate and producing xylitol as product.
- 2- The recombinant host cell according to claim 1, wherein the host cell produces D-arabitol from D-glucose, in particular under high osmotic pressure medium.
- 3- The recombinant host cell according to claim 2, wherein the host cell does not consume D-arabitol as a sole carbon source.
- 4- The recombinant host cell according to any one of claims 1-3, wherein the host cell is selected from bacteria, fungi and yeast.
- 5- The recombinant host cell according to claim 4, wherein the host cell is an osmophilic or osmotolerant yeast.
- 6- The recombinant host cell according to claim 4, wherein the host cell is *Pichia ohmeri*.
- 7- The recombinant host cell according to any one of claims 1-6, wherein the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) is from *E. coli* and/or *Ralstonia solanacearum*.
- 8- The recombinant host cell according to claim 7, wherein the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) comprises or consists in the sequence selected from the group of SEQ ID No 2 or 43.
- 9- The recombinant host cell according to claim 8, wherein the sequence encoding the NAD⁺-specific D-arabitol 4-oxidoreductase (EC 1.1.1.11) comprises or consists in the sequence of SEQ ID No 3 or SEQ ID No 42.

- 10- The recombinant host cell according to any one of claims 1-9, wherein the NADPH-specific xylitol dehydrogenase is a xylitol dehydrogenase from *Pichia stipitis* or *Gluconobacter oxydans* mutated for changing the cofactor specificity from NADH to NADPH.
- 11- The recombinant host cell according to claim 10, wherein the NADPH-specific xylitol dehydrogenase comprises or consists in the sequence of SEQ ID No 5 or 8.
- 12- The recombinant host cell according to claim 11, wherein the sequence encoding the NADPH-specific xylitol dehydrogenase comprises or consists in the sequence of SEQ ID No 6 or 9.
- 13- The recombinant host cell according to any one of claims 1-12, wherein the host cell is capable of producing a xylitol titer of at least 15 g/l in the supernatant after a 48 h culture.
- 14- The recombinant host cell according to any one of claims 1-13, wherein the host cell is a strain selected from strains I-4982, I-4960 and I-4981 deposited at the CNCM.
- 15- The recombinant host cell according to any one of claims 1-13, wherein the host cell comprises several copies of a sequence encoding a NAD⁺-specific D-arabitol 4-oxidoreductase and/or several copies of a sequence encoding the NADPH-specific xylitol dehydrogenase.
- 16- A method for producing xylitol comprising culturing a recombinant host cell according to any one of claims 1-15, and recovering xylitol.
- 17- Use of a recombinant host cell according to any one of claims 1-15 for producing xylitol.