



## (12) 发明专利

(10) 授权公告号 CN 101952452 B

(45) 授权公告日 2014. 07. 09

- (21) 申请号 200880126568. 0 7-27, 124 段 .  
US 20060117414 A1, 2006. 06. 01, 说明书第 137 段 .
- (22) 申请日 2008. 12. 11
- (30) 优先权数据  
61/007, 333 2007. 12. 11 US  
WO 2007136762 A2, 2007. 11. 29, 权利要求 34, 说明书第 2-7, 17, 18 页 .
- (85) PCT 国际申请进入国家阶段日  
2010. 08. 11  
Binderup et al.. Limited proteolysis of branching enzyme from Escherichia coli. 《Arch Bioch Biophys 》. 2000, 第 73-78 页 .
- (86) PCT 国际申请的申请数据  
PCT/US2008/086485 2008. 12. 11  
审查员 曹丙洲
- (87) PCT 国际申请的公布数据  
W02009/076559 EN 2009. 06. 18
- (73) 专利权人 合成基因组公司  
地址 美国加利福尼亚州
- (72) 发明人 P · G · 罗斯勒 陈游 刘博  
C · N · 多吉
- (74) 专利代理机构 北京纪凯知识产权代理有限公司 11245  
代理人 赵蓉民 张全信
- (51) Int. Cl.  
C12P 7/64 (2006. 01)
- (56) 对比文件  
US 20070087420 A1, 2007. 04. 19, 说明书第

权利要求书2页 说明书48页 附图5页

(54) 发明名称  
通过光合微生物分泌脂肪酸

(57) 摘要  
描述了将无机碳转化为分泌的脂肪酸的重组光合微生物。也描述了不需要收获细胞的情况下从培养基中回收分泌的脂肪酸的方法。

1. 重组蓝细菌的细胞培养物,所述蓝细菌被修饰为含有包括产生至少一种外源酰基-ACP 硫酯酶的至少一种重组表达系统的核酸分子,

其中所述酰基-ACP 硫酯酶优先地释放包含 6-20 个碳的脂肪酸链,

其中所述蓝细菌被进一步地修饰,使得编码酰基-ACP 合成酶的一个或多个基因失活或者下调,或者所述合成酶被抑制,

其中培养基提供无机碳实质上作为唯一的碳源,并且

其中所述蓝细菌分泌由所述酰基-ACP 硫酯酶释放的脂肪酸到所述培养基中。

2. 根据权利要求 1 所述的培养物,其中所述至少一种外源酰基-ACP 硫酯酶是 Fat B 硫酯酶。

3. 根据权利要求 1 所述的培养物,其中所述至少一种外源酰基-ACP 硫酯酶是源于萼距花属的 Fat B 硫酯酶。

4. 根据权利要求 1 所述的培养物,其中所述至少一种外源酰基-ACP 硫酯酶是 ChFatB2。

5. 根据权利要求 1 所述的培养物,其中所述重组蓝细菌被进一步修饰以产生外源  $\beta$ -酮脂酰合酶 (KAS)。

6. 根据权利要求 5 所述的培养物,其中所述外源 KAS 优先地产生具有所述硫酯酶对其具有优先活性的链长度的酰基-ACP。

7. 根据权利要求 1 所述的培养物,其中所述重组蓝细菌被进一步地修饰,使得编码  $\beta$ -氧化通路酶的一个或多个基因失活或者下调,或者所述酶被抑制。

8. 根据权利要求 1 所述的培养物,其中所述重组蓝细菌被进一步地修饰,使得编码参与糖类生物合成的酶的一个或多个基因失活或者下调,或者所述酶被抑制。

9. 根据权利要求 8 所述的培养物,其中参与糖类生物合成的所述酶是分支酶。

10. 将无机碳转化为脂肪酸的方法,所述方法包括:

温育权利要求 1 所述的培养物,使得其中的重组蓝细菌分泌所述脂肪酸到培养基中;  
和

从所述培养基中回收所述分泌的脂肪酸。

11. 根据权利要求 10 所述的方法,其中通过使所述培养基与颗粒吸附剂接触,从所述培养物中回收所述脂肪酸。

12. 根据权利要求 11 所述的方法,其中所述颗粒吸附剂在所述培养基中循环。

13. 根据权利要求 11 所述的方法,其中所述颗粒吸附剂被包含在固定床柱中。

14. 根据权利要求 13 所述的方法,其中所述培养基的 pH 在所述接触期间降低。

15. 根据权利要求 14 所述的方法,其中所述 pH 降低方法包括加入  $\text{CO}_2$ 。

16. 根据权利要求 15 所述的方法,其中所述培养基被再循环到所述培养物。

17. 根据权利要求 11 所述的方法,其中所述颗粒吸附剂是亲脂性的。

18. 根据权利要求 11 所述的方法,其中所述颗粒吸附剂是离子交换树脂。

19. 根据权利要求 1 所述的培养物,其中所述蓝细菌是选自下述的属:管孢藻属、色球藻属、蓝细菌属、蓝菌属、蓝杆藻属、蓝纤维藻属、Gloeobacter、粘球藻属、粘杆藻属、微囊藻属、原绿球藻属、原绿藻属、聚球藻属、集胞藻属、蓝囊胞菌属、色球藻属、蓝细菌属 (Stanieria)、异球藻属、拟色球藻属、黏囊藻属、宽球藻属、节螺藻属、博氏藻属、发毛针藻

属、吉特勒氏线状蓝细菌属、盐螺旋藻属、瘦鞘丝藻属、湖生蓝丝藻、鞘丝藻属、微鞘藻属、颤藻属、浮丝藻属、原绿发藻属、假鱼腥藻、螺旋藻属、斯塔氏蓝细菌属、束藻属、束毛藻属、灰线蓝细菌属、鱼腥藻属、项圈藻属、束丝藻属、眉藻属、蓝螺菌属、拟柱孢藻属、筒孢藻属、节球藻属、念珠藻属、胶须藻属、伪枝藻属、单歧藻属、拟绿胶蓝细菌属、费氏藻属、吉特勒氏蓝细菌属、Iyengariella、拟念珠藻属和真枝藻属。

20. 根据权利要求1所述的培养物,其中所述外源酰基-ACP硫酯酶是植物酰基-ACP硫酯酶。

21. 根据权利要求20所述的培养物,其中所述外源植物酰基-ACP硫酯酶源自拟南芥 (At);甘蓝型油菜 (*Brassica napus*) (Bn);Cinnamomum camphorum (Cc);黄灯笼辣椒 (Cch);萼距花 (Ch);披针叶萼距花 (CI);*Cuphea palustris* (Cp);芫荽 (Cs);红花 (Ct);*Cuphea wrightii* (Cw);油棕 (Eg);陆地棉 (Gh);山竹 (Gm);向日葵 (Ha);德国鸢尾 (Ig);鸢尾 (It);肉豆蔻 (Mt);小麦 (Ta);美洲榆 (Ua);*Cuphea calophylla* (Cc) 或加州月桂 (Uc)。

22. 根据权利要求1所述的培养物,其中所述外源酰基-ACP硫酯酶优先释放包含8、10、12、14、16或18个碳的脂肪酸链。

23. 根据权利要求5所述的培养物,其中所述外源 $\beta$ -酮脂酰合酶是 $\beta$ -酮脂酰合酶III或 $\beta$ -酮脂酰合酶IV。

24. 根据权利要求8所述的培养物,其中所述参与糖类生物合成的酶是葡聚糖合酶。

25. 根据权利要求9所述的培养物,其中所述分支酶是1,4- $\alpha$ -葡聚糖分支酶。

26. 根据权利要求11所述的方法,其中所述颗粒吸附剂的平均直径范围是0.5mm至30mm。

27. 根据权利要求11所述的方法,其中所述颗粒吸附剂包括聚乙烯、聚乙烯衍生物、聚苯乙烯、聚苯乙烯衍生物、聚酰胺、聚酰胺衍生物、聚酯、聚酯衍生物、聚氨酯、聚氨酯衍生物、聚丙烯酸酯、聚丙烯酸酯衍生物、聚硅氧烷、聚硅氧烷衍生物、多糖、多糖衍生物、玻璃或陶瓷材料。

28. 根据权利要求11所述的方法,其中所述颗粒吸附剂用无机化合物包被。

29. 根据权利要求28所述的方法,其中所述颗粒吸附剂用氢氧化铝、石墨、无烟煤或硅石包被。

30. 根据权利要求17所述的方法,其中所述颗粒吸附剂通过加入醚连接的烷基或者经过酯、硫代酸酯或酰胺键连接的酰基链进行修饰。

31. 根据权利要求30所述的方法,其中所述颗粒吸附剂通过加入链长度为8-30个碳的烷基进行修饰。

## 通过光合微生物分泌脂肪酸

[0001] 相关申请的交叉引用

[0002] 本申请要求在 2007 年 12 月 11 提交的临时申请 61/007,333 的权益。该申请的内容通过引用并入本文。

### 技术领域

[0003] 本申请涉及将无机碳转化为脂肪酸并分泌它们进入培养基的光合微生物、使用这种微生物生产脂肪酸的方法和其用途。脂肪酸可以直接地使用或者可以进一步地修饰成为替代形式如酯、还原形式如醇、或烃类,用于在不同工业中的应用,包括燃料和化学制品。

### 背景技术

[0004] 光合微生物——包括真核藻类和蓝细菌——包含各种脂类,包括极性脂类和中性脂类。极性脂类(例如,磷脂类、糖脂类、硫脂类)一般存在于结构膜中,而中性脂类(例如三酰甘油类、蜡酯类)积聚在细胞质油体内或者脂肪球内。大量研究工作专注于开发从光合微生物中产生脂基燃料和化学制品的方法。一般地,真核微藻类在营养充分的条件下生长,直至获得特定的细胞密度,在其后,细胞在营养缺乏的条件下进行生长,这经常导致中性脂类的积聚。细胞然后通过各种方法(例如,沉降,其可以通过加入絮凝剂来促进,然后进行离心)收获、干燥,然后脂类通过使用各种非极性溶剂从细胞中提取。细胞的收获和脂类的提取是成本密集型步骤。期望从光合微生物中获得脂类,而不需要进行细胞收获和提取。

[0005] PCT 公开号 W02007/136762 和 W02008/119082 描述了使用微生物生产生物燃料组分。这些文献公开了通过这些微生物生产脂肪酸衍生物——显然地,其是短链醇和长链醇、烃、脂肪醇和包括蜡、脂肪酸酯或者脂肪酯的酯。在描述脂肪酸生产的方面,提议作为这些衍生物的中间物,因此不分泌脂肪酸。而且,没有公开使用在包含无机碳作为主要碳源的培养基中生长的光合微生物将无机碳直接转化为分泌的脂肪酸。本发明利用光合微生物分泌脂肪酸进入培养基的功效,以回收这些有价值的化合物。

[0006] 本发明包括在光合微生物中表达异源酰基-ACP 硫酯酶(TE)基因。多种这些基因以及它们改变油料种子中的脂代谢的应用已经在先前描述。编码催化合成中的各个步骤和进一步代谢脂肪酸的蛋白质的基因已经被广泛地描述。

[0007] 植物酰基-ACP 硫酯酶的两个功能性类别(不饱和脂肪酸识别的 FatA 对饱和脂肪酸识别的 Fat B)可以基于氨基酸序列比对以及功能进行分组(cluster)。Fat A 显示对 18:1-ACP 明显的偏爱、对 18:0-和 16:0-ACP 具有微小的活性,而 Fat B 主要水解链长度在 8-16 个碳之间的饱和酰基-ACP。数个研究集中于改造植物硫酯酶使之具有完美的或者改变的底物专一性,作为适合特殊种子油类的策略。

[0008] 如在图 1 中显示,脂肪酸合成酶催化这样的重复循环,其中丙二酰基-酰基载体蛋白(ACP)与底物稠合——底物开始为乙酰-CoA——以形成乙酰乙酰基-ACP,释放 CO<sub>2</sub>。乙酰乙酰基-ACP 然后被还原、脱水和进一步还原为丁酰基-ACP,其自身随后可以与丙二酰

基-ACP 稠合,然后重复该循环,在每轮加入一个 2- 碳单元。游离脂肪酸的产生因此将被硫酯酶增强,其将从 ACP 释放脂肪酸自身,打破循环。也就是,酰基-ACP 被阻止重新进入循环。脂肪酸的产生也将通过升高脂肪酸合成酶水平和抑制导致脂肪酸降解或进一步代谢的任何酶来促进。

[0009] 图 2 更详细地描述了顺序形成越来越长链的酰基-ACP。如显示的,图 2 中列出的硫酯酶从 ACP 硫酯中释放脂肪酸。

[0010] 利用这个原理,Dehesh, K. 等, The Plant Journal(1996)9:167-172 描述了 " Production of high levels of octanoic(8:0)and decanoic(10:0)fattyacids in transgenic canola by overexpression of ChFatB2, a thioesterase cDNA from Cuphea hookeriana. " Dehesh, K. 等 Plant Physiology(1996)110:203-210 报道了 " Two novel thioesterases are key determinants of the bimodal distribution of acyl chain length of Cuphea palustris seed oil " 。

[0011] Voelker, T. 等, Science(1992)251:72-74 描述了 " Fatty acid biosynthesis redirected to medium chains in transgenic oilseed plants " 。 Voelker, T. 和 Davies, M., Journal of Bacteriology(1994)176:7320-7327 描述了 " Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase " 。

[0012] 发明公开内容

[0013] 本发明涉及能够分泌源于无机碳的脂肪酸进入培养基的重组光合微生物的产生。也提供了在不需要收获细胞的情况下从培养基中移走分泌的脂肪酸的方法。预期,这些改进将导致以较低的成本从光合微生物中生产脂基燃料和化学制品。另外,本发明使得能够生产所限定链长度的脂肪酸,因此使它们用于多种不同产物——包括燃料和化学制品的配制。

[0014] 二氧化碳(其与碳酸、碳酸氢盐和/或碳酸盐一起定义了术语“无机碳”) 在光合过程中转化为有机化合物。无机碳源包括任何递送无机碳的方式,任选地与不充当主要碳原料、但是仅作为混合物或载体(例如,散发自生物燃料(例如,乙醇)植物、能源植物、石油基精炼物以及大气和地下源)的化合物的任何其它组合相混合。

[0015] 本发明的一个实施方式涉及重组光合微生物的培养,所述生物包括编码至少一种外源酰基-ACP 硫酯酶的至少一种重组表达载体,其中所述至少一种外源酰基-ACP 硫酯酶优先地从这些 ACP 硫酯中释放包含 6 至 20 个碳的脂肪酸链。脂肪酸由作为它们的碳源的无机碳形成,并且所述培养物基本只包含无机碳作为碳源。外源硫酯酶的存在将增加期望的脂肪酸的分泌水平至少 2-4 倍。

[0016] 具体地,在一个实施方式中,本发明涉及重组光合微生物的细胞培养物,其中所述微生物已经被修饰,以含有包括产生至少一种外源酰基-ACP 硫酯酶的至少一种重组表达系统的核酸分子,其中所述酰基-ACP 硫酯酶优先地释放包含 6-20 个碳的脂肪酸链,并且其中所述培养基提供无机碳实质上作为唯一的碳源,和其中所述微生物分泌由酰基-ACP 硫酯酶释放的脂肪酸进入培养基。在可选的实施方式中,硫酯酶优先释放包含 6、7、8、9、10、11、12、13、14、15、16、17、18、19 或 20 个碳的脂肪酸链。

[0017] 在其它方面,本发明涉及通过温育这些培养物和从培养物中回收这些分泌的脂肪

酸来产生期望链长度的脂肪酸的方法。在一个实施方式中,回收使用固体颗粒吸附剂收获分泌的脂肪酸。如此回收的脂肪酸可以被进一步地合成修饰或直接用作生物燃料或化学制品的组分。

[0018] 附图简述

[0019] 图 1 是本领域中已知的脂肪酸合成途径的图。

[0020] 图 2 是本领域中已知的多种链长度的脂肪酸合成的更详细的图。

[0021] 图 3 是识别生产各种链长度脂肪酸的酶类别的脂肪酸生物合成的酶法概述。

[0022] 图 4 是来自培养基的脂肪酸的回收系统的示意图。

[0023] 图 5 显示基于图 4 中原理的实验系统。

[0024] 图 6 显示各种生物的代表性酰基-ACP 硫酯酶。

[0025] 实施发明的方式

[0026] 本发明提供分泌脂肪酸进入培养基中的光合微生物,和从培养基中吸附脂肪酸并收集它们以加工成为燃料和化学制品的方法。由此本发明消除或大大减少了对收获和提取细胞的需求,导致生产成本大量减少。

[0027] 图 2 是本发明的一个方面的概述。如在图 2 中显示,使用光合过程中的多个步骤,将二氧化碳转化为乙酰-CoA。乙酰-CoA 然后通过乙酰-CoA 羧化酶的作用转化为丙二酰-CoA。丙二酰-CoA 然后通过丙二酰 CoA:ACP 转酰酶的作用转化为丙二酰-ACP,所述转酰酶在脂肪酸合成酶逐渐作用时导致两个碳单元的连续加入。在本发明的一个实施方式中,通过提供适当的硫酯酶(图 2 中显示为 Fat B),所述过程在 6、或 8、或 10、或 12、或 14、或 16、或 18 个碳的碳链长度时基本停止。就在该实施方式中进行进一步转化为更长链脂肪酸来说,也可以收集细胞生物质。分泌的脂肪酸可以转化为各种其它形式,包括,例如甲基酯、烷、烯、 $\alpha$ -烯炔和脂肪醇。

[0028] 硫酯酶(Acyl-ACP TE)

[0029] 为了实现游离脂肪酸的分泌,提供生物至少一种用于至少一种硫酯酶的表达系统,所述硫酯酶优先起作用释放期望长度的脂肪酸。编码这种硫酯酶的多种基因在本领域中是可得的。这些中的一些是如下美国专利的主题:

[0030] 实例包括名称为"Plant medium-chain-preferring acyl-ACPthioesterases and related methods"的美国专利 5298421,其描述了从加州月桂(*Umbellularia californica*)未成熟的种子分离酰基-ACP 硫酯酶和编码它的基因。这些硫酯酶和它们的编码基因的其它来源包括:名称为"Plant thioesterase having preferential hydrolase activity toward C12acyl-ACP substrate"的美国专利 5,304,481;名称为"Plant thioesterases"的美国专利 5,344,771;名称为"Medium-chain thioesterases in plants"的美国专利 5,455,167;名称为"Plant thioesterases"的美国专利 5,512,482;名称为"Nucleotide sequences of soybean acyl-ACP thioesterase genes"的美国专利 5,530,186;名称为"Plant medium-chain thioesterases"的美国专利 5,639,790;名称为"C8and C10 medium-chain thioesterases in plants"的美国专利 5,667,997;名称为"Plant acyl-ACP thioesterase sequences"的美国专利 5,723,761;名称为"Plant thioesterases and use for modification offatty acid composition in plant seed oils"的美

国专利 5,807,893;名称为" Production of myristate in plant cells" 的美国专利 5,850,022;名称为" Middle chain-specific thioesterase genes from *Cuphea lanceolata*" 的美国专利 5,910,631;名称为" Specific for palmitoyl, stearoyl and oleoyl- $\alpha$ -thioesters nucleic acid fragments encoding acyl-ACP thioesterase enzymes and the use of these fragments in altering plant oil composition" 的美国专利 5,945,585;名称为" Engineering plant thioesterases for altered substrate specificity" 的美国专利 5,955,329;名称为" Nucleotide sequences of canola and soybean palmitoyl-ACP thioesterase genes and their use in the regulation of fatty acid content of the oils of soybean and canola plants" 的美国专利 5,955,650;和名称为" Acyl-ACP thioesterase nucleic acids from maize and methods of altering palmitic acid levels in transgenic plants therewith" 的美国专利 6,331,664。

[0031] 其它在如下公开文献中描述:

[0032] Dormann, P. 等, *Planta*(1993)189:425-432 描述" Characterization of two acyl-acyl carrier protein thioesterases from developing *Cuphea* seeds specific for medium-chain and oleoyl-acyl carrier protein"。Dormann, P. 等, *Biochimica Biophysica Acta*(1994)1212:134-136 描述" Cloning and expression in *Escherichia coli* of a cDNA coding for the oleoyl-acyl carrier protein thioesterase from coriander (*Coriandrum sativum* L.)"。Filichkin, S. 等, *European Journal of Lipid Science and Technology*(2006)108:979-990 描述" New FATB thioesterases from a high-laurate *Cuphea* species: Functional and complementation analyses"。Jones, A. 等, *Plant Cell*(1995)7:359-371 描述" Palmitoyl-acyl carrier protein (ACP) thioesterase and the evolutionary origin of plant acyl-ACP thioesterases"。Knutzon, D. S. 等, *Plant Physiology*(1992)100:1751-1758 描述" Isolation and characterization of two safflower oleoyl-acyl carrier protein thioesterase cDNA clones"。Slabaugh, M. 等, *The Plant Journal*(1998)13:611-620 描述" Condensing enzymes from *Cuphea wrightii* associated with medium chain fatty acid biosynthesis"。

[0033] 先前没有分离的编码这些酰基-ACP TE 的另外的基因可以从其种子油中天然包含大量中链脂肪酸的植物中分离,包括在樟科 (Lauraceae)、千屈菜科 (Lythraceae)、芸香科 (Rutaceae)、榆科 (Ulmaceae) 和蜡烛树科 (Vochysiaceae) 中的某些植物。一般地,由这些植物种子生产的脂肪酸被酯化为甘油并保留在细胞里面。包含这些产物的种子然后可以被收获并被加工以分离脂肪酸。也可以使用这些酶的其他来源例如细菌。

[0034] 已知来自植物的酰基-ACP TE 基于它们的氨基酸序列和它们对不同链长度和饱和度的酰基-ACP 的专一性可以被分成两大类。"Fat A"型的植物酰基-ACP TE 对于油酰-ACP 具有优先的活性,由此释放油酸——一种在羧基远端具有单个双键九碳的 18 碳脂肪酸。"Fat B"型的植物酰基-ACP TE 对于饱和的酰基-ACP 有优先的活性,并且可以具有宽的或窄的链长度专一性。例如,来自萼距花属 (*Cuphea*) 的不同种的 Fat B 酶已经显示从相应的酰基-ACP 释放长度为八个碳至长度为十六个碳范围的脂肪酸。在以下表 1 中列出

的是数种植物酰基-ACP TE 以及它们的优先底物。(脂肪酸通过标准速记符号 (standard shorthand notation) 命名, 其中在冒号之前的数字代表酰基链长度, 而在冒号之后的数字代表酰基链中的双键的数量)。

[0035] 表 1 植物酰基-ACP 硫酯酶

|        |                                        |             |
|--------|----------------------------------------|-------------|
| [0036] | 山竹( <i>Garcinia mangostana</i> )Fat A  | 18:1 和 18:0 |
|        | 红花( <i>Carthamus tinctorius</i> )Fat A | 18:1        |
|        | 芫荽( <i>Coriandrum sativum</i> ) FatA   | 18:1        |
|        | 萼距花( <i>Cuphea hookeriana</i> ) FatB1  | 16:0        |
|        | 萼距花 FatB2                              | 8:0 和 10:0  |
|        | <i>Cuphea wrightii</i> FatB1           | 12:0 至 16:0 |
|        | <i>Cuphea palustris</i> FatB1          | 8:0 和 10:0  |
|        | <i>Cuphea palustris</i> FatB2          | 14:0 和 16:0 |
|        | <i>Cuphea calophylla</i> FatB1         | 12:0 至 16:0 |
|        | 加州月桂 FatB1                             | 12:0        |
|        | 美洲榆( <i>Ulmus Americana</i> ) FatB1    | 8:0 和 10:0  |

[0037] 表 1 中列出的酶是示例性的, 编码酰基-ACP TE 的许多种另外的基因可以被分离并可以在本发明中使用, 包括但不限于如编码以下酰基-ACP TE 的那些基因 (通过 GenPept 登录号查阅) :

[0038] CAA52069.1, CAA52070.1, CAA54060.1, CAA85387.1, CAA85388.1, CAB60830.1, CAC19933.1, CAC19934.1, CAC39106.1, CAC80370.1, CAC80371.1, CAD32683.1, CAL50570.1, CAN60643.1, CAN81819.1, CA017726.1, CA042218.1, CA065585.1, CA068322.1, AAA33019.1, AAA33020.1, AAB51523.1, AAB51524.1, AAB51525.1, AAB71729.1, AAB71730.1, AAB71731.1, AAB88824.1, AAC49001.1, AAC49002.1, AAC49179.1, AAC49180.1, AAC49269.1, AAC49783.1, AAC49784.1, AAC72881.1, AAC72882.1, AAC72883.1, AAD01982.1, AAD28187.1, AAD33870.1, AAD42220.2, AAG35064.1, AAG43857.1, AAG43858.1, AAG43859.1, AAG43860.1, AAG43861.1, AAL15645.1, AAL77443.1, AAL77445.1, AAL79361.1, AAM09524.1, AAN17328.1, AAQ08202.1, AAQ08223.1, AAX51636.1, AAX51637.1, ABB71579.1, ABB71581.1, ABC47311.1, ABD83939.1, ABE01139.1, ABH11710.1, ABI18986.1, ABI20759.1, ABI20760.1, ABL85052.1, ABU96744.1, EAY74210.1, EAY86874.1, EAY86877.1, EAY86884.1, EAY99617.1, EAZ01545.1, EAZ09668.1, EAZ12044.1, EAZ23982.1, EAZ37535.1, EAZ45287.1, NP\_001047567.1, NP\_001056776.1, NP\_001057985.1, NP\_001063601.1, NP\_001068400.1, NP\_172327.1, NP\_189147.1, NP\_193041.1, XP\_001415703.1, Q39473, Q39513, Q41635, Q42712, Q9SQI3, NP\_189147.1, AAC49002, CAA52070.1, CAA52069.1, 193041.1, CAC39106, CA017726, AAC72883, AAA33020, AAL79361, AAQ08223.1, AAB51523, AAL77443, AAA33019, AAG35064 和 AAL77445。

[0039] 在本发明中有用的酰基-ACP TE 的另外来源包括: 拟南芥 (*Arabidopsis thaliana*) (At); 慢生型大豆根瘤菌 (*Bradyrhizobium japonicum*) (Bj); 甘蓝型油菜 (*Brassica napus*) (Bn); *Cinnamomum camphorum* (Cc); 黄灯笼辣椒 (*Capsicum chinense*) (Cch); 萼距花 (Ch); 披针叶萼距花 (*Cuphea lanceolata*) (CI); *Cuphea palustris* (Cp); 芫荽 (Cs); 红花 (Ct); *Cuphea wrightii* (Cw); 油棕 (*Elaeis guineensis*) (Eg); 陆地棉

(*Gossypium hirsutum*) (Gh); 山竹 (*Gm*); 向日葵 (*Helianthus annuus*) (Ha); 德国鸢尾 (*Iris germanica*) (Ig); 鸢尾 (*Iris tectorum*) (It); 肉豆蔻 (*Myristica fragrans*) (Mf); 小麦 (*Triticum aestivum*) (Ta); 美洲榆 (Ua); 和加州月桂 (Uc)。示例性 TE 与相应 NCBI 登录号一起在图 6 中显示。

[0040] 在一个实施方式中, 本发明考虑一种长度的中链脂肪酸的特定生产, 例如, 在重组光合微生物的一种培养物中主要产生 C8 脂肪酸。在另一种实施方式中, 本发明考虑两种或更多种不同长度脂肪酸的组合物的生产, 例如, 在重组光合微生物的一种培养物中生产 C8 和 C10 脂肪酸。

[0041] 以下说明的是操作这些本领域已知基因以构建合适的表达系统, 该表达系统导致在选择的重组光合生物中产生有效量的硫酯酶。在这种构建物中, 可以期望去除编码质体转运肽区的基因部分, 因为该区在原核生物中是不适当的。可选地, 如果表达发生在真核细胞中, 对宿主生物适当的质体转运肽编码区可以被替换。根据宿主, 也可以使用优选的密码子。

[0042] 其它修饰

[0043] 除了提供一种或多种合适的酰基 -ACP TE 基因的表达系统, 还可以在光合宿主中进行进一步的改变。例如, 可以修饰宿主以包括编码  $\beta$ -酮脂酰合成酶 (KAS) 的异源基因的表达系统, 所述  $\beta$ -酮脂酰合成酶优先地产生具有中等链长度的酰基 -ACP。这些 KAS 酶已经从数种植物进行了描述, 包括萼距花属 (*Cuphea*) 的各个种。参见: Dehesh, K. 等, *The Plant Journal* (1998) 15:383-390, 描述了 "KAS IV: a 3-ketoacyl-ACPsynthase from *Cuphea* sp. is a medium chain specific condensin enzyme." ; Slabaugh, M. 等, *The Plant Journal* (1998) 13:611-620, 并且将用于通过异源中链酰基 -ACP TE 增加识别和裂解合适长度的酰基 -ACP 分子的可用性。另一个实例是包含异源酰基 -ACP TE 基因的光合宿主细胞可以被进一步地修饰, 以包括编码多功能乙酰基 -CoA 羧化酶的异源基因、或编码多亚基型的乙酰基 -CoA 羧化酶的各种亚基的一组异源基因的表达系统。编码另外的酶或者脂肪酸生物合成路径的组分的其它异源基因也可以被引入并在包含酰基 -ACP TE 的宿主细胞中表达。

[0044] 光合微生物也可以被修饰, 使得编码  $\beta$ -氧化通路酶的一种或多种基因失活或者下调, 或者酶自身可以被抑制。这将阻止从酰基 -ACP 中释放的脂肪酸的降解, 因此增加分泌脂肪酸的产量。在期望的产物是中链脂肪酸的情况下, 编码优先使用这些链长度作为底物的酰基 -CoA 合成酶和 / 或酰基 -CoA 氧化酶的基因的失活或者下调将是有利的。编码中链专一性的酰基 -CoA 合成酶和 / 或中链专一性的酰基 -CoA 氧化酶的基因中使得酶的活性减小的突变对于增加分泌脂肪酸的产量也将是有效的。另外的修饰失活或者下调酰基 -ACP 合成酶基因、或者使基因或蛋白质失活。基因中的突变可以通过重组或者非重组的方法引入。这些酶和它们的基因是熟知的, 并且可以通过分裂、缺失、反义序列的生成、或者核糖酶的生成或者实施者已知的其它重组方法被专一地靶向。基因的失活也可以通过随机突变技术如 UV 实现, 并且生成的细胞用于筛选成功的突变体。蛋白质自身也可以通过细胞内生成合适抗体或者细胞内生成肽抑制剂所抑制。

[0045] 光合微生物也可以被修饰, 使得编码贮存糖或者聚羟基链烷酸酯 (polyhydroxyalkanoate, PHA) 生物合成途径酶的一个或多个基因已经失活或者下调, 或者

酶自身可以被抑制。实例包括参与糖原、淀粉、或者金藻昆布多糖合成的酶,包括葡聚糖合酶和分支酶。其它实例包括参与 PHA 生物合成的酶,如乙酰乙酰 -CoA 合酶和 PHA 合酶。

#### [0046] 表达系统

[0047] 通过引入合适的表达载体,使得蓝细菌和真核藻类中的异源基因能够表达。对于蓝细菌的转化,可以使用在蓝细菌中起作用的多种启动子,包括但不限于:通过加入异丙基  $\beta$ -D-1- 硫代半乳糖苷 (IPTG) 可诱导的 lac、tac 和 trc 启动子和衍生物,与转座子-或细菌染色体携带的抗生素抗性基因(新霉素磷酸转移酶、氯霉素乙酰转移酶、壮观霉素腺苷酸转移酶等)相关的启动子,与各种异源细菌和天然蓝细菌基因相关的启动子,病毒和噬菌体的启动子,和合成启动子。从蓝细菌分离、已经成功使用的启动子包括以下:

[0048] secA(分泌,通过细胞的氧化还原状态控制)

[0049] rbc(Rubisco 操纵子)

[0050] psaAB(PSI 反应中心蛋白;光调节的)

[0051] psbA(PSII 的 D1 蛋白;光-可诱导的)

[0052] 同样地,各种不同的转录终止子可以用于表达载体构建。可能的终止子的实例包括但不限于 psbA、psaAB、rbc、secA 和 T7 外被蛋白。

[0053] 表达载体通过标准的方法引入蓝细菌菌株,标准的方法包括但不限于天然的 DNA 摄取、接合、电穿孔、粒子轰击、和用玻璃珠、SiC 纤维或其它粒子磨损。载体可以是:1) 通过包括使同源重组能够进入染色体的侧翼序列,靶向整合入蓝细菌染色体;2) 通过包括使同源重组能够进入内源质粒的侧翼序列,靶向整合入内源蓝细菌质粒;或 3) 设计为使得表达载体在选择的主宿内复制。

[0054] 为了转化绿藻,可以使用在绿藻中起作用的多种基因启动子和终止子,包括但不限于来自衣藻属(*Chlamydomonas*) 和其它藻的启动子和终止子、来自病毒的启动子和终止子、和合成的启动子和终止子。

[0055] 表达载体通过标准方法引入绿藻菌株,标准方法包括但不限于电穿孔、粒子轰击、和用玻璃珠、SiC 纤维或其它粒子磨损。载体可以是:1) 通过包括使同源重组能够进入染色体的侧翼序列,靶向位置专一地整合入绿藻叶绿体染色体;2) 靶向整合入细胞(核定位的)的染色体。

[0056] 为了转化硅藻,在这些表达载体中可以使用在硅藻中起作用的多种基因启动子,包括但不限于:1) 来自海链藻属(*Thalassiosira*) 和其它不等鞭毛藻(heterokont algae) 的启动子、病毒启动子和合成启动子。适合用于表达载体的假微型海链藻(*Thalassiosira pseudonana*) 的启动子包括  $\alpha$ -微管蛋白启动子(SEQ ID NO:1)、 $\beta$ -微管蛋白启动子(SEQ ID NO:2) 和肌动蛋白启动子(SEQ ID NO:3)。适合用于表达载体的三角褐指藻(*Phaeodactylum tricornutum*) 的启动子包括  $\alpha$ -微管蛋白启动子(SEQ ID NO:4)、 $\beta$ -微管蛋白启动子(SEQ ID NO:5) 和肌动蛋白启动子(SEQ ID NO:6)。这些序列从公共数据库中可获得的相关生物的基因组序列推出,并且仅是示例性的可以使用的各种启动子。与这些和其它基因相关的终止子、或特定的异源基因可以用于终止转录和提供多聚腺苷酸的合适信号,并且可以以类似方式获得,或者是本领域已知的。

[0057] 表达载体通过标准方法引入硅藻菌株,标准方法包括但不限于电穿孔、粒子轰击、和用玻璃珠、SiC 纤维或其它粒子磨损。载体可以是:1) 通过包括使同源重组能够进入染色

体的侧翼序列,靶向位置专一地整合入硅藻叶绿体染色体;2) 靶向整合入细胞(核定位的)的染色体。

[0058] 宿主生物

[0059] 用于制备本发明的培养物的宿主细胞包括能够将无机碳转化成底物的任何光合生物,所述底物又被转化为脂肪酸衍生物。这些生物包括原核生物以及真核生物如藻类和硅藻类。

[0060] 宿主生物包括真核藻类和蓝细菌(蓝-绿藻类)。代表性藻类包括绿藻(green algae)((chlorophytes) 绿藻)、红藻、硅藻、绿枝藻类(prasinophytes)、灰色植物类(glaucophytes)、海产单细胞藻类(chlorarachniophytes)、裸藻类(euglenophytes)、有色植物类(chromophytes)和渦鞭毛藻类(dinoflagellates)。许多蓝细菌种类是已知的,并且已经使用分子生物学技术进行操作,包括单细胞蓝细菌集胞藻(*Synechocystis* sp.) PCC6803 和细长聚球藻(*Synechococcus elongates*) PCC7942,它们的基因组已经被全部测序。

[0061] 以下蓝细菌的属可以使用:一组包括

[0062] 管孢藻属(*Chamaesiphon*) 蓝纤维藻属(*Dactylococcopsis*) 原绿球藻属(*Prochlorococcus*)

[0063] 色球藻属(*Chroococcus*) *Gloeobacter* 原绿藻属(*Prochloron*)

[0064] 蓝细菌属(*Cyanobacterium*) 粘球藻属(*Gloeocapsa*) 聚球藻属(*Synechococcus*)

[0065] 蓝菌属(*Cyanobium*) 粘杆藻属(*Gloeotheca*) 集胞藻属(*Synechocystis*)

[0066] 蓝杆藻属(*Cyanothece*) 微囊藻属(*Microcystis*)

[0067] 另一组包括

[0068] 蓝囊胞菌属(*Cyanocystis*) 异球藻属(*Xenococcus*) 宽球藻属(*Pleurocapsa*)

[0069] 色球藻属(*Dermocarpella*) 拟色球藻属(*Chroococcidiopsis*)

[0070] 蓝细菌属(*Stanieria*) 黏囊藻属(*Myxosarcina*)

[0071] 还另一组包括

[0072] 节螺藻属(*Arthrospira*) 湖生蓝丝藻(*Limnothrix*) 假鱼腥藻(*Pseudanabaena*)

[0073] 博氏藻属(*Borzia*) 鞘丝藻属(*Lyngbya*) 螺旋藻属(*Spirulina*)

[0074] 发毛针藻属(*Crinalium*) 微鞘藻属(*Microcoleus*) 斯塔尔氏蓝细菌属(*Starria*)

[0075] 吉特勒氏线状蓝细菌属(*Geitlerinema*) 颤藻属(*Oscillatoria*) 束藻属(*Symploca*)

[0076] 盐螺旋藻属(*Halospirulina*) 浮丝藻属(*Planktothrix*) 束毛藻属(*Trichodesmium*)

[0077] 瘦鞘丝藻属(*Leptolyngbya*) 原绿发藻属(*Prochlorothrix*) 灰线蓝细菌属(*Tychonema*)

[0078] 还另一组包括

[0079] 鱼腥藻属(*Anabaena*) 蓝螺菌属(*Cyanospira*) 念珠藻属(*Nostoc*)

[0080] 项圈藻属(*Anabaenopsis*) 拟柱孢藻(*Cylindrospermopsis*) 胶须藻属(*Rivularia*)

[0081] 束丝藻属(*Aphanizomenon*) 筒孢藻属(*Cylindrospermum*) 伪枝藻属(*Scytonema*)

[0082] 眉藻属(*Calothrix*) 节球藻属(*Nodularia*) 单歧藻属(*Tolypothrix*)

[0083] 和另一组包括

[0084] 拟绿胶蓝细菌属(*Chlorogloeopsis*) 吉特勒氏蓝细菌属(*Geitleria*)

[0085] 拟念珠藻属(*Nostochopsis*) 费氏藻属(*Fischerella*)

[0086] *Iyengariella* 真枝藻属(*Stigonema*)

[0087] 另外,可以使用多种藻类,包括硅藻和绿藻。

[0088] 宿主菌株的期望性质包括:在 25-50℃有高的潜在的生长速度和脂质生产力,高光强度耐受性,在含盐或者盐水即广泛范围的水类型中生长、对高 O<sub>2</sub> 浓度生长抑制的耐性、丝状形态以有助于通过筛选收获;捕食抗性,被絮凝的能力(通过化学制品或者“按需自我絮凝”),优良的无机碳摄取特性,病毒或噬蓝藻体抗性,对游离脂肪酸或者与本发明方法有关的其它化合物的耐性,和进行代谢工程的能力。

[0089] 代谢工程通过电穿孔或者接合摄取 DNA 的能力促进,要求在事件基因置换(event gene replacement)或者基因敲除中缺少限制系统和有效的同源重组。

[0090] 脂肪酸吸附、去除和回收

[0091] 由以上描述的重组光合微生物分泌进入培养基的脂肪酸可以以多种方式回收。可以使用通过利用不混溶溶剂分开的简单分离方法。在一个实施方式中,可以使用颗粒吸附剂。取决于回收方法的设计,这些可以是亲脂性颗粒或者离子交换树脂。它们可以在分开的培养基中循环,然后收集,或者培养基可以经过固定床柱。例如,包含这些颗粒的色谱柱。然后利用合适的溶剂,脂肪酸从颗粒吸附剂中洗脱。可以进行蒸发溶剂、然后进一步处理分离的脂肪酸和脂类,以产生可以用于多种商业目的的化学制品和燃料。

[0092] 颗粒吸附剂可以具有范围为 0.5mm 至 30mm 的平均直径,其可以由各种材料制备,包括但不限于聚乙烯和衍生物、聚苯乙烯和衍生物、聚酰胺和衍生物、聚酯和衍生物、聚氨酯和衍生物、聚丙烯酸酯和衍生物、聚硅氧烷和衍生物、多糖和衍生物。某些玻璃和陶瓷材料也可以用作脂肪吸附体的固体支持组分。颗粒吸附剂的表面可以被修饰,使得它们更能够结合脂肪酸和脂类。这种修饰的一个实例是引入具有不同链长度的醚连接的烷基,链长度优选为 8-30 个碳。在另一个实例中,各种长度的酰基链可以经过酯、硫代酸酯或酰胺键连接到脂肪吸附体的表面。

[0093] 在一个实施方式中,颗粒吸附剂用已知结合脂肪酸和脂类的无机化合物包被。这些化合物的实例包括但不限于氢氧化铝、石墨、无烟煤和硅石。

[0094] 使用的颗粒也可以被磁化或以其它方式衍生化,以促进回收。例如,颗粒可以与结合对中的一个联接,并且吸附到包含相关结合伴侣的基底。

[0095] 脂肪酸可以通过使用合适的溶剂如己烷或乙醇从颗粒吸附剂中洗脱。颗粒吸附剂可以通过使它们返回培养基被再使用或在再生柱中使用。包含溶解的脂肪酸的溶剂然后被蒸发,留下纯化态的脂肪酸,用于进一步地转化为化学制品和燃料。颗粒吸附剂可以被设计为中性有浮力的或者阳性有浮力的,以增强在培养基中的循环。脂肪酸去除和回收的连续循环可以通过利用以上概述的步骤实施。回收的脂肪酸可以转化为可选的有机化合物,直接使用或者与其它组分混合。这种转化的化学方法在本领域中是熟知的,并且也考虑发展用于这种转化的生物学方法。

[0096] 本发明还考虑包括由本文描述的重组光合微生物产生的脂肪酸的多种组合物以及其用途。组合物可以包括脂肪酸自身、或者脂肪酸的进一步衍生物,如醇、烷和烯,其可以通过本领域中已知的任何方法以及通过开发转化的生物学方法,从微生物产生的脂肪酸生产。例如,脂肪酸可以通过催化氢化和催化脱氢转化为链烯烃。

[0097] 组合物可以用作,例如生物质原油。生物质原油可以通过精炼厂加工,通过包括加氢处理、脱羧、异构化和催化裂化和重整的工艺,其将组合物化合物转化为各种石油和石油

化学产品替代物——包括烷、烯烃和芳香类。通过包括酯交换作用和酯化作用的已建立的化学工艺,生物质原油也可以转化为酯基燃料,例如脂肪酸甲基酯(商业上已知为生物柴油)。

[0098] 另外,本领域技术人员可以考虑本发明脂肪酸和其衍生物的多种其它用途,其在本领域中是熟知的,例如,化学制品、肥皂、表面活性剂、去垢剂、润滑剂、保健食品、药品、化妆品等的生产。例如,本发明的脂肪酸的衍生物包括 C8 化学制品,如辛醇,其在用于化妆品和调料的酯的生产以及各种医学应用中使用,和辛烷,主要用作聚乙烯生产中的共聚单体。本发明脂肪酸的衍生物也可以包括 C10 化学制品,如癸醇,在增塑剂、表面活性剂和溶剂的制造中使用,和癸烯,在润滑剂的制造中使用。

[0099] 生物质原油是生物学产生的化合物或者不同的生物学产生化合物的混合物,它们可以用作精炼厂的原料,以代替或补充原油或其它形式石油。一般而言,但不是必须,这些原料已经通过生物学、化学、机械或热工艺预处理,以成为合适引入石油精炼厂的液态。

[0100] 本发明的脂肪酸可以是生物质原油、并进一步被加工为生物燃料组合物。生物燃料然后可以实现作为成品燃料或者燃料添加剂。

[0101] “成品燃料”被定义为化学化合物或者化学化合物的混合物(通过化学、热化学或者生物途径产生),其处于合适的化学和物理状态被直接用作发动机中的纯燃料或燃料添加剂。在多种情况中,但不总是,在发动机应用中使用的成品燃料的适宜性通过描述需要满足的必须的物理和化学性质的规范确定。发动机的一些实例是:内燃机、燃气涡轮机、蒸气涡轮机、外燃发动机和蒸汽锅炉。成品燃料的一些实例包括:压缩点火(柴油机)内燃机中使用的柴油,在航空涡轮机中使用的喷气燃料,在锅炉中使用以生成蒸气的燃料油或在外燃机中使用的燃料油,在弹性燃料发动机中使用的乙醇。燃料规范的实例是主要在美国使用的 ASTM 标准和主要在欧洲使用的 EN 标准。

[0102] “燃料添加剂”指由于各种原因与另一种燃料组合使用的化合物或组合物,所述原因包括但不限于符合使用生物燃料的任务、减少化石燃料来源的产品的消耗或者增强燃料或发动机的性能。例如,燃料添加剂可以用于改变凝固点/胶凝点、浊点、润滑性、粘度、氧化稳定性、点火性能、辛烷值和闪点。添加剂还可以作为抗氧化剂、破乳剂、增氧剂、热稳定性改良剂、十六烷值改进剂、稳定剂、低温流动改进剂、燃烧改良剂、消泡剂、消雾添加剂(anti-haze additives)、结冰抑制剂、喷油器清洁添加剂、抑烟剂、减阻添加剂、金属钝化剂、分散剂、去垢剂、破乳剂、染料、标记物、静电消除剂(static dissipaters)、杀菌剂和/或腐蚀抑制剂起作用。

[0103] 提供以下实施例以说明本发明、但不限制本发明。

[0104] 实施例 1

[0105] 由源自单细胞光自养蓝藻类细菌细长聚球藻 (*Synechococcus elongates*) PCC 7942 的菌株分泌脂肪酸

[0106] 修饰编码酰基-ACP 硫酯酶的萼距花 (*Cuphea hookeriana*) Fat B2 基因 (ChFatB2), 用于优化细长聚球藻 PCC 7942 中的表达。首先, 去除编码天然 ChFatB2 蛋白的质体转运肽区的基因部分。编码区的其余部分随后使用由 DNA2.0, Inc 提供的“Gene Designer”软件程序(版本 1.1.4.1)进行密码子优化。该 ChFatB2 基因的衍生物的核苷酸序列(此后称为 ChFatB2-7942)作为 SEQ ID NO:7 提供。该基因编码的蛋白质序列在 SEQ

ID NO :8 中提供。

[0107] 两种不同形式的 *trc* 启动子, *trc* (Egon, A. 等, *Gene* (1983) 25 :167-178) 和“增强”的 *trc*” (在此后称为 *trcE*, 来自 *TrcHis A*, Invitrogen) 被用于驱动细长集球藻 PCC 7942 中的 *ChFatB2-7942* 的表达。*Trc* 启动子通过由 *lacIq* 基因编码的 *Lac* 阻抑蛋白阻抑, 并且可以通过加入异丙基  $\beta$ -D-1- 硫代半乳糖苷 (IPTG) 诱导。*trcE* 启动子是旨在促进大肠杆菌 (*E. coli*) 中真核蛋白质的表达的 *trc* 的衍生物, 并且也可以被 IPTG 诱导。

[0108] 可操作地连接到 *trc* 或 *trcE* 的 *ChFatB2-7942* 与 *lacIq* 基因的融合片段被克隆入穿梭载体 pAM2314 (Mackey, S. R. 等, *Methods Mol. Biol.* (2007) 362 :115-129), 其使得细长聚球藻 PCC 7942 的转化能够经由双同源重组调节的整合进入染色体的“NS1”位置。包含 *trcE::ChFatB2-7942* 表达盒和 *lacIq* 基因的构建质粒被命名为 pSGI-YC01。SEQ ID NO :9 代表在 pSGI-YC01 的 NS1 重组位点之间并且包括 pSGI-YC01 的 NS1 重组位点的序列。包含 *trc::ChFatB2-7942* 表达盒和 *lacIq* 基因的构建质粒被命名为 pSGI-YC09。SEQ ID NO :10 代表在 pSGI-YC09 的 NS1 重组位点之间并且包括 pSGI-YC09 的 NS1 重组位点的序列。

[0109] 质粒 pSGI-YC01 和 pSGI-YC09 中的每一个与对照载体 pAM2314 一起被引入野生型细长聚球藻 PCC 7942 细胞中, 如 Golden 和 Sherman (*J. Bacteriol.* (1984) 158 :36-42) 所描述。在旋转振荡器上 (150rpm)、在 30°C, 用持续光照 ( $60 (\mu E m^{-2} sec^{-1})$ ), 重组和对照菌株都在 100ml 补充有壮观霉素 (5mg/L) 的 BG-11 培养基中预培养到晚对数期 ( $OD_{730nm} = 1.0$ )。培养物然后在 BG-11 中以起始  $OD_{730nm} = 0.4-0.5$  进行传代培养, 和培养过夜到  $OD_{730nm} = 0.7-0.9$ 。对于时间-进程研究, 60ml 等分试样的培养物转移入 250ml 烧瓶中, 并且如果合适的话, 通过加入 IPTG (终浓度 = 1mM) 诱导。培养物在 IPTG 诱导后的 0、48、96 和 168 小时取样, 然后使用 Millipore 真空过滤器歧管经过 Whatman® GF/F 过滤器进行过滤。滤液在螺旋端部培养管中收集, 用于气相色谱 (GC) 分析。

[0110] 使用液-液萃取, 从过滤的细胞培养物分离游离脂肪酸 (FFA)。5mL 的滤液与 125  $\mu L$  的 1M  $H_3PO_3$  和 0.25mL 的 5M NaCl 混合, 然后加入 2mL 的己烷并完全混合。对于 GC-FID 分析, 使用 40 : 1 的分流比, 将 0.2  $\mu L$  的己烷样品注射到 DB-FFAP 柱 (J&W Scientific, 15m  $\times$  250  $\mu m \times 0.25 \mu m$ ) 上, 温度曲线 (temperature profile) 为: 在 150°C 起始 0.5min, 然后以 15°C /min 加热至 230°C 并保持 7.1min (1.1mL/min He)。

[0111] 表明在 IPTG 诱导后 168 小时, 包含各种细长聚球藻菌株的培养物中的中链 FFA (8:0 和 10:0) 的水平 GC 分析结果示于表 1-1 中。

[0112] 表 1-1 在细长聚球藻的各种菌株中的中链脂肪酸的分泌

|        | 菌株         | 亲代菌株     | 加入的质粒     | 转基因                       | 脂肪酸 (mg/L) |      |
|--------|------------|----------|-----------|---------------------------|------------|------|
|        |            |          |           |                           | 8:0        | 10:0 |
| [0113] | SGC-YC2-5  | PCC 7942 | pAM2314   | 没有                        | ND         | ND   |
|        | SGC-YC1-2  | PCC 7942 | pSGI-YC01 | <i>trcE::ChFatB2-7942</i> | 1.5        | 3.5  |
|        | SGC-YC14-4 | PCC 7942 | pSGI-YC09 | <i>trc::ChFatB2-7942</i>  | 5.1        | 10.1 |

[0114] 说明: ND 表示“没有检测到” (< 1mg/L)。

[0115] 实施例 2

[0116] 由源自单细胞光自养蓝藻类细菌集胞藻 (*Synechocystis* sp.) PCC6803 的菌株分

泌脂肪酸

[0117]  $\text{trcE}::\text{ChFatB2-7942}$  和  $\text{trc}::\text{ChFatB2-7942}$  与  $\text{lacIq}$  基因的融合片段被克隆入穿梭载体 pSGI-YC03 (SEQ ID NO:11), 其使得集胞藻 PCC 6803 的转化能够经由双同源重组介导的整合进入染色体的 "RSI" 位点 (Williams, Methods Enzymol. (1988) 167:766-778)。包含  $\text{trcE}::\text{ChFatB2-7942}$  表达盒和  $\text{lacIq}$  基因的构建质粒被命名为 pSGTYC08。SEQ ID NO:12 代表在 pSGI-YC08 的 RSI 重组位点之间并包括 pSGI-YC08 的 RSI 重组位点的序列。包含  $\text{trc}::\text{ChFatB2-7942}$  表达盒和  $\text{lacIq}$  基因的构建质粒被命名为 pSGI-YC14。SEQ ID NO:13 代表在 pSGI-YC14 的 NS1 重组位点之间并且包括 pSGI-YC14 的 NS1 重组位点的序列。

[0118] 质粒 pSGI-YC08、pSGI-YC14 中的每一个和对照载体 pSGI-YC03 被引入野生型集胞藻 PCC 6803 细胞中, 如 Zang, X. 等, J Microbiol. (2007) 45:241-245 所描述。在旋转振荡器上 (150rpm)、在 30°C, 用持续光照 ( $60(\mu\text{E m}^{-2}\text{sec}^{-1})$ ), 重组和对照菌株都在 100ml 补充有卡那霉素 (10mg/L) 的 BG-11 培养基中预培养到晚对数期 ( $\text{OD}_{730\text{nm}} = 1.0$ )。培养物然后在 BG-11 中、以初始  $\text{OD}_{730\text{nm}} = 0.4-0.5$  进行传代培养, 并且培养过夜到  $\text{OD}_{730\text{nm}} = 0.7-0.9$ 。对于时间-进程研究, 60ml 等分试样的培养物转移入 250ml 烧瓶中, 并且如果合适的话, 通过加入 IPTG (终浓度 = 1mM) 进行诱导。培养物在 IPTG 诱导后的 0、72 和 144 小时取样, 然后使用 Millipore 真空滤器歧管经过 Whatman® GF/F 滤器进行过滤。滤液在螺旋端部培养管中收集, 用于气相色谱 (GC) 分析。通过液-液萃取, 从过滤的培养物上清液溶液中分离游离脂肪酸 (FFA)。对于每个样品, 2mL 过滤培养物用 50  $\mu\text{L}$  磷酸 (1M)、100  $\mu\text{L}$  的 NaCl (5M) 和 2mL 己烷的混合液萃取。使用 40:1 的分流比, 将 0.2  $\mu\text{L}$  样品注射到 DB-FFAP 柱 (J&WScientific, 15m $\times$ 250  $\mu\text{m}\times$ 0.25  $\mu\text{m}$ ) 上, 温度曲线为: 在 150°C 起始 0.5min、然后以 15°C/min 加热至 230°C 并保持 7.1min (1.1mL/min He)。

[0119] 表明在 IPTG 诱导后 144 小时、中链 FFA (8:0 和 10:0) 的水平 GC 分析结果在表 2-1 中显示。

[0120] 表 2-1 集胞藻的各种菌株中的中链脂肪酸分泌。

[0121]

| 菌株         | 亲代菌株     | 加入质粒      | 转基因                                | 脂肪酸(mg/L) |      |
|------------|----------|-----------|------------------------------------|-----------|------|
|            |          |           |                                    | 8:0       | 10:0 |
| SGC-YC9-8  | PCC 6803 | pSGI-YC03 | 没有                                 | ND        | ND   |
| SGC-YC10-5 | PCC 6803 | pSGI-YC08 | $\text{trcE}::\text{ChFatB2-7942}$ | 61.3      | 52.7 |
| SGC-YC16-2 | PCC 6803 | pSGI-YC14 | $\text{trc}::\text{ChFatB2-7942}$  | 2.7       | 5.8  |

[0122] 说明: ND 代表 "没有检测到" ( $< 1\text{mg/L}$ )。

[0123] 实施例 3

[0124] 由源自丝状蓝藻类细菌多变鱼腥藻 (*Anabaena variabilis*) ATCC29413 的菌株分泌脂肪酸

[0125] 分别使用来自 pSGI-YC14 和 pSGI-YC08 的引物 RS3-3F (SEQ IDNO:14) 和 4YC-rmBter-3 (SEQ ID NO:15), PCR 扩增  $\text{trc}::\text{ChFatB2-7942}$  和  $\text{trcE}::\text{ChFatB2-7942}$  与  $\text{lacIq}$  基因的融合片段, 然后克隆入穿梭载体 pEL17 中, 其使多变鱼腥藻 ATCC 29413 的转化能够经由双同源重组介导的整合进入染色体的  $\text{nifU1}$  基因座 (Lyons 和 Thiel,

J. Bacteriol. (1995) 177 :1570-1575)。构建的质粒对于 *trc::ChFatB2-7942* 和 *trcE::ChFatB2-7942* 分别命名为 pSGI-YC69 and pSGI-YC70。

[0126] 质粒 pSGI-YC69、pSGI-YC70 中的每一个与对照载体 pEL17 一起经由三次亲代接合,被引入野生型多变鱼腥藻 ATCC 29413 细胞中,如由 Elhai 和 Wolk (Methods Enzymol. (1988) 167 :747-754) 所描述的。在旋转振荡器上 (150rpm)、在 30℃,用持续光照 ( $60(\mu\text{E m}^{-2} \text{ sec}^{-1})$ ),重组菌株和对照菌株都在 100ml 补充有 5mM  $\text{NH}_4\text{Cl}$  和壮观霉素 (3mg/L) 的 BG-11 培养基中预培养到晚对数期 ( $\text{OD}_{730\text{nm}} = 1.0$ )。培养物然后在 BG-11 中、以初始  $\text{OD}_{730\text{nm}} = 0.4-0.5$  进行传代培养,并培养过夜到  $\text{OD}_{730\text{nm}} = 0.7-0.9$ 。对于时间-进程研究,60ml 等分试样的培养物转移入 250ml 烧瓶中,并且如果合适的话,通过加入 IPTG (终浓度 = 1mM) 进行诱导。培养物每隔 72 小时取样,然后使用 Millipore 真空过滤器歧管经过 Whatman® GF/F 过滤器进行过滤。滤液在螺旋端部培养管中收集,用于气相色谱 (GC) 分析,如在实施例 1 中描述的。

[0127] 实施例 4

[0128] 由源自包含失活的酰基-ACP 合成酶基因的细长聚球藻 PCC 7942 的菌株分泌脂肪酸

[0129] 细长聚球藻 PCC7942, synpcc7942\_0918 (Cyanobase 基因命名) 中的推定的酰基-ACP 合成酶基因经由用携带氯霉素抗性标记基因 *cat* (其编码氯霉素乙酰基转移酶) 的 1,741-bp DNA 序列替换其编码区的内部 422-bp 部分而断裂。引物对 918-15 (SEQ ID NO : 16)/918-13 (SEQ ID NO : 17) 和 918-25 (SEQ ID NO : 18)/918-23 (SEQ ID NO : 19) 被用于分别扩增对应于 synpcc7942\_0918 的编码区的 5' 部分 (1-480bp) 和 3' 部分 (903-1521bp) 的两个 DNA 片段。*cat* 片段用引物 NS21-3Cm (SEQ ID NO : 20) 和 *ter*-3Cm (SEQ ID NO : 21), 利用 PCR 由质粒 pAM1573 (Mackey 等, Methods Mol. Biol. 362 :115-29) 进行扩增,所述引物分别与引物 918-13 和 918-25 重叠。然后使用重组嵌合 PCR 技术扩增具有三种前述 PCR 片段以及引物 918-15 和 918-23 的完整断裂盒。生成的 2,840-bp 平端 PCR 片段 (SEQ ID NO : 22) 然后连接入 pUC19 (Yanisch-Perron 等, Gene 33 :103-119), 其已经用 *Hind*III 和 *Eco*RI 消化以去除多克隆位点,并且随后用 T4 DNA 聚合酶平端,以生成质粒 pSGTYC04。

[0130] 质粒 pSGI-YC04 被引入细长聚球藻 (*S. elongatus*) 株 SGC-YC1-2 中,其具有整合入 NS1 的 *trcE::ChFatB2-7942* 拷贝 (参见实施例 1)。生成的株被命名为 SGC-YC4-7。脂肪酸产生试验和 GC 分析如在实施例 1 中描述的进行。表明在 IPTG 诱导后 168 小时、各种细长聚球藻菌株的培养物中的 FFA 的水平 GC 分析结果在表 4-1 中显示。酰基-ACP 合成酶基因的失活可能对长链脂肪酸的分泌比中链脂肪酸的分泌有更大的影响。

[0131] 表 4-1 各种细长聚球藻菌株中的中链脂肪酸分泌

|        | 菌株         | 亲代菌株      | 加入的质粒     | 转基因                | 缺失                     | 脂肪酸(mg/L) |      |      |      |
|--------|------------|-----------|-----------|--------------------|------------------------|-----------|------|------|------|
|        |            |           |           |                    |                        | 8:0       | 10:0 | 16:0 | 16:1 |
| [0132] | SGC-YC2-5  | PCC 7942  | pAM2314   | 没有                 | 没有                     | ND        | ND   | ND   | 1.4  |
|        | ISGC-YC1-2 | PCC 7942  | pSGI-YC01 | trcE::ChFatB2-7942 | 没有                     | 1.4       | 4.2  | ND   | 1.6  |
|        | ISGC-YC4-7 | SGC-YC1-2 | pSGI-YC04 | trcE::ChFatB2-7942 | <i>synpcc7942_0918</i> | 1.0       | 3.1  | 1.1  | 3.9  |

[0133] 说明:ND 代表“没有检测到”( < 1mg/L)。

[0134] 实施例 5

[0135] 由源自包含失活的酰基-ACP 合成酶基因的集胞藻 PCC6803 的菌株分泌脂肪酸

[0136] 使用 PCR, 用引物 NB001 (SEQ ID NO:23) 和 NB002 (SEQ ID NO:24), 从基因组 DNA 扩增大约 1.7-kbp 的 DNA 片段, 所述 DNA 片段跨越来自集胞藻 PCC 6803 的酰基-ACP 合成酶编码基因 *slr1609* (Cyanobase 基因命名) 的上游区并进入其编码区。该片段被克隆入 pCR2.1 载体 (Invitrogen), 以产生质粒 pSGI-NB3, 并随后用限制酶 MfeI 切割。用引物 NB010 (SEQ ID NO:25) 和 NB011 (SEQ ID NO:26), 利用 PCR, 从质粒 pAM1573 扩增包含 *cat* 基因和相关调控序列的氯霉素抗性标记盒 (Andersson 等, *Methods Enzymol.* (2000) 305: 527-542) 以包含 MfeI 限制位点的侧翼。*cat* 基因表达盒然后插入 pSGI-NB3 的 MfeI 位点, 以生成 pSGI-NB5 (SEQ ID NO:27)。

[0137] 根据 Zang 等, *J Microbiology* (2007) 45:241-245, 将 pSGI-NB5 载体转化到包含 *trcE::ChFatB2-7942* 的集胞藻菌株 SGC-YC10-5 (参见实施例 1) 中。通过同源重组插入氯霉素抗性标记进入 *Slr1609* 基因通过 PCR 筛选插入和插入位点进行验证。生成的菌株命名为 SGC-NB10-4, 其在液体 BG-11 培养液中进行脂肪酸分泌的检测。所有的液体培养基生长条件使用旋转振荡器 (150rpm)、在 30°C, 用持续光照 ( $60 (\mu E m^{-2} sec^{-1})$ )。培养物在包含氯霉素和 / 或卡那霉素 (5mg/L) 的 BG-11 培养基中相应地接种并生长到足够密度 (最小  $OD_{730nm} = 1.6-2$ )。培养物然后用于接种在 250mL 聚碳酸酯烧瓶中的 100mL BG-11 培养基到  $OD_{730nm} = 0.4-0.5$ , 并温育过夜。45mL 的  $OD_{730nm} = 0.7-0.9$  的过夜培养物被加入新的 250mL 烧瓶, 用 1mM IPTG 诱导或者用作未诱导的对照。5mL 样品在诱导后 0、72 和 144 小时取样, 并如在实施例 2 中描述进行处理。

[0138] 通过液-液萃取, 从过滤培养物上清液溶液中分离游离脂肪酸 (FFA), 用于 GC/FID (火焰离子化检测器) 分析。对于每一个样品, 2mL 过滤培养物用 50 $\mu$ l 磷酸 (1M)、100 $\mu$ l NaCl (5M) 和 2mL 己烷的混合物萃取。使用 40:1 的分流比, 将 0.2 $\mu$ l 样品注射到 DB-FFAP 柱 (J&W Scientific, 15m $\times$ 250 $\mu$ m $\times$ 0.25 $\mu$ m) 上, 温度曲线为: 在 150°C 起始 0.5min、然后以 15°C/min 加热至 230°C 并保持 7.1min (1.1mL/min He)。

[0139] 表明在 144 小时后分泌的游离脂肪酸的水平 GC 结果在表 5-1 中显示。

[0140] 表 5.1 集胞藻的各菌株中的中链脂肪酸分泌。

[0141]

| 菌株         | 亲代菌株       | 加入的质粒     | 转基因                | 缺失              | 脂肪酸(mg/L) |      |
|------------|------------|-----------|--------------------|-----------------|-----------|------|
|            |            |           |                    |                 | 8:0       | 10:0 |
| SGC-YC10-5 | PCC 6803   | pSGI-YC08 | trcE::ChFatB2-7942 | 没有              | 58.3      | 67.7 |
| SGC-NB10-4 | SGC-YC10-5 | pSGI-NB5  | trcE::ChFatB2-7942 | <i>slr 1609</i> | 57.7      | 73.7 |

[0142] 说明:ND 代表“没有检测到”( < 1mg/L)。

[0143] 实施例 6

[0144] 集胞藻中的披针叶萼距花 Kas-IV 和向日葵 Kas-III 基因的表达

[0145] 合成包含功能操纵子的 DNA 片段,以使它以特定的顺序包含以下元件:trc 启动子,密码子被优化用于在细长聚球藻 PCC 7942 中表达的披针叶萼距花 3- 酮脂酰 - 酰基载体蛋白合酶 IV 基因 (C1Kas-IV,GenBank 登录号 CAC59946),和来自聚球藻 WH8102 的 rps14 终止子 (SEQ ID NO:28)。与各种侧翼限制酶识别位点在一起的这种完整功能操纵子的核苷酸序列在 SEQ ID NO:29 中提供。

[0146] 合成包含功能操纵子的另一种 DNA 片段,使得它以特定的顺序包含以下元件:trc 启动子,密码子优化用于在细长聚球藻 PCC 7942 和集胞藻 PCC 6803 中表达的向日葵 3- 酮脂酰 - 酰基载体蛋白合成酶 III 基因 (HaKas-III,GenBank 登陆号 ABP93352),和来自聚球藻 WH8102 的 rps14 终止子)。与各种侧翼限制酶识别位点在一起的这种功能操纵子的核苷酸序列在 SEQ ID NO:30 中提供。

[0147] 通过使用由 DNA2.0, Inc 提供的“Gene Designer”(版本 1.1.4.1) 软件程序进行密码子优化。包含 SEQ ID NO:29 中代表的密码子被修饰的 C1Kas-IV 基因的功能操纵子通过限制酶 SpeI 和 XbaI 消化,并在限制位点 SpeI 和 XbaI 之间插入质粒 pSGI-YC39 以形成质粒 pSGI-BL26,其使得功能操纵子能够在“RS2”重组位点整合入集胞藻 PCC 6803 染色体中 (Aoki 等,J. Bacteriol(1995)177:5606-5611)。包含在 SEQ IDNO:30 中代表的 DNA 片段的质粒 pSGTBL27 以相同方式构建。

[0148] 质粒 pSGI-BL43 包含 trcE 启动子、密码子优化的 C1Kas-IV 基因和 SEQ ID NO:31 中代表的 rps14 终止子,并通过插入来自 pTrcHis A(Invitrogen) 的 SpeI/NcoI trcE 片段进入 SpeI/NcoI- 消化的 pSGI-BL26 进行制备。另外的质粒 pSGTBL44 包含 trcE 启动子、优化的 C1Kas-IV 基因、细长聚球藻 PCC 7942 kaiBC 基因间区、优化的 HaKas-III 基因和 SEQ ID NO:32 中代表的 rps14 终止子,并通过插入经过 PCR 扩增产生的 BamHI/SacI 片段 (包含细长聚球藻 kaiBC 基因间区、HaKas-III 基因和 rps14 终止子) 进入 BglII/SacI- 消化的 pSGI-BL43 进行制备。用于生成包含 kaiBC 区、HaKas-III 和 rps14 终止子的 DNA 片段的 PCR 引物提供为 SEQ ID NO:33 和 SEQ ID NO:34。

[0149] 野生型集胞藻 PCC 6803 细胞和转基因的集胞藻菌株 SGC-YC10-5——其包含 ChFatB2-7942 基因——用质粒 pSGI-BL26、pSGI-BL27、pSGI-BL43 和 pSGI-BL44 转化,如由 Zang, X. 等 J. Microbiol. (2007) 45:241-245 描述。在旋转振荡器 (150rpm) 上、在 30℃, 用持续光照 ( $60(\mu E m^{-2} sec^{-1})$ ), 重组和野生型对照菌株在 20ml 的 BG-11 培养基中预培养到中间对数期 ( $OD_{730nm} = 0.7-0.9$ )。卡那霉素 ( $5\mu g/mL$ ) 和 / 或壮观霉素 ( $10\mu g/mL$ ) 被适当的包括在重组培养物中。培养物然后在 BG-11 中、以起始  $OD_{730nm} = 0.4-0.5$  进行传代培养, 并且培养过夜到  $OD_{730nm} = 0.7-0.9$ 。对于时间 - 进程研究, 45ml 等分试样的培养物转移入

250ml 烧瓶中,并且如果合适的话,通过加入 IPTG(终浓度 = 1mM) 进行诱导。培养物在 IPTG 诱导后的 0、72 和 144 小时取样,然后使用 Millipore 真空过滤器歧管经过 Whatman® GF/F 过滤器进行过滤。滤液在螺旋端部培养管中收集,用于气相色谱 (GC) 分析,如在实施例 2 中描述。

[0150] 显示在培养物接种后 144 小时的培养物上清液中的分泌的辛酸和癸酸的水平的结果在表 6-1 中显示。在显示的菌株中存在的 ClKas-IV 和 HaKas-III 基因是在 trc 启动子的控制之下。

[0151] 表 6-1 各种集胞藻菌株中的中链脂肪酸分泌 (以 mg/L)

[0152]

| 菌株         | 亲代菌株       | 加入质粒      | 转基因                                | 脂肪酸(mg/L) |      |
|------------|------------|-----------|------------------------------------|-----------|------|
|            |            |           |                                    | 8:0       | 10:0 |
| PCC 6803   | n/a        | n/a       | 没有                                 | ND        | ND   |
| SGC-YC10-5 | PCC 6803   | pSGI-YC08 | trcE-ChFatB2-7942                  | 69.8      | 68.4 |
| SGC-BL26-3 | PCC 6803   | pSGI-BL26 | trc-ClKas-IV                       | ND        | ND   |
| SGC-BL26-5 | SGC-YC10-5 | pSGI-BL26 | trcE-ChfatB2-7942<br>trc-ClKas-IV  | 69.5      | 71.9 |
| SGC-BL27-1 | PCC 6803   | pSGI-BL27 | trc-HaKas-III                      | ND        | ND   |
| SGC-BL27-2 | SGC-YC10-5 | pSGI-BL27 | trcE-ChFatB2-7942<br>trc-HaKas-III | 65.7      | 66.6 |

[0147] 说明:ND 代表“没有检测到”( < 1mg/L)

[0148] 为了更优化这些菌株中脂肪酸分泌的测量,在表 6-1 中显示的脂肪酸分泌数据被归一化为细胞培养物密度,测量为在 730nm 时的光密度 (OD<sub>730nm</sub>);这些数据在表 6-2 中提供。在本申请中描述的其它试验可以以类似的方式归一化。

[0149] 表 6-2 归一化的各集胞藻菌株中的中链脂肪酸分泌 (mg/L/OD<sub>730nm</sub>)

[0150]

| 菌株         | 亲代菌株       | 加入质粒      | 转基因                             | 脂肪酸  |      |
|------------|------------|-----------|---------------------------------|------|------|
|            |            |           |                                 | 8:0  | 10:0 |
| PCC 6803   | n/a        | n/a       | 没有                              | ND   | ND   |
| SGC-YC10-5 | PCC 6803   | pSGI-YC08 | trcE-ChFatB2-7942               | 11.7 | 11.4 |
| SGC-BL26-3 | PCC 6803   | pSGI-BL26 | trc-ClKas-IV                    | ND   | ND   |
| SGC-BL26-5 | SGC-YC10-5 | pSGI-BL26 | trcE-ChfatB2-7942 trc-ClKas-IV  | 11.7 | 12.1 |
| SGC-BL27-1 | PCC 6803   | pSGI-BL27 | trc-HaKas-III                   | ND   | ND   |
| SGC-BL27-2 | SGC-YC10-5 | pSGI-BL27 | trcE-ChFatB2-7942 trc-HaKas-III | 12.2 | 12.3 |

[0151] 说明:ND 表示“没有检测到”(< 1mg/L)。

[0152] 显示在培养物接种后 120 小时的另外菌株的培养物上清液中的分泌的辛酸和癸酸的水平的结果在表 6-3 中显示。在显示的菌株中存在的 ClKas-IV 和 HaKas-III 基因是在 trcE 启动子的控制之下。

[0153] 表 6-3 各集胞藻菌株的中链脂肪酸分泌 (以 mg/L)

[0154]

| 菌株               | 亲代菌株       | 加入的质粒     | 转基因                                       | 脂肪酸(mg/L) |      |
|------------------|------------|-----------|-------------------------------------------|-----------|------|
|                  |            |           |                                           | 8:0       | 10:0 |
| SGC-YC10-5       | PCC 6803   | pSGI-YC08 | trcE-ChFatB2-7942                         | 34.8      | 43.5 |
| SGC-BL44         | PCC 6803   | pSGI-BL44 | trcE-ClKAS-IV+HaKAS-III                   | ND        | ND   |
| SGC-YC10-5-B L43 | SGC-YC10-5 | pSGI-BL43 | trcE-ChFatB2-7942 trcE-ClKas-IV           | 40.0      | 48.1 |
| SGC- YC10-5-BL44 | SGC-YC10-5 | pSGI-BL44 | trcE-ChfatB2-7942 trcE-ClKAS-IV+HaKAS-III | 38.5      | 47.1 |

[0155] 说明:ND 代表“没有检测到”(< 1mg/L)。

[0156] 为了更优化这些菌株中脂肪酸分泌的测量,在表 6-1 中显示的脂肪酸分泌数据被归一化为细胞培养物密度,测量为在 730nm 的光密度 ( $OD_{730nm}$ );这些数据在表 6-4 中提供。

[0157] 表 6-4 归一化的各集胞藻菌株中的中链脂肪酸分泌 (mg/L/ $OD_{730nm}$ )

[0158]

| 菌株               | 亲代菌株       | 加入的质粒     | 转基因                                       | 脂肪酸 |      |
|------------------|------------|-----------|-------------------------------------------|-----|------|
|                  |            |           |                                           | 8:0 | 10:0 |
| SGC-YC10-5       | PCC 6803   | pSGI-YC08 | trcE-ChFatB2-7942                         | 6.8 | 8.5  |
| SGC-BL44         | PCC 6803   | pSGI-BL44 | trcE-ClKAS-IV+HaKAS-III                   | ND  | ND   |
| SGC-YC10-5-B L43 | SGC-YC10-5 | pSGI-BL43 | trcE-ChFatB2-7942 trcE-ClKas-IV           | 7.4 | 8.9  |
| SGC- YC10-5-BL44 | SGC-YC10-5 | pSGI-BL44 | trcE-ChfatB2-7942 trcE-ClKAS-IV+HaKAS-III | 8.3 | 10.2 |

[0159] 实施例 7

[0160] 引入异源酰基 -ACP 硫酯酶基因到硅藻中

[0161] 通过构建和利用包括可操作连接到在硅藻中起作用的基因调控区（启动子和终止子）的 ChFatB2 基因的表达载体，在多种硅藻（Bacillariophyceae）中表达编码对中链（8:0-10:0）酰基 -ACP 具有专一性的 ChFatB2 酶的衍生物的合成基因。在优选的实施方式中，优化该基因用于在特定硅藻种中表达，编码天然 ChFatB2 蛋白的质体转运肽区的基因部分用在硅藻中最佳起作用的质体转运肽替代。提供为 SEQ ID NO:35 的核苷酸序列代表已经被优化用于在假微型海链藻中表达的 ChFatB2 基因的合成衍生物，并且在其中基因的天然的质体转运肽编码区已经用与来自假微型海链藻（JGI 标识符 = jgi/Thaps3/40156/estExt\_gwp\_gwl.C\_chr\_40019）的叶绿体 ATP 合酶的偶联因子部分（CF1）的  $\gamma$  亚基相关的质体转运肽（包括偶联的信号序列）替换。由这个基因编码的蛋白质（在此后称为 ChFatB2-Thal）在 SEQ ID NO:36 中提供。

[0162] 为了产生假微型海链藻的表达载体，ChFatB2-Thal 基因被置于假微型海链藻  $\alpha$ -微管蛋白启动子和终止子调控序列之间。通过使用引物 PR1 (SEQ ID NO:37) 和 PR3 (SEQ ID NO:38)，由分离自假微型海链藻 CCMP 1335 的基因组 DNA 扩增  $\alpha$ -微管蛋白启动子，而通过使用引物 PR4 (SEQ ID NO:39) 和 PR8 (SEQ ID NO:40) 扩增  $\alpha$ -微管蛋白终止子。来自  $\alpha$ -微管蛋白启动子扩增子的 KpnI/BamHI 片段、来自  $\alpha$ -微管蛋白终止子的 BamHI/XbaI 片段和来自 KpnI/XbaI 切割 pUC118 的大片段（Vieira 和 Messing, Meth. Enzymol. (1987) 153:3-11）随后结合以形成 pSGI-PR5。来自 ChFatB2-Thal 基因的 NcoI/BamHI 片段然后插入 NcoI/BamHI 消化的 pSGI-PR5，以形成 pSGI-PR16。另外，合成编码来自诺尔斯氏链霉菌（*Streptomyces noursei*）(SEQ ID NO:41) 的诺尔丝菌素乙酰转移酶（NAT）的密码子优化的基因（Krugel 等, Gene (1993) 127:127-131），并且来自这种编码 NAT 的 DNA 分子的 NcoI/BamHI 片段被插入 pSGI-PR5 的大 NcoI/BamHI 片段以形成 pSGI-PR7，当引入假微型海链藻和其它硅藻时，其可以提供对抗生素诺尔丝菌素的抗性。

[0163] 通过粒子轰击，pSGI-PR16 和 pSGI-PR7 共同转化入假微型海链藻 CCMP 1335，如基本由 Poulsen 等，(J. Phycol. (2006) 42:1059-1065) 所描述。在 100mg/L 的诺尔丝菌素（ClonNAT，从 Werner BioAgents，德国获得）存在下，在琼脂平板上挑取转化的细胞。细胞中 ChFatB2-Thal 基因的存在通过使用 PCR 确证。在旋转振荡器上（150rpm）、在 18°C，用持续光照（60 ( $\mu$ E m<sup>-2</sup>sec<sup>-1</sup>），转化体在 ASW 液体培养基（Darley 和 Volcani, Exp. Cell Res. (1964) 58:334）中生长。在接种后 7 天移走样品，而培养基用于测试 FFA 的存在，如在实施例 1 中描述。

[0164] 虽然在这些特定的实验条件下没有检测到脂肪酸分泌，但是可以进行 ChFatB2-Thal 基因和硅藻宿主菌株的优化，以在硅藻中实现脂肪酸分泌，硅藻已知具有相对不透性的细胞壁。

[0165] 实施例 8

[0166] 通过绿藻分泌脂肪酸

[0167] 通过构建和利用包括可操作连接到在绿藻中起作用的基因调控区（启动子和终止子）的 ChFatB2 基因的表达载体，在多种绿藻（Chlorophyceae）中表达编码对中链（8:0-10:0）酰基 -ACP 具有专一性的 ChFatB2 酶的衍生物的合成基因。优化该基因在特定绿藻种中的表达，并且编码天然 ChFatB2 蛋白的质体转运肽区的基因部分用在绿藻中最佳起作用的质体转运肽代替。提供为 SEQ ID NO:42 的核苷酸序列代表 ChFatB2 基因的衍生

物,其被优化用于在莱茵衣藻 (*Chlamydomonas reinhardtii*) 中表达,并且在其中基因的天然质体转运肽编码区已经用与莱茵衣藻 (GenPept 登陆号 XP 001696335) 的偶联因子部分 (CF1) 的  $\gamma$  亚基相关的质体转运肽代替。由这个基因编码的蛋白质在 SEQ ID NO:42 中提供。

#### [0168] 实施例 9

#### [0169] 包含破裂 1,4- $\alpha$ -葡聚糖分支酶基因的集胞藻的菌株中脂肪酸的分泌

[0170] 使用 PCR,用引物 glgB-5 (SEQ ID NO:44) 和 glgB-3 (SEQ ID NO:45),从基因组 DNA 扩增 1.4-kbp 的 DNA 片段,所述 DNA 片段跨越集胞藻 PCC6803 的 1,4- $\alpha$ -葡聚糖分支酶基因 (glgB, Cyanobase 基因名称 = s110158) 的上游区并进入其编码区。该片段被克隆到 pCR4-Topo 载体 (Invitrogen) 以产生质粒 pSGI-BL32,并且随后用限制酶 *Ava*I 切割。包括 *aadA* 基因和相关调控序列的壮观霉素抗性标记盒通过 *Hind*III 从质粒 pSGI-BL27 消化。两个线性片段用 Quick Blunting™ 试剂盒 (NewEngland Biolabs) 处理。*aadA* 基因表达盒随后插入 pSGI-BL32 的 *Ava*I 位点以产生 SGI-BL33。插入 glgB 基因中并使 glgB 基因失活的 pSGI-BL33 部分提供为 SEQ ID NO:46。

[0171] 根据 Zang 等, *J. Microbiology* (2007) 45:241-245, pSGI-BL33 载体被转化到野生型集胞藻 PCC 6803 中和转化到包含 *trcE::ChFatB2-7942* 的集胞藻菌株 SGC-YC10-5 中 (参见实施例 1)。经由同源重组插入壮观霉素抗性标记到 S110158 (glgB) 基因通过 PCR 筛选插入和插入位点进行验证。验证的敲除菌株在液体 BG-11 培养基中测试脂肪酸的分泌。所有液体培养基生长条件使用旋转振荡器 (150rpm)、30°C 和用持续光照 ( $60 (\mu E m^{-2} sec^{-1})$ )。培养物在相应地包含壮观霉素 ( $10 \mu g/mL$ ) 和 / 或卡那霉素 ( $5 \mu g/mL$ ) 的 25mL 的 BG-11 培养基中接种,并生长至足够的密度 (最底  $OD_{730nm} = 1.6-2$ )。培养物然后用于在 250-mL 聚碳酸酯瓶中的 100mL BG-11 培养基中接种至  $OD_{730nm} = 0.4-0.5$ ,并温育过夜。 $OD_{730nm} = 0.5$  的 45mL 过夜培养物被加至新的 250-mL 烧瓶;一些培养物用 1mM IPTG 诱导或者用作未诱导的对照。样品 (0.5mL) 在诱导后的 0、72、144 和 216 小时采取,并如在实施例 2 中所描述的进行处理。

[0172] 通过液-液萃取,从过滤培养物上清液溶液中分离游离脂肪酸 (FFA),用于 GC/FID 分析。对于每一个样品,2mL 的过滤培养物用 50  $\mu L$  磷酸 (1M)、100  $\mu L$  NaCl (5M) 和 2mL 己烷的混合物萃取。使用 40:1 的分流比,将 0.2  $\mu L$  样品注射到 DB-FFAP 柱 (J&W Scientific, 15m $\times$ 250  $\mu m \times 0.25 \mu m$ ) 上,温度曲线为:在 150°C 起始 0.5min,然后以 15°C/min 加热至 230°C 并保持 7.1min (1.1mL/min He)。

[0173] 显示在 216 小时后的游离脂肪酸分泌水平的 GC 结果在表 9-1 中显示。

[0174] 表 9-1 各种集胞藻菌株中的中链脂肪酸分泌 (以 mg/L)

[0175]

| 菌株         | 亲代菌株       | 加入的质粒     | 缺失             | 转基因                      | 脂肪酸  |      |
|------------|------------|-----------|----------------|--------------------------|------|------|
|            |            |           |                |                          | 8:0  | 10:0 |
| PCC 6803   | n/a        | n/a       | 没有             | 没有                       | ND   | ND   |
| SGC-BL33-1 | PCC 6803   | pSGI-BL33 | S110158 (glgB) | 没有                       | ND   | ND   |
| SGC-YC10-5 | PCC 6803   | pSGI-YC08 | 没有             | <i>trcE-ChFatB2-7942</i> | 70.0 | 68.7 |
| SGC-BL33-2 | SGC-YC10-5 | pSGI-BL33 | S110158 (glgB) | <i>trcE-ChFatB2-7942</i> | 66.2 | 68.1 |

[0176] 说明:ND 代表“没有检测到”(< 1mg/L)

[0177] 为了更优化这些菌株中脂肪酸分泌的测量,在表 9-1 中显示的脂肪酸分泌数据被归一化为细胞培养物密度,测量为在 730nm 的光密度 ( $OD_{730nm}$ );这些数据在表 9-2 中提供。在本申请中描述的其它试验可以以类似的方式归一化。

[0178] 表 9-2 归一化的各种集胞藻菌株中的中链脂肪酸分泌 (mg/L/ $OD_{730nm}$ )

[0179]

| 菌株         | 亲代菌株       | 加入的质粒     | 缺失             | 转基因               | 脂肪酸  |      |
|------------|------------|-----------|----------------|-------------------|------|------|
|            |            |           |                |                   | 8:0  | 10:0 |
| PCC 6803   | n/a        | n/a       | 没有             | 没有                | ND   | ND   |
| SGC-BL33-1 | PCC 6803   | pSGI-BL33 | S110158 (glgB) | 没有                | ND   | ND   |
| SGC-YC10-5 | PCC 6803   | pSGI-YC08 | 没有             | trcE-ChFatB2-7942 | 9.8  | 9.7  |
| SGC-BL33-2 | SGC-YC10-5 | pSGI-BL33 | S110158 (glgB) | trcE-ChFatB2-7942 | 10.4 | 10.7 |

[0180] 说明:ND 代表“没有检测到”(< 1mg/L)

[0181] 实施例 10

[0182] 用疏水性吸附树脂从模式溶液中捕获游离脂肪酸

[0183] 通过在补充有 300mM NaCl 的 BG-11 培养基中溶解 75mg/L 辛酸和 75mg/L 癸酸并调节 pH 至 5.8 来配制掺料溶液 (spike solution)。在表 1 中列出的 50mg 的每种树脂称重放入 50mL 离心管中,与 1.0mL 的甲醇合并,并轻轻摇动。倒出过量的甲醇,并在 25 in Hg 真空下室温过夜,干燥树脂。50mL 的掺料溶液随后加入每种树脂,并在 31℃ 轻轻摇动温育 24 小时。在温育后,在 **Whatman®** GF/F 玻璃纤维过滤器上通过过滤去除树脂,而滤液通过气相色谱分析辛酸和癸酸含量,如在实施例 2 中描述。每种树脂对于辛酸和癸酸的容量随后通过在与每种树脂温育之前和之后的每种脂肪酸的浓度的差别测定。结果在以下表 10-1 中显示。

[0184] 表 10-1 数种商业可得的吸附剂的吸附容量

[0185]

| 描述                                    | 树脂类型           | 吸附容量 (mg/g) |      |         |
|---------------------------------------|----------------|-------------|------|---------|
|                                       |                | 辛酸          | 癸酸   | 总的游离脂肪酸 |
| Dowex Optipore® V503 (Dow Chemical)   | 交联后大孔聚苯乙烯二乙烯基苯 | 26.3        | 69.8 | 96.0    |
| Lewatit 1064 MD (LanXess)             | 大孔聚苯乙烯二乙烯基苯    | 1.1         | 46.7 | 47.8    |
| Zeolyst CBV 28014 (Zeolyst)           | 非常低氧化铝沸石       | 17.4        | 74.7 | 92.0    |
| Zeolyst CBV 901 (Zeolyst)             | 低氧化铝沸石         | 5.4         | 64.8 | 70.1    |
| Hisiv 3000 硅质岩 (UOP Honeywell)        | 疏水性硅质岩         | 15.3        | 23.7 | 39.1    |
| Lipidex 5000 (Packard Instrument Co.) | 烷基化葡聚糖凝胶       | 0.00        | 18.6 | 18.6    |
| Norit ROW 0.8 (Fluka)                 | 挤压的活性碳         | 40.2        | 71.8 | 112.1   |

[0186] 也研究了从疏水性吸附剂洗脱游离脂肪酸。Dowex® Optipore® V503、Zeolyst CBV 28014、Zeolyst CBV 901 和 Norit® ROW 对于每 mg 吸附剂用 1.0mL 的掺料溶液温育, 如上描述。在温育期后, 清洗吸附剂并用 0.1、0.5 或 1.0mL 甲醇 /mg 的吸附剂结合, 并在室温轻轻摇动 4 小时。甲醇洗脱液和吸附后峰通过气相色谱分析游离脂肪酸浓度。结果在以下表 10-2 中列出。

[0187] 表 10-2 甲醇中游离脂肪酸的解吸

[0188]

| mL MeOH/mg 树脂        | % 解吸      |           |          |
|----------------------|-----------|-----------|----------|
|                      | 0.1 mL/mg | 0.5 mL/mg | 1.0mL/mg |
| Dowex Optipore® V503 | 92%       | 84%       | 100%     |
| CBV 28014            | 53%       | 76%       | 84%      |
| CBV 901              | 78%       | 76%       | 57%      |
| Norit® ROW           | 44%       | 85%       | 77%      |

[0189] 利用 owex Optipore® V503 研究了 pH 对吸附容量的影响。40mg 的树脂与 40mL 用 150mg/L 辛酸和癸酸掺料 (spike) 并调节 pH 至 10.0、7.5、4.8 或 2.8 的 BG-11 培养基结合。

[0190] pH 10 掺料溶液用 5mM CAPS 缓冲。pH 7.5 和 pH 2.8 掺料溶液用 5mM 磷酸盐缓冲, 而 pH 4.8 掺料溶液由溶解的脂肪酸天然地缓冲, 加入 5mM NaCl 以维持稳定的导电性。掺料溶液与树脂一起温育, 如以上描述。游离脂肪酸浓度用从 Zen-bio 购买的酶法测定进行测量。结果在以下表 10-3 中显示。根据这些结果, 清楚的是, 游离脂肪酸的疏水性吸附在广泛的 pH 范围内是可能的。

[0191] 表 10-3 Dowex® Optipore® V503 在不同 pH 值的吸附容量

[0192]

| pH  | 吸附容量(mg FFA/g 树脂) |
|-----|-------------------|
| 10  | 42 ± 13           |
| 7.5 | 64 ± 4            |
| 4.8 | 172 ± 4           |
| 2.8 | 259 ± 1           |

[0193] 报道的值是两次重复试验的平均值 +/- 一个标准偏差。

[0194] 实施例 11

[0195] 从集胞藻菌株 SGC-YC10-5 的培养物体内捕获游离脂肪酸

[0196] 集胞藻菌株 SGC-YC10-5——其包含如在实施例 1 中描述的 ChFatB2-7942 基因——在具有和不具有Dowex® Optipore® V503 树脂的 BG-11 中培养。400mL 的新鲜培养物用 5mM IPTG 诱导,并在室温温育 1 小时以摄取诱导剂。培养物然后被分到 4 个 1000mL 具有 PTFE 排气帽的带阻板 Erlenmeyer 锥形瓶中。向这些锥形瓶中的两个加入大约 400mg 的Dowex® Optipore® V503。回收试验锥形瓶中的吸附树脂并且每日更换新鲜树脂共 10 天。回收的树脂用去离子水充分洗涤,并用 2mL 甲醇洗脱。每天从试验锥形瓶和对照锥形瓶取培养基样品。样品在 OD<sub>730nm</sub> 测量并在Whatman® GF/B 玻璃纤维过滤器上过滤,并通过气相色谱分析辛酸和癸酸含量,如在实施例 2 中先前描述的。结果在表 11-1 中显示。

[0197] 表 11-1 从集胞藻 SGC-YC10-5 培养物体内捕获游离脂肪酸

[0198]

|          | 平均比生长率 (d <sup>-1</sup> ) | 平均游离脂肪酸生产力 (mgL <sup>-1</sup> d <sup>-1</sup> ) |
|----------|---------------------------|-------------------------------------------------|
| 没有 Dowex | 0.090 ± 0.005             | 16 ± 0.8                                        |
| 有 Dowex  | 0.090 ± 0.010             | 31 ± 3                                          |

[0199] 报道的值是两次生物学重复的平均值, +/- 一个标准偏差。

[0200] 实施例 12

[0201] CO<sub>2</sub> 递送和产物回收的整合作为提高两者的效率和经济的手段

[0202] 以上表 10-3 显示了游离脂肪酸吸附容量和 pH 之间的清楚的关系。这种关系产生于电离形式的游离脂肪酸的提取的低效率。多种潜在的生产宿主要求比游离脂肪酸 pKa 显著高的 pH,以存活和繁殖。这种的一个极端实例是嗜碱蓝细菌,例如属于属聚球藻属、集胞藻属、螺旋藻属的那些和许多其它的种类,它们最佳生长优选在 pH9 和 pH11 之间。图 5 概述了本发明的实施方式,其中这个问题通过再循环培养物的部分来解决,所述培养物的部分首先通过容器——在容器中其与浓的 CO<sub>2</sub> 气体接触以降低 pH,然后通过固定吸附柱,在所述吸附柱中去质子的游离脂肪酸被捕获。

[0203] CO<sub>2</sub> 富集、游离脂肪酸耗尽的悬浮液然后返回到主体培养物。气-液接触器里面的压力可以被独立地控制,以提供流出吸附柱的流中稳定 pH。而且,柱后闪蒸器 (post-column flash vessel) 的压力可以被控制,以提供 CO<sub>2</sub> 的供给,其通过 PID 控制的 pH、溶解 CO<sub>2</sub>、排出气 CO<sub>2</sub>、或三者的任意组合,滴定测量主体培养物的 CO<sub>2</sub> 消耗速率。过量的 CO<sub>2</sub> 然后可以再循环。

[0204] 为了展示以上描述的本发明概念的证据,构建了一个试验系统,如在图 5 中显示的。

[0205] 容器 E-1 装满 4L pH 为 11.1、包含溶解在 100mM NaCl 中的 700mg/L 辛酸的掺料溶液。柱 C1 装满 45.2g 的Dowex® Optipore® V503 聚合树脂。树脂用两柱体积的甲醇活化,然后用三柱体积的 pH 11.1 的 100mM NaCl 洗涤。液-气接触容器 E2 然后装满 200mL 的掺料溶液和 34.7 磅/平方英寸 (psia) 的 CO<sub>2</sub>。当在 E-2 里面的掺料溶液的 pH 降低至 5-6 时(如通过在 E-2 里面包含的一片 pH 纸测定),蠕动泵 P-1 和 P-2 被设定到相同流速并开始装柱。阀 V-2 根据需要调节以增加柱压并阻止气泡形成。

[0206] 流体组分在周期间隔取 70-100mL,并通过从 Zen-Bio 购买的商业可得的游离脂肪酸测定进行辛酸测定。两个表面线性流速被评估:16.3cm/min 和 6.1cm/min。对于这两个流速,通过迂回的容器 E-2 进行对照试验,并且柱在 11.1 的 pH 直接装载。以下表 12-1 显示这个试验的结果。对于两个流速,柱动态结合量比 CO<sub>2</sub> 用于降低负载的 pH 时的结合量大大约 4 倍。

[0207] 表 12-1 有和没有 CO<sub>2</sub> 调节的负载酸化的动态结合量

|        |              |                            |                                     |
|--------|--------------|----------------------------|-------------------------------------|
| [0208] | 动态结合量 (mg/g) |                            |                                     |
|        | 流速 (cm/min)  | +34.7 psia CO <sub>2</sub> | 对照 (pH 11.1) 0 psia CO <sub>2</sub> |
|        | 6.1          | 43.5                       | 10.5                                |
|        | 16.3         | 7.2                        | 1.9                                 |

[0209] 实施例 13

[0210] 通过光合微生物分泌油酸

[0211] 通过构建和利用包括可操作连接到在宿主光合微生物中起作用的基因调控区(启动子和终止子)的Fat A 基因的表达载体,在不同光合微生物中表达编码对油酰基-ACP 具有专一性的 Fat A-型植物酰基-ACPT 酶的衍生物的合成基因。优化该基因在宿主光合微生物中的表达,并且编码天然 Fat A 蛋白的质体转运肽区的基因部分在蓝细菌中被去除表达,或者用在宿主真核光合微生物中有效起作用的质体转用肽替换。

[0212] 可以用于这个目的的基因包括但不限于:编码以下酰基-ACP TE 的那些(通过 GenPept 登录号查阅):NP\_189147.1、AAC49002、CAA52070.1、CAA52069.1、193041.1、CAC39106、CA017726、AAC72883、AAA33020、AAL79361、AAQ08223.1、AAB51523、AAL77443、AAA33019、AAG35064 和 AAL77445。

[0213] 以下是以上提到的所有序列的序列表。

[0214] SEQ ID NO:1

[0215] ATATCGTGGAGTATATCAATGGTGGGAGGTGTGGTGTAGTAGTTGCGAGCAAAGATGACACTTGGTAAACTGA

[0216] TGCGACGTGGATACTGCGACGAAGATTGGCCGTACACACGTCGGATTTGAATGAACATATGTGTTTTATTCAA

[0217] CCAATTTGACTAGTTTGAGGAACCTTCACGTGTTTCGCTCTCAAACCTTTGAGACAACAGCCTCCGAATCCAAAT

[0218] GAATGACTTTTAAACACAAGCTAGGAGCTGGTGATATATAATATGCTGGTTGTATGAAAGAGACTAATCGTGTG

[0219] AAATAAATGATGGCTCGCCCTAGTGAATGCTCCTCAGAGACGCTCATTCGTCCAAGTGTTCGTCACTTCTGTCA

[0220] TTGTTTCTCCGAGGCCAAGGTGGTCGAGTAGGTAGATACCAGCTATTCTCTTGCTTCTTTTACTTTATCTCCC

[0221] TCTACCAAAAACAGCACGTTATTATCTCCTTTCCATTCCACGCAATAACAAGAGGCAATCGGTAAAGAGGCACA

[0222] AACAAAGAGAAACAAAGACCCCGGCTGCTTCTCTCGTCCGTCCGCCGCCCTAAACTTCAAGTTTACTTCAAGTT  
[0223] CAATCTGTTTTTTGGCGCAAAAAGCGCGGTTGCTCCGCCGTCTCCGCACTTTTCAGTTCTCTGTGTCGAGGA  
[0224] CTGTTATCAACTTCCAAGATCTCCATCTCTTCTCCTATCCTCCCCTAACAAAGTACG  
[0225] SEQ ID NO :2  
[0226] GTTCAATGCCTTTGGTGTGTGTCGTCATAGGCACTTCGACTTTGCTCTTGGTTCCGTTATCCCAAACTTGAACG  
[0227] AGCGCCACGGTCCTCTCGGTTTGGTGGTATCCCAGGACCTCTCGTAGTTGATGCAGGGTTCAGAATCGAGATA  
[0228] ACTCATGTTGTGCTTTGTTGTTTTGTTGATTTTACCTTGCTTCCAGCTTTCGGTCTGTAATTACAGTGACACGC  
[0229] TGTACTAGAAATGATGTACGTTTGTATGGAATCTCTAAAAATTATGAGCTATTTATGAACACAGGAGTTCTCATCA  
[0230] ACTTTCATCGAAATCCGTAGGAGAATTCTAATGTCCTCTTCGGACGAGAGACAGACGTATCAGGAGTCACTTG  
[0231] AAGGTTCCAAGATTCTATCTTCATGAGGTCTGGATATGACAGTCTCGCTTCGAGGCAAGCCCTGTCAGTGTGA  
[0232] CCTTTTCGCGTCGTCAATAATTTTAGGAACGCAAGGATAGGGATTCTCCATAGTAAGGACTATTGTTTGACCCC  
[0233] TGAAACTTCAACCTTTACCCCAAGATGGGGCATTCTAAGTGAAAAACGTTTGTATGTATGCCCAATTCTCT  
[0234] ACACAGGAATAGGTATTGAATCACGTAGAAAAATGATCGTTGCGCCGCAAGCAACACACCGGCTCTCTTCGCC  
[0235] GCACTCTCTTCCAATCCAACAAAACAAACGCAACC  
[0236] SEQ ID NO :3  
[0237] ACGCAGATAGTGATATTTGCGTCACAGTCTCTTGTGTCATAGGAGAGGAGAACTAGAGAACAAAAAGCGTCA  
[0238] TGTAATAAATGTTGGATGTTGGCATGTGCTCCAGCCAGTATCCAAAACACCGAATTGTCGAGGTTGCTGAGCT  
[0239] TGCAGCACTCATGGCAACGGCTAATTTTCATATCTATGTTATCAATGTTATCTGTAACACTAATGCTAAGTAATG  
[0240] CGTCAACAACTTATCTCCTCCGGCTCTTCACTCCACTTCGCTGACGTGCTTTGATATTTTATCTGCTCTATTAT  
[0241] TCAAGTTGAATCTGCAGTTGAGGCATTCTCTAACTTAGCCGAGAAATCAAGACGGTGACTTTGAATTTACAAGT  
[0242] ACAGTTACGCTTACACAAGATACTTTCTCACAAAAAGATTCCGTTGGCTCCCACTGCGCATTGCTACTTGGT  
[0243] ACTATTTCCCATGTGGAACCTGGATTGGGGGAAAGAGGGAGTCTGAGTTTGTAAATGTACATTTGTTATTCCTT  
[0244] CATTATCGACAACATCACTAACTCATCGTGCATACAGAGAAAAACAATCTCCACTTTCTCAACAAAAGTGGCCA  
[0245] CAATGTGCCTCCGACACAGCCTCAAGAGCCGACCGATCGTTGCATTTTCACTCTCGAACACACACACACAC  
[0246] ACACCCACACACCACCACCTCTCTTTATCCAACC  
[0247] SEQ ID NO :4  
[0248] AGTCGGATTGAAAACAGCGAATGTACGCCATTCCAAAGCGCTCAGCAAAAGGAGACATATGCACACATCCAGC  
[0249] GGAAGTAAGTACGACACTTGAACAAGAGCATGACCTGTCAAAGCATGCTGCCATCGTCCGTTTCGCTTCTATTCC  
[0250] CAATGACACTTTGGTCACCAAGCACTTGA AAAACGGCAATCAGCAAAATAAGCGATAGACCTTGACCAACGGCAG  
[0251] CTTTCATCTTTTATGAACGGCAGATATTGTCATCCTCTTTTATCGATACAGCAAAACACGCAGAATTTCTGTCT  
[0252] CTTTCAAGACGACAAGCAGCAATTTTCGTACGCTGTCTAATTTATTGACTATGTTAGATAACACAACCTCTCAT  
[0253] GCGCTTTGAAAATCTGCTTACTTCACAGTAAAGAGACAAGCTCTTTGCACTGACTGCGACAGAGATGGAAAAAA  
[0254] GGAATTCTACCGCAATTGACAGACTGATGTGAAAACAGAGAGTAACCGTAAACAAGTACCGGTAAGTATGCGC  
[0255] GCAACCTTTACTTGTTCGTTGGCGTCTGTCAATTTGATGTACGCAGACTTGAAAAGTCGTTTCGCTCCATTGTG  
[0256] AAAAATATCATGCGACAACGTTTCAAGAGCGCGGTGCAATCGGTTTGCTTGTCTGATCCGCTGCTTTTT  
[0257] GAGCAACGACCTGCGGAGGACCACAATGATCTTTCTCTTGTGCTGAGAGCTAGTTCTATTACCTGTTCAATTAC  
[0258] CTGCTTTCTTGTATTACTCGAAGCTCTCGTTCTTCTATC  
[0259] SEQ ID NO :5  
[0260] TTTTGTAAATTCGCCACTACCTTTACGCAAGTAAGAACGTTTCATGCTGGAGTCGTGGACCAATCGTAAGGTATA

[0261] CGTTAGTCATACCGCGCCTGTACTATTTACGACACGAGAGAAAAGCCACTGCAGTTCTGGGATGGGATCAGATGC  
[0262] TTGCTCCTTTCACTGCGCTGGCAAACGTATGCTAGACACGACTCGGATCGGATATCGAAATCAAACGGCGGAG  
[0263] AATGGGTTTCGGATGACTGTCCGGAGCTACCTAGGAAAAGCTTCTTTTTCGTTTCGGACCACCAAGAGGGAAGCG  
[0264] CTGCCGTGTAATCGTGGCATAGGAAGCATCAGACGTATTTGTTTCGGATGAGATCACACCAGAACTAGCCAGGCAG  
[0265] CCAGCTAGCTATTTGTCATCTACAGATTTGAAACCAAACGTGGATACTAGAAAGCATGGGATTGACTGTGACTGT  
[0266] GATTTGTGTTGCACACTTTATACCTACCTCGACCTCGTACTTTGTGTAGTAGCAAAATGTGGATTGTGCGTTG  
[0267] AAATGTAGAAGGGTTTGGGGTTGACACGGGTTCAATCATATCCGGTACTCGAAAATGACCGCAACGATACTCA  
[0268] TCGATCGAGATACGGTGTACAGTAGACTACGTAGAAAACCTACGAGGAAGCAGATATGATTTTCCGGTCCGCA  
[0269] GCATCCACCCAGCCAACGTCCGCAAACAACCAACAACCTCGTCCCCCTTGTGTTCAAGATCTGCATTCCAT  
[0270] TGACAGCCTTTTCAACGAAAACCGTTCGCTCGTTTGATTCCATACGTCTTTGAATACCAACAGAAAAAT  
[0271] SEQ ID NO :6  
[0272] AAAGTATCAATAGCTTATTCAGATTTTGTGATGTTAGCCTACTTGTAAGCAGCGGAGGTCTGTCATGACGG  
[0273] TGTAGTGGCTGGTTTCGCTCCGAAATTAAGTTCTGGTTTATATCTCAACATAACTAGAGATAAAGTTACAGG  
[0274] CACGTTACTGTAAGTCCGAGATTGCTAATGCTTTGCTTCGGTGTCCGTAAAGCTTATGTTACTGTTCTAGATT  
[0275] AGAGTGGTATCCACGATTTTCAAACGAAAGTGACATATTGCGAATTGTGCAGTATCAGAAAATCTCCAAAGCAG  
[0276] GAGCATACATTAGTTTGGCCGTATTGCAACGAGTAGCTCTCCTGAAGATGCAAGTAATAGAGGCTGTGAGCGTG  
[0277] AATAATGAATTTGCCTGTTTGAAGCTGGGGATCACATCTCGTGCTCCCCAAAAGTCTCTCAGTAAATCAAGAA  
[0278] TGTTCCTATTTTGGAAAACATTGCTATTTATTTAGTTAACCGGCTTCGTCTCCCATTTAAATAAAGATTTTCA  
[0279] AAAATGACACCACCAACGTCCGCAAGATCACGATTGAGAGGATTCTTCTTTGTCCCAACCATGGATGACCTCT  
[0280] CCTATTAACACGTATATGAAGTACCGCTGCTGGTACCCGGAAAAGAGAGGACATTCCCTTGTGGGAGAGTCATCG  
[0281] ATGCGCTGCCAATCGAAAAAATGCCAAGGCGAGAAAAGCGCAGTTGTTCTTATAATCCAATTTTGAGTTTCA  
[0282] AGACATACTCGTTGCTACCTTCCCACCTTCCCAACCAAACTCGCAACC  
[0283] SEQ ID NO :7  
[0284] CCATGGCGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCC  
[0285] CCTCCTCGGACGTTTTTGCATCAGCTGCGCGACTGGAGTCGCTTGCTGACCGCCATCACAACAGTGTGTTGTCAA  
[0286] ATCTAAACGACCGGACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGT  
[0287] CTACTGTGCAGGACGGCTGGTGTTCGGTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACC  
[0288] GCTAGCATTGAAACGTTGATGAACCATCTCCAAGAAACCAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCT  
[0289] GGATGGTTTTTGGCCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAATAATGCAGATCAAAG  
[0290] TTAATCGATACCCGGCCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTGGCATG  
[0291] GGTCCGATTTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGCTCTGCTACGCGATGAT  
[0292] GAATCAAAAGACGCGTCGGTTGAGTAAGCTGCCGTATGAAGTTACCAAGAAATTGTTCCATTGTTTCGTTGATA  
[0293] GTCCCGTTATCGAGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGC  
[0294] CTGACGCCAGGTTGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAA  
[0295] AAGTATGCCTACGGAAGTCTGGAGACGAGGAACCTCTGCAGTCTCGCTCTGGAGTACCGCGTGAGTGTGGCC  
[0296] GTGATTCCGTGCTCGAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTTCGAGTCAATACCAACACCTC  
[0297] TTGCGGCTCGAAGATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGG  
[0298] GGCAATTAGTACCGGGAAAAACCTCCAATGGAAAACAGCGTCAGCTAATGATAGGATCC  
[0299] SEQ ID NO :8

[0300] MANGSAVSLKSGSLNTQEDTSSPPPRFTLHQLPDWSRLLTATTTVFVKSkrpDmHDrkskrpDMLVDSFGLES  
[0301] TVQDGLVFRQSFsIRsYEIGTDRTASIEtLMNHLqETSLNHCKSTGILLDGFGRtLEMCKRDLIWVVIKMQIKV  
[0302] NRYPAWGDTVEINTRFSRLGKIGMRDWLISDCNTGEILVRATSAYAMNqKTRRLSKLPYEVHQEIVPLFVDS  
[0303] PVIEDSDLKVHKFKVKtGDSIQKGLTPGWNDLVDNqHVSNNKYIGWILESMPTEVLETqELCSLALEYRRECOGR  
[0304] DSVLESVTAMDPSKVGVRsQYQHLLRLEDGTAIVNGATEWRPKNAGANGAISTGKTSNGNSVS  
[0305] SEQ ID NO :9  
[0306] CGCCGGGGCTGGCAGCTTAGTCCTGCGCAATCTCTACTACATCTGCCAACCCAGTGAAATTTTGATCTTTGCTG  
[0307] GCAGTAGTCGCCGAGTAGTGATGGCCGCCGAGTTGGCTATCGCTTGGTCAAGGGCGGCAGCAGCCTGCCGGTA  
[0308] CCTCTGCTGAAAAAGCGCTCCGCATGGATCTGACCAACATGATCATTGAGTTGCGCGTTTCCAATGCCTTCTC  
[0309] CAAGGGCGGCATTCCCCTGA CTGTTGAAGGCGTTGCCAATATCAAGATTGCTGGGAAGAACCAGCATCCACA  
[0310] ACGCGATCGAGCGGCTGCTTGGCAAAAACCGTAAGGAAATCGAGCAAATTGCCAAGGAGACCTCGAAGGCAAC  
[0311] TTGCGTGGTGTTTTAGCCAGCCTCACGCCGAGCAGATCAACGAGGACAAAATTGCCTTTGCCAAAAGTCTGCT  
[0312] GGAAGAGGCGGAGGATGACCTTGAGCAGCTGGGTCTAGTCCTCGATACGCTGCAAGTCCAGAACATTTCCGATG  
[0313] AGGTGGTTATCTCTCGGCTAGTGGACGCAAGCAGCGGGCTGATCTGCAGCGAGATGCCGAATTGCTGAAGCC  
[0314] GATGCCCAGGCTGCCTCTGCGATCCAAACGGCCGAAAATGACAAGATCACGGCCCTGCGTCGGATCGATCGCGA  
[0315] TGTAGCGATCGCCCAAGCCGAGGCCGAGCGCCGATTCAAGATGCGTTGACGCGGCGGAAGCGGTGGTGGCCG  
[0316] AAGCTGAAGCGGACATTGCTACCGAAGTCGCTCGTAGCCAAGCAGAACTCCCTGTGCAGCAGGAGCGGATCAAA  
[0317] CAGGTGCAGCAGCAACTTCAAGCCGATGTGATCGCCCCAGCTGAGGCAGCTTGTAACGGGCGATCGCGGAAGC  
[0318] GCGGGGGCGCCGCCCGTATCGTCGAAGATGAAAAGCTCAAGCGGAAGGGACCAACGGCTGGCGGAGGCTT  
[0319] GGCAGACCGCTGGTGCTAATGCCCGGACATCTTCCTGCTCCAGAAGCTCGAAATTCGAGCTCGGTACCATTTA  
[0320] CGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAAGG  
[0321] TGGTGAATGTGAAACCAGTAACGTTATACGATGTGCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC  
[0322] GTGGTGAAACCAGGCCAGCCAGTTTCTGCGAAAACGCGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTA  
[0323] CATTCCTCAACCGCTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGG  
[0324] CCTGACGCGCCGTCGCAAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTG  
[0325] TCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAACTTCTCGCGCAACGCGTCAGTGG  
[0326] GCTGATCATTAACATATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGT  
[0327] TATTTCTTGATGTCTCTGACCAGACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGC  
[0328] GTGGAGCATCTGGTCGCAATTGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCG  
[0329] TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGAAGGCGACT  
[0330] GGAGTGCCATGTCCGGTTTTCACAAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTT  
[0331] GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTC  
[0332] GGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAAACAGGATTTTC  
[0333] GCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTG  
[0334] TTGCCCGTCTCACTGGTGAAAAGAAAAACACCCTGGCGCCCAATACGCAAAACCGCCTCTCCCGCGCGTTGGC  
[0335] CGATTCAATTAATGCAGCTGGCAGACAGGTTTCCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTA  
[0336] AGTTAGCGGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGGTCAGGCA  
[0337] GCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTGCTGTGCTCAAGGCGCACTCCC  
[0338] GTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTA

[0339] ATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCGCCGCTGAGAAA  
[0340] AAGCGAAGCGGCACTGCTCTTTAACAATTTATCAGACAACTCTGTGTGGGCACTCGACCGGAATTATCGATTAAAC  
[0341] TTTATTATTAAAAATTAAAGAGGTATATATTAATGTATCGATTAAATAAGGAGGAATAAACCATGGCGAATGGT  
[0342] TCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCCCCTCCTCGGACGTT  
[0343] TTTGCATCAGCTGCCCGACTGGAGTCGCTTGCTGACCGCCATCACAACAGTGTTTGTCAAATCTAAACGACCGG  
[0344] ACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTACTGTGCAGGAC  
[0345] GGCCTGGTGTTCCGTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTAGCATTGAAAC  
[0346] GTTGATGAACCATCTCCAAGAAACCAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGGTTTTTGCC  
[0347] GCACATTGGAAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAAATGCAGATCAAAGTTAATCGATACCCG  
[0348] GCCTGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTTGGCATGGGTGCGATTGGCT  
[0349] GATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGCTGCGTACGCGATGATGAATCAAAAGACGC  
[0350] GTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTTGTCCATTGTTTCGTTGATAGTCCCGTTATCGAG  
[0351] GATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACGCCAGGTTG  
[0352] GAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAAGTATGCCTACGG  
[0353] AAGTCCTGAGACGCGAGGAACCTCTGCAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGATTCCGTGCTC  
[0354] GAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGCGGCTCGAAGA  
[0355] TGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAATTAGTACCG  
[0356] GGAAACCTCCAATGGAAACAGCGTCAGCTAATGATAGGATCCGAGCTCGAGATCTGCAGCTGGTACCATATGG  
[0357] GAATTGGAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAA  
[0358] GCGGTCTGATAAAACAGAAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACCTCAGAA  
[0359] GTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACCTGCCAGGCATCAAATAA  
[0360] AACGAAAGGCTCAGTCGAAAGACTGGGCCTTTGTTTTATCTGTTGTTTGTGCGTGAACGCTCTCCTGAGTAGG  
[0361] ACAAATCCGCCGGGAGCGGATTTGAACGTTGCGAAGCAACGGCCCGGAGGGTGGCGGGCAGGACGCCCGCCATA  
[0362] AACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTGTGCTTTCTACAAACTCTTTTGT  
[0363] TTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGGGGATCCGACTAGTAGGCCTCGAGGAATTCACGCG  
[0364] TACGTAGATCTCCGCGGCCGCGATCCTCTAGTATGCTTGTAACCGTTTGTGAAAAATTTTTAAATAAAAA  
[0365] AAGGGGACCTCTAGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGTCAATCAAAGGTCATCCACCGGA  
[0366] TCAGCTTAGTAAAGCCCTCGCTAGATTTAATGCGGATGTTGCGATTACTTCGCCAACTATTGCGATAACAAGA  
[0367] AAAAGCCAGCCTTTTCATGATATATCTCCCAATTTGTGTAGGGCTTATTATGCACGCTTAAAAATAATAAAGCA  
[0368] GACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCATCTAACGCTTGAGTTAAGCCGCGCCGCG  
[0369] AAGCGGCGTCGGCTTGAACGAATTGTTAGACATTATTTGCCGACTACCTTGGTGATCTCGCCTTTACGTTAGTG  
[0370] GACAAATTTCTTCCAACTGATCTGCGCGCGAGGCCAAGCGATCTTCTTCTGTCCAAGATAAGCCTGTCTAGCTT  
[0371] CAAGTATGACGGGTGATACTGGGCGCGCAGGCGCTCCATTGCCAGTCGGCAGCGACATCCTTCGGCGCGATT  
[0372] TTGCCGGTTACTGCGCTGTACCAAATGCGGGACAACGTAAGCACTACATTTGCTCATCGCCAGCCCAGTCGGG  
[0373] CGGCGAGTTCCATAGCGTTAAGGTTTTCAATTTAGCGCCTCAAATAGATCCTGTTTCAAGAACCGGATCAAAGAGTT  
[0374] CCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTGAGCAAGATAGCCAGATCAATGTGCG  
[0375] ATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTCGAGTTGCGGCTTAGC  
[0376] TGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTGCTCTCTC  
[0377] CAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTTCATCAAGCCTTACGGTCACC

[0378] GTAACCAGCAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCCAG  
[0379] CAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCACCG  
[0380] CTTCCTCATGATGTTTAACTTTGTTTTAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACATC  
[0381] AAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGGATGCCCCAGGCATAGACTGTACCCCCAAAAAACAGTC  
[0382] ATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCACCGCTGCGTTCGGTCAAGGTTCTGGACCAGTTGCGTG  
[0383] AGCGCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTCTGTCCTTCATCCGT  
[0384] TTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCTTGGCTGGCGAACGA  
[0385] GCGCAAGGTTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCCTTGCTGTCTTCTACGGCAAGGTGCTGTGCA  
[0386] CGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCGGTGGTGCTGACCCCGGAT  
[0387] GAAGTGGTTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTGTTCGCCAGCTTCTGTATGGAACGGGCAT  
[0388] GCGGATCAGTGAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTGATCACGGCACGATCATCGTGGGGGAGG  
[0389] GCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCCGAGAGCTTGGCAGCCAGCCTGCGCGAGCAGGGGAATTGA  
[0390] TCCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATTAATAATTGAGGACCTAGAGGTCCCTT  
[0391] TTTTATTTTAAAAATTTTTCACAAAACGGTTTACAAGCATAAAGCTCTAGAGTCGACCTGCAGGCATGCAAGC  
[0392] TTCGAGTCCCTGCTCGTCACGCTTTCAGGCACCGTGCCAGATATCGACGTGGAGTCGATCACTGTGATTGGCGA  
[0393] AGGGGAAGGCAGCGCTACCCAAATCGCTAGCTTGCTGGAGAAGCTGAAACAAACACGGGCATTGATCTGGCGA  
[0394] AATCCCTACCGGGTCAATCCGACTCGCCGCTGCGAAGTCTTAAGAGATAGCGATGTGACCGCGATCGCTTGTC  
[0395] AAGAATCCAGTGATCCCGAACCATAGGAAGGCAAGCTCAATGCTTGCTCGTCTTGAGGACTATCTAGATGTC  
[0396] TGTGGAACGCACATTTATTGCCATCAAGCCGATGGCGTTTCAGCGGGGTTTGGTCGGTACGATCATCGGCCGCT  
[0397] TTGAGCAAAAAGGCTTCAAACTGGTGGGCCTAAAGCAGCTGAAGCCAGTCGCGAGCTGGCCGAACAGCACTAT  
[0398] GCTGTCCACCGCGAGCGCCCTTCTTCAATGGCCTCGTCGAGTTTCATCACTCTGGGCCGATCGTGGCGATCGT  
[0399] CTTGGAAGGCGAAGGCGTTGTGGCGGCTGCTCGCAAGTTGATCGGCGCTACCAATCCGCTGACGGCAGAACCGG  
[0400] GCACCATCCGTGGTGATTTTGGTGTCAATATTGGCCGCAACATCATCCATGGCTCGGATGCAATCGAAACAGCA  
[0401] CAACAGGAAATTGCTCTCTGGTTTAGCCAGCAGAGCTAAGTGATTGGACCCCCACGATTCAACCCTGGCTGTA  
[0402] CGAATAAGGTCTGCATTCTTCAGAGAGACATTGCCATGCCGTGCTGCGATCGCCCTTCCAAGCTGCCCTTGCC  
[0403] CCGCTGTTTCGGGCTGGCAGCCCTGGCGTTGGGGCTGGCGACCGCTTGCCAAGAAAGCAGCGCTCCGCCGGCTG  
[0404] CCGGATC  
[0405] SEQ ID NO :10  
[0406] CGCCGGGGCTGGCAGCTTAGTCCTGCGCAATCTCTACTACATCTGCCAACCCAGTGAAATTTTGATCTTTGCTG  
[0407] GCAGTAGTCGCGCAGTAGTGATGGCCGCCGAGTTGGCTATCGCTTGGTCAAGGGCGGCAGCAGCCTGCGGGTA  
[0408] CCTCTGCTGGAAAAAGCGCTCCGCATGGATCTGACCAACATGATCATTGAGTTGCGCGTTTCCAATGCCCTTCTC  
[0409] CAAGGGCGGCATTCCCCTGACTGTTGAAGGCGTTGCCAATATCAAGATTGCTGGGGAAGAACCGACCATCCACA  
[0410] ACGCGATCGAGCGGCTGCTTGGCAAAAACCGTAAGGAAATCGAGCAAATTGCCAAGGAGACCTCGAAGGCAAC  
[0411] TTGCGTGGTGTTTTAGCCAGCCTCACGCCGGAGCAGATCAACGAGGACAAAATTGCCTTTGCCAAAAGTCTGCT  
[0412] GGAAGAGGCGGAGGATGACCTTGAGCAGCTGGGTCTAGTCTCGATACGCTGCAAGTCCAGAACATTTCCGATG  
[0413] AGGTGCGTTATCTCTCGGCTAGTGACGCAAGCAGCGGGCTGATCTGCAGCGAGATGCCGAATTGCTGAAGCC  
[0414] GATGCCAGGCTGCCTCTGCGATCCAAACGGCCGAAAATGACAAGATCACGGCCCTGCGTCGGATCGATCGCGA  
[0415] TGTAGCGATCGCCCAAGCCGAGGCCGAGCGCCGATTTCAGGATGCGTTGACGCGGCGCGAAGCGGTGGTGGCCG  
[0416] AAGCTGAAGCGGACATTGCTACCGAAGTCGCTCGTAGCCAAGCAGAACTCCCTGTGCAGCAGGAGCGGATCAAA

[0417] CAGGTGCAGCAGCAACTTCAAGCCGATGTGATCGCCCCAGCTGAGGCAGCTTGTAACGGGCGATCGCGAAGC  
[0418] GCGGGGGCCGCCGCCGTATCGTCGAAGATGGAAGCTCAAGCGGAAGGGACCAACGGCTGGCGGAGGCTT  
[0419] GGCAGACCGCTGGTGCTAATGCCCGCGACATCTTCTGCTCCAGAAGCTCGAAATTCGAGCTCGGTACCATTTA  
[0420] CGTTGACACCATCGAATGGTGAAAACTTTTCGGGTATGGCATGATAGCGCCCGAAGAGAGTCAATTCAGGG  
[0421] TGGTGAATGTGAAACCAGTAACGTTATACGATGTGCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGC  
[0422] GTGGTGAACCAGGCCAGCCACGTTTCTCGGAAAAACGCGGAAAAAGTGGAAGCGCGATGGCGGAGCTGAATTA  
[0423] CATTTCCCAACCGCTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGG  
[0424] CCCTGCACGCGCGTTCGCAAAATGTTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTG  
[0425] TCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGTGACAAATCTTCTCGCGCAACGCGTCAGTGG  
[0426] GCTGATCATTAATATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGT  
[0427] TATTTCTTGATGTCTCTGACCAGACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGC  
[0428] GTGGAGCATCTGGTCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGGCGCG  
[0429] TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGAAGGCGACT  
[0430] GGAGTGCCATGTCCGGTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTGCGATGCTGGTT  
[0431] GCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGTGCGCGTTGGTGCGGATATCTC  
[0432] GGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCACCATCAAAACAGGATTTTC  
[0433] GCCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGTGAAAGGCAATCAGCTG  
[0434] TTGCCCCGTCTCACTGGTGAAAAGAAAAACACCCTGGCGCCCAATACGCAAAACCGCCTCTCCCCGCGCGTTGGC  
[0435] CGATTTCATTAATGCAGCTGGCAGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTA  
[0436] AGTTAGCGGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCGTCAGGCA  
[0437] GCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCTGTGCTCGCTCAAGGCGCACTCCC  
[0438] GTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTTGACAATTA  
[0439] ATCATCCGGCTCGTATAATGTGTGGAATGTGAGCGGATAACAATTTACACAGGAAACAGACCATGGCGAATG  
[0440] GTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCCCCTCCTCGGACG  
[0441] TTTTTCATCAGCTGCCGACTGGAGTCGTTGCTGACCGCCATCACAACAGTGTGTTGTCAAATCTAAACGACC  
[0442] GGACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTACTGTGCAGG  
[0443] ACGGCTGGTGTTCCGTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTAGCATTGAA  
[0444] ACGTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGGTTTGG  
[0445] CCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAAAATGCAGATCAAAGTTAATCGATACC  
[0446] CGGCCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCGTTTGGGCAAAATGGCATGGGTGCGGATTGG  
[0447] CTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGTCTGCGTACGCGATGATGAATCAAAAGAC  
[0448] GCGTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATGTTCCATTGTTGTTGATAGTCCCGTTATCG  
[0449] AGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACGCCAGGT  
[0450] TGGAACGATCTGGATGTGAACCAACAGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAAGTATGCCTAC  
[0451] GGAAGTCTGGAGACGAGGAACCTCTGCAGTCTCGCTCTGGAGTACCGCGTGAGTGTGGCGGTGATTCCGTGC  
[0452] TCGAGTCCGTCACTGCGATGGACCCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGGCGCTCGAA  
[0453] GATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAATTAGTAC  
[0454] CGGGAAAACTCCAATGGAAACAGCGTCAGCTAATGATAGGATCCGAGCTCGAGATCTGCAGCTGGTACCATAT  
[0455] GGAATTCGAAGCTTGGCTGTTTTGGCGGATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGAG

[0456] AAGCGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACCTCAG  
[0457] AAGTGAACGCGGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACGCCAGGCATCAAAT  
[0458] AAAACGAAAGGCTCAGTCGAAAGACTGGGCCTTTCGTTTATCTGTTGTTGTGGTGAACGCTCTCCTGAGTA  
[0459] GGACAAATCCGCCGGGAGCGGATTGAACGTTGCGAAGCAACGGCCCCGAGGGTGGCGGGCAGGAAGCCCCCA  
[0460] TAAACTGCCAGGCATCAAATTAAGCAGAAGGCCATCCTGACGGATGGCCTTTTGTGCTTCTACAAACTCTTTT  
[0461] GTTTATTTTCTAAATACATTCAAATATGTATCCGCTCATGGGGATCCGACTAGTAGGCCTCGAGGAATTCACG  
[0462] CGTACGTAGATCTCCGCCGCCCGGATCCTCTAGTATGCTTGTAAACCGTTTGTGAAAAATTTTAAAAATAA  
[0463] AAAAGGGGACCTCTAGGGTCCCCAATTAATTAGTAATATAATCTATTAAAGGTCAATCAAAGGTCAATCCACCG  
[0464] GATCAGCTTAGTAAAGCCCTCGCTAGATTTAATGCGGATGTTGCGATTACTTCGCCAACTATTGCGATAACAA  
[0465] GAAAAAGCCAGCCTTTCATGATATATCTCCAATTTGTGTAGGGCTTATTATGCAAGCTTAAAAATAATAAAG  
[0466] CAGACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTGCATCTAACGCTTGAGTTAAGCCGCGCG  
[0467] CGAAGCGGCGTCGGCTTGAACGAATTGTAGACATTATTGCGGACTACCTTGGTGATCTCGCCTTTCACGTAG  
[0468] TGGACAAATTTCTCCAAGTATCTGCGCGCGAGGCCAAGCGATCTTCTTCTTGTCCAAGATAAGCCTGTCTAGC  
[0469] TTCAAGTATGACGGGTGATACTGGGCCGGCAGGCGCTCCATTGCCAGTCGGCAGCGACATCCTTCGGCGCGA  
[0470] TTTTGGCGGTTACTGCGCTGTACCAATGCGGGACAACGTAAGCACTACATTTGCTCATCGCCAGCCCAGTCG  
[0471] GGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTTTCAGGAACCGGATCAAAGAG  
[0472] TTCTCCGCCGCTGGACCTACCAAGGCAACGCTATGTTCTTCTTGTGTCAGCAAGATAGCCAGATCAATGT  
[0473] CGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCAATGCGCTGCCATTCTCCAATTCAGATTTCGCGCTTA  
[0474] GCTGGATAACGCCACGGAATGATGTCGTCGTGCACAACAATGGTGACTTCTACAGCGCGGAGAATCTCGCTCTC  
[0475] TCCAGGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAAGCCTTACGGTCA  
[0476] CCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTACAAATGTACGGCC  
[0477] AGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATACTTCGGCGATCAC  
[0478] CGCTTCCCTCATGATGTTTAACTTTGTTTATAGGGCGACTGCCCTGCTGCGTAACATCGTTGCTGCTCCATAACA  
[0479] TCAAACATCGACCCACGGCGTAACGCGCTTGTGCTTGGATGCCCGAGGCATAGACTGTACCCCAAAAAAACAG  
[0480] TCATAACAAGCCATGAAAAACGCCACTGCGCCGTTACCACCGCTGCGTTCCGGTCAAGGTTCTGGACCAGTTGCG  
[0481] TGAGCGCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGGGTTGCTGCTTCATCC  
[0482] GTTTCACCGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCTGTCTGGCTGGCGAAC  
[0483] GAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCCTTGTGTTCTTCTACGGCAAGGTGCTGTG  
[0484] CACGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCTGGTGTGACCCCGG  
[0485] ATGAAGTGGTTCGCATCCTCGGTTTTCTGGAAGGCGAGCATCGTTTGTTCGCCAGCTTCTGTATGGAACGGGC  
[0486] ATGCGGATCAGTGAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTGATCAGGCACGATCATCGTGGCGGA  
[0487] GGGCAAGGGCTCCAAGGATCGGGCCTTGATGTTACCCGAGAGCTTGGCACCCAGCCTGCGCGAGCAGGGGAATT  
[0488] GATCCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATATTACTAATTAATTGGGGACCTAGAGGTCCCC  
[0489] TTTTTTATTTTAAAAATTTTTTACAAAAACGGTTTACAAGCATAAAGCTCTAGAGTCGACCTGCAGGCATGCAA  
[0490] GCTTCGAGTCCCTGCTCGTCACGCTTTCAGGCACCGTGCCAGATATCGACGTGGAGTCGATCACTGTGATTGGC  
[0491] GAAGGGGAAGGCAGCGCTACCCAAATCGCTAGCTTGTGAGGAAGCTGAAACAAACCACGGGCATTGATCTGGC  
[0492] GAAATCCCTACCGGGTCAATCCGACTCGCCGCTGCGAAGTCTAAGAGATAGCGATGTGACCGCGATCGCTTG  
[0493] TCAAGAATCCAGTGATCCCGAACCATAGGAAGGCAAGCTCAATGCTTGCTCGTCTTGAGGACTATCTAGATG  
[0494] TCTGTGGAACGCACATTTATTGOCATCAAGCCGATGGCGTTCAGCGGGTTTGGTCCGTACGATCATCGGCCG

[0495] CTTTGAGCAAAAAGGCTTCAAACCTGGTGGGCCTAAAGCAGCTGAAGCCCAGTCGCGAGCTGGCCGAACAGCACT  
[0496] ATGCTGTCCACCGCGAGCGCCCTTCTTCAATGGCCTCGTCGAGTTTCATCACCTCTGGGCGGATCGTGGCGATC  
[0497] GTCTTGGAAGGCGAAGGCGTTGTGGCGGCTGCTCGCAAGTTGATCGGCGCTACCAATCCGCTGACGGCAGAACCC  
[0498] GGGCACCATCCGTGGTGATTTTGGTGTCATATTGGCCGCAACATCATCCATGGCTCGGATGCAATCGAAACAG  
[0499] CACAACAGGAAATTGCTCTCTGGTTTAGCCAGCAGAGCTAAGTGATTGGACCCCCACGATTCAACCCCTGGCTG  
[0500] TACGAATAAGGTCTGCATTCTTCAGAGAGACATTGCCATGCCCGTGCTGCGATCGCCCTTCCAAGCTGCCCTG  
[0501] CCCCCTGTTTTCGGGCTGGCAGCCCTGGCGTTGGGGCTGGCGACCGCTTGCCAAGAAAGCAGCGCTCCGCCGGC  
[0502] TGCCGGATC  
[0503] SEQ ID NO :11  
[0504] CTTCGGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAG  
[0505] GCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAAGGCCAGCAAAAAGGC  
[0506] CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCTGACGAGCATCACAAAAATC  
[0507] GACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTC  
[0508] GTGCGCTCTCCTGTTCCGACCCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCT  
[0509] TTCTCATAGCTCACGCTGTAGGTATCTCAGTTCCGGTGTAGGTGCTTCCGCTCCAAGCTGGGCTGTGTGCACGAAC  
[0510] CCCCCGTTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTA  
[0511] TCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA  
[0512] GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCG  
[0513] GAAAAAGAGTTGGTAGCTCTTGATCCGGCAAAACAAACCCCGCTGGTAGCGGTGGTTTTTTTGTGTGCAAGCAG  
[0514] CAGATTACGCGCAGAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGA  
[0515] CGAAAACTCAGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAA  
[0516] AATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG  
[0517] GCACCTATCTCAGCGATCTGTCTATTTTCGTTTCATCCATAGTTGCCTGACTCCCGCTCGTGTAGATAACTACGAT  
[0518] ACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTAT  
[0519] CAGCAATAAACAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT  
[0520] ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTAC  
[0521] AGGCATCGTGGTGTACGCTCGTCGTTTGGTATGGCTTCATTTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTA  
[0522] CATGATCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCC  
[0523] GCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATCTCTTACTGTCATGCCATCCGTAAGATGCTTTTC  
[0524] TGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGT  
[0525] CAATACGGGATAATACCGGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGAAAAACGTTCTTCGGGGCGA  
[0526] AAATCTCAAGGATCTTACCGCTGTTGAGATCCAGTTTCGATGTAACCCACTCGTGACCCCAACTGATCTTCAGC  
[0527] ATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAGGGAATAAGGG  
[0528] CGACACGGAATGTTGAATACTCATACTCTTCCTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTC  
[0529] ATGAGCGGATACATATTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT  
[0530] GCCACCTGACGTCTAAGAAACATTATTATCATGACATTAACTATAAAAAATAGGCGTATCACGAGGCCCTTTC  
[0531] GTCTCGCGGTTTTTCGGTGATGACGGTGAACCTCTGACACATGCAGCTCCCGGAGACGGTACAGCTTGTCTG  
[0532] TAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGGTCAGCGGGTGTGGCGGGTGTGGGGCTGGCTTAA  
[0533] CTATGCCGCATCAGAGCAGATTGTACTGAGAGTGACCAATAAAATTGTAAACGTTAATATTTGTAAAAATTCG

[0534] CGTTAAATTTTGTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAA  
[0535] GAATAGCCCGAGATAGGGTTGAGTGTGTTCAGTTTGGAAACAAGAGTCCACTATTAAGAAGCTGGACTCCAA  
[0536] CGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCAAGTTTTTTTGG  
[0537] GGTCGAGGTGCCGTAAAGCACTAAATCGGAACCTAAAGGGAGCCCCGATTAGAGCTTGACGGGAAAAGCCG  
[0538] GCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTAC  
[0539] GCTGCGCGTAACCACACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCTACTATGGTTGCTTTGACGTAT  
[0540] GCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAAATACCGCATCAGGCGCCATTCCGCATTAGGCTGCGCA  
[0541] ACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGG  
[0542] CGATTAAGTTGGGTAACGCCAGGGTTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGCCAAGCTATTGCT  
[0543] GAAGCGGAATCCCTGGTTAATGCGCGCGCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACGCTTCAA  
[0544] CCCGGCATTTTTAGAGCTAACCAAAATTTCTAAAACGGAAGAGTTATTGGCGATCGAAGCCCATCGCATGAGTC  
[0545] CCTATTTCCAGCGGGCCAATGATGTCCTCGTGGTATTGGATTGTATGATCCATGACATTGACCTGTTGCTGGAA  
[0546] TTGGTGGGTTTGGGAAGTGGTTAAACTGTCCGCCAGTGGCAGTCGGGCTTCTGGGTGAGGATATTTGGATTATGT  
[0547] CACCGTACGTTAGGCTTCTCCTCCGGCATTTGTGGCCACCCTCACCGCCAGTAAGGTCACCCATCGTAAAAATTC  
[0548] GTTCCATCGCCGCCCACTGCAAAAATTCCTCACCGAAGCGGATTTTCTCAATAACGAAATTTTGATCCATCGC  
[0549] CAAACCACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAAGGTTTA  
[0550] CACCAGTAATATTGAACCTCTCCACGCTGAATTAGAACATTTTATTTCATTGTGTTAGGGGAGGTGATCAACCT  
[0551] CAGTGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGACAGTCAG  
[0552] GAATGGCATGGGGGGGAAGTTGTGACAGAAATCAAGATGCCACCCTGGCCCTCAGTGGCAGTGTTTAAATCAA  
[0553] CTTAATTAATGCAATTATTGCGAGTTCAAACCTCGATAACTTTGTGAAATATTACTGTTGAATTAATCTATGACT  
[0554] ATTCAATACACCCCCCTAGCCGATCGCTGTTGGCTACCTCGCCGCGATCGCCTAAATCTCAGCGCCAAGAG  
[0555] TAGTTCCCTCAACACCAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGGAATTGTAACAGCCAACT  
[0556] ATGGCTTTGATGGTTATATGGTACCATATGCATGCGAGCTCAGATCTACCAGTTGTCTTTGGCGCAGCGCTTC  
[0557] CCACGCTGAGAGGGTGTAGCCCGTCACGGGTAACCGATATCGTGCACAGGCCTCTAGACCCGGGCTCGAGCTAG  
[0558] CAAGCTTGGCCGGATCCGGCCGATCCGGAGTTTGTAGAAAACGCAAAAAGGCCATCCGTCAGGATGGCTTTCTG  
[0559] CTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCTCCGGGCGGTTGCTCCGCAACGTTCA  
[0560] AATCCGCTCCCGGCGGATTTGTCTACTCAGGAGAGCGTTACCGACAAACAACAGATAAAACGAAAGGCCAG  
[0561] TCTTTGACTGAGCCTTTGTTTTATTGATGCCTGGCAGTTCCCTACTCTCGCATGGGGAGACCCACACTAC  
[0562] CATCGGCGCTACGGCGTTTCACTTCTGAGTTCCGCATGGGGTCAGGTGGGACCACCGCGCTACTGCCGCCAGGC  
[0563] AAATTCGTTTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACATCCGATGGCAAGAGGCCC  
[0564] GAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAGTTATCCCCCTCCATCA  
[0565] GGCAGTTTCCAGACATTACTCACCCGTCGCCACTCGTCAGCAAAGAAGCAAGCTTAGATCGACCTGCAGGGG  
[0566] GGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAAAATATATCATCATGAA  
[0567] CAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTTCTGC  
[0568] TCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTGGGCA  
[0569] ATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTA  
[0570] GCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATC  
[0571] AAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAAACAGCATTCAGGT  
[0572] ATTAGAAGAAATATCCTGATTACAGGTGAAAAATATTGTTGATGCGCTGGCAGTGTCTCGCGCGGTTGCATTGCA

[0573] TTCCTGTTTGTAATTGTCCTTTAAACAGCGATCGCGTATTTCTGCTCGCTCAGGCGCAATCACGAATGAATAAC  
[0574] GGTTTGGTTGATGCGAGTGATTTTGTATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCA  
[0575] TAAGCTTTGCCATTCTCACCGGATTCACTGCTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTGTACG  
[0576] AGGGGAAATTAATAGGTTGTATTGATGTTGGAAGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTA  
[0577] TGGAAGTGCCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTCAAAAATATGGTATTGATAATCCTGA  
[0578] TATGAATAAAATTGCAGTTTCATTTGATGCTCGATGAGTTTCTAATCAGAATTGGTTAATTGGTTGTAACACT  
[0579] GGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTTGCTGAGTTGAAGGATC  
[0580] AGATCACGCATCTTCCCGACAAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAATCACCAACTGGTCCACC  
[0581] TACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTCTGCTGGATGATGGGGCGATTACAGGCCTGGTATGA  
[0582] GTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGCCCCCCCCCTGCAGGTGATCTGGTAACCCAGCG  
[0583] CGGTTGCTACCAAGTAGTGACCGCTTCGTGATGCAAAATCCGCTGACGATATTCGGGCGATCGCTGCTGAATG  
[0584] CCATCGAGCAGTAACGTGGCGAATTCGGTACCGGTATGGATGGCACCGATGCGGAATCCCAACAGATTGCCTTT  
[0585] GACAACAATGTGGCCTGGAATAACCTGGGGGATTGTCCACCACCAACCCAAACGGGCCTACACTTCGGCTATTAG  
[0586] CACAGACACAGTGCAGAGTGTTTATGGCGTTAATCTGGAATAAACGATAACATTCCCATTGTTTTGCGTGGC  
[0587] CCATTTTTCCACCACCTTAATCCACAGATTTTCAGGTAATGCTTAACACGCGGGGAAATTGTACCCCGGTG  
[0588] ATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAAAACGGTAGTAATTACGGGCAATTTTGGTAATCG  
[0589] TTTAACCCAGGCACGGAGGGAGCGATTATCCCGTTTCGTAGGCACAGTGTTGGACAGTACTCCTTTGGAAA  
[0590] TGGTGGGACCCAACGGCCCGGTCACTGCGGTGGGTATTACCATTTGATAGTCTCAACCCCTACGTGGCCGGCAAT  
[0591] GGTCCCAAAATTTGTCCCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGGCTCCCTCTGGTTAGCCAC  
[0592] CAATCAAAATAACAGTGGCGGGGATTATATGGAGACCAAGCCCAATTTGTTTGGCAATTTACACCAGCGCG  
[0593] GTTTTTCCCGATGGCATTGCCAGTTTACTACCCACAGAAATTTGAACGGTATTTTCAACTCCAAGCGGAAGAT  
[0594] ATTACGGGACGGACAGTTATCCTAACCCAACTGGTGTGATTATGAAATTTCCGGCTTTGGTCTGGTGACGGT  
[0595] GTTGGGGCTGGCGGATTGGCCGGGGTTCAGGACAGCTATGACCTGACTTACATCGAAGATCATGACAACCTATT  
[0596] ACGACATTATCCTCAAAGGGGACGAAGCCGCGATTCCGCCAAATTAAGAGGGTTGCTTTGCCCTCCGAAGGGGAT  
[0597] TATTCGGCGGTTTATAATCCCGGTGGCCCCGGCAATGATCCAGAGAATGGTCCCCAAATTCGTAATCATGTCA  
[0598] TAGCTGTTTCTGTGTGAAATTTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTA  
[0599] AGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAA  
[0600] ACCTGTGCTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTTCGTATTGGGCGCT  
[0601] SEQ ID NO :12  
[0602] ATTGCTGAAGCGGAATCCCTGGTTAATGCCGCCGCCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACG  
[0603] CTTCAACCCGGCATTTTATAGACTAACCAAAATTTCTCAAAACGGAAGAGTTATTGGCGATCGAAGCCCATCGCA  
[0604] TGAGTCCCTATTCCAGCGGGCCAATGATGTCTCCGTGGTATTGGATTTGATGATCCATGACATTGACCTGTTG  
[0605] CTGGAATTGGTGGGTTCGGAAGTGGTTAAACTGTCCGCCAGTGGCAGTCGGGCTTCTGGGTGAGGATATTGGA  
[0606] TTATGTCACCGCTACGTTAGGCTTCTCCTCCGGCATTTGTGGCCACCTCACCGCCAGTAAGGTCAACCATCGTA  
[0607] AAATTCGTTCCATCGCCGCCCACTGCAAAAATTTCCCTCACCGAAGCGGATTTTCTCAATAACGAAATTTTGATC  
[0608] CATCGCCAAACACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAA  
[0609] GGTTTACACCAGTAATATTGAACCTCTCCACGCTGAATTAGAACATTTTATTCATTGTGTTAGGGGAGGTGATC  
[0610] AACCTCAGTGGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGAC  
[0611] AGTCAGGAATGGCATGGGGGGGAAGTTGTGACAGAATATCAAGATGCCACCTGGCCCTCAGTGCGAGTGTTTA

[0612] AATCAACTTAATTAATGCAATTATTGCGAGTTCAAACTCGATAACTTTGTGAAATATTACTGTTGAATTAATCT  
[0613] ATGACTATTCAATACACCCCCCTAGCCGATCGCCTGTTGGCCTACCTCGCCGCCGATCGCCTAAATCTCAGCGC  
[0614] CAAGAGTAGTTCCCTCAACACCAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGAATTGTAACAG  
[0615] CCAACTATGGCTTTGATGGTTATATGGTACCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
[0616] TAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTTCATGGCTGCGCCCCGACACCCGCCAACACCCGCTG  
[0617] ACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG  
[0618] TGTCAGAGGTTTTACCGTCATCACCGAAACGCGCGAGGCAGATCAATTCGCGCGGAAGGCGAAGCGGCA  
[0619] TGCATTTACGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGAAGAGAGTCA  
[0620] ATTCAGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCG  
[0621] TTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAAACGCGGGAAAAAGTGAAGCGGCGATGGCGGAG  
[0622] CTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTC  
[0623] CAGTCTGGCCCTGCACGCGCGTCCGAAATTTGTCGCGGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG  
[0624] TGGTGGTGTGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGC  
[0625] GTCAGTGGGCTGATCATTAACATATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT  
[0626] TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGC  
[0627] GACTGGGCGTGGAGCATCTGGTGCATTGGGTCAACAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTC  
[0628] TCGGCGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGA  
[0629] AGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTGCGA  
[0630] TGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCG  
[0631] GATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAACA  
[0632] GGATTTTTCGCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCA  
[0633] ATCAGCTGTTGCCCCGTCTCACTGGTGA AAAAGAAAAACCAACCTGGCGCCCAATACGCAAACCGCTCTCCCCGC  
[0634] GCGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCCCGACTGGAAAAGCGGGCAGTGAGCGCAACGCAA  
[0635] TTAATGTAAGTTAGCGGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGC  
[0636] GTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAATTGCTGTGCTCAAGGC  
[0637] GCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTT  
[0638] GACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGCGCG  
[0639] CTGAGAAAAAGCGAAGCGGCACTGCTCTTAAACAATTTATCAGACAATCTGTGTGGGCACTCGACCGGAATTAT  
[0640] CGATTAACTTTTATTATTAATAAAGAGGTATATTAATGTATCGATTAAATAAGGAGGAATAAACCATGG  
[0641] CGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCCCCTCCT  
[0642] CGGACGTTTTTGCATCAGCTGCCGACTGGAGTCGCTTGCTGACCGCCATCACAACAGTGTGTTGTCAAATCTAA  
[0643] ACGACCGGACATGCATGATCGGAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTACTG  
[0644] TGCAGGACGGCTGGTGTTCGTCATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTAGC  
[0645] ATTGAAACGTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGATGG  
[0646] TTTTGGCCGCACATTGGAAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAATGCAGATCAAAGTTAATC  
[0647] GATACCGGCCTGGGGAGATACGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTGGCATGGGTGCG  
[0648] GATTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGCTCTGCGTACGCGATGATGAATCA  
[0649] AAAGACGCGTCGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTTGTTCCATTGTTGTTGATAGTCCCG  
[0650] TTATCGAGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGACG

[0651] CCAGGTTGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGAAAAGTAT  
[0652] GCCTACGGAAGTCCCTGGAGACGCAGGAACCTCTGCAGTCTCGCTCTGGAGTACCGCCGTAGTGTGGCCGTGATT  
[0653] CCGTGCTCGAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTGCGAGTCAATACCAACACCTCTTGCGG  
[0654] CTCGAAGATGGGACCGCCATTGTGAACGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCAAT  
[0655] TAGTACCGGGAAAACCTCCAATGGAAACAGCGTCAGCTAATGATAGGATCCGAGCTCAGATCTACCAGGTTGTC  
[0656] CT'TGGCGCAGCGCTTCCCACGCTGAGAGGGTGTAGCCCGTCACGGGTAACCGATATCGTCGACAGGCCTCTAGA  
[0657] CCGGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTTGTAGAAACGCAAAAAGGCCATCCG  
[0658] TCAGGATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCCTCCGGGCGGT  
[0659] TGCTTCGCAACGTTCAAATCCGCTCCCGGCGGATTTGTCTACTCAGGAGAGCGTTCACCGACAAACAACAGAT  
[0660] AAAACGAAAGGCCAGTCTTTCGACTGAGCCTTTGTTTTATTGATGCCTGGCAGTTCCTACTCTCGCATGG  
[0661] GGAGACCCCACTACCATCGGCGCTACGGCGTTTCACTTCTGAGTTCGGCATGGGGTCAGGTGGGACCACCGC  
[0662] GCTACTGCCGCCAGGCAAAATCTGTTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACATC  
[0663] CGATGGCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTACCGTTTCCAGTAG  
[0664] TTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTCACCCGTCCGCCACTCGTCAGCAAAGAAGCAAGCTTA  
[0665] GATCGACCTGCAGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAAA  
[0666] AATATATCATCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCA  
[0667] ACGGGAAACGTTCTGCTCGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTC  
[0668] GCGATAATGTGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTTCTG  
[0669] AAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTAT  
[0670] GCCTCTTCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGA  
[0671] AAACAGCATTCAGGTATTAGAAGAATATCCTGATTGAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTG  
[0672] CGCCGGTTGCATTGATTCTGTTTGTAATTGTCCTTTTAAACAGCGATCGCGTATTTGCTCTCGCTCAGGCGCA  
[0673] ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAG  
[0674] TCTGGAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCATCATGGTGATTTCTCACTTGAT  
[0675] AACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCA  
[0676] GGATCTTGCCATCCTATGGAACCTGCCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATG  
[0677] GTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTT  
[0678] AATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTTTT  
[0679] GCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAAT  
[0680] CACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGCGA  
[0681] TTCAGGCTGGTATGAGTCAGCAACACCTTCTTCACGAGGCAGACCTCAGCGCCCCCCCCCCCCCTGCAGGTGGA  
[0682] TCTGGTAACCCAGCGCGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTGGGG  
[0683] CGATCGCTGCTGAATGCCATCGAGCAGTAACGTGGCGAATTCGGTACCGGTATGGATGGCACCGATGCGGAATC  
[0684] CCAACAGATTGCCTTTGACAACAATGTGGCCTGGAATAACCTGGGGGATTTGTCCACCACCACCAACGGGCCT  
[0685] ACATTTCGGCTATTAGCACAGACACAGTGCAGAGTGTATTGCGGTTAATCTGGAAAAAAGATAACATTCCC  
[0686] ATTGTTTTTGCGTGGCCCATTTTTCCCAACCACCTTAATCCACAGATTTTCAGGTAATGCTTAACACGGGGGA  
[0687] AATTGTCACCCCGGTGATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAACGGTAGTAATTACGG  
[0688] GCAATTTTGGTAATCGTTTAACCCCAGGCACGGAGGGAGCGATTTATCCCGTTTCCGTAGGCACAGTGTGGAC  
[0689] AGTACTCCTTTGGAAATGGTGGGACCCAACGGCCCGGTGAGTGGGTGGGTATTACCATTGATAGTCTCAACCC

[0690] CTACGTGGCCGGCAATGGTCCCAAAATTGTCGCCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGGCTC  
[0691] CCCTCTGGTTAGCCACCAATCAAAATAACAGTGGCGGGGATTTATATGGAGACCAAGCCCAATTTCTGTTTGGCA  
[0692] ATTTACACCAGCGCCGGTTTTTCCCCGATGGCATTGCCAGTTTACTACCCACAGAATTTGAACGGTATTTTCA  
[0693] ACTCCAAGCGGAAGATATTACGGGACGGACAGTTATCCTAACCCAACTGGTGTTGATTATGAAATTTCCCGCT  
[0694] TTGGTCTGGTGCAGGTGTTGGGGCTGGCGGATTTGGCCGGGTTTCAGGACAGCTATGACCTGACTTACATCGAA  
[0695] GATCATGACAACTATTACGACATTATCCTCAAAGGGGACGAAGCCGCAGTTCCGCAAAATTAAGAGGGTTGCTTT  
[0696] GCCCTCCGAAGGGGATTATTTCGGCGGTTTATAATCCCGGTGGCCCCGGCAATGATCCAGAGAATGGTCCCCCA  
[0697] SEQ ID NO :13  
[0698] ATTGCTGAAGCGGAATCCCTGGTTAATGCCGCGCCGATGCCAATTGCATTCTCCAAGTGGGGCACATTGAACG  
[0699] CTTCACCCCGGCATTTTTAGAGCTAACCAAAATTTCTCAAAACGGAAGAGTTATTGGCGATCGAAGCCCATCGCA  
[0700] TGAGTCCCTATTCCAGCGGGCCAATGATGTCTCCGTGGTATTGGATTTGATGATCCATGACATTGACCTGTTG  
[0701] CTGGAATTGGTGGGTTCCGAAGTGGTTAAACTGTCCGCCAGTGGCAGTCGGGCTTCTGGGTGAGGATATTTGGA  
[0702] TTATGTCAACGCTACGTTAGGCTTCTCCTCCGGCATTGTGGCCACCCTCACCGCCAGTAAGGTCACCCATCGTA  
[0703] AAATTCGTTCCATCGCCGCCACTGCAAAAATTCCTCACCGAAGCGGATTTTCTCAATAACGAAATTTTGATC  
[0704] CATCGCCAAACACCGCTGATTGGAGCGCGGACTATGGCCAGGTATTGTATCGCCAGGATGGTCTAATCGAAAA  
[0705] GGTTCACACAGTAATATTGAACCTCTCCACGCTGAATTAGAACATTTTATTCATTGTGTTAGGGGAGGTGATC  
[0706] AACCCTCAGTGGGGGGAGAACAGGCCCTCAAGGCCCTGAAGTTAGCCAGTTTAATTGAAGAAATGGCCCTGGAC  
[0707] AGTCAGGAATGGCATGGGGGGGAAGTTGTGACAGAATATCAAGATGCCACCCTGGCCCTCAGTGCGAGTGTTTA  
[0708] AATCAACTTAATTAATGCAATTATTGCGAGTTCAAACTCGATAACTTTGTGAAATATTACTGTTGAATTAATCT  
[0709] ATGACTATTCAATACACCCCCCTAGCCGATCGCTGTTGGCCTACCTCGCCGCCGATCGCTAAATCTCAGCGC  
[0710] CAAGAGTAGTTCCCTCAACACCAGTATTCTGCTCAGCAGTGACCTATTCAATCAGGAAGGGGGAATTGTAACAG  
[0711] CCAACTATGGCTTTGATGGTTATATGGTACCATATGGTGCACCTCTCAGTACAATCTGCTCTGATGCCGCATAGT  
[0712] TAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTGATGGCTGCGCCCCGACACCCGCCAACACCCGCTG  
[0713] ACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGTGACCGTCTCCGGGAGCTGCATG  
[0714] TGTGACAGGTTTTTCACCGTCATCACCGAAACGCGCGAGGCAGAGATCAATTGCGCGCGGAAGGCGAAGCGGCA  
[0715] TGCATTTACGTTGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCA  
[0716] ATTCAGGGTGGTGAATGTGAAACAGTAACGTTATACGATGTGCGAGAGTATGCCGGTGCTCTTTATCAGACCG  
[0717] TTTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAAACGCGGGAAAAAGTGAAGCGGCGATGGCGGAG  
[0718] CTGAATTACATTCCCAACCGCGTGGCACAACAACCTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTC  
[0719] CAGTCTGGCCCTGCACGCGCGTCGCAAAATGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG  
[0720] TGGTGGTGTGATGGTAGAACGAAGCGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGC  
[0721] GTCAGTGGGCTGATCATTAATATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT  
[0722] TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGC  
[0723] GACTGGGCGTGGAGCATCTGGTGCATTGGGTACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCT  
[0724] TCGGCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGA  
[0725] AGGCGACTGGAGTGCCATGTCCGTTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGA  
[0726] TGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGCG  
[0727] GATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCACCATCAACA  
[0728] GGATTTTCGCTGCTGGGGCAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCA

[0729] ATCAGCTGTTGCCCCGTCTCACTGGTGAAAAAGAAAAACCAACCTGGCGCCCAATACGCAAACCGCTCTCCCCGC  
[0730] GCGTTGGCCGATTTCATTAATGCAGCTGGCAGACAGGTTTCCCAGCTGGAAGCGGGCAGTGAGCGCAACGCAA  
[0731] TTAATGTAAGTTAGCGGAATTGATCTGGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGC  
[0732] GTCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTCGTGTGCTCAAGGC  
[0733] GCACTCCCGTTCTGGATAATGTTTTTTGCGCCGACATCATAACGGTTCTGGCAAATATTCTGAAATGAGCTGTT  
[0734] GACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGACCAT  
[0735] GGCGAATGGTTCTGCAGTCTCTTTGAAATCTGGAAGCTTGAATACGCAGGAGGATACTAGTTCCAGTCCCCCTC  
[0736] CTCGGACGTTTTTGCATCAGCTGCCCCAGCTGGAGTCGCTTGCTGACCGCCATCACAACAGTGTGTGTCAAATCT  
[0737] AAACGACCGGACATGCATGATCGAAAAAGCAAGCGCCAGATATGCTCGTCGATAGTTTCGGACTCGAGTCTAC  
[0738] TGTGCAGGACGGCCTGGTGTTCGGTCAATCCTTCAGCATCCGAAGCTACGAGATTGGTACGGACCGTACCGCTA  
[0739] GCATTGAAACGTTGATGAACCATCTCCAAGAAACAGTTTGAACCACTGCAAGAGCACGGGCATCCTGCTGGAT  
[0740] GGTTTTGGCCGCACATTGGAATGTGCAAGCGAGACTTGATCTGGGTGGTCATTAAATGCAGATCAAAGTTAA  
[0741] TCGATACCCGGCCTGGGGAGATACCGTTGAGATCAATACACGCTTTTCCCGTTTGGGCAAAATTGGCATGGGTC  
[0742] GCGATTGGCTGATCTCCGACTGCAACACCGGTGAGATCTTGGTCCGTGCAACGTCGTGCTACGCGATGATGAAT  
[0743] CAAAAGACGCGTCGGTTGAGTAAGCTGCCGTATGAAGTTCACCAAGAAATTGTTCCATTGTTTCGTTGATAGTCC  
[0744] CGTTATCGAGGATTCTGACCTCAAAGTCCACAAGTTTAAAGTCAAGACTGGCGATTCCATCCAGAAGGGCCTGA  
[0745] CGCCAGGTTGGAACGATCTGGATGTGAACCAACACGTTAGCAACGTTAAGTATATCGGCTGGATCTTGGAAAGT  
[0746] ATGCCTACGGAAGTCCTGGAGACGCAGGAACTCTGCAGTCTCGCTCTGGAGTACCGCCGTGAGTGTGGCCGTGA  
[0747] TTCCGTGCTCGAGTCCGTCACTGCGATGGACCCTAGCAAAGTGGGTGTTTCGAGTCAATACCAACACCTCTTGC  
[0748] GGCTCGAAGATGGGACCGCCATTGTGAACGGCGCGACCGAATGGCGCCCCAAAAATGCCGGCGCTAACGGGGCA  
[0749] ATTAGTACCGGAAAAACCTCCAATGGAAACAGCGTCAGCTAATGATAGGATCCGAGCTCAGATCTACCAGGTTG  
[0750] TCCTTGGCGCAGCGCTTCCACGCTGAGAGGGGTAGCCCGTCACGGGTAACCGATATCGTCGACAGGCCTCTA  
[0751] GACCCGGGCTCGAGCTAGCAAGCTTGGCCGGATCCGGCCGGATCCGGAGTTGTAGAAACGCAAAAAGGCCATC  
[0752] CGTCAGGATGGCCTTCTGCTTAATTTGATGCCTGGCAGTTTATGGCGGGCGTCTGCCCGCCACCCTCCGGGCC  
[0753] GTTGCTTCGCAACGTTCAAATCCGCTCCCGCGGATTTGTCTACTCAGGAGAGCGTTCACCGACAAACAACAG  
[0754] ATAAAACGAAAGGCCAGTCTTTTGCAGTGAACCTTTTCGTTTATTTGATGCCTGGCAGTTCCCTACTCTCGCAT  
[0755] GGGGAGACCCACACTACCATCGGCGCTACGGCGTTTCACTTCTGAGTTCCGGCATGGGGTCAGGTGGGACCACC  
[0756] GCGCTACTGCCGCCAGGCAAAATCTGTTTATTGAGCCGTTACCCACCTACTAGCTAATCCCATCTGGGCACA  
[0757] TCCGATGGCAAGAGGCCCGAAGGTCCCCCTCTTTGGTCTTGCGACGTTATGCGGTATTAGCTACCGTTTCCAGT  
[0758] AGTTATCCCCCTCCATCAGGCAGTTTCCCAGACATTACTCACCCGTCCGCCACTCGTCAGCAAAGAAGCAAGCT  
[0759] TAGATCGACCTGCAGGGGGGGGGGAAAGCCACGTTGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATA  
[0760] AAAATATATCATCATGAACAATAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATT  
[0761] CAACGGGAAACGTTCTTGTCTGAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGC  
[0762] TCGCGATAATGTGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTC  
[0763] TGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAACTGGCTGACGGAATTT  
[0764] ATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCCTGCGATCCCCGG  
[0765] GAAAACAGCATTCAGGTATTAGAAGATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTCC  
[0766] TGCGCCGGTTGCATTGATTCCTGTTTGTAAATGTTCCTTTTAAACAGCGATCGCGTATTTGCTCTCGCTCAGGCG  
[0767] CAATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACA

[0768] AGTCTGGAAGAAATGCATAAGCTTTTGCCATTCTCACCGATTTCAGTCGTCACATCATGGTGATTCTCACTTG  
[0769] ATAACCTTATTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATAC  
[0770] CAGGATCTTGCCATCCTATGGAAGTGCCTCGGTGAGTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATA  
[0771] TGGTATTGATAATCCTGATATGAATAAAATTGCAGTTTCATTGATGCTCGATGAGTTTTCTAATCAGAAATTGG  
[0772] TTAATTGGTTGTAACACTGGCAGAGCATTACGCTGACTTGACGGGACGGCGGCTTTGTTGAATAAATCGAACTT  
[0773] TTGCTGAGTTGAAGGATCAGATCACGCATCTTCCCGACAACGCAGACCGTTCCGTGGCAAAGCAAAAGTTCAAAA  
[0774] ATCACCAACTGGTCCACCTACAACAAAGCTCTCATCAACCGTGGCTCCCTCACTTTCTGGCTGGATGATGGGGC  
[0775] GATTACAGGCTGGTATGAGTCAGCAACACCTTCTTCAGGAGCAGACCTCAGCGCCCCCCCCCCCCCTGCAGGTC  
[0776] GATCTGGTAACCCAGCGCGGTTGCTACCAAGTAGTGACCCGCTTCGTGATGCAAAATCCGCTGACGATATTTCG  
[0777] GCGGATCGCTGCTGAATGCCATCGAGCAGTAACGTGGCGAATTCGGTACCGGTATGGATGGCACCGATGCGGAA  
[0778] TCCCAACAGATTGCCTTTGACAACAATGTGGCCTGGAATAACCTGGGGGATTTGTCCACCACCACCAACGGGC  
[0779] CTACACTTCGGCTATTAGCACAGACACAGTGCAGAGTGTATTATGGCGTTAATCTGGAACAAAAACGATAACATTC  
[0780] CCATTGTTTTTGGTGGCCCATTTTTCCACCACCCCTTAATCCACAGATTTTCAGGTAATGCTTAACACGGGG  
[0781] GAAATTGTACCCCGGTGATCGCCTCTTTGATTCCCAACAGTGAATACAACGAACGGCAACGGTAGTAATTAC  
[0782] GGGCAATTTTGGTAATCGTTTAACCCAGGCAACGAGGGAGCGATTTATCCCGTTTCCGTAGGCACAGTGTGG  
[0783] ACAGTACTCCTTTGGAATGGTGGGACCCAAACGGCCCGTTCAGTGGGTGGGTATTACCATTTGATAGTCTCAAC  
[0784] CCCTACGTGGCCGGCAATGGTCCCAAAATTTGTGCGCGCTAAGTTAGACCGCTTCAGTGACCTGGGGGAAGGGGC  
[0785] TCCCTCTGGTTAGCCACCAATCAAAATAACAGTGGCGGGGATTTATATGGAGACCAAGCCCAATTTCTGTTTGC  
[0786] GAATTTACACCAGCGCCGGTTTTTCCCGATGGCATTTGCCAGTTTACTACCCACAGAATTTGAACGGTATTTT  
[0787] CAACTCCAAGCGGAAGATATTACGGGACGGACAGTTATCCTAACCCAAACTGGTGTGATTATGAAATTTCCCGG  
[0788] CTTTGGTCTGGTGCAGGTGTTGGGGCTGGCGATTGTCGGGGGTTTCAGGACAGCTATGACCTGACTTACATCG  
[0789] AAGATCATGACAACTATTACGACATTATCCTCAAAGGGGACGAAGCCGCGAGTTCCGCAAAATTAAGAGGGTTGCT  
[0790] TTGCCCTCCGAAGGGGATTATTTCGGCGGTTTATAATCCCGGTGGCCCCGGCAATGATCCAGAGAATGGTCCCCC  
[0791] A  
[0792] SEQ ID NO :14  
[0793] ACCCTGGCCCTCAGTGCGAG  
[0794] SEQ ID NO :15  
[0795] TGCTTCTTTGCTGACGAGTGG  
[0796] SEQ ID NO :16  
[0797] GTGACTGGAACCGCCCTCG  
[0798] SEQ ID NO :17  
[0799] CCATCGAGCAGTAACGTGGCCGATAGTGACGCTAAACCAGGCTG  
[0800] SEQ ID NO :18  
[0801] CGAGTGGCGGACGGGTGAGTCTACGAGGGCGTGCAGAAGC  
[0802] SEQ ID NO :19  
[0803] CACCAAGTTGCCTTCACCGAC  
[0804] SEQ ID NO :20  
[0805] CAGCCTGGTTTAGCGTCACTATCGGCCACGTTACTGCTCGATGG  
[0806] SEQ ID NO :21

[0807] GCTTCTGCACGCCCTCGTAGACTCACCCGTCCGCCACTCG  
[0808] SEQ ID NO :22  
[0809] GTGACTGGAACCGCCCTCGCGCAACCCCGCGCCATTACGCCCCACGAACAGCAGCTTTTGCCAAACTGAAAAG  
[0810] CTATCGGATATCCAAAGCTTGTGCGAAATTTGGGACGTGCTGCCAGTCAATTTGGATCGATGCCGGCTTTGG  
[0811] TTGCACCCCATGCCAAACCAGCGATCACCTCAGTTATCAAGAATTGGCGATTAGATCCAAGCGTTTGAGCC  
[0812] GGACTGCTCGGCTGGGAGTGCCTACCTCCACAGCCGATGACTTTCCGCCCTCGCTTGGCGCAGTTTGCGGATAA  
[0813] CAGCCCCCGCTGGTTGATTGCTGACCAAGGCACTTGTGTCAGGGGCTGCCAATGCGGTGCGCGGCGCCCAAG  
[0814] CTGAAGTATCGGAGCTGCTCTACGCTTAGAGGACAGCGGTCGATCGGCTTGATTGTCGAAGACGCGGCGCTG  
[0815] CTGAAGAACTACAGCCTGGTTTAGCGTCACTATCGGCCACGTTACTGCTCGATGGCATTACAGCAGCGATCGCC  
[0816] CGAATATCGTCAGCGGATTTTGCATCAGGAAGCGGTCCTACTTGGTAGCAACCGCGCTGGGGTTACCAGATC  
[0817] CGTCGATCATATCGTCAATTATTACCTCCACGGGAGAGCCTGAGCAAACTGGCCTCAGGCATTTGAGAAGCAC  
[0818] ACGGTCACACTGCTTCCGGTAGTCAATAAACCGGTAAACCAGCAATAGACATAAGCGGCTATTTAACGACCCTG  
[0819] CCCTGAACCGACGACCGGGTGAATTTGCTTTGCAATTTCTGCCATTATCCGCTTATTATCACTTATTCAGGC  
[0820] GTAGCACCAGGCGTTTAAGGGCACCAATAACTGCCTTAAAAAATTACGCCCCGCCCTGCCACTCATCGCAGTA  
[0821] CTGTTGTAATTCATTAAGCATTCTGCCGACATGGAAGCCATCACAACGGCATGATGAACCTGAATCGCCAGCG  
[0822] GCATCAGCACCTTGTCGCCCTTGCGTATAATATTTGCCCATGGTGAAAACGGGGGCGAAGAAGTTGTCCATATTG  
[0823] GCCACGTTTAAATCAAACTGGTGAACTCACCCAGGGATTGGCTGAGACGAAAAACATATTTCTCAATAAAACC  
[0824] TTTAGGAAAATAGGCCAGGTTTTCACCGTAACAGCCACATCTTGCGAATATATGTGTAGAAACTGCCGAAAT  
[0825] CGTCGTGGTATTCACTCCAGAGCGATGAAAACGTTTCAGTTTGCTCATGGAAAACGGTGTAACAAGGGTGAACA  
[0826] CTATCCCATATCACCAGCTCACCGTCTTTCATTGCCATACGGAATTCGGATGAGCATTATCAGGCGGGCAAG  
[0827] AATGTGAATAAAGGCCGATAAACTTGTGCTTATTTTTCTTTACGGTCTTTAAAAAGGCCGTAATATCCAGCT  
[0828] GAACGGTCTGGTTATAGGTACATTGAGCAACTGACTGAAATGCCTCAAAATGTTCTTTACGATGCCATTGGGAT  
[0829] ATATCAACGGTGGTATATCCAGTGATTTTTTTCTCCATTTTAGCTTCCTTAGCTCCTGAAAATCTCGATAACTC  
[0830] AAAAAATACGCCCGGTAGTGATCTTATTTTCATTATGGTGAAAGTTGGAACCTCTTACGTGCCGATCAACGTCTC  
[0831] ATTTTCGCCAAAAGTTGGCCAGGGCTTCCCGGTATCAACAGGGACACCAGGATTTATTTATTTCTGCGAAGTGA  
[0832] TCTTCCGTACAGGTATTTATTCGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGA  
[0833] TAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGAAAATGTGCGCGGAACCCCTATTTGTTTATTTTTC  
[0834] TAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCTGATAAATGCTTCAATAATATTGAAAAAGGAA  
[0835] GAGTATGAGTATTCAACATTTCCGTGTGCGCCTTATTCCTTTTGTGCGCATTTTGCTTCTGTTTTGTCTC  
[0836] ACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGAT  
[0837] CTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCT  
[0838] GCTATGTGGCGCGGTATTTATCCCGTGTGACGGATCTAAGCTTGCTTCTTTGCTGACGAGTGGCGGACGGGTGAG  
[0839] TCTACGAGGGCGTGCAAGCAGTTTCGCGAGCAACCGCGAAGAAACGTCGCTTGATCGATACCTTCTTTGGC  
[0840] TTGAGTCAACGCTATGTTTGGCACGGCGCGCTGGCAAGGACTGGATTGCTGGCACTGAACCAATCCCCAGC  
[0841] CCAGCGCTCGCTGAGGGTGTCCGGATGTTGGCGCTAGCACCGTTGCATAAGCTGGGCGATCGCTCGTCTACG  
[0842] GCAAAGTACGAGAAGCCACGGGTGGCCGAATTCGGCAGGTGATCAGTGGCGGTGGCTCACTGGCACTGCACCTC  
[0843] GATACCTTCTTCGAAATGTTGGTGTGATTGCTGGTGGGTTATGGCTTGACAGAAACCTCACCAGTGCTGAC  
[0844] GGGGCGACGGCTTGGCACAACCTACGGGGTTCGGCCGGTCAGCCGATTCCAGGTACGGCGATTCCGATCGTCG  
[0845] ATCCTGAAACGAAGGAAAACCGACCCAGTGGCGATCGCGCTTGGTGCTGGCGAAAGGGCCGCAAATCATGCAG

[0846] GGCTACTTCAATAAACCCGAGGCGACCGCGAAAGCGATCGATGCCGAAGGTTGGTTTGACACCGGCGACTTAGG  
[0847] CTACATCGTCGGTGAAGGCAACTTGGTG  
[0848] SEQ ID NO :23  
[0849] CTCGAGCCCCCGTGCTATGACTAGC  
[0850] SEQ ID NO :24  
[0851] CTCGAGCCCGGAACGTTTTTTGTACCCC  
[0852] SEQ ID NO :25  
[0853] CAATTGGTCACACGGGATAATACCGCGCC  
[0854] SEQ ID NO :26  
[0855] CAATTGGTCGATCATATCGTCAATTATTACCTCCAC  
[0856] SEQ ID NO :27  
[0857] CCCCCGTGCTATGACTAGCGGCGATCGCCATACCGGCCACGACCATTTCGATTGGATCCCCAACGGCGGCCACA  
[0858] ACTTCCATGGCATTGAGATGCGGGGAATGATGTTCTAGACTCTGACGCACCAAAGCCAATTTTTGTTGATGGTT  
[0859] GCAATGGGGATGACTACTGTTCACTTTGCCCCAGCGTCAATGCCTAGACCTAGCAGTACCCCCAGGGCTGTGG  
[0860] TAGTGCCCCCACCACGCATTGCTTAGCACTAAGTAACTTTCCGGCATGTTCTGGGCTAACTGTGCGCCCCAC  
[0861] TGCAAAACCCTGCTGAAAAAGATGCTCCACCAGGGCCAACGGTAACGCTTGCCCTGTGGAAGACAGCGGGCGGG  
[0862] TTGTCCGTCTAGATTGATGACTGGCACCCTGGGGGAATGGGTAAACCAGAGTTAAATAAAATAAACCGGAGTAT  
[0863] GGAGGGCATCCACCAACGCTTTGGTGATGAACACTGGGGAAACCCAGAAATGAGGGGAGGTAAAGGATAGGTT  
[0864] GCCCCTGCCGTAGTTCCCTTGATTAAAAATTCGCATCGGCGATCGCCGTCAATTTTCGATCAGCGGGGGTTTTT  
[0865] ACCCGCCGAGAAATGCCCCGAATTAACCAGTTTCCGTAAAGCCCAACACACAGACAAACACCGGTGGACAGT  
[0866] GGCCATGGCGCTCAATCCAGGATAAAGCTTGGTCAGACTGGGTATAAACTGTCAACATATTTCTGCAAGAGTGG  
[0867] GCCCAATTGGGAAAATCAACCTCAAATCCATTGGAATAGCCTTTTTTCAACCGTAAAAATCCAACTTTCTCTCT  
[0868] TCCCTTCTTCCTTCCATCTGATTATGGTTACGCCAATTAACCTACCATTCCATCCATTGCCTGGCGGATATCTGG  
[0869] GCTATCACCGGAGAAAAATTTTGCCGATATTTGTGGCCCTCAACGATCGCCATAGTCATCCCCCGTAACTTTAAC  
[0870] CTATGCCCAATTGGTCACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAGAAC  
[0871] GTTCTTCGGGGCGAAAACCTCTCAAGGATCTTACCGCTGTGAGATCCAGTTTCGATGTAACCCACTCGTGACCCC  
[0872] AACTGATCTTCAGCATCTTTTACTTTTACCAGCGTTTCTGGGTGAGCAAAAAACAGGAAGGCAAAATGCCGCAAA  
[0873] AAAGGGAATAAGGGCGACACGGAAATGTGAATACTCATACTCTTCCCTTTTCAATATTATGAAGCATTTATC  
[0874] AGGGTTATGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACA  
[0875] TTTCCCGGAAAAGTGCCACCTGACGTCTAAGAAACCATTTATCATGACATTAACTATAAAAAATAGGCGTAT  
[0876] CACGAGGCCCTTTTCGTCTTCGAATAAATACCTGTGACGGAAGATCACTTCGCAGAATAAATAAATCCTGGTGTC  
[0877] CCTGTTGATACCGGGAAGCCCTGGGCCAACTTTTGGCGAAAATGAGACGTTGATCGGCACGTAAGAGGTTCCAA  
[0878] CTTTCACCATAATGAAATAAGATCACTACCGGGCGTATTTTTTGTAGTTATCGAGATTTTCAGGAGCTAAGGAAG  
[0879] CTAAATGGAGAAAAAATCACTGGATATACCACCGTTGATATATCCCAATGGCATCGTAAAGAACATTTTGAG  
[0880] GCATTTCACTGAGTTGCTCAATGTACCTATAACCAGACCGTTCAGCTGGATATTACGGCCTTTTTAAAGACCGT  
[0881] AAAGAAAAATAAGCACAAAGTTTTATCCGGCCTTTATTCACATTCTTGCCCGCCTGATGAATGCTCATCCGGAAT  
[0882] TCCGTATGGCAATGAAAGACGGTGAGCTGGTGATATGGGATAGTGTTACCCCTTGTTACACCGTTTTCCATGAG  
[0883] CAACTGAAACGTTTTTCATCGCTCTGGAGTGAATACCACGACGATTTCCGGCAGTTTCTACACATATATTCGCA  
[0884] AGATGTGGCGTGTTACGGTGAAAACCTGGCCTATTTCCCTAAAGGGTTTATTGAGAATATGTTTTTCGTCTCAG

[0885] CCAATCCCTGGGTGAGTTTCACCAAGTTTGTATTAAACGTGGCCAATATGGACAACCTCTTCGCCCCCGTTTTC  
[0886] ACCATGGGCAAAATATTATACGCAAGGCGACAAGGTGCTGATGCGCTGGCGATTTCAGGTTTCATCATGCCGTTTG  
[0887] TGATGGCTTCCATGTCCGCAGAATGCTTAATGAATTACAACAGTACTGCGATGAGTGGCAGGGCGGGCGTAAT  
[0888] TTTTTTAAGGCAGTTATTGGTGCCCTTAAACGCCTGGTGCTACGCCTGAATAAGTGATAATAAGCGGATGAATG  
[0889] GCAGAAATTCGAAAGCAAATTCGACCCGGTCGTCCGTTTCAGGGCAGGGTCGTTAAATAGCCGCTTATGTCTATT  
[0890] GCTGGTTTACCGTTTATTGACTACCGGAAGCAGTGTGACCGTGTGCTTCTCAAATGCCTGAGGCCAGTTTGCT  
[0891] CAGGCTCTCCCCGTGGAGGTAATAATTGACGATATGATCGACCAATTGCGGGAAGAAATTACAGCTTTTGCCGC  
[0892] TGGCCTACAGAGTTTAGGAGTTACCCCCCATCAACACCTGGCCATTTTCGCCGACAACAGCCCCCGGTGGTTTA  
[0893] TCGCCGATCAAGGCAGTATGTTGGCTGGAGCCGTCAACGCCGTCCGTTCTGCCCAAGCAGAGCGCCAGGAATTA  
[0894] CTCTACATCCTAGAAGACAGCAACAGCCGTACTTTAATCGCAGAAAATCGGCAAAACCTAAGCAAATTGGCCCT  
[0895] AGATGGCGAAACCATTGACCTGAAACTAATCATCCTCCTCACCAGTGAAGAAGTGGCAGAGGACAGCGCCATTTC  
[0896] CCCAATATAACTTTGCCAGGTCATGGCCCTAGGGGCGCGCAAAATCCCCACTCCCGTTCCCCGCCAGGAAGAA  
[0897] GATTTAGCCACCCTGATCTACACCTCCGGCACCACAGGACAACCCAAAGGGGTGATGCTCAGCCACGGTAATTT  
[0898] ATTGACCAAGTACGGGAATTGGATTCCGGTGATTATCCCCGCCCGCGATCAGGTGTTGAGCATTTTGCCCT  
[0899] GTTGGCACTCCCTAGAAAGAAGCGCCGAATATTTCTTCTTTCCCGGGGCTGCACGATGAACTACACCAGCATT  
[0900] CGCCATTTCAAGGGGGATGTGAAGGACATTAAACCCCATCACATTGTCCGTGTGCCCGGCTGTGGGAATCCCT  
[0901] CTACGAAGGGGTACAAAAACGTTCCGGCTAAGGGCGAATTCTGCAGATATCCATCACACTGGCGGCCGCTCG  
[0902] AGCATGCATCTAGAGGGCCCAATTGCGCCTATAGTGAGTGTATTACAATTCACTGGCCGTCGTTTTACAACGT  
[0903] CGTGACTGGGAAAACCTGGCGTTACCCAACTTAATCGCCTTGACGACATCCCCCTTTGCCAGCTGGCGTAA  
[0904] TAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAG  
[0905] CGGCGCATTAAGCGCGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCG  
[0906] CTCCTTTCCGCTTTCTTCCCTTCTCTTCTCGCCACGTTGCGCGGCTTTCCCGCTCAAGCTCTAAATCGGGGGCTC  
[0907] CCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACCTTGATTAGGGTGATGGTTACCGTAG  
[0908] TGGGCCATCGCCCTGATAGACGGTTTTTCCGCCCTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGT  
[0909] TCCAAACTGGAACAACACTCAACCCATCTCGGTCTATTCTTTGATTTATAAGGGATTTTGCGATTTCGGCC  
[0910] TATTGGTTAAAAATGAGCTGATTTAACAAAAATTTAACCGAATTTTAACAAAATTCAGGGCGCAAGGGCTGC  
[0911] TAAAGGAAGCGGAACACGTAGAAAGCCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGC  
[0912] TATCTGGACAAGGGAAAAACGAAGCGCAAAGAGAAAGCAGGTAGCTTGACGTGGGCTTACATGGCGATAGCTAG  
[0913] ACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCC  
[0914] TGCAAAGTAAACTGGATGGCTTTCTTGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGATCTGATCAAGAGAC  
[0915] AGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGC  
[0916] TATTCCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGATGCCGCGGTGTTCCGGCTGTCAGCGCAGGGG  
[0917] CGCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATC  
[0918] GTGGCTGGCCACGACGGGCGTTCCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGC  
[0919] TATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCTATCCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCT  
[0920] GATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAAACATCGCATCGA  
[0921] GCGAGCACGTACTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGAAGAAGAGCATCAGGGGCTCGCGC  
[0922] CAGCCGAACTGTTCCGCAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGTGTGACCCATGGCGATGCC  
[0923] TGCTTGCCGAATATCATGGTGAAAAATGGCCGCTTTTCTGGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGA

[0924] CCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC  
[0925] TCGTGCTTTACGGTATCGCCGCTCCCGATTTCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGA  
[0926] ATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCTTTTTTGGCGCATTTTGCCCTT  
[0927] CCTGTTTTTGCTCAGGAGAAAGCTGGTGAAGTAAAGATGCTGAAGATCAGTTGGGTGACGAGTGGGTTA  
[0928] CATCGAACTGGATCTCAACAGCGTAAGATCCTTGAGAGTTTTGCCCCGAAGAAGCTTTTCCAATGATGAGCA  
[0929] CTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCGGGCAAGAGCAACTCGGTGCGGCATA  
[0930] CACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGGCATGACAGTAAG  
[0931] AGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGAC  
[0932] CGAAGGAGCTAACCGCTTTTTTGACAACATGGGGGATCATGTAAGTCCCTTGATCGTTGGGAACCGGAGCTG  
[0933] AATGAAGCCATACCAAACGACGAGCGTGACACCAGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACCTATT  
[0934] AACTGGCGAACTACTTACTCTAGCTTCCCGCAACAATTAATAGACTGGATGGAGGCGGATAAAGTTGCAGGAC  
[0935] CACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTATTGCTGATAAATCTGGAGCGGTGAGCGTGGGTCTCGC  
[0936] GGTATCATTGACGACTGGGGCCAGATGGTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGAGTCAGGC  
[0937] AACTATGGATGAACGAAATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTTGGTAACTGTCAGACC  
[0938] AAGTTTACTCATATATACTTTAGATTGATTTAAAACCTCATTTTTTAATTTAAAAGGATCTAGGTGAAGATCCTT  
[0939] TTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGAT  
[0940] CAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAAACAAAAAACACCGCTACCAG  
[0941] CGGTGGTTTTGTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATA  
[0942] CCAAATACTGTTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCTACATACCT  
[0943] CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTTTACCGGGTTGGACTCAAGAC  
[0944] GATAGTTACCGGATAAGGCGCAGCGTCCGGCTGAACGGGGGGTTCGTGCACACAGCCAGCTTGGAGCGAACG  
[0945] ACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGA  
[0946] CAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGAAAAAGCCTGGTATC  
[0947] TTTATAGTCCTGTCCGGTTTTGCCACCTCTGACTTGAGCGTCGATTTTTTGATGCTCGTCAGGGGGCGGAGC  
[0948] CTATGGAAAAACGCCAGCAACGCGGCCCTTTTTACGGTTCCCTGGCCTTTTGCTGGCCTTTTGCTCACATGTTCTT  
[0949] TCCTGCGTTATCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGACGCC  
[0950] GAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCCAATACGAAACCGCCTCTCCCGCG  
[0951] CGTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAAT  
[0952] TAATGTGAGTTAGCTCACTCATTAGGCACCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA  
[0953] TTGTGAGCGGATAACAATTTACACAGGAAACAGCTATGACCATGATTACGCCAAGCTTGGTACCGAGCTCGGA  
[0954] TCCACTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTCTCGAG  
[0955] SEQ ID NO :28  
[0956] GATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATGC  
[0957] GGTGA  
[0958] SEQ ID NO :29  
[0959] GGATCCACTAGTCCTGAGGTGTTGACAAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACA  
[0960] ATTTACACAGGAAACAGACCATGGCCGTCGACTGCAACCAGCTCAAGAAGTCGAACTAAGAAAAAGCCTGC  
[0961] AATCAAACAGCGGCGGTGGTGGTTACCGGCATGGGTGTGGTGAAGTCCCTCGGGCATGAACCGGATGTGTTTT  
[0962] ACAACAATCTCCTGGATGGCGTGAGCGGCATTAGTGAGATCGAGAATTTGACTCGACGAGTTTCCCACTCGC

[0963] ATTGCCGGCGAAATCAAGAGTTTCAGCACCGACGGCTGGGTGCGGCCAAATTGAGCAAACGGATGGATAAAATT  
[0964] GATGCTGTATCTGCTCACCGCAGGCAAGAAAGCGCTGGCCGATGCGGGCATCACGGATGATGTGATGAAAGAGC  
[0965] TGGATAAACGCAAATGTGGAGTTCTGATTGGCAGTGGCATGGGCGGCATGAAGCTGTTCTACGATGCGCTCGAA  
[0966] GCCCTGAAGATTTCTGATCGAAAGATGAACCCATTCTGTGTGCCTTTTGGACCAAGAAATATGGGTAGCGCCAT  
[0967] GCTGGCTATGGATTTGGGGTGGATGGGGCCGAATTATAGTATTTCCACCGCGTGCGCAACCTCGAACTTCTGCA  
[0968] TCTTGAACGCGGCTAACACATTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCATT  
[0969] ATCCCTATCGGTTTGGGCGGCTTTGTTGCTTGCCGCGCTTGAGCCAACGCAATAACGACCCAACCAAGGCATC  
[0970] GCGCCCGTGGGACAGCAATCGCGATGGCTTCGTTCATGGGCGAGGGAGCCGGGTGCTGCTGTTGGAGGAGCTGG  
[0971] AACACGCGAAAAAGCGAGGCGGACAATCTATGCTGAGTTCTTGGGAGGGTCTTTACATGCGATGCCTACCAC  
[0972] ATGACGGAGCCTCACCCAGAGGGCGCAGGCGTGATCTTGTGTATCGAGAAGGCAATGGCTCAGGCAGGAGTCTC  
[0973] TCGCGAGGATGTTAACTACATTAATGCTCACGCAACGTCCACGCGGCTGGTGACATCAAGGAATACCAAGCTC  
[0974] TCGCCCATTTGTTTCGGCCAGAAGCTCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTCATCTGCTGGGT  
[0975] GCTGCCGTGGCGTCGAAGCTGTGACAGTCATTCAAGCCATCCGACCGGCTGGATTACCCCTAATCTGAACCT  
[0976] GGAAGACCCGACAAGGCCGTTGACGCAAAATTCCTCGTCGGACCGGAGAAAGAACGTCCTCAACGTTAAAGTCG  
[0977] GATTGAGCAATAGTTTCGGTTTGGTGGCCATAACTCTAGTATCCTGTTTGCACCTATAATTGATAATAGATC  
[0978] TGATCCGCTGTTGACCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATG  
[0979] CGGTGAGAGCTCACGTGTCTAGA  
[0980] SEQ ID NO :30  
[0981] GGATCCACTAGTCTGAGGTGTTGACAAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACA  
[0982] ATTTACACAGGAAACAGACCATGGCAAGCCGTGTTGTTGGTAAAGGTTGTAAACTCGTTGGATGTGGTAGTGC  
[0983] CGTCCCGAAGTTGGAGGTGAGTAACGACGACCTCAGTAAGATCGTGGATACTTCCGATGAATGGATTTCTGTTC  
[0984] GGACGGGAATCCGCAACCGGCGGGTGATTACTGGTAAGGATAAGATGACGGGGCTGGCGGTGAGGCAGCCAG  
[0985] AAAGCCCTGGAAATGGCTGAAGTCGATGCTGACGATGTGGACTTGCTCCTGTTGTGCACCTCCACCCAGATGA  
[0986] TCTCTTTGGAAGTGCGCCGCAATCCAGGCGGCACTCGGCTGCAAAGGAAACCTCTGGCATTTGATATTACAG  
[0987] CCGCTTGTAGCGGCTTCGTTCTGGGTCTGGTGAGTGCTTCTGCTATATCCGCGGCGGGGTTCAAGAACGTC  
[0988] CTGGTTATCGGCGGGACGCACTGAGCCGCTACGTCGATTGGAAGTACCGCGGCACATGCATTCTCTTTGGTGA  
[0989] CGCCGTGGCGCTGTGTTGGTCCAGGCGTGTGAGAGCGAGGACGACGGCGTCTTCGGGTTTGATCTGCATAGCG  
[0990] ATGGAGAGGGTTATCGCCACCTGCATACTGGGATCAAGGCGAACGAGGAGTTCGGGACGAACGGTTCCGTTGTG  
[0991] GATTTTCCGCCAAGCGCAGCAGCTACTCTTCCATCCAAATGAATGGGAAAGAAGTGTTCGGTTTCGCCGCGC  
[0992] CGTCGTGCCCCAGTCTATTGAGATCGCACTCGAGAACCGGGGCTCACAGCTTCTAGCATTGATTGGCTGCTGC  
[0993] TCCACCAAGCAAACCAACGAATCTTGGATGCGCTCGCAACGCGTCTGGAAATTCGCGCAGACCGCGTGATTAGT  
[0994] AACTTGGCTAATTACGGCAATACTTCTGCGCCAGCATTCCGTTGGCACTGGATGAAGCCGTGCGCAGCGGTAA  
[0995] GGTCAAACCCGTCAGACTATCGCAACTTCGGGGTTTGGAGCAGGCTTGACATGGGGCAGCGCATCATTCGCT  
[0996] GGAATTAATGATAGATCTGATCCGCTGTTGACCAACAGCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCT  
[0997] TTTTTTCATGCGCGGATGCGGTGAGAGCTCACGTGTCTAGA  
[0998] SEQ ID NO :31  
[0999] TGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGGAAACAGC  
[1000] GCCGCTGAGAAAAAGCGAAGCGGCACTGCTCTTAACAATTTATCAGACAATCTGTGTGGGCACTCGACCGGAA  
[1001] TTATCGATTAACTTTATTATTAATAAATAAAGAGGTATATTAATGTATCGATTAAATAAGGAGGAATAAACC

[1002] ATGGCCGTCGCACTGCAACCAGCTCAAGAACTCGCAACTAAGAAAAAGCCTGCAATCAAACAGCGGCGGTGGT  
[1003] GGTACCGGCATGGGTGTGGTGAAGCTCCCTCGGGCATGAACCGGATGTGTTTACAACAATCTCCTGGATGGCG  
[1004] TGAGCGGCATTAGTGAGATCGAGAATTTTGAAGCTCGACGAGTTTCCCACTCGCATTGCCGGCGAAATCAAGAGT  
[1005] TTCAGCACCGACGGCTGGGTGCGGCCAAATTGAGCAACGGATGGATAAATTGATGCTGTATCTGCTCACCGC  
[1006] AGGCAAGAAAGCGCTGGCCGATGCGGGCATCAAGGATGATGTGATGAAAGAGCTGGATAAACGCAATGTGGAG  
[1007] TTCTGATTGGCAGTGGCATGGGCGGCATGAAGCTGTTCTACGATGCGCTCGAAGCCCTGAAGATTTCTGTATCGA  
[1008] AAGATGAACCCATTCTGTGTGCCTTTTGCAGCCACGAATATGGGTAGCGCCATGCTGGCTATGGATTGGGGTG  
[1009] GATGGGGCCGAATTATAGTATTTCCACCGCGTGGCGAACCTCGAACTTCTGCATCTTGAACGCGGCTAACCA  
[1010] TTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCATTATCCCTATCGGTTTGGGCGGC  
[1011] TTTGTTGCTTGGCGCGCTTGAGCCAACGCAATAACGACCCAACCAAGGCATCGCGCCCGTGGGACAGCAATCG  
[1012] CGATGGCTTCGTTCATGGGCGAGGGAGCCGGGTGCTGCTGTTGGAGGAGCTGGAACACGCGAAAAAGCGAGGCG  
[1013] CGACAATCTATGCTGAGTTCTTGGGAGGGTCCCTTACATGCGATGCGCTACCACATGACGGAGCCTCACCCAGAG  
[1014] GCGCGAGGCGTGATCTGTGTATCGAGAAGGCAATGGCTCAGGCAGGAGTCTCTCGCGAGGATGTTAACTACAT  
[1015] TAATGCTCACGCAACGTCCACGCGGCTGGTGACATCAAGGAATACCAAGCTCTCGCCCATTTGTTTGGGCCAGA  
[1016] ACTCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTCATCTGCTGGGTGCTGCCGTTGGCGTGAAGCT  
[1017] GTGACAGTCATTCAAGCCATCCGACCGGCTGGATTACCCCTAATCTGAACCTGGAAGACCGGACAAGGCGT  
[1018] TGACGCAAAATTCCTCGTGGACCGGAGAAAGACGTCTCAACGTTAAAGTCGGATTGAGCAATAGTTTCGGTT  
[1019] TTGGTGGCCATAACTCTAGTATCCTGTTTGCACCCCTATAATTGATAATAGATCTGATCCGCTGTTGACCCAACA  
[1020] GCATGAGTCGTTATCCAAGGGGAGCTTCGGCTCCCTTTTTTCATGCGCGGATGCGGTGA  
[1021] SEQ ID NO :32  
[1022] CCTGAGGTGTTGACAATTAATCATCCGGCTCGTATAATGTGTGGAATTGTGAGCGGATAACAATTTACACAGG  
[1023] AAACAGCGCCGCTGAGAAAAAGCGAAGCGGCACTGCTCTTAAACAATTTATCAGACAATCTGTGTGGGCACTCG  
[1024] ACCGGAATTATCGATTAACTTTATTATTAATAAAGAGGTATATATTAATGTATCGATTAAATAAGGAGGA  
[1025] ATAAACCATGGCCGTCGCACTGCAACCAGCTCAAGAACTCGCAACTAAGAAAAAGCCTGCAATCAAACAGCGGC  
[1026] GCGTGGTGGTTACCGGCATGGGTGTGGTGAAGCTCCCTCGGGCATGAACCGGATGTGTTTACAACAATCTCCTG  
[1027] GATGGCGTGAGCGGCATTAGTGAGATCGAGAATTTTGAAGCTCGACGAGTTTCCCACTCGCATTGCCGGCGAAAT  
[1028] CAAGAGTTTACGACCGGACGGCTGGGTGCGGCCAAATTGAGCAACGGATGGATAAATTGATGCTGTATCTGC  
[1029] TCACCGCAGGCAAGAAAGCGCTGGCCGATGCGGGCATCACGGATGATGTGATGAAAGAGCTGGATAAACGCAAA  
[1030] TGTGGAGTTCTGATTGGCAGTGGCATGGGCGGCATGAAGCTGTTCTACGATGCGCTCGAAGCCCTGAAGATTTC  
[1031] GTATCGAAAGATGAACCCATTCTGTGTGCCTTTTGCAGCCACGAATATGGGTAGCGCCATGCTGGCTATGGATT  
[1032] TGGGTGGATGGGGCCGAATTATAGTATTTCCACCGCGTGGCGAACCTCGAACTTCTGCATCTTGAACGCGGCT  
[1033] AACCACATTATCCGTGGTGAAGCAGACATGATGCTCTGCGGCGGCTCCGATGCGGTCATTATCCCTATCGGTTT  
[1034] GGGCGGCTTTGTTGCTTGGCGCGCTTGAGCCAACGCAATAACGACCCAACCAAGGCATCGCGCCCGTGGGACA  
[1035] GCAATCGCGATGGCTTCGTCATGGGCGAGGGAGCCGGGTGCTGCTGTTGGAGGAGCTGGAACACGCGAAAAAG  
[1036] CGAGGCGGCAATCTATGCTGAGTTCTTGGGAGGGTCCCTTACATGCGATGCGCTACCACATGACGGAGCCTCA  
[1037] CCCAGAGGGCGCAGGCGTGATCTGTGTATCGAGAAGGCAATGGCTCAGGCAGGAGTCTCTCGCGAGGATGTTA  
[1038] ACTACATTAAATGCTCACGCAACGTCCACGCGGCTGGTGACATCAAGGAATACCAAGCTCTCGCCCATTTGTTTC  
[1039] GGCCAGAACTCGGAGCTGCGGGTCAATAGTACAAAGTCCATGATCGGTCATCTGCTGGGTGCTGCCGTTGGCGT  
[1040] CGAAGCTGTGACAGTCATTCAAGCCATCCGACCGGCTGGATTACCCCTAATCTGAACCTGGAAGACCGGACA

[1041] AGGCCGTTGACGCAAAATTCCTCGTCCGACCGGAGAAAGAACGTCTCAACGTTAAAGTCGGATTGAGCAATAGT  
[1042] TTCGGTTTTGGTGGCCATAACTCTAGTATCCTGTTTGACCCCTATAATTGATAATAGATCCTGTGCTTAACTGC  
[1043] TTTGTTGGTACTACCTGACTTCACCCCTCTTTTAAGATGGCAAGCCGTGTTGTTGGTAAAGGTTGTAAACTCGTT  
[1044] GGATGTGGTAGTGCCGTCCCGAAGTTGGAGGTGAGTAACGACGACCTCAGTAAGATCGTGGATACTTCCGATGA  
[1045] ATGGATTTCTGTTCCGACGGGAATCCGCAACCGGCGGGTGATTACTGGTAAGGATAAGATGACGGGGCTGGCGG  
[1046] TCGAGGCAGCCCAGAAAGCCCTGGAAATGGCTGAAGTCGATGCTGACGATGTGGACTTGCTCCTGTTGTGCACC  
[1047] TCCACCCAGATGATCTCTTTGGAAGTGCGCCGCAATCCAGGCGGCACTCGGCTGCAAAGGAAACCCCTCTGGC  
[1048] ATTTGATATTACAGCCGCTTGTAGCGGCTTCGTTCTGGGTCTGGTGAGTGCTTCCTGCTATATCCGCGGCGCG  
[1049] GGTTCAAGAACGTCCTGGTTATCGGCGCGGACGCACTGAGCCGCTACGTGATTGGACTGACCGCGGCACATGC  
[1050] ATTCTCTTTGGTGACGCGCTGGCGCTGTGTTGGTCCAGGCGTGTGAGAGCGAGGACGACGCGCTCTTCGGGTT  
[1051] TGATCTGCATAGCGATGGAGAGGGTTATCGCCACCTGCATACTGGGATCAAGGCGAACGAGGAGTTCCGGACGA  
[1052] ACGGTTCCGTTGTGGATTTTCCGCCAAGCGCAGCAGCTACTCTTCCATCCAAATGAATGGGAAAGAAGTGTTTC  
[1053] CGTTTCGCCTGCCGCGTGTGCCCCAGTCTATTGAGATCGCACTCGAGAACGCGGGCCTCACACGTTCTAGCAT  
[1054] TGATTGGCTGCTGCTCCACCAAGCAAACCAACGAATCTTGATGCCGTGCAACGCGTCTGGAAATTCGCCGAG  
[1055] ACCGCGTGATTAGTAACTTGGCTAATTACGGCAATACTTCTGCGCCAGCATTCCGTTGGCACTGGATGAAGCC  
[1056] GTGCGCAGCGTAAGGTCAAACCCGTCAGACTATCGCAACTTCGGGGTTTGAGAGGAGCTTGACATGGGGCAG  
[1057] CGGATCATTCGCTGGAATTAATGATAGATCTGATCCGCTGTTGACCCAACAGCATGAGTCGTTATCCAAGGGG  
[1058] AGCTTCGGCTCCCTTTTTTTCATGCGCGGATGCGG  
[1059] SEQ ID NO :33  
[1060] GTACGGGATCCCTGTGCTTAACTGCTTTGTTGGTACTACCTGACTTCACCCCTCTTTTAAGATGGCAAGCCGTGT  
[1061] TGTTGGTAAAGGTTG  
[1062] SEQ ID NO :34  
[1063] CACGTGAGCTCTCACCGCATCCGCGCATG  
[1064] SEQ ID NO :35  
[1065] TCATGAAGTTCCTTGTGCTCGCGTCTCAGCACTTGCAACTGCATCTGCTTTCACAACCAGTCCTGCCTCTTTC  
[1066] ACCACTGTCAGCAGTCCTTCGGTGAACAATGTGTTCCGACAGGAGGGAAATGCTCACAGGAACAGGAGAGCTAC  
[1067] CATTGTCATGGATGGAGCTAACGGAAGTGCAGTCAGTTTGAAAAGTGGGTCAATTGAATACGAGGAGGACACAA  
[1068] GTTCGTCGCCACCGCCCGTACATTCCCTTCACCAACTCCCTGATTGGAGCAGATTGCTCACTGCCATCACAACC  
[1069] GTTTTTGTAAAAGTAAGCGTCCGGATATGCATGATCGTAAGTCGAAAAGGCCGACATGCTCGTGGATAGTTT  
[1070] CGGGTTGGAGAGTACCGTTCAGGATGGACTCGTGTTCGGTCAAAGCTTTTCGATCCGTTTATGAGATTGGAA  
[1071] CTGATCGTACGGCTTCCATTGAGACTTTGATGAACCATCTTCAGGAGACTTCCCTCAACCATTTGTAAGAGTACA  
[1072] GGAATTTTGTGGATGGATTCCGACGCACACTCGAAATGTGTAAGCGGATTTGATTTGGGTCGTATTAAAAAT  
[1073] GCAGATCAAGGTTAATAGATACCCGGCTGGGGCGATACAGTAGAAATCAATACTAGGTTACGAGACTTTGGTA  
[1074] AGATCGGCATGGGTGAGATTGGCTCATTAGCGACTGCAATACCGGTGAGATCCTCGTCAGGGCAACCAGCGCC  
[1075] TACGCCATGATGAATCAGAAGACCCGAAGACTCTCGAAGCTTCGGTACGAGGTCCACCAAGAGATTGTCCCCCT  
[1076] TTTTGTGACTCCCCCGTAATTGAAGATTCGGATCTCAAGGTCCACAAATTCAAAGTTAAAACGGGTGACAGCA  
[1077] TCCAGAAGGGACTTACTCCTGGTTGGAACGACCTCGATGTGAACCAACATGTTTCGAACGTGAAATATATCGGC  
[1078] TGGATTCTTGAGAGTATGCCAACCGAGGTACTTGAGACGAGGAATTGTGCTCGTTGGCATTTGGAGTATCGTCG  
[1079] TGAGTGTGGGCGAGACTCAGTCCTCGAAAGTGTAAACAGCAATGGACCCAAAGCAAAGTTGGTGTTCGTTACAGT

[1080] ATCAACACCTCCTCCGTCTCGAGGATGGAACAGCCATTGTGAACGGGGCCACAGAGTGGAGGCCAAAGAACGCT  
[1081] GGGCGTAACGGAGCTATCTCCACAGGAAAGACCAGCAATGGTAACTCTGTGAGTTAATGATAGGATCC  
[1082] SEQ ID NO :36  
[1083] MKFLVVAVSALATASAFITSPASFTTVSSPSVNNVFGQEGNAHRNRRATIVMDGANGSAVSLKSGSLNTQEDTS  
[1084] SSPPPRTFHLQLPDWSRLTATITTVFKSKRPDMHDRKSKRPDMLVDSFGLESTVQDGLVFRQSFSIRSIEIGT  
[1085] DRTASIEITLMNHLQETSLNHCKSTGILLDFGRTLEMCKRDLIWVVIKMQIKVNRYPAWGDTVEINTRFSRLGK  
[1086] IGMGRDWLISDCNTGEILVRATSAYAMNQNKTRRLSKLPYEVHQEIVPLFVDSPIEDSDLKVHKFKVKTGDSI  
[1087] QKGLTPGWNLDVNQHVSNNKYIGWILESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVRISQY  
[1088] QHLLRLEDGTAIVNGATEWRPKNAGANGAISTGKTSNGNSVS  
[1089] SEQ ID NO :37  
[1090] CAGGATCCGGGGAGGTGTGGTGTAGT  
[1091] SEQ ID NO :38  
[1092] TAGGATCCAGTGGTGCCCATGGTACTTTGTTAGGGGAGGATAG  
[1093] SEQ ID NO :39  
[1094] CAGGATCCTCACTCTGTGCGCTGTTG  
[1095] SEQ ID NO :40  
[1096] CATCTAGAGAGGATTGATTTCCGAGTC  
[1097] SEQ ID NO :41  
[1098] ATGGGCACCACTCTCGACGACACGGCTTACCGCTACCGCACCAGTGTGCCGGGGACGCCGAGGCCATCGAGGC  
[1099] ACTGGATGGGTCTTACCAACCGACACCGTCTTCCGCGTACCGCCACCGGGGACGGCTTACCCCTGCGGGAGG  
[1100] TGCCGGTGGACCCGCCCTGACCAAGGTGTTCCCGACGACGAGTCGGACGACGAGTCGGACGACGGGGAGGAC  
[1101] GCGGACCCGGACTCCCGGACGTTCTGTCGCGTACGGGGACGACGGCGACCTGGCGGGCTTCTGTTGCTCTCGTA  
[1102] CTCCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGTGCCCCGGAGCACCGGGGGACGGGGTCCGGC  
[1103] GCGCGCTGATGGGGCTCGCGACGGAGTTGCCCCGAGCGGGGTGCCGGGCACCTCTGGCTGGAGGTCACCAAC  
[1104] GTCAACGACCGGGGATCCACGCGTACCGCGGATGGGGTTACCCCTCTGCGGCCTGGACACCGCCCTGTACGA  
[1105] CGGCACCGCCTCGGACGGCGAGCAGGCGCTCTACATGTCCATGCCCTGCCCTAA  
[1106] SEQ ID NO :42  
[1107] CCATGGCCGCTATGCTCGCCTCTAAGCAGGGCGCCTTTCATGGGCCGAGCTCCTTTGCCCCCGCCCCAAGGGC  
[1108] GTCGCCAGCCGGGCTCCCTGCAGGTGGTGGCCGGCGCCAAACGGCAGCGGGTGAGCCTGAAGTCGGGTTCCTT  
[1109] CAACACTCAGGAGGACACCTCGTCTCGCCCCCGCCGCGACGTTCTTGACCCAGCTGCCGGACTGGTCCCGCC  
[1110] TGCTGACGGCTATTACGACCGTGTTCGTGAAGTCGAAGCGCCCGACATGACGACCGCAAGAGCAAGCGGCCT  
[1111] GATATGCTGGTGGACAGCTTTGGCCTGGAGTCCACGGTGCAGGACGGCCTCGTGTTCGGCAAAGCTTCAGCAT  
[1112] CCGCAGCTACGAGATCGGCACGGACCGCACCGGTCGATCGAGACGCTCATGAACCACCTCCAGGAGACGTGCG  
[1113] TCAACCACTGCAAGTCCACCGGTATCTGCTGGACGGCTTTGGCCGACCCCTGGAGATGTGCAAGCGGGATCTG  
[1114] ATCTGGGTGGTGATCAAGATGCAGATCAAGGTGAACCGCTATCCCGCCTGGGGTGACACCGTCGAGATTAAAC  
[1115] CCGCTTCTCGCGCTGGGCAAGATCGGCATGGGGCGGACTGGCTGATCTCGGACTGCAACACTGGCGAGATCC  
[1116] TGGTCCGGGCCACGTCGGCCTACGCCATGATGAACCAGAAGACTCGGCGGCTGAGCAAGCTGCCTTACGAGGTG  
[1117] CATCAGGAGATCGTGCCGCTCTTCGTGGACAGCCCCGTGATCGAGGACAGCGATCTGAAGGTGCACAAGTTCAA  
[1118] GGTCAAGACCGGCGACAGCATCCAGAAGGGCTGACTCCCGGCTGGAACGACCTGGACGTGAACCAGCACGCTCT

[1119] CGAACGTGAAGTACATCGGCTGGATTCTGGAGTCGATGCCCCCGAGGTGCTGGAGACGCAGGAGCTGTGCTCC  
[1120] CTGGCGCTGGAGTATCGCCGCGAGTGGCCGCGACTCCGTGCTGGAGTCCGTCACCGCGATGGACCCGTCGAA  
[1121] GGTGGGTGTCCGAGCCAGTACCAACACCTGCTGCGCCTCGAGGACGGCACCGCCATTGTGAACGGCGCGACGG  
[1122] AGTGGCGGCCGAAGAACGCGGGCGCTAACGGCGCCATCTCCACGGGCAAGACCTCCAACGGCAACTCGGTGAGC  
[1123] TAATGATAGGATCC  
[1124] SEQ ID NO :43  
[1125] MAAMLASKQGA FMGRSSFAPAPKGVASRGS LQVVAGANGSAVSLKSGSLNTQEDTSSSPPPRTFLHQLPDWSRL  
[1126] LTAITTVFVKSKRPDMHDRKSKRPDMLVDSFGLESTVQDGLVFRQSFIRSIEIGTDRTAS IETLMNHLQETSL  
[1127] NHCKSTGILLDGFGR TLEMCKRDLI WVVIKMQIKVNRYPAWGDTVEINTRFSRLGKIGMRDWLISDCNTGEIL  
[1128] VRATSAYAMNQKTRRLSKLPYEVHQEIVPLFVDSPIEDSDLKVHKFKVKTGDSIQKGLTPGWNDLDVNQHVS  
[1129] NVKYIGWILESMPTEVLETQELCSLALEYRRECGRDSVLESVTAMDPSKVGVR SQYQHLLRLEDGTAIVNGATE  
[1130] WRPKNAGANGAISTGKTSNGNSVS  
[1131] SEQ ID NO :44  
[1132] GGTGGA AATGCCTATGTGTTAACG  
[1133] SEQ ID NO :45  
[1134] CGTAGGCAGTGTGCAACCAGGAGCC  
[1135] SEQ ID NO :46  
[1136] GGTGGA AATGCCTATGTGTTAACGGATCTACAAACCAGCACCAAACCTCTATTACGAACCCACGGTTTCCACT  
[1137] CTCCCCAACTGCAAGACTTGGGGCCCATTGATGTGGTTTTAACCCCGTCATTGGCATCAATATCCTCGGATTC  
[1138] CTGCCGGTGCTCAATGGCCAGAAAACCCCTGGAGCTTTGTGCGACTGTCCATCCCCAGGCGATCGTCCCCAC  
[1139] CTCTGGAGCCGCAGAATTGAACTATAGCGGTTTACTAACTAAAGTTTTACGTTTAGACGGCGATCTCAGTCAAT  
[1140] TTCCGCAGTCCCTAATTGACGAAGGGATACAAGCTTCCCTATGGGAACCCAGGTGGGAGTGCCCCCTCAATGTG  
[1141] CCCCATCCACCGTTGGCTAGGTTGGAATGTTCAAATCACTGTGCGGTGTGATGCTTGATAAATACAGTGAGCC  
[1142] AGGGAAAAC TGCAAAAAAGTGATAAAGTAGGTTAACTTGAATCAAAATCCTTTCTCCGCAGTCATAGCCAGG  
[1143] AGTAGGAAGATTACCAGCGAAGCAAGTTGTCTTCCCCTAGCTTTGGGCGGGCAAACCCCTTGCAGTATTGCCAA  
[1144] CGTCAAAAAATCACCATAGCCGAATGACCTACACCATCAACGCTGACCAAGTCCATCAGATTGTCCATAATCTT  
[1145] CACCACGATCCCTTTGAAGTGTTGGGCTGCCATCCCCCTCGGAGCTTTATGCTTGTAACCGTTTTGTGAAAAAA  
[1146] TTTTTAAAAATAAAAAAGGGGACCTCTAGGGTCCCCAATTAAATTAGTAATATAATCTATTAAAGGTCATTCAAAA  
[1147] GGTCATCCACCGGATCAGCTTAGTAAAGCCCTCGCTAGATTTTAATGCGGATGTTGCGATTACTTCGCCAACTA  
[1148] TTGCGATAACAAGAAAAAGCCAGCCTTT CATGATATATCTCCCAATTTGTGTAGGGCTTATATGCACGCTTAA  
[1149] AAATAATAAAAGCAGACTTGACCTGATAGTTTGGCTGTGAGCAATTATGTGCTTAGTG CATCTAACGCTTGAGT  
[1150] TAAGCCGCGCGCGAAGCGGCGTCGGCTTGAACGAATTGTTAGACATTATTTGCGGACTACCTTGGTGATCTCG  
[1151] CCTTTCACGTAGTGACAAAATTCTTCCAAC TGATCTGCGCGGAGGCCAAGCGATCTTCTTCTGTCCAAGATA  
[1152] AGCCTGTCTAGCTTCAAGTATGACGGGCTGATACTGGGCGGCGAGGCGCTCCATTGCCCAGTCGGCAGCGACAT  
[1153] CCTTCGGCGCGATTTTGTCCGGTTACTGCGCTGTACCAAAATGCGGGACAACGTAAGCACTACATTTGCTCATCG  
[1154] CCAGCCCAGTCGGGCGGCGAGTTCCATAGCGTTAAGGTTTCATTTAGCGCCTCAAATAGATCCTGTT CAGGAAC  
[1155] CGGATCAAAGAGTTCC TCCGCGCTGGACCTACCAAGGCAACGCTATGTTCTCTTGCTTTTGTGAGCAAGATAG  
[1156] CCAGATCAATGTGATCGTGGCTGGCTCGAAGATACCTGCAAGAATGTCATTGCGCTGCCATTCTCCAAATTGC  
[1157] AGTTCCGCTTAGCTGGATAACGCCACGGAATGATGTGCTGTCACAACAATGGTGACTTCTACAGCGCGGAG

[1158] AATCTCGCTCTCTCCAGGGAAGCCGAAGTTTCCAAAAGGTCGTTGATCAAAGCTCGCCGCGTTGTTTCATCAA  
[1159] GCCTTACGGTCACCGTAACCAGCAAATCAATATCACTGTGTGGCTTCAGGCCGCCATCCACTGCGGAGCCGTAC  
[1160] AAATGTACGGCCAGCAACGTCGGTTCGAGATGGCGCTCGATGACGCCAACTACCTCTGATAGTTGAGTCGATAC  
[1161] TTGGGCGATCACCGCTTCCTCATGATGTTAACTTTGTTTATGGGCGACTGCCCTGCTGCGTAACATCGTTGC  
[1162] TGCTCCATAACATCAAACATCGACCCACGGCGTAACGCGCTTGCTGCTTGATGCCCGACCGAGGCATAGACTG  
[1163] TACCCCAAAAAACAGTCATAACAAGCCATGAAAACCGCCACTGCGCCGTTACCAACCGTCGCTTCGGTCAAGG  
[1164] TTCTGGACCAAGTTGCGTGAGCGCATACGCTACTTGCATTACAGCTTACGAACCGAACAGGCTTATGTCCACTGG  
[1165] GTTCGTGCCTTCATCCGTTTCCACGGTGTGCGTCACCCGGCAACCTTGGGCAGCAGCGAAGTCGAGGCATTTCT  
[1166] GTCCTGGCTGGCGAACGAGCGCAAGGTTTCGGTCTCCACGCATCGTCAGGCATTGGCGGCCCTTGCTGTTCTTCT  
[1167] ACGGCAAGGTGCTGTGCACGGATCTGCCCTGGCTTCAGGAGATCGGAAGACCTCGGCCGTCGCGGCGCTTGCCG  
[1168] GTGGTGCTGACCCCGGATGAAGTGGTTGCGATCCTCGGTTTCTGGAAGGCGAGCATCGTTTGTTGCCCAGCT  
[1169] TCTGTATGGAACGGGCATGCGGATCAGTGAGGGTTTGCAACTGCGGGTCAAGGATCTGGATTTGATCACGGCA  
[1170] CGATCATCGTGCGGGAGGGCAAGGGCTCCAAGGATCGGGCTGGCACCCAGCCTGCGCGAGCAGGGGAATTGAT  
[1171] CCGGTGGATGACCTTTTGAATGACCTTTAATAGATTATATTACTAATTAATTGGGGACCTAGAGGTCCCTTT  
[1172] TTTATTTTAAAAATTTTTCACAAAACGGTTTACAAGCATAAAGCTTCGGGGACCACGGCAAGGTCAATCAATG  
[1173] GGTCATTGCTGCCTATTTACCCACGGCTGAAGCGGTAACGGTGTGCTTCCCACGATCGCCGGAAGTGATTA  
[1174] TGACCACGGTCCACCATCCCACTTTTTTGAATGCGTGTGGAGTTGGAAGAACGAAGAAATTATCAATTAAGA  
[1175] ATTACCGAAAAATGGCCACGAAAGGGTAATTTATGACCCCTATGGTTTTAAAACTCCCAAACTGACGGATTTTGA  
[1176] CCTCCATGTGTTTGGGGAAGGCAACCACCACCGTATTTACGAAAACTCGGTGCTCACCTGATGACGGTGGATG  
[1177] GAGTTAAAGGGGTTTATTTGCTGTGTGGGCCCCCAATGCCCGCAACGTTTCCATTTGGGGGATTTCAACAAC  
[1178] TGGGACGGCAGATTGCACCAAATGCGGAAACGCAACAACATGGTGTGGGAATTATTTATCCCTGAGTTGGGGGT  
[1179] GGGCACTTCTTATAAGTATGAGATTAAAACTGGGAAGGGCACATCTACGAAAAGACTGACCCCTACGGTTTTT  
[1180] ACCAAGAAGTACGCCCCAAAACCGCTTCCATTGTGGCAGACTTGGACGGTTACCAATGGCACGACGAAGATTGG  
[1181] TTGGAAGCTAGGCGCACCAGCGATCCCTGAGCAAACCGTTTCCGTTTACGAACTCCATTAGGCTCCTGGTT  
[1182] GCACACTGCCTACG

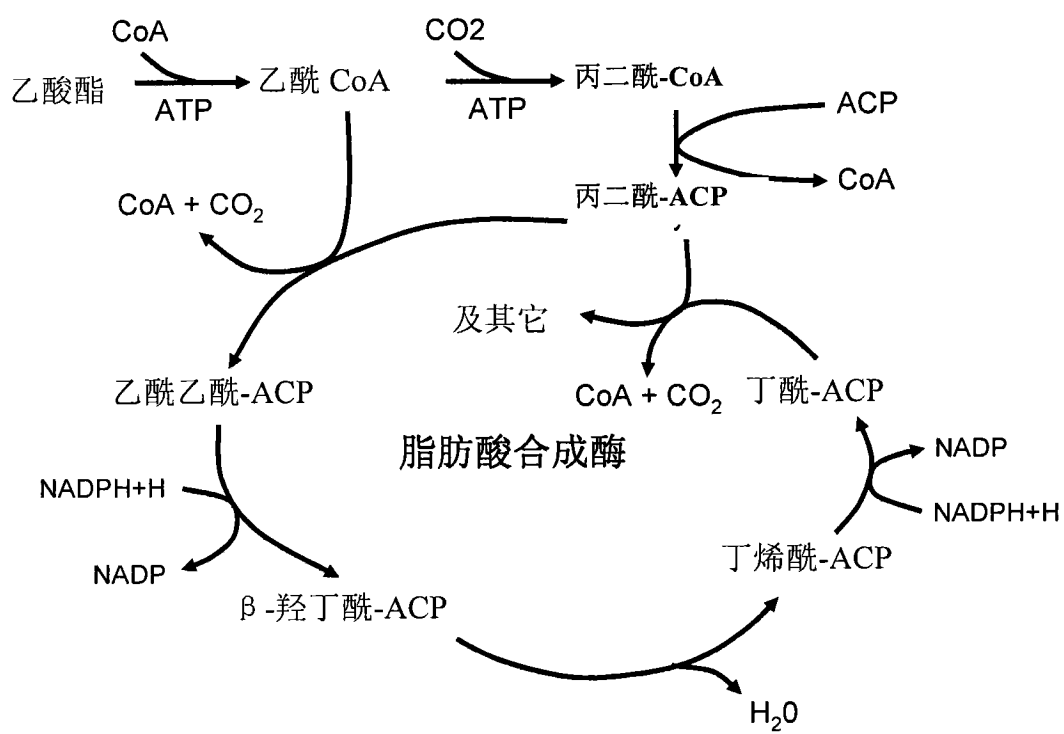


图 1

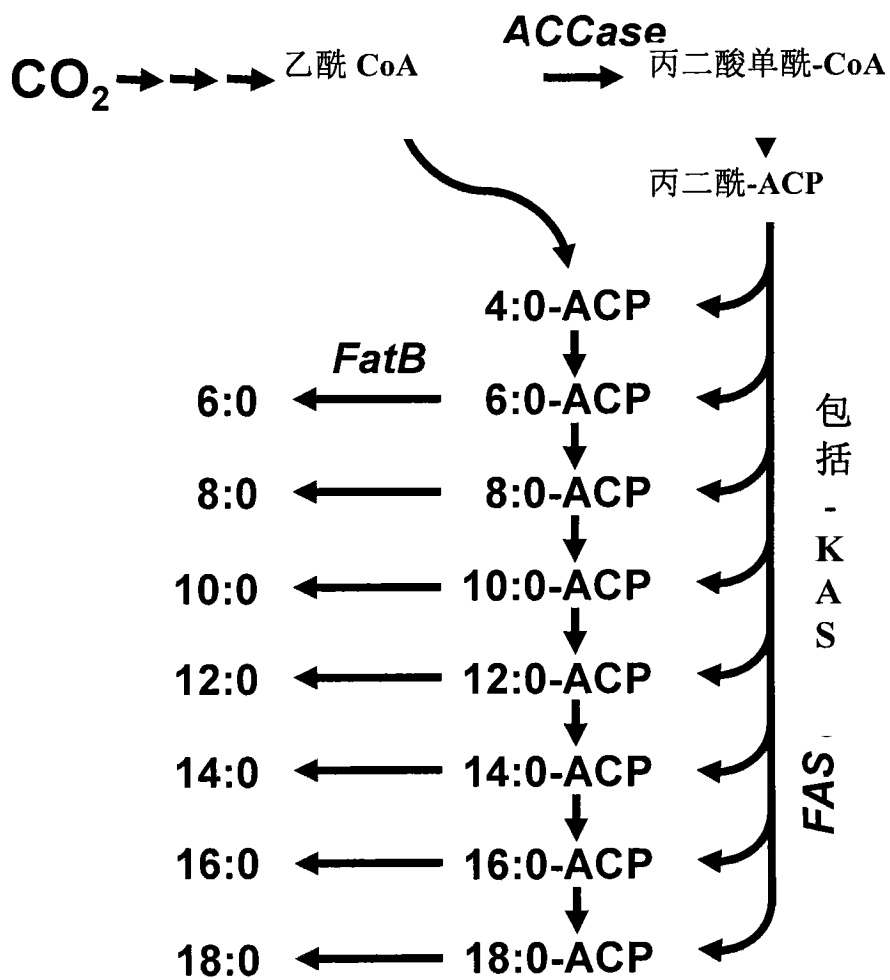


图 2

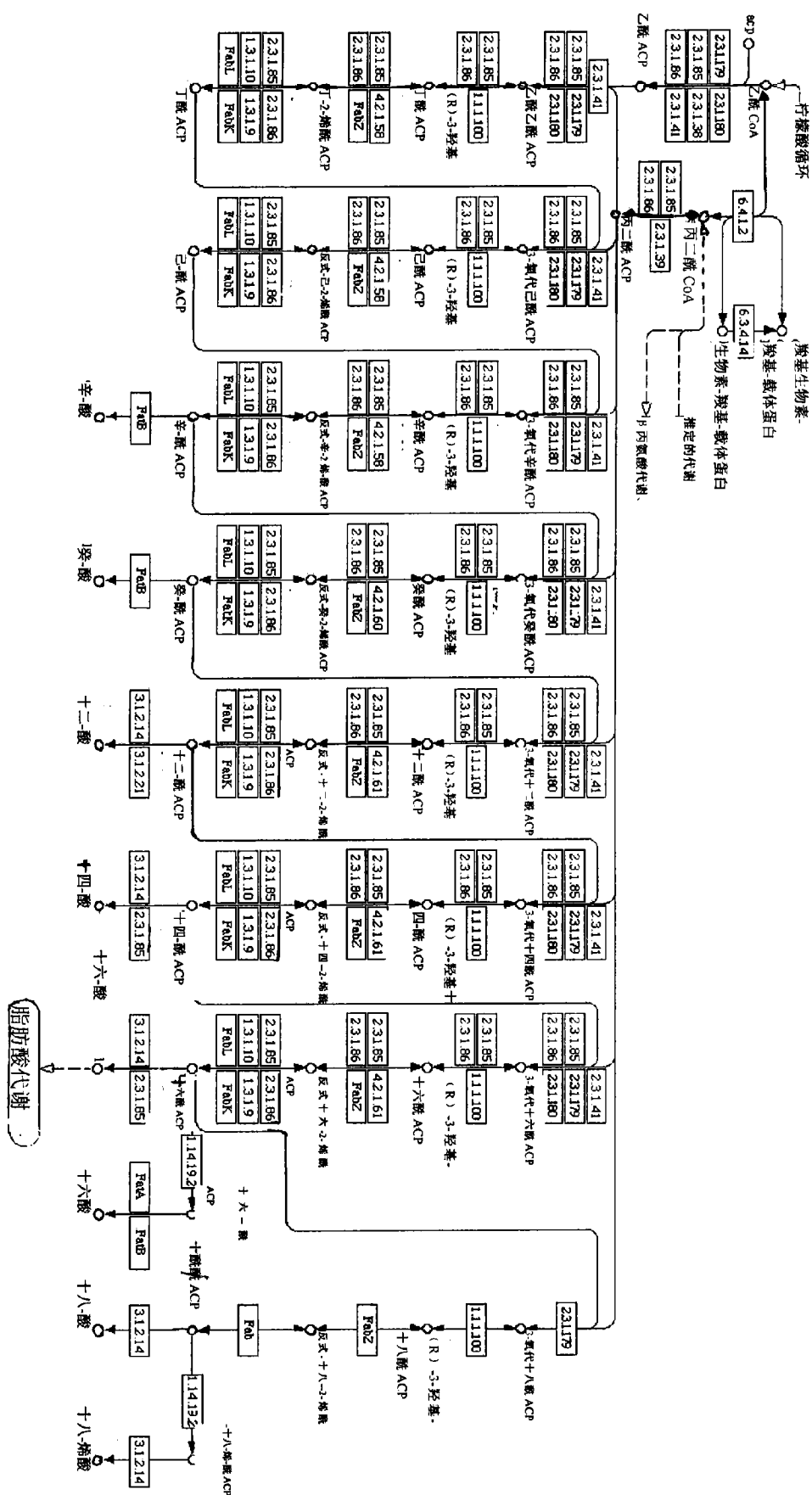


图 3

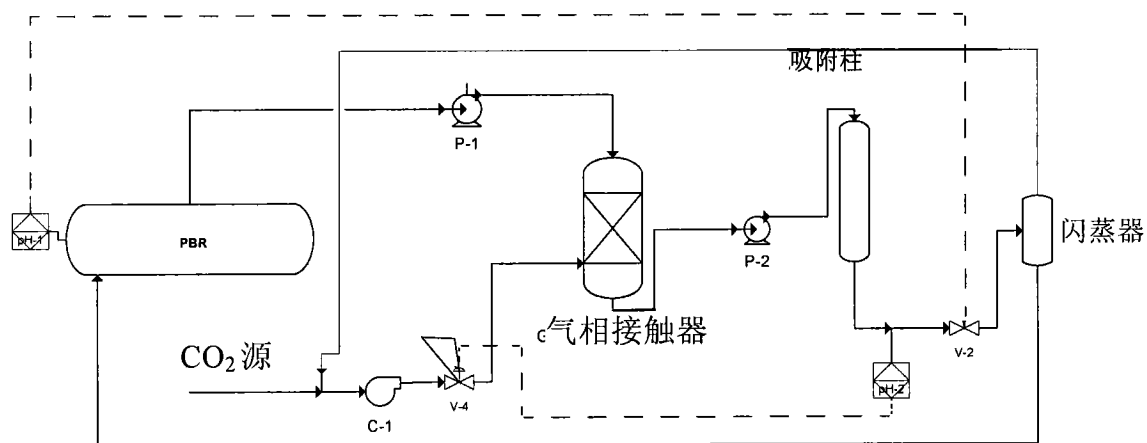
整合的 CO<sub>2</sub> 递送和产物回收示意图

图 4

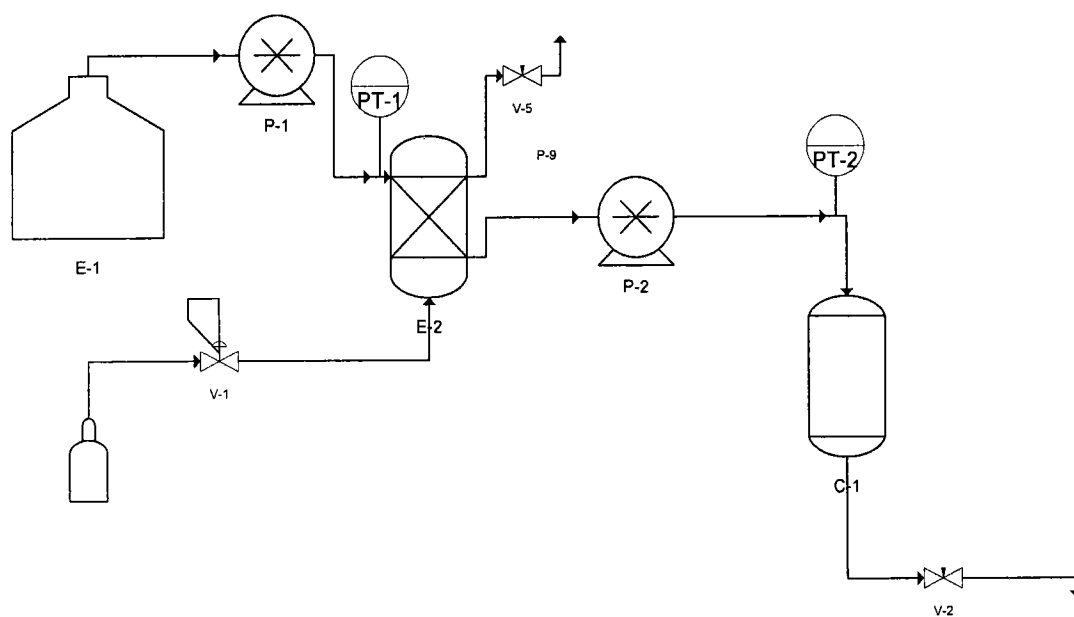
CO<sub>2</sub>-调节的疏水吸附柱装载的酸化

图 5

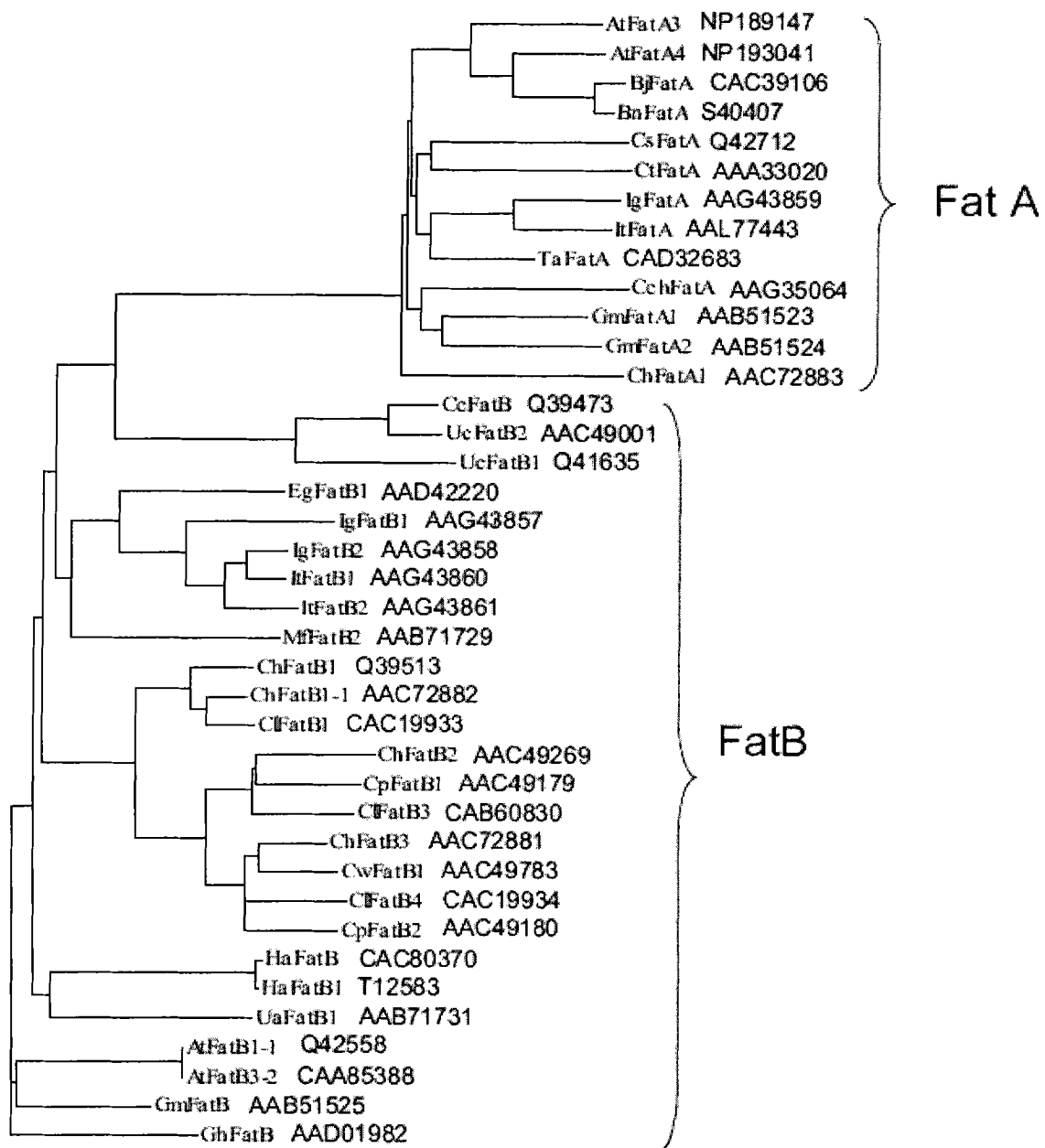


图 6