



(12) 发明专利

(10) 授权公告号 CN 113788884 B

(45) 授权公告日 2024.07.23

(21) 申请号 202110953383.4

(74) 专利代理机构 北京市柳沈律师事务所

11105

(22) 申请日 2012.10.05

专利代理人 张文辉

(65) 同一申请的已公布的文献号

申请公布号 CN 113788884 A

(51) Int.CI.

C07K 14/395 (2006.01)

(43) 申请公布日 2021.12.14

C07K 14/39 (2006.01)

(30) 优先权数据

C12N 15/81 (2006.01)

11184323.1 2011.10.07 EP

C12N 1/19 (2006.01)

12171006.5 2012.06.06 EP

C12N 15/67 (2006.01)

61/544,451 2011.10.07 US

C12R 1/645 (2006.01)

(62) 分案原申请数据

(56) 对比文件

201280060753.0 2012.10.05

CN 102471773 A, 2012.05.23

(73) 专利权人 隆扎有限公司

审查员 彭丹丹

地址 瑞士菲斯普

(72) 发明人 D.玛塔诺维克 B.加瑟 M.茂勒

权利要求书2页 说明书48页

R.普里尔霍佛 J.克雷恩 J.温格

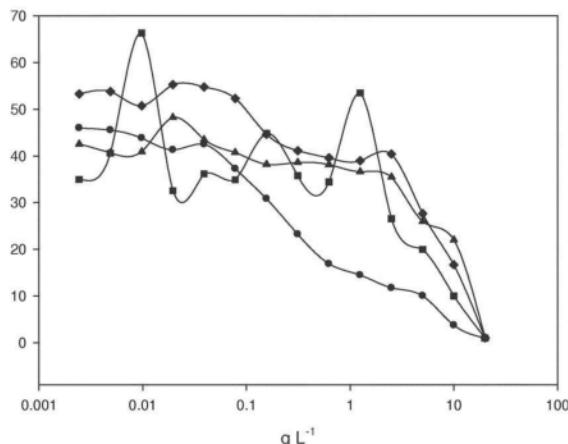
序列表16页 附图16页

(54) 发明名称

可调控启动子

(57) 摘要

一种通过培养包含表达构建体的重组真核细胞系来产生感兴趣蛋白质(POI)的方法，所述表达载体包含可调控的启动子和在所述启动子的转录调控下的编码POI的核酸分子，所述方法包括以下步骤：a)用阻遏所述启动子的基础碳源培养所述细胞系，b)用使所述启动子去阻遏的有限量的补充碳源培养所述细胞系，从而诱导POI以相比于天然pGAP启动子至少15%的转录速率产生，和c)产生并回收所述POI；以及另外地，分离的可调控的启动子和相应的表达系统。



1. 通过培养包含表达构建体的重组甲基营养型酵母来产生感兴趣蛋白质的方法,所述表达构建体包含可调控的启动子和在所述启动子的转录调控下的编码感兴趣蛋白质的核酸分子,所述方法包括以下步骤:

a) 用阻遏所述启动子的基础碳源培养所述重组甲基营养型酵母的细胞系,其中所述基础碳源是葡萄糖、甘油或其混合物,

b) 用使所述启动子去阻遏且完全诱导的无葡萄糖或有有限量葡萄糖的条件,培养所述细胞系,并

c) 产生并回收所述感兴趣蛋白质,

其中所述启动子由选自下组的核酸序列组成:

i) 由SEQ ID 1的核苷酸序列组成的pG1、由SEQ ID 2的核苷酸序列组成的pG3、由SEQ ID 4的核苷酸序列组成的pG4、由SEQ ID 3的核苷酸序列组成的pG6、由SEQ ID 5的核苷酸序列组成的pG7、或由SEQ ID 6的核苷酸序列组成的pG8;和

ii) 由SEQ ID 41的核苷酸序列组成的pG1a、由SEQ ID 42的核苷酸序列组成的pG1b、由SEQ ID 43的核苷酸序列组成的pG1c、由SEQ ID 44的核苷酸序列组成的pG1d、由SEQ ID 45的核苷酸序列组成的pG1e和由SEQ ID 46的核苷酸序列组成的pG1f。

2. 权利要求1的方法,其中所述启动子是功能活性的启动子,其为碳源可调控的启动子,所述启动子在重组真核细胞中,在完全诱导的状态下、并且是用化学性质确定的且无甲醇的给料培养基时,能以相比于所述细胞的天然pGAP启动子至少20%的转录速率表达感兴趣蛋白质。

3. 依照权利要求1的方法,其中所述步骤b)采用不提供葡萄糖或提供有限量的葡萄糖的补料培养基。

4. 依照权利要求1至2中任一项的方法,其中所述有限量为培养基中0-1g/L。

5. 依照权利要求1的方法,其中所述葡萄糖为生长限制性的,以保持特定生长速率(specific growth rate)在 0.02h^{-1} 至 0.2h^{-1} 的范围内。

6. 依照权利要求1的方法,其中所述葡萄糖为生长限制性的,以保持特定生长速率在 0.02h^{-1} 至 0.15h^{-1} 的范围内。

7. 权利要求1的方法,其中所述启动子是由SEQ ID 1的核苷酸序列组成的pG1。

8. 依照权利要求1的方法,其中所述启动子选自下组:由SEQ ID 1的核苷酸序列组成的pG1、由SEQ ID 41的核苷酸序列组成的pG1a、由SEQ ID 42的核苷酸序列组成的pG1b、由SEQ ID 43的核苷酸序列组成的pG1c、由SEQ ID 44的核苷酸序列组成的pG1d、由SEQ ID 45的核苷酸序列组成的pG1e和由SEQ ID 46的核苷酸序列组成的pG1f。

9. 依照权利要求1的方法,其中所述感兴趣蛋白质为异源蛋白质。

10. 依照权利要求9的方法,其中所述异源蛋白质选自治疗性蛋白质,包括抗体或其片段、酶和肽、蛋白质抗生素、毒素融合蛋白、碳水化合物-蛋白质偶联物、结构蛋白、调节蛋白、疫苗和疫苗样蛋白或颗粒、生长因子、激素和细胞因子。

11. 分离的核酸,其包含启动子,所述启动子由选自下组的核酸序列组成:

i) 由SEQ ID 1的核苷酸序列组成的pG1、由SEQ ID 2的核苷酸序列组成的pG3、由SEQ ID 3的核苷酸序列组成的pG6、由SEQ ID 5的核苷酸序列组成的pG7、或由SEQ ID 6的核苷酸序列组成的pG8;和

ii) 由SEQ ID 41的核苷酸序列组成的pG1a、由SEQ ID 42的核苷酸序列组成的pG1b、由SEQ ID 43的核苷酸序列组成的pG1c、由SEQ ID 44的核苷酸序列组成的pG1d、由SEQ ID 45的核苷酸序列组成的pG1e和由SEQ ID 46的核苷酸序列组成的pG1f。

12. 权利要求11的核酸，所述核酸包含功能活性的启动子，其为碳源可调控的启动子，所述启动子在重组真核细胞中，在完全诱导的状态下、并且是用化学性质确定的且无甲醇的给料培养基时，能以相比于所述细胞的天然pGAP启动子至少20%的转录速率表达感兴趣蛋白质。

13. 表达构建体，其包含依照权利要求12的核酸，其中所述启动子可操作地连接于在所述启动子转录控制下的编码感兴趣蛋白质的核苷酸序列，所述核酸与编码所述感兴趣蛋白质的核苷酸序列不天然关联。

14. 重组甲基营养型酵母，其包含权利要求13的表达构建体。

可调控启动子

[0001] 本申请是申请日为2012年10月05日的、发明名称“可调控启动子”、PCT申请号：PCT/EP2012/069757、专利申请号：201280060753.0的中国发明专利申请的分案申请。

技术领域

[0002] 本发明涉及可调控启动子和在真核细胞培养中在可调控启动子的控制下产生感兴趣蛋白质的方法。

背景技术

[0003] 已使用真核宿主实现重组蛋白质的成功产生。最突出的例子是酵母如酿酒酵母 (*Saccharomyces cerevisiae*)、巴斯德毕赤酵母 (*Pichia pastoris*) 或多形汉逊酵母 (*Hansenula polymorpha*)、丝状真菌如泡盛曲霉 (*Aspergillus awamori*) 或里氏木霉 (*Trichoderma reesei*)、或哺乳动物细胞如例如CHO细胞。尽管容易以高比率实现一些蛋白质的产生，但还有许多其他蛋白质仅在相当低的水平获得。

[0004] 基因在宿主生物体中的异源表达通常需要允许稳定转化宿主生物体的载体。载体会提供基因和临近编码序列5' 端的功能性启动子。由此，转录由该启动子序列调控和启动。迄今使用的大多数启动子源自编码通常以高浓度存在于细胞中的蛋白质的基因。

[0005] EP0103409A2披露了使用与糖酵解途径中特定酶的表达有关的酵母启动子，即牵涉丙酮酸激酶、磷酸丙糖异构酶、磷酸葡萄糖异构酶、磷酸甘油酸变位酶、己糖激酶1和2、葡糖激酶、磷酸果糖激酶、醛缩酶和糖酵解调控基因的表达的启动子。

[0006] WO 97/44470描述了来自解脂西洋蓍霉 (*Yarrowia lipolytica*) 的用于翻译延长因子1 (TEF1) 蛋白和用于核糖体蛋白S7的酵母启动子，其适于在酵母中异源表达蛋白质，而 EP1951877A1描述了将巴斯德毕赤酵母 TEF1 启动子用于产生异源蛋白质。

[0007] WO2005003310提供用于在酵母中表达感兴趣编码序列的方法，其使用来自产油酵母解脂西洋蓍霉的甘油醛-3-磷酸脱氢酶或磷酸甘油酸变位酶的启动子。

[0008] 源自牵涉巴斯德毕赤酵母的甲醇代谢途径的基因的启动子序列披露在US4808537 和 US4855231 (醇氧化酶AOX1、AOX2) 和 US6730499B1 (甲醛脱氢酶FLD1) 中。US20080153126A1 包括基于AOX1启动子的突变体启动子序列。

[0009] AOX1启动子仅应答甲醇而受到诱导并受其他碳源如葡萄糖或乙醇阻遏。甲醇的缺点在于其不适用于产生某些产物，因为其因毒性和易燃性而有潜在危险。因此，寻求AOX1启动子的备选物。

[0010] US2008299616A1引入苹果酸合酶 (MLS1) 基因的调控序列用于在巴斯德毕赤酵母中的异源基因表达，其在含有葡萄糖的培养基中受到阻遏，而在葡萄糖饥饿条件下或在存在乙酸盐时去阻遏。然而，不将该系统视为适用于有效的产生方法，因为MLS1启动子在去阻遏条件下较弱且活性较低。

[0011] Schöler 和 Schüller (Mol. Cell Biol. 1994 14(6):3613-22) 描述了异柠檬酸裂合酶基因ICL1的调控区，其在将细胞从发酵生长条件转移至非发酵生长条件后去阻遏。

[0012] WO2008063302A2描述了将源自巴斯德毕赤酵母的ADH1(醇脱氢酶)、EN01(烯醇化酶)和GUT1基因的新的可诱导启动子用于异源蛋白质的表达,CN1966688A中为巴斯德毕赤酵母omega 3-脂肪酸脱氢酶启动子序列,而W0002007117062A1中为巴斯德毕赤酵母来源的自动诱导(auto-inducible)的NPS启动子,其通过磷限制诱导。

[0013] WO2008128701A2描述了新的启动子的使用,其中源自巴斯德毕赤酵母的THI3(硫胺代谢)基因的启动子在含有硫胺的培养基中受到阻遏,并在硫胺耗尽时去阻遏。

[0014] US2009325241A1描述了一种采用木糖可诱导的启动子(FAS2启动子)在酵母细胞中产生乙醇的方法。

[0015] 期望提供改进的重组真核细胞系来产生能以高产率分离的发酵产物。因此,本发明的目标是提供适用于重组生产方法的备选的调控元件,其为简单且有效的。

发明内容

[0016] 所述目标由所要求保护的主题解决。

[0017] 依照本发明,提供一种通过培养包含表达构建体的重组真核细胞系来产生感兴趣蛋白质(POI)的方法,所述表达载体包含可调控的启动子和在所述启动子的转录调控下的编码POI的核酸分子,所述方法包括以下步骤:

[0018] a)用阻遏所述启动子的基础碳源培养所述细胞系,

[0019] b)用使所述启动子去阻遏的无或有限量的补充碳源培养所述细胞系,从而以相比于所述细胞的天然pGAP启动子至少15%的转录速率诱导所述POI的产生,和

[0020] c)产生并回收所述POI。

[0021] 所述培养步骤具体地包括在所述碳源的存在下,由此在包含所述碳源的培养基中,或在步骤b)中还在不存在补充碳源的情况下培养所述细胞系。

[0022] 对所述POI产生的诱导具体指对转录的诱导,具体包括所述POI的进一步的翻译和任选地表达。

[0023] 所述转录速率具体指在完全诱导所述启动子时获得的转录本的量。如果培养条件提供大概最大的诱导,例如以低于0.4g/L,优选低于0.04g/L,特定地低于0.02g/L的葡萄糖浓度,那么所述启动子视为去阻遏且完全诱导的。优选地,完全诱导的启动子显示相比于天然pGAP启动子至少15%,优选至少20%,更优选至少30%、40%、50%、60%、70%、80%、90%和至少100%或甚至更高为至少150%或至少200%的转录速率。所述转录速率可以例如通过报告基因如eGFP的转录本的量测定,如记载于下文实施例部分的,其显示在溶液中培养克隆时pG1启动子相比于天然pGAP启动子至少50%的相对高的转录速率。或者,所述转录速率可通过在微阵列上的转录强度测定,其中微阵列数据显示受阻遏和去阻遏状态之间表达水平的差异和在完全诱导状态中相比于天然pGAP启动子的高信号强度。这类微阵列数据具体地对于pG1显示超过200%的转录速率,对于pG3和pG4超过30%,对于pG6超过60%,对于pG7超过30%,对于pG8超过20%的转录速率(每个值均与天然pGAP相比)。发现现有技术中的启动子太弱而不适用于本发明的目的。

[0024] 具体地,所述天然pGAP启动子在所述重组真核细胞中以类似于在同一物种或菌株的天然真核细胞中的方式而具有活性(包括未经修饰(非重组)或重组真核细胞)。这类天然pGAP启动子通常理解为内源启动子,如此,是真核细胞同源性的,且充当标准或参照启动子

用于比较目的。

[0025] 例如,巴斯德毕赤酵母的天然pGAP启动子是巴斯德毕赤酵母中的未经修饰的内源性启动子序列,如用于控制中巴斯德毕赤酵母中GAPDH表达的,例如具有图13中显示的序列:巴斯德毕赤酵母(GS115)的天然pGAP启动子序列(SEQ ID 13)。如果将巴斯德毕赤酵母用作产生依照本发明的POI的宿主,那么将依照本发明的启动子的转录强度或速率与巴斯德毕赤酵母的这类天然pGAP启动子相比较。

[0026] 再举一个例子,酿酒酵母的天然pGAP启动子是酿酒酵母中未经修饰的内源性启动子序列,如用于控制酿酒酵母中GAPDH表达的。如果将酿酒酵母用作用于产生依照本发明的POI的宿主,那么将依照本发明的启动子的转录强度或速率与酿酒酵母的这类天然pGAP启动子相比较。

[0027] 因此,通常将依照本发明的启动子的相对转录强度或速率与用作产生POI的宿主的同一物种或菌株的细胞的天然pGAP启动子相比较。

[0028] 依照一个具体的实施方案,所述基础碳源不同于补充碳源,例如定量和/或定性上不同的。定量上的不同可以提供不同的条件来阻遏或去阻遏启动子活性。

[0029] 依照一个进一步的具体的实施方案,所述基础和补充碳源包含相同类型的分子或糖类,优选以不同的浓度。依照再一个具体的实施方案,所述碳源是两种或更多种不同碳源的混合物。

[0030] 可使用任何适用于真核细胞培养的有机碳。依照一个具体的实施方案,所述碳源是己糖如葡萄糖、果糖、半乳糖或甘露糖,二糖如蔗糖,醇如甘油或乙醇,或其混合物。

[0031] 依照一个具体的优选实施方案,所述基础碳源选自下组:葡萄糖、甘油、乙醇、或其混合物、和复合营养材料(complex nutrient material)。依照一个优选的实施方案,所述基础碳源是甘油。

[0032] 依照又一个具体的实施方案,所述补充碳源是己糖如葡萄糖、果糖、半乳糖和甘露糖,二糖如蔗糖,醇如甘油或乙醇,或其混合物。依照一个优选的实施方案,所述补充碳源是葡萄糖。

[0033] 具体地,所述方法可采用甘油作为基础碳源,葡萄糖作为补充碳源。

[0034] 可通过具体手段适宜地实现去阻遏条件。任选地,步骤b)采用提供无或有限量的补充碳源的给料培养基。

[0035] 具体地,所述给料培养基是化学性质确定的且无甲醇的。

[0036] 所述给料培养基可以液体形式或其他备选的形式如固体(例如片剂或其他持续释放手段)或气体(例如二氧化碳)添加到培养基中。然而,依照一个优选的实施方案,添加到细胞培养基的所述有限量的补充碳源可以甚至为0。优选地,补充碳源在培养基中的浓度为0-1g/L,优选低于0.6g/L,更优选低于0.3g/L,更优选低于0.1g/L,优选1-50mg/L,更优选1-10mg/L,具体地优选1mg/L或更低,如低于如用适宜标准测定法测量的检测限,例如在由生长细胞培养物消耗时测定为培养基中的残余浓度。

[0037] 在一个优选的方法中,所述有限量的补充碳源提供细胞培养物中低于检测限的残余量,如在生产阶段结束时或在发酵过程的输出中,优选在收获发酵产物时在发酵液中测定的。

[0038] 优选地,所述有限量的补充碳源是生长限制性的以保持范围在 0.02h^{-1} 至 0.2h^{-1} ,

优选 0.02h^{-1} 至 0.15h^{-1} 内的比生长速率。

[0039] 依照本发明的一个具体方面,所述启动子是巴斯德毕赤酵母启动子或其功能活性变体。

[0040] 在本文中,依照本发明的启动子应总是指本文中描述的序列及其功能活性变体。如下文详细解释的,这类变体包括源自巴斯德毕赤酵母以外的物种的同源物和类似物。

[0041] 依照本发明的方法可采用为巴斯德毕赤酵母的野生型启动子或其功能活性变体的启动子,例如能控制野生型或重组真核细胞中特定基因转录的,例如选定基因的野生型启动子,所述基因选自下组:G1 (SEQ ID 7),如编码(高亲和力)葡萄糖运输器的,G3 (SEQ ID 8)、G4 (SEQ ID 9),如编码线粒体醛脱氢酶的,G6 (SEQ ID 10)、G7 (SEQ ID 11),如编码主要易化物糖运输器家族(major facilitator sugar transporter family)的成员的,或G8 (SEQ ID 12),如编码Gti1_Pac2超家族成员的,或其功能活性变体。

[0042] 依照本发明,具体地提供了启动子或其功能活性变体,其将与野生型酵母细胞中的这类基因之一天然关联。

[0043] 依照一个具体的实施方案,所述细胞系选自下组:哺乳动物、昆虫、酵母、丝状真菌和植物细胞系,优选为酵母。

[0044] 具体地,酵母选自下组:毕赤酵母属(Pichia)、假丝酵母属(Candida)、球拟酵母属(Torulopsis)、Arxula、汉逊酵母属(Hansenula)、西洋蓍霉属(Yarrowia)、克鲁维酵母属(Kluyveromyces)、酵母属(Saccharomyces)、Komagataella,优选甲基营养型酵母。

[0045] 一种具体优选的酵母是巴斯德毕赤酵母、Komagataella pastoris、K.phaffii、或K.pseudopastoris。

[0046] 依照又一个具体的实施方案,所述启动子不与编码POI的核苷酸序列天然关联。

[0047] 具体地,POI是真核细胞蛋白质,优选为哺乳动物蛋白质。

[0048] 依照本发明产生的POI可以是多聚体蛋白,优选为二聚体或四聚体。

[0049] 依照本发明的一个方面,所述POI是重组或异源蛋白质,优选选自治疗性蛋白质包括抗体或其片段、酶和肽、蛋白质抗生素、毒素融合蛋白、糖-蛋白质偶联物、结构蛋白、调节蛋白、疫苗和疫苗样蛋白或颗粒、加工酶、生长因子、激素和细胞因子、或POI的代谢物。

[0050] 一种具体的POI是抗原结合分子如抗体或其片段。特异性的POI有抗体如单克隆抗体(mAb)、免疫球蛋白(Ig)或免疫球蛋白G类(IgG)、重链抗体(HcAb)、或其片段如片段-抗原结合(Fab)、Fd、单链可变片段(scFv)、或其工程化变体如例如Fv二聚体(双抗体)、Fv三聚体(三抗体)、Fv四聚体或小型抗体(minibody)和单域抗体如VH或VHH或V-NAR。

[0051] 依照一个具体的实施方案,发酵产物是使用POI、其代谢物或衍生物制备的。

[0052] 依照本发明的另一个方面,提供一种用于调控重组真核细胞中在碳源可调控启动子的转录控制下的POI表达的方法,所述启动子相比于所述细胞的天然pGAP启动子具有至少15%的转录强度,其中所述表达在限制碳源的条件下诱导。碳源可调控启动子优选相比于参照pGAP启动子具有至少20%的转录强度,具体地相比于天然pGAP启动子有如上文就转录速率而言描述的转录强度。因此,完全诱导的启动子相比于细胞的天然pGAP启动子具有优选至少15%,优选至少20%,更优选至少30%、40%、50%、60%、70%、80%、90%和至少100%或至少150%或至少200%的甚至更高的转录强度,如在选择用于产生POI的真核细胞中测定的。

[0053] 在一个优选的实施方案中,使用在去阻遏状态具有转录活性或转录强度的这类启动子,其在去阻遏状态中相比于阻遏状态为至少2倍,更优选至少5倍,甚至更优选至少10倍,更优选至少20倍,更优选至少30、40、50、或100倍。

[0054] 依照本发明的另一个方面,提供一种在重组真核细胞中在碳源可调控启动子的转录控制下产生POI的方法,其中所述启动子相比于所述细胞的天然pGAP启动子具有如上文所述的,即至少15%的转录强度。所述碳源可调控启动子相比于参照pGAP启动子优选具有至少20%,更优选至少30%、40%、50%、60%、70%、80%、90%和至少100%或至少150%或至少200%的甚至更高的转录强度。在一个优选的实施方案中,使用在去阻遏状态具有转录活性的这类启动子,其在去阻遏状态相比于阻遏状态为至少2倍,更优选至少5倍,甚至更优选至少10倍,更优选至少20倍,更优选至少30、40、50、或100倍。适宜地,将依照本发明的具体启动子用在这类方法中。

[0055] 在依照本发明的一个具体的优选的方法中,所述启动子是包含选自下组的核酸序列的可调控启动子:

[0056] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0057] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0058] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6);和

[0059] d) 源自a)、b)或c)的片段或变体,

[0060] 其中所述启动子是功能活性的启动子,其为碳源可调控的启动子,所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达POI。

[0061] 具体地,pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或pG8 (SEQ ID 6) 的变体是选自下组的功能活性变体:具有至少约60%核苷酸序列同一性的同源物,通过在序列内或序列远端之一或两端处一个或多个核苷酸的插入、缺失或取代对亲本核苷酸序列进行修饰可获得的同源物(优选具有至少200bp的核苷酸序列,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp,或至少1000bp),以及源自巴斯德毕赤酵母以外的物种的类似物。

[0062] 依照本发明的启动子的一些优选的功能活性变体是pG1、pG3、pG4、pG6、pG7或pG8启动子核苷酸序列中任一种的片段,优选包含启动子核苷酸序列3' 端的片段,例如源自所述启动子核苷酸序列之一的核苷酸序列,其具有具体的长度和在5' 末端区域的缺失,例如在5' 端核苷酸序列的切割,从而获得范围从3' 端到变化的5' 端的具体长度,如核苷酸序列长度至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp、或至少1000bp。

[0063] 已证实包含或组成为这类片段的例示性变体为功能活性的,所述片段例如具有范围为200至1000bp的具体长度,优选范围为250至1000bp,更优选范围为300至1000bp,例如包含3' 末端序列的片段。例如,pG1的功能活性变体选自下组:pG1a (SEQ ID 41) ,pG1b (SEQ ID 42) ,pG1c (SEQ ID 43) ,pG1d (SEQ ID 44) ,pG1e (SEQ ID 45) 和pG1f (SEQ ID 46) ,如此,一种核苷酸序列范围为300-1000bp,其包含3' 末端序列直至核苷酸1001。

- [0064] 依照本发明的另一个方面,提供分离的核酸,其包含选自下组的核酸序列:
- [0065] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或 pG8 (SEQ ID 6),
- [0066] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或 pG8 (SEQ ID 6),
- [0067] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或 pG8 (SEQ ID 6),和
- [0068] d) 源自a)、b)或c)的片段或变体,
- [0069] 其中所述核酸包含功能活性的启动子,其为碳源可调控的启动子,所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达POI。
- [0070] 具体地,pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或 pG8 (SEQ ID 6) 的变体是选自下组的功能活性变体:具有至少约60%核苷酸序列同一性的同源物,通过在序列内或序列远端之一或两端处一个或多个核苷酸的插入、缺失或取代对亲本核苷酸序列进行修饰可获得的同源物(优选具有至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp,或至少1000bp的核苷酸序列),以及源自巴斯德毕赤酵母以外的物种的类似物。
- [0071] 依照本发明启动子的一些优选的功能活性变体是pG1、pG3、pG6、pG7或pG8启动子核苷酸序列中任一种的片段,优选包含启动子核苷酸序列3' 端的片段,例如源自所述启动子核苷酸序列之一的核苷酸序列,其具有具体的长度和在5' 末端区域的缺失,例如在5' 端核苷酸序列的切割,从而获得范围从3' 端到变化的5' 端的具体长度,如核苷酸序列长度至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp,或至少1000bp。
- [0072] 已证实包含或组成为这类片段的例示性变体为功能活性的,所述片段例如具有范围为200至1000bp的具体长度,优选范围为250至1000bp,更优选范围为300至1000bp,例如包含3' 末端序列的片段。例如,pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和 pG1f (SEQ ID 46),如此,一种核苷酸序列范围为300-1000bp,其包含3' 末端序列直至核苷酸1001。
- [0073] 所述碳源可调控启动子优选具有如上文描述的转录强度,相比于参照pGAP启动子优选具有至少20%,更优选相比于天然pGAP启动子至少30%、40%、50%、60%、70%、80%、90%和至少100%或甚至更高的至少150%或至少200%的转录强度。在一个优选的实施方案中,使用在去阻遏状态具有转录活性的这类启动子,其在去阻遏状态相比于阻遏状态为至少2倍,更优选至少5倍,甚至更优选至少10倍,更优选至少20倍,更优选至少30、40、50、或100倍。适宜地,将依照本发明的具体启动子用在这类方法中。
- [0074] 依照本发明的再一个方面,提供一种表达构建体,其包含依照本发明的启动子可操作地连接于在所述启动子转录控制下的编码POI的核苷酸序列,所述启动子与POI的编码序列不天然关联。
- [0075] 本发明的另一个方面指包含依照本发明构建体的载体。
- [0076] 本发明的再一个方面指包含依照本发明的构建体或载体的重组真核细胞。
- [0077] 具体地,所述细胞选自下组:哺乳动物、昆虫、酵母、丝状真菌和植物细胞系,优选

为酵母。

[0078] 所述酵母可适宜地选自下组:毕赤酵母属(*Pichia*)、假丝酵母属(*Candida*)、球拟酵母属(*Torulopsis*)、*Arxula*、汉逊酵母属(*Hansenula*)、西洋蓍霉属(*Yarrowia*)、克鲁维酵母属(*Kluyveromyces*)、酵母属(*Saccharomyces*)、*Komagataella*,优选甲基营养型酵母。

[0079] 优选地,所述酵母是巴斯德毕赤酵母、*Komagataella pastoris*、*K. phaffii*、或*K. pseudopastoris*。

[0080] 依照一个具体的实施方案,采用一种细胞,其在存在过剩碳源的情况下相对于有限碳源的条件具有更高的比生长速率。

[0081] 本发明的又一个方面指将本发明的重组真核细胞用于产生POI。

[0082] 依照本发明的一个别的方面,提供一种方法来从真核细胞筛选或鉴定碳源可调控启动子,包括以下步骤:

[0083] a) 在存在碳源时,在细胞生长条件下以分批培养来培养真核细胞,

[0084] b) 进一步在存在有限量的补充碳源时,以补料分批培养来培养所述细胞,

[0085] c) 提供步骤a)和b)的细胞培养物的样品,并

[0086] d) 在所述样品中实施转录分析以鉴定出在步骤b)的细胞中比在步骤a)的细胞中显示更高转录强度的可调控启动子。

[0087] 所述更高的转录强度可通过完全诱导状态中的转录强度测定,其为例如在葡萄糖有限的恒化器(chemostat)培养条件下获得的,其在去阻遏状态相比于阻遏状态为至少2倍,更优选至少5倍,甚至更优选至少10倍,更优选至少20倍,更优选至少30、40、50、或100倍。

[0088] 优选地,所述转录分析是定量或半定量的,优选采用DNA微阵列、RNA测序和转录组分析。

附图说明

[0089] 图1:巴斯德毕赤酵母的启动子序列pG1 (SEQ ID 1)。

[0090] 图2:巴斯德毕赤酵母的启动子序列pG3 (SEQ ID 2)。

[0091] 图3:巴斯德毕赤酵母的启动子序列pG4 (SEQ ID 4)。

[0092] 图4:巴斯德毕赤酵母的启动子序列pG6 (SEQ ID 3)。

[0093] 图5:巴斯德毕赤酵母的启动子序列pG7 (SEQ ID 5)。

[0094] 图6:巴斯德毕赤酵母的启动子序列pG8 (SEQ ID 6)。

[0095] 图7:巴斯德毕赤酵母GS115基因组G1基因的编码序列 (SEQ ID 7)。

[0096] 图8:巴斯德毕赤酵母GS115基因组G3基因的编码序列 (SEQ ID 8)。

[0097] 图9:巴斯德毕赤酵母GS115基因组G4基因的编码序列 (SEQ ID 9)。

[0098] 图10:巴斯德毕赤酵母GS115基因组G6基因的编码序列 (SEQ ID 10)。

[0099] 图11:巴斯德毕赤酵母GS115基因组G7基因的编码序列 (SEQ ID 11)。

[0100] 图12:巴斯德毕赤酵母GS115基因组G8基因的编码序列 (SEQ ID 12)。

[0101] 图13:巴斯德毕赤酵母(GS115)的天然pGAP启动子序列 (SEQ ID 13)

#	名称	PAS*	PIPA*	GS115描述
pGAP	TDH3	PAS_chr2-1_0437	PIPA02510	甘油醛-3-磷酸脱氢酶

[0103] *PAS:巴斯德毕赤酵母GS115中的ORF名称;PIPA:巴斯德毕赤酵母菌株型DSMZ70382中的ORF名称

[0104] 图14:pG1(圆形)、pG3(三角形)、pG4(菱形)和pG6(正方形)启动子的去阻遏特性:对于pG1,在约0.04g葡萄糖/L或更低达到最大转录活性,而所有其他pG启动子在约4g/L或更低已达到最大转录活性。为了将不同启动子的相对诱导行为进行比较,通过将每个值除去相应启动子构建体的D20值将数据标准化。因此,数据是相对荧光值,而D20处数据点为1.0。

[0105] 图15:启动子序列pG1的功能活性变体;巴斯德毕赤酵母的pG1a-f(SEQ ID 41-46)。

具体实施方式

[0106] 如贯穿本说明书使用的具体术语具有以下含义。

[0107] 如本文中使用的,术语“碳源”意指适用作微生物能量来源的可发酵的碳底物,通常为糖源,如那些能被宿主生物或生产细胞系代谢的,具体为选自下组的来源:单糖、寡糖、多糖、醇包括甘油,以纯化形式或以原材料如复合营养材料提供。依照本发明,碳源可以单一碳源或不同碳源的混合物使用。

[0108] 如依照本发明使用的,“基础碳源”通常是适用于细胞生长的碳源,如真核细胞的营养物。所述基础碳源可在培养基,如基底培养基或复合培养基中提供,但还可在含有纯化碳源的化学性质确定的培养基中提供。基础碳源通常以提供细胞生长的量提供,特别是在培养过程中的生长阶段期间,例如以获得至少5g/L细胞干质量,优选至少10g/L细胞干质量或至少15g/L细胞干质量的细胞密度,例如在标准的传代培养步骤期间展现超过90%,优选超过95%的存活。

[0109] 依照本发明,基础碳源通常以过量或过剩量使用,这被理解为过量提供能源以增加生物质,例如在补料分批培养过程中细胞系的生长阶段期间。该过剩量具体是超过有限量的补充碳源以实现发酵液中的残余浓度,其为可测量的且通常比在供给有限量的补充碳源期间高至少10倍,优选至少50倍或至少100倍。

[0110] 术语“化学性质确定的”就细胞培养基如补料分批过程中的给料培养基而言,意指适用于生产细胞系的体外细胞培养的生长培养基,其中所有化学组分和肽均为已知的。通常而言,化学性质确定的培养基完全没有动物来源的组分且代表纯的且恒定的细胞培养环境。

[0111] 如依照本发明使用的“补充碳源”通常是促进生产细胞系的发酵产物生产的补充性底物,特别是在培养过程的生产阶段。所述生产阶段具体地在生长阶段之后,例如在分批、补料-分批和连续培养过程中。所述补充性碳源具体地可包含在补料分批过程的给料中。

[0112] “有限量”的碳源或“有限碳源”在本文中理解为保持生产细胞系在生产阶段或生产模式中所需要的碳源的量。可在补料-分批过程中采用这类有限量,其中碳源包含在给料培养基中并以低给料速率供应给培养物用于持续的能量投递以产生POI,而将生物质保持在低生长速率。给料培养基通常在细胞培养的生产阶段添加到发酵液。

[0113] 有限量的补充碳源可以例如通过细胞培养液中补充碳源的残余量测定,其在预确

定的阈值之下或甚至低于如在标准(糖)测定法中测量的检测限。残余量通常会在收获发酵产物时在发酵液中测定。

[0114] 有限量的补充碳源还可以通过限定补充碳源向发酵罐的平均给料速率测定,例如如通过在整个培养过程中,例如补料分批阶段按培养时间添加的量测定,来测定按时间计算的平均量。该平均给料速率保持较低以确保细胞培养物对补充碳源的完全使用,例如为 $0.6\text{ g L}^{-1}\text{h}^{-1}$ (g碳源每L初始发酵体积和h时间)至 $25\text{ g L}^{-1}\text{h}^{-1}$,优选为 $1.6\text{ g L}^{-1}\text{h}^{-1}$ 至 $20\text{ g L}^{-1}\text{h}^{-1}$ 。

[0115] 有限量的补充碳源还可以通过测量在生产过程之前和期间的比生长速率来测定,该比生长速率在生产阶段期间保持较低,例如在预确定的范围内,如在 0.02 h^{-1} 至 0.20 h^{-1} 的范围内,优选为 0.02 h^{-1} 至 0.15 h^{-1} 。

[0116] 如本文中使用的,术语“细胞系”指特定细胞类型的已有克隆,其已获得在延长的时段内增殖的能力。术语“宿主细胞系”指如用于表达内源或重组基因或代谢途径的产物以产生多肽或由这类多肽介导的细胞代谢物的细胞系。“生产宿主细胞系”或“生产细胞系”通常理解为即时可用于在生物反应器中培养以获得生产过程产物如POI的细胞系。术语“真核宿主”或“真核细胞系”意指任何真核细胞或生物,其可被培养以产生POI或宿主细胞代谢物。完全理解该术语不包括人。

[0117] 术语“表达”或“表达系统”或“表达盒”指含有处于可操作连接中的期望的编码序列和调控序列的核酸分子,从而用这些序列转化或转染的宿主能产生编码的蛋白质或宿主细胞代谢物。为了实现转化,表达系统可纳入载体中;然而,相关DNA还可以整合到宿主染色体中。表达可指分泌的或不分泌的表达产物,包括多肽或代谢物。

[0118] “表达构建体”或“载体”用于本文定义为在适宜的宿主生物体中克隆的重组核苷酸序列即重组基因的转录及其mRNA翻译所需要的DNA序列。表达载体通常包含用于在宿主细胞中自主复制的起点、可选择标志(例如氨基酸合成基因或赋予针对抗生素如新霉素、卡那霉素、G418或潮霉素的抗性的基因)、许多限制性酶切割位点、适宜的启动子序列和转录终止子,其组分可操作地连接在一起。如本文中使用的,术语“质粒”和“载体”包括自主复制的核苷酸序列以及基因组整合的核苷酸序列。

[0119] 如本文中使用的,术语“变体”在本发明的背景中指具有特定同源性或相似性的任何序列。变体启动子可以例如源自启动子序列pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或 pG8 (SEQ ID 6),其通过诱变以产生适用作重组细胞系中的启动子的序列。这类变体启动子可从突变体序列的库获得,其通过选出具有预确定特性的那些库成员。变体启动子可具有相同或甚至改进的特性,例如在诱导POI产生中有所改进,具有在阻遏和去阻遏条件下增加的差异效果。所述变体启动子还可以源自类似的序列,例如源自巴斯德毕赤酵母以外的真核生物物种或毕赤酵母属以外的属,如来自乳酸克鲁维酵母(*K. lactis*)、*Z. rouxii*、*P. stipitis*、多形汉逊酵母。具体地,可如此使用或作为亲本序列使用与类似于相应巴斯德毕赤酵母基因的基因天然关联的类似的启动子序列来产生其功能活性变体。具体地,

[0120] -与pG1类似的启动子特征在于其与类似于G1(高亲和力葡萄糖运输器;巴斯德毕赤酵母GS115描述:推定的运输器,糖搬运家族的成员;编码序列SEQ ID 7)的基因天然关联;

[0121] -与pG3类似的启动子特征在于其与类似于G3(编码序列SEQ ID 8)的基因天然关联;

[0122] -与pG4类似的启动子特征在于其与类似于G4(巴斯德毕赤酵母GS115:预测的线粒体醛脱氢酶;编码序列SEQ ID 9)的基因天然关联;

[0123] -与pG6类似的启动子特征在于其与类似于G6(编码序列SEQ ID 10)的基因天然关联;

[0124] -与pG7类似的启动子特征在于其与类似于G7(巴斯德毕赤酵母GS115:主要易化物糖运输器家族的成员;编码序列SEQ ID 11)的基因天然关联;

[0125] -与pG8类似的启动子特征在于其与类似于G8(巴斯德毕赤酵母GS115:Gti1_Pac2超家族的成员;编码序列SEQ ID 12)的基因天然关联。

[0126] 这类类似启动子序列或其功能活性变体的特性可使用标准技术测定。

[0127] 如本文中使用的核苷酸或启动子序列的“功能活性”变体意指通过在序列内或序列远端之一或两端处一个或多个核苷酸的插入、缺失或取代对亲本序列进行修饰所得到的序列,且所述修饰不影响(特别是不损害)该序列的活性。

[0128] 具体地,依照本发明的启动子序列的功能活性变体选自下组:

[0129] -具有至少60%核苷酸序列同一性的同源物,

[0130] -通过在序列内或序列远端之一或两端处一个或多个核苷酸的插入、缺失或取代对亲本核苷酸序列进行修饰可获得的同源物,优选具有(即包含或组成)至少200bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp或至少1000bp的核苷酸序列,和

[0131] -源自巴斯德毕赤酵母以外的物种的类似物。

[0132] 特别优选的功能活性变体是那些通过修饰源自依照本发明的启动子和/或启动子序列的片段,具有(即包含或组成)至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp或至少1000bp的核苷酸序列。

[0133] 依照本发明的启动子的一些优选的功能活性变体是pG1、pG3、pG4、pG6、pG7或pG8启动子核苷酸序列中任一种的片段,优选包含启动子核苷酸序列3'端的片段,例如源自所述启动子核苷酸序列之一的核苷酸序列,其具有特定的长度和在5'末端区域的缺失,例如在5'端核苷酸序列的切割,从而获得范围从3'端到变化的5'端的特定长度,如核苷酸序列长度为至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp或至少1000bp。

[0134] 已证实包含或组成为这类片段的例示性变体为功能活性的,所述片段例如具有范围为200至1000bp的特定长度,优选范围为250至1000bp,更优选范围为300至1000bp,例如包含3'末端序列的片段。例如,pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和pG1f (SEQ ID 46),如此,一种核苷酸序列范围为300-1000bp,其包含3'末端序列直至核苷酸1001。

[0135] 如本文中使用的,术语“可调控”就启动子而言指在真核细胞中存在过量碳源(营养底物)的情况下在分批培养的生长阶段受阻遏,而在生产细胞系的生产阶段,例如在减少碳量时,如在依照补料分批策略将生长限制碳源(营养底物)供给到培养物时去阻遏以

施加强启动子活性的启动子。在这一方面,术语“可调控”理解为“碳源限制可调控的”或“葡萄糖限制可调控的”,指通过碳消耗、减少、短缺或耗尽,或通过碳源的有限添加从而使其易于由细胞消耗对启动子的去阻遏。

[0136] 依照本发明的功能活性启动子是相对强的可调控启动子,其在细胞生长条件下(生长阶段)受到沉默或阻遏,而在生产条件下(生产阶段)活化或去阻遏,因此适用于通过限制碳源来诱导生产细胞系中的POI产生。因此,启动子的功能活性变体具有至少这类可调控特性。

[0137] 依照本发明的可调控启动子的强度指其转录强度,由以高或低频率在该启动子处发生的转录的启动效率代表。转录强度越高,转录在该启动子处的发生将更频繁。启动子强度是重要的,因为它确定了给定的mRNA序列的转录有多频繁,对一些基因相对于其他基因有效给出转录的更高的优先级,导致转录本的更高浓度。例如,编码大量需要的蛋白质的基因通常具有相对强的启动子。RNA聚合酶一次仅能实施一项转录任务,因此必须使其工作区分优先级以有效率。选择启动子强度中的差异以允许该优先性。依照本发明,可调控启动子在完全诱导状态相对较强,所述完全诱导状态通常理解为约最大活性的状态。相对强度通常针对标准启动子,如用作宿主细胞的细胞的相应pGAP启动子来测定。转录频率通常理解为转录速率,例如如通过在适宜的测定法例如RT-PCR或Northern印迹中转录本的量测定的。例如,在为巴斯德毕赤酵母的宿主细胞中测定依照本发明的启动子的转录强度并与巴斯德毕赤酵母的天然pGAP启动子相比较。

[0138] pGAP启动子启动编码甘油醛-3-磷酸脱氢酶(GAPDH)的gap基因的表达,其为存在于任何能在葡萄糖上生长的微生物中的构成性启动子。GAPDH(EC 1\2\1\12),作为糖酵解中的关键酶,在分解代谢和合成代谢性糖代谢中起着至关重要的作用。

[0139] 依照本发明的可调控启动子施加相对高的转录强度,其由相比于宿主细胞中的天然pGAP启动子(有时称为“同源pGAP启动子”)至少15%的转录速率或转录强度反映。优选地,相比于天然pGAP启动子,转录速率或强度为至少20%,在特别优选的情况下至少30%、至少40%、至少50%、至少60%、至少70%、至少80%、至少90%和至少100%或甚至更高,如至少150%或至少200%,例如在选定为宿主细胞用于产生POI的真核细胞中测定的。

[0140] 特别优选的是一种可调控启动子,其在诱导状态中至少具有pG1、pG3、PG4、pG6、pG7或pG8启动子之一的转录强度。采用pGAP启动子作为参照的比较性转录强度可通过标准手段测定,如通过测量转录本的量,例如采用微阵列,或者在细胞培养中,如通过测量重组细胞中相应基因表达产品的量。例示性的测试在实施例部分例示。

[0141] 具体地,依照本发明的启动子是碳源可调控的启动子,其具有差异性的启动子强度,如在存在葡萄糖和葡萄糖限制的情况下比较其强度的测试中所测定的,显示在相对高的葡萄糖浓度,优选至少10g/L,优选至少20g/L的浓度其仍受阻遏。具体地,依照本发明的启动子在有限葡萄糖浓度和完全诱导启动子的葡萄糖阈值浓度受到完全诱导,所述阈值低于20g/L,优选低于10g/L,低于1g/L,甚至低于0.1g/L或低于50mg/L,优选在低于40mg/L的葡萄糖浓度具有例如天然、同源pGAP启动子的至少50%的完全转录强度。

[0142] 优选地,通过在切换到低于预确定的碳源阈值的诱导条件时启动POI生产,并与阻遏状态中的强度比较来测定差异启动子强度。转录强度通常理解为在完全诱导状态中的强度,即在去阻遏条件下显示约最大活性。差异启动子强度为,例如依照相比于阻遏条件在去

阻遏条件下重组宿主细胞系中POI产生的效率或产率测定的,或通过转录本的量测定。依照本发明的可调控启动子具有优选的差异启动子强度,其在去阻遏状态中相比于阻遏状态为至少2倍,更优选至少5倍,甚至更优选至少10倍,更优选至少20倍,更优选至少30、40、50或100倍,还理解为倍数诱导。这类差异启动子强度可通过如由所附实施例例示的测试测定。

[0143] 现有技术中的启动子(MLS1启动子或ICL1启动子)表现为具有显著低于2倍诱导的差异启动子强度。这类现有技术启动子也不可用于工业POI生产,其相比于pGAP启动子标准具有约5%的启动子强度。这已在与依照本发明的启动子的直接比较中证实。

[0144] 术语“同源性”指两种或更多种核苷酸序列在相应位置处具有相同或保守的碱基对,其达到特定程度,高达接近100%的程度。同源序列通常具有至少约50%的核苷酸序列同一性,优选至少约60%的同一性,更优选至少约70%的同一性,更优选至少约80%的同一性,更优选至少约90%的同一性,更优选至少约95%的同一性。

[0145] 依照本发明的同源启动子序列优选与巴斯德毕赤酵母的pG1、pG3、pG4、pG6、pG7或pG8启动子核苷酸序列中的任一种在核苷酸序列的至少特定部分,如包含相应启动子核苷酸序列的3'区的部分,优选具有长达相应启动子核苷酸序列3'端的特定长度的部分,如具有至少200bp,优选至少250bp,优选至少300bp,更优选至少400bp,至少500bp,至少600bp,至少700bp,至少800bp,至少900bp或至少1000bp的长度的部分,和源自巴斯德毕赤酵母以外的物种的类似物具有特定同源性。具体地,至少那些部分优选在300-1000bp的范围内(包含相应启动子核苷酸序列3'末端序列)为同源的。

[0146] 类似序列通常源自其他物种或菌株。明确理解本发明的任一种源自巴斯德毕赤酵母以外物种的类似的启动子序列可以包含同源序列,即如本文中描述的具有特定同源性的序列。如此,术语“同源”还可包含类似的序列。在另一方面,理解本发明还指类似序列及其包含特定同源性的同源物。

[0147] “百分比(%)同一性”就基因的核苷酸序列而言定义为在比对序列并在必要时引入缺口以获取最大百分比序列同一性后,且不将任何保守取代视为序列同一性的一部分时,候选DNA序列中与DNA序列中的核苷酸相同的核苷酸的百分比。为测定百分比核苷酸序列同一性目的的比对可以本领域技术范围内的多种方式进行,例如使用公众可得到的计算机软件。本领域技术人员能够决定用于测量比对的适宜参数,包括对所比较序列全长获得最大比对所需的任何算法。

[0148] 如在本发明背景中使用的,术语“诱变”指提供核苷酸序列突变体的方法,例如经由一个或多个核苷酸的插入、缺失和/或取代,从而获得在非编码或编码区中具有至少一个变化的其变体。诱变可经由随机、半随机或定点突变获得。通常生成具有高基因多样性的大的随机化基因库,可依照特定期望的基因型或表型对其进行选择。

[0149] 如本文中使用的,术语“感兴趣的蛋白质(POI)”指通过重组技术在宿主细胞中产生的多肽或蛋白质。更具体地,所述蛋白质可以是不天然存在于宿主细胞中的多肽,即异源蛋白质,或者对于宿主细胞可以是天然的,即宿主细胞同源的蛋白质,但例如通过用含有编码POI的核酸序列的自身复制载体转化,或通过重组技术将编码POI的核酸序列的一个或多个拷贝整合到宿主细胞的基因组中,或通过对控制编码POI的基因的表达的一种或多种调控序列(例如启动子序列)的重组修饰来产生。在一些情况下,如本文中使用的,术语POI还指宿主细胞的任何代谢产物,如由重组表达的蛋白质介导的。

[0150] 如本文中使用的,术语“重组”意指“通过遗传工程制备或者是遗传工程的结果”。如此,“重组微生物”包含至少一种“重组核酸”。重组微生物具体地包含表达载体或克隆载体,或它已经遗传工程化以含有重组核酸序列。“重组蛋白质”通过在宿主中表达相应的重组核酸来产生。“重组启动子”是遗传工程化的非编码核苷酸序列,其适合用作如本文中描述的功能活性启动子。

[0151] 令人惊讶地,发现真核细胞能通过限制碳源的可获性来诱导POI的产生。发现碳饥饿条件触发强启动子活性的诱导,这是本领域前所未知的。如记载于US2008299616A1的在糖限制下去阻遏的巴斯德毕赤酵母MLS1启动子实际上对于POI生产是可比地较弱的可调控启动子。因此,令人惊讶地,可鉴定出巴斯德毕赤酵母的这类强可调控启动子并且能在真核生产细胞系中使用,具体地用于重组POI生产。

[0152] 尽管已测定巴斯德毕赤酵母GS115菌株的9.43Mbp基因组序列并披露在US20110021378A1中,但各序列如启动子序列的特性尚未详细研究。例如,如本文中描述的pG4序列(SEQ ID 4)在US20110021378A1中鉴定为启动子序列,然而,其可调控特性或其在碳饥饿条件下的使用是未知的。甚至令人惊讶地是,这类启动子能有效用于依照本发明的方法中。现有技术中的受调控启动子,如在工业规模POI生产中使用的,主要源自甲醇代谢途径且需要加入甲醇以诱导POI生产,这经常是不想要的。依照本发明的方法的优点在于,它可通过增强的表达提供增加的生产,且由于特定的启动子调控具有降低的污染风险,特别是在使用化学性质确定的无甲醇培养基时。

[0153] 发现依照本发明的可调控启动子仅在使用非常特定的适用于建立启动子阻遏和去阻遏条件的培养基时会发挥其可调节活性。例如,巴斯德毕赤酵母能在工业生产工艺条件下成功培养。首先采用在基础碳源如甘油上的分批培养,接着是以有限给料的补充碳源如葡萄糖的补料分批。在接近第一分批阶段结束时以及在受限生长条件下(例如使用有限量的补充碳源)取样品。使用DNA微阵列的转录组分析揭示在补充碳源上有强烈活性且在存在过剩碳时(即过量基础碳源)较弱或无活性的特定基因。至少6种启动子序列被鉴定为依照本发明的可调控启动子,即pG1(SEQ ID 1)、pG3(SEQ ID 2)、pG4(SEQ ID 4)、pG6(SEQ ID 3)、pG7(SEQ ID 5)和pG8(SEQ ID 6)。现有技术中的可比的MLS1或ICL1启动子仅为较弱的,其具有低于pG1启动子1/10的强度,且无可检测的调控。

[0154] 可在发酵工艺中验证在基础碳源上的阻遏性重组基因表达,和在受限的补充碳源上的强表达(即通过底物变化诱导)的特征。

[0155] 可从多种来源获得能用作依照本发明的调控序列的核苷酸序列,其会提供改进的重组蛋白质生产。依照本发明的启动子起源优选来自酵母细胞,最优选来自甲基营养型酵母如来自毕赤酵母属或来自巴斯德毕赤酵母的种,该启动子然后可用作亲本序列以产生适宜的变体,例如突变体或类似物。

[0156] 涵盖一系列酵母细胞,特别是毕赤酵母菌株的酵母细胞,可适用于获得负责在碳饥饿条件下的蛋白质生产的相应启动子序列,或在不同物种中的相应类似物。

[0157] 可采用标准技术生产鉴定的巴斯德毕赤酵母启动子的变体,包括功能活性变体如同源物和类似物。启动子可例如修饰为生成具有改变的表达水平和调节特性的启动子变体。

[0158] 例如,可通过诱变依照本发明的启动子序列来制备启动子库,其可用作亲本分子

例如来微调真核细胞中的基因表达,这通过分析变体在不同发酵策略下的表达并选出适宜的变体。可使用变体的合成库,例如以选出符合产生选定POI要求的启动子。这类变体可具有在真核宿主细胞中增加的表达效率和在耗尽碳源时的高表达。

[0159] 差异性发酵策略将区分生长阶段如依照本发明的步骤a), 和生产阶段如步骤b)。

[0160] 生长和/或生产能适宜地发生在分批模式、补料-分批模式或连续模式中。可使用任何适宜的生物反应器,包括分批、补料-分批、连续、搅拌池反应器或气升式反应器。

[0161] 提供在中试(pilot)或工业规模上的发酵工艺是有利的。工业工艺规模将优选采用至少10L的体积,具体地为至少50L,优选至少1m³,优选至少10m³,最优选至少100m³。

[0162] 优选工业规模中的生产条件,其指例如在100L至10m³或更大的反应器体积中的补料分批培养,采用数天的典型工艺时间,或在约50-1000L或更大的发酵罐体积中的连续过程,以约0.02-0.15h⁻¹的稀释速率。

[0163] 适宜的培养技术可涵盖在生物反应器中的培养,其开始于分批阶段,接着为在特定的高生长速率的较短的指数性补料分批阶段,再之后是在特定的低生长速率的补料分批阶段。另一项适宜的培养技术可涵盖分批阶段继之以在低稀释速率的连续培养阶段。

[0164] 本发明的一个优选的实施方案包括分批培养以提供生物质,接着是补料分批培养以得到高产率POI生产。

[0165] 例如,细胞系可在依照本发明的步骤a)中生长在甘油或葡萄糖上以获得生物质。

[0166] 优选在生物反应器中在获得至少1g/L细胞干重,更优选至少10g/L细胞干重,优选至少20g/L细胞干重的细胞密度的生长条件下培养依照本发明的宿主细胞系。提供在中试或工业规模上的这类生物质生产的产率是有利的。

[0167] 允许生物质累积的生长培养基,特别是基底生长培养基,通常包含碳源、氮源、硫源和磷源。通常而言,这类培养基还包含微量元素和维生素,且还可以包含氨基酸、蛋白胨或酵母提取物。

[0168] 优选的氮源包括NH₄H₂PO₄或NH₃或(NH₄)₂SO₄;

[0169] 优选的硫源包括MgSO₄或(NH₄)₂SO₄或K₂SO₄;

[0170] 优选的磷源包括NH₄H₂PO₄或H₃PO₄或NaH₂PO₄、KH₂PO₄、Na₂HPO₄或K₂HPO₄;

[0171] 其他典型的培养基组分包括KCl、CaCl₂,和微量元素如:Fe、Co、Cu、Ni、Zn、Mo、Mn、I、B;

[0172] 优选地所述培养基用维生素B₇补充;

[0173] 用于巴斯德毕赤酵母的典型生长培养基包含甘油或葡萄糖、NH₄H₂PO₄、MgSO₄、KCl、CaCl₂、生物素和微量元素。

[0174] 在生产阶段中,特别地使用仅仅具有有限量补充碳源的生产培养基。

[0175] 优选地,在具有适宜碳源的矿物培养基中培养宿主细胞系,由此进一步显著简化分离过程。优选的矿物培养基的一个例子是含以下的培养基:可利用的碳源(例如葡萄糖、甘油或甲醇),含大量元素(钾、镁、钙、铵、氯、硫、磷)和微量元素(酮、碘化物、锰、钼酸盐、钴、锌、和铁盐、和硼酸)的盐,以及任选地维生素或氨基酸,例如以补偿营养缺陷型。

[0176] 所述细胞在适于实现期望POI表达的条件下培养,根据表达系统和所表达蛋白的性质,例如该蛋白是否融合信号肽和该蛋白是否可溶或结合膜,所述POI能从细胞或培养基纯化。如熟练技术人员会理解的,培养条件会随着包括宿主细胞类型和采用的具体表达载

体在内的因素而变化。

[0177] 优选地,依照本发明的启动子对POI生产的诱导通过在有限量的补充碳源上培养细胞来控制,所述补充碳源为唯一的碳源和能量源。细胞在碳受限条件下生长得非常缓慢,但在可调控启动子的控制下产生高产率的POI。

[0178] 特定地,启动子活性在去阻遏状态中相比于阻遏状态中的差异为至少2倍,优选至少5倍,更优选至少10倍,更优选至少20倍,更优选至少30、40、50或100倍。

[0179] 通过选择依照本发明的适宜启动子序列,任选地与别的优选调控序列组合,可以提供在可比条件下相对于与生产细胞同源的GAP启动子、天然pGAP或从巴斯德毕赤酵母分离的GAP启动子至少相同,或至少约1.5倍,或至少约2倍,或至少约5倍、10倍或至少高达约15倍的活性(由启动子活性或转录强度代表,或由启动子强度调节)。

[0180] 典型的生产培养基包含补充碳源,和另外的NH₄H₂PO₄、MgSO₄、KCl、CaCl₂、生物素和微量元素。

[0181] 例如,添加到发酵的补充碳源的给料可包含具有高达50wt %可发酵糖的碳源。补充培养基的低给料速率将限制产物抑制对细胞生长的影响,如此基于底物提供的高产物产率将是可能的。

[0182] 优选地,发酵在从3至7.5的pH进行。

[0183] 典型的发酵时间为约24至120小时,温度在20°C至35°C的范围内,优选22-30°C。

[0184] 一般而言,如本文中所指的重组核酸或生物可通过本领域技术人员公知的重组技术产生。依照本发明,可采用常规的分子生物学、微生物学和本领域技术范围内的重组DNA技术。这类技术在文献中有完全解释。参见,例如Maniatis,Fritsch&Sambrook,"Molecular Cloning:A Laboratory Manual"(1982)。

[0185] 依照本发明的一个优选的实施方案,重组构建体通过将启动子和相关基因连接到载体中而获得。这些基因可通过使用这类载体转化宿主细胞而稳定整合到宿主细胞基因组中。

[0186] 表达载体可包括但不限于,克隆载体、经修饰的克隆载体和特定设计的质粒。如本发明中使用的优选表达载体可以是适用于在宿主细胞中表达重组基因的任何表达载体,且其根据宿主生物而选出。重组表达载体可以是能够复制或整合到宿主生物基因组中的任何载体,其亦称为宿主载体。

[0187] 在本发明中,优选使用源自pPUZZLE的质粒作为载体。

[0188] 适宜的表达载体通常包含适用于在真核宿主细胞中表达编码POI的DNA的别的调控序列。调控序列的例子包括操纵子、增强子、核糖体结合位点、和控制转录和翻译启动和终止的序列。所述调控序列可以可操作地连接于要表达的DNA序列。

[0189] 为了允许重组核苷酸序列在宿主细胞中的表达,所述表达载体可提供依照本发明的启动子临近于编码序列的5' 端,例如在信号肽基因的上游。由此,转录由该启动子序列调控和启动。

[0190] 信号肽可以是异源信号肽或天然和异源信号肽的杂合体,并且具体地可以是产生蛋白质的宿主生物异源或同源的。信号肽的功能是允许要分泌的POI进入内质网。它通常是指导蛋白质运输到质膜外的短(长3-60个氨基酸)肽链,由此使得易于分离并纯化该异源蛋白质。一些信号肽在运输蛋白质之后通过信号肽酶从蛋白质切掉。

[0191] 例示性的信号肽是来自酿酒酵母alpha-交配因子前原肽的信号序列和来自巴斯德毕赤酵母酸性磷酸酶基因(PH01)的信号肽。

[0192] 启动子序列理解为可操作地连接于编码序列,如果该启动子控制该编码序列转录的话。如果启动子序列与编码序列不天然关联,那么其转录在天然(野生型)细胞中不受该启动子控制,或者该序列与不同的连续序列重组合。

[0193] 为了证实相关序列的功能,可构建包含一种或多种调控元件的表达载体来驱动POI的表达,并将表达的产率与具有常规调控元件的构建体相比较。对实验规程的详细描述可见于下文实施例。鉴定出的基因可使用特异性核苷酸引物通过PCR从巴斯德毕赤酵母扩增,克隆到表达载体中并转化到真核细胞系中,例如使用酵母载体和巴斯德毕赤酵母的菌株,以用于高水平的生产各种不同的POI。为了估测依照本发明的启动子对如此生产的重组POI量的影响,可在摇瓶实验和补料分批或恒化器发酵中培养该真核细胞系,将其与包含常规的非碳源可调控的启动子(如例如相应细胞中的标准pGAP启动子)的菌株相比较。具体地,启动子的选择对于重组蛋白生产具有极大影响。

[0194] 用于微生物摄取重组DNA片段的优选的转化方法包括化学转化、电穿孔或通过原生质体形成(protoplastation)来转化。依照本发明的转化体可通过向宿主导入这类载体DNA,例如质粒DNA,并选择以高产率表达相关蛋白质或宿主细胞代谢物的转化体而获得。

[0195] POI可使用重组宿主细胞系生产,其通过培养转化体,如此在适宜的培养基中获得,从培养物分离表达的产物或代谢物,并任选地通过适宜方法将其纯化。

[0196] 依照本发明的转化体可通过向宿主导入这类载体DNA,例如质粒DNA,并选择以高产率表达POI或宿主细胞代谢物的转化体而获得。处理宿主细胞以使其能纳入外来DNA,其通过常规用于转化真核细胞的方法,如电脉冲方法、原生质体方法、醋酸锂方法和从其修改后的方法。优选地,巴斯德毕赤酵母通过电穿孔转化。

[0197] 依照本发明的优选宿主细胞系维持依照本发明采用的遗传特性,且生产水平保持较高,例如至少在 μg 水平,甚至在培养约20代,优选至少30代,更优选至少40代,最优选至少50代之后。稳定的重组宿主细胞在用于工业规模生产时视为极大的优势。

[0198] 依照本发明方法用于生产POI的几种不同办法是优选的。通过用携带编码相关蛋白的重组DNA和至少一种如上述的调控元件的表达载体来转化真核宿主细胞,制备经转化细胞的培养物,生长该培养物,诱导转录和POI生产,并回收发酵过程的产物,可表达、处理和任选地分泌物质。

[0199] 优选地,所述POI采用产生至少1mg/L,优选至少10mg/L,优选至少100mg/L,最优选至少1g/L产率的条件表达。

[0200] 优选地,测试依照本发明的宿主细胞的表达能力或产率,其通过以下测试:ELISA、活性测定法、HPLC、或其他适宜的测试。

[0201] 理解本文中公开的方法还可包括在允许POI表达的条件下培养所述重组宿主细胞,优选以分泌的形式或者作为细胞内产物。然后,可从细胞培养基分离重组产生的POI或宿主细胞代谢物,并通过本领域技术人员公知的技术进一步纯化。

[0202] 依照本发明产生的POI通常可使用本领域现有技术分离并纯化,包括提高期望的POI的浓度和/或降低至少一种杂质的浓度。

[0203] 如果POI是从细胞分泌的,那么可使用本领域现有技术从培养基将其分离并纯化。

重组表达产物从宿主细胞的分泌一般是有利的,原因包括有助于纯化过程,因为产物是从培养上清液回收而非从复杂的蛋白质混合物回收,后者是在破坏酵母细胞以释放胞内蛋白质时产生。

[0204] 培养的转化体细胞还可经超声处理或机械、酶法或化学破裂以获得含有期望POI的细胞提取物,从该提取物分离并纯化POI。

[0205] 可以使用用于获得重组多肽或蛋白质产物的分离和纯化方法,所述方法如利用溶解度差异的方法如盐析和溶剂沉淀,利用分子量差异的方法如超滤和凝胶电泳,利用电荷差异的方法如离子交换层析,利用特异性亲和力的方法如亲和层析,利用疏水性差异的方法如反相高效液相层析,和利用等电点差异的方法如等电聚焦。

[0206] 高度纯化的产物基本没有污染性蛋白质,且优选具有至少90%,更优选至少95%,或甚至至少98%,高达100%的纯度。纯化的产物可通过纯化细胞培养上清或从细胞残片获得。

[0207] 作为分离和纯化方法,优选以下标准方法:细胞破坏(如果POI胞内获得)、通过微孔过滤或切向流动滤器(Tangential Flow Filter,TFF)或离心进行的细胞(残片)分离和清洗、通过沉淀或热处理进行的POI纯化、通过酶消化进行的POI活化、通过层析如离子交换(IEX)、疏水相互作用层析(HIC)、亲和层析、大小排阻(SEC)或HPLC层析进行的POI纯化、通过超滤步骤进行的POI浓缩沉淀和清洗(POI precipitation of concentration and washing by ultrafiltration steps)。

[0208] 分离和纯化的POI可通过常规方法如Western印迹、HPLC、活性测定法或ELISA鉴定。

[0209] 所述POI可以是任何真核、原核或合成性多肽。它可以是分泌的蛋白质或胞内蛋白质。本发明还提供天然存在蛋白质的功能同源物、功能等同变体、衍生物和生物活性片段的重组生产。优选地,功能同源物等同或对应于序列且具有该序列的功能特征。

[0210] 本文中所指的POI可以是与真核宿主细胞同源或异源的产物,其优先用于治疗、预防、诊断、分析或工业用途。

[0211] 优选地,所述POI是异源重组多肽或蛋白质,其在真核细胞(优选酵母细胞)中优先作为分泌的蛋白质产生。优先产生的蛋白质的例子是免疫球蛋白、免疫球蛋白片段、抑酶肽(aprotinin)、组织因子途径抑制剂或其他蛋白酶抑制剂、和胰岛素或胰岛素前体、胰岛素类似物、生长激素、白介素、组织血纤维蛋白溶酶原激活剂、转化生长因子a或b、胰高血糖素、胰高血糖素样肽1(GLP-1)、胰高血糖素样肽2(GLP-2)、GRPP、Factor VII、Factor VIII、Factor XIII、血小板源性生长因子1、血清清蛋白、酶如脂肪酶或蛋白酶、或具有与天然蛋白质类似功能的功能同源物、功能等同变体、衍生物或生物活性片段。所述POI可以与天然蛋白质结构相似且可以源自天然蛋白质,其通过在天然蛋白质的C和N末端端之一或两端或侧链处添加一个或多个氨基酸,在天然氨基酸序列的一个或许多不同位点处一个或多个氨基酸的取代、在天然蛋白质的任一端或两端处或在氨基酸序列中的一个或几个位点处一个或多个氨基酸的缺失、或在天然氨基酸序列的一个或多个位点处一个或多个氨基酸的插入。这类修饰对于上述几种蛋白质是公知的。

[0212] POI可选自提供宿主细胞中生化反应的底物、酶、抑制剂或辅因子,其目的是获得所述生化反应或几个反应级联的产物,例如获得宿主细胞的代谢物。例示性的产物可以是

维生素,如核黄素、有机酸和醇,其可在表达依照本发明的重组蛋白或POI后以增加的产率获得。

[0213] 一般而言,表达重组产物的宿主细胞可以是适于POI重组表达的任何真核细胞。

[0214] 优选的哺乳动物细胞的例子是BHK、CHO (CHO-DG44、CHO-DUXB11、CHO-DUKX、CHO-K1、CHOK1SV、CHO-S)、HeLa、HEK293、MDCK、NIH3T3、NS0、PER.C6、SP2/0和VERO细胞。

[0215] 用作依照本发明的宿主细胞的优选的酵母细胞的例子包括但不限于,酵母属(例如酿酒酵母)、毕赤酵母属(例如巴斯德毕赤酵母或甲醇毕赤酵母(*P. methanolica*))、Komagataella属(*K. pastoris*、*K. pseudopastoris*或*K. phaffii*)、多形汉逊酵母或乳酸克鲁维酵母(*Kluyveromyces lactis*)。

[0216] 更新近的文献将巴斯德毕赤酵母划分并重命名为Komagataella *pastoris*、Komagataella *phaffii*和Komagataella *pseudopastoris*。本文中巴斯德毕赤酵母对于所有Komagataella *pastoris*、Komagataella *phaffii*和Komagataella *pseudopastoris*同义使用。

[0217] 优选的酵母宿主细胞源自甲基营养型酵母,如来自毕赤酵母或Komagataella的,例如巴斯德毕赤酵母或Komagataella *pastoris*或*K. phaffii*或*K. pseudopastoris*。宿主的例子包括酵母如巴斯德毕赤酵母。巴斯德毕赤酵母菌株的例子包括CBS 704 (=NRRL Y-1603=DSMZ 70382)、CBS 2612 (=NRRL Y-7556)、CBS 7435 (=NRRL Y-11430)、CBS 9173-9189 (CBS菌株:CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands), 和DSMZ 70877 (German Collection of Micro-organisms and Cell Cultures), 还有来自Invitrogen的菌株,如X-33、GS115、KM71和SMD1168。酿酒酵母菌株的例子包括W303、CEN.PK和BY系列(EUROSCARF保藏)。所有上文所述的菌株均已成功用于产生转化体并表达异源基因。

[0218] 依照本发明,优选的酵母宿主细胞如巴斯德毕赤酵母或酿酒酵母宿主细胞含有异源或重组的启动子序列,其可源自不同于生产宿主的巴斯德毕赤酵母或酿酒酵母菌株。在另一个特定的实施方案中,依照本发明的宿主细胞包含依照本发明的重组表达构建体,其包含起源于与宿主细胞相同的属、物种或菌株的启动子。

[0219] 所述启动子可以是依照本发明的启动子或在宿主细胞中显示转录活性的任何其他DNA序列,且可以源自编码宿主同源或异源的蛋白质的基因。优选地,所述启动子源自与宿主细胞同源的编码蛋白质的基因。

[0220] 例如,依照本发明的启动子可以源自酵母如酿酒酵母菌株,并且可用于在酵母中表达POI。一个特定优选的实施方案涉及源自巴斯德毕赤酵母的依照本发明的启动子用于在巴斯德毕赤酵母生产者宿主细胞系中产生重组POI的方法。核苷酸序列的同源起源协助其纳入同一属或物种的宿主细胞中,如此使得能够稳定生产POI,在工业制备工艺中可能以提高的产率。还有,可使用来自其他适宜酵母或其他真菌或其他生物如脊椎动物或植物的启动子的功能活性变体。

[0221] 如果POI是与宿主细胞同源的蛋白质,即天然存在于宿主细胞中的蛋白质,那么POI在宿主细胞中的表达可通过将其天然启动子序列用依照本发明的启动子序列交换来调控。

[0222] 该目的可例如通过用重组DNA分子转化宿主细胞来实现,所述重组DNA分子包含靶

基因的同源序列以允许位点特异性重组，包含适用于宿主细胞的启动子序列和可选择标志。应发生该位点特异性重组从而将启动子序列与编码POI的核苷酸序列可操作连接。这导致POI从依照本发明的启动子序列而非从天然启动子序列表达。

[0223] 在本发明的一个特定优选的实施方案中，所述启动子序列相对于POI的天然启动子序列具有增加的启动子活性。

[0224] 依照本发明，优选提供包含依照本发明的启动子序列可操作地连接于编码POI的核苷酸序列的巴斯德毕赤酵母宿主细胞系。

[0225] 依照本发明，还可以提供一种依照本发明的wildcard载体或宿主细胞，其包含依照本发明的启动子，且其即时可用于掺入编码POI的感兴趣的基因。如此，wildcard细胞系是预形成的宿主细胞系，其特征在于其表达能力。这是按照一种用于生成生产者细胞系来产生POI的创新的“wildcard”平台策略，例如使用位点特异性重组酶介导的盒交换。这类新的宿主细胞有助于在数天内将感兴趣的基因(GOI)，例如克隆到预确定的基因组表达热点以得到可再生的、高效的生产细胞系。

[0226] 依照一个优选的实施方案，依照本发明的方法采用编码POI的重组核苷酸序列，其在适用于整合到宿主细胞基因组中的质粒上以单拷贝或多拷贝每细胞提供。编码POI的重组核苷酸序列还可以再自主复制质粒上以单拷贝或多拷贝每细胞提供。

[0227] 依照本发明的优选方法采用质粒，其为真核表达载体，优选为酵母表达载体。表达载体可包括但不限于，克隆载体、经修饰的克隆载体和特定设计的质粒。如本发明中使用的优选的表达载体可以是任何适用于在宿主细胞中表达重组基因的表达载体且根据宿主生物进行选择。重组表达载体可以是能够复制或整合到宿主生物基因组中的任何载体(其亦称为宿主载体)，如酵母载体，其携带依照本发明的DNA构建体。优选的酵母表达载体用于在选自下组的酵母中表达：由汉逊酵母、毕赤酵母属、假丝酵母属和*Torulopsis*属代表的甲基营养型酵母。

[0228] 在本发明中，优选使用源自pPICZ、pGAPZ、pPIC9、pPICZalpha、pGAPZalpha、pPIC9K、pGAPHiS或pPUZZLE的质粒作为载体。

[0229] 依照本发明的一个优选的实施方案，通过将相关基因连接到载体中来获得重组构建体。这些基因可通过使用这类载体转化宿主细胞而稳定整合到宿主细胞基因组中。由所述基因编码的多肽可使用重组宿主细胞系生产，其通过培养转化体，如此在适宜的培养基中获得，从培养物分离表达的POI，并通过适合表达产物的方法将其纯化，具体为将POI与污染性蛋白质分开。

[0230] 表达载体可包含一种或多种表型可选择标志，例如编码赋予抗生素抗性或供应自养型需求的蛋白质的基因。酵母载体普遍含有来自酵母质粒的复制起点、自主复制序列(ARS)或者用于整合到宿主基因组中的序列、启动子区、用于多腺苷酸化的序列、用于转录终止的序列和可选择标志。

[0231] 用于分别连接DNA序列例如编码前体序列和/或POI的DNA序列、启动子和终止子，并将其插入含有整合或宿主复制所需信息的合适载体的规程是本领域技术人员公知的，例如由J.Sambrook等，“Molecular Cloning 2nd ed.”, Cold Spring Harbor Laboratory Press (1989) 描述的。

[0232] 将理解可构建使用依照本发明的调控元件和/或POI作为整合靶物的载体，其通过

首先制备含有编码调控元件和/或POI的完整DNA序列的DNA构建体,然后将此片段插入适宜的表达载体中,或通过连续插入含有各个元件(如信号、前导或异源蛋白)遗传信息的DNA片段继之以连接。

[0233] 依照本发明还可使用多克隆载体,其为具有多克隆位点的载体,其中可将期望的异源基因在多克隆位点掺入以提供表达载体。在表达载体中,将启动子置于POI的基因的上游并调控该基因的表达。在多克隆载体的情况下,由于在多克隆位点引入POI的基因,因此将启动子置于多克隆位点的上游。

[0234] 可通过确立的标准方法合成制备如提供以获得依照本发明的重组宿主细胞的DNA构建体,所述方法例如亚磷酰胺(phosphoramidite)方法。所述DNA构建体还可以具有基因组或cDNA起源,例如通过制备基因组或cDNA库并筛选编码本发明多肽的全部或部分的DNA序列来获得,所述筛选通过依照标准技术使用合成的寡核苷酸探针进行杂交来进行(Sambrook等,Molecular Cloning:A Laboratory Manual,Cold Spring Harbor,1989)。最后,DNA构建体可具有混合的合成和基因组起源,混合的合成和cDNA起源,或混合的基因组和cDNA起源,其通过依照标准技术将合成的、基因组或cDNA起源的片段(所述片段对应于完整DNA构建体的各个部分)适宜地退火来制备。

[0235] 在另一个优选的实施方案中,所述酵母表达载体能够稳定整合于酵母基因组中,例如通过同源重组。

[0236] 优选地,通过用依照本发明的调控元件和/或POI基因转化细胞获得的依照本发明的转化体宿主细胞可首先在以下条件培养,即有效生长至大量细胞而没有表达异源蛋白质的负担的条件。当该细胞系准备好用于POI表达时,选择培养技术来产生表达产物。

[0237] 下列定义的主题视为本发明的实施方案:

[0238] 1.一种通过培养包含表达构建体的重组真核细胞系来产生感兴趣蛋白质(POI)的方法,所述表达载体包含可调控的启动子和在所述启动子的转录调控下的编码POI的核酸分子,所述方法包括以下步骤:

[0239] a)用阻遏所述启动子的基础碳源培养所述细胞系,

[0240] b)用使所述启动子去阻遏的无补充碳源的条件或有有限量补充碳源的条件培养所述细胞系,从而诱导所述POI以相比于所述细胞的天然pGAP启动子至少15%的转录速率产生,并

[0241] c)产生并回收所述POI。

[0242] 2.依照项1的方法,其中所述基础碳源选自下组:葡萄糖、甘油、乙醇、和复合营养材料。

[0243] 3.依照项1或2的方法,其中所述补充碳源是诸如葡萄糖、果糖、半乳糖或甘露糖等己糖,诸如蔗糖等二糖,诸如甘油或乙醇等醇,或其混合物。

[0244] 4.依照项1至3中任一项的方法,其中所述基础碳源是甘油,且所述补充碳源是葡萄糖。

[0245] 5.依照项1至4中任一项的方法,其中步骤b)采用不提供或有有限量的所述补充碳源的补料培养基,优选为所述培养基中0-1g/L。

[0246] 6.依照项5的方法,其中所述给料培养基是化学性质确定的(chemically defined)且无甲醇的。

[0247] 7. 依照项1至6中任一项的方法,其中所述有限量的补充碳源为生长限制性的以保持范围在 0.02h^{-1} 至 0.2h^{-1} ,优选 0.02h^{-1} 至 0.15h^{-1} 内的比生长速率。

[0248] 8. 依照项7的方法,其中所述有限量的补充碳源提供细胞培养物中低于检测限的残余量。

[0249] 9. 依照项1至8中任一项的方法,其中所述启动子能控制野生型真核细胞中基因的转录,该基因选自下组:G1 (SEQ ID 7)、G3 (SEQ ID 8)、G4 (SEQ ID 9)、G6 (SEQ ID 10)、G7 (SEQ ID 11) 或 G8 (SEQ ID 12)、或其功能活性变体。

[0250] 10. 依照项9的方法,其中所述功能活性变体选自下组:具有至少约60%核苷酸序列同一性的同源物,通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物,优选包含或组成为至少200bp的核苷酸序列,以及源自巴斯德毕赤酵母(*Pichia pastoris*)以外的物种的类似物。

[0251] 11. 依照项9或10的方法,其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和 pG1f (SEQ ID 46)。

[0252] 12. 依照项1至11中任一项的方法,其中所述启动子是巴斯德毕赤酵母启动子或其功能活性变体。

[0253] 13. 依照项1至12中任一项的方法,其中所述细胞系选自下组:哺乳动物、昆虫、酵母、丝状真菌和植物细胞系,优选为酵母。

[0254] 14. 依照项13的方法,其中所述酵母选自下组:毕赤酵母属(*Pichia*)、假丝酵母属(*Candida*)、球拟酵母属(*Torulopsis*)、*Arxula*、汉逊酵母属(*Hansenula*)、西洋蓍霉属(*Yarrowia*)、克鲁维酵母属(*Kluyveromyces*)、酵母属(*Saccharomyces*)、*Komagataella*,优选甲基营养型酵母。

[0255] 15. 依照项14的方法,其中所述酵母是巴斯德毕赤酵母、*Komagataella pastoris*、*K. phaffii*或*K. pseudopastoris*。

[0256] 16. 依照项1至15中任一项的方法,其中所述启动子不与编码所述POI的核苷酸序列天然关联。

[0257] 17. 依照项1至16中任一项的方法,其中所述POI为异源蛋白质,优选选自治疗性蛋白质包括抗体或其片段、酶和肽、蛋白质抗生素、毒素融合蛋白、碳水化合物-蛋白质偶联物、结构蛋白、调节蛋白、疫苗和疫苗样蛋白或颗粒、加工酶、生长因子、激素和细胞因子、或POI的代谢物。

[0258] 18. 依照项1至17中任一项的方法,其中所述POI为真核蛋白,优选为哺乳动物蛋白。

[0259] 19. 依照项1至18中任一项的方法,其中所述POI为多聚体蛋白,优选为二聚体或四聚体。

[0260] 20. 依照项1至19中任一项的方法,其中所述POI为抗原结合分子如抗体或其片段。

[0261] 21. 依照项1至20中任一项的方法,其中使用所述POI、其代谢物或衍生物制备发酵产物。

[0262] 22. 用于调控POI在重组真核细胞中在碳源可调控启动子的转录控制下表达的方法,所述启动子相比于所述细胞的天然pGAP启动子具有至少15%的转录强度,其中所述表

达在限制所述碳源的条件下诱导。

[0263] 23. 在重组真核细胞中于碳源可调控启动子的转录控制下产生POI的方法, 其中所述启动子相比于所述细胞的天然pGAP启动子具有至少15%的转录强度。

[0264] 24. 依照项1至23中任一项的方法, 其中所述可调控启动子包含选自下组的核酸序列:

[0265] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0266] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0267] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6); 和

[0268] d) 源自a)、b) 或c) 的片段或变体,

[0269] 其中所述启动子是功能活性的启动子, 其为碳源可调控的启动子, 所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达POI。

[0270] 25. 依照项24的方法, 其中所述pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) 的变体是选自下组的功能活性变体: 具有至少约60%核苷酸序列同一性的同源物, 通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物, 优选包含或组成为至少200bp的核苷酸序列, 以及源自巴斯德毕赤酵母以外的物种的类似物。

[0271] 26. 依照项24或25的方法, 其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和pG1f (SEQ ID 46)。

[0272] 27. 一种分离的核酸, 其包含选自下组的核酸序列:

[0273] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0274] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6) ;

[0275] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6); 和

[0276] d) 源自a)、b) 或c) 的片段或变体,

[0277] 其中所述核酸包含功能活性的启动子, 其为碳源可调控的启动子, 所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达POI。

[0278] 28. 依照项27的核酸, 其中所述pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或pG8 (SEQ ID 6) 的变体是选自下组的功能活性变体: 具有至少约60%核苷酸序列同一性的同源物, 通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物, 优选具有至少200bp的核苷酸序列, 以及源自巴斯德毕赤酵母以外的物种的类似物。

[0279] 29. 依照项27或28的核酸, 其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和pG1f (SEQ

ID 46)。

[0280] 30. 一种表达构建体,其包含依照项27至29中任一项的核酸可操作地连接于在所述启动子转录控制下的编码POI的核苷酸序列,所述核酸与编码所述POI的核苷酸序列不天然关联。

[0281] 31. 包含依照项30的构建体的载体。

[0282] 32. 一种重组真核细胞,其包含项30的构建体,或项31的载体。

[0283] 33. 依照项31的细胞,其选自下组:哺乳动物、昆虫、酵母、丝状真菌和植物细胞系,优选为酵母。

[0284] 34. 依照项32的细胞,其中所述酵母选自下组:毕赤酵母属(*Pichia*)、假丝酵母属(*Candida*)、球拟酵母属(*Torulopsis*)、*Arxula*、汉逊酵母属(*Hansenula*)、西洋蓍霉属(*Yarrowia*)、克鲁维酵母属(*Kluyveromyces*)、酵母属(*Saccharomyces*)、*Komagataella*,优选甲基营养型酵母。

[0285] 35. 依照项34的细胞,其中所述酵母是巴斯德毕赤酵母、*Komagataella pastoris*、*K. phaffii*或*K. pseudopastoris*。

[0286] 36. 依照项32至35中任一项的细胞,其在存在过剩碳源的情况下相对于有限碳源条件具有更高的比生长速率。

[0287] 37. 从真核细胞鉴定出碳源可调控启动子的方法,包括以下步骤:

[0288] a) 在存在碳源的情况下,在细胞生长条件下于分批培养中培养真核细胞,

[0289] b) 进一步在存在有限量的补充碳源的情况下,在补料分批培养中培养所述细胞,

[0290] c) 提供步骤a) 和b) 的细胞培养物的样品,并

[0291] d) 在所述样品中实施转录分析以鉴定出在步骤b) 的细胞中比在步骤a) 的细胞中显示更高转录强度的可调控启动子。

[0292] 38. 依照项37的方法,其中所述转录分析是定量或半定量的,优选采用DNA微阵列、RNA测序和转录组分析。

[0293] 具体的例子涉及产生报告蛋白的重组生产巴斯德毕赤酵母细胞系的补料分批发酵,其采用甘油分批培养基和葡萄糖补料分批培养基。比较性启动子活性研究已证实可成功活化依照本发明的启动子来诱导重组蛋白生产。

[0294] 依照一个别的例子,将人血清清蛋白(HSA)作为POI在葡萄糖限制性诱导启动子的控制下产生,并测定HSA产率和基因拷贝数。

[0295] 依照另一个例子,实施在依照本发明的启动子控制下表达HSA的巴斯德毕赤酵母菌株的补料分批培养。发现在葡萄糖限制条件下对启动子活性的诱导相比于阻遏状态甚至超过120倍(对于pG1)和超过20倍(对于pG6)。

[0296] 别的例子涉及在pG1和pG6启动子的转录控制下表达猪羧肽酶B作为模式蛋白。

[0297] 再一个例子涉及抗体片段在pG1的转录控制下的表达。

[0298] 又一个例子证实了依照本发明的启动子变体的功能活性,如长度范围为300至1000bp的pG1片段。另外的实验已显示甚至更短的pG1片段在类似背景中也是功能活性的,如范围在200至1000bp的片段,或范围在250至1000的片段。

[0299] 参照以下实施例,前述说明将得到更全面的理解。然而,这类实施例仅代表实践本发明一个或多个实施方案的方法且不应解读为限制本发明的范围。

[0300] 实施例

[0301] 下文实施例例示了用于鉴定新的可调控启动子并分析其在巴斯德毕赤酵母中的表达特性的材料和方法。

[0302] 实施例1:在葡萄糖受限条件下在巴斯德毕赤酵母中鉴定多个被有效调控的强的基因

[0303] 为了鉴定出在葡萄糖受限条件下在巴斯德毕赤酵母中的多个被有效调控的强的基因及它们的相应启动子,使用微阵列进行基因表达模式的分析。将在甘油分批(过剩碳源)中生长的巴斯德毕赤酵母细胞与在以葡萄糖为生长限制的条件中(恒化器)培养的细胞(由此模拟通常以补料分批模式进行的蛋白质生产工艺的过程)相比较。

[0304] a) 菌株

[0305] 使用野生型巴斯德毕赤酵母菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures,Utrecht,The Netherlands),其能在没有补充物的最小培养基上生长。

[0306] b) 巴斯德毕赤酵母的培养

[0307] 发酵用Minifors反应器(Infors-HT, Switzerland)进行,最终工作体积为2.5L。

[0308] 使用以下培养基:

[0309] PTM₁痕量盐储液,每升含有

[0310] 6.0g CuSO₄ • 5H₂O, 0.08g NaI, 3.36g MnSO₄ • H₂O, 0.2g Na₂MoO₄ • 2H₂O, 0.02g H₃BO₃, 0.82g CoCl₂, 20.0g ZnCl₂, 65.0g FeSO₄ • 7H₂O, 0.2g生物素和5.0ml H₂SO₄(95% - 98%)。

[0311] 甘油分批培养基,每升含有

[0312] 2g柠檬酸一水合物(C₆H₈O₇ • H₂O), 39.2g甘油, 20.8g NH₄H₂PO₄, 0.5g MgSO₄ • 7H₂O, 1.6g KC₁, 0.022g CaCl₂ • 2H₂O, 0.8mg生物素和4.6ml PTM1痕量盐储液。添加HC1以将pH设成5。

[0313] 甘油补料分批培养基,每升含有

[0314] 632g甘油, 8g MgSO₄ • 7H₂O, 22g KC₁, 和0.058g CaCl₂ • 2H₂O。

[0315] 恒化器培养基,每升含有

[0316] 2g柠檬酸一水合物(C₆H₈O₇ • H₂O), 99.42g葡萄糖一水合物, 22g NH₄H₂PO₄, 1.3g MgSO₄ • 7H₂O, 3.4g KC₁, 0.02g CaCl₂ • 2H₂O, 0.4mg生物素和3.2ml PTM1痕量盐储液。添加HC1以将pH设成5。

[0317] 将溶解氧用搅拌器速度(500-1250rpm)控制为D0=20%。通风速率为60L h⁻¹空气,温度控制为25°C,且添加NH₄OH(25%)以控制pH设定点为5。

[0318] 为了开始发酵,将1.5L分批培养基无菌过滤到发酵罐中,并将巴斯德毕赤酵母以起始光密度(OD600)1进行接种(从YPG中,180rpm,28°C的过夜预培养物)。分批阶段达到约25h时干生物质浓度约20g/L,接着是使用葡萄糖培养基进行10h的指数性补料分批培养,产生约50g/L的干生物质浓度。然后,将体积减少至1.5L,并以0.15L h⁻¹的补料/收获速率开始恒化器培养,从而得到μ=0.1的恒定生长速率。发酵在恒化器开始后的50h时终止。

[0319] 该发酵进行3次以获得可靠的微阵列分析所需要的生物学重复份数。

[0320] 通过对培养上清液的HPLC分析来验证恒化器期间的碳受限条件(无可检测的残余

葡萄糖)。

[0321] c) 取样

[0322] 在甘油分批阶段结束时和葡萄糖恒化器的稳定态条件下取样品。在每一发酵期间一起完成常规采样以便进行光密度或酵母干重量的测定,定性显微镜检查和细胞活力分析。对于微阵列分析,取样品并如下处理:对于最佳猝灭,将9mL细胞培养液立即与4.5mL冰冷的5%酚(Sigma)溶液(在无水乙醇中)混合,并等分试样。将每2mL在预冷的收集管(GE healthcare, NJ)中离心(13200rpm 1分钟),完全除去上清液并将管储存在-80°C直至RNA纯化。

[0323] d) 用于微阵列杂交的RNA纯化和样品制备

[0324] 使用TRI试剂依照供应商说明书(Ambion, US)分离RNA。将细胞团粒重悬在TRI试剂中并用玻璃珠匀浆化,其以 5ms^{-1} 使用FastPrep 24(M.P.Biomedicals, CA)达40秒。在添加氯仿后,将样品离心并通过添加异丙醇从水相沉淀出总RNA。将团粒用70%乙醇清洗,干燥并重悬在无RNase的水中。通过使用Nanodrop 1000分光光度计(NanoDrop products, DE)测量OD260来测定RNA浓度。使用DNAfree Kit(Ambion, CA)从样品除去剩余DNA。将等同于10 μg RNA的样品体积在无RNase的水中稀释至50 μL ,然后添加DNase缓冲液I和rDNase I,并在37°C温育30分钟。在添加DNase灭活试剂后,将样品离心并将上清液转移至新管。如上文描述的再次测定RNA浓度。另外,使用RNA纳米芯片(Agilent)分析RNA完整性。为了监测从样品的扩增和标记到杂交的微阵列工作流,将Spike In Kit(Agilent, 产品编号:5188-5279)用作阳性对照。它含有来自腺病毒的10种不同的多腺苷酸化转录本,其与本身的RNA样品一起扩增、标记和共杂交。使用Quick Amp Labelling Kit(Agilent, Prod.No.: 5190-0444)将样品用Cy 3和Cy 5标记。因此,将500ng纯化的样品RNA稀释于8.3 μL 无RNase的水,添加2 μL Spike A或B,和1.2 μL T7启动子引物。将混合物在65°C变性10分钟并在冰上保持5分钟。然后,添加8.5 μL cDNA主混合物(每份样品:4 μL 5x第一链缓冲液,2 μL 0.1M DTT,1 μL 10mM dNTP混合物,1 μL MMLV-RT,0.5 μL RNase out),在40°C温育2小时,然后转移至65°C达15分钟并置于冰上5分钟。制备转录主混合物(每份样品:15.3 μL 无核酸酶的水,20 μL 转录缓冲液,6 μL 0.1M DTT,6.4 μL 50%PEG,0.5 μL RNase抑制剂,0.6 μL 无机磷酸酶,0.8 μL T7 RNA聚合酶,2.4 μL Cyanin 3或Cyanin 5)并添加到每个管,在40°C温育2小时。为了纯化获得的经标记的cRNA,使用RNeasy Mini Kit(Qiagen, 目录号74104)。将样品储存于-80°C。对cRNA浓度和标记效率的定量在Nanodrop分光光度计处完成。

[0325] e) 微阵列分析

[0326] 为了鉴定出葡萄糖受限的恒化器培养中被有效调控的强基因,将其3份生物学样品重复品与同一参照比较且是各在一次染色交换(dyeswap)中比较。参照行品通过甘油分批培养的样品以相等量组合而生成。

[0327] 将Gene Expression Hybridisation Kit(Agilent, Cat.No. 5188-5242)用于经标记的样品cRNA的杂交。对于杂交样品的制备,将每300ng cRNA(Cy3和Cy 5)和6 μL 10倍封闭试剂用无核酸酶的水稀释成终体积24 μL 。在添加1 μL 25倍片段化缓冲液后,将混合物在60°C温育30分钟。然后,添加25 μL GEx Hybridisation Buffer HI-RPM以停止反应。在以13,200rpm离心1分钟后,将样品在冰上变冷并立即用于杂交。使用自有设计的巴斯德毕赤酵母特异性寡核苷酸阵列(AMAD-ID: 026594, 8x15K定制阵列, Agilent)。微阵列杂交依照

Microarray Hybridisation Chamber User Guide(Agilent G2534A)完成。首先,将衬垫载玻片(gasket slide)移去盖并置于室(chamber)基底上,Agilent标签朝上。将样品(40 μ L每阵列)加载到8个正方形中每一个的中央。然后,将微阵列载玻片小心放置到衬垫载玻片上(Agilent标签朝下),将室盖放置在上面,并用夹具(clamp)固定。杂交在杂交炉中于65°C进行17小时。在扫描前,清洗微阵列芯片。因此,在浸没于清洗缓冲液1中时,将室拆除,使三明治式载玻片彼此脱离。将微阵列直接转移到具有清洗缓冲液1的另一个皿,清洗1分钟,转移至清洗缓冲液2(温度至少30°C)并再清洗1分钟。在通过将载玻片边与薄纸接触来干燥微阵列载玻片后,将其置于载玻片支持物中(Agilent标签朝上)。将载玻片支持物置于传送带(carousel)中并开始扫描。

[0328] f) 数据采集和微阵列数据的统计学评估

[0329] 将图像以50nm的分辨率用G2565AA微阵列扫描器(Agilent)扫描并输入Agilent Feature Extraction 9.5软件。将Agilent Feature Extraction 9.5用于点的强度的定量。然后,将原始均值点强度数据输入开放源软件R以用于进一步标准化和数据分析。

[0330] 对于数据预处理和标准化,使用R程序包limma、vsn和marray。强度数据不经背景校正而用VSN标准化,在标准化后将其转化成Cy5通道相对Cy3通道的log2比。使用limma程序包的lmfit和eBayes函数来计算差异表达。

[0331] 浏览微阵列数据中具有在阻遏态到诱导态之间表达水平中的高差异(倍数变化)以及在诱导态中的高信号强度的录入项,从而鉴定强表达的被有效调控的基因。选定的基因的列表显示于表1,倍数变化意指诱导态中的信号强度除以阻遏态中的信号强度。加入pGAP和pMLS1,pICL1的数据作为参照。

[0332] 表1:选择用于进一步表征的启动子和作为对照的pGAP、ICL1和MLS1的微阵列数据

名称	注释/酵母 同源物	基因标识号	倍数变化	强度*	%强度/转 录强度
pGAP	TDH3	PAS_chr2-1_0437	0.79	41052.5	100.0
pG1	-	PAS_chr1-3_0011	29.86	86312.9	210.2
pG3	YPR127W	PAS_chr4_0550	2.66	15644.4	38.1
pG4	-	PAS_chr4_0043	2.57	15664.8	38.2
pG6	ALD4	PAS_chr2-1_0853	2.10	26888.4	65.5
MLS1	MLS1	PAS_chr4_0191	0.81	1446.9	3.5
ICL1	ICL1	PAS_chr1-4_0338	1.71	2574.3	6.3

[0334]	pG7	HXT6	PAS_chr1-4_0570	3.3	13336.5	32.5
	pG8	SFL1	PAS_chr1-3_0165	2.1	9929.1	24.2

[0335] *绿通道中的诱导态

[0336] 实施例2:新鉴定的启动子在巴斯德毕赤酵母中的比较性启动子活性研究,使用eGFP作为胞内表达的报告基因

[0337] 为了分析新鉴定启动子在葡萄糖限制条件下的特性,如下实施摇瓶筛选:用含有

甘油作为碳源的丰富培养基完成24小时的预培养—模拟分批阶段(启动子的阻遏态),之后使用最小培养基和葡萄糖给料珠进行主要培养—模拟葡萄糖受限补料分批阶段(启动子的诱导态)。

[0338] a) 菌株和表达载体

[0339] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec; 150 (4) :519-29),该载体包含大肠杆菌的复制起点(pUC19)、抗生素抗性盒(赋予对新霉素抗性的Sh ble基因)以在大肠杆菌和酵母中选择、感兴趣基因(GOI)的表达盒(由多克隆位点和酿酒酵母CYC1转录终止子组成)、和用于整合到巴斯德毕赤酵母基因组(3' AOX1区)中的基因座。

[0340] b) 新鉴定的启动子pG1、pG3、pG4和pG6扩增和克隆到含有eGFP作为GOI的pPUZZLE表达载体中

[0341] 新鉴定的启动子序列及其相应基因(见实施例1)的列表显示于表2。将相应基因的5'非编码区直至起始密码子ATG的1000bp作为启动子序列通过PCR扩增(Phusion Polymerase, New England Biolabs),其使用表2中显示的引物。将这些序列克隆到pPUZZLE表达载体pPM1aZ10_eGFP中,得到pPM1aZ10_pG1_eGFP、pPM1aZ10_pG3_eGFP、pPM1aZ10_pG4_eGFP和pPM1aZ10_pG6_eGFP。另外,将含有普遍使用的甘油醛3-磷酸脱氢酶启动子(巴斯德毕赤酵母的pGAP,本文中SEQ ID 25)的载体pPM1aZ10_pGAP_eGFP用作参照。使用ApaI和SbfI限制性位点(见表2和3),将启动子插入eGFP基因的开始密码子的上游。通过Sanger测序验证启动子序列的正确性。

名称	靶物	序列	T _M	限制性位点	
[0342]	pG1_fw	pG1	SEQ ID 14 GATAGGGCCCCAACACATTGCT CCCCCTAGTCTC	70.8	ApaI
			SEQ ID 15 GATACCTGCAGGAAGGGTGGAA TTTTAAGGATCTTTAT		
	pG1_back	pG1	SEQ ID 16 GATAGGGCCCCAGCAATCCAGT AACCTTTCTGAAT	69.8	SbfI
			SEQ ID 17 GATACCTGCAGGTTGAGTTCAAT AAATTGTCCGGGA		
	pG3_fw	pG3	SEQ ID 18 GATAGGGCCCTGGACTGTTCAAT TTGAAGTCGATG	70.4	ApaI
			SEQ ID 19 GATACCTGCAGGGATAAAGGT AAGGGAAAAAAGCAA		
	pG3_back	pG3	SEQ ID 20 GATAGGGCCCAGACCAGCAGTT TAACTACGCAAATC	70.2	SbfI
			SEQ ID 21 GATACCTGCAGGCTTTCTTG GCAAGGAAAAATC		
	pG6_fw	pG6	SEQ ID 22 GATAGGGCCAATTGATTAAGTT CAGTGAAATTCAAAC	70.6	ApaI
			SEQ ID 23 GATACCTGCAGGATTATATTATGG GGAATAATGAAGAGAAGG		
	pG6_back	pG6		70.7	SbfI
	pG7_fw	pG7		69.1	ApaI
	pG7_back	pG7		70.9	SbfI

[0343]	pG8_fw	pG8	SEQ ID 24 GATAGGGCCCTGCACAACCATT GCCAGTAAGG	71.5	ApaI
	pG8_back	pG8	SEQ ID 25 GATACCTGCAGGTTTAGAAGA GGGAGAACTTAGATTGG	70.4	SbfI

[0344] 表2:用于启动子的PCR扩增的引物

启动子	5'引物	3'引物	克隆酶		片段
			5'	3'	
pG1	pG1_fw	pG1_back	ApaI	SbfI	988
pG3	pG3_fw	pG3_back	ApaI	SbfI	1011
pG4	pG4_fw	pG4_back	ApaI	SbfI	1022
pG6	pG6_fw	pG6_back	ApaI	SbfI	1022
pG7	pG7_fw	pG7_back	ApaI	SbfI	1022
pG8	pG8_fw	pG8_back	ApaI	SbfI	1022

[0346] 表3:扩增引物、克隆酶和克隆的启动子的长度

[0347] c) eGFP在巴斯德毕赤酵母中的表达用于分析启动子活性

[0348] 将所有质粒在3' AOX基因组整合区内用AscI线性化,接着电穿孔(2kV, 4ms, GenePulser, BioRad)到电感受态的巴斯德毕赤酵母中。

[0349] 对阳性转化体的选择在含有25μg/mL新霉素(Invivogen, CA)的YPD板(每升:10g酵母提取物,20g蛋白胨,20g葡萄糖,20g琼脂-琼脂)上进行。使用菌落PCR以确保转化质粒的存在。因此,通过将巴斯德毕赤酵母菌落热煮和冷冻各5分钟得到基因组DNA,并直接用适宜引物进行PCR。对于表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,将OD600为0.1的预培养物接种10ml YP培养基(每升:20g蛋白胨,10g酵母提取物)和2份葡萄糖给料珠(Kuhner, CH)中的主培养物。葡萄糖限制性生长条件由于这些给料珠的缓慢的葡萄糖释放动力学而实现,其由以下等式描述:(葡萄糖)= $1.63*t^{0.74}$ [mg/Disc]。在预培养结束时,以及接种主培养物后24和48小时取样品。细胞密度通过测量OD600测定,eGFP表达通过流式细胞术如Stadlmayr等(J.Biotechnology 2010Dec;150(4):519-29)中描述的分析。对于每份样品,分析10,000个细胞。使用未转化的巴斯德毕赤酵母野生型细胞测量巴斯德毕赤酵母的自身荧光,并从信号减去。相对的eGFP表达水平(荧光强度对细胞大小)显示为在构成性pGAP启动子控制下表达eGFP的克隆物的eGFP表达水平的百分数。

[0350] 使用启动子pG7和pG8完成别的类似的研究。如实施例2b中描述的完成克隆,但是将野生型巴斯德毕赤酵母菌株X-33(Invitrogen)用于转化pPM1aZ10_pG7_eGFP和

pPM1aZ10_pG8_eGFP。使用的引物和克隆片段列于表2和3。结果显示于表4。

	预培养		主培养		
	分批结束	标准差	48h	标准差	
[0351]	pG1	7.6	0.2	242.8	59.5
	pG3	-5.1	2.4	25.4	5.5
	pG4	-6.3	0.2	113.6	26.3
	pG6	3.3	0.8	158.9	146.9
	pG7	49.4	7.4	115.7	16.2
	pG8	0.8	4.1	36.1	21.1

[0352] 表4:在新启动子控制下表达eGFP的巴斯德毕赤酵母克隆物的筛选结果;显示的数据(荧光/细胞大小)与pGAP有关;

[0353] d) 在选定的表达eGFP的克隆物中测定eGFP基因拷贝数 (GCN)

[0354] 表达强度常与整合到巴斯德毕赤酵母基因组中的表达盒数目关联。因此,测定eGFP的基因拷贝数。使用DNeasy Blood&Tissue Kit (Quiagen, 目录号69504) 分离基因组DNA。使用定量PCR测定基因拷贝数。因此,使用SensiMix SYBR Kit (Bioline, QT605-05)。将Sensi Mix SYBR与引物和样品混合,并在实时PCR循环仪 (Rotor Gene, Qiagen) 中用于实时分析。引物的列表显示于表5。一式三份或一式四份地分析所有样品。将Rotor Gene软件用于数据分析。将肌动蛋白基因ACT1用作校正物。结果显示于表6。

[0355]	引物	靶物	序列	T _M	产物长度

			[°C]	
[0356]	PpACT1_Up	Act	SEQ ID 26 CCTGAGGCTTGTTCC ACCCATCT	61.3 148 bp
	PpACT1_Low	Act	SEQ ID 27 GGAACATAGTAGTACC ACCGGACATAACGA	61.4 148 bp
	PpeGFP_Up	GFP	SEQ ID 28 TCGCCGACCACTACCA GCAGAAA	61.4 124 bp
[0357]	PpeGFP_Low	GFP	SEQ ID 29 ACCATGTGATCGCGCT TCTCGTT	61.6 124 bp

表5:通过实时PCR进行拷贝数测定时用的引物

[0358]		% pGAP_eGFP荧光/大 小			
		GCN	预培养	主培养	
				24h	48h
	pG1#8	1	7.32	33.00	184.30
	pG1#9	1	7.73	33.96	303.21
	pG1#12	2	7.75	33.32	240.92
	pG6#48	1	3.45	2.07	56.11
	pG6#50	2	4.00	23.18	327.14
	pG6#53	1	2.52	9.78	93.51

[0359] 表6:在pG1和pG6控制下表达eGFP的选定的巴斯德毕赤酵母克隆物的筛选结果(与pGAP有关的荧光/细胞大小)和基因拷贝数;

[0360] e) 在一种eGFP克隆物的补料分批发酵中分析pG1启动子强度

[0361] 在DASGIP反应器中实施补料分批发酵,最终工作体积为0.7L。

[0362] 使用以下培养基:

[0363] PTM1痕量盐储液,每升含有

[0364] 6.0g CuSO₄ • 5H₂O, 0.08g NaI, 3.36g MnSO₄ • H₂O, 0.2g Na₂MoO₄ • 2H₂O, 0.02g

H_3BO_3 , 0.82g $CoCl_2$, 20.0g $ZnCl_2$, 65.0g $FeSO_4 \cdot 7H_2O$, 0.2g 生物素和 5.0ml H_2SO_4 (95% - 98%)。

[0365] 甘油分批培养基,每升含有

[0366] 2g 柠檬酸一水合物 ($C_6H_8O_7 \cdot H_2O$), 39.2g 甘油, 12.6g $NH_4H_2PO_4$, 0.5g $MgSO_4 \cdot 7H_2O$, 0.9g KCl , 0.022g $CaCl_2 \cdot 2H_2O$, 0.4mg 生物素和 4.6ml PTM1 痘量盐储液。添加 HCl 以将 pH 设成 5。

[0367] 葡萄糖补料分批培养基,每升含有

[0368] 464g 葡萄糖一水合物, 5.2g $MgSO_4 \cdot 7H_2O$, 8.4g KCl , 0.28g $CaCl_2 \cdot 2H_2O$, 0.34mg 生物素和 10.1mL PTM1 痘量盐储液。

[0369] 将溶解氧用搅拌器速度 (400-1200rpm) 控制为 $D_0 = 20\%$ 。通风速率为 24L h^{-1} 空气, 温度控制为 25°C, 且添加 NH_4OH (25%) 以控制 pH 设定点为 5。

[0370] 为了开始发酵, 将 400mL 分批培养基无菌过滤到发酵罐中, 并将巴斯德毕赤酵母克隆 pG1_eGFP#8 以起始光密度 (OD600) 1 进行接种 (来自预培养物)。约 25h 的分批阶段 (达到约 20g/L 的干生物质浓度) 之后, 是葡萄糖受限的补料分批, 其始于指数性给料 7h 并恒速给料 15g/L 达 13h, 从而产生约 100g/L 的最终干生物质浓度。在分批和补料分批阶段取样品, 并使用板阅读器 (Infinite 200, Tecan, CH) 分析 eGFP 表达。因此, 将样品稀释成光密度 (OD600) 为 5。结果显示于表 7, 其为相对荧光值/生物反应器 (FL/r)。

	pGAP_eGFP#2	pG1_eGFP#8		
	t [h]	FL/r	t [h]	FL/r
[0371]	-1.7	176.77	-0.38	131.95
	0.0	166.52	0.00	108.76
	0.5	199.59	0.28	100.35
	1.0	195.94	0.62	121.36
	1.5	173.68	1.12	161.16
	2.0	219.00	1.62	162.69
	3.0	321.14	2.12	148.34
	7.0	494.60	3.12	205.20
	19.1	1150.96	7.12	373.08
[0372]	20.0	1000.37	19.70	1745.65
			21.12	1831.52

[0373] 表 7: 在优化的补料分批发酵中在 pGAP 或 pG1 的控制下表达 eGFP 的两种不同巴斯德毕赤酵母克隆物的相对荧光/生物反应器。

[0374] 实施例 3: 新鉴定的启动子在巴斯德毕赤酵母中的比较性启动子活性研究, 使用人血清清蛋白 (HSA) 作为胞外表达的报告基因

[0375] 为了分析新鉴定的启动子在葡萄糖限制条件下的特性,如下实施摇瓶筛选:用含有甘油作为碳源的丰富培养基完成24小时的预培养—模拟分批阶段(启动子的阻遏态),之后使用最小培养基和葡萄糖给料珠进行主要培养—模拟葡萄糖受限的补料分批阶段(启动子的诱导态)。

[0376] a) 菌株和表达载体

[0377] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec; 150 (4): 519-29),基于新霉素抗性选择阳性转化体。对于人血清清蛋白(HSA)的分泌性表达,使用其天然分泌前导物。

[0378] b) 新鉴定的启动子pG1、pG3、pG4和pG6到自有表达载体中的扩增和克隆

[0379] 将实施例2b中扩增的4种启动子克隆到pPUZZLE表达载体pPM1aZ10_HSA中,得到pPM1aZ10_pG1_HSA、pPM1aZ10_pG3_HSA、pPM1aZ10_pG4_HSA和pPM1aZ10_pG6_HSA。另外,将含有普遍使用的甘油醛3-磷酸脱氢酶启动子(pGAP)的载体pPM1aZ10_pGAP_HSA用作参照。使用ApaI和SbfI限制性位点,将启动子插入HSA基因起始密码子的上游(见表3)。通过Sanger测序验证启动子序列的正确性。

[0380] c) HSA在巴斯德毕赤酵母中在新鉴定的葡萄糖限制性诱导启动子控制下的表达

[0381] 将所有质粒使用AscI限制酶线性化,接着电穿孔(使用针对巴斯德毕赤酵母的标准转化方案)到巴斯德毕赤酵母中。对阳性转化体的选择在含有25 μ g/mL新霉素的YPD板(每升:10g酵母提取物,20g蛋白胨,20g葡萄糖,20g琼脂-琼脂)上进行。使用菌落PCR以确保如实施例2c中描述的转化质粒的存在。

[0382] 对于HSA表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,用OD600为1的预培养物接种YP培养基(每升:20g蛋白胨,10g酵母提取物)和葡萄糖给料珠(Kuhner, CH)中的主培养物。葡萄糖限制性生长条件由于这些给料珠的缓慢的葡萄糖释放动力学而实现,其由以下等式描述:(葡萄糖)=1.63*t^{0.74}[mg/Disc]。在预培养结束时,以及接种主培养物后24和48小时取样品。生物质浓度通过测量OD600或细胞湿重测定。培养上清液中的HSA浓度通过人白蛋白ELISA定量套装(目录号E80-129,Bethyl Laboratories, TX, USA)依照供应商说明书手册进行定量。HSA标准品以400ng mL⁻¹的起始浓度使用。将样品相应地在样品稀释剂(50mM Tris-HCl, 140mM NaCl, 1% (w/v) BSA, 0.05% (v/v) Tween20, pH 8.0)中稀释。从每种构建体数个克隆物的筛选中获得的HSA滴度呈现于表8。

[0383]

克隆	HSA滴度[mg L ⁻¹]	
	预培养	主培养48h
pGAP_HSA #1	6.9	9.0
pGAP_HSA #2	9.0	8.6
pGAP_HSA #3	6.6	9.2
pGAP_HSA #4	18.9	20.4
pGAP_HSA #5	9.6	8.3
pGAP_HSA #6	10.8	8.8
pG1_HSA #19	0.6	6.9
pG1_HSA #20	0.6	6.7
pG1_HSA #21	0.1	7.0
pG1_HSA #22	-	-
pG1_HSA #23	1.3	13.5
pG1_HSA #24	1.1	13.7
pG1_HSA #25	0.5	8.9
pG1_HSA #26	0.5	9.2

[0384]	pG1_HSA #27	0.6	7.3
	pG1_HSA #28	0.6	6.1
	pG1_HSA #29	0.6	6.4
	pG1_HSA #30	0.6	7.1
	pG6_HSA #31	0.3	1.8
	pG6_HSA #32	0.3	1.7
	pG6_HSA #33	0.3	2.0
	pG6_HSA #34	0.4	2.0
	pG6_HSA #35	0.2	2.2
	pG6_HSA #36	0.3	2.5
	pG6_HSA #37	0.3	2.3
	pG6_HSA #38	0.2	1.5
	pG6_HSA #39	0.7	-
	pG6_HSA #40	0.2	2.4
	pG6_HSA #41	0.4	-
	pG6_HSA #42	-	1.9

[0385] 表8:在pGAP、pG1和pG6控制下表达HSA的巴斯德毕赤酵母克隆物的筛选结果

[0386] d) HSA基因拷贝数的测定

[0387] 如实施例2d中的实施基因组DNA分离和qPCR测量,其使用表9中给出的引物。结果显示于表10。

引物	靶物	序列	产物长度
PpACT1_U p	Act	SEQ ID 30	148 bp
		CCTGAGGCTTGTTCCACCCATCT	
PpACT1_L ow	Act	SEQ ID 31	148 bp
		GGAACATAGTAGTACCAACCGGACATAAC GA	
PpHSA_Up	HSA	SEQ ID 32	135 bp
		AAACCTAGGAAAAGTGGGCAGCAAATG	
PpHSA_Lo w	HSA	T	135 bp
		SEQ ID 33	
		ACTCTGTCACTTACTGGCGTTCTCATG	

[0390] 表9:通过实时PCR测定基因拷贝数时用的引物

克隆	GCN	HSA mgL ⁻¹ 主培养物	HSA mgL ⁻¹ GCN主培养物	均值	STDEV
pGAP_HSA#3	1	9.2	9.2	9.22	0.95
pGAP_HSA#4	2	20.4	10.2		
pGAP_HSA#5	1	8.3	8.3		
pG1_HSA#20	1	6.6	6.6		
pG1_HSA#21	1	7.0	7.0	6.81	0.13
pG1_HSA#23	2	13.5	6.8		
pG1_HSA#24	2	13.7	6.8		
克隆	GCN	HSA mgL ⁻¹ 主培养物	HSA mgL ⁻¹ GCN主培养物	均值	STDEV
pG6_HSA#36	1	2.5	2.5	2.07	0.52
pG6_HSA#37	1	2.3	2.3		
pG6_HSA#38	1	1.5	1.5		

[0392] 表10:在pGAP、pG1和pG6控制下表达HSA的选定的巴斯德毕赤酵母克隆物的筛选和基因拷贝数结果;

[0393] e) 在pG1和pG6启动子控制下表达HSA的巴斯德毕赤酵母菌株的补料分批培养

[0394] 在DASGIP生物反应器中实施发酵,最终工作体积为0.7L。菌株pG1_HSA#23有两个HSA基因拷贝,菌株pG6_HSA#36仅携带一个HSA基因拷贝。因此,培养在pGAP控制下表达HSA的两种不同的巴斯德毕赤酵母菌株(pGAP_HSA#3具有一个HSA基因拷贝,pGAP_HSA#4具有两个HSA基因拷贝)作为参照。所有发酵均一式两份地进行。

[0395] 使用以下培养基:

[0396] PTM1痕量盐储液,每升含有

[0397] 6.0g CuSO₄ • 5H₂O,0.08g NaI,3.36g MnSO₄ • H₂O,0.2g Na₂MoO₄ • 2H₂O,0.02g H₃BO₃,0.82g CoCl₂,20.0g ZnCl₂,65.0g FeSO₄ • 7H₂O,0.2g生物素和5.0ml H₂SO₄(95% - 98%)。

[0398] 甘油分批培养基,每升含有

[0399] 39.2g甘油,27.9g H₃PO₄(85%),7.8g MgSO₄ • 7H₂O,2.6g KOH,9.5g K₂SO₄,0.6g CaSO₄ • 2H₂O,0.4mg生物素和4.6ml PTM1痕量盐储液。在无菌过滤到发酵罐中后,将pH调整为5.85。

[0400] 葡萄糖补料分批培养基,每升含有

[0401] 550g葡萄糖一水合物,6.5g MgSO₄ • 7H₂O,10g KC1,0.35g CaCl₂ • 2H₂O,0.4mg生物素和12ml PTM1痕量盐储液。

[0402] 将溶解氧用搅拌器速度(400-1200rpm)控制为D0=20%。通风速率为24L h⁻¹空气,温度控制为25°C,且添加NH₄OH(25%)以控制pH设定点为5.85。

[0403] 为了开始发酵,将400mL分批培养基无菌过滤到发酵罐中,并将巴斯德毕赤酵母克隆物以起始光密度(OD600)1进行接种(从预培养物)。约25h的分批阶段达到约20g/L的干生物质浓度,接着是使用葡萄糖培养基的恒定补料分批(达100小时),从而产生约100g/L的最终干生物质浓度。在分批阶段pH为5.85,并在整个发酵中保持为5.85。在分批期间和补料分批阶段取样品。HSA浓度使用人白蛋白ELISA定量套装(Bethy1,目录号E80-129)量化,如实施例3c中记载的。生物质浓度和HSA滴度显示于表11,分批(pG1和pG6的阻遏条件)结束时和补料分批(pG1和pG6的诱导条件)结束时的产物产率(分泌的HSA量/生物质,HSA/YDM)在表12给出。由此能验证诱导策略。pG1和pG6在碳源过剩(在甘油分批中)时受阻遏,其显示几乎没有可检测的HAS,与pGAP驱动的克隆相反。pG1和pG6的诱导发生在补料分批阶段开始时切换到C限制性的条件时。pG1的诱导(HSA/YDM)相比于阻遏态为超过120倍,pG6的诱导相比于阻遏态为超过20倍,而对于pGAP几乎没有观察到变化(相比于分批阶段,HSA/YDM有3倍增加)。

[0404]	CLONE	发酵#	分批结束			补料分批结束		
			时间 [h]	YDM [gL ⁻¹]	HSA滴度 [mgL ⁻¹]	时间 [h]	YDM [gL ⁻¹]	HSA滴度 [mgL ⁻¹]
	pG1#23	A041	-1.1	24.7	0.5	99.6	125.3	328.6
	pG1#23	A048	-0.3	23.9	0.5	108.4	128.6	277.6
[0405]	pG6#36	A045	-0.1	23.5	0.3	104.7	125.2	21.8
	pG6#36	A049	-0.3	24.4	0.3	108.4	129.0	26.9
	pGAP#4	A044	-0.1	23.6	11.2	104.7	129.1	141.4
	pGAP#4	A051	-0.9	24.1	9.0	96.9	118.2	114.9
	pGAP#3	A050	-0.9	24.2	5.0	96.9	117.7	57.8

[0406] 表11:在pGAP、pG1或pG6控制下表达HSA的巴斯德毕赤酵母克隆物的7种发酵在分批结束和补料分批结束时的酵母干重量和HSA滴度

[0407]	GCN HSA	分批结束时		补料分批结束时			倍数诱导
		均值	SD	均值	SD		
		HSA/YDM		HSA/YDM			
	pG1#23	2	0.02	0.00	2.39	0.33	121.06
	pG6#36	1	0.01	0.00	0.19	0.02	21.39
	pGAP#4	2	0.42	0.07	1.03	0.09	3.16
	pGAP#3	1	0.21		0.49		

[0408] 表12:在pGAP、pG1或pG6控制下表达HSA的巴斯德毕赤酵母克隆物的7种发酵在分批结束和补料分批结束时的HSA滴度对酵母干重量

[0409] 实施例4:新鉴定的启动子在巴斯德毕赤酵母中在各种葡萄糖浓度下的比较性启动子活性研究,使用eGFP作为胞内表达的报告基因

[0410] 为了分析新鉴定的启动子在各种葡萄糖浓度的特性,如下实施摇瓶筛选:用含有甘油作为碳源的丰富培养基完成24小时的预培养(启动子的阻遏态),之后使用最小培养基和葡萄糖作为碳源进行主要培养(启动子的诱导态);

[0411] a) 菌株和表达载体

[0412] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre,Centraalbureau voor Schimmelcultures,Utrecht,The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec;150(4):519-29),基于新霉素抗性选择阳性转化体。

[0413] b) 新鉴定的启动子pG1、pG3、pG4和pG6扩增和克隆到含有eGFP作为GOI的pPUZZLE表达载体中

[0414] 扩增和克隆如实施例2中描述的完成。

[0415] c) eGFP在巴斯德毕赤酵母中表达用于分析启动子活性

[0416] 转化和克隆选择如实施例2中描述的完成。

[0417] 对于表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,用OD600为0.01的预培养物接种10ml YP培养基(每升:20g蛋白胨,10g酵母提取物)中的主培养物并以葡萄糖作为碳源。以从20至0.001g L⁻¹的各种浓度使用葡萄糖。

[0418] 在主培养物接种1-8小时后取样品。eGFP表达通过流式细胞术如Stadlmayr等(J.Biotechnology 2010Dec;150(4):519-29)中描述的分析,基于新霉素抗性选择阳性转化体。对于每份样品,分析10,000个细胞。使用未转化的巴斯德毕赤酵母野生型细胞来测量巴斯德毕赤酵母的自身荧光。

[0419] 实施例5:新鉴定的启动子在巴斯德毕赤酵母中的比较性启动子活性研究,使用猪羧肽酶B(CpB)作为胞内表达的报告基因

[0420] 为了分析新鉴定的启动子在葡萄糖限制条件下的特性,如下实施摇瓶筛选:用含有甘油作为碳源的丰富培养基完成24小时的预培养—模拟分批阶段(启动子的阻遏态),之后使用最小培养基和葡萄糖给料珠进行主要培养—模拟葡萄糖受限的补料分批阶段(启动子的诱导态);

[0421] a) 菌株和表达载体

[0422] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre,Centraalbureau voor Schimmelcultures,Utrecht,The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec;150(4):519-29),基于新霉素抗性选择阳性转化体。对于猪羧肽酶B(CpB)的分泌性表达,使用酵母alpha交配因子前导物。

[0423] b) 新鉴定的启动子pG1、pG3、pG4和pG6扩增和克隆到自有表达载体中

[0424] 将在实施例2b中扩增的两种启动子克隆到pPUZZLE表达载体pPM1aZ30_aMF_CpB

中,得到pPM1aZ30_pG1_aMF_CpB和pPM1aZ30_pG6_aMF_CpB。另外,将含有普遍使用的甘油醛3-磷酸脱氢酶启动子(pGAP)的载体pPM1dZ30_pGAP_CPB用作参照。使用ApaI和SbfI限制性位点,将启动子插入CpB基因起始密码子的上游。通过Sanger测序验证启动子序列的正确性。

[0425] c) 在巴斯德毕赤酵母中在新鉴定的葡萄糖限制性诱导启动子控制下的CpB表达
[0426] 将质粒使用SpeI或SapI限制酶线性化,接着电穿孔(使用针对巴斯德毕赤酵母的标准转化方案)到巴斯德毕赤酵母中。对阳性转化体的选择在含有25 μ g/mL新霉素的YPD板(每升:10g酵母提取物,20g蛋白胨,20g葡萄糖,20g琼脂-琼脂)上进行。使用菌落PCR以确保转化质粒的存在,如实施例2c中描述的。

[0427] 对于CpB表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)中作为预培养。在约24h后,用OD600为1的预培养物接种YP培养基(每升:20g蛋白胨,10g酵母提取物)和葡萄糖给料珠(Kuhner, CH)中的主培养物。葡萄糖限制性生长条件由于这些给料珠的缓慢的葡萄糖释放动力学而实现,其由以下等式描述:(葡萄糖)=1.63*t^{0.74}[mg/Disc]。在预培养结束时,以及接种主培养物后24和48小时取样品。生物质浓度通过测量OD600或细胞湿重测定。培养上清中的CpB浓度通过酶测定法基于马尿酰-L-精氨酸通过CpB到马尿酸的转化来定量。通过在反应开始时使用Hitachi U-2910分光光度计监测25°C在254nm处的吸收来测量反应动力学。将样品和标准用测定缓冲液(25mM Tris, 100mM HC1, pH 7.65)缓冲并使用活化缓冲液(0.01mgL⁻¹胰蛋白酶,300mM Tris, 1 μ M ZnCl₂, pH 7.65)活化。使用没有胰蛋白酶的活化缓冲液代替样品作为阴性对照。反应通过添加底物溶液(测定缓冲液中1mM马尿酰-L-精氨酸)开始。

[0428] d) 在pG6启动子控制下表达CpB的巴斯德毕赤酵母菌株的补料分批培养
[0429] 补料分批发酵如实施例3e中描述的完成。克隆物pPM1aZ10_pG6_CpB#4在分批中未产生可检出的CpB,而在补料分批结束时产生超过210mg/L的CpB。

[0430] 实施例6:新鉴定的启动子pG1和pG6在巴斯德毕赤酵母多拷贝克隆物中的比较性启动子活性研究,使用人血清清蛋白(HSA)作为胞外表达的报告基因

[0431] 为了分析新鉴定的启动子在葡萄糖限制条件下的特性,如下实施摇瓶筛选:用含有甘油作为碳源的丰富培养基完成24小时的预培养—模拟分批阶段(启动子的阻遏态),之后使用最小培养基和葡萄糖给料珠进行主要培养—模拟葡萄糖受限的补料分批阶段(启动子的诱导态);

[0432] a) 菌株和表达载体

[0433] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures,Utrecht,The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec; 150 (4):519-29),基于新霉素抗性选择阳性转化体。对于人血清清蛋白(HSA)的分泌性表达,使用其天然分泌前导物。

[0434] b) 新鉴定的启动子pG1和pG6扩增和克隆到自有表达载体中

[0435] 将在实施例2b中扩增的两种启动子克隆到pPUZZLE表达载体pPM1nZ30_HSA中,得到pPM1nZ30_pG1_HSA和pPM1nZ30_pG6_HSA。

[0436] 使用ApaI和SbfI限制性位点,将启动子插入HSA基因起始密码子的上游。通过

Sanger测序验证启动子序列的正确性。

- [0437] c) 在巴斯德毕赤酵母中在新鉴定的葡萄糖限制性诱导启动子控制下的HSA表达
- [0438] 将所有质粒使用AscI限制酶线性化,接着电穿孔(使用针对巴斯德毕赤酵母的标准转化方案)到巴斯德毕赤酵母中。对阳性转化体的选择在含有 $25\mu\text{g}/\text{mL}$ 新霉素的YPD板(每升:10g酵母提取物,20g蛋白胨,20g葡萄糖,20g琼脂-琼脂)上进行。基因拷贝数扩增如记载于Marx等(FEMS Yeast Res. 2009 Dec; 9(8): 1260-70)中的完成。使用菌落PCR以确保转化质粒的存在,如实施例2c中描述的。
- [0439] 对于HSA表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,用OD600为1的预培养物接种YP培养基(每升:20g蛋白胨,10g酵母提取物)和葡萄糖给料珠(Kuhner, CH)中的主培养物。葡萄糖限制性生长条件由于这些给料珠的缓慢的葡萄糖释放动力学而实现,其由以下等式描述:(葡萄糖) = $1.63 * t^{0.74} [\text{mg}/\text{Disc}]$ 。在预培养结束时,以及接种主培养物后24和48小时取样品。生物质浓度通过测量OD600或细胞湿重测定。培养上清中的HSA浓度通过人白蛋白ELISA定量套装(目录号E80-129,Bethyl Laboratories, TX, USA)依照供应商说明书手册进行定量。HSA标准品以 400ng mL^{-1} 的起始浓度使用。将样品相应地在样品稀释剂(50mM Tris-HCl, 140mM NaCl, 1% (w/v) BSA, 0.05% (v/v) Tween20, pH 8.0)中稀释。从筛选来自实施例3c的几个多拷贝克隆物和单拷贝克隆物得到的HSA滴度呈现于表13。

[0440]

克隆	HSA滴度 (mg/L)
pPM1aZ10_pG1_HSA#23	8.20
pPM1nZ30_pG1_HSA#C2	19.55
pPM1nZ30_pG1_HSA#4*1000	21.59
pPM1nZ30_pG1_HSA#5*1000	21.33
pPM1nZ30_pG1_HSA#X4	27.22
pPM1nZ30_pG1_HSA#X5	6.90
pPM1aZ10_pG6_HSA#36	1.55
pPM1nZ30_pG6_HSA#C6	14.12
pPM1nZ30_pG6_HSA#2*1000	15.85
pPM1nZ30_pG6_HSA#X5	11.52
pPM1nZ30_pG6_HSA#X8	7.87

[0441] 表13:在pGAP、pG1和pG6控制下表达HSA的巴斯德毕赤酵母多拷贝克隆物的筛选结果

[0442] d) HSA基因拷贝数的测定

[0443] 如实施例2d中的实施基因组DNA分离和qPCR测量,其使用表9中给出的引物。结果显示于表14。

	克隆	GCN	HSA mgL ⁻¹ 主培养物	HSA mgL ⁻¹ GCN 主培养物
[0444]	pPM1nZ30_pG1_HSA#4*1000	11	21.59	1.89
	pPM1nZ30_pG1_HSA#X4	17	27.22	1.64
	pPM1nZ30_pG6_HSA#C6	12	14.12	1.23
	pPM1nZ30_pG6_HSA#2*1000	6	15.85	2.50

[0445] 表14:在pGAP、pG1和pG6控制下表达HSA的选定的巴斯德毕赤酵母多拷贝克隆物的筛选和基因拷贝数结果

[0446] e) 在pG1和pG6启动子控制下表达HSA的多拷贝巴斯德毕赤酵母菌株的补料分批培养

[0447] 补料分批发酵如实施例3e中描述的完成。克隆物pPM1nZ30_pG1_HSA#4*1000和pPM1nZ30_pG6_HSA#C6在补料分批结束时分别达到1060和728mg/L HSA。

[0448] 实施例7:新鉴定的启动子pG1在巴斯德毕赤酵母中的比较性启动子活性研究,使用抗体片段(Fab)作为胞外表达的报告基因

[0449] 为了分析新鉴定的启动子在葡萄糖限制条件下的特性,如下实施摇瓶筛选:用含有甘油作为碳源的丰富培养基完成24小时的预培养—模拟分批阶段(启动子的阻遏态),之后使用最小培养基和葡萄糖给料珠进行主要培养—模拟葡萄糖受限的补料分批阶段(启动子的诱导态);

[0450] a) 菌株和表达载体

[0451] 将巴斯德毕赤酵母野生型菌株(CBS2612,CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures,Utrecht,The Netherlands)用作宿主菌株。菌株的转化使用称为pPUZZLE的自有载体进行(Stadlmayr等J.Biotechnol 2010Dec; 150 (4) :519-29),基于新霉素抗性选择阳性转化体。对于抗体Fab片段的分泌性表达,使用酵母alpha交配因子前导物。

[0452] b) 新鉴定的启动子pG1到自有表达载体中的扩增和克隆

[0453] 将在实施例2b中扩增的pG1启动子克隆到含有Fab作为GOI的pPUZZLE表达载体中,如实施例5b描述的。使用ApaI和SbfI限制性位点,将启动子插入Fab基因起始密码子的上游。通过Sanger测序验证启动子序列的正确性。

[0454] c) 在巴斯德毕赤酵母中在新鉴定的葡萄糖限制性诱导启动子pG1控制下的Fab表达

[0455] 将质粒使用SpeI或SapI限制酶线性化,接着电穿孔(使用针对巴斯德毕赤酵母的标准转化方案)到巴斯德毕赤酵母中。对阳性转化体的选择在含有25μg/mL新霉素的YPD板(每升:10g酵母提取物,20g蛋白胨,20g葡萄糖,20g琼脂-琼脂)上进行。使用菌落PCR以确保转化质粒的存在,如实施例2c中描述的。

[0456] 对于Fab表达筛选,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,用OD600为1的预培养物接种YP培养基(每升:20g蛋白胨,10g酵母提取物)和葡萄糖给料珠(Kuhner,CH)中的主培养

物。葡萄糖限制性生长条件由于这些给料珠的缓慢的葡萄糖释放动力学从而实现,其由以下等式描述:(葡萄糖) = $1.63 \cdot t^{0.74}$ [mg/Disc]。在预培养结束时,以及接种主培养物后24和48小时取样品。生物质浓度通过测量OD600或细胞湿重测定。通过ELISA使用在山羊中生成的抗人Kappa轻链(结合和游离)-碱性磷酸酶抗体来对Fab表达水平定量。来自筛选在pGAP和pG1控制下的几个表达Fab的克隆物的Fab滴度呈现于表15。

[0457]

	Fab (mg/L)
pPM1dZ30_pGAP_Fab#2	0.00
pPM1dZ30_pGAP_Fab#5	0.70
pPM1dZ30_pGAP_Fab#7	0.68
pPM1aZ30_pG1_Fab#2	2.02
pPM1aZ30_pG1_Fab#3	0.70
pPM1aZ30_pG1_Fab#4	1.10
pPM1aZ30_pG1_Fab#5	0.00
pPM1aZ30_pG1_Fab#6	0.56
pPM1aZ30_pG1_Fab#9	0.66
pPM1aZ30_pG1_Fab#10	1.80
pPM1aZ30_pG1_Fab#11	1.64
pPM1aZ30_pG1_Fab#12	2.31
pPM1aZ30_pG1_Fab#13	2.35
pPM1aZ30_pG1_Fab#14	2.27
pPM1aZ30_pG1_Fab#15	1.60
pPM1aZ30_pG1_Fab#16	1.45
pPM1aZ30_pG1_Fab#B9	2.89
pPM1aZ30_pG1_Fab#B10	2.32
pPM1aZ30_pG1_Fab#B11	6.45
pPM1aZ30_pG1_Fab#B12	3.24

	Fab (mg/L)
[0458]	pPM1aZ30_pG1_Fab#B13 2.57
	pPM1aZ30_pG1_Fab#B14 3.14
	pPM1aZ30_pG1_Fab#B15 3.23
	pPM1aZ30_pG1_Fab#B16 2.61
	pPM1aZ30_pG1_Fab#C1 10.58
	pPM1aZ30_pG1_Fab#C2 1.46
	pPM1aZ30_pG1_Fab#C3 12.38
	pPM1aZ30_pG1_Fab#C4 9.91
	pPM1aZ30_pG1_Fab#C5 1.96
	pPM1aZ30_pG1_Fab#C6 2.87
[0459]	pPM1aZ30_pG1_Fab#C7 7.03
	pPM1aZ30_pG1_Fab#C8 6.37

[0459] 表15:在pGAP和pG1控制下表达Fab的巴斯德毕赤酵母克隆物的筛选结果

[0460] d) 在pG1启动子控制下表达Fab的巴斯德毕赤酵母菌株的补料分批培养

[0461] 类似于实施例3e中描述的完成补料分批发酵,但使用如实施例2e中描述的葡萄糖补料分批。在补料分批结束时,克隆物pPM1aZ30_pG1_Fab#C4和pPM1aZ30_pG1_Fab#C7分别达到165和131mg/L Fab。

[0462] 实施例8:指数性补料分批发酵以控制新鉴定启动子在最大体积计量生产力处的比生长速率

[0463] 利用在新鉴定的启动子控制下表达报告基因的巴斯德毕赤酵母克隆物的恒化器培养来测定在不同生长速率的特定的和体积计量性的生产力。如Maurer等(Microb Cell Fact. 2006 Dec 11; 5:37)描述的,指数性补料分批发酵可用于以比生长速率生长巴斯德毕赤酵母克隆物以用于在整个给料阶段的改进的生产。由此,可优化空间-时间产率。应用优化的给料,补料分批阶段的空间-时间产率改进了超过35%。

[0464] 实施例9:测定启动子/转录强度:鉴定在不同葡萄糖浓度上启动子调控的比较性启动子活性研究,使用eGFP作为胞内表达的报告基因

[0465] 通过筛选在所述启动子控制下表达eGFP的克隆物来分析启动子的调控特性。因此,将单个菌落接种到液体YPG-Zeo培养基(每升:20g蛋白胨,10g酵母提取物,12.6g甘油和25mg新霉素)作为预培养。在约24h后,用OD600为0.01的预培养物接种10ml YP培养基(每升:20g蛋白胨,10g酵母提取物)和不同浓度的葡萄糖(20,10,5,2.5,1.25,0.625,0.313,0.156,0.078,0.039,0.020,0.010,0.005和0.002g/L)中的主培养物。在6小时后取样品并通过流式细胞术如Stadlmayr等(J Biotechnol. 2010 Dec; 150(4):519-29)描述的分析。对

每个细胞/数据点计算与细胞大小关联的荧光(前向散射乘1.5次方(forward scatter to the power of 1.5))并将其几何学均值用于比较在不同葡萄糖浓度中产生的eGFP表达水平。将在pGAP控制下表达eGFP的克隆物用作参照(巴斯德毕赤酵母的pGAP,本文中SEQ ID 25)。使用未转化的巴斯德毕赤酵母野生型细胞测量巴斯德毕赤酵母的自身荧光,并从信号减去。表16显示在约40mg/L葡萄糖或更低对pG1启动子的完全诱导,并将其转录强度与天然pGAP启动子比较。

[0466]	% pGAP	葡萄糖 (g/L)
14.7	20	
17.4	10	
23.7	5	
25.4	2.5	
28.2	1.25	
30.6	0.625	
36.9	0.3125	
44.5	0.15625	
50.9	0.078125	
56.2	0.0390625	
55.0	0.0195313	
57.5	0.0097656	
59.2	0.0048828	
59.6	0.0024414	

[0467] 表16:在pG1启动子控制下表达eGFP的巴斯德毕赤酵母克隆物在不同葡萄糖浓度(20-0.002g/L)的相对eGFP表达(相对于pGAP)

[0469] 实施例10:在葡萄糖浓度筛选测定法中现有技术的pICL1和pMLS1启动子与pG1的比较

[0470] 依照实施例9实施比较性启动子活性研究,其采用pICL1和pMLS1启动子作为参照以与依照本发明的pG1启动子比较。

[0471] 发现pICL1和pMLS1启动子两者的活性都非常弱,在高(D20:20g/L/阻遏)或低(D0.04:0.04g/L诱导=去阻遏)葡萄糖浓度没有显著差异。在任一情况下,活性都远低于在

相同背景中阻遏的pG1启动子的活性。结果显示于表17,其为相对于pGAP启动子的启动子活性%。

	D20/阻遏	D0.04/诱导
pG1#8	9.95 +/- 2.60	48.41 +/- 2.76
pICL1	2.68 +/- 1.78	5.07 +/- 0.90
pMLS1	-1.26 +/- 0.54	0.58 +/- 0.22

[0473] 表17:分别在pG1、pICL1和pMLS1控制下表达eGFP的菌株的相对荧光,所述菌株在含20g/L(D20)或0.04g/L(D0.04)葡萄糖的培养基中生长。

[0474] 实施例11:pG1变体的比较

[0475] 如实施例2a中描述的克隆pG1启动子的更短的变体,并类似于实施例2c中描述的进行筛选,只不过在使用24孔板(Whatman, UK, Art. Nr. 7701-5110)和给料珠1/4(12mm, Kuhner, CH)而非整个的规模缩减的装置中。将在pG1和pGAP控制下表达的克隆物用作对照。pG1及其变体的前向引物和长度列于表18。在pG1和pG1变体a-f控制下表达eGFP的细胞的相对荧光中没有显著差异。

启动子	引物	5'	3'	长度(bp)
pG1	GATAGGGCCCCAACACATTGCTCCCCCTAGTCTC SEQ ID 34	36	1001	988
pG1 a	GATAGGGCCCGGAATCTGTATTGTTAGAAAGAACG AGAG SEQ ID 35	143	1001	881
pG1 b	GATAGGGCCCCATATTCACTAGGTGTTCTTGCAC SEQ ID 36	338	1001	686
pG1 c	GATAGGGCCCCTGCAGATAGACTTCAAGATCTCAGG SEQ ID 37	509	1001	515
pG1 d	GATAGGGCCCGACCCGTTTCGTGACAAATT SEQ ID 38	632	1001	392
pG1 e	GATAGGGCCCCGGATAAGAGAATTGTTGATTAT SEQ ID 39	674	1001	350
pG1 f	GATAGGGCCCGCCTGCTCCATATTTTCCGG SEQ ID 40	719	1001	305

[0477] 表18:pG1及其变体:前向引物和pG1序列(SEQ ID 1)中的5'开始和3'端位置。pG1a-f的序列见图15(SEQ ID 41-46)。

[0478] 本发明提供以下内容:

[0479] 1. 一种通过培养包含表达构建体的重组真核细胞系来产生感兴趣蛋白质(POI)的方法,所述表达载体包含可调控的启动子和在所述启动子的转录调控下的编码POI的核酸分子,所述方法包括以下步骤:

[0480] a) 用阻遏所述启动子的基础碳源培养所述细胞系,

[0481] b) 用使所述启动子去阻遏的无补充碳源的条件或有有限量补充碳源的条件培养所述细胞系,从而诱导以相比于所述细胞的天然pGAP启动子至少15%的转录速率产生所述POI,并

[0482] c) 产生并回收所述POI。

[0483] 2. 依照项1的方法,其中所述基础碳源选自下组:葡萄糖、甘油、乙醇、其混合物、和复合营养材料。

[0484] 3. 依照项1或2的方法,其中所述补充碳源是诸如葡萄糖、果糖、半乳糖或甘露糖等己糖,诸如蔗糖等二糖,诸如甘油或乙醇等醇,或其混合物。

[0485] 4. 依照项1至3中任一项的方法,其中所述步骤b)采用不提供或提供有限量的所述补充碳源的补料培养基,优选所述有限量为所述培养基中0-1g/L。

[0486] 5. 依照项1至4中任一项的方法,其中所述有限量的补充碳源为生长限制性的,以保持比生长速率在 0.02h^{-1} 至 0.2h^{-1} ,优选 0.02h^{-1} 至 0.15h^{-1} 内。

[0487] 6. 依照项1至5中任一项的方法,其中所述启动子能控制野生型真核细胞中基因的转录,该基因选自下组:G1 (SEQ ID 7)、G3 (SEQ ID 8)、G4 (SEQ ID 9)、G6 (SEQ ID 10)、G7 (SEQ ID 11) 和 G8 (SEQ ID 12)、或其功能活性变体。

[0488] 7. 依照项6的方法,其中所述功能活性变体选自下组:具有至少约60%核苷酸序列同一性的同源物,通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物,优选具有至少200bp的核苷酸序列,以及源自除巴斯德毕赤酵母(*Pichia pastoris*)以外的物种的类似物。

[0489] 8. 依照项6或7的方法,其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和 pG1f (SEQ ID 46)。

[0490] 9. 依照项1至8中任一项的方法,其中所述细胞系选自下组:哺乳动物、昆虫、酵母、丝状真菌和植物细胞系,优选为酵母。

[0491] 10. 依照项1至9中任一项的方法,其中所述POI为异源蛋白质,优选选自治疗性蛋白质包括抗体或其片段、酶和肽、蛋白质抗生素、毒素融合蛋白、碳水化合物-蛋白质偶联物、结构蛋白、调节蛋白、疫苗和疫苗样蛋白或颗粒、加工酶、生长因子、激素和细胞因子、或POI的代谢物。

[0492] 11. 用于调控POI在重组真核细胞中在碳源可调控启动子的转录控制下表达的方法,所述启动子相比于所述细胞的天然pGAP启动子具有至少15%的转录强度,其中所述表达在限制所述碳源的条件下诱导。

[0493] 12. 在重组真核细胞中在碳源可调控启动子的转录控制下产生POI的方法,其中所述启动子相比于所述细胞的天然pGAP启动子具有至少15%的转录强度。

[0494] 13. 依照项1至12中任一项的方法,其中所述可调控启动子包含选自下组的核酸序列:

[0495] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；

[0496] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；

[0497] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；和

[0498] d) 源自a)、b)或c)的片段或变体，

[0499] 其中所述启动子是功能活性的启动子,其为碳源可调控的启动子,所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达所述POI。

[0500] 14. 依照项13的方法,其中所述pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG4 (SEQ ID 4)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)的变体是选自下组的功能活性变体:具有至少约60%核苷酸序列同一性的同源物,通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物,优选具有至少200bp的核苷酸序列,以及源自巴斯德毕赤酵母以外的物种的类似物。

[0501] 15. 依照项13或14的方法,其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和pG1f (SEQ ID 46)。

[0502] 16. 一种分离的核酸,其包含选自下组的核酸序列:

[0503] a) pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；

[0504] b) 与以下具有至少60%同源性的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；

[0505] c) 在严格条件下与以下杂交的序列:pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5)、或pG8 (SEQ ID 6)；和

[0506] d) 源自a)、b)或c)的片段或变体，

[0507] 其中所述核酸包含功能活性的启动子,其为碳源可调控的启动子,所述启动子在重组真核细胞中能以相比于所述细胞的天然pGAP启动子至少15%的转录速率表达POI。

[0508] 17. 依照项16的核酸,其中所述pG1 (SEQ ID 1)、pG3 (SEQ ID 2)、pG6 (SEQ ID 3)、pG7 (SEQ ID 5) 或pG8 (SEQ ID 6)的变体是选自下组的功能活性变体:具有至少约60%核苷酸序列同一性的同源物,通过在亲本核苷酸序列内或序列远端之一或两端处插入、缺失或取代一个或多个核苷酸而修饰得到的同源物,优选具有至少200bp的核苷酸序列,以及源自巴斯德毕赤酵母以外的物种的类似物。

[0509] 18. 依照项16或17的核酸,其中所述pG1的功能活性变体选自下组:pG1a (SEQ ID 41)、pG1b (SEQ ID 42)、pG1c (SEQ ID 43)、pG1d (SEQ ID 44)、pG1e (SEQ ID 45) 和pG1f (SEQ ID 46)。

[0510] 19. 一种表达构建体,其包含依照项16至18的核酸可操作地连接于在所述启动子转录控制下的编码POI的核苷酸序列,所述核酸与编码所述POI的核苷酸序列不天然关联。

[0511] 20. 一种重组真核细胞,其包含项19的构建体。

- [0512] 21. 从真核细胞鉴定出碳源可调控启动子的方法,包括以下步骤:
- [0513] a) 在存在碳源的情况下,在细胞生长条件下于分批培养中培养真核细胞,
- [0514] b) 进一步在存在有限量的补充碳源的情况下,在补料分批培养中培养所述细胞,
- [0515] c) 提供步骤a) 和b) 的细胞培养物的样品,并
- [0516] d) 在所述样品中实施转录分析以鉴定出在步骤b) 的细胞中比在步骤a) 的细胞中显示更高转录强度的可调控启动子。

- [0001] 序列表
[0002] <110> 隆扎有限公司(Lonza Ltd.)
[0003] <120> 可调控启动子
[0004] <130> L0001P
[0005] <160> 46
[0006] <170> PatentIn version 3.3
[0007] <210> 1
[0008] <211> 1001
[0009] <212> DNA
[0010] <213> 巴斯德毕赤酵母(Pichia pastoris)
[0011] <400> 1
[0012] atttccaccc ccatcccagt agaatgtagg gtccccaaac atttgctccc cctagtctcc 60
[0013] agggaaatgt aaaatatact gctaataagaa aacagtaaga cgctcagttg tcaggataat 120
[0014] tacgttcgac tgttagaaaa caggaatctg tattgttaga aagaacgaga gtttttacg 180
[0015] gcgcgcgccat attggccgt gtgaaaacag cttgaaaccc cactacttc aaaggttctg 240
[0016] ttgctataca cgaaccatgt ttaaccaacc tcgcctttaa cttgactgaa gtcatcggtt 300
[0017] aacaatcaag taccctagtc tgtctgaatg ctccttcca tattcagtag gtgtttcttg 360
[0018] cactttgca tgcactgcgg aagaattagc caatagcgcg tttcatatgc gcttttaccc 420
[0019] cctctttgtt caagcgaaaa atgcctgtaa gatttgggtgg gggtgtgagc cgtagctga 480
[0020] agtacaacag gctaattccc tgaaaaaaact gcagatagac ttcaagatct cagggattcc 540
[0021] cactatttgg tattctgata tgttttcct gatatgcata aaaactctaa tctaaaacctt 600
[0022] gaatctccgc tattttttt tttttttga tgaccccggtt ttcgtgacaa attaattcc 660
[0023] aacgggggtct tgtccggata agagaattttt gtttGattat ccgttcggat aaatggacgc 720
[0024] ctgctccata ttttccgggtt tattacccca cctggaaatgt cccagaattt tccggggattt 780
[0025] acggataata cggtgtctg gattaattaa tacgccaatgt cttacatccc gttcagttt 840
[0026] cgtgcgagta tgtcaataaa taaacaagat gagccaattt attggattttt ttgcagttt 900
[0027] accccgcacat agcttaggcat agccaagtgc tatgggtgtt agatgatgca cttggatgca 960
[0028] gtgagtttg gagtataaaa gatccttaaa attccaccct t 1001
[0029] <210> 2
[0030] <211> 1000
[0031] <212> DNA
[0032] <213> 巴斯德毕赤酵母(Pichia pastoris)
[0033] <400> 2
[0034] gtaaatagcg gcagcaatcc agtaaccttt tctgaatagc agaggcttaa ctaaaataat 60
[0035] gcccagggtt aaaaattcga aatttgcac caaaaataaa gacttgcgt tataagtctt 120
[0036] aacaaagtcc gcaattttgg agctaacggt ggcgggtgtt gggatattca ataatggtag 180
[0037] aatgttgctg cgggttatatg acagagcgtg aaacacactg aacaaggtaa atggaacaac 240
[0038] agcaattgca atatggggga ggatagtcaa gaacaaagca gcaatggcaa agtactgaat 300
[0039] attctccaaa gccaaaagggt ccagtggttt caacgacaaa gtctgtgg tatagctttg 360
[0040] gaacaaaagg acaccgaaaag actcgacagc gcccacaaat acagcgttgc agaagaacgca 420
[0041] attgattgct ccagagcttc taatagtcag aagatacccc aaaccccgaa gcaacgttag 480

[0042] cacatgacct aagaaccagg cgaagtgaag agtctggaaat aacgacaccc agtcagttt 540
 [0043] tcctgagctc ctgggtggat tggtagaagc atttgatttg ctggagtggtt ttttatttga 600
 [0044] agatgggttt gaagccattt ttgctaaaga gtcggagttt tgcttttagg gtttgttaag 660
 [0045] caaaggagga aaaactgcgc cgttgaagt cccaggttgtt ttcgcgtgtg aggccagcca 720
 [0046] gggaaagctt cttcggtac tttttttct tttgcagggtt ccggacggat taagcttcgg 780
 [0047] gtatgaggg gggcggttagc caattccgga cacaatattt cgtcgacgt agtcaccccg 840
 [0048] ccataaaatat acgcaggatt gaggtataaa catcgatagt ctttagtaatt aatacaattc 900
 [0049] agtggcgaat ttggcaacat gacgtaaggc ccactgttgtt ctataaaagg ggatgaattt 960
 [0050] tcattttttt gaggcctccc ggacaatttta ttgaactcaa 1000
 [0051] <210> 3
 [0052] <211> 1001
 [0053] <212> DNA
 [0054] <213> 巴斯德毕赤酵母 (*Pichia pastoris*)
 [0055] <400> 3
 [0056] agaccagcag tttaactacg caaatccaca ggaatttcta catcacaata ccaatggtaa 60
 [0057] taccacgacg tcaaggaatg gaaacgcacg cttggaggaa gacttcgtca acctcttcg 120
 [0058] gagtaaccgaa ggctaagaca ataagaagaa aaaaaaaaaa aaagcggtgg gggagggatt 180
 [0059] attaaataag gattatgtaa ccccagggtt ccgttctata catatttaag gattattttag 240
 [0060] gacaatcgat gaaatcgca tcaaactgga tgggagtttata gtgtccggat aatcgataaa 300
 [0061] atcatcttcg gaggagccgc ttgggtgggtt ggtgagagga gtgaaatatg tgtctcctca 360
 [0062] cccaagaatc gcgatatcag caccctgtgg gggacactat tggcctccct cccaaacctt 420
 [0063] cgatgtggta gtgttttattt atattgatta cattgattttt atagctaaac cctgcctgg 480
 [0064] tgcaagttga gctccgaattt ccaatattttttaaaatgcct gcaagataac ctcggatgg 540
 [0065] cgtccgaccc cgcttaattt ttttaactcc tttccaacga ggacttcgtt atttttgattt 600
 [0066] agggagttga gaaacgggggtt gtcttgatac ctcctcgatt tcagatccca cccccctctca 660
 [0067] gtcccaagtg ggaccccccctt cggccgtgaa atgcgcgcac ttttagttttt ttcgcattgt 720
 [0068] aacgcgggtt tccgtcaattt aaaagtgcac gactagggtt aactttacca tttttgtcgc 780
 [0069] actccgttcc ctcggaaatag ggggttagta attctgcgtt agtgcattttt ttaccccgcc 840
 [0070] aagggggggc gaaaagagac gacccatca cgcattctcc agtcgttcc tacgcctaca 900
 [0071] gcaccgacgt agttaactttt ctccatata taaagcaattt gccattcccc taaaaactttt 960
 [0072] aacctctgct ttttcttgat tttccttgc ccaaagaaaa g 1001
 [0073] <210> 4
 [0074] <211> 1000
 [0075] <212> DNA
 [0076] <213> 巴斯德毕赤酵母 (*Pichia pastoris*)
 [0077] <400> 4
 [0078] tgactgttc aatttgaagt cgatgttgc gatgtcaaga gagatgttca attatattttt 60
 [0079] tcatttgctt gttacactgg aaacgctact tttgttggcg gaaactctac cagtttggcc 120
 [0080] gtccatgttta acgtatgttgc tctggccgtt gaccgtttca acacgaacat aaccaatgtt 180
 [0081] aaatccactt acaggtctttt ttcataatggta ggcaattttt accttacttc tttggatgtt 240
 [0082] ccaagtgggg cttaaacgtt tggactaactt aatgttgcgtt ttgttacttcc aactccgag 300
 [0083] gtaaataaaag gattttttttt ggattcttcc aagtttgggtt ggaagttgtt acaggtttat 360

[0084] aagcatatcg tgcgcttgc cacaattgaa tcatttattt ttgcgagata catgaacaaa 420
 [0085] gtgtgaactg ggacccatta ctacaattcc cacgcaaccg ttgttcaaa gccccatattt 480
 [0086] tttgacaatt gtttcgttac acccccagg ttatgtacat cgcttgcaat gatgtgtgtc 540
 [0087] ccggaggatt ttccatattc agcttgaatt cgtatactca accaatatct gggggataac 600
 [0088] ttttatgtaa cctatacaaa tcaactatac tatttcacct ttgcaccaat catctccat 660
 [0089] ctgttaagt ttgccttct atatccctga ccctgacatc acccatgatt ccgctcaacg 720
 [0090] gtttcctct acatcgccc tctttggag agggtgtca gtttgacatt caaatttaccc 780
 [0091] cccgccatca cgcgcaaccg agaccgcacc cccgaattt cacaattac cccacaccct 840
 [0092] atactccacc actatgaggg ttattagaac tgatcacgta taaataccac cgcaagttcc 900
 [0093] caagggatcg tgttcttctt ctccaattgc aatcatattt ctgactctt ctagttcaga 960
 [0094] ttaattcctt tacacttgc ttttccctt acctttatcc 1000
 [0095] <210> 5
 [0096] <211> 1000
 [0097] <212> DNA
 [0098] <213> 巴斯德毕赤酵母 (Pichia pastoris)
 [0099] <400> 5
 [0100] aattgattaa gttcagtgaa attcaaacc gctatacaca acaggacaac tttgagttt 60
 [0101] gaaaaatccg atgttagtgta acggctagca cggccgcgtt tcaccggca gaccgggtt 120
 [0102] cgactcccg catcgaggtt tattttcca ttgcgttctt tagagtattc tcctcagcat 180
 [0103] gccccctga attttcctt tttccatgt gtccatttt tccactttt ttacagttt 240
 [0104] cctcgtatg ttaattggct acacaaaagc tgccacacga aacctaattc acgaaaaact 300
 [0105] atacagcctt cactaatccg tagcccccata atatgttgc cacgtgctgt tgggtactac 360
 [0106] ctgtagactc tcataccca ctccgtctt ctccaacaat taacgcagta ccgagattt 420
 [0107] tcagcagact caaattggc aaactctgta ttttcctt cccgcataat ttatgggtct 480
 [0108] caggcctcca cgttccgtt ttacttgaag aatattggct gcggaaaaag tggtaaggac 540
 [0109] aaccccttt taattggatc cagttttcc gaaatgttcc gatccgtacg tcatctccga 600
 [0110] agccgtacat ttcaactcaa tctacgtacg ttggactca gcgccttgg aattgccagg 660
 [0111] acagttact tgagttgata ttcccttgc gattgtgtc ttcttttcc aaaattttag 720
 [0112] gcttcgtttt aaaagtggaa tctggcgct agatcactt atgcctattt ttcacggaaa 780
 [0113] aataagtggt actatgcacc ccttaaacct aaagaaaaac ggaaaaattt ccccaaaacc 840
 [0114] tggtgatgtt ttgccttctt ttcttttcc ccgagtttt ctttttctt gtctgcaaaa 900
 [0115] ttccctctt gaccttagcg tccccggaaa aaattaacta cttaaggacc gaatgagccc 960
 [0116] cagctttcc ccttcttttcc attattcccc ataataataat 1000
 [0117] <210> 6
 [0118] <211> 1000
 [0119] <212> DNA
 [0120] <213> 巴斯德毕赤酵母 (Pichia pastoris)
 [0121] <400> 6
 [0122] ctgcacaacc attgccagta aggacgaaga gaaggccccaa ctacccaaaa ttccaggataa 60
 [0123] cgtcttcata ccatgcacgg acgcctacaa gacgctgtca agacatgcca acttcaacga 120
 [0124] agtgaacttt aacacattga tcggaaatt gaccaccaag ggaatgctgg ttgaggctgg 180
 [0125] aagcggtgcc agtgcctgta gggaaactgga ccgaaagttt agtaatgcat aagaggatat 240

[0126] atataggaat gcagtaataa tattagtacc cattaagtgg gctaaggccat tggaaggccg 300
 [0127] tctgactgat ggtgggttgc ttctcattta gatagtgcac ttgcaactac cgtctgagat 360
 [0128] ttagtttgat gtgaagctcc aggcacaaaa cagtataaga accttatctc cgcattattg 420
 [0129] ttcttcgtaa aaagtttgtg tgaagaaaca gggtagttg cgccatgttgc ttgtaatatg 480
 [0130] cgcataggat gggcatttgc cttcttcctt cggaaagagcc acaccgttag ctaaaaaagg 540
 [0131] acgcgcacatc accccaaaat agaatgtggg gaaataggac ggcacacttc ctctcaatca 600
 [0132] ctggacgtca gaaaaacaaa tgcgcattcg agtcaccctc cgtgataccc tccgtgatac 660
 [0133] cccctctccg tctattctga cagcgtctcc ccatgacgtt tcaatctact tagaaaaatg 720
 [0134] ttctttttt ttcccttcaa ttacacgtac tcattttctg caagggtctg gaggacatca 780
 [0135] ccaatctgcg actccataac ttagtcctga gtttatattt acgcttcac tgcgtgatag 840
 [0136] gaagaaaaag ttccacgaaa ttccccggcc aacttgcct tcggataag cagccactct 900
 [0137] cttctgccc atagtaagct tgcgcgaggc cccaaacttgg ccagaaactt taaatatgcc 960
 [0138] aaacaatctc ccccaatcta agttccctt cttctaaaaa 1000
 [0139] <210> 7
 [0140] <211> 1662
 [0141] <212> DNA
 [0142] <213> 巴斯德毕赤酵母 (*Pichia pastoris*)
 [0143] <400> 7
 [0144] atgtcctcgat ttttctaaa caaccaaaca gtaaagatga tgacgcctt ggaaagagct 60
 [0145] agagctctaa attttcaggc caaagtttat gataaaatttca caaaaactta caatatctac 120
 [0146] gctattgcaaa taacagccac cgtttctggat ctgtatgttgc gttttgatatt ttcttctgtg 180
 [0147] tcctcgatcg taagtcagga tcattacaga aactacttca accgtcccga cagtttgacg 240
 [0148] caaggggtta tcacccgcaag tatggctggat ggttcttctt tgggttcgtt attttcttct 300
 [0149] gacttccagg atatcttgg aagaagagtt gctctgcata tgtgcagtgt cctctggatt 360
 [0150] atcggggcca ttcttcaatg cgctgcacaa aaccaaggta tgctgatcgc agggagattt 420
 [0151] atttccggta tcgggtcgat gttgggtca gcttcagctc cagtttatttgc ttctgaagtt 480
 [0152] gctccagccaa agatttagagg aatgattggat ggatttatttca aattttctgt cactgtgggt 540
 [0153] atcatgataa tggatattat cggatattggat tgcactaca ttgacggcgt tgcatcattt 600
 [0154] agactggcctt ggggtttgc aatgggttcca ggtcttatttgc tttgggtcgat tttttttt 660
 [0155] ctccctgatcgat ctccaaatg gctggctaac cacaaccgtt gggaaagacgc agttgagttt 720
 [0156] attgctaatttgc ttgttgcataa aggtgacaga gaaaacggccg atgtgcgttgc gcaattggat 780
 [0157] gaagttcagg agcaactattt gattgacaaa gatgttttgc attttggat ctttgcatttgc 840
 [0158] tttaagaaag attgtatcaa acgtacccatc attggagtttgc cagctcaatgtt gttggcaacaa 900
 [0159] ctttgcgttgc ttaatgttgc aatgtactac gttgtgttgc tcttccaaat ggctgggtttt 960
 [0160] actggaaatg tggcggttgc atgcctca attcaatatttgc ttttgcatttgc ttttttttgc 1020
 [0161] gttccagctt tggatattttgc ggaccgtata ggcagacgc ccctactaat tgggtgggttgc 1080
 [0162] attttcatgtt gtattttggat gtttggatgttgc gcaatttttttgc ttttttttgc ttttttttgc 1140
 [0163] attgaaaattt tcagcggttgc tgatgttgc agaattacta ttccgttgc gcaacggcttgc 1200
 [0164] gcaacggccg gtttttttgc ctggatgttgc gcaatttttttgc ttttttttgc ttttttttgc 1260
 [0165] ggtatctgc tttgggttgc tggatgttgc attttcaacaa acagacaaaag agcaaaaggaa 1320
 [0166] gcaacggccg ctggatgttgc taactggatttgc ttcaactttgc cttggctat gttcgatgttgc 1380
 [0167] tcagccttgc gaaacatttgc atgaaagact tacatcattt ttggatgttgc ttttttttgc 1440

[0210] attctcagt gtttcggtaa cccagctgga gctgccatcg ctgctcatcc cagaatcaag 720
 [0211] aagattgctt tcaccggatc cactgcaaca ggccgtaaga tcatggaagc agccgctaaa 780
 [0212] tctaacctga aaaaagtcac tttgaaacta ggtggtaaat ctccaaacat tgtgttgaa 840
 [0213] gatgctgata tccagaagac tatccataac attattttgg gaatttctt caattctgg 900
 [0214] gaagtctgtt gtgcagggttc cagagtctac attcaagaca ctgtgtatga agaagtgc 960
 [0215] gaagccttca agaaggagac tgataacgtt aaggttggg gaccattcga agaagggtgc 1020
 [0216] ttccaagggc ctcagacctc tgagttgcaa cttAACAGAA tccttagtta catcaaacac 1080
 [0217] ggttaaggatg aaggtgcgtg tgtaattacc ggtggtaaa gataccgtaa ccgaggttac 1140
 [0218] tacattaagc ccacaatttt tgctgacgtt actgaagaca tgaagattgt caaggaggag 1200
 [0219] attttggtc ctgtggttac tatcactaag ttctctaccg tggatgaggt tggatggat 1260
 [0220] gccaacaaca ccaactatgg tctagctgt ggtattcaca caaacaactt gaacaaagcc 1320
 [0221] attgatgtt ccagtagaat caaggcgggt gtcgtttgga ttaacaccta caacgatttc 1380
 [0222] caccacatgg ttcccttcgg aggttatgga gaatctggta ttggcagaga gcttgggtgc 1440
 [0223] gaggcttgg ataactacac tcaagccaag gctatcagaa ttgcttacac tcctgaacat 1500
 [0224] aagtag 1506
 [0225] <210> 10
 [0226] <211> 1578
 [0227] <212> DNA
 [0228] <213> 巴斯德毕赤酵母(*Pichia pastoris*)
 [0229] <400> 10
 [0230] atgcttagaa cttctccagc tactaagaaa gctctcaagt cgccgattaa cgccttcaac 60
 [0231] gttgctgcct tgagattcta ctccctattc cctttgcagg ttccaaattac cttggccaaac 120
 [0232] ggttaagacct acaatcagcc aacagggttg tttatcaaca atgagttcgt tccttctaag 180
 [0233] caaggtaaga cctttgcgtt tttaaaccct tccactgagg aggagattac tcacgtctac 240
 [0234] gagtccagag aggacgcacgt tgagtttagcc gttgcagccg ctcaaaaggc tttcgactca 300
 [0235] acctggtcca cccaggaccc tgctgagaga ggttaaggct tgaacaagtt ggctgacctg 360
 [0236] atcgaggagc actctgagac ccttgcgcct atcgagtcct tggacaacgg taaggccatt 420
 [0237] tcctccgcta gaggtgatgt tggctgggt gtcgcctact tgaagtcctg tgccgggttgg 480
 [0238] gccgacaagg tttcggtag agttgtgaa accggaagct cccacttcaa ctacgttaga 540
 [0239] agagagccat tgggttttg tggcagatt atcccatgga actttcctt tctgatgtgg 600
 [0240] tcctggaaag ttggccagc tttggccact ggttaacactg ttgtcctgaa gacagccgag 660
 [0241] tctactcctc tgcgcgcct gtacgttcc caattggta aggaggccgg tatcccagct 720
 [0242] ggtgtccaca acattgtgtc cggtttcggt aagattactg gtgaagctat tgctactcat 780
 [0243] cctaagatca agaagggtgc cttcactggt tctaccgcca ctggcgtca catcatgaag 840
 [0244] gctgctcccg aatccaactt gaagaagggtt actttggagt tgggtggtaa atctccataac 900
 [0245] atcgtgttca acgatgctaa cattaagcaa gctgtcgcca acatcatcct cggtattttac 960
 [0246] tacaactctg gagaaggttt ttgtgctggt tccagagttt atgttcaatc cggtattttac 1020
 [0247] gacgagctt tggccgaatt caagactgct gctgagaatg tcaaggttgg taaccattc 1080
 [0248] gacgaggaca cttccaaagg tgctcaaacc tctcagcaac aattggagaa gattttgggt 1140
 [0249] ttgcgttggc gtggtaagaa ggacgggtgc actttgatca ctgggtggc cagatttaggt 1200
 [0250] gacaagggtt acttcgttca gccaactatc ttccggatgt ttacaccaga gatggagatt 1260
 [0251] gtcagatcggatg agatctttgg tcctgttgc actatcagca agtttgacac cattgtatgg 1320

[0252] gttgtcgacc ttgctaaca ctctcaatac ggtcttgctg ctggtatcca ctctgacat 1380
 [0253] atcaacaagg tcattgaacgt tgctgctaga atcaagtccg gtaccgtgt ggtcaacacc 1440
 [0254] tacaacgatt tccaccaaatt ggttccattc ggtggattt gccaatccgg tattggctgt 1500
 [0255] gagatgggtg ttgaagcttt ggaaaactac acccaataca aggctatccg tgtcaagatc 1560
 [0256] aaccacaaga acgagtaa 1578
 [0257] <210> 11
 [0258] <211> 1614
 [0259] <212> DNA
 [0260] <213> 巴斯德毕赤酵母 (Pichia pastoris)
 [0261] <400> 11
 [0262] atgagttcaa cagatatcca aggtgatcaa ggtgacaatg aaaagatata cgccatttag 60
 [0263] agcagtcctt ccaatgagca aataaaagat attcatgagg ctccggccga caacaaaagt 120
 [0264] gaactagaca tcccagtcaa acccaagggt tcctatatct tgggtctgt gttatgtctt 180
 [0265] ctagtcgcat tcgggtgtt cgtttcggt tgggataccg gtaccatctc aggttcgtt 240
 [0266] aacatgtctg actttacgag acgtttcgga cagtttaacg gtgaaacgta ttacctttct 300
 [0267] aaagttagag ttggtttaat tggttctatt ttcaacattt gttgtctat cggaggtgtc 360
 [0268] actctaggta aacttggtga cattgggtt agaaagaagg ctggatgtt cgtcatggtc 420
 [0269] atctatatgg tcgggtatattt gattcaaatt gcttccattt acaaattgtt ccaggatattt 480
 [0270] attggaagaa ttattgcagg tctggccgtc ggtgcagttt ccgttttattt cccatgttc 540
 [0271] atcagtgaga cttctcttaa acacatcaga gtttccttag tctctgtca ccaattaatg 600
 [0272] attacagccg gtattttctt gggtaactgtt accacttacg gaaccaagac ttacaccgac 660
 [0273] tccacccaaat ggagagttcc tttggattt tggttgcattt gggcatttctt gatgattgtt 720
 [0274] ggtatgacct tcatgccaga gtccccacgt ttcttgggtt aggttaacag agtcgacgag 780
 [0275] gctatgaagt ccattgccag agttaacaag gtctctatcg acgatccatc tgtctacaat 840
 [0276] gagatgagac ttatttctga cggattttag aaggagaagg aggctggtag cgtttcttgg 900
 [0277] ggtgaactgt tcactggtaa gccaaagatt ttctaccgtc tattgattgg tattttcatg 960
 [0278] caatcttgc aacaatttgc cggtaacaac tatttcttctt actacggAAC taccattttc 1020
 [0279] aaggctgtcg gattggacga ttcttccaa acttctatca ttcttgggtt tgtcaatttt 1080
 [0280] gcttccacat tcctaggat ctacaccatg gataaattt gtagaagaag aacactttt 1140
 [0281] ggagggttctg gagccatgtt tggttgggtt gtcattttca gttccgttgg tgtcaagtct 1200
 [0282] ctttatgaga acggtaagga tgatccatcc aaaccagcag gtaacgccat gattgtctt 1260
 [0283] acctgtctgt tcattttctt ctggatgtt acctgggtc cagggtttt cgtcggtgt 1320
 [0284] tctgaaacctt acccaacttag aatttagatcc aagggtatgg ccatcgctca aggttccaaat 1380
 [0285] tggctttggg gttccctcat tgccttcttcc actccattt aactcagggtc cattgatttc 1440
 [0286] gcctacgggtt acgtctttat gggatgtact ctgttgcct tcttcttgc tgcgttgc 1500
 [0287] gttcctgaaa ccaagggtct gtcgctggaa gacgttgcattt aagtctatga gaaccttacc 1560
 [0288] ttccggaaagag catatgcata cagccacacg attaaagaca agggcgccct ataa 1614
 [0289] <210> 12
 [0290] <211> 1032
 [0291] <212> DNA
 [0292] <213> 巴斯德毕赤酵母 (Pichia pastoris)
 [0293] <400> 12

- [0294] atggccctat ctcctaccta tcagggctac atatctacca ctggcgacgc gttgatcgtg 60
 [0295] atccaggcag ctctaaataa ccatttgaat cttctcccc gaagaccaag agaaagagag 120
 [0296] cgagatggc taatacgatc aggtAACgtA tttgttttg tcgagcaacg gtctcatatc 180
 [0297] aaacgatgga ccgatggtat cccctggct ccatctagag tccttggaaa gttcttgg 240
 [0298] tattggaaac tggacaagga taccccaaa aactcgcaaa gtgacgaaga tactgaggag 300
 [0299] gggagaaaga ggcgaaagac ttctgtggat gtaaccgatc caaataccag gcagttgg 360
 [0300] ggatcattgg tgacttccta tgacttcaaa gaggatggac ttattaaaaa aacactctcc 420
 [0301] ttgactttcc agaccggtgc taatgaagaa agggaaacag tgcacttgat tagttattat 480
 [0302] actccggaag atgtaacgaa ccatcgttt aacaggccgt ctgacaatcc atatctggcc 540
 [0303] aatatcactg tttcagagtc attattgact gccttgagag agagtaccct tggaggaaga 600
 [0304] gcaacgtctg atgacgagct ttcttagtc agaagtaact cgtagagta ccaagaggt 660
 [0305] ccaatgaaca tatctatgtc tttacctta tcaactccac tttcctgaa cacaggagta 720
 [0306] aactcaacta cccagctgca acagcaacaa ctacaacaac aacaacagca acagcaacag 780
 [0307] cagcagcaac aacagcagca acaacagcaa ccggtagcat ccctccaaa atttgatgga 840
 [0308] tccttctat tacaacaggg tgtaattcca gttcctcatt tcatggacca aaaaatggga 900
 [0309] agtagcaatt cgtggattaa caattggttt cgtccaaatt cgtcagaatc aaatgggcta 960
 [0310] tcggttatcg gacctcacaa gggatatgac gaacaaagtc cagcaacgag ttatactttg 1020
 [0311] aatgaacgtt ga 1032
 [0312] <210> 13
 [0313] <211> 491
 [0314] <212> DNA
 [0315] <213> 巴斯德毕赤酵母 (Pichia pastoris)
 [0316] <400> 13
 [0317] ctttttgta gaaatgtctt ggtgcctcg tccaaatcagg tagccatctc tgaaatatct 60
 [0318] ggctccgtt caactccgaa cgacctgctg gcaacgtaaa attctccggg gtaaaactta 120
 [0319] aatgtggagt aatggAACCA gaaacgtctc ttcccttctc tctccttcca ccgcccgtt 180
 [0320] ccgtccctag gaaattttac tctgctggag agcttcttct acggccccct tgcagcaatg 240
 [0321] ctcttcccag cattacgtt cggtaaaac ggaggtcgtg taccgcacct agcagcccc 300
 [0322] ggatggaaaa gtcccgccg tcgctggcaa taatagcggg cggacgcatt tcatgagatt 360
 [0323] attggaaacc accagaatcg aatataaaag gcaacacct ttcccaattt tggtttctcc 420
 [0324] tgacccaaag actttaaatt taatttattt gtccctattt caatcaattt aacaactatc 480
 [0325] acctgcaggc c 491
 [0326] <210> 14
 [0327] <211> 34
 [0328] <212> DNA
 [0329] <213> 人工
 [0330] <220>
 [0331] <223> 引物
 [0332] <400> 14
 [0333] gatagggccc caaacatttgc tccccctag tctc 34
 [0334] <210> 15
 [0335] <211> 39

- [0336] <212> DNA
[0337] <213> 人工
[0338] <220>
[0339] <223> 引物
[0340] <400> 15
[0341] gatacctgca ggaagggtgg aattttaagg atctttat 39
[0342] <210> 16
[0343] <211> 36
[0344] <212> DNA
[0345] <213> 人工
[0346] <220>
[0347] <223> 引物
[0348] <400> 16
[0349] gatagggccc cagcaatcca gtaaccttt ctgaat 36
[0350] <210> 17
[0351] <211> 36
[0352] <212> DNA
[0353] <213> 人工
[0354] <220>
[0355] <223> 引物
[0356] <400> 17
[0357] gatacctgca ggtaggttc aataaattgt ccggga 36
[0358] <210> 18
[0359] <211> 35
[0360] <212> DNA
[0361] <213> 人工
[0362] <220>
[0363] <223> 引物
[0364] <400> 18
[0365] gatagggccc tggactgttc aatttgaagt cgatg 35
[0366] <210> 19
[0367] <211> 37
[0368] <212> DNA
[0369] <213> 人工
[0370] <220>
[0371] <223> 引物
[0372] <400> 19
[0373] gatacctgca gggataaaag gtaaggaaa aaagcaa 37
[0374] <210> 20
[0375] <211> 36
[0376] <212> DNA
[0377] <213> 人工

- [0378] <220>
- [0379] <223> 引物
- [0380] <400> 20
- [0381] gatagggccc agaccaggcgtttaactacg caaatc 36
- [0382] <210> 21
- [0383] <211> 36
- [0384] <212> DNA
- [0385] <213> 人工
- [0386] <220>
- [0387] <223> 引物
- [0388] <400> 21
- [0389] gatacctgca ggcttttctt tgggcaagga aaaatc 36
- [0390] <210> 22
- [0391] <211> 39
- [0392] <212> DNA
- [0393] <213> 人工
- [0394] <220>
- [0395] <223> 引物
- [0396] <400> 22
- [0397] gatagggccc aattgattaa gttcagtcaaatttcaa 39
- [0398] <210> 23
- [0399] <211> 42
- [0400] <212> DNA
- [0401] <213> 人工
- [0402] <220>
- [0403] <223> 引物
- [0404] <400> 23
- [0405] gatacctgca ggatttatatt atgggaata atgaagagaa gg 42
- [0406] <210> 24
- [0407] <211> 33
- [0408] <212> DNA
- [0409] <213> 人工
- [0410] <220>
- [0411] <223> 引物
- [0412] <400> 24
- [0413] gatagggccc ctgcacaacc attgccagta agg 33
- [0414] <210> 25
- [0415] <211> 40
- [0416] <212> DNA
- [0417] <213> 人工
- [0418] <220>
- [0419] <223> 引物

- [0420] <400> 25
- [0421] gatacctgca gggttttaga agagggagaa ctttagattgg 40
- [0422] <210> 26
- [0423] <211> 24
- [0424] <212> DNA
- [0425] <213> 人工
- [0426] <220>
- [0427] <223> 引物
- [0428] <400> 26
- [0429] cctgaggctt tggccaccc atct 24
- [0430] <210> 27
- [0431] <211> 30
- [0432] <212> DNA
- [0433] <213> 人工
- [0434] <220>
- [0435] <223> 引物
- [0436] <400> 27
- [0437] ggaacatagt agtaccaccc gacataacga 30
- [0438] <210> 28
- [0439] <211> 22
- [0440] <212> DNA
- [0441] <213> 人工
- [0442] <220>
- [0443] <223> 引物
- [0444] <400> 28
- [0445] tcgccgacca ctaccaggcag aa 22
- [0446] <210> 29
- [0447] <211> 23
- [0448] <212> DNA
- [0449] <213> 人工
- [0450] <220>
- [0451] <223> 引物
- [0452] <400> 29
- [0453] accatgtgat cgcgcttctc gtt 23
- [0454] <210> 30
- [0455] <211> 24
- [0456] <212> DNA
- [0457] <213> 人工
- [0458] <220>
- [0459] <223> 引物
- [0460] <400> 30
- [0461] cctgaggctt tggccaccc atct 24

- [0462] <210> 31
[0463] <211> 30
[0464] <212> DNA
[0465] <213> 人工
[0466] <220>
[0467] <223> 引物
[0468] <400> 31
[0469] ggaacatagt agtaccacccg gacataacga 30
[0470] <210> 32
[0471] <211> 28
[0472] <212> DNA
[0473] <213> 人工
[0474] <220>
[0475] <223> 引物
[0476] <400> 32
[0477] aaaccttagga aaagtggca gcaaatgt 28
[0478] <210> 33
[0479] <211> 29
[0480] <212> DNA
[0481] <213> 人工
[0482] <220>
[0483] <223> 引物
[0484] <400> 33
[0485] actctgtcac ttactggcgt tttctcatg 29
[0486] <210> 34
[0487] <211> 34
[0488] <212> DNA
[0489] <213> 人工
[0490] <220>
[0491] <223> 引物
[0492] <400> 34
[0493] gatagggccc caaacatttg ctccccctag tctc 34
[0494] <210> 35
[0495] <211> 39
[0496] <212> DNA
[0497] <213> 人工
[0498] <220>
[0499] <223> 引物
[0500] <400> 35
[0501] gatagggccc ggaatctgta ttgttagaaa gaacgagag 39
[0502] <210> 36
[0503] <211> 36

- [0504] <212> DNA
- [0505] <213> 人工
- [0506] <220>
- [0507] <223> 引物
- [0508] <400> 36
- [0509] gatagggccc ccatattcag taggtgttcc ttgcac 36
- [0510] <210> 37
- [0511] <211> 36
- [0512] <212> DNA
- [0513] <213> 人工
- [0514] <220>
- [0515] <223> 引物
- [0516] <400> 37
- [0517] gatagggccc ctgcagatag acttcaagat ctcagg 36
- [0518] <210> 38
- [0519] <211> 32
- [0520] <212> DNA
- [0521] <213> 人工
- [0522] <220>
- [0523] <223> 引物
- [0524] <400> 38
- [0525] gatagggccc gaccccgtt tcgtgacaaa tt 32
- [0526] <210> 39
- [0527] <211> 37
- [0528] <212> DNA
- [0529] <213> 人工
- [0530] <220>
- [0531] <223> 引物
- [0532] <400> 39
- [0533] gatagggccc ccggataaga gaattttgtt tgattat 37
- [0534] <210> 40
- [0535] <211> 31
- [0536] <212> DNA
- [0537] <213> 人工
- [0538] <220>
- [0539] <223> 引物
- [0540] <400> 40
- [0541] gatagggccc gcctgctcca tattttccg g 31
- [0542] <210> 41
- [0543] <211> 859
- [0544] <212> DNA
- [0545] <213> 人工

- [0546] <220>
- [0547] <223> 启动子变体
- [0548] <400> 41
- [0549] ggaatctgta ttgttagaaa gaacgagagt ttttacggc gccccatat tggccgtgt 60
- [0550] gaaaacagct tgaaacccca ctacttcaa aggttctgtt gctatacacg aaccatgttt 120
- [0551] aaccaacctc gctttgact tgactgaagt catcggttaa caatcaagta ccctagtcg 180
- [0552] tctgaatgct cctttccata ttcaagtaggt gtttcttgca ctggcatg cactgcggaa 240
- [0553] gaattagcca atagcgcgtt tcatatgcgc ttttaccccc tctttgtca agcgcaaaat 300
- [0554] gcctgtaaga tttgggtgggg gtgtgagccg ttagctgaag tacaacaggc taattccctg 360
- [0555] aaaaaactgc agatagactt caagatctca gggattccca ctattggta ttctgatatg 420
- [0556] ttttcctga tatgcatcaa aactctaatt taaaacctga atctccgcta ttttttttt 480
- [0557] tttttgatg accccgtttt cggtacaaat taatttccaa cggggcttg tccggataag 540
- [0558] agaattttgt ttgattatcc gttcggataa atggacgcct gctccatatt tttccggtaa 600
- [0559] ttaccccacc tggaagtgcc cagaatttc cggggattac ggataatacg gtggctgga 660
- [0560] ttaattaata cgccaagtct tacattttgt tgcagtctcg tgcagttatg tgcaataata 720
- [0561] aacaagatga gccaatttat tggatttagtt gcagcttgac cccccatag ctggcatag 780
- [0562] ccaagtgcta tgggtgttag atgatgcact tggatgcagt gagtttgga gtataaaaaga 840
- [0563] tccttaaat tccaccctt 859
- [0564] <210> 42
- [0565] <211> 664
- [0566] <212> DNA
- [0567] <213> 人工
- [0568] <220>
- [0569] <223> 启动子变体
- [0570] <400> 42
- [0571] ccatattcag taggtgttcc ttgcactttt gcatgcactg cgaaagaatt agccaatagc 60
- [0572] gcgtttcata tgcgctttta cccctcttt tgtcaagcgc aaaatgcctg taagattttgg 120
- [0573] tgggggtgtg agccgttagc tgaagtacaa caggctaatt ccctaaaaaa actgcagata 180
- [0574] gacttcaaga tctcaggat tccactatt tggattctg atatgtttt cctgatatgc 240
- [0575] atcaaaaactc taatctaaaaa cctgaatctc cgctattttt ttttttttt tgatgacccc 300
- [0576] gtttcgtga caaattaatt tccaacgggg tcttgcggg ataagagaat ttgtttgat 360
- [0577] tatccgtcg gataaatgga cgcctgctcc atattttcc gtttattacc ccacctggaa 420
- [0578] gtgcccgaaa tttccgggg attacggata atacgggtt ctggattaat taatacgcca 480
- [0579] agtcttacat ttgttgcag tctcgtgcga gtatgtgcaa taataaacaa gatgagccaa 540
- [0580] ttatttggat tagtgcagc ttgacccgc catagctagg catagccaa tgctatgggt 600
- [0581] gttagatgat gcacttggat gcagttagtt ttggagtata aaagatcctt aaaattccac 660
- [0582] cctt 664
- [0583] <210> 43
- [0584] <211> 493
- [0585] <212> DNA
- [0586] <213> 人工
- [0587] <220>

- [0588] <223> 启动子变体
- [0589] <400> 43
- [0590] ctgcagatag acttcaagat ctcaggatt cccactattt ggtattctga tatgttttc 60
- [0591] ctgatatgca tcaaaactct aatctaaaac ctgaatctcc gctatTTTTT ttttttttt 120
- [0592] gatgacCCCG tttcgtgac aaattaattt ccaacgggt cttgtccgga taagagaatt 180
- [0593] ttgttgatt atccgttcgg ataaatggac gcctgctcca tattttccg gttattaccc 240
- [0594] cacctggaag tgcccagaat tttccggga ttacggataa tacgggtgtc tggattaatt 300
- [0595] aatacgccaa gtcttacatt ttgttgcagt ctcgtgcag tatgtcaat aataaacaag 360
- [0596] atgagccaat ttattggatt agttgcagct tgacCCGCC atagctaggc atagccaaGT 420
- [0597] gctatgggtg ttagatgatg cacttggatg cagtgagttt tggagtataa aagatccta 480
- [0598] aaattccacc ctt 493
- [0599] <210> 44
- [0600] <211> 370
- [0601] <212> DNA
- [0602] <213> 人工
- [0603] <220>
- [0604] <223> 启动子变体
- [0605] <400> 44
- [0606] gacCCGTT tcgtgacaaa ttaattcca acgggtctt gtccggataa gagaattttg 60
- [0607] ttgttattc cggtcgata aatggacgcc tgctccatat tttccgggtt attacccac 120
- [0608] ctggaaGTgc ccagaatttt ccggggatta cggataatac ggtggtctgg attaattaaat 180
- [0609] acgccaagtgc ttacatttt ttgcagtctc gtgcgagttat gtgcataataat aaacaagatg 240
- [0610] agccaaTTTA ttggatttagt tgcaGCTGA ccccgccata gctaggcata gccaaGTgct 300
- [0611] atgggtgtta gatgatgcac ttggatgcag ttagttttgg agtataaaag atcctaaaaa 360
- [0612] ttccaccctt 370
- [0613] <210> 45
- [0614] <211> 328
- [0615] <212> DNA
- [0616] <213> 人工
- [0617] <220>
- [0618] <223> 启动子变体
- [0619] <400> 45
- [0620] ccggataaga gaattttgtt tgattatccg ttccggataaa tggacgcctg ctccatattt 60
- [0621] ttccggTTT taccccacct ggaagtgcCcC agaattttcc ggggattacg gataatacgg 120
- [0622] tggtctggat taattaaatac gccaagtctt acatTTTTT gcagtctcgT gcgagttatgt 180
- [0623] gcaataataa acaagatgag ccaattttatt ggatttagttg cagcttgacc ccGCCatAGC 240
- [0624] taggcatAGC caagtgcTAT gggTgttaga ttagtgcactt ggatgcagtG agttttggag 300
- [0625] tataaaAGAT ccttaaaATT ccaccctt 328
- [0626] <210> 46
- [0627] <211> 283
- [0628] <212> DNA
- [0629] <213> 人工

- [0630] <220>
- [0631] <223> 启动子变体
- [0632] <400> 46
- [0633] gcctgctcca tattttccg gttattaccc cacctggaag tgcccagaat tttccgggga 60
- [0634] ttacggataa tacgggtggtc tggattaatt aatacgccaa gtcttacatt ttgttgcatg 120
- [0635] ctcgtgcgag tatgtgcaat aataaacaag atgagccaat ttattggatt agttgcagct 180
- [0636] tgaccccgcc atagcttaggc atagccaagt gctatgggtg ttagatgatg cacttggatg 240
- [0637] cagtgagttt tggagtataa aagatcctta aaattccacc ctt 283

ATTTCCACCCCCATCCCAGTAGAATGTAGGGTCCCCAACATTGCTCCCCCTAG
TCTCCAGGGAAATGTAAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGT
CAGGATAATTACGTCGACTGTAGTAAAACAGGAATCTGTATTGTTAGAAAGAACG
AGAGTTTTTACGGCGCCCATATTGGGCCGTGAAAACAGCTGAAACCCCCA
CTACTTCAAAGGTTCTGTTGCTATACACGAACCATGTTAACCAACCTCGCTTT
GACTTGACTGAAGTCATCGGTTAACATCAAGTACCCCTAGTCTGTCTGAATGCTCC
TTTCCATATTCACTAGGTGTTCTGCACCTTGATGCAGTCGGAAAGAATTAGC
CAATAGCGCGTTCATATGCGCTTTACCCCCCTTTGTCAAGCGAAAATGCCT
GTAAGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCCT
GAAAAAAACTGCAGATAGACTCAAGATCTCAGGGATTCCACTATTGGTATTCTG
ATATGTTTCCCTGATATGCATCAAAACTCTAATCTAAACCTGAATCTCCGCTATT
TTTTTTTTTTGATGACCCGTTTGTGACAAATTAAATTCCAACGGGGCTT
GTCCGGATAAGAGAATTTGTTGATTATCCGTTGGATAATGGACGCCGTGCTCC
ATATTTTCCGGTTATTACCCCACCTGGAAGTGCCAGAATTTCCGGGGATTACG
GATAATACGGTGGTCTGGATTAATTAAACGCCAAGTCTTACATTGTTGCAGTC
TCGTGCGAGTATGTGCAATAATAACAAAGATGAGCCAATTATTGGATTAGTTGCA
GCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCA
CTTGGATGCAGTGAGTTGGAGTATAAAAGATCCTAAAATTCCACCCCTT

图1

GTAAATAGCGGCAGCAATCCAGTAACCTTTCTGAATAGCAGAGCCTTAACAAAA
TAATGCCAGGGTAAAAAATTGAAATTGACACCAAAAATAAAGACTTGTGTTA
TAAGTCTAACAAAGTCCGCAATTGGAGCTAACGGTGGCGGTTGCTGGGATAT
TCAATAATGGTAGAATGTTGCTGCGGGTATATGACAGAGCGTGAAACACACTGAA
CAAGGTAATGGAACAACAGCAATTGCAATATGGGGAGGGATAGTCAAGAACAAA
GCAGCAATGGCAAAGTACTGAATATTCCAAAGCCAAAGGTCCAGTGGTTCA
ACGACAAAGTCTTGGTATAGCTTGGAACAAAAGGACACCGAAAGACTCGAC
AGCGCCCACAAATACAGCGTTGAGAACGAAATTGATTGCTCCAGAGCTTCTA
ATAGTCAGAAGATAACCCAAACCTCCGAGCAACGTTAGCACATGACCTAAGAAC
AGGCGAAGTGAAGAGTCTGGAATAACGACACCCAGTCAGTTTCTGAGCTCCT
GGTGGGATTGGTAGAACGATTGATTGCTTGGAGTGGTTTATTGAAGATGGT
GTTGAAGCCATTGTTGCTAAAGAGTCGGAGTTGCTTTAGGGTTGTTAACCAA
AGGAGGAAAAACTGCGCCGTTGAAGTCCCAGGTAGTTCGCGTGTGAGGCCAG
CCAGGGAAAGCTTCCTCGGTACTTTTTCTTGCAGGTTCCGGACGGATTAA
GCTTCGGGTTATGAGGGGGCGGTAGCCAATTCCGGACACAATTGCGTCGCA
GCTAGTCACCCCGCCATAAATATACGCAGGATTGAGGTAATAAACATCGATAGTCTT
AGTAATTAATACAATTCAAGTGGCGAACATTGGCAACATGACGTAAGGCCACTGTTG
TCTATAAAAGGGATGAATTTCATGTTTGAGGCCTCCGGACAATTATTGAA
CTCAA

图2

TGGACTGTTCAATTGAAGTCGATGCTGACGATGTCAAGAGAGATGCTCAATTATA
TTTGTCAATTGCTGGTTACACTGGAAACGCTACTTTGTTGGCGGAAACTCTACCA
GTTTGGCCGTCCATGTAAACGATGTCGTTCTGGGCCGTGACCGTTCAACACGAA
CATAAACCAATGACAAATCCACTTACAGGTCTAGTCATATGGAGGCAATTGGTACC
TTACTTCTTGATGTCCTAACGTGGTACTAACAAATGTCTCG
TTTGTCACTACAAACTCCGAGGTAAATAAAGGATTCTGTGGGATTCTCTCAAGTT
TGTTTGGAAAGTTGTAACAGGTTATAAGCATACTCGTGCCTGTCACAAATTGAAT
CATTATTGTTGCGAGATACATGAACAAAGTGTGAACGGGACCCATTACTACAAT
TCCCACGCAACCCTGTTCAAAGCCCATTGACAATTGTTCGTTACACC
CCCAGTTGATGTACATCGCTGCAATGATGTGTCCGGAGTATTTCATATT
CAGCTTGAATTCGTATACTCAACCAATATCTGGGGTATACTTTATGTAACCTATA
CAAATCAACTATACTATTCACCTTCGACCAATCATCTCCATTGTTAAGTTT
GCTTCCTATATCCCTGACCCCTGACATCACCATGATTCCGCTAACGGTTCTCCTC
TACATCGTCCCTCTTGAGAGGGTGTTCAGTTGACATTCAAATTACCCCCCGC
CATCACGCGCAACCGAGACCGCACCCCCGAATTTCACAAATTACCCACACCC
ATACTCCACCACTATGAGGGTTATTAGAACTGATCACGTATAAATACCACCGCAAG
TTCCAAGGGATCGTGTCTTCTCCAATTGCAATCATATTCTGACTCTTCTA
GTTCAAGATTAATTCTTACACTGCTTTCCCTACCTTATCC

图3

AGACCAGCAGTTAACTACGCAAATCCACAGGAATTCTACATCACAAATACCAATG
GTAATACCACGACGTCAAGGAATGGAAACGACGACTTGGAGGAAGACTTCGTCAA
CCTCTTGGAGTACCCGAGGCTAAGACAATAAGAAGAAAAAAAAAGAAAAGCG
GTGGGGGAGGGATTATTAAATAAGGATTATGTAACCCCAGGGTACCGTTCTATAC
ATATTAAAGGATTATTAGGACAATCGATGAAATCGGCATCAAACGGATGGAGT
ATAGTGTCCGGATAATCGGATAAACATCATCTTGCAGGGAGGCCGCTGGTTGG
TGAGAGGAGTGAAATATGTGTCTCCTCACCCAAGAATCGCGATATCAGCACCC
TGGGGGACACTATTGGCCTCCCTCCAAACCTCGATGTGGTAGTGCTTATTAT
ATTGATTACATTGATTACATAGCTAACCCCTGCCTGGTTGCAAGTTGAGCTCC
TTCCAATATTAGTAAATGCCTGCAAGATAACCTCGGTATGGCGTCCGACCCCC
TTAATTATTTAACTCCTTCCAACGAGGACTTCGTAATTTGATTAGGGAGTTGA
GAAACGGGGGGCTTGATACCTCCTCGATTCACTCCACCCCTCTCAGTCCC
AAGTGGGACCCCCCTCGGCCGTGAAATGCGCGCACTTAGTTTTTCGCATGTA
AACGCCGGTGTCCGTCAATTAAAAGTCGCAGACTAGGGTGAACTTACCATT
GTCGCACTCCGTCTCCTCGGAATAGGGGTGTAGTAATTCTGCAGTAGTGCAATT
TTACCCCCGCCAAGGGGGGGCGAAAAGAGACGACCTCATCACGCATTCTCCAGTC
GCTCTCTACGCCTACAGCACCGACGTAGTTAACCTCTGCTTTCTGATTTCCTGCC
GCCATTCCCCCTGAAAACTTAACCTCTGCTTTCTGATTTCCTGCCAAAGA
AAAG

图4

AATTGATTAAGTCAGTGAAATTCAAACCGCTATACACAACAGGACAACTTGAG
TTTAGAAAAATCCGATGTAGTGTAAACGGCTAGCACGGTCCGCTTCACCGGGCAG
ACCCGGGTTCGACTCCCGCATCGGAGTTATTTCCATTCGTTCTTAGAGTA
TTCTCCTCAGCATGCCCTGAATTTCCTTTCCATGTGTCCCATTTC
CTTTTTTACAGTTTCCTCGTGTAAATTGGCTACACAAAAGCTGCCACACGA
AACCTTAATCACGAAAAACTATACAGCCTCACTAATCCGTAGCCCCATAATATGT
TGTCCACGTGCTGTTGGGTACTACCTGTAGACTCTCATACCCCACCGTCTTCT
CCAACAATTAACGCAGTACCGAGATTATCAGCAGACTCAAATTGGGCAAACCTCT
GTATTTCCTGCCGCATAATTATGGGTCTCAGGCCTCCACGTTCTGTTA
CTTGAAGAATATTGGCTCGGAAAAAGTGGTAAGGACAACCCCCTTTAATTGGA
TCCAGTTTCGAAATGTTCCGATCCGTACGTACCTCCGAAGCCGTACATTTC
ACTCAATCTACGTAGCTTGGACTCAGCGCTCCTGGAATTGCCAGGACAGTTAC
TTGAGTTGATATTCCCTTAGATTGTGTGCTTCTTTCCAAAATTGAGGCTTCG
TTGAAAAGTGGAAATCTGGTCGCTAGATCACTCATGCCTATTTCACGGAAAAAA
TAAGTGGTACTATGCACCCCTAAACCTAAAGAAAAACGGAAAAATTACCCCCAAA
CCTGGTGATGTTTCGCCCCCTTCTTTATCCGAGTTTCTTTCTTGTCTG
CCAAATTCCCTCCTGACCTAGCGTCCCCGGAAAAAATTAACTACTTAAGGACCG
AATGAGCCCCAGCTTTCCCCTCTTCATTATTCCCCATAATATAAT

图5

CTGCACAACCATTGCCAGTAAGGACGAAGAGAAGGCCACTACCCAAAATTCAG
GATAACGTCTCATACCATGCAGCGACGCCCTACAAGACGCTGTCAAGACATGCCA
ACTTCAACGAAGTGAACTTAACACATTGATCGGGAAATTGACCACCAAGGGAAT
GCTGGTTGAGGCTGGAAGCGTTGCCAGTGCTGAGGGAACTGGACCGAAAGTT
TAGTAATGCATAAGAGGATATATAGGAATGCAGTAATAATATTAGTACCCATTAA
GTGGGCTAAGCCATTGGAAGGCCGTCTGACTGATGGTGGTCTTCTCATTAG
ATAGTGCATTGCAACTACCGTCTGAGATTGAGTTGATGTGAAGCTCCAGCGCC
AAAACAGTATAAGAACCTATCTCCGCATTATTGTTCTGCGTAAAAGTTGTGA
AGAAACAGGGTAGTTGCGCAGATTGAGTTGATATGCGCATAGGATGGTCATT
GACTTCTTCCTCGAAAGAGCCACACCGTTAGCTAAAAAGGACGCGCATCTACC
CCAAAATAGAATGTGGGAAATAGGACGCGCAACTCCTCTCAATCACTGGACGT
CAGAAAAACAAATGCGCAATCGAGTCACCCCTCCGTGATACCCTCCGTGATACCC
CTCTCCGTCTATTCTGACAGCGTCTCCCCATGACGTTCAATCTACTTAGAAAAGA
TTTCGTTTTTTTCTCAATTACACGATCTCATCTTCTGCAAGGGTCTGGAGGAC
ATCACCAATCTGCGACTCCATAACTTAGTCCTGAGTTATATTACGCTTCATCTGA
TGAGTAGGAAGAAAAAGTTCACGAAATTCCCCCGCCAATTGCCCTCGGAATA
AGCAGCCACTCTCCTCTGCCCATAGTAAGCTTGCAGGCCCCAACTTGGCC
AGAAACTTAAATATGCCAAACAATCTCCCCAATCTAAGTTCTCCCTTTCTAAAAA
A

图6

ATGTCCTCGTTTCTAAACAACCAAACAGTAAAGATGATGACGCCTGGAAAG
 AGCTAGAGCTCTAAATTTCAGGGCAAAGTTATGATAAATTCCAAAAACTTACAA
 TATCTACGCTATTGCAATAACAGCCACCCTGACTGATGTTGGAGGTGATA
 TTTCTTCTGTGTCCCTCGTCAAGTCAGGATCATTACAGAAACTACTTCAACCGT
 CCCGACAGTTGACGCAAGGGGTATCACCGCAAGTATGGCTGGAGGTTCTTCT
 TGGGTCGTTATTTCTGACTTCCAGGATATCTTGGAAGAAGAGTTGCTCTG
 CATATGTGCAGTGTCCCTGGATTATCGGGGCCATTCTCAATGCGCTGCACAAA
 ACCAAGGTATGCTGATCGCAGGGAGATTGATTCCGGTATCGGTGTCGGGTTGG
 TTCAGCTTCAGCTCCAGTCTATTGTTCTGAAGTTGCTCCAGCAAAGATTAGAGGAA
 TGATTGGAGGATTATTCATTTCTGTCACTGTGGGTATCATGATAATGTTTATA
 TCGGATATGGATGTCACTACATTGACGGCGTGCATCATTAGACTGGCCTGGGG
 TTTGCAAATGGTCCAGGTCTTATTCTTGTCGGTGTATTCTCCTCCTGAGT
 CTCCAAGATGGCTGGCTAACCAACCACAACCGCTGGGAAGACGCAGTTGAGGTTATTG
 CTAATGTTGTTGCAAAAGGTGACAGAGAAAACGCCATGTGCGTCTGCAATTGGA
 TGAAGTTCAAGGAGCAACTATTGATTGACAAAGATGCTTCTGATTTGGTTACCTG
 ATTTGTTAAGAAAGATTGTATCAAACGTACCTCATTGGAGTGTCAAGCTCAAGTG
 TGGCAACAACTTGTGGTATTAATGTTGCAATGTACTACGTTGTATCTCTCCA
 AATGGCTGGTTACTGGAAATGTGGCGTGGTATCGCCTCAATTCAATATGTT
 TGAATGTTGTTATGACTGTTCCAGCTTGTCTAATGGACCGTATAGGCAGACGA
 CCCCTACTAATTGGTGGTGGTATTTCATGTGTATTGGCTGTTGGAGTGGCAG
 GATTATTAGGCACCTACTCTGAACCAATTGAAAATTCAAGCGGTGATGATACTGTC
 AGAATTACTATTCTGACCAGCACAAGGCTGCAGCAAGGGGTGTTATTGCCTGTT
 CCTATCTATTCTGCTGCTCCTTGCTCCAACCTGGGTATCTGCATTGGTTAT
 GCCTCTGAAATTTCACAAACAGACAAAGAGCAAAGGGAGCAGCATTGCTGCCT
 CCGCTAACTGGATTTCAACTTGCCTTGGCTATGTTGCGCCATCAGCCTTAGA
 AACATTACATGGAAGACTTACATCATTGGAGTATTCGTTCTGCTTAACAATC
 CATGTTTCTTACAATTCCCAGAAACCAGAGGTAAGACTTGGAAAGAAATTGATCA
 AATGTTAAGGACAATATTCCAGCTGGAGAAGTGCTTCGTAACGTTCCAGATATGC
 CAATTTCACAAAGAGAAGGTAGTATCTACTGAGCATGCAGAAAATGCTCCAGC
 TCGTCCGAAAAAGCCTGATGGTTAGGAAGAGGAATCTGTATAA

图7

ATGGTAGTTGCAATCGAAGGTGGTACAGGCTAGGCCTATGAATCTTACTTGGA
AACCAACTCCAACCCCAATTGATGATGCAATTGAGACAATTAGATATGCTGTTGAG
GAAGCTGGTGTCAAGATACTTGAACGGAGGAGAGTTCTACAACCTTCCTCTGATT
CAAACCTGAATTGCAGTACATTCAAGGAATTGCAAAAAGGTACCCCGAGCTATAT
AAAAAGGTGAGTCTGTCGGTAAAAGGTGCTGTCAGTTGGTCGATGTGAGCCCCG
ATTCTTCCCCGGAGAACCTTGAAAAATCGATTCAAACATAACCAAACATTGCCG
AACAACTTCCTGCCAATTGGAGCCTGCTAGAATCGATAAACGTTACTCCATTGA
GGAGACAATAAAGAATCTCTCTAAGTTGTCGAAGATGGCAGAATTGGAGGTATT
TCACTTAGTGAAGTTGGTGCTGACACTATCAGAAGAGCTGCGAAAGTGGCTCCC
TCGCCTGTGTGGAAGTGGAGTTCTCTATTGACTAGAGATATTCTCATAATGGA
GTTCTTGCTGCTGTGAGGATTGAACATTCTATTATTGCCTACAGTCCCTGGG
AAGAGGATTTTGAUTGGAAAGATTCAATGACGATGAAGTTATTGAACACAATTGAA
ACTTGTTCACGGTTGAAAAAGATAGCCGACAAAAAAGGAGTCACATTGGCTCAAT
TGTCTCTGCGTGGTTACGAAAGTTGGAGATAAACACGTCAAGGTGCTCCTATT
CCAAGCTGCTCATCTCCTCGTAGAGTTGCAGAAAACACAAAAGAGATTCCATTGA
CTGATAGCGAGTTCCAGGAGATTACTGACTTGCAGAGTCGGTTCCAATCAAAGG
TGGTCGTTACAACAAAGCAAGTGAGGCTGTTCTAACGGTTAG

图8

ATGACATTGCTCCCTAGAATTGAGATTGACCTCCTAACGGATTGAAGTA
CACTCAACCATTGGGACTCTCATCAACAATGAGTTGTTGAAGGTGTAGAGGGAA
AAGCTCTTACCACTGATCAATCCTGTGATGAGACTAAAATAACCCAAGTTGGGA
AGCTTCTGCAGCGGATGTTGACCGTGCTGTTGATGCCGCTGAAGATGCTTCAAC
AACTCCGTATGGGCTACTCAGGACCCATTAGAGAGGGGAAAGCTGATGAACAAAT
TGGCAGACCTATCGATCGTACTCAACATCTGGCTGGTATCGAATCCATCGA
CAATGGTAAGGCCTACCTCTGCCAGGGTGTGTTACTCTTGCTGTCAACTAC
ATCAGATCCTGTGCTGGATGGGCCACAAGATTGGAAACGTTGTTGATTCCG
GAAACACCCACCTAACCTGGTAAAAGAGAGGCCATTGGGTGTGGACAAAT
TATCCCATGGAACCTTCCTCTCCTGATGTTGGCTTGGAAAGTTGGGACCTGCGCTG
GCCACAGGTAACACTGTTGTTGAAGACTGCCAGTCTACCCCTCTGTCGGGTT
TATACGTTGCCAAATTGATCAAGGAGGCCGGTTCCACCTGGTGTGGTTAACAT
TCTCAGTGGTTCGGTAAACCCAGCTGGAGCTGCCATCGCTGCTCATCCCAGAAC
AAGAAGATTGCTTCACCGGATCCACTGCAACAGGCCGTAAGATCATGGAAGCAG
CCGCTAAATCTAACCTGAAAAAAAGTCACTTGGAACTAGGTGGTAAATCTCCAAAC
ATTGTGTTGAAGATGCTGATATCCAGAAGACTATCCATAACATTATTTGGGAAT
CTTCTCAATTCTGGTGAAGTCTGTTGCAGGTTCCAGAGTCTACATTCAAGACA
CTGTGTATGAAGAAGTGCTGAAGCCTCAAGAAGGAGACTGATAACGTTAAGGT
TGGTGGACCATTGAAAGAAGGTGTCTCCAAGGGCCTCAGACCTCTGAGTTGCAA
CTTAACAGAACCTTAGTTACATCAAACACGGTAAGGATGAAGGTGCTCGTGTAAAT
TACCGGTGGTTCAAGATAACGTAACCGAGGTTACTACATTAAGCCCACAATTGG
CTGACGTTACTGAAGACATGAAGATTGTCAGGGAGGAGATTTGGTCTGTGGT
TACTACTAACAGTTCTACCGTGGATGAGGTTGTTGGATATGCCAACACACCA
ACTATGGCTAGCTGCTGGTATTCACACAAACAACTTGAACAAAGCCATTGATGTT
GCCAGTAGAATCAAGGCGGGTGTGTTGGATTAACACCTACAACGATTCCACC
ACATGGTCTTGGATAACTACACTCAAGCCAAGGCTATCAGAATTGCTTACACTCCTG
AACATAAGTAG

图9

ATGCTTAGAACTTCTCCAGCTACTAAGAAAGCTCTCAAGTCGCAGATTACGCC
TCAACGTTGCTGCCTTGAGATTCTACTCCTCATTGCCTTGCAGGTTCCAATTACCT
TGCCAAACGGTAAGACCTACAATCAGCCAACAGGTTGTTATCAACAATGAGTTC
GTTCCCTCTAAGCAAGGTAAAGACCTTGCTGTTAAACCCCTCCACTGAGGAGGA
GATTACTCACGTCTACGAGTCCAGAGAGGACGACGTTGAGTTAGCCGTTGCAGC
CGCTCAAAAGGCTTCGACTCAACCTGGTCCACCCAGGACCCTGCTGAGAGAGG
TAAGGTCTTGAACAAGTTGGCTGACCTGATCGAGGAGCACTCTGAGACCCTGCC
GCCATCGAGTCCTGGACAACGGTAAGGCCATTCCCTCCGCTAGAGGTGATGTTG
GTCTGGTTGTCGCCTACTTGAAGTCCTGTGCCGGTTGGGCCACAAGGTTTCG
GTAGAGTTGTTGAAACCGGAAGCTCCCACCTCAACTACGTTAGAAGAGAGGCCATT
GGGTGTTGTGGTCAGATTATCCCATGGAACCTTCCTCTGATGTGGCCTGG
AAAGTTGGTCCAGCTTGGCCACTGGTAACACTGTTGTCCTGAAGACAGCCGAGT
CTACTCCTCTGTCGCCCTGTACGTTCCAATTGGTCAAGGAGGCCGTATCCC
AGCTGGTGTCCACAACATTGTGTCCGGTTCGGTAAAGATTACTGGTGAAGCTATT
GCTACTCATCCTAAGATCAAGAAGGTTGCCTTCACTGGTTCTACCGCCACTGGTC
GTCACATCATGAAGGCTGCTGCCAATCCAACCTGAAGAAGGTTACTTGGAGTT
GGGTGGTAAATCTCTAACATCGTGTCAACGATGCTAACATTAGCAAGCTGTC
GCCAACATCATCCTCGGTATTTACTACAACCTGGAGAAGGTTGTTGTGCTGGTTC
CAGAGTTATGTTCAATCCGGTATTTACGACGAGCTTGGCCGAATTCAAGACTG
CTGCTGAGAATGTCAAGGTTGGTAACCCATTGACGAGGACACCTCCAAGGTGC
TCAAACCTCTCAGCAACAATTGGAGAAGATTGGGTTCGTTGAGCGTGGTAAG
AAGGACGGTGCTACTTGATTACTGGTGGTGGCAGATTAGGTGACAAGGGTTACT
TCGTCCAGCCAACATCTCGGTGATGTTACACCAGAGATGGAGATTGTCAAGGA
AGAGATCTTGGCCTGTTGTCACTATCAGCAAGTTGACACCATTGATGAGGTTG
TCGACCTTGCTAACGACTCTCAATACGGTCTGCTGCTGGTATCCACTCTGACGA
TATCAACAAGGTCAAGTGTGCTGTTGACTATCAGCAAGTTGACACCATTGATGAGGTTG
AACACCTACAACGATTCCACCAAATGGTCCATTGGTGGATTGGCCAATCCG
GTATTGGTCGTGAGATGGGTGTTGAAGCTTGGAAAACACACCCAAATACAAGGC
TATCCGTGTCAAGATCAACCACAAGAACGAGTAA

图10

ATGAGTTCAACAGATATCCAAGGTGATCAAGGTGACAATGAAAAGATATACGCCA
TTGAGAGCAGTCCCTCCAATGAGCAAATAAAAGATATTGAGGCTCCGGCCGA
CAACAAAAGTGAACTAGACATCCCAGTCAAACCCAAGGGTCTATATCTTGGTGT
CTGTGTTATGTCTTAGTCGCATCGTGGTTCTGTTGGTGGGATACCGG
TACCATCTCAGGTTCTGTTAACATGTCTGACTTACGAGACGTTCGGACAGTTA
ACGGTGAAACGTATTACCTTCTAAAGTGAGAGTTGGTTAATTGTTCTATTTCA
ACATTGGTTGTGCTATCGGAGGTGTCACTCTAGGTAAACTGGTGACATTGGGG
TAGAAAGAAGGCTTGATGTTGTCATGGTCATCTATGGTCGGTATTTGATTG
AAATTGCTCCATTGACAAATGGTACCAAGTATTGATGGAAAGAATTATTGCAGGT
CTGGCCGTCGGTGCAGTTCCGTTATCCCCATGTTCATCAGTGAGACTCTC
CTAAACACATCAGAGGTTCTTAGTCTCCTGCTACCAATTAATGATTACAGCCGGT
ATTTCTTGGGTTACTGTACCACTACGGAACCAAGACTTACACCGACTCCACCCA
ATGGAGAGTTCCCTGGGATTGTGTTCGCTTGGCCATTCTGATGATTGTTGGTA
TGACCTTCATGCCAGAGTCCCCACGTTCTGGTTGAGGTTAACAGAGTCGACGA
GGCTATGAAGTCCATTGCCAGAGTTAACAAAGGTCTCTACGACGATCCATCTGTC
TACAATGAGATGAGACTTATTCTGACGGTATTGAGAAGGAGAAGGAGGCTGGTA
GCGTTCTGGGTGAAGTCACTGGTAAGCCAAAGATTCTACCGTCTATTG
ATTGGTATTTCATGCAATCTTGCAACAATTGACCGTAACAACATTCTTCTAC
TACGGAACCTACCACTTCAAGGCTGTCGGATTGGACGATTCTTCAAACCTCTAT
CATTCTGGTGTCAATTGCTCACATTCTAGGTATCTACACCATGGATAA
ATTGGTAGAAGAAGAACACTTTAGGAGGTTCTGGAGCCATGGTTGTTGG
TCATTTCAAGTCCGTTGGTGTCAAGTCTCTTATGAGAACGGTAAGGATGATCCA
TCCAAACCAGCAGGTAAACGCCATGATTGTCTTCACCTGTCTGTTCAATTCTT
TGCATGTACCTGGGCTCCAGGTGTTCTGCGTTGTCTGAAACCTACCCACTT
AGAATTAGATCCAAGGGTATGCCATCGCTCAAGGTTCCAATTGGCTTGGGGTT
TCCTCATTGCCTTCTCACTCCATTATCTCAGGTGCCATTGATTGCGCTACGGT
TACGTCTTATGGGATGTACTCTGTTGCCTTCTTGTGACTTCTCGTTCT
GAAACCAAGGGTCTGCGCTGGAAGACGTTGATGAAGTCTATGAGAACCTACCT
TCGGAAGAGCATATGCATACAGCCACACGATTAAAGACAAGGGCGCCCTATAA

图11

ATGGCCCTATCTCCTACCTATCAGGGCTACATATCTACCACTGGCGACGCGTTGA
TCGTGATCCAGGCAGCTCTAAATAACCATTGAATCTTCTTCCCCGAAGACCAAGA
GAAAGAGAGCGAGATGGGCTAACGATCAGGTAACGTATTGTTTGTGAGC
AACGGTCTCATATCAAACGATGGACCGATGGTATCCCCTGGTCTCCATCTAGAGT
CCTTGGAAAGTTCTTGTATCGGGAACTGGACAAGGATACCCCAAAAACTCG
CAAAGTGACGAAGATACTGAGGAGGGAGAAAGAGGCAGAAGACTTCTGTGGAT
GTAACCGATCCAAATACCAGGCAGTTGGTGGATCATTGGTACTTCCTATGACT
TCAAAGAGGATGGACTTATTAAAAAAACACTCTCCTGACTTCCAGACC GG TGCT
AATGAAGAAAGGGAAACAGTGCACTTGATTAGTTATTACTCCGGAAGATGTAAC
GAACCATCGTTGAACAGGCCGTCTGACAATCCATATCTGGCCAATATCACTGTT
CAGAGTCATTATTGACTGCCTTGAGAGAGAGTACCCCTGGAGGAAGAGCAACGTC
TGATGACGAGCTTCTTAGTCAGAAGTAACCTCGTTAGAGTACCAAGAGGTACCAA
TGAACATATCTATGTCTTACCTTATCAACTCCACTTCCCTGAACACAGGAGTAA
ACTCAACTACCCAGCTGCAACAGCAACA ACTACAACAACAACAGCAACAGCA
ACAGCAGCAGCAACAAACAGCAGCAACAACAGCAACCCGGTAGCATCCCTCCAAA
ATTGATGGATCCTTCTATTACAACAGGGTGTATTCCAGTTCCCTCATGGAA
CCAAAAAAATGGGAAGTAGCAATTGTGGATTAACAATTGGTTCTGTCCAAATTGTT
CAGAATCAAATGGCTATCGTTATCGGACCTCACAGGGATATGACGAACAAAG
TCCAGCAACGAGTTACTTGAATGAACGTTGA

图12

CTTTTTGTAGAAATGTCTTGGTGCCTCGTCCAATCAGGTAGCCATCTCTGAAAT
ATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGG
TAAAACCTAAATGTGGAGTAATGGAACCAGAAACGTCTTCCCTCTCTCCCT
CCACCGCCCGTTACCGTCCCTAGGAAATTACTCTGCTGGAGAGCTTCTTCTAC
GGCCCCCTGCAGCAATGCTCTTCCCAGCATTACGTTGCAGGGTAAAACGGAGGT
CGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAA
TAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAAGAATCGAATAT
AAAAGGCGAACACCTTCCCAATTGGTTCTCCTGACCCAAAGACTTAAATT
AATTGATGGATCCTTCTATTCAATTGAACAACTACCTGCAGGCC

图13

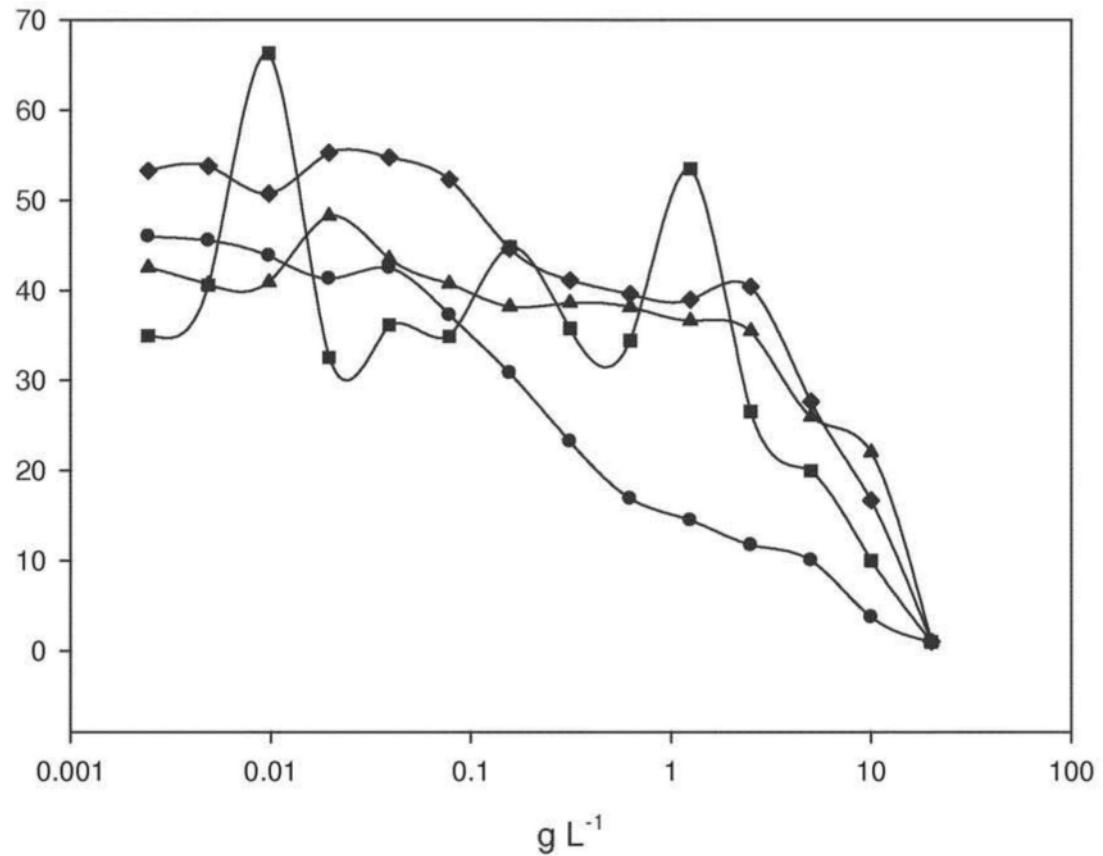


图14

pG1a (SEQ ID 41)

GGAATCTGTATTGTTAGAAAGAACGAGAGTTTACGGCGCCGCCATATTGGGC
CGTGTGAAAACAGCTGAAACCCCCACTACTTCAAAGGTTCTGTTGCTATACACGA
ACCATGTTAACCAACCTCGCTTGTACTGACTGAAGTCATCGGTTAACATCAA
GTACCCTAGTCTGTCTGAATGCTCCTTCCATATTCAAGTAGGTGTTCTGCACCT
TTGCATGCACTGCGGAAGAATTAGCCAATAGCGCGTTCATATGCGCTTTACCC
CCTCTTTGTCAAGCGAAAATGCCTGTAAGATTGGTGGGGGTGTGAGCCGTTA
GCTGAAGTACAACAGGCTAATTCCCTGAAAAAACTGCAGATAAGACTTCAAGATCTC
AGGGATTCCCACATTGGTATTCTGATATGTTTCTGATATGCATCAAAACTCT
AATCTAAAACCTGAATCTCCGCTATTTTTTTTTTGATGACCCGTTTCGT
GACAAATTAAATTCCAACGGGGCTTGTCCGGATAAGAGAATTGGTATTATC
CGTTCGGATAATGGACGCCTGCTCCATATTTCCGGTTATTACCCACCTGGAA
GTGCCAGAATTTCGGGGATTACGGATAATACGGTGGCTGGATTAAATA
CGCCAAGTCTTACATTGTTGCAGTCTCGTGCAGTATGTGCAATAATAACAAG
ATGAGCCAATTATTGGATTAGTTGCAGCTGACCCGCCATAGCTAGGCATAGC
CAAGTGCTATGGGTGTTAGATGATGCACTTGGATGCAGTGAGTTGGAGTATAA
AAGATCCTTAAAATTCCACCCCTT

图15

pG1b (SEQ ID 42)

CCATATTCACTAGGTGTTCTGCACCTTGCATGCACTGCGGAAGAATTAGCAA
TAGCGCGTTCATATGCGCTTACCCCTCTTGTCAAGCGAAAATGCCTGTA
AGATTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCCTGAA
AAAACGAGATAGACTCAAGATCTCAGGGATTCCCACATTGGTATTCTGATA
TGTTTCTGATATGCATCAAAACTCTAACCTAAACCTGAATCTCCGCTATTTT
TTTTTTTTGATGACCCGTTTCGTGACAAATTAAATTCCAACGGGGTCTGTC
CGGATAAGAGAATTGGTATTACCCACCTGGAAGTGCCCAGAATTCCGGGATTACGGA
TAATACGGTGGTCTGGATTAATTAAACGCCAAGTCTTACATTGGTGCAGTCTC
GTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTATTGGATTAGTTGCAGC
TTGACCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACT
TGGATGCAGTGAGTTGGAGTATAAAAGATCCTAAAATTCCACCCCTT

pG1c (SEQ ID 43)

CTGCAGATAGACTCAAGATCTCAGGGATTCCCACATTGGTATTCTGATATGTT
TTCCCTGATATGCATCAAAACTCTAACCTGAATCTCCGCTATTTTTTT
TTTTTGATGACCCGTTTCGTGACAAATTAAATTCCAACGGGGTCTGCTCCGGA
TAAGAGAATTGGTATTACCCACCTGGAAGTGCCCAGAATTCCGGGATTACGGATAATA
CCGGTTATTACCCACCTGGAAGTGCCCAGAATTCCGGGATTACGGATAATA
CGGTGGTCTGGATTAATTAAACGCCAAGTCTTACATTGGTGCAGTCTCGTGCG
AGTATGTGCAATAATAACAAGATGAGCCAATTATTGGATTAGTTGCAGCTTGAC
CCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACTGGAT
GCAGTGAGTTGGAGTATAAAAGATCCTAAAATTCCACCCCTT

图15(续)

pG1d (SEQ ID 44)

GACCCCGTTTCGTGACAAATTAAATTCCAACGGGGTCTTGTCCGGATAAGAGAA
TTTGTTGATTATCCGTCGGATAAATGGACGCCCTGCTCCATATTTCCGGTTAT
TACCCCACCTGGAAGTGCCAGAATTTCGGGGATTACGGATAATACGGTGGTC
TGGATTAATTAATACGCCAAGTCTTACATTTGTTGCAGTCTCGTGCAGTATGTG
CAATAATAACAAGATGAGCCAATTATTGGATTAGTTGCAGCTGACCCGCCAT
AGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACTTGGATGCAGTGAG
TTTGGAGTATAAAAGATCCTAAAATTCCACCCCTT

pG1e (SEQ ID 45)

CCGGATAAGAGAATTGTTGATTATCCGTCGGATAAATGGACGCCCTGCTCCAT
ATTTTCCGGTTATTACCCCCACCTGGAAGTGCCAGAATTTCGGGGATTACGG
ATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTGTTGCAGTCT
CGTGCAGTATGTGCAATAATAACAAGATGAGCCAATTATTGGATTAGTTGCAG
CTTGACCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCAC
TTGGATGCAGTGAGTTGGAGTATAAAAGATCCTAAAATTCCACCCCTT

pG1f (SEQ ID 46)

GCCTGCTCCATATTTCCGGTTATTACCCCCACCTGGAAGTGCCAGAATTTCGG
GGGATTACGGATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTT
GTTGCAGTCTCGTGCAGTATGTGCAATAATAACAAGATGAGCCAATTATTGG
TTAGTTGCAGCTTGACCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTA
GATGATGCACTTGGATGCAGTGAGTTGGAGTATAAAAGATCCTAAAATTCCAC
CCTT

图15(续)