



(12) 发明专利

(10) 授权公告号 CN 101128581 B

(45) 授权公告日 2012.04.18

(21) 申请号 200680005888.1

(22) 申请日 2006.02.23

(30) 优先权数据

A304/2005 2005.02.23 AT

(85) PCT申请进入国家阶段日

2007.08.23

(86) PCT申请的申请数据

PCT/AT2006/000079 2006.02.23

(87) PCT申请的公布数据

W02006/089329 EN 2006.08.31

(73) 专利权人 格拉茨科技大学

地址 奥地利格拉茨

专利权人 VTU 控股有限责任公司

(72) 发明人 弗朗兹·哈特纳 安东·格利德

(74) 专利代理机构 北京市柳沈律师事务所

11105

代理人 封新琴 巫肖南

(51) Int. Cl.

C12N 15/11 (2006.01)

C12N 15/63 (2006.01)

C07H 21/00 (2006.01)

(56) 对比文件

CN 1372001 A, 2002.10.02, 全文.

戴秀玉等人. 毕赤氏酵母醇氧化酶-2 基因启动子突变体的分离和鉴定. 遗传学报 27
7. 2000, 27(7), 641-646.

审查员 邢云龙

权利要求书 3 页 说明书 49 页

序列表 29 页 附图 5 页

(54) 发明名称

突变 AOX1 启动子

(57) 摘要

野生型巴斯德毕赤酵母 (*Pichiapastoris*) 醇氧化酶 1(AOX1) 启动子 (SEQIDNo. 1) 的突变巴斯德毕赤酵母 AOX1 启动子包含选自下组的至少一种突变:a) 转录因子结合位点 (TFBS), b) SEQIDNo. 1 的核苷酸 170 至 235(-784 至 -719)、核苷酸 170 至 191(-784 至 -763)、核苷酸 192 至 213(-762 至 -741)、核苷酸 192 至 210(-762 至 -744)、核苷酸 207 至 209(-747 至 -745)、核苷酸 214 至 235(-740 至 -719)、核苷酸 304 至 350(-650 至 -604)、核苷酸 364 至 393(-590 至 -561)、核苷酸 434 至 508(-520 至 -446)、核苷酸 509 至 551(-445 至 -403)、核苷酸 552 至 560(-402 至 -394)、核苷酸 585 至 617(-369 至 -337)、核苷酸 621 至 660(-333 至 -294)、核苷酸 625 至 683(-329 至 -271)、核苷酸 736 至 741(-218 至 -213)、核苷酸 737 至 738(-217 至 -216)、核苷酸 726 至 755(-228 至 -199)、核苷酸 784 至 800(-170 至 -154) 或核苷酸 823 至 861(-131 至 -93), 和它们的组合。

1. 野生型巴斯德毕赤酵母醇氧化酶 1 启动子,即 SEQ ID No. 1 的突变巴斯德毕赤酵母醇氧化酶 1 启动子,其包含 Seq ID No. 1 的核苷酸 170 至 235,对应于 -784 至 -719 中的至少一个突变用于在甲醇诱导条件下高表达。

2. 根据权利要求 1 的启动子,其特征在于所述启动子还包含 Seq ID No. 1 的核苷酸 694 至 723,对应于 -260 至 -231 和 / 或核苷酸 729 至 763,对应于 -225 至 -191 的突变和 / 或转录因子结合位点的突变和 / 或选自下组的至少一种突变:核苷酸 170 至 191,对应于 -784 至 -763、核苷酸 192 至 213,对应于 -762 至 -741、核苷酸 192 至 210,对应于 -762 至 -744、核苷酸 207 至 209,对应于 -747 至 -745、核苷酸 214 至 235,对应于 -740 至 -719、核苷酸 304 至 350,对应于 -650 至 -604、核苷酸 364 至 393,对应于 -590 至 -561、核苷酸 434 至 508,对应于 -520 至 -446、核苷酸 509 至 551,对应于 -445 至 -403、核苷酸 552 至 560,对应于 -402 至 -394、核苷酸 585 至 617,对应于 -369 至 -337、核苷酸 621 至 660,对应于 -333 至 -294、核苷酸 625 至 683,对应于 -329 至 -271、核苷酸 736 至 741,对应于 -218 至 -213、核苷酸 737 至 738,对应于 -217 至 -216、核苷酸 726 至 755,对应于 -228 至 -199、核苷酸 784 至 800,对应于 -170 至 -154 或核苷酸 823 至 861,对应于 -131 至 -93。

3. 根据权利要求 1 或 2 的启动子,其特征在于所述突变是缺失、取代和 / 或插入。

4. 根据权利要求 2 或 3 的启动子,其特征在于所述转录因子结合位点选自下组:Hap1、Hsf、Hap234、abaA、Stre、Rap1、Adr1、Mat1MC、Gcr1 和 QA-1F,其中所述转录因子结合位点 Hap1 包含 Seq ID No. 1 的核苷酸 54 至 58,Hsf 包含 Seq ID No. 1 的核苷酸 142 至 149 和 517 至 524,Hap234 包含 Seq ID No. 1 的核苷酸 196 至 200、206 至 210 和 668 至 672,abaA 包含 Seq ID No. 1 的核苷酸 219 至 224,Stre 包含 Seq ID No. 1 的核苷酸 281 至 285,Rap1 包含 Seq ID No. 1 的核苷酸 335 至 339,Adr1 包含 Seq ID No. 1 的核苷酸 371 至 377,Mat1MC 包含 Seq ID No. 1 的核苷酸 683 至 687,Gcr1 包含 Seq ID No. 1 的核苷酸 702 至 706 和 QA-1F 包含 Seq ID No. 1 的核苷酸 747 至 761。

5. 核酸分子,包含根据权利要求 1 至 4 任一项的至少一种突变巴斯德毕赤酵母醇氧化酶 1 启动子和编码蛋白质或肽或功能性核酸的至少一种核酸,其特征在于所述启动子和所述核酸可操作地连接在一起形成单或多拷贝表达盒。

6. 载体,其包含根据权利要求 1 至 4 任一项的突变巴斯德毕赤酵母醇氧化酶 1 启动子或根据权利要求 5 的核酸分子。

7. 细胞,其包含根据权利要求 1 至 4 任一项的至少一种突变巴斯德毕赤酵母醇氧化酶 1 启动子、根据权利要求 5 的至少一种核酸分子或根据权利要求 6 的至少一种载体。

8. 根据权利要求 7 的细胞,其特征在于所述细胞是真核细胞。

9. 根据权利要求 8 的细胞,其中所述真核细胞是酵母细胞。

10. 根据权利要求 9 的细胞,其中所述酵母细胞为甲基营养酵母细胞。

11. 根据权利要求 10 的细胞,其特征在于所述甲基营养酵母细胞选自下组:念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属。

12. 根据权利要求 11 的细胞,其特征在于所述选自毕赤酵母属的甲基营养酵母细胞是巴斯德毕赤酵母细胞。

13. 用于表达所选蛋白的试剂盒,所述试剂盒包含:

i) 根据权利要求 6 的载体,和

- ii) 能够在根据权利要求 1 至 4 任一项的启动子控制下表达所述蛋白的细胞。
14. 根据权利要求 13 的试剂盒,其特征在于所述细胞是酵母细胞。
15. 根据权利要求 14 的试剂盒,其中所述酵母细胞为甲基营养酵母细胞。
16. 根据权利要求 15 的试剂盒,其特征在于所述甲基营养酵母细胞选自下组:念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属。
17. 根据权利要求 16 的试剂盒,其特征在于所述选自毕赤酵母属的甲基营养酵母细胞是巴斯德毕赤酵母细胞。
18. 用于在细胞中表达重组蛋白、肽或功能性核酸的方法,所述方法包括以下步骤:
- 提供根据权利要求 5 的核酸分子或根据权利要求 6 的载体,所述核酸分子和载体包含根据权利要求 1 至 4 任一项的醇氧化酶 1 启动子和编码蛋白质、肽或功能性核酸的核酸,将所述启动子与所述核酸可操作地连接,
 - 将所述细胞用所述载体或所述核酸分子转化,
 - 在合适的培养基中培养转化的细胞,
 - 任选地诱导所述蛋白质、肽或功能性核酸的表达和
 - 分离所述表达的蛋白质、肽或功能性核酸。
19. 根据权利要求 18 的方法,其特征在于所述细胞是酵母细胞。
20. 根据权利要求 19 的方法,其中所述酵母细胞为甲基营养酵母细胞。
21. 根据权利要求 20 的方法,其特征在于所述甲基营养酵母细胞选自下组:念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属。
22. 根据权利要求 21 的方法,其特征在于所述选自毕赤酵母属的甲基营养酵母细胞是巴斯德毕赤酵母细胞。
23. 根据权利要求 5 的核酸分子、根据权利要求 6 的载体或根据权利要求 7 至 12 任一项的细胞用于表达蛋白质、肽或功能性核酸的用途。
24. 用于分离超表达克隆的方法,所述方法包括以下步骤:
- a) 将包含至少一种根据权利要求 1 至 4 中任一项的突变巴斯德毕赤酵母醇氧化酶 1 启动子和至少一种编码蛋白质或肽的核酸或功能性核酸和标记抗性基因的核酸分子或包含所述核酸分子的载体引入细胞,其中所述启动子和所述核酸可操作地连接在一起形成单或多拷贝表达盒,
 - b) 将步骤 a) 的细胞转移至培养基,所述培养基包含合适的选择标记、非阻抑碳源和甲醇,用于在诱导条件下超表达克隆的选择性生长,
 - c) 在所述培养基上培育来自步骤 b) 的细胞,
 - d) 分离获得自步骤 c) 的细胞的菌落,和
 - e) 通过测定所述细胞的表达率检测超表达克隆。
25. 根据权利要求 24 的方法,其特征在于选择标记是抗生素。
26. 根据权利要求 25 的方法,其中所述抗生素为 Zeocin 或遗传霉素。
27. 根据权利要求 24 至 26 中任一项的方法,其特征在于选择标记是 Zeocin,并且标记抗性基因是 sh ble 基因。
28. 根据权利要求 24 至 27 中任一项的方法,其特征在于所述细胞是酵母细胞。
29. 根据权利要求 28 的方法,其中所述酵母细胞为甲基营养酵母细胞。

30. 根据权利要求 29 的方法,其特征在於所述甲基营养酵母细胞选自下组:念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属。

31. 根据权利要求 30 的方法,其特征在於所述选自毕赤酵母属的甲基营养酵母细胞是巴斯德毕赤酵母细胞。

32. 根据权利要求 24 至 31 任一项的方法,其特征在於所述非阻抑碳源选自下组:丙氨酸、甘露糖醇、山梨糖醇、海藻糖、乳糖和它们的组合。

33. 根据权利要求 24 至 32 任一项的方法,其特征在於将核酸分子或载体通过转化,或通过原生质体融合,或通过粒子轰击来引入细胞。

34. 根据权利要求 33 的方法,其中所述转化为电穿孔或化学转化。

突变 AOX1 启动子

[0001] 本发明涉及突变巴斯德毕赤酵母 AOX1 启动子。

[0002] 酿酒酵母 (*S. cerevisiae*) 作为真核模式生物和产生体系一直以来在科学和生物技术使用上占据重要地位 (并且目前仍然占据重要地位)。上世纪另一种酵母获得了很大的关注:裂殖酵母 (fission yeast) 粟酒裂殖酵母 (*Schizosaccharomyces pombe*)。由于其只依靠裂殖的繁殖特性,粟酒裂殖酵母作为模式生物获得了极大的关注,并且时至今日在分子遗传学和细胞生物学方面,它连同酿酒酵母一起是最彻底研究的酵母菌种。在目前已知的超过 700 种不同的酵母菌种里,上述两种酵母仅能够为技术和科学应用提供有限的感兴趣的特性。自二十世纪七十年代或八十年代以来,将越来越多具有突出特性的酵母菌种研究用于生物技术和研究。这些所谓的非常规酵母 (non-conventional yeast; NCY) 或非酵母属酵母 (non-Saccharomyces yeast) (在这种情况下术语酵母属包括酵母粟酒裂殖酵母) 的开发出于几种理由:它们拥有医学重要性如白念珠菌 (*Candida albicans*) 或技术相关性 (technological relevance) 如具有在特殊底物 (例如 n- 烷烃、乳糖) 上的生长能力的 *Yarrowialipolytica* 和乳酸克鲁维酵母 (*Kluyveromyces lactis*)。例如大量地研究最普遍的人真菌病原体白念珠菌以揭示涉及病理的毒力因子的本质,因此白念珠菌成为致病酵母的模式生物。另一组良好建立的 NCY 是甲基营养酵母巴斯德毕赤酵母和多形汉逊酵母 (*Hansenula polymorpha*) (*Pichia angusta*), 其在重组蛋白产生和过氧化物酶体生物发生的研究方面优于酿酒酵母。这些仅是仍然具有技术或学术吸引力的非常规酵母中最突出的成员。到目前为止几个其它的菌种也是特别感兴趣的,而这个组将在今后几年迅速地成长。

[0003] 糖,自然界中种类最丰富的分子,为所有已知的酵母利用。尽管种与种之间的底物接受度 (acceptance) 存在巨大差异 (参见表 1), 葡糖 -6- 磷酸或果糖 -6- 磷酸到丙酮酸的转化是它们新陈代谢中共同的主题。总之,用于糖酵解途径的酶设备在不同的酵母之间存在巨大的变化。虽然酿酒酵母中大多数酶是,至少是部分地已知并表征的,而在 NCY 中仅描述了少数酶。在一些酵母中糖酵解所需的一些功能通过几种基因 / 酶介导,特别是那些在新陈代谢控制或调节中扮演额外角色和 / 或处于分支点的基因 / 酶,如葡糖激酶 / 己糖激酶、磷酸果糖激酶和甘油醛 -3- 磷酸脱氢酶。通常,差别调节的同工酶说明在变化的环境先决条件下的不同的功能。一些编码糖酵解酶的基因是组成性和高表达的,例如酿酒酵母 PGK1 (磷酸甘油酸激酶) 或巴斯德毕赤酵母 GAP 基因 (甘油醛 -3- 磷酸脱氢酶), 而其它酶是严格调节的如酿酒酵母的 ENO1 (烯醇化酶) 基因。

[0004] 表 1:选择的引起生物技术方面兴趣的具有葡萄糖和果糖以外相关商业底物的酵母

[0005]

酵母	能量代谢	选择的底物
酿酒酵母	Crabtree 阳性	蔗糖、麦芽糖、蜜三糖、乙醇
粟酒裂殖酵母	Crabtree 阳性	蔗糖、麦芽糖、蜜三糖
拜列氏接合糖酵母 (<i>Zygosaccharomyces bailii</i>)	Crabtree 阳性	乙酸、乙醇、甘油

<i>Yarrowia lipolytica</i>	Crabtree 阴性	n- 烷烃、脂肪酸、乙醇
树干毕赤酵母 (<i>Pichia stipitis</i>)	Crabtree 阴性	木糖
巴斯德毕赤酵母	Crabtree 阴性	甲醇、甘油
多形汉逊酵母	Crabtree 阴性	甲醇、甘油
许旺酵母 (<i>Schwannomyces occidentalis</i>)	Crabtree 阴性	淀粉、n- 烷烃、木糖、蔗糖、蜜 三糖、海藻糖、乳糖、乙醇
乳酸克鲁维酵母	Crabtree 阴性	乳糖、蔗糖、麦芽糖、蜜三糖、 乙醇、甘油、木糖醇、乳酸

[0006] 新陈代谢中丙酮酸的命运在酵母种类和培养条件之间变化显著。在酿酒酵母和其它所谓的 Crabtree 阳性酵母中,葡萄糖和相关的糖抑制呼吸作用。这导致丙酮酸通过丙酮酸脱羧酶转化成乙醇和 CO_2 , 甚至在高氧含量下,其也称为发酵。在多数 NCY 所属的 Crabtree 阴性酵母中,丙酮酸到乙醇的转化仅在缺氧条件下发生。在有氧条件下,丙酮酸通过丙酮酸脱氢酶和三羧酸 (TCA) 循环氧化成 CO_2 。TCA 循环对于细胞代谢是特别感兴趣的,由于它是将糖氧化为 CO_2 的唯一途径。氧化成 CO_2 导致 NADH 的产生,将其用作能量产生 (energy production)。此外 TCA 循环中间体是用于生物合成目的的代谢物的主要来源。由于中间体的去除,必须补充 TCA 循环以保持它的运行。酵母中主要的添补反应是丙酮酸羧化酶和乙醛酸循环。前者是生长于铵作为单独氮源上时的主要途径,而后者是生长于少于 3 个碳原子的碳源上时所需的。

[0007] 相对于这种突出的兴趣,关于 NCY 中涉及 TCA 循环的基因和酶几乎一无所知。由分解代谢反应产生的 NADH,在细胞溶胶或线粒体中,必需再氧化成 NAD^+ 以保持所述反应进行。在 Crabtree 阴性酵母 (例如巴斯德毕赤酵母) 中,在有氧条件下,NADH 主要通过呼吸链再氧化。在 Crabtree 阳性酵母,如酿酒酵母中的情况有显著的不同,其中呼吸作用和发酵共存。在有氧条件生长于葡萄糖上时,呼吸作用由葡萄糖阻抑而出现发酵。在这些条件下 NAD^+ 通过乙醇 (NADH 通过糖酵解产生) 或甘油的形成再生。酵母中的呼吸作用不同于每本生物化学教科书中描述的这种途径的动物范例。首先,一些酵母,如酿酒酵母和乳酸克鲁维酵母,缺乏呼吸链的复合物 I。在这些酵母中进行 NAD^+ 再生不用由外部或内部 NADH 脱氢酶泵送质子通过线粒体内膜。第二个主要的不同存在于 Crabtree 阴性酵母、真菌和植物中,是与细胞色素链的复合物 III 和 IV 平行的可选择的呼吸作用途径。可选择的呼吸作用由所谓的可选择的氧化酶介导,该氧化酶直接从泛醌池 (pool) 转移电子至氧气,而不将质子泵送通过线粒体内膜。

[0008] 用于生物合成目的的 NADPH 在戊糖磷酸途径 (PPP) 的氧化部分中产生。由这种途径提供的其它非常重要的代谢物是核糖 -5- 磷酸和赤藓糖 -4- 磷酸,分别用于核酸和核苷酸辅因子的合成和用于芳族氨基酸的合成。有关非常规酵母中涉及 PPP 的基因和它们相应的酶的信息中仍存在许多差距。少数酶从产朊念珠菌 (*Candida utilis*)、粟酒裂殖酵母和乳酸克鲁维酵母中分离。组成的和动力学的表征揭示了这些酶之间的几种不同。由于信息的缺乏,这些酵母中对于 PPP 的影响无法评估,然而已经说明的是例如糖酵解缺陷的乳酸克鲁维酵母的磷酸葡萄糖异构酶突变体能够在葡萄糖培养基中生长,与酿酒酵母相反。这个观察指出乳酸克鲁维酵母中戊糖磷酸途径的容量足够在作为碳源的葡萄糖上的生长。在甲基营养酵母中,可存在额外的转酮醇酶 (二羟丙酮合酶)。这种酶定位于过氧化物酶体中,并且通过带有二羟丙酮和甘油醛 -3- 磷酸形成的与木酮糖 -5- 磷酸的缩合将甲醛的同化作

用赋予细胞代谢。

[0009] 酵母作为单细胞真核生物提供吸引人的用于重组蛋白产生的表达体系。它们结合了细菌的正面特点 (pros), 如充分开发的遗传操作技术, 简单、安全并且因此廉价的 (大规模) 培养技术, 和作为主要益处的真核表达体系, 即真核蛋白加工。由于上述原因, 酿酒酵母多年以来在该领域占据重要地位, 导致大量蛋白质 (例如胰岛素、HBsAg、HSA) 在这种生物体中产生。由于高糖基化作用、壁膜间隙中分泌蛋白的保持、质粒不稳定性和低产物收率, 酿酒酵母表现出一些局限性。为了克服这种个别生物体的局限性, 开发了小组的非常规酵母作为宿主细胞用于异源基因表达。其中, 使用了乳酸克鲁维酵母、*Y. lipolytica* 和甲基营养酵母博伊丁氏念珠菌 (*Candida boidinii*)、多形汉逊酵母和巴斯德毕赤酵母, 但仅最后 2 个菌种获得了突出的商业兴趣。粟酒裂殖酵母展现一些接近高等真核生物的特征, 使这种酵母成为用于异源蛋白产生的非常有吸引力的宿主: (1) 转录起始机制与高等真核生物的更相似, (2) 一些哺乳动物启动子在粟酒裂殖酵母发挥功能, (3) RNA 剪接的能力, 突出显示在剪接体成分与哺乳动物的相似性方面, (4) 能够识别哺乳动物内质网保持信号 (retention signal) KDEL, (5) 糖蛋白中半乳糖残基的存在和 (6) 一些其它翻译后修饰如蛋白质的乙酰化和异戊二烯化相较于酵母细胞以更类似于哺乳动物的方式进行。考虑到粟酒裂殖酵母如在结构和功能基因组学中可靠的异源蛋白产生和高产量的应用, 几种上述的特征可能在不远的将来增加其在重组蛋白产生中的重要性。

[0010] 所有微生物拥有机制以使它们的新陈代谢适应环境中可用养分的最优利用。对这些环境约束因素快速而精确的适应是所有生物体中控制生长和其它生理参数的主要因素。对于酵母, 正如对于多数微生物一样, 葡萄糖是优选的碳源和能量来源。因此意料之中的是, 大体上, 但不排除其它可能, 自然界中最丰富的单糖葡萄糖是细胞的主要信使, 通过调节在转录控制水平的基因表达影响这些生物体生长和发育。基因组转录分析揭示了可观数量的基因由环境决定的葡萄糖水平调节。在葡萄糖利用中具有已知代谢功能如低亲和性葡萄糖转运蛋白和糖酵解酶的基因以及编码核糖体蛋白的基因由葡萄糖诱导。另一方面, 葡萄糖阻抑一大组基因, 包括涉及可选择的碳源利用、糖异生、乙醛酸循环、过氧化物酶体功能和呼吸作用的基因。呼吸作用阻抑 (Crabtree 效应) 仅在少数酵母菌种 (发酵酵母, Crabtree 阳性) 如酿酒酵母内发生, 而在多数酵母种内葡萄糖不阻抑呼吸作用 (Crabtree 阴性)。尽管经过过去 20 年对于葡萄糖阻抑的方法 (主要基于酵母酿酒酵母) 获得了广泛的认知, 但是没有完全了解其真实机制, 特别是葡萄糖感应和信号传导的上游部分。然而, 为了获得对本发明工作更好的理解, 以下简要描述酿酒酵母的碳分解代谢物阻抑的少数主要参与物。

[0011] SNF1 基因编码能够在酵母细胞高分子量复合物中发现的 Ser/Thr 蛋白激酶。SNF1 基因由复合物内部的构象改变来调节, 所述改变由 Snf1p 调节亚基的磷酸化引起。迄今为止鉴定了磷酸化并且因此活化 Snf1p 的 3 个上游激酶 (Pak1p、Elm1p 和 Tos3p)。其活性对于去阻抑多种葡萄糖阻抑的基因绝对是必需的。因此意料之中的是, Snf1p 或同源物 (homologue) 在真核生物中普遍是保守的。

[0012] 锌指蛋白 Mig1p 能够结合至多种葡萄糖阻抑的基因的启动子区。其最有可能通过募集通用阻抑物复合物 Ssn6 (Cyc8)-Tup1p 来发挥作用。Mig1p 的功能由蛋白激酶 Snf1 控制, 至今没有关于直接磷酸化的清楚证据。Mig1p 以非磷酸化形式定位于细胞核中。葡萄糖

减少引起 Mig1p 的磷酸化,接着易位至细胞质。当向细胞添加葡萄糖时,Mig1p 迅速地退回细胞核并且阻抑转录。

[0013] Adr1p 也属于锌指蛋白家族,并且发现 Adr1p 是过氧化物酶体蛋白的正效应物,而 ADH2 基因编码葡萄糖阻抑的醇脱氢酶 II。葡萄糖通过依赖环 AMP (cAMP) 的蛋白激酶在高 cAMP 水平下调 ADR1 的表达。主要调节效应出现在 mRNA 翻译水平,但同样观察到对转录和 mRNA 稳定性的调节效应,这依赖于所分析的酿酒酵母菌株。

[0014] 对于很多基因,包括许多涉及呼吸代谢的基因,转录由 Hap2/3/4/5 复合物在非发酵性 (non-fermentable) 碳源上活化。对于少数涉及呼吸作用的基因如 CYC1 (编码异-1-细胞色素 c) 和 COX6 (细胞色素 c 氧化酶亚基 VI),已经确定 Snf1 是在葡萄糖上生长之后去阻抑所必需的。当葡萄糖存在时阻抑 HAP4 的转录,虽然如此无法证明去阻抑中直接涉及 Hap4p 或 Snf1p。

[0015] Gcr1p 是糖酵解基因 (例如烯醇化酶、甘油醛-3-磷酸脱氢酶) 的主要转录激活蛋白。Gcr1p 连同普遍性转录因子 Rap1p 是糖酵解基因表达有关转录协调的主要项目 (principle item),并且对于高水平转录绝对是必要的。野生型和生长于不同碳源上的酿酒酵母 gcr1 突变体的基因组表达图式揭示了 53 个开读框 (ORF),包括糖酵解的基因,如依赖 Gcr1p 的基因。

[0016] 此处对一些转录因子以及 Snf1p 和 Mig1p 途径的描述应给出对葡萄糖阻抑网络中一些参与者的简略概述。应该注意存在用于葡萄糖阻抑的 Snf1p 途径之外的更多调节循环。尽管在过去的 20 年中对葡萄糖的感应和信号传导获得了广泛的认识,主要的问题仍遗留未解:葡萄糖信号的性质是什么,以及已知的信号传导途径如何调节和整合。

[0017] 有限数目的酵母种能够在作为唯一碳源和能量来源的甲醇上生长。它们属于以下四个属之一:毕赤酵母属 (*Pichia*)、汉逊酵母属 (*Hansenula*)、念珠菌属 (*Candida*) 和球拟酵母属 (*Torulopsis*),并且共有普遍性甲醇利用途径,所述普遍性甲醇利用途径在去阻抑或用甲醇诱导之后表达 (参见 1.3.1)。由于这种途径的起始反应在过氧化物酶体内区室化 (compartmentalise),也同样诱导这些细胞器。由于过氧化物酶体的强诱导,酵母博伊丁氏念珠菌、*PichiaMethanolica*、巴斯德毕赤酵母和多形汉逊酵母频繁地用于细胞生物学中来研究过氧化物酶体生物发生和功能。

[0018] 如上所述甲基营养酵母共享共同的甲醇利用途径。第一步是将甲醇氧化成为甲醛和过氧化氢,由醇氧化酶催化 (AOX, EC1.1.3.13)。将有毒的 H_2O_2 通过过氧化氢酶的作用分解为氧气和水。两种酶均汇集在过氧化物酶体中。甲醛通过两个连续的脱氢酶反应来氧化或通过与木酮糖-5-磷酸 (Xu5P) 的缩合在细胞代谢中同化。通过依赖谷胱甘肽 (GSH) 的甲醛脱氢酶和甲酸脱氢酶,将甲醛氧化成甲酸并且进一步氧化成二氧化碳,所述甲醛脱氢酶和甲酸脱氢酶均位于细胞溶胶中。将产生在这两种反应中的 NADH 用于产生在甲醇上生长的能量。缩合反应发生在过氧化物酶体内,并且由上述转酮醇酶二羟丙酮合酶催化。所得的 C3 化合物二羟丙酮 (DHA) 和甘油醛-3-磷酸 (GAP) 在细胞溶胶中进一步代谢。在 DHA 的磷酸化之后,果糖-1,6-二磷酸 (FBP) 通过二羟丙酮磷酸 (DHAP) 和 GAP 的醛缩酶反应形成。将 FBP 通过磷酸酶转成果糖-6-磷酸,并且将木酮糖-5-磷酸 (Xu5P) 在戊糖磷酸途径中再生。所产生 GAP 的三分之一进入用于细胞成分合成的糖异生途径。

[0019] 甲醇利用途径的关键酶醇氧化酶和甲酸脱氢酶在用甲醇诱导之后以非常高的

水平产生。醇氧化酶能够占可溶蛋白质总量的超过 30%，二羟丙酮合酶甲酸脱氢酶多至 20%。同样被诱导的过氧化物酶体能够占细胞容积的约 80%。开发几种甲醇利用基因的启动子序列用于重组蛋白产生。其中，这些强诱导型启动子是广泛使用巴斯德毕赤酵母和多形汉逊酵母作为蛋白质产生的宿主的主要原因。

[0020] 在多形汉逊酵母和博伊丁氏念珠菌中，一种基因编码醇氧化酶：MOX（甲醇氧化酶，多形汉逊酵母）和 AOX1（醇氧化酶，博伊丁氏念珠菌）。两种基因存在于两个毕赤酵母属的菌株巴斯德毕赤酵母（AOX1 和 AOX2）和 *P. Methanolica*（AUG1 和 AUG2，醇利用基因，或 MOD1 和 MOD2）中，并且 Aox1p 和 Aug1p 是主要的醇氧化酶。编码区的比较揭示了甲基营养酵母之间在氨基酸水平 73-85% 的相似性 [1]。巴斯德毕赤酵母 AOX1 和 AOX2 的 ORF（开读框）之间的同源性在核苷酸和氨基酸序列水平分别是 92% 和 97% [2, 3]。醇氧化酶是八聚体黄素蛋白，每个亚基含有一个非共价结合 FAD 或修饰的类似物（mFAD）。AOX 翻译发生在游离核糖体上，接着在翻译后输入到过氧化物酶体中。易位入过氧化物酶体由其 C 末端尽头的 PTS1（1 型过氧化物酶体靶向信号）序列靶向。Aox 寡聚体仅在输入过氧化物酶体基质后形成。

[0021] 在博伊丁氏念珠菌和巴斯德毕赤酵母中，无法在细胞溶胶中发现 Aox 寡聚体，二羟丙酮合酶与之相反，在易位到过氧化物酶体基质之前，二羟丙酮合酶在细胞溶胶中形成二聚体。不仅巴斯德毕赤酵母的醇氧化酶 1 的启动子序列是感兴趣的，还由于广泛的底物范围（短到中长度链的不饱和以及饱和伯醇）和在各种反应条件下的高稳定性，酶也是生物技术感兴趣的。所有醇氧化酶的调节发生在转录水平，并且最可能在转录起始阶段。尽管 AOX1 和 AOX2 受到相似的调节（在甘油或葡萄糖上检测不到 mRNA，在碳饥饿期（carbon starvation phase）可检测到 mRNA，在甲醇上大量的 mRNA），但它们的 5'-侧翼区不共有显著的同源性 [2, 4]。

[0022] 每个 AOX 基因座在相对于初级转录起始位点的位置 -43 呈现推定的 RNA 聚合酶结合位点（TATAAA；戈德堡-霍格内斯（Goldberg-Hogness）或 TATA 框）。两种巴斯德毕赤酵母 AOX mRNA 前导序列均非常富含 A 残基，并且通常对于酵母是长的（对于 AOX1 为 115 核苷酸（nt）和对于 AOX2 为 160nt）。ATG 起始密码子周围的翻译起始区（Kozak 序列 AOX1：CGAAACG ATG GCT，AOX2：GAGAAAA ATG GCC）与之前描述的酿酒酵母和高等真核生物的共有序列一致。巴斯德毕赤酵母和 *P. Methanolica* 中第二醇氧化酶基因的生理学作用仍不清楚。在这些菌株中 AOX1 或 AUG1 的破坏引起严重的生长缺陷（所谓的甲醇利用缓慢（methanol utilization slow；Mut^s）表型），而 aox2 和 aug2 菌株展现出相当于野生型菌株的生长速率。在 *P. Methanolica* 中观察到 9 种多形式的醇氧化酶，表现出 2 种基因产物 Aug1p 和 Aug2p 的随机寡聚化。AUG1 和 AUG2 受到区别性的调节：在碳饥饿和低甲醇浓度仅能检测到 Aug1p，而随着甲醇浓度的增加，Aug2p 对 Aug1p 的比率增加。随着 Aug2p 含量的提高向八聚体的转变是由于 AUG2 表达的增加，所述表达在转录水平调控。Aug1p 和 Aug2p 的两种同型八聚体（homooctamer）对甲醇的 K_m 值分别是大约 0.56 和 5.6mM。连同破坏 AUG1 引起在低甲醇浓度生长缺陷的发现 [5]，这些结果说明当在高甲醇浓度生长时，AUG2 对于 *P. Methanolica* 是有利条件。在巴斯德毕赤酵母中，既没有进一步详细地分析 AOX2 基因的作用，也没有找出对于保有（possessing）第二醇氧化酶基因的有利条件。由于实验室条件仅代表自由生活的微生物所面对的条件非常小的部分，在自然界中应该存在 AOX2 基因对

于巴斯德毕赤酵母具有选择的重要性的情况。

[0023] 博伊丁氏念珠菌 AOD1 和多形汉逊酵母 MOX 的表达在生长于作为唯一碳源的葡萄糖或乙醇上时被严格地阻抑,在甘油上时去阻抑,并且在甲醇上强烈地诱导。当葡萄糖和甲醇存在于培养基中时,也阻抑这两种酶的表达。如果甘油存在,甲醇能够诱导基因表达。在碳饥饿时去阻抑 AOD1 和 MOX 的转录而当乙醇存在时被阻抑 [6-9]。两种截然不同的调节机制负责乙醇或葡萄糖对甲醇利用代谢的阻抑 [10,11]。巴斯德毕赤酵母中的情况显著不同:当葡萄糖、乙醇或甘油存在于培养基(以非生长限制(non-growth limiting)的浓度)中时阻抑 AOX1。碳饥饿时的去阻抑和甲醇的诱导与 AOD1 和 MOX 相似。AOX1 表达去阻抑的碳源是例如山梨糖醇、甘露糖醇、海藻糖和丙氨酸 [12]。

[0024] 从甲醇到阻抑碳源如葡萄糖或乙醇的转换之后,在适应新碳源期间过氧化物酶体在几小时内降解。当适应葡萄糖或乙醇时,酵母泡中的蛋白水解降解再次遵循两种截然不同的机制,分别称为小自噬或大自噬。

[0025] 如上所述,甲基营养酵母巴斯德毕赤酵母和多形汉逊酵母广泛地用于重组蛋白的产生。到目前为止,多于 500 种蛋白质已在巴斯德毕赤酵母中产生。少数特征推动了它们的发展,所述特征带给它们在重组表达宿主之中的优势:1) 它们共有酵母在遗传操作和培养技术(实验室的和大规模的)方面的普遍优势;2) 生长至极高细胞密度的能力;和 3) 重组蛋白(分泌的或胞内的)的高水平产生。将编码甲醇利用途径反应的基因的强诱导型启动子开发用于重组蛋白产生。最广泛使用的分别是巴斯德毕赤酵母和多形汉逊酵母的醇氧化酶基因 AOX1 和 MOX 的启动子区。但也将甲醇利用途径基因的其它启动子区用以驱动重组蛋白产生:多形汉逊酵母和博伊丁氏念珠菌中的 FMD(甲酸脱氢酶)和 DAS1(二羟丙酮合酶)启动子,以及巴斯德毕赤酵母中的 FLD1(甲醛脱氢酶)启动子。后者在甲胺作为唯一氮源及葡萄糖作为碳源时也能够诱导。用于外源基因组成性表达的启动子也是可用的:巴斯德毕赤酵母中的 GAP(甘油醛-3-磷酸脱氢酶)启动子元件和多形汉逊酵母中的 PMA1(编码质膜 H⁺-腺苷三磷酸酶)启动子。为巴斯德毕赤酵母(例如 HIS4)和多形汉逊酵母(例如 LEU2 和 URA3)开发了几种营养缺陷型的宿主菌株/标记基因组合。显性选择标记也是可用的(例如 Zeocin[™]、G418 抗性)。基因整合到甲基营养酵母中主要(如果非排外)通过同源整合完成。携带 ARS(自主复制序列)区的载体也是可用的,但是如果将选择压力释放,它们通常十分不稳定,这导致它们有限的技术应用。在巴斯德毕赤酵母中,将外源基因位点特异整合至 AOX1 或 HIS4 基因座。其它可能的整合位点是例如 GAP 基因座(用于 GAP 表达)或任何其它选择标记基因座(例如 ADE1、URA3、ARG4 和 LEU2)。在多形汉逊酵母中表达盒以头尾排列随机整合,导致具有高拷贝数(多至 100)的有丝分裂稳定的整合体。然而,高拷贝数通常不产生高水平表达。具有强烈影响的额外因素是:整合盒的结构、待表达蛋白质的性质和结构以及整合位点。特别是整合盒结构对于基因剂量的效果具有强烈的影响。如何最优化表达盒和基因剂量的进一步讨论在 [13,14] 中给出。甲基营养酵母属于 Crabtree 阴性酵母组,因此当生长于有氧条件下时,乙醇产生在非常低的水平发生。由于这个事实,这些酵母能够在发酵罐培养(fermenter culture)中生长至非常高的细胞密度得到高产物产率。当生物反应器中的甲醇浓度在生长限制范围(growth-limiting sphere)时,能够将 AOX1 驱动蛋白的产生进一步增加 3-5 倍。在标准条件下巴斯德毕赤酵母仅分泌少量内源蛋白的事实使每种分泌的重组蛋白都在培养基中最丰富。分泌能够作为下游纯化方法中重要

的第一步。对于蛋白质分泌,酿酒酵母 MF α 1 (交配因子 α) 前原前导序列 (prepro leader sequence) 和源自酸性磷酸酶 (PH01) 的序列广泛用于巴斯德毕赤酵母和多形汉逊酵母中。在一些情况下,使用具有天然分泌信号的植物、真菌和哺乳动物蛋白质获得充分的分泌。如上所述,酵母能够进行翻译后修饰如二硫键形成、信号序列加工 (例如 MF α 1 的前原前导序列)、脂质添加以及 N- 和 O- 连接的糖基化。虽然在哺乳动物细胞中产生由多种糖组成的非常复杂的 N- 和 O- 连接的寡糖结构 (例如 N- 乙酰氨基葡萄糖、半乳糖和唾液酸),但是多数酵母产生缺乏一些糖本体 (sugar entity) 如半乳糖或唾液酸的高甘露糖型结构。这些非哺乳动物结构能够导致治疗应用上严重的问题,主要由于它们的高潜在免疫原性。与酿酒酵母相反,在多形汉逊酵母和巴斯德毕赤酵母中,超甘露糖化 (hypermannosylation) 较少,并且无超免疫原性的末端 α -1,3- 连接的甘露糖并入 N- 连接寡糖。为了克服免疫原性的问题 (和一些其它问题如在血流中的低稳定性),已经开始努力将源自酵母的寡糖结构人源化,如最近的文献揭示的,特别是在巴斯德毕赤酵母中。迄今为止,绝大多数研究问题和商业方法依赖于公知的酵母酿酒酵母。由于对非常规酵母日益增长的认识,和大规模发酵和糖基化问题方面明显的优势,多形汉逊酵母和巴斯德毕赤酵母迅速成为优选的酵母。工业上实施的几种生产方法强调了所述事实。

[0026] 在 W002/081650 中,公开对 AOX1 启动子区的鉴定,其可以用于突变型 AOX1 启动子的构建。因为其中公开的 AOX1 启动子缺失的序列区非常长,能够观察到截然不同的启动子调节序列的累积效果而非单一效果。然而,这种方法将无法开发强烈增强的启动子。特别当通过缺失或复制原始启动子的部分构建具有增强特性的新启动子时,需要准确的关于调节序列范围的知识。

[0027] 本发明的一个目的是提供具有增强性质的改进的 AOX1 启动子,以促进蛋白质产生中的下游加工,从而增加时空产率 (time-space-yield) 并且帮助提升产物质量。

[0028] 本发明的另一个目的是提供载体或宿主菌株中的强 AOX1 启动子,所述载体或宿主菌株部分或全部将遭遇葡萄糖阻抑。优势是具有在高葡萄糖浓度存在下驱动强表达的启动子。

[0029] 本发明的另一个目的是提供 AOX1 启动子,所述 AOX1 启动子允许使用减少量的甲醇或不用甲醇来产生蛋白质。这种启动子将在工业生产方法上具有重要的影响。由于安全问题使用甲醇作为诱导物的生产植物需要特殊设备。这与巴斯德毕赤酵母应用于多种较低特殊性的生产植物中矛盾。另外,甲醇存在下的蛋白质稳定性能够阻碍基于甲醇的蛋白质表达的诱导。这对于生产强力的工业蛋白质较不关键,但是对于例如分泌的治疗蛋白却成为主要问题。

[0030] 这种启动子的构建需要野生型巴斯德毕赤酵母 AOX1 启动子的具体部分 (例如调节元件、转录因子结合位点) 的知识,当将所述 AOX1 启动子以某种方式突变时,显示对表达行为的影响。所以本发明的一个目的是鉴定这些部分并且因此提供创造具有增强特征的 AOX1 启动子的方法。

[0031] 因此本发明涉及野生型巴斯德毕赤酵母醇氧化酶 1 (AOX1) 启动子 (SEQ ID No. 1) 的突变巴斯德毕赤酵母 AOX1 启动子,包含选自下组的至少一种突变:

[0032] a) 转录因子结合位点 (TFBS),

[0033] b) Seq ID No. 1 的核苷酸 170 至 235 (-784 至 -719)、核苷酸 170 至 191 (-784

至-763)、核苷酸 192 至 213(-762 至 -741)、核苷酸 192 至 210(-762 至 -744)、核苷酸 207 至 209(-747 至 -745)、核苷酸 214 至 235(-740 至 -719)、核苷酸 304 至 350(-650 至 -604)、核苷酸 364 至 393(-590 至 -561)、核苷酸 434 至 508(-520 至 -446)、核苷酸 509 至 551(-445 至 -403)、核苷酸 552 至 560(-402 至 -394)、核苷酸 585 至 617(-369 至 -337)、核苷酸 621 至 660(-333 至 -294)、核苷酸 625 至 683(-329 至 -271)、核苷酸 736 至 741(-218 至 -213)、核苷酸 737 至 738(-217 至 -216)、核苷酸 726 至 755(-228 至 -199)、核苷酸 784 至 800(-170 至 -154) 或核苷酸 823 至 861(-131 至 -93), 和它们的组合。全部的描述中括号内的 (负) 数反映启动子相对于翻译起始密码子 (例如 ATG) 的相应位置。例如, 包含 N_x GACTATGN $_y$ 的核酸序列中 "ATG" 的 "A" 对应于位置 +1, 而 "ATG" 的 "A" 之前的 "T" 相对于位置 -1。

[0034] 根据本发明, 突变 AOX1 启动子在转录因子结合位点和 / 或以上所述范围的核酸序列之一中包含至少一种突变。已证明 AOX1 启动子的这些区域特别适合于修饰所述启动子从而改变其特征。当然也可以引入上述突变的组合以增强 AOX1 启动子的特性 (例如选自 a) 的两种 TFBS 突变, 一种选自 a) 的 TFBS 突变和一种选自 b) 的突变, 一种选自 a) 的突变和两种选自 b) 的突变)。例如, TFBS 的突变可以和 Seq ID No. 1 的核苷酸 737 至 738(-217 至 -216) 和 / 或核苷酸 207 至 209(-747 至 -745) 内的突变结合。巴斯德毕赤酵母中在 AOX1 启动子控制下的蛋白质表达一般通过添加甲醇来诱导并且由培养基中葡萄糖的存在来抑制。为了增强或减少所述培养基添加剂对蛋白质表达的影响, 启动子优选在上述启动子区突变。突变 AOX1 启动子产生感兴趣的蛋白的有效性 (efficacy) 依赖于整合到宿主染色体上载体的量 (即拷贝) 改变。特别地证实了多拷贝菌株展现出增强的启动子效果。由于毕赤酵母属菌株的抗生素抗性依赖于整合到所述宿主染色体中抗生素抗性盒 (引入到宿主的载体优选包含抗生素抗性盒, 以使宿主能够在包含抗生素作为选择标记的培养基上 / 中生长) 的数目, 多拷贝菌株可以通过在选择性琼脂平板上施用增加的抗生素浓度 (范围在 10 μ g/ml 至 10mg/ml 内, 优选 50 μ g/ml 至 1000 μ g/ml; 依赖于所用抗生素; 例如, 遗传霉素: 0.1 至 10mg/ml, 优选 0.2 至 5mg/ml, 特别是 0.25 至 4mg/ml, zeocin: 10 至 5000 μ g/ml, 优选 50 至 3000 μ g/ml, 特别是 100 至 2000 μ g/ml) 以增加选择压力 (例如 [14]; Scorer, C. A. et al. (1994) Bio/Technology 12:181-184)。然而, 发现含有抗生素抗性盒多重性的细胞生长不仅仅依赖于抗生素的浓度, 而且还依赖于时间。因此, 相较于单拷贝菌株, 在含有相同浓度抗生素的培养基上多拷贝菌株能够在较短时间内生长成可检测的菌落。这种特性使本领域技术人员能够在单拷贝菌株开始生长之前检测和分离多拷贝菌株。例如, 含有一个单拷贝抗生素抗性盒的菌株在 72 小时于琼脂平板上生长成可检测的菌株, 而含有多于一个拷贝所述盒的相同的菌株在 24 至 48 小时生长成相同的大小。

[0035] 特别是携带具有突变的 AOX1 启动子的多拷贝菌株显示出令人惊讶的增强的表达率 (expression rate), 所述突变在核苷酸 694 至 723(-260 至 -231) 内和核苷酸 737 至 738(-217 至 -216) 内。

[0036] 为了增加宿主在甲醇存在下的蛋白质表达效率, 优选将 AOX1 启动子 (SEQ ID No. 1) 的核苷酸 170 至 235(-784 至 -719) 突变。假设带有突变 AOX1 启动子的质粒仅整合到宿主染色体 / 基因组中一次 (单拷贝突变), 在这个区域的突变将蛋白质表达增加至野生型 AOX1 启动子的 120 至 130%。然而, 在所有其它上述区域内的突变减少或不影响甲醇诱导蛋白表达的功效。与此相反, 在野生型 AOX1 启动子的启动子区的突变 (如上所列) —

依赖于所述突变—在去阻抑条件下导致增加和减少的蛋白质表达（例如参见表 13，实施例 1）。

[0037] 然而，含有多于一个拷贝的突变 AOX1 启动子的重组菌株导致在去阻抑和甲醇诱导条件下具有增强活性的菌株（多拷贝菌株，例如参见图 7，实施例 2）。具体而言，在去阻抑条件下和 / 或在甲醇诱导下，含有以下突变的多拷贝菌株相较于在野生型 AOX1 启动子控制下的蛋白质表达显示增加的表达：所述突变在 SEQ ID No. 1 的核苷酸 694 和 723 内 (d6)，在 SEQ ID No. 1 的核苷酸 694 和 723 (-260 和 -231) 内 (d6)，在 SEQ ID No. 1 的核苷酸 694 和 723 (-260 和 -231) 和核苷酸 304 和 350 (-650 和 -604) 内 (d2d6)；在 TFBS 内，特别是在 Rap1、Gcr1、QA-1F、Hsf_1、Adr1、Hsf_2、Mat1MC、abaA 和 Hap2345 内。在去阻抑条件下，这些多拷贝菌株中的一些显示出相较于野生型启动子控制下的表达增加大约 10 倍的蛋白质表达。甲醇作为诱导物存在下使用根据本发明的启动子时，表达效率增强多于 5 倍。因此，这些突变，特别是当以多拷贝形式存在于宿主中时是蛋白质表达优选使用的。两种或两种以上上述突变的组合可以进一步增强启动子强度（参见例如图 7，实施例 2）。

[0038] 转录因子结合位点可以用实验方法鉴定（例如通过迁移率变动 (mobilityshift) 或足迹分析 (footprint analysis)）或通过与已知转录因子结合位点的序列比较（例如通过计算机分析，[15]）。

[0039] 影响所述启动子强度和特征的启动子区的知识可用以设计具有截然不同性质的启动子（在去阻抑条件下的高蛋白质表达和 / 或在甲醇存在下的高蛋白质表达）。此外如果将这些突变启动子一倍或多倍地整合到宿主的基因组中，这些性质可能增强或改变（例如参见实施例 1 至 3）。

[0040] 然而，在一些情况下，应该降低而不是增加启动子活性。特别是共表达调节蛋白如激酶、磷酸化酶和辅助蛋白例如蛋白伴侣、蛋白质二硫键异构酶、顺反异构酶、折叠酶、蛋白质二硫键异构酶和蛋白酶，在许多情况下必需低于主要产物，所述调节蛋白可能在野生型启动子或根据本发明的增强启动子下由细胞产生。特别地，证实了两种不同产物（例如辅助蛋白和主要产物）的组合表达分别在活性增加的 AOX1 启动子和活性减少的 AOX1 启动子（与野生型活性比较）的控制下是有利的，因为主要和次级产物的表达率相差甚至比使用野生型 AOX1 启动子更多。优选通过缺失激活子结合位点例如 HSF 或 HAP 或通过将阻抑物结合位点插入到野生型 AOX1 启动子中来获得减少的表达。因此，使用具有减少的活性的 AOX1 启动子防止细胞蛋白质表达机构 (protein expression machinery) 的过载，其结果将是主要产物的产量减少。例如，BessettePH 等 (PNAS USA (1999) 96 :13703-13708) 可显示可以通过硫氧还蛋白的共表达来显著增加活性多肽的表达。

[0041] 根据优选实施方案，启动子还包含 Seq ID No. 1 的核苷酸 694 至 723 (-260 至 -231) 和 / 或核苷酸 729 至 763 (-225 至 -191) 内的突变。

[0042] 影响这些核苷酸范围的突变与上述突变的结合导致更加增强的启动子活性。例如，在去阻抑和诱导条件下，与相同条件下野生型启动子控制的表达相比，实施 Seq ID No. 1 的核苷酸 694 至 723 (-260 至 -231) 和核苷酸 737 至 738 (-217 至 -216) 的双突变导致启动子显示较高的表达水平。当将包含启动子的核酸以多于一个拷贝引入细胞时（产生多拷贝克隆），可增强这种双突变的效果。

[0043] 突变优选是缺失、取代、插入、倒位和 / 或倍增 (multiplication)。

[0044] 为了改变巴斯德毕赤酵母野生型 AOX1 启动子的特性,几种突变类型是可能的。可以将包含上述区域(转录因子结合位点(TFBS)、Seq ID No. 1 的核苷酸 170 至 235(-784 至 -719)、170 至 191(-784 至 -763)、192 至 213(-762 至 -741)、192 至 210(-762 至 -744)、207 至 209(-747 至 -745)、214 至 235(-740 至 -719)、304 至 350(-650 至 -604)、364 至 393(-590 至 -561)、434 至 508(-520 至 -446)、509 至 551(-445 至 -403)、552 至 560(-402 至 -394)、585 至 617(-369 至 -337)、621 至 660(-333 至 -294)、625 至 683(-329 至 -271)、694 至 723(-260 至 -231)、729 至 763(-225 至 -191)、736 至 741(-218 至 -213)、737 至 738(-217 至 -216)、726 至 755(-228 至 -199)、784 至 800(-170 至 -154) 或 823 至 861(-131 至 -93)) 的启动子延伸(stretch) 部分或全部缺失、部分或全部用其它核苷酸或核酸序列取代、通过插入单核苷酸或核酸序列破坏、部分或全部倒位或倍增。所有这些突变导致启动子活性的改变,因为所述突变影响结构特征和/或例如转录因子的识别/结合部分。然而,这些改变可能导致启动子相较于野生型启动子增加或减少的活性。

[0045] 众所周知在现有技术中倍增/重复特定核酸延伸可能增加启动子活性。许多真核启动子特别是酵母启动子对基因表达的调节,涉及结合在启动子内部的转录因子之间的多种相互作用。多种位点对于甚至最小的顺式作用元件的功能是必需的。在酵母细胞中,上游激活子序列(UAS) 对于转录是必需的。它们以任一方向并在相对于 TATA 框和转录起始位点可变的距离上起作用,但是与高等真核生物中的增强子相反,它们必需在这些基础元件(basal element) 的下游。UAS 是几种转录激活子的靶物。

[0046] 酵母细胞中多数阻抑现象得自转录因子的失活和缺失。然而,也能够鉴定一些负调节位点(上游阻抑序列(URS))。

[0047] 基于巴斯德毕赤酵母 AOX2 启动子的缺失分析,发现三种调节区,两种负作用区(URS1 和 URS2) 和一种正作用结构域(UAS) [3]。对于多形汉逊酵母 MOX 启动子,还描述了两种上游激活序列(UAS1 和 UAS2) 和一种阻抑物结合位点(URS1) [8]。相应序列还能够在 AOX1 启动子上鉴定(核苷酸 585 至 614(-369 至 -340) 和 725 至 756(-229 至 -198),其显示与 AOX2UAS 的相似性 [3],以及核苷酸 622 至 656(-332 至 -298) [8])。这些核酸延伸的倍增(2、3、4、5、6 或 7 倍的 UAS) 可能导致具有增强强度的启动子,产生更加有力的蛋白表达。因此构建包含多 UAS 的启动子,优选包含与 AOX2 和 MOXUAS 相似的上述序列区,或其它多序列延伸(例如上述范围的核酸序列) 也落入本发明的范围,并且认为是优选的实施方案。激活序列通常在启动子的几百碱基对之内。例如,多数激活序列在增强的启动子的大约 200 至 400 碱基对之内。启动子的更上游通常含有另外的增强子和转录因子结合位点。

[0048] 可通过本领域技术人员已知的标准方法(例如 Molecular Cloning :A Laboratory Manual(Third Edition), J. Sambrook and D. Russell, 2001, ColdSpring Harbor Laboratory Press) 引入 AOX1 启动子的至少一种突变。

[0049] 根据本发明的优选实施方案,转录因子结合位点(TFBS) 选自下组:Hap1、Hsf、Hap234、abaA、Stre、Rap1、Adr1、Mat1MC、Gcr1 和 QA-1F。

[0050] 这些 TFBS 的至少一种的突变导致具有不同的特征的突变启动子(参见实施例 2)。

[0051] 优选地,转录因子结合位点(TFBS) Hap1 包含 Seq ID No. 1 的核苷酸 54(-900) 至 58(-896), Hsf 包含 Seq ID No. 1 的核苷酸 142(-812) 至 149(-805) 和 517(-437) 至 524(-430), Hap234 包含 Seq ID No. 1 的核苷酸 196(-758) 至 200(-754)、206(-748) 至

210(-744) 和 668(-286) 至 672(-282), abaA 包含 Seq ID No. 1 的核苷酸 219(-735) 至 224(-730), Stre 包含 Seq ID No. 1 的核苷酸 281(-673) 至 285(-669), Rap1 包含 Seq ID No. 1 的核苷酸 335(-619) 至 339(-615), Adr1 包含 Seq ID No. 1 的核苷酸 371(-583) 至 377(-577), Mat1MC 包含 Seq ID No. 1 的核苷酸 683(-271) 至 687(-267), Gcr1 包含 Seq ID No. 1 的核苷酸 702(-252) 至 706(-248) 和 QA-1F 包含 Seq ID No. 1 的核苷酸 747(-207) 至 761(-193)。

[0052] 这些 TFBS 可以用实验方法或通过使用计算机程序的辅助与其它启动子(例如真核生物的启动子)已知 TFBS 的比较来鉴定(例如参见实施例 1)。

[0053] 表 2 提供突变 AOX1 启动子对表达蛋白质、肽或功能性核酸的影响的总结(与野生型活性相比)。

[0054] 表 2 :野生型 AOX1 启动子突变体对表达蛋白质、肽或功能性核酸的影响

[0055]

突变	单拷贝克隆		多拷贝克隆	
	去阻抑条件 ¹	甲醇诱导条件 ¹	去阻抑条件 ¹	甲醇诱导条件 ¹
ΔHap1		+		
ΔHsf_1	-	+	+	+
Δ1	+	+		
ΔHap2345_1		+		+
ΔHap2345_2	-	+		+
ΔabaA	-		+	+
ΔStre	+	+		
Δ2		-		
ΔRap1	-		+	+

[0056]

	单拷贝克隆		多拷贝克隆	
$\Delta 3$	-	-		
ΔAdr1	-	-	+	+
$\Delta 4$	-	-		
ΔHsf_2				
$\Delta 5$	-	-		
$\Delta \text{Hap2345}_3$			+	+
ΔMat1MC				
$\Delta 6$	+	+	+	+
$\Delta 6^*$	+	-	+	+
ΔGcr1	-	+	+	+
$\Delta 7$	-	-		
$\Delta \text{QA-1F}$	+		+	+
$\Delta \text{QA-1F}_{\text{zus}}$	+	-		
$\Delta \text{Hsf}_2_{\text{dHap2345}_1}$		+	+	+
$\Delta \text{Hsf}_2_{\text{dHap2345}_1\text{zus}}$		+		
$\Delta \text{Hsf}_2_{\text{Mat1MC}}$	-		+	+
$\Delta 8$	-	-		
$\Delta 9$	-	-		
$\Delta 2\Delta 6$		-	+	+
$\Delta 736-41$		-		
$\Delta 737-38$				
$\Delta \text{InD-d4m}$				
$\Delta \text{D-d4}$				
$\Delta 1-1$				
$\Delta 1-2$				
$\Delta 1-3$				
$\Delta 1\text{-SacI}$				

[0057] ¹ 相较于野生型 AOX1 启动子的表达率：- 减少，+ 增加

[0058] 本发明另一个方面涉及核酸分子，所述核酸分子包含根据本发明的突变巴斯德毕

赤酵母醇氧化酶 1(AOX1) 启动子和编码蛋白质、肽的核酸或功能性核酸,其中启动子和所述核酸可操作地连接在一起。

[0059] 突变 AOX1 启动子能够连接至编码蛋白质(例如酶)、肽(例如激素)或功能性核酸(例如 siRNA)的基因。将所得核酸片段引入生物体中时,所述生物体优选酵母,特别巴斯德毕赤酵母菌株,所述核酸片段可用于表达例如蛋白质。所述核酸分子的构建对本领域技术人员是熟知的,并且能够使用标准分子生物学方法(例如 Molecular Cloning: A Laboratory Manual(Third Edition), J. Sambrook and D. Russell, 2001, Cold Spring Harbor Laboratory Press; manual "Pichia Expression Kit", Invitrogen Corp.) 进行。

[0060] “可操作地连接”指将第一序列置于足够接近第二序列的位置,从而使第一序列能够对第二序列或在第二序列控制下的区域施加影响。例如,可将启动子可操作地与基因连接从而使基因在启动子控制下表达,其通常连接至基因的 5'。通常核心启动子将在自翻译起始位点的几百碱基对之内。大约 30bp 下游通常存在下游启动子元件。

[0061] 本发明的另一个方面涉及载体,所述载体包含根据本发明的突变巴斯德毕赤酵母醇氧化酶 1(AOX1) 启动子或如上所述的核酸分子。

[0062] 为了将突变启动子引入到宿主中,将所述启动子任选地与编码蛋白质、肽或功能性核酸的核酸可操作地连接,所述宿主优选是甲基营养酵母菌株(例如巴斯德毕赤酵母菌株),所述启动子必须在载体中提供,其可能用于转化所述宿主。例如,所述载体可以是酵母附加体质粒(YEp)、酵母整合型质粒(YIp)或酵母人工染色体。这种载体通常包含用于载体在大肠杆菌中繁殖的复制起点(如果需要在微生物宿主中扩增)和选择标记、用于重组蛋白在酵母中表达的启动子和终止子以及用于酵母的选择标记。非整合型载体还包含自主复制序列(ARS),所述自主复制序列确保细胞中载体的稳定性(例如 Myers, A. M., et al. (1986) Gene 45:299-310)。不含有 AR 序列的整合型载体包含与基因组区域同源的序列区。可选地,可将线性 DNA,例如源自 PCR 的线性 DNA 用于转化。

[0063] 本发明的另一个方面涉及细胞,所述细胞包含如本文所公开的至少一种突变巴斯德毕赤酵母醇氧化酶 1(AOX1) 启动子、至少一种核酸片段或至少一种载体。可通过例如电穿孔将含有突变 AOX1 启动子的核酸分子(例如载体,其中将启动子可操作地与编码蛋白质的核酸连接)引入宿主中。在以单拷贝或多拷贝引入所述宿主中之后,或作为单拷贝或多拷贝自主复制质粒存在于细胞中之后,将所述核酸分子整合到染色体中。如果使用几种突变启动子,能够将它们全部与一种单独的基因(所述基因编码蛋白质或功能性核酸(例如核酶、反义 RNA 等)、同样的蛋白质或不同的蛋白质(例如将一种启动子变体与选择标记连接并且将另一种突变启动子与应该表达的另一种蛋白连接))连接。因此在本发明的范围内,优选产生单拷贝菌株,其包含与编码蛋白质、肽或功能性核酸的核酸可操作连接的一个拷贝 AOX1 启动子;以及多拷贝菌株,其包含与编码蛋白质、肽或功能性核酸的核酸可操作连接的多于一个拷贝,优选至少 2、3、4、5、6、7、8、9、10、15 或 20 拷贝的 AOX1 启动子。

[0064] 根据本发明的优选实施方案,所述细胞是真核细胞,具体为酵母细胞,优选甲基营养酵母细胞。

[0065] 优选地,甲基营养酵母细胞选自下组:念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属,特别巴斯德毕赤酵母细胞。

[0066] 可以将 AOX1 启动子及由其突变的变体功能性地 (functionally) 引入非常多种不同的酵母细胞中, 包括甲基营养 (例如巴斯德毕赤酵母) 和非甲基营养 (例如酿酒酵母) 细胞。启动子, 特别是 AOX1 和 MOX 启动子向其它生物体的可转移性对于本领域技术人员是已知的。尽管底物特异性和一些调节特征在不同酵母 (例如巴斯德毕赤酵母、多形汉逊酵母和酿酒酵母) 之中有所区别, 但已证明了外源启动子的识别 (例如 Raschke, W. C., et al., Gene, 1996. 177:163-7; Pereira, G. G. and C. P. Hollenberg, Eur J Biochem, 1996. 238:181-91)。例如, 在酿酒酵母中识别多形汉逊酵母 MOX 启动子, 在葡萄糖存在下阻抑并且在碳源限制时去阻抑。类似地, 可将 AOX1 启动子用于多形汉逊酵母中, 并且按照与 MOX 启动子相同的方式调节。与 AOX1 启动子紧密相关的 ZZA1 启动子也能够成功地用在酿酒酵母中。

[0067] 本发明的另一个方面涉及用于表达所选蛋白质或转录为功能性 RNA 的试剂盒, 所述试剂盒包含

[0068] i) 如上所定义的载体, 和

[0069] ii) 能够在根据本发明的启动子控制下表达所述蛋白质或功能性 RNA 的细胞。

[0070] 根据本发明的载体能够在用于表达所选蛋白质或转录功能性 RNA (例如核酶、反义 RNA、RNAi) 的试剂盒中使用。

[0071] 根据本发明的优选实施方案, 所述细胞是酵母细胞, 优选甲基营养酵母细胞。

[0072] 优选的甲基营养酵母细胞选自下组: 念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属, 特别巴斯德毕赤酵母细胞。

[0073] 本发明的另一个方面涉及用于在细胞中表达重组蛋白、肽或功能性核酸的方法, 所述方法包含以下步骤:

[0074] - 提供载体或核酸分子, 其包含根据本发明的 AOX1 启动子和编码蛋白质、肽或功能性核酸的核酸, 将所述启动子与所述核酸可操作地连接,

[0075] - 用所述载体或所述核酸分子转化所述细胞,

[0076] - 在合适的培养基中培养转化的细胞,

[0077] - 任选地诱导所述蛋白质、肽或功能性核酸的表达和

[0078] - 分离所述表达的蛋白质、肽或功能性核酸。

[0079] 根据本发明的优选实施方案, 所述细胞是酵母细胞, 优选甲基营养酵母细胞。

[0080] 优选地, 甲基营养酵母细胞选自下组: 念珠菌属、汉逊酵母属、毕赤酵母属和球拟酵母属, 特别巴斯德毕赤酵母细胞。本发明的另一个方面涉及根据本发明的核酸分子、载体或细胞用于表达蛋白质、肽或功能性核酸的用途。

[0081] 将根据本发明的核酸分子、载体或细胞用于表达蛋白质、肽或功能性核酸时, 选择满足表达所需条件的合适 AOX1 启动子是有利的 (例如在去阻抑条件下 (= 不向培养基添加葡萄糖) 或在甲醇诱导条件下的高表达或组成性表达)。合适的突变 AOX1 启动子可在表 2 帮助下选择。

[0082] 本发明的另一个方面涉及用于分离超表达克隆 (super expression clone) 的方法, 所述方法包含以下步骤:

[0083] a) 将包含突变甲醇诱导型启动子的核酸分子或载体引入细胞, 所述甲醇诱导型启动子优选 AOX1 启动子, 其与编码蛋白质或功能性核酸的核酸和抗性标记基因可操作地连

接,

[0084] b) 将步骤 a) 的细胞转移至培养基,所述培养基包含合适的选择标记、非阻抑碳源和甲醇用于在诱导条件下超表达克隆的选择性生长;或所述培养基包含合适的选择标记和非阻抑碳源,而不含甲醇用于超表达克隆在去阻抑条件下的选择性生长,

[0085] c) 在所述培养基上培育来自步骤 b) 的细胞,

[0086] d) 分离获得自步骤 c) 的细胞的菌落和

[0087] e) 通过测定所述细胞的表达率检测超表达克隆。

[0088] 超或高表达克隆的构建需要使本领域技术人员能够分离这些克隆的方法,所述克隆含有包含突变甲醇诱导型启动子的载体或核酸。本文提供这样的方法。所述方法的第一步是将包含启动子的核酸(例如载体)引入合适的细胞,所述细胞能够调节所述启动子。启动子本身可以通过遗传工程或通过化学(例如亚硫酸氢盐、亚硝酸盐、延胡索酸、胍)或物理(例如辐射,特别是UV辐射)诱变来进行突变。在进一步的步骤中将含有所述突变启动子的细胞转移至培养基,优选直接或经由液体培养基转移至固体培养基,所述培养基包含抗生素(例如Zeocin)和山梨糖醇(或另外的非阻抑碳源,例如[12]所述,特别是丙氨酸、甘露糖醇或海藻糖)用于在去阻抑条件下高表达克隆的生长,并且另外包含甲醇,如果期望发现在诱导条件下的高表达克隆。通过将葡萄糖与甲醇一起包含于培养基,可分离葡萄糖非阻抑的和甲醇诱导的转化体(为了防止甲醇挥发,可以将培养基在培育期间贮藏在甲醇饱和或包含甲醇的气氛中)。细胞在合适的培养基中或培养基上培养之后,将所述细胞从所述培养基中分离并且可以用于进一步分析(例如测定准确的表达率,分离启动子以分析与野生型启动子相比启动子核酸序列中的改变)。根据本发明和例如[12]公开的方法中使用的非阻抑碳源优选以0.1至10%的量使用,优选以0.2至5%的量,更优选以0.3至3%的量,特别以0.5至1%的量。优选的非阻抑碳源选自下组:丙氨酸、甘露糖醇、山梨糖醇、海藻糖、乳糖和它们的组合。

[0089] 合适的抗性标记基因的选择依赖于用来选择转化体的标记。例如,如果使用Zeocin作为标记,待引入载体的在突变AOX1启动子控制下的抗性标记基因是Sh ble基因。如果核酸编码蛋白质或肽,所得的/所表达的蛋白质可以是融合蛋白。特别有利的是提供在突变AOX1启动子的控制下的抗性标记基因,因为在这种情况下抗性标记基因产物的表达率也依赖于突变启动子的启动子强度和行为。例如,负责核酸产物高表达的强启动子也将增加抗性标记基因产物的表达率。这种克隆对于具有呈现启动子强度减少的启动子的克隆具有选择优势。这使得在细胞转化的再生之后能够直接选择超表达克隆。

[0090] 表达率优选通过以下方法测定:如凝胶电泳(例如SDS-PAGE)、抗体结合(例如ELISA)、定量(反转录酶)PCR(例如实时PCR)、酶活性(例如如果表达的蛋白质是酶)或荧光测定(蛋白质具有特征发射光谱如绿色荧光蛋白)。

[0091] 通过转化细胞在含有非阻抑碳源的培养基中/上的选择性生长来选择在不存在其它阻抑C源(AOX1启动子的情况下是葡萄糖)时显示表达增加的启动子(转化体)。在不存在其它阻抑C源(AOX1启动子的情况下是葡萄糖)而存在诱导物(例如甲醇)时,通过转化细胞在含有非阻抑碳源和诱导物(例如甲醇)的培养基中/上的选择性生长来选择显示表达增加的启动子(转化体)。诱导物也可以是非阻抑碳源。通过将多拷贝与如上所述的调节选择结合来选择超表达克隆,所述多拷贝导致针对抗生素(例如Zeocin)的较高

的抗性或培养基主要成分（例如 Leu、His、Arg、Ura）的较高生产力。

[0092] 用于根据本发明的方法中的培养基组合物可以直接从制造商或分销商由与巴斯德毕赤酵母有关的试剂盒、细胞和载体来获得（例如 Invitrogen）。培养基中的甲醇浓度可以优选 0.05 至 15%，更优选 0.1 至 10%，特别是 0.3 至 5%。在科学文献中，描述了用于不同培养条件的不同甲醇浓度。例如，摇瓶可以含有 1% 或更少的甲醇 (Guama MM, et al. (1997) *Biotech. Bioeng.* 56 :279-286), 发酵方法可以含有 0.5% 的甲醇 (Damasceno LM, et al. (2004) *Protein Expr Purif* 37 :18-26 ;Hellwig S., et al. (2001) *Biotechnol Bioeng* 74 :344-352 ;Hellwig S., et al. (1999) *Biotechnol Appl Biochem* 30 :267-275)。

[0093] 多拷贝克隆的增强表达可能不仅依赖于细胞中多于一个拷贝的突变启动子的存在，还由于缺少几种转录因子，因为这些因子可以结合至所述细胞中高数量的转录因子结合位点。这能够通过比较在甲醇诱导条件下的表达率与去阻抑条件下的表达率来证明，其中能够发现增强的表达率不仅是细胞中突变 AOX1 启动子拷贝数的效果（非线性效果）。例如，菌株 d6*F10 显示这种特征。

[0094] 用于分离超表达克隆的培养基可以包含另外的培养基成分如亮氨酸、尿嘧啶、精氨酸、组氨酸和 / 或腺嘌呤，并且山梨糖醇可以用葡萄糖交换以鉴定启动子变体，所述启动子变体在葡萄糖存在下与野生型启动子变体相比显示出减少的阻抑。

[0095] 当使用营养缺陷型菌株时，可以将细胞转移至包含山梨糖醇（或其它非阻抑碳源）并且含有单独培养基成分（例如亮氨酸、尿嘧啶、精氨酸、组氨酸和腺嘌呤）的培养基上，用于在使用营养缺陷标记 (auxotrophy marker) 的去阻抑条件下超表达克隆的选择性生长（步骤 b）。

[0096] 在包含 AOX1 启动子的载体中普遍使用的 P(TEF)-Zeo 抗性标记导致 zeocin 抗性蛋白的组成性表达，并且因此允许通过对较高浓度抗生素的抗性分离来多拷贝克隆。所述新方法使这种效果与调节特征结合以检测启动子和多拷贝克隆，其导致在一些可控制的调节环境下的较高表达（例如去阻抑表达、诱导表达等）。这使得检测新启动子和克隆成为可能，所述新启动子具有改变的调节性质，所述克隆中的多拷贝克隆导致在这种特殊调节条件下增强的表达。

[0097] “超表达克隆”是在突变启动子的控制下与在野生型启动子控制下相比表达较多蛋白质或功能性核酸的表达克隆，或与使用具有通常所用的启动子 - 选择标记组合的载体例如 P(TEF)-Zeo 相比表达较多蛋白质或功能性核酸的表达克隆。与相同蛋白质或肽或功能性核酸在野生型启动子控制下的表达率相比，根据本发明的“超表达克隆”的表达率可以增加至少 20%，优选至少 50%，更优选至少 100%，特别是至少 500%（平均值加上两至三倍的标准差）。“超表达克隆”可以优选包含多于一个拷贝根据本发明的突变启动子或核酸分子。可选地，“超表达克隆”也可以命名为“高表达克隆”。

[0098] 根据本发明“甲醇诱导型启动子”是其活性由培养基中存在的甲醇调节的启动子。这些启动子优选是 AOX1（来自巴斯德毕赤酵母）或 MOX（来自多形汉逊酵母）启动子或源自甲基营养酵母的任何其它甲醇诱导型和葡萄糖阻抑的启动子例如 FMD, FLD, DAS（例如参见表 6, 实施例 1）。

[0099] 根据本发明的优选实施方案，选择标记是抗生素，优选 zeocin。

[0100] 待用于培养基中的选择标记取决于细胞的哪种分子特征能够用于将含有核酸或

载体的细胞区别于不含所述核酸或载体的细胞,所述核酸或载体包含突变的或野生型甲醇诱导型启动子。因此选择标记可以是抗生素(抗生素抗性基因可存在于引入所述细胞的载体或核酸)。为了补偿一些菌株的营养缺陷,培养基中的选择标记可以是物质如亮氨酸、尿嘧啶、精氨酸、组氨酸和腺嘌呤,取决于营养缺陷的类型。

[0101] 优选的核酸分子、载体和细胞是根据本发明的核酸、载体和细胞。

[0102] 根据本发明的优选实施方案,将核酸分子或载体通过转化引入细胞,所述转化通过本领域技术人员公知的标准方法,优选电穿孔、化学转化、原生质体融合或通过粒子轰击(参见例如 Current Protocols in Molecular Biology, John Wiley&Sons, Edited by: Fred M. Ausubel et al.; Molecular Cloning: A Laboratory Manual (Third Edition), J. Sambrook and D. Russell, 2001, Cold Spring Harbor Laboratory Press) 进行。

[0103] 本发明通过以下的图和实施例来进一步描述而不受其限制。

[0104] 图 1 显示用甲醇诱导之前 (A) 和诱导 24 (B) 和 72 (C) 小时之后,微量表达 GFP-Zeo 的巴斯德毕赤酵母菌株的 SDS-PAGE。样品如实施例 1h) 中所述制备。泳道 1 是 X-33 (阴性对照),泳道 2-4 是 X-33GFP-Zeo 菌株 Mut^SA9、D2 和 E2,泳道 5 是 X-33d6*F10。所有 GFP-Zeo 克隆中在 42kDa 存在强条带。

[0105] 图 2 显示 AOX1 启动子区和一些转录因子结合位点中序列缺失的概况。将 delta1-9 区通过重叠延伸 PCR 缺失。

[0106] 图 3 显示去阻抑(碳饥饿)之后微量 AOX1 启动子变体的荧光强度条形图。将细胞以微量生长在 1% 葡萄糖上。数据表现了 4 个独立测量的平均值 ± 标准偏差 (SD)。RFU: 相对荧光单位; WT: 具有在野生型 AOX1 启动子控制下的 GFP-Zeo 的巴斯德毕赤酵母菌株 GFP-Zeo D2; D1-D9: 在 GFP-Zeo 基因之前具有缺失构建体 AOX1 Δ1-Δ9 的巴斯德毕赤酵母菌株; EX: 激发波长; EM: 发射波长。

[0107] 图 4 显示在甲醇诱导之后微量 AOX1 启动子变体的荧光强度条形图。将细胞以微量生长在 1% 葡萄糖上。数据表现了 4 个独立测量的平均值 ± 标准偏差。RFU: 相对荧光单位; WT: 具有在野生型 AOX1 启动子控制下的 GFP-Zeo 的巴斯德毕赤酵母菌株 GFP-Zeo D2; D1-D9: 在 GFP-Zeo 基因之前具有缺失构建体 AOX1 Δ1-Δ9 的巴斯德毕赤酵母菌株; EX: 激发波长; EM: 发射波长。

[0108] 图 5 显示微量所选 AOX1 启动子变体的荧光强度条形图。显示在去阻抑和诱导条件下具有野生型和 Δ6 启动子变体的单拷贝菌株和多拷贝菌株的表达水平。数据表现 4 个独立测量的平均值 ± 标准偏差。WT: 具有野生型 AOX1 启动子的单拷贝 GFP-Zeo 菌株 (GFP-Zeo D2), D6: 单拷贝 AOX1 Δ6* 克隆; WT_E2: 具有野生型 AOX1 启动子的多拷贝 GFP-Zeo 克隆; D6*F10: 多拷贝 AOX1 Δ6* 克隆 (X-33d6F10)。

[0109] 图 6 显示巴斯德毕赤酵母菌株在具有不同 Zeocin™ 浓度的 MD 和 MDM 琼脂平板上液滴测试 (drop test) 的结果。将细胞生长在 BMD (1%) 培养基上直到 OD₅₉₅ 为 1.5, 分 10 步稀释至最终稀释率 10⁵, 并且使用 48 针复制器 (48pin replicator) 转移至琼脂平板。图片顶部的数字表示对于所有平板均相同的稀释因子。如上所述制备 MD 培养基。在 MDM-Zeo 平板中添加甲醇至终浓度大约 0.5%。添加 Zeocin™ 至终浓度分别为 100、200 和 500 μg/ml。X-33: 巴斯德毕赤酵母 X-33, A9: 巴斯德毕赤酵母 GFP-Zeo Mut^SA9, D2: 巴斯德毕赤酵母 GFP-Zeo D2, E2: 巴斯德毕赤酵母 GFP-Zeo E2。

[0110] 图 7 显示几种多拷贝菌株与对照菌株相比的表达水平 ;a) 在去阻抑条件下的活性 ;b) 在甲醇诱导之后的活性。

[0111] 图 8 显示 $\Delta 6^*$ 多拷贝菌株在去阻抑和诱导条件下与对照菌株相比的表达水平。

[0112] 实施例

[0113] 实施例 1 :

[0114] 材料和方法 :

[0115] a) DNA 制备 / 纯化试剂盒 :

[0116] 根据提供的操作手册使用几种商业上能够获得的 DNA 制备和纯化试剂盒 (参见表 3)。

[0117] 表 3 :DNA 制备和纯化试剂盒

[0118]

试剂盒	制造商
Easy-DNA™ Kit	Invitrogen Corp., Carlsbad, CA, USA
QIAprep® Spin Miniprep Kit	QIAGEN GmbH, Hilden, Germany
Wizard® Plus SV Minipreps DNA Purification System	Promega GmbH, Mannheim, Germany
GenElute™ High Performance (HP) Plasmid Midiprep Kit Germany	Sigma-Aldrich Handels GmbH, Vienna, Austria
QIAquick® Gel Extraction Kit	QIAGEN GmbH, Hilden, Germany
Quantum Prep™ Freeze N Squeeze DNA Gel Extraction Spin Columns	Bio-Rad Laboratories GmbH, Vienna, Austria
QIAquick® PCR Purification Kit	QIAGEN GmbH, Hilden, Germany

[0119] b) TOPO® 克隆

[0120] 根据提供的操作手册进行 TOPO® 克隆 (用于克隆到 pCR® 4Blunt- TOPO® 中和用于克隆到 pCR® -Blunt II- TOPO® 中)。总是将 4 μ l 的 PCR 产物用于克隆。根据上述规程 (protocol) 将每个克隆反应的 2 和 4 μ l 转化到 One Shot® 化学感受态大肠杆菌 TOP10F' 细胞中 (Invitrogen Corp., Carlsbad, CA, USA)。

[0121] c) 大肠杆菌转化

[0122] 根据 SEM (简单而有效的方法) 规程将连接反应物和质粒转化到大肠杆菌中 [16]。将化学感受态大肠杆菌 TOP10F' 细胞用于转化。

[0123] d) 巴斯德毕赤酵母转化

[0124] 感受态巴斯德毕赤酵母细胞的制备：使用期望的巴斯德毕赤酵母宿主菌株的单菌落接种 300ml 带挡板宽颈 Erlenmeyer 瓶 (baffled wide-necked Erlenmeyer flask) 内的 50ml YPD (2% 葡萄糖) 中。在 30℃、60% 湿度和 130rpm (Pilot **Shake®** RC-2TE) 过夜温育之后，使用一定体积的这种预培养物 (pre-culture) 接种 2l 带挡板宽颈 Erlenmeyer 摇瓶中的 200ml YPD (2% 葡萄糖) 至 595nm 处的光密度 (OD_{595}) 为大约 0.1。将培养物在与预培养相同的条件下生长至光密度为 1.0 至 1.5。将细胞在 4℃ 和 4000rpm 沉淀 10 分钟，并且在 200ml 用冰预冷的无菌水中重悬浮。将这种流程重复 3 次，分别将细胞在 100ml 水、10ml 1M 山梨糖醇和 0.5ml 1M 山梨糖醇中重悬浮。

[0125] 将 10 μ g 期望的质粒在 300 μ l 终体积中用 BglIII 和 NotI (各 50u) 线性化过夜。限制酶切消化之后，将 DNA 根据标准规程 [16] 在 EtOH 和 0.3M 乙酸钠中沉淀。将 DNA 溶于 11 μ l 无菌 ddH₂O，并且使用 MF-Millipore™ Membrane Filter (参见表 12) 脱盐大约 1-2 小时。如果将 PCR 产物用于转化，如上所述从 EtOH 沉淀开始加工大约 4-5 μ g DNA。

[0126] 对于每个转化，将 80 μ l 制备的细胞与 10 μ g 如上所述的 DNA 混合，并且在冰上培育 5 分钟。将混合物转移至用冰预冷的电转化杯 (electrotransformation cuvette) (Bio-Rad)，并且在 200 Ω 、25 μ F 和 2.5kV 进行脉冲。立刻添加 1ml 用冰预冷的 1M 山梨糖醇。将悬浮液转移至无菌的 12ml PP 管 (Greiner, Frickenhausen, Germany, #184261) 并且在 30℃ 不摇动温育 2 小时。在此再生阶段之后，将等分试样 (aliquot) 涂布在选择平板上。对于在诱导条件下具有高表达的转化体的选择，将细胞涂布在 MSM-Zeo 平板上，所述 MSM-Zeo 平板含有基本培养基和山梨糖醇 (或任何其它非阻抑碳源) 甲醇和 zeocin。对于在去阻抑条件下显示高表达的克隆的选择，可将细胞涂布在缺乏甲醇的基本山梨糖醇 zeo 培养基平板上。在含有甲醇的选择平板中包括葡萄糖，使其能够检测葡萄糖非阻抑表达克隆和它们的启动子。

[0127] e) 菌落 PCR：

[0128] 将期望的毕赤酵母属菌株的单菌落在 100 μ l 小管 (microtube) 内的 100 μ l ddH₂O 中重悬浮，并且在 95℃ 加热 5 至 10 分钟。在 13, 200rpm 离心 1 分钟后，将 10 μ l 上清液用作 PCR 反应的模板。将 5 μ l 这种第一轮 PCR 用作第二轮的模板。将 5 μ l 第二轮 PCR 用于对照凝胶。PCR 反应根据所提供的操作手册在缓冲条件下，在终体积 50 μ l 中含有 10pmol 的每种引物 (AOX1_col 和 GFPprev)、200 μ M 的每种 dNTP 和 2.5 单位的 Hot Star **Taq®** DNA 聚合酶 (QIAGEN) 或 Taq DNA 聚合酶 (Promega)。为了测序，使用 **QIAquick®** PCR Purification Kit 纯化第二 PCR 产物。

[0129] 表 4：菌落 PCR 的温度程序

[0130]

温度	Taq	Hot Star Taq®	循环
95°C	5 分钟	15 分钟	1
95°C	30 秒	30 秒	
57°C	30 秒	30 秒	
72°C	1 分 30 秒	1 分 30 秒	30
72°C	10 分钟	10 分钟	1

[0131] f) 巴斯德毕赤酵母基因组 DNA 分离：

[0132] 将期望的巴斯德毕赤酵母菌株在无菌的 12ml PP 管内的 5ml YPD 中在旋转桶上于 30°C 生长过夜，直至终 OD₅₉₅ 为 5-10。根据所提供规程使用 Easy-DNA™ Kit 将 1.5ml 培养物用于 DNA 分离。

[0133] g) 蛋白质试验 (protein assay)：

[0134] 溶液中蛋白质浓度的测量已长期在生物化学中使用。它们的主要应用之一是将多种生物化学方法归一化为总蛋白量，如本案例中用于氧消耗率 (oxygen consumption rate)。最常使用的测定蛋白质浓度的方法是 Bradford、Lowry 和 BCA™ 方法。考虑到敏感性、动力范围 (dynamic range) 和对特定试剂的相容性，这些方法具有一定的局限性。在这三种试验中，Bradford 和 Lowry 比 BCA™ 更加可靠和可重复。另一方面在去污剂和 / 或还原剂存在时，Lowry 和 Bradford 具有严重的局限性，其导致高空白值。因此在化学溶解之后，BCA™ 试验是优选的方法。如根据说明书操作手册 (Pierce Biotechnology Inc.)，在化学细胞溶解之后以 Y-Per® 和 BSA 作为标准物，使用 BCA™ 试验测定蛋白质浓度，因此以下仅简要描述主要步骤。将 200 μl 培养物在 4000rpm 和 4°C 离心 5 分钟。弃去上清液之后，将沉淀在 100 μl Y-Per® 中通过上下吹吸 (pipetting up and down) 来重悬浮。将悬浮液在 Thermomixer 中 1.5ml 小管内于室温和 600rpm 温育 20 分钟。在 13, 200rpm 和室温将细胞碎片沉淀 10 分钟之后，将上清液转移至新的小管中，并且贮藏在 4°C 用于 BCA™ 试验或 SDS-PAGE。将 25 μl 样品在微量培养板的孔中与 200 μl BCA™ 工作试剂 (试剂 A : 试剂 B = 50:1) 混合，充分震荡 30 秒，并且用平板密封物 (Promega) 盖紧。在 37°C 温育 30 分钟并冷却至室温之后，使用 Spectramax Plus384 平板读数器 (plate reader) 测定在 562nm 的吸收。如有必要，在 BCA 试验之前将样品用 ddH₂O 稀释。

[0135] h) SDS-PAGE：

[0136] 如上面部分描述的使用 Y-Per® 作为试剂通过化学细胞裂解来制备用于 SDS-PAGE 的样品。将 10 μl 溶解产物与 10 μl 2x SSB (sigma 样品缓冲液) 混合，并且在 95°C 温育 5-10 分钟，并且将 15 μl 的这种混合物加样到蛋白质凝胶上。使用 180V 进行 1 小时电泳，并且使用 Coomassie™ 蓝染色检测蛋白质条带。

[0137] 表 5：用于 SDS-PAGE 的凝胶制备

[0138]

	积层胶 (4%)	分离胶 (12%)
ddH ₂ O	3.05ml	3.35ml
30% 丙烯酰胺 / 双	650 μ l	4ml
0.5M Tris-HCl pH6.8	1.25ml	
1.5M Tris-HCl pH8.8		2.5ml
10% (w/v) SDS	50 μ l	100 μ l
TEMED	5 μ l	10 μ l
10% APS	25 μ l	50 μ l

[0139] i) 葡萄糖试验：

[0140] 使用无去蛋白质作用的葡萄糖-UV 己糖激酶方法测定葡萄糖浓度 (DIPRO med Handels GmbH, Weigelsdorf, Austria, Prod. no. D590522)。将 50 μ l 毕赤酵母属培养物转移至 PCR 微量培养板并且在 4000rpm 离心 5 分钟。将 10 μ l 上清液添加至 UV-Star 微量培养板内的 190 μ l 己糖激酶试剂中并且在室温培育 15 分钟。培育之后使用 Spectramax Plus384 平板读数器在 340nm 测定吸收。

[0141] j) 液滴测试：

[0142] 将巴斯德毕赤酵母菌株生长在 BMD (1%) 中至终 OD₅₉₅ 为 1.5, 并且分 10 步稀释至终稀释率 10⁵。使用 48 针复制器转移至琼脂平板上。将平板在 30°C 温育直到菌落出现 (在 MD 平板上通常是 2 天)。

[0143] k) 序列比对：

[0144] 所有序列比对使用 INRA 主页上的 MultAlin 进行 (Institut National de la Recherche Agronomique, Paris, France) (prodes.toulouse.inra.fr/multalin/multalin.html) [17], 或使用 European Bioinformatics Institute 的 ClustalW (EBI, www.ebi.ac.uk/clustalw) [18] 进行。对于使用 MultAlin 的序列比较, 总是将 DNA 序列相似性矩阵用于比较。

[0145] 甲醇利用途径基因和多数过氧化物酶体基因以相对于葡萄糖阻抑、碳饥饿时去阻抑和通过甲醇诱导类似的方式调节。使用已定义的一组转录因子 (抑制子以及诱导子) 的相似转录调节应该是这种调节模式的原因。在这些启动子区内的转录因子结合位点应该显示一些保守区域。共调节的基因的启动子区之间的多重序列比对应该揭示涉及调节一致基因的转录因子的保守结合位点。已报导甲基营养酵母巴斯德毕赤酵母、多形汉逊酵母和博伊丁氏念珠菌的几种基因是共调节的并且已分离了它们的启动子序列 (表 6)。

[0146] 表 6: 来自甲基营养酵母巴斯德毕赤酵母、多形汉逊酵母和博伊丁氏念珠菌的甲醇利用途径基因或过氧化物酶体基因的共调节基因

[0147]

酵母	基因	酶	Genbank 登录号	文献
巴斯德毕赤酵母	AOX1	醇氧化酶		www.invitrogen.com
	AOX2	醇氧化酶	X79871	
	ZZA1	醇氧化酶	S62281	
多形汉逊酵母	FLD1	甲醛脱氢酶	AF066054	
	MOX	甲醇氧化酶	A11156	
	DAS	二羟丙酮合酶	A11168	
	CAT	过氧化氢酶	X56501	
博伊丁氏念珠菌	AOD1	醇氧化酶	M81702	

	FLD1	甲醛脱氢酶	AB085186	
	FDH1	甲酸脱氢酶	AB035095	
	DAS1	二羟丙酮合酶	AB035094	
	PMP20	过氧化物酶体膜蛋白	AB035096	
	PMP47	过氧化物酶体膜蛋白	AB035097	
	CTA1	过氧化氢酶	AB064338	

[0148] 1) 转录因子分析：

[0149] 使用 Genomatix Software GmbH Servers 上的 GenomatixSuite1.6.1 内的 MatInspector Release professional6.1Jan.2003 进行转录因子分析 [15]。使用 Matrix Family Library Version3.1.1April2003group ALL fungi.lib(www.genomatix.de) 将来自 pPICZB 的 PAOX1 序列搜索转录因子结合位点。

[0150] m) 引物：

[0151] 表 7：所述实施例中使用的引物列表（由 MWG Biotech AG,Ebersberg,Germany 合成）

[0152]

SEQ ID No.	名称	序列	T _m [°C]
2	P(AOX1) forw	AAGGTACCAGATCTAACATCCAAAGACGAAAG	70
3	P(AOX1) rev	CTAGCCATGGTTGAATTCTTTTCAATAATTAGTTGTTTTTG	67
4	GFPZeo forw	GAAAGAATTCAACCATGGCTAGCAAAGGAG	70
5	GFPZeo rev	GATGATGGTCTAGAACGTGTCTAGTCCTGCTCCTC	70
6	AOX1TT forw	GACACGTCTAGACCATCATCATCATCATCATTG	67
7	AOX1TT rev	ATAGCGGCCGCACAAACGAAGGTCTC	72
8	AOX1 Δ 1forw	CAACACCCACTTTAGGCTACTAACACCATGACTTTATTAG	71
9	AOX1 Δ 1rev	GTTAGTAGCCTAAAGTGGGTGTTGAGGAGAAGAG	70
10	AOX1 Δ 2forw	GTTCATGTTTGTAGATGAGGGCTTTCTGAGTG	67
11	AOX1 Δ 2rev	GCCCTCATCTACAAACATGAACCTCGCCAG	71
12	AOX1 Δ 3forw	GAGGGCTTTCCCAAATGGCCCAAACTG	70
13	AOX1 Δ 3rev	CCATTGGGAAAGCCCTCATCTGGAGTG	70
14	AOX1 Δ 4forw	CGGCCAGTTGTTGGTATTGATTGACGAATGC	69
15	AOX1 Δ 4rev	CAATACCAACAACCTGGCCGTTAGCATTTTC	71
16	AOX1 Δ 5forw	GCTTCTGAACCTTGCTCTCCACATTGTATGCTTC	68
17	AOX1 Δ 5rev	GTGGAGACAAGGTTTCAGAAGCGATAGAGAGAC	68
18	AOX1 Δ 6forw	GTCTCCACACTGCTGATAGCCTAACGTTTC	66
19	AOX1 Δ 6rev	GGCTATCAGCAGTGTGGAGACAATGCATAATCATC	71
20	AOX1 Δ 7forw	GGAATACTGCTCTAACCCTACTTGACAGC	65
21	AOX1 Δ 7rev	GTAGGGGTTAGAGCAGTATTTCCACCAGAATC	67
22	AOX1 Δ 8forw	CTTGACAGCAAGCTGCCCTGTCTTAAACC	66
23	AOX1 Δ 8rev	GGGCAGCTTGCTGTCAAGTAGGGGTTAG	68
24	AOX1 Δ 9forw	CTGTCTTAAACCTTACTGGTTCCAATTGACAAGC	68
25	AOX1 Δ 9rev	GGAACCAGTAAGGTTTAAGACAGGGCAGC	69
26	423forw	GATACACTAGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCC	87*
27	1372forw	GTGAAGGTGATGTACATACGGAAGCTTACCCTTAAATTTATTGTC	81*
28	2325forw	CGTGGCCGAGGAGCAGGACTGACACGTTCTAGACCATCATC	86*
29	AOX1_col	TCCAAAGACGAAAGGTTGAATG	72
30	GFPprev	CCGTATGTAGCATCACCTTCACC	74

[0153] * 使用方程式 2 (QuikChange®多位点定向诱变试剂盒) 计算 T_m

[0154] 实施例 1.1 : 报告构建体的克隆

[0155] 使用 GFP-Zeo 作为由 AOX1 启动子变体驱动的基因表达的报告分子 (reporter)。构建 ATG 起始密码子周围的序列以满足用于在酵母中高表达基因的 Kozak 共有序列的最低要求。为了改变 GFP-Zeo 基因之前的启动子区, 通过重叠延伸 PCR 插入 EcoRI 限制位点 (表 8)。

[0156] 表 8 : 用于本实施例的 AOX1 序列 (源自 pPICZ) 与巴斯德毕赤酵母菌株 NRRLY-11430 (Genbank AN:U96967, [2]) 的 AOX1 序列之间翻译起始位点和周围序列的比较。将 EcoRI 限制位点用下划线标记, 并且在位置 -3 和 +4 用粗体字母标记最低 Kozak 要求。

[0157]

			-3	+1	+4	
P(AOX1)-GFP	AAAACAACTA	ATTATTgaAa	gaattc	AACc	ATGGCTAgCa	
AOX1 (U96967)	AAAACAACTA	ATTATTcgA-	-----	AACg	ATGGCTAtCc	

[0158] 使用 10ng 载体 pPICZ-B ARS1 作为模板来扩增基于 PCR 产生的报告体系组分 P(AOX1)。反应也在合适缓冲条件中的 50 μ l 终体积中含有 10pmol 的每种引物 (分别是 P(AOX1)forw 和 P(AOX1)rev)、200 μ M 的每种 dNTP 和 2.5U Synergy™ 聚合酶。

[0159] 将 AOX1TT 类似于 AOX1 启动子进行扩增。在这个反应中使用 AOX1TTforw 和 AOX1TTrev 作为引物。两种 PCR 反应均在循环变温器中进行 30 个循环 (95°C, 1 分钟; 55°C, 30 秒; 68°C, 2 分 30 秒), 以及在 95°C 5 分钟的初始变性步骤和在 68°C 10 分钟的最终延伸步骤。将 2 μ l 第一轮 PCR 用于与以上相同条件的第二轮扩增。唯一的区别是延伸温度增加至 72°C。

[0160] 使用 25ng 的载体 pTracer™-CMV2 作为模板扩增 GFP-Zeo [19]。反应也在合适缓冲条件中的 50 μ l 终体积中含有 10pmol 的每种引物 (分别是 GFP-Zeoforw 和 GFP-Zeo rev)、200 μ M 的每种 dNTP 和 2.5U Synergy™ 聚合酶。PCR 反应在循环变温器中 (参见表 8) 进行 30 个循环 (95°C, 1 分钟; 55°C, 30 秒; 72°C, 2 分 30 秒), 以及在 95°C 5 分钟的初始变性步骤和在 72°C 10 分钟的最终延伸步骤。

[0161] 在重叠延伸 PCR 之前将所有 PCR 产物通过琼脂糖凝胶电泳纯化。反应在合适缓冲条件中的 50 μ l 终体积中含有 10ngP(AOX1)、5ngAOX1TT 和如上制备的作为模板的 GFP-Zeo、200 μ M 的每种 dNTP 和 2.5U Synergy™ 聚合酶。PCR 反应在循环变温器中 (参见表 8) 进行 30 个循环 (95°C, 1 分钟; 53°C, 50 秒; 68°C, 3 分 30 秒), 以及在 95°C 5 分钟的初始变性步骤和在 68°C 10 分钟的最终延伸步骤。10 个循环之后添加 10 μ l 的混合物, 所述混合物含有合适的缓冲条件中的 10pmol 的外部引物 P(AOX1)forw 和 AOX1TTrev、还有 200 μ M 的每种 dNTP 和 2.5U Synergy™ 聚合酶。在这次添加之后如程序设定继续 PCR。将具有大约 2.4kb 期望大小的所得 PCR 产物在琼脂糖凝胶上纯化。将纯化的产物克隆到 pCR® 4Blunt-TOPO®

载体中并测序。测序揭示在报告构建体内的 4 个突变和 1 个缺失。

[0162] 碱基对缺失位点存在于原始启动子序列的位置 -15。因为这个位置在所有 pPICZ 载体 (A、B 和 C; 在 SfuI 限制位点内) 的多克隆位点之内, 缺失不应该影响启动子活性并且因此不正确。第一个突变 (T→C) 存在于启动子区位置 -828。其它三个突变存在于 GFP-Zeo 编码序列内分别在位置 +122、+507 和 +1075。

[0163] 位置 +122 的 G→A 转变将 Gly 的 GGA 密码子变成 GAA 密码子, 导致 G41A 氨基酸改变。位置 +507 的 T→C 转变是沉默突变, 仅改变 R169 的密码子。位置 +1075 的最后的突变 (T→C) 将 TGA 终止密码子变成精氨酸密码子 CGA。在构建 pAOX 载体之后, 将突变 -828、+122 和 +1075 使用 **QuikChange®** 多位点定向诱变试剂盒修复。不改变位置 +507 的沉默突变和多聚接头中的突变, 因为其不引入稀有密码子。

[0164] 通过使用 KpnI 和 NotI 从 **pCR®4Blunt- TOPO®** 载体切除 PAOX1-GFP-Zeo-AOX1TT 片段并且将其插入 **pBlueScript® SK-** 载体的 KpnI 和 NotI 位点之间, 来构建 pAOX。

[0165] 使用 **QuikChange®** 多位点定向诱变试剂盒 (Stratagene, Amsterdam, The Netherlands) 修正存在于 AOX1 启动子和 GFP-Zeo 序列中的突变。根据提供的操作手册进行 PCR 反应, 其在合适的缓冲条件中的 25 μ l 终体积中含有 100ng pAOX、100ng 的诱变引物 (分别是 423forw、1372forw 和 2325forw) 和 1 μ l **QuikChange®** 多酶混合物, PCR 反应在循环变温器中进行 30 个循环 (95°C, 1 分钟; 55°C, 1 分钟; 65°C, 10 分 30 秒), 以及在 95°C 1 分钟的初始变性步骤。根据提供的操作手册进行 DpnI 消化和化学转化入大肠杆菌 XL10- **GOLD®** (Invitrogen Corp.) 细胞。所有 3 种突变的改正通过测序鉴定。

[0166] 实施例 1.2 : AOX1 启动子缺失的构建

[0167] 使用 P(AOX1)forw 作为正向引物和 AOX n rev (n = 1...9) 作为反向引物合成 AOX1 启动子的左臂。使用 10pmol 的 AOX n forw (n = 1...9) 作为正向引物和 P(AOX1)rev 作为反向引物合成右臂。使用 12ng 载体 pAOX 作为模板和 10pg 的每种引物合成全部的臂。反应同样在合适的缓冲条件中, 在 50 μ l 的终体积中含有 10pmol 的每种引物、200 μ M 的每种 dNTP 和 0.6U Pwo DNA 聚合酶。PCR 在循环变温器中进行 30 个循环 (95°C, 1 分钟; 55°C, 1 分钟; 68°C, 1 分 30 秒), 以及在 95°C 5 分钟的初始变性步骤和在 68°C 10 分钟的最终延伸步骤。在将所有臂用作重叠延伸 PCR 的模板之前将其在琼脂糖凝胶上纯化。

[0168] 表 9 : 用于启动子缺失的重叠引物对和臂长

[0169]

构建体	左臂		右臂	
	内部引物	臂长[bp]	内部引物	臂长[bp]
PAOX1Δ1	AOXΔ1 rev	184	AOXΔ1 forw	738
PAOX1Δ2	AOXΔ2 rev	315	AOXΔ2 forw	624
PAOX1Δ3	AOXΔ3 rev	374	AOXΔ3 forw	578
PAOX1Δ4	AOXΔ4 rev	519	AOXΔ4 forw	421
PAOX1Δ5	AOXΔ5 rev	636	AOXΔ5 forw	290
PAOX1Δ6	AOXΔ6 rev	708	AOXΔ6 forw	247
PAOX1Δ7	AOXΔ7 rev	742	AOXΔ7 forw	209
PAOX1Δ8	AOXΔ8 rev	794	AOXΔ8 forw	171
PAOX1Δ9	AOXΔ9 rev	833	AOXΔ9 forw	115

[0170] 反应在合适的缓冲条件中的 50 μl 终体积中含有如上所述制备的作为模板的 10ng 每种臂、200 μM 的每种 dNTP 和 0.6U Pwo DNA 聚合酶。PCR 在循环变温器中进行 30 个循环 (95℃, 45 秒; 60℃, 45 秒; 68℃, 2 分钟), 以及在 95℃ 5 分钟的初始变性步骤和在 68℃ 10 分钟的最终延伸步骤。在 10 个循环之后添加 20 μl 混合物, 所述混合物含有在合适缓冲条件中的 10pmol 外部引物 P(AOX1) forw 和 P(AOX1) rev, 还有 200 μM 的每种 dNTP 和 0.6UPwo DNA 聚合酶。在这次添加之后如程序设定继续 PCR。

[0171] 将具有大约 898-947bp 期望大小的所得 PCR 产物在琼脂糖凝胶上纯化并且克隆到 pCR®4Blunt-TOPO® (Δ2、Δ4、Δ5、Δ7 和 Δ8) 或 pCR®-Blunt II-TOPO® 载体 (Δ1、Δ3、Δ6 和 Δ9) 中并且测序。

[0172] 通过使用 BglIII 和 EcoRI 从 TOPO® 载体切除 P_{AOX1} Δ 片段并且将它们插入到 pAOX 载体的 BglIII 和 EcoRI 位点之间取代野生型 AOX1 启动子, 来构建 pAOX Δ 载体。所得载体通过测序验证。

[0173] 实施例 1.3: 巴斯德毕赤酵母转化和转化体分析

[0174] 如前所述进行巴斯德毕赤酵母转化。通过在 MSM-Zeo 琼脂平板上以等分试样涂布转化和再生的毕赤酵母属细胞来选择 P_{AOX1} (或 P_{AOX1} Δ)-GFP-Zeo-AOX1TT 的整合。

[0175] 将巴斯德毕赤酵母菌株在深孔平板中培育, 所述深孔平板每孔含有 300 μl BMD(1%), 在 28℃、320rpm 和 80% 湿度于室温生长 60 小时。这段时间之后, 取 50 μl 用于 GFP 荧光的测定。通过添加 250 μl BMM2/ 孔然后再温育 72 小时来进行诱导。通过在 10、24 和 48 小时之后添加 50 μl BMM10 补充甲醇。甲醇诱导 72 小时后再次测量 GFP 荧光。

[0176] 分析报告酶的表达通过 Spectramax Gemini XS 平板读数器中具有在 395nm 的激发和在 507nm 的发射的 GFP 荧光检测来分析巴斯德毕赤酵母中 GFP-Zeo 的表达。将如上所述在深孔平板中培养的 50 μl 巴斯德毕赤酵母培养物在 FIA 微量滴定板中用 ddH₂O 稀释 1+3。由于有限的样品量, 只进行一次测量。所有给出的平均值 ± 标准偏差由至少 3 个不同培养物 (孔) 计算。

[0177] 如果将整合盒整合到 AOX1 基因座而不取代 AOX1 基因, 重组的毕赤酵母属菌株能够在甲醇上以野生型速率生长, 而通过双交换取代 AOX1 基因导致在甲醇上低得多的生长速率。将这两种生长表型分别称为甲醇利用正型 (methanol utilization plus) (Mut^+) 和甲醇利用缓慢型 (Mut^s)。为了分析甲醇利用表型, 使用 96 针复制器将巴斯德毕赤酵母小规模培养物转移到 MM 和 MD 琼脂平板上并且在 30°C 温育 2 天。2 天之后, 如果毕赤酵母属菌株具有 Mut^+ 表型则在两个平板上均出现菌落, 而具有 Mut^s 表型的菌株仅在 MD 平板上出现菌落。

[0178] 将由 pAOX 或 pAOX Δ 质粒之一的转化得到的所有毕赤酵母属菌株通过菌落 PCR 分析, 并且缺失构建体还通过测序保证启动子序列在报告基因之前 (GFP-Zeo)。

[0179] 实施例 1.4 : AOX1 启动子的定向进化

[0180] 虽然在基因编码区上的 PCR 诱变得到的良好的发展和建立, 然而对于在启动子区的诱变一无所知。由于知识的缺乏进行几种诱变条件: 为了最小化突变谱内的偏好, 使用两种不同的聚合酶, Taq DNA 聚合酶和 **Mutazyme®** DNA 聚合酶 (Stratagene Inc.)。由于完全缺乏在启动子序列进化的突变频率方面的知识的事实, 测试几种突变频率 (理论上 1 至 ~14/kb)。

[0181] 使用 Hot Star **Taq®** DNA 聚合酶诱变: 根据 [20] 在 100 μ l 反应体积中于启动子序列上进行诱变 PCR。反应含有在合适缓冲条件中的 5U Hot Star **Taq®** DNA 聚合酶、12ng pAOX、40pmol 的每种引物 (P(AOX1)forw 和 P(AOX1)rev) 和 dNTPs (200 μ M dGTP、200 μ M dATP、1mM dTTP、1mM dCTP)。将 $MgCl_2$ 浓度增加至 7mM (通常为 3mM) 以改变聚合酶的错误率。PCR 在循环变温器中进行 30 个循环 (95°C, 45 秒; 55°C, 45 秒; 72°C, 1 分 30 秒), 以及在 95°C 15 分钟的初始变性步骤和在 72°C 10 分钟的最终延伸步骤。

[0182] 根据提供的操作手册在 50 μ l 的终体积中在启动子序列上进行 **GeneMorph®** 随机诱变试剂盒的操作。使用不同量的载体 pAOX 作为模板 (参见表 10)。使用 12.5pmol 的每种引物, P(AOX1)forw 和 P(AOX1)rev。PCR 反应在循环变温器中进行 30 个循环 (95°C, 30 秒; 55°C, 30 秒; 68°C, 1 分 30 秒), 以及在 95°C 1 分钟的初始变性步骤和在 68°C 10 分钟的最终延伸步骤。

[0183] 表 10 : **GeneMorph®** PCR 反应中所用的模板量

[0184]

No.	突变频率	pAOX 量	期望的突变 /kb
1	低 - 中	12ng	~ 3 或更低
2	中	1.2ng	3-7
3	中 - 高	120pg	~ 7 或更高

[0185] 进行具有上述条件 (Taq, 3x **GeneMorph®**) 的第一轮诱变。为了获得更高的诱变频率, 使用 **GeneMorph®** 反应 #3 作为模板用于第二轮 PCR。使用 Taq 和 **GeneMorph®** #2 和 #3 条件。

[0186] 在转化进入巴斯德毕赤酵母 X-33GFP-Zeo Mut^s A9 细胞之前, 将所有 PCR 反应如前所述地沉淀和脱盐。使用标准转化和再生方法。通过如下方法来选择在葡萄糖培养基中诱导的启动子: 在含有 100-500 μ g/ml Zeocin™ 的 MD 琼脂平板上涂布 150 μ l 等分试样的转化细胞悬浮液并且在 30°C 温育 2 天。

[0187] 实施例 1.5 : 结果与讨论

[0188] I) 报告体系的表征

[0189] 迄今为止,多种 GFP 变体在分子生物学中使用。尽管仅在少数点突变上有所不同,它们的特征相差甚远。除了改进的折叠性质 (folding property),它们的荧光光谱以及它们的量子产率并且因此在强度上都大不相同。依赖于它们的激发最大值 (excitation maximum),能够将绿色荧光蛋白分成两个主要组:在 395nm 具有激发最大值和 470nm 的较小最大值 (minor maximum) 的野生型 GFP 变体,和在 480-490nm 具有激发最大值的红移 GFP 变体。根据其氨基酸序列,环 -3-GFP (cycle-3-GFP) 属于野生型 GFP 变体组,在 395nm 具有激发最大值。

[0190] 为了在巴斯德毕赤酵母中表达时控制光谱性质,测定荧光光谱。GFP-Zeo 中环 -3-GFP 的总激发最大值是 395nm,而在 478nm 的第二最大值是渐消失的。发射光谱揭示 510nm 的发射最大值。操作手册建议的两种激发波长中,395nm 是优选的并且将其用于所有进一步的测量。

[0191] 自吸收在荧光光谱学中是非常频繁的现象。在高浓度荧光团 (fluorophor) 时,能够将处于与吸收 (激发) 光谱重叠的区域的发射的质子吸收 (辐射能量转移)。如果将发生自吸收 (发射内部滤光效果 (emission inner filter effect)),将观察到较低的荧光强度。这导致启动子活性的低估。没有内部滤光效果的情况下,荧光强度以线性方式随荧光团增加而增加。因此测试表达 GFP-Zeo 的巴斯德毕赤酵母细胞增加体积的荧光活性。

[0192] 在细胞水平至多 3000RFU 无发射内部滤光效果是能够检测的。不能够评估在细胞内由 GFP 的积累引起的自吸收。检测到在全部 72 小时诱导阶段荧光的线性增加。由于该原因,细胞内的内部滤光效果看来是不可能的。因此细胞核内 GFP-Zeo 的积累对于其定量没有问题。本研究中测定的单拷贝启动子活性的范围内未发生内部滤光效果。由于缺乏自吸收,启动子活性的低估不太可能发生。其他人观察到的内部滤光效果最可能的原因是使用不同的 GFP 变体和低得多的 Stokes 位移,并且因此重叠激发和发射光谱。当比较 GFP 表达实验的结果时必须需要小心。使用几种 GFP 变体,所述 GFP 变体具有截然不同的光谱性质,还有最优化的密码子使用和因此在不同表达宿主中非常不同的表达水平,使不同实验室结果的可比性变得复杂。

[0193] II) 微量 AOX1 启动子活性

[0194] 小规模培养微生物细胞 (例如酵母、细菌) 通常以摇瓶培养进行。在摇瓶中接种和培养大微生物文库是劳动和时间密集型的而导致高成本。近年来开发了使用深孔微量滴定板的微量培养体系作为备选方案。由于平行操作例如 96 或 384 个菌株 / 培养物并且需要较少材料,微量滴定体系在劳力、时间并且因此在成本强度方面优于摇瓶培养。由于几种原因,微量滴定体系主要的缺点:小样品体积和低通气效率是较不相关的:(1) 分析系统中的技术进步致使大量化合物较低检测限制,这使得所需样品体积非常低;(2) 用于在深孔微量滴定板中培育的方法和设备也已改进。已经在少数研究中证明,微量滴定板中的通气速率和因此的生长条件与摇瓶相似。也已经使用环 -3-GFP 作为报告蛋白质证明对于酿酒酵母的 GAL1 启动子的实时研究与摇瓶研究一致。

[0195] 如上所述在深孔微量滴定板中研究 AOX1 启动子驱动的 GFP-Zeo 表达。细胞在葡萄糖上生长之后,接着是使用甲醇做为碳源和能量源的诱导阶段。巴斯德毕赤酵母细胞中使用甲醇的 AOX1 启动子诱导导致 GFP 荧光的快速增加,所述巴斯德毕赤酵母细胞具

有 PAOX1-GFP-Zeo-AOX1TT 表达盒。GFP 荧光以线性方式增加直到 72 小时。如果添加甲醇, GFP-Zeo 的表达将继续。如果不添加甲醇, 甲醇通过蒸发和消耗在 24 小时内耗尽并且 GFP-Zeo 表达减少至去阻抑水平。

[0196] GFP-Zeo 荧光的增加也和 SDS-PAGE 显示的 GFP-Zeo 蛋白质一致。在甲醇诱导下, 出现的约 42kDa 的蛋白质条带随荧光增加变得更强。42kDa 处的强条带存在于全部 GFP-Zeo 克隆中, 而在阴性对照 (X-33 野生型) 中无条带出现。同样在 X-33d6*F10 的样品中, 72 小时甲醇诱导之后发现强条带 (图 1C, 泳道 5)。尽管未归一化, 42kDa 条带的强度和合适的荧光水平之间清楚的相关性是可以估计的。

[0197] III) 转录因子结合位点

[0198] 如前所述, 几种转录因子结合位点的共有序列是已知的。AOX1 启动子序列的序列分析揭示了几种推定的转录因子结合位点, 具有少数引起特别兴趣的结果。在热激因子和胁迫应答元件基序之中, 发现少数通常已知涉及葡萄糖调节的转录因子结合位点。最感兴趣的结合位点总结于表 11 和图 2。

[0199] 表 11 : 在 AOX1 启动子序列内存在的转录因子 (TF) 结合位点。大写字母的碱基对表示核心序列 (core sequence) (基质中 4 种最保守、连续性的残基), 以下划线标记的碱基对表示高信息含量 ($C_i > 60$, 最大值为 100)。

[0200]

TF 基质	位置 (5)*	缺失 变体	核心 相似性	基质 相似性	序列	Seq ID No.
HAP1. 01	52-66 (-902 至 -888)		1. 000	0. 802	ctgtggatgt <u>CGG</u> At	31
HSF. 01	135 至 155 (-819 至 -799)		1. 000	0. 784	AGAAgaggagtgga ggtcctg	32
HAP234. 01	193 至 205 (-761 至 -749)	Δ 1	1. 000	0. 895	caageCCAAtaac	33
HAP234. 01	203 至 215 (-751 至 -739)	Δ 1	1. 000	0. 923	gagctCCAAtcaa	34
ABAA. 01	213 至 227 (-741 至 -727)	Δ 1	1. 000	0. 949	ctcgctCATTccaat	35
STRE. 01	279 至 287 (-675 至 -667)		1. 000	1. 000	ccAGGGggg	36
RAP1. 01	332 至 346 (-622 至 -608)	Δ 2	1. 000	0. 845	tacACCCgaacatca	37
ADR1. 01	371 至 377 (-583 至 -577)	Δ 3	1. 000	1. 000	tGGGGtc	38
HSF. 03	516 至 536 (-438 至 -418)	Δ 4	1. 000	0. 862	AGAAacttccaaaag tcggc	39
HAP234. 01	665 至 677 (-289 至 -277)	Δ 5	1. 000	0. 883	atcatCCAaaaag	40
MAT1MC. 0 1	680 至 690 (-274 至 -264)		1. 000	0. 901	tgcaTTGTctc	41
GCR1. 02	699 至 713 (-255 至 -241)	Δ 6	1. 000	0. 872	atgCTTCaagattc	42

QA-1F. 01	743 至 763 (-211 至 -191)	Δ 7	0.785	0.775	acagttaaatttTGATc atga	43
-----------	----------------------------	-----	-------	-------	---------------------------	----

[0201] * 所给的位置相对于 GFP-Zeo 基因的翻译起始位点 (ATG) 标记 ; 推定的转录因子结合位点的核心序列以大写字母显示。

[0202] c 表示与互补链的同源性。

[0203] IV) 甲醇调节基因中的调节序列

[0204] 文献中描述的几种序列涉及甲醇诱导基因的调节。基于巴斯德毕赤酵母 AOX2 启动子的缺失分析描述三种调节区域, 两种负作用区域 (negative acting region) (URS1 和 URS2, 上游阻抑序列) 和正作用结构域 (positive acting domain) (UAS, 上游活化序列) [3]。对于多形汉逊酵母 MOX 启动子, 也描述了两种上游激活序列 (UAS1 和 UAS2) 和一种阻抑物结合位点 (URS1) [8]。

[0205] V) AOX1 启动子的缺失构建体

[0206] 基于转录因子分析和多序列比对, 选择 9 种启动子区用于如前所述的通过重叠延伸 PCR 的缺失。将 AOX1 启动子缺失构建体克隆到 pAOX 载体中以取代“野生型 AOX1”启动子 5' 至报告基因 GFP-Zeo。将质粒线性化并且整合到巴斯德毕赤酵母基因组中。

[0207] 表 12 : AOX1 启动子构建体中缺失的序列

[0208]

构建体	位置*		序列	Seq ID No.
	5'端	3'端		
PAOX1Δ1	170 -784	235 -719	ttgccatcgaaaaaccagcccagttattgggctt gattggagctcgcctcattccaattccttcta	44
PAOX1Δ2	304 (-650)	350 (-604)	ttattccgaatgcaacaagctccgcattacaccc gaacatcactcc	45
PAOX1Δ3	364 (-590)	393 (-561)	ctgagtgtgggggtcaaatagtttcatgttc	46
PAOX1Δ4	509 (-445)	551 (-403)	gtcaaaaagaaacttccaaaagtcggcataccg ttgtcttgt	47
PAOX1Δ5	625 (-329)	683 (-271)	ccggtgcacctgtgccgaaacgcaaattgggga aacaccgcgtttttggatgattatgca	48
PAOX1Δ6	694 (-260)	723 (-231)	attgtatgctccaagattctggtgggaat	49
PAOX1Δ7	729 (-225)	763 (-191)	tgatagcctaacgttcatgatcaaaatttaactgt	50
PAOX1Δ8	784 (-170)	800 (-154)	aatatataaacagaagg	51
PAOX1Δ9	823 (-131)	861 (-93)	ttttttatcatcattattagcttactttcataattgcg	52

[0209] * 所给的位置相对于 Seq ID No. 1 标记

[0210] 如上所述分析整合体的 GFP-Zeo 表达和正确的启动子序列在 GFP-Zeo 基因之前的整合。进一步详细分析单拷贝整合体在不同碳源中的微量 GFP-Zeo 表达水平。在全部构建体（缺失和野生型）中，只要葡萄糖或甘油存在于培养基中（有或无甲醇），则不能检测到 GFP 荧光。碳饥饿（其代表去阻抑条件）时，检测到 GFP 荧光的轻微增加。与野生型相比，一些启动子变体显示显著的区别（图 3）。在去阻抑条件下在 6 种构建体（ $\Delta 3$ 、 $\Delta 4$ 、 $\Delta 5$ 、 $\Delta 7$ 、 $\Delta 8$ 和 $\Delta 9$ ，参见图 3）中发现显著较低的启动子活性。 $\Delta 1$ 具有野生型活性，而构建体 $\Delta 2$ 和 $\Delta 6^*$ 得到显著较高的 GFP-Zeo 表达。后者的表达水平显著高于野生型水平。

[0211] 甲醇诱导时，所有变体显示显著减少的启动子活性，仅有一个例外： $\Delta 1$ 产生比野生型高约 20% 的活性。所有其它变体活性的减少是相当显著的，可参见图 4。

[0212] 根据甲醇诱导野生型活性归一化的所有变体和野生型构建体的启动子活性总结于表 13。

[0213] 表 13：微量 AOX1 启动子变体的荧光强度。数据表示 4 个独立测量的平均 SD。WT 启动子在 72 小时甲醇诱导之后的荧光强度（100%）是 987 ± 81 。只要葡萄糖存在于培养基中则检测不到荧光。

[0214]

构建体	相对荧光强度[%]	
	去阻抑	甲醇
PAOX1	2.8 ± 0.1	100
PAOX1 $\Delta 1$	3.0 ± 0.5	120 ± 12
PAOX1 $\Delta 2$	4.4 ± 0.8	40 ± 3
PAOX1 $\Delta 3$	0.7 ± 0.2	68 ± 8
PAOX1 $\Delta 4$	1.9 ± 0.1	72 ± 4
PAOX1 $\Delta 5$	0.23 ± 0.04	30 ± 4
PAOX1 $\Delta 6^*$	9.1 ± 0.6	42 ± 2
PAOX1 $\Delta 7$	2.2 ± 0.4	31.3 ± 0.5
PAOX1 $\Delta 8$	0.3 ± 0.2	17.1 ± 0.7
PAOX1 $\Delta 9$	1.3 ± 0.1	61 ± 3

[0215] 构建体 $\Delta 8$ 中 TATA 框的缺失导致启动子严重的破坏，在去阻抑和诱导条件时分别大约 90% 和 80% 严重降低的活性。通过消除实验测定 (Ellis, S. B., et al., Mol. Cell. Biol. (1985) 5:1111-1121) 的转录起始 (transcription initiation start) ($\Delta 9$)，未观察到对于表达水平的这种强烈影响。其是甲醇诱导之后最好的缺失构建体之一。和期望的一样，TATA 框对于转录水平具有强烈的影响。相反转录起始似乎不像 TATA 框一样重要。距离 TATA 框一定距离内的另一个区域可能在将原有的转录起始缺失之后作为转录起始起作用。能够推测这种缺失在表达过程的几个阶段（例如转录起始、mRNA 稳定性、翻译起始）的效

果,因为缺失改变了 mRNA 的 5' 端。

[0216] 仅两种构建体 $\Delta 2$ 和 $\Delta 6^*$ 在去阻抑之后显示显著较高的表达水平。将 Rap1p 和 Gcr1p 推定的转录因子结合位点包括在缺失的序列之中。此外,推定的 QA-1F 转录因子结合位点与 $\Delta 6^*$ 的缺失序列非常接近。值得注意的是,已知 Rap1p 和 Gcr1p 结合位点存在于启动子序列中时以协同方式 (synergistic manner) 起作用 [21]。依赖于普遍性转录因子 Rap1p 的结合位点和合适的转录因子的序列背景,Rap1p 具有不同的细胞功能(例如端粒结构、杂交 (mating)、翻译、糖酵解) [22-24]。如前所述,Gcr1p 是调节和协调糖酵解基因的主要元件,并且对于酿酒酵母中的高水平表达是绝对必需的。Rap1p 和 Gcr1p 的结合位点存在于糖酵解基因上游激活序列 (UAS) 核心区域的附近处并且通过 Rap1p 将 DNA 弯曲来减轻 Gcr1p 结合。另一方面,相邻的 Rap1p 结合位点不是依赖 Gcr1p 的基因活化的绝对必要条件。看起来当较高数量的 CT 框 (CT-box) 存在时,Gcr1p 能够促进与其结合位点的结合。虽然描述了 Rap1p 与 Gcr1p 以及 Gcr1p 与 Gcr1p 清楚的相互作用,但是提出了一些其它因子与 Gcr1p 和 / 或 Rap1p 相互作用调节复合物的活性。在过去三十年中获得了诱导机制的广泛认识。

[0217] 在 AOX1 启动子序列中找不到如上所述的功能性 UAS 中 Gcr1p 和 Rap1p 结合位点的基本上紧密的接近性。相反,两种结合位点距离 367bp。在推定的 Gcr1p 结合位点中,其核心序列 CTTCC 在 AOX1 启动子序列中以 2 倍存在,但它们之中没有与 Rap1p 结合位点或另外的 CTTCC 基序直接相邻。因此存在于许多糖酵解基因中的这两种结合位点的协同作用看起来是不可能的。由于 Rap1p 和 Gcr1p 的推定作用是在去阻抑条件下用于 AOX1 的阻抑物蛋白,用于这种推定的新细胞功能的两种蛋白质(相互)作用的新形式是可能的。

[0218] 通过观察无甲醇诱导时具有非常高的 GFP-Zeo 表达的多拷贝菌株,强调在碳饥饿时的阻抑中涉及 $\Delta 6^*$ 缺失(包括推定的 Gcr1p 结合位点)。 $\Delta 1$ - $\Delta 9$ 系列最佳克隆的 GFP-Zeo 表达,称为巴斯德毕赤酵母 X-33d6*F10,示于图 5。在这种 $\Delta 6^*$ 多拷贝菌株中去阻抑之后的 GFP-Zeo 表达(60 小时,微量)比单拷贝野生型启动子菌株 (X-33GFP-Zeo D2) 在甲醇诱导之后的表达高大约 10%。甲醇诱导之后巴斯德毕赤酵母 X-33d6*F10 的表达水平也远高于具有野生型启动子的多拷贝菌株 (X-33GFP-Zeo E2)。

[0219] 巴斯德毕赤酵母 AOX1 和 DAS1 和多形汉逊酵母 MOX 启动子区启动酿酒酵母中报告酶 β -半乳糖苷酶(大肠杆菌的 lacZ) 的表达 [9]。酿酒酵母中这些基因的调节模式与它们的天然宿主相似:葡萄糖阻抑基因表达。在碳饥饿条件时轻微地去阻抑表达,并且甘油作为碳源诱导表达。在酿酒酵母中 AOX1 和 DAS1 调节区域控制下的 β -半乳糖苷酶表达水平与得自具有强酿酒酵母 CYC1(组成型)和 GAL2(诱导型)启动子的表达水平可比较 [9]。已经证明酿酒酵母中由 MOX 启动子驱动的表达也由乙醇、甲醇和油酸诱导。另外非常重要的发现是启动子的去阻抑 / 诱导中涉及 Adr1p。Adr1p 是酿酒酵母中 ADH2(醇脱氢酶 2) 和一些过氧化物酶体蛋白的正效应物 [25],当培养基中缺乏葡萄糖时 Adr1p 也是 MOX 启动子的正效应物。

[0220] 如前所述由于当甘油存在时 MOX 的去阻抑,AOX1 和 MOX 基因的调节模式在它们的天然宿主中显著不同。使用多形汉逊酵母中的 AOX1 启动子区揭示在异源宿主中甘油不阻抑 AOX1 启动子 [26]。因此,异源 AOX1 启动子的调节看起来如同源的 MOX 启动子一样。这导致如下提示:巴斯德毕赤酵母和多形汉逊酵母之间在调节模式中显著的不同是由于在这

两种酵母中对不同碳源的总体转录反应。言下之意是尽管甘油和葡萄糖阻抑机制是与在巴斯德毕赤酵母中（部分）相同的，但在多形汉逊酵母（和酿酒酵母中一样）的情况不同并且甘油不使用葡萄糖阻抑机制。

[0221] 存在于 AOX1 启动子序列中的三种推定的 HAP2/3/4/5 结合位点中的两种在 $\Delta 1$ 缺失构建体内，而第三种在 $\Delta 5$ 中。 $\Delta 1$ 的序列缺失导致甲醇诱导时启动子活性的增加，但未观察到对于去阻抑启动子水平的影响。相反， $\Delta 5$ 的缺失导致在去阻抑以及诱导条件下启动子活性的严重减少。在 $\Delta 1$ 缺失中，发现推定的构巢曲霉 (*Aspergillus nidulans*) abaA 结合位点。abaA 基因产物是转录激活子，所述转录激活子涉及构巢曲霉中的分生孢子梗（无性繁殖器）发育 [27]。因为所有推定的结合位点是可能的激活子序列 [27]，它们的缺失应该对 $\Delta 5$ 构建体中存在的表达水平具有负效应。由于两种缺失均非常长的事实，另外的结合位点可能是造成观察到的效果的原因。 $\Delta 1$ 的缺失对表达水平具有相反的效果的事实说明推定的结合位点之一是阻抑物基序，或存在另外的结合位点，其超越了缺失推定的 HAP 和 abaA 结合位点的效果由此增加表达水平。

[0222] 但是，已知 HAP 复合物在酿酒酵母中负责正向调节涉及呼吸作用和能量代谢的基因。呼吸作用的调节由氧水平以及存在于培养基中的碳源控制，二者均由 Hap 复合物介导。在发酵酵母酿酒酵母中，葡萄糖阻抑几种基因和因此的呼吸链功能以及柠檬酸循环。呼吸作用的葡萄糖阻抑部分通过 Hap 复合物介导，即通过只要葡萄糖存在时 Hap4p 的缺失。相反，氧依赖调节似乎由 Hap1p 调节 [28]。复合物从呼吸作用酵母乳酸克鲁维酵母中分离 Hap 复合物基因的同源物。涉及呼吸作用的基因在呼吸作用酵母中组成型表达，即使存在葡萄糖。迄今为止，已经显示来自 Hap 复合物的几乎每种呼吸链基因均独立地调控 [29]。Hap 复合物的作用看起来在调整碳和氮的同化作用中，如同样在酿酒酵母 [30] 和构巢曲霉 [29] 中发现的一样。

[0223] 甲醇利用途径中的第一步是氧消耗，主要由巴斯德毕赤酵母中的 AOX1 基因产物催化。通过氧气调节涉及能量代谢的大多数基因和几乎每种编码氧消耗酶的基因，主要通过 Hap1p 和 / 或 Hap2/3/4/5p [28]。当生长在作为唯一能量源和碳源的甲醇上时，甲醇利用途径导致碳同化作用和能量产生。AOX1 启动子调节中涉及 Hap 复合物识别基序 TTCCAA 直觉上讲得通。

[0224] 包括第二推定 HSF 结合位点的 $\Delta 4$ 构建体导致去阻抑和诱导之后启动子活性减少 30%。因此 HSF 是去阻抑和诱导条件下 AOX1 基因表达的普通的增强子。在酿酒酵母中，几种胁迫条件如热休克、氧化胁迫和葡萄糖饥饿导致 HSF 的活化。同样已经证明“代谢总开关 (metabolic master switches)”之一的蛋白激酶 Snf1p 涉及磷酸化并且因此在碳饥饿时的 HSF 活化 [31]。因此在葡萄糖饥饿（有或无诱导）时出现 AOX1 的完全活化中涉及 HSF。

[0225] 现有技术中公开了 AOX1 启动子的表达研究，其使用截短形式 (truncated version) 以及具有缺失序列的变体 [32,33]。

[0226] 表 14: Inan 等的启动子研究的结果 [32,33]；使用 0.5% 甲醇作为碳源进行诱导，使用 0.5% 甲醇和 0.5% 乙醇阻抑；起始位置表示 AOX1 启动子中相对于翻译起始位点 (ATG) 的序列 5' 端

[0227]

启动子片段	相对于 SEQ ID No.1 的缺失	相对活性[%]	
		诱导	阻抑
InanABCDEF	-	100	3.1 ± 0.3
Inan_BCDEF	7 至 152 (-947 至-802)	76 ± 5	1.9 ± 0.2
Inan__CDEF	1 至 292 (-947 至-661)	49 ± 4	2.2 ± 0.5
Inan___DEF	1 至 432 (-947 至-521)	14 ± 3	1.3 ± 0
Inan____EF	1 至 559 (-947 至-394)	24 ± 7	1.8 ± 0
Inan_____F	1 至 798 (-947 至-245)	7 ± 2	1.8 ± 0.2
InanA_CDEF	153 至 292 (-801 至-661)	63 ± 3	2.1 ± 0.2
InanAB_DEF	293 至 432 (-660 至-521)	109 ± 12	3.8 ± 0.4
InanABC_EF	433 至 559 (-520 至-394)	128 ± 6	5.0 ± 0.6
InanABCD_F	560 至 798 (-393 至-245)	16 ± 1	0.8 ± 0.2

[0228] 起始于 153(-801) (表 14) 的构建体 Inan_BCDEF 揭示了 153(-801) 上游至少一种激活蛋白的结合位点。这种激活子结合位点的候选物是用 MatInspector 发现的互补链上 Hap1p(52 至 66、-902 至 -888) 和 HSF(135 至 155、-819 至 -799) 的结合位点。在 SacI 限制位点 (210-215(-744 至 -739)) 的截短导致启动子几乎达到野生型启动子的活性 (GeoffLin Cereghino, Poster, Sixth Meeting on "Current Topics in Gene Expression and Proteomics", San Diego, October 19-22, 2003)。为了用 SacI 截短的启动子构建体 (pHWG0, GeoffLin Cereghino, poster) 达到野生型启动子水平, 阻抑物蛋白的第二结合位点可以存在于 210(-744) 的上游, 其缺失对于启动子活性具有相同影响但方向相反。阻抑物蛋白的位置是 169(-784) 和 210(-744) 之间, 因为 $\Delta 1$ 构建体 ($\Delta 169(-784)$ 至 $234(-719)$) 含有阻抑物结合位点。 $\Delta 1$ 的缺失导致启动子活性增加 20% (表 14), 其在通过缺失激活蛋白结合位点而减少的范围之内。

[0229] 通过与 $\Delta 4$ ($\Delta 508(-445)$ 至 $550(-403)$) 的比较, 阻抑物结合位点的位置能够进一步精确至 433(-520) 和 508(-445) 之间的序列, 因为 $\Delta 4$ 缺失包括正作用转录因子, 516 至

536(-438 至 -418) 处的 HSF。如果正作用 HSF(如果其是 HSF) 位于提出的区域之内,能够推测在 433 和 508(-520 和 -445) 之间的阻抑物结合位点的较强效果。如果 HSF 的结合位点位于 508 和 536(-445 和 -418) 之间的区域,另外的激活子结合位点位于 536 和 560(-418 和 -393) 之间。如果不是,则可能是相同的结合位点。因为 InanABCD_F(Δ 560 至 709(-393 至 -245)) 变体仅具有 16% 的野生型活性,同样 Δ 5 构建体(624 至 682(-329 至 -271)) 导致大约 70% 野生型水平的减少。和期望的一样,将 InanB 片段从全长启动子缺失(产生 InanA_CDEF) 以及从 Inan_BCDEF 中缺失(产生 Inan_CDEF) 分别导致较长片段 63 和 64% 的减少。相反,尽管从全长启动子缺失 C 片段导致启动子活性大约 10% 的增加,从截短的 Inan_CDEF 片段缺失导致 49 至 14% 的减少(表 14)。解释是依赖于它们的结合位点背景的转录因子的协同结合。最后的激活蛋白结合位点位于 713 和 760 之间(-24 至 -194)(Geoff Lin Cereghino, Poster San Diego)。此外,通过 Δ 7 构建体(Δ 729 至 763, -225 至 -191), 激活子的位置可以精确到下游 729(-225)。

[0230] 总而言之,发现对于 AOX1 启动子的表达水平具有强烈影响的几个区域。结合来自本文提供的实施例和来自其他作者的所有已知调节位点,排除含有 TATA 框和转录起始位点的区域,至少 10 个调节位点存在于 P_{AOX1} 启动子序列上。

[0231] 提供的数据揭示 AOX1 启动子的配合调节(orchestral regulation): 几种因素对于结合至 DNA 的最大表达水平是必需的。在诱导条件下,几种正作用转录因子(激活子)结合至 DNA,而多数阻抑物蛋白不结合导致高水平表达。尽管是去阻抑的,启动子活性仅达到诱导水平的小部分($\sim 3\%$)。这最可能是由于较少的激活子和较多的阻抑物蛋白结合至启动子区。在阻抑条件下,能够假设无激活子结合至 DNA 和几种阻抑物结合至 DNA 在阻抑条件下具有进一步增加的阻抑物/激活子比率。

[0232] 已经证明对于葡萄糖阻抑的酿酒酵母的 ADH2(醇脱氢酶 2) 启动子,其结合紧邻核小体的激活蛋白(例如 Adr1p),导致不稳定并且因此导致在去阻抑时染色质的重排。重排发生在 TATA 框和转录起始位点的区域因此增加它们的可接近性。由于较高的可接近性,出现稳定的前起始复合物的形成,因此将启动子活性增加至基础水平。在增强 P_{AOX1} 驱动表达的几种转录因子的结合中,可以假设至少对于去阻抑的相似机制。将所有数据和假设放在一起, AOX1 启动子的调节是非常复杂的,并且几种(正和负作用)转录因子的推定结合位点揭示了高协同的机制,其能够整合多种信号用于 AOX1 启动子的调节。

[0233] VI) AOX1 启动子的 PCR 诱变

[0234] 此处已经证明在转录因子结合位点的核心序列内的特定突变导致它们的效应力(effector force) 的显著改变。假定少数激活子和阻抑物蛋白作用于 AOX1 启动子以导致其非常强的调节(在葡萄糖条件下几乎无活性,在甲醇中非常高的活性)。因此 AOX1 启动子的随机诱变应该导致几种启动子变体,所述启动子变体具有破坏的或减少的阻抑物结合位点活性。进行具有不同突变率的一系列 PCR 反应。将所得启动子变体转化到巴斯德毕赤酵母 GFP-Zeo Mut^sA9 菌株中,其中将 AOX1 基因用 GFP-Zeo 菌株取代。由诱变的启动子变体取代野生型 AOX1 启动子应该以特定的比率发生。在 MD-Zeo 琼脂平板上筛选当培养基中存在葡萄糖时具有较高表达率的启动子变体。

[0235] 涂布到含有 100 μ g/ml ZeocinTM 的 MD 琼脂平板上导致平板带有巴斯德毕赤酵母细胞的斑点并且无明显的单菌落。似乎选择压力不足以阻抑野生型菌株的生长。尽管当葡

葡萄糖存在时在巴斯德毕赤酵母 GFP-Zeo Mut^SA9 菌株中不能检测到荧光,少数 GFP-Zeo 蛋白可能在赋予 ZeocinTM 抗性的细胞中表达。为了测试用于 GFP-Zeo Mut^SA9 菌株生长抑制的较高的 ZeocinTM 浓度,进行如前所述的液滴测试。

[0236] 能够在图 6 中清楚地看出,增加至 200 μ g/ml 不减少巴斯德毕赤酵母菌株的细胞生存力(与 100 μ g/ml 相比),但增加至 500 μ g/ml 则减少,所述巴斯德毕赤酵母菌株携带在 AOX1 启动子控制下的 GFP-Zeo 基因。期望的是启动子诱变应该仅导致轻微增加的表达水平,因此 500 μ g/ml ZeocinTM 的选择压力似乎太高。最终选择 350 μ g/ml 用于诱变启动子变体的所有进一步筛选。

[0237] 由于涉及许多启动子区的非常复杂的转录调节,使用高诱变率的随机诱变方法是有利的。

[0238] 实施例 2:启动子缺失的产生

[0239] 基于实施例 1 的结果,产生第二代缺失变体。相对于第一系列,在这些新的缺失构建体中,仅将推定的转录因子结合位点的小并且特异性的序列延伸(5-15bp)缺失(表 15)。

[0240]

[0241]

表 15: 特异性转录因子结合位点的缺失在去阻抑(葡萄糖饥饿)和甲醇诱导时对于表达水平的影响。同样提供突变 $\Delta 1$ - $\Delta 9$ 以及单突变的组合。所有数字是在相同条件下与野生型启动子活性比较的相对启动子活性。

缺失	Bereich von Seq		区域	缺失(粗体、下划线标记)和相邻的 5 个核苷酸(5'和 3')	正效应
	ID	No. 1			
Δ Hap1	-900 至 -896	Inan A	<u>GCCATCCGACATCCA</u>	诱导条件下增加的表达	
Δ Hsf_1	-812 至 -805	Inan A	<u>GGACCTCCACTCCTCTTC</u>	诱导条件下产生多拷贝菌株、增加的表达	
Δ 1	-784 至 -719	Inan B	CCACTTTTGCCATCGAAAAACCAGCCCACTTATGGGC <u>TTGATTGGAGCTCGCTCATTCCAATTCCTTCTATTAGG</u>	诱导条件下增加的表达	
Δ Hap2345_1	-758 至 -754	Inan B, Δ 1	<u>CAGTTATTGGGCTTG</u>	产生多拷贝菌株	
Δ Hap2345_2	-748 至 -744	Inan B, Δ 1	<u>GCTTGATTGGAGCTC</u>	产生多拷贝菌株	
Δ abaA	-735 至 -730	Inan B, Δ 1	<u>TCGCTCATTCCAATTC</u>	产生多拷贝菌株	
Δ Stre	-673 至 -669	Inan B	<u>TGGCCCCCTGGCGA</u>	诱导条件下增加的表达	
Δ 2	-650 至 -604	Inan C	TTTGTTTATTTCGGAATGCAACAAGCTCCGCATTACAC <u>CCGAACATCACTCCAGATG</u>	去阻抑条件下较高的表达	
Δ Rap1	-619 至 -615	Inan C, Δ 2	<u>ATTACACCCGAACAT</u>	产生多拷贝菌株	
Δ 3	-590 至 -561	Inan C	<u>GCCTTCTGAGTGTGGGGTCAAAATAGTTTCATGTTCCCC</u> AA		
Δ Adr1	-583 至 -577	Inan C, Δ 3	<u>GAGTGTGGGGTCAAAATA</u>	产生多拷贝菌株	
Δ 4	-445 至 -403	Inan D	<u>AGTTGACAAGACAAACGGTATGCCGACTTTTGGAAGTT</u> <u>TCTTTTIGACTTGGT</u>		
Δ Hsf_2	-437 至 -430	Inan D, Δ 4	<u>AAAAAGAAACTTCCAAAA</u>		

[0242]

Δ5	-329 至 -271	Inan E	<u>GAACCCCGGTGCACCTGTGCCGAAACGCAAAATGGGGA</u> <u>AACACCCGCCTTTTGGGATGATTATGCATTGTC</u>	产生多拷贝菌株
ΔHap2345_3	-286 至 -282	Inan E, Δ5	<u>CGCTTTTGGGATGAT</u>	
ΔMat1MC	-271 至 -267	Inan E, Δ5°	<u>TATGCATTGCTCTCCA</u>	
Δ6	-260 至 -231	Inan E & F	<u>TCCACATTGTATGCTTCCAAGAATCTGGTGGGAATACT</u> GC	
Δ6*	-260 至 -231	Inan E & F	<u>TCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACT</u> GC	去阻抑条件下较高的表 达，去阻抑条件下产生具 有高表达的超克隆
ΔGcr1	-252 至 -248	Inan E, Δ6	<u>TAGCCTAACGTT</u>	产生多拷贝菌株
Δ7	-225 至 -191	Inan F	<u>GTATGCTTCCAAGAT</u> <u>ACTGCTGATAGCCTAACGTTTCATGATCAAAATTTAACT</u> <u>GTCTAA</u>	
ΔQA-1F	-207 至 -193	Inan F, Δ7	<u>TTCATGATCAAAATTAACTGTCT</u>	去阻抑条件下产生多拷贝 菌株、增加的活性
ΔQA-1Fzus	-218 至 -213 -207 至 -193	Inan F, Δ7	<u>ATAGCCTAACGTTTCATGATCAAAATTTAACTGTCT</u>	去阻抑条件下 增加的活性
ΔHsf_2_dHap2345_1	-758 至 -754 -437 至 -430	Inan B, Δ1 Inan D, Δ4	<u>CAGTTATTGGGCTTG</u> <u>AAAAAGAAACTTCCAAAA</u>	产生多拷贝菌株
ΔHsf_2_dHap2345_1	-758 至 -754 -747 至 -745	Inan B, Δ1 Inan B, Δ1	<u>CAGTTATTGGGCTTGATTGGAGCT</u> <u>AAAAAGAAACTTCCAAAA</u>	
zus	-437 至 -430	Inan D, Δ4	<u>AAAAAGAAACTTCCAAAA</u>	
ΔHsf_2_Mat1MC	-437 至 -430 -271 至 -267	Inan D, Δ4 Inan E, Δ5°	<u>AAAAAGAAACTTCCAAAA</u> <u>TATGCATTGCTCTCCA</u>	产生多拷贝菌株
Δ8	-170 至 -154	Inan F	<u>ACAGCAATATATAAACAGAAAGGAGCT</u>	
Δ9	-131 至 -93	Inan F	<u>ACCTTTTTTTATCATCATATTAGCTTACTTTCATAA</u>	

[0243] 材料和方法：
[0244] a) 诱变

产生多拷贝菌株

Δ2Δ6	-650 至 -604	Inan C	<u>TTGCGACTGG</u>
	-260 至 -231	Inan E & F	<u>TTTGTATTCCGAATGCAACAAGCTCCGCATTACAC</u> <u>CCGAACATCACTCCAGATG</u> <u>TCCACATTGTATGCTTCCCAAGATTCTGGTGGGAATACT</u>
Δ736-41 Δ737-38 ΔInD-d4m	-218 至 -213	Inan F, Δ7	GC <u>ATAGCCTAAACGTTTCAT</u>
	-217 至 -216	Inan F, Δ7	<u>TAGCCTAACGTT</u>
	-402 至 -394	Inan D	<u>CTTGTTTGGTATTGATTGA</u>
	-520 至 -446	Inan D	<u>CTTGGAACCTAATAATGACAAAAGCGTGATCTCATCCAA</u> <u>GATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCC</u> <u>AGTTGCTTGG</u>
Δ1-1 Δ1-2 Δ1-3 Δ1-SacI	-784 至 -763	Inan B, Δ1	<u>CCACITTTTGCCCATCGAAAAACCAGCCCAGTTA</u>
	-762 至 -741	Inan B, Δ1	<u>AGCCCCAGTTATTGGGCTTGATTTGGAGCTCGCT</u>
	-740 至 -719	Inan B, Δ1	<u>GGAGCTCGCTCATTTCCAAATTCCTTCTATTAGG</u>
	-762 至 -744	Inan B, Δ1	<u>CCACITTTTGCCCATCGAAAAACCAGCCCAGTTATTGGGC</u> <u>TTGATTGGAGCTC</u>
Ohi, et al, AOX2 UAS	-228 至 -199	Inan F, Δ7	<u>AATACTGCTGATAGCCTAACGTTTCATGATCAAAATAAT</u> AC
	-369 至 -337	Inan E	<u>TAATCTCATTAATGCTTAGCGCAGTCTCTCTATCGCTT</u> TAATC
Ohi, et al, AOX2 UAS	-333 至 -294	Inan E	<u>TTCTGAACCCCGGTGCACCTGTGCCGAAACGCAAAATG</u> <u>GGGAAACACCCCGC</u>

[0245] 使用根据 Wang 等 [34] 的两阶段定位诱变规程导入所有缺失。在第一步中, 评估两种单独的反应 (一种用于正向引物而一种用于反向引物) (100ngpA0X 模板, 15pmol 引物, dATP、dTTP、dCTP 和 dGTP 各 200 μ M, 2.5UPfuUltra™ 聚合酶, 总体积 50 μ l, 在合适的缓冲条件中)。将这两个 PCR 反应的 25 μ l 组合并进行第二 PCR 反应步骤。

[0246] 将 1 μ l DpnI 限制酶 (10u/ μ l) 添加至 30 μ l 第二 PCR 反应步骤中并且在 37°C 温育 1 小时。将 1-5 μ l DpnI 消化的 PCR 反应转化到电感受态大肠杆菌细胞 [16] 中并且在 SOC 培养基中的 1 小时再生时间之后涂布于 LB-Amp 平板上。

[0247] 表 16 : 用于转录因子结合位点缺失的定位诱变的引物

[0248]

缺失	名称	序列(5'-->3')	SEQ ID No.
Hap1	Hap1fw	GAATGAAACCTTTTTGCCATAT CCACAGGTCCATTCTCAC	53
	Hap1rv	GAATGGACCTGTGGATATGGC AAAAAGGTTTCATTCAACC	54
Hsf_1	Hsf_1fw	CCGTTGCAAACGCAGGACCTC TTCTCCTCAACACCCAC	55
	Hsf_1rv	GTGTTGAGGAGAAGAGGTCCT GCGTTTGCAACGGTCTG	56
Hap2345_1	Hap2345_1fw	CGAAAAACCAGCCCAGTTGCT TGATTGGAGCTCGCTCATTCC	57
	Hap2345_1rv	GAGCGAGCTCCAATCAAGCAA CTGGGCTGGTTTTTCGATG	58
Hap2345_2	Hap2345_2fw	CAGCCCAGTTATTGGGCTTGA GCTCGCTCATTCCAATTCC	59
	Hap2345_2rv	GGAATTGGAATGAGCGAGCTC AAGCCCAATAACTGGGCTG	60
ABAA	ABAAfw	GGCTTGATTGGAGCTCGCTAAT TCCTTCTATTAGGCTAC	61
	ABAArv	GTAGCCTAATAGAAGGAATTAG CGAGCTCCAATCAAGCC	62
Stre_1	Stre_1fw	GCCTGTCTATCCTGGCCGGCGA GGTTCATGTTTGTTCATTTC	63
	Stre_1rv	CAAACATGAACCTCGCCGGCC AGGATAGACAGGCTAATAAAG	64

[0249]

缺失	名称	序列(5'-->3')	SEQ ID No.
Rap1	Rap1fw	GCAACAAGCTCCGCATTACAA CATCACTCCAGATGAGG	65
	Rap1rv	CCTCATCTGGAGTGATGTTGTA ATGCGGAGCTTGTTGC	66
Adr1	Adr1fw	CCAGATGAGGGCTTTCTGAGT GAAATAGTTTCATGTTCCC	67
	Adr1rv	GGGAACATGAAACTATTTCACT CAGAAAGCCCTCATCTGG	68
Hsf_2	Hsf_2fw	GCCAGTTGGTCAAAAACAAAA GTCGGCATAACCGTTTGTC	69
	Hsf_2rv	CGGTATGCCGACTTTTGTTTTT GACCAACTGGCCGTTAGC	70
Hap2345_3	Hap2345_3fw	CAAATGGGGAAACACCCGCTT ATGATTATGCATTGTCTCCAC	71
	Hap2345_3rv	GAGACAATGCATAATCATAAGC GGTGTTTCCCCATTTGCG	72
Mat1MC	Mat1MCfw	GCTTTTTTGGATGATTATGCCTC CACATTGTATGCTTCCAAG	73
	Mat1MCrv	CTTGGAAGCATAACAATGTGGA GGCATAATCATCCAAAAAGC	74
Gcr1	Gcr1fw	CATTGTCTCCACATTGTATGAA GATTCTGGTGGGAATACTGC	75
	Gcr1rv	GTATTCCCACCAGAATCTTCAT ACAATGTGGAGACAATGC	76
QA-1F	QA-1Ffw	GCTGATAGCCTAACGTTTCATGT TCTAACCCCTACTTGACAGC	77
	QA-1Frv	GTCAAGTAGGGGTTAGAACAT GAACGTTAGGCTATCAGCAG	78
736-741	d736-41fw	GGAATACTGCTGATAGCTTCAT GATCAAAATTAACTGTTC	79
	d736-41rv	GTAAATTTTGATCATGAAGCT ATCAGCAGTATTCCCACC	80
737-738	d737-38fw	GGAATACTGCTGATAGCCACGT TCATGATCAAAATTAACTG	81
	d737-38rv	GTAAATTTTGATCATGAACGT GGCTATCAGCAGTATTCC	82

[0250] b) 巴斯德毕赤酵母转化和克隆的表征：

[0251] 制备如上所述构建的质粒并如实施例 1 所述转化到巴斯德毕赤酵母中。

[0252] 结果和讨论：

[0253] 观察到实施例 2 的短突变对表达水平的强烈影响，如已经描述的实施例 1 的较大

缺失一样,其中所有突变对于启动子活性具有显著的正或负效应。特异性转录因子结合位点的短缺失对于启动子活性具有强烈的影响并且给出关于独立调节位点(例如转录因子结合位点)调节性质的更加精确的信息。特别感兴趣的是 Gcr1,因为其结合位点包括在 $\Delta 6$ 缺失中。pAOX $\Delta 6$ 缺失突变体启动子区的测序和巴斯德毕赤酵母克隆基因组 DNA 的菌落 PCR 产物揭示了启动子区中额外的缺失(缺失 SEQ ID No. 1 的核苷酸 737 至 738(-217 至 -216))。由于这种启动子变体导致增强的启动子活性,因而在去阻抑条件下产生较高的表达率,所以能够将额外的突变导入根据本发明的启动子以增加在这些条件下的蛋白质表达。

[0254] 具有额外缺失(缺失 SEQ ID No. 1 的核苷酸 736 至 741(-218 至 -213))的 QA-1F 克隆的启动子活性与不具有这种额外缺失的 Δ QA-1F 启动子相比显著不同:活性从 $\sim 30\%$ (去阻抑)和 $\sim 100\%$ (诱导)的野生型活性(AOX1 Δ QA-1F,参见表 15)分别改变至 $\sim 140\%$ 和 $\sim 70\%$ (AOX1 Δ QA-1F_{zus},参见表 15)。这 6 个核苷酸的额外缺失看起来对启动子活性具有剧烈的影响。因此通过如上所述的定位诱变规程导入携带这种突变($\Delta 736-741$)的新启动子变体。在这个区域两种突变偶然(accidentally)并且独立地出现两次,这导致去阻抑条件下启动子活性的增加。值得注意的是虽然在两种构建体中存在第二种并且最可能是负影响的突变,但是存在启动子活性的增加。

[0255] 与单缺失相似通过重叠延伸 PCR 产生 $\Delta 2$ 和 $\Delta 6$ 的组合($\Delta 2 \Delta 6$)。表 17 清楚地说明缺失两种片段导致在去阻抑以及诱导条件下启动子活性的非常大的减少。由于与前述的 $\Delta 6^*$ 构建体相比,在这种构建体中不存在额外的 TA 缺失,这个结果支持以下推测:偶然发生的额外突变($\Delta 737-38$)是碳饥饿时启动子活性增加的原因。

[0256] 几种缺失导致启动子活性显著的减少(例如 Hsf,还有 Hap1 和 Hap2345_1)。这些推定的结合位点是序列重复的出众的靶物,所述序列重复应导致启动子活性的增加。

[0257] 有趣的是,在通过定位诱变产生的 $\Delta 736-741$ 变体的 4 种克隆中的 2 种中,在位置 552 至 560(-402 至 -394)存在 9 核苷酸(TTGGTATTG)的新缺失。缺失存在于不同区域的效果也存在于 Δ Hsf_2 构建体中。由于局部序列同源性这种效果是期望的。因此这种额外的缺失区域($\Delta 552-560$ 、 $\Delta 737-38$ 和 $\Delta 736-41$)和紧密邻近的序列(上游和下游 5bp)也是推定的转录因子结合位点,并且因此是用于缺失和重复的高度感兴趣的靶物。缺失变体 $\Delta 736-41$ 导致在甲醇诱导条件下表达水平更大的减少。

[0258] 多拷贝菌株:

[0259] 在多数情况下,产生多拷贝菌株导致 GFP-Zeo 超表达菌株。在许多情况下,这些菌株主要在甲醇诱导条件下具有高于 d6^{*}F10 菌株的表达水平。多拷贝菌株的产生可以使用几种构建体完成,特别使用 $\Delta 6^*$ 构建体,双缺失 d2d6,包括 $\Delta 1$ 、 $\Delta 2$ 和 $\Delta 6^*$ 缺失的构建体,以及例如 Gcr1、Rap1、abaA、Hap2345_1,还有例如 QA-1F、Adr1、Hsf_2_Mat1MC 和 Hsf_2_Hap2345_1(参见图 7)。在这些菌株中,存在在诱导条件下与 d6^{*}F10 菌株相比较高的表达率。相反 d6^{*}F10 菌株能够在去阻抑条件下比已知产生的任何其它菌株产生更多的 GFP-Zeo。将 $\Delta 6^*$ 构建体重复地转化至巴斯德毕赤酵母中导致高数量的多拷贝菌株,所述多拷贝菌株具有与 d6^{*}F10 菌株相当的活性,特别在去阻抑条件下(图 8)。

[0260] 使用野生型启动子构建体,观察到与使用启动子变体相比低得多的多拷贝菌株(例如 E2 菌株)的频率。尽管分析了多 2-4 倍的转化体,最佳转化体 E2 的表达水平仅是单

拷贝转化体的两倍。总而言之转化启动子变体导致较高的多拷贝菌株频率,并且这些菌株是比多拷贝野生型启动子菌株更多产的。

[0261] 实施例 3 :可供选择的报告蛋白

[0262] 为了测试对于其它基本上良好表达的和工业相关的蛋白质(例如酶)的所有 GFP-Zeo 结果的实用性,将一些启动子变体克隆到这些报告酶(例如 PaHNL5 α 和 HRP)之前。

[0263] 克隆:

[0264] 将启动子变体克隆到载体 pPICZ α B-HRP-WT[35] 和 pGAPZ A-PaHNL5 α 中。为了 pPICZ α B-HRP-WT 中的启动子交换,将 NdeI 限制位点通过定位诱变(100ng 载体作为模板,引物 NdeI PICZfor 和 NdeI PICZrev—参见表 18)插入到启动子的 5' 端。将所得载体称为 pPICZ α B-NdeI-HRP-WT。

[0265] 表 18 :用于在 pPICZ α B-HRP-WT 中引入 NdeI 限制位点和用于启动子交换的定位诱变的引物

[0266]

名称	序列 (5' → 3')	Seq ID No.
NdeI PICZfor	GAGATCAGATCTAACATATGCCAAAGAC	83
	GAAAGGTTG	
NdeI PICZrev	CAACCTTTCGTCTTTGGCATATGTTAGAT	84
	CTGATCTC	
AOX1NDE1	AAACATATGAGATCTAACATCCAAAGAC	85
	GAAAGG	
AOX1rev	TGGTTGAATTCTTTCAATAATTAGTTG	86

[0267] 对于 pGAPZ A-PaHNL5 α 表达克隆,首先将 PaHNL5 α 基因从 pHIL-D2 载体 (Glieder, A., et al. Angew. Chemie Int. Ed. 2003) 克隆至 pGAPZ A 载体中,产生质粒 pGAPZA-PaHNL5 α 。可以直接在 EcoRI/BglIII 消化 pGAPZA-PaHNL5 α 和 pAOX Δ 质粒之后将启动子变体克隆到 pGAPZ A-PaHNL5 α 中。为了 pPICZ α B-NdeI-HRP-WT 中的交换,将启动子变体通过 PCR 扩增,所述 PCR 使用引物 AOX1NDE1 和 AOX1rev(参见表 18, 10ng pAOX Δ , 10pmol 引物 AOX1NDE1 和 AOX1rev, 200 μ M 的每种 dNTP, 0.6u Phusion™ 聚合酶在合适的缓冲条件中,并且总体积为 50 μ l)。将 PCR 产物和 pPICZ α B-NdeI-HRP-WT 质粒使用 NdeI/HindIII 限制位点克隆。

[0268] 转化、生长和酶试验 (enzyme assay):

[0269] 对于巴斯德毕赤酵母转化,将所有 HRP 载体通过 NdeI 线性化,并且将所有 PaHNL5 α 质粒通过 BglIII 线性化。如实施例 1 中所述进行转化。同样如实施例 1 中所述进行巴斯德毕赤酵母菌株的生长,仅有少数不同。将初始 BMD(1%) 的量增加至 350 μ l,并且在 60 小时之后取 100 μ l 培养物用于离心 (4000rpm、4°C、10min)。完全如实施例 1 中所述进行甲醇诱导。

[0270] 取 50 μ l (去阻抑) 或 10 μ l (诱导) 来自离心的上清液进行 HNL 试验,并且两种情况下均取 15 μ l 用于 HRP 试验。

[0271] HRP 试验 (根据 [35]):

[0272] 将 15 μ l 上清液添加至 PS 微量滴定板内 50mM NaOAc 缓冲液 pH4.5 中的 150 μ l 1mM ABTS/2.9mM H_2O_2 中。在 Spectramax Plus384 平板读数器 (Molecular Devices, Sunnyvale, CA, USA) 中于 405nm 跟踪吸收 5 分钟。

[0273] HNL 试验 (根据 [36]):

[0274] 将 50 μ l 或 10 μ l 上清液添加至 UV-Star 微量滴定板中 100 或 140 μ l 0.05M 磷酸盐-柠檬酸盐缓冲液 pH5.0 中。通过添加 50 μ l 0.06M 杏仁腈 (mandelonitrile) 溶液 (溶于 0.003M 磷酸盐-柠檬酸盐缓冲液, pH3.5) 来起始反应, 并且在 Spectramax Plus384 平板读数器 (Molecular Devices, Sunnyvale, CA, USA) 中于 280nm 跟踪 5 分钟。

[0275] 结果和讨论:

[0276] 使用可供选择的报导蛋白 PaHNL5 α 和 HRP 的结果清楚地显示使用 GFP-Zeo 检测的启动子活性的可转移性 (表 17)。

[0277] 由于 HRP 试验的较低的敏感性, 在去阻抑条件的表达水平低于检测极限。因此在去阻抑条件下不能测定 HRP 表达。

[0278] 表 17: 具有可供选择的报告酶的几种 AOX1 启动子变体的启动子活性 (将在相同条件下 (分别是去阻抑和诱导) 与野生型启动子的活性相比的相对活性引用在括号内)

[0279]

启动子	GFP-Zeo		PaHNL5 α		HRP	
	去阻抑	甲醇	去阻抑 [mU/min]	甲醇 [mU/min]	去阻抑 [mU/min]	甲醇 [mU/min]
P(AOX1)	27.3 (100%)	987 (100%)	2.58 (100%)	69.5 (100%)	n.d.	20.3 (100%)
P(AOX1) Δ 1	29.5 (108%)	1188 (120%)	2.37 (92%)	100 (144%)	n.d.	26.9 (132%)
P(AOX1) Δ 2	43.0 (157%)	399 (40%)	n.d.	91.7 (132%)	n.d.	9.6 (47%)
P(AOX1) Δ 6*	89.9 (329%)	422 (42%)	8.65 (335%)	51.7 (74%)	n.d.	17.5 (86%)
P(AOX1) Δ 2 Δ 6	9.9 (36%)	336 (34%)	1.29 (50%)	37.5 (54%)	n.d.	9.9 (49%)

[0280] n. d. 无法检测

[0281] 为了将多拷贝选择转移至可供选择的报告体系, 将 AOX1 启动子变体克隆到合适的 HRP 和 PaHNL5 α 质粒中的 Zeocin 抗性基因之前, 由此取代 TEF1 启动子。

[0282] 实施例 4: 可供选择的报告蛋白 GFP

[0283] 为了测试具有 GFP 的启动子变体, 将实施例 1 和 2 中所述的启动子变体克隆到环 -3-GFP 基因之前。

[0284] 克隆:

[0285] 将载体 pAOX 中环 -3-GFP 中的内部 BamHI 和 XhoI 限制位点通过定位诱变缺失, 所述定位诱变使用引物 Bam-del-f 和 Xho-del-f (表 19) 和 100ng 载体作为模板。使用引物

GFP-Zeo forw(Seq. ID No. 4, 表 7, 10pmol) 和 wtGFP-XhoI-r(表 19, 10pmol) 和 Phusion™ 聚合酶, 在合适的条件下通过 PCR 从所得质粒 (10ng) 扩增 GFP 片段。可以使用 EcoRI/XhoI 的限制切割并且使用 T4DNA 连接酶连接将所得的 PCR 产物克隆到载体 pPICZ B 中。将所得质粒命名为 pPICZ-GFP。

[0286] 将所有启动子变体克隆到 pPICZ-GFP 中能够直接在 BglIII/EcoRI 消化 pPICZ-GFP 和 pAOX Δ 质粒后进行。

[0287] 表 19 : 用于载体 pAOX 中环 -3-GFP 定位诱变和其 GFP 片段扩增的引物

[0288]

名称	序列(5'→3')	Seq. ID No.
Bam-del-f	cgccacaacattgaagatggtccgttcaactagcagaccattatc	87
Xho-del-f	ggaaacattctcggacacaaaacttgagtacaactataactcacacaatg	88
wtGFP-XhoI-r	atctcgagttacttgtacaattcatccatgcatgtgtaatccc	89

[0289] 转化、生长和 GFP 检测 :

[0290] 为了巴斯德毕赤酵母的转化, 将所有质粒通过 BglIII 线性化。如实施例 1 中所述进行转化。在转化和 2 小时的再生阶段之后, 将细胞涂布在含有 100 μg/ml Zeocin 的 YPD-Zeo 琼脂平板上。

[0291] 巴斯德毕赤酵母的生长、甲醇诱导和 GFP 荧光的测量完全如实施例 1 中所述进行。

[0292] 结果和讨论 :

[0293] 再一次, 使用 GFP 作为报告系统的结果显示使用 GFP-Zeo 检测的启动子活性的可转移性 (表 20)。

[0294] 多拷贝菌株 :

[0295] 如实施例 1 和 2 中所述, 使用 Zeocin 作为选择标记的多拷贝菌株的发生是非常普遍的。能够通过将选择平板上 Zeocin 的浓度分别增加至 500 和 1000 μg/ml 来极大地增加多拷贝菌株的频率。

[0296] 表 20 : 具有 GFP 和 GFP-Zeo 作为报告基因的几种 AOX1 启动子变体与相同条件 (分别是去阻抑和诱导) 下野生型启动子相比的相对启动子活性

[0297]

启动子变体	GFP		GFP-Zeo	
	菌株 No.	甲醇 RFU	菌株 No.	甲醇 RFU
WT	E1	100 %	D2	100 %
Δ Hap1	C9	89 %	A2	84 %
Δ 1	4E6	79 %	A9	134 %
Δ 1-3	8-F12	75 %	D5	67 %
Δ 2	G12	37 %	F2	40%
Δ Rap1	D6	27 %	B9	34 %
Δ 3	H3	26 %	H2	70 %
Δ Adr1	A9	50 %	A2	56 %
Δ 4	C7	66 %	H9	71 %
Δ 5	38E6	28 %	D4	31 %
Δ Mat1MC	6C2	31 %	F6	32 %
Δ 6	37F5	79 %	H3	91 %
Δ 6*	E11	23 %	A5	40 %
Δ Gcr1	A9	60 %	A2	55 %
Δ 7	D12	38 %	A7	25 %
Δ QA-1F	7A3	61 %	E2	61 %
Δ QA-1Fzus	7A6	15 %	H7	25 %
Δ 8	E1	11 %	H1	17 %
Δ 9	3E5	23 %	A12	61 %
Δ 2 Δ 6	4B10	22 %	F3	21 %
Δ 736-41	5A7	8.8 %	C6	6 %
Δ 737-38	1G11	5.0 %	A3	8 %
多拷贝菌株				
Δ 1-3	8B10	400 %		
Δ 6	37A3	650 %		

[0298] 实施例 5 :使用 GFP 的充分系 (sufficiency series)

[0299] 为了在不含有几乎全部转录因子结合位点的体系中测试小部分的 AOX1 启动子, 将 AOX1 启动子在位置 -176 和 -194TATA 框之前切去少数碱基对, 其产生基础启动子元件 AOX176 和 AOX194(表 21)。为了允许后续将启动子元件克隆到基础启动子片段之前以及克隆基础启动子, 将 BspTI 和 EcoRI 限制位点分别插入 5' 和 3' 端。

[0300] 表 21 :将添加到基础启动子变体之前的基础 AOX1 启动子元件 AOX176 和 AOX194 和启动子片段 737 和 201-214 的序列。将限制位点 BspTI 和 EcoRI 加下划线。

[0301]

名称	序列(5'→3')	Seq. ID No.
AOX176	<u>CTTAAGGACAGCAATATATAAACAGAAGGAAG</u> CTGCCCTGTCTTAAACCTTTTTTTTATCATCAT TATTAGCTTACTTTTCATAATTGCGACTGGTTCC AATTGACAAGCTTTTGATTTTAACGACTTTTAA CGACAACCTTGAGAAGATCAAAAAACAACCTAAT TATTGAAAGAATTC	90
AOX194	<u>CTTAAGTGTCTTAACCCCTACTTGACAGCAATA</u> TATAAACAGAAGGAAGCTGCCCTGTCTTAAAC CTTTTTTTTTATCATCATTATTAGCTTACTTTTCAT AATTGCGACTGGTTCCAATTGACAAGCTTTTG ATTTTAACGACTTTTAACGACAACCTTGAGAAG ATCAAAAAACAACCTAATTATTGAAAGAATTC	91
737	TAGCCTAACGTT	92
201-214	CATGATCAAAATTT	93

[0302] 克隆：

[0303] 分别使用引物 AOX1basalrv (表 21, 10pmol) 和 AOXbasalfwn (10pmol, AOX194) 和 AOX176fw (10pmol, AOX176) 从 pAOX (10ng) 扩增基础 AOX1 元件。使用 Phusion™ 聚合酶 (0.6u) 在合适条件下在总体积 50 μl 中进行 PCR。

[0304] 如上所述使用引物 AOX1basalrv 和 737-38AOX176 通过 PCR 扩增启动子变体 AOX176-737。如上所述使用引物 AOX1basalrv 和 201-214AOX176 通过 PCR 扩增启动子变体 AOX176-201-214。

[0305] 可使用 BglIII/EcoRI 的限制切割并且使用 T4DNA 连接酶的连接将所得 PCR 产物克隆到载体 pPICZ-GFP 中, 由此取代野生型 AOX1 启动子。

[0306] BglIII BspTI

EcoRI

[0307] ~~~~~ ~~~~~

~~~~~

[0308] AGATCTCGAC TTAAGCAATC GTCTTACTTT CTAACCTTTTC TTACCTTTTA CATTTCAGCA  
ATATATATAT ATATTTCAAG GATATACCGA ATTCTCTAGAGCTGAATTCGTTAG CAGAATGAAA  
GATTGAAAAG AATGGAAAAT GTAAAGTCGT TATATATATA TATAAAGTTC CTATATGGCT TAAG

[0309] 将 4 种寡核苷酸 Leu2basal1f、Leu2basal2f、Leu2basal1r 和 Leu2basal2r (各 25pmol) 混入 20 μl 的总体积, 加热至 95℃ 持续 2 分钟, 并且缓慢冷却至室温。将 3 μl 的混合物与 159ng 的 pPICZ-GFPBglIII/EcoRI 片段于 16℃ 连接 6 小时。在转化入大肠杆菌之后, 将所得载体称为 pLeu2basal-GFP。

[0310] 如上所述使用引物 LEU2basalrv 和 737-38Leu2, 和 pLeu2basal-GFP 作为模板通过 PCR 扩增启动子变体 Leu2-737。可使用 BglIII/EcoRI 的限制切割并且使用 T4DNA 连接酶的连接将所得 PCR 产物克隆入载体 pPICZ-GFP, 由此取代野生型 AOX1 启动子。将所得质粒称为 pLeu2-GFP-737。

[0311] 表 22 : 用于产生基础启动子元件和充分构建体 (sufficiency construct) 的引物

[0312]

| 名称            | 序列(5'→3')                                                     | Seq ID<br>No. |
|---------------|---------------------------------------------------------------|---------------|
| AOX1basalrv   | TTTGAATTCTTTCAATAATTAGTTGTTTTTTG                              | 94            |
| AOX176fw      | TTAGATCTCGACTTAAGGACAGCAATATATAA<br>ACAGAAGGAAG               | 95            |
| AOX1basalfwn  | TTAGATCTCGACTTAAGTGTTCTAACCCCTAC<br>TTGACAG                   | 96            |
| 737-38AOX176  | AAAGATCTTAGCCTAACGTTCTTAAGGACAGC<br>AATATATAAACAGAAGGAAG      | 97            |
| 201-214AOX176 | AAAGATCTCATGATCAAAATTTCTTAAGGACA<br>GCAATATATAAACAGAAGGAAG    | 98            |
| LEU2basal1f   | GATCTCGACTTAAGCAATCGTCTTACTTTCTA<br>ACTTTTCTTACCTTTTACATTTCAG | 99            |
| LEU2basal2f   | CAATATATATATATATTTCAAGGATATACCG                               | 100           |
| LEU2basal1r   | AATTCGGTATATCCTTGAAATATATATATATT<br>GCTGAAATGTAAAAG           | 101           |
| LEU2basal2r   | GTAAGAAAAGTTAGAAAGTAAGACGATTGCT<br>TAAGTCGA                   | 102           |
| LEU2basalrv   | GGTTGAATTCGGTATATCCTTG                                        | 103           |
| 737-38Leu2    | AAAGATCTTAGCCTAACGTTCTTAAGCAATCG<br>TCTTACTTTCTAAC            | 104           |

[0313] 转化、生长和 GFP 检测：

[0314] 对于巴斯德毕赤酵母转化，将所有质粒通过 BamHI 线性化。如实施例 1 中所述进行转化。在转化和 2 小时再生阶段之后，将细胞涂布在含有 100 μg/ml Zeocin 的 YPD-Zeo 琼脂平板上。

[0315] 巴斯德毕赤酵母的生长、甲醇诱导和 GFP 荧光的测定完全如实施例 1 中所述的进行。

[0316] 结果和讨论：

[0317] 这个实验显示实施例 1 和 2 中鉴定的小元件的添加能够用于增加基础启动子元件的启动子强度，所述基础启动子元件源自 AOX1 启动子或源自酿酒酵母 LEU2 启动子。

[0318] 多拷贝菌株：

[0319] 转化之后存在的多拷贝菌株的发生和频率与实施例 4 中所述完全相同。质粒中线性化的不同位点对多拷贝菌株的产生不具有任何影响。

[0320] 表 23 :不添加和添加假定作为调节物结合位点的小 AOX1 启动子片段之后基础启动子元件的启动子活性。将 GFP 用作报告蛋白。显示单拷贝以及多拷贝菌株。

[0321]

|                     | 单拷贝        |            | 多拷贝        |           |
|---------------------|------------|------------|------------|-----------|
|                     | 去阻抑<br>RFU | 甲醇<br>RFU  | 去阻抑<br>RFU | 甲醇<br>RFU |
| pAOX176-GFP         | n.d.       | 22.1 ± 0.2 | -          | -         |
| pAOX194-GFP         | n.d.       | 16.9 ± 2.9 | -          | -         |
| pAOX176-GFP-737     | n.d.       | 21.2 ± 1.1 | 59 ± 6     | 265 ± 38  |
| pAOX176-GFP-201-214 | 69.6 ± 6.0 | 44.9 ± 4.8 | -          | -         |
| pLeu2basal-GFP      | n.d.       | 11.3 ± 3.6 | -          | -         |
| pLeu2-GFP-737       | n.d.       | 19.2 ± 1.9 | 55 ± 5     | 138 ± 11  |

[0322] 参考：

[0323] Nakagawa, T., et al. (1999) *Yeast* 15(12), 1223-30.

[0324] Koutz, P., et al. (1989) *Yeast* 5(3), 167-77.

[0325] Ohi, H., et al. (1994) *Mol. Gen. Genet.* 243(5), 489-99.

[0326] Cregg, J. M., et al. (1989) *Mol. Cell. Biol.* 9(3), 1316-23.

[0327] Nakagawa, T., et al. (2002) *Yeast* 19(12), 1067-73.

[0328] Sakai, Y., et al. (1998) *J. Bacteriol.* 180(22), 5885-5890.

[0329] Genu, V., et al. (2003) *Eur. J. Biochem.* 270(11), 2467-2475.

[0330] Goedecke, S., et al. (1994) *Gene* 139(1), 35-42.

[0331] Tschopp, J. F., et al. (1987) *Nucleic Acids Res.* 15(9), 3859-76.

[0332] Parpinello, G., et al. (1998) *J. Bacteriol.* 180(11), 2958-67.

[0333] Alamae, T. et al. (1994) *Yeast* 10(11), 1459-66.

[0334] Inan, M. et al. (2001) *J. Biosci. Bioeng.* 92(6), 585-589.

[0335] Sreekrishna, K., et al. (1997) *Gene* 190(1), 55-62.

[0336] Romanos, M., et al. (1998) *Methods Mol. Biol.* 103(55-72).

[0337] Quandt, K., et al. (1995) *Nucleic Acids Res.* 23(23), 4878-84.

[0338] Ausubel, F. M., et al., *Current Protocols in Molecular Biology*, John Wiley and Sons, New York, 2003.

[0339] Corpet, F. (1988) *Nucleic Acids Res.* 16(22), 10881-90.

[0340] Chenna, R., et al. (2003) *Nucleic Acids Res.* 31(13), 3497-500.

[0341] Bennett, R. P., et al. (1998) *BioTechniques* 24(3), 478-82.

[0342] Farinas, E. T., et al. (2001) *Adv. Synth. Catal.* 343(6), 601-606.

[0343] Huie, M. A. et al. (1996) *Yeast* 12(4), 307-17.

[0344] Lopez, M. C., et al. (1998) *PNAS U. S. A.* 95(24), 14112-7.

[0345] Zeng, X., et al. (1997) *Genetics* 147(2), 493-505.

- [0346] Del Vescovo, V., et al. (2004) J. Mol. Biol. 338 (5), 877-93.
- [0347] Simon, M., et al. (1992) Yeast 8 (4), 303-9.
- [0348] Raschke, W. C., et al. (1996) Gene 177 (1-2), 163-7.
- [0349] Andrianopoulos, A. et al. (1994) Mol. Cell. Biol. 14 (4), 2503-15.
- [0350] Kwast, K. E., et al. (1998) J. Exp. Biol. 201 (Pt8) (1177-95.
- [0351] Bourgarel, D., et al. (1999) Mol. Microbiol. 31 (4), 1205-15.
- [0352] Dang, V. D., et al. (1996) J. Bacteriol. 178 (7), 1842-9.
- [0353] Hahn, J. S. et al. (2004) J. Biol. Chem. 279 (7), 5169-76.
- [0354] W002/081650
- [0355] US6, 699, 691
- [0356] Wang, W. et al. (1999) Biotechniques 26 (4), 680-2.
- [0357] Morawski, B., et al. (2000) Protein Eng 13 (5), 377-84.
- [0358] Weis, R., et al. (2004) FEMS Yeast Res. 5, 179-189.
- [0359] 序列表：
- [0360] Seq ID No. 1：巴斯德毕赤酵母启动子的 AOX1
- [0361]

**ggtagcagatcctaacaacgacgaaagggtgaatgaaaccttttgccatccgacatccacagggtccattct  
cacacataagtgcacaacgcaacaggaggggatacactagcagcagaccgttgcaaacgcaggacctccactcctc  
ttctcctcaaacaccacttttgccatcgaaaaaccagcccagttattgggcttgattggagctcgctcattccaattcctct  
attaggctactaacaccatgactttattagcctgtctatcctggccccctggcgagggtcatgtttgtttatttcgaatgc  
aacaagctccgcattacaccgaacatcactccagatgagggctttctgagtgtgggtcaaatagttcatgttccca  
aatggcccaaaactgacagtttaaacgctgtcttggaaacctaatgacaaaagcgtgatctcatccaagatgaactaa  
gttgggtcggtgaaatgctaacggccagttggcaaaaagaaacttccaaaagtcggcataccgtttgtcttgggtat  
tgattgacgaatgctcaaaaataatctcatlaaatgcttagcgcagttctctatcgcttctgaaccccgggtgcacctgtgc  
cgaaacgcaaatggggaacaccgcttttggatgattatgcattgtctccacattgtatgettccaagattctgggtggg  
aatactgctgalagcctaacggtcatgatcaaaatttaactgttctaaccctacttgacagcaatatataacagaagga  
agctgccctgtcttaaaccttttttcatcattattagcttacttcataattgcgactgggtccaattgacaagctttgat  
ttaaaccgactttaacgacaacttgagaagatcaaaaaaacctaatattgaagaattcaacc**

|        |                                                                    |     |
|--------|--------------------------------------------------------------------|-----|
| [0001] | 序列表                                                                |     |
| [0002] | <110> 格拉茨科技大学 (Technische Universität Graz)                        |     |
| [0003] | <120> 突变 AOX1 启动子                                                  |     |
| [0004] | <130>R 47071                                                       |     |
| [0005] | <140>PCT/AT2006/000079                                             |     |
| [0006] | <141>2006-02-23                                                    |     |
| [0007] | <150>AT A 304/2005                                                 |     |
| [0008] | <151>2005-02-23                                                    |     |
| [0009] | <160>137                                                           |     |
| [0010] | <170>PatentIn version 3.3                                          |     |
| [0011] | <210>1                                                             |     |
| [0012] | <211>953                                                           |     |
| [0013] | <212>DNA                                                           |     |
| [0014] | <213> 巴斯德毕赤酵母                                                      |     |
| [0015] | <220>                                                              |     |
| [0016] | <223>AOX1 启动子                                                      |     |
| [0017] | <400>1                                                             |     |
| [0018] | ggtaccagat ctaacatcca aagacgaaag gttgaatgaa acctttttgc catccgacat  | 60  |
| [0019] | ccacaggtcc attctcacac ataagtgcc aacgcaacag gaggggatac actagcagca   | 120 |
| [0020] | gaccgttgca aacgcaggac ctccactcct cttctcctca acaccactt ttgccatga    | 180 |
| [0021] | aaaaccagcc cagttattgg gcttgattgg agctcgetca ttccaattcc ttctattagg  | 240 |
| [0022] | ctactaacac catgacttta ttagcctgtc tatectggcc cccctggcga ggttcatgtt  | 300 |
| [0023] | tgttttatttc cgaatgcaac aagctccgca ttacaccga acatcactcc agatgagggc  | 360 |
| [0024] | tttctgagtg tggggtcaaa tagtttcatg tteccaaat ggcccaaac tgacagtta     | 420 |
| [0025] | aacgctgtct tggaacctaa tatgacaaaa gcgtgatctc atccaagatg aactaagttt  | 480 |
| [0026] | ggttcgttga aatgctaacg gccagttggt caaaaagaaa cttccaaaag tgggcatacc  | 540 |
| [0027] | gtttgtcttg ttggtattg attgacgaat gctcaaaaat aatctcatta atgcttagcg   | 600 |
| [0028] | cagtctctct atcgcttctg aaccccggtg cacctgtgcc gaaacgcaaa tggggaaaca  | 660 |
| [0029] | cccgtttttt ggatgattat gcattgtctc cacattgtat gcttccaaga ttctggtggg  | 720 |
| [0030] | aatactgctg atagcctaac gttcatgac aaaatttaac tgttctaacc cctacttgac   | 780 |
| [0031] | agcaatatat aaacagaagg aagctgccct gtcttaaacc ttttttttta tcattcattat | 840 |
| [0032] | tagcttactt tcataattgc gactgggtcc aattgacaag cttttgattt taacgacttt  | 900 |
| [0033] | taacgacaac ttgagaagat caaaaaacaa ctaattattg aaagaattca acc         | 953 |
| [0034] | <210>2                                                             |     |
| [0035] | <211>32                                                            |     |
| [0036] | <212>DNA                                                           |     |
| [0037] | <213> 人工的                                                          |     |
| [0038] | <220>                                                              |     |

|        |                                                |    |
|--------|------------------------------------------------|----|
| [0039] | <223> 引物                                       |    |
| [0040] | <400>2                                         |    |
| [0041] | aaggtaccag atctaacatc caaagacgaa ag            | 32 |
| [0042] | <210>3                                         |    |
| [0043] | <211>42                                        |    |
| [0044] | <212>DNA                                       |    |
| [0045] | <213> 人工的                                      |    |
| [0046] | <220>                                          |    |
| [0047] | <223> 引物                                       |    |
| [0048] | <400>3                                         |    |
| [0049] | ctagccatgg ttgaattctt tcgaataatt agttgttttt tg | 42 |
| [0050] | <210>4                                         |    |
| [0051] | <211>30                                        |    |
| [0052] | <212>DNA                                       |    |
| [0053] | <213> 人工的                                      |    |
| [0054] | <220>                                          |    |
| [0055] | <223> 引物                                       |    |
| [0056] | <400>4                                         |    |
| [0057] | gaaagaattc aaccatggct agcaaaggag               | 30 |
| [0058] | <210>5                                         |    |
| [0059] | <211>34                                        |    |
| [0060] | <212>DNA                                       |    |
| [0061] | <213> 人工的                                      |    |
| [0062] | <220>                                          |    |
| [0063] | <223> 引物                                       |    |
| [0064] | <400>5                                         |    |
| [0065] | gatgatggtc tagaacgtgt cagtcctgct cctc          | 34 |
| [0066] | <210>6                                         |    |
| [0067] | <211>34                                        |    |
| [0068] | <212>DNA                                       |    |
| [0069] | <213> 人工的                                      |    |
| [0070] | <220>                                          |    |
| [0071] | <223> 引物                                       |    |
| [0072] | <400>6                                         |    |
| [0073] | gacacgttct agaccatcat catcatcatc attg          | 34 |
| [0074] | <210>7                                         |    |
| [0075] | <211>26                                        |    |
| [0076] | <212>DNA                                       |    |
| [0077] | <213> 人工的                                      |    |

|        |                                             |    |
|--------|---------------------------------------------|----|
| [0078] | <220>                                       |    |
| [0079] | <223> 引物                                    |    |
| [0080] | <400>7                                      |    |
| [0081] | atagcggccg cacaaacgaa ggtctc                | 26 |
| [0082] | <210>8                                      |    |
| [0083] | <211>40                                     |    |
| [0084] | <212>DNA                                    |    |
| [0085] | <213> 人工的                                   |    |
| [0086] | <220>                                       |    |
| [0087] | <223> 引物                                    |    |
| [0088] | <400>8                                      |    |
| [0089] | caacacccac tttaggctac taacaccatg actttattag | 40 |
| [0090] | <210>9                                      |    |
| [0091] | <211>34                                     |    |
| [0092] | <212>DNA                                    |    |
| [0093] | <213> 人工的                                   |    |
| [0094] | <220>                                       |    |
| [0095] | <223> 引物                                    |    |
| [0096] | <400>9                                      |    |
| [0097] | gttagtagcc taaagtgggt gttgaggaga agag       | 34 |
| [0098] | <210>10                                     |    |
| [0099] | <211>32                                     |    |
| [0100] | <212>DNA                                    |    |
| [0101] | <213> 人工的                                   |    |
| [0102] | <220>                                       |    |
| [0103] | <223> 引物                                    |    |
| [0104] | <400>10                                     |    |
| [0105] | gttcatgttt gtagatgagg gctttctgag tg         | 32 |
| [0106] | <210>11                                     |    |
| [0107] | <211>30                                     |    |
| [0108] | <212>DNA                                    |    |
| [0109] | <213> 人工的                                   |    |
| [0110] | <220>                                       |    |
| [0111] | <223> 引物                                    |    |
| [0112] | <400>11                                     |    |
| [0113] | gccctcatct acaaacatga acctcgccag            | 30 |
| [0114] | <210>12                                     |    |
| [0115] | <211>28                                     |    |
| [0116] | <212>DNA                                    |    |

|        |                                      |    |
|--------|--------------------------------------|----|
| [0117] | <213> 人工的                            |    |
| [0118] | <220>                                |    |
| [0119] | <223> 引物                             |    |
| [0120] | <400>12                              |    |
| [0121] | gagggctttc ccaaattggcc caaaactg      | 28 |
| [0122] | <210>13                              |    |
| [0123] | <211>28                              |    |
| [0124] | <212>DNA                             |    |
| [0125] | <213> 人工的                            |    |
| [0126] | <220>                                |    |
| [0127] | <223> 引物                             |    |
| [0128] | <400>13                              |    |
| [0129] | ccatttggga aagccctcat ctggagtg       | 28 |
| [0130] | <210>14                              |    |
| [0131] | <211>31                              |    |
| [0132] | <212>DNA                             |    |
| [0133] | <213> 人工的                            |    |
| [0134] | <220>                                |    |
| [0135] | <223> 引物                             |    |
| [0136] | <400>14                              |    |
| [0137] | cggccagttg ttggtattga ttgacgaatg c   | 31 |
| [0138] | <210>15                              |    |
| [0139] | <211>29                              |    |
| [0140] | <212>DNA                             |    |
| [0141] | <213> 人工的                            |    |
| [0142] | <220>                                |    |
| [0143] | <223> 引物                             |    |
| [0144] | <400>15                              |    |
| [0145] | caataccaac aactggccgt tagcatttc      | 29 |
| [0146] | <210>16                              |    |
| [0147] | <211>33                              |    |
| [0148] | <212>DNA                             |    |
| [0149] | <213> 人工的                            |    |
| [0150] | <220>                                |    |
| [0151] | <223> 引物                             |    |
| [0152] | <400>16                              |    |
| [0153] | gcttctgaac cttgtctcca cattgtatgc ttc | 33 |
| [0154] | <210>17                              |    |
| [0155] | <211>32                              |    |

|        |                                       |    |
|--------|---------------------------------------|----|
| [0156] | <212>DNA                              |    |
| [0157] | <213> 人工的                             |    |
| [0158] | <220>                                 |    |
| [0159] | <223> 引物                              |    |
| [0160] | <400>17                               |    |
| [0161] | gtggagacaa ggttcagaag cgatagagag ac   | 32 |
| [0162] | <210>18                               |    |
| [0163] | <211>29                               |    |
| [0164] | <212>DNA                              |    |
| [0165] | <213> 人工的                             |    |
| [0166] | <220>                                 |    |
| [0167] | <223> 引物                              |    |
| [0168] | <400>18                               |    |
| [0169] | gtctccacac tgctgatage ctaacgttc       | 29 |
| [0170] | <210>19                               |    |
| [0171] | <211>35                               |    |
| [0172] | <212>DNA                              |    |
| [0173] | <213> 人工的                             |    |
| [0174] | <220>                                 |    |
| [0175] | <223> 引物                              |    |
| [0176] | <400>19                               |    |
| [0177] | ggctatcage agtgtggaga caatgcataa tcac | 35 |
| [0178] | <210>20                               |    |
| [0179] | <211>30                               |    |
| [0180] | <212>DNA                              |    |
| [0181] | <213> 人工的                             |    |
| [0182] | <220>                                 |    |
| [0183] | <223> 引物                              |    |
| [0184] | <400>20                               |    |
| [0185] | ggaatactgc tctaaccctt acttgacagc      | 30 |
| [0186] | <210>21                               |    |
| [0187] | <211>32                               |    |
| [0188] | <212>DNA                              |    |
| [0189] | <213> 人工的                             |    |
| [0190] | <220>                                 |    |
| [0191] | <223> 引物                              |    |
| [0192] | <400>21                               |    |
| [0193] | gtaggggtta gagcagtatt cccaccagaa tc   | 32 |
| [0194] | <210>22                               |    |

|        |                                                  |    |
|--------|--------------------------------------------------|----|
| [0195] | <211>29                                          |    |
| [0196] | <212>DNA                                         |    |
| [0197] | <213> 人工的                                        |    |
| [0198] | <220>                                            |    |
| [0199] | <223> 引物                                         |    |
| [0200] | <400>22                                          |    |
| [0201] | cttgacagca agctgccctg tcttaaacc                  | 29 |
| [0202] | <210>23                                          |    |
| [0203] | <211>28                                          |    |
| [0204] | <212>DNA                                         |    |
| [0205] | <213> 人工的                                        |    |
| [0206] | <220>                                            |    |
| [0207] | <223> 引物                                         |    |
| [0208] | <400>23                                          |    |
| [0209] | gggcagcttg ctgtcaagta ggggttag                   | 28 |
| [0210] | <210>24                                          |    |
| [0211] | <211>34                                          |    |
| [0212] | <212>DNA                                         |    |
| [0213] | <213> 人工的                                        |    |
| [0214] | <220>                                            |    |
| [0215] | <223> 引物                                         |    |
| [0216] | <400>24                                          |    |
| [0217] | ctgtcttaaa ccttactggt tccaattgac aagc            | 34 |
| [0218] | <210>25                                          |    |
| [0219] | <211>29                                          |    |
| [0220] | <212>DNA                                         |    |
| [0221] | <213> 人工的                                        |    |
| [0222] | <220>                                            |    |
| [0223] | <223> 引物                                         |    |
| [0224] | <400>25                                          |    |
| [0225] | ggaaccagta aggtttaaga cagggcagc                  | 29 |
| [0226] | <210>26                                          |    |
| [0227] | <211>44                                          |    |
| [0228] | <212>DNA                                         |    |
| [0229] | <213> 人工的                                        |    |
| [0230] | <220>                                            |    |
| [0231] | <223> 引物                                         |    |
| [0232] | <400>26                                          |    |
| [0233] | gatacactag cagcagaccg ttgcaaacgc aggacctcca ctcc | 44 |

|        |                                                     |    |
|--------|-----------------------------------------------------|----|
| [0234] | <210>27                                             |    |
| [0235] | <211>47                                             |    |
| [0236] | <212>DNA                                            |    |
| [0237] | <213> 人工的                                           |    |
| [0238] | <220>                                               |    |
| [0239] | <223> 引物                                            |    |
| [0240] | <400>27                                             |    |
| [0241] | gtgaaggtga tgctacatac ggaaagctta cccttaaatt tatttgc | 47 |
| [0242] | <210>28                                             |    |
| [0243] | <211>41                                             |    |
| [0244] | <212>DNA                                            |    |
| [0245] | <213> 人工的                                           |    |
| [0246] | <220>                                               |    |
| [0247] | <223> 引物                                            |    |
| [0248] | <400>28                                             |    |
| [0249] | cgtggccgag gagcaggact gacacgttct agaccatcat c       | 41 |
| [0250] | <210>29                                             |    |
| [0251] | <211>22                                             |    |
| [0252] | <212>DNA                                            |    |
| [0253] | <213> 人工的                                           |    |
| [0254] | <220>                                               |    |
| [0255] | <223> 引物                                            |    |
| [0256] | <400>29                                             |    |
| [0257] | tccaaagacg aaaggttgaa tg                            | 22 |
| [0258] | <210>30                                             |    |
| [0259] | <211>23                                             |    |
| [0260] | <212>DNA                                            |    |
| [0261] | <213> 人工的                                           |    |
| [0262] | <220>                                               |    |
| [0263] | <223> 引物                                            |    |
| [0264] | <400>30                                             |    |
| [0265] | ccgtatgtag catcaccttc acc                           | 23 |
| [0266] | <210>31                                             |    |
| [0267] | <211>15                                             |    |
| [0268] | <212>DNA                                            |    |
| [0269] | <213> 人工的                                           |    |
| [0270] | <220>                                               |    |
| [0271] | <223>A0X1 的转录因子结合位点                                 |    |
| [0272] | <400>31                                             |    |

|        |                         |    |
|--------|-------------------------|----|
| [0273] | ctgtggatgt cggat        | 15 |
| [0274] | <210>32                 |    |
| [0275] | <211>21                 |    |
| [0276] | <212>DNA                |    |
| [0277] | <213> 人工的               |    |
| [0278] | <220>                   |    |
| [0279] | <223>A0X1 的转录因子结合位点     |    |
| [0280] | <400>32                 |    |
| [0281] | agaagaggag tggaggtcct g | 21 |
| [0282] | <210>33                 |    |
| [0283] | <211>13                 |    |
| [0284] | <212>DNA                |    |
| [0285] | <213> 人工的               |    |
| [0286] | <220>                   |    |
| [0287] | <223>A0X1 的转录因子结合位点     |    |
| [0288] | <400>33                 |    |
| [0289] | caagcccaat aac          | 13 |
| [0290] | <210>34                 |    |
| [0291] | <211>13                 |    |
| [0292] | <212>DNA                |    |
| [0293] | <213> 人工的               |    |
| [0294] | <220>                   |    |
| [0295] | <223>A0X1 的转录因子结合位点     |    |
| [0296] | <400>34                 |    |
| [0297] | gagctccaat caa          | 13 |
| [0298] | <210>35                 |    |
| [0299] | <211>15                 |    |
| [0300] | <212>DNA                |    |
| [0301] | <213> 人工的               |    |
| [0302] | <220>                   |    |
| [0303] | <223>A0X1 的转录因子结合位点     |    |
| [0304] | <400>35                 |    |
| [0305] | ctcgetcatt ccaat        | 15 |
| [0306] | <210>36                 |    |
| [0307] | <211>9                  |    |
| [0308] | <212>DNA                |    |
| [0309] | <213> 人工的               |    |
| [0310] | <220>                   |    |
| [0311] | <223>A0X1 的转录因子结合位点     |    |

|        |                       |    |
|--------|-----------------------|----|
| [0312] | <400>36               |    |
| [0313] | ccagggggg             | 9  |
| [0314] | <210>37               |    |
| [0315] | <211>15               |    |
| [0316] | <212>DNA              |    |
| [0317] | <213> 人工的             |    |
| [0318] | <220>                 |    |
| [0319] | <223>A0X1 的转录因子结合位点   |    |
| [0320] | <400>37               |    |
| [0321] | tacacccgaa catca      | 15 |
| [0322] | <210>38               |    |
| [0323] | <211>7                |    |
| [0324] | <212>DNA              |    |
| [0325] | <213> 人工的             |    |
| [0326] | <220>                 |    |
| [0327] | <223>A0X1 的转录因子结合位点   |    |
| [0328] | <400>38               |    |
| [0329] | tggggtc               | 7  |
| [0330] | <210>39               |    |
| [0331] | <211>20               |    |
| [0332] | <212>DNA              |    |
| [0333] | <213> 人工的             |    |
| [0334] | <220>                 |    |
| [0335] | <223>A0X1 的转录因子结合位点   |    |
| [0336] | <400>39               |    |
| [0337] | agaaacttcc aaaagtcggc | 20 |
| [0338] | <210>40               |    |
| [0339] | <211>13               |    |
| [0340] | <212>DNA              |    |
| [0341] | <213> 人工的             |    |
| [0342] | <220>                 |    |
| [0343] | <223>A0X1 的转录因子结合位点   |    |
| [0344] | <400>40               |    |
| [0345] | atcatccaaa aag        | 13 |
| [0346] | <210>41               |    |
| [0347] | <211>11               |    |
| [0348] | <212>DNA              |    |
| [0349] | <213> 人工的             |    |
| [0350] | <220>                 |    |

|        |                                                                    |    |
|--------|--------------------------------------------------------------------|----|
| [0351] | <223>A0X1 的转录因子结合位点                                                |    |
| [0352] | <400>41                                                            |    |
| [0353] | tgcatgtgtct c                                                      | 11 |
| [0354] | <210>42                                                            |    |
| [0355] | <211>15                                                            |    |
| [0356] | <212>DNA                                                           |    |
| [0357] | <213> 人工的                                                          |    |
| [0358] | <220>                                                              |    |
| [0359] | <223>A0X1 的转录因子结合位点                                                |    |
| [0360] | <400>42                                                            |    |
| [0361] | atgcttccaa gattc                                                   | 15 |
| [0362] | <210>43                                                            |    |
| [0363] | <211>21                                                            |    |
| [0364] | <212>DNA                                                           |    |
| [0365] | <213> 人工的                                                          |    |
| [0366] | <220>                                                              |    |
| [0367] | <223>A0X1 的转录因子结合位点                                                |    |
| [0368] | <400>43                                                            |    |
| [0369] | acagttaaatt tttgatcatg a                                           | 21 |
| [0370] | <210>44                                                            |    |
| [0371] | <211>66                                                            |    |
| [0372] | <212>DNA                                                           |    |
| [0373] | <213> 人工的                                                          |    |
| [0374] | <220>                                                              |    |
| [0375] | <223> 从 A0X1 启动子删除的核酸片段                                            |    |
| [0376] | <400>44                                                            |    |
| [0377] | tttgccatcg gaaaaaccagc ccagttattg ggcttgattg gagctcgctc attccaattc | 60 |
| [0378] | cttcta                                                             | 66 |
| [0379] | <210>45                                                            |    |
| [0380] | <211>47                                                            |    |
| [0381] | <212>DNA                                                           |    |
| [0382] | <213> 人工的                                                          |    |
| [0383] | <220>                                                              |    |
| [0384] | <223> 从 A0X1 启动子删除的核酸片段                                            |    |
| [0385] | <400>45                                                            |    |
| [0386] | ttattttccga atgcaacaag ctccgcatta caccgaaca tcactcc                | 47 |
| [0387] | <210>46                                                            |    |
| [0388] | <211>30                                                            |    |
| [0389] | <212>DNA                                                           |    |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
| [0390] | <213> 人工的                                                         |    |
| [0391] | <220>                                                             |    |
| [0392] | <223> 从 AOX1 启动子删除的核酸片段                                           |    |
| [0393] | <400>46                                                           |    |
| [0394] | ctgagtgtgg ggtcaaataag ttcatgttc                                  | 30 |
| [0395] | <210>47                                                           |    |
| [0396] | <211>43                                                           |    |
| [0397] | <212>DNA                                                          |    |
| [0398] | <213> 人工的                                                         |    |
| [0399] | <220>                                                             |    |
| [0400] | <223> 从 AOX1 启动子删除的核酸片段                                           |    |
| [0401] | <400>47                                                           |    |
| [0402] | gtcaaaaaga aacttccaaa agtcggcata ccgtttgtct tgt                   | 43 |
| [0403] | <210>48                                                           |    |
| [0404] | <211>59                                                           |    |
| [0405] | <212>DNA                                                          |    |
| [0406] | <213> 人工的                                                         |    |
| [0407] | <220>                                                             |    |
| [0408] | <223> 从 AOX1 启动子删除的核酸片段                                           |    |
| [0409] | <400>48                                                           |    |
| [0410] | ccggtgcacc tgtgccgaaa cgcaaattggg gaaacacccg ctttttggat gattatgca | 59 |
| [0411] | <210>49                                                           |    |
| [0412] | <211>30                                                           |    |
| [0413] | <212>DNA                                                          |    |
| [0414] | <213> 人工的                                                         |    |
| [0415] | <220>                                                             |    |
| [0416] | <223> 从 AOX1 启动子删除的核酸片段                                           |    |
| [0417] | <400>49                                                           |    |
| [0418] | attgtatgct tccaagattc tgggtgggaat                                 | 30 |
| [0419] | <210>50                                                           |    |
| [0420] | <211>35                                                           |    |
| [0421] | <212>DNA                                                          |    |
| [0422] | <213> 人工的                                                         |    |
| [0423] | <220>                                                             |    |
| [0424] | <223> 从 AOX1 启动子删除的核酸片段                                           |    |
| [0425] | <400>50                                                           |    |
| [0426] | tgatagccta acgttcatga tcaaaattta actgt                            | 35 |
| [0427] | <210>51                                                           |    |
| [0428] | <211>17                                                           |    |

|        |                                             |    |
|--------|---------------------------------------------|----|
| [0429] | <212>DNA                                    |    |
| [0430] | <213> 人工的                                   |    |
| [0431] | <220>                                       |    |
| [0432] | <223> 从 AOX1 启动子删除的核酸片段                     |    |
| [0433] | <400>51                                     |    |
| [0434] | aatatataaa cagaagg                          | 17 |
| [0435] | <210>52                                     |    |
| [0436] | <211>39                                     |    |
| [0437] | <212>DNA                                    |    |
| [0438] | <213> 人工的                                   |    |
| [0439] | <220>                                       |    |
| [0440] | <223> 从 AOX1 启动子删除的核酸片段                     |    |
| [0441] | <400>52                                     |    |
| [0442] | tttttttatac atcattatta gcttactttc ataattgcg | 39 |
| [0443] | <210>53                                     |    |
| [0444] | <211>40                                     |    |
| [0445] | <212>DNA                                    |    |
| [0446] | <213> 人工的                                   |    |
| [0447] | <220>                                       |    |
| [0448] | <223> 引物                                    |    |
| [0449] | <400>53                                     |    |
| [0450] | gaatgaaacc tttttgccat atccacaggt ccattctcac | 40 |
| [0451] | <210>54                                     |    |
| [0452] | <211>40                                     |    |
| [0453] | <212>DNA                                    |    |
| [0454] | <213> 人工的                                   |    |
| [0455] | <220>                                       |    |
| [0456] | <223> 引物                                    |    |
| [0457] | <400>54                                     |    |
| [0458] | gaatggacct gtggatatgg caaaaaggtt tcattcaacc | 40 |
| [0459] | <210>55                                     |    |
| [0460] | <211>38                                     |    |
| [0461] | <212>DNA                                    |    |
| [0462] | <213> 人工的                                   |    |
| [0463] | <220>                                       |    |
| [0464] | <223> 引物                                    |    |
| [0465] | <400>55                                     |    |
| [0466] | ccgttgcaaa cgcaggacct cttctcctca acacccac   | 38 |
| [0467] | <210>56                                     |    |

|        |                                               |    |
|--------|-----------------------------------------------|----|
| [0468] | <211>38                                       |    |
| [0469] | <212>DNA                                      |    |
| [0470] | <213> 人工的                                     |    |
| [0471] | <220>                                         |    |
| [0472] | <223> 引物                                      |    |
| [0473] | <400>56                                       |    |
| [0474] | gtgttgagga gaagaggtcc tgcgtttgca acggtctg     | 38 |
| [0475] | <210>57                                       |    |
| [0476] | <211>42                                       |    |
| [0477] | <212>DNA                                      |    |
| [0478] | <213> 人工的                                     |    |
| [0479] | <220>                                         |    |
| [0480] | <223> 引物                                      |    |
| [0481] | <400>57                                       |    |
| [0482] | cgaaaaacca gccagttgc ttgattggag ctcgctcatt cc | 42 |
| [0483] | <210>58                                       |    |
| [0484] | <211>40                                       |    |
| [0485] | <212>DNA                                      |    |
| [0486] | <213> 人工的                                     |    |
| [0487] | <220>                                         |    |
| [0488] | <223> 引物                                      |    |
| [0489] | <400>58                                       |    |
| [0490] | gagcgagctc caatcaagca actgggctgg tttttcgatg   | 40 |
| [0491] | <210>59                                       |    |
| [0492] | <211>40                                       |    |
| [0493] | <212>DNA                                      |    |
| [0494] | <213> 人工的                                     |    |
| [0495] | <220>                                         |    |
| [0496] | <223> 引物                                      |    |
| [0497] | <400>59                                       |    |
| [0498] | cagccagtt attgggcttg agctcgctca ttccaattcc    | 40 |
| [0499] | <210>60                                       |    |
| [0500] | <211>40                                       |    |
| [0501] | <212>DNA                                      |    |
| [0502] | <213> 人工的                                     |    |
| [0503] | <220>                                         |    |
| [0504] | <223> 引物                                      |    |
| [0505] | <400>60                                       |    |
| [0506] | ggaattggaa tgagcgagct caagcccaat aactgggctg   | 40 |

|        |                                                |    |
|--------|------------------------------------------------|----|
| [0507] | <210>61                                        |    |
| [0508] | <211>39                                        |    |
| [0509] | <212>DNA                                       |    |
| [0510] | <213> 人工的                                      |    |
| [0511] | <220>                                          |    |
| [0512] | <223> 引物                                       |    |
| [0513] | <400>61                                        |    |
| [0514] | ggcttgattg gagctcgcta attccttcta ttaggctac     | 39 |
| [0515] | <210>62                                        |    |
| [0516] | <211>39                                        |    |
| [0517] | <212>DNA                                       |    |
| [0518] | <213> 人工的                                      |    |
| [0519] | <220>                                          |    |
| [0520] | <223> 引物                                       |    |
| [0521] | <400>62                                        |    |
| [0522] | gtagcctaataa agaaggaatt agcgagctcc aatcaagcc   | 39 |
| [0523] | <210>63                                        |    |
| [0524] | <211>42                                        |    |
| [0525] | <212>DNA                                       |    |
| [0526] | <213> 人工的                                      |    |
| [0527] | <220>                                          |    |
| [0528] | <223> 引物                                       |    |
| [0529] | <400>63                                        |    |
| [0530] | gcctgtctat cctggccggc gaggttcatg tttgtttatt tc | 42 |
| [0531] | <210>64                                        |    |
| [0532] | <211>42                                        |    |
| [0533] | <212>DNA                                       |    |
| [0534] | <213> 人工的                                      |    |
| [0535] | <220>                                          |    |
| [0536] | <223> 引物                                       |    |
| [0537] | <400>64                                        |    |
| [0538] | caaacatgaa cctcgccggc caggatagac aggctaataa ag | 42 |
| [0539] | <210>65                                        |    |
| [0540] | <211>38                                        |    |
| [0541] | <212>DNA                                       |    |
| [0542] | <213> 人工的                                      |    |
| [0543] | <220>                                          |    |
| [0544] | <223> 引物                                       |    |
| [0545] | <400>65                                        |    |

|        |                                             |    |
|--------|---------------------------------------------|----|
| [0546] | gcaacaagct ccgcattaca acatcactcc agatgagg   | 38 |
| [0547] | <210>66                                     |    |
| [0548] | <211>38                                     |    |
| [0549] | <212>DNA                                    |    |
| [0550] | <213> 人工的                                   |    |
| [0551] | <220>                                       |    |
| [0552] | <223> 引物                                    |    |
| [0553] | <400>66                                     |    |
| [0554] | cctcatctgg agtgatgttg taatgcggag cttgttgc   | 38 |
| [0555] | <210>67                                     |    |
| [0556] | <211>40                                     |    |
| [0557] | <212>DNA                                    |    |
| [0558] | <213> 人工的                                   |    |
| [0559] | <220>                                       |    |
| [0560] | <223> 引物                                    |    |
| [0561] | <400>67                                     |    |
| [0562] | ccagatgagg gctttctgag tgaaatagtt tcatgttccc | 40 |
| [0563] | <210>68                                     |    |
| [0564] | <211>40                                     |    |
| [0565] | <212>DNA                                    |    |
| [0566] | <213> 人工的                                   |    |
| [0567] | <220>                                       |    |
| [0568] | <223> 引物                                    |    |
| [0569] | <400>68                                     |    |
| [0570] | gggaacatga aactatttca ctcagaaagc cctcatctgg | 40 |
| [0571] | <210>69                                     |    |
| [0572] | <211>39                                     |    |
| [0573] | <212>DNA                                    |    |
| [0574] | <213> 人工的                                   |    |
| [0575] | <220>                                       |    |
| [0576] | <223> 引物                                    |    |
| [0577] | <400>69                                     |    |
| [0578] | gccagttggc caaaaacaaa agtcggcata ccgtttgtc  | 39 |
| [0579] | <210>70                                     |    |
| [0580] | <211>40                                     |    |
| [0581] | <212>DNA                                    |    |
| [0582] | <213> 人工的                                   |    |
| [0583] | <220>                                       |    |
| [0584] | <223> 引物                                    |    |

|        |                                                |    |
|--------|------------------------------------------------|----|
| [0585] | <400>70                                        |    |
| [0586] | cggtatgccg acttttgttt ttgaccaact ggccgtagc     | 40 |
| [0587] | <210>71                                        |    |
| [0588] | <211>42                                        |    |
| [0589] | <212>DNA                                       |    |
| [0590] | <213> 人工的                                      |    |
| [0591] | <220>                                          |    |
| [0592] | <223> 引物                                       |    |
| [0593] | <400>71                                        |    |
| [0594] | caaatgggga aacacccgct tatgattatg cattgtctcc ac | 42 |
| [0595] | <210>72                                        |    |
| [0596] | <211>41                                        |    |
| [0597] | <212>DNA                                       |    |
| [0598] | <213> 人工的                                      |    |
| [0599] | <220>                                          |    |
| [0600] | <223> 引物                                       |    |
| [0601] | <400>72                                        |    |
| [0602] | gagacaatgc ataatcataa gcgggtgttt ccccatattgc g | 41 |
| [0603] | <210>73                                        |    |
| [0604] | <211>41                                        |    |
| [0605] | <212>DNA                                       |    |
| [0606] | <213> 人工的                                      |    |
| [0607] | <220>                                          |    |
| [0608] | <223> 引物                                       |    |
| [0609] | <400>73                                        |    |
| [0610] | gcttttttga tgattatgcc tccacattgt atgcttccaa g  | 41 |
| [0611] | <210>74                                        |    |
| [0612] | <211>41                                        |    |
| [0613] | <212>DNA                                       |    |
| [0614] | <213> 人工的                                      |    |
| [0615] | <220>                                          |    |
| [0616] | <223> 引物                                       |    |
| [0617] | <400>74                                        |    |
| [0618] | cttgaagca tacaatgtgg aggcataatc atccaaaaag c   | 41 |
| [0619] | <210>75                                        |    |
| [0620] | <211>42                                        |    |
| [0621] | <212>DNA                                       |    |
| [0622] | <213> 人工的                                      |    |
| [0623] | <220>                                          |    |

|        |                                                |    |
|--------|------------------------------------------------|----|
| [0624] | <223> 引物                                       |    |
| [0625] | <400>75                                        |    |
| [0626] | cattgtctcc acattgtatg aagattctgg tgggaatact gc | 42 |
| [0627] | <210>76                                        |    |
| [0628] | <211>40                                        |    |
| [0629] | <212>DNA                                       |    |
| [0630] | <213> 人工的                                      |    |
| [0631] | <220>                                          |    |
| [0632] | <223> 引物                                       |    |
| [0633] | <400>76                                        |    |
| [0634] | gtattccac cagaatcttc atacaatgtg gagacaatgc     | 40 |
| [0635] | <210>77                                        |    |
| [0636] | <211>42                                        |    |
| [0637] | <212>DNA                                       |    |
| [0638] | <213> 人工的                                      |    |
| [0639] | <220>                                          |    |
| [0640] | <223> 引物                                       |    |
| [0641] | <400>77                                        |    |
| [0642] | gctgatagcc taacgttcat gttctaacc ctacttgaca gc  | 42 |
| [0643] | <210>78                                        |    |
| [0644] | <211>41                                        |    |
| [0645] | <212>DNA                                       |    |
| [0646] | <213> 人工的                                      |    |
| [0647] | <220>                                          |    |
| [0648] | <223> 引物                                       |    |
| [0649] | <400>78                                        |    |
| [0650] | gtcaagtagg ggtagaaca tgaacgttag gctatcagca g   | 41 |
| [0651] | <210>79                                        |    |
| [0652] | <211>41                                        |    |
| [0653] | <212>DNA                                       |    |
| [0654] | <213> 人工的                                      |    |
| [0655] | <220>                                          |    |
| [0656] | <223> 引物                                       |    |
| [0657] | <400>79                                        |    |
| [0658] | ggaatactgc tgatagcttc atgatcaaaa tttactgtt c   | 41 |
| [0659] | <210>80                                        |    |
| [0660] | <211>40                                        |    |
| [0661] | <212>DNA                                       |    |
| [0662] | <213> 人工的                                      |    |

|        |                                               |    |
|--------|-----------------------------------------------|----|
| [0663] | <220>                                         |    |
| [0664] | <223> 引物                                      |    |
| [0665] | <400>80                                       |    |
| [0666] | gttaaatttt gatcatgaag ctatcagcag tattcccacc   | 40 |
| [0667] | <210>81                                       |    |
| [0668] | <211>42                                       |    |
| [0669] | <212>DNA                                      |    |
| [0670] | <213> 人工的                                     |    |
| [0671] | <220>                                         |    |
| [0672] | <223> 引物                                      |    |
| [0673] | <400>81                                       |    |
| [0674] | ggaatactgc tgatagccac gttcatgac aaaatttaac tg | 42 |
| [0675] | <210>82                                       |    |
| [0676] | <211>40                                       |    |
| [0677] | <212>DNA                                      |    |
| [0678] | <213> 人工的                                     |    |
| [0679] | <220>                                         |    |
| [0680] | <223> 引物                                      |    |
| [0681] | <400>82                                       |    |
| [0682] | gttaaatttt gatcatgaac gtggctatca gcagtattcc   | 40 |
| [0683] | <210>83                                       |    |
| [0684] | <211>37                                       |    |
| [0685] | <212>DNA                                      |    |
| [0686] | <213> 人工的                                     |    |
| [0687] | <220>                                         |    |
| [0688] | <223> 引物                                      |    |
| [0689] | <400>83                                       |    |
| [0690] | gagatcagat ctaacatatg ccaaagacga aaggttg      | 37 |
| [0691] | <210>84                                       |    |
| [0692] | <211>37                                       |    |
| [0693] | <212>DNA                                      |    |
| [0694] | <213> 人工的                                     |    |
| [0695] | <220>                                         |    |
| [0696] | <223> 引物                                      |    |
| [0697] | <400>84                                       |    |
| [0698] | caacctttcg tctttggcat atgttagatc tgatctc      | 37 |
| [0699] | <210>85                                       |    |
| [0700] | <211>34                                       |    |
| [0701] | <212>DNA                                      |    |

|        |                                                       |    |
|--------|-------------------------------------------------------|----|
| [0702] | <213> 人工的                                             |    |
| [0703] | <220>                                                 |    |
| [0704] | <223> 引物                                              |    |
| [0705] | <400>85                                               |    |
| [0706] | aaacatatga gatctaacaat ccaaagacga aagg                | 34 |
| [0707] | <210>86                                               |    |
| [0708] | <211>27                                               |    |
| [0709] | <212>DNA                                              |    |
| [0710] | <213> 人工的                                             |    |
| [0711] | <220>                                                 |    |
| [0712] | <223> 引物                                              |    |
| [0713] | <400>86                                               |    |
| [0714] | tggttgaatt ctttcaataa ttagttg                         | 27 |
| [0715] | <210>87                                               |    |
| [0716] | <211>46                                               |    |
| [0717] | <212>DNA                                              |    |
| [0718] | <213> 人工的                                             |    |
| [0719] | <220>                                                 |    |
| [0720] | <223> 引物                                              |    |
| [0721] | <400>87                                               |    |
| [0722] | cgccacaaca ttgaagatgg ttccgticac ctagcagacc attatc    | 46 |
| [0723] | <210>88                                               |    |
| [0724] | <211>49                                               |    |
| [0725] | <212>DNA                                              |    |
| [0726] | <213> 人工的                                             |    |
| [0727] | <220>                                                 |    |
| [0728] | <223> 引物                                              |    |
| [0729] | <400>88                                               |    |
| [0730] | ggaaacattc tcggacacaa acttgagtac aactataact cacacaatg | 49 |
| [0731] | <210>89                                               |    |
| [0732] | <211>44                                               |    |
| [0733] | <212>DNA                                              |    |
| [0734] | <213> 人工的                                             |    |
| [0735] | <220>                                                 |    |
| [0736] | <223> 引物                                              |    |
| [0737] | <400>89                                               |    |
| [0738] | atctcgagtt acttgtacaa ttcattcatg ccatgtgtaa tccc      | 44 |
| [0739] | <210>90                                               |    |
| [0740] | <211>178                                              |    |

|        |                                                                   |     |
|--------|-------------------------------------------------------------------|-----|
| [0741] | <212>DNA                                                          |     |
| [0742] | <213> 人工的                                                         |     |
| [0743] | <220>                                                             |     |
| [0744] | <223>A0X1 启动子片段                                                   |     |
| [0745] | <400>90                                                           |     |
| [0746] | cttaaggaca gcaatatata aacagaagga agctgccctg tcttaaacct tttttttat  | 60  |
| [0747] | catcattatt agcttacttt cataattgcg actggttcca attgacaagc ttttgatttt | 120 |
| [0748] | aacgactttt aacgacaact tgagaagatc aaaaaacaac taattattga aagaattc   | 178 |
| [0749] | <210>91                                                           |     |
| [0750] | <211>195                                                          |     |
| [0751] | <212>DNA                                                          |     |
| [0752] | <213> 人工的                                                         |     |
| [0753] | <220>                                                             |     |
| [0754] | <223>A0X1 启动子片段                                                   |     |
| [0755] | <400>91                                                           |     |
| [0756] | cttaagtgtt ctaacccta cttgacagca atatataaac agaaggaagc tgccctgtct  | 60  |
| [0757] | taaacccttt tttttatcat cattattagc ttactttcat aattgcgact ggttccaatt | 120 |
| [0758] | gacaagcttt tgattttaac gacttttaac gacaacttga gaagatcaaa aaacaactaa | 180 |
| [0759] | ttattgaaag aattc                                                  | 195 |
| [0760] | <210>92                                                           |     |
| [0761] | <211>12                                                           |     |
| [0762] | <212>DNA                                                          |     |
| [0763] | <213> 人工的                                                         |     |
| [0764] | <220>                                                             |     |
| [0765] | <223>A0X1 启动子片段                                                   |     |
| [0766] | <400>92                                                           |     |
| [0767] | tagcctaacg tt                                                     | 12  |
| [0768] | <210>93                                                           |     |
| [0769] | <211>14                                                           |     |
| [0770] | <212>DNA                                                          |     |
| [0771] | <213> 人工的                                                         |     |
| [0772] | <220>                                                             |     |
| [0773] | <223>A0X1 启动子片段                                                   |     |
| [0774] | <400>93                                                           |     |
| [0775] | catgatcaaa attt                                                   | 14  |
| [0776] | <210>94                                                           |     |
| [0777] | <211>32                                                           |     |
| [0778] | <212>DNA                                                          |     |
| [0779] | <213> 人工的                                                         |     |

|        |                                                               |    |
|--------|---------------------------------------------------------------|----|
| [0780] | <220>                                                         |    |
| [0781] | <223> 引物                                                      |    |
| [0782] | <400>94                                                       |    |
| [0783] | tttgaattct ttcaataatt agttgttttt tg                           | 32 |
| [0784] | <210>95                                                       |    |
| [0785] | <211>43                                                       |    |
| [0786] | <212>DNA                                                      |    |
| [0787] | <213> 人工的                                                     |    |
| [0788] | <220>                                                         |    |
| [0789] | <223> 引物                                                      |    |
| [0790] | <400>95                                                       |    |
| [0791] | ttagatctcg acttaaggac agcaatatat aaacagaagg aag               | 43 |
| [0792] | <210>96                                                       |    |
| [0793] | <211>39                                                       |    |
| [0794] | <212>DNA                                                      |    |
| [0795] | <213> 人工的                                                     |    |
| [0796] | <220>                                                         |    |
| [0797] | <223> 引物                                                      |    |
| [0798] | <400>96                                                       |    |
| [0799] | ttagatctcg acttaagtgt tctaaccctt acttgacag                    | 39 |
| [0800] | <210>97                                                       |    |
| [0801] | <211>52                                                       |    |
| [0802] | <212>DNA                                                      |    |
| [0803] | <213> 人工的                                                     |    |
| [0804] | <220>                                                         |    |
| [0805] | <223> 引物                                                      |    |
| [0806] | <400>97                                                       |    |
| [0807] | aaagatctta gcctaacgtt cttaaggaca gcaatatata aacagaagga ag     | 52 |
| [0808] | <210>98                                                       |    |
| [0809] | <211>54                                                       |    |
| [0810] | <212>DNA                                                      |    |
| [0811] | <213> 人工的                                                     |    |
| [0812] | <220>                                                         |    |
| [0813] | <223> 引物                                                      |    |
| [0814] | <400>98                                                       |    |
| [0815] | aaagatctca tgatcaaaat ttcttaagga cagcaatatata taaacagaag gaag | 54 |
| [0816] | <210>99                                                       |    |
| [0817] | <211>57                                                       |    |
| [0818] | <212>DNA                                                      |    |

|        |                                                                  |    |
|--------|------------------------------------------------------------------|----|
| [0819] | <213> 人工的                                                        |    |
| [0820] | <220>                                                            |    |
| [0821] | <223> 引物                                                         |    |
| [0822] | <400>99                                                          |    |
| [0823] | gatctcgact taagcaatcg tcttactttc taactttttct tacctttttac atttcag | 57 |
| [0824] | <210>100                                                         |    |
| [0825] | <211>31                                                          |    |
| [0826] | <212>DNA                                                         |    |
| [0827] | <213> 人工的                                                        |    |
| [0828] | <220>                                                            |    |
| [0829] | <223> 引物                                                         |    |
| [0830] | <400>100                                                         |    |
| [0831] | caatatatat atatatttca aggatatacc g                               | 31 |
| [0832] | <210>101                                                         |    |
| [0833] | <211>49                                                          |    |
| [0834] | <212>DNA                                                         |    |
| [0835] | <213> 人工的                                                        |    |
| [0836] | <220>                                                            |    |
| [0837] | <223> 引物                                                         |    |
| [0838] | <400>101                                                         |    |
| [0839] | aattcggtat atccttgaaa tatatatata tattgctgaa atgtaaaag            | 49 |
| [0840] | <210>102                                                         |    |
| [0841] | <211>39                                                          |    |
| [0842] | <212>DNA                                                         |    |
| [0843] | <213> 人工的                                                        |    |
| [0844] | <220>                                                            |    |
| [0845] | <223> 引物                                                         |    |
| [0846] | <400>102                                                         |    |
| [0847] | gtaagaaaag ttagaaagta agacgattgc ttaagtcca                       | 39 |
| [0848] | <210>103                                                         |    |
| [0849] | <211>22                                                          |    |
| [0850] | <212>DNA                                                         |    |
| [0851] | <213> 人工的                                                        |    |
| [0852] | <220>                                                            |    |
| [0853] | <223> 引物                                                         |    |
| [0854] | <400>103                                                         |    |
| [0855] | ggttgaattc ggtatatcct tg                                         | 22 |
| [0856] | <210>104                                                         |    |
| [0857] | <211>46                                                          |    |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
| [0858] | <212>DNA                                                          |    |
| [0859] | <213> 人工的                                                         |    |
| [0860] | <220>                                                             |    |
| [0861] | <223> 引物                                                          |    |
| [0862] | <400>104                                                          |    |
| [0863] | aaagatctta gcctaacggt cttaagcaat cgtcttactt tctaac                | 46 |
| [0864] | <210>105                                                          |    |
| [0865] | <211>15                                                           |    |
| [0866] | <212>DNA                                                          |    |
| [0867] | <213> 人工的                                                         |    |
| [0868] | <220>                                                             |    |
| [0869] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0870] | <400>105                                                          |    |
| [0871] | gccatccgac atcca                                                  | 15 |
| [0872] | <210>106                                                          |    |
| [0873] | <211>18                                                           |    |
| [0874] | <212>DNA                                                          |    |
| [0875] | <213> 人工的                                                         |    |
| [0876] | <220>                                                             |    |
| [0877] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0878] | <400>106                                                          |    |
| [0879] | ggacctccac tcctcttc                                               | 18 |
| [0880] | <210>107                                                          |    |
| [0881] | <211>76                                                           |    |
| [0882] | <212>DNA                                                          |    |
| [0883] | <213> 人工的                                                         |    |
| [0884] | <220>                                                             |    |
| [0885] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0886] | <400>107                                                          |    |
| [0887] | ccacttttgc catcgaaaaa ccagcccagt tattgggctt gattggagct cgctcattcc | 60 |
| [0888] | aattccttct attagg                                                 | 76 |
| [0889] | <210>108                                                          |    |
| [0890] | <211>15                                                           |    |
| [0891] | <212>DNA                                                          |    |
| [0892] | <213> 人工的                                                         |    |
| [0893] | <220>                                                             |    |
| [0894] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0895] | <400>108                                                          |    |
| [0896] | cagtattgg gcttg                                                   | 15 |

|        |                                                                 |    |
|--------|-----------------------------------------------------------------|----|
| [0897] | <210>109                                                        |    |
| [0898] | <211>15                                                         |    |
| [0899] | <212>DNA                                                        |    |
| [0900] | <213> 人工的                                                       |    |
| [0901] | <220>                                                           |    |
| [0902] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [0903] | <400>109                                                        |    |
| [0904] | gcttgattgg agctc                                                | 15 |
| [0905] | <210>110                                                        |    |
| [0906] | <211>16                                                         |    |
| [0907] | <212>DNA                                                        |    |
| [0908] | <213> 人工的                                                       |    |
| [0909] | <220>                                                           |    |
| [0910] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [0911] | <400>110                                                        |    |
| [0912] | tcgctcattc caattc                                               | 16 |
| [0913] | <210>111                                                        |    |
| [0914] | <211>15                                                         |    |
| [0915] | <212>DNA                                                        |    |
| [0916] | <213> 人工的                                                       |    |
| [0917] | <220>                                                           |    |
| [0918] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [0919] | <400>111                                                        |    |
| [0920] | tggccccctt ggcga                                                | 15 |
| [0921] | <210>112                                                        |    |
| [0922] | <211>57                                                         |    |
| [0923] | <212>DNA                                                        |    |
| [0924] | <213> 人工的                                                       |    |
| [0925] | <220>                                                           |    |
| [0926] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [0927] | <400>112                                                        |    |
| [0928] | tttgttttatt tccgaatgca acaagctccg cattacaccc gaacatcact ccagatg | 57 |
| [0929] | <210>113                                                        |    |
| [0930] | <211>15                                                         |    |
| [0931] | <212>DNA                                                        |    |
| [0932] | <213> 人工的                                                       |    |
| [0933] | <220>                                                           |    |
| [0934] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [0935] | <400>113                                                        |    |

|        |                                                            |    |
|--------|------------------------------------------------------------|----|
| [0936] | attacacccg aacat                                           | 15 |
| [0937] | <210>114                                                   |    |
| [0938] | <211>40                                                    |    |
| [0939] | <212>DNA                                                   |    |
| [0940] | <213> 人工的                                                  |    |
| [0941] | <220>                                                      |    |
| [0942] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                           |    |
| [0943] | <400>114                                                   |    |
| [0944] | gctttctgag tgtggggtca aatagtttca tgttccccaa                | 40 |
| [0945] | <210>115                                                   |    |
| [0946] | <211>17                                                    |    |
| [0947] | <212>DNA                                                   |    |
| [0948] | <213> 人工的                                                  |    |
| [0949] | <220>                                                      |    |
| [0950] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                           |    |
| [0951] | <400>115                                                   |    |
| [0952] | gagtgtgggg tcaaata                                         | 17 |
| [0953] | <210>116                                                   |    |
| [0954] | <211>53                                                    |    |
| [0955] | <212>DNA                                                   |    |
| [0956] | <213> 人工的                                                  |    |
| [0957] | <220>                                                      |    |
| [0958] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                           |    |
| [0959] | <400>116                                                   |    |
| [0960] | agttgacaag acaaacggtg tgccgacttt tggaagtttc tttttgactt ggt | 53 |
| [0961] | <210>117                                                   |    |
| [0962] | <211>18                                                    |    |
| [0963] | <212>DNA                                                   |    |
| [0964] | <213> 人工的                                                  |    |
| [0965] | <220>                                                      |    |
| [0966] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                           |    |
| [0967] | <400>117                                                   |    |
| [0968] | aaaaagaaac ttccaaaa                                        | 18 |
| [0969] | <210>118                                                   |    |
| [0970] | <211>69                                                    |    |
| [0971] | <212>DNA                                                   |    |
| [0972] | <213> 人工的                                                  |    |
| [0973] | <220>                                                      |    |
| [0974] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                           |    |

|        |                                                                   |    |
|--------|-------------------------------------------------------------------|----|
| [0975] | <400>118                                                          |    |
| [0976] | gaaccccggt gcacctgtgc cgaaacgcaa atggggaaac acccgctttt tggatgatta | 60 |
| [0977] | tgcattgtc                                                         | 69 |
| [0978] | <210>119                                                          |    |
| [0979] | <211>15                                                           |    |
| [0980] | <212>DNA                                                          |    |
| [0981] | <213> 人工的                                                         |    |
| [0982] | <220>                                                             |    |
| [0983] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0984] | <400>119                                                          |    |
| [0985] | cgcttttttg atgat                                                  | 15 |
| [0986] | <210>120                                                          |    |
| [0987] | <211>15                                                           |    |
| [0988] | <212>DNA                                                          |    |
| [0989] | <213> 人工的                                                         |    |
| [0990] | <220>                                                             |    |
| [0991] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [0992] | <400>120                                                          |    |
| [0993] | tatgcattgt ctcca                                                  | 15 |
| [0994] | <210>121                                                          |    |
| [0995] | <211>40                                                           |    |
| [0996] | <212>DNA                                                          |    |
| [0997] | <213> 人工的                                                         |    |
| [0998] | <220>                                                             |    |
| [0999] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [1000] | <400>121                                                          |    |
| [1001] | tccacattgt atgcttccaa gattctggtg ggaatactgc                       | 40 |
| [1002] | <210>122                                                          |    |
| [1003] | <211>12                                                           |    |
| [1004] | <212>DNA                                                          |    |
| [1005] | <213> 人工的                                                         |    |
| [1006] | <220>                                                             |    |
| [1007] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                  |    |
| [1008] | <400>122                                                          |    |
| [1009] | tagcctaacg tt                                                     | 12 |
| [1010] | <210>123                                                          |    |
| [1011] | <211>15                                                           |    |
| [1012] | <212>DNA                                                          |    |
| [1013] | <213> 人工的                                                         |    |

|        |                                                    |    |
|--------|----------------------------------------------------|----|
| [1014] | <220>                                              |    |
| [1015] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                   |    |
| [1016] | <400>123                                           |    |
| [1017] | gtatgcttcc aagat                                   | 15 |
| [1018] | <210>124                                           |    |
| [1019] | <211>45                                            |    |
| [1020] | <212>DNA                                           |    |
| [1021] | <213> 人工的                                          |    |
| [1022] | <220>                                              |    |
| [1023] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                   |    |
| [1024] | <400>124                                           |    |
| [1025] | actgctgata gcctaacggtt catgatcaaa atttaactgt tctaa | 45 |
| [1026] | <210>125                                           |    |
| [1027] | <211>25                                            |    |
| [1028] | <212>DNA                                           |    |
| [1029] | <213> 人工的                                          |    |
| [1030] | <220>                                              |    |
| [1031] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                   |    |
| [1032] | <400>125                                           |    |
| [1033] | ttcgatgatca aaatttaact gttct                       | 25 |
| [1034] | <210>126                                           |    |
| [1035] | <211>36                                            |    |
| [1036] | <212>DNA                                           |    |
| [1037] | <213> 人工的                                          |    |
| [1038] | <220>                                              |    |
| [1039] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                   |    |
| [1040] | <400>126                                           |    |
| [1041] | atagcctaac gttcatgatc aaaatttaac tgttct            | 36 |
| [1042] | <210>127                                           |    |
| [1043] | <211>27                                            |    |
| [1044] | <212>DNA                                           |    |
| [1045] | <213> 人工的                                          |    |
| [1046] | <220>                                              |    |
| [1047] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                   |    |
| [1048] | <400>127                                           |    |
| [1049] | acagcaatat ataaacagaa ggaagct                      | 27 |
| [1050] | <210>128                                           |    |
| [1051] | <211>49                                            |    |
| [1052] | <212>DNA                                           |    |

|        |                                                                 |    |
|--------|-----------------------------------------------------------------|----|
| [1053] | <213> 人工的                                                       |    |
| [1054] | <220>                                                           |    |
| [1055] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [1056] | <400>128                                                        |    |
| [1057] | accttttttt ttatcatcat tattagctta ctttcataat tgcgactgg           | 49 |
| [1058] | <210>129                                                        |    |
| [1059] | <211>16                                                         |    |
| [1060] | <212>DNA                                                        |    |
| [1061] | <213> 人工的                                                       |    |
| [1062] | <220>                                                           |    |
| [1063] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [1064] | <400>129                                                        |    |
| [1065] | atagcctaac gttcat                                               | 16 |
| [1066] | <210>130                                                        |    |
| [1067] | <211>85                                                         |    |
| [1068] | <212>DNA                                                        |    |
| [1069] | <213> 人工的                                                       |    |
| [1070] | <220>                                                           |    |
| [1071] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [1072] | <400>130                                                        |    |
| [1073] | cttggaacct aatatgacaa aagcgtgac tcaccaaga tgaactaagt ttggttcggt | 60 |
| [1074] | gaaatgctaa cggccagttg cttgg                                     | 85 |
| [1075] | <210>131                                                        |    |
| [1076] | <211>32                                                         |    |
| [1077] | <212>DNA                                                        |    |
| [1078] | <213> 人工的                                                       |    |
| [1079] | <220>                                                           |    |
| [1080] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [1081] | <400>131                                                        |    |
| [1082] | ccacttttgc catcgaaaaa ccagcccagt ta                             | 32 |
| [1083] | <210>132                                                        |    |
| [1084] | <211>32                                                         |    |
| [1085] | <212>DNA                                                        |    |
| [1086] | <213> 人工的                                                       |    |
| [1087] | <220>                                                           |    |
| [1088] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                                |    |
| [1089] | <400>132                                                        |    |
| [1090] | agcccagtta ttgggcttga ttggagctcg ct                             | 32 |
| [1091] | <210>133                                                        |    |

|        |                                                          |    |
|--------|----------------------------------------------------------|----|
| [1092] | <211>32                                                  |    |
| [1093] | <212>DNA                                                 |    |
| [1094] | <213> 人工的                                                |    |
| [1095] | <220>                                                    |    |
| [1096] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                         |    |
| [1097] | <400>133                                                 |    |
| [1098] | ggagctcgct cattccaatt ccttctatta gg                      | 32 |
| [1099] | <210>134                                                 |    |
| [1100] | <211>51                                                  |    |
| [1101] | <212>DNA                                                 |    |
| [1102] | <213> 人工的                                                |    |
| [1103] | <220>                                                    |    |
| [1104] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                         |    |
| [1105] | <400>134                                                 |    |
| [1106] | ccacttttgc catcgaaaaa ccagcccagt tattgggctt gattggagct c | 51 |
| [1107] | <210>135                                                 |    |
| [1108] | <211>40                                                  |    |
| [1109] | <212>DNA                                                 |    |
| [1110] | <213> 人工的                                                |    |
| [1111] | <220>                                                    |    |
| [1112] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                         |    |
| [1113] | <400>135                                                 |    |
| [1114] | aatactgctg atagcctaac gttcatgatc aaaataatac              | 40 |
| [1115] | <210>136                                                 |    |
| [1116] | <211>43                                                  |    |
| [1117] | <212>DNA                                                 |    |
| [1118] | <213> 人工的                                                |    |
| [1119] | <220>                                                    |    |
| [1120] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                         |    |
| [1121] | <400>136                                                 |    |
| [1122] | taatctcatt aatgcttagc gcagtctctc tategcttta atc          | 43 |
| [1123] | <210>137                                                 |    |
| [1124] | <211>50                                                  |    |
| [1125] | <212>DNA                                                 |    |
| [1126] | <213> 人工的                                                |    |
| [1127] | <220>                                                    |    |
| [1128] | <223> 含有待删除的转录因子结合位点的 AOX1 启动子片段                         |    |
| [1129] | <400>137                                                 |    |
| [1130] | ttctgaacce cgggtgcacct gtgccgaaac gcaaattgggg aaacacccgc | 50 |

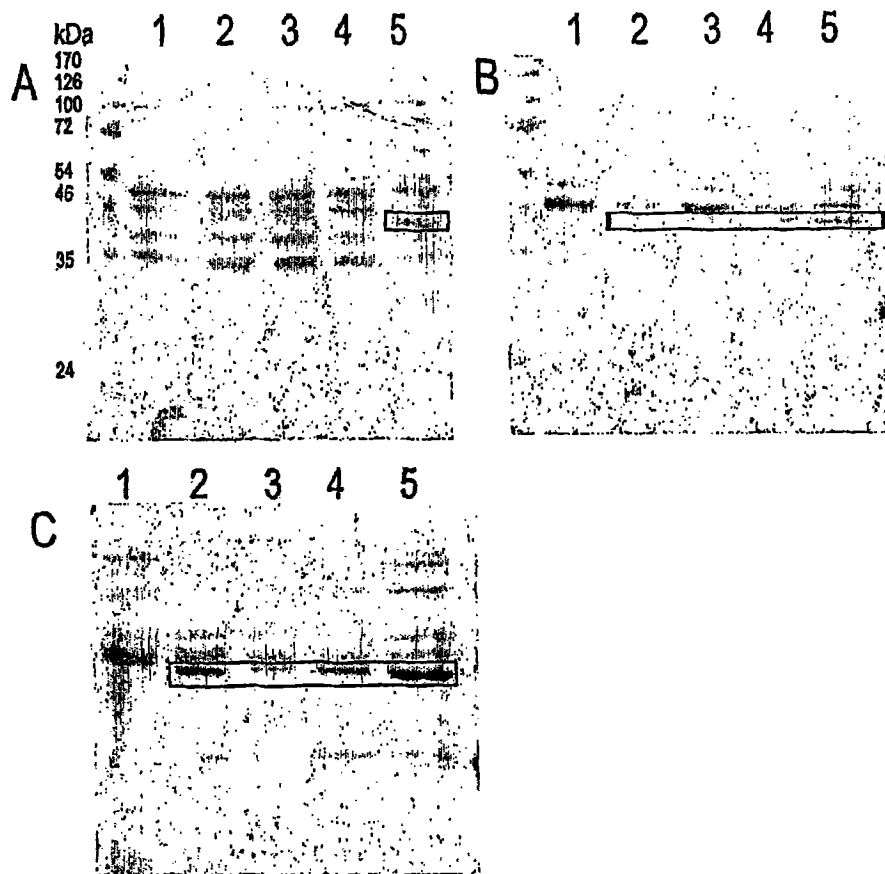


图 1

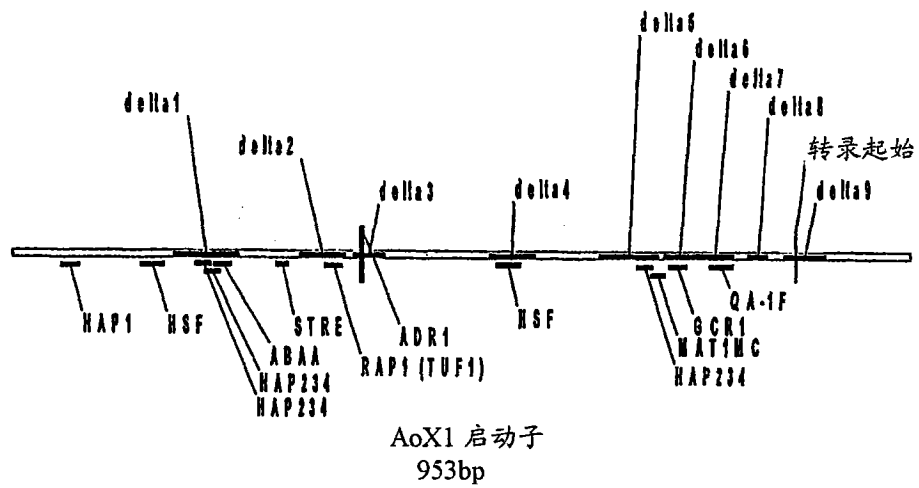


图 2

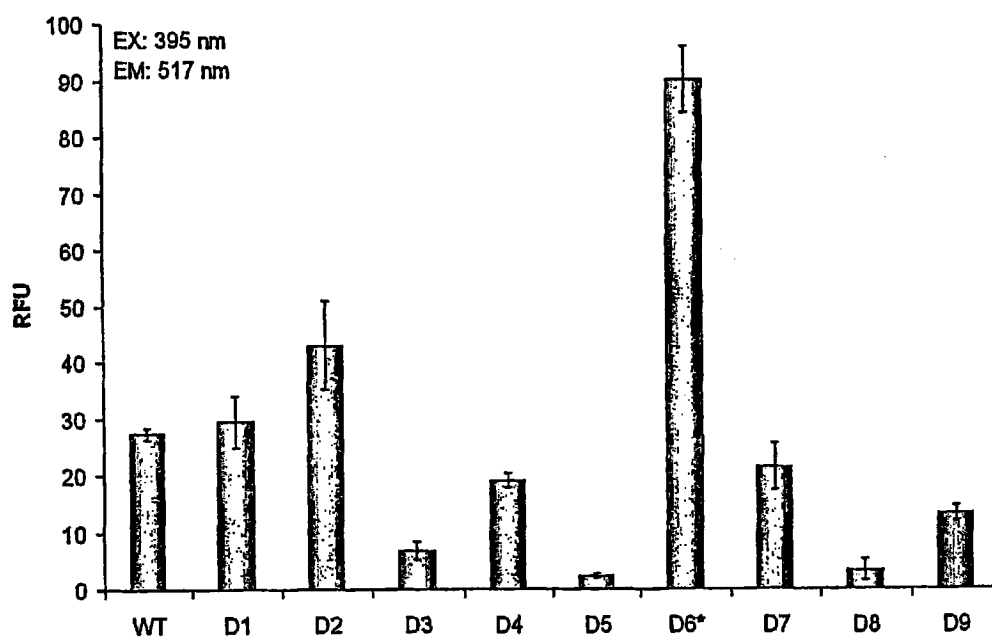


图 3

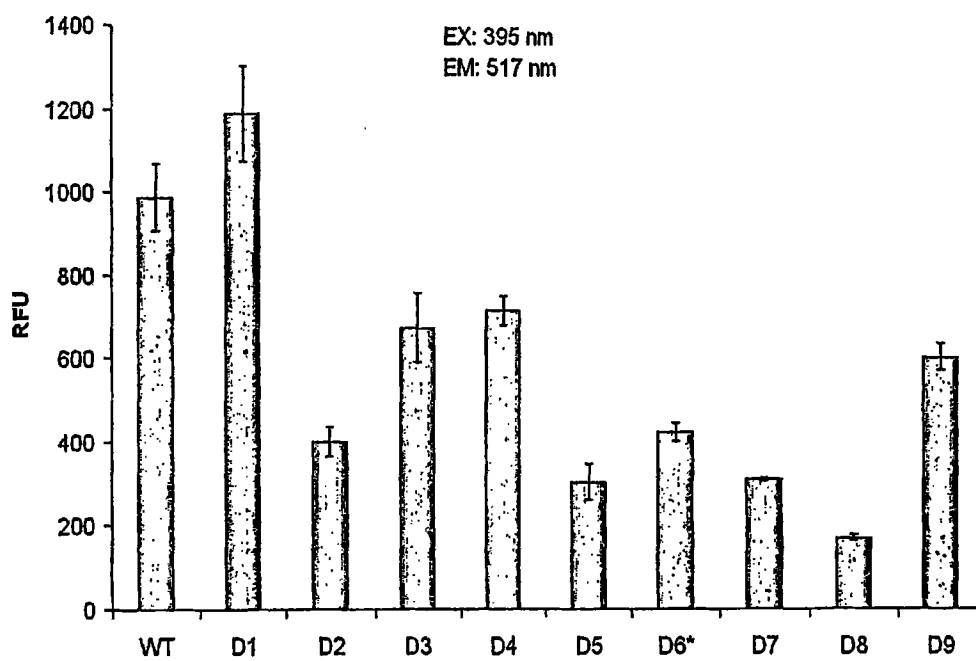


图 4

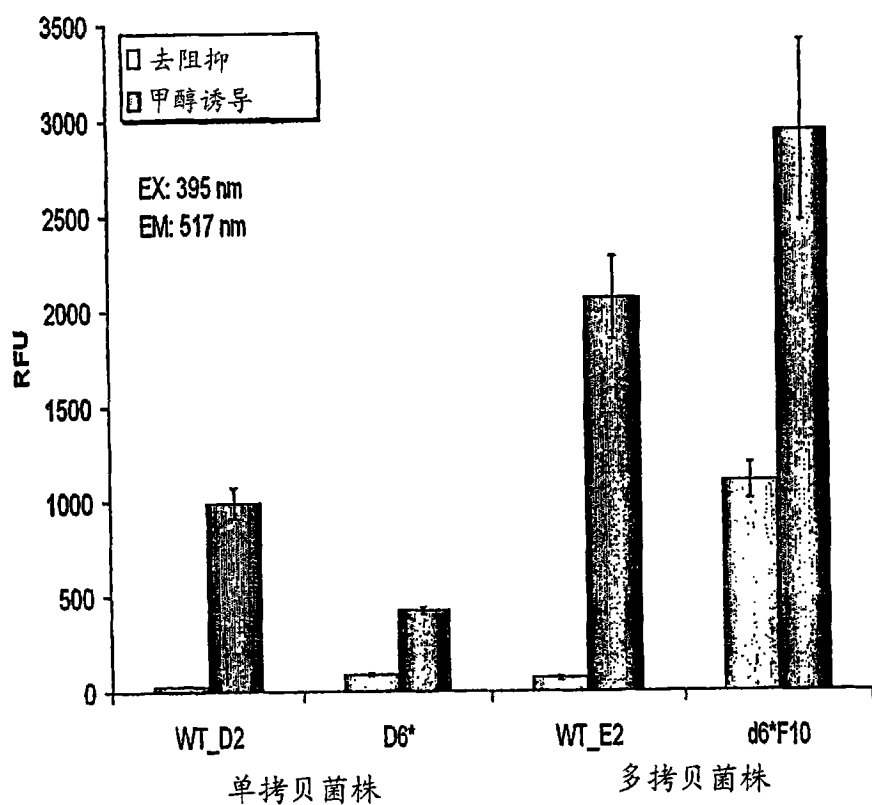


图 5

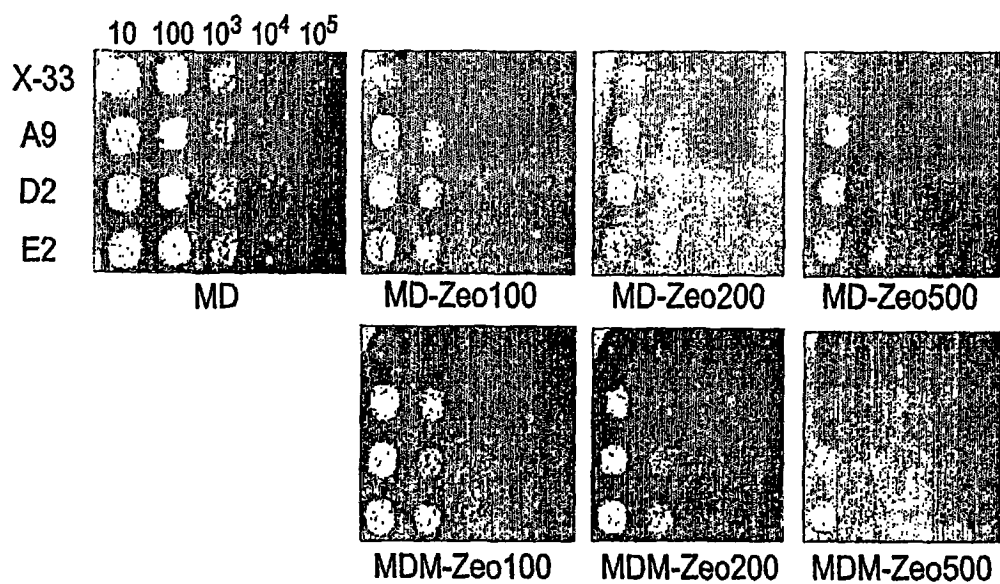


图 6

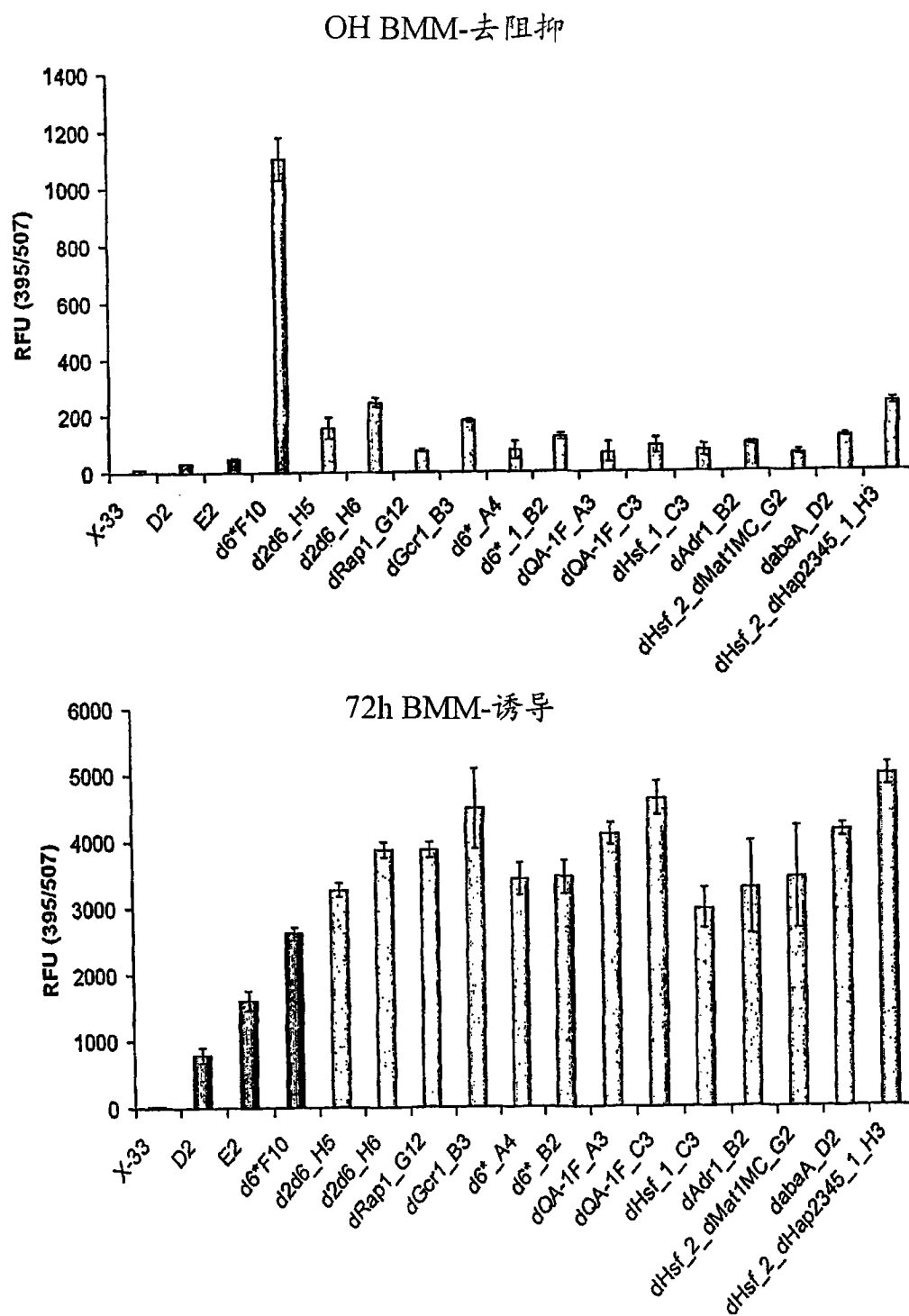


图 7

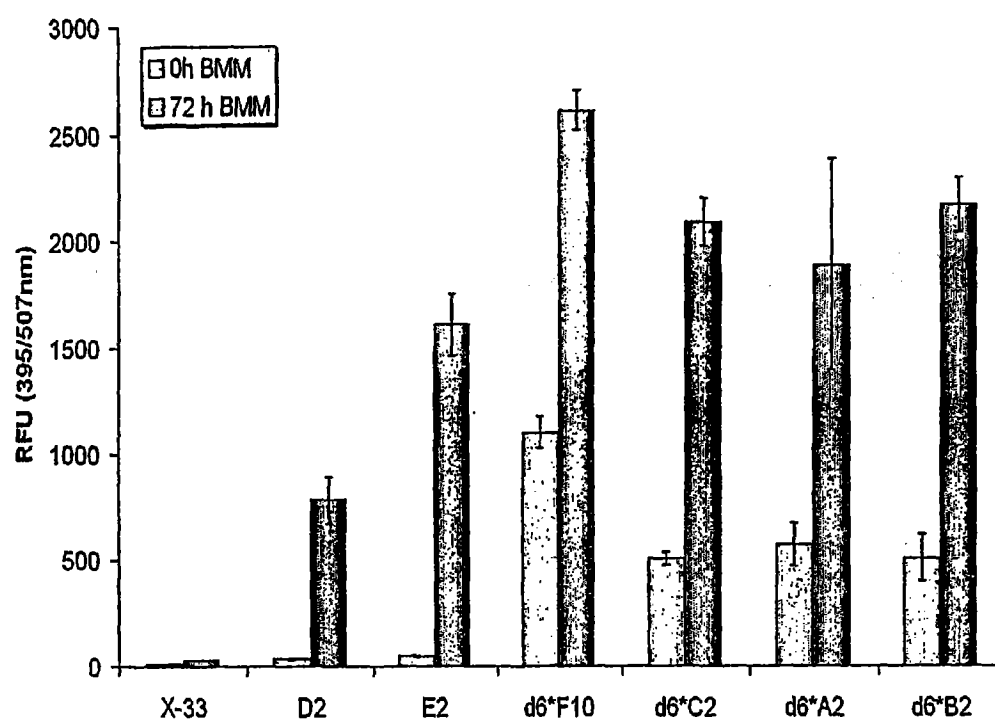


图 8