



(12)发明专利申请

(10)申请公布号 CN 111433219 A

(43)申请公布日 2020.07.17

(21)申请号 201880078452.8

(22)申请日 2018.12.04

(30)优先权数据

62/594,930 2017.12.05 US

(85)PCT国际申请进入国家阶段日

2020.06.04

(86)PCT国际申请的申请数据

PCT/US2018/063717 2018.12.04

(87)PCT国际申请的公布数据

W02019/112982 EN 2019.06.13

(71)申请人 合成基因组股份有限公司

地址 美国加利福尼亚州

(72)发明人 I·阿加维

(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 姚远

(51)Int.Cl.

C07K 14/405(2006.01)

C12N 1/12(2006.01)

C12N 1/36(2006.01)

C12N 9/16(2006.01)

C12N 9/22(2006.01)

C12N 15/10(2006.01)

C12N 15/63(2006.01)

C12Q 1/44(2006.01)

权利要求书6页 说明书39页

序列表16页 附图5页

按照条约第19条修改的权利要求书6页

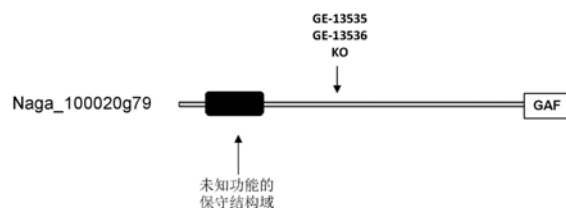
按照条约第19条修改的声明或说明1页

(54)发明名称

通过基因修饰信号传导蛋白提高藻类脂质生产力的

(57)摘要

本发明提供编码包含GAF结构域的多肽的基因表达减弱的突变微生物,其中与衍生其的野生型微生物相比,所述突变微生物具有更高的脂质生产力和/或表现出增加的碳与脂质分配。还提供使用所述突变微生物、引导RNA和用于产生突变微生物的核酸构建体产生脂质的方法。



1. 一种突变微生物,其编码包含GAF结构域的多肽的基因表达减弱,其中所述突变微生物:

- a) 比对照微生物多生产至少约25%的脂质;和/或
 - b) 相对于所述对照微生物,表现出增加的碳与脂质分配;
- 当所述突变微生物和对照微生物在相同条件下培养时。

2. 根据权利要求1所述的突变微生物,其中所述GAF结构域是GAF2结构域。

3. 根据权利要求2所述的突变微生物,其中所述GAF2结构域与SEQ ID NO:1所示的氨基酸序列具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的同一性。

4. 根据权利要求2所述的突变微生物,其中所述GAF2结构域包含SEQ ID NO:1所示的氨基酸序列或其保守变体。

5. 根据权利要求1所述的突变微生物,其中所述多肽包含与SEQ ID NO:2所示的氨基酸序列具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列。

6. 根据权利要求1所述的突变微生物,其中所述多肽包含SEQ ID NO:2所示的氨基酸序列或其保守变体。

7. 根据权利要求1所述的突变微生物,其中所述突变微生物包含针对长短鞭毛体或藻类物种中Naga_100020g79基因座或同线基因座中的基因或影响所述基因表达的一种或多种突变。

8. 根据权利要求7所述的突变微生物,其中所述突变微生物包含针对包含开放阅读框的基因或影响包含开放阅读框的基因的表达的突变,所述开放阅读框与SEQ ID NO:3具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的同一性。

9. 根据权利要求7所述的突变微生物,其中所述一种或多种突变存在于以下核酸中或影响以下核酸的表达:编码包含与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的核酸;和/或包含开放阅读框的核酸,所述开放阅读框包含与SEQ ID NO:3或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的核苷酸序列。

10. 根据权利要求7所述的突变微生物,其中所述一种或多种突变在Naga_100020g79基因座中的基因的GAF结构域中。

11. 根据权利要求1至10中任一项所述的突变微生物,其中所述对照微生物是野生型微生物。

12. 根据权利要求1至11中任一项所述的突变微生物,其中所述突变微生物比对照微生物

物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的脂肪酸甲酯可衍生的脂质(FAME脂质)。

13. 根据权利要求12所述的突变微生物,其中当在包含硝酸盐作为唯一氮源的培养基中培养时,所述突变微生物比对照微生物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的脂肪酸甲酯可衍生的脂质(FAME脂质)。

14. 根据权利要求12或13所述的突变微生物,其中所述突变微生物是藻类,并且当在光合自养条件下培养时,比对照微生物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的FAME脂质。

15. 根据权利要求12至14中任一项所述的突变微生物,其中所述微生物比对照微生物多产生至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%或至少约120%的贮存脂质。

16. 根据权利要求15所述的突变微生物,其中所述贮存脂质是三酰基甘油酯(TAG)。

17. 根据权利要求1至16中任一项所述的突变微生物,其中所述突变微生物表现出比所述对照微生物的FAME/TOC比高至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约100%、至少约120%、至少约140%、至少约160%、至少约180%、至少约200%、至少约220%、至少约240%、至少约260%、至少约280%、至少约300%、至少约320%、至少约340%、至少约360%、至少约380%或至少约400%的FAME/TOC比。

18. 根据权利要求17所述的突变微生物,其中所述突变微生物表现出至少约0.35的FAME/TOC比。

19. 根据权利要求1至18中任一项所述的突变微生物,其中脂质生产力通过分批、半连续或连续的生产力测定来确定。

20. 根据权利要求1至18中任一项所述的突变微生物,其中所述相同条件包含在包含小于2mM铵的培养基中培养所述突变微生物和对照微生物。

21. 根据权利要求20所述的突变微生物,其中所述相同条件包含在包含硝酸盐作为基本上唯一的氮源的培养基中培养所述突变微生物和对照微生物。

22. 根据权利要求1至21中任一项所述的突变微生物,其中所述相同条件包含在相对于所述对照微生物富营养物的培养基中培养所述突变微生物和对照微生物。

23. 根据权利要求1至22中任一项所述的突变微生物,其中所述突变微生物是经典来源的突变或基因工程突变。

24. 根据权利要求23所述的突变微生物,其中所述突变微生物在编码包含GAF结构域的多肽的基因或影响其表达的基因中具有突变,导致与所述基因在对照微生物中的表达相比,所述编码包括GAF结构域的多肽的基因的表达降低。

25. 根据权利要求23所述的突变微生物,其中所述突变是敲低突变。

26. 根据权利要求25所述的突变微生物,其中所述敲低突变是使用Cas/CRISPR系统产生的。

27. 根据权利要求23所述的突变微生物,其中所述突变微生物包含靶向编码具有GAF结构域的多肽的基因或影响其表达的基因的RNAi构建体、核酶构建体或反义构建体。

28. 根据权利要求24所述的突变微生物,其中所述突变微生物在编码包含GAF结构域的多肽的基因或影响其表达的基因中具有敲除突变。

29. 根据权利要求28所述的突变微生物,其中所述敲除突变是通过位点定向同源重组产生的。

30. 根据权利要求28或29所述的突变微生物,其中所述敲除突变通过部分或全部缺失、截短、移码或插入突变来破坏基因。

31. 根据权利要求28至30中任一项所述的突变微生物,其中所述敲除突变是使用大范围核酸酶、锌指核酸酶、类转录激活因子效应物核酸酶 (TALEN) 系统和/或Cas/CRISPR系统产生的。

32. 根据权利要求26、28、30或31中任一项所述的突变微生物,其中所述突变微生物包含Cas/CRISPR介导的插入所述基因中。

33. 根据权利要求1至32中任一项所述的突变微生物,其中所述突变微生物包含至少一种另外的基因修饰,所述基因修饰赋予除草剂抗性、毒素抗性、增强的生长特性、增强的光合效率、增强的脂质产生或积累以及特定脂质的产生。

34. 根据权利要求1至33中任一项所述的突变微生物,其中所述突变微生物是藻类或长鞭毛体物种。

35. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自以下组成的组的藻类物种:曲壳藻属 (Achnanthes)、茧形藻属 (Amphiprora)、双眉藻属 (Amphora)、纤维藻属 (Ankistrodesmus)、星胞藻属 (Asteromonas)、黄金色藻属 (Boekelovia)、波利氏藻属 (Bolidomonas)、包特氏菌属 (Borodinella)、气球藻属 (Botrydium)、葡萄藻属 (Botryococcus)、片球藻属 (Bracteococcus)、角毛藻属 (Chaetoceros)、四鞭藻属 (Carteria)、衣藻属 (Chlamydomonas)、绿球藻属 (Chlorococcum)、绿梭藻属 (Chlorogonium)、小球藻属 (Chlorella)、蓝隐藻属 (Chroomonas)、金球藻属 (Chrysosphaera)、球钙板藻属 (Cricosphaera)、隐甲藻属 (Crypthecodinium)、隐藻属 (Cryptomonas)、小环藻属 (Cyclotella)、链带藻属 (Desmodesmus)、杜氏藻属 (Dunaliella)、后棘藻属 (Elipsoidon)、圆石藻属 (Emiliania)、独球藻属 (Eremosphaera)、衣迪斯藻属 (Ernodesmus)、眼虫藻属 (Euglena)、真眼点藻属 (Eustigmatos)、披刺藻属

(*Franceia*)、脆杆藻属(*Fragilaria*)、拟脆杆藻属(*Fragilaropsis*)、丽丝藻属(*Gloeothamnion*)、红球藻属(*Haematococcus*)、菱板藻属(*Hantzschia*)、异弯藻属(*Heterosigma*)、膜胞藻属(*Hymenomonas*)、等鞭金藻属(*Isochrysis*)、鳞孔藻属(*Lepocinclis*)、微芒藻属(*Micractinium*)、蒜头藻属(*Monodus*)、单针藻属(*Monoraphidium*)、微球藻属(*Nannochloris*)、微拟球藻属(*Nannochloropsis*)、舟形藻属(*Navicula*)、新绿藻属(*Neochloris*)、肾鞭藻属(*Nephrochloris*)、肾藻属(*Nephroselmis*)、菱形藻属(*Nitzschia*)、棕鞭藻属(*Ochromonas*)、鞘藻属(*Oedogonium*)、卵囊藻属(*Oocystis*)、蚝球藻属(*Ostreococcus*)、拟小球藻属(*Parachlorella*)、雪衣藻属(*Parietochloris*)、杜氏亚属盐藻属(*Pascheria*)、巴夫藻属(*Pavlova*)、普莱格门(*Pelagomonas*)、褐指藻属(*Phoeodactylum*)、噬菌体属(*Phagus*)、皮级微绿藻属(*Picochlorum*)、扁藻属(*Platymonas*)、颗石藻属(*Pleurochrysis*)、肋球藻属(*Pleurococcus*)、原壁菌属(*Prototheca*)、拟绿球藻属(*Pseudochlorella*)、拟新绿藻属(*Pseudoneochloris*)、拟角星鼓藻属(*Pseudostaurastrum*)、塔胞藻属(*Pyramimonas*)、桑椹藻属(*Pyrobotrys*)、栅列藻属(*Scenedesmus*)、裂衣藻属(*Schizochlamydeella*)、骨条藻属(*Skeletonema*)、螺旋藻属(*Spyrogyra*)、裂丝藻属(*Stichococcus*)、四球藻属(*Tetrachlorella*)、四月藻属(*Tetraselmis*)、海链藻属(*Thalassiosira*)、黄丝藻属(*Tribonema*)、无隔藻属(*Vaucheria*)、鲜绿球藻属(*Viridiella*)、魏氏藻属(*Vischeria*)和团藻属(*Volvox*)。

36. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自由以下组成的组的长短鞭毛体:硅藻、黄绿藻、黄藻、褐藻、金藻或针胞藻。

37. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自由网粘菌纲(*Labyrinthulomycete*)物种组成的组的长短鞭毛体,所述物种为网粘菌属(*Labyrinthula*)、类网粘菌属(*Labyrinthuloides*)、破囊壶菌属(*Thraustochytrium*)、裂殖壶菌属(*Schizochytrium*)、不动壶菌属(*Aplanochytrium*)、橙黄壶菌属(*Aurantiochytrium*)、椭圆壶菌属(*Oblongichytrium*)、日本壶菌属(*Japonochytrium*)、戴普罗弗属(*Diplophrys*)或吾肯氏壶菌属(*Ulkenia*)。

38. 根据权利要求34所述的突变微生物,其中所述突变微生物是长短鞭毛体藻类。

39. 根据权利要求38所述的突变微生物,其中所述突变微生物是黄绿藻。

40. 根据权利要求39所述的突变微生物,其中所述突变微生物是选自由以下组成的组的属的物种:小绿藻属、绿囊藻属(*Chlorobryopsis*)、椭圆球藻属、真眼点藻属、角绿藻属、拟单胞藻属、蒜头藻属、微拟球藻属、假拟小椿藻属、拟角星鼓藻属、假四角藻属和魏氏藻属。

41. 根据权利要求38所述的突变微生物,其中所述突变微生物是硅藻。

42. 根据权利要求41所述的突变微生物,其中所述突变微生物是选自由以下组成的组的属的物种:双眉藻属、角毛藻属、小环藻属、脆杆藻属、拟脆杆藻属、菱板藻属、舟形藻属、菱形藻属、褐指藻属或海链藻属。

43. 一种产生脂质的方法,包含在培养基中培养根据权利要求1至42中任一项所述的突变微生物以产生脂质。

44. 根据权利要求43所述的方法,进一步包含从所述微生物、所述培养基或两者分离脂质。

45. 根据权利要求43或44所述的方法,其中所述微生物使用分批、连续或半连续培养条件进行培养。

46. 根据权利要求43至45中任一项所述的方法,其中所述微生物是藻类,并且所述培养是在光合自养条件下进行的。

47. 根据权利要求43至45中任一项所述的方法,其中所述微生物是网粘菌纲(Labyrinthulomycete),并且所述培养是在异养条件下进行的。

48. 根据权利要求43至47中任一项所述的方法,其中所述微生物在光生物反应器或发酵罐中培养。

49. 根据权利要求43至47中任一项所述的方法,其中所述微生物在池塘、运河、海基生长容器、沟渠、水道或渠道中生长。

50. 一种CRISPR系统的引导RNA,其中所述引导RNA包含对应于SEQ ID NO:4的序列。

51. 根据权利要求50所述的引导RNA,其中所述引导RNA是嵌合引导。

52. 根据权利要求50所述的引导RNA,其中所述引导RNA不包括tracr序列。

53. 一种用于同源重组的核酸构建体,其包含来自或邻近以下基因的核苷酸序列:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的藻类基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3和/或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

54. 一种用于表达反义RNA、shRNA、微小RNA或核酶的核酸构建体,其包含与以下基因的至少一部分互补的核苷酸序列:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3和/或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

55. 一种编码CRISPR系统的引导RNA的核酸分子,其中所述引导RNA包含以下基因的至少一部分:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的藻类基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

56. 一种用于产生根据权利要求1至42中任一项所述的突变微生物的方法,其包含将减弱包含GAF结构域的多肽表达的一种或多种突变和/或一种或多种试剂引入所述微生物。

57. 根据权利要求57所述的方法, 其中所述一种或多种突变影响包含SEQ ID NO:1和/或SEQ ID NO:2的氨基酸序列的多肽的表达; 位于Naga_100020g79基因座的基因的表达; 或包含开放阅读框的基因的表达, 所述开放阅读框包含SEQ ID NO:3的核苷酸序列。

58. 根据权利要求57所述的方法, 其中所述一种或多种试剂选自由以下组成的组: 反义RNA、RNAi、shRNA、微小RNA、核酶、大范围核酸酶、锌指核酸酶、Cas/CRISPR系统的组分和/或类转录激活因子效应物核酸酶 (TALEN) 系统的组分。

通过基因修饰信号传导蛋白提高藻类脂质生产力

[0001] 相关申请的交叉引用

[0002] 本申请根据35U.S.C.§119(e) 要求2017年12月5日提交的美国序列第62/594,930号的优先权,其全部内容通过引用以其整体并入本文。

[0003] 序列表的并入

[0004] 所附序列表中的材料特此通过引用并入本申请。名为SGI2120_1W0_序列表,txt的所附序列表文本文件创建于2018年11月19日,大小为34kb。可以在使用Windows OS的计算机上使用Microsoft Word访问该文件。

技术领域

[0005] 本发明涉及具有增加的脂质生产力的突变微生物,例如藻类和长短鞭毛体,以及它们用于生产脂质的方法。

[0006] 通过增加脂质生物合成来提高脂质生产力的各种尝试集中在操纵编码用于氮同化或脂质代谢的酶的基因或编码参与脂质储存的多肽的基因。例如,US2014/0162330公开了一种三角褐指藻(*Phaeodactylum tricornutum*)藻株,其中硝酸还原酶(NR)基因已被基于RNAi的敲低减弱;Trentacoste等人((2013)《美国科学学院学报(Proc.Natl.Acad.Sci.USA)》110:19748-19753)公开了用靶向预计参与脂质分解代谢的Thaps3_264297基因的RNAi构建体转化的硅藻;并且W02011127118公开了用编码油质蛋白(脂质存储蛋白)的基因以及编码二酰基甘油转移酶(DGAT)的基因转化衣藻属(*Chlamydomonas*)。尽管在每种情况下,基于显微镜或用亲脂性染料染色来确定脂质产生增加,但没有提供通过所操纵的细胞的定量脂质产生,也没有测定随时间的脂质生产力。

[0007] Daboussi等人2014(《自然通讯(Nature Comm.)》5:3881)报道了破坏三角褐指藻中的UGPase基因(其被认为是为海带多糖(一种贮藏碳水化合物)合成提供前体),导致脂质积累增加。然而,没有生化数据展示表明海带多糖含量受到影响(或甚至存在),并且没有报道脂质和生物质生产力。类似地,几个小组已经报道了衣藻属无淀粉突变体中脂质积累的增加(Wang等人,2009《真核细胞(Eukaryotic Cell)》8:1856-1868;Li等人2010《代谢工程(Metab Eng.)》12:387-391),然而,实际测量脂质生产力的连续报道的结论是这些藻株在光合自养条件下生长时生长受到损害(Siaut等人,2011《BMC生物技术(BMC Biotechnol.)》11:7;Davey等人2014《真核细胞(Eukaryot Cell)》13:392-400)。这些报告得出的结论是,野生型亲本藻株实际上实现了最高的脂质生产力(以每天每升TAG测量)。

[0008] W0 2011/097261和US20120322157报道了编码抑制蛋白的被称为“SN03”的基因在衣藻属中过表达时在营养充足的条件下具有增加脂质产生的作用。然而,观察到SN03基因的过表达导致未鉴定的极性脂质的出现,其未被定量,并且不导致甘油三酯(TAG)的增加。Boyle等人((2012)《生物化学杂(J.Biol.Chem.)》287:15811-15825)描述了另一种被鉴定为潜在调节应激诱导的脂质生物合成的多肽。在衣藻属中敲除编码“SQUAMOUS”结构域多肽的NRR1基因导致在氮耗尽下的脂质生物合成相对于野生型细胞减少;然而,没有获得显示脂质产生增加的突变体。US 2010/0255550建议在藻类细胞中过表达推定的转录因子

(TF1、TF2、TF3、TF4和TF5)来增加脂质产生,但没有公开此类藻株。

[0009] 美国专利申请公开第US 2017/005803号公开了编码包括锌Cys结构域的调节子的基因,当在包括硝酸盐的培养基中培养时,该基因的弱化导致突变藻类中脂质生产力增加。突变藻类在培养物中显示生长,以野生型细胞的至少80%的速率积累生物质,同时产生多达野生型祖代藻株两倍的脂质。美国专利申请公开第US 2017/0121742号公开了突变藻类,其编码具有溴结构域和TAZ锌指结构域的多肽的基因表达减弱,所述突变藻类显示了相对于野生型藻类,脂质生产力提高并且生物质生产力的降低最小。

发明内容

[0010] 在一方面,本文提供突变微生物,其编码多肽的基因表达减弱,所述多肽包括称为cGMP特异性磷酸二酯酶、腺苷酸环化酶和转录激活因子FhlA (GAF) 结构域的结构域,例如GAF2结构域(例如SEQ ID NO:1),在一些实施例中,当突变微生物和对照微生物在相同条件下培养时,其比对照微生物产生更多的脂质和/或相对于对照微生物表现出增加的碳与脂质分配。在一些实施例中,对照微生物是野生型微生物。在一些实施例中,突变微生物比对照微生物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%(例如,多25%、50%、100%、150%、200%或250%中的至少任一种)的脂肪酸甲酯可衍生的脂质(FAME脂质)。在一些实施例中,本文提供的突变微生物表现出比对照微生物的FAME/TOC比高至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约100%、至少约120%、至少约140%、至少约160%、至少约180%、至少约200%、至少约220%、至少约240%、至少约260%、至少约280%、至少约300%、至少约320%、至少约340%、至少约360%、至少约380%或至少约400%(例如,高50-400%)的FAME与TOC比。

[0011] 如上所述的基因的“减弱的表达”包括,例如,基因的表达降低,使得产生降低水平的(其可能是不可检测的水平)由所述基因编码的功能性多肽。表达减弱还包括其中产生突变(如缺失、截短或移码)多肽的表达,使得突变多肽相对于非突变多肽具有降低或改变的功能。表达减弱可以是编码多肽的基因突变的结果,或者可以是设计为降低编码多肽的基因表达的构建体,例如靶向基因的RNAi构建体的表达或递送的结果。

[0012] 在一些实施例中,通过传统诱变或通过基因工程技术产生突变微生物。在一些实施例中,突变微生物可以在编码包括GAF(例如GAF2)结构域的多肽的基因或影响其表达的基因中具有突变,与对照微生物中基因的表达水平相比,所述突变导致编码包括GAF结构域的多肽的基因的表达水平降低。在一些实施例中,使用诱导双链断裂的一种或多种试剂产生一种或多种突变。在一些实例中,所述试剂是大范围核酸酶、锌指核酸酶、类转录激活因子效应物核酸酶(TALEN)系统和/或Cas核酸酶。在一些实施例中,导致编码GAF结构域的基因的表达减弱的突变是插入突变。

[0013] 在一些实施例中,突变微生物是任何真核微生物,并且在说明性实施例中,突变微生物是长短鞭毛体或藻类。在一些实例中,突变微生物是长短鞭毛体藻类,例如硅藻或黄绿

藻 (Eustigmatophyte) 物种, 并且可以是例如硅藻属的物种, 例如茧形藻属 (Amphiprora)、双眉藻属 (Amphora)、角毛藻属 (Chaetoceros)、小环藻属 (Cyclotella)、脆杆藻属 (Fragilaria)、拟脆杆藻属 (Fragilaropsis)、菱板藻属 (Hantzschia)、舟形藻属 (Navicula)、菱形藻属 (Nitzschia)、褐指藻属 (*Phaeodactylum*)、褐指藻属 (*Phaeodactylum*)、褐指藻属 (*Phaeodactylum*)、骨条藻属 (Skeletonema) 或海链藻属 (Thalassiosira)。在一些实例中, 突变藻类是黄绿藻, 例如属于诸如以下的属的黄绿藻: 小绿藻属 (Chloridella)、绿囊藻属 (Chlorobptryx)、椭圆藻属 (Ellipsoidion)、真眼点藻属 (Eustigmatos)、角绿藻属 (Goniochloris)、拟单胞藻属 (Monodopsis)、蒜头藻属 (Monodus)、微拟球藻属 (Nannochloropsis)、假拟小椿藻属 (Pseudocharaciopsis)、拟角星鼓藻属 (Pseudostaruastrum)、假四角藻属 (Pseudotetraëdriella) 和魏氏藻属 (Vischeria)。

[0014] 本文还提供包含本文提供的突变体的生物物质。进一步提供本文提供的突变体的提取物。提取物可以是粗提取物或部分纯化、纯化或精制的提取物, 其可以包括细胞组分的任何组合, 包括但不限于膜、脂质、蛋白质、碳水化合物、可溶性分子和不溶性分子。例如, 提取物可以任选地是经历一种或多种处理的提取物, 所述处理例如但不限于选择性沉淀、高温或低温处理、过滤或离心。

[0015] 还包括使用本文公开的突变微生物生产脂质的方法。例如, 本文提供的突变微生物可以分批、半连续或连续培养的方式培养, 以产生一种或多种脂质。所述方法可以包括从培养物中 (例如从细胞、培养基或整个培养物中) 分离一种或多种脂质。培养基可以是富氮的或可以是氮受限的或贫氮的。在一些实施例中, 如本文所提供的用于培养突变微生物以产生脂质的培养基可以包括硝酸盐作为基本上唯一的氮源。在一些实例中, 所述方法可包括在光合自养条件下培养藻类突变体。

[0016] 还包括用于表达引导RNA的DNA分子; 靶向编码影响脂质产生的含GAF2结构域的蛋白质的基因的引导RNA; 和使用基因工程技术产生突变微生物的核酸构建体。靶向GAF基因 (例如GAF2基因) 的引导RNA可与包括GAF结构域的基因的编码区具有同源性, 或可与内含子、5'UTR、3'UTR或5'UTR上游基因区的序列具有同源性。

[0017] 当结合附图阅读本发明的以下详细描述时, 本发明的这些和其它目的和特征将变得更加显而易见。

附图说明

[0018] 图1是由位于卡迪斯微拟球藻 (*N. gaditana*) 基因座Naga_100020g79的基因 (由SEQ ID NO:2表示) 编码的蛋白质的示意图。方框表示功能未知的保守结构域 (在SEQ ID NO:2所示基因产物的氨基酸残基109-390处) 和GAF2结构域 (在SEQ ID NO:2所示基因产物的氨基酸残基822-1047处)。标记GE-13535、GE13536 K0的箭头显示使用CRISPR/Cas9插入潮霉素抗性标志物来产生敲除藻株GE-13535和GE-13536的位置。该图未按比例绘制。

[0019] 图2是展示使用硝酸盐作为氮源分批培养生长的突变株GE-13535和GE-13536以及野生型藻株WT-3730的FAME/TOC比的图。误差条代表两个生物学复制的标准偏差。图中使用的符号: 菱形代表野生型WT-3730, 圆圈代表敲除突变体GE-13535, 以及三角形代表敲除突

变体GE-13536。

[0020] 图3是描绘在分批测定第7天从野生型WT-3730 (无模式) 和敲除藻株GE-13535 (虚线) 和GE-13536 (斜线) 分离的脂质中存在的各种链长的脂肪酸的量的柱状图。

[0021] 图4至6提供描绘在分批测定中卡迪斯微拟球藻野生型藻株和GE-13536敲除藻株生产力的图, 其中培养基含有硝酸盐 (NO_3) 作为唯一氮源或补充有铵 (NO_3+NH_4)。图4展示八天的测定中每日的FAME生产力; 图5展示八天的测定中每日的TOC生产力; 图6提供测定每天的FAME/TOC比。误差条代表除第7天外的两个生物学复制的标准偏差。图中使用的符号: 空心菱形代表在仅硝酸盐的培养基PM074中培养的野生型WT-3730, 实心菱形代表在补充有铵PM124的硝酸盐培养基中培养的野生型WT-3730, 空心三角形代表在仅硝酸盐的培养基PM074中培养的敲除突变体GE-13536, 以及实心三角形代表在补充有铵PM124的硝酸盐培养基中培养的敲除突变体GE-13536。

具体实施方式

[0022] 定义

[0023] 除非另有定义, 否则本文所用的所有技术和科学术语具有与本发明所属领域的普通技术人员通常理解的相同的含义。在发生冲突的情况下, 以包括定义的本申请为准。除非上下文另有要求, 否则单数术语应包括复数, 并且复数术语应包括单数。

[0024] 所有标题都是为了读者的方便, 并不以任何方式限制本发明。对本发明的方面或实施例的引用不一定指示所描述的方面不可以与本发明的其它所描述的方面或本发明的其它方面的特征组合。出于所有目的, 本文提及的所有出版物、专利和其它参考文献通过引用整体并入本文, 如同每个单独的出版物或专利申请均被明确地和单独地指出通过引用并入。

[0025] 在本文中在短语如“A和/或B”中使用的术语“和/或”旨在包括“A和B”、“A或B”、“A”和“B”。

[0026] 当在数值或范围的列表之前时, 术语“约”、“大约”等是指该列表或范围中的每个单独的值, 如同该列表或范围中的每个单独的值紧随该术语之后。该术语表示相同指代的值与之精确相同、接近或相似。在一些用法中, “约”或“大约”可以表示在规定值的10%之内、5%之内或2.5%之内。

[0027] “任选的”或“任选地”表示随后描述的事件或状况可以发生或不能发生, 并且该描述包括事件或状况发生的情况以及事件或状况没有发生的情况。例如, 短语任选地组合物可以包含组合表示组合物可以包含不同分子的组合, 或者可以不包含组合, 使得该描述既包括组合又不包括组合 (即, 组合的各个成员)。范围可以在本文中表示为从约一个特定值和/或至约另一特定值。当表达此类范围时, 另一方面包括从一个特定值和/或至另一特定值。类似地, 当通过使用先行词约或大约将值表示为近似值时, 将理解该特定值形成另一方面。还将理解的是, 每个范围的端点相对于另一端点并且独立于另一端点都是重要的。另外, 范围 (例如90-100%) 意味着包括范围本身以及该范围内的每个独立值, 如同每个值被单独列出。除非另外特别指出, 否则本申请内提供的所有范围均包括该范围的上端和下端的值。术语“组合的”或“组合”或“结合”可以指一起施用的试剂的物理组合, 或者参考例如时间和/或物质性, 同时使用两种或更多种试剂。

[0028] 在未阐明在特定上下文中“基本上”的预期含义的情况下,该术语用于包括次要和无关的偏差,该偏差对在本发明主题的上下文中认为重要的特征并不重要。

[0029] 术语“基因”被广泛地用来指编码多肽或表达的RNA的核酸分子(通常为DNA,但任选为RNA)的任何片段。因此,基因包括编码表达的RNA的序列(其可以包括多肽编码序列或例如功能性RNA,如核糖体RNA、tRNA、反义RNA、微小RNA、短发夹RNA、核酶等)。基因可进一步包含其表达需要或影响其表达的调节序列,以及与处于其天然状态的蛋白质或RNA编码序列相关的序列,例如内含子序列、5'或3'非翻译序列等。在一些实例中,“基因”可以仅指DNA或RNA分子的蛋白质编码部分,其可以包括或不包括内含子。基因的长度优选为大于50个核苷酸,长度更优选为大于100个核苷酸,并且例如长度可以为50个核苷酸至500,000个核苷酸,例如长度为100个核苷酸至100,000个核苷酸或长度为约200个核苷酸至约50,000个核苷酸,或长度为约200个核苷酸至约20,000个核苷酸。基因可以从多种来源获得,包括从目标来源克隆或从已知或预测的序列信息合成。

[0030] 术语“核酸”或“核酸分子”是指DNA或RNA的片段(例如,mRNA),并且还包括具有修饰的主链的核酸(例如,肽核酸、锁核酸和其它修饰的核酸或核酸类似物(例如膦肽核酸,Efimov和Chakhmakhcheva (2005)《分子生物学方法(Methods Mol Biol.)》288:147-163))或修饰的或非天然存在的核碱基。核酸分子可以是双链的、部分双链的或单链的;包含基因或其一部分的单链核酸分子可以是编码(有义)链或非编码(反义)链。尽管核酸的序列可以以DNA的形式展示,但是本领域普通技术人员认识到相应的RNA序列将具有相似的序列,其中胸腺嘧啶被尿嘧啶代替,即“t”被“u”代替。

[0031] 核酸分子可以“源自”指定的来源,其包括从指定的来源(全部或部分)分离核酸片段。核酸分子也可以通过例如从指定的多核苷酸源直接克隆、PCR扩增或人工合成而源自指定的源,或基于与指定的多核苷酸源(例如可以是生物体物种)相关的序列。源自特定来源或物种的基因或核酸分子还包括相对于来源核酸分子具有序列修饰的基因或核酸分子。例如,源自来源的基因或核酸分子(例如,特定参考基因)可以包括相对于来源基因或核酸分子的一种或多种突变,所述突变是非故意的或故意引入的,并且如果故意引入一种或多种突变,包括取代、缺失或插入,则可以通过细胞或核酸的随机或靶向突变、通过扩增或其它基因合成或分子生物学技术、或通过化学合成或其任何组合来引入序列改变。源自编码功能性RNA或多肽的参考基因或核酸分子的基因或核酸分子可以编码与参考或源功能性RNA或多肽或其功能性片段具有至少约50%、至少约55%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的功能性RNA或多肽。例如,源自编码功能性RNA或多肽的参考基因或核酸分子的基因或核酸分子可以编码与参考或源功能性RNA或多肽或其功能性片段具有至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的序列同一性的功能性RNA或多肽。

[0032] 如本文所用,将“分离的”核酸或蛋白质从其自然环境或核酸或蛋白质自然存在的背景中去除。例如,将分离的蛋白质或核酸分子从在其天然或自然环境中与其相关的细胞或生物体中去除。在某些情况下,分离的核酸或蛋白质可以被部分或基本上纯化,但是分离不需要特定的纯化水平。因此,例如,分离的核酸分子可以是已经从其自然整合的染色体、基因组或附加体切除的核酸序列。

[0033] “纯化的”核酸分子或核苷酸序列或蛋白质或多肽序列基本上不含细胞物质和细胞组分。例如,纯化的核酸分子或蛋白质可以基本上不含除缓冲液或溶剂之外的化学物质。“基本上不含”并不旨在表示不可检测到除新核酸分子之外的其它组分。

[0034] 术语“天然存在的”和“野生型”是指在自然界中最常观察到的自然存在的形式,并且因此被任意称为“野生型”。例如,天然存在的或野生型核酸分子、核苷酸序列或蛋白质可以存在于天然来源中或与天然来源分离,并且未被人为操纵进行有意修饰。“野生型”还可以指在一个或多个特定核苷酸位置处的序列,或在一个或多个特定密码子位置处的序列,或在一个或多个特定氨基酸位置处的序列。

[0035] 如本文所用,“减弱”是指数量、程度、强度或浓度降低。减弱的基因表达可以指所述基因的转录或编码蛋白的翻译、折叠或组装的显著降低的量和/或速率。作为非限制性实例,减弱的基因可以是突变或破坏的基因(例如,通过部分或全部缺失、截短、移码或插入突变破坏的基因),其不编码完整的功能性开放阅读框或由于基因调节序列的改变或破坏而具有降低的表达。减弱的基因也可以是由构建体靶向的基因,所述构建体降低基因的表达,例如反义RNA、微小RNA、RNAi分子或核酶。减弱的基因表达可以是被消除(例如减少到不显著或不可检测的量)的基因表达。减弱的基因表达也可以是导致RNA或蛋白质不完全起作用或不起作用的基因表达,例如,减弱的基因表达可以是导致截短的RNA和/或多肽的基因表达。

[0036] “外源核酸分子”或“外源基因”是指已被引入(“转化”)到细胞中的核酸分子或基因。转化的细胞可以被称为重组细胞,其中可以引入其它外源基因。如果用核酸分子转化的细胞的后代继承了外源核酸分子,则也称为“转化”。相对于被转化的细胞,外源基因或核酸分子可以源自不同的物种(并且因此是“异源的”),或者可以源自相同的物种(并且因此是“同源的”)。“内源”核酸分子、基因或蛋白质是当其存在于宿主中或由宿主天然产生时的天然核酸分子、基因或蛋白质。

[0037] 本文使用术语“天然”来指示核酸序列或氨基酸序列,因为它们天然存在于宿主中。本文使用术语“非天然的”来指示在宿主中不是天然存在的核酸序列或氨基酸序列。因此,“非天然”核酸分子是宿主细胞中非天然存在的核酸分子,例如,非天然核酸分子对于其被引入的宿主细胞或微生物是外源的,并且相对于宿主细胞或微生物可以是异源的。另外,已经从细胞中取出、经历实验室操作、并引入或再引入宿主细胞中,使得其在基因组中的序列或位置相对于其在未操作的生物体中的位置不同(即,与其在未转化的生物体中不并置或不可操作地连接的序列并置或可操作地连接)的核酸序列或氨基酸序列被认为是“非天然的”。非天然基因还包括宿主微生物内源的基因,该基因可操作地连接到已经重组入宿主基因组中的一种或多种异源调节序列。

[0038] “重组”或“工程化”核酸分子是已经通过人为操纵而改变的核酸分子。作为非限制性实例,重组核酸分子包括以下任何核酸分子:1)已在体外部分或完全合成或修饰,例如,使用化学或酶促技术(例如,通过使用化学核酸合成,或通过使用用于核酸分子复制、聚合、消化(核酸外切或核酸内切)、连接、逆转录、转录、碱基修饰(包括例如甲基化)、整合或重组(包括同源和位点特异性重组)的酶);2)包括本质上不相连的相连核苷酸序列;3)已使用分子克隆技术进行工程化,使得其相对于天然存在的核酸分子序列缺少一个或多个核苷酸;和/或4)已使用分子克隆技术进行操作,使得其相对于天然存在的核酸序列具有一个或多

个序列改变或重排。作为非限制性实例，cDNA是重组DNA分子，如同通过体外聚合酶反应产生，或已连接接头或已整合到载体如克隆载体或表达载体中的任何核酸分子。

[0039] 如本文所用，术语“重组蛋白”是指通过基因工程产生的蛋白质，而不管氨基酸是否不同于野生型蛋白的氨基酸。

[0040] 当应用于生物体时，术语重组、工程化或基因工程化是指已通过将异源或外源重组核酸序列引入生物体而操作的生物体(例如非天然核酸序列)，并且包括基因敲除、靶向突变、基因置换和启动子置换、缺失、破坏或插入，以及将转基因或合成基因或核酸序列引入生物体。即，重组、工程化或基因工程化是指已经被人为干预改变的生物体。重组或基因工程化生物也可以是已引入基因“敲低”构建体的生物体。此类构建体包括但不限于RNAi、微小RNA、shRNA、siRNA、反义和核酶构建体。还包括其基因组已被大范围核酸酶、锌指核酸酶、TALEN或Cas/CRISPR系统的活性改变的生物体。可以将外源或重组核酸分子整合到重组/基因工程化生物体的基因组中，或者在其它情况下，可以不整合到宿主基因组中。如本文所用，“重组微生物”或“重组宿主细胞”包括本发明的重组微生物的后代或衍生物。因为由于突变或环境影响而在后代中可能发生某些修饰，所以后代或衍生物实际上可能与亲本细胞不同，但仍包括在本文所用术语的范围内。

[0041] 术语“启动子”是指能够结合细胞中的RNA聚合酶并启动下游(3'方向)编码序列的转录的核酸序列。启动子包括以高于背景的可检测水平启动转录所必需的最小数目的碱基或元件。启动子可包括转录起始位点以及负责RNA聚合酶结合的蛋白质结合结构域(共有序列)。真核启动子通常但不总是含有“TATA”框和“CAT”框。原核启动子可以含有-10和-35原核启动子共有序列。来自多种不同来源的大量启动子，包括组成型、诱导型和阻遏型启动子在本领域中是众所周知的。代表性来源包括例如藻类、病毒、哺乳动物、昆虫、植物、酵母和细菌细胞类型，并且来自这些来源的合适的启动子容易获得，或者可以基于在线公开获得的序列合成制备，或例如来自保藏中心如ATCC以及其它商业或个人来源。启动子可以是单向的(在一个方向上启动转录)或双向的(在任一方向上启动转录)。启动子可以是组成型启动子、阻遏型启动子或诱导型启动子。除了其中RNA聚合酶结合以启动转录的基因近端启动子之外，启动子区域还可以包括可以在约1kb、约2kb、约3kb、约4kb、约5kb或更多基因转录起始位点之内的基因上游的其它序列，其中其它序列可影响下游基因的转录速率，并任选影响启动子对发育、环境或生化(例如，代谢)条件的反应性。

[0042] 当提及多核苷酸、基因、核酸、多肽或酶使用时，术语“异源的”是指来自宿主生物体物种之外的源或源自宿主生物体物种之外的源的多核苷酸、基因、核酸、多肽或酶。相反，本文使用“同源的”多核苷酸、基因、核酸、多肽或酶表示源自宿主生物体物种的多核苷酸、基因、核酸、多肽或酶。当提及基因调节序列或用于维持或操纵基因序列的辅助核酸序列(例如，启动子、5'非翻译区、3'非翻译区、多聚腺苷酸添加序列、内含子序列、剪接位点、核糖体结合位点、内部核糖体进入序列、基因组同源区、重组位点等)时，“异源的”表示调节序列或辅助序列不与其调节或辅助核酸序列在构建体、基因组、染色体或附加体中并置的基因天然相关。因此，在自然状态下(即在非基因工程化生物体的基因组中)与和其没有可操作连接的基因可操作地连接的启动子在本文中被称为“异源启动子”，即使该启动子可以源自与其连接的基因相同的物种(或在某些情况下的相同的生物体)。

[0043] 如本文所用，术语“蛋白质”或“多肽”旨在涵盖单数“多肽”以及多个“多肽”，并且

是指由通过酰胺键(也称为肽键)线性连接的单体(氨基酸)组成的分子。术语“多肽”是指两个或更多个氨基酸的任何一条或多条链,并且不指产物的特定长度。因此,“多肽”的定义中包括肽、二肽、三肽、寡肽、“蛋白质”、“氨基酸链”或用于指代两个或更多个氨基酸的一条或多条链的任何其它术语,并且术语“多肽”可以代替或与这些术语中的任一种互换使用。

[0044] 通常在基因或物种名称后的括号中提供的基因和蛋白质登录号是序列记录的唯一标识符,该序列记录可在由美国国立卫生研究院维护的美国国家生物技术信息中心(NCBI)网站(ncbi.nlm.nih.gov)上公开获得。“GenInfo标识符”(GI)序列标识号特定于核苷酸或氨基酸序列。如果序列以任何方式发生变化,则会分配一个新的GI号。序列修订历史记录工具可用于跟踪出现在特定基因库(GenBank)记录中的序列的各种GI号、版本号和更新日期。基于登录号和GI号来检索和获得核酸或基因序列或蛋白质序列是在例如细胞生物学、生物化学、分子生物学和分子遗传学领域中众所周知的。基因位点标识符是指在Corteggiani Carpinelli等人(2014)《分子植物(Mol Plant)》7:323-335中描述的已公开基因组,并且可在CRIBI基因组学微拟球藻基因组门户网站(CRIBI Genomics Nannochloropsis genome portal)(nannochloropsis.org)上在线获得。

[0045] 如本文所用,关于核酸或多肽序列的术语“同一性百分比”或“同源性”定义为在以最大同一性百分比对准序列并在必要时引入缺口以实现最大同源性百分比后,候选序列中与已知多肽相同的核苷酸或氨基酸残基的百分比。对于多肽序列,不应将N末端或C末端插入或缺失解释为影响同源性,并且少于约65、少于约60、少于约50、少于约40、少于约30、少于约20或少于约10个氨基酸残基的多肽序列内部缺失和/或插入不应被解释为影响所比较的氨基酸(蛋白质)序列的同源性。对于核酸序列,5'末端或3'末端的插入或缺失不应被解释为影响同源性,并且少于约200、少于约180、少于约150、少于约120、少于约100、少于约90、少于约80、少于约70、少于约60、少于约50、少于约40或少于约30个核苷酸的多肽序列内部缺失和/或插入不应被解释为影响所比较的核酸序列的同源性。可以使用被定制用于序列相似性搜索的blastp、blastn、blastx、tblastn和tblastx程序使用的算法通过BLAST(基本局部比对搜索工具)分析来确定核苷酸或氨基酸序列水平的同源性或同一性(Altschul(1997)《核酸研究(Nucleic Acids Res.)》25,3389-3402和Karlin(1990)《美国科学院学报(Proc.Natl.Acad.Sci.USA)》87,2264-2268)。BLAST程序所用的方法是首先考虑查询序列和数据库序列之间的相似片段(有或没有缺口),然后评估所有被鉴定匹配的统计显著性,并且最后仅概括那些满足预选显著性阈值的匹配。有关序列数据库相似性检索中基本问题的讨论,请参见Altschul(1994),《自然遗传学(Nature Genetics)》6,119-129。直方图、描述、对准、期望(即,用于报告与数据库序列匹配的统计显著性阈值)、截止值、矩阵和过滤器(低复杂度)的检索参数可以处于默认设置。Blastp、blastx、tblastn和tblastx使用的默认评分矩阵是BLOSUM62矩阵(Henikoff(1992)《美国科学院学报(Proc.Natl.Acad.Sci.USA)》89,10915-10919),建议用于长度超过85的查询序列(核苷酸碱基或氨基酸)。

[0046] 对于设计用于比较核苷酸序列的blastn,评分矩阵由M(即,一对匹配残基的奖励分数)与N(即,错配残基的惩罚分数)之比设置,其中M和N的默认值分别可以为+5和-4。四个blastn参数可以如下调整:Q=10(缺口产生罚分);R=10(缺口延伸罚分);wink=1(在查询中的每个Winkth位置产生字命中);和gapw=16(设置其中产生缺口对准的窗口宽度)。用于

比较氨基酸序列的等效Blastp参数设置可以是:Q=9;R=2;wink=1;和gapw=32。在GCG软件包10.0版中可用的序列之间的Bestfit比较可以使用DNA参数GAP=50(缺口产生罚分)和LEN=3(缺口延伸罚分),并且蛋白质比较中的等效设置可以是GAP=8和LEN=2。

[0047] 因此,当提及本发明的多肽或核酸序列时,包括与全长多肽或核酸序列或其片段具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%、至少约99%或约100%(例如50%、75%或90%中的至少任一种)序列同一性的序列同一性,其包含整个蛋白质的至少约100个、至少约125个、至少约150个(例如,至少100个)或更多个氨基酸残基的连续序列;此类序列的变体,例如,其中至少一个氨基酸残基已经在所公开的含有插入和取代的序列的N-和/或C-末端和/或之内插入。所考虑的变体可以另外或替代地包括含有通过例如同源重组或定点或PCR诱变导致的预定突变的那些,和其它物种的相应多肽或核酸,包括但不限于本文所述的那些,含有插入和取代的多肽或核酸家族的等位基因或其它天然存在的变体;和/或衍生物,其中多肽已经通过取代、化学、酶或其它合适的方法用除含有插入和取代的天然存在的氨基酸之外的部分(例如,可检测部分如酶)共价修饰。

[0048] 如本文所用,短语“保守氨基酸取代”或“保守突变”是指一个氨基酸被具有共有特性的另一氨基酸替代。定义单个氨基酸之间共有特性的一种功能方法是分析同源生物体相应蛋白质之间氨基酸变化的标准化频率(Schulz(1979)《蛋白质结构原理(Principles of Protein Structure)》,斯普林格出版社(Springer-Verlag))。根据此类分析,可以定义氨基酸组,其中一个组内的氨基酸优先相互交换,并且因此,它们彼此对整体蛋白质结构的影响最为相似(Schulz(1979)《蛋白质结构原理(Principles of Protein Structure)》,斯普林格出版社)。以这种方式定义的氨基酸基团的实例可以包括:“带电/极性基团”,包括Glu、Asp、Asn、Gln、Lys、Arg和His;“芳族或环状基团”,包括Pro、Phe、Tyr和Trp;以及“脂族基团”,包括Gly、Ala、Val、Leu、Ile、Met、Ser、Thr和Cys。在每个组中,也可以鉴定亚组。例如,带电/极性氨基酸的组可以细分为亚组,包括:包含Lys、Arg和His的“带正电的亚组”;包含Glu和Asp的“带负电的亚组”;以及包含Asn和Gln的“极性亚组”。在另一个实例中,芳族或环状基团可以细分为亚组,包括:包含Pro、His和Trp的“氮环亚组”;以及包含Phe和Tyr的“苯基亚组”。在另一个进一步的实例中,脂族基团可以细分为亚组,包括:包含Val、Leu和Ile的“大脂族非极性亚组”;包含Met、Ser、Thr和Cys的“脂族弱极性亚组”;以及包含Gly和Ala的“小残基亚组”。保守突变的实例包括上述亚组内的氨基酸的氨基酸取代,例如但不限于:用Arg代替Lys,或反之亦然,使得可以保持正电荷;用Asp代替Glu,或反之亦然,使得可以保持负电荷;用Thr代替Ser,或反之亦然,使得可以保持游离的-OH;和用Asn代替Gln,或反之亦然,使得可以保持游离的-NH₂。“保守变体”是包括一个或多个氨基酸的多肽,所述氨基酸已被具有共有特性如属于上述相同的氨基酸组或亚组的氨基酸替换以代替参考多肽的一个或多个氨基酸(例如,其序列已在出版物或序列数据库中公开或其序列已由核酸测序确定的多肽)。

[0049] 如本文所用,“表达”包括至少在RNA产生水平上的基因表达,并且“表达产物”包括所得产物,例如表达基因的多肽或功能性RNA(例如核糖体RNA、tRNA、反义RNA、微小RNA、shRNA、核酶等)。术语“增加的表达”包括基因表达的改变以促进mRNA产生的增加和/或多肽表达的增加。[基因产物的]“产生的增加”包括与多肽的天然产生或酶活性相比,多肽表达

量、多肽的酶活性水平或两者的组合的增加。

[0050] 本发明的一些方面包括特定多核苷酸序列的表达的部分、实质或完全缺失、沉默、失活或下调。可以使基因部分、基本上或完全缺失、沉默、失活，或者它们的表达可以被下调，以影响由它们编码的多肽执行的活性，例如酶的活性。通过插入破坏基因功能和/或表达的核酸序列（例如病毒插入、转座子诱变、大范围核酸酶工程化、同源重组、CRISPR/cas系统或本领域已知的其它方法），可以使基因部分、基本上或完全缺失、沉默、失活或下调。术语“消除(eliminate)”、“消除(elimination)”和“敲除”可以与术语“缺失”、“部分缺失”、“实质缺失”或“完全缺失”互换使用。在某些实施例中，可以使用cas/CRISPR系统通过定点同源重组或靶向整合或突变来工程化目标微生物以敲除特定目标基因。在仍其它实施例中，可以使用利用cas/CRISPR系统、RNAi或反义DNA(asDNA)构建体的基因靶向插入或突变来部分、基本上或完全沉默、失活或下调目标特定基因。

[0051] 某些核酸分子或特定多核苷酸序列的这些插入、破坏、缺失、碱基修饰或其它修饰可以理解为涵盖“基因修饰”或“转化”，使得微生物的所得藻株或宿主细胞可以理解为是“基因修饰的”、“基因工程化的”或“转化的”。

[0052] 如本文所用，“上调的”或“上调”包括目标基因或核酸分子的表达或酶活性的增加，例如与在其它方面相同但尚未上调的基因或酶中的表达或活性相比，基因表达或酶活性的增加。

[0053] 如本文所用，“下调的”或“下调”包括目标基因或核酸分子的表达或酶活性的降低，例如与在其它方面相同但尚未下调的基因或酶中的表达或活性相比，基因表达或酶活性的降低。

[0054] 如本文所用，“突变体”是指在基因中具有突变的生物体，所述突变是传统诱变的结果，例如，使用 γ 辐射、UV或化学诱变剂。本文所用的“突变体”还指由于基因工程而具有改变的基因结构或表达的重组细胞，作为非限制性实例，其可包括过表达，包括在不同的时间、生物或环境调节下和/或在与天然发生的程度不同的程度上的基因表达，和/或在重组细胞中非天然表达的基因的表达；同源重组，包括敲除和敲入（例如，用编码具有比野生型多肽更高或更低活性的多肽和/或显性失活多肽的基因进行基因置换）；通过RNAi、反义RNA或核酸酶等进行的基因弱化；以及使用大范围核酸酶、锌指核酸酶、TALEN和/或CRISPR技术进行的基因组工程化等。本文所述的突变体是通过人为干预产生的，并且因此不是天然存在的生物体。目标突变生物体典型地具有与缺少突变的相应野生型或祖代藻株不同的表型，其中可以通过生长测定、产物分析、光合特性、生化测定等评估表型。当提及基因时，“突变体”表示该基因相对于天然或野生型基因具有至少一个碱基（核苷酸）变化、缺失或插入。突变（一个或多个核苷酸的改变、缺失和/或插入）可以在基因的编码区域内，或者可以在内含子、3'UTR、5'UTR或启动子区域内，例如，在翻译起始位点的约1kb、约2kb或约3kb、约4kb或约5kb内。例如，如本文所公开的基因表达减弱的突变体可具有突变，其可为转录起始位点的基因5'区域中的一个或多个核碱基变化和/或一个或多个核碱基缺失和/或一个或多个核碱基插入，例如，在非限制性实例中，在已知或推定的转录起始位点的2kb内、1.5kb内、1kb内或0.5kb内，或在翻译起始位点的3kb内、2.5kb内、2kb内、1.5kb内、1kb内或0.5kb内。作为非限制性实例，突变基因可以是在启动子区域内具有可以增加或降低该基因的表达的突变、插入和/或缺失的基因；可以是具有导致产生非功能蛋白、截短蛋白、显性失活蛋白

和/或无蛋白的缺失的基因;和/或可以是具有一个或多个导致所编码蛋白质的氨基酸的变化或导致基因转录本的异常剪接的点突变的基因等。

[0055] 多肽的保守结构域包括在“cd”(保守结构域)数据库、COG数据库、pfam数据库、SMART数据库、PRK数据库、TIGRFAM数据库或本领域已知的其它数据库中鉴定的那些。由美国国立卫生研究院赞助的美国国家生物技术信息中心网站包括保守结构域数据库(CDD),其被描述为“蛋白质注释资源”,由古老结构域和全长蛋白质的一组经过良好注释的多个序列对准模型组成。这些可以作为位置特定的评分矩阵(PSSMs)得到,用于通过RPS-BLAST快速识别蛋白质序列中的保守结构域。CDD内容物包括NCBI定义的结构域,其使用3D结构信息来明确定义域边界并提供对序列/结构/功能关系的深入了解,以及从多个外部源数据库(Pfam、SMART、COG、PRK、TIGRFAM)导入的域模型。可以在序列中搜索保守结构域,例如,在NCBI的cdd数据库中。参见Marchler-Bauer等人(2015)《核酸研究(Nucleic Acids Res.)》43(D) 222-226。

[0056] 术语“Pfam”是指由Pfam联盟维护并且可在多个赞助的万维网站上获得的蛋白质结构域和蛋白质家族的大集合。Pfam的最新版本是基于UniProt蛋白质数据库版本2012_06的Pfam 30.0(2016年6月)。使用多个序列对准和隐马尔可夫模型(HMM)识别Pfam结构域和家族。Pfam-A家族或结构域分配是通过使用蛋白质家族的代表性成员和基于种子对准的剖面隐马尔可夫模型的策划种子对准产生的高质量分配。(除非另有说明,否则所查询的蛋白质与Pfam结构域或家族的匹配是Pfam-A匹配。)然后,将所有属于该家族的识别序列用于自动产生该家族的完全对准(Sonnhammer(1998)《核酸研究(Nucleic Acids Research)》26,320-322;Bateman(2000)《核酸研究(Nucleic Acids Research)》26,263-266;Bateman(2004)《核酸研究(Nucleic Acids Research)》32,数据库期刊,D138-D141;Finn(2006)《核酸研究(Nucleic Acids Research)》数据库期刊34,D247-251;Finn(2010)《核酸研究(Nucleic Acids Research)》数据库期刊38,D211-222)。通过使用任何相关网站(例如pfam.xfam.org)访问Pfam数据库,可以使用HMMER同源性搜索软件(例如HMMER2,HMMER3或更高版本,在线可得)针对HMM查询蛋白质序列。将查询的蛋白质标识为pfam家族(或具有特定的Pfam结构域)的重要匹配是二进制值大于或等于Pfam结构域的采集阈值的那些。期望值(e值)也可以用作在Pfam中包含查询的蛋白质或确定查询的蛋白质是否具有特定Pfam域的标准,其中低e值(远小于1.0,例如小于0.1或小于或等于0.01)表示匹配是由于偶然导致的低概率。

[0057] “cDNA”是包含mRNA分子的核苷酸序列的至少一部分的DNA分子,不同之处在于该DNA分子代替存在于mRNA序列中的核苷胸腺嘧啶或T代替尿苷或U。cDNA可以是双链的、部分双链的或单链的,并且可以是例如mRNA序列的互补序列。在优选的实例中,cDNA不包括出现在该cDNA所对应的天然存在的基因中的一个或多个内含子序列(即,出现在生物体的基因组中的基因)。例如,cDNA可以具有与天然存在的基因的内含子的下游序列并置的天然存在的基因的内含子的上游序列,其中上游和下游序列在自然DNA分子中不并置(即,这些序列不会在自然存在的基因中并置)。cDNA可以通过mRNA分子的逆转录产生,或者可以例如通过化学合成和/或通过使用一种或多种限制酶、一种或多种连接酶、一种或多种聚合酶(包括但不限于可用于聚合酶链反应(PCR)的耐高温聚合酶)、一种或多种重组酶等基于cDNA序列的知识合成,其中cDNA序列的知识可以任选基于从基因组序列中鉴定编码区或从序列中编

译出多个部分cDNA。

[0058] “对照细胞”或“对照微生物”是从其直接或间接衍生突变微生物(基因工程化或诱变微生物)的野生型细胞或微生物,或者是与提及的突变细胞或微生物基本相同的细胞或微生物(即,具有相同属和种,优选相同藻株),不同之处在于对照细胞或微生物不具有导致主题微生物具有的脂质产生增加的突变。例如,如果突变微生物的以下表达减弱:(1) 编码包括GAF(例如,GAF2)结构域的多肽的基因;(2) 编码具有GAF(例如GAF2)结构域的多肽的基因,其与SEQ ID NO:1具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%或至少约95%的同一性;或(3) 编码多肽的基因,其与SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%或至少约95%的同一性,对照细胞可以与突变微生物基本相同,不同之处在于对照微生物的根据1)、2)或3)的基因表达不减弱。

[0059] 如本文所用,“相同条件”或“相同培养条件”表示基本相同的条件,即,可能存在的参考条件之间的任何差异都是微小的,并且与对本发明而言是重要的微生物的功能或特性(包括脂质生产或生物质生产)无关。

[0060] 如本文所用,“脂质(lipid或lipids)”是指脂肪、蜡、脂肪酸、脂肪酸衍生物(如脂肪醇、蜡酯、烷烃和烯烃)、固醇、单甘油酯、甘油二酯、甘油三酯、磷脂、鞘脂、糖脂和甘油酯。“FAME脂质”或“FAME”是指具有可衍生为脂肪酸甲酯的酰基部分的脂质,例如单酰基甘油酯、二酰基甘油酯、三酰基甘油酯(TAG)、蜡酯和膜脂质如磷脂、半乳糖脂等。在一个实例中,脂质生产力可以以毫克/升(mg/L)计的FAME生产力来评估,并且对于藻类,可以报告为克/平方米/天(g/m²/天)。在一个实例中,本文提供的半连续测定可以通过考虑入射辐射面积将mg/L值转换为g/m²/天(例如,半连续过程测定(SCPA)烧瓶架孔可以是1 1/2”x 33/8”或0.003145m²),并且培养物的体积可以为550ml。为了获得以g/m²/天计的生产力值,可以将mg/L值乘以每日稀释率(30%)和转换系数0.175。除非另有说明,否则将脂质或其亚类(例如TAG或FAME)以百分比提及时,该百分比为重量百分比。

[0061] “生物质”是指无论是活细胞还是死细胞的细胞质量,并且可以例如使用本领域已知的方法以吸出的沉淀物重量来评估,但是更优选为干重(例如,培养样品或沉淀细胞的冻干物)、无灰干重(AFDW)或总有机碳(TOC)。在允许生长的条件下,在培养物的生长期间,生物质会增加,并且例如在分批培养中可称为“生物质积累”。在经历稳定或规则稀释的连续或半连续培养中,在培养物稀释期间会除去以其它方式积累在培养物中的产生的生物质。因此,这些培养物的每日生物质生产力(生物质的增加)也可以称为“生物质积累”。可以将生物质生产力评估为以毫克/升(mg/L)计的TOC生产力,并且对于藻类,可以报告为克/平方米/天(g/m²/天)。在本文提供的半连续测定中,通过考虑入射辐射面积(SCPA烧瓶架孔为1 1/2”x 33/8”或0.003145m²)和培养物的体积(550ml)将mg/L值转换为g/m²/天。为了获得以g/m²/天计的生产力值,将mg/L值乘以每日稀释率(30%)和转换系数0.175。除非另有说明,否则生物质以百分比表示时,该百分比是重量百分比。

[0062] 在本公开的上下文中,“氮源”是可以被主题微生物吸收并代谢并掺入生物分子中以进行生长和繁殖的氮源。例如,在本发明的上下文中,不能被微生物吸收和/或代谢以使其生长的包括氮的化合物(例如,含氮的生物缓冲液,例如Hepes,Tris等)不被视为氮源。

[0063] 如本文所用,“还原氮”是化学形式为铵盐、氨、尿素、酰胺或氨基酸(例如,可以被正在培养的微生物吸收并代谢以提供掺入生物分子的氮源从而支持生长的氨基酸)的氮。可以是氮源的氨基酸的实例可以包括但不限于谷氨酸、谷氨酰胺、组氨酸、脯氨酸、赖氨酸、精氨酸、天冬酰胺、丙氨酸和甘氨酸。在可以存在于微生物培养基中的氮源的上下文中,“非还原氮”是指在被微生物同化为有机化合物之前必须还原的硝酸盐或亚硝酸盐。

[0064] “(培养基中)唯一的氮源”与“基本上唯一的氮源”互换使用,并且表示没有其它可以被生物体代谢的氮源(即提供可以被生物体吸收并被生物体掺入生物分子中的氮源,例如蛋白质和核酸)被有意地添加到培养基中,或者没有其它可以被生物体利用的氮源以足以显著增加在参考培养基中培养的生物体或细胞的生长的量存在。在整个本申请中,为简洁起见,术语“仅硝酸盐”用于表征其中硝酸盐是生物体可用于支持生长的唯一氮源的培养基。

[0065] 类似地,“(在培养基中)唯一的碳源”与“基本上唯一的碳源”可互换使用,并且表示没有其它可被微生物代谢的碳源(即,用于能源或作为生物分子产生的碳源)以足以显著增加在参考培养基中培养的微生物或细胞的生产力、生长或繁殖的量存在,或者可以以大于掺入生物分子中的碳的5%的百分比掺入生物分子如由微生物或细胞产生的脂质中。

[0066] “富氮”条件是指其中不能通过向培养基中添加额外的氮(呈可被微生物使用的形式)来赋予进一步的生长或繁殖益处的培养基条件。类似地,“富营养物”条件是指其中没有营养物被限制生长或繁殖的培养基条件,也就是说,当培养基富含营养物时,向培养基中添加其它营养物不会导致生长或繁殖率的改善。在“富营养物”的上下文中,“营养物”包括,作为非限制性实例,磷酸盐、硫、铁和任选的二氧化硅,但不包括可被生物体用作能源的碳源,例如糖或有机酸。

[0067] 本文公开用于操纵、测定、培养和分析微生物的方法。本文阐述的本发明还利用本领域已知的用于细胞培养、微生物转化、基因工程和生化分析的标准方法、技术和试剂。尽管与本文描述的那些类似或等同的方法和材料可以用于本发明的实践或测试中,但是下面描述了合适的方法和材料。材料、方法和实例仅是说明性的,并不旨在是限制性的。根据详细描述和权利要求,本发明的其它特征和优点将显而易见。

[0068] 具有增加的脂质生产力的突变微生物

[0069] cGMP特异性磷酸二酯酶、腺苷酸环化酶和转录激活因子FhlA(“GAF”)结构域存在于来自多种微生物的多种蛋白质中,并以其中存在至少一个这些结构域的某些蛋白质命名。在一些实施例中,本公开提供突变微生物,其具有至少一种编码具有GAF结构域的多肽的基因的减弱和/或改变的表达和/或功能(包括但不限于此类基因的失活和/或缺失),和/或具有GAF结构域本身的多肽的减弱和/或改变的表达和/或功能。还包括突变微生物,其具有影响编码具有影响脂质产生的GAF结构域的多肽的基因的表达和/或功能的改变的基因或蛋白质的表达或功能。在各种实施例中,与不具有编码多肽和/或具有GAF结构域的多肽的基因的表达或功能的此类减弱和/或改变的对照微生物相比,此类突变微生物可产生更多的脂质和/或表现出碳与脂质分配的增加。GAF2结构域是GAF结构域家族的一个子集,典型地被表征为属于pfam PF13185。在一些实施例中,上面提及的突变微生物的GAF结构域是GAF2结构域。在一些实施例中,编码GAF2结构域的基因定位于卡迪斯微拟球藻(*Nannochloropsis gaditana*)第14号染色体上的Naga_100020g79基因座,或另一物种中的

同线基因座。在一些实施例中,编码包括GAF2结构域的多肽的基因包括GAF2结构域,其具有与SEQ ID NO:1具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列。在其它实施例中,突变微生物的编码具有GAF(例如GAF2)结构域的多肽的基因表达减弱,所述多肽包括SEQ ID NO:1所列的氨基酸序列或其保守变体。在其它实施例中,突变微生物的编码具有GAF结构域(例如GAF2结构域)的多肽的基因表达减弱,所述多肽具有SEQ ID NO:1所列的氨基酸序列。编码GAF2结构域的基因可以编码包括GAF2结构域并与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的多肽。

[0070] 因此,在一些实施例中,突变微生物具有编码具有GAF2结构域的多肽的基因的减弱和/或改变的表达和/或功能。在一些实施例中,GAF2结构域可以被表征为pfam PF13185,其二进制值大于20.0或大于28.0(聚集截止)并且e值小于0.01或小于0.001。

[0071] 在一些实施例中,与对照微生物相比,本文提供的产生更多脂质和/或表现出增加的碳与脂质分配的突变微生物可以具有编码多肽的基因的减弱和/或改变的表达和/或功能,所述多肽包含与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列。在其它实施例中,突变微生物的编码包括SEQ ID NO:2所列氨基酸序列或其保守变体的多肽的基因表达减弱。在一些实施例中,突变微生物的编码具有SEQ ID NO:2所列氨基酸序列的多肽的基因表达减弱。

[0072] 在一些实施例中,当突变微生物和对照微生物都在基本相同的条件下培养时,相对于对照微生物,本文提供的突变微生物(例如,通过传统诱变或基因工程化获得的微生物)多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的脂质。例如,当突变微生物和对照微生物都在其中对照微生物培养物产生生物质的基本相同的条件下培养时,相对于对照微生物,突变微生物可以产生多约25%至约250%、多约25%至约225%、多约25%至约200%、多约25%至约175%、多约25%至约150%、多约25%至约125%、多约50%至约250%、多约50%至约225%、多约50%至约200%、多约50%至约175%、多约50%至约150%、多约50%至约125%、多约75%至约250%、多约75%至约225%、多约75%至约200%、或多约75%至约175%、多约75%至多约150%或多约75%至约125%(例如多25-250%)的脂质。相对于对照微生物,培养条件可以是富氮的,并且可以是富营养物的。在一些实施例中,对照微生物是从其直接或间接衍生突变体的相同物种的野生型微生物,并且相对于野生型微生物,培养条件是富氮的,并且可以是富营养物的。

[0073] 与对照微生物相比,本文提供的突变体产生更多脂质或显示更大的碳与脂质分配的培养条件可以是分批、半连续或连续条件,所述条件可以是富氮的、氮受限的(例如,其中氮源以小于约5mM、小于约4mM、小于约3mM、小于约2mM、小于约1mM或小于约0.5mM的浓度存在)或贫氮的(基本上不含氮源)。在一些实施例中,编码包括GAF2结构域的多肽或与SEQ ID

N0:2具有至少50%、60%、70%、80%、85%、90%或95%同一性的基因表达减弱的突变微生物产生比对照藻株更多脂质的条件是其中硝酸盐作为培养基中的基本上唯一氮源存在的条件。

[0074] 在一些实施例中,例如在其中硝酸盐是基本上唯一氮源并且突变微生物和对照微生物两者都产生生物质的条件下,本文所公开的突变微生物产生比对照微生物多至少约15%、至少20%、至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%或至少约120%的贮存脂质,例如三酰基甘油酯(TAG)。

[0075] 在一些实施例中,当突变微生物和对照微生物在支持对照微生物生长和繁殖的基本上相同的条件下,即在对照微生物培养物产生生物质的条件下培养时,本文提供的突变微生物可在至少约三天的培养期内,例如在至少约四天、至少约五天、至少约六天、至少约七天、至少约八天、至少约九天、至少约十天、至少约十一天、至少约十二天、至少约十三天、至少约十四天、至少约十五天、至少约二十天、至少约三十天或至少约六十天的培养期内,显示比对照微生物更大的脂质生产力。在一些实例中,其中本文提供的突变微生物相对于对照微生物多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的脂质的培养期可以是少于180天、少于120天或少于90天,其中突变体可在该时期内具有更高的平均每日脂质生产力。例如,本文提供的突变微生物在三至90天、三至60天、三至三十天或三至十五天的培养期期间可比对照微生物多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的脂质。例如,本文提供的突变微生物在约五至约90天、约五至约60天、约五至约三十天、或约五至约十五天、或约七至约90天、约七至约60天、约七至约三十天、约七至约二十天或约七至约十五天的培养期期间可比对照微生物多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的脂质。

[0076] 微生物产生的脂质的量可以通过在培养期期间的任何点,例如在培养期间结束时或在培养期期间的间隔,如每天、每隔一天等,取出培养样品来测定。生产力可以是体积生产力,例如,培养物的生产力可以表示为每毫升或每升培养物的重量,并且可以是每日生产力(例如,mg/升/天或g/升/天),例如,历经多天培养(例如,至少约三天、四天、五天、六天、七天、八天、九天、十天、十一天、十二天、十三天、十四天、十五天或更多天)的平均每日生产力,或者可以是在培养物中在限定的时间段内每单位体积产生的总量。优选在培养期期间多次测量生产力,例如,至少约两次或至少约三次,并且可以每天、每隔一天、每三天等进行

评估。

[0077] 生物质生产力可以例如通过测量总有机碳(TOC)或通过其它方法如测量干重或无灰干重(AFDW)来评估。测量TOC的方法是本领域已知的(例如美国专利第8,835,149号)并在本文提供。测量AFDW的方法也是公知的,并且可以在例如通过引用整体并入本文的美国专利第8,940,508号中找到。

[0078] 测量由微生物产生的脂质的量的方法也是本领域公知的,并在本文实例中提供。例如,总的可提取脂质可以根据Folch等人(1957)《生物化学杂志(J.Biol.Chem.)》226:497-509;Bligh&Dyer(1959)《生理生化杂志(Can.J.Biochem.Physiol.)》37:911-917;或Matyash等人(2008)《脂质研究杂志(J.Lipid Res.)》49:1137-1146确定,例如,并且还可以使用傅立叶变换红外光谱(FT-IR)评估以脂质存在的生物质的百分比(Pistorius等人(2008)《生物技术和生物工程(Biotechnol&Bioengin.)》103:123-129)。例如,美国专利第8,207,363号和W02011127118中提供了FAME和TAG的重量分析的其它参考文献,将其各自通过引用整体并入本文。

[0079] 在一些实施例中,当在相同的培养条件下培养突变微生物和对照微生物时,相对于对照微生物,本文提供的突变微生物在培养期长度内或每天多产生平均至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的FAME脂质,其中相对于对照微生物,培养条件在至少约三天、至少约四天、至少约五天、至少约七天、至少约十天、至少约十二天或至少约十四天内是富氮的,并且优选地是富营养物的培养条件,并且可以是其中突变微生物和对照微生物都产生生物质的培养条件。

[0080] 培养条件可包括在培养基中培养,所述培养基包括少于约5mM、少于约4.5mM、少于约4mM、少于约3.5mM、少于约3mM、少于约2.5mM、少于约2mM、少于约1.5mM、少于约1mM、少于约0.5mM(例如,少于5mM)的或基本上不含还原氮源,如铵。例如,铵浓度可以为约0至约5mM、约0至约4.5mM、约0至约4.0mM、约0至约3.5mM、约0至约3mM、约0至约2.5mM、约0至约2.0mM、约0至约1.5mM、约0至约1.0mM或约0至约0.5mM(例如0-5mM)的浓度。铵浓度可以是约0.2至约3mM、0.2至约2.5mM、约0.2至约2mM、约0.2至约1.5mM、约0.2至约1mM、约0.3至约2.5mM、约0.3至约2mM、约0.3至约1.5mM或约0.3至约1mM(例如0.2-3mM)的浓度。在进一步的实例中,铵浓度可以是约0.4mM至约2.5mM、约0.4mM至约2mM或约0.4mM至约1.5mM(例如0.4-2.5mM)的浓度。可替代地或另外,培养条件可包括在包括硝酸盐作为基本上唯一氮源的培养基中培养。在一些实例中,对照微生物是野生型微生物,例如,从其直接或间接地衍生突变微生物的野生型微生物。

[0081] 在一些实施例中,本文提供的突变微生物是当在包括少于约5mM、少于约4.5mM、少于约4mM、少于约3.5mM、少于约3mM、少于约2.5mM铵、少于约2.0mM、少于约1.5mM铵、少于约1.0mM铵、少于约0.5mM铵(例如少于5mM)或基本上不含铵并且包括例如至少约1.0mM、至少约2.0mM、至少约3.0mM、至少约4.0mM、至少约5.0mM、至少约6.0mM、至少约7.0mM、至少约8.0mM、至少约9.0mM或至少约10.0mM硝酸盐(例如至少1.0mM)的培养基中在光合自养条件下培养时,比对照藻类多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少

50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%的FAME脂质的藻类微生物。例如,铵浓度可以为约0至约5mM、约0至约4.5mM、约0至约4.0mM、约0至约3.5mM、约0至约3mM、约0至约2.5mM、约0至约2.0mM、约0至约1.5mM、约0至约1.0mM或约0至约0.5mM(例如0-5mM)的浓度。铵浓度可以是约0.2至约3mM、0.2至约2.5mM、约0.2至约2mM、约0.2至约1.5mM、约0.2至约1mM、约0.3至约2.5mM、约0.3至约2mM、约0.3至约1.5mM或约0.3至约1mM(例如0.2-3mM)的浓度。在进一步的实例中,铵浓度可以是约0.4mM至约2.5mM、约0.4mM至约2mM或约0.4mM至约1.5mM(例如0.4-2.5mM)的浓度。在一些实例中,培养条件可以包括基本上没有铵,并且在一些实例中,可以包括基本上没有还原氮作为氮源。在一些实例中,培养物包括硝酸盐作为氮源,其可以任选地基本上是培养基中的唯一氮源。光合自养条件可以处于昼夜循环。昼夜循环的光周期可以是任何长度,并且可以是例如约四小时至约二十二小时,并且可以是例如约六小时至约二十小时,例如约八小时至约十八小时/二十四小时循环。微生物可以暴露于自然光或人造光或其组合。可用光的强度可在整个光周期内变化。

[0082] 在其中对照微生物和突变微生物都产生生物质的相同条件下,相对于培养的对照微生物,本文提供的突变微生物可具有更大的碳与脂质分配。相对于对照微生物具有增加的碳与脂质分配的突变体可具有增加的碳与总可提取脂质、与总中性脂质、与甘油三酯和/或与FAME可衍生的脂质的分配。在一些实例中,本文提供的突变微生物可具有比对照微生物高至少约50%、至少60%、至少70%、至少80%、至少90%、至少100%、至少120%、至少140%、至少160%、至少180%、至少200%、至少220%、至少240%、至少260%、至少280%、至少300%、至少320%、至少340%、至少360%、至少380%或至少400%的所产生的FAME可衍生脂质(“FAME”)的量与所产生的生物质(例如TOC或无灰干重(AFDW))之比。例如,当突变微生物和对照微生物都在其中对照微生物培养物产生生物质的基本相同的条件下培养时,相对于对照微生物,突变微生物的所产生的FAME可衍生脂质(“FAME”)的量与所产生的生物质(例如TOC或无灰干重(AFDW))之比可以具有高约50%至高约400%、高约50%至高约375%、高约50%至高约350%、高约50%至高约325%、高约50%至高约300%、高约50%至高约275%、高约50%至高约250%、高约50%至高约225%、高约75%至高约400%、高约75%至高约375%、高约75%至高约350%、高约75%至高约325%、高约75%至高约300%、高约75%至高约275%、高约75%至高约250%或高约75%至高约225%(例如高50-400%)的脂质生产力。脂质和生物质的产生和/或生产可以例如通过重量分析来评估,如本领域已知的和本文实例中所显示的。例如,当突变微生物和对照微生物都在相同条件下培养时,本文提供的突变微生物的FAME与TOC之比可比对照微生物的FAME/TOC比高至少约50%、至少60%、至少70%、至少80%、至少90%、至少100%、至少120%、至少140%、至少160%、至少180%、至少200%、至少220%、至少240%、至少260%、至少280%、至少300%、至少320%、至少340%、至少360%、至少380%或至少约400%。在一些实施例中,突变微生物中的FAME/TOC比相对于对照微生物高约50-400%。

[0083] 在各种实例中,当在富氮条件下,例如富营养物的条件下培养时,本文提供的突变微生物的FAME/TOC比相对于对照微生物可以是例如至少约0.30、至少约0.35、至少约0.40、

至少约0.45、至少约0.50、至少约0.55、至少约0.60、至少约0.65、至少约0.70、至少约0.75、至少约0.80、至少约0.85或至少约0.90(例如0.30、0.50、0.75或0.90中的至少任一个)。培养条件可以包括例如包括小于约5mM、小于约4.5mM、小于约4mM、小于约3.5mM、小于约3mM、小于约2.5mM、小于约2mM、小于约1.5mM、小于约1.0mM或小于约0.5mM(例如,小于5mM)的铵,并且在一些实例中可以包括至少约1.0mM、至少约2.0mM、至少约3.0mM、至少约4.0mM、至少约5.0mM、至少约6.0mM、至少约7.0mM、至少约8.0mM、至少约9.0mM或至少约10.0mM(例如,至少1.0mM)的硝酸盐的培养基。在一些实例中,培养条件可以包括基本上没有铵,并且在一些实例中,可以包括基本上没有还原氮作为氮源。在一些实例中,培养物包括硝酸盐作为氮源,其可以任选地基本上是培养基中的唯一氮源。在一个说明性实施例中,突变微生物表现出至少约0.35的FAME/TOC比。

[0084] 将本文提供的具有增加的脂质产生的突变体的特性与对照微生物的相同特性进行比较,所述对照微生物可以是与所述突变体相同物种的野生型生物体,优选地是过量产生脂质的突变体的祖代藻株。可替换地,对照微生物可以是与突变微生物基本上相同的微生物,除了对照微生物不具有导致更高脂质生产力的突变以外。例如,对照微生物可以是基因工程化的微生物或传统突变的生物体,其已被进一步突变或工程化以产生如本文所公开的具有增加的脂质生产力和/或增加的脂质分配的突变体。

[0085] 在一些实例中,对照微生物可以是与突变微生物基本上相同的微生物,除了对照微生物在调节脂质诱导的基因(即,编码具有GAF2结构域的多肽的基因、编码包括与SEQ ID NO:1具有至少50%同一性的氨基酸序列的多肽的基因、和/或编码与SEQ ID NO:2具有至少50%同一性的多肽的基因,其突变导致增加的脂质产生)中不具有突变以外。还将具有破坏、减弱或其它直接或间接遗传操作的基因(导致脂质诱导调节基因的结构或表达改变)的过量产生脂质的突变体的性质与不具有破坏、减弱或其它直接或间接遗传操作的脂质诱导调节基因导致脂质诱导调节基因的结构或表达改变的对照细胞的相同性质进行比较(无论细胞是否是“野生型”)。例如,对照细胞可以是包括一种或多种非天然基因的重组细胞或在正在评估其效果的除脂质诱导调节基因以外的基因中突变的细胞等。

[0086] 突变微生物可以是任何真核微藻株,例如任何以下属的任何物种:曲壳藻属(Achnanthes)、茧形藻属(Amphiprora)、双眉藻属(Amphora)、纤维藻属(Ankistrodesmus)、星胞藻属(Asteromonas)、黄金色藻属(Boekelovia)、波利氏藻属(Bolidomonas)、包特氏菌属(Borodinella)、气球藻属(Botrydium)、葡萄藻属(Botryococcus)、片球藻属(Bracteococcus)、角毛藻属(Chaetoceros)、四鞭藻属(Carteria)、衣藻属(Chlamydomonas)、绿球藻属(Chlorococcum)、绿梭藻属(Chlorogonium)、小球藻属(Chlorella)、蓝隐藻属(Chroomonas)、金球藻属(Chrysosphaera)、球钙板藻属(Cricosphaera)、隐甲藻属(Cryptothecodinium)、隐藻属(Cryptomonas)、小环藻属(Cyclotella)、链带藻属(Desmodesmus)、杜氏藻属(Dunaliella)、后棘藻属(Elipsoïdon)、圆石藻属(Emiliania)、独球藻属(Eremosphaera)、衣迪斯藻属(Ernodesmus)、眼虫藻属(Euglena)、真眼点藻属(Eustigmatos)、披刺藻属(Franceia)、脆杆藻属(Fragilaria)、拟脆杆藻属(Fragilaropsis)、丽丝藻属(Gloeothamnion)、红球藻属(Haematococcus)、菱板藻属(Hantzschia)、异弯藻属(Heterosigma)、膜胞藻属(Hymenomonas)、等鞭金藻属(Isochrysis)、鳞孔藻属(Lepocinclis)、微芒藻属(Micractinium)、蒜头藻属(Monodus)、

单针藻属(Monoraphidium)、微球藻属(Nannochloris)、微拟球藻属(Nannochloropsis)、舟形藻属(Navicula)、新绿藻属(Neochloris)、肾鞭藻属(Nephrochloris)、肾藻属(Nephroselmis)、菱形藻属(Nitzschia)、棕鞭藻属(Ochromonas)、鞘藻属(Oedogonium)、卵囊藻属(Oocystis)、蚝球藻属(Ostreococcus)、拟小球藻属(Parachlorella)、雪衣藻属(Parietochloris)、杜氏亚属盐藻属(Pascheria)、巴夫藻属(Pavlova)、普莱格门(Pelagomonas)、褐指藻属(*Phaeodactylum*)、噬菌体属(Phagus)、皮级微绿藻属(Picochlorum)、扁藻属(Platymonas)、颗石藻属(Pleurochrysis)、肋球藻属(Pleurococcus)、原壁菌属(Prototheca)、拟绿球藻属(Pseudochlorella)、拟新绿藻属(Pseudoneochloris)、拟角星鼓藻属(Pseudostaurastrum)、塔胞藻属(Pyramimonas)、桑椹藻属(Pyrobotrys)、栅列藻属(Scenedesmus)、裂衣藻属(Schizochlamydeella)、骨条藻属(Skeletonema)、螺旋藻属(Spyrogyra)、裂丝藻属(Stichococcus)、四球藻属(Tetrachlorella)、四月藻属(Tetraselmis)、海链藻属(Thalassiosira)、黄丝藻属(Tribonema)、无隔藻属(Vaucheria)、鲜绿球藻属(Viridiella)、魏氏藻属(Vischeria)和团藻属(Volvox)。特别合适的物种的非限制性实例包括例如硅藻,例如以下属中的任一种的物种:双眉藻属、角毛藻属、小环藻属、脆杆藻属、拟脆杆藻属、菱板藻属、蒜头藻属、舟形藻属、菱形藻属、褐指藻属或海链藻属,或者黄绿藻,例如真眼点藻属、微拟球藻属、拟角星鼓藻属或魏氏藻属。

[0087] 在一些实例中,重组藻类是绿藻,即绿色植物界绿藻门的藻类成员,包括但不限于以下纲的任一种的微藻:绿藻纲(Chlorophyceae)、四片藻纲(Chlorodendrophyceae)、横裂甲藻纲(Pedinophyceae)、真绿藻纲(Pleurostrophyceae)、绿枝藻纲(Prasinophyceae)和共球藻纲(Trebouxiophyceae)。在一些实例中,本文提供的重组藻类可以是本身为绿藻纲、绿枝藻纲、共球藻纲或四片藻纲中任一种的成员的物种,例如以下的任一种的物种:星胞藻属、纤维藻属、四鞭藻属、衣藻属、绿球藻属、绿梭藻属、金球藻属、链带藻属、杜氏藻属、红球藻属、单针藻属、新绿藻属、鞘藻属、普莱格门、肋球藻属、桑椹藻属、栅列藻属、团藻属、微胞藻属、蚝球藻属、绿枝藻属、谢氏菌属、四月藻属、葡萄藻属、小球藻属、独球藻属、披刺藻属、微芒藻属、微球藻属、卵囊藻属、拟小球藻属、皮级微绿藻属、原壁菌属或拟绿球藻属。在各种实例中,本文提供的重组藻类可以是共球藻纲的物种或藻种,例如但不限于葡萄藻属、小球藻属、独球藻属、披刺藻属、微芒藻属、微球藻属、卵囊藻属、拟小球藻属、皮级微绿藻属、原壁菌属或拟绿球藻属。

[0088] 在一些实例中,重组藻类是长短鞭毛体藻类,并且可以属于硅藻(硅藻门)、黄绿藻、黄藻、褐藻、金藻或针胞藻。在一些实例中,突变藻类属于硅藻门或黄绿藻属,例如但不限于以下属:茧形藻属、双眉藻属、角毛藻属、小环藻属、脆杆藻属、拟脆杆藻属、菱板藻属、蒜头藻属、微拟球藻属、舟形藻属、菱形藻属、褐指藻属、褐指藻属、拟角星鼓藻属、魏氏藻属、褐指藻属、骨条藻属或海链藻属。在一些实例中,突变藻类是黄绿藻,并且属于选自以下组成的组的属:小绿藻属、绿囊藻属、椭球藻属、真眼点藻属、角绿藻属、拟单胞藻属、蒜头藻属、微拟球藻属、假拟小椿藻属、拟角星鼓藻属、假四角藻属和魏氏藻属。在一些实例中,突变藻类细胞是微拟球藻属物种。

[0089] 可替代地,本文提供的突变微生物可以是本身为海藻网粘菌纲微生物的长短鞭毛体,例如Labrinthulid或破囊壶菌属的成员,例如以下属中任一种的物种:网粘菌属

(Labryinthula)、类网粘菌属(Labryinthuloides)、破囊壶菌属(Thraustochytrium)、裂殖壶菌属(Schizochytrium)、不动壶菌属(Aplanochytrium)、橙黄壶菌属(Aurantiochytrium)、椭圆壶菌属(Oblongichytrium)、日本壶菌属(Japonochytrium)、戴普罗弗属(Diplophrys)和吾肯氏壶菌属(Ulkenia)。

[0090] 突变体可以是自发突变体、传统衍生突变体或调节基因表达减弱的工程化突变体,所述调节基因例如其表达影响多种其它基因如编码转录因子或转录激活因子的基因的表达的基因。

[0091] 调节脂质产生的基因表达减弱的突变微生物可以是“敲除”突变体,例如,其中多肽的阅读框被破坏,使得不产生功能性蛋白质。例如,基因可以在阅读框中包括插入、缺失或突变,其导致不产生功能性蛋白质。在各种实例中,敲除突变可以通过将序列插入基因中,例如插入基因的编码区中来产生,所述序列通常但不必须包括选择性标记基因。此类插入可以通过使用将供体片段整合到靶基因座中的cas/CRISPR系统来进行,或者可以通过例如同源重组来进行。此类插入可以破坏开放阅读框和/或剪接信号,或产生无功能的融合蛋白或截短的蛋白。在其它实例中,突变微生物可以是“敲低”突变体,其中基因的表达相对于野生型细胞的表达水平降低但未消除,例如降低5%或更低至95%或更高,例如5%至95%或10%至90%。敲低可以是其中突变、插入或缺失发生在基因的非编码区如基因的5'或3'区的突变体,或者可以通过在细胞中表达降低靶基因表达的构建体如RNAi、核酶或反义构建体来实现。除CRISPR系统之外,同源重组、转座元件或随机整合可以用于产生插入突变体(敲低或敲除)。

[0092] 本文提供的突变微生物可以通过靶向目标微生物的内源基因来设计,所述内源基因编码包括本文公开的GAF(例如GAF2)结构域的多肽。可以使用生物信息学方法、分子生物学技术及其组合在目标微生物中鉴定此类基因。例如,编码包括GAF(例如GAF2)结构域的多肽的基因可以使用印迹杂交、通过杂交筛选cDNA文库或PCR,例如使用简并探针和/或引物来鉴定。在公共或专有数据库中可获得的基因组序列可以通过进行序列匹配(例如,blast程序,如blastp、blastn和tblastn(针对翻译的核苷酸序列查询的蛋白质序列))或分析编码的氨基酸序列的结构域结构的多种程序中的任一种来检索。例如,HMMER提供用于分析由基因编码的结构和功能结构域的在线软件,所述软件可用于扫描基因组序列,包括例如hmmsearch和hmmscan。可以在线进行此类检索。通过构建系统发育树以确定蛋白质之间的关系,程序如MUSCLE和halign也可用于检索蛋白质如本文公开的蛋白质(例如,含GAF(例如,GAF2)结构域的多肽)的直系同源物。基因靶向可利用在目标微生物基因组中鉴定的序列。不必解析基因的完整结构以靶向基因进行减弱。例如,使用本文公开的方法,包括但不限于基因组编辑(使用大范围核酸酶、锌指核酸酶、TALEN或Cas/CRISPR系统)、RNAi构建体、反义构建体、同源重组构建体和核酶构建体,仅基因序列的一部分可以用于基因减弱构建体和技术。

[0093] 在一些实施例中,突变微生物可被进一步工程化或诱变以具有至少一种额外的基因修饰,所述基因修饰赋予除草剂抗性、毒素抗性、增强的生长特性、增强的光合效率、增强的脂质产生或积累以及特定脂质的产生。

[0094] 基因减弱

[0095] 本文提供的调节脂质生物合成的基因表达减弱的突变微生物是通过人为干预,例

如通过传统诱变或基因工程化产生的突变体。用于产生微生物藻株的突变体的方法是本领域众所周知的。例如,本文提供的突变微生物可以通过任何可行的诱变方法产生的突变体,所述诱变方法包括但不限于UV照射、 γ 照射或化学诱变。筛选具有增加的脂质产生的突变体可以例如通过用如本领域已知的亲脂性染料如尼罗红或氟硼吡咯(BODIPY)染色(例如,Cabanelas等人(2015)《生物资源技术(Bioresource Technology)》184:47-52),或通过使用本领域已知的或本文公开的分析生物化学方法定量由藻株产生的脂质进行。

[0096] 在一个方面,本公开提供基因修饰的生物体,例如具有一种或多种减弱天然存在的脂质调节基因表达的基因修饰或突变的微生物,所述天然存在的脂质调节基因为例如编码具有GAF结构域(例如GAF2结构域)的多肽的天然存在的基因,所述基因在野生型生物体中与SEQ ID NO:1具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性;编码多肽的天然存在的基因,所述基因在野生型生物体中与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性;位于长短鞭毛体或藻类物种Naga_100020g79基因座或同线基因座的天然存在的基因;和/或具有编码区的天然存在的基因,所述基因与SEQ ID NO:3具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性。如本文所用,“减弱”或“改变”基因(例如,“脂质调节基因”)的表达和/或功能是指以减少正常表达的全功能蛋白质的产生、表达和/或功能的任何方式降低或消除基因的表达。用于减弱基因如脂质调节基因的手段包括例如同源重组构建体;CRISPR系统,包括引导RNA、Cas9或其它Cas酶,以及任选地,用于插入靶位点的供体片段;RNAi构建体,包括shRNA、反义RNA构建体;核酶构建体;TALENs,锌指核酸酶;和大范围核酸酶。例如,在一些实施例中,可以通过例如同源重组和/或通过双链断裂诱导剂的活性介导的插入或基因置换来破坏基因,所述双链断裂诱导剂为例如大范围核酸酶(参见例如WO 2012/017329 (US 2013/0164850) 和US2016/0272980)、锌指核酸酶(Perez-Pinera等人(2012)《化学生物学现状(Curr.Opin.Chem.Biol.)》16:268-277;WO 2012/017329 (US 2013/0164850);和US2012/0324603)、TALEN(WO 2014/207043 (US 2016/0130599);WO 2014/076571 (US 2016/0272980))或CRISPR系统的cas蛋白(例如Cas9蛋白)(参见例如US 8,697,359;US 8,795,965;US 8,889,356;US 2016/0304893;US 2016/0090603;US 2014/0068797)。本领域普通技术人员将理解,其它破坏方法是本领域已知的并且在此将是合适的。

[0097] 在一些实施例中,突变微生物具有一种或多种突变,所述突变针对位于长短鞭毛体或藻类物种Naga_100020g79基因座或同线基因座的基因或影响其表达。在一些实施例中,突变微生物具有针对具有开放阅读框的基因或影响具有开放阅读框的基因的表达的突变,所述开放阅读框与SEQ ID NO:3具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性。在一些实施例中,突变微生物具有一种或多种突变,其存在于或影响以下的表达:编码包含与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、

至少97%、至少98%或至少约99%同一性的氨基酸序列的多肽的核酸；和/或具有包含与SEQ ID NO:3具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的核苷酸序列的开放阅读框的核酸。在一些实施例中，突变微生物在Naga_100020g79基因座或与其同线的染色体基因座中的基因（本文中称为STP-4551基因或其同源物）的GAF2结构域中具有一种或多种突变。

[0098] 工程化为脂质调节基因表达减弱的重组微生物可具有被破坏的脂质调节基因，其包括至少一个降低或消除所述基因表达的插入、突变或缺失，使得不产生全功能脂质调节基因或以比不包括被破坏的脂质调节基因的对照微生物产生的量低的量产生全功能脂质调节基因。例如，在一些实施例中，一种或多种突变（一个或多个核苷酸的改变、缺失和/或插入）可以在基因的编码区中，或者可以在基因的转录或非转录部分中的内含子、3'UTR、5'UTR或启动子区中，例如，在转录起始位点的约2kb内或在翻译起始位点的约3kb内。在一些实施例中，例如，如本文所公开的基因表达减弱的突变微生物可具有一种或多种突变，其可为转录起始位点的基因5'区域中的一个或多个核碱基变化和/或一个或多个核碱基缺失和/或一个或多个核碱基插入，例如，在非限制性实例中，在已知或推定的转录起始位点的约2kb内、1.5kb内、1kb内或0.5kb内，或在翻译起始位点的约3kb内、2.5kb内、2kb内、1.5kb内、1kb内或约0.5kb内。作为非限制性实例，突变基因可以是在启动子区内具有增加或降低基因的表达的突变、插入或缺失的基因；可以是具有导致产生无功能蛋白质、截短蛋白质、显性失活蛋白质或无蛋白质的缺失的基因；可以是一个或多个导致编码的蛋白质的氨基酸改变或导致基因转录物的异常剪接等的点突变的基因。

[0099] 本文提供的多产生至少约25%、至少30%、至少35%、至少40%、至少45%、至少50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少100%、至少110%、至少120%、至少130%、至少140%、至少150%、至少160%、至少170%、至少180%、至少190%、至少200%、至少210%、至少220%、至少230%、至少240%或至少约250%脂质的突变体也可以是基因工程化突变体，例如，其中编码具有GAF结构域（例如GAF2结构域）的多肽的基因，或位于Naga_100020g79基因座或其直系同源物的基因（例如，编码具有GAF结构域的多肽的基因，其与SEQ ID NO:1具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性，或者与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性的编码多肽的基因）已经通过同源重组被靶向用于敲除、敲低或基因置换（例如，具有可以编码相对于野生型多肽具有降低的活性的多肽的基因的突变形式）的突变体。例如，可以通过定点定向同源重组工程化目标微生物藻株，以将序列插入基因组位点中，并且从而改变基因和/或其表达，或将启动子插入宿主微生物的遗传基因座中以影响特定基因或一组基因在该位点的表达。

[0100] 例如，通过同源重组进行的基因敲除、基因敲低或基因置换可以通过转化核酸（例如，DNA）片段来进行，所述核酸（例如，DNA）片段包括与要改变的基因组区域同源的序列，其中同源序列被外源序列中断，所述外源序列通常是允许选择整合的构建体的选择标记基因。在外源序列或突变的基因序列的任一侧上的基因组同源侧翼序列的长度可以是例如至

少约50、至少100、至少200、至少300、至少400、至少500、至少600、至少700、至少800、至少900、至少1,000、至少1,200、至少1,500、至少1,750或至少约2,000个核苷酸。可以在载体中提供基因敲除或基因“敲入”构建体,其中外源序列侧接有靶基因序列,所述构建体可以任选地线性化,例如在经历同源重组的区域之外,或者可以在不在载体背景下的线性片段提供,例如,敲除或敲入构建体可以是分离的或合成的片段,包括但不限于PCR产物。在一些情况下,分裂标记系统可用于通过同源重组产生基因敲除,其中可引入两个DNA片段,所述片段可再生选择标志物并通过三个交叉事件破坏目标基因座(Jeong等人(2007)《联邦医学院微生物学快报(FEMS Microbiol Lett)》273:157-163)。

[0101] 本文所提及并且最近由Hsu等人(《细胞(Cell)》157:1262-1278,2014)评论的CRISPR系统,除cas核酸酶多肽或复合物外,还包括通过与靶位点序列互补与基因组靶位点相互作用的靶向RNA(通常表示为“crRNA”),与cas多肽复合的反式激活(“tracr”)RNA,并且还包括(通过互补)结合靶向crRNA的区域。本公开设想使用可以共转化为宿主藻株(或在宿主藻株中表达)的两个RNA分子(“crRNA”和“tracrRNA”),所述宿主藻株表达或被cas蛋白转染以用于基因组编辑,或使用包括与靶序列互补的序列以及与cas蛋白相互作用的序列的单个引导RNA。也就是说,在一些策略中,本文使用的CRISPR系统可包含两个单独的RNA分子(RNA多核苷酸:“tracr-RNA”和“targeter-RNA”或“crRNA”,见下文),并且在本文中称为“靶向双分子DNA的RNA”或“靶向两分子DNA的RNA”。可替代地,如实例中所示,靶向DNA的RNA还可以包括与cas蛋白相互作用的反式激活序列(除了靶向同源(“cr”)序列之外),即靶向DNA的RNA可以是单个RNA分子(单个RNA多核苷酸),并且在本文中称为“嵌合引导RNA”、“单引导RNA”或“sgRNA”。术语“靶向DNA的RNA”和“gRNA”是包含性的,均指靶向双分子DNA的RNA和靶向单分子DNA的RNA(即sgRNA)。单分子引导RNA和两个RNA系统均已在文献中和例如在通过引用整体并入本文的美国专利申请公开第US 2014/0068797号中进行了详细描述。本文提出的方法和组合物的一些实施例包括引导RNA,其具有与位于Naga_100020g79基因座的基因中的靶序列相对应的序列,例如具有SEQ ID NO:5所示序列的靶序列。在一些实施例中,引导RNA是嵌合引导。在其它实施例中,引导RNA不包括tracr序列。任何cas蛋白均可用于本文的方法,例如,Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9(也称为Csn1和Csx12)、Cas10、Cbf1、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Cpf1、Csb1、Csb2、Csb3、Csx17、Csx14、Csx10、Csx16、CsaX、Csx3、Csx1、Csx15、Csf1、Csf2、Csf3、Csf4、C2c1、C2c2、C2c3及其同源物或其修饰形式。作为非限制性实例,cas蛋白可以是Cas9蛋白,例如酿脓葡萄球菌(*Staphylococcus pyogenes*)、嗜热链球菌(*S. thermophilus*)、肺炎链球菌(*S. pneumonia*)、金黄色葡萄球菌(*S. aureus*)或脑膜炎奈瑟氏球菌(*Neisseria meningitidis*)的Cas9蛋白。还考虑了在通过引用整体并入本文的美国专利申请公开第US 2014/0068797号中以SEQ ID NO:1-256和795-1346提供的Cas9蛋白,和可以结合来自一个以上Cas9蛋白的结构域以及已鉴定出的Cas9蛋白的变体和突变体的嵌合Cas9蛋白。

[0102] Cas核酸酶活性切割靶DNA以产生双链断裂。这些断裂然后由细胞以两种方式中的一种修复:非同源末端连接或同源定向修复。在非同源末端连接(NHEJ)中,通过将断裂末端彼此直接连接来修复双链断裂。在这种情况下,没有新的核酸材料插入该位点,尽管某些核酸材料可能丢失,导致缺失或改变,通常导致突变。在同源定向修复中,将可与切割的靶DNA

序列具有同源性的供体多核苷酸(有时称为“供体DNA”或“编辑DNA”)用作修复切割的靶DNA序列的模板,导致将遗传信息从供体多核苷酸转移到靶DNA中。这样,可以将新的核酸材料插入/复制到位点中。由于NHEJ和/或同源定向修复(例如使用供体DNA分子)而对靶DNA进行的修饰可导致例如基因校正、基因置换、基因标记、转基因插入、核苷酸缺失、基因破坏、基因突变等。

[0103] 在一些情况下,通过定点修饰多肽(例如,cas核酸酶、锌指核酸酶、大范围核酸酶或TALEN)切割DNA可用于通过切割靶DNA序列并允许细胞在不存在外源提供的供体多核苷酸的情况下修复所述序列来添加、删除或改变靶DNA中核酸材料的序列。此类NHEJ事件可导致切割末端重新结合的位点发生突变(“错配”),从而导致基因破坏。

[0104] 如果将靶向DNA的RNA与供体DNA共同施用于表达cas核酸酶的细胞,则可以使用主题方法向靶DNA序列(例如,通过插入诱变“敲除”或“敲入”编码蛋白质(例如选择标志物、报告基因和/或任何目标蛋白质)的核酸)、siRNA、miRNA等添加(即插入或替换)核酸材料以修饰核酸序列(例如,引入突变)等。

[0105] 供体DNA可以在特定的实施例中包括基因调节序列(例如,启动子),其可以使用CRISPR靶向插入在基因的编码区的上游和基因的假定近端启动子区的上游,例如,脂质调节基因的编码区起始ATG上游至少约50bp、至少100bp、至少120bp、至少150bp、至少200bp、至少250bp、至少300bp、至少350bp、至少400bp、至少450bp或至少500bp。供体DNA可以包括可能干扰天然启动子的序列,例如选择标志物或任何方便的序列。插入脂质调节基因开放阅读框的起始ATG上游的额外序列(例如,在脂质调节基因5'UTR或转录起始位点的上游)可以降低甚至消除内源性脂质调节基因的表达。可替代地或另外,天然脂质调节基因可以使其内源性启动子全部或部分地由弱或不同调节的启动子或非启动子序列替代。

[0106] 在一些实例中,引入宿主细胞以产生高效基因组编辑细胞系的核酸分子编码相对于相应的野生型酶突变的Cas9酶,使得突变的Cas9酶缺乏切割含有靶序列的靶多核苷酸的一条或两条链的能力。例如,来自酿脓葡萄球菌的Cas9的RuvC I催化结构域中的天冬氨酸至丙氨酸取代(D10A)将Cas9从切割两条链的核酸酶转化为切口酶(切割一条单链的酶)。使Cas9成为切口酶的突变的其它实例包括但不限于H840A、N854A和N863A。在一些实施例中,Cas9切口酶可以与一个或多个引导序列例如两个引导序列组合使用,所述引导序列分别靶向DNA靶的有义和反义链。这种组合允许两条链都被切开并用于诱导NHEJ。可以使用两个切口酶靶(非常接近但靶向DNA的不同链)诱导诱变的NHEJ。使用在交错位置切割相反藻株的酶对基因座的此类靶向还可以减少非靶标的切割,因为必须精确且特异性地切割两条链以实现基因组突变。在另外的实例中,可以在细胞中表达其切割DNA的能力受损的突变Cas9酶,其中还引入了靶向靶基因的转录或翻译起始位点上游序列的一个或多个引导RNA。在这种情况下,cas酶可以结合靶序列并阻断靶基因的转录(Qi等人(2013)《细胞(Cell)》152:1173-1183)。这种基因表达的CRISPR干扰可以称为RNAi,并且在Larson等人(2013)《自然协议(Nat.Protoc.)》8:2180-2196中进行了详细描述。在一些情况下,cas多肽如Cas9多肽是融合多肽,其包含例如:i) Cas9多肽(其可以任选地是如上所述的变体Cas9多肽);和b) 共价连接的异源多肽(也称为“融合伴侣”)。异源核酸序列可以与另一核酸序列(例如,通过基因工程化)连接以产生编码嵌合多肽的嵌合核苷酸序列。在一些实施例中,Cas9融合多肽通过将Cas9多肽与提供亚细胞定位的异源序列融合来产生(即,异源序列是亚细胞定位序列,例

如用于靶向细胞核的核定位信号(NLS);用于靶向线粒体的线粒体定位信号;用于靶向叶绿体的叶绿体定位信号;ER保留信号等)。在一些实施例中,异源序列可以提供标签(即,异源序列是可检测的标记),以易于追踪和/或纯化(例如,荧光蛋白,例如绿色荧光蛋白(GFP)、YFP、RFP、CFP、mCherry、tdTomato等;血凝素(HA)标签;FLAG标签;Myc标签等)。

[0107] 宿主细胞可以用例如载体构建体进行基因工程化(例如,转导或转化或转染),所述载体构建体可以是例如用于同源重组的载体,所述载体包括与宿主细胞脂质调节基因位点的一部分或其邻近区域同源的核酸序列,或者可以是用于表达以下的任何一种或组合的表达载体:cas蛋白(例如,Cas9蛋白)、CRISPR嵌合引导RNA、crRNA和/或tracrRNA、RNAi构建体(例如shRNA)、反义RNA或核酶。载体可以是例如质粒、病毒颗粒、噬菌体等形式。用于表达多肽或RNA以进行基因组编辑的载体也可以设计成例如通过同源重组整合到宿主中。含有本文所述的多核苷酸序列(例如,与宿主脂质调节基因序列具有同源性的序列(包括在脂质调节基因编码序列上游和下游的序列),以及任选的选择标志物或报告基因)的载体可以用来转化合适的宿主以引起脂质调节基因的减弱。

[0108] 在一些实例中,重组微生物可以具有降低但未消除的脂质调节基因的表达,并且重组微生物可以具有例如约25%至约250%或更高的脂质产生增加。例如,相对于对照微生物,脂质产生增加可以为增加约25%至增加约250%、增加约25%至增加约225%、增加约25%至增加约200%、增加约25%至增加约175%、增加约25%至增加约150%、增加约25%至增加约125%、增加约50%至增加约250%、增加约50%至增加约225%、增加约50%至增加约200%、增加约50%至增加约175%、增加约50%至增加约150%、增加约50%至增加约125%、增加约75%至增加约250%、增加约75%至增加约225%、增加约75%至增加约200%、或增加约75%至增加约175%、增加约75%至增加约150%、或增加约75%至增加约125%(例如增加25-250%)。本文提供的基因修饰的微生物可以在一些实例中包括用于减弱脂质调节基因的表达的核酸构建体,例如,编码具有GAF结构域,例如GAF2结构域的多肽的基因,其与SEQ ID NO:1具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性;或编码包含与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列的多肽的基因。例如,宿主微生物可以包括用于表达RNAi分子、核酶或反义分子的构建体,其降低编码具有GAF结构域的多肽的脂质调节基因的表达,所述基因与SEQ ID NO:1具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性,或者降低编码多肽的脂质调节基因的表达,所述基因与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性。在一些实例中,本文提供的重组微生物可包括至少一种引入的用于降低脂质调节基因的表达的(外源或非天然)构建体。

[0109] 在一些实例中,可以选择工程化藻株(例如,基因工程化藻株)来表达脂质调节基因,所述表达相对于不包括使用本领域已知的方法如RNA-Seq或逆转录-PCR(RT-PCR)来减弱但不消除脂质调节基因表达的基因修饰的对照细胞降低。如本文所提供的基因工程化藻

株可以被工程化以包括用于通过减少编码脂质调节序列的基因的mRNA的量、稳定性或可翻译性来减弱基因表达的构建体。例如,可以使用本领域已知的方法,用靶向脂质调节基因的mRNA的反义RNA、RNAi或核酶构建体转化微生物如藻类或长短鞭毛体藻株。例如,可以将包括基因的全部或部分转录区域的反义RNA构建体引入微生物以降低基因表达(Shroda等人(1999)《植物细胞(The Plant Cell)》11:1165-78;Ngiam等人(2000)《应用与环境微生物学(Appl. Environ. Microbiol.)》66:775-782;Ohnuma等人(2009)《原生质(Protoplasma)》236:107-112;Lavaud等人(2012)《美国科学公共图书馆(PLoS One)》7:e36806)。可替代地或另外,可以将靶向具有GAF(例如,GAF2)结构域的基因的RNAi构建体(例如,编码短发夹RNA的构建体)引入微生物如藻类或长短鞭毛体中以降低脂质调节基因的表达(参见例如Cerruti等人(2011)《真核细胞(Eukaryotic Cell)》(2011) 10:1164-1172;Shroda等人(2006)《当代遗传学(Curr. Genet.)》49:69-84)。

[0110] 核酶是以位点特异性方式切割核酸的RNA-蛋白质复合物。核酶具有拥有核酸内切酶活性的特定催化结构域。例如,美国专利第5,354,855号(通过引用整体并入本文)报道了某些核酶可以以比已知的核糖核酸酶更大的序列特异性并接近DNA限制酶的序列特异性作为核酸内切酶。可以将催化RNA构建体(核酶)设计为与编码本文所提供的基因的mRNA碱基配对以切割mRNA靶标。在一些实例中,核酶序列可以整合在反义RNA构建体中以介导靶标的切割。可以考虑各种类型的核酶,它们的设计和用途是本领域已知的并且例如描述于Haseloff等人(1988)《自然(Nature)》334:585-591。核酶通过互补碱基对相互作用退火至位点而靶向给定序列。所述靶向需要两段同源性。这些同源序列段侧接于上文定义的催化核酶结构。每一段同源序列的长度可在7至15个核苷酸之间变化。定义同源序列的唯一要求是,在靶RNA上,它们被特定的序列(切割位点)分开。对于锤头状核酶,切割位点是靶RNA上的二核苷酸序列,其是尿嘧啶(U),然后是腺嘌呤、胞嘧啶或尿嘧啶(A、C或U)(Thompson等人(1995)《核酸研究(Nucl Acids Res)》23:2250-68)。统计上在任何给定的RNA中出现这种二核苷酸的频率为16分之3。因此,对于给定的1,000个碱基的靶信使RNA,统计上可能有187个二核苷酸切割位点。

[0111] 已经详细讨论了核酶指导的RNA切割活性的一般设计和优化(Haseloff和Gerlach(1988)《自然(Nature)》334:585-591;Symons(1992)《生物化学年鉴(Ann Rev Biochem)》61:641-71;Chowrira等人(1994)《生物化学杂志(J Biol Chem)》269:25856-64;Thompson等人(1995)同上),均通过引用整体并入。设计和测试用于有效切割靶RNA的核酶是本领域技术人员众所周知的过程。设计和测试核酶的科学方法的实例在Chowrira等人,(1994)同上,Lieber和Strauss(1995)《分子细胞生物学(Mol Cell Biol.)》15:540-51中描述,其各自通过引用并入。用于下调给定基因的操作和优选序列的鉴定是制备和测试给定序列的问题,并且是本领域技术人员已知的常规实践的“筛选”方法。RNAi构建体的使用例如在上文引用的文献以及US 2005/0166289和WO 2013/016267(均通过引用并入本文)中描述。通过表达RNAi构建体,例如RNAi短发夹(sh)构建体,将与靶基因同源的双链RNA递送至细胞或在细胞中产生。所述构建体可包括与靶基因相同的序列,或与靶基因序列具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%的同一性。所述构建体可具有至少约20、至少30、至少40、至少50、至少100、至少200、至少300、至少400、至少500、至少600、至少700、

至少800、至少900或至少约1kb的与靶基因同源的序列。可以使用选择用于RNAi构建体例如产生shRNA的构建体的连续或诱导表达的启动子来工程化表达载体。

[0112] 用于基因减弱的核酸构建体,例如核酶、RNAi或反义构建体,可包括至少约十五、至少二十、至少三十、至少四十、至少五十或至少六十个核苷酸,其与待工程化的微生物的内源性脂质调节基因序列的至少约一部分具有至少约80%的同一性(例如至少85%、至少90%、至少95%或至少约99%的同一性)或互补性。用于基因减弱的核酸构建体,例如核酶、RNAi或反义构建体,可包括至少约十五、至少二十、至少三十、至少四十、至少五十或至少六十个核苷酸,其与天然存在的基因的序列具有至少约80%,如至少约95%或约100%的同一性或互补性,所述天然存在的基因为例如编码多肽的基因,其与内源性脂质调节基因具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%或至少85%、至少90%或至少约95%的序列同一性,所述内源性脂质调节基因为例如位于Naga_100020g79基因座的基因(SEQ ID NO:3)或编码具有GAF结构域(如GAF2)的多肽的基因(例如SEQ ID NO 1)。例如,用于基因减弱的核酸构建体,例如核酶、RNAi或反义构建体,可包括至少约十五、至少二十、至少三十、至少四十、至少五十或至少六十个核苷酸,其与天然存在的脂质调节基因,如本文提供的任何基因的序列具有至少约80%的同一性或互补性。核苷酸序列的长度可以为例如至少约30个核苷酸至至少约3千碱基,例如长度为至少约30个核苷酸至至少约50个核苷酸,长度为至少约50个核苷酸至至少约100个核苷酸,长度为至少约100个核苷酸至至少约500个核苷酸,长度为至少约500个核苷酸至至少约1kb,长度为至少约1kb至至少约2kb,或长度为至少约2kb至至少约5kb。例如,反义序列的长度可以是至少约100个核苷酸至至少约1kb。例如,用于基因减弱的核酸构建体,例如核酶、RNAi或反义构建体,可以包括至少约十五、至少二十、至少三十、至少四十、至少五十、至少六十或至少约100个核苷酸,其与内源性脂质调节基因或其部分具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%或至少85%,例如至少86%、至少87%、至少88%、至少89%、至少90%、至少91%、至少92%、至少93%、至少94%或至少约95%的同一性或互补性。

[0113] 反义、RNAi或核酶构建体中使用的启动子可以是在宿主生物体中具有功能并且适合于将靶基因表达降低至所需量的所需表达水平的任何启动子。在藻类和长短鞭毛体中起作用的启动子是本领域已知的并且在本文中公开。可以使用任何可行的方法,包括本文公开的任何方法,将构建体转化为藻类。用于减弱脂质调节基因表达的核酸分子转化的重组生物体或微生物,例如但不限于反义、RNAi或核酶构建体,可以具有本文所述的脂质调节突变体的特性,包括例如相对于不包括导致基因表达减弱的外源核酸分子的宿主生物体或微生物,减少的叶绿素、增加的光合作用效率和增加的培养物生产力。

[0114] 核酸分子和构建体

[0115] 本文还提供编码多肽的核酸分子,其包括与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列。可替代地或另外,本文提供的核酸分子可包括与SEQ ID NO:2的氨基酸序列具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的序列,和/或由与SEQ ID NO:3具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少

95%、至少96%、至少97%、至少98%或至少约99%同一性的核苷酸序列编码的序列,本文提供的核酸分子可以包括编码属于pfam PF13185的GAF结构域,如GAF2结构域的氨基酸序列例如,由核酸分子编码的多肽可以包括具有与SEQ ID NO:4具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列的GAF结构域。

[0116] 在各种实例中的核酸分子可以是或包含缺少天然存在的基因中存在的一个或多个内含子的cDNA,或者可替代地,可以包括天然存在的基因中不存在的一个或多个内含子。各种实例中的核酸分子可以具有与天然存在的基因不是100%相同的序列。例如,核酸分子可以包括相对于天然存在的基因的突变,所述突变降低编码的多肽的活性或降低所述基因编码的mRNA或蛋白质的表达。

[0117] 各种实例中的核酸分子可包含与编码多肽的序列可操作连接的异源启动子,所述多肽包括与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列和/或与SEQ ID NO:3或SEQ ID NO:4具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的序列。可替代地或另外,核酸分子可以包含载体,所述载体包括编码多肽的序列,所述多肽包括与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列和/或与SEQ ID NO:3或SEQ ID NO:4具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的序列。

[0118] 本公开还提供被设计用于减弱编码GAF结构域,如GAF2结构域的基因的表达的构建体。在各种实例中,所述构建体可以是或包含编码CRISPR系统的引导RNA的序列、RNAi构建体、反义构建体、核酶构建体或用于同源重组的构建体,例如具有一个或多个与编码本文公开的天然存在的GAF(例如,GAF2)结构域的基因具有同源性的核苷酸序列,和/或在衍生所述基因的天然基因组中与其相邻的序列的构建体。例如,所述构建体可以包括编码具有GAF结构域的多肽的基因的至少一部分,例如,与编码多肽的基因的至少一部分同源的序列,所述多肽包括与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列。

[0119] 用于基因减弱的构建体可以包括,例如,基因的编码区、内含子、5'UTR、启动子区或3'UTR的至少一部分,所述基因编码具有GAF结构域,如GAF2的多肽,或与SEQ ID NO:2在有义或反义方向上具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的多肽,或者与SEQ ID NO:3在有义或反义方向上具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的基因的一部分。

[0120] 在进一步的实例中,可以设计构建体以用于引导RNA(例如,CRISPR系统的引导

RNA)的体外或体内表达,所述引导RNA被设计成靶向天然存在的基因,所述基因的序列与SEQ ID NO:3的至少一部分具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性,和/或编码具有GAF结构域,如GAF2的多肽,所述结构域包含与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列,和/或可以包括与编码具有GAF结构域(例如,GAF2)的多肽的基因的一部分同源的序列,包括例如内含子、5'UTR、启动子区域和/或3'UTR。

[0121] 在又进一步的实例中,用于减弱编码含有GAF结构域的多肽的基因的表达的构建体可以是引导RNA或反义寡核苷酸,其中所述序列与编码在反义取向上具有GAF结构域的多肽的基因的转录区具有同源性。

[0122] 用于使GAF结构域编码基因或编码与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的多肽的基因的表达减弱的核酸构建体可以包括例如天然存在的GAF结构域编码基因和/或编码与SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的多肽的基因和/或与SEQ ID NO:3的一部分具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的基因的序列的至少约17个、至少18个、至少19个、至少20个、至少21个、至少22个、至少23个、至少24个或至少约25个核苷酸。

[0123] 在一个实例中,本文提供一种核酸分子,其与SEQ ID NO:3或SEQ ID NO:4的至少一部分具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、或至少约75%、至少约80%、至少约85%、至少约90%或至少约95%(例如50%、70%、90%或95%中的至少一种)同一性,其中核酸分子编码CRISPR系统的引导RNA。核酸分子可包括例如天然存在的含有GAF结构域的基因的序列(例如但不限于SEQ ID NO:3)的至少约17、至少约18、至少约19、至少约20、至少约21、至少约22、至少约23、至少约24或至少约25个核苷酸(例如,17-25个核苷酸中的至少任何一个)。

[0124] 另外,本文提供反义、核酶或RNAi构建体,其包括具有GAF结构域(例如,GAF2)的基因的至少一部分,或与SEQ ID NO:2具有至少约65%同一性的多肽和/或与SEQ ID NO:3的一部分具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的多肽,其中启动子如异源启动子可操作地连接至GAF结构域基因序列,并且GAF结构域基因序列处于反义方向。

[0125] 此外,本文提供用于同源重组的构建体,其包括来自或邻近天然存在的藻类基因的核苷酸序列,所述核苷酸序列编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列的多肽;或位于Naga_100020g79基因座的基因;和/或包含ORF的基因,ORF包含与SEQ ID NO:3和/或SEQ ID

N0:4具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的序列。在一些实施例中,核苷酸序列与异源核酸序列并置,所述异源核酸序列在非限制性实例中可以是选择标志物或可检测标记基因。在一些实例中,用于同源重组的构建体包括两个核酸序列,其来自或邻近于天然存在的藻类基因,所述藻类基因编码具有与SEQ ID N0:1或SEQ ID N0:2具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的氨基酸序列的多肽;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,ORF包含与SEQ ID N0:3和/或SEQ ID N0:4具有至少约50%、至少55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的序列,其中两个序列侧接于用于插入基因位点的异源序列。

[0126] 本领域技术人员将理解,多种转化方法可用于微生物的基因转化,并且因此可用于本发明的方法。“稳定转化”旨在表示引入生物体中的核酸构建体整合到生物体的基因组中,或者是稳定的附加型构建体的一部分,并且能够被其后代遗传。“瞬时转化”旨在表示将多核苷酸引入生物体,并且不整合到基因组中,或者另外不被后续世代确立并稳定地遗传。

[0127] 基因转化可以导致转基因、来自核或质体的构建体的稳定插入和/或表达,并且在某些情况下可以导致转基因的瞬时表达。转化方法也可用于引入引导RNA或编辑DNA。据报道,微藻的基因转化成功地用于30种以上不同的微藻株,所述藻株至少属于约22种绿藻、红藻和褐藻、硅藻、眼虫和腰鞭毛虫物种(参见例如Radakovits等人,《真核细胞(Eukaryotic Cell)》,2010;和Gong等人,《工业微生物学与生物技术杂志(J.Ind.Microbiol.Biotechnol.)》,2011)。此类有用的转化方法的非限制性实例包括在玻璃珠或碳化硅晶须存在下的细胞搅动,如例如Dunahay,《生物技术(Biotechniques)》,15(3):452-460,1993;Kindle,《美国科学院学报(Proc.Natl.Acad.Sci.U.S.A.)》,1990;Michael和Miller,《植物杂志(Plant J.)》,13,427-435,1998所报道的。电穿孔技术已成功用于几种微藻物种的基因转化,包括微拟球藻属(参见,例如,Chen等人,《植物学杂志(J.Phycol.)》,44:768-76,2008)、小球藻属(参见例如Chen等人,《当代遗传学(Curr.Genet.)》,39:365-370,2001;Chow和Tung,《植物细胞报告(Plant Cell Rep.)》Vol.18,No.9,778-780,1999)、衣藻属(Shimogawara等人,《遗传学(Genetics)》,148:1821-1828,1998)、杜氏藻属(Sun等人,《分子生物技术(Mol.Biotechnol.)》,30(3):185-192,2005)。微粒轰击法(也称为微粒轰击、基因枪转化或生物弹轰击)已成功用于几种藻类物种,包括例如硅藻物种如褐指藻属(Apt等人,《分子和一般遗传学(Mol.Gen.Genet.)》,252:572-579,1996)、小环藻属和舟形藻属(Dunahay等人,《植物学杂志(J.Phycol.)》,31:1004-1012,1995)、细柱藻属(Fischer等人,《植物学杂志(J.Phycol.)》,35:113-120,1999)和角毛藻属(Miyagawa-Yamaguchi等人,《植物学研究(Phycol.Res.)》59:113-119,2011),以及绿藻物种,如小球藻属(El-Sheekh,《植物生物学(Biologia Plantarum)》,Vol.42,No.2:209-216,1999)和团藻属物种(Jakobiak等人,《原生生物(Protist)》,155:381-93,2004)。此外,农杆菌介导的基因转移技术也可用于微藻的基因转化,如例如Kumar,《植物科学(Plant Sci.)》,166(3):731-738,2004和Cheney等人,《植物学杂志(J.Phycol.)》,Vol.37,Suppl.11,2001中所报道的。

[0128] 如本文所述的转化载体或构建体通常将包含赋予靶宿主细胞例如藻类细胞可选择的或可评分的表型的标记基因,或可以与包括标志物的构建体共转化。已经成功开发了多种选择标记,用于有效分离藻类的基因转化体。常见的选择标记包括抗生素抗性、荧光标记和生化标记。几种不同的抗生素抗性基因已成功地用于选择微藻转化体,包括杀稻瘟菌素(blastocidin)、博来霉素(参见,例如,Apt等人,1996,同上;Fischer等人,1999,同上;Fuhrmann等人,《植物杂志(Plant J.)》,19:353-61,1999,Lumbreras等人,《植物杂志(Plant J.)》,14(4):441-447,1998;Zaslavskaia等人,《植物学杂志(J. Phycol.)》,36:379-386,2000)、壮观霉素(Cerutti等人,《遗传学(Genetics)》,145:97-110,1997;Doetsch等人,《遗传学现状(Curr. Genet.)》,39:49-60,2001;Fargo,《分子细胞生物学(Mol. Cell. Biol.)》,19:6980-90,1999)、链霉素(Berthold等人,《原生生物(Protist)》,153:401-412,2002)、巴龙霉素(Jakobiak等人,《原生生物(Protist)》,同上;Sizova等人,《基因(Gene)》,277:221-229,2001)、诺尔丝菌素(Zaslavskaia等人,2000,同上)、G418(Dunahay等人,1995,同上;Poulsen和Kroger,《欧洲生物化学学会联盟快报(FEBS Lett.)》,272:3413-3423,2005,Zaslavskaia等人,2000,同上)、潮霉素(Berthold等人,2002,同上)、氯霉素(Poulsen和Kroger,2005,同上)等。用于微藻类如衣藻属的其它选择标记可以是提供对以下的抗性的标记:卡那霉素和阿米卡星抗性(Bateman,《分子和一般遗传学(Mol. Gen. Genet.)》263:404-10,2000)、新霉素(zeomycin)和腐草霉素(例如ZEOCIN™腐草霉素D1)抗性(Stevens,《分子和一般遗传学(Mol. Gen. Genet.)》251:23-30,1996)以及巴龙霉素和新链丝菌素(neomycin)抗性(Sizova等人,2001,同上)。已使用的其它荧光或发色标记包括萤光素酶(Falciatore等人,《海洋生物技术杂志(J. Mar. Biotechnol.)》,1:239-251,1999;Fuhrmann等人,《植物分子生物学(Plant Mol. Biol.)》,2004;Jarvis和Brown,《遗传学现状(Curr. Genet.)》,19:317-322,1991)、 β -葡萄糖醛酸酶(Chen等人,2001,同上;Cheney等人,2001,同上;Chow和Tung,1999,同上;El-Sheekh,1999,同上;Falciatore等人,1999,同上;Kubler等人,《海洋生物技术杂志(J. Mar. Biotechnol.)》,1:165-169,1994)、 β -半乳糖苷酶(Gan等人,《应用植物学杂志(J. Appl. Phycol.)》,15:345-349,2003;Jiang等人,《植物细胞报告(Plant Cell Rep.)》,21:1211-1216,2003;Qin等人,《高技术快报(High Technol. Lett.)》,13:87-89,2003)和绿色荧光蛋白(GFP)(Cheney等人,2001,同上;Ender等人,《植物细胞(Plant Cell)》,2002;Franklin等人,《植物杂志(Plant J.)》,2002;56,148,210)。

[0129] 本领域技术人员将容易理解,可有效地使用各种已知的启动子序列以用于根据本发明的微藻物种的转化系统。例如,通常用于在微藻中驱动转基因表达的启动子包括花椰菜花叶病毒启动子35S(CaMV35S)的各种版本,所述花椰菜花叶病毒启动子35S已用于鞭毛藻和绿藻门(Chow等人,《植物细胞报告(Plant Cell Rep.)》,18:778-780,1999;Jarvis和Brown,《遗传学现状(Curr. Genet.)》,317-321,1991;Lohuis和Miller,《植物杂志(Plant J.)》,13:427-435,1998)。据报道,来自猿猴病毒的SV40启动子在几种藻类中也有活性(Gan等人,《应用植物学杂志(J. Appl. Phycol.)》,151:345-349,2003;Qin等人,《水生生物学(Hydrobiologia)》398-399,469-472,1999)。来自衣藻属的RBCS2(二磷酸核糖羧化酶,小亚基)(Fuhrmann等人,《植物杂志(Plant J.)》,19:353-361,1999)和PsaD(光系统I复合物的丰富蛋白;Fischer和Rochaix,《欧洲生物化学学会联盟快报(FEBS Lett.)》581:5555-

5560, 2001) 的启动子也可以是有用的。HSP70A/RBCS2和HSP70A/ β 2TUB (微管蛋白) 的融合启动子 (Schroda等人,《植物杂志 (Plant J.)》, 21:121-131, 2000) 也可用于改善转基因的表达, 其中HSP70A启动子当位于其它启动子上游时可以用作转录激活子。在例如硅藻物种中, 也可以在编码硅藻岩藻黄质-叶绿素a/b结合蛋白的fcp基因 (Falciatore等人,《海洋生物技术 (Mar. Biotechnol.)》, 1:239-251, 1999; Zaslavskaja等人,《植物学杂志 (J. Phycol.)》36:379-386, 2000) 或编码黄绿藻紫黄质-叶绿素a/b结合蛋白的vcp基因 (参见美国专利第8,318,482号, 通过引用并入本文) 的启动子的控制下实现目标基因的高水平表达。如果需要的话, 诱导型启动子可以在转基因微藻中提供快速且严格控制的基因表达。例如, 可以将编码硝酸盐还原酶的NR基因的启动子区域用作此类诱导型启动子。NR启动子活性通常被铵抑制, 并在铵被硝酸盐替代时被诱导 (Poulsen和Kroger,《欧洲生物化学学会联盟快报 (FEBS Lett)》272:3413-3423, 2005), 因此当在铵/硝酸盐存在下生长微藻细胞时, 可以关闭或打开基因表达。可用于本文提供的构建体和转化系统中的其它藻类启动子包括在美国专利第8,883,993号; 美国专利申请公开第US 2013/0023035号; 美国专利申请公开第US 2013/0323780号; 和美国专利申请公开第US 2014/0363892号中公开的那些, 所述专利全部通过引用整体并入本文。

[0130] 宿主细胞可以是未转化的细胞, 或者已经被至少一种核酸分子转染的细胞。例如, 经工程化以使脂质调节基因表达减弱的藻类宿主细胞可以进一步包括一个或多个基因, 其可以赋予任何期望的特性, 例如但不限于增加目标生物分子如一种或多种蛋白质、色素、醇或脂质的生产。

[0131] 制备脂质的方法

[0132] 本文还提供通过培养如本文所提供的突变微生物来生产脂质的方法, 在相同条件下培养时, 所述突变微生物相对于对照细胞具有增加的脂质生产力。所述方法包括在合适的培养基中培养如本文提供的突变微生物以产生脂质, 并从培养物中回收生物质或至少一种脂质。在一些实例中, 微生物可以是藻类, 并且培养物可以是光合自养培养物。培养可以分批、半连续或连续方式进行。

[0133] 在一些实例中, 突变微生物可以在包含少于约5mM铵、少于约4.5mM铵、少于约4mM铵、少于约3.5mM铵、少于约3mM铵、少于约2.5mM铵、少于约2mM铵、少于约1.5mM铵、少于或等于约1mM铵、少于或等于约0.5mM (例如, 少于5mM) 或基本上不含铵的培养基中培养。培养基可以包括例如约0至约5mM铵、约0至约4.5mM铵、约0至约4.0mM铵、约0至约3.5mM铵、约0至约3mM铵、约0至约2.5mM铵、约0至约2.0mM铵、约0至约1.5mM铵、约0至约1.0mM铵、约0至约0.5mM铵、约0.2至约3mM铵、约0.2至约2.5mM铵、约0.2至约2mM铵、约0.2至约1.5mM铵、约0.2至约1mM铵、约0.3至约2.5mM铵、约0.3至约2mM铵、约0.3至约1.5mM铵、约0.3至约1mM铵、约0.4mM至约2.5mM铵、约0.4至约2mM铵、约0.4mM至约1.5mM (例如0-5mM) 铵。可以在包括硝酸盐的培养基中培养微生物, 在一些实例中, 硝酸盐可以是培养基中基本上唯一的氮源, 或者还可以存在少于约5mM铵、少于约4.5mM铵、少于约4mM铵、少于约3.5mM铵、少于约3mM铵、少于约2.5mM铵、少于约2mM铵、少于约1.5mM铵、少于或等于约1mM铵、少于或等于约0.5mM (例如少于5mM), 或者基本上不存在铵。可替代地或另外, 培养基可以包含尿素, 在一些实例中, 尿素可以是培养基中基本上唯一的氮源。

[0134] 产生脂质的微生物可以在任何合适的容器中培养, 包括烧瓶或生物反应器。在一

些实例中,突变微生物是藻类,并且暴露于光,持续至少一部分培养期,其中藻类可以暴露于人造或自然光(或补充有人造光的自然光)。在其对弱光的响应中被解除调节的包含突变藻类的培养物可以在光/暗循环中进行培养,所述光/暗循环可以是例如自然或程序性的光/暗循环,并且作为示例性实例,可以提供十二个小时的光照/十二个小时的黑暗,十四个小时的光照/十个小时的黑暗,十六个小时光照/八个小时的黑暗,等等。可替代地,可以在连续光照下培养藻类突变体。

[0135] 培养是指通过使用选择和/或对照条件有意培养一个或多个细胞的生长(例如,细胞大小、细胞含量和/或细胞活性增加)和/或繁殖(例如,通过有丝分裂增加细胞数量)。生长和繁殖的结合可以被称为增殖。本文提供的微生物可以被培养至少约五、至少约六、至少约七、至少约八、至少约九、至少约十、至少约十一、至少约十二、至少约十三、至少约十四或至少约十五天,或至少约一、二、三、四、五、六、七、八、九或十周或更长时间。可以在相对于对照藻类富营养物的培养基中进行培养。

[0136] 可用于培养重组微生物的选择和/或对照条件的非限制性实例可包括使用确定的培养基(具有已知特性,例如pH、离子强度和/或碳源)、指定的温度、氧张力、二氧化碳水平、生物反应器的生长等或它们的组合。在一些实施例中,微生物或宿主细胞可以使用光和还原碳源混合营养生长。可替代地,微生物或宿主细胞可以光养培养。当光养生长时,藻株可以有利的地使用光作为能源。无机碳源,如CO₂或碳酸氢盐可用于微生物合成生物分子。如本文所用,“无机碳”包括不能被生物体用作可持续能源的含碳化合物或分子。通常,“无机碳”可以是不能进一步氧化以用于可持续能源也不能用作生物体降低能量的来源的CO₂(二氧化碳)、碳酸、碳酸氢盐、碳酸盐、碳酸氢盐等或其组合的形式。光合自养的微生物生长可以在其中无机碳基本上是唯一碳源的培养基上生长。例如,在其中无机碳基本上是唯一碳源的培养物中,培养基中可能提供的任何有机(还原)碳分子或有机碳化合物无法被细胞吸收和/或代谢以用于能量,和/或不以足以为细胞培养物的生长和增殖提供可持续能量的量存在。

[0137] 根据本发明的方法有用的微生物和宿主细胞可以在世界各地的不同位置和环境找到。用于脂质和/或其它产品最佳繁殖和产生的特定生长培养基可以变化,并且可以进行优化以促进产物如脂质、蛋白质、色素、抗氧化剂等等的生长、繁殖或生产。在一些情况下,某些微生物藻株可能因为特定的微生物或宿主细胞藻株存在一些抑制性组分或缺乏一些基本的营养需求而无法在特定的生长培养基中生长。

[0138] 固体和液体生长培养基通常可从多种来源获得,如制备适用于多种微生物藻株的特定培养基的说明一样。例如,各种淡水和盐水介质可以包括Barsanti (2005)《藻类:解剖学、生物化学与生物技术 (Algae: Anatomy, Biochemistry & Biotechnology)》, CRC出版社中描述的用于培养藻类的培养基和方法的那些。藻类培养基配方也可以在各种藻类培养集合的网站上找到,作为非限制性实例包括《藻类的UTEX培养集合 (UTEX Culture Collection of Algae)》;《藻类和原生动物的培养集合 (Culture Collection of Algae and Protozoa)》;和Katedra Botaniky。

[0139] 培养方法可任选地包括在微生物中诱导一种或多种基因的表达和/或调节代谢途径。诱导表达可以包括向培养物中添加营养物或化合物,从培养基中去除一种或多种组分,增加或减少光和/或温度,和/或促进目标基因表达的其它操作。此类操作可很大程度上取

决于与目标基因可操作连接的(异源)启动子的性质。

[0140] 在本发明的一些实施例中,可以在配备有人工光源和/或具有一个或多个对包括阳光在内的光足够透明以使得能够进行、促进和/或维持可接受的微生物生长和增殖的壁的光生物反应器中培养具有增加的脂质生产力的微生物。为了生产脂肪酸产物或甘油三酯,可以在摇瓶、试管、小瓶、微量滴定皿、皮氏培养皿等或其组合中另外或替代地培养光合微生物或宿主细胞。

[0141] 另外地或可替代地,突变或重组的光合微生物或宿主细胞可以在池塘、运河、海基生长容器、沟渠、水道、渠道等或其组合中生长。在此类系统中,温度可以不受调节,或者可以采用各种加热或冷却方法或装置。与标准生物反应器一样,无机碳源(例如但不限于CO₂、碳酸氢盐、碳酸盐等)包括但不限于空气、富含CO₂的空气、烟道气等或其组合可被供应给培养物。当供应烟道气和/或除CO₂以外可能还含有CO的其它无机源时,可能需要预处理此类源,使得引入(光)生物反应器的CO水平不构成就微生物的生长、增殖和/或存活而言的危险和/或致死剂量。

[0142] 突变微生物可任选地包括一种或多种编码用于产生产物例如但不限于脂质的多肽的非天然基因。

[0143] 所述方法包括培养如本文提供的突变微生物,例如当在相同条件下培养以产生生物质或脂质时相对于对照细胞具有增加的脂质生产力的如本文提供的突变微生物。可以通过本领域普通技术人员已知的回收方法从培养物中回收脂质,例如通过全培养物提取,例如使用有机溶剂,或通过首先分离从中提取脂质的生物质(参见,例如,Hussein等人,《应用生物化学与生物技术(Appl.Biochem.Biotechnol.)》175:3048-3057;Grima等人(2003)《生物技术进展(Biotechnol.Advances)》20:491-515)。在一些情况下,可以通过细胞的均质化来增强脂肪酸产物的回收(Gunerken等人(2015)《生物技术进展(Biotechnol.Advances)》33:243-260)。例如,可以通过在升高的温度和/或压力下用溶剂萃取藻类来从藻类中分离出脂质如脂肪酸、脂肪酸衍生物和/或甘油三酯,如于2012年2月29日提交的标题为“从藻类中溶剂提取产物”的共同待决、共同转让的美国专利公开第US 2013/0225846号所述,其通过引用整体并入本文。可以通过例如离心或过滤来收获生物质。可以干燥和/或冷冻生物质。可以从生物质如各种脂质或一种或多种蛋白质中分离其它产物。本发明还包括包含脂质调节突变体的生物质的藻类生物质,例如本文公开的任一种,例如但不限于脂质调节突变体,其包括在与SEQ ID NO:1具有至少约50%、至少约55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的编码具有GAF(例如GAF2)结构域的多肽的基因中的突变,或在与SEQ ID NO:1具有至少约50%、至少约55%、至少60%、至少65%、至少70%、至少75%、至少80%、至少85%、至少90%、至少95%、至少96%、至少97%、至少98%或至少约99%同一性的编码多肽的基因中的突变。

[0144] 还考虑细胞提取物,包括粗制细胞裂解物或已经从中分离出一种或多种组分的细胞裂解物。可以任选地使用一种或多种分离步骤,例如但不限于离心、沉淀、过滤或色谱法来产生提取物。在各种非限制性实例中,可以从沉降的固体、颗粒、膜或聚集的细胞材料中倾析、过滤或除去细胞裂解物以产生提取物。细胞提取物可以可替代地或另外地任选是使用一种或多种也可能存在于提取物中的洗涤剂、表面活性剂或溶剂的提取产物。提取物可

以任选地包括一种或多种盐、缓冲剂、稳定剂或抗氧化剂。

[0145] 如本领域普通技术人员将理解的,在本文中还可能预期其它实施例。在以下实例中进一步描述某些实施例。这些实施例仅作为实例提供,并不旨在以任何方式限制权利要求的范围。

[0146] 本文引用的所有参考文献通过引用整体并入。所有标题都是为了读者的方便,并不以任何方式限制本发明。对本发明的方面或实施例的引用不一定指示所描述的方面不可以与本发明的其它所描述的方面或本发明的其它方面的特征组合。

[0147] 实例

[0148] 实例中使用的培养基

[0149] PM074是通过向1升终体积的Instant Ocean盐溶液(35g/L)(水生生态系统(Aquatic Eco Systems),阿波普卡,佛罗里达州)中添加1.3ml PROLINE®F/2藻类饲料部分A(水生生态系统)和1.3ml PROLINE®F/2藻类饲料部分B(水生生态系统)制成的10X F/2的富氮(“仅硝酸盐”)培养基。脯氨酸A和脯氨酸B一起包括8.8mM NaNO₃、0.361mM NaH₂PO₄·H₂O、10X F/2痕量金属和10X F/2维生素(Guillard (1975) “用于饲养海洋无脊椎动物的浮游植物培养(Culture of phytoplankton for feeding marine invertebrates)”,《海洋无脊椎动物培养(Culture of Marine Invertebrate Animals)》(编辑:Smith W.L.和Chanley M.H.) 普莱纽姆出版社(Plenum Press),纽约,美国,pp 26-60)。

[0150] PM124培养基是补充有5mM铵和10mM HEPES且pH为8.0的PM074。它是通过在PM074配方中加入10mls的1M HEPES pH 8和5mls的NH₄Cl制成的(终体积为1L)。通过调节PM074的铵浓度并添加其它Hepes缓冲剂制得其它铵水平受控的培养基。

[0151] 相对于野生型微拟球藻属,PM074和PM124培养基是富氮的和富营养物的。

[0152] 实例1

[0153] 影响脂质生物合成的含GAF结构域的多肽的鉴定

[0154] 创建卡迪斯微拟球藻的转基因藻株,其中使用CRISPR技术通过定向诱变在野生型卡迪斯微拟球藻藻株(命名为WT-3730)中功能性切除或敲除位于Naga_100020g79基因座的基因,所述基因编码信号转导蛋白中发现的含有未知功能的保守结构域和GAF2结构域(Pfam ID:PF13185;由SEQ ID NO:1表示)的蛋白质。在Naga_100020g79基因座处的卡迪斯微拟球藻基因的编码区的核苷酸序列提供为SEQ ID NO:3,并且SEQ ID NO:2代表编码蛋白的氨基酸序列。所述基因被命名为STP-4551基因。STP-4551的GAF2结构域的氨基酸序列提供为SEQ ID NO:1。

[0155] 为了产生敲除突变体,开发了一种高效的微拟球藻Cas9编辑器品系(卡迪斯微拟球藻藻株GE-6791),基本上如共同待决的申请US 2017/0073695“用于高效体内基因组编辑的组合物和方法(Compositions and Methods for High Efficiency In Vivo Genome Editing)”中所公开的。表达编码酿脓链球菌Cas9核酸酶的基因并且还包括杀稻瘟菌素抗性基因和绿色荧光蛋白基因(TurboGFP, Evrogen)的藻株GE-6791,通过用引导RNA靶向来修饰以除去GFP基因。通过电穿孔进行转化,基本上如美国专利申请公开US 2014/0220638中所公开的,所述专利通过引用并入本文。所得藻株GE-13038包括Cas9基因和杀稻瘟菌素抗性基因,并且是GFP⁻(不表达GFP)。将GE-13038藻株用作由嵌合引导RNA和供体DNA转化的宿主以进行插入敲除。因此,工程化并分离微拟球藻藻株,所述藻株包括藻株Cas9表达盒,其

含有来自酿脓链球菌的针对卡迪斯微拟球藻进行密码子优化的Cas9基因(SEQ ID NO:5), 其包括编码N端核定位信号的序列(SEQ ID NO:6), 然后是FLAG标签(SEQ ID NO:7)和肽接头(一起提供为SEQ ID NO:8), 由卡迪斯微拟球藻RPL24启动子(SEQ ID NO:9)驱动并由卡迪斯微拟球藻双向终止子2(SEQ ID NO:10)终止, 和选择标记表达盒, 其含有来自土曲霉的针对卡迪斯微拟球藻进行密码子优化的杀稻瘟菌素S脱氨酶基因(SEQ ID NO:11), 由卡迪斯微拟球藻TCTP启动子(SEQ ID NO:12)驱动并且然后是EIF3终止子(SEQ ID NO:13)。

[0156] 位于Naga_100020g79基因座(具有SEQ ID NO:3的序列的ORF;图1)的STP-4551基因是使用Cas9介导的基因组编辑靶向破坏的基因之一。简而言之, 将潮霉素抗性表达盒靶向插入基因, 并制备DNA构建体(SGI-DNA, 拉荷亚, 加利福尼亚州)以产生引导RNA, 其中DNA分子包括在T7启动子下游工程化的嵌合引导序列。嵌合引导序列包括与位于酿脓链球菌Cas9 PAM序列(NGG)上游的位于Naga_100020g79基因座的基因内的序列同源的21bp靶序列(SEQ ID NO:4), 并且还包括反式激活CRISPR RNA(tracr RNA)序列。通过首先制备由互补DNA寡核苷酸(SEQ ID NO:14和SEQ ID NO:15)组成的DNA模板来合成嵌合的引导序列, 所述DNA模板在退火时会产生双链构建体, 其具有紧接在引导靶序列上游的T7启动子, 然后是tracr配偶和tracr序列。将两个DNA寡核苷酸退火以产生双链DNA模板, 根据制造商的说明, 使用MEGAscriptTMT7试剂盒(生命科技公司(Life Technologies), 卡尔斯巴德, 加利福尼亚州#AM1354M)在体外转录反应中使用所述模板以合成引导RNA。根据制造商的方案, 使用Zymo-SpinTMV-E柱(Zymo研究公司(Zymo Research) #C1024-25)纯化所得的RNA。

[0157] 用于插入位于Naga_100020g79基因座的基因中的供体片段包括选择标记盒, 所述选择标记盒包括在卡迪斯微拟球藻EIF3启动子(SEQ ID NO:17)下游的潮霉素抗性基因(HygR, SEQ ID NO:16), 并且然后是卡迪斯微拟球藻双向终止子2(SEQ ID NO:10), 其中整个启动子-潮霉素抗性基因终止子序列侧接有5' (SEQ ID NO:18)和3' (SEQ ID NO:19)末端的27个碱基对识别序列以产生称为“Hyg抗性盒”(SEQ ID NO:20)的DNA片段。

[0158] 为了靶向敲除位于Naga_100020g79基因座的基因, 使用5 μ g靶向位于Naga_100020g79基因座的基因的纯化的嵌合引导RNA和1 μ g可选择的供体DNA(Hyg抗性盒Naga_100020g79 SEQ ID NO:20)通过电穿孔转化Cas9编辑器品系GE-6791, 基本上如US 2014/0220638中所述。电穿孔后, 将细胞铺板在含有潮霉素的PM124琼脂培养基上以选择掺入潮霉素抗性盒的转化体。将转化体贴到新鲜的板上, 并通过菌落PCR进行筛选, 以将供体片段插入到位于Naga_100020g79基因座的基因中。

[0159] 为了进行菌落PCR筛选, 将来自待筛选菌落的少量细胞悬浮于100 μ l的5%Chelex 100树脂(BioRad)/TE溶液中, 并将悬浮液于99 $^{\circ}$ C煮沸10分钟, 然后将试管短暂旋转。将一微升裂解物上清液添加到PCR反应混合物中, 其中PCR混合物和反应根据制造商的QIAGEN快速循环PCR主混合物规程(手册可在Qiagen网站上获得)设置并执行。用于检测供体片段插入靶基因座(Naga_100020g79)中的引物为SEQ ID NO:21和SEQ ID NO:22。基于PCR的菌落筛选鉴定了两个位于Naga_100020g79基因座的STP-4551基因的敲除藻株GE-13535和GE-13536。

[0160] 实例2

[0161] 批次生产力测定中的Naga_100020g79敲除突变体

[0162] 在富氮培养基PM074中在批次生产力测定中评估具有编码含GAF2结构域的多肽的

破坏基因的敲除突变藻株GE-13535和GE-13536,所述培养基包含8.8mM硝酸盐作为细胞可用的唯一氮源且不存在任何可以支持藻类生长的还原碳源(即,在光合自养条件下进行生产力测定)。接种后,在分批测定中,将工程化的敲除菌株和野生型藻株WT-3730在含有175mL PM074培养基的75cm²矩形组织培养瓶中以三份培养物生长7天。在这些条件下,大约在第3天,培养基中的氮开始变得有限,在整个测定的其余时间内,培养基中的氮浓度持续下降。将烧瓶以其最窄的“宽度”尺寸相对于LED灯放置。用不透明的白色塑料遮盖培养烧瓶,以提供21.1cm²的矩形开口,以使辐照度到达培养物中。将入射辐照度编入16小时光照:8小时黑暗循环中,其中在4小时的时段内辐照度从0到1200uE线性斜升,并且然后辐照度从1200到0uE线性斜降。每天将去离子水添加到培养物中以替代蒸发损失。通过设定在25℃的水浴调节培养物的温度。在第0天以0.5的OD₇₃₀接种培养物,并在第3、5和7天取出样品(5mL)以评估细胞密度、作为脂质量度的脂肪酸甲酯(FAME)和总有机碳(TOC)。在光照循环结束前30分钟进行采样。

[0163] 在这些测定中,通过测量代表脂质的脂肪酸甲酯(FAME)和代表生物质的总有机碳(TOC)来评估碳与脂质分配表型;并使用FAME/TOC比来评估藻株与野生型3730藻株相比是否具有增加的碳与脂质分配。

[0164] 对使用GeneVac HT-4X干燥的2mL样品进行FAME分析。向每个干燥的颗粒中添加以下物质:500μL的500mM KOH的甲醇溶液,200μL的含0.05%二丁基羟基甲苯的四氢呋喃,40μL的2mg/mL C11:0游离脂肪酸/C13:0甘油三酯/C23:0脂肪酸甲酯内标混合物和500μL玻璃珠(直径425-600μm)。将小瓶用敞口的PTFE隔膜衬里的盖子盖住,并置于SPEX GenoGrinder中以1.65krpm持续7.5分钟。然后将样品在80℃加热5分钟并使其冷却。为了进行衍生化,将500μL的10%三氟化硼的甲醇溶液添加到样品中,然后在80℃加热30分钟。使管冷却,然后加入2mL庚烷和500μL的5M NaCl。将样品以2K rpm涡旋5分钟,并且最后以1K rpm离心3分钟。使用Gerstel MPS自动进样器对庚烷层进行采样。定量使用80μg的C23:0FAME内标。在Agilent 7890A气相色谱系统上,使用J&W科技(J&W Scientific)127-3212DB-FFAP,10m×100μm×100nm柱和FID检测器,在260℃下运行样品。流速为500μL/分钟,使用H₂作为具有恒定流量控制的载气。将烘箱在100℃下放置0.98分钟,然后以15.301℃/分钟至230℃,并保持1.66分钟。入口含有4mm玻璃棉填充衬管(安捷伦(Agilent)P/N 5183-4647),并设定在250℃,并使用40:1的分流比。进样量为900nL。通过用去离子水将2mL细胞培养物稀释至20mL的总体积来确定总有机碳(TOC)。每次测量将三次进样注入岛津(Shimadzu)TOC-Vcsj分析仪,以测定总碳(TC)和总无机碳(TIC)。将燃烧炉设定为720℃,并且通过从TC中减去TIC来确定TOC。4点校准范围为2ppm至200ppm,对应于非稀释培养物的20-2000ppm,其中相关系数 $r^2 > 0.999$ 。

[0165] 测定的结果提供在表1和2以及图2和3中。表1和图2展示了在批次生产力测定中在包括8.8mM硝酸盐作为细胞可利用的唯一氮源的富氮培养基PM074中生长的野生型卡迪斯微拟球藻藻株WT-3730的FAME/TOC比,以及在培养的第3、5和7天,突变藻株(GE-13535和GE-13536)的FAME/TOC比。在整个生产力运行中,野生型藻株具有的FAME/TOC比接近0.2,而突变体GE-13535和GE-13536的FAME/TOC比徘徊在0.4左右,因而导致相对于野生型增加95-131%(取决于培养日)。表2和图3展示了运行第5天的FAME曲线;与野生型藻株相比,所述突变体具有高水平的16:0和16:1脂肪酸和低水平的EPA(20:5)。

[0166] 表1野生型和突变藻株的FAME/TOC比

[0167]	日	WT	s.d.	GE-13535	s.d.	增加%	GE-13536	s.d.	增加%
	3	0.175	0.0043	0.341	<0.0000	94.86	0.401	0.0013	129.14
	5	0.180	0.0038	0.397	0.0051	120.56	0.416	0.0034	131.11
	7	0.211	0.0009	0.473	0.0033	124.76	0.477	0.0080	126.07

[0168] 表2.脂肪酸占总脂肪酸的百分比

[0169]		C14:0	C16:0	C16:1	C18:1	C18:1	C18:2	C20:4	C20:5	总FAME
	WT-3730	4.9	27.5	28.9	2.5	0.6	1.9	4.1	28.9	100.0
	WT-3730	5.3	26.3	26.3	3.2	0.6	1.9	3.9	30.7	100.0
	GE-13535	3.1	44.1	34.9	9.2	0.9	0.8	1.6	3.2	100.0
	GE-13536	2.9	42.9	35.0	8.7	0.9	0.9	1.9	3.5	100.0

[0170] 由于两个突变藻株GE-13535和GE-13536源自相同的转化事件,并且因此在Naga_100020g79基因座中含有相同的突变,因此只选择一个突变藻株 (GE-13536) 进行进一步测试。在分批测定中突变藻株GE-13536使用两种不同的氮源培养,一组两份培养物含有8.8mM硝酸盐 (PM074培养基), 另一组两份培养物含有5mM铵和8.8mM硝酸盐 (PM124培养基), 以及在相同条件下的两份野生型培养物,并评估每种条件下与野生型相比的生产力。表3和图4展示,当在仅硝酸盐培养基上生长时,GE-13536比野生型藻株产生显著更多的FAME,显示在第7天的FAME比野生型多约97%;在第5天的FAME比野生型多约116%;以及在第3天的FAME比野生型多122%。当在仅硝酸盐培养基上生长时,与野生型相比,GE-13536还表现出TOC降低(表3和图5)。在测定过程中,通过TOC测量的在仅硝酸盐培养基上培养的GE-13536突变藻株的生物质约为在相同条件下培养的野生型藻株的生物质的62-87%。在仅硝酸盐培养基上生长时,与在相同培养基上生长的野生型藻株相比,在实验过程中观察到GE-13536的FAME/TOC比增加了约154-214% (表5和图6)。当铵存在于培养基中时,所述突变藻株具有与野生型藻株相似的FAME、TOC和FAME/TOC比。因此,在仅硝酸盐培养基上诱导Naga_100020g79敲除突变体进行脂质生物合成,但当培养基中存在铵时则不能诱导。

[0171] 表3野生型和突变藻株GE-13536产生的FAME

	日	平均 s.d. WT(NO3)	平均 s.d. WT (NO3+NH4)	增加%	平均 s.d. GE-13536 (NO3)	增加%	平均 s.d. GE-13536 (NO3+NH4)	增加%
[0172]	3	97.05 +3.15	109.59 +5.56	12.92	215.18 +3.79	121.72	95.09 +1.11	-2.02
	5	150.44 +4.10	163.91 +3.37	8.95	324.58 +9.25	115.75	145.84 +2.24	-3.06
	7	194.59 +10.27	242.05 +15.10	24.39	377.67 ±	94.08	211.41 +4.72	8.64

[0173] 表4.野生型和突变藻株GE-13536产生的TOC

	平均 s.d. 日 WT(NO3)	平均 s.d. WT (NO3+NH4)	增加%	平均 s.d. GE-13536 (NO3)	增加%	平均 s.d. GE-13536 (NO3+NH4)	增加%
[0174]							
3	338.85 ±16.33	405.35 ±23.41	19.62	294.85 +0.92	-12.99	336.85 ±8.13	-0.59
5	528.6 ±25.46	603 ±30.69	14.07	389.55 +6.72	-26.31	537.4 ±17.96	1.66
7	724 ±41.01	821.5 ±41.58	13.47	448.6 ±	-38.04	742.5 ±16.83	2.56

[0175] 表5.野生型和突变藻株GE-13536的FAME/TOC比

	平均 日 WT(NO3)	平均 WT (NO3+NH4)	增加%	平均 GE-13536 (NO3)	增加 %	平均 GE-13536 (NO3+NH4)	增加 %
[0176]							
3	0.287 ±0.0045	0.270 ±0.0019	-5.92	0.7298 ±0.0106	154.29	0.28 ±0.00	-2.44
5	0.285 ±0.0060	0.272 ±0.0083	-4.56	0.8331 ±0.0094	192.32	0.27 ±0.00	-5.26
7	0.269 ±0.0010	0.295 ±0.0035	9.67	0.8445 ±	213.94	0.29 ±0.00	7.81

[0177] 尽管已经根据实例和优选实施例描述了某些实施例,但是应当理解,本领域技术人员将想到各种变化和修改。因此,意图是所附权利要求覆盖落入所附权利要求的范围内的所有此类等同变化。

[0178] 尽管已经参考以上实例描述了本发明,但是应当理解,修改和变型包含在本发明的精神和范围内。因此,本发明仅由所附权利要求限定。

序列表

<110> 合成基因组股份有限公司

I • 阿加维

<120> 通过基因修饰信号传导蛋白提高藻类脂质生产力

<130> SGI2120-1W0

<150> US 62/594,930

<151> 2017-12-04

<160> 22

<170> PatentIn版3.5

<210> 1

<211> 91

<212> PRT

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> STP-4551的GAF2结构域

<400> 1

Leu Ser Gln Trp Ile Tyr Tyr Ser Arg Asn Phe Ala Phe Phe Glu Gly

1 5 10 15

Gln Gly Met Pro Gly Arg Val Phe Arg Ser Gln Val Pro Glu Tyr Leu

20 25 30

Glu Asn Ile Ala Ser Ile Asp Ala Lys Leu Phe Leu Arg Arg Asp Gly

35 40 45

Ala Gln Val Ala Gly Val His Ser Ser Phe Gly Val Pro Leu Val Gln

50 55 60

Asp Asp Ser Val Ile Val Leu Val Phe Tyr Ser Asn His Asp Ser Ser

65 70 75 80

Ser His Ile Thr Gln Glu Leu Leu Asp Phe Val

85 90

<210> 2

<211> 1050

<212> PRT

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> STP-4551蛋白

<400> 2

Met Ala Ser Ala His Gly Thr Ser Gly Ala Asn Val Met Leu Leu Ser

1	5	10	15
Gly Gly Met Gly Asp Gly Lys Ala Phe Arg Glu Ala Ser Cys Pro Cys			
20	25	30	
Pro Pro Phe Gln His Ser Ser Pro Thr Thr Lys Leu Tyr Gln Asn Gln			
35	40	45	
Gly Gly Arg Gly Glu Gln Thr Met Pro Val Thr His Leu His Val Leu			
50	55	60	
Pro Pro Arg Val Gln Lys Val Ser Ile Leu Ser Asn Asn Phe Pro Gln			
65	70	75	80
Leu Gln Pro Lys Ala Gly Ser Tyr Cys Cys Lys Cys Gly Cys Pro Arg			
85	90	95	
Ala Asp Val Arg Leu Thr Cys Cys Asn Ser Val Val His Gly Arg Cys			
100	105	110	
Val Tyr Pro Trp Pro Val Thr Arg Cys Pro Ser Cys Ser Met Asp Gly			
115	120	125	
Lys Glu Met Ala Phe Glu Val Leu Leu Pro Glu Pro Tyr Thr Arg Asn			
130	135	140	
Gly Ala Thr Leu Ser Leu Ser His Gln Glu Leu Ser Lys Lys Arg Gly			
145	150	155	160
Gly Arg Trp Ser Glu Ala Glu Cys Ala Phe Ala Glu Leu Val Met Glu			
165	170	175	
Gly Phe Gly Ser Gly Met Leu Pro Leu Arg Asp Gly Ser Arg Leu Gly			
180	185	190	
Thr Phe Leu Cys Ser Leu Leu Gln Cys Thr Ser Ala Arg Leu Ser Ser			
195	200	205	
Lys Leu Arg Thr Gly Lys Arg Asn Phe Arg Tyr Leu Val Asp Glu Gly			
210	215	220	
Ile Ser Ala Glu Ser Ile Phe Arg Tyr Glu Gln Glu Gln Arg Gln Leu			
225	230	235	240
Ser Arg Arg Glu Glu Leu Phe Leu His Tyr Leu Ser Pro Glu Glu Ala			
245	250	255	
Ala Leu Ala Ser Arg Ser Leu Gln Leu Glu Trp Arg Leu Gln Phe Leu			
260	265	270	
His Leu Gly Lys Gln Val Gly Val Arg Val Ser Asn Leu Tyr Glu Trp			
275	280	285	
Asp Ala Phe Leu Lys Thr Val Val Pro Ala Ala Pro Lys Ala Val Gly			
290	295	300	
Gly Ala Gly Gly Pro Gln Ala Leu Ser Gly Arg Lys Pro Glu Gly Asn			
305	310	315	320

Ala Glu Asp Asp Thr Thr Ser Pro Pro Thr Phe Ser Arg Lys Arg Arg		
325	330	335
Thr Ile Gly Pro Thr Val Ala Gln Cys Pro Asp Asn His Leu Gln Arg		
340	345	350
Val Arg Gly Ala Arg Gln Ala Ser Ala Glu Gln Asp Gly Ala Leu Pro		
355	360	365
Pro Ser Ser Leu Ala Thr Arg Gly Leu Gly Val Leu Met Leu Arg Leu		
370	375	380
Ala Ala Asp Lys His Arg Asn Phe Thr Val Gln Asn Gly Val Met Ala		
385	390	395
Ala Thr Ser Ile Glu Asp Leu Leu Leu Leu Leu Arg Gln Ala Pro Gly		
405	410	415
Asp Glu Cys Ser Val Leu Ala Arg Ala Gly Asn Gly Ala Val Leu Phe		
420	425	430
Val Ser Val Gly Ile Lys Gly Leu Leu Gly Trp Gly Asp Ala Asp Leu		
435	440	445
Val Gly Thr Pro Leu Ile Asp Ala Lys Ser Arg Gly Gly Cys Ala Thr		
450	455	460
Gly Ser Gly Phe Gln Gly Arg Leu Gly Arg Asn Gln Asn Gly Val Ile		
465	470	475
His Gly Glu Asp Gly Ser Val Leu Gln Ala Ala Val Arg Gln Ser Gln		
485	490	495
Val His Thr Ala Ser Val Leu Glu Arg Leu Gly Asn Pro Gln Leu Pro		
500	505	510
Val Gln Glu His Asp Leu Met Met Val Pro Leu Val Leu Arg Leu Tyr		
515	520	525
Arg Arg Asp Arg Thr Leu Val Arg Arg Gln Leu Thr Val Phe Ala Thr		
530	535	540
Glu Asp Phe Leu Leu Leu His Gly Ser Leu Ala Ala Pro Thr Pro Pro		
545	550	555
Val His Arg Ala Gln Gly Arg Arg Arg Glu Asn Val His Arg Pro Thr		
565	570	575
Lys Gln Cys Gly Tyr Gln Glu Gly Phe Tyr Leu Ser Gln His Glu Gly		
580	585	590
Gln His Thr Asp Asp Tyr Gln Pro Tyr Gln Gln Gln Gln Gln Gln Gln		
595	600	605
Tyr His Glu Glu His Gln Gly Gln Gln Glu Lys Gln Glu Pro Gln Gly		
610	615	620
Pro Gln Gln Trp Arg His Pro His Asp Val Gly His Val Glu Thr Arg		

625	630	635	640
Gln Gly Lys Glu Arg Asn Asp Phe Ser Gly His Leu Ser Lys Ser Ser			
	645	650	655
Leu Thr Leu Ser Leu Arg Asn Leu Ala Leu Ala Pro Ser His Gln Tyr			
	660	665	670
Pro Arg Val Tyr Arg Asn Thr Ser Phe Thr Ser Thr Leu Pro Leu Cys			
	675	680	685
Arg Ala Ser Val Glu Arg Met Gln Ala Ser Ser Leu Ser Thr Asn Tyr			
	690	695	700
Val Gln Pro Ala His Leu Ser Tyr Glu Gln Val Val His Glu Glu Gly			
705	710	715	720
Gly Ala Leu Pro Leu Pro Glu Ser Ile Ile Gly Ser Arg Gly Pro Gln			
	725	730	735
Ala His Ser Ile Thr Thr Ser Glu His Glu Ser Ser Val Pro Ile Ala			
	740	745	750
Glu Gly Gly His Asp Met Asn Ser Cys Ile Leu His Gly Ala Ala Leu			
	755	760	765
Leu Gly Arg Arg Ser Glu Val Ala Pro Thr Leu Val Leu Glu Ser Asn			
	770	775	780
Met Asp Val Glu Ala Arg Val Glu Gly Ala Ile Glu Arg Leu Glu Lys			
785	790	795	800
Trp Gly Ser Gln Leu Ser Glu Glu Gly Tyr Asn Ile Ala Thr Ser Ala			
	805	810	815
Asp Asp Thr Ile Val Pro Arg Lys Gly Glu Ser Gly Leu Ala Glu Ala			
	820	825	830
Glu Val Ile Glu Trp Ala Thr Asp Glu Asp Val Arg Leu Ala Leu Gly			
	835	840	845
Glu His Ala Arg Glu Pro Thr Pro Gln Met Gln Arg Asp Ala Cys Gly			
	850	855	860
Gly Ala Tyr Asp Leu Cys Tyr Ser Pro Pro Ser Thr Phe Glu Glu Thr			
865	870	875	880
Met Arg Ala Phe Leu Leu Gly Leu Pro Ala Gln Cys Phe Gln Ser Ala			
	885	890	895
Asp Ile Trp Val Pro Phe Arg Asp Pro Arg Thr Lys Arg Leu Ser Leu			
	900	905	910
Gln Phe Gly Ser Gly Leu Ala Leu Arg Gly Asp Leu Ser Gln Trp Ile			
	915	920	925
Tyr Tyr Ser Arg Asn Phe Ala Phe Phe Glu Gly Gln Gly Met Pro Gly			
	930	935	940

Arg Val Phe Arg Ser Gln Val Pro Glu Tyr Leu Glu Asn Ile Ala Ser			
945	950	955	960
Ile Asp Ala Lys Leu Phe Leu Arg Arg Asp Gly Ala Gln Val Ala Gly			
	965	970	975
Val His Ser Ser Phe Gly Val Pro Leu Val Gln Asp Asp Ser Val Ile			
	980	985	990
Val Leu Val Phe Tyr Ser Asn His Asp Ser Ser Ser His Ile Thr Gln			
	995	1000	1005
Glu Leu Leu Asp Phe Val Arg Arg Ala Thr Ala Ser Trp Lys Ile			
	1010	1015	1020
Gln Thr Thr Phe Pro Val Pro Lys Ala Leu Thr Ala Asp Pro Asp			
	1025	1030	1035
His Gln Pro Ala Val Val Arg Ser Ala Leu Leu Arg			
	1040	1045	1050

<210> 3

<211> 3153

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> STP-4551开放阅读框 (ORF)

<400> 3

```

atggcgctccg cgcatggcac gtccggcgcg aatgttatgc tcttgagcgg tggcatggga 60
gatggcaaag catttcggga ggcctcttgc cttgtcctc cgtttcaaca ctccagtccg 120
acaaccaagc tttacaaaa ccaaggtggg cgtggtgaac aaactatgcc tgtaacacac 180
ctacatgtac tgccgcctcg cggtcagaag gtgtccatat tgtcgaacaa tttccccag 240
ttgcagccga aggcgggctc ctactgctgc aagtgcggtt gtccacgggc tgatgtcagg 300
ctgacatgtt gcaatagtgt cgtccatggc cgttgtgtat acccatggcc cgtgactcgc 360
tgccccgagct gtagcatgga cgggaaagag atggcttttg aagtgccttct gccggagcct 420
tatacccgca acggcgcaac cttgtcgtc tctcaccagg agttgagtaa aaagcgtggg 480
gggcggtggt cagaagctga gtgtgcattt gcggagctgg ttatggaggg gttcggttcc 540
ggtatgttgc ccctgaggga cggcagtcgc cttggcacct tctgtgtag cttgcttcaa 600
tgcacttcag cgcgcttgc ctccaagtta cgcacgggca agcgaaattt tcggtatttg 660
gtggatgagg gcatttctgc cgaaagcatc ttccgatatg aacaggaaca acgccaactc 720
tcgcggcgcg aggagctctt cctgcattat ctctcccccg aggaggccgc actggcctct 780
cgcagtctgc aactggaatg gcgcctccag ttctgcacc tcggcaaaca agtgggcgta 840
cgggtaagca atttatacga atgggacgct tttctgaaga ccgtggtccc agccgccccg 900
aaagccgtgg gtggtgcagg tggccccag gccctctctg ggcggaagcc cgaagggaac 960
gcagaggatg acaccacgtc gccgccgacc ttttctcgga agaggcggac gatcggtcct 1020

```

```

accgttgac agtgccctga caatcacctt cagcgtgtgc gtggcgcccg gcaggcatcg 1080
gccgagcaag atggcgctct accaccgtct tccctcgca ctcgaggatt aggcgtcctg 1140
atgctacgtc tggccgcgga caaacaccgg aacttcacgg tgcaaaacgg ggtcatggcc 1200
gccacttcaa tcgaggactt gcttctcctc ctgcggcaag ctcttgga cgaatgttca 1260
gtcttagctc gggccgggaa cggagctgtc ctcttcgtct ccgttggtat taaaggcttg 1320
cttgggtggg gcgatgccga cctgggtggc actccgctca tcgatgcgaa gagcagagga 1380
ggatgtgcga cagggtccgg cttccaaggt cgtctcggca ggaaccagaa tggggtcatc 1440
catggagaag acgggagtgt tctgcaggcg gccgttcgtc aaagccaagt tcacactgcc 1500
agtgtgcttg agcgctggg caatccgcag ctgccagttc aagagcatga tttgatgatg 1560
gtgcccctgg ttttacgcct gtatcgtcgg gaccgtactt tagtgcgccg gcagctcaca 1620
gtgttcgcca cggaagactt ccttctcctc cacggtcttc tggcagctcc gactcctccg 1680
gtgcatcgag ctcaaggacg gagacgtgaa aacgtccatc ggccaactaa gcagtgtggc 1740
taccaggaag gattttacct gtcgcagcat gaaggacagc aactgatga ctatcaacct 1800
tatcagcagc agcaacagca gcaataccac gaggagcacc aaggccagca agaaaaaca 1860
gaaccgcaag gtccctcagca atggcggcac ccgcatgacg ttggacacgt tgaacaagg 1920
cagggcaagg agagaaatga tttttcaggc cacttgctga aaagctcgct gacattgtct 1980
cttcggaatc tagcgcttgc tccatcacat caataccctc gtgtctaccg caatacttct 2040
ttcacctcga ccttaccctt ttgtcgagca tctgtgaaa ggatgcaggc atctagttta 2100
tcgaccaatt atgtgcagcc tgcccatcta agctacgagc aggttggttca tgaagagggc 2160
ggggtcttcc ccttacctga gtcgattatc ggctcccgtg gaccgcaagc gcatagcatt 2220
accacgtccg agcatgagag tagtgtaccg atcgtgagg gtggtcatga tatgaatagc 2280
tgcatcttcc atggcgcggc tctactgggc cgccgcagtg aggttgcccc aacattagtc 2340
ctggagtcca acatggacgt ggaagcacgg gtggagggtg ccatcgagcg gttggaaaag 2400
tgggggtctc aattaagtga ggagggttat aatatcgcaa caagcgaga tgacactatc 2460
gtgcctcgta aaggagaaag cgggctggcc gaggcggagg tcatcgagtg ggcgacggat 2520
gaggacgtga ggctggcgct cggagagcat gctcgtgagc ccacgcctca aatgcagcgt 2580
gatgcctgtg gaggagctta cgacctctgc tactctcctc catccacctt tgaggagaca 2640
atgcgcgcgt ttttgctcgg tctgcctgcc cagtgttcc aaagcgctga tatctgggtg 2700
ccattccggg accctcgcac taagcggtt tcgttcaat ttggcagcgg gttagctctt 2760
cgaggagact tatcacagtg gatttactac tctcgaaact ttgccttttt tgaagggcaa 2820
ggaatgccgg ggcgagtctt tcgctcgcag gtgcccagat acctggagaa tatcgctcc 2880
atcgacgcca aacttttctt acgtcgagat ggggcgcaag tcgctggagt ccaactcatc 2940
ttcgggtgcc ccctcgtgca ggatgactca gtgattgttt tggcttctta ctgaaccac 3000
gactcgtcat ctcatatcac ccaggagctg ctagacttcg ttcgtcgagc cacagcatct 3060
tggaataatcc agaccacgtt tcccgtacca aaggccttga cggcagatcc cgaccatcaa 3120
ccagctgtcg tgcgtcagc tttattaagg tag 3153

```

<210> 4

<211> 21

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> STP-4551基因中的靶序列

<400> 4

acgctgaagg tgattgtcag g 21

<210> 5

<211> 4101

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 针对微拟球藻属优化的Cas9基因密码子

<400> 5

gacaagaagt actccatcgg gctggacatc gggacgaact ccgtgggatg ggccgtgatc 60
acagacgaat acaaggtgcc ttccaagaag ttcaaggtgc tggggaacac ggacagacac 120
tccatcaaga agaacctcat cggggccttg ctcttcgact ccggagaaac cgccgaagca 180
acgcgattga aaagaaccgc cagaagacga tacacacgac ggaagaaccg catctgtctac 240
ctccaggaga tcttcagcaa cgagatggcc aaggtggacg actcgttctt tcatcgcctg 300
gaggagagct tcctggtgga ggaagacaag aaacatgagc gccacccgat ctccgggaac 360
atcgtggacg aagtggccta ccacgagaaa taccacacga tctaccactt gcgcaagaaa 420
ctcgtggact ccacggacaa agcggacttg cggttgatct acttggcctt ggcccacatg 480
atcaaatttc ggggccactt cctgatcgag ggcgacttga atcccgacaa ttccgacgtg 540
gacaagctct tcatccagct ggtgcagacc tacaaccagc tcttcgagga gaaccccatc 600
aatgcctccg gagtggacgc caaagccatc ttgtccgccc gattgtccaa atccagacgc 660
ttggagaact tgatcgcaca acttcctggc gagaagaaga acggcctctt cggcaacttg 720
atcgcgctgt cgctgggatt gacgcctaac ttcaagtcca acttcgactt ggccgaggac 780
gccaaagtgc aactgtccaa ggacacctac gacgacgacc tcgacaacct gctggcccaa 840
attggcgacc aatacgcgga cttgtttttg gcggccaaga acttgagcga cgccatcttg 900
ttgagcgaca tcttgcgctg gaatacggag atcaccaaag cccctttgtc cgcctctatg 960
atcaagcggc acgacgagca ccaccaagac ttgacctgt tgaaagccct cgtgcggcaa 1020
caattgcccg agaagtacaa ggagatcttc ttcgaccagt ccaagaacgg gtacgccggc 1080
tacatcgacg gaggagcctc ccaagaagag ttctacaagt tcatcaagcc catcctggag 1140
aagatggacg gcaccgagga gttgctcgtg aagctgaacc gcgaagactt gttgcgaaaa 1200
cagcggacgt tcgacaatgg cagcatcccc caccaaatcc atttgggaga gttgcacgcc 1260
atcttgcgac ggcaagagga cttctacccg ttctgaagg acaaccgca gaaaatcgag 1320
aagatcctga cgttcagaat cccctactac gtgggaccct tggcccagg caattcccgg 1380

```

tttgcattgga tgacgcgcaa aagcgaagag acgatcaccc cctggaactt cgaagaagtg 1440
gtcgacaaaag gagcatccgc acagagcttc atcgagcgaa tgacgaactt cgacaagaac 1500
ctgccccaacg agaaggtgtt gcccaagcat tcgctgctgt acgagtactt cacggtgtac 1560
aacgagctga ccaaggtgaa gtacgtgacc gagggcatgc gcaaaccgcg gttcctgtcg 1620
ggagagcaaaa agaaggccat tgtggacctg ctgttcaaga ccaaccggaa ggtgaccgtg 1680
aaacagctga aagaggacta cttcaagaag atcgagtgcg tcgactccgt ggagatctcc 1740
ggcgtggagg accgattcaa tgcctccttg ggaacctacc atgacctcct gaagatcatc 1800
aaggacaagg acttcctgga caacgaggag aacgaggaca tcctggagga catcgtgctg 1860
accctgaccc tgttcgagga ccgagagatg atcgaggaaac ggttgaaaac gtacgcccac 1920
ttgttcgacg acaaggtgat gaagcagctg aaacgccgcc gctacaccgg atggggacga 1980
ttgagccgca aactgattaa tggaattcgc gacaagcaat ccggaagac catcctggac 2040
ttcctgaagt ccgacgggtt cgccaaccgc aacttcatgc agtcatcca cgacgactcc 2100
ttgaccttca aggaggacat ccagaaggcc caagtgtccg gacaaggaga ctcttgcac 2160
gagcacatcg ccaatttggc cggatcccc gcaatcaaaa aaggcatctt gcaaaccgtg 2220
aaagtggctg acgaactggt gaagtgatg ggacggcaca agcccagaa catcgtgatc 2280
gaaatggccc gcgagaacca aaccacccaa aaaggacaga agaactcccg agagcgcatg 2340
aagcggatcg aagaggcat caaggagttg ggctcccaga tcctgaagga gcatcccgtg 2400
gagaataccc aattgcaaaa cgagaagctc tacctctact acctccagaa cgggcgggac 2460
atgtacgtcg accaagagct ggacatcaac cgcctctccg actacgatgt ggatcatatt 2520
gtgccccaga gcttcctcaa ggacgacagc atcgacaaca aggtcctgac gcgcagcgac 2580
aagaaccggg gcaagtctga caatgtgcct tccgaagaag tcgtgaagaa gatgaagaac 2640
tactggcggc agctgctcaa cgccaagctc atcacccaac ggaagttcga caacctgacc 2700
aaggccgaga gaggaggatt gtccgagttg gacaaagccg gcttcattaa acgccaactc 2760
gtggagaccc gccagatcac gaagcacgtg gcccaaatct tggactcccg gatgaacacg 2820
aaatacgacg agaatgacaa gctgatccgc gaggtgaagg tgatcacgct gaagtccaag 2880
ctggtgagcg acttcggaa ggacttccag ttctacaagg tgcgggagat caacaactac 2940
catcacgccc atgacgccta cctgaacgcc gtggtcggaa ccgccctgat caagaaatac 3000
cccaagctgg agtccgaatt cgtgtacgga gattacaagg tctacgacgt gcggaagatg 3060
atcgcgaaagt ccgagcagga gatcggcaaa gccaccgcca agtacttctt ttactccaac 3120
atcatgaact tcttcaagac cgagatcacg ctgccaacg gcgagatccg caagcgcccc 3180
ctgatcgaga ccaacggcga gacgggagag attgtgtggg acaaaggaag agattttgcc 3240
acagtgcgca aggtgctgtc catgcctcag gtgaacatcg tgaagaagac cgaggtgcaa 3300
acaggagggt tttccaaaaga gtccattttg cctaagagga attccgacaa gctcatcgcc 3360
cgcaagaagg actgggaccc caagaagtac gggggcttcg actccccac ggtggcctac 3420
tccgtgttgg tgggtggcaa agtggagaaa gggaagagca agaagctgaa atccgtgaag 3480
gagttgctcg gaatcacgat catggaacga tcgtcgttcg agaaaaacc catcgacttc 3540
ctcgaagcca aagggtacaa agaggtgaag aaggacctga tcatcaagct gcccaagtac 3600
tcctgttcg agctggagaa cggccgcaag cggatgctgg cctccgccg ggaactgcag 3660
aaagggaacg aattggcctt gccctccaaa tacgtgaact tcctctactt ggctcccat 3720

```

tacgaaaagc tcaaaggatc ccctgaggac aatgagcaga agcaactctt cgtggaacaa 3780
cacaagcact acctggacga gatcatcgag cagatcagcg agttctccaa gcgcgtgac 3840
ctcgcgcgacg ccaacctgga caaggtgctc tccgcctaca acaagcacccg cgacaagcct 3900
atccgcgagc aagccgagaa tatcattcac ctgtttaccc tgacgaattt gggagcccct 3960
gccgccttta aatactttga caccaccatc gaccgcaaaa gatacacctc caccaaggaa 4020
gtcttggacg ccaccctcat ccaccagtcc atcacgggcc tctacgagac gcgcatcgac 4080
ctctcccaat tgggcggcga c 4101

<210> 6

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 核定位序列

<400> 6

cccaagaaaa agcggaaggt cggc 24

<210> 7

<211> 24

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> FLAG标签

<400> 7

gactacaagg atgacgatga caag 24

<210> 8

<211> 147

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 核定位序列-肽接头-FLAG标签

<400> 8

atgcccaga aaaagcggaa ggtcggcgac tacaaggatg acgatgacaa gttggagcct 60
 ggagagaagc cctacaaatg ccctgagtgc ggaaagagct tcagccaatc tggagccttg 120
 acccggcacac aacgaacgca tacacga 147

<210> 9

<211> 1000

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> RPL24启动子

<400> 9

aataagcata catcatatga atacaattca gcttaaattt atcatacaaa gatgtaagt 60
 cagcgtgggt ctgtaacgat cgggcgtaat ttaagataat gcgaggacc gggggaggtt 120
 ttggaacgga atgaggaatg ggtcatggcc cataataata atatgggtt ggtcgccctcg 180
 cacagcaacc gtacgtgcga aaaaggaaca gatccattta ataagttgaa cgttattctt 240
 tcctatgcaa tgcgtgtatc ggaggcgaga gcaagtcata ggtggctgcg cacaataatt 300
 gagtctcagc tgagcgccgt ccgcggttg tgtgagtgg catcctctc ccggcctatc 360
 gctcacatcg cctctcaatg gtggtggtg ggcctgatat gacctcaatg ccgaccata 420
 ttaaaaccca gtaaagcatt caccaacgaa cgaggggctc ttttgtgtgt gttttgagta 480
 tgattttaca cctctttgtg catctctctg gtcttccttg gttcccgtag tttgggcatc 540
 atcactcacg cttccctcga ccttcgttct tcctttacaa ccccgacaca ggtcagagtt 600
 ggagtaatca aaaaaggggt gcacgaatga gatacatag attttgacag atatcctttt 660
 actggagagg gttcaaggga tcaaatgaac agcgggctt ggcaatctag ggagggatcg 720
 gaggttggca gcgagcgaaa gcgtgtccat ctttttggt gtcacacctc acgaaccaac 780
 tgtttagcagg ccagcacaga tgacatacga gaatctttat tatatcgtag accttatgtg 840
 gatgaccttt ggtgctgtgt gtctggcaat gaacctgaag gcttgatagg gaggtggctc 900
 ccgtaaacc tttgtccttt ccacgtgag tctccccgc actgtccttt atacaaattg 960
 ttacagtcac ctgcaggcgg tttttctttg gcaggcaaag 1000

<210> 10

<211> 317

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> 双向终止子2

<400> 10

agtgatgcgg cctttaggaa acaccacaaa agtaattgac aatctcagga acgatctgcg 60
 tgtttacagc ttcccaaata acaattatac cacgtacaa aagggttta atgtatctca 120
 caaattcttc taataggtac agcttctcaa attgggtgta tgatgtgaca cttcgtctca 180

cacacgtcac gataattcag cgtatggctt cccttcatca cattcacgca aacttctaca 240
 caaccctggg cataatttctt gtgttggcaa cactcccga atcgattctg cacacaatgg 300
 ttcattcaat gattcaa 317

<210> 11

<211> 399

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 来自土曲霉的针对卡迪斯微拟球藻进行密码子优化的杀稻瘟菌素S脱氨酶基因

<400> 11

atggccaagc ctttatccca agaggaatcc acgctgatcg aacgtgcaac tgcgaccatc 60
 aacagcatac ctattagcga ggactactcg gtggccagtg cagccctctc gtccgacggt 120
 cggatcttta ccggcgtgaa tgtatatcat ttcaccggag ggccatgcgc ggagctcgtg 180
 gtccctcgaa cggccgctgc ggctgctgcc ggaaatctga cgtgcatagt ggccatcggg 240
 aacgaaaacc gcggcattct gtctccgtgc gggcgatgtc ggcaggtgct gcttgacttg 300
 caccggggga tcaaggcaat tgtcaaagat tccgatgggc agcccacagc ggttggcatc 360
 agggagttgc ttccctctgg ctacgtctgg gagggttga 399

<210> 12

<211> 999

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> TCTP启动子

<400> 12

cgtgcaggtg tacagattga aggaacaat ggagatatct ttggcagttg aaaaccgtgt 60
 tcgaatcatg cttttctact ctccaactga gacgaaattt atagcgccat gtcgcttctg 120
 actaccaggc ttaggaaggc ctcatcacia gctggatcgg ttcgaattaa gcaggcactg 180
 aagccaagct tgcaagacag ccacctttta attccctcaa aacactttct caattcagcc 240
 cggtaaatat gccgattcac agcggccaag atagagggga ggtagcaag aatgttgca 300
 tccctcccca gtcgttgcc tgcacacaac ctaggccttc acctttccat ggaaaattga 360
 gaagtgaata ttggttttct tacggcatat cagatgaaat catgaccct aaacatgaag 420
 agctgcaggc aaaacacctg ctctggacga gcacgatgaa atctcgagaa cccgccgtac 480
 ttcagttgat cccgcatgat gacggccgcc attgaaataa gccacctcac tttattctag 540
 caccgatttc caccgttggt agggccgaac gaggacaatt tcgtgcgaaa caagcacgaa 600
 cacgcacacg attagtagta cagacgagca gatcgatggc atgcggcacg gtctcgcgtt 660

ctcggcgacc aggacaacgg agcagaggga ggcctgccga gttccgaggg gcatttttagt 720
ccaaaattgt gttgacacgt gaacaagtgg ctgaaaaga ggaaggaaat gcctggggtt 780
cccttcgaga gcgggaaactc gcttgtgcgt catcctagct acccatggtc cctttgtggg 840
ggaggctgtt tcgtcctacc gaatgtgtgg cgctccatgc atcttctgcc tcccaaacca 900
ccaacatgag cacgcgaagg aaggagaaaa aagtggccgc aacgttctct tctcatattt 960
attgtctcat cacaaacata ggtacataat acaacaatc 999

<210> 13

<211> 318

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> EIF3终止子

<400> 13

ggcactgtaa ccccggttcc gctcgacgaa ggctgggagc gccctttcgg tgggataaaa 60
tggatgcttt accgctgcgc ttcggctgag gaagagagaa atgcgagcgg ggatcggggt 120
cctagaaacg aagaaaggag aacaagttcc tggccaaaga aaaacaagac aaataccctc 180
tccaggcctg ggccattac ttttttttgc tgtttcttat acctgcactc gtgctttctct 240
agtctgtcga gaccttacct gatcttcctc cctccatcgc tccccgcccc ccccatccga 300
gcaaccgtcg accatacg 318

<210> 14

<211> 121

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> STP-4551引导RNA有义寡核苷酸

<400> 14

taatacgact cactataacg ctgaaggtga ttgtcagggt tttagagcta gaaatagcaa 60
gttaaaaataa ggctagtccg ttatcaactt gaaaaagtgg caccgagtcg gtgctttttt 120
t 121

<210> 15

<211> 121

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> STP-4551引导RNA反向互补寡核苷酸

<400> 15

aaaaaaagca cgcactcggg gccacttttt caagttgata acgcactagc cttattttta 60
cttgctatatt ctagctctaa aacctgaca atcaccttca gcgttatagt gagtcgtatt 120
a 121

<210> 16

<211> 1029

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> Hyg抗性基因

<400> 16

atggggaaga aaccggaact gaccgctacg tccgtggaga aattccttat tgagaagtgc 60
gactctgtct ccgacttgat gcaactgagc gagggagagg agagtagggc gttctcgttt 120
gacgtagggg gtcggggata cgtgttgagg gttaatagtt gtgcggacgg gttctacaag 180
gatcggtatg tctaccgtca tttcgctcc gccgctctcc ccataccaga ggtactggac 240
attggggagt ttagcgaatc tctcacgtac tgcatctcgc gccgagccca gggagtgcac 300
ttgcaagatc tgcccgaac tgaattgcct gccgttttgc aaccgtggc cgaggccatg 360
gacgcgatcg ctgccgcaga tctgtctcag acgtccggct ttggaccttt tgggccccag 420
ggcatcgggc agtacacgac ctggcgagac ttcatctgcg ccattgccga tcctcacgtc 480
tatcattggc agacagtcac ggatgacacc gtgtctgcat ccgtggccca agcactggac 540
gaactcatgt tgtgggccga ggattgccct gaggtcaggc acctggtgca cgcggatttc 600
ggcagcaata acgtacttac agacaatggc cggattactg ctgtcatcga ctggtccgaa 660
gcgatgtttg gtgatagcca atacgaagtg gcgaacatat tcttctggcg tccctggttg 720
gcgtgcatgg agcagcagac acgctacttt gaacggaggc acccgagct ggccggctcc 780
ccacgactcc gcgcctatat gttgcgtatc ggactcgatc agctttacca gtctctcgtc 840
gacggcaact tcgacgacgc cgcgtgggag cagggccgct gcgacgcgat agtccgcagc 900
ggggtctggg cgggtgggtcg gacccaaatc gcacgccggt cggctgcggt gtggacagac 960
ggctgtgttg aggtgcttgc ggactcgggc aaccgtaggc cgagcacccg accgcgtgca 1020
aaggagtga 1029

<210> 17

<211> 1000

<212> DNA

<213> 卡迪斯微拟球藻

<220>

<221> misc_feature

<223> EIF3启动子

<400> 17

```
tcataatcaa agatgagcca gccacgaagc taccggagaa ttctgtaaga aaaatgttta 60
aagttgaaaa tgctaacagt gaagtgatat ctttttttaa tggagtgttg aggtgaagtc 120
tagcatcgta ggggaaaaca ggattctgtg tcttcattc tactccttga taaagcgaag 180
aaatccgaca aaaccaaaga gattgttcaa gtttaagatt tgtaagcgta caactatgaa 240
cttcttctct ttgtaggcct gagtggtcgt atgcatacga ttcataagat gaatcagtat 300
cgctggattt tgcttaggag taaagcaca ctaagaaaat atgctgcctg gcaggcatcc 360
tgagacatga ggcaagcgac gtagcaattg aatcctaatt taagccaggg catctgtatg 420
actctgttag ttaattgatg aaccaatgag ctttaaaaaa aaatcgttgc gcgtaatgta 480
gttttaattc tccgccttga ggtgcggggc catttcggac aaggttcttt ggacggagat 540
ggcagcatgt gtcccttctc caaattggtc cgtgtggtag ttgagatgct gccttaaaat 600
tctgtcgggt catcctgcct tcgcattcac tcctttcgag ctgtcgggtt cctcacgagg 660
cctccgggag cggattgcgc agaaaggcga cccggagaca cagagacat acaccgacta 720
aattgcactg gacgatacgg catggcgacg acgatggcca agcattgcta cgtgattatt 780
cgccttgtca ttcagggaga aatgatgaca tgtgtgggac ggtctttaca tgggaagagg 840
gcatgaaaat aacatggcct ggcgggatgg agcgtcacac ctgtgtatgc gttcgatcca 900
caagcaactc accatttgcg tcggggcctg tctccaatct gctttaggct acttttctct 960
aathtagcct attctataca gacagagaca cacagggatc 1000
```

<210> 18

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 27个核苷酸的5' ID序列

<400> 18

```
tccacagccc gaacccatga gagagaa 27
```

<210> 19

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 27个核苷酸的3' ID序列

<400> 19

gcccgaatcg agttgatggc ccgcaaa 27

<210> 20

<211> 2400

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 具有侧翼ID序列的HygR盒

<400> 20

```
tccacagccc gaacccatga gagagaatca taatcaaaga tgagccagcc acgaagctac 60
cggagaattc tgtaagaaaa atgtttaaag ttgaaaatgc taacagtga gtgatatcct 120
tttttaatgg agtggtgagg tgaagtctag catcgtaggg gaaaacagga ttctgtgtct 180
tccattctac tccttgataa agcgaagaaa tccgacaaaa ccaaagagat tgttcaagtt 240
taagatttgt aagcgtacaa ctatgaactt cttctctttg taggcctgag tggtcgtatg 300
catacgattc atgaagtga tcagtatcgc tggattttgc ttaggagtaa agcacaacta 360
agaaaatatg ctgcctggca ggcatcctga gacatgaggg aagcgacgta gcaattgaat 420
cctaatttaa gccagggcat ctgtatgact ctgttagtta attgatgaac caatgagctt 480
taaaaaaaaa tcgttgcgcg taatgtagtt ttaattctcc gccttgaggt gcggggccat 540
ttcggacaag gttctttgga cggagatggc agcatgtgtc cttctccaa attggtccgt 600
gtggttagtt agatgctgcc ttaaaattct gctcggctcat cctgccttcg cattcactcc 660
tttcgagctg tcgggttcct cacgaggcct ccgggagcgg attgcgcaga aaggcgaccc 720
ggagacacag agaccataca ccgactaaat tgcactggac gatacggcat ggcgacgacg 780
atggccaagc attgctacgt gattattcgc cttgtcattc agggagaaat gatgacatgt 840
gtgggacggc ctttacctgg gaagagggca tgaaaataac atggcctggc gggatggagc 900
gtcacacctg tgtatgcgtt cgatccacaa gcaactcacc atttgcgtcg gggcctgtct 960
ccaatctgct ttaggctact tttctctaatt ttagcctatt ctatacagac agagacacac 1020
agggatcatg gggaagaaac cggaactgac cgctacgtcc gtggagaaat tccttattga 1080
gaagttcgac tctgtctccg acttgatgca actgagcgag ggagaggaga gtagggcggt 1140
ctcgtttgac gtagggggtc ggggatacgt gttgagggtt aatagttgtg cggacgggtt 1200
ctacaaggat cggtatgtct accgtcattt cgcctccgcc gctctcccca taccagaggt 1260
actggacatt ggggagttta gcgaatctct cacgtactgc atctcgcgcc gagcccaggg 1320
agtgacgttg caagatctgc ccgaaactga attgcctgcc gttttgcaac ccgtggccga 1380
ggccatggac gcgatcgctg ccgcagatct gtctcagacg tccggctttg gaccttttgg 1440
gccccagggc atcgggcagt acacgacctg gcgagacttc atctgcgcca ttgccgatcc 1500
```

tcacgtctat cattggcaga cagtcattgga tgacaccgtg tctgcatccg tggcccaagc 1560
 actggacgaa ctcatgttgt gggccgagga ttgccctgag gtcaggcacc tggcgcacgc 1620
 ggatttcggc agcaataacg tacttacaga caatggctcg attactgctg tcatcgactg 1680
 gtccgaagcg atgtttggtg atagccaata cgaagtggcg aacatattct tctggcgtcc 1740
 ctggttggcg tgcatggagc agcagacacg ctactttgaa cggaggcacc cggagctggc 1800
 cggctcccca cgactccgcg cctatatgtt gcgtatcgga ctcgatcagc tttaccagtc 1860
 tctcgtcgac ggcaacttcg acgacgccgc gtgggcgcag ggccgctgcg acgcgatagt 1920
 ccgcagcggg gctgggacgg tgggtcggac ccaaatcgca cgccggtcgg ctgcggtgtg 1980
 gacagacggc tgtgttgagg tgcttgcgga ctggggcaac cgtaggccga gcacccgacc 2040
 gcgtgcaaag gaggattga atcattgaat gaaccattgt gtgcagaatc gatttcggga 2100
 gtgttgccaa cacaagaaat atgcccaggg ttgtgtagaa gtttgctga atgtgatgaa 2160
 gggaagccat acgctgaatt atcgtgacgt gtgtgagacg aagtgtcaca tcatacacc 2220
 aatttgagaa gctgtaccta ttagaagaat ttgtgagata cattaaacc cttttgttac 2280
 gtggtataat tgttatttgg gaagctgtaa acacgcagat cgttcctgag attgtcaatt 2340
 acttttgtgg tgtttcctaa aggccgcac actgcccga tcgagttgat ggcccga 2400

<210> 21

<211> 20

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 检测供体片段插入Naga_100020g79的靶基因座中的正向引物

<400> 21

accctgtcgc acatcctcct 20

<210> 22

<211> 22

<212> DNA

<213> 人工序列

<220>

<223> 合成的

<220>

<221> misc_feature

<223> 检测供体片段插入Naga_100020g79的靶基因座的反向引物

<400> 22

ttctgaagac cgtggtccca gc 22

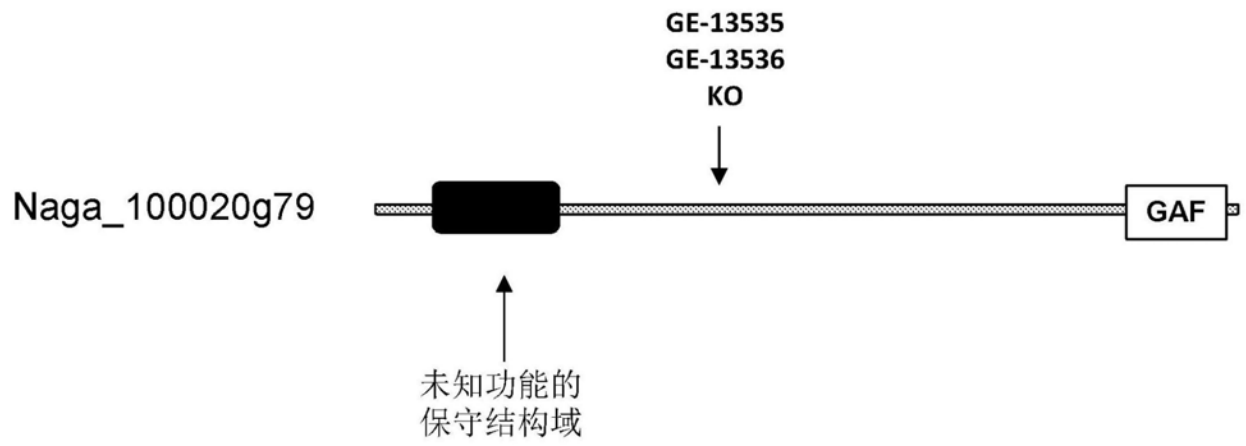


图1

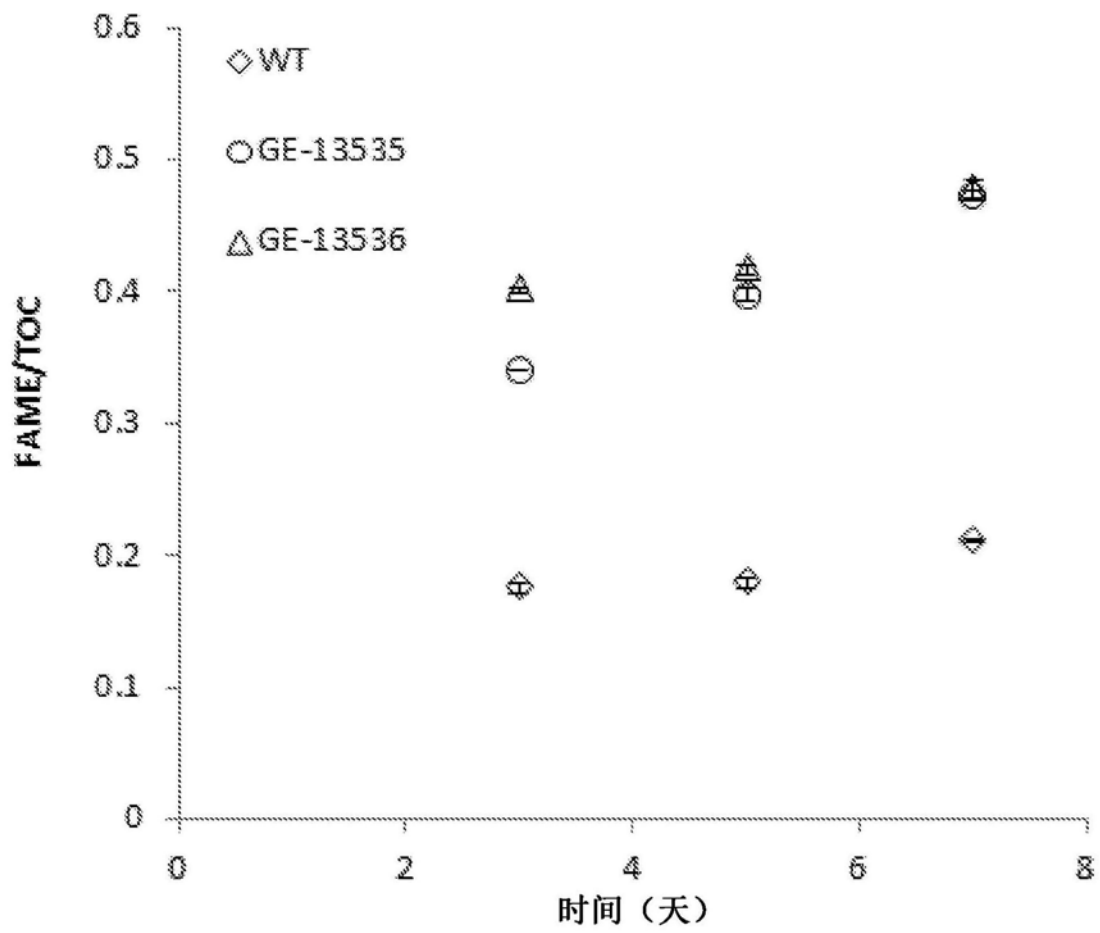


图2

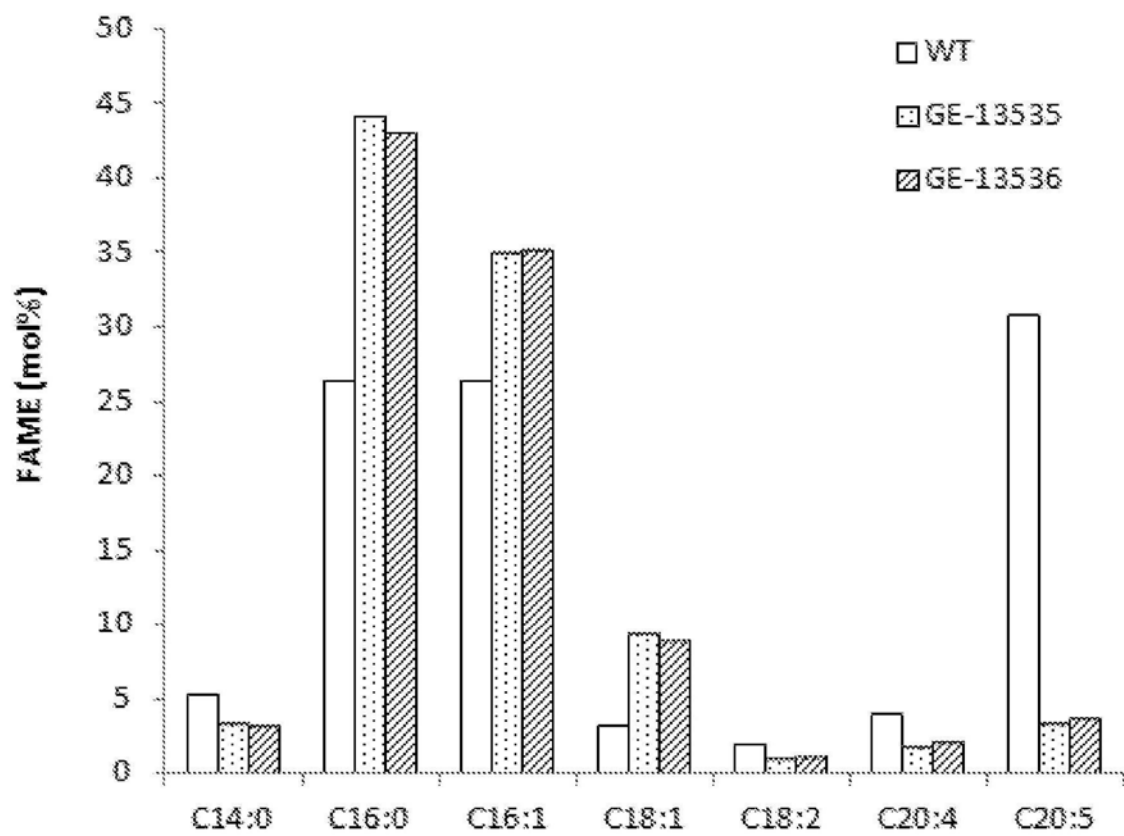


图3

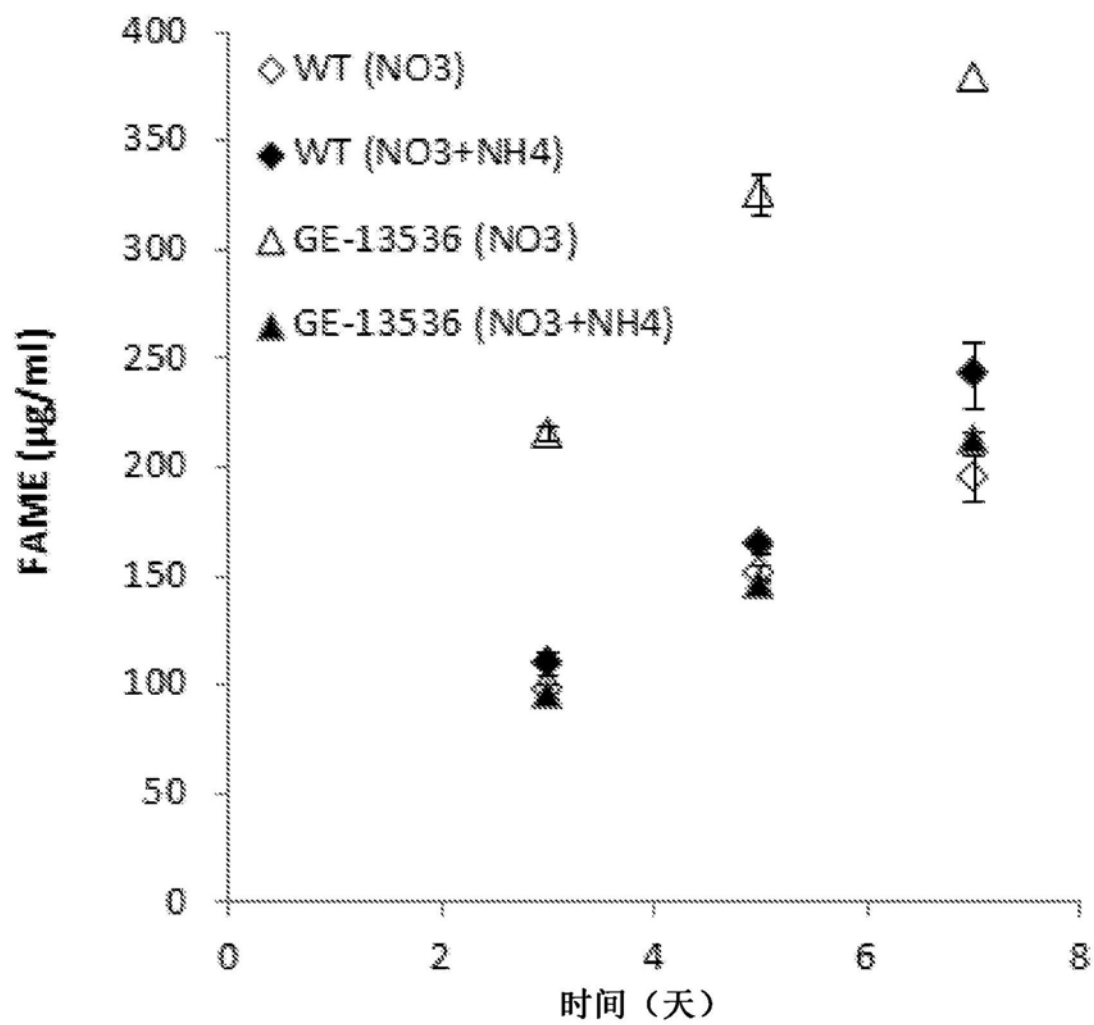


图4

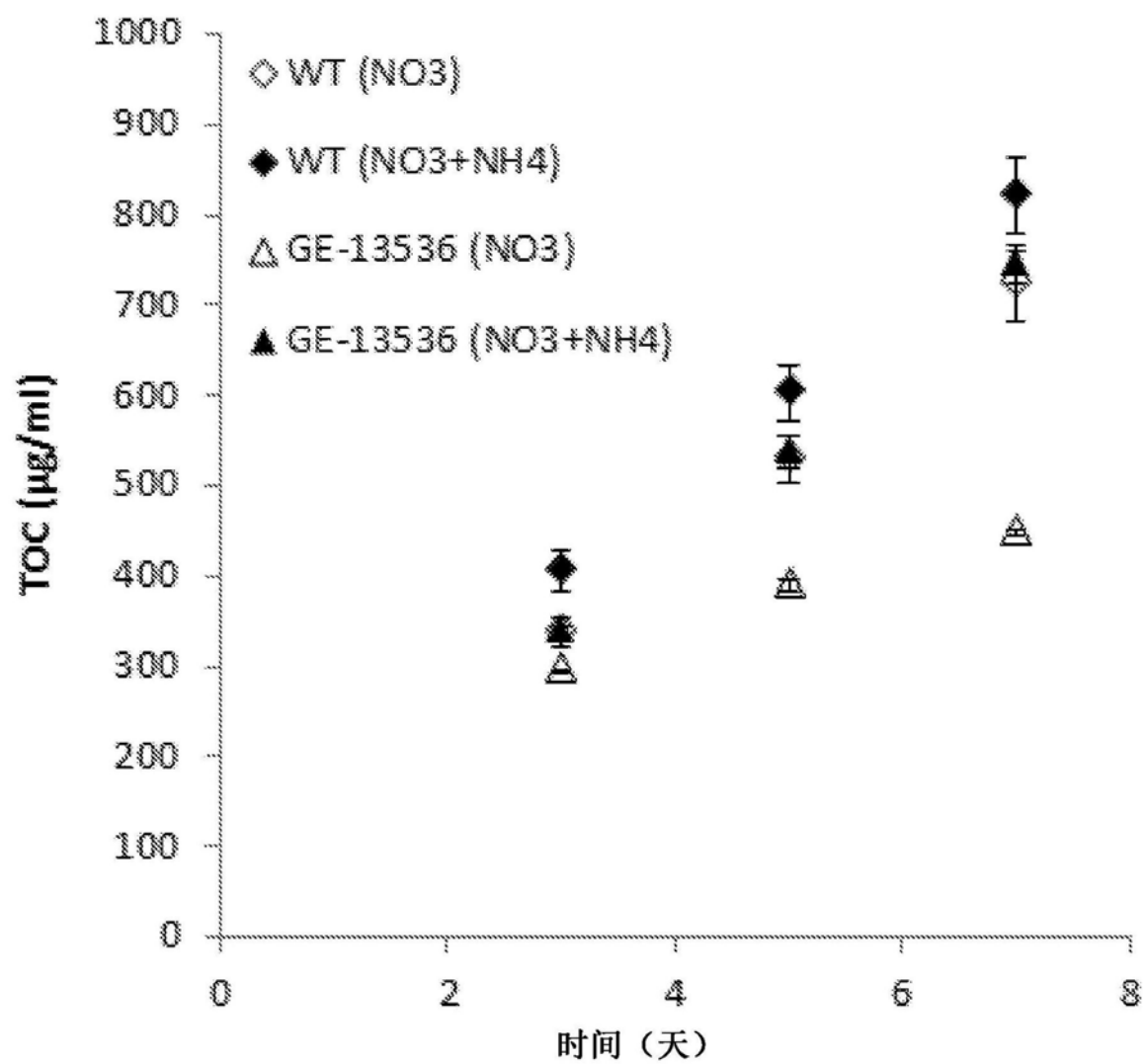


图5

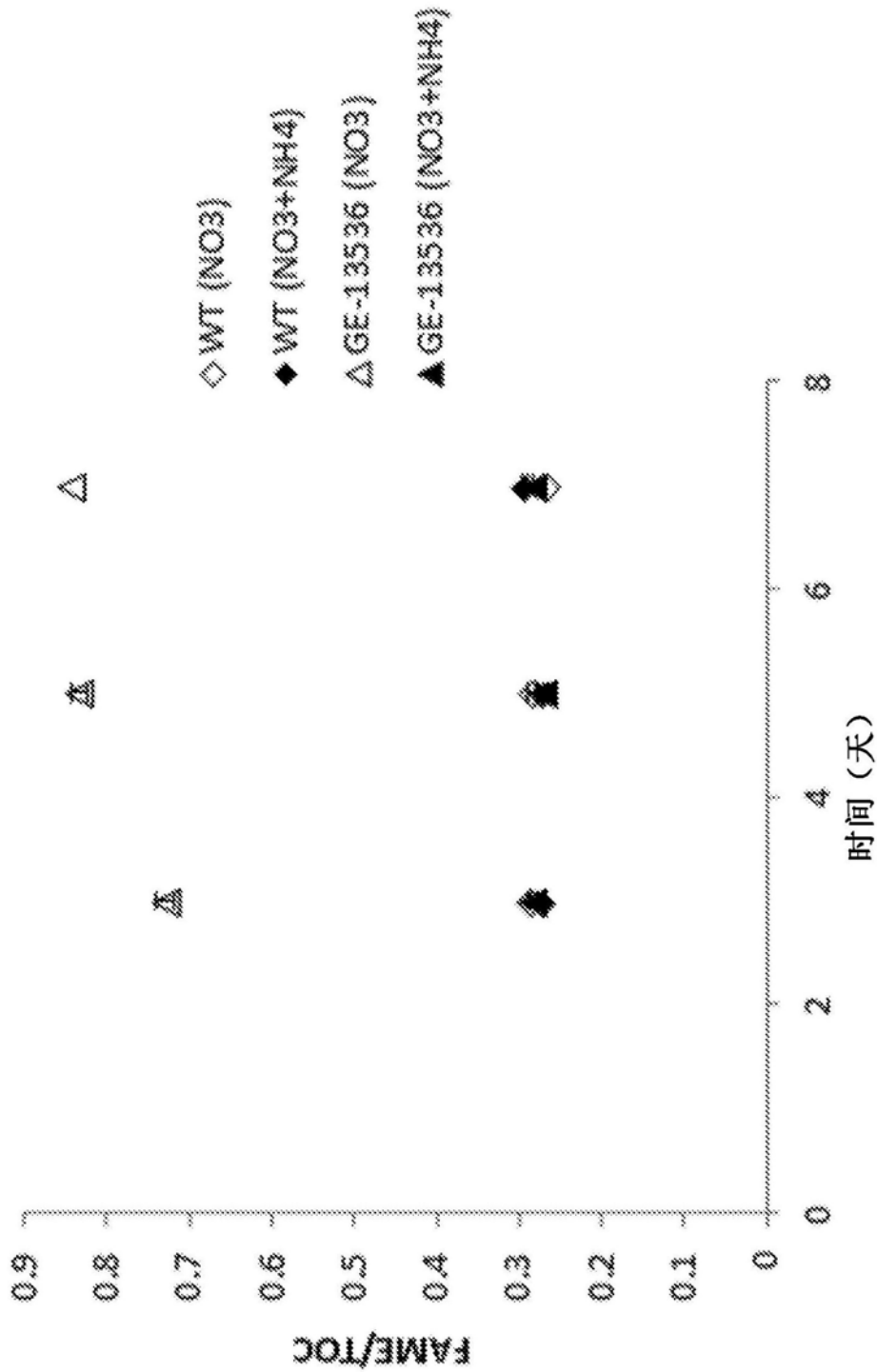


图6

1. 一种突变微生物,其编码包含GAF结构域的多肽的基因表达减弱,其中所述突变微生物:

- a) 比对照微生物多生产至少约25%的脂质;和/或
 - b) 相对于所述对照微生物,表现出增加的碳与脂质分配;
- 当所述突变微生物和对照微生物在相同条件下培养时。

2. 根据权利要求1所述的突变微生物,其中所述GAF结构域是GAF2结构域。

3. 根据权利要求2所述的突变微生物,其中所述GAF2结构域与SEQ ID NO:1所示的氨基酸序列具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的同一性。

4. 根据权利要求2所述的突变微生物,其中所述GAF2结构域包含SEQ ID NO:1所示的氨基酸序列或其保守变体。

5. 根据权利要求1所述的突变微生物,其中所述多肽包含与SEQ ID NO:2所示的氨基酸序列具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列。

6. 根据权利要求1所述的突变微生物,其中所述多肽包含SEQ ID NO:2所示的氨基酸序列或其保守变体。

7. 根据权利要求1所述的突变微生物,其中所述突变微生物包含针对长短鞭毛体或藻类物种中Naga_100020g79基因座或同线基因座中的基因或影响所述基因表达的一种或多种突变。

8. 根据权利要求7所述的突变微生物,其中所述突变微生物包含针对包含开放阅读框的基因或影响包含开放阅读框的基因的表达的突变,所述开放阅读框与SEQ ID NO:3具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%的同一性。

9. 根据权利要求7所述的突变微生物,其中所述一种或多种突变存在于以下核酸中或影响以下核酸的表达:编码包含与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的核酸;和/或包含开放阅读框的核酸,所述开放阅读框包含与SEQ ID NO:3或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的核苷酸序列。

10. 根据权利要求7所述的突变微生物,其中所述一种或多种突变在Naga_100020g79基因座中的基因的GAF结构域中。

11. 根据权利要求1至10中任一项所述的突变微生物,其中所述对照微生物是野生型微生物。

12. 根据权利要求1至11中任一项所述的突变微生物,其中所述突变微生物比对照微生物

物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的脂肪酸甲酯可衍生的脂质(FAME脂质)。

13. 根据权利要求12所述的突变微生物,其中当在包含硝酸盐作为唯一氮源的培养基中培养时,所述突变微生物比对照微生物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的脂肪酸甲酯可衍生的脂质(FAME脂质)。

14. 根据权利要求12或13所述的突变微生物,其中所述突变微生物是藻类,并且当在光合自养条件下培养时,比对照微生物多产生至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%、至少约120%、至少约130%、至少约140%、至少约150%、至少约160%、至少约170%、至少约180%、至少约190%、至少约200%、至少约210%、至少约220%、至少约230%、至少约240%或至少约250%的FAME脂质。

15. 根据权利要求12至14中任一项所述的突变微生物,其中所述微生物比对照微生物多产生至少约15%、至少约20%、至少约25%、至少约30%、至少约35%、至少约40%、至少约45%、至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约100%、至少约110%或至少约120%的贮存脂质。

16. 根据权利要求15所述的突变微生物,其中所述贮存脂质是三酰基甘油酯(TAG)。

17. 根据权利要求1至16中任一项所述的突变微生物,其中所述突变微生物表现出比所述对照微生物的FAME/TOC比高至少约50%、至少约60%、至少约70%、至少约80%、至少约90%、至少约100%、至少约120%、至少约140%、至少约160%、至少约180%、至少约200%、至少约220%、至少约240%、至少约260%、至少约280%、至少约300%、至少约320%、至少约340%、至少约360%、至少约380%或至少约400%的FAME/TOC比。

18. 根据权利要求17所述的突变微生物,其中所述突变微生物表现出至少约0.35的FAME/TOC比。

19. 根据权利要求1至18中任一项所述的突变微生物,其中脂质生产力通过分批、半连续或连续的生产力测定来确定。

20. 根据权利要求1至18中任一项所述的突变微生物,其中所述相同条件包含在包含小于2mM铵的培养基中培养所述突变微生物和对照微生物。

21. 根据权利要求20所述的突变微生物,其中所述相同条件包含在包含硝酸盐作为基本上唯一的氮源的培养基中培养所述突变微生物和对照微生物。

22. 根据权利要求1至21中任一项所述的突变微生物,其中所述相同条件包含在相对于所述对照微生物富营养物的培养基中培养所述突变微生物和对照微生物。

23. 根据权利要求1至22中任一项所述的突变微生物,其中所述突变微生物是经典来源的突变或基因工程突变。

24. 根据权利要求23所述的突变微生物,其中所述突变微生物在编码包含GAF结构域的多肽的基因或影响其表达的基因中具有突变,导致与所述基因在对照微生物中的表达相比,所述编码包括GAF结构域的多肽的基因的表达降低。

25. 根据权利要求23所述的突变微生物,其中所述突变是敲低突变。

26. 根据权利要求25所述的突变微生物,其中所述敲低突变是使用Cas/CRISPR系统产生的。

27. 根据权利要求23所述的突变微生物,其中所述突变微生物包含靶向编码具有GAF结构域的多肽的基因或影响其表达的基因的RNAi构建体、核酶构建体或反义构建体。

28. 根据权利要求24所述的突变微生物,其中所述突变微生物在编码包含GAF结构域的多肽的基因或影响其表达的基因中具有敲除突变。

29. 根据权利要求28所述的突变微生物,其中所述敲除突变是通过位点定向同源重组产生的。

30. 根据权利要求28或29所述的突变微生物,其中所述敲除突变通过部分或全部缺失、截短、移码或插入突变来破坏基因。

31. 根据权利要求28至30中任一项所述的突变微生物,其中所述敲除突变是使用大范围核酸酶、锌指核酸酶、类转录激活因子效应物核酸酶 (TALEN) 系统和/或Cas/CRISPR系统产生的。

32. 根据权利要求26、28、30或31中任一项所述的突变微生物,其中所述突变微生物包含Cas/CRISPR介导的插入所述基因中。

33. 根据权利要求1至32中任一项所述的突变微生物,其中所述突变微生物包含至少一种另外的基因修饰,所述基因修饰赋予除草剂抗性、毒素抗性、增强的生长特性、增强的光合效率、增强的脂质产生或积累以及特定脂质的产生。

34. 根据权利要求1至33中任一项所述的突变微生物,其中所述突变微生物是藻类或长鞭毛体物种。

35. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自以下组成的组的藻类物种:曲壳藻属 (Achnanthes)、茧形藻属 (Amphiprora)、双眉藻属 (Amphora)、纤维藻属 (Ankistrodesmus)、星胞藻属 (Asteromonas)、黄金色藻属 (Boekelovia)、波利氏藻属 (Bolidomonas)、包特氏菌属 (Borodinella)、气球藻属 (Botrydium)、葡萄藻属 (Botryococcus)、片球藻属 (Bracteococcus)、角毛藻属 (Chaetoceros)、四鞭藻属 (Carteria)、衣藻属 (Chlamydomonas)、绿球藻属 (Chlorococcum)、绿梭藻属 (Chlorogonium)、小球藻属 (Chlorella)、蓝隐藻属 (Chroomonas)、金球藻属 (Chrysosphaera)、球钙板藻属 (Cricosphaera)、隐甲藻属 (Crypthecodinium)、隐藻属 (Cryptomonas)、小环藻属 (Cyclotella)、链带藻属 (Desmodesmus)、杜氏藻属 (Dunaliella)、后棘藻属 (Elipsoidon)、圆石藻属 (Emiliania)、独球藻属 (Eremosphaera)、衣迪斯藻属 (Ernodesmus)、眼虫藻属 (Euglena)、真眼点藻属 (Eustigmatos)、披刺藻属

(*Franceia*)、脆杆藻属(*Fragilaria*)、拟脆杆藻属(*Fragilaropsis*)、丽丝藻属(*Gloeothamnion*)、红球藻属(*Haematococcus*)、菱板藻属(*Hantzschia*)、异弯藻属(*Heterosigma*)、膜胞藻属(*Hymenomonas*)、等鞭金藻属(*Isochrysis*)、鳞孔藻属(*Lepocinclis*)、微芒藻属(*Micractinium*)、蒜头藻属(*Monodus*)、单针藻属(*Monoraphidium*)、微球藻属(*Nannochloris*)、微拟球藻属(*Nannochloropsis*)、舟形藻属(*Navicula*)、新绿藻属(*Neochloris*)、肾鞭藻属(*Nephrochloris*)、肾藻属(*Nephroselmis*)、菱形藻属(*Nitzschia*)、棕鞭藻属(*Ochromonas*)、鞘藻属(*Oedogonium*)、卵囊藻属(*Oocystis*)、蚝球藻属(*Ostreococcus*)、拟小球藻属(*Parachlorella*)、雪衣藻属(*Parietochloris*)、杜氏亚属盐藻属(*Pascheria*)、巴夫藻属(*Pavlova*)、普莱格门(*Pelagomonas*)、褐指藻属(*Phoeodactylum*)、噬菌体属(*Phagus*)、皮级微绿藻属(*Picochlorum*)、扁藻属(*Platymonas*)、颗石藻属(*Pleurochrysis*)、肋球藻属(*Pleurococcus*)、原壁菌属(*Prototheca*)、拟绿球藻属(*Pseudochlorella*)、拟新绿藻属(*Pseudoneochloris*)、拟角星鼓藻属(*Pseudostaurastrum*)、塔胞藻属(*Pyramimonas*)、桑椹藻属(*Pyrobotrys*)、栅列藻属(*Scenedesmus*)、裂衣藻属(*Schizochlamydeella*)、骨条藻属(*Skeletonema*)、螺旋藻属(*Spyrogyra*)、裂丝藻属(*Stichococcus*)、四球藻属(*Tetrachlorella*)、四月藻属(*Tetraselmis*)、海链藻属(*Thalassiosira*)、黄丝藻属(*Tribonema*)、无隔藻属(*Vaucheria*)、鲜绿球藻属(*Viridiella*)、魏氏藻属(*Vischeria*)和团藻属(*Volvox*)。

36. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自由以下组成的组的长短鞭毛体:硅藻、黄绿藻、黄藻、褐藻、金藻或针胞藻。

37. 根据权利要求34所述的突变微生物,其中所述突变微生物是选自由网粘菌纲(*Labyrinthulomycete*)物种组成的组的长短鞭毛体,所述物种为网粘菌属(*Labyrinthula*)、类网粘菌属(*Labyrinthuloides*)、破囊壶菌属(*Thraustochytrium*)、裂殖壶菌属(*Schizochytrium*)、不动壶菌属(*Aplanochytrium*)、橙黄壶菌属(*Aurantiochytrium*)、椭圆壶菌属(*Oblongichytrium*)、日本壶菌属(*Japonochytrium*)、戴普罗弗属(*Diplophrys*)或吾肯氏壶菌属(*Ulkenia*)。

38. 根据权利要求34所述的突变微生物,其中所述突变微生物是长短鞭毛体藻类。

39. 根据权利要求38所述的突变微生物,其中所述突变微生物是黄绿藻。

40. 根据权利要求39所述的突变微生物,其中所述突变微生物是选自由以下组成的组的属的物种:小绿藻属、绿囊藻属(*Chlorobryopsis*)、椭圆球藻属、真眼点藻属、角绿藻属、拟单胞藻属、蒜头藻属、微拟球藻属、假拟小棒藻属、拟角星鼓藻属、假四角藻属和魏氏藻属。

41. 根据权利要求38所述的突变微生物,其中所述突变微生物是硅藻。

42. 根据权利要求41所述的突变微生物,其中所述突变微生物是选自由以下组成的组的属的物种:双眉藻属、角毛藻属、小环藻属、脆杆藻属、拟脆杆藻属、菱板藻属、舟形藻属、菱形藻属、褐指藻属或海链藻属。

43. 一种产生脂质的方法,包含在培养基中培养根据权利要求1至42中任一项所述的突变微生物以产生脂质。

44. 根据权利要求43所述的方法,进一步包含从所述微生物、所述培养基或两者分离脂质。

45. 根据权利要求43或44所述的方法,其中所述微生物使用分批、连续或半连续培养条件进行培养。

46. 根据权利要求43至45中任一项所述的方法,其中所述微生物是藻类,并且所述培养是在光合自养条件下进行的。

47. 根据权利要求43至45中任一项所述的方法,其中所述微生物是网粘菌纲(Labyrinthulomycete),并且所述培养是在异养条件下进行的。

48. 根据权利要求43至47中任一项所述的方法,其中所述微生物在光生物反应器或发酵罐中培养。

49. 根据权利要求43至47中任一项所述的方法,其中所述微生物在池塘、运河、海基生长容器、沟渠、水道或渠道中生长。

50. 一种CRISPR系统的引导RNA,其中所述引导RNA包含对应于SEQ ID NO:4的序列。

51. 根据权利要求50所述的引导RNA,其中所述引导RNA是嵌合引导。

52. 根据权利要求50所述的引导RNA,其中所述引导RNA不包括tracr序列。

53. 一种用于同源重组的核酸构建体,其包含来自或邻近以下基因的核苷酸序列:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的藻类基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3和/或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

54. 一种用于表达反义RNA、shRNA、微小RNA或核酶的核酸构建体,其包含与以下基因的至少一部分互补的核苷酸序列:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3和/或SEQ ID NO:4具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

55. 一种编码CRISPR系统的引导RNA的核酸分子,其中所述引导RNA包含以下基因的至少一部分:编码具有与SEQ ID NO:1或SEQ ID NO:2具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的氨基酸序列的多肽的天然存在的藻类基因;位于Naga_100020g79基因座的基因;和/或包含ORF的基因,所述ORF包含与SEQ ID NO:3具有至少约50%、至少约55%、至少约60%、至少约65%、至少约70%、至少约75%、至少约80%、至少约85%、至少约90%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%同一性的序列。

56. 一种用于产生根据权利要求1至42中任一项所述的突变微生物的方法,其包含将减弱包含GAF结构域的多肽表达的一种或多种突变和/或一种或多种试剂引入所述微生物。

57. 根据权利要求56所述的方法, 其中所述一种或多种突变影响包含SEQ ID NO:1和/或SEQ ID NO:2的氨基酸序列的多肽的表达; 位于Naga_100020g79基因座的基因的表达; 或包含开放阅读框的基因的表达, 所述开放阅读框包含SEQ ID NO:3的核苷酸序列。

58. 根据权利要求57所述的方法, 其中所述一种或多种试剂选自由以下组成的组: 反义RNA、RNAi、shRNA、微小RNA、核酶、大范围核酸酶、锌指核酸酶、Cas/CRISPR系统的组分和/或类转录激活因子效应物核酸酶 (TALEN) 系统的组分。

[0001] 申请人修改了原权利要求57的引用关系,使其引用在先的权利要求56。申请人恭敬地认为此修改没有超出原始公开的范围,符合专利法第33条的规定。