



(12) 发明专利申请

(10) 申请公布号 CN 102725306 A

(43) 申请公布日 2012. 10. 10

(21) 申请号 201080061191. 2

E · 巴塔各里亚

(22) 申请日 2010. 11. 12

D · A · 彼瓦杜阿尔特

(30) 优先权数据

PCT/NL2009/050683 2009. 11. 12 NL

(74) 专利代理机构 北京派特恩知识产权代理事
务所(普通合伙) 11270

代理人 武晨燕 张颖玲

(85) PCT申请进入国家阶段日

2012. 07. 11

(51) Int. Cl.

C07K 14/38 (2006. 01)

(86) PCT申请的申请数据

PCT/NL2010/050755 2010. 11. 12

(87) PCT申请的公布数据

W02011/059329 EN 2011. 05. 19

(71) 申请人 乌得勒支大学控股有限公司

地址 荷兰乌得勒支

申请人 高等教育、科学研究及疾病护理协会
荷兰科学技术基金会
莱顿大学

(72) 发明人 帕特丽夏·安·凡凯克

罗纳德·彼得·德弗里斯

雅各布·菲瑟 A·F·J·拉姆

C·A·M·J·J·范德亨德尔

权利要求书 2 页 说明书 13 页 附图 9 页

(54) 发明名称

戊糖转运蛋白及其用途

(57) 摘要

本发明涉及从粗碳源生产生物燃料、蛋白、肽和其他增值化合物。本发明人鉴定编码新戊糖转运蛋白,特别是L-阿拉伯糖和/或D-木糖的转运蛋白的基因。通过 xlnR 和 araR 调控黑曲霉 (*Aspergillus niger*) 基因有助于鉴定这些基因和它们的底物特异性。本发明提供新戊糖转运蛋白以及它们的编码核酸。本发明还提供(过量)表达转运蛋白的宿主细胞及其工业应用,例如生物燃料生产。

1. 一种将至少部分木质纤维素粗碳源转化为增值化合物的方法,所述方法包括在所述粗碳源的存在下培养宿主细胞,所述宿主细胞表达编码多肽的核酸序列,所述多肽选自下列物质组成的组:

a) 多肽,其具有的氨基酸序列与图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列至少 80%相同,并且表现出体外和 / 或体内戊糖转运活性;

b) 多肽,其与图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列相同;以及

c) 多肽片段,该多肽片段为如 a) 或 b) 所定义的多肽的片段,包括图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段,并且表现出体外和 / 或体内戊糖转运活性。

2. 根据权利要求 1 所述的方法,其中所述宿主细胞表达图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示序列的至少 200 个,优选 300 个连续氨基酸的片段。

3. 根据权利要求 1 或 2 所述的方法,其中所述粗碳源包括果胶和 / 或半纤维素。

4. 根据任何权利要求 3 所述的方法,其中所述粗碳源包括阿拉伯聚糖、阿拉伯半乳糖、木聚糖和 / 或木葡聚糖,优选阿拉伯糖和 / 或木糖,更优选 L-阿拉伯糖和 / 或 D-木糖。

5. 根据前述权利要求中任一项所述的方法,其中所述粗碳源选自植物生物质、草本原料、农业残余物、林业残余物、城市固体废物、废纸、以及纸浆和造纸厂残余物所组成的组。

6. 根据权利要求 1-5 中任一项所述的方法,其中所述增值化合物为生物燃料,优选生物乙醇。

7. 根据前述权利要求中任一项所述的方法,其中所述宿主细胞还包括至少一种核酸分子,所述核酸分子编码参与阿拉伯糖和 / 或木糖代谢的酶。

8. 根据权利要求 7 所述的方法,其中所述宿主细胞包括至少一种编码酶的基因,所述酶选自自由 L-核酮糖激酶、L-核酮糖 -5-P 4-差向异构酶、L-阿拉伯糖 -异构酶、大肠杆菌 araBAD 操纵子编码的酶、NAD(P)H-依赖性木糖还原酶 (XR)、NAD⁺-依赖性木糖醇脱氢酶 (XDH)、木糖异构酶 (XI) 以及木酮糖激酶所组成的组。

9. 一种重组宿主细胞,包括分离的核酸序列,所述分离的核酸序列编码如权利要求 1 或 2 所定义的多肽,并且其中所述宿主细胞还包括至少一种核酸分子,所述核酸分子编码参与至少一种戊糖,优选阿拉伯糖和 / 或木糖代谢的酶。

10. 根据权利要求 9 所述的宿主细胞,其中所述核酸序列与图 1B、2B、3B、4B、5B、6B、7B 或 8B 所示的核酸序列至少 90%,优选至少 95%相同。

11. 根据权利要求 10 所述的宿主细胞,其中所述核酸序列与图 1B、2B、3B 或 4B 所示的核酸序列至少 95%相同。

12. 根据权利要求 9-11 中任一项所述的宿主细胞,其中,包括至少一种编码酶的基因,所述酶选自自由 L-核酮糖激酶、L-核酮糖 -5-P 4-差向异构酶、L-阿拉伯糖 -异构酶、大肠杆菌 araBAD 操纵子编码的酶、NAD(P)H-依赖性木糖还原酶 (XR)、NAD⁺-依赖性木糖醇脱氢酶 (XDH)、木糖异构酶 (XI) 以及木酮糖激酶所组成的组。

13. 根据权利要求 9-12 中任一项所述的宿主细胞,其中所述宿主细胞为真菌,优选选自自由曲霉属、木霉属、酵母属、*Chrysosporium lucknowense* C1、克鲁维酵母属、汉逊酵母

属、毕赤酵母属以及耶罗威亚酵母属所组成的组。

14. 根据权利要求 13 所述的宿主细胞,其中,所述宿主细胞为酿酒酵母或黑曲霉宿主细胞。

15. 多肽在提高宿主细胞的至少一种戊糖的吸收和 / 或利用中的用途,所述多肽选自下列物质组成的组:

a) 多肽,其具有的氨基酸序列与图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列至少 80% 相同,并且表现出体外和 / 或体内戊糖转运活性;

b) 多肽,其与图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列相同;以及

c) 多肽片段,该多肽片段为如 a) 或 b) 所定义的多肽的片段,其包括图 1A、2A、3A、4A、5A、6A、7A 或 8A,优选图 1A、2A、3A 或 4A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段,并且表现出体外和 / 或体内戊糖转运活性。

16. 根据权利要求 15 所述的用途,其中所述戊糖为阿拉伯糖,优选 L-阿拉伯糖;和 / 或木糖,优选 D-木糖。

17. 根据权利要求 15 或 16 所述的用途,其中所述宿主细胞为酵母细胞或丝状真菌。

18. 根据权利要求 16 所述的用途,其中所述宿主细胞选自由曲霉属、木霉属、酵母属、*Chrysosporium lucknowense* C1、克鲁维酵母属、汉逊酵母属、毕赤酵母属以及耶罗威亚酵母属所组成的组。

19. 根据权利要求 18 所述的用途,其中所述宿主细胞为酿酒酵母或黑曲霉宿主细胞。

戊糖转运蛋白及其用途

技术领域

[0001] 本发明涉及从粗碳源生产生物燃料和其他增值化合物。特别地,本发明涉及通过宿主细胞将至少部分木质纤维素粗碳源转化为增值化合物的方法,以及用于这类方法的宿主细胞。

背景技术

[0002] 粗碳源(主要是植物生物质)的利用越来越受到业界的关注,不仅关于建立的发酵,而且还关于诸如生物乙醇等新产物。从可再生原料生产乙醇是对使用化石燃料的可持续替代方式。用于运输燃料的生物乙醇可以通过木质纤维素原料如农业和林业废弃物或者能源作物的发酵以可持续的方式生产。其他增值化合物包括蛋白,如酶和肽。例如甜菜生产糖所剩余的糖蜜可用于生产蛋白和肽。参见 Siqueira 等人,(2008) *Bioresour. Technol.* 99(17):8156-63; Alriksson 等人,(2009) *Appl. Environ. Microbiol.* 75(8):2366-74; Peixoto-Nogueira Sde C 等人,(2009) *J. Ind. Microbiol. Technol.* 36(1):149-55。He 等人公开通过基因修饰的酿酒酵母 (*Saccharomyces cerevisiae*) 从糖蜜生产麦角固醇(2007, *Appl. Microbiol. Biotechnol.* 75:55-60); Ghazi 等人描述的甜菜糖浆和糖蜜作为低成本原料用于低聚果糖的酶促生产(2006, *J. Agric. Food Chem.* 54(8):2964-8)。

[0003] 完全的底物利用、抑制剂耐受性以及乙醇生产率是选择微生物发酵的重要方面。已知有酵母酿酒酵母 (*S. cerevisiae*) 菌株满足后两个条件。然而,代谢工程需要获得能够发酵半纤维素中最丰富的戊糖(如 L-阿拉伯糖和 D-木糖)的菌株。虽然 L-阿拉伯糖比 D-木糖小的分数存在,但是为了总体经济性,也需要将其高效地转化为乙醇。此外,L-阿拉伯糖转化为乙醇会降低与酵母竞争的污染生物所用的碳源。

[0004] 因此,从例如麦秸、玉米秸秆、甘蔗渣和木材水解产物可持续生产生物燃料乙醇需要己糖和戊糖两部分的发酵。人们正致力于将木质纤维素水解产物发酵为乙醇。虽然已设计的修饰的酿酒酵母菌株,能够在 D-木糖上生长并发酵 D-木糖,但是它们的生长速度仍然较慢。先前,通过从树干毕赤酵母 (*Pichia stipitis*) 异源表达 NAD(P)H-依赖性木糖还原酶(XR)和 NAD⁺-依赖性木糖醇脱氢酶(XDH)实现了木糖的利用。然而,这两种酶之间的辅因子不平衡导致较低的乙醇收率和产率。由于低酶表达以及木糖醇对 XI 抑制,在试图表达细菌木糖异构酶(XI)基因中也获得有限的结果。木糖利用还可受限于转运、低木酮糖激酶水平以及低水平的戊糖磷酸途径。

[0005] 鉴于目前的重点是用于例如生物燃料生产的戊糖到乙醇的发酵,这里亟需开发改进的宿主酵母菌株。此外,通过酵母以及其他工业丝状真菌(如曲霉(*Aspergillus*)或木霉(*Trichoderma*))提高戊糖的利用,会改善对粗碳底物的发酵,从而扩大了表现出修饰的戊糖转运的宿主细胞的可能应用。在宿主细胞于例如 D-葡萄糖和 D-木糖上共培养的情况下,通常优选使用一种糖。例如,首先使用 D-葡萄糖,然后代谢 D-木糖。因为不同糖的不同生长速度,通常导致双相生长曲线。表达异源转运蛋白(例如戊糖转运蛋白)或者其他

启动子（例如组成型启动子）控制下的同源转运蛋白的菌株可以导致同时利用两种（或更多种）糖（例如 D- 葡萄糖和 D- 木糖）的细胞，从而改善生长。

发明内容

[0006] 因此本发明人着手鉴定编码新戊糖转运蛋白，特别是 L- 阿拉伯糖或 D- 木糖的转运蛋白的基因。据发现通过 xlnR（木糖反应正作用调节子）和 araR（L- 阿拉伯糖反应正作用调节子）调控黑曲霉（*Aspergillus niger*）基因有助于鉴定这些基因和它们的底物特异性。这通过比较来自调节子缺失菌株与转移至相关戊糖的参考菌株的微阵列数据来确定。选择转录水平具有大于 4.5 倍变化的基因进行进一步的生化分析。这样确定了 8 种新的多肽序列及其编码核酸。更特别地，他们确定了转运蛋白 An08g01720（在本文中简称为 1720）、An03g01620（1620）、An11g01100（1100）、An06g00560（0560）、An02g08230（8230）、An07g00780（0780）、An13g02590（2590）以及 An03g02190（2190）。这些蛋白分别由基因 An08g01720、An03g01620、An11g01100、An06g00560、An02g08230、An07g00780、An13g02590 以及 An03g02190 编码。新的转运蛋白的蛋白和 cDNA 序列可见图 1-8。

[0007] 因此本发明涉及一种多肽，其选自由下列物质组成的组：a) 多肽，其具有的氨基酸序列与图 1A、2A、3A、4A、5A、6A、7A 或 8A 所示的氨基酸序列至少 80% 相同，并且表现出体外和 / 或体内戊糖转运活性；b) 多肽，其与图 1A、2A、3A、4A、5A、6A、7A 或 8A 所示的氨基酸序列相同；以及 c) 多肽片段，如 a) 或 b) 所定义的，该多肽片段包括图 1A、2A、3A、4A、5A、6A、7A 或 8A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段，并且表现出体外和 / 或体内戊糖转运活性。优选地，多肽选自由下列物质所组成的组：a) 多肽，其具有的氨基酸序列与图 1A、2A、3A 或 4A 所示的氨基酸序列至少 80% 相同，并且表现出体外和 / 或体内戊糖转运活性；b) 多肽，其与图 1A、2A、3A 或 4A 所示的氨基酸序列相同；以及 c) 多肽片段，如 a) 或 b) 所定义的，该多肽片段包括图 1A、2A、3A 或 4A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段，并且表现出体外和 / 或体内戊糖转运活性。更优选地，所述多肽具有 L- 阿拉伯糖和 / 或 D- 木糖转运活性。在一个实施例中，戊糖转运活性为 L- 阿拉伯糖转运活性。在另一个实施例中，戊糖转运活性为 D- 木糖转运活性。

[0008] 木糖和阿拉伯糖转运蛋白是本领域已知的。例如，Leandro 等人（2006 *Biochem. J.* 395 :543-549）公开了来自酵母中型假丝酵母（*Candida intermedia*）的两种葡萄糖 / 木糖转运蛋白基因。可获得关于改良真菌在戊糖上的生长的各种报道，例如 Bengtsson 等人（2009, *Biotechnol. Biofuels* 5 ;2 :9）、Krahulec 等人（2009 *Biotechnol. J.* 4 : 684-694）以及 Rundquist 等人（2009, *Appl. Microbiol. Biotechnol.* 82 :123-130）的报道。W02008/080505 涉及来自酵母树干毕赤酵母的阿拉伯糖转运蛋白以及其在从生物质生产生物化学品中的应用。W02009/008756 公开了用编码来自酵母的特异性 L- 阿拉伯糖转运蛋白的核酸序列转化的宿主细胞，以及所述宿主细胞在生产生物燃料中的应用。W02007/018442 涉及编码木糖的活性转运蛋白的中型假丝酵母基因以及修饰的表达该基因的酵母细胞。然而，根据本发明的特异性戊糖转运蛋白在本领域中未被描述或提出。

[0009] 在一个实施例中，所述多肽包括图 1A、2A、3A、4A、5A、6A、7A 或 8A 所示序列的至少 200 个或 300 个连续氨基酸的片段。所述片段的特征在于其表现出体外和 / 或体内戊糖转运活性，特别是阿拉伯糖或木糖转运活性。戊糖转运活性可以通过本领域已知的方法容易

地测定。例如,这包括使用放射性标记(例如¹⁴C)的戊糖和/或己糖底物。参见Walsh等人(1994 J. Bacteriol. 176, 953-958)。

[0010] 优选地,所述多肽序列表现出与图1A、2A、3A、4A、5A、6A、7A或8A所示氨基酸序列的至少90%、优选至少95%相同,并且表现出体外和/或体内戊糖转运活性。更优选地,该序列与图1A、2A、3A、4A、5A、6A、7A或8A所述序列之一的96%、97%、98%或99%相同。在具体方面,本发明提供具有与图1A、2A、3A、4A、5A、6A、7A或8A所示氨基酸序列相同序列的多肽。所述多肽可以源自任何微生物,优选(丝状)真菌,如黑曲霉或者其他(工业用)真菌,例如选自其他曲霉属、粗糙脉孢菌(*Neurospora crassa*)、稻瘟病菌(*Magnaporthe grisea*)、木霉属、青霉属(*Penicillium species*)、镰刀菌属(*Fusarium species*)、*Chrysosporium lucknowensis*(C 1)、篮状菌属(*Talaromyces sp.*)、嗜热真菌属(*Thermomyces sp.*)、腐质霉属(*Humicola sp.*)、酵母属(*Saccharomyces sp.*)、克鲁维酵母属(*Kluyveromyces sp.*)、汉逊酵母属(*Hansenula sp.*)、毕赤酵母属(*Pichia sp.*)以及耶罗威亚酵母属(*Yarrowia sp.*)。

[0011] 本发明还涵盖包括一个或多个氨基酸改变(例如缺失、取代和/或插入)的变异或突变多肽,其基本上保持完整的转运活性。在一个实施例中,变异包括一个或多个保守氨基酸取代。当然,激活突变特别引人注目。

[0012] 另一方面涉及融合蛋白,其包括本文所述的作为第一片段的转运蛋白多肽,以及作为第二片段的所关注的异源多肽。所述第一片段可位于所述第二片段的N端或C端。所关注的示例性的多肽包括糖传感蛋白、信号传导途径组分、戊糖转化代谢酶(定位在细胞内)以及靶向序列。在一个实施例中,一起提供有转运肽与序列,所述序列选自质膜靶向序列、增加在质膜的周转(turnover)的序列以及提高菌丝中的正确定位的序列所组成的组。本发明还提供能够选择性结合至本发明的戊糖转运蛋白的抗体或其功能性片段。本领域技术人员能够利用本领域已知的方法产生这样的抗体(片段)。参见例如Ed Harlow, Cold Spring Harbor Laboratory(冷泉港实验室);David Lane, Imperial Cancer Research Fund Laboratories(帝国癌症研究基金会的实验室);编写的“Antibodies :A Laboratory Manual(抗体:实验室手册)”, ISBN 978-087969314-5。

[0013] 本发明的多肽可以利用本文所公开的分离的核酸序列来提供。所述核酸序列通常为cDNA序列。在一个实施例中,本发明提供与An08g01720(图4B)、An03g01620(图3B)或An11g01100(图2B)至少85%同源的阿拉伯糖转运蛋白基因。在另一个实施例中,本发明提供的表现出与An06g00560(图1B)、An02g08230(图7B), An07g00780(图5B)或An13g02590(6B)的至少85%相同的木糖转运蛋白基因。在一个实施例中,所述核酸序列与图1B、2B、3B、4B、5B、6B、7B或8B所示的核酸序列至少85%、优选至少90%相同。更优选地,所述序列至少91%、92%、93%、94%、95%、96%、97%、98%或99%相同。在具体方面,所述核酸序列由图1B、2B、3B、4B、5B、6B、7B或8B所示的核酸序列组成。

[0014] 所述核酸可以是较大的核酸分子(如表达载体)的一部分。优选允许在宿主细胞(例如酵母宿主细胞)中表达所编码的戊糖转运蛋白的表达载体。例如,可以使用pRS系列质粒(Silkorski RS, Hieter P 1989 A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*)、pYES系列质粒(Invitrogen, Carlsbad CA, USA)或pYEX系列质粒(Clontech,

CA, USA)。所述载体可以包括一个或多个常规元素,例如本领域技术人员已知的抗生素抗性标记、转录增强子等。

[0015] 应当理解,本文所提供的戊糖转运蛋白以及它们的编码基因具有许多生物技术和工业应用。同源表达允许通过基因破坏或过量表达来改善黑曲霉 (*A. niger*) 中的戊糖的吸收/利用。高效启动子,例如葡糖淀粉酶、木聚糖酶、甘油醛-三磷酸以及本领域技术人员已知的其他启动子可用作适合在最佳工艺条件下表达的启动子。在除黑曲霉以外的宿主细胞中,以与同源表达相似的方式异源表达转运蛋白基因可以导致增强戊糖(L-阿拉伯糖、D-木糖)的吸收并提高戊糖的利用(例如在生物燃料生产或任何其他类型的应用中)。

[0016] 编码和允许本文所公开的戊糖转运蛋白表达的载体有利地用来改变宿主细胞的戊糖吸收/利用。在一个实施例中,本发明涉及基因工程宿主细胞,其提供有编码多肽的分离的核酸(优选是载体的一部分),所述多肽选自由下列物质所组成的组:a)多肽,其具有的氨基酸序列与图1A、2A、3A、4A、5A、6A、7A或8A所示的氨基酸序列至少80%相同,并且表现出体外和/或体内戊糖转运活性;b)多肽,其与图1A、2A、3A、4A、5A、6A或7A所示的氨基酸序列相同;以及c)多肽片段,如a)或b)所定义的,其包括图1A、2A、3A、4A、5A、6A或7A所示的氨基酸序列中的至少100个连续氨基酸的片段,并且表现出体外和/或体内戊糖转运活性。特别地,本发明提供一种重组宿主细胞,其包括如上文所定义的编码多肽的分离的核酸序列,并且其中所述宿主细胞还包括至少一种核酸分子,所述核酸分子编码参与至少一种戊糖如阿拉伯糖和/或木糖代谢的酶。

[0017] 宿主细胞可以是任何合适的原核或真核生物。在一个实施例中,宿主细胞为真菌宿主细胞。优选地,宿主细胞为酵母细胞和丝状真菌,如酿酒酵母和黑曲霉。所关注的其他宿主细胞包括曲霉属、木霉属、青霉属、镰刀菌属、还有子囊真菌 *Chrysosporium lucknowense* Cl、酵母属、克鲁维酵母属、汉逊酵母属、毕赤酵母属以及耶罗威亚酵母属。其他有用的细胞包括担子菌,例如栓菌属(*Trametes* sp.),如变色栓菌(*T. versicolor*)。

[0018] 在一个具体实施例中,本发明还提供一种真菌宿主细胞,优选丝状真菌,将其进行基因修饰以降低本发明的多肽的至少一种基因编码的表达。这可以通过相应基因的缺失或破坏来实现,例如通过同源重组。

[0019] 宿主细胞可以具有其他额外的组分,如编码蛋白的至少一种核酸分子,所述蛋白参与戊糖的代谢,特别是木糖和/或阿拉伯糖的代谢。实例包括L-核酮糖激酶、L-核酮糖-5-P 4-差向异构酶和L-阿拉伯糖-异构酶。优选地,所述核酸分子编码参与细菌中阿拉伯糖和/或木糖代谢的蛋白。适当使用一种或多种大肠杆菌(*E. coli*)*araBAD*操纵子编码的酶。用于异源表达的其他有用的酶包括来自树干毕赤酵母的(P)H-依赖性木糖还原酶(XR)和NAD+-依赖性木糖醇脱氢酶(XDH),细菌木糖异构酶(XI)基因和木酮糖激酶。

[0020] 本发明还涉及一种将木质纤维素粗碳源转化为增值化合物的方法,所述方法包括,在所述粗碳源的存在下,培养如上文所述的宿主细胞并允许戊糖转运蛋白的表达。特别地,本发明提供一种将至少部分木质纤维素粗碳源转化为增值化合物的方法,所述方法包括,在所述粗碳源的存在下培养宿主细胞,所述宿主细胞表达编码多肽的核酸序列,所述多肽选自由下列物质组成的组:a)多肽,其具有的氨基酸序列与图1A、2A、3A、4A、5A、6A、7A或8A,优选图1A、2A、3A或4A所示的氨基酸序列至少80%相同,并且表现出体外和/或体内戊糖转运活性;b)多肽,其与图1A、2A、3A、4A、5A、6A、7A或8A,优选图1A、2A、3A或4A所示的

氨基酸序列相同；以及 c) 多肽片段，如 a) 或 b) 所定义的，其包括图 1A、2A、3A、4A、5A、6A、7A 或 8A、优选图 1A、2A、3A 或 4A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段，并且表现出体外和 / 或体内戊糖转运活性。

[0021] 优选地，木质纤维素粗碳源包括果胶和 / 或半纤维素。粗碳源通常在己糖（主要是葡萄糖）的存在下，优选包括阿拉伯聚糖、阿拉伯半乳聚糖、木聚糖和 / 或木葡聚糖，优选阿拉伯糖和 / 或木糖，更优选 L-阿拉伯糖和 / 或 D-木糖，或者上述物质的任何组合。阿拉伯聚糖是多糖，其主要是阿拉伯糖的聚合物。木聚糖（CAS 编号：9014-63-5）是用来描述在植物细胞壁和一些藻类中发现的各种高度复杂的多糖的通用术语。木聚糖是由木糖单元组成的多糖。植物细胞壁木聚糖在侧链的数目和组成上极为不同，取决于植物物种，但是许多包括连接至木糖骨架的阿拉伯糖。在一些绿藻的细胞壁中发现的木聚糖，特别是在一些大型管藻目 (macrophytic siphonous genera) 中代替纤维素。相似地，在一些红藻中，木聚糖代替内部纤维状细胞壁层纤维素。木聚糖是常用的饲料原料中最重要抗营养因子之一。木葡聚糖是存在于所有维管植物的初生细胞壁中的半纤维素。在许多双子叶植物中，木葡聚糖是初生细胞壁中最丰富的半纤维素。木葡聚糖具有 $\beta 1 \rightarrow 4$ -连接的葡萄糖残基的骨架，大部分用 1-6 连接的木糖侧链取代。木糖残基通常用半乳糖残基加帽，有时随后通过岩藻糖残基加帽，但是也可以用阿拉伯糖加帽。植物科之间木葡聚糖的具体结构不同。阿拉伯半乳聚糖是由阿拉伯糖和半乳糖组成的混合多糖，根据植物物种的不同阿拉伯糖和半乳糖的比例不同。

[0022] 有利地，所述粗碳源通常为植物生物质或其衍生的组合物，如半纤维素水解产物。例如，木质纤维素粗碳源可以选自由植物生物质、草本原料、农业残余物、林业残余物、城市固体废物、废纸、以及纸浆和造纸厂残余物所组成的组。在一些实施例中，木质纤维素原料为干酒糟 (distiller' s dried grains) 或含有可溶物的干酒糟。在一些实施例中，干酒糟或含有可溶物的干酒糟来源于玉米。

[0023] 本发明的方法和宿主细胞有利地用于包括戊糖与葡萄糖的共同利用的方法中，即葡萄糖和木糖的同时利用，葡萄糖和阿拉伯糖的同时利用，或者葡萄糖、木糖和阿拉伯糖的同时利用。这可以通过使用特定启动子以调节一种或多种新戊糖转运蛋白的表达来实现。可以有利地提供在特定的启动子控制下带有至少一种编码己糖（例如葡萄糖）转运蛋白基因的宿主细胞，所述特定的启动子可以与控制戊糖转运蛋白表达的启动子相同或不同。合适的启动子包括组成型或诱导型启动子（在特定条件下表达）。考虑到可以利用戊糖但是依次利用葡萄糖 / 木糖 / 阿拉伯糖的酵母菌株，这个实施方案特别引人注目。如果同时利用碳源，过程时间可以加快，这会提高过程效率并降低成本。生长底物可以是纯糖（己糖和戊糖）的任何混合物或者含有葡萄糖和戊糖的底物，特别是上文所提到的那些。因此，本发明还提供一种宿主细胞，其包括一种或多种新戊糖转运蛋白，还包括至少一种己糖转运蛋白，优选葡萄糖转运蛋白。在一个具体方面，将戊糖和己糖转运蛋白置于相同启动子的控制下。在另一个具体方面，将戊糖和己糖转运蛋白置于不同启动子的控制下。

[0024] 所述增值化合物可以是任何期望或有用的生物化学品，如生物燃料、有机酸、蛋白质物质、固醇等。优选地，所述增值化合物为生物燃料，更优选为生物乙醇。由于增加了代谢通量以及后续的乙醇生产，因此通过（异源）表达本发明的戊糖转运蛋白，可以提高乙醇的收率和产率。因此，本发明还涉及一种提供生物乙醇的方法，所述方法包括在宿主细胞中

根据本发明的核酸的表达,所述宿主细胞利用戊糖并表达至少一种图 1 所示的新戊糖转运蛋白。

[0025] 本发明还提供一种本发明的多肽、核酸分子、表达载体和 / 或宿主细胞在提高宿主细胞,优选真菌宿主细胞如酵母的戊糖吸收和 / 或利用中的用途。在一个优选实施例中,所述宿主细胞为重组工业酿酒酵母菌株,例如乙醇红色重组子 (Ethanol red recombinant)。在另一个优选实施例中,所述宿主细胞为丝状真菌。所述戊糖为阿拉伯糖,优选 L-阿拉伯糖;或者木糖,优选 D-木糖。

[0026] 在一个实施例中,本发明提供一种 An06g00560 和 / 或 An08g01720 的用途,或者表达该基因或其功能性同系物以增强 L-阿拉伯糖吸收和 / 或 L-阿拉伯糖利用的宿主细胞的用途。优选地,由于 An08g01720 具有低的 K_m 4.9mM (与 75mM 相比) 而使用 An08g01720。在另一个具体实施例中,本发明提供一种 An03g01620、An03g02190 和 / 或 An11g01100 的用途,或者表达该基因或其功能性同系物以增强 L-阿拉伯糖和 / 或 D-木糖吸收和 / 或利用的宿主细胞的用途。据发现 An11g01100 优选转运木糖,而 An03g01620 优选转运阿拉伯糖。因为 An11g01100 对 D-木糖具有比对 L-阿拉伯糖高的亲和性,其特别用于增加 D-木糖的吸收和 / 或利用。相比之下,An03g01620 和 An03g02190 对 L-阿拉伯糖具有较高的亲和性。然而,当任一种糖或两种糖存在时,两者均适合。

[0027] 此外,本发明提供一种重组宿主细胞,其包括编码多肽的分离的核酸(优选是载体的一部分),所述多肽选自下列物质组成的组:a) 多肽,其具有的氨基酸序列与图 1A 或 4A 所示的氨基酸序列至少 80% 相同,并且表现出体外和 / 或体内 L-阿拉伯糖转运活性;b) 多肽,其与图 1A 或 4A 所示的氨基酸序列相同;以及 c) 多肽片段,如 a) 或 b) 所定义的,其包括图 1A 或 4A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段,并且表现出体外和 / 或体内 L-阿拉伯糖转运活性。优选地,所述宿主细胞还包括至少一种编码参与 L-阿拉伯糖代谢的蛋白的核酸分子。实例包括 L-核酮糖激酶、L-核酮糖 -5-P₄- 差向异构酶和 L-阿拉伯糖 - 异构酶。此外,本发明还提供一种重组宿主细胞,其包括编码多肽的分离的核酸(优选是载体的一部分),所述多肽选自下列物质组成的组:a) 多肽,其具有的氨基酸序列与图 2A、3A 或 8A 所示的氨基酸序列至少 80% 相同,并且表现出体外和 / 或体内 L-阿拉伯糖和 D-木糖转运活性;b) 多肽,其与图 2A、3A 或 8A 所示的氨基酸序列相同;以及 c) 多肽片段,如 a) 或 b) 所定义的,其包括图 2A、3A 或 8A 所示的氨基酸序列中的至少 100 个连续氨基酸的片段,并且表现出体外和 / 或体内 L-阿拉伯糖和 D-木糖转运活性。优选地,所述宿主细胞还包括至少一种编码参与 L-阿拉伯糖和 / 或 D-木糖代谢的蛋白的核酸分子,例如选自自由 L-核酮糖激酶、L-核酮糖 -5-P₄- 差向异构酶、L-阿拉伯糖 - 异构酶、大肠杆菌 araBAD 操纵子编码的酶、NAD(P)H- 依赖性木糖还原酶 (XR)、来自树干毕赤酵母的 NAD⁺- 依赖性木糖醇脱氢酶 (XDH)、细菌木糖异构酶 (XI) 基因以及木酮糖激酶所组成的组。在一个具体方面,所述宿主细胞包括至少一种编码参与 L-阿拉伯糖代谢的蛋白的(外源)核酸分子,以及至少一种编码参与 D-木糖代谢的蛋白的(外源)核酸分子。

[0028] 另一具体方面涉及一种重组宿主细胞,其包括编码多肽的分离的核酸(优选是载体的一部分),所述多肽选自下列物质组成的组:a) 多肽,其具有的氨基酸序列与图 2A 所示的氨基酸序列至少 80% 相同,并且表现出体外和 / 或体内 D-木糖转运活性;b) 多肽,其与图 2A 所示的氨基酸序列相同;以及 c) 多肽片段,如 a) 或 b) 所定义的,其包括图 2A 所示的

氨基酸序列中的至少 100 个连续氨基酸的片段,并且表现出体外和 / 或体内 D- 木糖转运活性。优选地,所述宿主细胞还包括至少一种编码参与 D- 木糖代谢的蛋白的核酸分子,例如选自自由 L- 核酮糖激酶、L- 核酮糖 -5-P 4- 差向异构酶、大肠杆菌 araBAD 操纵子编码的酶、NAD(P)H- 依赖性木糖还原酶 (XR)、来自树干毕赤酵母的 NAD⁺- 依赖性木糖醇脱氢酶 (XDH)、细菌木糖异构酶 (XI) 基因以及木酮糖激酶所组成的组。

附图说明

[0029] 图 1 :A) 黑曲霉戊糖转运蛋白 An06g00560 的氨基酸序列 ;B) An06g00560 基因的核苷酸序列 ;

[0030] 图 2 :A) 黑曲霉戊糖转运蛋白 An11g01100 的氨基酸序列 ;B) An11g01100 基因的核苷酸序列 ;

[0031] 图 3 :A) 黑曲霉戊糖转运蛋白 An03g01620 的氨基酸序列 ;B) An03g01620 基因的核苷酸序列 ;

[0032] 图 4 :A) 黑曲霉戊糖转运蛋白 An08g01720 的氨基酸序列 ;B) An08g01720 基因的核苷酸序列 ;

[0033] 图 5 :A) 黑曲霉戊糖转运蛋白 An07g00780 的氨基酸序列 ;B) An07g00780 基因的核苷酸序列 ;

[0034] 图 6 :A) 黑曲霉戊糖转运蛋白 An13g02590 的氨基酸序列 ;B) An13g02590 基因的核苷酸序列 ;

[0035] 图 7 :A) 黑曲霉戊糖转运蛋白 An02g08230 的氨基酸序列 ;B) An02g08230 基因的核苷酸序列 ;

[0036] 图 8 :A) 黑曲霉戊糖转运蛋白 An03g02190 的氨基酸序列 ;B) An03g02190 基因的核苷酸序列。

具体实施方式

[0037] 实验部分

[0038] 材料和方法

[0039] 菌株和生长条件

[0040] 使用的所有黑曲霉菌株均来源于黑曲霉 N400 (= CBS120.49), 并且如表 1 所述。30°C 下, 在含有 2% 果糖、pH6.0 的完全培养基 (de Vries, R. P., K. Burgers, P. J. I. van de Vondervoort, J. C. Frisvad, R. A. Samson, and J. Visser. 2004. A new black *Aspergillus* species, *A. vadensis*, is a promising host for homologous and heterologous protein production (一种新的黑曲霉菌物种 (*A. vadensis*), 是用于同源和异源蛋白质生产的很有前景的宿主). *Appl. Environ. Microbiol.* (应用和环境微生物学), 70 :3954-3959.), 旋转振荡器为 250rev./min 的条件下培养预培养物。为了利于含有营养缺陷标记的菌株的生长, 在培养基中加入必需的补充物。16h 后, 收获菌丝体并用不含碳源的 MM 洗涤, 在含有 25mM L- 阿拉伯糖或 D- 木糖的 50ml MM 加入 1gr (湿重) 的菌丝体的等分试样并额外温育 2 小时, 然后收获。通过吸过滤器来收获菌丝体, 用不含碳源的 MM 洗涤, 在纸之间干燥并在液氮中冷冻。在 -70°C 下储存该菌丝体样品。

[0041] 30℃下,在含有2% (w/v) 麦芽糖的YP[1% (w/v) Bacto 酵母提取物 /2% (w/v) Bacto 蛋白胨] 上培养酵母菌株 EB. VW4000 (参考菌株)。使用的其他酿酒酵母菌株来源于菌株 EB. VW4000, 并且用基于 pYEX-BX (Clontech) 的质粒转化。根据快速和简单的 TRAF0 协议 (参考 TRAF0) 进行酵母细胞的质粒转化。30℃下,于 250rev./min 的旋转振荡器中,在补充了 0.1% (w/v) 酪蛋白氨基酸以及 0.2mg/l 色氨酸、20mg/l 亮氨酸、20mg/l 组氨酸、122mg/l 尿苷的 YNB[0.67% (w/v) Difco 酵母氮源] 中培养酵母菌株。使用如本文所述的碳源。除非另有说明,使用 0.5mM CuSO₄ 诱导从 CUP1 启动子表达。

[0042] 通过将 x¹⁴C-1-D- 木糖加入固体培养基 (2% Difco 琼脂) 上来进行 ¹⁴C-D- 木糖平板筛选 (plate screens)。在聚碳酸酯膜上的限定位置上接种 10³ 个酵母细胞,并且在 30℃ 下培养 2 天。为了避免产生任何 ¹⁴C- 二氧化碳,将平板置于还包括基于 NaOH 的碳阱的大型密封玻璃容器中。利用 x 胶片放射自显影使得能够检测已转运放射性标记的 D- 木糖的克隆。

[0043] 分子生物学方法

[0044] 除非另有说明,根据标准程序 (Sambrook 等., 1989) 来进行一般方法如 PCR、连接、消化、大肠杆菌 (*Escherichia coli*) (DH5 α) 的转化、质粒 DNA 分离以及凝胶电泳。根据供应商的说明书,利用 **TRIZOL**® 试剂 (Life Technologies) 从粉状菌丝体分离总 RNA。将 3 μ g 的总 RNA 与 3.3 μ l 的 6M 己二醛、10 μ l 的 DMSO 和 2 μ l 的 0.1M 磷酸钠缓冲液 (pH7), 在 20 μ l 的总体积中于 50℃ 下温育 1h 以使得 RNA 变性来进行 RNA 印迹分析 (Northern-blot analysis)。利用 0.01M 磷酸钠缓冲液 (pH7) 在 1.5% 琼脂糖凝胶上分离 RNA 样品,并且通过毛细管印迹转移至 Hybond-N 滤器 (Amersham Biosciences)。68℃ 下,在 30mM NaCl、3mM 柠檬酸钠和 0.5% (w/v) SDS 的严格条件进行印迹 (Southern blots) 的洗涤。根据制造商的说明书 (Roche) 进行探针的 DIG- 标记、它们的杂交和检测。

[0045] 通过 PCR 利用作为模版的 cDNA 文库来产生酵母的表达构建体。使用如表 2 所述的寡核苷酸。所使用的 cDNA 文库来自萌发时间进程,以及如 VanKuyk 等 *Biochem. J.* (2004) 379, 375-383 和 De Groot 等 . (2007) *Food Technol. Biotechnol.* 45 :134-138 所述来制备在 L- 阿拉伯糖或 D- 木糖上培养的菌丝体。根据供应商的说明书,利用 CloneMiner cDNA 库构建方法 (Invitrogen) 构建萌发时间进程库。通过在 CM- 葡萄糖培养基中预生长 16 小时并转移至具有聚碳酸酯滤器的琼脂平板的菌丝体构建分生孢子接合库 (conidiation library)。用第二聚碳酸酯滤器覆盖菌丝体以抑制分生孢子接合,或者没有第二滤器下温育并培养 8h 或 27h。汇集来自不同条件的等量 RNA 并用于构建每个库。将 cDNA 克隆入供体载体 pDONOR222 以产生进入文库。因为不能从 cDNA 文库获得基因 An13g02590, 所以利用 **Superscript**® 一步 RT-PCR 系统 (Invitrogen, Paisley, UK) 从 RNA 扩增。按照制造商的说明书,将来自重复 PCR 的产物克隆入 pGEMT-easy (Promega, Wisconsin, USA) 或 pJET (Fermentas, Ontario, Canada)。测序来自每个反应的多个克隆体 (Macrogen, South Korea), 并且将序列与公开的菌株 NRRL3122 (CBS 513.88)、ATCC1015 (Pei 等 . 2007. *Nat. Biotechnol.* 25, 221-231、以及 Baker SE (2006, *Med. Mycol.* 44 :Suppl 1 :S17-21) 的黑曲霉基因组序列进行比对。序列比对用来确定 PCR 错误、菌株差异和未剪接的内含子。将无 PCR 错误的含有完全加工的 cDNA 的基因克隆入 pYEX-BX (Clontech, California, USA), 用于酿酒酵母中的分析。

[0046] 微阵列分析

[0047] 通过用 BioArray 高收率 RNA 转录标记试剂盒 (ENZO) 或者 Affymetrix 真核生物单周靶标记和对照试剂包, 分别标记 20g 或 2g 的总 RNA 来产生生物素标记的反义 cRNA。利用 Agilent2100 生物分析仪检测 cRNA 的质量。将标记的 cRNA 杂交至 Affymetrix 黑曲霉基因芯片 (Affymetrix, Santa Clara, CA)。将 CBS513.88 (13) 的注释基因组的编码序列用作序列模版。

[0048] 从基因的 3' 端开始, 用 600-bp 片段设计寡核苷酸探针。探针集合由遍布整个芯片的 12 对 (匹配和不匹配) 25-bp 寡核苷酸探针组成。洗涤和染色的自动化过程后, 通过使用 Affymetrix 基因芯片操作系统 (Affymetrix GeneChip Operating System) 软件从扫描的阵列计算表达的绝对值。将微阵列套件 Affymetrix 版本 5.1 (Microarray Suite Affymetrix version 5.1) (Affymetrix Inc., Santa Clara, CA)、Spotfire DecisionSite (Spotfire, Inc. Somerville, MA)、Gene-Data Expressionist Analyst V Pro 2.0.18 (GeneData, Basel, Switzerland) 以及 R 统计环境 (www.r-project.org) 用于数据分析。将阵列与 3 个在麦芽糖上生长的 7 日龄的夹心培养物的外周的分别获得的 RNA 样品杂交。因为样品之间的相关性为 0.982, 并且发现平均信号比值为 0.044, 这就决定了可用生物学重复进行所有其他杂交。

[0049] 利用 Affymetrix 基因芯片操作系统处理 Affymetrix DAT 文件。通过利用 Affymetrix 标准化尺度至靶强度值 100 (TGT-100) 来从 CEL 文件产生 CHP 文件。

[0050] 功能分析

[0051] 以 24 小时的间隔通过 Perkin Elmer X 测量法来测量生长 (OD590), 来测定菌株 EB. VW4000 的己糖转运缺陷互补。用 Excel (Microsoft, USA) 将来自每个条件的 6-8 次重复的各个数据转化为平均 OD590 和标准差, 并将其作图使得能够比较菌株。

[0052] 酵母菌株培养 16-20h (约 D600, 2.0) 以便进行糖-转运实验 (零反式-内流测定)。将细胞通过离心 (在 4000xg 下 10min) 制粒, 用冰冷的 0.1M 磷酸盐缓冲液 (pH6.5) 洗涤并重悬以提供 0.1M 磷酸盐缓冲液 (pH6.5) 中的 10% 湿重 / 体积的悬浮液。将细胞保持在冰上直至被需要。根据 Walsh 等在 30°C 下 5s 温育期间, 测定 ¹⁴C- 标记的 D- 葡萄糖、D- 果糖、D- 甘露糖和 D- 木糖的零反式-内流。若在 pH5 下进行的实验, 则使用 0.1M 苯二甲酸 (pH5.0) 代替磷酸盐缓冲液。使用 Enzfitter 软件 (版本 1.05; Biosoft) 通过非线性回归分析确定不同单糖的转运蛋白的表观动力学参数。

[0053] 如 Serrano, R. (1977) Eur. J. Biochem. 80, 97-102; Santos, E (1982) Arch. Biochem. Biophys. 216, 625-660 先前所述, 利用 Titralab 模型 (Radiometer, Copenhagen, Denmark) 通过记录酵母悬浮液中的 pH 变化来监测糖转运期间的质子吸收。在 30°C 下, 利用磁力搅拌器混合悬浮液。从各实验开始通过 10mM HCl 的脉冲将 pH 降低至 5.0-5.1。

[0054] 转移实验:

[0055] 预培养: 16h CM 2% 果糖

[0056] 转移: 将约 1gr (湿重) 菌丝体转移至含有 25mM L-阿拉伯糖或木糖的 50ml MM 中

[0057] 表 1 本研究中使用的菌株

[0058]

[0059]

物种	名称	描述	基因型	参考文献
黑曲霉	N402	野生型		Bos 等. 1988
黑曲霉	UU-A033.21	araR 缺失菌株	nicA1, leuA1, δ argB:: δ araR, pyrA6	Battaglia 等,
黑曲霉	UU-A062.10	xlnR 缺失菌株	nicA1, leuA1, δ argB, pyrA6:: δ xlnR	Battaglia 等,
酿酒酵母	EB.VW.4000	己糖转运负菌株	MAT δ leu2-3, 112 ura3-52 trp1-289 his3- δ 1 MAL2-8° SUC2 hxt17 δ hxt13 δ ::loxP hxt15 δ ::loxP hxt16 δ :	(Wieczorke 等, FEBS Lett. 1999 Dec 31; 464 (3) :123-8.)
			δ loxP hxt14 δ ::loxP hxt12 δ ::loxP hxt9 δ ::loxP hxt11 δ ::loxP hxt10 δ ::loxP hxt8 δ ::loxP hxt514 δ ::loxP hxt2 δ ::loxP hxt367 δ ::loxP gal2 δ	

[0060] 表 2. 本研究中使用的寡核苷酸。用于克隆的限制性位点以斜体标明。

[0061]

基因	寡核苷酸名称	序列 5'-3'
An08g017 20	An08g01720up_SalI	<i>GTCGACATGCGTCTCTCCCCAGCA TG</i>
	An08g01720dw_SalI	<i>ATATGTCGACTCAACTCACTTCATT GTGGGTTCG</i>
An03g016 20	An03g01620up_BglII	<i>ATATAGATCTATGTATCGCATTTCGA ATATCTACG</i>
	An03g01620dw_EcoRI	<i>ATATGAATTCTCACGCCATTTCGTC ATGG</i>
An11g011 00	An11g01100up2_SmaI	<i>CCCGGGCATCATGGCTATCGGCAA</i>
	An11g01100dw_EcoRI	<i>ATATGAATTCAAGCAATCTTATCCG GAGTAG</i>
An06g005 60	An06g00560up_BamH I	<i>ATATGGATCCTCAACATGGGTATGG GTGC</i>
	An06g00560dw_NsiI	<i>ATATATGCATTACGCCGAGGGAGG AGTC</i>
An13g025 90	An13g02590up1_Bam HI	<i>ATATGGATCCAGAATGCTCATTTC ACTACCG</i>
	An13g02590up2_Bam HI	<i>ATATGGATCCAAGCTATGGAGAAC TTCGCTG</i>
	An13g02590dw2_Eco RI	<i>GAATTCATATCAGTTTTGTACATCC GCC</i>
	An13g02590dw_EcoRI	<i>ATATGAATTCTTTAACCATCATTAC</i>

[0062]

		ACGGAG
An02g082 30	An02g08230up_BamH I	GGATCCTATCCATCGGTGTCTCAAG AT
	An02g08230dw_ EcoRI	GAATTCACACACTCCGTCATGGTC AC
An07g007 80	An07g00780up1_Bam HI	ggatccAACCATGTCTGAGCCTAAGA
	An07g00780dw1_Eco RI	gaattcGCGGGATAGCCACCA

[0063] 戊糖转运蛋白对酵母生长的影响

[0064] 研究含有示例性新戊糖转运蛋白的酵母菌株关于在戊糖的存在下对生长的影响。如表 3 所示,观察到在 L-阿拉伯糖和 D-木糖的存在下,表达 0560 或 1100 转运蛋白的酿酒酵母表现出降低的生长。我们相信降低的生长是由于不利用戊糖的酿酒酵母菌株细胞内(无节制的)戊糖(或它们的代谢物)累积的毒性作用。有趣的是,在一定程度上均观察到两种糖对于两种蛋白的影响。

[0065] 表 3. 显示含有编码新戊糖转运蛋白的基因的宿主细胞中增加的戊糖吸收的毒性作用的结果。使用的生长培养基为 YNB+2% 琼脂+0.5mM CuSO₄。糖的浓度如表中所示。++++ = 生长非常良好, +++ = 生长良好, ++ = 生长不良, + = 生长非常不良, +/- = 生长隐约可见, - = 未生长。

	糖			菌株		
	麦芽糖	L-阿拉伯糖	D-木糖	对照	0560	1100
	50mM			++++	++++	++++
	50mM	25mM		++++	+	+++
	50mM	50mM		++++	+/-	++
[0066]	50mM	100mM		++++	+/-	+
	50mM	500mM		++++	-	+/-
	50mM		25mM	++++	+++	++
	50mM		50mM	++++	++	+
	50mM		100mM	++++	++	+
	50mM		500mM	++++	++	+

[0067] 戊糖转运活性测量

[0068] 酵母菌株生长 16-20h(约 D600, 2.0) 以便进行糖 - 转运实验(零反式 - 内流测定)。将细胞通过离心(在 4000g 下 10min) 制粒,用冰冷的 0.1M 磷酸盐缓冲液(pH6.5) 洗涤并重悬以提供 0.1M 磷酸盐缓冲液(pH6.5) 中的 10% 湿重 / 体积的悬浮液。将细胞保持在冰上直至被需要。根据 Walsh 等人(1994, J. Bacteriol. 176, 953-958) 在 30°C 下 5s 温

育期间,测定 ^{14}C - 标记的 L- 阿拉伯糖或 D- 木糖的零反式 - 内流。结果如下文表 4 所示。

[0069] 表 4. 基于在零 - 反式 - 内流测定中获得的数据,通过新戊糖转运蛋白的阿拉伯糖和木糖转运蛋白的 K_m 值。TND = 转运未检测

[0070]

基因	阿拉伯糖	木糖
An06g00560	75mM	TND
An03g01620	7.5mM	23mM
An03g02190	2.6mM	9mM
An11g01100	200mM	138mM
An08g01720	4.9mM	TND

[0071] 从表 4 的 K_m 值明显可见, An06g00560 和 An08g01720 均以可检测的水平转运 L- 阿拉伯糖而不是 D- 木糖, 因此这两种转运蛋白对 L- 阿拉伯糖具有特异性。An08g01720 具有比 An06g00560 高约 15 倍的对 L- 阿拉伯糖的亲合性 (4.9mM 与 75mM 相比)。An03g01620、An03g02190 和 An11g01100 均转运所测试的两种戊糖。An11g01100 对 D- 木糖具有比对 L- 阿拉伯糖高的亲合性, 虽然所获得的两种糖的 K_m 值在同一数量级 (即 100-200mM)。An03g01620 和 An03g02190 对 L- 阿拉伯糖的亲合性 (K_m 在 1-10mM 范围中) 比对 D- 木糖的亲合性高, 各自的 K_m 值高约 3 倍。

[0072] 因此 L- 阿拉伯糖的 K_m 值的范围为 2.6mM-200mM 以及 D- 木糖的 K_m 值的范围为 9mM-138mM, 这组转运蛋白能够转运广泛浓度范围的戊糖 L- 阿拉伯糖和 D- 木糖。

MGMGAGAEAVSGKRAELAGNRVGRGLMSSKKTFAIALFASLGGFVYGYNQ
GMFSEILTMNSFIKATDGYAARTGLKQGMLTSILELGAWVGTLLNGYLADAL
GRRQTVVVAVVIFCVGVIVQACTKNAGYVFAGRFTGLGVGNLSMIVPLYNAE
LAPPEIRGSLVAVQQLSITFGIMVSFWIGYGTNYIGGTGDGQSIAAWEVPVCIQ
VLPALILAAGMVLFPQSPRHLMNQGREEECLQTLARLRDAPTDDILVRIEYL
EIKSLKMFEEETAKKKYPQYQDGSFKNFMIGFYDYLSLVTNPSLFRKRTTVAC
LVMLFQQWNGINAINYYAPQVFEGLLELGGNTTSLLATGVAGIFEFVFTIPAVL
WVDNIGRKKTLLAGAIGMAVCHFICAGLIGSYEGTFGEHKSAGWATVAFVWI
FIINFAYSWGPCAWIVVSEVFPLSMRAKGVSIGGSSNWLNNFGVGLATSPFIA
ASTYGTFFIFGCITVVGAIYVWFFVPETKGRRTLEEMDELFGSEGMAAEDEALR
QRIDRDIGLTALLHGEMIDSSEKVPEKAVHADNVVLTPSA

图 1A0560 蛋白

ATGGGTGCGGGCGCTGAGGCCGTGTCCGGCAAGCGTGCTGAGCTTGCCGG
CAATCGCGTCGGATGGCGTGGGCTTATGAGCAGCAAGAAAACATTTGCGA
TTGCGCTGTTTTCGCTCGCTGGGTGGTTTTGTCTACGGATAACAACCAGGGA
ATGTTCTCCGAAATTTCTGACTATGAATTTCTTTCATCAAGGCGACCGACGG
ATATGCAGCGCGAACTGGTCTTAAGCAGGGTATGCTTACATCGATCTTGG
AACTGGGAGCGTGGGTGGTACACTGTTGAACGGTTATCTCGCCGATGCT
CTGGGCCGTCGCCAAACTGTCGTTGTCCGCGTGGTCATCTTCTGCGTCGG
TGTGATCGTCCAGGCTTGTACGAAAAACGCTGGCTATGTTTTTCGCTGGTC
GGTTCGTTACCGGTCTGGGAGTCGGTAATCTGAGTATGATTGTTCCCTCTG
TATAACGCAGAGTTGGCTCCCCCGAAATTCGTGGCTCCCTCGTCGCCGT
GCAGCAGCTGTCCATCACCTTCGGTATCATGGTCAGCTTCTGGATCGGCT
ACGGTACCAACTATATTGGAGGGACGGGTGATGGCCAGTCCATTGCGGCC
TGGGAGGTCCCCGTGTGCATCCAAGTCCTGCCGGCCCTGATCCTCGCTGC
CGGCATGGTTCTATTCATGCCTCAATCGCCTCGTCATCTGATGAACCAGG
GACGGGAAGAGGAATGTCTGCAGACGCTTGCACGACTTCGCGATGCCCC
ACTGATGATATTTCTGGTCCGCATTGAGTACTTGGAGATCAAGTCGCTCAA
GATGTTTGAGGAAGAGACCGCTAAGAAGAAGTACCCTCAATATCAGGATG
GATCTTTCAAGAGTAACTTTATGATCGGTTTCTACGACTACCTCTCGCTC
GTCACCAACCGTTCGCTGTTCAAGCGGACTACAGTTGCGTGCTTGGTCAT
GCTTTTCCAACAATGGAACGGAATCAACGCCATTAACACTACGCGCCTC
AAGTCTTCGAAGGCCTTGAATTTGGGAGGAAATACGACATCTCTATTGGCT
ACCGGTGTTGCTGGCATTTTCGAGTTTGTCTTACAATCCCGGCCGTGCT
CTGGGTGACAAATATTGGTCGTAAGAAGACCCTGCTTGCCGGAGCCATTG
GCATGGCAGTCTGCCATTTTCATCTGTGCAGGCCTCATCGGCTCCTATGAG
GGGACCTTTGGGGAGCACAAGAGCGCGGATGGGCGACAGTGGCATTTGT
GTGGATCTTCATCATTAACTTTGCCTACTCTTGGGGTCCGTGCGCGTGGA
TTGTCTGCTCCGAAGTTTCCCGCTGAGTATGCGTGCCAAGGGTGTCACT
ATTGGCGGCAGCAGCAACTGGCTCAACAACCTTCGGTGTCCGTCTGGCCAC
CTCTCCCTTCATCGCGGCCTCTACCTACGGTACCTTCATCTTCTTCGGCT
GCATTACCGTTCGTCGGAGCGATTTACGTGTGGTTCTTCGTTCCCGAGACC
AAGGGGCGCACCCCTGGAAGAGATGGACGAACTGTTCCGGCTCTGAGGGCAT
GGCAGCCGAAGACGAGGCGCTCCGGCAACGGATTGATCGCGACATCGGGT
TGACGGCACTGCTCCACGGCGAGATGATTGACTCTAGCGAGAAGGTTCCC
GAAAAGGCGGTCCATGCGGACGATGTGATGCTGACTCCTCCCTCGGCGTA
A

图 1B0560cDNA

MAIGNLYFIAAIAVVGGLFGFDISSMSAIIETDAYLCYFNQAPVTTYDDDGKRV
 CQGSPASVQGGITASMAGGSWLGLSISGFISDRLGRRRTAIQIGSIIWCIGSIIVCA
 SQNIPMLIVGRIINGLSVGICSAQVPVYISEIAPPTKRGRVVLQQWAIWGLI
 MFYVSYGCSFIKGTAAFRIPWGLQMIPAVLLFLGMMLLPESPRWLARKDRWE
 ECHAVLTLVHGQGDPSFPVQREYEEIKSMCEFERQNADVSYLELTKPNMLN
 RTHVGVFVQIWSQLTGMNVMMYYITYVFAMAGLKGNNNLISSSIQYVINVC
 MTPALVWGDQWGRRPTFLIGSLFMMIWMYINAGLMASYGHPAPPGLNNV
 EAESWVIHGAPSKAVIASTYLFVASYAISFGPASWVYPPELPLRVRGKATALC
 TSANWAFNFALSYPVPAFVNIQWKVYILFGVFCTAMFLHIFFFPETTGKTL
 EEVEAIFTDPNGIPYIGTPAWKTKNEYSRGAHIEEVGFEDKKVAGGQTIHQE
 VTATPKIA

图 2A1100 蛋白

1 atggetateg gcaatcttta ctccattgag gccatgcccg tcgtcggcgg tggtctgttc
 61 ggtttcgata tctcgtgat gtcggccatc atcgagaccg atgcctatct ctgttacttc
 121 aaccaggctc ctgtcactta cgatgatgat ggcaagaggg tctgtcaggg ccccagcgcg
 181 agtgtgcagg gtggtatcac cgctccatg gctgggtggt cctgggtggg ctctgtgatc
 241 tgggtttca tctcggacag gcttggctgt cgtactgeca ttcagatcgg tccatcacc
 301 tgggtcattg gatccatcat tgtctgtgcc tcccagaaca tcccattgct gatcgtcggg
 361 cgtatcatca acggtctgag tgtgggtatc tctccgctc aggtgccagt gtatattcg
 421 gagattgctc ctccaacaa gcgtggctgt gtcgtcggtc tgcaacaatg ggcattacc
 481 tggggatccc tgatcatggt ctactctcc tatggatgca gcttcatcaa gggtaaggcg
 541 gcctccgga tccctgggg tctgcagatg atccctgccc tgcattggt cctgggtatg
 601 atgctcctgc ctgagtcacc ccgtggctg gcaaggcaagg accgatggga ggagtgccac
 661 gctgttttga cctcgtcca cggtcaggga gaccgagct ctcccttctg gcagcgtgaa
 721 tatgaagaga tcaagagcat gtgcgagttt gaggcggaaa accgagatgt ctctacctc
 781 gagctgttca ageccaacat gcttaaccgt acccatgtgg gtgtttctgt tcagatctgg
 841 tctcagttga ctggaatgaa cgtcatgatg tactacatca cctactctt tgccatggcc
 901 ggcttgaaag gtaacaacaa ctgtatctcc tccagtatcc agtacgtgat caacgtgtgc
 961 atgactgtgc cggctctggt gtgggggtgat cagtggggcc gtcgcccagc ctcttgatc
 1021 ggttccctct tcatgatgat ctggatgtac attaatgctg gctcgtatggc cagctacggt
 1081 catcccgcgc cggccggcgg tctcaacaac gtggaagccc agtcttgggt catccacggc
 1141 ggcgccagca aggctgtcat tgccagtacc tacctcttct tagcctcata cccatctcc
 1201 ttcggccccg ccagctgggt gtaccgccc gaactcttcc ctctgcgtgt gcggggcaag
 1261 gctaccgccc tctgcacttc agccaactgg gcctcaact tgcctcag ctatcttgc
 1321 cccccggcat ttgtcaacat ccagtgggaag gctacatcc tctcgggtgt ctctgtact
 1381 gccatgttct tgcacatttt ctctctttt cccgagacca cgggtaagac cctggaagag
 1441 gtcgaggcca tcttcaactga tccaatggt attcgtaca tcggtactcc cgcctggaag
 1501 acaaagaacg agtactcgcg cgggtcacac attgaggagg ttggcttga agatgagaag
 1561 aagggtgctg gtgggcagac taccaccag gaggtcacgg ctactccgga taagattgct
 1621 tga

//

图 2B1100cDNA

MYRISNIYVLAGFGTIGGALFGFDVSSMSAWIGTDQYLEYFNHPDSDLQGGIT
 ASMSAGSFAGALAAGFISDRIGRRYSLMLACCIWVIGAAIQCSAQNV AHLVAG
 RVISGLSVGITSSQVCVYLAELAPARIRGRIVGIQQWAIEWGMLIMYLISYGCG
 QGLAGAASFRVSWG VQGIPALILLAALPFFPESPRWLASKERWEEALDTLALL
 HAKGDRNDPVVQVEYEEVQEAARLAQEAKDISFFSLFGPKIWKRTL CGVSAQ
 VWQQLLGGNVAMYYVVYIFNMAGMSGNTTLYSSAIQYVIFLVTTGTILPFVDR
 IGRRLLLLTGSVLCMACHFAIAGLMSRHHVDSVDGNANLKWSITGPPGKG
 VIACSYIFVAVYGFTWAPVAWIYASEVFPLKYRAKGVGLSAAGNWIFNFALAY
 FVAPAFTNIQWKTYIIFGVFCTVMTFHVFFFPETARRSLEDIDL MFETDMKP
 WKTHQIHDRFGEEVERHKHKDMADQEKG VVSTHDEMA
 //

图 3A1620 蛋白

1 atgtatcgca ttctgaatat ctacgtcctg gctggettcg gcaccattgg cgggtgccttg
 61 ttccgcttcg atgtcagctc catgagtgcc tggattggca cggatcagta tctggaatac
 121 ttcaaccacc cagattccga cttgcaaggt ggtattaccg cttccatgct tgctggetcc
 181 ttgcccgtg cactcgctgc cgcttcata tcagaccgaa ttggctcctg ttactccctt
 241 atgctagcgt gttcatctg ggtcatcgcc gctgcgattc aatgtagcgc ccaaaatgct
 301 gcccacctgg tcgcccggag agttatcagc ggctgtccg tcggtattac atctcaccag
 361 gtctgcgttt atctcgtga actagctccg gctcggattc gtggctgtat cgtcggcatt
 421 cagcaatggg ccatcgaatg gggtatgctc atcatgtatc tgatctcgtc cggatgcggc
 481 caaggctcgg ctggagcagc ttctttccga gtatcatggg gtgttcaggg tateccggcc
 541 ctcactctac ttcccgcatt acccttcttc ccgaatctc ctogttggct ggcaagtaaa
 601 gagcgtggg aggaggcttt ggatacttta ggcctgctgc atgccaaggg cgaccgcaac
 661 gatcccgtgg tccaagtggg atacgaagaa gtgcaggaag cagcacgcac tgcccaggag
 721 gccaaggaca ttctttctt tcaactgttc ggaccaaga ttggaagcg aacctttgct
 781 ggtgtcagcg cccaggtctg gcagcagttg ctccggaggc atgttgcgat gtactacgct
 841 gtgtacatct tcaacatggt tggcatgtct ggaaacacca cctgtactc gtcagccatc
 901 cagtacgta tcttctagt cacaacgggt actattctgc catttgttga tcgtatcgg
 961 cgacggcttt tgctctcac tggatccgct ctctgcatgg cttgccactt tgctattgct
 1021 ggtcttatgg cttcactgg tcacatggt gactcggctg atggcaatgc gaacctgaag
 1081 tggtaatta cgggaccacc cggcaaaggt gtcacgctc gctcgtacat ctctgttccc
 1141 gtttatgggt tcactgggc cccagtagca tggatctatg cttctgaagt ctccccctt
 1201 aagtaccgtg caaagggtgt cggctctca gctcgggta actggatttt caacttcgct
 1261 ttggcgtact ttgtggctcc tgcttccacc aacattcagt ggaagactta tatcatctc
 1321 ggtgtctttt gcaactgtat gacgttccat gtgttcttt tctaccggga gactgctcga
 1381 cgatcgttg aagacattga cttgatgttt gagacggaca tgaagccctg gaagacgcac
 1441 cagattcatg accgcttcgg cgaagaagtt gagcgacaca agcacaagga catggctgac
 1501 caagaaaagg gagttgtgct gaccatgac gaaatggcgt ga

图 3B1620cDNA

MRLSPAWYQFLVGVFASLGSFLYGYDLGVIAEVIACGSFISRFNLNDTESGLV
 VSMFTAGAFFGAAGFAGPSGDKLGRRWTTTVGCVLFCCLGGGLQTGARTVAYLY
 SGRFFAGLGVGFLTMIIPLYQAEICHPDIRGRVTALQQFMLGVGSLCAAWISY
 GTYIGFSETNDTQWQLPLGLQIAPAVFLGLLIMLFPESPRWLIDHGQHEKGLK
 TLAMLHAHGNEEDPWVRAEFNQIQESIVYEHENEAKSYKELFTSRSSFRRFL
 LCCSLQASVQMTGVSAIQYYSVTIYGQIGISGDKTLQYQAINSIALVAQFLCILF
 IDRVGRRWSLIWGNLGNMVTFIVACILLAQFPPEHNTGAHWGFIIMTWLYN
 FFSCTCGPLSWIIPAEVFDTRTRRSKGVSLATMTSYAFNTMIGQVTPIAMESVR
 YRYYFLFIICNFTNAVFFWLLLPETKKLPLEEMNYLFSNSPWIVAGKRKEDYV
 PHDLQRRLEEEAEKREVFATHNEVS

图 4A1720 蛋白

1 atgctctct cccagcatg gtatcaatc ctagtaggg ttttgcctc gctaggttca
 61 ttcttatatg gatacgattt aggtgtgatt gctgaagtc tegctgogg gtcgttcata
 121 tccagattea atttgaatga cactgaatct ggacttgtgg tgcgatggt cactgcaggt
 181 gctttctttg gtcagcgtt tgcaggtcca agtggcgaca agctaggcag acgatggacc
 241 attactgtcg gttgtgttct cttctgcctc ggaggtggtc ttcagactgg agcacggact
 301 gtcgctacc tetatagcgg aaggttcttt gctggactgg gggteggctt ccttactatg
 361 attatccctc tttaccagc ggagatctgc catccggaca taegtgtctg tgttactgct
 421 ctgcagcaat tcatgctagg agttgggtct ctatgtgccg cttggatctc ataccggaca
 481 tatattggat ttccgagac caatgacaca caatggcagc tacccttgg gctgcaaat
 541 gcaccagctg tgttctggg acttttgatc atgctttcc ctgagtcacc acggtggctc
 601 attgatcagc gccagcacga gaaaggcttg aagacattgg caatgcttca tgcccacggc
 661 aacgaagagg atccttgggt tctgtgggaa ttaaccaga tcaagaaag cattgtttac
 721 gagcatgaga acgaagccaa gtctacaag gaactcttca cctcaagatc gctctccgt
 781 cgctgttctc tttgttgctc attacaggca tctgtacaga tgaccggcgt gctctgcaatc
 841 caatactact ctgtaacct ctatggctag attggcatta gtggcgacaa aacgctacag
 901 tateagcta tcaactccat catcgccctt gtagctcaat tctatgcat tctcttcaat
 961 gaccgagctg gtcgcccctg gagtttgatc tggggtaacc taggaaacat ggtaacgttt
 1021 attgttgcct gcattctggt agcacagttc cccccagaaa gccacaacac cggagccac
 1081 tggggcttta tcatcatgac ttggctatac aacttctct tctctgtcac atgtggccct
 1141 ctctctgga tcatccagc agaggtatc gacaccgaa cggggtcgaa gggcgcttctg
 1201 ttggctacta tgacatcgt tgcattcaac acgatgattg ggcaggtgac gccatgtcc
 1261 atggagagtg tgcgctatcg gtactattc ctctttatca tctgcaactc caagaacgg
 1321 gtgttctttt ggetgttct tctgagacg aagaaattgc cgttgaaga aatgaactac
 1381 ctgttctcga attctccgtg gattgtggcg ggtaagcggg aggaggatta tgttctcag
 1441 gatttacagc gcaggcttga ggaggaggcg gagaagagag aggtgttctg gaaccacaat
 1501 gaagtgaagt ga

//

图 4B1720cDNA

MSEPKNQPVEEPTIPPEEGVKGWLTVAGSFCALFSSFGFLNAIGVFQTTYQQT
 SLKDYDASDISWIFAVQLALMWAPGPLWGRMIDTYGPIPVLWPCSILCVLGLC
 MTSLAHEYYQIFLAQGLCFGIGAGGVFTSAMICVGGQWFVRRRGLATGIAVSGS
 SLGGVIFPIFHDRVINDIGFYGAVRYTALFVGIMLAIACLLIRSRLPRKEWNGK
 AAWVDLTLKDTAFGLYTAGAFFIMWGLWAPFDYISSMAENAGFSSTLALYLI
 SIINAASIFGRLIPPQLADVFGHFVNLTLCCFGTGVSMMLCLWLFPNYHPSHAGI
 IVFSAVYGFVSGAVVSLMMPVAKVGDQLQTLGQRFQTFQLIMSVSCLTGLPIM
 GAILEKQNYTDYSGLQLFGACCGILGSVLIGAATFLLRKMRNTSKI

图 5A0780 蛋白

1 atgtctgagc ctaagaacca gccggtagag gagccaacca tcccgccca agaaggggtc
 61 aagggctggc tgacggtggc cggaagcttt tgcgccctat tctctcctt tgggttcctg
 121 aatgctatcg gactctcca aacaacatat caacagacgt cgctcaaaga ttacgaecgcg
 181 tccgacatat catggatttt cgcggtgcag ctggccctca tgtggggogcc gggaccgcta
 241 tggggccgca tgattgacac ttatggcccg atcccggtcc tctggccatg cagcattctc
 301 tgtgtgctag gactttgtat gactagcctg gcgcaacgagt actatcaaat attccttgcg
 361 caaggectct gcttcggcat cggagccggt ggtgttttca ctccggccat gatttgcgct
 421 ggccaatggt tegtcccgcc cggaggcctt gcaactggca ttgcggtatc gggaagtca
 481 cttgggggtg ttatcttcc catattccat gaccgggtca tcaatgacat cgctctctat
 541 ggtgccttc ggtatacagc gctattcgtc ggaattatgc tggcaatagc ctgcctaetg
 601 atccgcagcc gcttccgog caaagagtgg aacggtaaag cagcgtgggt cgacctgact
 661 ctgctgaagg ataccgcatt tggattgtac accgctggcg ctttttctat catgtggggt
 721 ctctgggccc ctctcgacta tattccagc atggcagaga aecggggatt tctctcgact
 781 cttgctctct acctgatttc tatcatcaat gccgctcca tcttcggacg tctcatcct
 841 cctcagctag cagacgtgtt tggtaacttc aacgtcctga cactctgctg ctccggtact
 901 ggagtgtcaa tgctctgct tggcttctt tcaactate accttcaaca tgcgggcatt
 961 atcgtctct cggcggctca tggcttcgtc agtggcgcgg ttgtctcct tatgatgcca
 1021 tgtgttgcca aagtgggaga tctgcaaacc ctgggtcagc gatttggaaac attcagcta
 1081 atcatgtcag taagctgctt gaccgggtg ccgatcatgg gagccatact cgaanaagca
 1141 aattacacgg attactcagg gttgcagctt ttgggtgctt gctgtgggat attagggtca
 1201 gttttgatcg gagcagcgac gttcttctc aggaaaatgc ggaatacag caaaatatga
 //

图 5B0780cDNA

MENFADLFAKGVFLGGYISNTYGRKWCIFSMNLYALGSAAVVSSKTDAQILT
 GRALHYTYLGMQLAVIPTTLAELAPRNVRGAMGVLYWLSIKIGGLVVTGIARG
 TSGISTNAAWQIPFGLLIVIPFICIWLWVHIPESPRWLLLRDRQEEALQALKRY
 RSKTTPQHEIQEDLDEMAQKVTAQLQKKSFKDLFIKGNRERTFVVAANFFQ
 QASGQAFASQYGTLFVKQLKSVNAFSVTLGTNATDIGALIISTALSDDVVGRM
 LFHISGFLETAALMTMGGLGTADTSNTAAKEGIVAMLLLYSFGWSLAWAPLV
 YVLGAELPSAGLREMTLRIAYTVKLVTEFAVTFSYPYLETADDPGHVDIGGKL
 GFYGSLSAVSVIFGYFFIPETRKLELEDINKKYESTESEMEQKVEYQITLDSI
 YPCAPIPLRVNDG

图 6A 2590 蛋白

1 atggagaact tcgctgattt gtttgcgaaa ggtgtcttcc taggaggta tattagcaat
 61 acctatggtc ggaaatgggtg catatttagc atgaatctct acgcactagg cagtgcagca
 121 gtagtggtga gtcctaaaa acacgcacaa atattaacag gtcgagcatt acactacatc
 181 tacctaggaa tgcaactage cgtaatccca accacctcg ccgaactegc cccccggaac
 241 gtcgcggcg ccatggggtt cctatactgg ctcagcatca aaatcggagg cctagtagta
 301 accggcatcg cagcaggcac atccggcacc tcaacaaacg ccgcatggca aatccccttc
 361 ggcttgattc tcgtgattcc atttatttgc atctggctcg tctggataat acccgagtct
 421 ccgcatggc tgcctctccg cgaccgacaa gaagaagcac tgcaagcgtt gaaacgatac
 481 cgatcgaana ccacacccca acatgagata caagaagatc tcgacgaaat ggccgagaag
 541 gtcacggccc agctgcagaa gaaaagcttc aaggatctct ttatcaaggg aaatcgtgaa
 601 cgtacgttcg ttgttgcgc tgcgaacttc ttcagcagg cttcggggca ggcgtttgcg
 661 agtcagtatg ggacgctttt cgtcaagcag ttgaaatcgg tcaatgcgtt tagtgttacg
 721 ctgggtaacta atgogactga tattggggcg ttgattattt cgacggcgtg gactgatgta
 781 gtgggacgga gaatgctctt ccacataage ggcttctcog aaacggcggc cttaatgaca
 841 atgggcggtc tcggtaccg agacactagt aacactgctg ccaaagaagg catcgtcgtc
 901 atgctcttc tttacagctt tggttggtct cttgcctggg cgcggttggg gtatgtgctt
 961 ggtgcccagt tgcctctgc ggggttgaga gaaatgactc tgcggatcgc gtatactgtg
 1021 aagctgggta ctgagttcgc tgtgacattt tcgtatcctt acctcgaac ccgagacgat
 1081 cccggccatg tggacattgg tggtaagctg ggattcattt atgggtcgtc gtctgctgct
 1141 tcagtcatat ttgggtactt ctccatccca gaaacgagga agtttagatt ggaggatata
 1201 aataaaaagt atgagtcgtc aacggagagt gaaatggagc agaaggttga gtatcagatt
 1261 actetagatt ctatctatcc ttgtgcgccg atctctctcc gtgtaaatga tggtaa

图 6B 2590 cDNA

MGRYFTIGLAAFAATGSYLFQYDSGVM TDVIESKNFLAFFNTTQTSSIIIGAINS
 TFSGGACIGALQGGLTMDRFGRKFTIQMGAFICMIGAILQSSAKNLAMILVGRI
 LAGWAVGLMSMSVPVYQAEVAHPRSRGFIIGLAQQMIGVGFIVSTWVGYGSL
 HAPNTSEFQWRFLAFQAVPAVLLVIGMFFMPESPRYLIEKERYEEMKILRR
 LHFDGTNEDWIQTEYNEIKTTIEAEKAVTVPGWLIMFRVPQWRTRLMHGIAV
 QVFTQMTGVNVVNYQTIMYNALGITGNRNTLVAGIYNCVGPITNFIFFFLD
 RVGRRKPMFLGTIAITIALVCEAALYSQNL DGTTRKGYSIGGVFFIFCITVIFSL
 FGPCSWVYMAEVMPMQIRGRGNFATAIGNWAVSTLWNQVSPIALDKIQWK
 FYFVFAAWNVCITLPTVYFFFKETNQSLEEIDLLFGGRALGMLPEDVAAKG
 GTVETEKAEETGVSVINVEHTSTV

图 7A 8230 蛋白

1 atgggtcgat actttacaat tggcctggct gcatttgag ctacggggtc atatctgttc
 61 ggatacgaca gcggtggtat gacggacgta atcgagtcta aaaactttct ggctttcttc
 121 aacaccaccc agacatcacc catcatcggg getatcaaca gcacatttcc ggggggtgct
 181 tgcattgggt cgctccaggg agggctgacg atggatcgct tcggacgaaa gttcactate
 241 caaatgggtg ctttcatatg catgattggc gctattctcc agtcgtcagc gaagaatttg
 301 gctatgattc ttgttggctg cattctggcc ggctgggctg ttggtctcat gtcgatgtct
 361 gttccagtgt atcaagcaga ggctcctcat ccccttccg gtgggttcat cctcgtctt
 421 gcccagcaga tgatcgggtg tggattcatt gtcagtacct gggctcggeta cggctccctt
 481 catgcaccga ataccagcga gttccaatgg cgattcccgc ttgcctttca ggcagtcctt
 541 gcagttctcc tggttatagg catgttttc atgcccagat cacctcggta cctgatogaa
 601 aaggaacgct acgaggaagc catgaaaatc ctccggagac ttcacttcca cggcaccaac
 661 gaggactgga ttcagactga gtacaacgag atcaaaaacca ccattgaagc cgaaaaagct
 721 gtcacagttc cggctggct catcatgttt cgtgtgcccc aatggcgaac ccgactcatg
 781 cacgggatcg cgtccaagt attcaaccag atgacagggc tcaacgtcgt gaactactat
 841 caaacaatca tgtacaatgc cctcggcacc actggttaacc gcaacacct cgtagcagga
 901 atctacaact gctcggctcc catcacgaa ttcattttca tcttttctct tctcagcga
 961 gttggggccc gcaagcccat gttgttcgga acaatcgeta tcaccattgc cctggtctgc
 1021 gaagcagctc tctactcgca gaacctggac ggcaactcga agggatacag tctcgggtgt
 1081 gttttcttca tctctcact cactgttacc tttctctgt catttggctc ttgcagttgg
 1141 gtctacatgg cagaagtaat gccaatgcaa atcccgggc gaggaaacgc ctccgccact
 1201 gcaatcggca actgggcccgt aagcactctc tggaaaccagg tctcccgat cgcgctggac
 1261 aaaatcagc ggaagtttta tttcgtgtt gcggcgtgga acgtatgcat cacctccca
 1321 accgtctact tcttcttfaa agagacaaat caaaaatac tcgaagagat cgatctactc
 1381 tttgggtgct gtgccctggg aatgcttct gaggatgtgg ctgccaaggg aggaacagtg
 1441 gagactgaga aggcagagga gacgggggta tcggttatta atgttgagca tacgagtact
 1501 gtgtgacct gacggagtgt gtgaattc

//

图 7B 8230 cDNA

MRLPKIKMPKRKPQDELNGDVLQQEAQATGPNQKSEEGIVDTPIPLLTWRS
 FLMGIFVSMGGFLFGYDTGQISGFLEMPNQLQRYGQQQADGTYFNSNARSL
 IVALLSIGTLIGALIAAPIADRVGRKWSISGWSAMVCVGITIQISSPFGKQWYQVA
 MGRWVAGLVGALSLLVPMYQAETGPRHIRGSLVSTYQLFITLGFVANCINF
 GTEARNDTGSWRIPMGITYIWAMILGFGIALFPESARYDYRHGREAKAARTLS
 RISSWKEQERGEITWSHLFHAPRMKYRVAVGVALQALQQLTGANYFFYYGTT
 IFRGAGISNSYVTQMILGGVNFVGTFLGLYLIENYGRRRSLITGALWFMVFCM
 VFASVGHFSLDHENPERTHTAGVVMVVFACLFILGFSTWGPVMVWTIAELY
 PSEFRARAMSLATASNWLWNFLAFFTFITSAIDFRLGYVVFAGCLFLAAGLV
 YVAVIEGRGRTLEEIDTMYVMKVPPWKELQVCISGHRSI

图 8A 2190 蛋白

ATGAGATTACCAAAAATTTAAAATGCCTCGTAAGAGGGCCGCAAGATGAACTCA
ATGGTGATGTGCTGCAGCAAGAGGGCGCAAGCTACCGGCCCAATGGCCAGA
AATCTGAAGAGGGGATTGTGGATACACCTATCCCTCTGCTCACGTGGCGTTC
TTTCTTAATGGGCATTTTCGTGTCAATGGGTGGTTTCCTCTTCGGCTACGAC
ACTGGCCAAATTTTCAGGCTTCTTGGAGATGCCGAACCTTCTCCAAAGATACG
GACAACAGCAAGCTGACGGAACATACTATTTCTCCAATGCCCGTTCCGGGTCT
GATTGTGCGCTGCTTTCAATAGGTACCTTGATCGGCGCCCTTATTGCAGCT
CCCATCGCAGATCGTGTTGGCCGAAATGGTCCATCAGCGGCTGGAGTGCC
ATGGTGTGTGTCGGTATCACCATACAAATATCTTCCCCTTTCGGAAAGTGGT
ACCAGGTAGCCATGGGTCTGTTGGGTCTGCTCGGCGTTGGGGCTCTAT
CATTGCTTGTTCATGTACCAAGCGGAACTGGACCGAGGCACATTCGTGG
ATCTCTTGTAGCACTTATCAGCTTTTCATCACCCCTTGGAATCTTCGTGCCA
ACTGTATCAACTTCGGAACAGAAGCTCGAAATGATACAGGCTCCTGGCGTAT
TCCCATGGGCATCACCTACATCTGGGCCATGATCCTGGGCTTCGGAATTGCA
CTATTCCCCGAAAGTGCCCGGTACGACTATAGACATGGAAGGGAGGCGAAA
GCAGCTCGTACCCTATCGAGAATAAGTTCGTGGAAGGAACAGGAACGCGGA
GAGATAACCTGGTCTCATCTTTTCCATGCACCCCGCATGAAGTATCGCGTCG
CAGTGGGAGTCGCCCTTCAAGCACTGCAACAGTTAACGGGAGCGAATTACTT
TTTCTACTACGGGACGACGATCTTTCGAGGCGCCGGGATTTCAAATTCCTAC
GTAACACAAATGATTCTTGGCGGTGTGAATTTTCGGAACGACGTTCCTAGGGC
TGTACCTGATCGAGAACTACGGTCGTGTCGATCCCTGATAACTGGAGCATT
GTGGATGTTTGTGTTTTCATGGTTPTCGCCTCCGTGGGCCATTTCTCTCTAG
ACCATGAAAACCCAGAACGGACACACACAGCTGGCGTGGTCATGGTTGTTTT
CGCCTGTCTTTTCATCTCGGTTTCGCATCAACATGGGGTCCCATGGTGTGG
ACCATTATCGCAGAGCTGTATCCCTCCGAATTCGTCGCCGCGCCATGTCTC
TCGCCACAGCATCCAATTGGCTCTGGAACCTCCTTCTCGCCTTTTTACCCCC
CTTCATCACCAAGTGCATTGACTTCCGCCTCGGCTACGTCTTTGCCGGCTGC
CTCTTCCTTGCGGCGGGCCTGGTTTACGTGGCAGTCATTGAAGGTGCGTGGCC
GTACATTGGAAGAGATCGATACCATGTACGTTATGAAGGTGCCGCCCTGGAA
AGAGCTCCAAGTATGTATTTCCGGACATCGATCCATATGA

图 8B 2190 cDNA