[19] 中华人民共和国国家知识产权局

[12] 发明专利说明书

专利号 ZL 03802659.7

[51] Int. Cl.
C12N 15/61 (2006.01)
C12N 15/54 (2006.01)
C12N 1/14 (2006.01)
C12N 1/19 (2006.01)
C12P 7/10 (2006.01)
C12P 7/20 (2006.01)

[45] 授权公告日 2009 年 1 月 7 日

[11] 授权公告号 CN 100448996C

[51] Int. Cl. (续)
C12P 7/46 (2006.01)
C12P 7/54 (2006.01)
C12P 7/56 (2006.01)
C12P 13/04 (2006.01)
C12P 35/00 (2006.01)

[22] 申请日 2003.1.23 [21] 申请号 03802659.7

[30] 优先权

[32] 2002. 1.23 [33] EP [31] 02075266.3

[86] 国际申请 PCT/NL2003/000049 2003.1.23

[87] 国际公布 WO2003/062430 英 2003.7.31

[85] 进入国家阶段日期 2004.7.23

[73] 专利权人 皇家奈达尔科股份有限公司

地址 荷兰贝亨奥普佐姆

[72] 发明人 胡贝图斯·约翰内斯·玛丽·奥普登坎普

哈里·拉曼厄德·哈尔汉吉

克里斯蒂安·范德德里夫特

雅各布斯·托马斯·普龙克

[56] 参考文献

WO 9624667 A1 1996.8.15

US 5935837 A 1999.8.10

US 5858764 A 1999.1.12

审查员 刘红霞

[74] 专利代理机构 永新专利商标代理有限公司

代理人 林晓红

权利要求书 3 页 说明书 26 页 附图 1 页

[54] 发明名称

戊糖的发酵

[57] 摘要

本发明涉及用来自厌氧真菌的编码真核生物的木糖异构酶的核酸序列转化的宿主细胞。 当该编码真核生物木糖异构酶的序列被表达时，赋予宿主细胞将木糖转化为木酮糖的能力，木酮糖进一步被宿主细胞代谢。 因此，宿主细胞能够以木糖作为碳源生长。 宿主细胞优选为真核微生物如酵母或丝状真菌。 本发明还涉及制备发酵产物如乙醇的方法，其中本发明的宿主细胞利用木糖生长并产生发酵产物。 本发明还涉及编码可以从厌氧真菌中获得的真核生物木糖异构酶和木酮糖激酶的核酸序列。

1、以包含编码 *Piromyces* 木糖异构酶的核苷酸序列的核酸构建体转化的真核微生物宿主细胞，所述木糖异构酶包含 SEQ ID NO.1 的氨基酸序列，其中所述的核酸构建体经转化宿主细胞后赋予该宿主细胞在作为碳源的木糖上生长的能力。

2、按照权利要求 1 的转化的宿主细胞，其中所述宿主细胞是酵母。

3、按照权利要求 2 的转化的宿主细胞，其中所述酵母为属于糖酵母属(*Saccharomyces*)、克鲁维氏酵母属(*Kluyveromyces*)、假丝酵母属(*Candida*)、毕赤酵母属(*Pichia*)、裂殖糖酵母属(*Schizosaccharomyces*)、汉逊氏酵母属(*Hansenula*)、克氏酵母属(*Kloechera*)、许旺氏酵母属(*Schwanniomyces*)或耶氏酵母属(*Yarrowia*)之一的酵母。

4、按照权利要求 3 的转化的宿主细胞，其中的酵母属于啤酒糖酵母(*S.cerevisiae*)、*S.bulderi*、*S.barnetti*、微小糖酵母(*S.exiguus*)、葡萄汁糖酵母(*S.uvarum*)、糖化糖酵母(*S.diastaticus*)、乳克鲁维氏酵母(*K.lactis*)、马克斯克鲁维氏酵母(*K.marxianus*)或脆壁克鲁维氏酵母(*K.fragilis*)之一。

5、按照权利要求 1 的转化的宿主细胞，其中所述宿主细胞是丝状真菌。

6、按照权利要求 5 的转化的宿主细胞，其中所述丝状真菌为属于曲霉属(*Aspergillus*)、木霉属(*Trichoderma*)、腐质菌属(*Humicola*)、孢霉属(*Acremonium*)、镰刀霉属(*Fusarium*)或青霉菌属(*Penicillium*)之一的丝状真菌。

7、按照前述权利要求中任一项的转化的宿主细胞，其中编码木糖异构酶的核苷酸序列被可操作地连接于在所述宿主细胞中引起木糖异构酶充分表达的启动子，以赋予所述宿主细胞将木糖异构为木酮糖的能力。

8、按照权利要求 7 的转化的宿主细胞，其中所述启动子对宿主细胞中的分解代谢物阻遏不敏感。

9、按照权利要求 1 的转化的宿主细胞，其中所述宿主细胞包含一种遗传修饰，该遗传修饰产生选自如下一组的特征：

(a) 增加木糖转运到宿主细胞内；

(b) 木酮糖激酶活性增强；

(c) 戊糖磷酸途径的流量增加；

(d) 对分解代谢物阻遏的敏感性降低；

(e) 对乙醇、渗透性或有机酸的耐受力增强，和

(f) 副产物生成下降。

10、按照权利要求 9 的转化的宿主细胞，其中所述遗传修饰由过度表达内源基因、表达异源基因、或其组合组成，且其中所述基因选自编码己糖或戊糖转运蛋白、木酮糖激酶、来自戊糖磷酸途径的酶、糖酵解酶、或 ethanologenic 酶的基因。

11、按照权利要求 9 的转化的宿主细胞，其中所述遗传修饰由灭活内源基因组成，其中该基因选自由编码己糖激酶的基因、糖酵母 *MIG1* 和 *MIG2* 基因、及其杂交的同系物组成的组。

12、按照权利要求 1 的转化的宿主细胞，其中所述宿主细胞表达一种或多种酶，所述的酶赋予所述宿主细胞产生乳酸、乙酸、琥珀酸、氨基酸、1，3-丙二醇、乙烯、甘油、β-内酰胺抗生素和头孢菌素的能力。

13、按照权利要求 12 的转化的宿主细胞，其中所述宿主细胞含有引起乙醇脱氢酶活性下降的遗传修饰。

14、生产乙醇的方法，其中该方法包含用权利要求 1 至 13 中任一项的转化的宿主细胞发酵含有木糖原料的培养基的步骤，其中所述宿主细胞将木糖发酵为乙醇。

15、按照权利要求 14 的方法，其中所述方法还包括回收所述乙醇的步骤。

16、按照权利要求 14 或 15 的方法，其中所述培养基还含有葡萄糖原料。

17、按照权利要求 14 的方法，其中乙醇容积产率至少为 0.5g 乙醇/升/小时。

18、按照权利要求 14 或 15 的方法，其中乙醇产量至少为基于葡萄糖或木糖的理论产量的 50%。

19、生产一种发酵产物的方法，所述发酵产物选自乳酸、乙酸、琥珀酸、氨基酸、1，3-丙二醇、乙烯、甘油、β-内酰胺抗生素和头孢菌素，其中该方法包括用权利要求 12 或 13 的转化的宿主细胞发酵含有木糖原料的培养基的步骤，其中所述宿主细胞将木糖发酵为发酵产物。

20、按照权利要求 19 的方法，其中所述方法还包括回收所述发酵产物的步骤。

21、按照权利要求 19 的方法，其中所述培养基还含有葡萄糖原料。

戊糖的发酵

技术领域

本发明涉及用编码真核生物木糖异构酶的核酸序列转化的宿主细胞。木糖异构酶在宿主细胞中被表达，赋予宿主细胞将木糖转化为木酮糖的能力。宿主细胞被用于通过发酵含有木糖的培养基来制备乙醇和其他发酵产物的方法中。本发明还涉及编码真核生物木糖异构酶的核酸序列。

背景技术

在过去的数十年中，大规模地消耗传统燃料(石油燃料)已经造成严重的污染。此外，对世界石油储备是有限的认识以及环境意识提高，已经刺激了研究替代燃料如乙醇的可行性的新动机，乙醇能够降低 CO_2 生成的 60-90%。虽然生物来源的乙醇可以通过发酵从许多不同来源获得的己糖来产生，但是迄今用作燃料乙醇的工业生产的底物是蔗糖和玉米淀粉。这些底物的缺点是十分昂贵。

扩大燃料乙醇的生产要求能够采用较低成本的原料。目前，仅仅来自植物生物的木质素纤维素原料可以被大量获得来替代用于乙醇生产的作物。木质素纤维素材料的主要可发酵糖为葡萄糖和木糖，分别占木质素纤维素的 40%和 25%。但是，大部分能够进行乙醇发酵的酵母如啤酒酵母(*Saccharomyces cerevisiae*)不能够以木糖作为碳源。此外，还不知晓哪些微生物能够以高乙醇产量和高乙醇生产率的方式地将木糖发酵为乙醇。为了能够从木质素纤维素水解产物中商业化地生产乙醇，需要具有这两种特性的微生物。因此，本发明的目的在于提供能够进行乙醇发酵和利用木糖作为碳源的酵母。

D-木糖可以被许多微生物代谢，如肠细菌、一些酵母和真菌。在大部分利用木糖的细菌中，木糖直接被木糖(葡萄糖)异构酶(XI)异构为 D-木酮糖。但是，丝状真菌和酵母不具有进行这种一步异构化的能力，首先通过木糖还原酶(XR)作用将木糖还原为木糖醇，然后木糖醇被木糖醇脱氢酶(XDH)转化为木酮糖。第一步需要 NAD(P)H 作为辅助因子，而第二步需要 NAD^+。随后，在生成的木酮糖被木酮糖激酶(XK)磷酸化后进入戊糖磷酸途径(PPP)。在具有严格 NADPH 依赖性的木糖还原酶(XR)的生物体中，木糖厌氧发酵为乙醇是不可能的。这是因为木糖醇脱氢酶(XDH)是严格 NAD^+依赖的，导致氧化还原失衡(即 NAD^+损耗)。为了解决厌氧条件下的氧化还原失衡，生物体生成副产物如甘油和木糖醇。类似地，与由葡萄糖产生β-内酰胺相比，由木糖需氧产生β-内酰胺也受到负面影响。导致这些低产量的可能原因还有，与采用葡萄糖相比，该路径中需要相对较多形式为 NADPH 的还原性等价物(W.M. van Gulik 等,生物技术生物工程(Biotechnol. Bioeng.)，第 68 卷，第 6 期，2000 年 6 月 20 日)。

在过去数年中，人们已经作了许多努力来将木糖代谢引入啤酒酵母和类似酵母中，如 Zaldivar 等讨论(2001，应用微生物生物学(Appl. Microbiol. Biotechnol.)，56：17-34)。一种方法涉及至少编码木糖(醛糖)还原酶和木糖醇脱氢酶的基因，如毕赤氏酵母(*Pichia stipitis*)的 *XYL1* 和 *XYL2* 在啤酒酵母中表达(US 5,866,382；WO 95/13362 和 WO 97/42307)。虽然这种方法能够使啤酒酵母在木糖中生长，其通常是低乙醇生产率和产量以及高木糖醇生成，主要由于 XR 和 XDH 之间的氧化还原失衡。

在啤酒酵母或相关酵母或在丝状真菌中，XI 的表达将避免氧化还原失衡和随后的木糖醇生成和分泌。来自多种细菌的木糖异构酶基因已经被插入到啤酒酵母中，但是嗜温性原核生物的 XI 在啤酒酵母中的表达不能产生活性 XI(Amore 和 Hollenberg，1989，核酸研究(Nucleic Acids Res.)17：7515；Amore 等，1989，应用微生物生物技术(Appl. Microbiol.

Biotechnol.)30：351-357；Chan 等 1986，生物技术通讯(Biotechnol. Lett)8：231-234；Chan 等 1989，应用微生物生物技术(Appl. Microbiol. Biotechnol.)31：524-528；Ho 等，1983, Fed.Proc.Fed. Am.Soc.Exp.Bio. 42：2167；Hollenberg，1987，EBC 啤酒酵母讨论会，赫尔辛基(芬兰)，24-25，1986 年 11 月；Sarthy 等，1987，应用环境微生物学(Appl. Environ. Microbiol.)53：1996-2000；Ueng 等，1985，生物技术通讯(Biotechnol. Lett.)7：153-158)。然而，在啤酒酵母中表达的两种来自嗜热细菌的 XI 在 85℃下具有 1μmol/min/mg 的特异性活性(Bao 等，1999，微生物学报 39：49-54；Walfridson 等，1996，应用环境微生物学(Appl. Environ. Microbiol.)61：4184-4190)。但是，在啤酒酵母的生理温度下(20-35℃)仅仅只保留这种活性的极少部分，对于由木糖进行有效的乙醇发酵是不够的。因此，仍然需要编码 XI 的核苷酸，该核酸能够在酵母中被表达来提供生理条件下足够以木糖作为碳源的 XI 活性。

发明描述

定义

木糖异构酶

“木糖异构酶”(EC5.3.1.5.)在此定义为催化 D-木糖直接异构化为 D-木酮糖和反之亦然的酶。这种酶也称之为 D-木糖酮异构酶。一些木糖异构酶还能够催化 D-葡萄糖和 D-果糖之间的转化，因而有时是指葡萄糖异构酶。木糖异构酶需要镁作为辅助因子。本发明的木糖异构酶还通过本文下面公开的其氨基酸序列进一步限定。同样地，木糖异构酶可以通过编码该酶的核苷酸序列以及通过本文下面描述的与编码木糖异构酶的参比核苷酸序列杂交的核苷酸序列限定。

木糖异构酶活性单位(U)在此定义为 37℃下在含有 50mM 磷酸缓冲液(pH7.0)、10mM 木糖和 10mM $MgCl_2$ 的反应混合物中每分钟生成 1 nmol 木酮糖时所需的酶的量。生成的木酮糖通过 Dische 和 Borenfreund

的方法(1951，生物活性杂志 192：583-587)或实施例中描述的 HPLC 确定。

序列相同性和相似性

序列相同性在此定义为两条或两条以上氨基酸(多肽或蛋白质)序列或两条或两条以上核酸(多核苷酸)序列之间的关系，通过比较序列来确定。在本领域，视情况而定，“相同性”还指氨基酸或核酸序列之间的序列相关程度，由这些序列的链之间的匹配确定。两条氨基酸序列之间的“相似性”通过一种多肽的氨基酸序列及其保守氨基酸取代与第二种多肽的序列比较来确定。“相同性”和“相似性”可以容易地由已知方法计算，包括但不限于在(计算分子生物学(Computational Molecular Biology)，Lesk,A.M.,编辑，牛津大学出版社，纽约，1988；生物计算：信息学和基因组计划(Biocomputing: Informatics and Genome Projects)，Smith,D.W.,编辑，学院出版社，纽约，1993；序列数据的计算分析，第一部分(Computer Analysis of Sequence Data, Part I)，Griffin,A.M.和 Griffin,H.G.编辑，Hurnana 出版社，新泽西，1994；分子生物学中的序列分析(Sequence Analysis in Molecular Biology)，von Heine,G.，学院出版社，1987；和序列分析引物(Sequence Analysis Primer)，Gribskov, M.和 Devereux, J.，编辑，M Stockton 出版社，纽约，1991；和 Carillo, H. 和 Lipman,D., SIAM 应用数学杂志(SIAM J. Applied Math.)48：1073(1988)中描述的方法。

确定相同性的优选方法被设计来在被测试的序列间产生最大匹配。确定相同性和相似性的方法被编写在公众可以获得的计算机程序中。确定两条序列之间的相同性和相似性的优选计算机程序方法包括如GCG程序包(Devereux,J.等，核苷酸研究(Nucleic Acids Research)12(1)：387(1994))、BestFit、BLASTP、BLASTN 和 FASTA(Altschul,S.F.等，分子生物学杂志(J. Mol. Biol.) 215：403-410(1990)。BLAST X 程序是公众可以从 NCBI 或其他来源(BLADT Manual，Altschul,S.,等，NCBI NLM NIH Bethesda，马里兰 20894；Altschul,S.,等，分子生物学杂志(J. Mol. Biol.) 215：403-410(1990)获得。众所周知的 Smith Waterman 算法也可以用于确

定相同性。

用于多肽序列比较的优选参数包括如下：算法：Needleman 和 Wunsch，分子生物学杂志(J. Mol. Biol.)，48：443-453(1970)；比较矩阵：Hentikoff 和 Hentikoff，美国国家科学院汇编(Proc. Natl. Acad. Sci. USA.)89：10915-10919(1992)的 BLOSSUM62；缺口罚分：12；和缺口长度罚分：4。用于这些参数的程序是公众可以得到的威斯康星州麦迪逊的 Genetics Computer Group 的“Ogap”程序。前述的参数是氨基酸比较的罚分参数(以及缺口结束不罚分)。

核苷酸比较的优选参数包括如下：算法：Needleman 和 Wunsch，分子生物学杂志(J. Mol. Biol.)，48：443-453(1970)；比较矩阵：匹配=+10，错配=0；缺口罚分：50；缺口长度罚分：3。可以获得威斯康星州麦迪逊的 Genetics Computer Group 的 Gap 程序。以上给出的是核酸比较的罚分参数。

任选地，在确定氨基酸相似程度中，技术人员应当还会考虑到所谓的“保守”氨基酸取代，这对于技术人员来说是显然的。保守氨基酸取代是指具有相似侧链的残基的可互换性。例如，一组具有脂族侧链的氨基酸为甘氨酸、丙氨酸、缬氨酸、亮氨酸和异亮氨酸；一组具有脂族羟基侧链的氨基酸为丝氨酸和苏氨酸；一组具有含有酰胺侧链的氨基酸为天冬酰胺酸和谷氨酰胺，一组具有芳香族侧链的氨基酸为苯基丙氨酸、酪氨酸和色氨酸；一组具有碱性侧链的氨基酸为赖氨酸、精氨酸和组氨酸；和一组具有含硫侧链的氨基酸为半胱氨酸和甲硫氨酸。优选的保守氨基酸取代组为：缬氨酸-亮氨酸-异亮氨酸、苯基丙氨酸-酪氨酸、赖氨酸-精氨酸、丙胺酸-缬氨酸和天冬酰胺酸-谷氨酰胺。本文公开的氨基酸序列取代变体是那些在公开的序列中至少一个残基被去除并且一个不同的残基被插入到其位置上的氨基酸序列。优选地，氨基酸变化是保守的。每一个天然氨基酸的优选保守取代如下：丙胺酸取代丝氨酸；精氨酸取代赖氨酸；天冬酰胺取代谷氨酰胺或组氨酸；天冬氨酸取代谷氨酸；半

胱氨酸取代丝氨酸或丙氨酸；谷氨酰胺取代天冬酰胺；谷氨酸取代天冬氨酸；甘氨酸取代脯氨酸；组氨酸取代天冬酰胺或谷氨酰胺；异亮氨酸取代亮氨酸或苏氨酸；苏氨酸取代丝氨酸；色氨酸取代酪氨酸；酪氨酸取代色氨酸或苯丙氨酸；和缬氨酸取代异亮氨酸或亮氨酸。

核酸序列杂交

本发明的编码木糖异构酶或木酮糖激酶的核苷酸序列还可以通过其在适度的或优选为严格的杂交条件下分别与 SEQ ID NO.2 或 SEQ ID NO.4 的核苷酸序列杂交的能力来限定。严格的杂交条件在此定义为使至少约 25 个核苷酸、优选约 50、75 或 100 个核苷酸、更优选约 200 或更多个核苷酸的核酸序列在约 65℃的温度下、在含有约 1M 盐、优选 6 x SSC 或任何其他含有相似离子强度的溶液中杂交，在 65℃下在含有约 0.1M 或更低浓度的盐、优选 0.2 x SSC 或任何其他具有相似离子强度的溶液中洗涤的条件。优选地，进行杂交过夜，即至少 10 小时，优选洗涤进行至少 1 小时并至少更换洗涤溶液 2 次。这些条件通常使具有约 90%或更高序列相同性的序列特异性地杂交。

适度条件在此定义为使至少约 50 个核苷酸、优选约 200 或更多个核苷酸的核苷酸序列在约 45℃的温度下、在含有约 1M 盐、优选 6 x SSC 或任何其他含有相似离子强度的溶液中杂交，在室温下在含有约 1M 或更低浓度的盐、优选 6 x SSC 或任何其他具有相似离子强度的溶液中洗涤的条件。优选地，进行杂交过夜，即至少 10 小时，优选洗涤进行至少 1 小时并至少更换洗涤溶液 2 次。这些条件通常使达到 50%的序列相同性的序列特异性地杂交。本领域技术人员将能够改进这些杂交条件来特异性地确定相同性在 50%和 90%之间变化的序列。

可操作地连接

如本文所用术语“可操作地连接”是指多核苷酸元件根据功能关系连接。当一条核酸与另一条核酸序列处于一种功能关系中，该核酸即为“可操作地连接”。例如，如果启动子或增强子影响了编码序列的转录，

则为可操作地连接到该编码序列上。可操作地连接是指被连接的 DNA 序列通常是毗邻的，并且当需要将两个蛋白质的编码区结合在一起时，被连接的 DNA 序列是毗邻并在阅读框中。

启动子

本文所用术语“启动子”是指能够控制一种或多种基因转录的核酸片段，位于基因的转录起始位点的转录方向的上游，可以通过存在 NDA 依赖的 RNA 聚合酶结合位点、转录起始位点和任何其他 DNA 序列，包括但不限于转录因子结合位点、抑制子和激活子蛋白结合位点和本领域技术人员已知直接或间接地调控该启动子的转录量的任何其他核苷酸序列来确定。“组成型”启动子是在大多数环境的和发育条件下具有活性的启动子。“可诱导的”启动子是在环境和发育调控下具有活性的启动子。

发明详述

在第一个方面，本发明涉及能够将木糖异构为木酮糖的转化的宿主细胞。通过用包含编码木糖异构酶的核苷酸序列的核酸构建体转化宿主细胞来赋予该宿主细胞将木糖异构为木酮糖的能力。被转化的宿主细胞异构木糖为木酮糖的能力是直接将木糖异构为木酮糖。这应当理解为在由木糖异构酶催化的单个反应中将木糖异构为木酮糖，不同于木糖通过木糖醇过渡两步转化为木酮糖，其分别由木糖还原酶和木糖醇脱氢酶催化。

编码木糖异构酶的核苷酸序列优选以活性形式在被转化的宿主细胞中被表达。因此，核苷酸序列在宿主细胞中表达生成具有特定活性的木糖异构酶，在 25℃下每毫克蛋白至少 10U 木糖异构酶活性，优选 25℃下每毫克蛋白至少 20、25、30、50、100、200 或 300U。在被转化的宿主细胞中表达的木糖异构酶特异活性在此定义为宿主细胞的无细胞溶解产物如酵母的无细胞溶解产物中的每毫克蛋白的木糖异构酶活性单位量。木糖异构酶活性的确定、蛋白质含量和无细胞溶解产物的制备在实施例 1

中描述。另外，特异性活性可以如在实施例 4 中所指出确定。相应地，核苷酸序列在宿主细胞中表达生成具有特异活性的木糖异构酶，在 30℃下每毫克蛋白质至少 50U，优选至少 100、200、500、或 750U 的木糖异构酶活性。

优选地，核苷酸序列在宿主细胞中表达生成具有木糖 K_m 小于 50、40、30 或 25mM 的木糖异构酶，更优选地，木糖 K_m 约 20mM 或更小。

编码木糖异构酶的核苷酸序列可选自由下面构成的组：

（a）编码包含与 SEQ ID NO.1 的氨基酸序列具有至少 40、45、49、50、53、55、60、70、80、90、95、97、98 或 99%的序列相同性的氨基酸序列的多肽的核苷酸序列；

（b）包含与 SEQ ID NO.2 的核苷酸序列具有至少 40、50、55、56、57、60、70、80、90、95、97、98 或 99%的序列相同性的核苷酸序列的核苷酸序列；

（c）其互补链与(a)或(b)的核酸分子序列杂交的核苷酸序列；

（d）其序列由于遗传密码子的简并性区别于(c)的核酸分子的核苷酸序列。

核苷酸序列优选编码真核生物木糖异构酶，即氨基酸序列与天然存在于真核微生物中的木糖异构酶的序列相同的木糖异构酶。真核生物木糖异构酶的表达提高了木糖异构酶在真核宿主细胞如酵母中以活性形式被表达的可能性，与嗜温原核生物的木糖异构酶不同。更优选地，核苷酸序列编码植物木糖异构酶(如来自大麦(*Hordeum vulgare*))或真菌木糖异构酶(如来自担子菌类(*Basidiomycete*))。但是，更优选地，编码厌氧真菌木糖异构酶的核苷酸序列进一步提高在真核宿主细胞、特别在酵母中以酶催化活性形式表达的可能性。更加优选的是编码厌氧真菌的木糖异构酶的核苷酸序列，该厌氧真菌属于 *Neocallimastix*、*Caecomyces*、*Piromyces*、*Orpinomyces* 或 *Ruminomyces* 家族。

用编码木糖异构酶的核苷酸序列转化的宿主细胞优选为能够主动或

被动地将木糖转运到细胞里面的宿主。宿主细胞优选地含有主动糖酵解、戊糖磷酸途径和优选地含有木酮糖激酶活性使得从木糖异构得到的木酮糖可以被代谢为丙酮酸。宿主优选还含有将丙酮酸转化为期望的代谢产物如乙醇、乙烯或乳酸的酶。优选的宿主细胞为天然能够进行乙醇发酵、优选为厌氧乙醇发酵的宿主细胞。宿主细胞优选还具有对乙醇和有机酸如乳酸、乙酸或蚁酸以及糖代谢产物如糠醛和羟甲基糠醛的高耐受力。宿主细胞的这些特征或活性的任何之一可以是天然存在于宿主细胞或通过遗传修饰被引入或被改进。适合的宿主细胞是微生物如细菌或真菌，但是，最适合作为宿主细胞的是酵母或丝状真菌。

酵母在此被定义为真核微生物，包括主要以单细胞形式生长的所有种类的真菌亚门(Alexopoulos,C.J.,1962，真菌学导论(Introductory Mycology)，John Wiley&Sons,Inc,纽约)。真菌可能通过单细胞菌体芽殖生长或者通过机体融合生长。作为宿主细胞的优选酵母属于酵母属*(Saccharomyces)*、克鲁维酵母属*(Kluyveromyces)*、假丝酵母属*(Candida)*、毕赤酵母属*(Pichia)*、粟酒裂殖酵母属*(Schizosaccharomyces)*、汉逊氏酵母属*(Hansenula)*、克氏酵母属*(Kloechera)*、许旺氏酵母属*(Schwanniomyces)*和耶氏酵母属*(Yarrowia)*。优选地，酵母能够进行厌氧发酵、更优选进行厌氧乙醇发酵。

丝状真菌在此被定义为真核微生物，包括所有丝状形式的真菌亚门。这些真菌特征为由壳多糖、纤维素和其他多糖复合物组成的营养菌丝体。本发明的丝状真菌从形态学、生理学和遗传学上区别于酵母。丝状真菌的营养性生长是通过菌丝增长，大多数丝状真菌的碳代谢是专性需氧的。作为宿主细胞的优选丝状真菌属于曲霉属(*Aspergillus*)、木霉属(*Trichoderma*)、腐质菌属(*Humicola*)、孢霉属(*Acremonium*)、镰刀霉属(*Fusarium*)和青霉菌属(*Penicillium*)。

在过去数年中，有人建议引入各种生物体用于从作物糖类中产生生物性乙醇。但是，实际上，所有主要生物性乙醇生产方法中一直采用酵

母属的酵母作来产生乙醇。这是由于酵母属用于工业生产的许多具有吸引力的特点，即对酸性、醇和渗透性的高耐受力，能够厌氧生长，以及其高乙醇发酵能力。作为宿主细胞的优选酵母种类包括啤酒酵母、*S.Bulderi*、*S.barnetti*、*S.exiguus*、*S.uvarum*、*S.diastaticus*、乳酸克鲁维酵母(*K.lactis*)、*K.marxianus*、脆壁克鲁维酵母(*K.fragilis*)。

宿主细胞用下文进一步限定的核酸构建体转化，含有单个但是优选含有多个核酸构建体的拷贝。核酸构建体被保持为游离状态，因此包含用于自动复制的序列，如 ARS 序列。合适的游离核酸构建体可能如基于酵母 2μ 或 pKD1(Fleer 等，1991，生物技术(Biotechnology) 9：968-975)质粒。但是，优选地，核酸构建体以一个或多个拷贝被整合到宿主细胞的基因组中。整合到宿主细胞的基因组可以通过非常规重组随机发生，但是优选地通过在真菌分子遗传学领域熟知的同源重组将核酸构建体整合到宿主细胞基因组(见如 WO 90/14423、EP-A-0481008、EP-A-0635574 和 US 6,265,186)。

在按照本发明的优选转化的宿主细胞中，核酸构建体赋予宿主细胞以木糖作为碳源、优选作为唯一的碳源，优选在厌氧条件下生长的能力，优选地，被转化的宿主实质上不产生木糖醇，如产生的木糖醇低于检测极限或者如基于摩尔浓度低于被消耗的碳的 5、2、1%。被转化的宿主细胞能够以木糖作为唯一碳源以至少 0.01、0.02、0.05、0.1 或 $0.2h^{-1}$ 的速度生长。因此，本发明的转化的宿主细胞表达具有上述限定的特异活性水平的木糖异构酶。

宿主细胞还包含遗传修饰，产生一种或多种选自由下列构成的组的特征：(a)增加木糖转运到宿主细胞内；(b)木酮糖激酶活性增强；(c)戊糖磷酸路径的流量增加；(d)对代谢产物阻遏的灵敏度降低；(e)对乙醇、渗透性或有机酸的耐受力增强，和(f)副产物生成下降。副产物被理解为是指含碳分子而不是期望的发酵产物，包括如木糖醇、甘油和/或乙酸。这种遗传修饰可以通过常规诱变和筛选和/或选择期望突变来引入。另一方

面，遗传修饰包括过度表达内源基因和/或表达异种基因和/或灭活内源基因。基因优选选自编码己糖或戊糖转运子的基因；木酮糖激酶如来自啤酒酵母(*XKS1* Deng 和 Ho，1990 应用生物化学生物技术(Appl. Biochem. Biotechnol.) 24-25：193-199)或 Piromyces(*xylB*，即 SEQ ID NO.4)的木酮糖激酶基因；戊糖磷酸途径的酶如转醛酶(*TAL1*)或转酮酶(*TKL1*)(见如 Meinander 等，1995，药物毒理学，增刊(Pharmacol. Toxicol. Suppl.) 2：45)、糖酵解酶、ethanologenic 酶如乙醇脱氢酶。优选灭活的内源基因包括己糖激酶基因如啤酒酵母 *HXK2* 基因(见 Diderich 等，2001，应用环境微生物学(Appl. Environ. Microbiol.)67：1587-1593)；啤酒酵母 *MIG1* 或 *MIG2* 基因；(非特异)醛糖还原酶基因如啤酒酵母 *GRE3* 基因(Träff 等，2001，应用环境微生物学(Appl. Environ. Microbiol.)67：5668-5674)；参与甘油代谢的酶的基因如啤酒酵母甘油磷酸脱氢酶 1 和/或 2 基因；或这些基因在其他宿主种类中的(杂交)同系物。对宿主细胞的木糖发酵的更加优选的修饰在 Zaldivar 等中评述(2001,同上)。

在另一方面，本发明涉及用于生产非乙醇发酵产物的被转化的宿主细胞。这种非乙醇的发酵产物大体上包括任何批量或少数可以通过真核微生物如酵母或丝状真菌产生的化学物质。这种分解产物包括如乳酸、乙酸、琥珀酸、氨基酸、1，3-丙二醇、乙烯、甘油、β-内酰胺抗生素和头孢菌素

如上述的用本发明的核酸构建体转化宿主细胞和宿主细胞优选为酵母的其他遗传修饰可以通过本领域熟知的方法进行。这种方法如可以从标准手册中知晓，如 Sambrook 和 Russel(2001)“分子克隆：实验室手册(第 3 版)，冷泉巷实验室，冷泉巷实验室出版社，或 F.Ausubel 等，编辑，“现有的分子生物学操作”，Green Publishing and Wiley Interscience，纽约(1987)。用于真菌宿主细胞的转化和遗传修饰的方法可以从如 EP-A-0635574、WO98/46772、WO99/60102 和 WO00/37671 中知晓。

在另一方面，本发明涉及包含编码如上定义的木糖异构酶的核苷酸

序列的核酸构建体，用于转化如上定义的宿主细胞。在核酸构建体中，编码木糖异构酶的核苷酸如下定义地被可操作地连接到启动子上来控制和起始核苷酸序列在宿主细胞中转录。启动子优选地能够在宿主细胞中引起足够的木糖异构酶表达，来赋予宿主细胞将木糖异构为木酮糖的能力。优选地，如上定义启动子在宿主细胞中产生特异的木糖异构酶活性。用于本发明的核酸构建体中的启动子包括组成型和诱导型的启动子以及工程启动子。在本发明中使用的优选启动子另外还将对代谢产物(葡萄糖)阻遏不敏感和/或优选不需要用木糖诱导。具有这些特征的启动子可以广泛地获得，是本领域技术人员已知的。这种启动子的适合例子包括如糖分解基因的酵母启动子，如酵母磷酸果糖激酶(PPK)、丙糖磷酸异构酶(TPI)、甘油醛-3-磷酸脱氢酶(GPD、TDH3 或 GAPDH)、丙酮酸激酶(PYK)、磷酸甘油酸激酶(PGK)启动子；关于这些启动子更加详细的内容可见(WO93/03159)。其他有用的启动子为核糖体蛋白编码基因启动子、乳糖酶基因启动子(LAC4)、乙醇脱氢酶启动子(ADH1，ADH4 等等)，和烯醇化酶启动子(ENO)。其他组成型和诱导型的启动子和增强子或上游激活序列是本领域技术人员已知的。如果需要，用于本发明的核酸构建体的启动子可以被修饰，来影响其控制特性。优选地，为了表达木糖异构酶而用于核酸构建体的启动子与表达木糖异构酶的宿主细胞是同源的。

在核酸构建体中，编码木糖异构酶的核苷酸序列的 3'端优选被连接到转录终止序列。优选地，终止序列在选择宿主细胞如选择酵母种类中是可操作的。无论如何，终止子的选择不是苛刻的，可以如来自任何的酵母基因，来自非酵母的真核生物的基因有时也可以起作用。转录终止序列优选地还包含多聚腺苷酸信号。

任选地，选择性标记可存在于核酸构建体中。在本文中所用的术语“标记”是指编码一种用于选择或筛选的特征或表型的基因，宿主细胞含有该标记。标记基因可以是抗生素抗性基因，适合的抗生素被用于从未被转化的细胞中选择被转化细胞。适合的抗生素抗性标记的例子包括

如二氢叶酸还原酶、潮霉素-B-磷酸转移酶、3'-O-磷酸转移酶II(卡那霉素、新霉素和 G418 抗性)。虽然抗生素抗性标记最方便于多倍体宿主细胞的转化，但是优选地，可以使用非抗生素抗性标记，如营养缺陷型标记(URA3、TRP1、LEU2)或非洲酒酵母(*S.pombe)* TP1 基因(在 Russell P R，1985，基因(Gene) 40：125-130 中描述)

在一个优选实施方案中，用核酸构建体转化的宿主细胞无标记基因。用于构建无标记基因的重组微生物宿主细胞的方法公开在 EP-A-0635574 中，基于采用双向标记如 *A.nidulans* 的 amdS(乙酰胺酶)基因或酵母 URA3 和 LYS2 基因。另外，可筛选的标记如绿色萤光蛋白、lacZ、萤光素酶、氯霉素乙酰转移酶、β-葡糖苷酸酶可以被并入本发明的核酸构建体，用来筛选被转化的细胞。

存在于本发明的核酸构建体中的其他可选择元件包括但不限于一条或多条先导序列、增强子、整合因子和/或报道基因、内含子序列、着丝粒(centromers)、端粒(telomer)和/或基质附着(MAR)序列。本发明的核酸构建体还包含用于自动复制的序列，如 ARS 序列。适合的游离核酸构建体可以如基于酵母 2μ 或 pKD1(Fleer 等，1991，生物技术(Bioteclmology) 9：968-975)质粒。另外，核酸构建体还包含用于整合序列，优选通过同源重组。因此，这种序列是与宿主细胞基因组中整合靶位点同源的序列。本发明的核酸构建体以本身已知的方式被提供，通常包括限制和连接核苷酸/核苷酸序列方法，对此可参考标准手册，如 Sambrook 和 Russel(2001)"分子克隆：实验室手册(第 3 版)，冷泉巷实验室，冷泉巷实验室出版社"，Green Publishing and Wiley Interscience，纽约(1987)。

在另一方面，本发明涉及包含编码木糖异构酶的核苷酸序列的核酸分子。该核酸分子优选地选自由下面构成的组：

（a）编码包含与 SEQ ID NO.1 的氨基酸序列具有至少 40、45、49、50、53、55、60、70、80、90、95、97、98 或 99%的序列相同性的氨基酸序列的多肽的核酸分子；

（b）包含与 SEQ ID NO.2 的核苷酸序列具有至少 40、50、55、56、57、60、70、80、90、95、97、98 或 99%的序列相同性的核苷酸序列的核酸分子；

（c）其互补链与(a)或(b)的核酸分子序列杂交的核酸分子；

（d）其序列由于遗传密码子的简并性区别于(c)的核酸分子的核酸分子。

或者，(a)的核酸分子可以编码包含与 SEQ ID NO.1 的氨基酸序列具有至少 67、68、69、70、80、90、95、97、98、或 99%的序列相似性的氨基酸序列的多肽。(c)的核苷酸序列优选地在适当条件下，优选在如本文上述限定的严紧条件下杂交。优选地，核酸分子来自真核生物，更优选来自真核微生物如真菌，最优选来自厌氧真菌如上述的厌氧真菌。

还有一方面，本发明涉及包含编码木酮糖激酶优选 D-木酮糖激酶的核苷酸序列的核酸分子。D-木酮糖激酶(D-xylulose kinase)(EC2.7.1.17；也称作为 D-木酮糖激酶(D-xylulokinase)在此定义为催化 D-木酮糖转化为木酮糖-5-磷酸的酶。核酸分子优选地选自由下列构成的组：

（a）编码包含与 SEQ ID NO.3 的氨基酸序列具有至少 45、47、48、49、50、55、60、70、80、90、95、97、98 或 99%的序列相同性的氨基酸序列的多肽的核酸分子；

（b）包含与 SEQ ID NO.4 的核苷酸序列具有至少 30、37、38、39、40、50、60、70、80、90、95、97、98 或 99%的序列相同性的核苷酸序列的核酸分子；

（c）其互补链与(a)或(b)的核酸分子序列杂交的核酸分子；

（d）其序列由于遗传密码子的简并性区别于(c)的核酸分子的核苷酸序列。

或者，(a)的核酸分子可以编码包含与 SEQ ID NO.3 的氨基酸序列具有至少 64、65、66、70、80、90、95、97、98、或 99%的序列相似性的氨基酸序列的多肽。(c)的核苷酸序列优选地在适当条件下，优选在如本

文上述限定的严紧条件下杂交。优选地，核酸分子来自真核生物，更优选来自真核微生物如真菌，最优选来自厌氧真菌如上述的厌氧真菌。

在另一个方面，本发明涉及本发明的转化的宿主细胞用于发酵碳源的发酵方法，糖源包括木糖原料如木糖。除了木糖原料，发酵液中的碳源还包括葡萄糖原料。木糖或葡萄糖原料可以是木糖或葡萄糖本身或者是任何包含木糖或葡萄糖部分的碳水化合物低聚体或多聚物，如木质素纤维素、木聚糖、纤维素、淀粉等。对于从这种碳水化合物中释放木糖或葡萄糖部分，将适合的糖酶(如木聚糖酶、葡聚糖酶、淀粉酶等等)加入到发酵液中或者可以通过转化宿主细胞产生。在后一种情形中，被转化的宿主细胞通过基因工程加工来产生和分泌这些糖酶。在一种优选方法中，被转化的宿主细胞发酵木糖和葡萄糖，优选同时地发酵，在此情况下，优选使用对葡萄糖阻遏不敏感的转化的宿主细胞来防止二次生长。除了以木糖(和葡萄糖)来源作为碳源，发酵液还包含转化的宿主细胞生长所需的适当成分。微生物如酵母生长所用的发酵液的组成在本领域是已知的。

发酵过程是产生发酵产物如乙醇、乳酸、乙酸、琥珀酸、氨基酸、1，3-丙二醇、乙烯、甘油、β-内酰胺抗生素如青霉素G和青霉素V及其可发酵的衍生物和头孢菌素的过程。发酵过程可以是需氧或厌氧的发酵过程。厌氧发酵过程在此定义为在缺乏氧气或者实质上无氧被消耗时如低于 5mmol/L/h 进行发酵过程，其中有机分子作为电子供体和电子受体。缺乏氧气时，在糖酵解和生物量形成中生成的NADH不能被氧化磷酸化氧化。为了解决该问题，许多微生物利用丙酮酸或其衍生物作为电子和氢的受体，从而再生NAD^+。因此，在优选厌氧发酵过程中，丙酮酸被用作电子(和氢受体)，被还原为发酵产物如乙醇、乳酸、1，3-丙二醇、乙烯、乙酸或琥珀酸。

发酵过程优选在最适合于被转化的宿主细胞的温度下进行。因此，对于大多数酵母或真菌宿主细胞来说，发酵过程在低于 38℃的温度下进

行。对于酵母或丝状真菌宿主细胞来说，发酵过程优选在低于 35、33、30 或 28℃和高于 20、22 和 25℃的温度下进行。

优选过程为生成乙醇的过程，其中该过程包含步骤：(a)用上面限定的转化宿主细胞发酵含有木糖原料的培养基，其中宿主细胞将木糖发酵为乙醇；和任选地，(b)回收乙醇。发酵液还可以包含葡萄糖原料，其也被发酵为乙醇。在该过程中，容量乙醇生产率优选至少为 0.5、1.0、1.5、2.0、2.5、3.0、5.0 或 10.0g 每升每小时。在该过程中，基于木糖和/或葡萄糖的乙醇产量优选至少为 50、60、70、90、95 或 98%。乙醇产量在此定义为理论产量的百分比，对于葡萄糖和木糖，产量为 0.51g 乙醇/g 葡萄糖或木糖。

在另一方面，本发明涉及生产发酵产物的方法，发酵产物选自由乳酸、乙酸、琥珀酸、氨基酸、1，3-丙二醇、乙烯、甘油、β-内酰胺抗生素和头孢菌素构成的组。该方法优选包含步骤(a)用本文上述限定的转化宿主细胞发酵含有木糖原料的培养基，其中宿主细胞将木糖发酵为发酵产物；和任选地，(b)回收发酵产物。在一种优选方法中，发酵液还可以包含葡萄糖原料。

附图说明

附图 1.啤酒酵母转化子在含有以 25mM 半乳糖和 100mM 木糖作为碳源的培养基中的生长曲线。转化子 pYes 含有无插入片段的酵母表达载体。转化子 14.3、16.2 和 16.2.2 用含有 *Piromyces sp.* E2 木糖异构酶编码序列的 pYES 载体转化。

具体实施方式

实施例 1

克隆 *Piromyces* 木聚糖酶异构酶和木酮糖激酶的 cDNA

厌氧真菌 *Piromyces sp.* E2(ATCC 76762)分离自印度象排泄物，在 39

℃、N_2/CO_2(80%/20%)条件下，在添加了各种碳源的 M2 培养基(24)中厌氧生长。所用的碳源为微晶纤维素(微晶的纤维素类型 PH105，Serva，德国)、果糖和木糖(均为 0.5%，w/v)。在生长停止后，通过氢的生成来判断，细胞通过离心(15,000 x g，4℃，15min)或通过尼龙纱(孔径 30μm)来收获细胞。

无细胞提取物的制备

真菌细胞用去离子水洗涤来去除培养基成分。通过在液氮中冷冻细胞接着在研钵中用玻璃珠(直径 0.10-0.11mm)研磨来制备无细胞提取物。将 Tris/HCl 缓冲液(100mM，pH7.0)加入到粉末中(1：1，w/v)，加热 15min 后离心该悬浮液(18,000 x g，4℃，15min)。澄清的上清液用作胞内酶的来源。

酶分析

在 37℃下，在含有 50mM 磷酸缓冲液(pH7.0)、10mM 木糖、10mM $MgCl_2$ 和适量无细胞提取物的反应混合液中分析木糖异构酶的活性。生成的木酮糖的量通过半胱氨酸-咔唑方法(9)确定。木酮糖激酶和木糖还原酶活性按照 Witteveen 等的描述(28)进行分析。一个活性单位定义为在分析条件下每分钟生成 1nmol 木酮糖的酶含量。生成的木酮糖通过 Dische 和 Borenfreund 的方法(Dische 和 Borenfreund，1995，生物学化学杂志 192：583-587)或者通过采用 Biorad HPX-87N 柱在 80℃下操作和用 0.01M Na_2HPO_4 作为洗脱液以 0.6ml/min 洗脱的 HPLC 来确定。用折光率检测器在 60℃的内部温度下检测木糖和木酮糖。

特异性活性表示为单位/mg 蛋白。可以以牛 γ 球蛋白作为标准物用 Bio-Rad 蛋白试剂(Bio-Rad 实验室，CA，美国)来测定蛋白。

Piromyces sp. E2 cDNA 文库的随机测序

采用先前描述的构建在载体 λZAP II 上的 cDNA 文库。通过用 ExAssist 辅助噬菌体(Stratagene，La Jolla，CA，美国)大片段切除，将该文库的等分试样被转移到 pBluescript Sk-克隆上。随机筛选的克隆用 M13

反向引物测序来获得 5'端的序列。不完整的 cNDAs 用于合成探针，该探针用于再筛选文库。为了获得全长序列，从 PUC18 生成亚克隆。用具有 dRhodamine 终止循环测序简便反应的 DNA 测序试剂盒(Perkin-Elmer Applied Biosystems)的 ABI prism 310 自动测序仪进行测序。

结果

对从厌氧真菌 *Piromyces sp.* E2 的 cDNA 文库随机选择的亚克隆进行测序，得到序列分别与木糖异构酶和 D-木酮糖激酶具有极高同源性的两个克隆(pH97 和 pAK44)。详细地分析这两个克隆。

克隆 pH97 不含有完整的 ORF，因而用根据克隆 pH97 的序列信息设计的探针重新筛选 cDNA 文库。这样得到具有 1669bp 插入片段的克隆 pR3。可以确定编码与木糖异构酶极其相似的 437 个氨基酸的蛋白质的 ORF。虽然 5'为转录区仅含有 4bp，推定的起始甲硫氨酸残基与已知木糖异构酶序列的排列十分吻合。3'非转录的区域长度为 351bp，具有高 AT 含量，这厌氧真菌的典型特征。ORF 含有对于与底物反应(催化的三联体 His102、Asp105、Asp340 和 Lys235)和镁的结合(Glu232)(14，26)十分重要的氨基酸序列。此外，存在两个作为木糖异构酶的典型模板(残基 185-194 和 230-237)(20)。*Piromyces sp.* E2 木糖异构酶(XylA)与流感嗜血杆菌(*Haemophilus influenza*) (52%相同性，68%相似性)和大麦属(49%相同性，67%相似性)的酶的同源性最高。从 cDNA 推导的多肽对应 49.395Da 的分子量和 5.2 的计算 pI。

第二个克隆 pAK44 具有 2041bp 插入片段，含有编码 494 个氨基酸的蛋白质的 ORF，蛋白质的分子量为 53.158Da，pI 为 5.0。第一个甲硫氨酸之前有 111bp 的 5'端非转录区，而 3'端非转录区包含 445bp。这两个区域都富含 AT。BLAST 和 FASTA 检索表明与木酮糖激酶具有高相似性。这两个由 Rodriguez-Pefia 等(22)定义的磷酸一致区域被发现位于局部排列的位点 6-23 和 254-270 上。而且，在 Prosite 数据库中描述的糖激酶的该家族的特征也被确定(131-145 和 351-372)。*Piromyces sp.* E2 木酮糖激

酶(XylB)与流感嗜血杆菌 XylB 蛋白的同源性最高(46%相同性，64%相似性)。

实施例 2

构建酵母表达载体

在啤酒酵母中表达 *Piromyces sp.* E2 的木糖异构酶

Piromyces sp. E2 的 cDNA 和 *pfu* 聚合酶(Stratagene)用于 PCR 反应中。用木糖异构酶基因的 5'和 3'端的序列设计引物，还含有 Sfi Ⅰ和 Xba Ⅰ限制性位点。PCR 产物被克隆到 pPICZα 载体中(Invitrogen，Carlsbad，CA，美国)。为了得到木糖异构酶基因，用 EcoR Ⅰ和 Xba Ⅰ消化 pPICZα 载体。将消化产物连接到 pYes2 载体上(Invitrogen)。具有木糖异构酶基因的 pYes2 质粒被转化到啤酒酵母中(stam BJ1991，由 Beth Jones，UvA 赠送)。该菌株的表型为：matα、leu2、trp1、ura3-251、prb1-1122 和 pep4-3。

将转化子放到 SC 平板上(0.67%YNB 溶液+0.05%L-Leu+0.05%L-Trp+2%葡萄糖+2%琼脂)。未被转化的细胞不能在这些平板上生长。

诱导

被转化的啤酒酵母细胞在 25℃下、在葡萄糖培养基中生长 72h(棉子糖可以用作葡萄糖的替代物)。回收细胞，悬浮在用半乳糖替代葡萄糖的 SC 培养基中。诱导 8h 后，回收细胞，用玻璃珠(直径 0.10-0.11mm)和“破裂缓冲液”(breaking buffer) (50mM 磷酸缓冲液+5%甘油+蛋白酶抑制剂)溶解细胞。溶解后离心混合物(18,000 x g，4℃，15min)。澄清的上清液用上述方法(实施例 1)确定木糖异构酶活性。在 37℃下测定到 10U/mg 蛋白的活性。

实施例 3

被转化的酵母菌株在木糖上生长

培养基成分

啤酒酵母在具有下列成分的SC培养基中生长：0.67%(w/v)酵母氮基(yeast nitrogen base)；0.01%(w/v)L-色氨酸；0.01%(w/v)L-亮氨酸以及葡萄糖或半乳糖或木糖或这些底物的组合(见下文)。对于琼脂平板，培养基中添加2%(w/v)细菌琼脂。

生长试验

啤酒酵母菌株BJ1991(表型：matα、leu2、trp1、ura3-251、prb1-1122、pep4-3)用无插入片段的pYes2转化，三种被选择的转化子(16.2.1；16.2.2和14.3)含有具有*Piromyces sp.* E2木糖异构酶基因的pYes2，在具有10mM葡萄糖作为碳源的SC-琼脂平板上生长。当使克隆可视化时，单个克隆用于接种具有100mM木糖和25mM半乳糖作为碳源的液体SVC-培养基。通过测定600nm下在LKB Ultrospec K 分光光度计中光密度的升高来监控生长。

结果

生长试验的结果汇集在附图 1 中。用无插入片段的 pYes2 转化的BJ1991 菌株培养物的OD_{600}升高达到80h。在此之后观察到逐步下降。这是由于酵母细胞的聚集造成，常常在生长结束时出现。这三种转化子的培养物在80h后并不停止生长，继续升高至少到150h。

实施例 4

为 *Piromyces sp.* E2 木糖异构酶在啤酒酵母中组成型地表达而构建新的改进的酵母表达载体

pPICZα 载体含有编码木糖异构酶的 *Piromyces sp.* E2 的基因，作为含有 $Vent_R$ DNA 聚合酶(New England Biolab)的 PCR 的模板。用编码木糖异构酶的基因的5'和3'端序列设计引物，包括 EcoR Ⅰ 和 Spe Ⅰ 位点。此外，设计引物来去除 pPICZα 构建体上的 Xba Ⅰ 位点，替换为终止密码子(TAA)。最终产物被设计成保存原始开放阅读框，而不增加 pPICZα 构建

体上的氨基酸(组氨酸和 c-真菌标记(c-Myc tags))。用 EcoR Ⅰ和 Spe Ⅰ剪切 PCR 产物。终产物被克隆到由 pYES2(Invitrogen)产生的载体上。在该载体中，pYES2 上的 *GAL1* 启动子用 *TPI1* 替换来确保木糖异构酶的组成型表达，从而消除培养基中对半乳糖的需求。*TPI1* 启动子从改进的质粒 pYX012(R&D system)上克隆得到。该启动子被剪切成 Nhe Ⅰ-EcoR Ⅰ片段。*TPI1* 启动子和木糖异构酶的编码基因的 PCR 产物都被连接到用 Spe Ⅰ和 Xba Ⅰ剪切的 pYES2 上。用该质粒转化啤酒酵母菌株 CEN.PK113-5D(由法兰克福的 Peter Kõtter 赠送)。该菌株的表型为：MatA ura3-52。在以 2%葡萄糖作为碳源的矿物质培养基平板上(Verduyn 等：苯甲酸对酵母中的代谢通量的影响：对呼吸和乙醇发酵的调控的持续培养研究 (1992) 酵母 (Yeast) 8(7)：501-17)筛选转化子。未被转化的细胞不能在这些平板上生长。转化子在缺少碳源的衡化培养中依赖葡萄糖/木糖的混合物生长。在这些条件下生长的转化子表现出很高的木糖异构酶活性(30℃下 800 单位/mg)，酶活性根据由 Kersters-Hildersson 等开发的特异性的酶分析方法(用 D-山梨醇脱氢酶酶促分析 D-木糖异构酶的动力学特征。酶微生物技术(Enz. Microb. Technol.) 9(1987)145-148)测定。木糖异构酶在被转化的啤酒酵母菌株的无细胞提取物中的体外活性依赖于二价阳离子(Mg2+或 Co2+)，测定到约 20mM 的相对低的木糖 Km 值。

序列表

<110> 皇家奈达尔科股份有限公司

<120> 戊糖的发酵

<130> piromyces 木糖异构酶

<140> BO 44829
<141> 2001-12-31

<160> 4

<170> PatentIn Ver. 2.1

<210> 1
<211> 437
<212> PRT
<213> Piromyces sp.

<400> 1

```
Met Ala Lys Glu Tyr Phe Pro Gln Ile Gln Lys Ile Lys Phe Glu Gly
  1               5                  10                  15

Lys Asp Ser Lys Asn Pro Leu Ala Phe His Tyr Tyr Asp Ala Glu Lys
             20                  25                  30

Glu Val Met Gly Lys Lys Met Lys Asp Trp Leu Arg Phe Ala Met Ala
         35                  40                  45

Trp Trp His Thr Leu Cys Ala Glu Gly Ala Asp Gln Phe Gly Gly Gly
     50                  55                  60

Thr Lys Ser Phe Pro Trp Asn Glu Gly Thr Asp Ala Ile Glu Ile Ala
 65                  70                  75                  80

Lys Gln Lys Val Asp Ala Gly Phe Glu Ile Met Gln Lys Leu Gly Ile
                 85                  90                  95

Pro Tyr Tyr Cys Phe His Asp Val Asp Leu Val Ser Glu Gly Asn Ser
            100                 105                 110

Ile Glu Glu Tyr Glu Ser Asn Leu Lys Ala Val Val Ala Tyr Leu Lys
        115                 120                 125

Glu Lys Gln Lys Glu Thr Gly Ile Lys Leu Leu Trp Ser Thr Ala Asn
    130                 135                 140

Val Phe Gly His Lys Arg Tyr Met Asn Gly Ala Ser Thr Asn Pro Asp
145                 150                 155                 160

Phe Asp Val Val Ala Arg Ala Ile Val Gln Ile Lys Asn Ala Ile Asp
                165                 170                 175

Ala Gly Ile Glu Leu Gly Ala Glu Asn Tyr Val Phe Trp Gly Gly Arg
            180                 185                 190

Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys Glu
        195                 200                 205

His Met Ala Thr Met Leu Thr Met Ala Arg Asp Tyr Ala Arg Ser Lys
    210                 215                 220

Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro Thr
225                 230                 235                 240
```

```
Lys His Gln Tyr Asp Val Asp Thr Glu Thr Ala Ile Gly Phe Leu Lys
                245                 250                 255

Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn His
            260                 265                 270

Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala Val
        275                 280                 285

Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr Gln
    290                 295                 300

Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Gln Tyr Glu Leu Val
305                 310                 315                 320

Gln Ala Trp Met Glu Ile Ile Arg Gly Gly Gly Phe Val Thr Gly Gly
                325                 330                 335

Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu Asp
            340                 345                 350

Ile Ile Ile Ala His Val Ser Gly Met Asp Ala Met Ala Arg Ala Leu
        355                 360                 365

Glu Asn Ala Ala Lys Leu Leu Gln Glu Ser Pro Tyr Thr Lys Met Lys
    370                 375                 380

Lys Glu Arg Tyr Ala Ser Phe Asp Ser Gly Ile Gly Lys Asp Phe Glu
385                 390                 395                 400

Asp Gly Lys Leu Thr Leu Glu Gln Val Tyr Glu Tyr Gly Lys Lys Asn
                405                 410                 415

Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala Ile
            420                 425                 430

Val Ala Met Tyr Gln
        435
```

<210> 2
<211> 1669
<212> DNA
<213> Piromyces sp.

<400> 2

```
gtaaatggct aaggaatatt tcccacaaat tcaaaagatt aagttcgaag gtaaggattc 60
taagaatcca ttagccttcc actactacga tgctgaaaag gaagtcatgg gtaagaaaat 120
gaaggattgg ttacgtttcg ccatggcctg gtggcacact ctttgcgccg aaggtgctga 180
ccaattcggt ggaggtacaa agtctttccc atggaacgaa ggtactgatg ctattgaaat 240
tgccaagcaa aaggttgatg ctggtttcga aatcatgcaa aagcttggta ttccatacta 300
ctgtttccac gatgttgatc ttgtttccga aggtaactct attgaagaat acgaatccaa 360
ccttaaggct gtcgttgctt acctcaagga aaagcaaaag gaaaccggta ttaagcttct 420
ctggagtact gctaacgtct tcggtcacaa gcgttacatg aacggtgcct ccactaaccc 480
agactttgat gttgtcgccc gtgctattgt tcaaattaag aacgccatag acgccggtat 540
tgaacttggt gctgaaaact acgtcttctg gggtggtcgt gaaggttaca tgagtctcct 600
taacactgac caaaagcgtg aaaaggaaca catggccact atgcttacca tggctcgtga 660
ctacgctcgt tccaagggat tcaagggtac tttcctcatt gaaccaaagc caatggaacc 720
aaccaagcac caatacgatg ttgacactga aaccgctatt ggtttcctta aggcccacaa 780
cttagacaag gacttcaagg tcaacattga agttaaccac gctactcttg ctggtcacac 840
tttcgaacac gaacttgcct gtgctgttga tgctggtatg ctcggttcca ttgatgctaa 900
ccgtggtgac taccaaaacg gttgggatac tgatcaattc ccaattgatc aatacgaact 960
cgtccaagct tggatggaaa tcatccgtgg tggtggtttc gttactggtg gtaccaactt 1020
cgatgccaag actcgtcgta actctactga cctcgaagac atcatcattg cccacgtttc 1080
```

```
tggtatggat gctatggctc gtgctcttga aaacgctgcc aagctcctcc aagaatctcc 1140
atacaccaag atgaagaagg aacgttacgc ttccttcgac agtggtattg gtaaggactt 1200
tgaagatggt aagctcaccc tcgaacaagt ttacgaatac ggtaagaaga acggtgaacc 1260
aaagcaaact tctggtaagc aagaactcta cgaagctatt gttgccatgt accaataagt 1320
taatcgtagt taaattggta aaataattgt aaaatcaata aacttgtcaa tcctccaatc 1380
aagtttaaaa gatcctatct ctgtactaat taaatatagt acaaaaaaaa atgtataaac 1440
aaaaaaaagt ctaaaagacg gaagaattta atttagggaa aaaataaaaa taataataaa 1500
caatagataa atcctttata ttaggaaaat gtcccattgt attattttca tttctactaa 1560
aaaagaaagt aaataaaaca caagaggaaa ttttcccttt ttttttttt tgtaataaat 1620
tttatgcaaa tataaatata aataaaataa taaaaaaaaa aaaaaaaaa             1669
```

<210> 3
<211> 494
<212> PRT
<213> Piromyces sp.

<400> 3

```
Met Lys Thr Val Ala Gly Ile Asp Leu Gly Thr Gln Ser Met Lys Val
  1               5                  10                  15

Val Ile Tyr Asp Tyr Glu Lys Lys Glu Ile Ile Glu Ser Ala Ser Cys
            20                  25                  30

Pro Met Glu Leu Ile Ser Glu Ser Asp Gly Thr Arg Glu Gln Thr Thr
        35                  40                  45

Glu Trp Phe Asp Lys Gly Leu Glu Val Cys Phe Gly Lys Leu Ser Ala
    50                  55                  60

Asp Asn Lys Lys Thr Ile Glu Ala Ile Gly Ile Ser Gly Gln Leu His
65                  70                  75                  80

Gly Phe Val Pro Leu Asp Ala Asn Gly Lys Ala Leu Tyr Asn Ile Lys
                85                  90                  95

Leu Trp Cys Asp Thr Ala Thr Val Glu Glu Cys Lys Ile Ile Thr Asp
            100                 105                 110

Ala Ala Gly Gly Asp Lys Ala Val Ile Asp Ala Leu Gly Asn Leu Met
        115                 120                 125

Leu Thr Gly Phe Thr Ala Pro Lys Ile Leu Trp Leu Lys Arg Asn Lys
    130                 135                 140

Pro Glu Ala Phe Ala Asn Leu Lys Tyr Ile Met Leu Pro His Asp Tyr
145                 150                 155                 160

Leu Asn Trp Lys Leu Thr Gly Asp Tyr Val Met Glu Tyr Gly Asp Ala
                165                 170                 175

Ser Gly Thr Ala Leu Phe Asp Ser Lys Asn Arg Cys Trp Ser Lys Lys
            180                 185                 190

Ile Cys Asp Ile Ile Asp Pro Lys Leu Leu Asp Leu Leu Pro Lys Leu
        195                 200                 205

Ile Glu Pro Ser Ala Pro Ala Gly Lys Val Asn Asp Glu Ala Ala Lys
    210                 215                 220

Ala Tyr Gly Ile Pro Ala Gly Ile Pro Val Ser Ala Gly Gly Gly Asp
225                 230                 235                 240

Asn Met Met Gly Ala Val Gly Thr Gly Thr Val Ala Asp Gly Phe Leu
                245                 250                 255
```

```
Thr Met Ser Met Gly Thr Ser Gly Thr Leu Tyr Gly Tyr Ser Asp Lys
            260                 265                 270

Pro Ile Ser Asp Pro Ala Asn Gly Leu Ser Gly Phe Cys Ser Ser Thr
        275                 280                 285

Gly Gly Trp Leu Pro Leu Leu Cys Thr Met Asn Cys Thr Val Ala Thr
    290                 295                 300

Glu Phe Val Arg Asn Leu Phe Gln Met Asp Ile Lys Glu Leu Asn Val
305                 310                 315                 320

Glu Ala Ala Lys Ser Pro Cys Gly Ser Glu Gly Val Leu Val Ile Pro
                325                 330                 335

Phe Phe Asn Gly Glu Arg Thr Pro Asn Leu Pro Asn Gly Arg Ala Ser
            340                 345                 350

Ile Thr Gly Leu Thr Ser Ala Asn Thr Ser Arg Ala Asn Ile Ala Arg
        355                 360                 365

Ala Ser Phe Glu Ser Ala Val Phe Ala Met Arg Gly Gly Leu Asp Ala
    370                 375                 380

Phe Arg Lys Leu Gly Phe Gln Pro Lys Glu Ile Arg Leu Ile Gly Gly
385                 390                 395                 400

Gly Ser Lys Ser Asp Leu Trp Arg Gln Ile Ala Ala Asp Ile Met Asn
                405                 410                 415

Leu Pro Ile Arg Val Pro Leu Leu Glu Glu Ala Ala Ala Leu Gly Gly
            420                 425                 430

Ala Val Gln Ala Leu Trp Cys Leu Lys Asn Gln Ser Gly Lys Cys Asp
        435                 440                 445

Ile Val Glu Leu Cys Lys Glu His Ile Lys Ile Asp Glu Ser Lys Asn
    450                 455                 460

Ala Asn Pro Ile Ala Glu Asn Val Ala Val Tyr Asp Lys Ala Tyr Asp
465                 470                 475                 480

Glu Tyr Cys Lys Val Val Asn Thr Leu Ser Pro Leu Tyr Ala
                485                 490
```

<210> 4
<211> 2041
<212> DNA
<213> Piromyces sp.

<400> 4

```
attatataaa ataactttaa ataaaacaat ttttatttgt ttatttaatt attcaaaaaa 60
aattaaagta aaagaaaaat aatacagtag aacaatagta ataatatcaa aatgaagact 120
gttgctggta ttgatcttgg aactcaaagt atgaaagtcg ttatttacga ctatgaaaag 180
aaagaaatta ttgaaagtgc tagctgtcca atggaattga tttccgaaag tgacggtacc 240
cgtgaacaaa ccactgaatg gtttgacaag ggtcttgaag tttgttttgg taagcttagt 300
gctgataaca aaaagactat tgaagctatt ggtatttctg gtcaattaca cggttttgtt 360
cctcttgatg ctaacggtaa ggctttatac aacatcaaac tttggtgtga tactgctacc 420
gttgaagaat gtaagattat cactgatgct gccggtggtg acaaggctgt tattgatgcc 480
cttggtaacc ttatgctcac cggtttcacc gctccaaaga tcctctggct caagcgcaac 540
aagccagaag ctttcgctaa cttaaagtac attatgcttc cacacgatta cttaaactgg 600
aagcttactg gtgattacgt tatggaatac ggtgatgcct ctggtaccgc tctcttcgat 660
tctaagaacc gttgctggtc taagaagatt tgcgatatca ttgacccaaa acttttagat 720
```

```
ttacttccaa agttaattga accaagcgct ccagctggta aggttaatga tgaagccgct 780
aaggcttacg gtattccagc cggtattcca gtttccgctg gtggtggtga taacatgatg 840
ggtgctgttg gtactggtac tgttgctgat ggtttcctta ccatgtctat gggtacttct 900
ggtactcttt acggttacag tgacaagcca attagtgacc cagctaatgg tttaagtggt 960
ttctgttctt ctactggtgg atggcttcca ttactttgta ctatgaactg tactgttgcc 1020
actgaattcg ttcgtaacct cttccaaatg gatattaagg aacttaatgt tgaagctgcc 1080
aagtctccat gtggtagtga aggtgtttta gttattccat tcttcaatgg tgaaagaact 1140
ccaaacttac caaacggtcg tgctagtatt actggtctta cttctgctaa caccagccgt 1200
gctaacattg ctcgtgctag tttcgaatcc gccgttttcg ctatgcgtgg tggtttagat 1260
gctttccgta agttaggttt ccaaccaaag gaaattcgtc ttattggtgg tggttctaag 1320
tctgatctct ggagacaaat tgccgctgat atcatgaacc ttccaatcag agttccactt 1380
ttagaagaag ctgctgctct tggtggtgct gttcaagctt tatggtgtct taagaaccaa 1440
tctggtaagt gtgatattgt tgaactttgc aaagaacaca ttaagattga tgaatctaag 1500
aatgctaacc caattgccga aaatgttgct gtttacgaca aggcttacga tgaatactgc 1560
aaggttgtaa atactctttc tccattatat gcttaaattg ccaatgtaaa aaaaaatata 1620
atgccatata attgccttgt caatacactg ttcatgttca tataatcata ggacattgaa 1680
tttacaaggt ttatacaatt aatatctatt atcatattat tatacagcat ttcattttct 1740
aagattagac gaaacaattc ttggttcctt gcaatataca aaatttacat gaatttttag 1800
aatagtctcg tatttatgcc caataatcag gaaaattacc taatgctgga ttcttgttaa 1860
taaaaacaaa ataaataaat taaataaaca aataaaaatt ataagtaaat ataaatatat 1920
aagtaatata aaaaaaaagt aaataaataa ataaataaat aaaaattttt tgcaaatata 1980
taaataaata aataaaatat aaaaataatt tagcaaataa attaaaaaaa aaaaaaaaaa 2040
a                                                                  2041
```

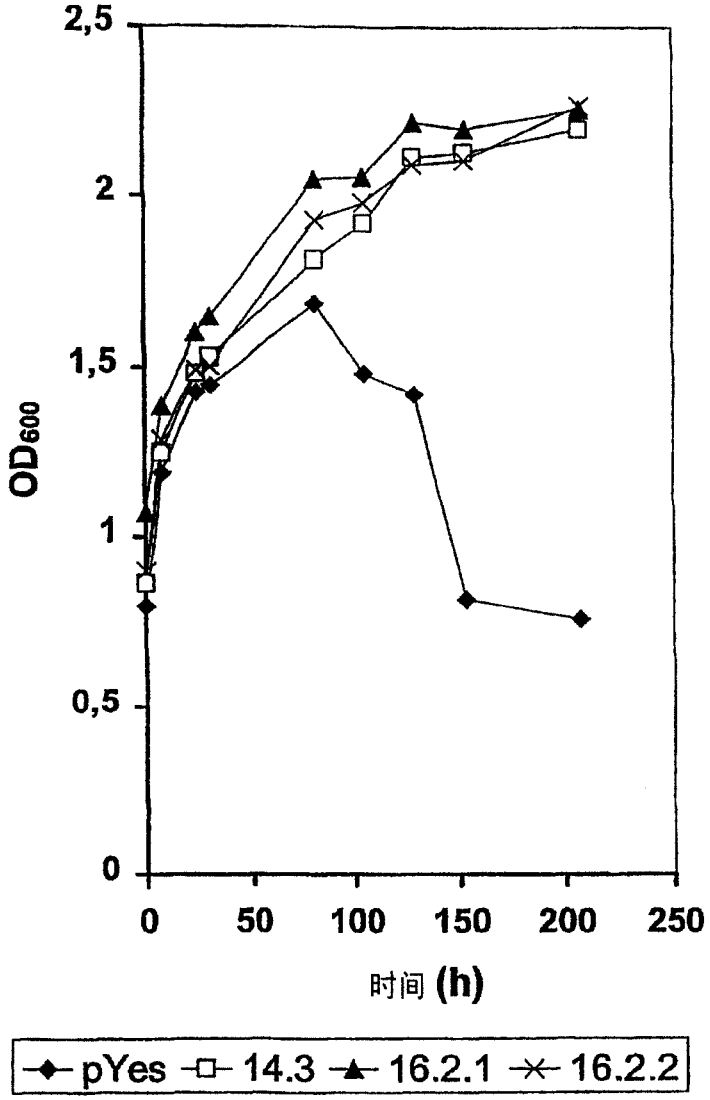

图1