



(12) 发明专利申请

(10) 申请公布号 CN 114341351 A

(43) 申请公布日 2022. 04. 12

(21) 申请号 202080040385.8

(22) 申请日 2020.04.24

(30) 优先权数据

62/838,770 2019.04.25 US

(85) PCT国际申请进入国家阶段日

2021.11.30

(86) PCT国际申请的申请数据

PCT/US2020/029946 2020.04.24

(87) PCT国际申请的公布数据

W02020/219972 EN 2020.10.29

(71) 申请人 非凡食品有限公司

地址 美国加利福尼亚州

(72) 发明人 B·罗伊乔杜里 S·尚卡尔

(74) 专利代理机构 北京律盟知识产权代理有限公司 11287

代理人 沈锦华

(51) Int.Cl.

C12N 9/10 (2006.01)

C12N 15/63 (2006.01)

C12P 21/02 (2006.01)

权利要求书3页 说明书36页

序列表36页 附图31页

(54) 发明名称

用于产生含血红素蛋白的菌株和方法

(57) 摘要

本文献涉及用于产生蛋白质的材料和方法。在一个方面中，本文献提供一种细胞，其包含：第一外源核酸构建体，其包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶 (ALAS) 蛋白质进行编码的核苷酸序列，其中所述ALAS包含第一血红素响应基序HRM中的突变和第二HRM中的突变；和第二外源核酸构建体，其包含对血红素结合蛋白进行编码的核苷酸序列，其中包含对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

1. 一种细胞,其包括:

第一外源核酸构建体,其包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白质进行编码的核苷酸序列,其中所述ALAS包括至少第一血红素响应基序HRM,且其中所述ALAS包括所述第一HRM中的突变;和

第二外源核酸构建体,其包括对血红素结合蛋白进行编码的核苷酸序列,

其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

2. 根据权利要求1所述的细胞,其中所述细胞为真菌细胞或为酵母细胞。

3. 根据权利要求2所述的细胞,其中酵母细胞为巴斯德毕赤酵母细胞。

4. 根据权利要求1至3中任一权利要求所述的细胞,其中所述第一HRM中的所述突变为半胱氨酸到不同氨基酸的突变。

5. 根据权利要求1至4中任一权利要求所述的细胞,其中所述ALAS蛋白质包括第二HRM,且其中所述ALAS蛋白质包括所述第二HRM中的突变。

6. 根据权利要求5所述的细胞,其中所述第二HRM中的所述突变为半胱氨酸到不同氨基酸的突变。

7. 根据权利要求5所述的细胞,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述不同氨基酸相同。

8. 根据权利要求5所述的细胞,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述不同氨基酸不相同。

9. 根据权利要求4至8中任一权利要求所述的细胞,其中所述不同氨基酸选自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、谷氨酸、天冬氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。

10. 根据权利要求4至8中任一权利要求所述的细胞,其中所述不同氨基酸选自非极性脂肪族氨基酸、芳香族氨基酸、极性不带电氨基酸或带正电氨基酸。

11. 根据权利要求4至9中任一权利要求所述的细胞,其中所述不同氨基酸为丝氨酸。

12. 根据权利要求4至9中任一权利要求所述的细胞,其中所述不同氨基酸为丙氨酸。

13. 根据权利要求4至9中任一权利要求所述的细胞,其中所述不同氨基酸为苯丙氨酸。

14. 根据权利要求4至9中任一权利要求所述的细胞,其中所述不同氨基酸为组氨酸。

15. 根据权利要求1至14中任一权利要求所述的细胞,其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。

16. 根据权利要求1至15中任一权利要求所述的细胞,其中所述ALAS蛋白质包括与SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

17. 根据权利要求1至16中任一权利要求所述的细胞,其中所述血红素结合蛋白选自由以下组成的群组:球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。

18. 根据权利要求1至16中任一权利要求所述的细胞,其中所述血红素结合蛋白为豆血红蛋白。

19. 根据权利要求1至16中任一权利要求所述的细胞,其中所述血红素结合蛋白包括与SEQ ID NO:1-27中任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

20. 一种在细胞中产生血红素结合蛋白的方法,其包括:

表达第一外源核酸构建体,所述第一外源核酸构建体包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶 (ALAS) 蛋白质进行编码的核苷酸序列,其中所述ALAS包括至少第一血红素响应基序HRM,且其中所述ALAS包括所述第一HRM中的突变;和

表达第二外源核酸构建体,所述第二外源核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列,

其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

21. 根据权利要求20所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的相应方法高至少5%的效价产生所述血红素结合蛋白。

22. 根据权利要求20或权利要求21所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的相应方法高至少10%的效价产生所述血红素结合蛋白。

23. 根据权利要求20所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的相应方法高至少5%的效价产生所述血红素结合蛋白。

24. 根据权利要求20所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的相应方法高至少10%的效价产生所述血红素结合蛋白。

25. 根据权利要求20至24中任一权利要求所述的方法,其中所述方法在不添加甲醇的情况下进行。

26. 根据权利要求20至25中任一权利要求所述的方法,其中所述第一HRM中的所述突变为半胱氨酸到不同氨基酸的突变。

27. 根据权利要求20至26中任一权利要求所述的方法,其中所述ALAS蛋白质包括第二HRM,且其中所述ALAS蛋白质包括所述第二HRM中的突变。

28. 根据权利要求26所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的相应方法高至少5%的效价产生所述血红素结合蛋白。

29. 根据权利要求26所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的相应方法高至少10%的效价产生所述血红素结合蛋白。

30. 根据权利要求27至29中任一权利要求所述的方法,其中所述第二HRM中的所述突变为半胱氨酸到不同氨基酸的突变。

31. 根据权利要求30所述的方法,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述不同氨基酸相同。

32. 根据权利要求30所述的方法,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述不同氨基酸不相同。

33. 根据权利要求26至32中任一权利要求所述的方法,其中所述不同氨基酸选自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、天冬氨酸、谷氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。

34. 根据权利要求26至33中任一权利要求所述的方法,其中所述不同氨基酸为丝氨酸。

35. 根据权利要求26至33中任一权利要求所述的方法,其中所述不同氨基酸为丙氨酸。

36. 根据权利要求26至33中任一权利要求所述的方法,其中所述不同氨基酸为苯丙氨酸。

酸。

37. 根据权利要求26至33中任一权利要求所述的方法,其中所述不同氨基酸为组氨酸。

38. 根据权利要求20至37中任一权利要求所述的方法,其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。

39. 根据权利要求20至38中任一权利要求所述的方法,其中所述ALAS蛋白质包括与SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

40. 根据权利要求20至39中任一权利要求所述的方法,其中所述血红素结合蛋白选自由以下组成的群组:球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。

41. 根据权利要求20至39中任一权利要求所述的方法,其中所述血红素结合蛋白为豆血红蛋白。

42. 根据权利要求20至39中任一权利要求所述的方法,其中异源血红素结合蛋白包括与SEQ ID NO:1-27中任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

用于产生含血红素蛋白的菌株和方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2019年4月25日提交的第62/838,770号美国临时申请的优先权,所述申请的全文通过引用并入本文中。

[0003] 用电子方式提交的文本文件的描述

[0004] 与本申请一起以电子方式提交的文本文件的内容的全文通过引用并入本文中:序列表的计算机可读格式副本,文件名:38767-0158W01.txt,记录日期,2020年4月24日,文件大小约69千字节。

技术领域

[0005] 本公开大体上涉及DNA构建体和使用此类DNA构建体对细胞(例如,酵母细胞(例如甲基营养型酵母细胞))进行基因工程改造的方法。

背景技术

[0006] 例如巴斯德毕赤酵母(*Pichia pastoris*)的细胞常用于表达重组蛋白。本文提供可用于在细胞(例如真菌细胞,例如曲霉(*Aspergillus*)细胞、木霉(*Trichoderma*)细胞或酵母细胞(例如甲基营养型酵母细胞))中高效表达一或多种蛋白质的构建体。

发明内容

[0007] 本文献至少部分地基于对氨基乙酰丙酸合成酶(ALAS)中可引起血红素结合蛋白表达增加的突变的识别。本文所描述的突变型ALAS蛋白质可用于在例如毕赤酵母属(*Pichia*)中高效表达血红素结合蛋白。

[0008] 在一个方面中,本文提供一种甲基营养型酵母细胞,其包含:第一外源核酸构建体,其包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白进行编码的核苷酸序列,其中所述ALAS包含第一血红素响应基序(HRM)中的第一突变;和第二外源核酸构建体,其包括对血红素结合蛋白进行编码的核苷酸序列,其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

[0009] 实施方案可具有以下特征中的一或多个。所述甲基营养型酵母细胞可为毕赤酵母属细胞、假丝酵母属(*Candida*)细胞、汉逊酵母属(*Hansenula*)细胞或球拟酵母属(*Torulopsis*)细胞。所述甲基营养型酵母细胞可为嗜甲醇毕赤酵母(*Pichia methanolica*)细胞、巴斯德毕赤酵母细胞、博伊丁假丝酵母(*Candida boidinii*)细胞或多形汉逊酵母(*Hansenula polymorpha*)细胞。所述甲基营养型酵母细胞可为巴斯德毕赤酵母细胞。所述第一突变可为半胱氨酸到丝氨酸突变。所述第一突变可为半胱氨酸到丙氨酸突变。所述ALAS可包含第二HRM中的第二突变。所述第二突变可为半胱氨酸到丝氨酸突变。所述第二突变可为半胱氨酸到丙氨酸突变。所述第一外源核酸构建体可包含与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。所述第一外源核酸构建体可包含与SEQ ID NO:

28中的核酸序列具有至少95%序列一致性的核酸序列。所述ALAS蛋白质可包含与SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。所述ALAS蛋白质可包含与SEQ ID NO:29中的氨基酸序列具有至少95%序列一致性的氨基酸序列。所述血红素结合蛋白可选自由以下组成的群组:球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。所述血红素结合蛋白可选自由以下组成的群组:雄性球蛋白、血绿蛋白、细胞球蛋白、无脊椎动物血红蛋白(erythrocrutorin)、黄素血红蛋白、球蛋白E、球蛋白X、球蛋白Y、血红蛋白、组胺球蛋白、豆血红蛋白、肌血球素、神经球蛋白、非共生血红蛋白、原球蛋白和截短血红蛋白。所述血红素结合蛋白可为非共生血红蛋白。所述血红素结合蛋白可为豆血红蛋白。所述血红素结合蛋白可包含与SEQ ID NO:1-27中的任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。所述甲基营养型酵母细胞可进一步包含第三核酸构建体,其包括对转录因子进行编码的核苷酸序列,其中所述第三核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件或第三启动子元件。所述第一启动子元件可包含所述转录因子的识别序列。所述第二外源核酸构建体可以可操作地连接到第二启动子元件,且其中所述第二启动子元件包含所述转录因子的识别序列。所述第三核酸构建体可以可操作地连接到所述第三启动子元件,且其中所述第三启动子元件包含所述转录因子的识别序列。所述甲基营养型酵母细胞可进一步包含第四核酸构建体,其包括对参与血红素生物合成的蛋白质进行编码的核苷酸序列,其中所述第四核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件、所述第三启动子元件或第四启动子元件。所述参与血红素生物合成的蛋白质可选自由以下组成的群组:ALA脱水酶、胆色素原脱氨酶、UPG III合成酶、UPG III脱羧酶、CPG氧化酶、PPG氧化酶和亚铁螯合酶。所述第一外源核酸构建体可为异源核酸构建体。所述第二外源核酸构建体可为异源核酸构建体。所述血红素结合蛋白可为外源血红素结合蛋白。所述血红素结合蛋白可为异源血红素结合蛋白。

[0010] 在另一方面中,本文提供一种在甲基营养型酵母细胞中产生血红素结合蛋白的方法,其包含:表达第一外源核酸构建体,所述第一外源核酸构建体包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白进行编码的核苷酸序列,其中所述ALAS包含第一血红素响应基序(HRM)中的第一突变;和表达第二外源核酸构建体,所述第二外源核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列,其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

[0011] 实施方案可具有以下特征中的一或多个。所述ALAS可包含第二HRM中的第二突变。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少5%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少10%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少15%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少20%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变的对应方法高至少5%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变的对应方法高至少10%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变的对应方法高至少15%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变的对应方法高至少20%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述

第一突变和所述第二突变的对应方法高至少5%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少10%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少15%的效价产生所述血红素结合蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少20%的效价产生所述血红素结合蛋白。所述方法可在不添加甲醇的情况下进行。所述甲基营养型酵母细胞可为毕赤酵母属细胞、假丝酵母属细胞、汉逊酵母属细胞或球拟酵母属细胞。所述甲基营养型酵母细胞可为嗜甲醇毕赤酵母细胞、巴斯德毕赤酵母细胞、博伊丁假丝酵母细胞或多形汉逊酵母细胞。所述甲基营养型酵母细胞可为巴斯德毕赤酵母细胞。所述第一突变可为半胱氨酸到丝氨酸突变。所述第一突变可为半胱氨酸到丙氨酸突变。所述第二突变可为半胱氨酸到丝氨酸突变。所述第二突变可为半胱氨酸到丙氨酸突变。所述第一外源核酸构建体可包含与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。所述第一外源核酸构建体可包含与SEQ ID NO:28中的核酸序列具有至少95%序列一致性的核酸序列。所述ALAS蛋白质可包含与SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。所述ALAS蛋白质可包含与SEQ ID NO:29中的氨基酸序列具有至少95%序列一致性的氨基酸序列。所述血红素结合蛋白可选自由以下组成的群组：球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。所述血红素结合蛋白可选自由以下组成的群组：雄性球蛋白、血绿蛋白、细胞球蛋白、无脊椎动物血红蛋白、黄素血红蛋白、球蛋白E、球蛋白X、球蛋白Y、血红蛋白、组胺球蛋白、豆血红蛋白、肌血球素、神经球蛋白、非共生血红蛋白、原球蛋白和截短血红蛋白。所述血红素结合蛋白可为非共生血红蛋白。所述血红素结合蛋白可为豆血红蛋白。异源血红素结合蛋白可包含与SEQ ID NO:1-27中的任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。所述方法可进一步包含表达第三核酸构建体，所述第三核酸构建体包括对转录因子进行编码的核苷酸序列，其中所述第三核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件或第三启动子元件。所述第一启动子元件可包含所述转录因子的识别序列。所述第二外源核酸构建体可以可操作地连接到第二启动子元件，且其中所述第二启动子元件包含所述转录因子的识别序列。所述第三核酸构建体可以可操作地连接到所述第三启动子元件，且其中所述第三启动子元件包含所述转录因子的识别序列。所述方法可进一步包含表达第四核酸构建体，所述第四核酸构建体包括对参与血红素生物合成的蛋白质进行编码的核苷酸序列，其中所述第四核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件、所述第三启动子元件或第四启动子元件。所述参与血红素生物合成的蛋白质可选自由以下组成的群组：ALA脱水酶、胆色素原脱氨酶、UPG III合成酶、UPG III脱羧酶、CPG氧化酶、PPG氧化酶和亚铁螯合酶。所述第一外源核酸构建体可为异源核酸构建体。所述第二外源核酸构建体可为异源核酸构建体。所述血红素结合蛋白可为外源血红素结合蛋白。所述血红素结合蛋白可为异源血红素结合蛋白。

[0012] 在另一方面中，本文提供一种巴斯德毕赤酵母细胞，其包含：第一外源核酸构建体，其包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶 (ALAS) 蛋白进行编码的核苷酸序列，其中所述ALAS包含第一血红素响应基序 (HRM) 中的第一突变和第二HRM中的第二突变；和第二外源核酸构建体，其包括对豆血红蛋白进行编码的核苷酸序列，其中包括对豆血红蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第

一启动子元件,或包括对豆血红蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到第二启动子元件。

[0013] 实施方案可包含以下特征中的一或多个。所述第一突变可为半胱氨酸到丝氨酸突变。所述第二突变可为半胱氨酸到丝氨酸突变。

[0014] 在另一方面中,本文提供一种产生豆血红蛋白的方法,所述方法包含:表达第一外源核酸构建体,所述第一外源核酸构建体包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白进行编码的核苷酸序列,其中所述ALAS包含第一血红素响应基序(HRM)中的第一突变和第二HRM中的第二突变;和表达第二外源核酸构建体,所述第二外源核酸构建体包括对豆血红蛋白进行编码的核苷酸序列,其中包括对豆血红蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件,或包括对豆血红蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到第二启动子元件。

[0015] 实施方案可具有以下特征中的一或多个。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少5%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少10%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少15%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一外源核酸构建体的对应方法高至少20%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少5%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少10%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少15%的效价产生所述豆血红蛋白。所述方法可以比缺少所述第一突变和所述第二突变的对应方法高至少20%的效价产生所述豆血红蛋白。所述方法可在不添加甲醇的情况下进行。

[0016] 在另一方面中,本文提供一种巴斯德毕赤酵母细胞,其包含第一外源核酸构建体,所述第一外源核酸构建体包括对可操作地连接到第一启动子元件的与SEQ ID NO:29具有至少90%序列一致性的蛋白质进行编码的核苷酸序列,其中所述核酸在对应于SEQ ID NO:29的位置12的位置中编码丝氨酸残基,且其中所述核酸在对应于SEQ ID NO:29的位置39的位置中编码丝氨酸残基。

[0017] 实施方案可包含以下特征中的一或多个。所述巴斯德毕赤酵母细胞可进一步包含第二外源核酸构建体,其包括对与SEQ ID NO:1-27中的任一个具有至少90%序列一致性的蛋白质进行编码的核苷酸序列,其中所述第二外源核酸构建体可操作地连接到所述第一启动子元件,或所述第二外源核酸构建体可操作地连接到第二启动子元件。所述巴斯德毕赤酵母细胞可进一步包含第二外源核酸构建体,其包括对与SEQ ID NO:4具有至少90%序列一致性的蛋白质进行编码的核苷酸序列,其中所述第二外源核酸构建体可操作地连接到所述第一启动子元件,或所述第二外源核酸构建体可操作地连接到第二启动子元件。

[0018] 除非另外定义,否则本文使用的所有技术和科学术语都具有与本发明所属领域的普通技术人员通常所理解相同的含义。尽管与本文所描述的方法和材料类似或等效的方法和材料可以用于实践本发明,但是下文描述了合适的方法和材料。本文所提及的所有公开案、专利申请、专利和其它参考文献的全文都通过引用并入。如有冲突,以本说明书(包含定

义)为准。另外,材料、方法和实例仅仅是说明性的并且不打算限制性的。

[0019] 在附图和以下描述中阐述本公开的一或多个实施例的细节。本公开的其它特征、目标和优点将根据所述描述和图式以及权利要求书显而易见。权利要求书中的词语“包括”根据专利法中的标准惯例可经“基本上由……组成”或“由……组成”替代。

附图说明

[0020] 图1提供描绘血红素生物合成途径中所涉及的步骤的示意图。

[0021] 图2提供血红素b的结构。

[0022] 图3提供氯化血红素的结构。

[0023] 图4提供示例性血红素结合蛋白 (SEQ ID NO:1-27) 的序列。

[0024] 图5提供突变型ALAS的示例性构建体示意图。

[0025] 图6提供野生型ALAS的核酸序列 (SEQ ID NO:28) 和氨基酸序列 (SEQ ID NO:29)。

[0026] 图7提供示例性突变型ALAS的核酸序列 (SEQ ID NO:30) 和氨基酸序列 (SEQ ID NO:31)。

[0027] 图8提供SEQ ID NO:32-66的序列。

具体实施方式

[0028] 本文献涉及用于蛋白质产生的材料和方法。确切地说,本文献涉及用于产生血红素和血红素结合蛋白的材料和方法。

[0029] 例如巴斯德毕赤酵母的甲基营养型酵母常用于产生重组蛋白。毕赤酵母属菌株通常能够在使用甲醇作为唯一碳源的情况下生长。应理解,尽管巴斯德毕赤酵母已重新分类为驹形氏酵母属 (*Komagataella*) 菌种,例如法夫驹形氏酵母 (*Komagataella phaffii*)、巴斯德驹形氏酵母 (*Komagataella pastoris*) 或假巴斯德驹形氏酵母 (*Komagataella pseudopastoris*),但术语‘巴斯德毕赤酵母’仍在使用中,且可以指任何适当的驹形氏酵母属菌种。通常,巴斯德毕赤酵母的实验室菌株为法夫驹形氏酵母。

[0030] 结合血红素的蛋白质包含细胞色素、过氧化氢酶、肌血球素和血红蛋白等。氨基乙酰丙酸合成酶 (ALAS) 是催化血红素生物合成 (血红素b生物合成途径的示意图参见图1) 中的第一步骤的酶 (EC 2.3.1.37),其催化甘氨酸和琥珀酰辅酶A (succinyl-CoA) 转化为氨基乙酰丙酸。经由血红素生物合成的此步骤,ALAS转运到线粒体中。氨基乙酰丙酸通过其它酶 (例如,ALA脱水酶 (ALAD)、胆色素原脱氨酶 (PBGD)、尿卟啉原III合成酶 (UPG3S)、尿卟啉原III脱羧酶 (UPG3D)、粪卟啉原氧化酶 (COPROX)、原卟啉原IX氧化酶 (PROTOX) 和/或亚铁螯合酶 (FC)) 转化为血红素b (图2)。在一些情况下,ALAS的作用可为对血红素生物合成途径进行速率限制。其它血红素 (例如,血红素o、血红素a、血红素c) 可经由酶对血红素b的作用而形成。氯化血红素 (图3) 为原卟啉IX (例如,血红素b的原卟啉) 与三价铁 (+3氧化态) 和氯配体的络合物。

[0031] 如本文所使用,“血红素生物合成途径的中间产物”是指以下中的一或多个:δ-氨基乙酰丙酸 (d-ALA)、胆色素原、羟甲基胆素、尿卟啉原III、粪卟啉原III、原卟啉原IX或原卟啉IX。在一些实施例中,血红素生物合成途径的中间产物可选自由以下组成的群组:δ-氨基乙酰丙酸 (d-ALA)、胆色素原、羟甲基胆素、尿卟啉原III、粪卟啉原III、原卟啉原IX和原卟

啉IX。在一些实施例中,血红素生物合成途径的中间产物可选自由以下组成的群组:胆色素原、羟甲基胆素、尿卟啉原III、粪卟啉原III、原卟啉原IX或原卟啉IX。

[0032] 一些ALAS蛋白质到线粒体的转运可受ALAS蛋白质内所含有的一或多个血红素调控基序(HRM;有时也被称为血红素响应基序)影响。许多HRM包含C-P基序,且半胱氨酸通常是血红素的轴向配体。在一些实施例中,HRM具有R/L/N/A/C/S/H/I/G/Q-C-P-L/V/I/F/C-L/M/P/V (SEQ ID NO:32) 的序列(例如,HRM可具有A-C-P-F-V (SEQ ID NO:33)、H-C-P-V-V (SEQ ID NO:34)、I-C-P-F-M (SEQ ID NO:35) 或G-C-P-V-V (SEQ ID NO:36) 的序列;参见例如图8)。一些生物体的ALAS序列中具有3个HRM,而其它生物体(例如甲基营养型酵母,例如巴斯德毕赤酵母)具有两个HRM。一般来说,HRM可被称为HRM1、HRM2、HRM3等,此取决于蛋白质序列中HRM出现的位置,例如蛋白质序列中第一个出现的HRM(从N端到C端读取)将被称为HRM1。HRM中半胱氨酸(例如到丝氨酸或丙氨酸)的突变可增加ALAS到线粒体的转运(参见例如冈萨雷斯-多明格斯(González-Domínguez)等人,酵母(Yeast).2001年1月15日;18(1):41-8.(PubMed ID (PMID) 11124700);宗像(Munakata)等人,生物化学杂志(JBiochem).2004年8月;136(2):233-8.(PMID 15496594);戴利(Dailey)等人,生物化学杂志(Biochem J).2005年3月1日;386(Pt 2):381-6.(PMID 15482256)),其中的每一个的全文通过引用并入本文中)。不受任何特定理论束缚,认为血红素与野生型ALAS的HRM结合可以负反馈方式抑制ALAS到线粒体的转运。还认为血红素参与一些ALAS蛋白质的降解(参见例如久保田(Kubota)等人,生物化学杂志(J Biol Chem).2016年9月23日;291(39):20516-29.数字对象标识符:10.1074/jbc.M116.719161.电子出版2016年8月5日.(PMID 27496948),其全文通过引用并入本文中)。

[0033] 应了解,“第一HRM”可以是蛋白质序列的任何部分中的HRM;“第一HRM”可以(但不一定是)蛋白质序列中第一个出现的HRM(HRM1)。在一些实施例中,第一HRM为HRM1,且第二HRM为HRM2。在一些实施例中,第一HRM为HRM1,且第二HRM为HRM3。在一些实施例中,第一HRM为HRM2,且第二HRM为HRM1。在一些实施例中,第一HRM为HRM2,且第二HRM为HRM3。在一些实施例中,第一HRM为HRM3,且第二HRM为HRM1。在一些实施例中,第一HRM为HRM3,且第二HRM为HRM2。

[0034] 在一些实施例中,核酸中的突变可以是插入、缺失或取代。在一些实施例中,核酸中的突变可以是取代(例如,鸟苷到胞嘧啶突变)。在一些实施例中,编码序列(例如,编码蛋白质的序列)中的取代可以是沉默突变(例如,编码相同氨基酸)。在一些实施例中,编码序列中的取代可以是非同义突变(例如,错义突变或无义突变)。在一些实施例中,编码序列中的取代可以是错义突变(例如,编码不同氨基酸)。在一些实施例中,编码序列中的取代可以是无义突变(例如,编码提前终止密码子)。在一些实施例中,核酸中的突变可以是缺失。应理解,突变可用于使用例如CRISPR、TALEN和/或锌指核酸酶来改变内源核酸。

[0035] 在一些实施例中,蛋白质序列中的突变可以是插入、缺失或取代。应理解,编码蛋白质的核酸中的突变可引起蛋白质序列中的突变。在一些实施例中,蛋白质序列中的突变为取代(例如,半胱氨酸到丝氨酸突变、半胱氨酸到丙氨酸突变、半胱氨酸到缬氨酸突变、半胱氨酸到亮氨酸突变、半胱氨酸到异亮氨酸突变、半胱氨酸到甘氨酸突变、半胱氨酸到苯丙氨酸突变、半胱氨酸到苏氨酸突变、半胱氨酸到甲硫氨酸突变、半胱氨酸到色氨酸突变、半胱氨酸到酪氨酸突变、半胱氨酸到天冬酰胺突变、半胱氨酸到谷氨酰胺突变、半胱氨酸到脯

氨酸突变、半胱氨酸到精氨酸突变、半胱氨酸到组氨酸突变、半胱氨酸到赖氨酸突变、半胱氨酸到天冬氨酸突变、或半胱氨酸到谷氨酸突变)。在一些实施例中,蛋白质序列中的突变可以是缺失。

[0036] 在一些实施例中,如本文所描述的ALAS蛋白质可包含至少一个HRM中的突变;例如第一HRM、第二HRM、第三HRM、第四HRM、第五HRM等。

[0037] 在一些实施例中,如本文所描述的ALAS蛋白质可包含至少一个HRM(例如,第一HRM)中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含单个HRM(例如,第一HRM或第二HRM)中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含至少一个HRM(例如,第一HRM)的半胱氨酸中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含单个HRM(例如,第一HRM或第二HRM)中的半胱氨酸中的突变。

[0038] 在一些实施例中,如本文所描述的ALAS蛋白质可包含至少两个HRM(例如,第一HRM和第二HRM)中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含两个HRM(例如,第一HRM和第二HRM)中的每一个中的突变。在一些实施例中,第一HRM中的突变和第二HRM中的突变可以相同(例如,都是半胱氨酸到丝氨酸或丙氨酸突变)。在一些实施例中,第一HRM中的突变和第二HRM中的突变可以不同(例如,一个是半胱氨酸到丝氨酸突变且一个是半胱氨酸到丙氨酸突变)。在一些实施例中,如本文所描述的ALAS蛋白质可包含至少两个HRM(例如,第一HRM和第二HRM)的半胱氨酸中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含两个HRM(例如,第一HRM和第二HRM)中的每一个中的半胱氨酸中的突变。在一些实施例中,第一HRM中的突变和第二HRM中的突变可以相同(例如,都是半胱氨酸到丝氨酸或丙氨酸突变)。在一些实施例中,第一HRM中的突变和第二HRM中的突变可以不同(例如,一个是半胱氨酸到丝氨酸突变且一个是半胱氨酸到丙氨酸突变)。

[0039] 在一些实施例中,如本文所描述的ALAS蛋白质可包含至少三个HRM(例如,第一HRM、第二HRM和第三HRM)中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含三个HRM(例如,第一HRM、第二HRM和第三HRM)中的每一个中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含至少三个HRM(例如,第一HRM、第二HRM和第三HRM)的半胱氨酸中的突变。在一些实施例中,如本文所描述的ALAS蛋白质可包含三个HRM(例如,第一HRM、第二HRM和第三HRM)中的每一个中的半胱氨酸中的突变。在一些实施例中,第一HRM、第二HRM和第三HRM中的每一个中的突变可以相同(例如,都是半胱氨酸到丝氨酸或丙氨酸突变)。在一些实施例中,三个HRM中的一个或两个中的突变可以不同。

[0040] 在一些实施例中,取代(突变为另一氨基酸)可以是HRM中的氨基酸(例如,半胱氨酸)的取代。在一些实施例中,第一HRM中的突变可以是第一HRM中的氨基酸(例如,半胱氨酸)的取代。在一些实施例中,第二HRM中的突变可以是第二HRM中的氨基酸(例如,半胱氨酸)的取代。在一些实施例中,第三HRM中的突变可以是第三HRM中的氨基酸(例如,半胱氨酸)的取代。取代可以是任何适当的取代。在一些实施例中,另一氨基酸可选自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。在一些实施例中,另一氨基酸可选自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、天冬氨酸、谷氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。在一些实施例中,另一氨基酸是

选自非极性脂肪族氨基酸(例如,甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸或缬氨酸)、芳香族氨基酸(例如,苯丙氨酸、色氨酸或酪氨酸)、极性不带电氨基酸(例如,丝氨酸、苏氨酸、天冬酰胺或谷氨酰胺)或带正电氨基酸(精氨酸、组氨酸或赖氨酸)。在一些实施例中,所述另一氨基酸为丝氨酸。在一些实施例中,所述另一氨基酸为丙氨酸。在一些实施例中,所述另一氨基酸为苯丙氨酸。在一些实施例中,所述另一氨基酸为天冬氨酸。在一些实施例中,所述另一氨基酸为组氨酸。

[0041] 出人意料地,当在两个HRM中的每一个中具有半胱氨酸到丝氨酸突变的ALAS蛋白质与外源血红素结合蛋白共表达时,所述外源血红素结合蛋白的效价可显著增加。

[0042] 一般来说,“效价”为溶液中物质的量的测量值。如本文所使用,除非另外规定,否则血红素结合蛋白的“效价”是指多肽的总量,无论其是否与血红素结合。蛋白质的效价可以通过适合的方法测量,例如高效液相色谱(HPLC或UPLC)、液相色谱-质谱(LC-MS)、酶联免疫吸附分析(ELISA)、酶活性测量、铁测量技术(例如原子吸收光谱、LC-MS或紫外线和/或可见光光谱)。

[0043] ALAS中的突变(例如本文所描述的那些突变)可用于增加血红素产生。在一些实施例中,血红素结合蛋白的效价可增加。在一些实施例中,豆血红蛋白(LegH)的效价可增加。本公开的材料和方法可适用于增加血红素的产生。

[0044] 本文所描述的材料和方法可应用于许多行业中。举例来说,血红素蛋白可用于食品中。作为另一实例,可出于研究目的(例如,为了研究药物代谢)而产生血红素蛋白(例如,血红蛋白、细胞色素P450)。血红素蛋白也可用于工业中(例如,过氧化氢酶、漆酶和/或过氧化酶用于例如洗涤清洁剂、纸浆漂白、木质素降解等用途)。血红素蛋白的另一潜在应用是在生物催化中(例如,细胞色素P450、脂肪氧合酶和/或漆酶可用于生产药物和日用化学品)。一些血红素蛋白可以用作治疗剂(例如,一氧化氮合成酶或作为血液替代物的部分)。在一些情况下,血红素蛋白可用于电子目的(例如,用于产生可再生电)。作为另一实例,血红素疗法可以是患有急性卟啉症的治疗选择,所述卟啉症是由于不能产生血红素而引起的八种遗传疾病的群组。

[0045] 本文提供允许对细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))进行基因工程改造以产生突变型ALAS的核酸构建体(有时也称为核酸分子)。另外,本文提供允许对细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))进行基因工程改造以增加血红素结合蛋白的表达的核酸构建体。在一些实施例中,本文提供允许对细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))进行基因工程改造以增加在不存在诱导分子的情况下诱导型启动子对血红素结合蛋白的表达的核酸构建体。

[0046] 本文提供允许对细胞进行基因工程改造的核酸构建体。A细胞可以是任何适当的细胞。举例来说,细胞可以是细菌细胞(例如大肠杆菌(*E. coli*)细胞、枯草杆菌(*B. subtilis*)细胞或乳酸乳球菌(*Lactococcus lactis*)细胞)、真菌细胞、藻类细胞、植物细胞、昆虫细胞或哺乳动物细胞。在一些实施例中,细胞可以是真菌细胞。在一些实施例中,细胞可以是丝状真菌细胞。在一些实施例中,细胞可以是曲霉或木霉细胞。在一些实施例中,细胞可以是酵母细胞。酵母细胞的非限制性实例包含毕赤酵母属(例如,嗜甲醇毕赤酵母、巴斯德毕赤酵母)、假丝酵母属(例如,博伊丁假丝酵母)细胞、汉逊酵母属(例如,多形汉

逊酵母)细胞、球拟酵母属细胞和酵母属(*Sacharomyces*) (例如,酿酒酵母(*Sacharomyces cerevisiae*))细胞。在一些实施例中,细胞可以是甲基营养型酵母细胞。甲基营养型酵母细胞的非限制性实例包含毕赤酵母属细胞、假丝酵母属细胞、汉逊酵母属细胞和球拟酵母属细胞。在一些实施例中,细胞可以是毕赤酵母属细胞或酵母属细胞。在一些实施例中,细胞为真核细胞。在一些实施例中,细胞为真菌细胞。在一些实施例中,细胞为植物细胞。在一些实施例中,细胞为藻类细胞。在一些实施例中,细胞为酵母细胞。在一些实施例中,细胞为酿酒酵母细胞。在一些实施例中,细胞为甲基营养型酵母细胞。虽然本文中例示使用毕赤酵母属菌种(例如巴斯德毕赤酵母)的方法,但可使用其它细胞,例如毕赤酵母属的其它菌种或来自假丝酵母属、汉逊酵母属、毕赤酵母属和球拟酵母属中的任一个的菌种。甲基营养型酵母的菌种的非限制性实例包含嗜甲醇毕赤酵母、巴斯德毕赤酵母、博伊丁假丝酵母和多形汉逊酵母。

[0047] 因此,在一个方面中,本文提供用于表达蛋白质的材料和方法。在一些实施例中,本文提供一种细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞)),其包含第一核酸构建体,所述第一核酸构建体包括对ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,第一核酸编码可操作地连接到第一启动子元件的ALAS蛋白质。在一些实施例中,ALAS蛋白质包含至少1个HRM(例如1、2、3或更多个HRM)。在一些实施例中,ALAS蛋白质包含至少2个HRM(例如2、3或更多个HRM)。在一些实施例中,ALAS蛋白质包含至少3个HRM(例如3或更多个HRM)。在一些实施例中,ALAS蛋白质包含第一HRM中的突变(有时也称为“第一HRM中的第一突变”)。在一些实施例中,ALAS蛋白质包含第二HRM中的突变(有时也称为“第二HRM中的第二突变”)。在一些实施例中,ALAS蛋白质包含第一HRM中的突变和第二HRM中的突变。在一些实施例中,第一HRM为HRM1。在一些实施例中,第二HRM为HRM2。在一些实施例中,ALAS蛋白质包含第三HRM中的突变(有时也称为“第三HRM中的第三突变”)。在一些实施例中,ALAS蛋白质包含第一HRM中的突变、第二HRM中的突变和第三HRM中的突变。

[0048] 在一些实施例中,第一HRM中的突变为取代。在一些实施例中,第一HRM中的突变为半胱氨酸取代。在一些实施例中,第一HRM中的突变为半胱氨酸到丝氨酸突变。在一些实施例中,第一HRM中的突变为半胱氨酸到丙氨酸突变。在一些实施例中,第一HRM中的突变为半胱氨酸到苯丙氨酸突变。在一些实施例中,第一HRM中的突变为半胱氨酸到天冬氨酸突变。在一些实施例中,第一HRM中的突变为半胱氨酸到组氨酸突变。

[0049] 在一些实施例中,第二HRM中的突变为取代。在一些实施例中,第二HRM中的突变为半胱氨酸取代。在一些实施例中,第二HRM中的突变为半胱氨酸到丝氨酸突变。在一些实施例中,第二HRM中的突变为半胱氨酸到丙氨酸突变。在一些实施例中,第二HRM中的突变为半胱氨酸到苯丙氨酸突变。在一些实施例中,第二HRM中的突变为半胱氨酸到天冬氨酸突变。在一些实施例中,第二HRM中的突变为半胱氨酸到组氨酸突变。

[0050] 在一些实施例中,第三HRM中的突变为取代。在一些实施例中,第三HRM中的突变为半胱氨酸取代。在一些实施例中,第三HRM中的突变为半胱氨酸到丝氨酸突变。在一些实施例中,第三HRM中的突变为半胱氨酸到丙氨酸突变。在一些实施例中,第三HRM中的突变为半胱氨酸到苯丙氨酸突变。在一些实施例中,第三HRM中的突变为半胱氨酸到天冬氨酸突变。在一些实施例中,第三HRM中的突变为半胱氨酸到组氨酸突变。

[0051] 在一些实施例中,第一HRM中的突变对应于SEQ ID NO:29中的残基12中的突变。在一些实施例中,第二HRM中的突变对应于SEQ ID NO:29中的残基39中的突变。在一些实施例中,第一HRM中的突变是对应于SEQ ID NO:29中的残基12中的半胱氨酸到丝氨酸突变的半胱氨酸到丝氨酸突变。在一些实施例中,第一HRM中的突变是对应于SEQ ID NO:29中的残基12中的半胱氨酸到丙氨酸突变的半胱氨酸到丙氨酸突变。在一些实施例中,第一HRM中的突变是对应于SEQ ID NO:29中的残基12中的半胱氨酸到苯丙氨酸突变的半胱氨酸到苯丙氨酸突变。在一些实施例中,第一HRM中的突变是对应于SEQ ID NO:29中的残基12中的半胱氨酸到组氨酸突变的半胱氨酸到组氨酸突变。在一些实施例中,第二HRM中的突变是对应于SEQ ID NO:29中的残基39中的半胱氨酸到丝氨酸突变的半胱氨酸到丝氨酸突变。在一些实施例中,第二HRM中的突变是对应于SEQ ID NO:29中的残基39中的半胱氨酸到丙氨酸突变的半胱氨酸到丙氨酸突变。在一些实施例中,第二HRM中的突变是对应于SEQ ID NO:29中的残基39中的半胱氨酸到苯丙氨酸突变的半胱氨酸到苯丙氨酸突变。在一些实施例中,第二HRM中的突变是对应于SEQ ID NO:29中的残基39中的半胱氨酸到组氨酸突变的半胱氨酸到组氨酸突变。

[0052] 在一些实施例中,第一核酸包含对应于SEQ ID NO:28中的核苷酸位置35处的鸟苷到胞嘧啶突变的突变。在一些实施例中,第一核酸包含对应于SEQ ID NO:28中的核苷酸位置34处的胸腺嘧啶到鸟苷突变和SEQ ID NO:28中的核苷酸位置35处的鸟苷到胞嘧啶突变的突变。在一些实施例中,第一核酸包含对应于SEQ ID NO:28中的核苷酸位置116处的鸟苷到胞嘧啶突变的突变。在一些实施例中,第一核酸包含对应于SEQ ID NO:28中的核苷酸位置115处的胸腺嘧啶到鸟苷突变和SEQ ID NO:28中的核苷酸位置116处的鸟苷到胞嘧啶突变的突变。应了解,半胱氨酸到丝氨酸突变可以通过除本文中具体公开的那些核苷酸以外的核苷酸处的突变来实现。应了解,半胱氨酸到丙氨酸突变可以通过除本文中具体公开的那些核苷酸以外的核苷酸处的突变来实现。在一些实施例中,第一核酸构建体包含SEQ ID NO:30。在一些实施例中,ALAS蛋白质具有SEQ ID NO:31的序列。(参见例如图7)。

[0053] 如本文所使用,“可操作地连接”意味着启动子或其它表达元件相对于核酸编码序列以引导或调控所述编码序列的表达的方式(例如,框内)定位。

[0054] 不同于参考蛋白质序列的蛋白质序列中(例如,与参考ALAS蛋白质序列(例如SEQ ID NO:29)相比,另一生物体的ALAS蛋白质序列中的“对应”氨基酸位置(或取代)可以通过对所关注蛋白质序列进行序列比对来识别。应理解,在一些情况下,蛋白质比对中存在空位。类似地,不同于参考核酸序列的核酸序列中(例如,与参考ALAS核酸序列(例如SEQ ID NO:28)相比,另一生物体的ALAS核酸序列中的“对应”核酸位置(或取代)可以通过对所关注核酸序列进行序列比对来识别。应理解,在一些情况下,核酸比对中存在空位。如本文所使用,“相对于”参考序列的核苷酸或氨基酸位置可以是参考序列中的对应核苷酸或氨基酸位置。

[0055] 本文所描述的方法中使用的核酸分子通常为DNA,但在适当情形下可使用RNA分子。如本文所使用,“外源”是指从例如相同或不同生物体引入到细胞中的任何核酸序列或以合成方式产生的核酸(例如,经密码子优化的核酸序列)。举例来说,外源核酸可以是引入到一种微生物(例如,一个种属或菌种的甲基营养型酵母)中的来自另一种属或菌种的甲基营养型酵母的核酸;然而,外源核酸还可以是在尽管存在对应的原生核酸序列的情况下仍

作为额外拷贝以重组方式引入到甲基营养型酵母中的来自甲基营养型酵母的核酸,或以重组方式引入到甲基营养型酵母中的与所述甲基营养型酵母的原生序列相比含有一或多个突变、插入或缺失的来自甲基营养型酵母的核酸。举例来说,巴斯德毕赤酵母含有对ALAS进行编码的内源核酸;(例如以重组方式引入到巴斯德毕赤酵母中的)巴斯德毕赤酵母ALAS核酸的额外拷贝被视为外源的。类似地,“外源”蛋白质是由外源核酸编码的蛋白质。

[0056] 在一些情况下,外源核酸可以是异源核酸。如本文所使用,“异源”核酸是指并非生物体原生的任何核酸序列(例如,异源核酸可以是引入到一种微生物(例如,一个种属或菌种的甲基营养型酵母)中的来自另一种属或菌种的甲基营养型酵母的核酸(无论其是否已经密码子优化))。类似地,“异源”蛋白质是由异源核酸编码的蛋白质。

[0057] 当核酸分子的任何部分(例如启动子序列或所编码蛋白质的序列)对宿主生物体为外源的时,所述核酸分子被视为对所述宿主生物体为外源的。当核酸分子的任何部分(例如启动子序列或所编码蛋白质的序列)对于宿主生物体为异源的时,所述核酸分子被视为对所述宿主生物体为异源的。

[0058] 在一些实施例中,本文献提供一种细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞)),其包括如本文所公开的第一核酸构建体和第二核酸构建体,所述第二核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列。在一些实施例中,第二核酸编码可操作地连接到第一启动子元件的血红素结合蛋白。在一些实施例中,第二核酸编码可操作地连接到第二启动子元件的血红素结合蛋白。在一些实施例中,血红素结合蛋白可以是外源血红素结合蛋白。在一些实施例中,血红素结合蛋白可以是异源血红素结合蛋白。在一些实施例中,血红素结合蛋白可选自由以下组成的群组:球蛋白(Pfam数据库中的PF00042)、细胞色素(例如,细胞色素P450、细胞色素a、细胞色素b、细胞色素c)、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。在一些实施例中,球蛋白可选自由以下组成的群组:雄性球蛋白、血绿蛋白、细胞球蛋白、无脊椎动物血红蛋白、黄素血红蛋白、球蛋白E、球蛋白X、球蛋白Y、血红蛋白(例如, β 血红蛋白、 α 血红蛋白)、组胺球蛋白、豆血红蛋白、肌血球素、神经球蛋白、非共生血红蛋白、原球蛋白和截短血红蛋白(例如,HbN、HbO、G1b3、蓝藻球蛋白)。在一些实施例中,血红素结合蛋白可以是非共生血红蛋白。在一些实施例中,血红素结合蛋白可以是豆血红蛋白。在一些实施例中,血红素结合蛋白可以是大豆血红蛋白(LegH)。LegH的参考氨基酸序列在图4中提供为SEQ ID NO:4。LegH为结合到血红素的蛋白质,其产生415nm下的特征吸收和独特的红色。LegH蛋白(也称为LGB2)天然存在于大豆的根瘤中(参见例如UniprotKB登录号P02236)。还参见W0 2014/110539和W02014/110532,其中每一者的全文通过引用并入本文中。在一些实施例中,血红素结合蛋白可具有与SEQ ID NO:1-27(图4)中的任一个中所阐述的氨基酸序列至少70%(例如至少75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%)一致的氨基酸序列。在一些实施例中,血红素结合蛋白可具有SEQ ID NO:1-27中的任一个中所阐述的氨基酸序列。在一些实施例中,血红素结合蛋白使用血红素作为辅因子。如本文所使用,“辅因子”为直接参与酶催化的分子或离子。在一些实施例中,血红素结合蛋白可以是或可以不是酶。在一些实施例中,血红素结合蛋白可以是或可以不是血红素生物合成途径的一部分。在一些实施例中,血红素结合蛋白可以是或可以不是ALAS。在一些实施例中,血红素结合蛋白可以是或可以不是亚铁螯合酶。在一些实施例中,血红素结合蛋白可以是或可以不是粪卟

啉原氧化酶。在一些实施例中，血红素结合蛋白可以是或可以不是ALAD、PBGD、UPG3S、UPG3D、COPROX、PROTOX或FC。在一些实施例中，血红素结合蛋白具有与SEQ ID NO:1-27中的任一个中的氨基酸序列至少70%（例如至少75%、80%、85%、90%或95%）一致的氨基酸序列。在一些实施例中，血红素结合蛋白为源于细菌的血红素结合蛋白、源于酵母的血红素结合蛋白、源于藻类的血红素结合蛋白、源于真菌的血红素结合蛋白或源于植物的血红素结合蛋白。

[0059] 如本文所使用，“源于细菌的蛋白质”、“源于酵母的蛋白质”、“源于藻类的蛋白质”、“源于真菌的蛋白质”或“源于植物的蛋白质”是指蛋白质的直接来源，并且可以意味着分别在细菌、酵母、藻类、真菌或植物中产生的任何蛋白质，这与所述蛋白质是否在所述细菌、酵母、藻类、真菌或植物中天然表达无关。

[0060] 本文提供产生ALAS蛋白质的方法。本文还提供使用本文所描述的细胞（例如真菌细胞，例如曲霉细胞、木霉细胞或酵母细胞（例如甲基营养型酵母细胞））中的任一个来产生ALAS蛋白质的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生ALAS蛋白质的方法。在一些实施例中，本文提供包括以下的方法：表达第一核酸构建体，所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中，与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比，ALAS蛋白质的效价可增加至少5%（例如至少6%、7%、8%、9%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%或更多）。在一些实施例中，与表达不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比，ALAS蛋白质的效价可增加至少5%（例如至少6%、7%、8%、9%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%或更多）。在一些实施例中，本文提供包括以下的方法：表达第一核酸构建体，所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中，与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比，ALAS蛋白质的效价可增加至少5%（例如至少6%、7%、8%、9%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%或更多）。在一些实施例中，与表达不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比，ALAS蛋白质的效价可增加至少5%（例如至少6%、7%、8%、9%、10%、15%、20%、25%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%或更多）。

[0061] 本文提供产生四吡咯或其衍生物的方法。本文还提供使用本文所描述的细胞（例如真菌细胞，例如曲霉细胞、木霉细胞或酵母细胞（例如甲基营养型酵母细胞））中的任一个来产生四吡咯或其衍生物的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生四吡咯或其衍生物的方法。在一些实施例中，本文提供产生四吡咯或其衍生物的方法，其包括表达第一核酸构建体，所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中，与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比，四吡咯或其衍生物的效价可

增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,四吡咯或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,本文提供产生四吡咯或其衍生物的方法,其包括表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的第二突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,四吡咯或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,四吡咯或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,四吡咯或其衍生物为二氢卟吩 (例如,叶绿素) 或其衍生物。在一些实施例中,四吡咯或其衍生物为前咕啉、钴-前咕啉、咕啉 (例如,维生素B12) 或其衍生物。在一些实施例中,四吡咯或其衍生物为卟啉或其衍生物。在一些实施例中,四吡咯或其衍生物为血红素或其衍生物。在一些实施例中,四吡咯或其衍生物为血红素B。在一些实施例中,四吡咯或其衍生物为血红素的代谢产物 (例如胆红素或其衍生物)。

[0062] 本文提供产生血红素生物合成途径的中间产物、血红素 (例如,血红素B、血红素o、血红素a、血红素c)、咕啉 (例如,维生素B12)、叶绿素或其衍生物的方法。本文还提供使用本文所描述的细胞 (例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞 (例如甲基营养型酵母细胞)) 中的任一个来产生血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的方法。在一些实施例中,本文提供产生血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的方法,其包括表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,本文提供产生血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的方法,其包括

表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达对不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素生物合成途径的中间产物、血红素、咕啉、叶绿素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,血红素生物合成途径的中间产物、血红素或其衍生物为血红素。在一些实施例中,血红素生物合成途径的中间产物、血红素或其衍生物为血红素B。

[0063] 本文提供产生血红素(例如,血红素B、血红素o、血红素a、血红素c)或其衍生物的方法。本文还提供使用本文所描述的细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的任一个来产生血红素或其衍生物的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生血红素或其衍生物的方法。在一些实施例中,本文提供产生血红素或其衍生物的方法,其包括表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达对不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,本文提供产生血红素或其衍生物的方法,其包括表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列。在一些实施例中,与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,与表达对不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素或其衍生物的效价可增加至少5% (例如至少10%、20%、30%、40%、50%、60%、70%、80%、90%、100%、120%、140%、150%、160%、180%、200%、220%、240%、250%、280%、300%、320%、340%、350%或更多)。在一些实施例中,血红素或其衍生物为血红素B。

[0064] 本文提供产生血红素结合蛋白的方法。本文还提供使用本文所描述的细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的任一个来

产生血红素结合蛋白的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生血红素结合蛋白的方法。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列。在本文所描述的任一种方法的一些实施例中,所述方法允许血红素结合蛋白的效价增加。在一些实施例中,与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素结合蛋白的效价可增加至少5%(例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与表达不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素结合蛋白的效价可增加至少5%(例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,血红素结合蛋白的效价可增加至少5%(例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与表达不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,血红素结合蛋白的效价可增加至少5%(例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。

[0065] 本文提供产生四吡咯结合蛋白的方法。本文还提供使用本文所描述的细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的任一个来产生四吡咯结合蛋白的方法。本文还提供使用本文所描述的核酸构建体中的任一个来产生四吡咯结合蛋白的方法。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对四吡咯结合蛋白进行编码的核苷酸序列。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对四吡咯结合蛋白进行编码的核苷酸序列。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括

第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对四吡咯结合蛋白进行编码的核苷酸序列。在一些实施例中,本文提供包括以下的方法:表达第一核酸构建体,所述第一核酸构建体包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列;和表达第二核酸构建体,所述第二核酸构建体包括对四吡咯结合蛋白进行编码的核苷酸序列。在本文所描述的任一种方法的一些实施例中,所述方法允许四吡咯结合蛋白的效价增加。在一些实施例中,与缺少包括对包括第一HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,四吡咯结合蛋白的效价可增加至少5% (例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与表达对不包括第一HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,四吡咯结合蛋白的效价可增加至少5% (例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与缺少包括对包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核苷酸序列的核酸构建体的对应方法相比,四吡咯结合蛋白的效价可增加至少5% (例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,与表达对不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法相比,四吡咯结合蛋白的效价可增加至少5% (例如至少6%、7%、8%、9%、10%、11%、12%、13%、14%、15%、16%、17%、18%、19%、20%、21%、22%、23%、24%、25%、30%、35%、40%、45%、50%或更多)。在一些实施例中,四吡咯为二氢卟吩 (例如,叶绿素) 或其衍生物。在一些实施例中,四吡咯为前咕啉、钴-前咕啉、咕啉 (例如,维生素B12) 或其衍生物。在一些实施例中,四吡咯为卟啉或其衍生物。在一些实施例中,四吡咯为血红素或其衍生物。在一些实施例中,四吡咯为血红素B。在一些实施例中,四吡咯为血红素的代谢产物 (例如胆红素或其衍生物)。

[0066] 如本文所使用,“对应方法”是除了所确认的差异以外在各个方面都与参考方法基本上一致的方法。举例来说,表达对不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸的对应方法将在所有方面 (例如,细胞的基因组成、培养的温度和时间等) 相同,除了所述对应方法将表达对不包括第一HRM中的突变和第二HRM中的突变的ALAS蛋白质进行编码的核酸。

[0067] 对细胞 (例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞 (例如甲基营养型酵母细胞)) 进行基因工程改造通常包含将重组核酸分子 (也称为核酸构建体) 引入到所述细胞中。如本文所描述,重组核酸分子通常包含对可操作地连接到至少一个启动子元件 (例如,诱导型或组成型启动子元件) 的蛋白质 (例如,参与血红素生物合成的蛋白质、血红素结合蛋白或转录因子) 进行编码的外源核酸。在一些实施例中,重组核酸分子可包含可操作地连接到同一或不同启动子元件 (例如,可操作地连接到编码第一蛋白质的第一核酸的第一启动子和可操作地连接到编码第二蛋白质的第二核酸的第二启动子,或可操作地连接到编码第一蛋白质的第一核酸和编码第二蛋白质的第二核酸的启动子) 的两个或更多个蛋白质编码序列的线性序列。在一些情况下,包含至少一个可操作地连接到编码蛋白质的核苷酸

序列的启动子的重组核酸分子可被称为盒。

[0068] 重组核酸可包含表达元件。表达元件包含引导和调控核酸编码序列的表达的核酸序列。表达元件的一个实例为启动子序列。表达元件还可包含内含子、增强子序列、反应元件或调节核酸的表达的诱导型元件。表达元件可以是细菌、酵母、昆虫、哺乳动物或病毒来源的,并且载体可以含有来自不同来源的元件的组合。

[0069] 可利用适当的寡核苷酸(例如引物)对,使用任何数目的扩增技术(参见例如PCR引物:实验指南(PCR Primer:A Laboratory Manual),1995,迪芬巴赫(Dieffenbach)和德弗克斯勒(Dveksler)编,冷泉港实验室出版社(Cold Spring Harbor Laboratory Press),纽约州冷泉港(Cold Spring Harbor,NY);以及第4,683,195号、第4,683,202号、第4,800,159号和第4,965,188号美国专利)来检测核酸。已经开发了对原始PCR方法的许多修改,并且可以用于检测所选择的核酸。

[0070] 甲醇利用通常经由通过醇氧化酶的作用将甲醇转化成甲醛来引发。巴斯德毕赤酵母含有两个醇氧化酶基因AOX1和AOX2。与不具有降低的醇氧化酶活性的菌株相比,具有降低的醇氧化酶活性的菌株(“甲醇慢速利用”或MutS型菌株)可产生更多的从AOX1启动子表达的重组蛋白。两个AOX基因都突变且完全缺乏醇氧化酶活性的菌株无法使甲醇代谢,但仍可以通过甲醇诱导来自AOX1启动子的表达。这些菌株保留使用其它碳源用于生长的能力,但在添加甲醇后仍表达来自AOX1启动子的异源蛋白质。因为这些菌株不会使甲醇代谢(“甲醇不利用”或Mut-型菌株),所以诱导蛋白质表达需要少得多的甲醇,并且携带这些突变的菌株避免了与大规模发酵中的甲醇给料相关的问题。参见例如基鲁沃鲁(Chiruvolu)等人,1997,酶与微生物技术(Enzyme Microb.Technol.),21:277-83。

[0071] 适合的转录因子和编码转录因子的核酸(例如,编码转录因子的外源核酸)包含例如来自巴斯德毕赤酵母的Mxr1。代表性的巴斯德毕赤酵母Mxr1核酸序列可见于例如GenBank登录号DQ395124,而代表性的巴斯德毕赤酵母Mxr1蛋白质序列可见于例如GenBank登录号ABD57365。在一些实施例中,转录因子为来自法夫驹形氏酵母的Mit1序列(参见例如UniParc登录号UPI0001A4D18B)。适合的转录因子也可存在于多形汉逊酵母中(例如,Adr1序列;例如核酸序列参见GenBank登录号AE0102000005,碱基858873至862352,且氨基酸序列参见GenBank登录号ESX01253)和博伊丁假丝酵母中(例如,Trm1序列;例如核酸序列参见GenBank登录号AB365355且氨基酸序列参见GenBank登录号BAF99700;和Trm2序列;例如核酸序列参见GenBank登录号AB548760且氨基酸序列参见GenBank登录号BAJ07608)。

[0072] 例如Mxr1的转录因子可以低水平正常表达。在一些实施例中,需要将外源核酸(例如,转录因子)置于可诱导的启动子的控制下。

[0073] 毕赤酵母属中甲醇调控的转录因子可结合到AOX1启动子且与Mxr1协同作用以活化来自AOX1启动子的转录。在一些实施例中,两个甲醇调控的转录因子(例如,Mxr1和Mit1)可以可操作地连接到甲醇诱导型启动子元件。

[0074] 存在可以在对细胞(例如真菌细胞,例如曲霉、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))进行基因工程改造时使用的多种诱导型启动子。适合的甲醇诱导型启动子包含pAOX1(例如,pAOX1,例如巴斯德毕赤酵母AOX1的启动子(参见例如GenBank登录号U96967.1的启动子),或2019年4月17日提交的第62/835,338号美国临时专利申请(其全文

通过引用并入本文)中所描述的启动子)以及其它甲醇诱导型启动子或来自那些启动子的启动子元件。这些包含(但不限于)pAOX2启动子(例如,来自法夫驹形氏酵母或巴斯德驹形氏酵母的pAOX2启动子(参见例如GenBank登录号X79871.1))、来自例如博伊丁假丝酵母的醇氧化酶(AOD1)启动子(参见例如GenBank登录号E06147.1)、来自多形汉逊酵母的醇氧化酶(MOX)启动子(参见例如GenBank登录号AJ313360.1)、来自嗜甲醇毕赤酵母的MOD1或MOD2启动子(参见例如雷蒙德(Raymond)等人,1998,酵母(Yeast),14:11-23;和中川(Nakagawa)等人,1999,酵母,15:1223-30)、来自巴斯德毕赤酵母的DHAS启动子(参见例如GenBank登录号FJ752551的启动子)或来自其的启动子元件、来自巴斯德驹形氏酵母的甲醛脱氢酶(FLD1)启动子(参见例如GenBank登录号KJ755994.1)或来自巴斯德毕赤酵母的PEX8启动子(参见例如克兰蒂(Kranthi)等人,2010,酵母,27:705-11)。通常,这些启动子可以由甲醇诱导。适合的组成型启动子和组成型启动子元件包含(但不限于)来自转录延伸因子EF-1 α 基因(TEF1)的巴斯德毕赤酵母启动子(或其部分),其以组成型方式强转录。也可使用其它适合的组成型启动子(或来自其的启动子元件),包含(但不限于)来自巴斯德驹形氏酵母的甘油醛-3-磷酸脱氢酶(GAPDH)启动子(参见例如GenBank登录号U62648.1的启动子)、来自潜在糖基磷脂酰肌醇(GPI)锚定蛋白的启动子、来自法夫驹形氏酵母的GCW14p(PAS_chr1-4_0586)(参见例如GenBank登录号XM_002490678的启动子)和来自巴斯德驹形氏酵母的3-磷酸甘油酸激酶基因(PGK1)的启动子(参见例如GenBank登录号AY288296)。应了解,启动子的选择可受表达系统影响。举例来说,对于在法夫驹形氏酵母中的表达,可选择法夫驹形氏酵母启动子,而对于在博伊丁假丝酵母中的表达,可选择博伊丁假丝酵母启动子。然而,在一些情况下,来自一种生物体(例如,法夫驹形氏酵母)的启动子可适合于用于另一种生物体(例如,博伊丁假丝酵母或巴斯德驹形氏酵母)中。此外,应注意,可组合诱导型(例如甲醇诱导型)和组成型启动子(或来自其的启动子元件)的组合以进一步增加可操作地与其连接的核酸中的任一个的表达。

[0075] 如本文所描述的经编码蛋白质中的任一个可以可操作地连接到诱导型启动子元件(例如,甲醇诱导型启动子元件)或组成型启动子元件。诱导型启动子和来自其的元件在上文论述。组成型启动子和组成型启动子元件为所属领域中已知的。举例来说,来自巴斯德毕赤酵母的常用组成型启动子为来自转录延伸因子EF-1 α 基因(TEF1)的启动子或其部分,其以组成型方式强转录。然而,可使用其它组成型启动子或来自其的启动子元件,包含(但不限于)来自巴斯德驹形氏酵母的甘油醛-3-磷酸脱氢酶(GAPDH)启动子(参见例如GenBank登录号U62648.1的启动子)、来自潜在糖基磷脂酰肌醇(GPI)锚定蛋白的启动子、来自法夫驹形氏酵母的GCW14p(PAS_chr1-4_0586)(参见例如GenBank登录号XM_002490678的启动子)或来自巴斯德驹形氏酵母的3-磷酸甘油酸激酶基因(PGK1)的启动子(参见例如GenBank登录号AY288296的启动子)。

[0076] 在一些实施例中,本文中的细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的任一个可进一步包含第三核酸构建体,所述第三核酸构建体包括对可操作地连接到第一启动子元件、第二启动子元件或第三启动子元件的第三蛋白质进行编码的核苷酸序列。在一些实施例中,本文中的细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的任一个可进一步包含第四核酸构建体,所述第四核酸构建体包括对可操作地连接到第一启动子元件、第二启动子元

件、第三启动子元件或第四启动子元件的第四蛋白质进行编码的核苷酸序列。在一些实施例中,第三蛋白质可以是转录因子。在一些实施例中,第四蛋白质可以是转录因子。在一些实施例中,本文中的启动子元件中的任一个(例如,第一启动子元件、第二启动子元件、第三启动子元件或第四启动子元件)可含有转录因子的一或多个识别序列。因此,在一些实施例中,可以构建反馈环以使得转录因子驱动转录因子的额外拷贝的表达以及ALAS和血红素结合蛋白中的一或多个的表达。在一些实施例中,转录因子可以是Mxr1(参见例如第9,938,327号美国专利,其全文通过引用并入)。在一些实施例中,第三蛋白质可以是参与血红素生物合成的蛋白质。在一些实施例中,第四蛋白质可以是参与血红素生物合成的蛋白质。在一些实施例中,参与血红素生物合成的蛋白质可选自由以下组成的群组: δ 氨基乙酰丙酸脱水酶(ALAD)、胆色素原脱氢酶(PBGD)、尿卟啉原III合成酶(UPG3S)、尿卟啉原III脱羧酶(UPG3D)、粪卟啉原氧化酶(COPROX)、原卟啉原IX氧化酶(PROTOX)和/或亚铁螯合酶(FC)。在一些实施例中,参与血红素生物合成的蛋白质可选自由以下组成的群组: δ 氨基乙酰丙酸脱水酶(ALAD)、胆色素原脱氢酶(PBGD)、尿卟啉原III合成酶(UPG3S)、尿卟啉原III脱羧酶(UPG3D)、粪卟啉原氧化酶(COPROX)和/或原卟啉原IX氧化酶(PROTOX)。

[0077] 在酿酒酵母中的先前研究鉴别出ALA脱水酶和胆色素原脱氢酶为血红素生物合成中的速率限制酶(参见例如霍夫曼(Hoffman)等人,2003,生物化学与生物物理研究通讯(Biochem.Biophys.Res.Comm.),310(4):1247-53)。然而,来自甘油醛-3-磷酸脱氢酶(GAP)启动子的个别血红素酶在巴斯德毕赤酵母中的异源表达未能克服与含有血红素的重组蛋白的表达相关的局限性(参见克赖纳(Krainer)等人,2015,微生物细胞工厂(Microb.Cell Fact.),13;14:4)。含有重组血红素的蛋白质在巴斯德毕赤酵母中的表达可以通过由甲醇诱导型启动子共表达血红素生物合成途径的一或多种酶实现,但是应了解,参与血红素生物合成途径的基因中的一或多个可以由一或多个组成型启动子表达(参见例如第9,938,327号美国专利,其全文通过引用并入)。

[0078] 另外,应注意,如本文所描述的对可操作地连接到启动子元件的第一蛋白质(例如ALAS蛋白质)进行编码的第一核酸可以与对可操作地连接到启动子元件的第二蛋白质(例如血红素结合蛋白)进行编码的第二核酸在物理上分离(也就是说,第一和第二核酸可以是完全分离的分子)。替代地,对可操作地连接到启动子元件的第一蛋白质进行编码的第一核酸和对可操作地连接到启动子元件的第二蛋白质进行编码的第二核酸可包含在同一核酸构建体中。在一些实施例中,对可操作地连接到启动子元件的第一蛋白质进行编码的第一核酸可以与对可操作地连接到启动子元件的第二蛋白质进行编码的第二核酸邻接。所属领域的技术人员应了解,如果编码第二蛋白质的第二核酸分子与编码所关注蛋白质的第一核酸邻接,那么可以使用单一启动子或来自其的启动子元件来驱动两个或所有基因(例如编码第一蛋白质以及第二蛋白质的核酸)的转录。

[0079] 将核酸引入到细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中的方法为所属领域中已知的,并且包含(但不限于)转导、电穿孔、基因枪微粒递送和化学转化。

[0080] 此外,培养细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))的方法为所属领域中已知的。参见例如毕赤酵母属方案(Pichia Protocols),分子生物学方法(Methods In Molecular Biology),389,格瑞(Cregg)编,

2007,第2版。在一些情况下,可能需要将甲醇引入或添加到培养基中,但如本文所证实,甲醇并非获得一或多种所关注蛋白质的高水平的高效表达所必需的。在一些情况下(例如当表达对参与血红素生物合成的酶进行编码的一或多种核酸时),可能需要向培养基补充铁或其药理学上或代谢上可接受的(或GRAS)盐。

[0081] 本文所描述的重组核酸分子可以稳定地整合到细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))的基因组中,或可以由复制感受态质粒在染色体外表达。实现两者的方法为已知的且在所属领域中使用。

[0082] 本文所提供的方法还可以包含纯化所表达的蛋白质。如本文所使用,“富集的”蛋白质是按干重计占生产细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))的质量的至少5%(例如至少6%、7%、8%、9%、10%、15%、20%、25%、30%、35%、40%、45%、50%、55%、60%或更多),或按干重计占生产细胞溶解物(例如,不包含细胞壁或细胞膜物质)的质量的至少10%(例如至少15%、20%、25%、30%、35%、40%、45%、50%、55%、60%、65%、70%、75%、80%、90%、95%或99%)的蛋白质。如本文所使用,“纯化的”蛋白质是已从其天然附有的细胞组分分离或纯化的蛋白质。通常,当蛋白质按干重计为至少60%(例如至少65%、70%、75%、80%、85%、90%、95%或99%)且不含其天然缔合的蛋白质和天然存在的分子时,将其视为“纯化的”。

[0083] 如本文所使用,核酸可包含DNA和RNA,并且包含含有一或多种核苷酸类似物或主链修饰的核酸。核酸可以是单链或双链的,此通常取决于其预期用途。还提供不同于给定序列的核酸和多肽。核酸和多肽可以与给定核酸或多肽序列具有至少50%序列一致性(例如,至少55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%或99%序列一致性)。

[0084] 在计算序列一致性百分比时,对两个序列进行比对并且确定两个序列之间核苷酸或氨基酸残基的一致匹配的数目。一致匹配的数目除以比对区的长度(即,比对的核苷酸或氨基酸残基的数目)并且乘以100,得到序列一致性百分比值。应了解,比对区的长度可以是一个或两个序列的一部分到长达最短序列的全长大小。还应了解,单一序列可与多于一个其它序列比对,且因此,可在每一经比对区上具有不同的序列一致性百分比值。

[0085] 可以使用计算机程序ClustalW和默认参数来执行两个或更多个序列的比对以确定序列一致性百分比,这允许跨核酸或多肽序列的整个长度进行所述核酸或多肽序列的比对(全局比对)。陈娜(Chenna)等人,2003,核酸研究(Nucleic Acids Res.),31(13):3497-500。ClustalW计算查询序列与一或多个目标序列之间的最佳匹配,且对其进行比对以使得可确定身份、相似性和差异。可以将一或多个残基的空位插入到查询序列、目标序列或这两者中以使序列比对最大化。对于核酸序列的快速成对比对,可使用默认参数(即,字长:2;窗口大小:4;评分方法:百分比;顶面对角线数目:4;和空位罚分:5);对于多个核酸序列的比对,可使用以下参数:空位开放罚分:10.0;空位延伸罚分:5.0;和权重转换:是。对于多肽序列的快速成对比对,可使用以下参数:字长:1;窗口大小:5;评分方法:百分比;顶面对角线数目:5;和空位罚分:3。对于多肽序列的多序列比对,可使用以下参数:权重矩阵:blosum;空位开放罚分:10.0;空位延伸罚分:0.05;亲水性空位:开启;亲水性残基:Gly、Pro、Ser、Asn、Asp、Gln、Glu、Arg和Lys;和残基特定空位罚分:开启。ClustalW可以在例如万维网上(World Wide Web)的贝勒医学院搜索启动器(Baylor College of Medicine Search

Launcher) 网站或欧洲生物信息研究所 (European Bioinformatics Institute) 网站运行。

[0086] 可以将改变引入到核酸分子中,从而引起所编码的多肽的氨基酸序列的改变。举例来说,可以使用诱变(例如,定点诱变、PCR介导的诱变、转座子诱变、化学诱变、UV诱变或辐射诱发的诱变)或通过化学合成具有改变的核酸分子而将此类改变引入到核酸编码序列中。此类核酸改变可引起一或多个氨基酸残基处的保守和/或非保守氨基酸取代。“保守氨基酸取代”是其中一个氨基酸残基被具有类似侧链的另一氨基酸残基替换的氨基酸取代(参见例如戴霍夫(Dayhoff)等人,1978,蛋白质序列和结构地图集(Atlas of Protein Sequence and Structure),5(增刊3):345-352,其提供氨基酸取代的频率表),且非保守取代是其中氨基酸残基被不具有类似侧链的氨基酸残基替换的取代。核酸和/或多肽序列可以如本文所描述经修饰以改进一或多种特性,例如(但不限于)增加的表达(例如转录和/或翻译)、更紧密的调控、去调控、分解代谢物阻遏丧失、经修改的特异性、分泌、热稳定性、溶剂稳定性、氧化稳定性、蛋白酶抗性、催化活性和/或颜色。

[0087] 如本文所使用,“经分离”核酸分子是不含将核酸的一端或两端天然侧接在所述经分离核酸分子所来源的生物体的基因组中的序列的核酸分子(例如,通过PCR或限制性核酸内切酶消化产生的cDNA或基因组DNA片段)。为了操纵方便或产生融合核酸分子,通常将此类经分离核酸分子引入到载体(例如克隆载体或表达载体)中,下文更详细地论述。另外,经分离核酸分子可包含经工程改造的核酸分子,例如重组或合成核酸分子。。

[0088] 如本文所描述的载体可以引入到宿主细胞(例如真菌细胞,例如曲霉细胞、木霉细胞或酵母细胞(例如甲基营养型酵母细胞))中。如本文所使用,“宿主细胞”是指向其中引入核酸的特定细胞,并且还包含携带载体的此类细胞的后代。A宿主细胞可以是任何原核或真核细胞。举例来说,核酸可以在细菌细胞(例如大肠杆菌)中或在昆虫细胞、酵母或哺乳动物细胞(例如中国仓鼠卵巢细胞(CHO)或COS细胞)中表达。其它适合的宿主细胞是所属领域的技术人员已知的。在活体内和活体外将核酸引入到宿主细胞中的许多方法是所属领域的技术人员熟知的,并且包含(但不限于)电穿孔、磷酸钙沉淀、聚乙二醇(PEG)转化、热休克、脂质体转染、微注射和病毒介导的核酸转移。

[0089] 核酸可以使用所属领域中常规的技术进行分离。举例来说,可以使用包含(但不限于)重组核酸技术和/或聚合酶链反应(PCR)的任何方法来分离核酸。通用PCR技术描述于例如PCR引物:实验指南,迪芬巴赫和德弗克斯勒编,冷泉港实验室出版社,1995中。重组核酸技术包含例如限制酶消化和连接,其可以用于分离核酸。经分离核酸还可以通过化学方式以单一核酸分子或一系列寡核苷酸的形式合成。

[0090] 可以通过已知方法,例如DEAE离子交换、凝胶过滤和羟基磷灰石色谱从天然来源(例如生物样品)纯化多肽。还可以例如通过在表达载体中表达核酸来纯化多肽。另外,可以通过化学合成获得经纯化的多肽。多肽的纯化程度可以使用任何适当的方法来测量,例如柱色谱、聚丙烯酰胺凝胶电泳或HPLC分析。

[0091] 还提供含有核酸(例如编码多肽的核酸)的构建体或载体。构建体或载体(包含表达构建体或载体)是可商购的或可通过所属领域中常规的重组DNA技术产生。含有核酸的构建体或载体可以具有可操作地连接到此类核酸的表达元件,并且可以进一步包含序列,例如编码可选标记物(例如抗生素抗性基因)的序列。含有核酸的构建体或载体可以编码嵌合或融合多肽(即,可操作地连接到异源多肽的多肽,所述异源多肽可以在所述多肽的N端或C

端处)。代表性异源多肽包含可以用于纯化经编码多肽的异源多肽(例如6xHis标签、谷胱甘肽S-转移酶(GST))。

[0092] 还可以使用杂交来检测核酸。核酸之间的杂交详细论述于萨姆布鲁克(Sambrook)等人(1989, 分子克隆:实验指南(Molecular Cloning: A Laboratory Manual), 第2版, 冷泉港实验室出版社, 纽约州冷泉港; 章节7.37-7.57、9.47-9.57、11.7-11.8和11.45-11.57)中。萨姆布鲁克等人公开了用于小于约100个核苷酸的寡核苷酸探针的适合的南方墨点条件(章节11.45-11.46)。可以使用章节11.46中所提供的公式来计算长度小于100个核苷酸的序列与第二序列之间的 T_m 。萨姆布鲁克等人另外还公开了用于大于约100个核苷酸的寡核苷酸探针的南方墨点条件(参见章节9.47-9.54)。可以使用萨姆布鲁克等人的章节9.50-9.51中所提供的公式来计算长度大于100个核苷酸的序列与第二序列之间的 T_m 。

[0093] 含有核酸的膜进行预杂交和杂交的条件以及洗涤含有核酸的膜以去除过量和非特异性结合的探针的条件可在杂交的严格度中起重要作用。适当时,可在中等或高严格度条件下进行此类杂交和洗涤。举例来说,可以通过降低洗涤溶液中的盐浓度和/或通过增加进行洗涤的温度来使洗涤条件更严格。简单地借助于实例,高严格度条件通常包含在65°C下在0.2X SSC中洗涤膜。

[0094] 另外,解释杂交量可例如受经标记的寡核苷酸探针的比活性、探针所杂交的模板核酸上的探针结合位点的数目和放射自显影图或其它检测介质的暴露量影响。所属领域的一般技术人员将容易了解,尽管可以使用任何数目的杂交和洗涤条件来检查探针核酸分子与固定目标核酸的杂交,但检查在相同杂交、洗涤和暴露条件下探针与目标核酸的杂交更为重要。优选地,目标核酸在同一膜上。

[0095] 如果与一种核酸的杂交相比于与另一种核酸的杂交高至少5倍(例如,至少6倍、7倍、8倍、9倍、10倍、20倍、50倍或100倍),那么认为核酸分子与所述核酸杂交而非与另一种核酸杂交。可以在膜上直接定量杂交的量,或使用例如磷光成像仪或密度计(分子动力学(Molecular Dynamics)公司,加利福尼亚州森尼维耳市(Sunnyvale, CA))从放射自显影图定量杂交的量。

[0096] 可以使用抗体来检测多肽。使用抗体来检测多肽的技术包含酶联免疫吸附分析(ELISA)、蛋白质印迹、免疫沉淀和免疫荧光。抗体可以是多克隆或单克隆的。可以使用所属领域中众所周知的方法产生对多肽具有特异性结合亲和力的抗体。抗体可以使用所属领域中已知的方法附着到固体载体,例如微量滴定板。在多肽存在下,形成抗体-多肽复合物。

[0097] 通常使用可检测标记实现检测(例如扩增产物、杂交复合物或多肽的检测)。术语“标记”意图涵盖直接标记以及间接标记的使用。可检测标记包含酶、辅基、荧光材料、发光材料、生物发光材料和放射性材料。

[0098] 本文中描述可用于产生缺少用于选择的序列(即,缺少可选标记物)的菌株的方法。这些方法包含使用圆形质粒DNA载体和线性DNA序列;所述圆形质粒DNA载体含有选择标记物和DNA复制起点(也称为自主复制序列(ARS)),并且所述线性DNA序列含有通过同源重组整合到基因组(例如毕赤酵母属基因组)中的序列。所述线性DNA分子另外可包含对一或多种所关注蛋白质进行编码的核酸序列,所述蛋白质例如(但不限于)ALAS、血红素结合蛋白或第三蛋白质(例如,转录因子或参与血红素生物合成的蛋白质)。

[0099] 细胞(例如,毕赤酵母属细胞)可以用DNA分子和通过圆形质粒上的可选标记物的

存在所选择的转化体两者来进行转化。接着可以使用例如PCR来筛选线性DNA分子整合到基因组中的转化体。一旦鉴别出正确整合有不含标记物的线性DNA分子的转化体,就可以在不存在对圆形质粒的选择的情况下使细胞生长。因为携带标记物的质粒在不存在选择的情况下并未稳定地维持,所以在放宽选择之后通常非常快速地丢失质粒。所得菌株在不存在用于选择的异源序列的情况下携带整合的线性DNA。因此,此方法可以用于构建缺少可选标记物(例如异源选择标记物)而对重组蛋白产率几乎没有影响的菌株(例如毕赤酵母属菌株)。

[0100] 根据本公开,在本领域的技术范围内可以采用常规的分子生物学、微生物学、生物化学和重组DNA技术。此类技术在文献中进行了充分解释。将在以下实例中进一步描述本公开的材料和方法,所述实例不限制权利要求书中所描述的方法和物质组合物的范围。

[0101] 将在以下实例中进一步描述本公开的材料和方法,所述实例不限制权利要求书的范围。

[0102] 示例性实施例

[0103] 实施例1为一种细胞,其包括:

[0104] 第一外源核酸构建体,其包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白质进行编码的核苷酸序列,其中所述ALAS包括至少第一血红素响应基序(HRM),且其中所述ALAS包括所述第一HRM中的突变;和

[0105] 第二外源核酸构建体,其包括对血红素结合蛋白进行编码的核苷酸序列,

[0106] 其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

[0107] 实施例2为根据实施例1所述的细胞,其中所述细胞为真菌细胞。

[0108] 实施例3为根据实施例2所述的细胞,其中所述细胞为曲霉或木霉细胞。

[0109] 实施例4为根据实施例1至3中任一项所述的细胞,其中所述细胞为酵母细胞。

[0110] 实施例5为根据实施例4所述的细胞,其中所述酵母细胞为甲基营养型酵母细胞。

[0111] 实施例6为根据实施例5所述的细胞,其中所述甲基营养型酵母细胞为毕赤酵母属细胞、假丝酵母属细胞、汉逊酵母属细胞或球拟酵母属细胞。

[0112] 实施例7为根据实施例5至6中任一项所述的细胞,其中所述甲基营养型酵母细胞为嗜甲醇毕赤酵母细胞、巴斯德毕赤酵母细胞、博伊丁假丝酵母细胞或多形汉逊酵母细胞。

[0113] 实施例8为根据实施例5至7中任一项所述的细胞,其中所述甲基营养型酵母细胞为巴斯德毕赤酵母细胞。

[0114] 实施例9为根据实施例1至8中任一项所述的细胞,其中所述第一HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0115] 实施例10为根据实施例1至9中任一项所述的细胞,其中所述ALAS蛋白质包括第二HRM,且其中所述ALAS蛋白质包括所述第二HRM中的突变。

[0116] 实施例11为根据实施例10所述的细胞,其中所述第二HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0117] 实施例12为根据实施例10或实施例11所述的细胞,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述另一氨基酸相同。

[0118] 实施例13为根据实施例10或实施例11所述的细胞,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述另一氨基酸不相同。

[0119] 实施例14为根据实施例10至13中任一项所述的细胞,其中所述ALAS蛋白质包括第三HRM,且其中所述ALAS蛋白质包括所述第三HRM中的突变。

[0120] 实施例15为根据实施例14所述的细胞,其中所述第三HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0121] 实施例16为根据实施例15所述的细胞,其中对于所述第一HRM中的所述突变、所述第二HRM中的所述突变和所述第三HRM中的所述突变,所述另一氨基酸相同。

[0122] 实施例17为根据实施例9至16中任一项所述的细胞,其中所述另一氨基酸选自自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、谷氨酸、天冬氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。

[0123] 实施例18为根据实施例9至16中任一项所述的细胞,其中所述另一氨基酸选自非极性脂肪族氨基酸、芳香族氨基酸、极性不带电氨基酸或带正电氨基酸。

[0124] 实施例19为根据实施例18所述的细胞,其中所述非极性脂肪族氨基酸选自自由以下组成的群组:甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸和缬氨酸。

[0125] 实施例20为根据实施例18所述的细胞,其中所述芳香族氨基酸选自自由以下组成的群组:苯丙氨酸、色氨酸和酪氨酸。

[0126] 实施例21为根据实施例18所述的细胞,其中所述极性不带电氨基酸选自自由以下组成的群组:极性不带电氨基酸丝氨酸、苏氨酸、天冬酰胺或谷氨酰胺。

[0127] 实施例22为根据实施例18所述的细胞,其中所述带正电氨基酸选自自由以下组成的群组:精氨酸、组氨酸和赖氨酸。

[0128] 实施例23为根据实施例9至22中任一项所述的细胞,其中所述另一氨基酸为丝氨酸。

[0129] 实施例24为根据实施例9至22中任一项所述的细胞,其中所述另一氨基酸为丙氨酸。

[0130] 实施例25为根据实施例9至22中任一项所述的细胞,其中所述另一氨基酸为苯丙氨酸。

[0131] 实施例26为根据实施例9至22中任一项所述的细胞,其中所述另一氨基酸为组氨酸。

[0132] 实施例27为根据实施例1至26中任一项所述的细胞,其中所述第一HRM为HRM1。

[0133] 实施例28为根据实施例10至27中任一项所述的细胞,其中所述第二HRM为HRM2。

[0134] 实施例29为根据实施例1至26中任一项所述的细胞,其中所述第一HRM为HRM2。

[0135] 实施例30为根据实施例10至27中任一项所述的细胞,其中所述第二HRM为HRM1。

[0136] 实施例31为根据实施例1至30中任一项所述的细胞,其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。

[0137] 实施例32为根据实施例1至30中任一项所述的细胞,其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少95%序列一致性的核酸序列。

[0138] 实施例33为根据实施例1至30中任一项所述的细胞,其中所述第一外源核酸构建体包括SEQ ID NO:30中的核酸序列。

[0139] 实施例34为根据实施例1至32中任一项所述的细胞,其中所述ALAS蛋白质包括与

SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

[0140] 实施例35为根据实施例1至32中任一项所述的细胞,其中所述ALAS蛋白质包括与SEQ ID NO:29中的氨基酸序列具有至少95%序列一致性的氨基酸序列。

[0141] 实施例36为根据实施例1至35中任一项所述的细胞,其中所述ALAS蛋白质包括SEQ ID NO:31中的氨基酸序列。

[0142] 实施例37为根据实施例1至36中任一项所述的细胞,其中所述血红素结合蛋白选自自由以下组成的群组:球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。

[0143] 实施例38为根据实施例1至36中任一项所述的细胞,其中所述血红素结合蛋白选自自由以下组成的群组:雄性球蛋白、血绿蛋白、细胞球蛋白、无脊椎动物血红蛋白、黄素血红蛋白、球蛋白E、球蛋白X、球蛋白Y、血红蛋白、组胺球蛋白、豆血红蛋白、肌血球素、神经球蛋白、非共生血红蛋白、原球蛋白和截短血红蛋白。

[0144] 实施例39为根据实施例1至36中任一项所述的细胞,其中所述血红素结合蛋白为非共生血红蛋白。

[0145] 实施例40为根据实施例1至36中任一项所述的细胞,其中所述血红素结合蛋白为豆血红蛋白。

[0146] 实施例41为根据实施例1至40中任一项所述的细胞,其中所述血红素结合蛋白包括与SEQ ID NO:1-27中任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

[0147] 实施例42为根据实施例1至41中任一项所述的细胞,其进一步包括第三核酸构建体,所述第三核酸构建体包括对转录因子进行编码的核苷酸序列,其中所述第三核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件或第三启动子元件。

[0148] 实施例43为根据实施例42所述的细胞,其中所述第一启动子元件包括所述转录因子的识别序列。

[0149] 实施例44为根据实施例42或43中任一项所述的细胞,其中所述第二外源核酸构建体可操作地连接到第二启动子元件,且其中所述第二启动子元件包括所述转录因子的识别序列。

[0150] 实施例45为根据实施例42或43中任一项所述的细胞,其中所述第三核酸构建体可操作地连接到所述第三启动子元件,且其中所述第三启动子元件包括所述转录因子的识别序列。

[0151] 实施例46为根据实施例1至45中任一项所述的细胞,其进一步包括第四核酸构建体,所述第四核酸构建体包括对参与血红素生物合成的蛋白质进行编码的核苷酸序列,其中所述第四核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件、所述第三启动子元件或第四启动子元件。

[0152] 实施例47为根据实施例46所述的细胞,其中所述参与血红素生物合成的蛋白质选自自由以下组成的群组:ALA脱水酶、胆色素原脱氨酶、UPG III合成酶、UPG III脱羧酶、CPG氧化酶、PPG氧化酶和亚铁螯合酶。

[0153] 实施例48为根据实施例1至47中任一项所述的细胞,其中所述第一外源核酸构建体为异源核酸构建体。

[0154] 实施例49为根据实施例1至48中任一项所述的细胞,其中所述第二外源核酸构建

体为异源核酸构建体。

[0155] 实施例50为根据实施例1至47中任一项所述的细胞,其中所述血红素结合蛋白为外源血红素结合蛋白。

[0156] 实施例51为根据实施例1至47或50中任一项所述的细胞,其中所述血红素结合蛋白为异源血红素结合蛋白。

[0157] 实施例52为一种在细胞中产生血红素结合蛋白的方法,其包括:

[0158] 表达第一外源核酸构建体,所述第一外源核酸构建体包括对可操作地连接到第一启动子元件的氨基乙酰丙酸合成酶(ALAS)蛋白质进行编码的核苷酸序列,其中所述ALAS包括至少第一血红素响应基序(HRM),且其中所述ALAS包括所述第一HRM中的突变;和

[0159] 表达第二外源核酸构建体,所述第二外源核酸构建体包括对血红素结合蛋白进行编码的核苷酸序列,

[0160] 其中包括对所述血红素结合蛋白进行编码的核苷酸序列的所述第二外源核酸构建体可操作地连接到所述第一启动子元件或可操作地连接到第二启动子元件。

[0161] 实施例53为根据52所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的对应方法高至少5%的效价产生所述血红素结合蛋白。

[0162] 实施例54为根据实施例52或实施例53所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的对应方法高至少10%的效价产生所述血红素结合蛋白。

[0163] 实施例55为根据实施例52或实施例53所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的对应方法高至少15%的效价产生所述血红素结合蛋白。

[0164] 实施例56为根据实施例52或实施例53所述的方法,其中所述方法以比缺少所述第一外源核酸构建体的对应方法高至少20%的效价产生所述血红素结合蛋白。

[0165] 实施例57为根据实施例52所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的对应方法高至少5%的效价产生所述血红素结合蛋白。

[0166] 实施例58为根据实施例52所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的对应方法高至少10%的效价产生所述血红素结合蛋白。

[0167] 实施例59为根据实施例52所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的对应方法高至少15%的效价产生所述血红素结合蛋白。

[0168] 实施例60为根据实施例52所述的方法,其中所述方法以比缺少所述第一HRM中的所述突变的对应方法高至少20%的效价产生所述血红素结合蛋白。

[0169] 实施例61为根据实施例52至60中任一项所述的方法,其中所述方法在不添加甲醇的情况下进行。

[0170] 实施例62为根据实施例52至61中任一项所述的方法,其中所述细胞为真菌细胞。

[0171] 实施例63为根据实施例62所述的方法,其中所述细胞为曲霉或木霉细胞。

[0172] 实施例64为根据实施例62至63中任一项所述的方法,其中所述细胞为酵母细胞。

[0173] 实施例65为根据实施例64所述的方法,其中所述酵母细胞为甲基营养型酵母细胞。

[0174] 实施例66为根据实施例65所述的方法,其中所述甲基营养型酵母细胞为毕赤酵母属细胞、假丝酵母属细胞、汉逊酵母属细胞或球拟酵母属细胞。

[0175] 实施例67为根据实施例65至66中任一项所述的方法,其中所述甲基营养型酵母细

胞为嗜甲醇毕赤酵母细胞、巴斯德毕赤酵母细胞、博伊丁假丝酵母细胞或多形汉逊酵母细胞。

[0176] 实施例68为根据实施例51至53中任一项所述的方法,其中所述甲基营养型酵母细胞为巴斯德毕赤酵母细胞。

[0177] 实施例69为根据实施例52至68中任一项所述的方法,其中所述第一HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0178] 实施例70为根据实施例52至69中任一项所述的方法,其中所述ALAS蛋白质包括第二HRM,且其中所述ALAS蛋白质包括所述第二HRM中的突变。

[0179] 实施例71为根据实施例70所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的对应方法高至少5%的效价产生所述血红素结合蛋白。

[0180] 实施例72为根据实施例70所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的对应方法高至少10%的效价产生所述血红素结合蛋白。

[0181] 实施例73为根据实施例70所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的对应方法高至少15%的效价产生所述血红素结合蛋白。

[0182] 实施例74为根据实施例70所述的方法,其中所述方法以比缺少所述第一HRM和第二HRM中的所述突变的对应方法高至少20%的效价产生所述血红素结合蛋白。

[0183] 实施例75为根据实施例70至74中任一项所述的方法,其中所述第二HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0184] 实施例76为根据实施例70至75中任一项所述的方法,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述另一氨基酸相同。

[0185] 实施例77为根据实施例70至75中任一项所述的方法,其中对于所述第一HRM中的所述突变和所述第二HRM中的所述突变,所述另一氨基酸不相同。

[0186] 实施例78为根据实施例70至77中任一项所述的细胞,其中所述ALAS蛋白质包括第三HRM,且其中所述ALAS蛋白质包括所述第三HRM中的突变。

[0187] 实施例79为根据实施例78所述的细胞,其中所述第三HRM中的所述突变为半胱氨酸到另一氨基酸的突变。

[0188] 实施例80为根据实施例79所述的细胞,其中对于所述第一HRM中的所述突变、所述第二HRM中的所述突变和所述第三HRM中的所述突变,所述另一氨基酸相同。

[0189] 实施例81为根据实施例69至80中任一项所述的方法,其中所述另一氨基酸选自由以下组成的群组:精氨酸、组氨酸、赖氨酸、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺、硒代半胱氨酸、甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸、天冬氨酸、谷氨酸、苯丙氨酸、色氨酸、酪氨酸和缬氨酸。

[0190] 实施例82为根据实施例69至80中任一项所述的方法,其中所述另一氨基酸选自非极性脂肪族氨基酸、芳香族氨基酸、极性不带电氨基酸或带正电氨基酸。

[0191] 实施例83为根据实施例82所述的方法,其中所述非极性脂肪族氨基酸选自由以下组成的群组:甘氨酸、脯氨酸、丙氨酸、异亮氨酸、亮氨酸、甲硫氨酸和缬氨酸。

[0192] 实施例84为根据实施例82所述的方法,其中所述芳香族氨基酸选自由以下组成的群组:苯丙氨酸、色氨酸和酪氨酸。

[0193] 实施例85为根据实施例82所述的细胞,其中所述极性不带电氨基酸选自由以下组

成的群组：极性不带电氨基酸丝氨酸、苏氨酸、天冬酰胺或谷氨酰胺。

[0194] 实施例86为根据实施例82所述的方法，其中所述带正电氨基酸选自由以下组成的群组：精氨酸、组氨酸和赖氨酸。

[0195] 实施例87为根据实施例69至86中任一项所述的方法，其中所述另一氨基酸为丝氨酸。

[0196] 实施例88为根据实施例69至86中任一项所述的方法，其中所述另一氨基酸为丙氨酸。

[0197] 实施例89为根据实施例69至86中任一项所述的方法，其中所述另一氨基酸为苯丙氨酸。

[0198] 实施例90为根据实施例69至86中任一项所述的方法，其中所述另一氨基酸为组氨酸。

[0199] 实施例91为根据实施例52至90中任一项所述的细胞，其中所述第一HRM为HRM1。

[0200] 实施例92为根据实施例70至91中任一项所述的细胞，其中所述第二HRM为HRM2。

[0201] 实施例93为根据实施例52至90中任一项所述的细胞，其中所述第一HRM为HRM2。

[0202] 实施例94为根据实施例70至91中任一项所述的细胞，其中所述第二HRM为HRM1。

[0203] 实施例95为根据实施例52至94中任一项所述的方法，其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少90%序列一致性的核酸序列。

[0204] 实施例96为根据实施例52至95中任一项所述的方法，其中所述第一外源核酸构建体包括与SEQ ID NO:28中的核酸序列具有至少95%序列一致性的核酸序列。

[0205] 实施例97为根据实施例52至96中任一项所述的方法，其中所述第一外源核酸构建体包括SEQ ID NO:30中的核酸序列。

[0206] 实施例98为根据实施例52至97中任一项所述的方法，其中所述ALAS蛋白质包括与SEQ ID NO:29中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

[0207] 实施例99为根据实施例52至98中任一项所述的方法，其中所述ALAS蛋白质包括与SEQ ID NO:29中的氨基酸序列具有至少95%序列一致性的氨基酸序列。

[0208] 实施例100为根据实施例52至99中任一项所述的方法，其中所述ALAS蛋白质包括SEQ ID NO:31中的氨基酸序列。

[0209] 实施例101为根据实施例52至100中任一项所述的方法，其中所述血红素结合蛋白选自由以下组成的群组：球蛋白、细胞色素、细胞色素c氧化酶、木质素酶、过氧化氢酶和过氧化酶。

[0210] 实施例102为根据实施例52至100中任一项所述的方法，其中所述血红素结合蛋白选自由以下组成的群组：雄性球蛋白、血绿蛋白、细胞球蛋白、无脊椎动物血红蛋白、黄素血红蛋白、球蛋白E、球蛋白X、球蛋白Y、血红蛋白、组胺球蛋白、豆血红蛋白、肌血球素、神经球蛋白、非共生血红蛋白、原球蛋白和截短血红蛋白。

[0211] 实施例103为根据实施例52至100中任一项所述的方法，其中所述血红素结合蛋白为非共生血红蛋白。

[0212] 实施例104为根据实施例52至100中任一项所述的方法，其中所述血红素结合蛋白为豆血红蛋白。

[0213] 实施例105为根据实施例52至100中任一项所述的方法，其中异源血红素结合蛋白

包括与SEQ ID NO:1-27中任一个中的氨基酸序列具有至少90%序列一致性的氨基酸序列。

[0214] 实施例106为根据实施例52至105中任一项所述的方法,其进一步包括表达第三核酸构建体,所述第三核酸构建体包括对转录因子进行编码的核苷酸序列,其中所述第三核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件或第三启动子元件。

[0215] 实施例107为根据实施例106所述的方法,其中所述第一启动子元件包括所述转录因子的识别序列。

[0216] 实施例108为根据实施例107所述的方法,其中所述第二外源核酸构建体可操作地连接到第二启动子元件,且其中所述第二启动子元件包括所述转录因子的识别序列。

[0217] 实施例109为根据实施例107至108中任一项所述的方法,其中所述第三核酸构建体可操作地连接到所述第三启动子元件,且其中所述第三启动子元件包括所述转录因子的识别序列。

[0218] 实施例110为根据实施例52至109中任一项所述的方法,其进一步包括表达第四核酸构建体,所述第四核酸构建体包括对参与血红素生物合成的蛋白质进行编码的核苷酸序列,其中所述第四核酸构建体可操作地连接到所述第一启动子元件、所述第二启动子元件、所述第三启动子元件或第四启动子元件。

[0219] 实施例111为根据实施例110所述的方法,其中所述参与血红素生物合成的蛋白质选自自由以下组成的群组:ALA脱水酶、胆色素原脱氨酶、UPG III合成酶、UPG III脱羧酶、CPG氧化酶、PPG氧化酶和亚铁螯合酶。

[0220] 实施例112为根据实施例52至111中任一项所述的方法,其中所述第一外源核酸构建体为异源核酸构建体。

[0221] 实施例113为根据实施例52至112中任一项所述的方法,其中所述第二外源核酸构建体为异源核酸构建体。

[0222] 实施例114为根据实施例52至113中任一项所述的方法,其中所述血红素结合蛋白为外源血红素结合蛋白。

[0223] 实施例115为根据实施例52至111或114中任一项所述的方法,其中所述血红素结合蛋白为异源血红素结合蛋白。

[0224] 实例

[0225] 实例1

[0226] 聚合酶链反应

[0227] 使用Phusion高保真PCR预混液(新英格兰生物实验室(New England Biolabs),目录号M0531)、各0.6μM的正向引物和反向引物以及10至50ng的模板DNA从基因组DNA或质粒DNA模板扩增所关注基因。反应条件如下:

[0228]	1 次循环	初始变性	98°C	1 分钟
	25 次循环	变性	98°C	10 秒
		粘接	60°C	20 秒
		扩展	72°C	30 秒/kb
[0229]	1 次循环	最终扩展	72°C	5 分钟
		保持	4°C	一直

[0230] 使用引物(ATGGAGTTTGTCTGCCCGTCAG; SEQ ID NO:65)和

(CTACAATCTGACTCCTGATGAGGTTTC;SEQ ID NO:66)从基因组DNA扩增巴斯德毕赤酵母ALAS基因(KEGG标识符PAS_chr2-1_0716,可访问京都基因与基因组百科全书(Kyoto Encyclopedia of Genes and Genomes;KEGG)网站)。在实例中,wtALAS表示ALAS的此原生序列(SEQ ID NO:28)。

[0231] 实例2

[0232] ALAS克隆和诱变

[0233] 遵循制造商的建议,将PCR产物使用NucleoSpin胶/PCR纯化试剂盒(NucleoSpin Gel and PCR Clean-Up)(宝生物工程(Takara Bio),目录号740609)进行纯化,且使用Zero Blunt TOPO PCR克隆试剂盒(赛默飞世尔科技(Thermo Fisher Scientific),目录号K280020)在pCR-BluntII-TOPO载体中进行克隆。使用两组引物(set1:CGTCAGTCCATGAATGCCTCTCCCTTTGTCAGGTCAACTTC;SEQ ID NO:37和GAAGTTGACCTGACAAAGGGAGAGGCATTCATGGACTGACG;SEQ ID NO:38,set2:GCTGCTACTGCTAGTCATTCTCCCGTGGTTGGCCCTG;SEQ ID NO:39和CAGGGCCAACCGGAGAAATGACTAGCAGTAGCAGC;SEQ ID NO:40)和QuikChange II XL定点诱变试剂盒(安捷伦(Agilent),目录号200521)对所得载体(40ng经纯化的质粒)进行定点诱变,以产生携带位置12和39处的半胱氨酸残基到丝氨酸残基突变的mutALAS(SEQ ID NO:30)。在实例中,除非另外规定,否则mutALAS确切地表示这两个突变。反应条件如下:

[0234]	1 次循环	初始变性	95°C	1 分钟
	18 次循环	变性	95°C	50 秒
		粘接	60°C	50 秒
		扩展	68°C	11 分钟
	1 次循环	最终扩展	68°C	7 分钟
		保持	4°C	一直

[0235] 通过分别使用引物set1(SEQ ID NO:37、SEQ ID NO:38)和set2(SEQ ID NO:39、SEQ ID NO:40)对携带wtALAS的载体进行定点诱变来执行单一半胱氨酸到丝氨酸突变C12S或C39S。

[0236] 实例3

[0237] ALAS(野生型或突变型、C12S、C39S)基因整合盒的构建

[0238] 将wtALAS或mutALAS的整合盒设计成使得基因整合在来自巴斯德毕赤酵母的甲醇诱导型醇氧化酶1(AOX1)启动子元件控制下且在后面紧接来自巴斯德毕赤酵母FDH1基因的转录终止序列的翻译停止信号之前发生。线性构建体含有3'端一半的启动子元件,接着是ALAS基因(野生型或突变型),接着是FDH1转录终止子。此后紧接选择盒,所述选择盒含有来自棉阿舒囊霉(Ashbya gossypii)的pTEF启动子元件、来自构巢曲霉(Aspergillus nidulans)的乙酰胺酶基因(amdS)和来自棉阿舒囊霉的TEF终止子。最后,构建体含有5'端一半的启动子元件(参见例如图5)。

[0239] 使用重叠PCR产生线性构建体[3'pAOX1-ALAS(wt/mut)-FDH1tt-pTEF-amdS-TEFtt-5'pAOX1)。用于扩增线性构建体的引物:

[0240]	重叠 PCR	引物序列	SEQ ID NO:
	PCR1	AAACGCTGTCTTGGAACCTAATATGAC	41
		GACGGGCGACAACTCCATCGTTTCGAATAATTAGTTG	42
	PCR2	CAACTAATTATTTCGAAACGATGGAGTTTGTGCGCCCGTCAG	43
		AATTAAATACATTTCAACTACAATCTGACTCCTGATGAGGTTTCG	44
	PCR3	CCTCATCAGGAGTCAGATTGTAGTTGAAATGTATTTAATTG	45
		AAACTGTCAGTTTTGGGCCATTG	46

[0241] 如上文所指出进行个别PCR反应条件。PCR产物使用NucleoSpin胶/PCR纯化试剂盒(宝生物工程, 目录号740609)进行纯化。通过在第2部分PCR反应中以等摩尔比率混合三种经纯化的PCR扩增子, 使用0.5U铂Pfx DNA聚合酶(赛默飞世尔科技, 目录号11708039)、1x扩增缓冲液、0.3mM dNTP和1mM MgCl₂来进行PCR产物的最终重叠。PCR条件为:

[0242] 第1部分:

[0243]	1 次循环	初始变性	94°C	2 分钟
	15 次循环	变性	94°C	15 秒
		粘接	58°C	30 秒
		扩展	68°C	5 分钟
		保持	4°C	

[0244] 在第1部分之后, 向反应混合物中加入引物AAACGCTGTCTTGGAACCTAATATGAC (SEQ ID NO:41) 和AAACTGTCAGTTTTGGGCCATTG (SEQ ID NO:46) (0.3μM最终浓度) 以及1.5U Pfx, 且继续到第2部分。

[0245] 第2部分:

[0247]	1 次循环	初始变性	94°C	2 分钟
	20 次循环	变性	94°C	15 秒
		粘接	58°C	30 秒
		扩展	68°C	5 分钟
		保持	4°C	一直

[0248] 示例性突变型核酸(例如SEQ ID NO:30)和蛋白质(例如SEQ ID NO:31)的序列提供于图7中。

[0249] 在后接来自巴斯德毕赤酵母FDH1基因的转录终止序列的经修饰pAOX1控制下, 在自主复制载体(panARS)中构建wtALAS和mutALAS的表达质粒。所述载体赋予对G418(遗传霉素(Geneticin))的抗性。遵循制造商的建议, 使用Gibson Assembly预混液(新英格兰生物实验室, 目录号E2611L)进行插入序列GFP、wtALAS和mutALAS在载体中的克隆。

[0250] 实例4

[0251] 巴斯德毕赤酵母转化感受态细胞的制备

[0252] 巴斯德毕赤酵母(法夫驹形氏酵母)的所选菌株在25ml YPD培养基中生长到中期指数生长期(约20D)。通过以930×g离心15分钟来收集细胞。将细胞沉淀物再悬浮于2ml的80%YPD和200mM HEPES的溶液(pH 6.8)中。添加75μl的1M DTT。再悬浮的细胞沉淀物在30°C下以100rpm混合25分钟。将40ml体积的冰冷无菌水添加到悬浮液中, 并且通过以1125×g离心15分钟来收集细胞并且放在冰上。将细胞沉淀物再悬浮于40ml冰冷的水中, 且如前所述收集以进行两次额外洗涤步骤。然后将细胞沉淀物再悬浮于20ml冰冷的1M山梨糖醇中, 且如前所述通过离心进行收集。将最终细胞沉淀物悬浮于0.3ml冰冷的无菌1M山梨糖醇中,

等分试样且在-80℃下冷冻。

[0253] 实例5

[0254] 向巴斯德毕赤酵母中的转化

[0255] 对于基因组整合,使用具有GenePulser (拜耳雷德 (Bio-Rad)) 设置的1mm间距GenePulser电击杯(拜耳雷德,目录号1652083)以1.15kV将100至300ng的线性化DNA转化到30μl电转感受态巴斯德毕赤酵母细胞中。将1ml YPD/1M山梨糖醇(1:1体积/体积)立即添加到所述细胞中。在以100rpm进行振荡的情况下,在30℃下回收细胞3小时。将100μl回收混合物涂覆在含有5mM乙酰胺的酵母碳底板(天惠华(Teknova),目录号Y5216)上。将板在30℃下培育48小时。将个别克隆株划线到含有乙酰胺的酵母碳底板上以获得单一菌落,且使用经分离菌落进行菌落PCR或gDNA prep以确认基因整合到染色体和序列整合构建体中。

[0256] 对于质粒DNA转化,遵循与基因组整合相同的步骤,不同之处在于在细胞电穿孔和回收之后,将50至100μl回收混合物涂覆在含有300μg/ml G418(遗传霉素)的YPD板上。

[0257] 实例6

[0258] 菌株St2、St3、St5和St6的构建

[0259] 高产量亲本菌株(St1)具有受甲醇诱导型强启动子pAOX1(醇氧化酶1)控制的预先存在重组ALAS,以及其它受pAOX1驱动的血红素酶、碳响应性转录因子MxR1和多个LegH拷贝。低产量亲本LegH菌株(St4)缺少重组MxR1、ALAS和其它血红素酶,氨基乙酰丙酸脱水酶(ALAD)除外。

[0260] 感受态St1细胞(表1;高LegH效价菌株)用线性盒(wtALAS和mutALAS的线性盒)中的每一个进行转化,且基于其在含有乙酰胺作为唯一氮源的琼脂板上生长的能力来选择含有amdS选择盒的转化体。将所得菌株(分别具有wtALAS和mutALAS的整合盒的St2和St3,表1)纯化、分离,且通过菌落PCR和测序来验证受pAOX1驱动的wtALAS或mutALAS的存在。类似地,将感受态St4细胞(低LegH效价菌株)用wtALAS和mutALAS的线性盒进行转化,分别获得St5和St6。

[0261] 表1.具有整合于基因组中的LegH的菌株

[0262]	亲本	添加的重组 wtALAS 的 1 个额外拷贝	添加的重组 mutALAS 的 1 个额外拷贝
	高 LegH 效价菌株(菌株 St1)	菌株 St2	菌株 St3
	低 LegH 效价菌株(菌株 St4)	菌株 St5	菌株 St6

[0263] 实例7

[0264] ALAS基因拷贝数目分析

[0265] 通过基于探针的qPCR测量不同菌株中的ALAS基因拷贝数目。简单来说,基因组DNA(20ng)用1×PrimeTime基因表达主混合物(整合DNA技术(Integrated DNA Technologies),目录号1055770)进行扩增,且在实时qPCR CFX96机器(拜耳雷德)中进行PrimeTime qPCR探针分析。相对于肌动蛋白进行qPCR归一化。遵循双 Δ Ct分析方法来计算菌株之间所关注基因的相对拷贝数目。引物和探针序列为:

[0266]	基因	引物序列	SEQ ID. NO.	探针序列	SEQ ID. NO.
	ALAS	GCTCTCCAACAGCAG AGATAC	47	6-FAM/AAGCCCAAA/Zen/CCTCCG ACATTGCTA/3IABkFQ	51
[0267]		GTCCATACGGATCGG AGAAAC	48		
	肌动蛋白	AGCAACATCCCTGAT TCCG	49	HEX/TCGCCGTAA/Zen/GTTCTTGG TT TAGACGTTC/3IABkFQ	52
		ATGCGTACCTTCAAT CCTGG	50		

[0268] 如根据相对于肌动蛋白水平归一化的ALAS的qPCR所测量,经工程改造的菌株(表1)含有相同数目的额外ALAS基因拷贝,一个wtALAS的拷贝或一个mutALAS的拷贝。因此,针对所得菌株的表型的任何差异(例如,St2与St3和St5与St6),排除ALAS基因剂量。

[0269] 实例8

[0270] 血红素酶、MxR1、LegH和Mb基因的PCR检测

[0271] 通过对pAOX1驱动的重组血红素生物合成途径酶、MxR1和LegH基因的PCR来进行菌株表征。PCR的正向引物序列为对pAOX1具有特异性的TAGCGCAGTCTCTCTATCGCTTC (SEQ ID NO:53)。反向引物对如下所示的每一所关注基因具有特异性。

	基因	反向引物	扩增子大小	SEQ ID NO:
[0272]	5-氨基乙酰丙酸合成酶	CACTGGGTTGTGCACATTGG	1995 bp	54
	δ-氨基乙酰丙酸脱水酶	ACAATATTCTTCTCTGCCGC	1268 bp	55
	胆色素原脱氢酶	TTGATCTCGTCAAGAATGCG	1358 bp	56
	尿卟啉原 III 合成酶	TAGGTGCCACAACCTTTTGGTTTC	1102 bp	57
	尿卟啉原脱羧酶	GATCCAATGCGATGACATTCTTGT	1430 bp	58
	粪卟啉原 III 氧化酶	ACCTGCAATAACTCCTCTTCTCTG	1301 bp	59
	原卟啉原氧化酶	CCACTGAGGGTAGCCGAATC	2027 bp	60
	亚铁螯合酶	GGGCTCTGAAAACTCTTTTGG	1486 bp	61
	MxR1	GCATGTCTCAATAACAGATCTCGACGG	629 bp	62
	LegH	AAGCCTCTTGTTTTTCTGTAAATGCAC	382 bp	63
	牛 Mb	TGATGGCGTCCGAGATGAACTC	688 bp	64

[0273] 下方示出这些反应的结果。

[0274]	重组基因(受 pAOX1 控制)	St1	St2	St3	St4	St5	St6	St7
	5-氨基乙酰丙酸合成酶	x	x	x	-	x	x	x
	δ-氨基乙酰丙酸脱水酶	x	x	x	x	x	x	x
	胆色素原脱氢酶	x	x	x	-	-	-	x
	尿卟啉原 III 合成酶	x	x	x	-	-	-	x
	尿卟啉原脱羧酶	x	x	x	-	-	-	x
	粪卟啉原 III 氧化酶	x	x	x	-	-	-	x
	原卟啉原氧化酶	x	x	x	-	-	-	x
[0275]	亚铁螯合酶	x	x	x	-	-	-	x
	MxR1	x	x	x	-	-	-	x
	LegH	x	x	x	x	x	x	-
	Mb (牛肌血球素)	-	-	-	-	-	-	x

[0276] 实例9

[0277] 菌株St4、St5和St6的摇瓶培养

[0278] 在以200rpm进行振荡的情况下,在30℃下将菌株接种到生长培养基(1%酵母提取物、2%蛋白胨、补充有1%甘油)中过夜。第二天,用补充有1%甲醇和1%右旋糖的YP培养基将过夜的培养物稀释到0.5至0.7的OD600。使培养物生长48小时,并且在4℃下通过以约4000g离心15分钟来进行收集。

[0279] 不同于St1,低LegH效价菌株St4不具有预先存在的重组ALAS拷贝。与St5(wtALAS)相比,整合mutALAS拷贝(St6)引起约30%的LegH效价改进。效价计算是基于如通过液相色谱方法所测量的LegH含量,如实例13中所描述。

[0280] 相对LegH效价展示于表2中。

[0281] 表2

[0282]		St5	St6
	相对LegH效价	1.00	1.32

[0283] 实例10

[0284] 菌株St1、St2和St3的2L培养

[0285] 菌株St1、St2和St3在30℃下在2L发酵罐中在含有右旋糖作为主要碳源的培养基中生长。不使用甲醇。在含有预先存在的重组ALAS的高LegH效价菌株(St1)的背景下,过表达mutALAS的St3的LegH效价与亲本St1相比改进>30%。当与wtALAS(St2中的wtALAS)相比时,muALAS(St3中的muALAS)引起20%的LegH效价改进。效价计算是基于如通过液相色谱方法所测量的LegH含量,如实例13中所描述。

[0286] 相对LegH效价展示于表3中。

[0287] 表3

[0288]		St1	St2	St3
	相对LegH效价	1.00	1.10	1.37

[0289] 因此,使两个ALAS HRM突变且在具有不同LegH效价的毕赤酵母属菌株中过表达突变型ALAS(muALAS)引起LegH效价的进一步改进。此外,这表明muALAS对LegH效价的改进与菌株基因组成(MxR1和除ALAS外的血红素酶)和甲醇存在无关。

[0290] 实例11

[0291] muALAS改进血红素水平

[0292] 基于反相高效液相色谱对总血红素的定量分析表明,具有muALAS的菌株与含有wtALAS的菌株相比累积更高血红素水平。表4中示出菌株St3(muALAS)与菌株St2(wtALAS)中的血红素定量。另外,muALAS增加多个菌株中的血红素荷载。

[0293] 表4

[0294]		St2	St3
	相对血红素效价	1.00	2.7

[0295] 实例12

[0296] muALAS的ALAS蛋白质水平

[0297] 一般认为血红素以反馈方式调控ALAS水平,从而调控其自身的水平。在蛋白质层级方面,当菌株在具有右旋糖的2L发酵罐中生长时,St3(muALAS)中的ALAS水平与St2(wtALAS)相比高3倍(表5),如通过鸟枪质谱所定量。

[0298] 表5

[0299]	St2	St3
相对ALAS效价	1.00	3.3

[0300] 实例13

[0301] 豆血红蛋白的定量

[0302] 细胞培养液样品向下沉淀 (以 $4000 \times g$ 、 4°C 、30分钟) 并且倾析。然后将沉淀物样品用溶解缓冲液 (150mM NaCl、50mM磷酸钾、pH 7.4) 稀释四次。将300 μL 各再悬浮液分配到每孔具有120 μL 珠粒 (锆/二氧化硅珠粒 (0.5mm)) 的96孔深板中以进行细胞溶解。用微型珠粒搅拌器进行溶解3分钟, 接着将板在冰上冷却5分钟, 且接着进行另外2分钟的珠粒研磨。然后将板快速离心 (以 $4000 \times g$ 、 4°C 、30分钟)。通过0.2 μm 滤板过滤上清液 (以 $4000 \times g$ 、 4°C 、60分钟)。

[0303] 将经过滤的溶解物负载到具有尺寸排阻柱 (Acquity BEH SEC柱, 200 $\text{\AA}$, 1.7 μm , $4.6 \times 150\text{mm}$) 的UHPLC上。方法参数: 1) 移动相: 5mM NaCl, 50mM磷酸钾 (pH 7.4);

[0304] 2) 流动速率: 0.3mL/min; 3) 注入体积: 10 μL ; 4) 运行时间: 15分钟; 5) 样品盘温度: 4°C 。使用在280nm和415nm下的吸光度, 用纯化的LegH标准构建校准曲线。利用谷到谷峰值积分方法使用峰面积进行定量。280nm下的吸光度与所存在的多肽的量成比例, 且415nm下的吸光度与所存在的血红素的量成比例。在两个波长下在相同洗脱时间发现峰的情况下, 检测含有血红素的蛋白质。

[0305] 实例14

[0306] mutALAS改进牛肌血球素的水平

[0307] 除LegH外, 还评估牛肌血球素的产量。通过将黄牛 (*Bos taurus*) (NM_173881.2) 的肌血球素cDNA整合在含有于pAOX1控制下整合的血红素酶的重组拷贝和MxR1的菌株中而产生菌株St7 (St7表征描述于实例8中)。

[0308] 如实例3中所描述构建用于表达GFP (对照)、wtALAS和mutALAS的三种表达质粒。使用如实例5中所描述的方法, 通过在菌株S7中转化分别过表达GFP、wtALAS和mutALAS的这三种质粒而产生三种菌株St8、St9和St10, 如表6中所示。通过使用实例9中所描述的方法 (不同之处在于生长培养基补充有300 $\mu\text{g}/\text{ml}$ G418) 使转化体生长, 获得wtALAS或mutALAS的游离型表达。

[0309] 表6

	质粒表达	宿主菌株	宿主菌株表达	所得菌株
[0310]	1 GFP	St7	Mb	St8
	2 wtALAS			St9
	3 mutALAS (C12S、C39S)			St10

[0311] 使细胞沉淀, 且执行通过LC-MS进行的鸟枪蛋白质组学以定量Mb。当相对于菌株St8 (GFP) 中的肌血球素水平归一化时, St9和St10中的平均肌血球素水平如表7中所示。

[0312] 表7

[0313]	菌株	归一化的肌血球素水平
--------	----	------------

St8 (GFP)	1
St9 (wtALAS)	1.60
St10 (mutALAS)	2.58

[0314] 实例15

[0315] 到其它氨基酸的突变

[0316] 除半胱氨酸到丝氨酸突变以外,还评估额外突变。通过在后接FDH1转录终止序列的经修饰pA0X1控制下,在自主复制载体 (panARS) 中进行基因合成和克隆来构建mutALAS变体的表达质粒。此外,使用如实例5中所描述的方法在菌株St1中转化这些质粒,且当使用实例9中所描述的方法 (不同之处在于生长培养基补充有300 μ g/ml G418) 进行培养时,所得转化体产生LegH。这些菌株所产生的LegH的效价相对于wtALAS所产生的LegH的效价归一化,如通过实例13中所描述的方法所测定。结果展示于表8中。

[0317] 表8

变体	归一化的LegH效价倍数
C12, C39 (wtALAS)	1
C12S, C39S (mutALAS)	1.14
C12A, C39A	1.22
C12D, C39D	0.91
C12F, C39F	1.17
C12H, C39H	1.12

[0319] 实例16

[0320] 巴斯德毕赤酵母中的单一半胱氨酸到丝氨酸突变

[0321] 除评估巴斯德毕赤酵母中的ALAS的双突变 (C12S、C39S) 以外,还评估单一半胱氨酸到丝氨酸突变。如实例2中所描述,通过定点诱变在wtALAS中产生单一突变。如实例3中所描述构建用于表达wtALAS、mutALAS和单一ALAS突变 (C12S和C39S) 的表达质粒,且使用如实例5中所描述的方法在菌株St1中转化所述表达质粒。所得转化体在使用实例9中描述的方法 (不同之处在于生长培养基补充有300 μ g/ml G418) 进行培养时产生LegH。这些菌株所产生的LegH的效价相对于wtALAS所产生的LegH的效价归一化,如通过实例13中所描述的方法所测定。结果展示于表9中。

[0322] 表9

变体	归一化的LegH效价倍数
C12, C39 (wtALAS)	1
C12S, C39S	1.27
C12S	1.03
C39S	1.01

[0324] 其它实施例

[0325] 应了解,尽管本发明已经结合其具体实施方式进行了描述,但是前述描述旨在说明而非限制本发明的范围,本发明的范围由所附权利要求书的范围限定。其它方面、优点和修改在所附权利要求书的范围内。

序列表

<110> 非凡食品有限公司 (Impossible Foods Inc.)

<120> 用于产生含血红素蛋白的菌株和方法

<130> 38767-0158W01

<150> 62/838,770

<151> 2019-04-25

<160> 66

<170> PatentIn version 3.5

<210> 1

<211> 161

<212> PRT

<213> 绿豆 (*Vigna radiata*)

<400> 1

```

Met Thr Thr Thr Leu Glu Arg Gly Phe Thr Glu Glu Gln Glu Ala Leu
1           5           10           15
Val Val Lys Ser Trp Asn Val Met Lys Lys Asn Ser Gly Glu Leu Gly
          20           25           30
Leu Lys Phe Phe Leu Lys Ile Phe Glu Ile Ala Pro Ser Ala Gln Lys
          35           40           45
Leu Phe Ser Phe Leu Arg Asp Ser Thr Val Pro Leu Glu Gln Asn Pro
          50           55           60
Lys Leu Lys Pro His Ala Val Ser Val Phe Val Met Thr Cys Asp Ser
65           70           75           80
Ala Val Gln Leu Arg Lys Ala Gly Lys Val Thr Val Arg Glu Ser Asn
          85           90           95
Leu Lys Lys Leu Gly Ala Thr His Phe Arg Thr Gly Val Ala Asn Glu
          100          105          110
His Phe Glu Val Thr Lys Phe Ala Leu Leu Glu Thr Ile Lys Glu Ala
          115          120          125
Val Pro Glu Met Trp Ser Pro Ala Met Lys Asn Ala Trp Gly Glu Ala
          130          135          140
Tyr Asp Gln Leu Val Asp Ala Ile Lys Tyr Glu Met Lys Pro Pro Ser
145          150          155          160
Ser

```

<210> 2

<211> 133

<212> PRT

<213> 极端嗜酸甲烷氧化细菌 (*Methylophilum infernum*)

<400> 2

Met Ile Asp Gln Lys Glu Lys Glu Leu Ile Lys Glu Ser Trp Lys Arg
 1 5 10 15
 Ile Glu Pro Asn Lys Asn Glu Ile Gly Leu Leu Phe Tyr Ala Asn Leu
 20 25 30
 Phe Lys Glu Glu Pro Thr Val Ser Val Leu Phe Gln Asn Pro Ile Ser
 35 40 45
 Ser Gln Ser Arg Lys Leu Met Gln Val Leu Gly Ile Leu Val Gln Gly
 50 55 60
 Ile Asp Asn Leu Glu Gly Leu Ile Pro Thr Leu Gln Asp Leu Gly Arg
 65 70 75 80
 Arg His Lys Gln Tyr Gly Val Val Asp Ser His Tyr Pro Leu Val Gly
 85 90 95
 Asp Cys Leu Leu Lys Ser Ile Gln Glu Tyr Leu Gly Gln Gly Phe Thr
 100 105 110
 Glu Glu Ala Lys Ala Ala Trp Thr Lys Val Tyr Gly Ile Ala Ala Gln
 115 120 125
 Val Met Thr Ala Glu
 130

<210> 3

<211> 139

<212> PRT

<213> 风产液菌 (Aquifex aeolicus)

<400> 3

Met Leu Ser Glu Glu Thr Ile Arg Val Ile Lys Ser Thr Val Pro Leu
 1 5 10 15
 Leu Lys Glu His Gly Thr Glu Ile Thr Ala Arg Met Tyr Glu Leu Leu
 20 25 30
 Phe Ser Lys Tyr Pro Lys Thr Lys Glu Leu Phe Ala Gly Ala Ser Glu
 35 40 45
 Glu Gln Pro Lys Lys Leu Ala Asn Ala Ile Ile Ala Tyr Ala Thr Tyr
 50 55 60
 Ile Asp Arg Leu Glu Glu Leu Asp Asn Ala Ile Ser Thr Ile Ala Arg
 65 70 75 80
 Ser His Val Arg Arg Asn Val Lys Pro Glu His Tyr Pro Leu Val Lys
 85 90 95
 Glu Cys Leu Leu Gln Ala Ile Glu Glu Val Leu Asn Pro Gly Glu Glu
 100 105 110
 Val Leu Lys Ala Trp Glu Glu Ala Tyr Asp Phe Leu Ala Lys Thr Leu

115	120	125
Ile Thr Leu Glu Lys Lys Leu Tyr Ser Gln Pro		
130	135	
<210> 4		
<211> 145		
<212> PRT		
<213> 大豆 (Glycine max)		
<400> 4		
Met Gly Ala Phe Thr Glu Lys Gln Glu Ala Leu Val Ser Ser Ser Phe		
1 5 10 15		
Glu Ala Phe Lys Ala Asn Ile Pro Gln Tyr Ser Val Val Phe Tyr Thr		
20 25 30		
Ser Ile Leu Glu Lys Ala Pro Ala Ala Lys Asp Leu Phe Ser Phe Leu		
35 40 45		
Ser Asn Gly Val Asp Pro Ser Asn Pro Lys Leu Thr Gly His Ala Glu		
50 55 60		
Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Gly Gln Leu Lys Ala Asn		
65 70 75 80		
Gly Thr Val Val Ala Asp Ala Ala Leu Gly Ser Ile His Ala Gln Lys		
85 90 95		
Ala Ile Thr Asp Pro Gln Phe Val Val Val Lys Glu Ala Leu Leu Lys		
100 105 110		
Thr Ile Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Ser Ser		
115 120 125		
Ala Trp Glu Val Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala		
130 135 140		
Phe		
145		
<210> 5		
<211> 162		
<212> PRT		
<213> 大麦 (Hordeum vulgare)		
<400> 5		
Met Ser Ala Ala Glu Gly Ala Val Val Phe Ser Glu Glu Lys Glu Ala		
1 5 10 15		
Leu Val Leu Lys Ser Trp Ala Ile Met Lys Lys Asp Ser Ala Asn Leu		
20 25 30		
Gly Leu Arg Phe Phe Leu Lys Ile Phe Glu Ile Ala Pro Ser Ala Arg		
35 40 45		

Gln Met Phe Pro Phe Leu Arg Asp Ser Asp Val Pro Leu Glu Thr Asn
 50 55 60
 Pro Lys Leu Lys Thr His Ala Val Ser Val Phe Val Met Thr Cys Glu
 65 70 75 80
 Ala Ala Ala Gln Leu Arg Lys Ala Gly Lys Ile Thr Val Arg Glu Thr
 85 90 95
 Thr Leu Lys Arg Leu Gly Gly Thr His Leu Lys Tyr Gly Val Ala Asp
 100 105 110
 Gly His Phe Glu Val Thr Arg Phe Ala Leu Leu Glu Thr Ile Lys Glu
 115 120 125
 Ala Leu Pro Ala Asp Met Trp Gly Pro Glu Met Arg Asn Ala Trp Gly
 130 135 140
 Glu Ala Tyr Asp Gln Leu Val Ala Ala Ile Lys Gln Glu Met Lys Pro
 145 150 155 160
 Ala Glu
 <210> 6
 <211> 1153
 <212> PRT
 <213> 稻瘟病菌 (Magnaporthe oryzae)
 <400> 6
 Met Asp Gly Ala Val Arg Leu Asp Trp Thr Gly Leu Asp Leu Thr Gly
 1 5 10 15
 His Glu Ile His Asp Gly Val Pro Ile Ala Ser Arg Val Gln Val Met
 20 25 30
 Val Ser Phe Pro Leu Phe Lys Asp Gln His Ile Ile Met Ser Ser Lys
 35 40 45
 Glu Ser Pro Ser Arg Lys Ser Ser Thr Ile Gly Gln Ser Thr Arg Asn
 50 55 60
 Gly Ser Cys Gln Ala Asp Thr Gln Lys Gly Gln Leu Pro Pro Val Gly
 65 70 75 80
 Glu Lys Pro Lys Pro Val Lys Glu Asn Pro Met Lys Lys Leu Lys Glu
 85 90 95
 Met Ser Gln Arg Pro Leu Pro Thr Gln His Gly Asp Gly Thr Tyr Pro
 100 105 110
 Thr Glu Lys Lys Leu Thr Gly Ile Gly Glu Asp Leu Lys His Ile Arg
 115 120 125
 Gly Tyr Asp Val Lys Thr Leu Leu Ala Met Val Lys Ser Lys Leu Lys
 130 135 140
 Gly Glu Lys Leu Lys Asp Asp Lys Thr Met Leu Met Glu Arg Val Met

145	150	155	160
Gln Leu Val Ala Arg	Leu Pro Thr Glu Ser	Lys Lys Arg Ala Glu Leu	
	165	170	175
Thr Asp Ser Leu Ile Asn	Glu Leu Trp Glu Ser	Leu Asp His Pro Pro	
	180	185	190
Leu Asn Tyr Leu Gly Pro	Glu His Ser Tyr Arg	Thr Pro Asp Gly Ser	
	195	200	205
Tyr Asn His Pro Phe Asn	Pro Gln Leu Gly Ala	Ala Gly Ser Arg Tyr	
	210	215	220
Ala Arg Ser Val Ile Pro	Thr Val Thr Pro Pro	Gly Ala Leu Pro Asp	
225	230	235	240
Pro Gly Leu Ile Phe Asp	Ser Ile Met Gly Arg	Thr Pro Asn Ser Tyr	
	245	250	255
Arg Lys His Pro Asn Asn	Val Ser Ser Ile Leu	Trp Tyr Trp Ala Thr	
	260	265	270
Ile Ile Ile His Asp Ile	Phe Trp Thr Asp Pro	Arg Asp Ile Asn Thr	
	275	280	285
Asn Lys Ser Ser Ser Tyr	Leu Asp Leu Ala Pro	Leu Tyr Gly Asn Ser	
	290	295	300
Gln Glu Met Gln Asp Ser	Ile Arg Thr Phe Lys	Asp Gly Arg Met Lys	
305	310	315	320
Pro Asp Cys Tyr Ala Asp	Lys Arg Leu Ala Gly	Met Pro Pro Gly Val	
	325	330	335
Ser Val Leu Leu Ile Met	Phe Asn Arg Phe His	Asn His Val Ala Glu	
	340	345	350
Asn Leu Ala Leu Ile Asn	Glu Gly Gly Arg Phe	Asn Lys Pro Ser Asp	
	355	360	365
Leu Leu Glu Gly Glu Ala	Arg Glu Ala Ala Trp	Lys Lys Tyr Asp Asn	
	370	375	380
Asp Leu Phe Gln Val Ala	Arg Leu Val Thr Ser	Gly Leu Tyr Ile Asn	
385	390	395	400
Ile Thr Leu Val Asp Tyr	Val Arg Asn Ile Val	Asn Leu Asn Arg Val	
	405	410	415
Asp Thr Thr Trp Thr Leu	Asp Pro Arg Gln Asp	Ala Gly Ala His Val	
	420	425	430
Gly Thr Ala Asp Gly Ala	Glu Arg Gly Thr Gly	Asn Ala Val Ser Ala	
	435	440	445
Glu Phe Asn Leu Cys Tyr	Arg Trp His Ser Cys	Ile Ser Glu Lys Asp	
450	455	460	

Ser	Lys	Phe	Val	Glu	Ala	Gln	Phe	Gln	Asn	Ile	Phe	Gly	Lys	Pro	Ala
465					470					475					480
Ser	Glu	Val	Arg	Pro	Asp	Glu	Met	Trp	Lys	Gly	Phe	Ala	Lys	Met	Glu
					485					490					495
Gln	Asn	Thr	Pro	Ala	Asp	Pro	Gly	Gln	Arg	Thr	Phe	Gly	Gly	Phe	Lys
					500					505					510
Arg	Gly	Pro	Asp	Gly	Lys	Phe	Asp	Asp	Asp	Asp	Leu	Val	Arg	Cys	Ile
					515					520					525
Ser	Glu	Ala	Val	Glu	Asp	Val	Ala	Gly	Ala	Phe	Gly	Ala	Arg	Asn	Val
					530										540
Pro	Gln	Ala	Met	Lys	Val	Val	Glu	Thr	Met	Gly	Ile	Ile	Gln	Gly	Arg
545					550					555					560
Lys	Trp	Asn	Val	Ala	Gly	Leu	Asn	Glu	Phe	Arg	Lys	His	Phe	His	Leu
					565					570					575
Lys	Pro	Tyr	Ser	Thr	Phe	Glu	Asp	Ile	Asn	Ser	Asp	Pro	Gly	Val	Ala
					580					585					590
Glu	Ala	Leu	Arg	Arg	Leu	Tyr	Asp	His	Pro	Asp	Asn	Val	Glu	Leu	Tyr
					595					600					605
Pro	Gly	Leu	Val	Ala	Glu	Glu	Asp	Lys	Gln	Pro	Met	Val	Pro	Gly	Val
					610					615					620
Gly	Ile	Ala	Pro	Thr	Tyr	Thr	Ile	Ser	Arg	Val	Val	Leu	Ser	Asp	Ala
625					630					635					640
Val	Cys	Leu	Val	Arg	Gly	Asp	Arg	Phe	Tyr	Thr	Thr	Asp	Phe	Thr	Pro
					645					650					655
Arg	Asn	Leu	Thr	Asn	Trp	Gly	Tyr	Lys	Glu	Val	Asp	Tyr	Asp	Leu	Ser
					660					665					670
Val	Asn	His	Gly	Cys	Val	Phe	Tyr	Lys	Leu	Phe	Ile	Arg	Ala	Phe	Pro
					675					680					685
Asn	His	Phe	Lys	Gln	Asn	Ser	Val	Tyr	Ala	His	Tyr	Pro	Met	Val	Val
					690					695					700
Pro	Ser	Glu	Asn	Lys	Arg	Ile	Leu	Glu	Ala	Leu	Gly	Arg	Ala	Asp	Leu
705					710					715					720
Phe	Asp	Phe	Glu	Ala	Pro	Lys	Tyr	Ile	Pro	Pro	Arg	Val	Asn	Ile	Thr
					725					730					735
Ser	Tyr	Gly	Gly	Ala	Glu	Tyr	Ile	Leu	Glu	Thr	Gln	Glu	Lys	Tyr	Lys
					740					745					750
Val	Thr	Trp	His	Glu	Gly	Leu	Gly	Phe	Leu	Met	Gly	Glu	Gly	Gly	Leu
					755					760					765
Lys	Phe	Met	Leu	Ser	Gly	Asp	Asp	Pro	Leu	His	Ala	Gln	Gln	Arg	Lys

770	775	780
Cys Met Ala Ala Gln Leu Tyr Lys Asp Gly Trp Thr Glu Ala Val Lys		
785	790	795
Ala Phe Tyr Ala Gly Met Met Glu Glu Leu Leu Val Ser Lys Ser Tyr		800
805	810	815
Phe Leu Gly Asn Asn Lys His Arg His Val Asp Ile Ile Arg Asp Val		
820	825	830
Gly Asn Met Val His Val His Phe Ala Ser Gln Val Phe Gly Leu Pro		
835	840	845
Leu Lys Thr Ala Lys Asn Pro Thr Gly Val Phe Thr Glu Gln Glu Met		
850	855	860
Tyr Gly Ile Leu Ala Ala Ile Phe Thr Thr Ile Phe Phe Asp Leu Asp		
865	870	875
Pro Ser Lys Ser Phe Pro Leu Arg Thr Lys Thr Arg Glu Val Cys Gln		
885	890	895
Lys Leu Ala Lys Leu Val Glu Ala Asn Val Lys Leu Ile Asn Lys Ile		
900	905	910
Pro Trp Ser Arg Gly Met Phe Val Gly Lys Pro Ala Lys Asp Glu Pro		
915	920	925
Leu Ser Ile Tyr Gly Lys Thr Met Ile Lys Gly Leu Lys Ala His Gly		
930	935	940
Leu Ser Asp Tyr Asp Ile Ala Trp Ser His Val Val Pro Thr Ser Gly		
945	950	955
Ala Met Val Pro Asn Gln Ala Gln Val Phe Ala Gln Ala Val Asp Tyr		
965	970	975
Tyr Leu Ser Pro Ala Gly Met His Tyr Ile Pro Glu Ile His Met Val		
980	985	990
Ala Leu Gln Pro Ser Thr Pro Glu Thr Asp Ala Leu Leu Leu Gly Tyr		
995	1000	1005
Ala Met Glu Gly Ile Arg Leu Ala Gly Thr Phe Gly Ser Tyr Arg		
1010	1015	1020
Glu Ala Ala Val Asp Asp Val Val Lys Glu Asp Asn Gly Arg Gln		
1025	1030	1035
Val Pro Val Lys Ala Gly Asp Arg Val Phe Val Ser Phe Val Asp		
1040	1045	1050
Ala Ala Arg Asp Pro Lys His Phe Pro Asp Pro Glu Val Val Asn		
1055	1060	1065
Pro Arg Arg Pro Ala Lys Lys Tyr Ile His Tyr Gly Val Gly Pro		
1070	1075	1080

His Ala	Cys Leu Gly Arg Asp	Ala Ser Gln Ile Ala	Ile Thr Glu
1085	1090	1095	
Met Phe	Arg Cys Leu Phe Arg	Arg Arg Asn Val Arg	Arg Val Pro
1100	1105	1110	
Gly Pro	Gln Gly Glu Leu Lys	Lys Val Pro Arg Pro	Gly Gly Phe
1115	1120	1125	
Tyr Val	Tyr Met Arg Glu Asp	Trp Gly Gly Leu Phe	Pro Phe Pro
1130	1135	1140	
Val Thr	Met Arg Val Met Trp	Asp Asp Glu	
1145	1150		
<210>	7		
<211>	530		
<212>	PRT		
<213>	尖孢镰刀菌 (Fusarium oxysporum)		
<400>	7		
Met Lys Gly Ser Ala Thr Leu Ala Phe Ala Leu Val Gln Phe Ser Ala			
1	5	10	15
Ala Ser Gln Leu Val Trp Pro Ser Lys Trp Asp Glu Val Glu Asp Leu			
	20	25	30
Leu Tyr Met Gln Gly Gly Phe Asn Lys Arg Gly Phe Ala Asp Ala Leu			
	35	40	45
Arg Thr Cys Glu Phe Gly Ser Asn Val Pro Gly Thr Gln Asn Thr Ala			
	50	55	60
Glu Trp Leu Arg Thr Ala Phe His Asp Ala Ile Thr His Asp Ala Lys			
65	70	75	80
Ala Gly Thr Gly Gly Leu Asp Ala Ser Ile Tyr Trp Glu Ser Ser Arg			
	85	90	95
Pro Glu Asn Pro Gly Lys Ala Phe Asn Asn Thr Phe Gly Phe Phe Ser			
	100	105	110
Gly Phe His Asn Pro Arg Ala Thr Ala Ser Asp Leu Thr Ala Leu Gly			
	115	120	125
Thr Val Leu Ala Val Gly Ala Cys Asn Gly Pro Arg Ile Pro Phe Arg			
	130	135	140
Ala Gly Arg Ile Asp Ala Tyr Lys Ala Gly Pro Ala Gly Val Pro Glu			
145	150	155	160
Pro Ser Thr Asn Leu Lys Asp Thr Phe Ala Ala Phe Thr Lys Ala Gly			
	165	170	175
Phe Thr Lys Glu Glu Met Thr Ala Met Val Ala Cys Gly His Ala Ile			
	180	185	190

Gly Gly Val His Ser Val Asp Phe Pro Glu Ile Val Gly Ile Lys Ala		
195	200	205
Asp Pro Asn Asn Asp Thr Asn Val Pro Phe Gln Lys Asp Val Ser Ser		
210	215	220
Phe His Asn Gly Ile Val Thr Glu Tyr Leu Ala Gly Thr Ser Lys Asn		
225	230	235
Pro Leu Val Ala Ser Lys Asn Ala Thr Phe His Ser Asp Lys Arg Ile		
	245	250
Phe Asp Asn Asp Lys Ala Thr Met Lys Lys Leu Ser Thr Lys Ala Gly		
	260	265
Phe Asn Ser Met Cys Ala Asp Ile Leu Thr Arg Met Ile Asp Thr Val		
	275	280
Pro Lys Ser Val Gln Leu Thr Pro Val Leu Glu Ala Tyr Asp Val Arg		
290	295	300
Pro Tyr Ile Thr Glu Leu Ser Leu Asn Asn Lys Asn Lys Ile His Phe		
305	310	315
Thr Gly Ser Val Arg Val Arg Ile Thr Asn Asn Ile Arg Asp Asn Asn		
	325	330
Asp Leu Ala Ile Asn Leu Ile Tyr Val Gly Arg Asp Gly Lys Lys Val		
	340	345
Thr Val Pro Thr Gln Gln Val Thr Phe Gln Gly Gly Thr Ser Phe Gly		
	355	360
Ala Gly Glu Val Phe Ala Asn Phe Glu Phe Asp Thr Thr Met Asp Ala		
370	375	380
Lys Asn Gly Ile Thr Lys Phe Phe Ile Gln Glu Val Lys Pro Ser Thr		
385	390	395
Lys Ala Thr Val Thr His Asp Asn Gln Lys Thr Gly Gly Tyr Lys Val		
	405	410
Asp Asp Thr Val Leu Tyr Gln Leu Gln Gln Ser Cys Ala Val Leu Glu		
	420	425
Lys Leu Pro Asn Ala Pro Leu Val Val Thr Ala Met Val Arg Asp Ala		
	435	440
Arg Ala Lys Asp Ala Leu Thr Leu Arg Val Ala His Lys Lys Pro Val		
450	455	460
Lys Gly Ser Ile Val Pro Arg Phe Gln Thr Ala Ile Thr Asn Phe Lys		
465	470	475
Ala Thr Gly Lys Lys Ser Ser Gly Tyr Thr Gly Phe Gln Ala Lys Thr		
	485	490
Met Phe Glu Glu Gln Ser Thr Tyr Phe Asp Ile Val Leu Gly Gly Ser		

500	505	510
Pro Ala Ser Gly Val Gln Phe Leu Thr Ser Gln Ala Met Pro Ser Gln		
515	520	525
Cys Ser		
530		
<210> 8		
<211> 358		
<212> PRT		
<213> 禾谷镰刀菌 (<i>Fusarium graminearum</i>)		
<400> 8		
Met Ala Ser Ala Thr Arg Gln Phe Ala Arg Ala Ala Thr Arg Ala Thr		
1 5 10 15		
Arg Asn Gly Phe Ala Ile Ala Pro Arg Gln Val Ile Arg Gln Gln Gly		
20 25 30		
Arg Arg Tyr Tyr Ser Ser Glu Pro Ala Gln Lys Ser Ser Ser Ala Trp		
35 40 45		
Ile Trp Leu Thr Gly Ala Ala Val Ala Gly Gly Ala Gly Tyr Tyr Phe		
50 55 60		
Tyr Gly Asn Ser Ala Ser Ser Ala Thr Ala Lys Val Phe Asn Pro Ser		
65 70 75 80		
Lys Glu Asp Tyr Gln Lys Val Tyr Asn Glu Ile Ala Ala Arg Leu Glu		
85 90 95		
Glu Lys Asp Asp Tyr Asp Asp Gly Ser Tyr Gly Pro Val Leu Val Arg		
100 105 110		
Leu Ala Trp His Ala Ser Gly Thr Tyr Asp Lys Glu Thr Gly Thr Gly		
115 120 125		
Gly Ser Asn Gly Ala Thr Met Arg Phe Ala Pro Glu Ser Asp His Gly		
130 135 140		
Ala Asn Ala Gly Leu Ala Ala Ala Arg Asp Phe Leu Gln Pro Val Lys		
145 150 155 160		
Glu Lys Phe Pro Trp Ile Thr Tyr Ser Asp Leu Trp Ile Leu Ala Gly		
165 170 175		
Val Cys Ala Ile Gln Glu Met Leu Gly Pro Ala Ile Pro Tyr Arg Pro		
180 185 190		
Gly Arg Ser Asp Arg Asp Val Ser Gly Cys Thr Pro Asp Gly Arg Leu		
195 200 205		
Pro Asp Ala Ser Lys Arg Gln Asp His Leu Arg Gly Ile Phe Gly Arg		
210 215 220		
Met Gly Phe Asn Asp Gln Glu Ile Val Ala Leu Ser Gly Ala His Ala		

225	230	235	240
Leu Gly Arg Cys His Thr Asp Arg Ser Gly Tyr Ser Gly Pro Trp Thr			
	245	250	255
Phe Ser Pro Thr Val Leu Thr Asn Asp Tyr Phe Arg Leu Leu Val Glu			
	260	265	270
Glu Lys Trp Gln Trp Lys Lys Trp Asn Gly Pro Ala Gln Tyr Glu Asp			
	275	280	285
Lys Ser Thr Lys Ser Leu Met Met Leu Pro Ser Asp Ile Ala Leu Ile			
	290	295	300
Glu Asp Lys Lys Phe Lys Pro Trp Val Glu Lys Tyr Ala Lys Asp Asn			
305	310	315	320
Asp Ala Phe Phe Lys Asp Phe Ser Asn Val Val Leu Arg Leu Phe Glu			
	325	330	335
Leu Gly Val Pro Phe Ala Gln Gly Thr Glu Asn Gln Arg Trp Thr Phe			
	340	345	350
Lys Pro Thr His Gln Glu			
	355		
<210> 9			
<211> 122			
<212> PRT			
<213> 卵配衣藻 (Chlamydomonas eugametos)			
<400> 9			
Met Ser Leu Phe Ala Lys Leu Gly Gly Arg Glu Ala Val Glu Ala Ala			
1	5	10	15
Val Asp Lys Phe Tyr Asn Lys Ile Val Ala Asp Pro Thr Val Ser Thr			
	20	25	30
Tyr Phe Ser Asn Thr Asp Met Lys Val Gln Arg Ser Lys Gln Phe Ala			
	35	40	45
Phe Leu Ala Tyr Ala Leu Gly Gly Ala Ser Glu Trp Lys Gly Lys Asp			
	50	55	60
Met Arg Thr Ala His Lys Asp Leu Val Pro His Leu Ser Asp Val His			
65	70	75	80
Phe Gln Ala Val Ala Arg His Leu Ser Asp Thr Leu Thr Glu Leu Gly			
	85	90	95
Val Pro Pro Glu Asp Ile Thr Asp Ala Met Ala Val Val Ala Ser Thr			
	100	105	110
Arg Thr Glu Val Leu Asn Met Pro Gln Gln			
	115	120	
<210> 10			

<211> 121

<212> PRT

<213> 梨形四膜虫 (*Tetrahymena pyriformis*)

<400> 10

```

Met Asn Lys Pro Gln Thr Ile Tyr Glu Lys Leu Gly Gly Glu Asn Ala
1           5           10           15
Met Lys Ala Ala Val Pro Leu Phe Tyr Lys Lys Val Leu Ala Asp Glu
           20           25           30
Arg Val Lys His Phe Phe Lys Asn Thr Asp Met Asp His Gln Thr Lys
           35           40           45
Gln Gln Thr Asp Phe Leu Thr Met Leu Leu Gly Gly Pro Asn His Tyr
           50           55           60
Lys Gly Lys Asn Met Thr Glu Ala His Lys Gly Met Asn Leu Gln Asn
65           70           75           80
Leu His Phe Asp Ala Ile Ile Glu Asn Leu Ala Ala Thr Leu Lys Glu
           85           90           95
Leu Gly Val Thr Asp Ala Val Ile Asn Glu Ala Ala Lys Val Ile Glu
           100          105          110
His Thr Arg Lys Asp Met Leu Gly Lys
           115          120

```

<210> 11

<211> 117

<212> PRT

<213> 大草履虫 (*Paramecium caudatum*)

<400> 11

```

Met Ser Leu Phe Glu Gln Leu Gly Gly Gln Ala Ala Val Gln Ala Val
1           5           10           15
Thr Ala Gln Phe Tyr Ala Asn Ile Gln Ala Asp Ala Thr Val Ala Thr
           20           25           30
Phe Phe Asn Gly Ile Asp Met Pro Asn Gln Thr Asn Lys Thr Ala Ala
           35           40           45
Phe Leu Cys Ala Ala Leu Gly Gly Pro Asn Ala Trp Thr Gly Arg Asn
50           55           60
Leu Lys Glu Val His Ala Asn Met Gly Val Ser Asn Ala Gln Phe Thr
65           70           75           80
Thr Val Ile Gly His Leu Arg Ser Ala Leu Thr Gly Ala Gly Val Ala
           85           90           95
Ala Ala Leu Val Glu Gln Thr Val Ala Val Ala Glu Thr Val Arg Gly
           100          105          110

```

Asp Val Val Thr Val

115

<210> 12

<211> 147

<212> PRT

<213> 黑曲霉 (*Aspergillus niger*)

<400> 12

Met Pro Leu Thr Pro Glu Gln Ile Lys Ile Ile Lys Ala Thr Val Pro

1 5 10 15

Val Leu Gln Glu Tyr Gly Thr Lys Ile Thr Thr Ala Phe Tyr Met Asn

20 25 30

Met Ser Thr Val His Pro Glu Leu Asn Ala Val Phe Asn Thr Ala Asn

35 40 45

Gln Val Lys Gly His Gln Ala Arg Ala Leu Ala Gly Ala Leu Phe Ala

50 55 60

Tyr Ala Ser His Ile Asp Asp Leu Gly Ala Leu Gly Pro Ala Val Glu

65 70 75 80

Leu Ile Cys Asn Lys His Ala Ser Leu Tyr Ile Gln Ala Asp Glu Tyr

85 90 95

Lys Ile Val Gly Lys Tyr Leu Leu Glu Ala Met Lys Glu Val Leu Gly

100 105 110

Asp Ala Cys Thr Asp Asp Ile Leu Asp Ala Trp Gly Ala Ala Tyr Trp

115 120 125

Ala Leu Ala Asp Ile Met Ile Asn Arg Glu Ala Ala Leu Tyr Lys Gln

130 135 140

Ser Gln Gly

145

<210> 13

<211> 165

<212> PRT

<213> 玉米 (*Zea mays*)

<400> 13

Met Ala Leu Ala Glu Ala Asp Asp Gly Ala Val Val Phe Gly Glu Glu

1 5 10 15

Gln Glu Ala Leu Val Leu Lys Ser Trp Ala Val Met Lys Lys Asp Ala

20 25 30

Ala Asn Leu Gly Leu Arg Phe Phe Leu Lys Val Phe Glu Ile Ala Pro

35 40 45

Ser Ala Glu Gln Met Phe Ser Phe Leu Arg Asp Ser Asp Val Pro Leu

50	55	60
Glu Lys Asn Pro Lys Leu Lys Thr His Ala Met Ser Val Phe Val Met		
65	70	75
Thr Cys Glu Ala Ala Ala Gln Leu Arg Lys Ala Gly Lys Val Thr Val		
85	90	95
Arg Glu Thr Thr Leu Lys Arg Leu Gly Ala Thr His Leu Arg Tyr Gly		
100	105	110
Val Ala Asp Gly His Phe Glu Val Thr Gly Phe Ala Leu Leu Glu Thr		
115	120	125
Ile Lys Glu Ala Leu Pro Ala Asp Met Trp Ser Leu Glu Met Lys Lys		
130	135	140
Ala Trp Ala Glu Ala Tyr Ser Gln Leu Val Ala Ala Ile Lys Arg Glu		
145	150	155
Met Lys Pro Asp Ala		
165		
<210> 14		
<211> 169		
<212> PRT		
<213> 粳稻(<i>Oryza sativa</i> subsp. <i>japonica</i>)		
<400> 14		
Met Ala Leu Val Glu Gly Asn Asn Gly Val Ser Gly Gly Ala Val Ser		
1	5	10
Phe Ser Glu Glu Gln Glu Ala Leu Val Leu Lys Ser Trp Ala Ile Met		
20	25	30
Lys Lys Asp Ser Ala Asn Ile Gly Leu Arg Phe Phe Leu Lys Ile Phe		
35	40	45
Glu Val Ala Pro Ser Ala Ser Gln Met Phe Ser Phe Leu Arg Asn Ser		
50	55	60
Asp Val Pro Leu Glu Lys Asn Pro Lys Leu Lys Thr His Ala Met Ser		
65	70	75
Val Phe Val Met Thr Cys Glu Ala Ala Ala Gln Leu Arg Lys Ala Gly		
85	90	95
Lys Val Thr Val Arg Asp Thr Thr Leu Lys Arg Leu Gly Ala Thr His		
100	105	110
Phe Lys Tyr Gly Val Gly Asp Ala His Phe Glu Val Thr Arg Phe Ala		
115	120	125
Leu Leu Glu Thr Ile Lys Glu Ala Val Pro Val Asp Met Trp Ser Pro		
130	135	140
Ala Met Lys Ser Ala Trp Ser Glu Ala Tyr Asn Gln Leu Val Ala Ala		

145	150	155	160
Ile Lys Gln Glu Met Lys Pro Ala Glu			
	165		
<210> 15			
<211> 160			
<212> PRT			
<213> 拟南芥(<i>Arabidopsis thaliana</i>)			
<400> 15			
Met Glu Ser Glu Gly Lys Ile Val Phe Thr Glu Glu Gln Glu Ala Leu			
1	5	10	15
Val Val Lys Ser Trp Ser Val Met Lys Lys Asn Ser Ala Glu Leu Gly			
	20	25	30
Leu Lys Leu Phe Ile Lys Ile Phe Glu Ile Ala Pro Thr Thr Lys Lys			
	35	40	45
Met Phe Ser Phe Leu Arg Asp Ser Pro Ile Pro Ala Glu Gln Asn Pro			
50	55	60	
Lys Leu Lys Pro His Ala Met Ser Val Phe Val Met Cys Cys Glu Ser			
65	70	75	80
Ala Val Gln Leu Arg Lys Thr Gly Lys Val Thr Val Arg Glu Thr Thr			
	85	90	95
Leu Lys Arg Leu Gly Ala Ser His Ser Lys Tyr Gly Val Val Asp Glu			
	100	105	110
His Phe Glu Val Ala Lys Tyr Ala Leu Leu Glu Thr Ile Lys Glu Ala			
	115	120	125
Val Pro Glu Met Trp Ser Pro Glu Met Lys Val Ala Trp Gly Gln Ala			
	130	135	140
Tyr Asp His Leu Val Ala Ala Ile Lys Ala Glu Met Asn Leu Ser Asn			
145	150	155	160
<210> 16			
<211> 147			
<212> PRT			
<213> 豌豆(<i>Pisum sativum</i>)			
<400> 16			
Met Gly Phe Thr Asp Lys Gln Glu Ala Leu Val Asn Ser Ser Trp Glu			
1	5	10	15
Ser Phe Lys Gln Asn Leu Ser Gly Asn Ser Ile Leu Phe Tyr Thr Ile			
	20	25	30
Ile Leu Glu Lys Ala Pro Ala Ala Lys Gly Leu Phe Ser Phe Leu Lys			
	35	40	45

Asp Thr Ala Gly Val Glu Asp Ser Pro Lys Leu Gln Ala His Ala Glu
 50 55 60
 Gln Val Phe Gly Leu Val Arg Asp Ser Ala Ala Gln Leu Arg Thr Lys
 65 70 75 80
 Gly Glu Val Val Leu Gly Asn Ala Thr Leu Gly Ala Ile His Val Gln
 85 90 95
 Arg Gly Val Thr Asp Pro His Phe Val Val Val Lys Glu Ala Leu Leu
 100 105 110
 Gln Thr Ile Lys Lys Ala Ser Gly Asn Asn Trp Ser Glu Glu Leu Asn
 115 120 125
 Thr Ala Trp Glu Val Ala Tyr Asp Gly Leu Ala Thr Ala Ile Lys Lys
 130 135 140
 Ala Met Thr
 145
 <210> 17
 <211> 145
 <212> PRT
 <213> 豇豆 (*Vigna unguiculata*)
 <400> 17
 Met Val Ala Phe Ser Asp Lys Gln Glu Ala Leu Val Asn Gly Ala Tyr
 1 5 10 15
 Glu Ala Phe Lys Ala Asn Ile Pro Lys Tyr Ser Val Val Phe Tyr Thr
 20 25 30
 Thr Ile Leu Glu Lys Ala Pro Ala Ala Lys Asn Leu Phe Ser Phe Leu
 35 40 45
 Ala Asn Gly Val Asp Ala Thr Asn Pro Lys Leu Thr Gly His Ala Glu
 50 55 60
 Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Ala Gln Leu Arg Ala Ser
 65 70 75 80
 Gly Gly Val Val Ala Asp Ala Ala Leu Gly Ala Val His Ser Gln Lys
 85 90 95
 Ala Val Asn Asp Ala Gln Phe Val Val Val Lys Glu Ala Leu Val Lys
 100 105 110
 Thr Leu Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Gly Thr
 115 120 125
 Ala Val Glu Leu Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala
 130 135 140
 Tyr
 145

<210> 18

<211> 154

<212> PRT

<213> 黄牛 (Bos taurus)

<400> 18

```

Met Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Ala Trp Gly
1           5           10           15
Lys Val Glu Ala Asp Val Ala Gly His Gly Gln Glu Val Leu Ile Arg
           20           25           30
Leu Phe Thr Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys
           35           40           45
His Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys
           50           55           60
His Gly Asn Thr Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys
65           70           75           80
Gly His His Glu Ala Glu Val Lys His Leu Ala Glu Ser His Ala Asn
           85           90           95
Lys His Lys Ile Pro Val Lys Tyr Leu Glu Phe Ile Ser Asp Ala Ile
           100          105          110
Ile His Val Leu His Ala Lys His Pro Ser Asp Phe Gly Ala Asp Ala
           115          120          125
Gln Ala Ala Met Ser Lys Ala Leu Glu Leu Phe Arg Asn Asp Met Ala
           130          135          140
Ala Gln Tyr Lys Val Leu Gly Phe His Gly
145           150

```

<210> 19

<211> 154

<212> PRT

<213> 欧亚野猪 (Sus scrofa)

<400> 19

```

Met Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Val Trp Gly
1           5           10           15
Lys Val Glu Ala Asp Val Ala Gly His Gly Gln Glu Val Leu Ile Arg
           20           25           30
Leu Phe Lys Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys
           35           40           45
His Leu Lys Ser Glu Asp Glu Met Lys Ala Ser Glu Asp Leu Lys Lys
           50           55           60
His Gly Asn Thr Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys

```

65	70	75	80
Gly His His Glu Ala Glu Leu Thr Pro Leu Ala Gln Ser His Ala Thr			
	85	90	95
Lys His Lys Ile Pro Val Lys Tyr Leu Glu Phe Ile Ser Glu Ala Ile			
	100	105	110
Ile Gln Val Leu Gln Ser Lys His Pro Gly Asp Phe Gly Ala Asp Ala			
	115	120	125
Gln Gly Ala Met Ser Lys Ala Leu Glu Leu Phe Arg Asn Asp Met Ala			
	130	135	140
Ala Lys Tyr Lys Glu Leu Gly Phe Gln Gly			
145	150		
<210>	20		
<211>	154		
<212>	PRT		
<213>	马 (Equus caballus)		
<400>	20		
Met Gly Leu Ser Asp Gly Glu Trp Gln Gln Val Leu Asn Val Trp Gly			
1	5	10	15
Lys Val Glu Ala Asp Ile Ala Gly His Gly Gln Glu Val Leu Ile Arg			
	20	25	30
Leu Phe Thr Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys			
	35	40	45
His Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys			
	50	55	60
His Gly Thr Val Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys			
65	70	75	80
Gly His His Glu Ala Glu Leu Lys Pro Leu Ala Gln Ser His Ala Thr			
	85	90	95
Lys His Lys Ile Pro Ile Lys Tyr Leu Glu Phe Ile Ser Asp Ala Ile			
	100	105	110
Ile His Val Leu His Ser Lys His Pro Gly Asp Phe Gly Ala Asp Ala			
	115	120	125
Gln Gly Ala Met Thr Lys Ala Leu Glu Leu Phe Arg Asn Asp Ile Ala			
	130	135	140
Ala Lys Tyr Lys Glu Leu Gly Phe Gln Gly			
145	150		
<210>	21		
<211>	152		
<212>	PRT		

<213> 本氏烟草 (*Nicotiana benthamiana*)

<400> 21

```

Met Ser Ser Phe Thr Glu Glu Gln Glu Ala Leu Val Val Lys Ser Trp
1           5           10           15
Asp Ser Met Lys Lys Asn Ala Gly Glu Trp Gly Leu Lys Leu Phe Leu
          20           25           30
Lys Ile Phe Glu Ile Ala Pro Ser Ala Lys Lys Leu Phe Ser Phe Leu
          35           40           45
Lys Asp Ser Asn Val Pro Leu Glu Gln Asn Ala Lys Leu Lys Pro His
          50           55           60
Ser Lys Ser Val Phe Val Met Thr Cys Glu Ala Ala Val Gln Leu Arg
65           70           75           80
Lys Ala Gly Lys Val Val Val Arg Asp Ser Thr Leu Lys Lys Leu Gly
          85           90           95
Ala Thr His Phe Lys Tyr Gly Val Ala Asp Glu His Phe Glu Val Thr
          100          105          110
Lys Phe Ala Leu Leu Glu Thr Ile Lys Glu Ala Val Pro Glu Met Trp
          115          120          125
Ser Val Asp Met Lys Asn Ala Trp Gly Glu Ala Phe Asp Gln Leu Val
          130          135          140
Asn Ala Ile Lys Thr Glu Met Lys
145          150

```

<210> 22

<211> 132

<212> PRT

<213> 枯草芽孢杆菌 (*Bacillus subtilis*)

<400> 22

```

Met Gly Gln Ser Phe Asn Ala Pro Tyr Glu Ala Ile Gly Glu Glu Leu
1           5           10           15
Leu Ser Gln Leu Val Asp Thr Phe Tyr Glu Arg Val Ala Ser His Pro
          20           25           30
Leu Leu Lys Pro Ile Phe Pro Ser Asp Leu Thr Glu Thr Ala Arg Lys
          35           40           45
Gln Lys Gln Phe Leu Thr Gln Tyr Leu Gly Gly Pro Pro Leu Tyr Thr
          50           55           60
Glu Glu His Gly His Pro Met Leu Arg Ala Arg His Leu Pro Phe Pro
65           70           75           80
Ile Thr Asn Glu Arg Ala Asp Ala Trp Leu Ser Cys Met Lys Asp Ala
          85           90           95

```

Met Asp His Val Gly Leu Glu Gly Glu Ile Arg Glu Phe Leu Phe Gly
 100 105 110

Arg Leu Glu Leu Thr Ala Arg His Met Val Asn Gln Thr Glu Ala Glu
 115 120 125

Asp Arg Ser Ser
 130

<210> 23

<211> 131

<212> PRT

<213> 谷氨酸棒状杆菌 (*Corynebacterium glutamicum*)

<400> 23

Met Thr Thr Ser Glu Asn Phe Tyr Asp Ser Val Gly Gly Glu Glu Thr
 1 5 10 15

Phe Ser Leu Ile Val His Arg Phe Tyr Glu Gln Val Pro Asn Asp Asp
 20 25 30

Ile Leu Gly Pro Met Tyr Pro Pro Asp Asp Phe Glu Gly Ala Glu Gln
 35 40 45

Arg Leu Lys Met Phe Leu Ser Gln Tyr Trp Gly Gly Pro Lys Asp Tyr
 50 55 60

Gln Glu Gln Arg Gly His Pro Arg Leu Arg Met Arg His Val Asn Tyr
 65 70 75 80

Pro Ile Gly Val Thr Ala Ala Glu Arg Trp Leu Gln Leu Met Ser Asn
 85 90 95

Ala Leu Asp Gly Val Asp Leu Thr Ala Glu Gln Arg Glu Ala Ile Trp
 100 105 110

Glu His Met Val Arg Ala Ala Asp Met Leu Ile Asn Ser Asn Pro Asp
 115 120 125

Pro His Ala
 130

<210> 24

<211> 124

<212> PRT

<213> 集胞藻属 (*Synechocystis* sp.)

<400> 24

Met Ser Thr Leu Tyr Glu Lys Leu Gly Gly Thr Thr Ala Val Asp Leu
 1 5 10 15

Ala Val Asp Lys Phe Tyr Glu Arg Val Leu Gln Asp Asp Arg Ile Lys
 20 25 30

His Phe Phe Ala Asp Val Asp Met Ala Lys Gln Arg Ala His Gln Lys

35					40					45						
Ala	Phe	Leu	Thr	Tyr	Ala	Phe	Gly	Gly	Thr	Asp	Lys	Tyr	Asp	Gly	Arg	
50					55					60						
Tyr	Met	Arg	Glu	Ala	His	Lys	Glu	Leu	Val	Glu	Asn	His	Gly	Leu	Asn	
65					70					75					80	
Gly	Glu	His	Phe	Asp	Ala	Val	Ala	Glu	Asp	Leu	Leu	Ala	Thr	Leu	Lys	
85					90					95						
Glu	Met	Gly	Val	Pro	Glu	Asp	Leu	Ile	Ala	Glu	Val	Ala	Ala	Val	Ala	
100					105					110						
Gly	Ala	Pro	Ala	His	Lys	Arg	Asp	Val	Leu	Asn	Gln					
115					120											
<210>	25															
<211>	183															
<212>	PRT															
<213>	聚球藻属 (Synechococcus sp.)															
<400>	25															
Met	Asp	Val	Ala	Leu	Leu	Glu	Lys	Ser	Phe	Glu	Gln	Ile	Ser	Pro	Arg	
1	5				10					15						
Ala	Ile	Glu	Phe	Ser	Ala	Ser	Phe	Tyr	Gln	Asn	Leu	Phe	His	His	His	
20				25					30							
Pro	Glu	Leu	Lys	Pro	Leu	Phe	Ala	Glu	Thr	Ser	Gln	Thr	Ile	Gln	Glu	
35				40					45							
Lys	Lys	Leu	Ile	Phe	Ser	Leu	Ala	Ala	Ile	Ile	Glu	Asn	Leu	Arg	Asn	
50					55					60						
Pro	Asp	Ile	Leu	Gln	Pro	Ala	Leu	Lys	Ser	Leu	Gly	Ala	Arg	His	Ala	
65					70					75					80	
Glu	Val	Gly	Thr	Ile	Lys	Ser	His	Tyr	Pro	Leu	Val	Gly	Gln	Ala	Leu	
85				90					95							
Ile	Glu	Thr	Phe	Ala	Glu	Tyr	Leu	Ala	Ala	Asp	Trp	Thr	Glu	Gln	Leu	
100					105					110						
Ala	Thr	Ala	Trp	Val	Glu	Ala	Tyr	Asp	Val	Ile	Ala	Ser	Thr	Met	Ile	
115					120					125						
Glu	Gly	Ala	Asp	Asn	Pro	Ala	Ala	Tyr	Leu	Glu	Pro	Glu	Leu	Thr	Phe	
130					135					140						
Tyr	Glu	Trp	Leu	Asp	Leu	Tyr	Gly	Glu	Glu	Ser	Pro	Lys	Val	Arg	Asn	
145					150					155					160	
Ala	Ile	Ala	Thr	Leu	Thr	His	Phe	His	Tyr	Gly	Glu	Asp	Pro	Gln	Asp	
165				170					175							
Val	Gln	Arg	Asp	Ser	Arg	Gly										

180

<210> 26

<211> 118

<212> PRT

<213> 地木耳 (*Nostoc commune*)

<400> 26

```

Met Ser Thr Leu Tyr Asp Asn Ile Gly Gly Gln Pro Ala Ile Glu Gln
1           5           10           15
Val Val Asp Glu Leu His Lys Arg Ile Ala Thr Asp Ser Leu Leu Ala
          20           25           30
Pro Val Phe Ala Gly Thr Asp Met Val Lys Gln Arg Asn His Leu Val
          35           40           45
Ala Phe Leu Ala Gln Ile Phe Glu Gly Pro Lys Gln Tyr Gly Gly Arg
          50           55           60
Pro Met Asp Lys Thr His Ala Gly Leu Asn Leu Gln Gln Pro His Phe
65           70           75           80
Asp Ala Ile Ala Lys His Leu Gly Glu Arg Met Ala Val Arg Gly Val
          85           90           95
Ser Ala Glu Asn Thr Lys Ala Ala Leu Asp Arg Val Thr Asn Met Lys
          100          105          110
Gly Ala Ile Leu Asn Lys

```

115

<210> 27

<211> 136

<212> PRT

<213> 巨大芽孢杆菌 (*Bacillus megaterium*)

<400> 27

```

Met Arg Glu Lys Ile His Ser Pro Tyr Glu Leu Leu Gly Gly Glu His
1           5           10           15
Thr Ile Ser Lys Leu Val Asp Ala Phe Tyr Thr Arg Val Gly Gln His
          20           25           30
Pro Glu Leu Ala Pro Ile Phe Pro Asp Asn Leu Thr Glu Thr Ala Arg
          35           40           45
Lys Gln Lys Gln Phe Leu Thr Gln Tyr Leu Gly Gly Pro Ser Leu Tyr
          50           55           60
Thr Glu Glu His Gly His Pro Met Leu Arg Ala Arg His Leu Pro Phe
65           70           75           80
Glu Ile Thr Pro Ser Arg Ala Lys Ala Trp Leu Thr Cys Met His Glu
          85           90           95

```

Ala Met Asp Glu Ile Asn Leu Glu Gly Pro Glu Arg Asp Glu Leu Tyr
 100 105 110
 His Arg Leu Ile Leu Thr Ala Gln His Met Ile Asn Ser Pro Glu Gln
 115 120 125
 Thr Asp Glu Lys Gly Phe Ser His
 130 135

<210> 28

<211> 1683

<212> DNA

<213> 巴斯德必赤酵母 (*Pichia pastoris*)

<400> 28

atggagtttg	tcgcccgtca	gtccatgaat	gcctgtccct	ttgtcaggtc	aacttctacc	60
caccatttga	agaagttggc	agcaaacagt	tctctagctg	ctactgctag	tcattgtccc	120
gtggttggcc	ctgctctcca	acagcagaga	tactactctc	aaccttccaa	gccagcccaa	180
gcccaaacct	ccgacattgc	tactgggatc	aagaaggatg	tttctccgat	ccgtatggac	240
tctaataaaa	ccgcctttga	ttacaatgga	atgtatgagt	ctgatcttgc	gaataaacgt	300
aaagataact	cgtatcgtta	tttcaataac	atcaaccgtc	tagccaagga	gtttcccaag	360
gcacatcgcc	agaccgaaga	tgacaagggtg	accgtctggt	gctctaacga	ctacttagga	420
atgggtaggc	atcctgagat	tatcaaaacc	atgaaggcta	ccatggacaa	gtacggttcc	480
ggagcaggag	gaactaggaa	cattgcaggt	cataaccacg	ccgctatcaa	tttggaagc	540
gagttggctt	gcttgaacaa	gaaggaagcg	gctctggtgt	tttcatcatg	tttcatagct	600
aacgatgcaa	tcatctcggt	gttgggacaa	aaaatcaaaa	atttggtcat	tttctctgac	660
cagtcgaatc	atgcttccat	gatattgggt	gtgcgtaact	ccaaagcgaa	gaagcacatc	720
ttcaagcaca	acaatttgaa	ggatctggag	tcgcagttag	ctcagtaccc	caagtcgact	780
cctaaactga	tcgccttcga	gtcagtttac	tctatgtgtg	gatctgtggc	tccattgag	840
aagatttgcg	atttggctaa	aaggtaggt	gccctcacct	tcttgatga	agttcatgct	900
gttggaaatg	atggtcctca	tggacagggt	gtagctgagc	atttggactt	tgatctgcat	960
ttacagtctg	gaatcgccag	tcctagcgtg	gtggacaaac	gcaccatatt	ggatcgtgtc	1020
gacatgatta	ctggtacttg	cggaaagtca	tttggtagctg	ttggagggtta	cgttgctggt	1080
agtgccaaacc	taattgattg	gttaagatcc	tatgcgccag	gtttcatttt	cactaccaca	1140
cttcctcctg	ctatcatggc	tggtacagcc	acttctgttc	gtattgttag	ggccgacatt	1200
gaggccccgt	tcaagcaaca	gcttaatact	cgctacgtca	aagactcatt	tgaaaacctt	1260
ggatttccag	tcatttccaaa	cccaagtcac	attgttctctg	ttctagttgg	aaatgctgca	1320
gatgccaaaga	aggcatccga	tatgttaatg	aacaaacacc	gtatttatgt	tcaagctatt	1380
aactacccta	ctgtgcctgt	cggatgaagaa	cgactaagga	ttactcctac	tccaggtcat	1440
ggaaaggaga	tttgtgacca	gctgatcagc	gctgtcgacg	atgtttttac	tgagcttaat	1500
ttaccaagaa	tcaacaaatg	gcagtcccaa	ggtggtcatt	gcggtgttgg	tgatgctaata	1560
tacgtaccag	aaccaaatct	gtggactcag	gaccagctca	gcttgacaaa	ccaagacttg	1620
cactccaatg	tgcacaaccc	agtgattgag	cagatcgaaa	cctcatcagg	agtcagattg	1680

tag	1683
<210> 29	
<211> 560	
<212> PRT	
<213> 巴斯德必赤酵母(Pichia pastoris)	
<400> 29	
Met Glu Phe Val Ala Arg Gln Ser Met Asn Ala Cys Pro Phe Val Arg	
1 5 10 15	
Ser Thr Ser Thr His His Leu Lys Lys Leu Ala Ala Asn Ser Ser Leu	
20 25 30	
Ala Ala Thr Ala Ser His Cys Pro Val Val Gly Pro Ala Leu Gln Gln	
35 40 45	
Gln Arg Tyr Tyr Ser Gln Pro Ser Lys Pro Ala Gln Ala Gln Thr Ser	
50 55 60	
Asp Ile Ala Thr Gly Ile Lys Lys Asp Val Ser Pro Ile Arg Met Asp	
65 70 75 80	
Ser Asn Glu Thr Ala Phe Asp Tyr Asn Gly Met Tyr Glu Ser Asp Leu	
85 90 95	
Ala Asn Lys Arg Lys Asp Asn Ser Tyr Arg Tyr Phe Asn Asn Ile Asn	
100 105 110	
Arg Leu Ala Lys Glu Phe Pro Lys Ala His Arg Gln Thr Glu Asp Asp	
115 120 125	
Lys Val Thr Val Trp Cys Ser Asn Asp Tyr Leu Gly Met Gly Arg His	
130 135 140	
Pro Glu Ile Ile Lys Thr Met Lys Ala Thr Met Asp Lys Tyr Gly Ser	
145 150 155 160	
Gly Ala Gly Gly Thr Arg Asn Ile Ala Gly His Asn His Ala Ala Ile	
165 170 175	
Asn Leu Glu Ser Glu Leu Ala Cys Leu Asn Lys Lys Glu Ala Ala Leu	
180 185 190	
Val Phe Ser Ser Cys Phe Ile Ala Asn Asp Ala Ile Ile Ser Leu Leu	
195 200 205	
Gly Gln Lys Ile Lys Asn Leu Val Ile Phe Ser Asp Gln Ser Asn His	
210 215 220	
Ala Ser Met Ile Leu Gly Val Arg Asn Ser Lys Ala Lys Lys His Ile	
225 230 235 240	
Phe Lys His Asn Asn Leu Lys Asp Leu Glu Ser Gln Leu Ala Gln Tyr	
245 250 255	
Pro Lys Ser Thr Pro Lys Leu Ile Ala Phe Glu Ser Val Tyr Ser Met	

260	265	270
Cys Gly Ser Val Ala Pro Ile Glu Lys Ile Cys Asp Leu Ala Lys Arg		
275	280	285
Tyr Gly Ala Leu Thr Phe Leu Asp Glu Val His Ala Val Gly Met Tyr		
290	295	300
Gly Pro His Gly Gln Gly Val Ala Glu His Leu Asp Phe Asp Leu His		
305	310	315
Leu Gln Ser Gly Ile Ala Ser Pro Ser Val Val Asp Lys Arg Thr Ile		
325	330	335
Leu Asp Arg Val Asp Met Ile Thr Gly Thr Cys Gly Lys Ser Phe Gly		
340	345	350
Thr Val Gly Gly Tyr Val Ala Gly Ser Ala Asn Leu Ile Asp Trp Leu		
355	360	365
Arg Ser Tyr Ala Pro Gly Phe Ile Phe Thr Thr Thr Leu Pro Pro Ala		
370	375	380
Ile Met Ala Gly Thr Ala Thr Ser Val Arg Ile Val Arg Ala Asp Ile		
385	390	395
Glu Ala Arg Ile Lys Gln Gln Leu Asn Thr Arg Tyr Val Lys Asp Ser		
405	410	415
Phe Glu Asn Leu Gly Ile Pro Val Ile Pro Asn Pro Ser His Ile Val		
420	425	430
Pro Val Leu Val Gly Asn Ala Ala Asp Ala Lys Lys Ala Ser Asp Met		
435	440	445
Leu Met Asn Lys His Arg Ile Tyr Val Gln Ala Ile Asn Tyr Pro Thr		
450	455	460
Val Pro Val Gly Glu Glu Arg Leu Arg Ile Thr Pro Thr Pro Gly His		
465	470	475
Gly Lys Glu Ile Cys Asp Gln Leu Ile Ser Ala Val Asp Asp Val Phe		
485	490	495
Thr Glu Leu Asn Leu Pro Arg Ile Asn Lys Trp Gln Ser Gln Gly Gly		
500	505	510
His Cys Gly Val Gly Asp Ala Asn Tyr Val Pro Glu Pro Asn Leu Trp		
515	520	525
Thr Gln Asp Gln Leu Ser Leu Thr Asn Gln Asp Leu His Ser Asn Val		
530	535	540
His Asn Pro Val Ile Glu Gln Ile Glu Thr Ser Ser Gly Val Arg Leu		
545	550	555
560		
<210> 30		
<211> 1683		

<212> DNA

<213> 人工

<220>

<223> 突变序列

<400> 30

```

atggagtttg tcgcccgtca gtccatgaat gcctctccct ttgtcaggtc aacttctacc      60
caccatttga agaagttggc agcaaacagt tctctagctg ctactgctag tcattctccc      120
gtggttggcc ctgctctcca acagcagaga tactactctc aaccttccaa gccagcccaa      180
gccccaaacct ccgacattgc tactgggata aagaaggatg tttctccgat ccgtatggac      240
tctaataaaa ccgcctttga ttacaatgga atgtatgagt ctgatcttgc gaataaacgt      300
aaagataact cgtatcgtaa tttcaataac atcaaccgtc tagccaagga gtttcccaag      360
gcacatcgcc agaccgaaga tgacaagggt accgtctggt gctctaacga ctacttagga      420
atgggtaggc atcctgagat tatcaaaacc atgaaggcta ccatggacaa gtacggttcc      480
ggagcaggag gaactaggaa cattgcaggt cataaccacg ccgctatcaa tttggaaagc      540
gagttggctt gcttgaacaa gaaggaagcg gctctggtgt tttcatcatg tttcatagct      600
aacgatgcaa tcatctcggt gttgggacaa aaaatcaaaa atttggtcat tttctctgac      660
cagtcgaatc atgcttccat gatattgggt gtgcgtaact ccaaagcgaa gaagcacatc      720
ttcaagcaca acaatttgaa ggatctggag tcgcagttag ctcatgacct caagtcgact      780
cctaaactga tcgccttcga gtcagtttac tctatgtgtg gatctgtggc tccattagag      840
aagatttgcg atttggctaa aaggtagcgt gccctcacct tcttgatga agttcatgct      900
gttggaaatg atggtcctca tggacagggt gtagctgagc atttggaact tgatctgcat      960
ttacagtctg gaatcgccag tcctagcgtg gtggacaaac gcacatatt ggatcgtgtc     1020
gacatgatta ctggtacttg cggaaagtca tttggtactg ttggagggtta cgttgctggt     1080
agtgccaaacc taattgattg gttaagatcc tatgcgccag gtttcatttt cactaccaca     1140
cttctctctg ctatcatggc tggtagagcc acttctgttc gtattgttag ggccgacatt     1200
gaggcccgta tcaagcaaca gcttaatact cgctacgtca aagactcatt tgaaaacctt     1260
ggtattccag tcattccaaa cccaagtcac attgttctct ttctagttgg aaatgctgca     1320
gatgccaaaga aggcattccg tatgttaatg aacaaacacc gtatttatgt tcaagctatt     1380
aactacccta ctgtgcctgt cggatgaaga cgactaagga ttactcctac tccaggtcat     1440
ggaaaggaga tttgtgacca gctgatcagc gctgtcgacg atgtttttac tgagcttaat     1500
ttaccaagaa tcaacaaatg gcagtcccaa ggtgggtcatt gcggtgttgg tgatgctaata     1560
tacgtaccag aacccaatct gtggactcag gaccagctca gcttgacaaa ccaagacttg     1620
cactccaatg tgcacaaccc agtgattgag cagatcgaaa cctcatcagg agtcagattg     1680
tag                                                                 1683

```

<210> 31

<211> 560

<212> PRT

<213> 人工

<220>

<223> 突变序列

<400> 31

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

290	295	300
Gly Pro His Gly Gln Gly Val Ala Glu His Leu Asp Phe Asp Leu His		
305	310	315
Leu Gln Ser Gly Ile Ala Ser Pro Ser Val Val Asp Lys Arg Thr Ile		320
	325	330
Leu Asp Arg Val Asp Met Ile Thr Gly Thr Cys Gly Lys Ser Phe Gly		335
	340	345
Thr Val Gly Gly Tyr Val Ala Gly Ser Ala Asn Leu Ile Asp Trp Leu		350
	355	360
Arg Ser Tyr Ala Pro Gly Phe Ile Phe Thr Thr Thr Leu Pro Pro Ala		365
370	375	380
Ile Met Ala Gly Thr Ala Thr Ser Val Arg Ile Val Arg Ala Asp Ile		
385	390	395
Glu Ala Arg Ile Lys Gln Gln Leu Asn Thr Arg Tyr Val Lys Asp Ser		400
	405	410
Phe Glu Asn Leu Gly Ile Pro Val Ile Pro Asn Pro Ser His Ile Val		415
	420	425
Pro Val Leu Val Gly Asn Ala Ala Asp Ala Lys Lys Ala Ser Asp Met		430
	435	440
Leu Met Asn Lys His Arg Ile Tyr Val Gln Ala Ile Asn Tyr Pro Thr		445
450	455	460
Val Pro Val Gly Glu Glu Arg Leu Arg Ile Thr Pro Thr Pro Gly His		
465	470	475
Gly Lys Glu Ile Cys Asp Gln Leu Ile Ser Ala Val Asp Asp Val Phe		480
	485	490
Thr Glu Leu Asn Leu Pro Arg Ile Asn Lys Trp Gln Ser Gln Gly Gly		495
	500	505
His Cys Gly Val Gly Asp Ala Asn Tyr Val Pro Glu Pro Asn Leu Trp		510
	515	520
Thr Gln Asp Gln Leu Ser Leu Thr Asn Gln Asp Leu His Ser Asn Val		525
	530	535
His Asn Pro Val Ile Glu Gln Ile Glu Thr Ser Ser Gly Val Arg Leu		540
545	550	555
		560

<210> 32

<211> 5

<212> PRT

<213> 人工

<220>

<223> 示例性HRM序列

<220>
<221> misc_feature
<222> (1) .. (1)
<223> Xaa可为R/L/N/A/C/S/H/I/G/Q
<220>
<221> misc_feature
<222> (4) .. (4)
<223> Xaa可为L/V/I/F/C
<220>
<221> misc_feature
<222> (5) .. (5)
<223> Xaa可为L/M/P/V
<400> 32
Xaa Cys Pro Xaa Xaa
1 5
<210> 33
<211> 5
<212> PRT
<213> 人工
<220>
<223> 示例性HRM序列。
<400> 33
Ala Cys Pro Phe Val
1 5
<210> 34
<211> 5
<212> PRT
<213> 人工
<220>
<223> 示例性HRM序列。
<400> 34
His Cys Pro Val Val
1 5
<210> 35
<211> 5
<212> PRT
<213> 人工
<220>
<223> 示例性HRM序列。

<400> 35	
Ile Cys Pro Phe Met	
1 5	
<210> 36	
<211> 5	
<212> PRT	
<213> 人工	
<220>	
<223> 示例性HRM序列。	
<400> 36	
Gly Cys Pro Val Val	
1 5	
<210> 37	
<211> 41	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 37	
cgtcagtcca tgaatgcctc tccctttgtc aggtcaactt c	41
<210> 38	
<211> 41	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 38	
gaagttgacc tgacaaaggg agaggcattc atggactgac g	41
<210> 39	
<211> 37	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 39	
gctgctactg ctagtcattc tcccgtggtt ggccctg	37
<210> 40	
<211> 37	
<212> DNA	

<213> 人工	
<220>	
<223> 合成引物	
<400> 40	
cagggccaac cacgggagaa tgactagcag tagcagc	37
<210> 41	
<211> 27	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 41	
aaacgctgtc ttggaacctataatgac	27
<210> 42	
<211> 38	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 42	
gacgggacgac aaactccatc gtttcgaata attagttg	38
<210> 43	
<211> 40	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 43	
caactaatta ttcgaaacga tggagtttgt cgcccgtcag	40
<210> 44	
<211> 45	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 44	
aattaaatac atttcaacta caatctgact cctgatgagg tttcg	45
<210> 45	
<211> 42	

<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 45	
cctcatcagg agtcagattg tagttgaaat gtatttaatt tg	42
<210> 46	
<211> 24	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 46	
aaactgtcag ttttgggcca ttg	24
<210> 47	
<211> 21	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 47	
gctctccaac agcagagata c	21
<210> 48	
<211> 21	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 48	
gtccatacgg atcggagaaa c	21
<210> 49	
<211> 19	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 49	
agcaacatcc ctgattccg	19
<210> 50	

<211>	20	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	50	
atgcgtacct tcaatcctgg		20
<210>	51	
<211>	24	
<212>	DNA	
<213>	人工	
<220>		
<223>	PCR探针. A1用6FAM标记. A9用Zen标记. A21用3IABkFQ标记	
<400>	51	
aagcccaaac ctccgacatt gcta		24
<210>	52	
<211>	28	
<212>	DNA	
<213>	人工	
<220>		
<223>	PCR探针. T1用HEX标记. A9用Zen标记. C28用3IABkFQ标记	
<400>	52	
tcgccgtaag ttcttggttt agacgttc		28
<210>	53	
<211>	23	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	53	
tagcgcagtc tctctatcgc ttc		23
<210>	54	
<211>	20	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	54	
cactgggttg tgcacattgg		20

<210>	55	
<211>	20	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	55	
acaatattct tctctgccgc		20
<210>	56	
<211>	20	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	56	
ttgatctcgt caagaatgcg		20
<210>	57	
<211>	23	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	57	
taggtgccac aacttttggt ttc		23
<210>	58	
<211>	24	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	58	
gatccaatgc gatgacattc ttgt		24
<210>	59	
<211>	24	
<212>	DNA	
<213>	人工	
<220>		
<223>	合成引物	
<400>	59	

acctgcaata actcctcttc tctg	24
<210> 60	
<211> 20	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 60	
ccactgaggg tagccgaatc	20
<210> 61	
<211> 22	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 61	
gggctctgaa aaactctttt gg	22
<210> 62	
<211> 27	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 62	
gcatgtctca ataacagatc tcgacgg	27
<210> 63	
<211> 27	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 63	
aagcctcttg tttttctgta aatgcac	27
<210> 64	
<211> 22	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	

<400> 64	
tgatggcgtc cgagatgaac tc	22
<210> 65	
<211> 21	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 65	
atggagtttg tcgcccgtca g	21
<210> 66	
<211> 27	
<212> DNA	
<213> 人工	
<220>	
<223> 合成引物	
<400> 66	
ctacaatctg actcctgatg aggtttc	27

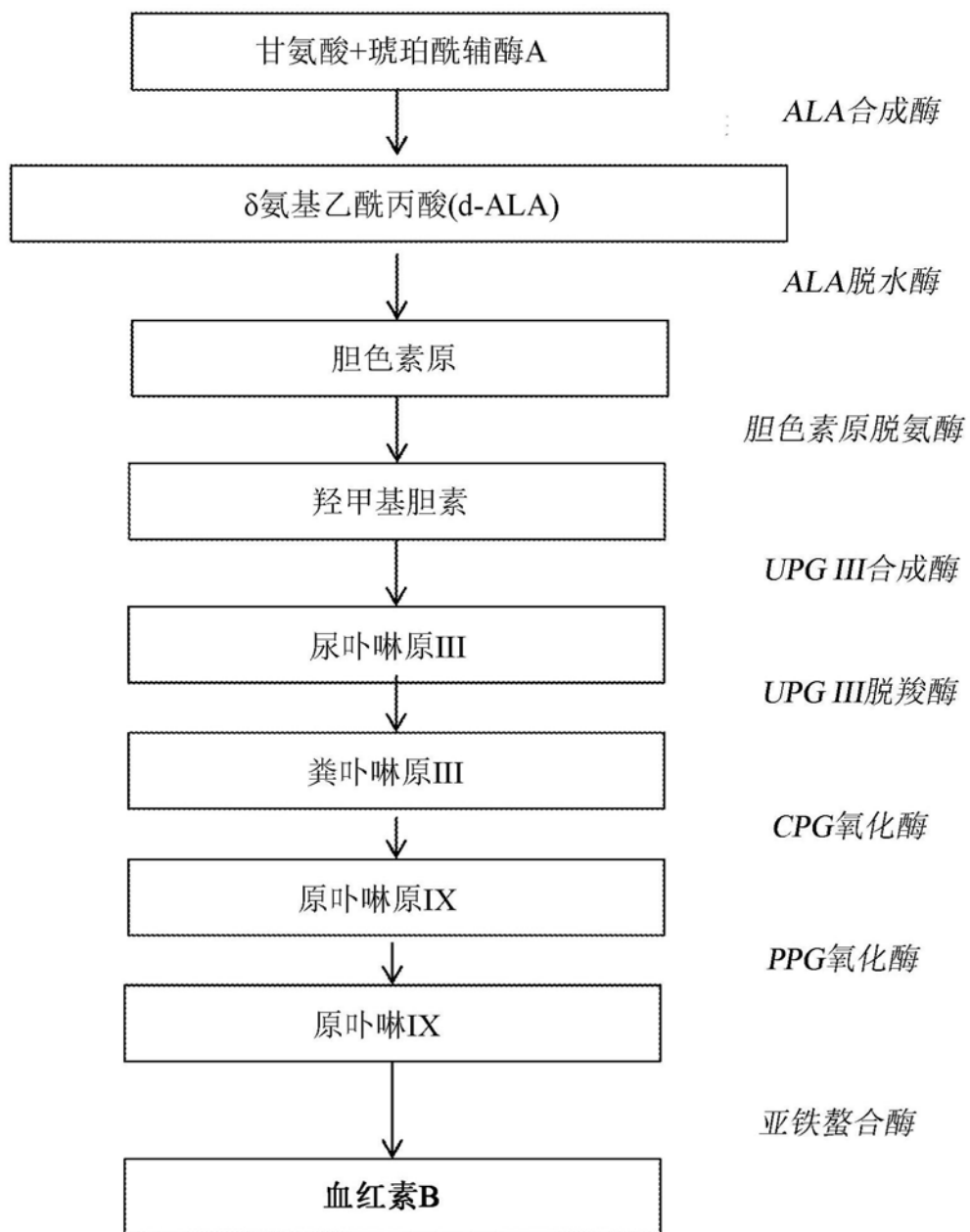


图1

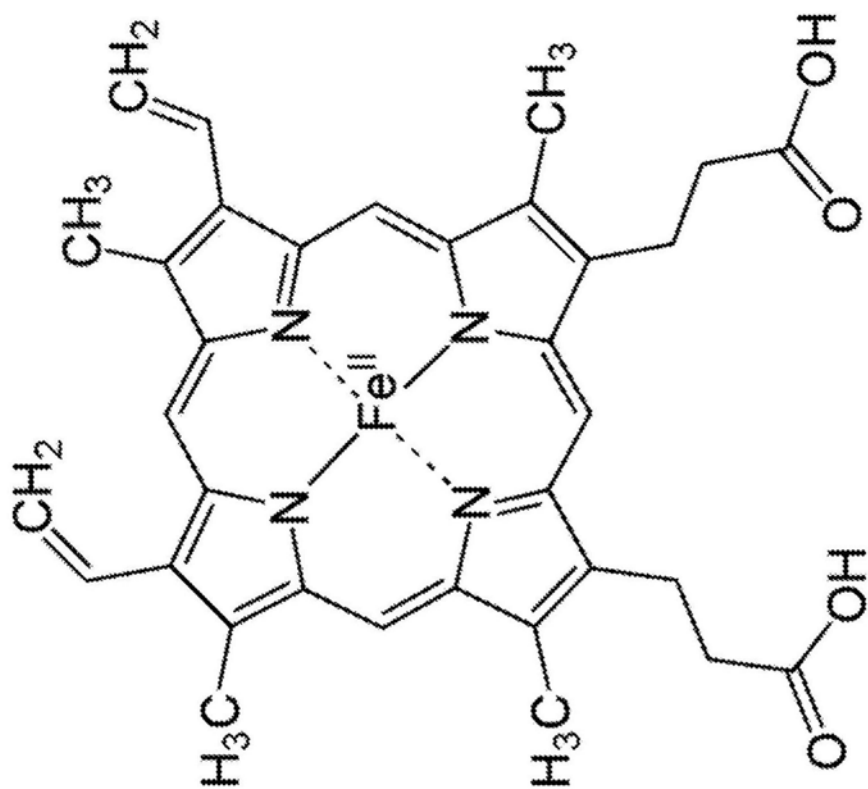


图2

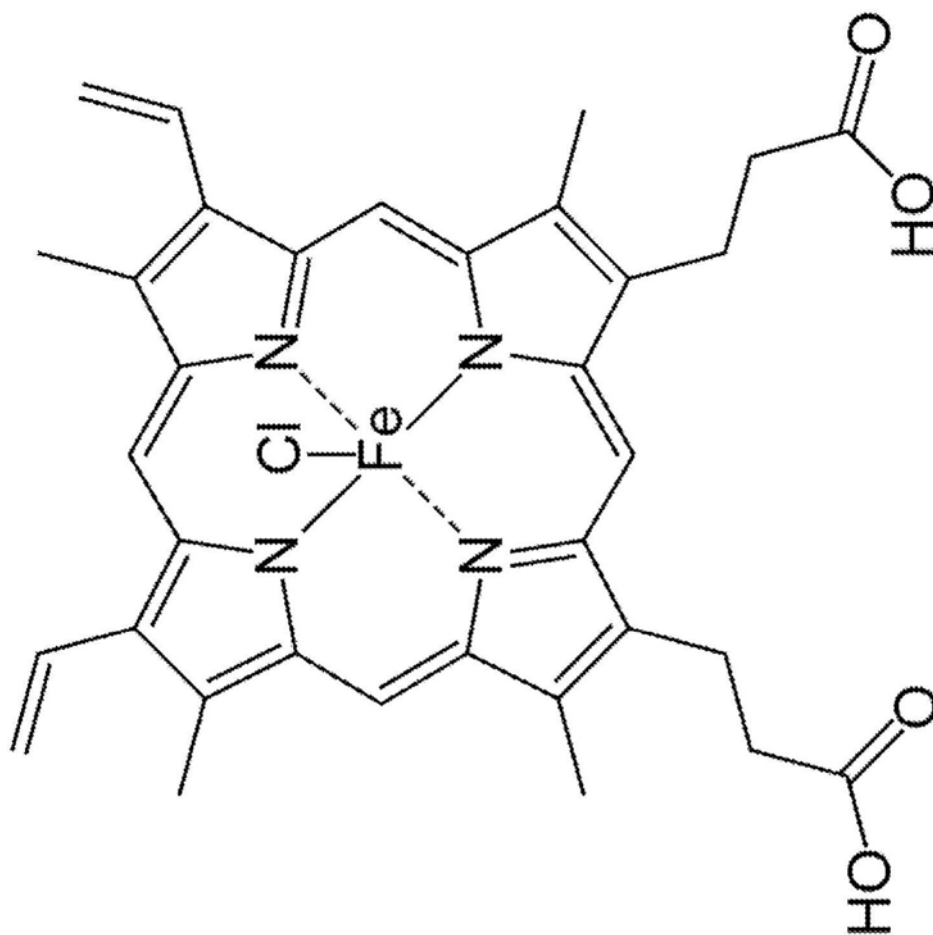


图3

<110> 非凡食品有限公司(Impossible Foods Inc.)

<210> 1

<211> 161

<212> PRT

<213> 绿豆(*Vigna radiata*)

<400> 1

Met	Thr	Thr	Thr	Leu	Glu	Arg	Gly	Phe	Thr	Glu	Glu	Gln	Glu	Ala	Leu	
1				5					10					15		
Val	Val	Lys	Ser	Trp	Asn	Val	Met	Lys	Lys	Asn	Ser	Gly	Glu	Leu	Gly	
			20					25					30			
Leu	Lys	Phe	Phe	Leu	Lys	Ile	Phe	Glu	Ile	Ala	Pro	Ser	Ala	Gln	Lys	
		35					40					45				
Leu	Phe	Ser	Phe	Leu	Arg	Asp	Ser	Thr	Val	Pro	Leu	Glu	Gln	Asn	Pro	
	50					55					60					
Lys	Leu	Lys	Pro	His	Ala	Val	Ser	Val	Phe	Val	Met	Thr	Cys	Asp	Ser	
65					70					75					80	
Ala	Val	Gln	Leu	Arg	Lys	Ala	Gly	Lys	Val	Thr	Val	Arg	Glu	Ser	Asn	
			85						90					95		
Leu	Lys	Lys	Leu	Gly	Ala	Thr	His	Phe	Arg	Thr	Gly	Val	Ala	Asn	Glu	
			100					105					110			
His	Phe	Glu	Val	Thr	Lys	Phe	Ala	Leu	Leu	Glu	Thr	Ile	Lys	Glu	Ala	
		115					120					125				
Val	Pro	Glu	Met	Trp	Ser	Pro	Ala	Met	Lys	Asn	Ala	Trp	Gly	Glu	Ala	
	130					135					140					
Tyr	Asp	Gln	Leu	Val	Asp	Ala	Ile	Lys	Tyr	Glu	Met	Lys	Pro	Pro	Ser	
145					150					155					160	
Ser																

<210> 2

<211> 133

<212> PRT

<213> 极端嗜酸甲烷氧化细菌(*Methylococcus inferorum*)

<400> 2

Met	Ile	Asp	Gln	Lys	Glu	Lys	Glu	Leu	Ile	Lys	Glu	Ser	Trp	Lys	Arg	
1				5					10					15		
Ile	Glu	Pro	Asn	Lys	Asn	Glu	Ile	Gly	Leu	Leu	Phe	Tyr	Ala	Asn	Leu	
			20					25					30			
Phe	Lys	Glu	Glu	Pro	Thr	Val	Ser	Val	Leu	Phe	Gln	Asn	Pro	Ile	Ser	
		35					40					45				
Ser	Gln	Ser	Arg	Lys	Leu	Met	Gln	Val	Leu	Gly	Ile	Leu	Val	Gln	Gly	

图4

50		55		60
Ile Asp Asn Leu Glu Gly	Leu Ile Pro Thr Leu	Gln Asp Leu Gly Arg		
65	70	75	80	
Arg His Lys Gln Tyr Gly	Val Val Asp Ser His Tyr	Pro Leu Val Gly		
	85	90	95	
Asp Cys Leu Leu Lys Ser	Ile Gln Glu Tyr Leu Gly	Gln Gly Phe Thr		
	100	105	110	
Glu Glu Ala Lys Ala Ala	Trp Thr Lys Val Tyr Gly	Ile Ala Ala Gln		
	115	120	125	
Val Met Thr Ala Glu				
130				

<210> 3

<211> 139

<212> PRT

<213> 风产液菌(Aquifex aeolicus)

<400> 3

Met Leu Ser Glu Glu Thr	Ile Arg Val Ile Lys Ser Thr	Val Pro Leu
1	5	10
Leu Lys Glu His Gly Thr	Glu Ile Thr Ala Arg Met Tyr	Glu Leu Leu
	20	25
Phe Ser Lys Tyr Pro Lys	Thr Lys Glu Leu Phe Ala Gly	Ala Ser Glu
	35	40
Glu Gln Pro Lys Lys Leu	Ala Asn Ala Ile Ile Ala Tyr	Ala Thr Tyr
	50	55
Ile Asp Arg Leu Glu Glu	Leu Asp Asn Ala Ile Ser Thr	Ile Ala Arg
65	70	75
Ser His Val Arg Arg Asn	Val Lys Pro Glu His Tyr Pro	Leu Val Lys
	85	90
Glu Cys Leu Leu Gln Ala	Ile Glu Glu Val Leu Asn Pro Gly	Glu Glu
	100	105
Val Leu Lys Ala Trp Glu	Glu Ala Tyr Asp Phe Leu Ala Lys	Thr Leu
	115	120
Ile Thr Leu Glu Lys Lys	Leu Tyr Ser Gln Pro	
130	135	

<210> 4

<211> 145

<212> PRT

<213> 大豆(Glycine max)

<400> 4

Met Gly Ala Phe Thr Glu	Lys Gln Glu Ala Leu Val Ser	Ser Ser Phe
1	5	10
Glu Ala Phe Lys Ala Asn	Ile Pro Gln Tyr Ser Val Val	Phe Tyr Thr
	20	25
		30

图4(接上页)

```

Ser Ile Leu Glu Lys Ala Pro Ala Ala Lys Asp Leu Phe Ser Phe Leu
   35           40           45
Ser Asn Gly Val Asp Pro Ser Asn Pro Lys Leu Thr Gly His Ala Glu
   50           55           60
Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Gly Gln Leu Lys Ala Asn
   65           70           75           80
Gly Thr Val Val Ala Asp Ala Ala Leu Gly Ser Ile His Ala Gln Lys
           85           90           95
Ala Ile Thr Asp Pro Gln Phe Val Val Val Lys Glu Ala Leu Leu Lys
           100           105           110
Thr Ile Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Ser Ser
           115           120           125
Ala Trp Glu Val Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala
   130           135           140
Phe
145

```

<210> 5

<211> 162

<212> PRT

<213> 大麦(*Hordeum vulgare*)

<400> 5

```

Met Ser Ala Ala Glu Gly Ala Val Val Phe Ser Glu Glu Lys Glu Ala
  1           5           10           15
Leu Val Leu Lys Ser Trp Ala Ile Met Lys Lys Asp Ser Ala Asn Leu
           20           25           30
Gly Leu Arg Phe Phe Leu Lys Ile Phe Glu Ile Ala Pro Ser Ala Arg
           35           40           45
Gln Met Phe Pro Phe Leu Arg Asp Ser Asp Val Pro Leu Glu Thr Asn
           50           55           60
Pro Lys Leu Lys Thr His Ala Val Ser Val Phe Val Met Thr Cys Glu
   65           70           75           80
Ala Ala Ala Gln Leu Arg Lys Ala Gly Lys Ile Thr Val Arg Glu Thr
           85           90           95
Thr Leu Lys Arg Leu Gly Gly Thr His Leu Lys Tyr Gly Val Ala Asp
           100           105           110
Gly His Phe Glu Val Thr Arg Phe Ala Leu Leu Glu Thr Ile Lys Glu
           115           120           125
Ala Leu Pro Ala Asp Met Trp Gly Pro Glu Met Arg Asn Ala Trp Gly
   130           135           140
Glu Ala Tyr Asp Gln Leu Val Ala Ala Ile Lys Gln Glu Met Lys Pro
   145           150           155           160
Ala Glu

```

<210> 6

图4(接上页)

<211> 1153

<212> PRT

<213> 稻瘟病菌(Magnaporthe oryzae)

<400> 6

Met	Asp	Gly	Ala	Val	Arg	Leu	Asp	Trp	Thr	Gly	Leu	Asp	Leu	Thr	Gly
1				5					10					15	
His	Glu	Ile	His	Asp	Gly	Val	Pro	Ile	Ala	Ser	Arg	Val	Gln	Val	Met
			20					25					30		
Val	Ser	Phe	Pro	Leu	Phe	Lys	Asp	Gln	His	Ile	Ile	Met	Ser	Ser	Lys
		35					40					45			
Glu	Ser	Pro	Ser	Arg	Lys	Ser	Ser	Thr	Ile	Gly	Gln	Ser	Thr	Arg	Asn
	50					55					60				
Gly	Ser	Cys	Gln	Ala	Asp	Thr	Gln	Lys	Gly	Gln	Leu	Pro	Pro	Val	Gly
65					70					75					80
Glu	Lys	Pro	Lys	Pro	Val	Lys	Glu	Asn	Pro	Met	Lys	Lys	Leu	Lys	Glu
				85					90					95	
Met	Ser	Gln	Arg	Pro	Leu	Pro	Thr	Gln	His	Gly	Asp	Gly	Thr	Tyr	Pro
			100					105					110		
Thr	Glu	Lys	Lys	Leu	Thr	Gly	Ile	Gly	Glu	Asp	Leu	Lys	His	Ile	Arg
		115					120					125			
Gly	Tyr	Asp	Val	Lys	Thr	Leu	Leu	Ala	Met	Val	Lys	Ser	Lys	Leu	Lys
	130					135					140				
Gly	Glu	Lys	Leu	Lys	Asp	Lys	Thr	Met	Leu	Met	Glu	Arg	Val	Met	
145					150					155					160
Gln	Leu	Val	Ala	Arg	Leu	Pro	Thr	Glu	Ser	Lys	Lys	Arg	Ala	Glu	Leu
			165						170					175	
Thr	Asp	Ser	Leu	Ile	Asn	Glu	Leu	Trp	Glu	Ser	Leu	Asp	His	Pro	Pro
			180					185					190		
Leu	Asn	Tyr	Leu	Gly	Pro	Glu	His	Ser	Tyr	Arg	Thr	Pro	Asp	Gly	Ser
		195					200					205			
Tyr	Asn	His	Pro	Phe	Asn	Pro	Gln	Leu	Gly	Ala	Ala	Gly	Ser	Arg	Tyr
	210					215					220				
Ala	Arg	Ser	Val	Ile	Pro	Thr	Val	Thr	Pro	Pro	Gly	Ala	Leu	Pro	Asp
225					230					235					240
Pro	Gly	Leu	Ile	Phe	Asp	Ser	Ile	Met	Gly	Arg	Thr	Pro	Asn	Ser	Tyr
			245						250					255	
Arg	Lys	His	Pro	Asn	Asn	Val	Ser	Ser	Ile	Leu	Trp	Tyr	Trp	Ala	Thr
			260					265					270		
Ile	Ile	Ile	His	Asp	Ile	Phe	Trp	Thr	Asp	Pro	Arg	Asp	Ile	Asn	Thr
		275					280					285			
Asn	Lys	Ser	Ser	Ser	Tyr	Leu	Asp	Leu	Ala	Pro	Leu	Tyr	Gly	Asn	Ser
	290					295					300				
Gln	Glu	Met	Gln	Asp	Ser	Ile	Arg	Thr	Phe	Lys	Asp	Gly	Arg	Met	Lys
305					310					315					320
Pro	Asp	Cys	Tyr	Ala	Asp	Lys	Arg	Leu	Ala	Gly	Met	Pro	Pro	Gly	Val
			325						330					335	
Ser	Val	Leu	Leu	Ile	Met	Phe	Asn	Arg	Phe	His	Asn	His	Val	Ala	Glu
			340					345					350		

图4(接上页)

Asn	Leu	Ala	Leu	Ile	Asn	Glu	Gly	Gly	Arg	Phe	Asn	Lys	Pro	Ser	Asp
		355					360					365			
Leu	Leu	Glu	Gly	Glu	Ala	Arg	Glu	Ala	Ala	Trp	Lys	Lys	Tyr	Asp	Asn
	370					375					380				
Asp	Leu	Phe	Gln	Val	Ala	Arg	Leu	Val	Thr	Ser	Gly	Leu	Tyr	Ile	Asn
385					390					395					400
Ile	Thr	Leu	Val	Asp	Tyr	Val	Arg	Asn	Ile	Val	Asn	Leu	Asn	Arg	Val
				405					410					415	
Asp	Thr	Thr	Trp	Thr	Leu	Asp	Pro	Arg	Gln	Asp	Ala	Gly	Ala	His	Val
			420						425				430		
Gly	Thr	Ala	Asp	Gly	Ala	Glu	Arg	Gly	Thr	Gly	Asn	Ala	Val	Ser	Ala
		435					440					445			
Glu	Phe	Asn	Leu	Cys	Tyr	Arg	Trp	His	Ser	Cys	Ile	Ser	Glu	Lys	Asp
	450					455					460				
Ser	Lys	Phe	Val	Glu	Ala	Gln	Phe	Gln	Asn	Ile	Phe	Gly	Lys	Pro	Ala
465					470					475					480
Ser	Glu	Val	Arg	Pro	Asp	Glu	Met	Trp	Lys	Gly	Phe	Ala	Lys	Met	Glu
				485					490					495	
Gln	Asn	Thr	Pro	Ala	Asp	Pro	Gly	Gln	Arg	Thr	Phe	Gly	Gly	Phe	Lys
			500					505					510		
Arg	Gly	Pro	Asp	Gly	Lys	Phe	Asp	Asp	Asp	Asp	Leu	Val	Arg	Cys	Ile
		515					520					525			
Ser	Glu	Ala	Val	Glu	Asp	Val	Ala	Gly	Ala	Phe	Gly	Ala	Arg	Asn	Val
	530					535					540				
Pro	Gln	Ala	Met	Lys	Val	Val	Glu	Thr	Met	Gly	Ile	Ile	Gln	Gly	Arg
545					550					555					560
Lys	Trp	Asn	Val	Ala	Gly	Leu	Asn	Glu	Phe	Arg	Lys	His	Phe	His	Leu
			565						570					575	
Lys	Pro	Tyr	Ser	Thr	Phe	Glu	Asp	Ile	Asn	Ser	Asp	Pro	Gly	Val	Ala
			580					585					590		
Glu	Ala	Leu	Arg	Arg	Leu	Tyr	Asp	His	Pro	Asp	Asn	Val	Glu	Leu	Tyr
		595					600					605			
Pro	Gly	Leu	Val	Ala	Glu	Glu	Asp	Lys	Gln	Pro	Met	Val	Pro	Gly	Val
	610					615					620				
Gly	Ile	Ala	Pro	Thr	Tyr	Thr	Ile	Ser	Arg	Val	Val	Leu	Ser	Asp	Ala
625					630					635					640
Val	Cys	Leu	Val	Arg	Gly	Asp	Arg	Phe	Tyr	Thr	Thr	Asp	Phe	Thr	Pro
			645						650					655	
Arg	Asn	Leu	Thr	Asn	Trp	Gly	Tyr	Lys	Glu	Val	Asp	Tyr	Asp	Leu	Ser
			660					665					670		
Val	Asn	His	Gly	Cys	Val	Phe	Tyr	Lys	Leu	Phe	Ile	Arg	Ala	Phe	Pro
		675					680					685			
Asn	His	Phe	Lys	Gln	Asn	Ser	Val	Tyr	Ala	His	Tyr	Pro	Met	Val	Val
	690					695					700				
Pro	Ser	Glu	Asn	Lys	Arg	Ile	Leu	Glu	Ala	Leu	Gly	Arg	Ala	Asp	Leu
705					710					715					720
Phe	Asp	Phe	Glu	Ala	Pro	Lys	Tyr	Ile	Pro	Pro	Arg	Val	Asn	Ile	Thr
			725						730					735	
Ser	Tyr	Gly	Gly	Ala	Glu	Tyr	Ile	Leu	Glu	Thr	Gln	Glu	Lys	Tyr	Lys

图4(接上页)

			740					745				750			
Val	Thr	Trp	His	Glu	Gly	Leu	Gly	Phe	Leu	Met	Gly	Glu	Gly	Gly	Leu
		755					760					765			
Lys	Phe	Met	Leu	Ser	Gly	Asp	Asp	Pro	Leu	His	Ala	Gln	Gln	Arg	Lys
		770				775					780				
Cys	Met	Ala	Ala	Gln	Leu	Tyr	Lys	Asp	Gly	Trp	Thr	Glu	Ala	Val	Lys
785					790					795					800
Ala	Phe	Tyr	Ala	Gly	Met	Met	Glu	Glu	Leu	Leu	Val	Ser	Lys	Ser	Tyr
			805						810					815	
Phe	Leu	Gly	Asn	Asn	Lys	His	Arg	His	Val	Asp	Ile	Ile	Arg	Asp	Val
			820					825					830		
Gly	Asn	Met	Val	His	Val	His	Phe	Ala	Ser	Gln	Val	Phe	Gly	Leu	Pro
		835					840					845			
Leu	Lys	Thr	Ala	Lys	Asn	Pro	Thr	Gly	Val	Phe	Thr	Glu	Gln	Glu	Met
	850					855					860				
Tyr	Gly	Ile	Leu	Ala	Ala	Ile	Phe	Thr	Thr	Ile	Phe	Phe	Asp	Leu	Asp
865					870					875					880
Pro	Ser	Lys	Ser	Phe	Pro	Leu	Arg	Thr	Lys	Thr	Arg	Glu	Val	Cys	Gln
			885						890					895	
Lys	Leu	Ala	Lys	Leu	Val	Glu	Ala	Asn	Val	Lys	Leu	Ile	Asn	Lys	Ile
			900					905					910		
Pro	Trp	Ser	Arg	Gly	Met	Phe	Val	Gly	Lys	Pro	Ala	Lys	Asp	Glu	Pro
		915					920					925			
Leu	Ser	Ile	Tyr	Gly	Lys	Thr	Met	Ile	Lys	Gly	Leu	Lys	Ala	His	Gly
	930					935					940				
Leu	Ser	Asp	Tyr	Asp	Ile	Ala	Trp	Ser	His	Val	Val	Pro	Thr	Ser	Gly
945					950					955					960
Ala	Met	Val	Pro	Asn	Gln	Ala	Gln	Val	Phe	Ala	Gln	Ala	Val	Asp	Tyr
			965						970					975	
Tyr	Leu	Ser	Pro	Ala	Gly	Met	His	Tyr	Ile	Pro	Glu	Ile	His	Met	Val
			980					985					990		
Ala	Leu	Gln	Pro	Ser	Thr	Pro	Glu	Thr	Asp	Ala	Leu	Leu	Leu	Gly	Tyr
		995					1000					1005			
Ala	Met	Glu	Gly	Ile	Arg	Leu	Ala	Gly	Thr	Phe	Gly	Ser	Tyr	Arg	Glu
	1010					1015					1020				
Ala	Ala	Val	Asp	Asp	Val	Val	Lys	Glu	Asp	Asn	Gly	Arg	Gln	Val	Pro
1025					1030					1035					1040
Val	Lys	Ala	Gly	Asp	Arg	Val	Phe	Val	Ser	Phe	Val	Asp	Ala	Ala	Arg
			1045						1050				1055		
Asp	Pro	Lys	His	Phe	Pro	Asp	Pro	Glu	Val	Val	Asn	Pro	Arg	Arg	Pro
			1060					1065				1070			
Ala	Lys	Lys	Tyr	Ile	His	Tyr	Gly	Val	Gly	Pro	His	Ala	Cys	Leu	Gly
		1075					1080					1085			
Arg	Asp	Ala	Ser	Gln	Ile	Ala	Ile	Thr	Glu	Met	Phe	Arg	Cys	Leu	Phe
	1090					1095					1100				
Arg	Arg	Arg	Asn	Val	Arg	Arg	Val	Pro	Gly	Pro	Gln	Gly	Glu	Leu	Lys
1105				1110						1115					1120
Lys	Val	Pro	Arg	Pro	Gly	Gly	Phe	Tyr	Val	Tyr	Met	Arg	Glu	Asp	Trp
			1125						1130					1135	

图4(接上页)

Gly Gly Leu Phe Pro Phe Pro Val Thr Met Arg Val Met Trp Asp Asp
 1140 1145 1150
 Glu

<210> 7
 <211> 530
 <212> PRT
 <213> 尖孢镰刀菌(*Fusarium oxysporum*)

<400> 7
 Met Lys Gly Ser Ala Thr Leu Ala Phe Ala Leu Val Gln Phe Ser Ala
 1 5 10 15
 Ala Ser Gln Leu Val Trp Pro Ser Lys Trp Asp Glu Val Glu Asp Leu
 20 25 30
 Leu Tyr Met Gln Gly Gly Phe Asn Lys Arg Gly Phe Ala Asp Ala Leu
 35 40 45
 Arg Thr Cys Glu Phe Gly Ser Asn Val Pro Gly Thr Gln Asn Thr Ala
 50 55 60
 Glu Trp Leu Arg Thr Ala Phe His Asp Ala Ile Thr His Asp Ala Lys
 65 70 75 80
 Ala Gly Thr Gly Gly Leu Asp Ala Ser Ile Tyr Trp Glu Ser Ser Arg
 85 90 95
 Pro Glu Asn Pro Gly Lys Ala Phe Asn Asn Thr Phe Gly Phe Phe Ser
 100 105 110
 Gly Phe His Asn Pro Arg Ala Thr Ala Ser Asp Leu Thr Ala Leu Gly
 115 120 125
 Thr Val Leu Ala Val Gly Ala Cys Asn Gly Pro Arg Ile Pro Phe Arg
 130 135 140
 Ala Gly Arg Ile Asp Ala Tyr Lys Ala Gly Pro Ala Gly Val Pro Glu
 145 150 155 160
 Pro Ser Thr Asn Leu Lys Asp Thr Phe Ala Ala Phe Thr Lys Ala Gly
 165 170 175
 Phe Thr Lys Glu Met Thr Ala Met Val Ala Cys Gly His Ala Ile
 180 185 190
 Gly Gly Val His Ser Val Asp Phe Pro Glu Ile Val Gly Ile Lys Ala
 195 200 205
 Asp Pro Asn Asn Asp Thr Asn Val Pro Phe Gln Lys Asp Val Ser Ser
 210 215 220
 Phe His Asn Gly Ile Val Thr Glu Tyr Leu Ala Gly Thr Ser Lys Asn
 225 230 235 240
 Pro Leu Val Ala Ser Lys Asn Ala Thr Phe His Ser Asp Lys Arg Ile
 245 250 255
 Phe Asp Asn Asp Lys Ala Thr Met Lys Lys Leu Ser Thr Lys Ala Gly
 260 265 270
 Phe Asn Ser Met Cys Ala Asp Ile Leu Thr Arg Met Ile Asp Thr Val
 275 280 285
 Pro Lys Ser Val Gln Leu Thr Pro Val Leu Glu Ala Tyr Asp Val Arg

图4(接上页)

```

      290              295              300
Pro Tyr Ile Thr Glu Leu Ser Leu Asn Asn Lys Asn Lys Ile His Phe
305              310              315              320
Thr Gly Ser Val Arg Val Arg Ile Thr Asn Asn Ile Arg Asp Asn Asn
      325              330              335
Asp Leu Ala Ile Asn Leu Ile Tyr Val Gly Arg Asp Gly Lys Lys Val
      340              345              350
Thr Val Pro Thr Gln Gln Val Thr Phe Gln Gly Gly Thr Ser Phe Gly
      355              360              365
Ala Gly Glu Val Phe Ala Asn Phe Glu Phe Asp Thr Thr Met Asp Ala
      370              375              380
Lys Asn Gly Ile Thr Lys Phe Phe Ile Gln Glu Val Lys Pro Ser Thr
385              390              395              400
Lys Ala Thr Val Thr His Asp Asn Gln Lys Thr Gly Gly Tyr Lys Val
      405              410              415
Asp Asp Thr Val Leu Tyr Gln Leu Gln Gln Ser Cys Ala Val Leu Glu
      420              425              430
Lys Leu Pro Asn Ala Pro Leu Val Val Thr Ala Met Val Arg Asp Ala
      435              440              445
Arg Ala Lys Asp Ala Leu Thr Leu Arg Val Ala His Lys Lys Pro Val
      450              455              460
Lys Gly Ser Ile Val Pro Arg Phe Gln Thr Ala Ile Thr Asn Phe Lys
465              470              475              480
Ala Thr Gly Lys Lys Ser Ser Gly Tyr Thr Gly Phe Gln Ala Lys Thr
      485              490              495
Met Phe Glu Glu Gln Ser Thr Tyr Phe Asp Ile Val Leu Gly Gly Ser
      500              505              510
Pro Ala Ser Gly Val Gln Phe Leu Thr Ser Gln Ala Met Pro Ser Gln
      515              520              525
Cys Ser
530

```

<210> 8

<211> 358

<212> PRT

<213> 禾谷镰刀菌(*Fusarium graminearum*)

<400> 8

```

Met Ala Ser Ala Thr Arg Gln Phe Ala Arg Ala Ala Thr Arg Ala Thr
 1              5              10              15
Arg Asn Gly Phe Ala Ile Ala Pro Arg Gln Val Ile Arg Gln Gln Gly
      20              25              30
Arg Arg Tyr Tyr Ser Ser Glu Pro Ala Gln Lys Ser Ser Ser Ala Trp
      35              40              45
Ile Trp Leu Thr Gly Ala Ala Val Ala Gly Gly Ala Gly Tyr Tyr Phe
      50              55              60
Tyr Gly Asn Ser Ala Ser Ser Ala Thr Ala Lys Val Phe Asn Pro Ser
65              70              75              80

```

图4(接上页)

```

Lys Glu Asp Tyr Gln Lys Val Tyr Asn Glu Ile Ala Ala Arg Leu Glu
      85                      90                      95
Glu Lys Asp Asp Tyr Asp Asp Gly Ser Tyr Gly Pro Val Leu Val Arg
      100                    105                    110
Leu Ala Trp His Ala Ser Gly Thr Tyr Asp Lys Glu Thr Gly Thr Gly
      115                    120                    125
Gly Ser Asn Gly Ala Thr Met Arg Phe Ala Pro Glu Ser Asp His Gly
      130                    135                    140
Ala Asn Ala Gly Leu Ala Ala Ala Arg Asp Phe Leu Gln Pro Val Lys
145                      150                      155                      160
Glu Lys Phe Pro Trp Ile Thr Tyr Ser Asp Leu Trp Ile Leu Ala Gly
      165                    170                    175
Val Cys Ala Ile Gln Glu Met Leu Gly Pro Ala Ile Pro Tyr Arg Pro
      180                    185                    190
Gly Arg Ser Asp Arg Asp Val Ser Gly Cys Thr Pro Asp Gly Arg Leu
      195                    200                    205
Pro Asp Ala Ser Lys Arg Gln Asp His Leu Arg Gly Ile Phe Gly Arg
      210                    215                    220
Met Gly Phe Asn Asp Gln Glu Ile Val Ala Leu Ser Gly Ala His Ala
225                      230                      235                      240
Leu Gly Arg Cys His Thr Asp Arg Ser Gly Tyr Ser Gly Pro Trp Thr
      245                    250                    255
Phe Ser Pro Thr Val Leu Thr Asn Asp Tyr Phe Arg Leu Leu Val Glu
      260                    265                    270
Glu Lys Trp Gln Trp Lys Lys Trp Asn Gly Pro Ala Gln Tyr Glu Asp
      275                    280                    285
Lys Ser Thr Lys Ser Leu Met Met Leu Pro Ser Asp Ile Ala Leu Ile
      290                    295                    300
Glu Asp Lys Lys Phe Lys Pro Trp Val Glu Lys Tyr Ala Lys Asp Asn
305                      310                      315                      320
Asp Ala Phe Phe Lys Asp Phe Ser Asn Val Val Leu Arg Leu Phe Glu
      325                    330                    335
Leu Gly Val Pro Phe Ala Gln Gly Thr Glu Asn Gln Arg Trp Thr Phe
      340                    345                    350
Lys Pro Thr His Gln Glu
      355

```

<210> 9

<211> 122

<212> PRT

<213> 卵配衣藻(*Chlamydomonas eugametos*)

<400> 9

```

Met Ser Leu Phe Ala Lys Leu Gly Gly Arg Glu Ala Val Glu Ala Ala
  1                      5                      10                      15
Val Asp Lys Phe Tyr Asn Lys Ile Val Ala Asp Pro Thr Val Ser Thr
      20                    25                    30
Tyr Phe Ser Asn Thr Asp Met Lys Val Gln Arg Ser Lys Gln Phe Ala

```

图4(接上页)

```

          35          40          45
Phe Leu Ala Tyr Ala Leu Gly Gly Ala Ser Glu Trp Lys Gly Lys Asp
  50          55          60
Met Arg Thr Ala His Lys Asp Leu Val Pro His Leu Ser Asp Val His
65          70          75          80
Phe Gln Ala Val Ala Arg His Leu Ser Asp Thr Leu Thr Glu Leu Gly
          85          90          95
Val Pro Pro Glu Asp Ile Thr Asp Ala Met Ala Val Val Ala Ser Thr
          100          105          110
Arg Thr Glu Val Leu Asn Met Pro Gln Gln
          115          120

```

<210> 10

<211> 121

<212> PRT

<213> 梨形四膜虫(*Tetrahymena pyriformis*)

<400> 10

```

Met Asn Lys Pro Gln Thr Ile Tyr Glu Lys Leu Gly Gly Glu Asn Ala
  1          5          10          15
Met Lys Ala Ala Val Pro Leu Phe Tyr Lys Lys Val Leu Ala Asp Glu
          20          25          30
Arg Val Lys His Phe Phe Lys Asn Thr Asp Met Asp His Gln Thr Lys
          35          40          45
Gln Gln Thr Asp Phe Leu Thr Met Leu Leu Gly Gly Pro Asn His Tyr
          50          55          60
Lys Gly Lys Asn Met Thr Glu Ala His Lys Gly Met Asn Leu Gln Asn
65          70          75          80
Leu His Phe Asp Ala Ile Ile Glu Asn Leu Ala Ala Thr Leu Lys Glu
          85          90          95
Leu Gly Val Thr Asp Ala Val Ile Asn Glu Ala Ala Lys Val Ile Glu
          100          105          110
His Thr Arg Lys Asp Met Leu Gly Lys
          115          120

```

<210> 11

<211> 117

<212> PRT

<213> 大草履虫(*Paramecium caudatum*)

<400> 11

```

Met Ser Leu Phe Glu Gln Leu Gly Gly Gln Ala Ala Val Gln Ala Val
  1          5          10          15
Thr Ala Gln Phe Tyr Ala Asn Ile Gln Ala Asp Ala Thr Val Ala Thr
          20          25          30
Phe Phe Asn Gly Ile Asp Met Pro Asn Gln Thr Asn Lys Thr Ala Ala
          35          40          45

```

图4(接上页)

```

Phe Leu Cys Ala Ala Leu Gly Gly Pro Asn Ala Trp Thr Gly Arg Asn
 50          55          60
Leu Lys Glu Val His Ala Asn Met Gly Val Ser Asn Ala Gln Phe Thr
65          70          75          80
Thr Val Ile Gly His Leu Arg Ser Ala Leu Thr Gly Ala Gly Val Ala
          85          90          95
Ala Ala Leu Val Glu Gln Thr Val Ala Val Ala Glu Thr Val Arg Gly
          100          105          110
Asp Val Val Thr Val
          115

```

<210> 12

<211> 147

<212> PRT

<213> 黑曲霉(*Aspergillus niger*)

<400> 12

```

Met Pro Leu Thr Pro Glu Gln Ile Lys Ile Ile Lys Ala Thr Val Pro
 1          5          10          15
Val Leu Gln Glu Tyr Gly Thr Lys Ile Thr Thr Ala Phe Tyr Met Asn
          20          25          30
Met Ser Thr Val His Pro Glu Leu Asn Ala Val Phe Asn Thr Ala Asn
          35          40          45
Gln Val Lys Gly His Gln Ala Arg Ala Leu Ala Gly Ala Leu Phe Ala
          50          55          60
Tyr Ala Ser His Ile Asp Asp Leu Gly Ala Leu Gly Pro Ala Val Glu
65          70          75          80
Leu Ile Cys Asn Lys His Ala Ser Leu Tyr Ile Gln Ala Asp Glu Tyr
          85          90          95
Lys Ile Val Gly Lys Tyr Leu Leu Glu Ala Met Lys Glu Val Leu Gly
          100          105          110
Asp Ala Cys Thr Asp Asp Ile Leu Asp Ala Trp Gly Ala Ala Tyr Trp
          115          120          125
Ala Leu Ala Asp Ile Met Ile Asn Arg Glu Ala Ala Leu Tyr Lys Gln
          130          135          140
Ser Gln Gly
145

```

<210> 13

<211> 165

<212> PRT

<213> 玉米(*Zea mays*)

<400> 13

```

Met Ala Leu Ala Glu Ala Asp Asp Gly Ala Val Val Phe Gly Glu Glu
 1          5          10          15
Gln Glu Ala Leu Val Leu Lys Ser Trp Ala Val Met Lys Lys Asp Ala

```

图4(接上页)

		20						25				30					
Ala	Asn	Leu	Gly	Leu	Arg	Phe	Phe	Leu	Lys	Val	Phe	Glu	Ile	Ala	Pro		
		35						40				45					
Ser	Ala	Glu	Gln	Met	Phe	Ser	Phe	Leu	Arg	Asp	Ser	Asp	Val	Pro	Leu		
	50					55					60						
Glu	Lys	Asn	Pro	Lys	Leu	Lys	Thr	His	Ala	Met	Ser	Val	Phe	Val	Met		
65					70					75					80		
Thr	Cys	Glu	Ala	Ala	Ala	Gln	Leu	Arg	Lys	Ala	Gly	Lys	Val	Thr	Val		
				85					90					95			
Arg	Glu	Thr	Thr	Leu	Lys	Arg	Leu	Gly	Ala	Thr	His	Leu	Arg	Tyr	Gly		
			100					105						110			
Val	Ala	Asp	Gly	His	Phe	Glu	Val	Thr	Gly	Phe	Ala	Leu	Leu	Glu	Thr		
		115						120				125					
Ile	Lys	Glu	Ala	Leu	Pro	Ala	Asp	Met	Trp	Ser	Leu	Glu	Met	Lys	Lys		
	130					135					140						
Ala	Trp	Ala	Glu	Ala	Tyr	Ser	Gln	Leu	Val	Ala	Ala	Ile	Lys	Arg	Glu		
145					150					155					160		
Met	Lys	Pro	Asp	Ala													
				165													

<210> 14

<211> 169

<212> PRT

<213> 粳稻(*Oryza sativa* subsp. *japonica*)

<400> 14

Met	Ala	Leu	Val	Glu	Gly	Asn	Asn	Gly	Val	Ser	Gly	Gly	Ala	Val	Ser		
1				5				10					15				
Phe	Ser	Glu	Glu	Gln	Glu	Ala	Leu	Val	Leu	Lys	Ser	Trp	Ala	Ile	Met		
		20						25					30				
Lys	Lys	Asp	Ser	Ala	Asn	Ile	Gly	Leu	Arg	Phe	Phe	Leu	Lys	Ile	Phe		
		35				40						45					
Glu	Val	Ala	Pro	Ser	Ala	Ser	Gln	Met	Phe	Ser	Phe	Leu	Arg	Asn	Ser		
	50					55					60						
Asp	Val	Pro	Leu	Glu	Lys	Asn	Pro	Lys	Leu	Lys	Thr	His	Ala	Met	Ser		
65					70					75				80			
Val	Phe	Val	Met	Thr	Cys	Glu	Ala	Ala	Ala	Gln	Leu	Arg	Lys	Ala	Gly		
				85				90						95			
Lys	Val	Thr	Val	Arg	Asp	Thr	Thr	Leu	Lys	Arg	Leu	Gly	Ala	Thr	His		
			100					105					110				
Phe	Lys	Tyr	Gly	Val	Gly	Asp	Ala	His	Phe	Glu	Val	Thr	Arg	Phe	Ala		
		115				120						125					
Leu	Leu	Glu	Thr	Ile	Lys	Glu	Ala	Val	Pro	Val	Asp	Met	Trp	Ser	Pro		
	130					135					140						
Ala	Met	Lys	Ser	Ala	Trp	Ser	Glu	Ala	Tyr	Asn	Gln	Leu	Val	Ala	Ala		
145					150					155					160		
Ile	Lys	Gln	Glu	Met	Lys	Pro	Ala	Glu									
				165													

图4(接上页)

<210> 15
 <211> 160
 <212> PRT
 <213> 拟南芥(*Arabidopsis thaliana*)

<400> 15
 Met Glu Ser Glu Gly Lys Ile Val Phe Thr Glu Glu Gln Glu Ala Leu
 1 5 10 15
 Val Val Lys Ser Trp Ser Val Met Lys Lys Asn Ser Ala Glu Leu Gly
 20 25 30
 Leu Lys Leu Phe Ile Lys Ile Phe Glu Ile Ala Pro Thr Thr Lys Lys
 35 40 45
 Met Phe Ser Phe Leu Arg Asp Ser Pro Ile Pro Ala Glu Gln Asn Pro
 50 55 60
 Lys Leu Lys Pro His Ala Met Ser Val Phe Val Met Cys Cys Glu Ser
 65 70 75 80
 Ala Val Gln Leu Arg Lys Thr Gly Lys Val Thr Val Arg Glu Thr Thr
 85 90 95
 Leu Lys Arg Leu Gly Ala Ser His Ser Lys Tyr Gly Val Val Asp Glu
 100 105 110
 His Phe Glu Val Ala Lys Tyr Ala Leu Leu Glu Thr Ile Lys Glu Ala
 115 120 125
 Val Pro Glu Met Trp Ser Pro Glu Met Lys Val Ala Trp Gly Gln Ala
 130 135 140
 Tyr Asp His Leu Val Ala Ile Lys Ala Glu Met Asn Leu Ser Asn
 145 150 155 160

<210> 16
 <211> 147
 <212> PRT
 <213> 豌豆(*Pisum sativum*)

<400> 16
 Met Gly Phe Thr Asp Lys Gln Glu Ala Leu Val Asn Ser Ser Trp Glu
 1 5 10 15
 Ser Phe Lys Gln Asn Leu Ser Gly Asn Ser Ile Leu Phe Tyr Thr Ile
 20 25 30
 Ile Leu Glu Lys Ala Pro Ala Ala Lys Gly Leu Phe Ser Phe Leu Lys
 35 40 45
 Asp Thr Ala Gly Val Glu Asp Ser Pro Lys Leu Gln Ala His Ala Glu
 50 55 60
 Gln Val Phe Gly Leu Val Arg Asp Ser Ala Ala Gln Leu Arg Thr Lys
 65 70 75 80
 Gly Glu Val Val Leu Gly Asn Ala Thr Leu Gly Ala Ile His Val Gln
 85 90 95
 Arg Gly Val Thr Asp Pro His Phe Val Val Val Lys Glu Ala Leu Leu

图4(接上页)

```

          100          105          110
Gln Thr Ile Lys Lys Ala Ser Gly Asn Asn Trp Ser Glu Glu Leu Asn
          115          120          125
Thr Ala Trp Glu Val Ala Tyr Asp Gly Leu Ala Thr Ala Ile Lys Lys
          130          135          140
Ala Met Thr
145

```

<210> 17

<211> 145

<212> PRT

<213> 豇豆(*Vigna unguiculata*)

<400> 17

```

Met Val Ala Phe Ser Asp Lys Gln Glu Ala Leu Val Asn Gly Ala Tyr
 1          5          10          15
Glu Ala Phe Lys Ala Asn Ile Pro Lys Tyr Ser Val Val Phe Tyr Thr
          20          25          30
Thr Ile Leu Glu Lys Ala Pro Ala Ala Lys Asn Leu Phe Ser Phe Leu
          35          40          45
Ala Asn Gly Val Asp Ala Thr Asn Pro Lys Leu Thr Gly His Ala Glu
          50          55          60
Lys Leu Phe Gly Leu Val Arg Asp Ser Ala Ala Gln Leu Arg Ala Ser
65          70          75          80
Gly Gly Val Val Ala Asp Ala Ala Leu Gly Ala Val His Ser Gln Lys
          85          90          95
Ala Val Asn Asp Ala Gln Phe Val Val Val Lys Glu Ala Leu Val Lys
          100          105          110
Thr Leu Lys Glu Ala Val Gly Asp Lys Trp Ser Asp Glu Leu Gly Thr
          115          120          125
Ala Val Glu Leu Ala Tyr Asp Glu Leu Ala Ala Ala Ile Lys Lys Ala
          130          135          140
Tyr
145

```

<210> 18

<211> 154

<212> PRT

<213> 黄牛(*Bos taurus*)

<400> 18

```

Met Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Ala Trp Gly
 1          5          10          15
Lys Val Glu Ala Asp Val Ala Gly His Gly Gln Glu Val Leu Ile Arg
          20          25          30
Leu Phe Thr Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys
          35          40          45

```

图4(接上页)

```

His Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys
 50                      55                      60
His Gly Asn Thr Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys
 65                      70                      75                      80
Gly His His Glu Ala Glu Val Lys His Leu Ala Glu Ser His Ala Asn
                      85                      90                      95
Lys His Lys Ile Pro Val Lys Tyr Leu Glu Phe Ile Ser Asp Ala Ile
                      100                      105                      110
Ile His Val Leu His Ala Lys His Pro Ser Asp Phe Gly Ala Asp Ala
                      115                      120                      125
Gln Ala Ala Met Ser Lys Ala Leu Glu Leu Phe Arg Asn Asp Met Ala
                      130                      135                      140
Ala Gln Tyr Lys Val Leu Gly Phe His Gly
145                      150

```

<210> 19

<211> 154

<212> PRT

<213> 欧亚野猪(*Sus scrofa*)

<400> 19

```

Met Gly Leu Ser Asp Gly Glu Trp Gln Leu Val Leu Asn Val Trp Gly
 1                      5                      10                      15
Lys Val Glu Ala Asp Val Ala Gly His Gly Gln Glu Val Leu Ile Arg
                      20                      25                      30
Leu Phe Lys Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys
                      35                      40                      45
His Leu Lys Ser Glu Asp Glu Met Lys Ala Ser Glu Asp Leu Lys Lys
 50                      55                      60
His Gly Asn Thr Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys
 65                      70                      75                      80
Gly His His Glu Ala Glu Leu Thr Pro Leu Ala Gln Ser His Ala Thr
                      85                      90                      95
Lys His Lys Ile Pro Val Lys Tyr Leu Glu Phe Ile Ser Glu Ala Ile
                      100                      105                      110
Ile Gln Val Leu Gln Ser Lys His Pro Gly Asp Phe Gly Ala Asp Ala
                      115                      120                      125
Gln Gly Ala Met Ser Lys Ala Leu Glu Leu Phe Arg Asn Asp Met Ala
                      130                      135                      140
Ala Lys Tyr Lys Glu Leu Gly Phe Gln Gly
145                      150

```

<210> 20

<211> 154

<212> PRT

<213> 马(*Equus caballus*)

图4(接上页)

<400> 20

```

Met Gly Leu Ser Asp Gly Glu Trp Gln Gln Val Leu Asn Val Trp Gly
 1          5          10          15
Lys Val Glu Ala Asp Ile Ala Gly His Gly Gln Glu Val Leu Ile Arg
          20          25          30
Leu Phe Thr Gly His Pro Glu Thr Leu Glu Lys Phe Asp Lys Phe Lys
          35          40          45
His Leu Lys Thr Glu Ala Glu Met Lys Ala Ser Glu Asp Leu Lys Lys
          50          55          60
His Gly Thr Val Val Leu Thr Ala Leu Gly Gly Ile Leu Lys Lys Lys
65          70          75          80
Gly His His Glu Ala Glu Leu Lys Pro Leu Ala Gln Ser His Ala Thr
          85          90          95
Lys His Lys Ile Pro Ile Lys Tyr Leu Glu Phe Ile Ser Asp Ala Ile
          100          105          110
Ile His Val Leu His Ser Lys His Pro Gly Asp Phe Gly Ala Asp Ala
          115          120          125
Gln Gly Ala Met Thr Lys Ala Leu Glu Leu Phe Arg Asn Asp Ile Ala
          130          135          140
Ala Lys Tyr Lys Glu Leu Gly Phe Gln Gly
145          150

```

<210> 21

<211> 152

<212> PRT

<213> 本氏烟草(*Nicotiana benthamiana*)

<400> 21

```

Met Ser Ser Phe Thr Glu Glu Gln Glu Ala Leu Val Val Lys Ser Trp
 1          5          10          15
Asp Ser Met Lys Lys Asn Ala Gly Glu Trp Gly Leu Lys Leu Phe Leu
          20          25          30
Lys Ile Phe Glu Ile Ala Pro Ser Ala Lys Lys Leu Phe Ser Phe Leu
          35          40          45
Lys Asp Ser Asn Val Pro Leu Glu Gln Asn Ala Lys Leu Lys Pro His
          50          55          60
Ser Lys Ser Val Phe Val Met Thr Cys Glu Ala Ala Val Gln Leu Arg
65          70          75          80
Lys Ala Gly Lys Val Val Val Arg Asp Ser Thr Leu Lys Lys Leu Gly
          85          90          95
Ala Thr His Phe Lys Tyr Gly Val Ala Asp Glu His Phe Glu Val Thr
          100          105          110
Lys Phe Ala Leu Leu Glu Thr Ile Lys Glu Ala Val Pro Glu Met Trp
          115          120          125
Ser Val Asp Met Lys Asn Ala Trp Gly Glu Ala Phe Asp Gln Leu Val
          130          135          140
Asn Ala Ile Lys Thr Glu Met Lys
145          150

```

图4(接上页)

<210> 22

<211> 132

<212> PRT

<213> 枯草芽孢杆菌(*Bacillus subtilis*)

<400> 22

```

Met Gly Gln Ser Phe Asn Ala Pro Tyr Glu Ala Ile Gly Glu Glu Leu
 1          5          10          15
Leu Ser Gln Leu Val Asp Thr Phe Tyr Glu Arg Val Ala Ser His Pro
          20          25          30
Leu Leu Lys Pro Ile Phe Pro Ser Asp Leu Thr Glu Thr Ala Arg Lys
          35          40          45
Gln Lys Gln Phe Leu Thr Gln Tyr Leu Gly Gly Pro Pro Leu Tyr Thr
          50          55          60
Glu Glu His Gly His Pro Met Leu Arg Ala Arg His Leu Pro Phe Pro
65          70          75          80
Ile Thr Asn Glu Arg Ala Asp Ala Trp Leu Ser Cys Met Lys Asp Ala
          85          90          95
Met Asp His Val Gly Leu Glu Gly Glu Ile Arg Glu Phe Leu Phe Gly
          100          105          110
Arg Leu Glu Leu Thr Ala Arg His Met Val Asn Gln Thr Glu Ala Glu
          115          120          125
Asp Arg Ser Ser
          130

```

<210> 23

<211> 131

<212> PRT

<213> 谷氨酸棒状杆菌(*Corynebacterium glutamicum*)

<400> 23

```

Met Thr Thr Ser Glu Asn Phe Tyr Asp Ser Val Gly Gly Glu Glu Thr
 1          5          10          15
Phe Ser Leu Ile Val His Arg Phe Tyr Glu Gln Val Pro Asn Asp Asp
          20          25          30
Ile Leu Gly Pro Met Tyr Pro Pro Asp Asp Phe Glu Gly Ala Glu Gln
          35          40          45
Arg Leu Lys Met Phe Leu Ser Gln Tyr Trp Gly Gly Pro Lys Asp Tyr
          50          55          60
Gln Glu Gln Arg Gly His Pro Arg Leu Arg Met Arg His Val Asn Tyr
65          70          75          80
Pro Ile Gly Val Thr Ala Ala Glu Arg Trp Leu Gln Leu Met Ser Asn
          85          90          95
Ala Leu Asp Gly Val Asp Leu Thr Ala Glu Gln Arg Glu Ala Ile Trp
          100          105          110
Glu His Met Val Arg Ala Ala Asp Met Leu Ile Asn Ser Asn Pro Asp

```

图4(接上页)

```

          115                      120                      125
Pro His Ala
      130

<210> 24
<211> 124
<212> PRT
<213> 集胞藻属(Synechocystis sp.)

<400> 24
Met Ser Thr Leu Tyr Glu Lys Leu Gly Gly Thr Thr Ala Val Asp Leu
 1          5          10          15
Ala Val Asp Lys Phe Tyr Glu Arg Val Leu Gln Asp Asp Arg Ile Lys
      20          25          30
His Phe Phe Ala Asp Val Asp Met Ala Lys Gln Arg Ala His Gln Lys
      35          40          45
Ala Phe Leu Thr Tyr Ala Phe Gly Gly Thr Asp Lys Tyr Asp Gly Arg
      50          55          60
Tyr Met Arg Glu Ala His Lys Glu Leu Val Glu Asn His Gly Leu Asn
65          70          75          80
Gly Glu His Phe Asp Ala Val Ala Glu Asp Leu Leu Ala Thr Leu Lys
      85          90          95
Glu Met Gly Val Pro Glu Asp Leu Ile Ala Glu Val Ala Ala Val Ala
      100          105          110
Gly Ala Pro Ala His Lys Arg Asp Val Leu Asn Gln
      115          120

<210> 25
<211> 183
<212> PRT
<213> 聚球藻属(Synechococcus sp.)

<400> 25
Met Asp Val Ala Leu Leu Glu Lys Ser Phe Glu Gln Ile Ser Pro Arg
 1          5          10          15
Ala Ile Glu Phe Ser Ala Ser Phe Tyr Gln Asn Leu Phe His His His
      20          25          30
Pro Glu Leu Lys Pro Leu Phe Ala Glu Thr Ser Gln Thr Ile Gln Glu
      35          40          45
Lys Lys Leu Ile Phe Ser Leu Ala Ala Ile Ile Glu Asn Leu Arg Asn
      50          55          60
Pro Asp Ile Leu Gln Pro Ala Leu Lys Ser Leu Gly Ala Arg His Ala
65          70          75          80
Glu Val Gly Thr Ile Lys Ser His Tyr Pro Leu Val Gly Gln Ala Leu
      85          90          95
Ile Glu Thr Phe Ala Glu Tyr Leu Ala Ala Asp Trp Thr Glu Gln Leu
      100          105          110

```

图4(接上页)

```

Ala Thr Ala Trp Val Glu Ala Tyr Asp Val Ile Ala Ser Thr Met Ile
      115                      120                      125
Glu Gly Ala Asp Asn Pro Ala Ala Tyr Leu Glu Pro Glu Leu Thr Phe
      130                      135                      140
Tyr Glu Trp Leu Asp Leu Tyr Gly Glu Glu Ser Pro Lys Val Arg Asn
145                      150                      155                      160
Ala Ile Ala Thr Leu Thr His Phe His Tyr Gly Glu Asp Pro Gln Asp
                      165                      170                      175
Val Gln Arg Asp Ser Arg Gly
                      180

```

<210> 26

<211> 118

<212> PRT

<213> 地木耳(*Nostoc commune*)

<400> 26

```

Met Ser Thr Leu Tyr Asp Asn Ile Gly Gly Gln Pro Ala Ile Glu Gln
 1      5      10      15
Val Val Asp Glu Leu His Lys Arg Ile Ala Thr Asp Ser Leu Leu Ala
      20      25      30
Pro Val Phe Ala Gly Thr Asp Met Val Lys Gln Arg Asn His Leu Val
      35      40      45
Ala Phe Leu Ala Gln Ile Phe Glu Gly Pro Lys Gln Tyr Gly Gly Arg
 50      55      60
Pro Met Asp Lys Thr His Ala Gly Leu Asn Leu Gln Gln Pro His Phe
65      70      75      80
Asp Ala Ile Ala Lys His Leu Gly Glu Arg Met Ala Val Arg Gly Val
      85      90      95
Ser Ala Glu Asn Thr Lys Ala Ala Leu Asp Arg Val Thr Asn Met Lys
      100      105      110
Gly Ala Ile Leu Asn Lys
      115

```

<210> 27

<211> 136

<212> PRT

<213> 巨大芽孢杆菌(*Bacillus megaterium*)

<400> 27

```

Met Arg Glu Lys Ile His Ser Pro Tyr Glu Leu Leu Gly Gly Glu His
 1      5      10      15
Thr Ile Ser Lys Leu Val Asp Ala Phe Tyr Thr Arg Val Gly Gln His
      20      25      30
Pro Glu Leu Ala Pro Ile Phe Pro Asp Asn Leu Thr Glu Thr Ala Arg
      35      40      45
Lys Gln Lys Gln Phe Leu Thr Gln Tyr Leu Gly Gly Pro Ser Leu Tyr

```

图4(接上页)

50		55		60	
Thr	Glu	Glu	His	Gly	His
65					Pro
Glu	Ile	Thr	Pro	Ser	Arg
					Ala
Ala	Met	Asp	Glu	Ile	Asn
					Leu
His	Arg	Leu	Ile	Leu	Thr
					Ala
Thr	Asp	Glu	Lys	Gly	Phe
130					

图4(接上页)

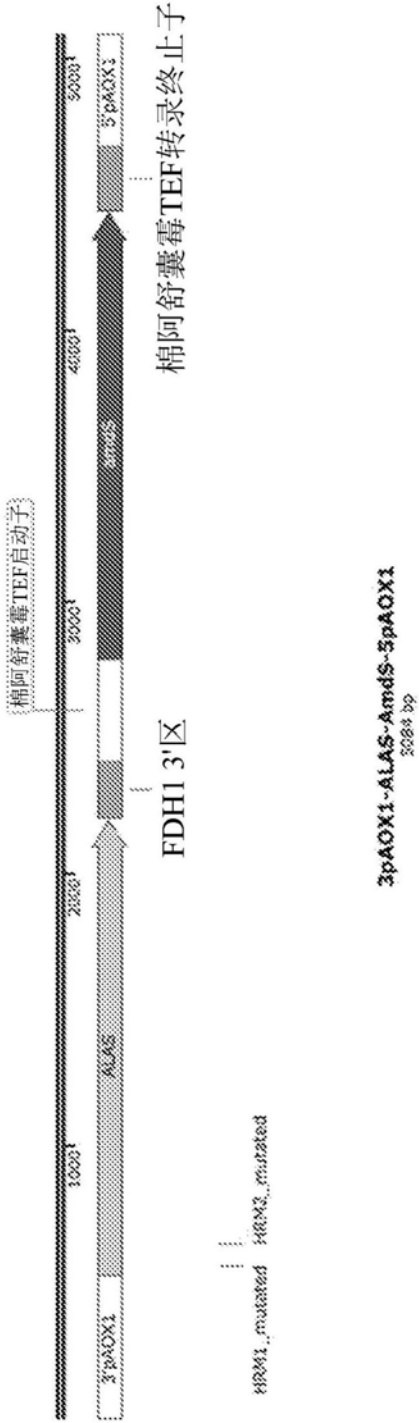


图5

SEQ ID NO: 28

```

atggagtttg tgcgccgtca gtccatgaat gectgtccct ttgtcaggtc aacttctacc      60
caccatttga agaagttggc agcaaacagt tctctagctg ctactgctag tcattgtccc      120
gtggttggtc ctgctctcca acagcagaga tactactctc aaccttccaa gccagcccaa      180
gccccaaact ccgacattgc tactgggatc aagaaggatg tttctccgat ccgtatggac      240
tctaataaaa ccgcctttga ttacaatgga atgtatgagt ctgatcttgc gaataaacgt      300
aaagataact cgtatcgtta tttcaataac atcaaccgtc tagccaagga gtttcccaag      360
gcacatcgcc agaccgaaga tgacaagggt accgtctggt gctctaacga ctacttagga      420
atgggtaggc atcctgagat tatcaaaacc atgaaggcta ccattggacaa gtacgggttcc      480
ggagcaggag gaactaggaa cattgcaggt cataaccacg ccgctatcaa tttggaaaagc      540
gagttggctt gcttgaacaa gaaggaagcg gctctggtgt tttcatcatg tttcatagct      600
aacgatgcaa tcattctcgtt gttgggacaa aaaatcaaaa atttggtcat tttctctgac      660
cagtcgaatc atgcttccat gatattgggt gtgcgtaact ccaaagcgaa gaagcacatc      720
ttcaagcaca acaatttgaa ggatctggag tcgcagttag ctacgtacc ccaagtcgact      780
cctaaactga tgcctctcga gtcagtttac tctatgtgtg gatctgtggc tcccattgag      840
aagatttgcg atttggctaa aaggtacggt gccctcacct tcttgatga agttcatgct      900
gttggaatgt atggtcctca tggacagggt gtagctgagc atttggactt tgatctgcat      960
ttacagtctg gaatcgccag tctagcgtg gtggacaaac gcaccatatt ggatcgtgtc      1020
gacatgatta ctggtacttg cggaaagtca tttggtactg ttggagggtta cgttgcgtgt      1080
agtgccaaac taattgattg gtttaagatc tatgcgccag gtttcatttt cactaccaca      1140
cttcctctct ctatcatggc tggtaacagc acttctgttc gtattgttag ggccgacatt      1200
gaggcccgta tcaagcaaca gcttaatact cgctacgtca aagactcatt tgaaaacctt      1260
ggtattccag tcattccaaa cccaagtcac atgttccctg ttctagttgg aaatgctgca      1320
gatgccaaag aggcattccg tatgttaatg aacaaacacc gtatttatgt tcaagctatt      1380
aactacccta ctgtgctgtg cgggtgaagaa cgactaagga ttactcctac tccaggtcac      1440
ggaaaggaga tttgtgacca gctgatcagc gctgtcgacg atgtttttac tgagcttaat      1500
ttaccaagaa tcaacaaatg gcagtcacca ggtggtcatt gcggtgttgg tgatgctaata      1560
tacgtaccag aacccaatct gtggactcag gaccagctca gcttgacaaa ccaagacttg      1620
cactccaatg tgcacaacc cagtattgag cagatcgaaa cctcatcagg agtcagattg      1680
tag

```

SEQ ID NO: 29

10	20	30	40	50
MEFVARQSMN	ACPFVRSTST	HHLKKLAANS	SLAATASHCP	VVGPALQQQR
60	70	80	90	100
YYSQPSKPAQ	AQTSDIATGI	KKDVSPIRMD	SNETAFDYNG	MYESDLANKR
110	120	130	140	150
KDNSYRYFNN	INRLAKEFPK	AHRQTEDDKV	TVWCSNDYLG	MGRHPEIIKT
160	170	180	190	200
MKATMDKYGS	GAGGTRNIAG	HNHAAINLES	ELACLNKKEA	ALVFSSCFIA
210	220	230	240	250
NDAIISLLGQ	KIKNLVIFSD	QSNHASMILG	VRNSKAKKHI	FKHNNLKDLE
260	270	280	290	300
SQLAQYPKST	PKLIAFESVY	SMCGSVAPIE	KICDLAKRYG	ALTFLDEVHA
310	320	330	340	350
VGMYGPHGQG	VAEHLDFDLH	LQSGIASPSV	VDKRTILDRV	DMITGTCGKS
360	370	380	390	400
FGTVGGYVAG	SANLIDWLRS	YAPGFIFTTT	LPPAIMAGTA	TSVRIVRADI
410	420	430	440	450

图6

EARIKQQLNT	RYVKDSFENL	GIPVIPNPSH	IVPVLVGNA	DAKKASDMLM
460	470	480	490	500
NKHRIYVQAI	NYPTVPVGEE	RLRITPTPGH	GKEICDQLIS	AVDDVFTELN
510	520	530	540	550
LPRINKWQSQ	GGHCGVGDN	YVPEPNLWTQ	DQLSLTNQDL	HSNVHNPVIE
560				
QIETSSGVRL				

图6(接上页)

示例性突变型ALAS基因和蛋白质的序列

基因序列(SEQ ID NO: 30):

ATGGAGTTTGTGCGCCGTCAGTCCATGAATGCCTCTCCCTTTGTCAGGTCAACTTCTACCCACC
ATTTGAAGAAGTTGGCAGCAAACAGTTCTCTAGCTGCTACTGCTAGTCATTCTCCCGTGGTTGG
CCCTGCTCTCCAACAGCAGAGATACTACTCTCAACCTTCCAAGCCAGCCCAAGCCCAAACCTCC
GACATTGCTACTGGGATCAAGAAGGATGTTTCTCCGATCCGTATGGACTCTAATGAAACCGCCT
TTGATTACAATGGAATGTATGAGTCTGATCTTGCGAATAAACGTAAAGATAACTCGTATCGTTA
TTTCAATAACATCAACCGTCTAGCCAAGGAGTTTCCCAAGGCACATCGCCAGACCGAAGATGAC
AAGGTGACCGTCTGGTGCTCTAACGACTACTTAGGAATGGGTAGGCATCCTGAGATTATCAAAA
CCATGAAGGCTACCATGGACAAGTACGGTTCCGGAGCAGGAGGAAGTAGGAACATTGCAGGTCA
TAACCACGCCGCTATCAATTTGGAAGCGAGTTGGCTTGCTTGAACAAGAAGGAAGCGGCTCTG
GTGTTTTTCATCATGTTTCATAGCTAACGATGCAATCATCTCGTTGTTGGGACAAAAAATCAAAA
ATTTGGTCATTTTCTCTGACCAGTCGAATCATGCTTCCATGATATTGGGTGTGCGTAACTCCAA
AGCGAAGAAGCACATCTTCAAGCACACAATTTGAAGGATCTGGAGTCGCAGTTAGCTCAGTAC
CCCAAGTCGACTCCTAAACTGATCGCCTTCGAGTCAGTTTACTCTATGTGTGGATCTGTGGCTC
CCATTGAGAAGATTTGCGATTTGGCTAAAAGGTACGGTGCCCTCACCTTCTTGATGAAGTTCA
TGCTGTTGGAATGTATGGTCCTCATGGACAGGGTGTAGCTGAGCATTTGGACTTTGATCTGCAT
TTACAGTCTGGAATCGCCAGTCCTAGCGTGGTGGACAAACGCACCATATTGGATCGTGTGCGACA
TGATTACTGGTACTTGCGGAAAGTCATTTGGTACTGTTGGAGGTTACGTTGCTGGTAGTGCCAA
CCTAATTGATTGGTTAAGATCCTATGCGCCAGGTTTCATTTTCACTACCACACTTCCTCCTGCT
ATCATGGCTGGTACAGCCACTTCTGTTCGTATTGTTAGGGCCGACATTGAGGCCCGTATCAAGC
AACAGCTTAATACTCGCTACGTCAAAGACTCATTTGAAAACCTTGGTATTCCAGTCATTCCAAA
CCCAAGTCACATTGTTCCCTGTTCTAGTTGGAAATGCTGCAGATGCCAAGAAGGCATCCGATATG
TTAATGAACAAACACCGTATTTATGTTCAAGCTATTAACCTACTGTGCCTGTCGGTGAAG
AACGACTAAGGATTACTCCTACTCCAGGTCATGGAAAGGAGATTTGTGACCAGCTGATCAGCGC
TGTCGACGATGTTTTTACTGAGCTTAATTTACCAAGAATCAACAAATGGCAGTCCCAAGGTGGT
CATTGCGGTGTTGGTGATGCTAATTACGTACCAGAACCAATCTGTGGACTCAGGACCAGCTCA
GCTTGACAAACCAAGACTTGCACTCCAATGTGCACAACCCAGTGATTGAGCAGATCGAAACCTC
ATCAGGAGTCAGATTGTAG

图7

突变的密码子(野生型中的TGT到突变型中的TCT)以粗体和双下划线指示。

蛋白质序列(SEQ ID NO: 31):

MEFVARQSMNA**S**PFVIRSTSTHHLKKLAANSSLAATASH**S**PVVGPALQQORYYSQPSKPAQAQTS
DIATGIKKDVSPIRMDSNETAFDYNGMYESDLANKRKDNSYRYFNNINRLAKEFPKAHRQTEDD
KVTVWCSNDYLGMRHPEIIKTMKATMDKYGSGAGGTRNIAGHNHAAINLESELACLNKKEAAL
VFSSCFIANDAIISLLGQKIKNLVIFSDQSNHASMILGVRNSKAKKHIFKHNNLKDLESQLAQY
PKSTPKLIAFESVYSMCGSVAPIEKICDLAKRYGALTFLDEVHAVGMYGPHGQGVAEHLDFDLH
LQSGIASPSVVDKRTILDRVDMITGTCGKSFGTVGGYVAGSANLIDWLRSYAPGFIFTTTLPPA
IMAGTATSVRIVRADIEARIKQQLNTRYVKDSFENLGIPVIPNPSHIVPVLVGNAADAKKASDM
LMNKHRIYVQAINYPTVPVGEERLRITPTPGHGKEICDQLISAVDDVFTELNLPRINKWQSQGG
HCGVGDANYVPEPNLWTQDQLSLTNQDLHSNVHNPVIEQIETSSGVRL

长度: 560 aa

分子量: 61.5 kDa

wtALAS中存在的两个半胱氨酸残基突变为丝氨酸(粗体和双下划线):
C12S和C39S。

图7(接上页)

SEQ ID NO: 32
R/L/N/A/C/S/H/I/G/Q-C-P-L/V/I/F/C-L/M/P/V

SEQ ID NO: 33
A-C-P-F-V

SEQ ID NO: 34
H-C-P-V-V

SEQ ID NO: 35
I-C-P-F-M

SEQ ID NO: 36
G-C-P-V-V

SEQ ID NO: 37
CGTCAGTCCATGAATGCCTCTCCCTTTGTCAGGTCAACTTC

SEQ ID NO: 38
GAAGTTGACCTGACAAAGGGAGAGGCATTCATGGACTGACG

SEQ ID NO: 39
GCTGCTACTGCTAGTCATTCTCCCGTGGTTGGCCCTG

SEQ ID NO: 40
CAGGGCCAACCACGGGAGAATGACTAGCAGTAGCAGC

SEQ ID NO: 41
AAACGCTGTCTTGGAACCTAATATGAC

SEQ ID NO: 42
GACGGGCGACAACTCCATCGTTTCGAATAATTAGTTG

SEQ ID NO: 43
CAACTAATTATTCGAAACGATGGAGTTTGTCGCCCCGTCAG

SEQ ID NO: 44
AATTAAATACATTTCAACTACAATCTGACTCCTGATGAGGTTTCG

SEQ ID NO: 45
CCTCATCAGGAGTCAGATTGTAGTTGAAATGTATTTAATTG

SEQ ID NO: 46
AAACTGTCAGTTTTGGGCCATTTG

图8

SEQ ID NO: 47
GCTCTCCAACAGCAGAGATAC

SEQ ID NO: 48
GTCCATACGGATCGGAGAAAC

SEQ ID NO: 49
AGCAACATCCCTGATTCCG

SEQ ID NO: 50
ATGCGTACCTTCAATCCTGG

SEQ ID NO: 51
6-FAM/AAGCCCAAA/Zen/CCTCCGA CATTGCTA/3IABkFQ

SEQ ID NO: 52
HEX/TCGCCGTAA/Zen/GTTCTTGGTT TAGACGTTC/3IABkFQ

SEQ ID NO: 53
TAGCGCAGTCTCTCTATCGCTTC

SEQ ID NO: 54
CACTGGGTTGTGCACATTGG

SEQ ID NO: 55
ACAATATTCTTCTCTGCCGC

SEQ ID NO: 56
TTGATCTCGTCAAGAATGCG

SEQ ID NO: 57
TAGGTGCCACAACCTTTTGGTTTC

SEQ ID NO: 58
GATCCAATGCGATGACATTCTTGT

SEQ ID NO: 59
ACCTGCAATAACTCCTCTTCTCTG

SEQ ID NO: 60
CCACTGAGGGTAGCCGAATC

SEQ ID NO: 61
GGGCTCTGAAAACTCTTTTGG

图8(接上页)

SEQ ID NO: 62
GCATGTCTCAATAACAGATCTCGACGG

SEQ ID NO: 63
AAGCCTCTTGTTTTTCTGTAAATGCAC

SEQ ID NO: 64
TGATGGCGTCCGAGATGAACTC

SEQ ID NO: 65
ATGGAGTTTGTCGCCCCGTCAG

SEQ ID NO: 66
CTACAATCTGACTCCTGATGAGGTTTC

图8(接上页)