



(12) 发明专利申请

(10) 申请公布号 CN 115190912 A

(43) 申请公布日 2022. 10. 14

(21) 申请号 202080097713.8

M • R • 拉科尔 T • D • 埃里驰

(22) 申请日 2020.12.28

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

(30) 优先权数据

专利代理师 张小勇

62/955,014 2019.12.30 US

63/058,169 2020.07.29 US

(85) PCT国际申请进入国家阶段日
2022.08.29

(51) Int.Cl.

C12N 15/10 (2006.01)

C12N 9/22 (2006.01)

C12N 15/63 (2006.01)

(86) PCT国际申请的申请数据

PCT/US2020/067138 2020.12.28

(87) PCT国际申请的公布数据

W02021/138247 EN 2021.07.08

(71) 申请人 生命编辑制药股份有限公司

地址 美国北卡罗来纳

(72) 发明人 A • B • 克拉维利 P • 博登

T • D • 博文 M • 寇伊勒

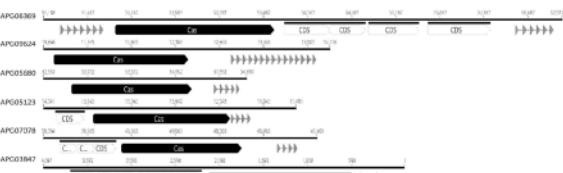
权利要求书23页 说明书100页
序列表245页 附图1页

(54) 发明名称

RNA指导核酸酶及其活性片段与变体以及使用
方法

(57) 摘要

本发明提供了用于结合至所关注的目标序列的组合物及方法。该组合物可用于切割或修饰所关注的目标序列、检测所关注的目标序列和修饰所关注的序列的表达。组合物包含RNA指导核酸酶多肽、CRISPR RNA、反式活化的CRISPR RNA、指导RNA和编码其的核酸分子。还提供了包含该核酸分子的载体及宿主细胞。还提供了用于结合所关注的目标序列的CRISPR系统,其中该CRISPR系统包含RNA指导核酸酶多肽及一种或多种指导RNA。



1. 一种核酸分子, 包含编码RNA指导核酸酶 (RGN) 多肽的多核苷酸, 其中所述多核苷酸包含编码RGN多肽的核苷酸序列, 所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列;

其中当与能够与DNA分子的目标DNA序列杂交的指导RNA (gRNA) 结合时, 所述RGN多肽能够以RNA指导的序列特异性方式结合所述目标DNA序列, 且

其中编码RGN多肽的所述多核苷酸可操作地连接至与所述多核苷酸异源的启动子。

2. 如权利要求1所述的核酸分子, 其中所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

3. 如权利要求1所述的核酸分子, 其中所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。

4. 如权利要求1至3中任一项所述的核酸分子, 其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

5. 如权利要求4所述的核酸分子, 其中所述RGN多肽能够在结合时切割所述目标DNA序列。

6. 如权利要求1至3中任一项所述的核酸分子, 其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

7. 如权利要求6所述的核酸分子, 其中所述RGN多肽能够在结合时切割所述目标DNA序列。

8. 如权利要求7所述的核酸分子, 其中通过所述RGN多肽的切割产生双链断裂。

9. 如权利要求7所述的核酸分子, 其中通过所述RGN多肽的切割产生单链断裂。

10. 如权利要求6至9中任一项所述的核酸分子, 其中所述RGN多肽与碱基编辑多肽可操作地融合。

11. 如权利要求10所述的核酸分子, 其中所述碱基编辑多肽为脱氨酶。

12. 如权利要求6至11中任一项所述的核酸分子, 其中所述目标DNA序列位于与前间隔序列邻近基序 (PAM) 相邻的位置。

13. 如权利要求1至12中任一项所述的核酸分子, 其中所述RGN多肽包含一种或多种核定位信号。

14. 如权利要求1至13中任一项所述的核酸分子, 其中所述RGN多肽针对在真核细胞中的表达而被密码子优化。

15. 一种包含如权利要求1至14中任一项所述的核酸分子的载体。

16. 如权利要求15所述的载体, 还包含至少一个核苷酸序列, 所述至少一个核苷酸序列编码能够与所述目标DNA序列杂交的所述gRNA。

17. 如权利要求16所述的载体, 其中所述RGN多肽包含与SEQ ID NO: 11具有至少90%序列同一性的氨基酸序列, 且所述gRNA包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 116具有至少90%序列同一性的CRISPR重复序列。

18. 如权利要求16所述的载体, 其中所述RGN多肽包含与SEQ ID NO: 11具有至少95%序列同一性的氨基酸序列, 且所述gRNA包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 116具有至少95%序列同一性的CRISPR重复序列。

19. 如权利要求16所述的载体, 其中所述RGN多肽包含与SEQ ID NO: 11具有100%序列

同一性的氨基酸序列,且所述gRNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

20.如权利要求16所述的载体,其中所述gRNA包含tracrRNA。

21.如权利要求20所述的载体,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;
和

i)与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

22.如权利要求20所述的载体,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;
和

i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

23. 如权利要求20所述的载体,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR

RNA,所述CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中所述gRNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

24.如权利要求20至23中任一项所述的载体,其中所述gRNA为单指导RNA。

25.如权利要求20至23中任一项所述的载体,其中所述gRNA为双指导RNA。

26.一种包含如权利要求1至14中任一项所述的核酸分子或如权利要求15至25中任一项所述的载体的细胞。

27.一种制备RGN多肽的方法,包含在所述RGN多肽被表达的条件下,培养如权利要求26所述的细胞。

28.一种制备RGN多肽的方法,包含将异源核酸分子引入细胞内,所述异源核酸分子包含编码RNA指导核酸酶(RGN)多肽的核苷酸序列,所述RNA指导核酸酶(RGN)多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列;

其中当与能够与DNA分子的目标DNA序列杂交的指导RNA(gRNA)结合时,所述RGN多肽以RNA指导的序列特异性方式结合所述目标DNA序列;

且在所述RGN多肽被表达的条件下,培养所述细胞。

29.如权利要求28所述的方法,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

30.如权利要求28所述的方法,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。

31.如权利要求28至30中任一项所述的方法,还包括纯化所述RGN多肽。

32.如权利要求28至30中任一项所述的方法,其中所述细胞还表达一种或多种指导RNA,其与所述RGN多肽结合,以形成RGN核糖核蛋白复合物。

33.如权利要求32所述的方法,还包括纯化所述RGN核糖核蛋白复合物。

34.一种分离的RNA指导核酸酶(RGN)多肽,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列;和

其中当与能够与DNA分子的目标DNA序列杂交的指导RNA(gRNA)结合时,所述RGN多肽能够以RNA指导的序列特异性方式结合所述目标DNA序列。

35.如权利要求34所述的分离的RGN多肽,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

36.如权利要求34所述的分离的RGN多肽,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。

37. 如权利要求34至36中任一项所述的分离的RGN多肽,其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

38. 如权利要求37所述的分离的RGN多肽,其中所述RGN多肽能够在结合时切割所述目标DNA序列。

39. 如权利要求34至36中任一项所述的分离的RGN多肽,其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

40. 如权利要求39所述的分离的RGN多肽,其中所述RGN多肽能够在结合时切割所述目标DNA序列。

41. 如权利要求40所述的分离的RGN多肽,其中通过所述RGN多肽的切割产生双链断裂。

42. 如权利要求40所述的分离的RGN多肽,其中通过所述RGN多肽的切割产生单链断裂。

43. 如权利要求39至42中任一项所述的分离的RGN多肽,其中所述RGN多肽与碱基编辑多肽可操作地融合。

44. 如权利要求43所述的分离的RGN多肽,其中所述碱基编辑多肽为脱氨酶。

45. 如权利要求34至44中任一项所述的分离的RGN多肽,其中所述目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

46. 如权利要求34至45中任一项所述的分离的RGN多肽,其中所述RGN多肽包含一种或多种核定位信号。

47. 一种用于结合DNA分子的目标DNA序列的系统,所述系统包含:

a) 能够与所述目标DNA分子杂交的一种或多种指导RNA,或包含编码所述一种或多种指导RNA(gRNA)的一种或多种核苷酸序列的一种或多种多核苷酸;和

b) 包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽,或包含编码所述RGN多肽的核苷酸序列的多核苷酸;

其中编码所述一种或多种指导RNA的至少一个所述核苷酸序列及编码所述RGN多肽的所述核苷酸序列可操作地连接至与所述核苷酸序列异源的启动子;且

其中所述一种或多种指导RNA能够与所述RGN多肽形成复合物,以便将所述RGN多肽导向至与所述DNA分子的所述目标DNA序列结合。

48. 一种用于结合DNA分子的目标DNA序列的系统,所述系统包含:

a) 能够与所述目标DNA分子杂交的一种或多种指导RNA,或包含编码所述一种或多种指导RNA(gRNA)的一种或多种核苷酸序列的一种或多种多核苷酸;和

b) 包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽;

其中所述一种或多种指导RNA能够与所述目标DNA序列杂交,且

其中所述一种或多种指导RNA能够与所述RGN多肽形成复合物,以便将所述RGN多肽导向至与所述DNA分子的所述目标DNA序列结合。

49. 如权利要求48所述的系统,其中编码所述一种或多种指导RNA的至少一个所述核苷酸序列可操作地连接至与所述核苷酸序列异源的启动子。

50. 如权利要求47至49中任一项所述的系统,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

51. 如权利要求47至49中任一项所述的系统,其中所述RGN多肽包含与SEQ ID NO:2、4、

1、3和5至109中任一具有100%序列同一性的氨基酸序列。

52. 如权利要求47至51中任一项所述的系统,其中未发现所述RGN多肽与所述一种或多种指导RNA在自然界中彼此复合。

53. 如权利要求47至51中任一项所述的系统,其中所述目标DNA序列为真核目标DNA序列。

54. 如权利要求47至53中任一项所述的系统,其中所述RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

55. 如权利要求47至53中任一项所述的系统,其中所述RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

56. 如权利要求47至53中任一项所述的系统,其中所述RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

57. 如权利要求47至53中任一项所述的系统,其中所述一种或多种指导RNA包含tracrRNA。

58. 如权利要求57所述的系统,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的

CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

59. 如权利要求57所述的系统,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

60. 如权利要求57所述的系统,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA

还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

61.如权利要求57至60中任一项所述的系统,其中所述一种或多种指导RNA为单指导RNA (sgRNA)。

62.如权利要求57至60中任一项所述的系统,其中所述一种或多种指导RNA为双指导RNA。

63.如权利要求47至62中任一项所述的系统,其中所述目标DNA序列位于细胞内。

64.如权利要求63所述的系统,其中所述细胞为真核细胞。

65.如权利要求64所述的系统,其中所述真核细胞为植物细胞。

66.如权利要求64所述的系统,其中所述真核细胞为哺乳动物细胞。

67.如权利要求64所述的系统,其中所述真核细胞为昆虫细胞。

68.如权利要求63所述的系统,其中所述细胞为原核细胞。

69.如权利要求47至68中任一项所述的系统,其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

70.如权利要求69所述的系统,其中当被转录时,所述一种或多种指导RNA能够与所述目标DNA序列杂交,且所述指导RNA能够与所述RGN多肽形成复合物,以导向所述目标DNA序列的切割。

71.如权利要求47至68中任一项所述的系统,其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

72.如权利要求71所述的系统,其中当被转录时,所述一种或多种指导RNA能够与所述目标DNA序列杂交,且所述指导RNA能够与所述RGN多肽形成复合物,以导向所述目标DNA序列的切割。

73.如权利要求72所述的系统,其中所述RGN多肽能够产生双链断裂。

74.如权利要求72所述的系统,其中所述RGN多肽能够产生单链断裂。

75.如权利要求71至74中任一项所述的系统,其中所述RGN多肽可操作地连接至碱基编辑多肽。

76.如权利要求75所述的系统,其中所述碱基编辑多肽为脱氨酶。

77.如权利要求71至76中任一项所述的系统,其中所述目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

78.如权利要求47至77中任一项所述的系统,其中所述RGN多肽包含一种或多种核定位信号。

79.如权利要求47至78中任一项所述的系统,其中所述RGN多肽针对在真核细胞中的表达而被密码子优化。

80.如权利要求47至79中任一项所述的系统,其中包含编码所述一种或多种指导RNA的所述核苷酸序列的多核苷酸及包含编码RGN多肽的所述核苷酸序列的所述多核苷酸位于一个载体上。

81.如权利要求47至80中任一项所述的系统,其中所述系统还包含一种或多种供体多核苷酸或包含编码所述一种或多种供体多核苷酸的一种或多种核苷酸序列的一种或多种多核苷酸。

82. 一种药物组合物, 包含如权利要求1至14中任一项所述的核酸分子、如权利要求15至25中任一项所述的载体、如权利要求26所述的细胞、如权利要求34至46中任一项所述的分离的RGN多肽、或如权利要求47至81中任一项所述的系统和药学上可接受的载剂。

83. 一种结合DNA分子的目标DNA序列的方法, 包含将如权利要求47至81中任一项所述的系统递送至所述目标DNA序列或包含所述目标DNA序列的细胞。

84. 如权利要求83所述的方法, 其中所述RGN多肽或所述指导RNA还包含可检测标记, 从而允许对所述目标DNA序列的检测。

85. 如权利要求83所述的方法, 其中所述指导RNA或所述RGN多肽还包含表达调控子, 从而调控所述目标DNA序列的或受所述目标DNA序列的转录控制下的基因的表达。

86. 一种切割DNA分子的目标DNA序列的方法, 包含将如权利要求47至81中任一项所述的系统递送至所述目标DNA序列或包含所述目标DNA序列的细胞。

87. 如权利要求86所述的方法, 其中经修饰的目标DNA序列包含异源DNA在所述目标DNA序列内的插入。

88. 如权利要求86所述的方法, 其中经修饰的目标DNA序列包含至少一个核苷酸自所述目标DNA序列的缺失。

89. 如权利要求86所述的方法, 其中经修饰的目标DNA序列包含所述目标DNA序列中的至少一个核苷酸的突变。

90. 一种用于结合DNA分子的目标DNA序列的方法, 所述方法包括:

a) 通过组合以下, 以在体外将RNA指导核酸酶 (RGN) 核糖核苷酸复合物在适合形成所述RGN核糖核苷酸复合物的条件下进行组装:

i) 能够与所述目标DNA序列杂交的一种或多种指导RNA; 和

ii) RGN多肽, 所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列; 和

b) 使所述目标DNA序列或包含所述目标DNA序列的细胞与所述在体外组装的RGN核糖核苷酸复合物接触;

其中所述一种或多种指导RNA与所述目标DNA序列杂交, 从而将所述RGN多肽导向至与所述目标DNA序列结合。

91. 如权利要求90所述的方法, 其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

92. 如权利要求90所述的方法, 其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

93. 如权利要求92所述的方法, 其中所述目标DNA序列位于与前间隔序列邻近基序 (PAM) 相邻的位置。

94. 如权利要求90至93中任一项所述的方法, 其中所述RGN多肽或所述指导RNA还包含可检测标记, 从而允许对所述目标DNA序列的检测。

95. 如权利要求90至93中任一项所述的方法, 其中所述指导RNA或所述RGN多肽还包含表达调控子, 从而允许对所述目标DNA序列的表达的调控。

96. 一种用于切割和/或修饰DNA分子的目标DNA序列的方法, 包含使所述DNA分子与以下接触:

a) RNA指导核酸酶(RGN)多肽,其中所述RGN包含与SEQ ID NO:2、4、1、3和5至109中任一具有至少90%序列同一性的氨基酸序列;和

b) 能够将(a)的所述RGN靶向至所述目标DNA序列的一种或多种指导RNA;

其中所述一种或多种指导RNA与所述目标DNA序列杂交,从而将所述RGN多肽导向至与所述目标DNA序列结合,并且发生所述目标DNA序列的切割和/或修饰。

97. 如权利要求96所述的方法,其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

98. 如权利要求96所述的方法,其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

99. 如权利要求98所述的方法,其中通过所述RGN多肽的切割产生双链断裂。

100. 如权利要求98所述的方法,其中通过所述RGN多肽的切割产生单链断裂。

101. 如权利要求98至100中任一项所述的方法,其中所述RGN多肽可操作地连接至碱基编辑多肽。

102. 如权利要求101所述的方法,其中所述碱基编辑多肽为脱氨酶。

103. 如权利要求98至102中任一项所述的方法,其中所述目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

104. 如权利要求96至103中任一项所述的方法,其中经修饰的目标DNA序列包含异源DNA在所述目标DNA序列内的插入。

105. 如权利要求96至103中任一项所述的方法,其中经修饰的目标DNA序列包含至少一个核苷酸自所述目标DNA序列的缺失。

106. 如权利要求96至103中任一项所述的方法,其中经修饰的目标DNA序列包含所述目标DNA序列中的至少一个核苷酸的突变。

107. 如权利要求90至106中任一项所述的方法,其中所述RGN包含与SEQ ID NO:2、4、1、3和5至109中任一具有至少95%序列同一性的氨基酸序列。

108. 如权利要求90至106中任一项所述的方法,其中所述RGN包含与SEQ ID NO:2、4、1、3和5至109中任一具有100%序列同一性的氨基酸序列。

109. 如权利要求90至108中任一项所述的方法,其中未发现所述RGN多肽与所述一种或多种指导RNA在自然界中彼此复合。

110. 如权利要求90至109中任一项所述的方法,其中所述目标DNA序列为真核目标DNA序列。

111. 如权利要求90至110中任一项所述的方法,其中所述RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

112. 如权利要求90至110中任一项所述的方法,其中所述RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

113. 如权利要求90至110中任一项所述的方法,其中所述RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

114. 如权利要求90至110中任一项所述的方法, 其中所述一种或多种指导RNA包含tracrRNA。

115. 如权利要求114所述的方法, 其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列; 和

i) 与SEQ ID NO:128具有至少90%序列同一性的tracrRNA, 其中所述一种或多种指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

116. 如权利要求114所述的方法,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

117. 如权利要求114所述的方法,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA

还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;
和

i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

118.如权利要求114至117中任一项所述的方法,其中所述一种或多种指导RNA为单指导RNA (sgRNA)。

119.如权利要求114至117中任一项所述的方法,其中所述一种或多种指导RNA为双指导RNA。

120.如权利要求83至119中任一项所述的方法,其中所述目标DNA序列位于细胞内。

121.如权利要求120所述的方法,其中所述细胞为真核细胞。

122.如权利要求121所述的方法,其中所述真核细胞为植物细胞。

123.如权利要求121所述的方法,其中所述真核细胞为哺乳动物细胞。

124.如权利要求121所述的方法,其中所述真核细胞为昆虫细胞。

125.如权利要求120所述的方法,其中所述细胞为原核细胞。

126.一种包含如权利要求96至119中任一项所述的方法修饰的目标DNA序列的细胞。

127.如权利要求126所述的细胞,其中所述细胞为真核细胞。

128. 如权利要求127所述的细胞,其中所述真核细胞为植物细胞。
129. 一种包含如权利要求128所述的细胞的植物。
130. 一种包含如权利要求128所述的细胞的种子。
131. 如权利要求127所述的细胞,其中所述真核细胞为哺乳动物细胞。
132. 如权利要求131所述的细胞,其中所述哺乳动物细胞为人细胞。
133. 如权利要求132所述的细胞,其中所述人细胞为免疫细胞。
134. 如权利要求133所述的细胞,其中所述人细胞为干细胞。
135. 如权利要求134所述的细胞,其中所述干细胞为经诱导的多能干细胞。
136. 如权利要求127所述的细胞,其中所述真核细胞为昆虫细胞。
137. 如权利要求126所述的细胞,其中所述细胞为原核细胞。
138. 一种包含实施方案127及131-135中任一项所述的细胞和药学上可接受的载剂的药物组合物。
139. 一种用于检测样本中的DNA分子的目标DNA序列的试剂盒,所述试剂盒包含:
- a) 包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽或包含编码所述RGN多肽的核苷酸序列的多核苷酸,其中当与能够与所述目标DNA序列杂交的指导RNA结合时,所述RGN多肽能够以RNA指导的序列特异性方式结合及切割DNA分子的所述目标DNA序列;
- b) 所述指导RNA或包含编码所述指导RNA的核苷酸序列的多核苷酸;和
- c) 不与所述指导RNA杂交的检测单链DNA(ssDNA)。
140. 如权利要求139所述的试剂盒,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。
141. 如权利要求139所述的试剂盒,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。
142. 如权利要求139至141中任一项所述的试剂盒,其中编码所述指导RNA的至少一个所述核苷酸序列及编码所述RGN多肽的所述核苷酸序列可操作地连接至与所述核苷酸序列异源的启动子。
143. 如权利要求139至142中任一项所述的试剂盒,其中未发现所述RGN多肽与所述一种或多种指导RNA在自然界中彼此复合。
144. 如权利要求139至142中任一项所述的试剂盒,其中所述目标DNA序列为真核目标DNA序列。
145. 如权利要求139至144中任一项所述的试剂盒,其中所述检测ssDNA包含荧光团/淬灭剂对。
146. 如权利要求139至144中任一项所述的试剂盒,其中所述检测ssDNA包含荧光共振能量转移(FRET)对。
147. 如权利要求139至146中任一项所述的试剂盒,其中所述RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列,且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。
148. 如权利要求139至146中任一项所述的试剂盒,其中所述RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列,且所述一种或多种指导RNA包含CRISPR

RNA,所述CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

149.如权利要求139至146中任一项所述的试剂盒,其中所述RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列,且所述一种或多种指导RNA包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

150.如权利要求139至146中任一项所述的试剂盒,其中所述一种或多种指导RNA包含tracrRNA。

151.如权利要求150所述的试剂盒,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

152. 如权利要求150所述的试剂盒,其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

b) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

c) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

d) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的

CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

153.如权利要求150所述的试剂盒,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;
和

i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中所述一种或多种指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

154.如权利要求150至153中任一项所述的试剂盒,其中所述一种或多种指导RNA为单指导RNA(sgRNA)。

155.如权利要求150至153中任一项所述的试剂盒,其中所述一种或多种指导RNA为双指导RNA。

156.如权利要求139至155中任一项所述的试剂盒,其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

157.如权利要求139至155中任一项所述的试剂盒,其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

158. 如权利要求157所述的试剂盒,其中通过所述RGN多肽的切割产生双链断裂。

159. 如权利要求157所述的试剂盒,其中通过所述RGN多肽的切割产生单链断裂。

160. 如权利要求157至159中任一项所述的试剂盒,其中所述目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

161. 一种检测样本中的DNA分子的目标DNA序列的方法,所述方法包括:

a) 使所述样本与以下接触:

i) 包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽,其中当与能够与所述目标DNA序列杂交的指导RNA结合时,所述RGN多肽能够以RNA指导的序列特异性方式结合及切割DNA分子的所述目标DNA序列;

ii) 所述指导RNA;和

iii) 不与所述指导RNA杂交的检测单链DNA(ssDNA);和

b) 测量通过所述RGN切割所述检测ssDNA产生的可检测信号,从而检测所述目标DNA序列。

162. 如权利要求161所述的方法,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

163. 如权利要求161所述的方法,其中所述RGN多肽包含与SEQ ID NO:2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。

164. 如权利要求161至163中任一项所述的方法,其中所述样本包含来自细胞裂解物的DNA分子。

165. 如权利要求161至163中任一项所述的方法,其中所述样本包含细胞。

166. 如权利要求165所述的方法,其中所述细胞为真核细胞。

167. 如权利要求161至163中任一项所述的方法,其中包含所述目标DNA序列的所述DNA分子通过存在于包含RNA的样本中的RNA模板分子的逆转录产生。

168. 如权利要求167所述的方法,其中所述RNA模板分子为RNA病毒。

169. 如权利要求168所述的方法,其中所述RNA病毒为冠状病毒。

170. 如权利要求169所述的方法,其中所述冠状病毒为蝙蝠SARS样冠状病毒、SARS-CoV或SARS-CoV-2。

171. 如权利要求167至170中任一项所述的方法,其中包括RNA的样本来源于包含细胞的样本。

172. 如权利要求161至171中任一项所述的方法,其中所述检测ssDNA包含荧光团/淬灭剂对。

173. 如权利要求161至171中任一项所述的方法,其中所述检测ssDNA包括荧光共振能量转移(FRET)对。

174. 如权利要求161至173中任一项所述的方法,其中所述方法还包含在步骤a)的接触之前或与步骤a)的接触一起扩增所述样本中的核酸。

175. 一种切割单链DNA(ssDNA)的方法,所述方法包括使核酸群体与以下接触,其中所述群体包含DNA分子,所述DNA分子包含目标DNA序列及多个非目标ssDNA:

a) 包含与SEQ ID NO:2、4、1、3和5至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽,其中当与能够与所述目标DNA序列杂交的指导RNA结合时,

所述RGN多肽能够以RNA指导的序列特异性方式结合及切割所述目标DNA序列;和

b) 所述指导RNA;

其中所述RGN多肽切割所述多个非目标ssDNA。

176. 如权利要求175所述的方法, 其中所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有至少95%序列同一性的氨基酸序列。

177. 如权利要求175所述的方法, 其中所述RGN多肽包含与SEQ ID NO: 2、4、1、3和5至109中任一个具有100%序列同一性的氨基酸序列。

178. 如权利要求175至177中任一项所述的方法, 其中所述核酸群体在细胞裂解物内。

179. 如权利要求175至178中任一项所述的方法, 其中包含所述目标DNA序列的所述DNA分子通过RNA模板分子的逆转录产生。

180. 如权利要求161至179中任一项所述的方法, 其中未发现所述RGN多肽与所述指导RNA在自然界中彼此复合。

181. 如权利要求161至180中任一项所述的方法, 其中所述RGN多肽包含与SEQ ID NO: 11具有至少90%序列同一性的氨基酸序列且所述指导RNA包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 116具有至少90%序列同一性的CRISPR重复序列。

182. 如权利要求161至180中任一项所述的方法, 其中所述RGN多肽包含与SEQ ID NO: 11具有至少95%序列同一性的氨基酸序列且所述指导RNA包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 116具有至少95%序列同一性的CRISPR重复序列。

183. 如权利要求161至180中任一项所述的方法, 其中所述RGN多肽包含与SEQ ID NO: 11具有100%序列同一性的氨基酸序列且所述指导RNA包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 116具有100%序列同一性的CRISPR重复序列。

184. 如权利要求161至180中任一项所述的方法, 其中所述指导RNA包含tracrRNA。

185. 如权利要求184所述的方法, 其中所述tracrRNA选自以下组成的组:

a) 与SEQ ID NO: 121具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 111具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO: 2具有至少90%序列同一性的氨基酸序列;

b) 与SEQ ID NO: 123具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 113具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO: 4具有至少90%序列同一性的氨基酸序列;

c) 与SEQ ID NO: 120具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 110具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO: 1具有至少90%序列同一性的氨基酸序列;

d) 与SEQ ID NO: 122具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 112具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO: 3具有至少90%序列同一性的氨基酸序列;

e) 与SEQ ID NO: 124具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含CRISPR RNA, 所述CRISPR RNA包含与SEQ ID NO: 114具有至少90%序列同一性的CRISPR重复序列, 且其中所述RGN多肽包含与SEQ ID NO: 5具有至少90%序列同一性的氨基酸序列;

f) 与SEQ ID NO: 125具有至少90%序列同一性的tracrRNA, 其中所述指导RNA还包含

CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;
和

i)与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

186.如权利要求184所述的方法,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;
和

i)与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中所述指导RNA还包含

CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

187.如权利要求184所述的方法,其中所述tracrRNA选自以下组成的组:

a)与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

b)与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

c)与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

d)与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

e)与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

f)与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中所述指导RNA还包含CRISPR RNA,所述CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中所述RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

188.如权利要求184至187中任一项所述的方法,其中所述指导RNA为单指导RNA(sgRNA)。

189.如权利要求184至187中任一项所述的方法,其中所述指导RNA为双指导RNA。

190.如权利要求161至189中任一项所述的方法,其中所述目标DNA序列位于为单链的所述DNA分子的区域内。

191.如权利要求161至189中任一项所述的方法,其中所述目标DNA序列位于为双链的所述DNA分子的区域内。

192.如权利要求191所述的方法,其中所述目标DNA序列通过所述RGN多肽的切割产生双链断裂。

193. 如权利要求191所述的方法, 其中所述目标DNA序列通过所述RGN多肽的切割产生单链断裂。

194. 如权利要求191至193中任一项所述的方法, 其中所述目标DNA序列位于与前间隔序列邻近基序 (PAM) 相邻的位置。

195. 一种治疗疾病的方法, 所述方法包括对需要治疗的个体施用有效量的实施方案82或138的药物组合物。

196. 如权利要求195所述的方法, 其中所述疾病与因果突变关联, 且所述有效量的所述药物组合物校正所述因果突变。

RNA指导核酸酶及其活性片段与变体以及使用方法

[0001] 相关申请的交叉引用

[0002] 本申请要求2019年12月30日提交的第62/955,014号美国临时申请及2020年7月29日提交的第63/058,169号美国临时申请的优先权,上述每一篇专利申请都通过引用整体并入本文。

[0003] 关于序列表的声明

[0004] 与本申请关联的序列表已通过ASCII格式代替纸质拷贝来提供,在此通过引用并入本说明书。该ASCII拷贝被命名为L103438_1180W0_(0077_8)_SL,其大小为558,899字节,其创建于2020年12月17日,且以电子方式通过EFS-Web提交。

发明领域

[0005] 本发明与分子生物学及基因编辑领域有关。

[0006] 发明背景

[0007] 靶向基因组编辑或修饰正迅速成为基础及应用研究的重要工具。最初的方法涉及例如大范围核酸酶(meganuclease)、锌指融合蛋白或TALEN之类的工程核酸酶,这需要产生具有对每一种特定目标序列而言特异的工程化、可程序化、序列特异的DNA结合域的嵌合核酸酶。RNA指导核酸酶(例如,CRISPR-Cas细菌系统的成簇规律间隔的短回文重复序列(CRISPR)-关联(Cas)蛋白)通过将核酸酶与指导RNA(指导RNA与特定目标序列特异性杂交)复合而允许特异性序列的靶向。相较于为每一个目标序列产生嵌合核酸酶,产生目标特异性指导RNA的成本更低且更有效。这种RNA指导核酸酶可用于通过序列特异性断裂的引入来编辑基因组,该断裂通过易错的非同源末端连接(NHEJ)被修复,以在特异性基因组位置引入突变。作为另一种选择,可通过同源导向修复将异源DNA引入基因组位点。RNA指导核酸酶(RGN)也可当与脱氨酶融合时用于碱基编辑,或用于检测特异性核苷酸序列。

[0008] 发明概述

[0009] 提供了用于结合所关注的目标序列的组合物及方法。该组合物可用于切割或修饰所关注的目标序列、检测所关注的目标序列、以及修饰所关注的序列的表达。组合物包括RNA指导核酸酶(RGN)多肽、CRISPR RNA(crRNA)、反式活化的CRISPR RNA(tracrRNA)、指导RNA(gRNA)、编码其的核酸分子、包括核酸分子的载体及宿主细胞、以及包括RGN、gRNA及检测单链DNA的试剂盒。还提供了用于结合所关注的目标序列的CRISPR系统,其中该CRISPR系统包括RNA指导核酸酶多肽以及一种或多种指导RNA。因此,本文所公开的方法涉及用于结合所关注的目标序列,并且在一些实施方案中,用于切割或修饰所关注的目标序列。所关注的目标序列可例如由于非同源末端连接、引入的供体序列的同源导向修复或碱基编辑而被修饰。还提供了使用检测单链DNA检测DNA分子的目标DNA序列的方法和试剂盒。

[0010] 附图简要说明

[0011] 图1显示本发明的代表性RGN的细菌基因组基因座。

[0012] 详述

[0013] 受益于前述描述及相关附图中呈现的教导,本发明所属领域的技术人员将想到本

文中阐述的本发明的许多修改及其他实施方案。因此,应该明白,本发明不限于所公开的具体实施方案,且修改以及其他实施方案预期被包括于所附实施方案的范畴内。虽然本文采用特定术语,但这些术语仅以一般性及描述性意义使用,而非出于限制性目的。

[0014] I. 概述

[0015] RNA指导核酸酶(RGN)允许在基因组内的特异性位点的靶向操纵,并且在基因靶向的背景下有利于治疗及研究应用。在各种生物体(包括哺乳动物)中,例如,通过刺激非同源末端连接和同源重组,RNA指导核酸酶已被用于基因组工程。本文所述的组合物及方法有利于在多核苷酸中产生单链或双链断裂、修饰多核苷酸、检测多核苷酸内的特定定位点、或修饰特定基因的表达。

[0016] 本文所公开的RNA指导核酸酶可通过修饰目标序列来改变基因表达。在具体实施方案中,RNA指导核酸酶通过指导RNA(gRNA)作为成簇规律间隔的短回文重复序列(CRISPR)RNA指导核酸酶系统的局部(part)而被导向至目标序列。RGN之所以被认为是“RNA指导”的,是因为指导RNA与RNA指导核酸酶形成复合物,以将RNA指导核酸酶导向至与目标序列结合,并且在一些实施方案中,在该目标序列处引入单链或双链断裂。在目标序列被切割后,断裂可以被修复,使得该目标序列的DNA序列在修复过程期间被修饰。因此,本文提供了使用RNA指导核酸酶修饰宿主细胞的DNA中的目标序列的方法。例如,RNA指导核酸酶可用于修饰真核细胞或原核细胞的基因组基因座处的目标序列。

[0017] II. RNA指导核酸酶

[0018] 本文提供了RNA指导核酸酶。术语“RNA指导核酸酶(RGN)”是指以序列特异性方式结合至特定目标核苷酸序列且通过与多肽复合并与该目标序列杂交的指导RNA分子而被导向至该目标核苷酸序列的多肽。虽然RNA指导核酸酶有能力在结合时切割目标序列,但术语“RNA指导核酸酶”还涵盖能够结合至目标序列但不切割目标序列的无核酸酶活性的RNA指导核酸酶。通过RNA指导核酸酶切割目标序列可导致单链或双链断裂。仅能切割双链核酸分子的单链的RNA指导核酸酶在本文中被称为切口酶。

[0019] 本发明的RGN为2类CRISPR-Cas系统的成员。更具体地说,它们为V型CRISPR-Cas系统的成员。V型CRISPR-Cas系统广义上被定义为含有负责使用该指导RNA来靶向dsDNA(双链DNA)的单效应子核酸酶的系统;另外,该单效应子核酸酶含有造成催化活性的割裂RuvC核酸酶域(Jinek等人,2014,Science doi:10.1126/science.1247997;Zetsche等人,2015,Cell doi:10.1016/j.cell.2015.09.038;Shmakov等人,2017,Nat Rev Microbiol doi:10.1038/nrmicro.2016.184;Yan等人,2018,Science doi:10.1126/science.aav7271;Harrington等人,2018,Science doi:10.1126/science.aav4294;它们中的每一篇的全部内容通过引用并入本文)。大多数V型效应子也可靶向ssDNA(单链DNA),而常常没有PAM要求(Zetsche等人,2015;Yan等人,2018;Harrington等人,2018)。

[0020] 该V-A型特征蛋白(signature protein)为Cas12a。其长度为1,000-1,400个氨基酸,且除了RuvC域,还具有几个域,包括具有识别叶(recognition lobe)的楔入域(wedge domain)(Yamano等人,(2016)Cell 165:949-962)。相对地,相较于大多数其他V型系统,V-U型系统的大小较小(长度为500-700个氨基酸)。V-U[′]也具有割裂RuvC域及带正电荷的桥式螺旋(Shmakov等人,2017)。尽管Cas12a与cas1、cas2且偶尔与cas4共定位,但该V-U蛋白常常不具有有效应子蛋白编码的辅助Cas蛋白,(Shmakov等人,2017)。基于V-U型系统与其他V

型成员之间的这些不同,Shmakov等人(2017)建议在确定功能性时,该V-U型系统应当接收新型/子型命名。

[0021] 例如,Cas14酶的长度为400-700个氨基酸(Harrington等人,2018)。在第一次公开时,这些系统被称为(tout)来自Vs型的规范Cas12效应子蛋白的单独的Cas酶。Yan等人后来的公开已于该V型命名规则内将Cas14a、-b和-c称为V-F子型。Cas14a及b与为V-U3型的c2c10最密切相关。Cas14c与分别为V-U2及V-U4型的c2c8及c2c9最密切相关(Harrington等人,2018;Yan等人,2018)。该Cas14 RGN的基因组基因座与辅助Cas蛋白关联,且该tracrRNA在该Cas14与该重复序列-间隔序列阵列之间被编码。与能够处理来自含有多个指导RNA的单转录物的个别指导的Cas12a不同,这些系统不能处理它们自己的指导RNA(Harrington等人,2018)。

[0022] 除APG06369外,本发明的全部RGN都含有割裂RuvC域。然而,本发明的许多RGN具有唯一基因座排列,暗示这些RGN对2类CRISPR-Cas分类系统为新的。本发明的RGN所来源自的基因座(见实施例1中的表1)中没有含有Cas1或Cas2。

[0023] 如本文所公开的,APG07339、APG09624、APG03003、APG05405、APG09777、APG05680、APG02119、APG03285、APG04998和APG07078为未用辅助基因编码且除了crRNA还可要求tracrRNA的独立Cas效应子。基于本文的公开,这些CRISPR-Cas系统需要接收新分类。另外,谱系分析揭示可将这些RGN分组为三个不同子类。一个子类含有APG07078。第二子类含有APG05680及APG03285。第三子类含有APG07339、APG09624、APG03003、APG05405、APG09777、APG02119和APG04998。

[0024] APG06369为唯一效应子核酸酶,该唯一效应子核酸酶缺少可区别RuvC域且坐落于具有不规范辅助基因、以前从未见过的CRISPR基因座中。APG06369具有四个辅助基因(该四个辅助蛋白如SEQ ID NO:178-181所示),它们中没有一个具有注释域(annotated domain)或功能。APG06369为唯一Cas蛋白。

[0025] 在谱系上,APG03847、APG05625、APG03759、APG05123和APG03524形成含有效应子核酸酶的唯一RuvC的支系(clade)。这些RGN具有至多3个辅助基因:一个为HNH核酸内切酶,一个为HTH转录调节子(regulator),而第三个具有未知功能或域。APG03847的辅助蛋白如SEQ ID NO:182、183和184所示。APG05625的辅助蛋白如SEQ ID NO:185、186和187所示。APG03524的辅助蛋白如SEQ ID NO:188、189和190所示。APG03759及APG05123的辅助蛋白分别如SEQ ID NO:191及192所示。它们在它们的基因座处具有唯一CRISPR重复序列排列,其中与APG03847、APG05625、APG03759、APG05123和APG03524关联的重复序列对齐(flush with)大量蛋白的编码序列。此为CRISPR-Cas系统的极不寻常特征,且暗示一种不需要前导序列的CRISPR表达形式。此CRISPR表达形式不同于迄今已知的任何系统。

[0026] 本文所公开的RNA指导核酸酶包括表1中显示的RNA指导核酸酶,其氨基酸序列如SEQ ID NO:1至109所示,且其活性片段或变体保持以RNA指导的序列特异性方式结合至目标核苷酸序列的能力。在一些实施方案中,RGN的此活性片段或变体能够切割单链或双链目标序列。在一些实施方案中,本发明的RGN的活性变体包含与如SEQ ID NO:1至109所示的氨基酸序列中任一个具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的氨基酸序列。在某些实施方案中,本发明的RGN的活性片段包含如SEQ ID NO:1至109所示的氨基酸序

列中任一个的至少50、100、150、200、250、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050个或更多个连续氨基酸残基。本文提供的RNA指导核酸酶可包含至少一个核酸酶域(例如,DNase、RNase域)及至少一个RNA识别域和/或RNA结合域,以与指导RNA相互作用。可在本文提供的RNA指导核酸酶中发现的另外域包括但不限于:DNA结合域、解旋酶域、蛋白质-蛋白质相互作用域及二聚化域。在具体实施方案中,本文提供的RNA指导核酸酶可与DNA结合域、解旋酶域、蛋白质-蛋白质相互作用域及二聚化域中的一种或多种包括至少70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的序列同一性。

[0027] 在各种实施方案中,目标核苷酸序列被本文提供的RNA指导核酸酶结合,并与RNA指导核酸酶关联的指导RNA杂交。如果该多肽具有核酸酶活性,则可由RNA指导核酸酶随后切割该目标序列。术语“切割(cleave)”或“切割(cleavage)”是指目标核苷酸序列的主链内的至少一个磷酸二酯键的水解,其可导致该目标序列内的单链或双链断裂。在各种实施方案中,本发明公开的RGN可切割多核苷酸内的核苷酸,起核酸内切酶的作用,或者,本发明公开的RGN可为核酸外切酶,从多核苷酸的末端(5'末端和/或3'末端)移除连续的核苷酸。在一些实施方案中,所公开的RGN可在多核苷酸的任何位置内切割目标序列的核苷酸,且由此起核酸内切酶及核酸外切酶二者的作用。通过本发明公开的RGN对目标多核苷酸的切割可导致交错的断裂或钝末端。

[0028] 在一些实施方案中,为结合至和/或切割目标多核苷酸,RGN需要至少一个RGN辅助蛋白的表达或存在。在这些实施方案中的一些实施方案中,RGN需要如SEQ ID NO:178-192所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG06369(SEQ ID NO:11)或其变体或片段的特定实施方案中,活性需要如SEQ ID NO:178-181所示的至少一个RGN辅助蛋白或其活性变体或片段。在这些实施方案中其中该RGN为APG03847(SEQ ID NO:12)或其变体或片段的一些实施方案中,活性需要如SEQ ID NO:182-184所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG05625(SEQ ID NO:13)或其变体或片段的某些实施方案中,活性需要如SEQ ID NO:185-187所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG03524(SEQ ID NO:16)或其变体或片段的一些实施方案中,活性需要如SEQ ID NO:188-190所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG03759(SEQ ID NO:14)或其变体或片段的特定实施方案中,活性需要如SEQ ID NO:191所示的RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG05123(SEQ ID NO:15)或其变体或片段的某些实施方案中,活性需要如SEQ ID NO:192所示的RGN辅助蛋白或其活性变体或片段。

[0029] 在一些实施方案中,本发明公开的RNA指导核酸酶可为来源于细菌物种或古菌物种的野生型序列。在一些实施方案中,RNA指导核酸酶可为野生型多肽的变体或片段。例如,可修饰该野生型RGN以改变核酸酶活性或改变PAM特异性。在一些实施方案中,RNA指导核酸酶不是天然存在的。

[0030] 在某些实施方案中,RNA指导核酸酶起切口酶的作用,仅切割目标核苷酸序列的单链。这样的RNA指导核酸酶具有单一作用的核酸酶域。在特定实施方案中,切口酶能够切割正链或负链。在这些实施方案中的一些实施方案中,额外的核酸酶域已经突变,使得核酸酶活性降低或消除。

[0031] 在一些实施方案中, RNA指导核酸酶完全缺少核酸酶活性, 且在本文中被称为无核酸酶活性 (nuclease-dead) 或核酸酶失活 (nuclease inactive)。本领域中用于将突变引入氨基酸序列内的任何已知方法 (诸如PCR介导的诱变及定点诱变) 可用于产生无切口酶或无核酸酶活性的RGN。例如参见第2014/0068797号美国公开以及第9,790,490号美国专利; 其每一篇的全部内容通过引用而被并入本文。

[0032] 缺少核酸酶活性的RNA指导核酸酶可用于将融合多肽、多核苷酸或小分子载荷递送至特定基因组位置。在这些实施方案中的一些实施方案中, RGN多肽或指导RNA可与可检测标记融合, 以允许特定序列的检测。作为非限制性实例, 无核酸酶活性的RGN可与可检测标记 (例如, 荧光蛋白) 融合且靶向至与疾病关联的特定序列, 以允许该疾病关联序列的检测。

[0033] 在一些实施方案中, 无核酸酶活性的RGN可靶向至特定基因组位置, 以改变期望序列的表达。在一些实施方案中, 通过干扰所靶向的基因组区域内的RNA聚合酶或转录因子的结合, 无核酸酶活性的RNA指导核酸酶与目标序列的结合导致该目标序列的或受该目标序列的转录控制的基因的表达减少。在其他实施方案中, RGN (例如, 无核酸酶活性的RGN) 或其复合的指导RNA还包含表达调控子 (modulator), 其在与目标序列结合时用来阻遏或活化该目标序列或受该目标序列的转录控制的基因的表达。在这些实施方案中的一些实施方案中, 表达调控子通过表观遗传机制调控该目标序列或所调节基因的表达。

[0034] 在一些实施方案中, 无核酸酶活性的RGN或仅有切口酶活性的RGN可靶向至特定基因组位置, 以通过与碱基编辑多肽 (例如, 脱氨酶多肽) 或其活性变体或片段融合 (这直接化学修饰核碱基 (例如, 直接使核碱基脱氨基)) 来修饰目标多核苷酸的序列, 导致从一种核碱基转化为另一种核碱基。该碱基编辑多肽可在其N端侧或C端侧与该RGN融合。另外, 该碱基编辑多肽可通过肽接头而与该RGN融合。有利于此类组合物及方法的脱氨酶多肽的非限制性实例包括胞苷脱氨酶或腺嘌呤脱氨酶 (诸如, Gaudelli 等人, (2017) Nature 551:464-471、第2017/0121693号及第2018/0073012号美国专利公开、第W0/2018/027078号国际专利公开中描述的腺苷碱基编辑器, 或者第W0 2020/139873号国际专利公开及第62/785,391号 (2018年12月27日提交申请)、第62/932,169号 (2019年11月7日提交申请) 和第63/077,089号 (2020年9月11日提交申请) 的美国临时专利申请中公开的脱氨酶中任一个, 上述每一篇的全部内容通过引用而被并入本文)。此外, 本领域中已知RGN与碱基编辑酶之间的某些融合蛋白也可包含至少一个尿嘧啶稳定多肽, 其通过脱氨酶增加胞苷、脱氧胞苷或胞嘧啶与核酸分子中的胸苷、脱氧胸苷或胸腺嘧啶的突变率。尿嘧啶稳定多肽的非限制性实例包括2020年7月15日提交申请的第63/052,175号美国临时专利申请公开的尿嘧啶稳定多肽及尿嘧啶糖基化酶抑制剂 (UGI) 域 (SEQ ID NO:137), 该尿嘧啶糖基化酶抑制剂 (UGI) 域可增加碱基编辑效率。在特定实施方案中, 本公开内容提供一种包含本文描述的RGN或其变体、脱氨酶和视情况而定的至少一个尿嘧啶稳定多肽 (诸如, UGI) 的融合蛋白。在某些实施方案中, 融合至该碱基编辑多肽的RGN为切割该碱基编辑多肽不发挥作用的DNA链的切口酶 (例如, 脱氨酶)。与多肽或域融合的RNA指导核酸酶可通过接头而被分开或连接。如本文所使用的, 术语“接头”是指连接两个分子或部分 (例如, 核酸酶的结合域及切割域) 的化学基团或分子。在一些实施方案中, 接头连接RNA指导核酸酶的gRNA结合域与诸如脱氨酶的碱基编辑多肽。在一些实施方案中, 接头连接无核酸酶活性的RGN与脱氨酶。通常情况下, 接头位于两

个基团、分子或其他部分之间或侧边有两个基团、分子或其他部分,且通过共价键而被连接到每一个,由此连接二者。在一些实施方案中,接头为一个氨基酸或多个氨基酸(例如,肽或蛋白质)。在一些实施方案中,接头为有机分子、基团、聚合物或化学部分。在一些实施方案中,接头的长度为5-100个氨基酸,例如,长度为5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、30-35、35-40、40-45、45-50、50-60、60-70、70-80、80-90、90-100、100-150、或150-200个氨基酸。也考虑了更长或更短的接头。

[0035] 在各种实施方案中,本公开内容提供本发明公开的RNA指导核酸酶,包含至少一个核定位信号(NLS),以增强该RGN向细胞核的运输。核定位信号为本领域中已知的且一般包含一段碱性氨基酸(参见,例如,Lange等人,J.Biol.Chem. (2007) 282:5101-5105)。在特定实施方案中,RGN包含2、3、4、5、6或更多个核定位信号。核定位信号可为异源NLS。有用于本发明公开的RGN的核定位信号的非限制性实例为SV40大T抗原、核质素及c-Myc的核定位信号(参见,例如,Ray等人,(2015) Bioconjug Chem 26 (6):1004-7)。在特定实施方案中,RGN包含如SEQ ID NO:149或150所示的NLS序列。该RGN可在其N端、C端、或在N端及C端二者包含一种或多种NLS序列。例如,该RGN可在N端区域处包含两个NLS序列,而在C端区域处包含四个NLS序列。

[0036] 本领域中已知的将多肽定位于特定亚细胞位置的其他定位信号序列也可用于靶向该RGN,包括但不限于:质体定位序列、线粒体定位序列和靶向质体及线粒体二者的双靶向信号序列(参见,例如,Nassoury及Morse (2005) Biochim Biophys Acta 1743:5-19; Kunze及Berger (2015) Front Physiol dx.doi.org/10.3389/fphys.2015.00259; Herrmann及Neupert (2003) IUBMB Life 55:219-225; Soll (2002) Curr Opin Plant Biol 5:529-535; Carrie及Small (2013) Biochim Biophys Acta 1833:253-259; Carrie等人,(2009) FEBS J 276:1187-1195; Silva-Filho (2003) Curr Opin Plant Biol 6:589-595; Peeters及Small (2001) Biochim Biophys Acta 1541:54-63; Murcha等人,(2014) J Exp Bot 65:6301-6335; Mackenzie (2005) Trends Cell Biol 15:548-554; Glase等人,(1998) Plant Mol Biol 38:311-338)。

[0037] 在某些实施方案中,本发明公开的RNA指导核酸酶包含促进该RGN的细胞摄取的至少一个细胞穿透域。细胞穿透域为本领域中已知的且一般包含:数段带正电荷的氨基酸残基(即,聚阳离子细胞穿透域)、交替极性的氨基酸残基及非极性氨基酸残基(即,两亲性细胞穿透域)、或疏水性氨基酸残基(即,疏水性细胞穿透域)(参见,例如,Milletti F. (2012) Drug Discov Today 17:850-860)。细胞穿透域的非限制性实例为来自人免疫缺陷病毒1的反式活化转录激活子(TAT)。

[0038] 核定位信号、质体定位信号、线粒体定位信号、双靶向定位信号和/或细胞穿透域可位于该RNA指导核酸酶的氨基端(N端)、羧基端(C端)、或内部位置中。

[0039] 在某些实施方案中,本发明公开的RGN可通过接头肽而直接或间接地与诸如切割域、脱氨酶域、或表达调控域的效应域融合。此域可位于该RNA指导核酸酶的N端、C端或内部位置中。在这些实施方案中的一些实施方案中,融合蛋白的RGN组成为无核酸酶活性的RGN。

[0040] 在一些实施方案中,RGN融合蛋白包含切割域,该切割域为能够切割多核苷酸(即, RNA、DNA或RNA/DNA杂合体)的任何域,并且包括但不限于限制性核酸内切酶和归巢核酸内切酶(homing endonuclease),诸如IIS型核酸内切酶(例如,FokI)(参见,例如,Belfort等

人, (1997) *Nucleic Acids Res.* 25:3379-3388; Linn 等人, (eds.) *Nucleases* (核酸酶), 冷泉港实验室出版社 (Cold Spring Harbor Laboratory Press), 1993)。

[0041] 在一些实施方案中, RGN 融合蛋白包含脱氨酶域, 该脱氨酶域将核碱基脱去氨基, 导致一种核碱基转化为另一种核碱基, 且包括但不限于胞苷脱氨酶或腺嘌呤脱氨酶碱基编辑器 (参见, 例如, Gaudelli 等人, (2017) *Nature* 551:464-471、第 2017/0121693 号及第 2018/0073012 号美国专利公开、第 9,840,699 号美国专利和第 W0/2018/027078 号国际公开、第 PCT/US2019/068079 号国际专利申请和第 62/785,391 号 (2018 年 12 月 27 日提交申请) 及第 62/932,169 号 (2019 年 11 月 7 日提交申请) 美国临时专利申请)。

[0042] 在一些实施方案中, RGN 融合蛋白的效应域可为表达调控域, 该表达调控域为用来上调或下调转录的域。该表达调控域可为表观遗传修饰域、转录阻遏域或转录活化域。

[0043] 在这些实施方案中的一些实施方案中, RGN 融合蛋白的表达调控子包含表观遗传修饰域, 该表观遗传修饰域共价修饰 DNA 或组蛋白以改变组蛋白结构和/或染色体结构, 而不改变该 DNA 序列, 引起基因表达的变化 (即, 上调或下调)。表观遗传修饰的非限制性实例包括赖氨酸残基的乙酰化或甲基化、精氨酸甲基化、丝氨酸及苏氨酸磷酸化和组蛋白的赖氨酸泛素化 (ubiquitination) 及 SUMO 化 (sumoylation) 和 DNA 中胞嘧啶残基的甲基化和羟甲基化。表观遗传修饰域的非限制性实例包括组蛋白乙酰基转移酶域、组蛋白去乙酰基酶域、组蛋白甲基转移酶域、组蛋白去甲基化酶域、DNA 甲基转移酶域及 DNA 去甲基化酶域。

[0044] 在一些实施方案中, 融合蛋白的表达调控子包含转录阻遏域, 该转录阻遏域与诸如 RNA 聚合酶及转录因子的转录控制元件和/或转录调节蛋白相互作用, 以减少或终止至少一个基因的转录。转录阻遏域为本领域中已知且包括但不限于类 Sp1 阻遏物、I κ B 及 Kr $\ddot{u}$ ppel 关联盒 (KRAB) 域。

[0045] 在一些实施方案中, 融合蛋白的表达调控子包含转录活化域, 该转录活化域与诸如 RNA 聚合酶及转录因子的转录控制元件和/或转录调节蛋白相互作用, 以增加或活化至少一个基因的转录。转录活化域为本领域中已知且包括但不限于单纯疱疹病毒 VP16 活化域及 NFAT 活化域。

[0046] 在一些实施方案中, 本发明公开的 RGN 多肽包含可检测标记或纯化标签。该可检测标记或纯化标签可直接地或通过接头肽间接地位于该 RNA 指导核酸酶的 N 端、C 端或内部位置中。在这些实施方案中的一些实施方案中, 融合蛋白的 RGN 组成为无核酸酶活性 RGN。在其他实施方案中, 融合蛋白的 RGN 组成为具切口酶活性 RGN。

[0047] 可检测标记为可直观的或可以其他方式观察的分子。可检测标记可与该 RGN 融合为融合蛋白 (例如, 荧光蛋白), 也可与 RGN 多肽缀合的、可直观地或通过其他手段检测的小分子。可与本发明公开的 RGN 融合为融合蛋白的可检测标记包括任何可检测蛋白域, 包括但不限于可用特异性抗体检测的荧光蛋白或蛋白域。荧光蛋白的非限制性实例包括绿色荧光蛋白 (例如, GFP、EGFP、ZsGreen1) 及黄色荧光蛋白 (例如, YFP、EYFP、ZsYellow1)。小分子可检测标记的非限制性实例包括放射性标记, 诸如, ^3H 以及 ^{35}S 。

[0048] 在一些实施方案中, 本发明公开的 RGN 多肽包含纯化标签, 该纯化标签可采用的任何分子, 以从混合物 (例如, 生物样本、培养基) 分离蛋白质或融合蛋白。纯化标签的非限制性实例包括生物素、myc、麦芽糖结合蛋白 (MBP) 和谷胱甘肽-S-转移酶 (GST)。

[0049] III. 指导 RNA

[0050] 本公开内容提供指导RNA及编码其的多核苷酸。术语“指导RNA”是指核苷酸序列，该核苷酸序列与目标核苷酸序列具有足够互补性，以与该目标序列杂交且将关联RNA指导核酸酶的序列特异性结合导向至该目标核苷酸序列。因此，RGN的各指导RNA为一种或多种RNA分子（一般而言，一种或两种），其可与该RGN结合且指导RGN来与特定目标核苷酸序列结合，且在该RGN具有切口酶或核酸酶活性的那些实施方案中，还切割该目标核苷酸序列。在一些实施方案中，指导RNA包含CRISPR RNA (crRNA) 且在一些实施方案中，包含反式活化的CRISPR RNA (tracrRNA)。包含crRNA及tracrRNA二者的天然指导RNA一般包含两个分开的RNA分子，此两个分开的RNA分子通过该crRNA的重复序列及该tracrRNA的抗重复序列彼此杂交。

[0051] 在一些实施方案中，CRISPR阵列内的天然直接重复序列在28至37个碱基对的长度范围内。在一些实施方案中，CRISPR阵列内的天然直接重复序列的长度在约23bp至约55bp（例如，23bp至55bp）的范围内。在一些实施方案中，CRISPR阵列内的间隔序列的长度在约32至约38bp的范围内。在一些实施方案中，CRISPR阵列内的间隔序列的长度在约21bp至约72bp（例如，21bp至72bp）的范围内。在一些实施方案中，本文所公开的CRISPR阵列包含少于50个单位的CRISPR重复序列-间隔序列。该CRISPR被转录为被称为初级CRISPR转录物的长转录物的局部，其包含该CRISPR阵列的大部分。该初级CRISPR转录物被Cas蛋白切割，以产生crRNA，或者在一些情况中，产生前体crRNA (pre-crRNA)，其被额外的Cas蛋白进一步处理为成熟的crRNA。成熟的crRNA包含间隔序列及CRISPR重复序列。在其中将前体crRNA处理为成熟（或经处理）crRNA的一些实施方案中，成熟涉及移除约1至约6个或更多个5'、3'或5'及3'核苷酸。出于基因组编辑或靶向所关注的特定目标核苷酸序列的目的，在前体crRNA分子成熟期间移除的这些核苷酸对于产生或设计指导RNA不是必需的。本发明公开的每一个RGN蛋白 (SEQ ID NO:1-109) 的共有重复序列分别示于SEQ ID NO:201-309中。APG07339 (SEQ ID NO:1)、APG09624 (SEQ ID NO:2)、APG03003 (SEQ ID NO:3)、APG05405 (SEQ ID NO:4)、APG09777 (SEQ ID NO:5)、APG05680 (SEQ ID NO:6)、APG06369 (SEQ ID NO:11)、APG03847 (SEQ ID NO:12)、APG05625 (SEQ ID NO:13) 和APG03524 (SEQ ID NO:16) 中的每一个的经处理crRNA重复序列分别公开于SEQ ID NO:110-119中。

[0052] CRISPR RNA (crRNA) 包含间隔序列及CRISPR重复序列。“间隔序列”为与所关注的目标核苷酸序列直接杂交的核苷酸序列。该间隔序列被工程化为与所关注的目标序列完全地或部分地互补。在各种实施方案中，间隔序列可包含约8个核苷酸至约30个核苷酸或更多个核苷酸。例如，该间隔序列的长度可为约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30个或更多个核苷酸。在一些实施方案中，间隔序列的长度为约10至约26个核苷酸，或长度为约12至约30个核苷酸。在特定实施方案中，间隔序列的长度为约30个核苷酸。在一些实施方案中，当使用适合的比对 (alignment) 算法进行最佳比对时，间隔序列与其对应的目标序列之间的互补性程度为约或大于约50%、约60%、约70%、约75%、约80%、约81%、约82%、约83%、约84%、约85%、约86%、约87%、约88%、约89%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或更高。在特定实施方案中，间隔序列不含使用本领域中已知的任何适合的多核苷酸折叠算法可预测的二级结构，多核苷酸折叠算法包括但不限于mFold (参见，例如，Zuker及Stiegler (1981) *Nucleic Acids Res.* 9:133-148) 及

RNAfold (参见,例如,Gruber等人,(2008) Cell 106 (1):23-24)。

[0053] CRISPR RNA重复序列包含核苷酸序列,该核苷酸序列或者独自地或者与所杂交 tracrRNA配合形成通过该RGN分子识别的结构。在各种实施方案中,CRISPR RNA重复序列可包含约8个核苷酸至约30个核苷酸或更多个核苷酸。例如,CRISPR重复序列的长度可为约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30个或更多个核苷酸。在一些实施方案中,CRISPR重复序列的长度为约21个核苷酸。在一些实施方案中,当使用适合的比对算法进行最佳比对时,CRISPR重复序列与其对应的tracrRNA序列之间的互补性程度为约或大于约50%、约60%、约70%、约75%、约80%、约81%、约82%、约83%、约84%、约85%、约86%、约87%、约88%、约89%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或更高。在特定实施方案中,CRISPR重复序列包含SEQ ID NO:110至119、139、141、143、146和201至309的核苷酸序列的任一个或其活性变体或片段,其当被包含于指导RNA内时能够将本文所提供的关联RNA指导核酸酶的序列特异性结合导向至所关注的目标序列。在某些实施方案中,野生型序列的活性CRISPR重复序列变体包含与如SEQ ID NO:110至119、139、141、143、146和201至309所示的核苷酸序列中任一个具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核苷酸序列。在某些实施方案中,野生型序列的活性CRISPR重复序列片段包含如SEQ ID NO:110至119、139、141、143、146和201至309所示的核苷酸序列中任一个的至少5、6、7、8、9、10、11、12、13、14、15、16、17、18、19或20个连续核苷酸。

[0054] 在某些实施方案中,crRNA不是天然存在的。在这些实施方案中的一些实施方案中,特异性CRISPR重复序列在自然界中(in nature)不与该工程化间隔序列连接,且该CRISPR重复序列被认为与该间隔序列是异源的。在某些实施方案中,间隔序列为非天然存在的工程化序列。

[0055] 在一些实施方案中,指导RNA还包含tracrRNA分子。反式活化的CRISPR RNA或tracrRNA分子包含核苷酸序列,该核苷酸序列包含具有与crRNA的CRISPR重复序列杂交的足够互补性的区域,其在本文中被称作抗重复序列区域。在一些实施方案中,tracrRNA分子还包含具二级结构(例如,茎-环)的区域,或在与其对应的crRNA杂交时形成二级结构。在特定实施方案中,tracrRNA的与CRISPR重复序列完全地或部分地互补的区域处于分子的5'末端,且该tracrRNA的3'末端包含二级结构。此二级结构区域一般包含被发现与该抗重复序列相邻、包括连结(nexus)发夹的几个发夹结构。该tracrRNA的3'末端处常常存在端发夹,其结构及数量可变,但常常包括富含GC的Rho独立转录终止子发夹,其后在3'末端处具有一串U。参见,例如,Briner等人,(2014) Molecular Cell 56:333-339、Briner及Barrangou (2016) Cold Spring Harb Protoc;doi:10.1101/pdb.top090902及第2017/0275648号美国专利公开,上述每一篇的全部内容通过引用并入本文。

[0056] 在各种实施方案中,与CRISPR重复序列完全地或部分地互补的tracrRNA的抗重复序列区域包含约6个核苷酸至约30个核苷酸或更多个核苷酸。例如,tracrRNA抗重复序列与该CRISPR重复序列之间的碱基配对区域的长度可为约6、约7、约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25、约26、约27、约28、约29、约30个或更多个核苷酸。在特定实施方案中,与CRISPR重复序列完全地或部分地

互补的tracrRNA的抗重复序列区域的长度为约10个核苷酸。在一些实施方案中,当使用适合的比对算法进行最佳比对时,CRISPR重复序列与其对应的tracrRNA抗重复序列之间的互补性程度为约或大于约50%、约60%、约70%、约75%、约80%、约81%、约82%、约83%、约84%、约85%、约86%、约87%、约88%、约89%、约90%、约91%、约92%、约93%、约94%、约95%、约96%、约97%、约98%、约99%或更高。

[0057] 在各种实施方案中,整个tracrRNA可包含约60个核苷酸至多约210个核苷酸。例如,该tracrRNA的长度可为约60、约65、约70、约75、约80、约85、约90、约95、约100、约105、约110、约115、约120、约125、约130、约135、约140、约150、约160、约170、约180、约190、约200、约210个或更多个核苷酸。在特定实施方案中,tracrRNA的长度为约100至约201个核苷酸,长度为包括约95、约96、约97、约98、约99、约100、约105、约106、约107、约108、约109个及约100个核苷酸。在某些实施方案中,tracrRNA的长度为约96个核苷酸。

[0058] 在特定实施方案中,tracrRNA包含SEQ ID NO:120至128、140、142、145、147和148的核苷酸序列的任一个或其活性变体或片段,其当被包含于指导RNA内时能够将本文提供的关联RNA指导核酸酶的序列特异性结合导向至所关注的目标序列。在某些实施方案中,野生型序列的活性tracrRNA序列变体包含与如SEQ ID NO:120至128、140、142、145、147和148所示的核苷酸序列的任一个具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性的核苷酸序列。在某些实施方案中,野生型序列的活性tracrRNA序列片段包含如SEQ ID NO:120至128、140、142、145、147和148所示的核苷酸序列的任一个的至少5、10、15、20、25、30、35、40、45、50、55、60、65、70、75、80个或更多个连续核苷酸。

[0059] 当两个序列在严格条件下彼此杂交时,可认为该两个多核苷酸序列实质上互补。同样,如果与RGN结合的指导RNA在严格条件下与目标序列结合,则认为该RGN以序列特异性方式结合至该特定目标序列。“严格条件”或“严格杂交条件”旨在指两个多核苷酸序列彼此杂交至可检测程度高于其他序列(例如,至少比背景高2倍)的条件。严格条件为序列依赖性的,且在不同环境下将会不同。通常情况下,严格条件将是这样的条件,其中:在pH 7.0至8.3,盐类浓度小于约1.5M Na离子,通常情况下约0.01至1.0M Na离子浓度(或其它盐类),且针对短序列(例如,10至50个核苷酸),温度为至少约30℃,而针对长序列(例如,大于50个核苷酸),温度为至少约60℃。通过添加诸如甲酰胺的去稳定剂也可达到严格条件。示例性的低严格条件包括在37℃下以30至35%甲酰胺、1M NaCl、1%SDS(十二烷基硫酸钠)的缓冲溶液杂交及在50至55℃以1X至2X SSC洗涤(20X SSC=3.0M NaCl/0.3M柠檬酸三钠)。示例性的中等严格条件包括在37℃下在40至45%甲酰胺、1.0M NaCl、1%SDS中杂交及在55至60℃下在0.5X至1X SSC中洗涤。示例性的高严格条件包括在37℃下在50%甲酰胺、1M NaCl、1%SDS中杂交及在60至65℃下在0.1X SSC中洗涤。视情况,洗涤缓冲液可包含约0.1%至约1%的SDS。杂交持续时间一般小于约24小时,通常约4至约12小时。洗涤时间的持续时间至少为足以达到平衡的时间长度。

[0060] T_m 为50%的互补目标序列与完全匹配的序列杂交的温度(在所定义的离子强度和pH下)。对于DNA-DNA杂合体, T_m 可以由Meinkoth和Wahl (1984) Anal. Biochem. 138:267-284中的等式: $T_m = 81.5^\circ\text{C} + 16.6(\log M) + 0.41(\%GC) - 0.61(\%form) - 500/L$ 大致估计;其中M为单价阳离子的莫耳浓度,%GC为鸟苷和胞嘧啶核苷酸在该DNA中的百分比,%form为甲酰胺

在杂交溶液中的百分比,及L为碱基对中的杂合体的长度。一般而言,严格条件被选择为比在所定义的离子强度和pH下的特异性序列及其互补序列的热解链温度(T_m)低约5℃。然而,极度严格条件可采用在比该热解链温度(T_m)低1、2、3或4℃下的杂交和/或洗涤;中等严格条件可采用在比该热解链温度(T_m)低6、7、8、9或10℃的温度下的杂交和/或洗涤;低严格条件可采用在比该热解链温度(T_m)低11、12、13、14、15或20℃的温度下的杂交和/或洗涤。本领域技术人员将明白,使用该等式、杂交及洗涤组合物以及要求的 T_m ,杂交和/或洗涤溶液的严格性的变化被固有地描述了。核酸杂交的充分指南可在Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 (Elsevier, New York); 和Ausubel等人, eds. (1995) *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley-Interscience, New York) 中得到。参见Sambrook等人, (1989) *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York)。

[0061] 术语“序列特异性”也可是指相较于与随机化背景序列的结合,以更高频率与目标序列结合。

[0062] 在一些实施方案中,例如,其中该指导RNA包含crRNA及tracrRNA二者,该指导RNA可为单指导RNA或双指导RNA系统。单指导RNA包含单RNA分子上的crRNA及tracrRNA,而双指导RNA系统包含存在于两个不同的RNA分子上的crRNA及tracrRNA,其通过该crRNA的CRISPR重复序列的至少一部分及该tracrRNA的至少一部分而彼此杂交,其可与该crRNA的CRISPR重复序列完全地或部分地互补。在其中指导RNA为单指导RNA的那些实施方案中的一些实施方案中,crRNA与tracrRNA被接头核苷酸序列分开。一般而言,为避免二级结构在该接头核苷酸序列的核苷酸内的形成或避免包含该接头核苷酸序列的核苷酸的二级结构的形成,接头核苷酸序列为不包括互补碱基的核苷酸序列。在一些实施方案中,crRNA与tracrRNA之间的接头核苷酸序列的长度为至少3、至少4、至少5、至少6、至少7、至少8、至少9、至少10、至少11、至少12个或更多个核苷酸。在特定实施方案中,单指导RNA的接头核苷酸序列的长度为至少4个核苷酸。在某些实施方案中,接头核苷酸序列为如SEQ ID NO:136所示的核苷酸序列。在其他实施方案中,接头核苷酸序列的长度为至少6个核苷酸。

[0063] 可化学合成或通过体外转录合成该单指导RNA或双指导RNA。用于确定RGN与指导RNA之间的序列特异性结合的测定为本领域中所知、并且包括但不限于所表达RGN与该指导RNA之间的体外结合测定,其可用可检测标记(如生物素)标记并用于下拉检测测定,其中,该指导RNA:RGN复合物是通过该可检测标记(例如,利用链霉抗生物素蛋白珠)捕获的。具与该指导RNA无关的序列或结构的对照指导RNA可用作该RGN与RNA之非特异性结合的阴性对照。在某些实施方案中,指导RNA为SEQ ID NO:129至135及310中任一个,其中该间隔序列可为任何序列且被指示为poly-N序列。

[0064] 在某些实施方案中,指导RNA可作为RNA分子而被引入目标细胞、细胞器或胚胎中。可以体外转录或化学合成该指导RNA。在其他实施方案中,将编码该指导RNA的核苷酸序列引入该细胞、细胞器或胚胎中。在这些实施方案中的一些实施方案中,编码该指导RNA的核苷酸序列可操作地被连接至启动子(例如, RNA聚合酶III启动子)。该启动子可为天然启动子或与该指导RNA编码的核苷酸序列异源。

[0065] 在各种实施方案中,如本文所述,指导RNA可作为核糖核蛋白复合物而被引入目标细胞、细胞器或胚胎中,其中该指导RNA与RNA指导核酸酶多肽结合。

[0066] 指导RNA通过该指导RNA与该目标核苷酸序列的杂交而将关联RNA指导核酸酶导向至所关注的特定目标核苷酸序列。目标核苷酸序列可包含DNA、RNA或二者的组合,且可为单链或双链的。目标核苷酸序列可为基因组DNA(即,染色体DNA)、质粒DNA、或RNA分子(例如,信使RNA、核糖体RNA、转运RNA、微RNA、小干扰RNA)。该目标核苷酸序列可在体外或在细胞中被RNA指导核酸酶结合(而在一些实施方案中,被切割)。该RGN所靶向的染色体序列可为细胞核、质体或线粒体染色体序列。在一些实施方案中,目标核苷酸序列在该目标基因组中是唯一的。

[0067] 在一些实施方案中,目标核苷酸序列与前间隔序列邻近基序(PAM)相邻。在某些实施方案中,双链目标序列的切割依赖于PAM的存在,而单链目标序列的切割不依赖于PAM。前间隔序列邻近基序一般在距该目标核苷酸序列约1至约10个核苷酸内,包括距该目标核苷酸序列约1、约2、约3、约4、约5、约6、约7、约8、约9、或约10个核苷酸。该PAM可为该目标序列的5'或3'。在一些实施方案中,该PAM为本发明公开的RGN的目标序列的5'。一般而言,该PAM是约3-4个核苷酸的共有序列,但在特定实施方案中,PAM的长度可为1、2、3、4、5、6、7、8、9个或更多个核苷酸。在一些实施方案中,PAM为该目标序列的5',且富含T。

[0068] 在一些实施方案中,RGN分别与包含SEQ ID NO:110至119、139、141、143、146和201至309中任一个所示的CRISPR重复序列或其活性变体或片段及SEQ ID NO:120至128、140、142、145、147和148中任一个分别所示的tracrRNA序列或其活性变体或片段的指导序列结合。该RGN系统被进一步描述于本说明书的实施例1及表1中。

[0069] 本领域中众所周知,PAM序列对给定核酸酶的特异性受酶浓度的影响(参见,例如,Karvelis等人,(2015) Genome Biol 16:253),其可以通过改变用于表达该RGN的启动子或被递送至该细胞、细胞器或胚胎的核糖核蛋白复合物的量来修饰。

[0070] 在通过该RGN的结合及切割依赖于PAM序列的这些实施方案中,在识别其对应的PAM序列时,RGN可在特异性切割位点切割该目标核苷酸序列。如本文所使用的,切割位点是由目标核苷酸序列内的两个特定核苷酸组成,其之间的核苷酸序列被RGN切割。该切割位点可包含在5'或3'方向上自该PAM起的第1和第2、第2和第3、第3和第4、第4和第5、第5和第6、第7和第8、或第8和第9个核苷酸。在一些实施方案中,切割位点可在5'或3'方向上自该PAM起的10、11、12、13、14、15、16、17、18、19或20个以上的核苷酸。在一些实施方案中,切割位点距离该PAM 4个核苷酸。在其他实施方案中,切割位点距离该PAM至少15个核苷酸。因RGN可切割目标核苷酸序列,导致交错的末端,所以,在一些实施方案中,基于该多核苷酸的正(+)链上的两个核苷酸离开该PAM的距离及该多核苷酸的负(-)链上的两个核苷酸离开该PAM的距离来定义该切割位点。

[0071] IV. 编码RNA指导核酸酶、CRISPR RNA和/或tracrRNA的核苷酸

[0072] 本公开内容提供包含本发明公开的CRISPR RNA、tracrRNA和/或sgRNA的多核苷酸及包含编码本发明公开的RNA指导核酸酶、CRISPR RNA、tracrRNA和/或sgRNA的核苷酸序列的多核苷酸。本发明公开的多核苷酸包括包含或编码CRISPR重复序列的那些多核苷酸,该CRISPR重复序列包含SEQ ID NO:110至119、139、141、143、146和201至309的核苷酸序列中任一个或其活性变体或片段,其当被包含于指导RNA内时能够将关联RNA指导核酸酶的序列

特异性结合导向至所关注的目标序列。还公开了包含或编码tracrRNA的多核苷酸,该tracrRNA包含SEQ ID NO:120至128、140、142、145、147和148的核苷酸序列中任一个或其活性变体或片段,其当被包含于指导RNA内时能够将关联RNA指导核酸酶的序列特异性结合导向至所关注的目标序列。还提供编码RNA指导核酸酶的多核苷酸,该RNA指导核酸酶包含如SEQ ID NO:1至109所示的氨基酸序列中任一个及其活性片段或变体,其保持以RNA指导的序列特异性方式结合至目标核苷酸序列的能力。

[0073] 术语“多核苷酸”的使用不旨在将本公开内容局限于包含DNA的多核苷酸,但可考虑这样的DNA多核苷酸。本领域技术人员将认识到多核苷酸可包含核糖核苷酸(RNA)及核糖核苷酸与脱氧核糖核苷酸的组合。这样的脱氧核糖核苷酸和核糖核苷酸包括天然存在的分子及合成类似物二者。这些包括例如肽核酸(PNA)、PNA-DNA嵌合体(chimer)、锁核酸(LNA)和硫代磷酸酯连接序列。本文所公开的多核苷酸还涵盖所有形式的序列,包括但不限于单链形式、双链形式、DNA-RNA杂合体、三链螺旋结构、茎环结构、环状形式(例如,包括环状RNA)等等。

[0074] 在一些实施方案中,编码RGN的核酸分子可针对在所关注的生物体中的表达而被密码子优化。“密码子优化的”编码序列为使其密码子使用频率设计成模拟优选密码子使用频率或特定宿主细胞的转录条件的多核苷酸编码序列。由于核酸水平上的一种或多种密码子的改变使得翻译的氨基酸序列未改变,所以该特定宿主细胞或生物体中的表达被增强。核酸分子可以全部或部分地被密码子优化。密码子表和提供一大范围生物体的偏好信息的其他参考文献在本领域中是可得(参见,例如,Campbell及Gowri (1990) Plant Physiol. 92:1-11,有关植物优选密码子使用的讨论)。本领域中用于合成植物优选基因的方法是可得。参见,例如,第5,380,831号和第5,436,391号美国专利及Murray等人,(1989) Nucleic Acids Res. 17:477-498,它们通过引用并入本文。

[0075] 在一些实施方案中,可在表达盒中提供编码本文中提供的RGN、crRNA、tracrRNA和/或sgRNA的多核苷酸,以用于在体外表达或在所关注的细胞、细胞器、胚胎或生物体中表达。该盒可包括5'和3'调节序列,该5'和3'调节序列可操作地连接至编码本文中提供的允许该多核苷酸表达的RGN、crRNA、tracrRNA和/或sgRNA的多核苷酸。该盒可额外含有至少一种额外基因或基因元件,以共转化至该生物体内。如果包括额外基因或元件,则该组成被可操作地连接。术语“可操作地连接”旨在表达两个或更多个元件之间的功能性连接。例如,启动子与所关注的编码区域(例如,对RGN、crRNA、tracrRNA和/或sgRNA编码的区域)之间的可操作地连接为允许所关注的编码区域表达的功能性连接。可操作地连接的元件可为连续的或非连续的。当用于指两个蛋白质编码区域的连接时,通过可操作地连接意指该编码区域在相同的阅读框中。作为另一种选择,可在多个表达盒上提供额外基因或元件。例如,编码本发明公开的RGN的核苷酸序列可存在于一个表达盒上,而编码crRNA、tracrRNA或完整指导RNA的核苷酸序列可在分开的表达盒上。这样的表达盒被提供有多个限制性位点和/或重组位点,以使该多核苷酸的插入受调节区域的转录调节。该表达盒可额外含有选择性标记基因。

[0076] 该表达盒在该5'-3'转录方向上可包括:转录(而在一些实施方案中,翻译)起始区域(即,启动子)、本发明的RGN-、crRNA-、tracrRNA-和/或sgRNA-编码多核苷酸和在所关注的生物体中起作用的转录(而在一些实施方案中,翻译)终止区域(即,终止区域)。本发明的

启动子能够在宿主细胞中导向或驱动编码序列的表达。该调节区域(例如,启动子、转录调节区域和翻译终止区域)可与该宿主细胞或彼此为内源的或异源的。如本文所使用的,关于序列的“异源”为源自外来物种的序列,或者,如果来自相同物种,则为通过精心的人为干预从其在组合物和/或基因组基因座中的天然形式被实质上修饰的序列。如本文中所使用的,嵌合基因包含与转录起始区域可操作地连接的编码序列,该转录起始区域与该编码序列是异源的。

[0077] 方便的终止区域可从根癌农杆菌(*A. tumefaciens*)的Ti质粒获得,诸如,章鱼碱合酶及胭脂碱合酶终止区域。也可参见Guerineau等人,(1991) *Mol. Gen. Genet.* 262:141-144; Proudfoot (1991) *Cell* 64:671-674; Sanfacon等人,(1991) *Genes Dev.* 5:141-149; Mogen等人,(1990) *Plant Cell* 2:1261-1272; Munroe等人,(1990) *Gene* 91:151-158; Ballas等人,(1989) *Nucleic Acids Res.* 17:7891-7903; 和Joshi等人,(1987) *Nucleic Acids Res.* 15:9627-9639。

[0078] 额外调节信号包括但不限于:转录起始初始位点(transcriptional initiation start site)、操纵基因、激活子、增强子、其他调节元件、核糖体结合位点、起始密码子、终止信号等等。参见,例如,第5,039,523号和第4,853,331号美国专利;EPO 0480762A2; Sambrook等人,(1992) *Molecular Cloning: A Laboratory Manual*, ed. Maniatis等人,(冷泉港实验室出版社,冷泉港,纽约(Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.)), 后称“Sambrook 11”; Davis等人, eds. (1980) *Advanced Bacterial Genetics* (冷泉港实验室出版社,冷泉港,纽约((Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.))) 以及其中引用的参考文献。

[0079] 在制备表达盒时,可操纵各种DNA片段,以便在适宜的取向上及恰当的情况下在适宜的阅读框中对该DNA序列进行提供。为此,可以运用衔接蛋白或接头来连接该DNA片段,或者可涉及其他操纵以对方便的限制位点、除去多余的DNA、除去限制位点等进行提供。为此目的,可涉及体外诱变、引物修复、限制、退火、重新置换,例如,转换和颠换。

[0080] 很多启动子可用于本发明的实施。可基于期望结局来选择启动子。该核酸可以与组成型、诱导型、生长阶段特异性、细胞类型特异性、组织优选、组织特异性的启动子或其他启动子组合,用于所关注的生物体中的表达。参见,例如,WO 99/43838中及第8,575,425号;第7,790,846号;第8,147,856号;第8,586,832号;第7,772,369号;第7,534,939号;第6,072,050号;第5,659,026号;第5,608,149号;第5,608,144号;第5,604,121号;第5,569,597号;第5,466,785号;第5,399,680号;第5,268,463号;第5,608,142号;以及第6,177,611号美国专利中所示的启动子;它们通过引用并入本文。

[0081] 对于在植物中的表达,组成型启动子也包括CaMV 35S启动子(Odell等人,(1985) *Nature* 313:810-812); 水稻肌动蛋白(McElroy等人,(1990) *Plant Cell* 2:163-171); 泛素(Christensen等人,(1989) *Plant Mol. Biol.* 12:619-632及Christensen等人,(1992) *Plant Mol. Biol.* 18:675-689); pEMU (Last等人,(1991) *Theor. Appl. Genet.* 81:581-588); 和MAS (Velten等人,(1984) *EMBO J.* 3:2723-2730)。

[0082] 诱导型启动子的实例为:可通过缺氧或冷应激诱导的Adh1启动子、可通过热应激诱导的Hsp70启动子、可由光诱导的PPDK启动子及PEP羧化酶(pepcarboxylase)启动子。同样有用的为化学诱导型的启动子,诸如,安全剂诱导的In2-2启动子(第5,364,780号美国专

利)、生长素诱导的且是绒毡层特异性的但在愈合组织中也有活性的Axig1激活子(PCT US 01/22169)、类固醇反应性启动子(参见,例如,Schena等人,(1991) Proc.Natl.Acad.Sci.USA 88:10421-10425及McNellis等人,(1998) Plant J.14(2):247-257中的雌激素诱导的ERE启动子和糖皮质激素诱导型启动子)和四环素诱导型和四环素阻遏型启动子(参见,例如,Gatz等人,(1991) Mol.Gen.Genet.227:229-237及第5,814,618号和第5,789,156号美国专利),它们通过引用并入本文。

[0083] 可采用组织特异性或组织优选的启动子来靶向特定组织内的表达构建体的表达。在某些实施方案中,组织特异性或组织优选的启动子在植物组织中是有活性的。在植物中受发育控制的启动子的实例包括在诸如叶、根、果实、种子或花的某些组织中优选地起始转录的启动子。“组织特异性”启动子为仅在某些组织中起始转录的启动子。与基因的组成型表达不同,组织特异性表达是几种水平的基因调节相互作用的结果。因此,来自同源或密切相关的植物物种的启动子可优选地用于在特定组织中实现转基因的有效且可靠的表达。在一些实施方案中,表达包含组织优选的启动子。“组织优选的”启动子是优选地在,但不一定完全在或仅在某些组织中起始转录的启动子。

[0084] 在一些实施方案中,编码RGN、crRNA和/或tracrRNA的核酸分子包含细胞类型特异性启动子。“细胞类型特异性”启动子为主要在一种或多种器官的某些细胞类型中驱动表达的启动子。例如,其中在植物中起作用的细胞类型特异性启动子可为主要活性的植物细胞的一些实例包括BETL细胞、根、叶中的维管细胞、柄细胞和茎细胞。核酸分子还可包括细胞类型优选的启动子。“细胞类型优选的”启动子为主要驱动大部分在,但不一定完全在或仅在一种或多种器官中的某些细胞类型中的表达的启动子。例如,其中在植物中起作用的细胞类型优选的启动子可具优选活性的植物细胞的一些实例包括BETL细胞、根、叶中的维管细胞、柄细胞和茎细胞。

[0085] 编码该RGN、crRNA、tracrRNA和/或sgRNA的核酸序列可与例如通过用于体外mRNA合成的噬菌体RNA聚合酶识别的启动子序列可操作地连接。在此类实施方案中,可纯化该体外转录的RNA,以用于本文所述的方法。例如,该启动子序列可为T7、T3或SP6启动子序列,或T7、T3或SP6启动子序列的变体。在此类实施方案中,可纯化所表达蛋白和/或RNA,以用于本文所述的基因组修饰方法。

[0086] 在某些实施方案中,编码该RGN、crRNA、tracrRNA和/或sgRNA的多核苷酸也可与多腺苷酸化信号(例如,SV40 polyA信号及植物中起作用的其他信号)和/或至少一个转录终止序列连接。另外,编码该RGN的序列也可与编码至少一个核定位信号、至少一个细胞穿透域和/或能够将蛋白质运输到特定亚细胞位置的至少一个信号肽的序列连接,如本文中别处所述。

[0087] 编码该RGN、crRNA、tracrRNA和/或sgRNA的多核苷酸可存在于一个载体或多个载体中。“载体”是指用于将核酸转移、递送或引入宿主细胞内的多核苷酸组合物。适合的载体包括质粒载体、噬菌粒、粘粒、人工/微型染色体、转座子和病毒载体(例如,慢病毒载体、腺相关病毒载体、杆状病毒载体)。该载体可包含额外的表达控制序列(例如,增强子序列、Kozak序列、多腺苷酸化序列、转录终止序列)、选择性标记序列(例如,抗生素抗性基因)、复制起点等等。额外信息可在“Current Protocols in Molecular Biology”Ausubel等人, John Wiley&Sons,纽约,2003;或“Molecular Cloning:A Laboratory Manual”Sambrook&

Russell, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., 3rd edition, 2001中找到。

[0088] 该载体也可包含用于选择转化细胞的选择性标记基因。对于转化的细胞或组织的选择采用选择性标记基因。标记基因包括：编码抗生素抗性的基因，诸如，编码新霉素磷酸转移酶II (NEO) 和潮霉素磷酸转移酶 (HPT) 的基因；以及对诸如草铵膦、溴苯腈、咪唑啉酮及2,4-二氯苯氧乙酸盐 (2,4-D) 的除草性化合物赋予抗性的基因。

[0089] 在一些实施方案中，包含编码该RGN多肽的序列的表达盒或载体还可包含编码crRNA和/或tracrRNA或者被组合以产生指导RNA的crRNA及tracrRNA的序列。编码该crRNA和/或tracrRNA的序列可与至少转录控制序列可操作地连接，用于该crRNA和/或tracrRNA在所关注的生物体或宿主细胞中的表达。例如，编码该crRNA和/或tracrRNA的多核苷酸可与通过RNA聚合酶III (Pol III) 识别的启动子序列可操作地连接。适合的Pol III启动子的实例包括但不限于哺乳动物U6、U3、H1及7SL RNA启动子及水稻U6及U3启动子。

[0090] 如所指示的，包含编码该RGN、crRNA、tracrRNA和/或sgRNA的核苷酸序列的表达构建体可用于转化所关注的生物体。用于转化的方法涉及将核苷酸构建体引入所关注的生物体内。通过“引入”旨在将该核苷酸构建体引至该宿主细胞，由此使得该构建体进入该宿主细胞内部。本发明的方法不需要一种将核苷酸构建体引至宿主生物体，但是该核苷酸构建体进入宿主生物体的至少一个细胞内部的特定方法。该宿主细胞可为真核细胞或原核细胞。在特定实施方案中，真核宿主细胞为植物细胞、哺乳动物细胞或昆虫细胞。用于将核苷酸构建体引入植物及其他宿主细胞内的方法为本领域中已知，其包括但不限于：稳定转化方法、瞬时转化方法及病毒介导方法。

[0091] 这些方法得到转化的生物体，诸如植物，包括：整株植物以及植物器官（例如，叶、茎、根等）、种子、植物细胞、繁殖体、胚胎及其后代。植物细胞可被分化，也可不被分化（例如，愈伤组织、悬浮培养细胞、原生质体、叶细胞、根细胞、韧皮部细胞、花粉）。

[0092] “转基因生物体”或“转化生物体”或“稳定转化的”生物体或细胞或组织是指已并入或整合了编码本发明的RGN、crRNA和/或tracrRNA的多核苷酸的生物体。应当认识到，其他外源或内源核酸序列或DNA片段也可被并入该宿主细胞内。农杆菌介导及基因枪介导的转化仍然为用于植物细胞转化而主要采用的两种方法。然而，宿主细胞的转化可通过感染、转染、显微注射、电穿孔、显微投影 (microprojection)、基因枪或粒子轰击、电穿孔、二氧化硅/碳纤维、超声介导的、PEG介导的、磷酸钙共沉淀、聚阳离子DMSO技术、DEAE葡聚糖程序和病毒介导的、脂质体介导的等等进行。编码RGN、crRNA和/或tracrRNA的多核苷酸的病毒介导的引入包括逆转录病毒、慢病毒、腺病毒和腺相关病毒介导的引入及表达，以及花椰菜花叶病毒、双生病毒和RNA植物病毒的使用。

[0093] 转化操作流程以及用于将多肽或多核苷酸序列引入植物内的操作流程可随针对转化所靶向的宿主细胞的类型（例如，单子叶植物或双子叶植物细胞）而不同。用于转化的方法为本领域中已知，且包括第8,575,425号、第7,692,068号、第8,802,934号、第7,541,517号美国专利中阐述的那些方法，这些专利中的每一篇都通过引用并入本文。也可参见Rakoczy-Trojanowska, M. (2002) Cell Mol Biol Lett. 7:849-858; Jones等人, (2005) Plant Methods 1:5; Rivera等人, (2012) Physics of Life Reviews 9:308-345; Bartlett等人, (2008) Plant Methods 4:1-12; Bates, G.W. (1999) Methods in Molecular Biology

111:359-366;Binns及Thomashow(1988),Microbiology 42:575-606中的Annual Reviews;Christou,P.(1992)The Plant Journal 2:275-281;Christou,P.(1995)Euphytica 85:13-27;Tzfira等人,(2004)TRENDS in Genetics 20:375-383;Yao等人,(2006)Journal of Experimental Botany 57:3737-3746;Zupan和Zambryski(1995)Plant Physiology 107:1041-1047;Jones等人,(2005)Plant Methods 1:5。

[0094] 转化可导致核酸稳定或瞬时并入细胞内。“稳定转化”旨在表达引入宿主细胞内的核苷酸构建体整合至该宿主细胞的基因组内,且能够被其后代遗传。“瞬时转化”旨在表达多核苷酸被引入该宿主细胞内而不整合至该宿主细胞的基因组内。

[0095] 用于叶绿体转化的方法为本领域中已知。参见,例如,Svab等人,(1990)Proc.Natl.Acad.Sci.USA 87:8526-8530;Svab及Maliga(1993)Proc.Natl.Acad.Sci.USA 90:913-917;Svab及Maliga(1993)EMBO J.12:601-606。该方法取决于含有选择性标记的DNA的粒子枪递送及通过同源重组将该DNA靶向至该质体基因组。另外,质体转化可通过细胞核编码的及质体导向的RNA聚合酶的组织优选表达来反式活化(transactivation)沉默质体携带的转基因来完成。McBride等人(1994)在Proc.Natl.Acad.Sci.USA 91:7301-7305已报导了此系统。

[0096] 已经被转化的细胞可按照传统方式生长成转基因生物体,诸如,植物。参见,例如,McCormick等人,(1986)Plant Cell Reports 5:81-84。然后,可以使这些植物生长,且用相同转化品系或不同品系授粉,且鉴定出具有期望表型特征的组成型表达的所得杂合体。可生长两代或更多代,以确保稳定维持并遗传期望表型特征的表达,然后,收获种子,以确保已经达成期望表型特征的表达。以此方式,本发明提供具有稳定地并入其基因组内的本发明的核苷酸构建体(例如,本发明的表达盒)的转化种子(也称为“转基因种子”)。

[0097] 作为另一种选择,可将已被转化的细胞引入生物体内。这些细胞可源自该生物体,其中该细胞以离体措施被转化。

[0098] 本文提供的序列可用于转化任何植物物种,包括但不限于单子叶植物及双子叶植物。所关注的植物的实例包括但不限于:玉蜀黍(玉米)、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦和油菜、芸苔属物种(Brassica sp.)、苜蓿、黑麦、小米、红花、花生、甘薯、木薯、咖啡、椰子、菠萝、柑橘树、可可、茶、香蕉、鳄梨、无花果、番石榴、芒果、橄榄、木瓜、腰果、澳洲胡桃、杏仁、燕麦、蔬菜、观赏植物以及针叶树。

[0099] 蔬菜包括但不限于:西红柿、莴苣、绿豆、青豆、豌豆和诸如黄瓜(cucumber)、哈密瓜(cantaloupe)及香瓜(musk melon)的黄瓜(Cucumis)属的成员。观赏植物包括但不限于:杜鹃花、绣球花、芙蓉、玫瑰、郁金香、水仙、矮牵牛、康乃馨、猩猩木和菊花。优选地,本发明的植物为农作物(例如,玉米、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦、油菜等)。

[0100] 如本文所使用的,术语“植物”包括:植物细胞、植物原生质体、植物可再生所自的植物细胞组织培养物、植物愈伤组织、植物块和在植物或植物的诸如胚胎、花粉、胚珠、种子、叶子、花、枝、果实、仁、穗、玉米穗轴、壳、茎、根、根尖、花药等等的局部中是完整的植物细胞。谷物旨在表达出于种植或繁殖该物种之外的目的而由商业种植者生产的成熟种子。再生植物的后代、变体及突变体也包括于本发明的范围内,前提条件是这些局部包含所引

入的多核苷酸。还提供了保持了本文中公开的序列的经处理的植物产品或副产物,例如,包括豆粕。

[0101] 编码该RGN、crRNA和/或tracrRNA的多核苷酸也可用于转化任何原核物种,包括但不限于:古菌和细菌(例如,芽孢杆菌属(*Bacillus* sp.)、克雷伯氏菌属(*Klebsiella* sp.)、链霉菌属(*Streptomyces* sp.)、根瘤菌属(*Rhizobium* sp.)、埃希氏菌属(*Escherichia* sp.)、假单胞菌属(*Pseudomonas* sp.)、沙门氏菌属(*Salmonella* sp.)、志贺氏杆菌属(*Shigella* sp.)、弧菌属(*Vibrio* sp.)、耶尔森菌属(*Yersinia* sp.)、支原体菌属(*Mycoplasma* sp.)、农杆菌属(*Agrobacterium*)、乳酸乳杆菌属(*Lactobacillus* sp.))。

[0102] 编码该RGN、crRNA和/或tracrRNA的多核苷酸可用于转化任何真核物种,包括但不限于:动物(例如,哺乳动物、昆虫、鱼类、鸟类和爬行动物)、真菌、变形虫、藻类和酵母。

[0103] 传统的基于病毒和非病毒的基因转移方法可用于将核酸引入哺乳动物细胞或目标组织中。这样的方法可用于将编码CRISPR系统组成的核酸施用培养中的细胞、或宿主生物体中的细胞。非病毒载体递送系统包括:DNA质粒、RNA(例如,本文描述的载体的转录物)、裸核酸和与诸如脂质体的递送载剂复合的核酸。病毒载体递送系统包括DNA及RNA病毒,其在递送至细胞后具有附加型或整合的基因组。有关基因治疗程序的综述,参见Anderson, *Science* 256:808-813(1992); Nabel&Feigner, *TIBTECH* 11:211-217(1993); Mitani&Caskey, *TIBTECH* 11:162-166(1993); Dillon, *TIBTECH* 11:167-175(1993); Miller, *Nature* 357:455-460(1992); Van Brunt, *Biotechnology* 6(10):1149-1154(1988); Vigne, *Restorative Neurology and Neuroscience* 8:35-36(1995); Kremer&Perricaudet, *British Medical Bulletin* 51(1):31-44(1995); Haddada等人,于Current Topics in Microbiology and Immunology, Doerfler及Bohm(编者)(1995); 和Yu等人, *Gene Therapy* 1:13-26(1994)。

[0104] 核酸的非病毒递送方法包括脂质体转染(lipofection)、核转染、显微注射、基因枪、病毒体、脂质体、免疫脂质体、聚阳离子或脂质:核酸缀合物、裸DNA、人工病毒颗粒和DNA的药剂增强摄取。例如,第5,049,386号、第4,946,787号和第4,897,355号美国专利中描述了脂质体转染,且脂质体转染试剂是市售的(例如,TransfectamTM及LipofectinTM)。适合用于多核苷酸的有效受体识别脂质体转染的阳离子及中性脂质包括Feigner的W0 91/17424; W0 91/16024中的那些阳离子和中性脂质。递送可为至细胞(例如,在体外或离体施用)或目标组织(例如,体内施用)。包括靶向的脂质体诸如免疫脂质复合物的脂质:核酸复合物的制备为本领域技术人员所熟知(参见,例如,Crystal, *Science* 270:404-410(1995); Blaese等人, *Cancer Gene Ther.* 2:291-297(1995); Behr等人, *Bioconjugate Chem.* 5:382-389(1994); Remy等人, *Bioconjugate Chem.* 5:647-654(1994); Gao等人, *Gene Therapy* 2:710-722(1995); Ahmad等人, *Cancer Res.* 52:4817-4820(1992); 第4,186,183号、第4,217,344号、第4,235,871号、第4,261,975号、第4,485,054号、第4,501,728号、第4,774,085号、第4,837,028号和第4,946,787号美国专利)。

[0105] 使用基于RNA或DNA病毒的系统来递送核酸利用高度进化的过程将病毒靶向至体内的特异性细胞且将该病毒载荷运输至该细胞核。病毒载体可被直接施用患者(体内),或者该病毒载体可用于体外处理细胞,并且可视情况将经修饰的细胞施用患者(离体)。传统的基于病毒的系统可包括用于基因转移的逆转录病毒、慢病毒、腺病毒、腺关联及单纯疱疹

病毒载体。用逆转录病毒、慢病毒和腺相关病毒基因转移方法,在宿主基因组中的整合是可能的,常常导致所插入的转基因的长期表达。另外,在许多不同细胞类型及目标组织中已经观察到高转导效率。

[0106] 可通过并入外来包膜蛋白而改变逆转录病毒的向性(tropism),从而扩展目标细胞的潜在目标群体。慢病毒载体为能够转导或感染非分裂细胞并且通常情况下产生高病毒滴度的逆转录病毒载体。因此,逆转录病毒基因转移系统的选择将依赖于该目标组织。逆转录病毒载体由顺式作用长末端重复序列构成,该顺式作用长末端重复序列具有达到6-10kb外来序列的包装能力。最小顺式作用LTR足够用于载体的复制及包装,然后,使用其将治疗基因整合至该目标细胞内,以提供永久转基因表达。广泛使用的逆转录病毒载体包括:基于鼠白血病病毒(MuLV)、长臂猿白血病病毒(GaLV)、猿猴免疫缺陷病毒(SIV)、人免疫缺陷病毒(HIV)和其组合的载体(参见,例如,Buchscher等人,J.Viral.66:2731-2739(1992);Johann等人,J.Viral.66:1635-1640(1992);Sommnerfelt等人,Viral.176:58-59(1990);Wilson等人,J.Viral.63:2374-2378(1989);Miller等人,1.Viral.65:2220-2224(1991);PCT/US94/05700)。

[0107] 在优选瞬时表达的应用中,可以使用基于腺病毒的系统。基于腺病毒的载体能够在许多细胞类型中有非常高的转导效率,并且不需要细胞分裂。用这样的载体,已经获得了高滴度及高表达水平。此载体可在相对简单的系统中大量产生。腺相关病毒(“AAV”)载体也可例如在核酸及肽的体外产生中用于用目标核酸转导细胞,且用于体内及离体基因治疗程序(参见,例如,West等人,Virology 160:38-47(1987);第4,797,368号美国专利;WO 93/24641;Katin,Human Gene Therapy 5:793-801(1994);Muzyczka,1.Clin.Invest.94:1351(1994))。重组AAV载体的构建描述于很多出版物中,包括第5,173,414号美国专利;Tratschin等人,Mol.Cell.Biol.5:3251-3260(1985);Tratschin等人,Mol.Cell.Biol.4:2072-2081(1984);Hermonat&Muzyczka,PNAS 81:6466-6470(1984);和Samulski等人,1.Viral.63:03822-3828(1989)。包装细胞通常情况下用于形成能够感染宿主细胞的病毒颗粒。这样的细胞包括包装腺病毒的293细胞和包装逆转录病毒的 Ψ J2细胞或PA317细胞。

[0108] 基因治疗中使用的病毒载体通常通过产生将核酸载体包装于病毒颗粒内的细胞系而被产生。该载体通常情况下含有用于包装且随后整合至宿主内所需的最小病毒序列,其他病毒序列被用于要表达的多核苷酸的表达盒置换。缺失的病毒功能通常情况下通过该包装细胞系以反式提供。例如,基因治疗中使用的AAV载体通常情况下仅具有来自包装及整合至该宿主基因组内所需的AAV基因组的ITR序列。病毒DNA被包装于细胞系中,该细胞系含有辅助质粒,其编码其他AAV基因,即,rep和cap,但缺少ITR序列。

[0109] 也可用腺病毒作为辅助病毒来感染该细胞系。该辅助病毒促进AAV载体的复制及来自该辅助质粒的AAV基因的表达。由于缺少ITR序列,未以显著量包装该辅助质粒。腺病毒的污染可通过例如热处理来降低,腺病毒对该热处理比AAV更敏感。将核酸递送至细胞的额外方法为本领域技术人员已知。参见,例如,US20030087817,其通过引用并入本文。

[0110] 在一些实施方案中,用本文描述的一种或多种载体瞬时地或非瞬时地转染宿主细胞。在一些实施方案中,细胞如其天然存在于个体中那样被转染。在一些实施方案中,被转染的细胞取自个体。在一些实施方案中,细胞来源于取自个体的细胞,诸如细胞系。用于组织培养的各种细胞系为本领域中已知。细胞系的实例包括但不限于:C8161、CCRF-CEM、

MOLT、mIMCD-3、NHDF、HeLaS3、Huh1、Huh4、Huh7、HUVEC、HASMC、HEKn、HEKa、MiaPaCel 1、Panel、PC-3、TF1、CTLL-2、CIR、Rat6、CVI、RPTE、A10、T24、182、A375、ARH-77、Calu1、SW480、SW620、SKOV3、SK-UT、CaCo2、P388D1、SEM-K2、WEHI-231、HB56、TIB55、lurkat、145.01、LRMB、Bcl-1、BC-3、IC21、DLD2、Raw264.7、NRK、NRK-52E、MRC5、MEF、Hep G2、HeLa B、HeLa T4、COS、COS-1、COS-6、COS-M6A、BS-C-1猴肾上皮细胞、BALB/3T3小鼠胚胎成纤维细胞、3T3 Swiss、3T3-L1、132-d5人胎儿成纤维细胞；10.1小鼠成纤维细胞、293-T、3T3、721、9L、A2780、A2780ADR、A2780cis、A172、A20、A253、A431、A-549、ALC、B16、B35、BCP-I细胞、BEAS-2B、bEnd.3、BHK-21、BR 293、BxPC3、C3H-10T1/2、C6/36、Ca1-27、CHO、CHO-7、CHO-IR、CHO-K1、CHO-K2、CHO-T、CHO Dhfr-/-、COR-L23、COR-L23/CPR、COR-L235010、CORL23/R23、COS-7、COV-434、CML T1、CMT、CT26、D17、DH82、DU145、DuCaP、EL4、EM2、EM3、EMT6/AR1、EMT6/AR10.0、FM3、H1299、H69、HB54、HB55、HCA2、HEK-293、HeLa、Hepalclc7、HL-60、HMEC、HT-29、lurkat、1Y细胞、K562细胞、Ku812、KCL22、KG1、KY01、LNCap、Ma-Me1 1-48、MC-38、MCF-7、MCF-10A、MDA-MB-231、MDA-MB-468、MDA-MB-435、MDCKII、MDCKII、MOR/0.2R、MONO-MAC 6、MTD-1A、MyEnd、NCI-H69/CPR、NCI-H69/LX10、NCI-H69/LX20、NCI-H69/LX4、NIH-3T3、NALM-1、NW-145、OPCN/OPCT细胞系、Peer、PNT-1A/PNT2、RenCa、RIN-5F、RMA/RMAS、Saos-2细胞、Sf-9、SkBr3、T2、T-47D、T84、THP1细胞系、U373、U87、U937、VCaP、Vero细胞、WM39、WT-49、X63、YAC-1、YAR和其转基因品种。细胞系可从本领域技术人员已知的各种来源获得(参见,例如,美国典型菌种保藏中心(American Type Culture Collection)(ATCC)(马纳沙斯,VA.))。

[0111] 在一些实施方案中,使用用本文描述的一种或多种载体转染的细胞建立包含一种或多种来源于载体的序列的新细胞系。在一些实施方案中,使用用如本文描述的CRISPR系统的组成瞬时转染(诸如,通过一种或多种载体的瞬时转染,或用RNA转染)并通过CRISPR复合物的活性修饰的细胞建立包含含有该修饰但缺少任何其他外源序列的细胞的新细胞系。在一些实施方案中,使用用本文描述的一种或多种载体瞬时或非瞬时转染的细胞或来源于这样的细胞的细胞系来评估一种或多种测试化合物。

[0112] 在一些实施方案中,使用本文描述的一种或多种载体来产生非人的转基因动物或转基因植物。在一些实施方案中,转基因动物为哺乳动物,诸如,小鼠、大鼠或兔子。

[0113] V. 多肽及多核苷酸的变体及片段

[0114] 本公开内容提供天然存在的(即,野生型)RNA指导核酸酶的活性变体及片段,其氨基酸序列如SEQ ID NO:1至109所示;以及天然存在的CRISPR重复序列的活性变体和片段,诸如,SEQ ID NO:110至119、139、141、143、146和201至309所示的序列;和天然存在的tracrRNA的活性变体和片段,诸如,SEQ ID NO:120至128、140、142、145、147和148所列的序列;和编码其的多核苷酸。还提供诸如SEQ ID NO:178-192所示的序列的天然存在的RGN辅助蛋白的活性变体及片段。

[0115] 尽管与所关注的多核苷酸或多肽相较下可以改变变体或片段的活性,但该变体和片段应保持所关注的多核苷酸或多肽的功能。例如,当与所关注的多核苷酸或多肽相较时,变体或片段可具有增加的活性、降低的活性、不同的活性谱或活性的任何其他改变。

[0116] 诸如本文所公开的天然存在的RGN多肽的片段及变体将保持序列特异性的RNA指导的DNA结合活性。在特定实施方案中,诸如本文所公开的天然存在的RGN多肽的片段及变

体将保持核酸酶活性(单链或双链)。

[0117] 当指导RNA(包含tracrRNA)的局部以序列特异性方式结合至RNA指导核酸酶(与指导RNA复合的)且将RNA指导核酸酶(与指导RNA复合的)指导至目标核苷酸序列时,诸如本文所公开的天然存在的CRISPR重复序列的片段及变体将保持该能力。

[0118] 诸如本文所公开的天然存在的tracrRNA的片段及变体当作为指导RNA(包含CRISPR RNA)的局部时将保持以序列特异性方式将RNA指导核酸酶(与该指导RNA复合的)指导至目标核苷酸序列的能力。

[0119] 诸如本文所公开的天然存在的RGN辅助蛋白的片段及变体(当RGN系统(即,RGN蛋白及指导RNA)的局部时)将保持允许该RGN系统以序列特异性方式与目标核苷酸序列结合的能力。

[0120] 术语“片段”是指本发明的多核苷酸或多肽序列的部分。“片段”或“生物活性部分”包括多核苷酸,该多核苷酸包含足够数量的连续核苷酸,以保持该生物活性(即,当包含于指导RNA内时,结合RNA至并以序列特异性方式将RGN导向至目标核苷酸序列)。“片段”或“生物活性部分”包括多肽,该多肽包含足够数量的连续氨基酸残基,以保持该生物活性(即,当与指导RNA复合时,以序列特异性方式与目标核苷酸序列结合)。该RGN蛋白的片段包括由于替代的下游初始位点的使用而比全长序列短的那些。RGN蛋白的生物活性部分可为包含例如SEQ ID NO:1至109的10、25、50、100、150、200、250、300、350、400、450、500、550、600个或更多个连续氨基酸残基的多肽。可通过重组技术制备这些生物活性部分且评估其序列特异性的RNA指导的DNA结合活性。CRISPR重复序列的生物活性片段可包含SEQ ID NO:110至119、139、141、143、146和201至309中任一个的至少8个连续氨基酸。CRISPR重复序列的生物活性部分可为包含例如SEQ ID NO:110至119、139、141、143、146和201至309中任一个的8、9、10、11、12、13、14、15、16、17、18、19或20个连续核苷酸的多核苷酸。tracrRNA的生物活性部分可为包含例如SEQ ID NO:120至128、140、142、145、147和148中任一个的8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25、26、27、28、29、30、35、40、45、50、55、60、65、70、75、80个或更多个连续核苷酸的多核苷酸。该RGN辅助蛋白的片段包括由于替代的下游初始位点的使用而比该全长序列短的那些。RGN辅助蛋白的生物活性部分可为包含例如SEQ ID NO:178至192的10、25、50、100、150、200个或更多个连续氨基酸残基的多肽。可通过重组技术制备这些生物活性部分且评估其生物活性。

[0121] 一般而言,“变体”旨在指实质上相似的序列。对于多核苷酸,变体包含于天然多核苷酸内的一种或多种内部位点处的一种或多种核苷酸的缺失和/或添加和/或在天然多核苷酸中一种或多种位点处的一种或多种核苷酸的置换。如本文所使用的,“天然”或“野生型”多核苷酸或多肽分别包含天然存在的核苷酸序列或氨基酸序列。对于多核苷酸,保守变体包括因为该基因密码的简并性而编码所关注基因的天然氨基酸序列的那些序列。与例如用下面概述的聚合酶链式反应(PCR)及杂交技术一样,可用众所周知的分子生物学技术鉴定诸如这些的天然存在的等位基因变体。变体多核苷酸还包括合成取得的多核苷酸,诸如通过使用定点诱变产生的但其仍然编码所关注的多肽或多核苷酸的那些。一般而言,本文所公开的特定多核苷酸的变体与如通过本文别处描述的序列比对程序及参数所确定的那个特定多核苷酸具有至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高序列同一性。

[0122] 本文所公开的特定多核苷酸(即,该参考多核苷酸)的变体还可通过比较通过变体多核苷酸所编码的多肽与通过该参考多核苷酸所编码的多肽之间的序列同一性百分比来评估。可使用本文别处描述的序列比对程序及参数来计算任何两个多肽之间的序列同一性百分比。当通过两个多肽共有的序列同一性百分比的比较来评估本文所公开的任何给定的多核苷酸对时(其编码该两个多肽),该两个编码多肽之间的序列同一性百分比为至少约40%、45%、50%、55%、60%、65%、70%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的同一性。

[0123] 在特定实施方案中,本发明公开的多核苷酸编码RNA指导核酸酶多肽,该RNA指导核酸酶多肽包含的氨基酸序列与SEQ ID NO:1至109中任一个的氨基酸序列具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的同一性。

[0124] 本发明的RGN多肽的生物活性变体可相差少至约1-15个氨基酸残基、少至约1-10个(诸如,约6-10个)、少至5个、少至4个、少至3个、少至2个、或少至1个氨基酸残基。在具体实施方案中,多肽可包含N端或C端截短,该截短可至少包含从该多肽的N端或C端缺失10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、150、200、250、300、350、400、450、500、550、600个或更多个氨基酸。

[0125] 在某些实施方案中,本发明公开的多核苷酸编码RNA指导核酸酶辅助多肽,该RNA指导核酸酶辅助多肽包含与SEQ ID NO:178至192中任一个的氨基酸序列具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的同一性的核苷酸序列。

[0126] 本发明的RGN辅助多肽的生物活性变体可相差少至约1-15个氨基酸残基、少至约1-10个(诸如,约6-10个)、少至5个、少至4个、少至3个、少至2个、或少至1个氨基酸残基。在具体实施方案中,多肽可包含N端或C端截短,其可至少包含从该多肽的N端或C端缺失10、15、20、25、30、35、40、45、50、55、60、65、70、75、80、85、90、95、100、150、200个或更多个氨基酸。

[0127] 在某些实施方案中,本发明公开的多核苷酸包含或编码CRISPR重复序列,该CRISPR重复序列包含与如SEQ ID NO:110至119、139、141、143、146和201至309所示的核苷酸序列中任一个具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的同一性的核苷酸序列。

[0128] 本发明公开的多核苷酸可包含或编码tracrRNA,该tracrRNA包含与如SEQ ID NO:120至128、140、142、145、147和148所示的核苷酸序列中任一个具有至少40%、45%、50%、55%、60%、65%、70%、75%、80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或更高的同一性的核苷酸序列。

[0129] 本发明的CRISPR重复序列或tracrRNA的生物活性变体可相差少至约1-15个核苷酸、少至约1-10个(诸如,约6-10个)、少至5个、少至4个、少至3个、少至2个、或少至1个核苷酸。在具体实施方案中,多核苷酸可包括5'或3'截短,该截短可至少包含从该多核苷酸的5'或3'端缺失10、15、20、25、30、35、40、45、50、55、60、65、70、75、80个或更多个核苷酸。

[0130] 应当认识到,可对本文提供的RGN多肽、CRISPR重复序列及tracrRNA进行修饰,从而产生变体蛋白质及多核苷酸。可通过定点诱变技术的应用来引入人设计的变化。作为另一种选择,在结构上和/或功能上与本文所公开的序列有关的天然的、尚未知晓的或尚未被鉴定的多核苷酸和/或多肽也可被认为落入本发明的范围内。可在不改变该RGN蛋白功能的非保守区域中进行保守氨基酸置换。作为另一种选择,可进行改良该RGN的活性的修饰。

[0131] 变体多核苷酸和蛋白质还涵盖自诸如DNA混编(DNA shuffling)的诱变及重组程序取得的序列及蛋白质。用这种程序,操纵本文所公开的一种或多种不同的RGN蛋白质(例如,SEQ ID NO:1至109),以产生具有期望特性的新RGN蛋白质。以这种方式,重组多核苷酸文库是自包含序列区域的有关序列多核苷酸的群体产生的,该序列区域具有实质的序列同一性且可在体外或体内同源重组。例如,使用这种措施,编码所关注域的序列基序可在本文提供的RGN序列与其他已知RGN基因之间混编,以获得编码具有所关注的改良特性(诸如在酶的情况下,增加的 K_m)的蛋白质的新基因。这种DNA混编的策略是本领域中已知的。参见,例如,Stemmer(1994)Proc.Natl.Acad.Sci.USA 91:10747-10751;Stemmer(1994)Nature 370:389-391;Cramer等人,(1997)Nature Biotech.15:436-438;Moore等人,(1997)J.Mol.Biol.272:336-347;Zhang等人,(1997)Proc.Natl.Acad.Sci.USA 94:4504-4509;Cramer等人,(1998)Nature 391:288-291;和第5,605,793号及第5,837,458号美国专利。“混编的”核酸为通过混编程序(诸如本文阐述的任何混编程序)产生的核酸。混编的核酸为通过例如以人工方式及视情况而定的递归方式(实体或虚拟地)重组两个或更多个核酸(或字符串)产生的。一般而言,混编过程中使用一种或多种筛选步骤来鉴定所关注的核酸;此筛选步骤可在任何重组步骤之前或之后进行。在一些(但不是全部)混编实施方案中,期望在选择之前进行多轮重组,以增加要筛选的池的多样性。重组及选择的整个过程可视情况递归地重复。依赖于背景,混编可是指重组及选择的整个过程,或者,作为另一种选择,可仅是指该整个过程的重组部分。

[0132] 如本文在两个多核苷酸或多肽序列的背景中所使用的“序列同一性”或“同一性”涉及到当在指定的比较窗上比对最大对应性时为相同的两个序列中的残基。当使用与蛋白质有关的序列同一性百分比时,应当认识到,不相同的残基位置常常通过保守氨基酸置换而不同,其中氨基酸残基置换具有类似化学性质(例如,电荷或疏水性)的其他氨基酸残基,且因此不改变分子的功能性质。当序列在保守置换方面不同时,可上调序列同一性百分比,以校正该置换的保守性质。通过这种保守置换而不同的序列被称为具有“序列相似性”或“相似性”。用于进行此调整的手段为本领域技术人员所熟知。通常情况下,此涉及将保守置换作为一部分错配而非完全错配进行评分,从而增加序列同一性百分比。因此,例如,当对相同的氨基酸给予1分,而对非保守置换给予零分时,对保守置换给予0与1之间的得分。例如,如在程序PC/GENE(Intelligenetics,Mountain View,California)中所实施的那样,计算保守置换的得分。

[0133] 如本文所使用的,“序列同一性百分比”是指通过在比较窗上比较两个最佳比对序列所确定的值,其中该多核苷酸序列在该比较窗中的部分可包含相较于用于该两个序列的最佳比对的参考序列(不包含添加或缺失)的添加或缺失(即,空位)。通过确定在两个序列中出现相同的核酸碱基或氨基酸残基的位置的数目来求得匹配位置的数目;将匹配位置的数目除以比较窗中的位置总数;且将该结果乘以100,求得序列同一性百分比,来计算该百

分比。

[0134] 除非另有说明,否则本文提供的序列同一性/相似性值是指使用GAP版本10使用以下参数获得的值:使用GAP权重50及长度权重3及nwsgapdna.cmp评分矩阵的核苷酸序列的同一性%及相似性%;使用GAP权重8及长度权重2及BLOSUM62评分矩阵的氨基酸序列的同一性%及相似性%;或其任何等效程序。通过“等效程序”是指:对于所涉及的任何两个序列,当与GAP版本10产生的对应比对相较时,产生具有相同的核苷酸或氨基酸残基匹配及相同的序列同一性百分比的比对的任何序列比较程序。

[0135] 当为了进行相似性评分而使用所定义的氨基酸置换矩阵(例如,BLOSUM62)、空位存在罚分及空位延伸罚分比对两个序列,以达到那个序列对可能的最高分数时,两个序列被“最佳比对”。氨基酸置换矩阵及其在定量两个序列之间的相似性中的使用在本领域中是众所周知的,并且被描述于例如Dayhoff等人,(1978)“A model of evolutionary change in proteins(蛋白质中的进化变迁模型)”中;“Atlas of Protein Sequence and Structure(蛋白质序列及结构的图谱)”,Vol.5,Suppl.3(ed.M.O.Dayhoff),pp.345-352;Natl.Biomed.Res.Found.,Washington,D.C.;和Henikoff等人,(1992)Proc.Natl.Acad.Sci.USA 89:10915-10919中。BLOSUM62矩阵常常用作序列比对操作流程中的默认评分置换矩阵。空位存在罚分是针对单氨基酸空位在其中一个比对序列中的引入而施加的,而空位延伸罚分是针对被插入已经打开的空位处的每一个额外空氨基酸位置而施加的。该比对通过比对开始和结束处的每一个序列的氨基酸位置定义,并且可视情况通过在一个或两个序列中插入一个空位或多个空位来定义,以便达到最高可能分数。尽管可以手动完成最佳比对和评分,但是可通过使用计算机实施的比对算法(例如Altschul等人(1997)Nucleic Acids Res.25:3389-3402中描述并且在美国国家生物技术信息中心(National Center for Biotechnology Information)网站(在ncbi.nlm.nih.gov的万维网)向大众开放的gapped BLAST 2.0)而促进该过程。可以使用例如可通过www.ncbi.nlm.nih.gov获得的和Altschul等人在(1997)Nucleic Acids Res.25:3389-3402中描述的PSI-BLAST来准备包括多重比对的最佳比对。

[0136] 关于与参考序列最佳比对的氨基酸序列,氨基酸残基“对应于”该参考序列中在该比对中与该残基配对的位置。该“位置”由数字标示,该数字基于其相对于N端的位置依序鉴定该参考序列中的每一个氨基酸。归因于在确定最佳比对时必须考虑的缺失、插入、截短、融合等,所以,一般而言,通过简单地从N端计数所确定的测试序列中的氨基酸残基数目不一定与其在该参考序列中的对应位置的数目相同。例如,在所比对的测试序列中存在缺失的情况下,将不存在与缺失位点处的参考序列中的位置对应的氨基酸。当在所比对的参考序列中存在插入时,该插入将不对应于该参考序列中的任何氨基酸位置。在截短或融合的情况下,该参考序列或所比对的序列中可存在不对应于相应序列中的任何氨基酸的氨基酸段(stretch)。

[0137] VI. 抗体

[0138] 还涵盖针对本发明的RGN多肽或包含本发明该RGN多肽的核糖核蛋白,包括具有SEQ ID NO:1至109所述的氨基酸序列中任一个或其活性变体或其片段的那些RGN多肽或核糖核蛋白,或本发明该RGN辅助蛋白,包括具有SEQ ID NO:178至192所示的氨基酸序列中任一个或其活性变体或片段的那些RGN辅助蛋白的抗体。产生抗体的方法在本领域中是众所

周知的(参见,例如,Harlow及Lane(1988)Antibodies:A Laboratory Manual,冷泉港实验室,冷泉港,纽约(Cold Spring Harbor Laboratory,Cold Spring Harbor,N.Y.);和第4,196,265号美国专利)。这些抗体可用于试剂盒中,用于RGN多肽或核糖核蛋白的检测及分离。因此,本公开内容提供了包含特异性结合至本文描述的多肽或核糖核蛋白的抗体的试剂盒,该多肽或核糖核蛋白包括例如具有SEQ ID NO:1至109或178至192中任一个的序列的多肽。

[0139] VII.用于结合所关注的目标序列的系统 and 核糖核蛋白复合物及其制备方法

[0140] 本公开内容提供一种用于结合所关注的目标序列的系统,其中该系统包含至少一个指导RNA或编码该至少一个指导RNA的核苷酸序列以及至少一个RNA所指导核酸酶或编码该至少一个RNA指导核酸酶的核苷酸序列。该指导RNA与所关注的目标序列杂交,且也与该RGN多肽形成复合物,从而将该RGN多肽导向至与该目标序列结合。在这些实施方案中的一些实施方案中,RGN包含SEQ ID NO:1至109中任一个的氨基酸序列或其活性变体或片段。在各种实施方案中,指导RNA包含CRISPR重复序列,该CRISPR重复序列包含SEQ ID NO:110至119、139、141、143、146和201至309中任一个的核苷酸序列或其活性变体或片段。在一些特定实施方案中,指导RNA包含tracrRNA,该tracrRNA包含SEQ ID NO:120至128、140、142、145、147和148中任一个的核苷酸序列或其活性变体或片段。该系统的指导RNA可为单指导RNA或双指导RNA。在一些特定实施方案中,系统包含与该指导RNA异源的RNA所指导核酸酶,其中未发现该RGN与指导RNA在自然界中彼此复合(即,彼此结合)。

[0141] 在一些实施方案中,为结合至和/或切割目标多核苷酸,系统还包含至少一个RGN辅助蛋白。在这些实施方案中的一些实施方案中,系统还包含如SEQ ID NO:178-192所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG06369(SEQ ID NO:11)或其变体或片段的特定实施方案中,系统还可包含如SEQ ID NO:178-181所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG03847(SEQ ID NO:12)或其变体或片段的这些实施方案中的一些实施方案中,系统还可包含如SEQ ID NO:182-184所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG05625(SEQ ID NO:13)或其变体或片段的某些实施方案中,系统还可包含如SEQ ID NO:185-187所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG03524(SEQ ID NO:16)或其变体或片段的一些实施方案中,系统还可包含如SEQ ID NO:188-190所示的至少一个RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG03759(SEQ ID NO:14)或其变体或片段的特定实施方案中,系统还可包含如SEQ ID NO:191所示的RGN辅助蛋白或其活性变体或片段。在其中该RGN为APG05123(SEQ ID NO:15)或其变体或片段的某些实施方案中,系统还可包含如SEQ ID NO:192所示的RGN辅助蛋白或其活性变体或片段。

[0142] 本文提供的用于结合所关注的目标序列的系统可为核糖核蛋白复合物,其是与至少一个蛋白质结合的RNA的至少一个分子。本文提供的核糖核蛋白复合物包含作为该RNA组成的至少一个指导RNA及作为该蛋白质组成的RNA所指导核酸酶。此类核糖核蛋白复合物可自天然表达RGN多肽的细胞或生物体中纯化,且已经被工程化,以表达对所关注的目标序列特异的特定指导RNA。作为另一种选择,该核糖核蛋白复合物可自己用编码RGN多肽及指导RNA的多核苷酸转化且在允许该RGN多肽和指导RNA表达的条件下培养的细胞或生物体中纯化。因此,提供用于制造RGN多肽或RGN核糖核蛋白复合物的方法。此类方法包括:在该RGN多

肽(而在一些实施方案中,指导RNA)被表达的条件下,培养包含编码RGN多肽的核苷酸序列的细胞,而且在一些实施方案中,培养包含编码指导RNA的核苷酸序列的细胞。然后,可自所培养的细胞的裂解物中纯化该RGN多肽或RGN核糖核蛋白。

[0143] 用于自生物样本的裂解物中纯化RGN多肽或RGN核糖核蛋白复合物的方法在本领域中是已知的(例如,尺寸排阻和/或亲和性层析法、2D-PAGE、HPLC、反相层析法、免疫沉淀法)。在一些特定方法中,该RGN多肽为重组产生的且包含纯化标签以帮助其纯化,其包括但不限于:谷胱甘肽-S-转移酶(GST)、几丁质结合蛋白(CBP)、麦芽糖结合蛋白、硫氧还蛋白(TRX)、聚(NANP)、串连亲和纯化(TAP)标签、myc、AcV5、AU1、AU5、E、ECS、E2、FLAG、HA、nus、Softag 1、Softag 3、Strep、SBP、Glu-Glu、HSV、KT3、S、S1、T7、V5、VSV-G、6xHis、10xHis、生物素羧基运载蛋白(BCCP)及钙调蛋白。一般而言,所标记的RGN多肽或RGN核糖核蛋白复合物是使用固定化金属亲和性层析法纯化。应当理解,可单独地或组合地使用本领域中已知的其他类似方法,包含其他形式的层析法或例如免疫沉淀法。

[0144] “分离的”或“纯化的”多肽或其生物活性部分基本上或实质上不含在其天然存在的环境中发现的、正常情况下与该多肽相伴或交互作用的组成。因此,分离的或纯化的多肽当通过重组技术被产生时基本上不含其他细胞物质或培养基,或当被化学合成时基本上不含化学前体或其他化学物质。基本上不含细胞物质的蛋白质包括具有少于约30%、20%、10%、5%、或1%(以干重计)的污染蛋白质的蛋白质制剂。当重组产生本发明的蛋白质或其生物活性部分时,培养基最佳地呈现少于约30%、20%、10%、5%、或1%(以干重计)的化学前体或所关注的非蛋白质化学物质。

[0145] 本文所提供的用于结合和/或切割所关注的目标序列的特定方法涉及使用体外组装的RGN核糖核蛋白复合物。RGN核糖核蛋白复合物的体外组装可使用本领域中已知的方法,其中允许该RGN多肽与该指导RNA结合的条件下使RGN多肽与指导RNA接触。如本文中所使用的,“接触(contact、contacting、contacted)”是指在适合进行期望反应的条件下,将期望反应的组成放在一起。该RGN多肽可自生物样本、细胞裂解物或培养基中纯化,通过体外转换产生,或被化学合成。该指导RNA可自生物样本、细胞裂解物或培养基中纯化,在体外被翻译,或被化学合成。可使该RGN多肽及指导RNA在溶液(例如,缓冲盐溶液)中产生接触以允许该RGN核糖核蛋白复合物的体外组装。

[0146] VIII. 结合、切割或修饰目标序列的方法

[0147] 本公开内容提供用于结合、切割和/或修饰所关注的目标核苷酸序列的方法。该方法包括向该目标序列或包含该目标序列的细胞、细胞器或胚胎,递送包含至少一个指导RNA或编码该至少一个指导RNA的多核苷酸及至少一个RGN多肽或编码该至少一个RGN多肽的多核苷酸的系统。在这些实施方案中的一些实施方案中,RGN包含SEQ ID NO:1到109的氨基酸序列中任一个或其活性变体或片段。在各种实施方案中,指导RNA包含CRISPR重复序列,该CRISPR重复序列包含SEQ ID NO:110至119、139、141、143、146和201至309的核苷酸序列中任一个或其活性变体或片段。在特定实施方案中,指导RNA包含tracrRNA,该tracrRNA包含SEQ ID NO:120至128、140、142、145、147和148的核苷酸序列中任一个或其活性变体或片段。该系统的指导RNA可为单指导RNA或双指导RNA。该系统的RGN可为无核酸酶活性的RGN,具有切口酶活性,也可融合多肽。在一些实施方案中,融合多肽包含碱基编辑多肽,例如,胞苷脱氨酶或腺苷脱氨酶。在其他实施方案中,RGN融合蛋白包含逆转录酶。在其他实施方

案中,RGN融合蛋白包含多肽,该多肽募集功能性核酸修复复合物的成员,诸如,核苷酸切除修复(NER)或转录偶联-核苷酸切除修复(TC-NER)途径的成员(Wei等人,2015,PNAS USA 112(27):E3495-504;Troelstra等人,1992,Cell 71:939-953;Marnef等人,2017,J Mol Biol 429(9):1277-1288),如2020年1月27日提交申请的第62/966,203号美国临时专利申请中所描述的,且其全部内容通过引用并入本文。在一些实施方案中,RGN融合蛋白包括CSB(van den Boom等人,2004,J Cell Biol 166(1):27-36;van Gool等人,1997,EMBO J 16(19):5955-65;其实例如SEQ ID NO:138所示),CSB为该TC-NER(核苷酸切除修复)途径的成员且在其他成员的募集中产生功用。在进一步实施方案中,RGN融合蛋白包含CSB的活性域,诸如,包含SEQ ID NO:138的氨基酸残基356-394的CSB的酸性域(Teng等人,2018,Nat Commun 9(1):4115)。

[0148] 在特定实施方案中,RGN和/或指导RNA对于该RGN和/或指导RNA(或编码该RGN及指导RNA中的至少一个的多核苷酸)被引入到的细胞、细胞器或胚胎为异源的。

[0149] 在一些实施方案中,该方法还需要递送至少一个RGN辅助蛋白或编码该至少一个RGN辅助蛋白的多核苷酸,以便该RGN结合至和/或切割目标多核苷酸。在这些实施方案中的一些实施方案中,方法还需要递送如SEQ ID NO:178-192所示的至少一个RGN辅助蛋白或其活性变体或片段或编码该至少一个RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG06369(SEQ ID NO:11)或其变体或片段的特定实施方案中,该方法还包含递送如SEQ ID NO:178-181所示的至少一个RGN辅助蛋白或其活性变体或片段或编码该至少一个RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG03847(SEQ ID NO:12)或其变体或片段的这些实施方案中的一些实施方案中,该方法还包含递送如SEQ ID NO:315-317所示的至少一个RGN辅助蛋白或其活性变体或片段或编码该至少一个RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG05625(SEQ ID NO:13)或其变体或片段的某些实施方案中,该方法还包含递送如SEQ ID NO:185-187所示的至少一个RGN辅助蛋白或其活性变体或片段或编码该至少一个RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG03524(SEQ ID NO:16)或其变体或片段的一些实施方案中,该方法还包含递送如SEQ ID NO:188-190所示的至少一个RGN辅助蛋白或其活性变体或片段或编码该至少一个RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG03759(SEQ ID NO:14)或其变体或片段的特定实施方案中,该方法还包含递送如SEQ ID NO:191所示的RGN辅助蛋白或其活性变体或片段或编码该RGN辅助蛋白或其活性变体或片段的多核苷酸。在其中该RGN为APG05123(SEQ ID NO:15)或其变体或片段的某些实施方案中,该方法还包含递送如SEQ ID NO:192所示的RGN辅助蛋白或其活性变体或片段或编码该RGN辅助蛋白或其活性变体或片段的多核苷酸。

[0150] 在其中该方法包括递送编码指导RNA和/或RGN多肽的多核苷酸的那些实施方案中,细胞或胚胎可之后在该指导RNA和/或RGN多肽被表达的条件下培养。在各种实施方案中,该方法包括使目标序列与RGN核糖核蛋白复合物接触。该RGN核糖核蛋白复合物可包含无核酸酶活性或具有切口酶活性的RGN。在一些实施方案中,核糖核蛋白复合物的RGN为包含碱基编辑多肽的融合多肽。在某些实施方案中,该方法包括将RGN核糖核蛋白复合物引入包含目标序列的细胞、细胞器或胚胎内。该RGN核糖核蛋白复合物可为已自生物样本被纯化、重组产生并随后纯化、或如本文所描述的体外组装的复合物。在其中与该目标序列或细

胞、细胞器或胚胎接触的RGN核糖核蛋白复合物已经在体外被组装的那些实施方案中,该方法还可包含该复合物在与该目标序列、细胞、细胞器或胚胎接触之前的体外组装。

[0151] 可使用本领域中已知的、包括但不限于电穿孔的任何方法将纯化的或体外组装的RGN核糖核蛋白复合物引入细胞、细胞器或胚胎内。作为另一种选择,可使用本领域中已知的任何方法(例如,电穿孔)将编码或包含该指导RNA的RGN多肽和/或多核苷酸引入细胞、细胞器或胚胎内。

[0152] 在递送至该目标序列或包含该目标序列的细胞、细胞器或胚胎,或与该目标序列或包含该目标序列的细胞、细胞器或胚胎接触后,该指导RNA将该RGN导向至以序列特异性方式与该目标序列结合。在其中该RGN具有核酸酶活性的那些实施方案中,RGN多肽在结合时切割所关注的目标序列。该目标序列随后可通过诸如非同源末端连接或用所提供的供体多核苷酸的同源导向修复的内源修复机制而被修饰。

[0153] 测量RGN多肽与目标序列的结合的方法在本领域中是已知的,且包括:染色质免疫沉淀测定、凝胶迁移率变动测定、DNA下拉测定、报导子测定(reporter assay)、微孔板捕获及检测测定。同样,测量目标序列的切割或修饰的方法在本领域中是已知的,且包括体外或体内切割测定,其中在为了促进降解产物的检测而将恰当标签(例如,放射性同位素、荧光物质)附接至该目标序列的情况下或在未为了促进降解产物的检测而将恰当标签(例如,放射性同位素、荧光物质)附接至该目标序列的情况下,使用PCR、测序或凝胶电泳来确认切割。作为另一种选择,可使用切口触发的指数扩增反应(NTEXPAR)测定(参见,例如,Zhang等人(2016) Chem. Sci. 7:4951-4957)。可使用Surveyor测定来评估体内切割(Guschin等人(2010) Methods Mol Biol 649:247-256)。

[0154] 在一些实施方案中,方法涉及使用与一个以上的指导RNA复合的单一类型RGN。该一个以上的指导RNA可靶向单基因的不同区域,也可靶向多个基因。

[0155] 在其中未提供供体多核苷酸的那些实施方案中,通过RGN多肽引入的双链断裂可通过非同源末端连接(NHEJ)修复过程修复。由于NHEJ的易错性质,该双链断裂的修复可导致对该目标序列的修饰。如本文所使用的,关于核酸分子的“修饰”是指该核酸分子的核苷酸序列的变化,其可为一种或多种核苷酸的缺失、插入或置换或它们的组合。该目标序列的修饰可导致改变的蛋白质产物的表达或编码序列的灭活。

[0156] 在其中存在供体多核苷酸的那些实施方案中,供体多核苷酸中的供体序列可在所引入的双链断裂的修复历程期间被整合至该目标核苷酸序列内或与该目标核苷酸序列交换,导致外源供体序列的引入。因此,供体多核苷酸包含期望被引入所关注的目标序列内的供体序列。在一些实施方案中,供体序列改变该原始目标核苷酸序列,使得该新整合的供体序列将不被该RGN识别及切割。该供体序列的整合可通过在该供体多核苷酸内包括与该目标核苷酸序列侧翼的序列具有基本序列同一性、本文中被称作“同源臂”的侧翼序列来增强,以允许同源导向的修复过程。在一些实施方案中,同源臂具有至少50个碱基对、100个碱基对及多达2000个碱基对或更多个碱基对的长度,且与其在该目标核苷酸序列内的对应序列具有至少90%、至少95%或更高序列同源性。

[0157] 在其中该RGN多肽引入双链交错断裂的那些实施方案中,供体多核苷酸可包含侧翼为相容突出部的供体序列,以允许在该双链断裂的修复期间通过非同源修复过程将供体序列直接连接至包含突出部的经切割的目标核苷酸序列。

[0158] 在其中该方法涉及使用为切口酶(即,仅能够切割双链多核苷酸中的单链)的RGN的那些实施方案中,该方法可包含引入靶向相同的或重叠的目标序列且切割该多核苷酸的不同链的两个RGN切口酶。例如,可将仅切割双链多核苷酸的正(+)链的RGN切口酶与仅切割双链多核苷酸的负(-)链的第二RGN切口酶一起引入。

[0159] 在各种实施方案中,提供一种用于结合目标核苷酸序列且检测该目标序列的方法,其中该方法包括将至少一个指导RNA或编码该至少一个指导RNA的多核苷酸及至少一个RGN多肽或编码该至少一个RGN多肽的多核苷酸引入细胞、细胞器或胚胎中;表达该指导RNA和/或RGN多肽(如果编码序列被引入),其中该RGN多肽为无核酸酶活性的RGN且还包含可检测标记,并且该方法还包含检测该可检测标记。该可检测标记可与该RGN融合为融合蛋白(例如,荧光蛋白),也可与为该RGN多肽缀合或被并入该RGN多肽内、可通过视觉或通过其他手段检测的小分子。

[0160] 本文还提供用于在目标序列的调节下调控所关注的目标序列或基因的表达的方法。这些方法包括:将至少一个指导RNA或编码该至少一个指导RNA的多核苷酸及至少一个RGN多肽或编码该至少一个RGN多肽的多核苷酸引入细胞、细胞器或胚胎内;表达该指导RNA和/或RGN多肽(如果编码序列被引入),其中该RGN多肽为无核酸酶活性的RGN。在这些实施方案中的一些实施方案中,无核酸酶活性的RGN为包含如本文所描述的表达调控域(即,表观遗传修饰域、转录活化域或转录阻遏域)的融合蛋白。

[0161] 本公开内容还提供用于结合和/或修饰所关注的目标核苷酸序列的方法。这些方法包括递送系统至该目标序列或包含该目标序列的细胞、细胞器或胚胎,该系统包含至少一个指导RNA或编码该至少一个指导RNA的多核苷酸及至少一个融合多肽,该至少一个融合多肽包含本发明的RGN及碱基编辑多肽(例如,胞苷脱氨酶或腺苷脱氨酶)或编码该融合多肽的多核苷酸。

[0162] 本领域技术人员将理解,本发明公开的任一方法可用于靶向单个目标序列或多个目标序列。因此,这些方法包括与多个不同的指导RNA组合地使用单RGN多肽,其可靶向单个基因和/或多个基因内的多个不同的序列。本文还涵盖其中与多个不同的RGN多肽组合地引入多个不同的指导RNA的方法。这些指导RNA及指导RNA/RGN多肽系统可靶向单个基因和/或多个基因内的多个不同的序列。

[0163] 在一个方面,本发明提供含有上述方法及组合物中所公开的任何一种或多种元件的试剂盒。在一些实施方案中,该试剂盒包含载体系统及使用该试剂盒的说明。在一些实施方案中,载体系统包含:(a)第一调节元件,该第一调节元件与编码该crRNA序列的DNA系列及用于在指导序列的上游插入经编码的crRNA序列的一种或多种插入位点可操作地连接,其中当表达时,该指导序列在真核细胞中将CRISPR复合物导向至与目标序列序列特异性结合,其中该CRISPR复合物包含与(a)指导RNA多核苷酸复合的CRISPR酶;和/或(b)第二调节元件,该第二调节元件与酶编码序列可操作地连接,该酶编码序列编码包含核定位序列的所述CRISPR酶。

[0164] 在一些实施方案中,试剂盒包含与用于插入载体的指导序列对应的一种或多种寡核苷酸,以便可操作地连接该指导序列与调节元件。在一些实施方案中,试剂盒包含同源重组模板多核苷酸。在一个方面,本发明提供用于使用CRISPR系统的一种或多种元件的方法。本发明的CRISPR复合物提供用于修饰目标多核苷酸的有效手段。本发明的CRISPR复合物具

有种类繁多的功用,包括在多种细胞类型中修饰(例如,缺失、插入、易位、灭活、活化、碱基编辑)目标多核苷酸。就其本身而言,本发明的CRISPR复合物在例如基因治疗、药物筛选、疾病诊断以及预后中具有广泛的应用。示例性的CRISPR复合物包含与指导序列复合的CRISPR酶,该指导序列与该目标多核苷酸内的目标序列杂交。

[0165] IX. 目标多核苷酸

[0166] 在一个方面,本发明提供修饰真核细胞中的目标多核苷酸的方法,其可为在体内、离体或在体外。在一些实施方案中,该方法包括:自人或非人动物或植物(包括微藻类)取样细胞或细胞群体;和修饰该细胞或这些细胞。培养可以在任何阶段离体发生。甚至可以将该细胞或这些细胞重新引入该非人动物或植物(包含微藻类)中。

[0167] 使用自然变异性,植物育种者将关于诸如产量、质量、均匀性、耐寒性以及抗害虫性的可期望质量的大多数有用基因组合起来。这些可期望质量还包括生长、日长度偏好、温度要求、花卉或生殖发育的起始日期、脂肪酸含量、抗虫性、抗病性、线虫抗性、真菌抗性、除草剂抗性、对各种环境因素(包括干旱、热、湿、冷、风及包括高盐度的不利土壤条件)的耐受性。这些有用基因的来源包括天然或外来品种、原生种(heirloom variety)、野生植物近源种和例如以诱变剂处理植物材料的诱导突变。使用本发明,向植物育种者提供诱导突变的新工具。因此,本领域技术人员可以分析基因组以获得有用基因的来源,且在具有所期望特征或性状的品种中采用本发明以诱导有用基因的增加,同时比先前的诱变剂更精确,且因此加速并改良植物育种计划。

[0168] RGN系统的目标多核苷酸可为对真核细胞是内源或外源的任何多核苷酸。例如,该目标多核苷酸可为驻留于真核细胞的细胞核中的多核苷酸。该目标多核苷酸可为编码基因产物(例如,蛋白质)的序列或非编码序列(例如,调节多核苷酸或无用DNA)。在一些实施方案中,该目标序列与PAM(前间隔序列邻近基序)相关联;即,该CRISPR复合物识别的短序列。该PAM的精确序列和长度要求随所使用的CRISPR酶而不同(并且在一些实施方案中,RGN不需要PAM系列),但该PAM通常情况下是与该前间隔序列(即,该目标序列)相邻的2-5个碱基对序列。

[0169] CRISPR复合物的目标多核苷酸可包括很多疾病关联基因及多核苷酸以及信号传导生化途径相关基因及多核苷酸。目标多核苷酸的实例包括与信号传导生化途径相关联的序列,例如,信号传导生化途径相关基因或多核苷酸。目标多核苷酸的实例包括疾病关联基因或多核苷酸。“疾病关联”基因或多核苷酸是指:与非疾病对照的组织或细胞相较下,在来源于感染疾病的组织的细胞中以异常水平或以异常形式产出转录或翻译产物的任何基因或多核苷酸。其可为一种变得以异常高水平表达的基因;其可为一种变得以异常低水平表达的基因,其中改变的表达与疾病的发生和/或进展相关。疾病关联基因也指具有突变或基因变异的基因,该具有突变或基因变异的基因为直接造成疾病病因(例如,因果突变),或为与造成疾病病因(例如,因果突变)的基因呈连锁不平衡(linkage disequilibrium)。该转录或翻译产物可为已知的或未知的,且还可处于正常或异常水平。疾病关联基因及多核苷酸的实例可自在万维网上找到的约翰·霍普金斯大学(马里兰州巴尔的摩)麦库席克-内森斯遗传医学研究所(McKusick-Nathans Institute of Genetic Medicine)和美国国家医学图书馆(马里兰州贝塞斯达)(National Library of Medicine(Bethesda,Md.))的国家生物技术信息中心(National Center for Biotechnology Information)找到。

[0170] 虽然CRISPR系统在对于其相对容易靶向所关注的基因组序列方面格外有用,但仍存在该RGN做什么可解决因果突变的问题。一种措施为在RGN(优选地,该RGN的无活性或切口酶变体)与碱基编辑酶或碱基编辑酶(诸如,胞苷脱氨酶或腺苷脱氨酶基编辑器)的活性域之间产生融合蛋白(第9,840,699号美国专利,该美国专利通过引用并入本文)。在一些实施方案中,这些方法包括使DNA分子与:(a)包含本发明的RGN及诸如脱氨酶的碱基编辑多肽的融合蛋白;和(b)将(a)的融合蛋白靶向至该DNA链的目标核苷酸序列的gRNA接触;其中该DNA分子以有效量并且在适于核碱基脱氨化的条件下与该融合蛋白及gRNA接触。在一些实施方案中,该目标DNA序列包含与疾病或病症关联的序列,并且其中该核碱基的脱氨化导致与疾病或病症无关的序列。在一些实施方案中,该目标DNA序列驻留于农作物的等位基因中,其中所关注的性状的特定等位基因导致具有较低农艺价值的植物。该核碱基的脱氨化导致改良性状并增加植物的农艺价值的等位基因。

[0171] 在一些实施方案中,该DNA序列包含与疾病或病症关联的T→C或A→G点突变,并且其中突变体C或G碱基的脱氨化导致与疾病或病症无关的序列。在一些实施方案中,该脱氨化校正与该疾病或病症关联的序列中的点突变。

[0172] 在一些实施方案中,与该疾病或病症关联的序列编码蛋白质,且其中该脱氨化将终止密码子引入与该疾病或病症关联的序列中,导致编码蛋白质的截短。在一些实施方案中,该接触在易患有、患有或诊断患有该疾病或病症的个体体内进行。在一些实施方案中,疾病或病症为与该基因组中的点突变或单碱基突变关联的疾病。在一些实施方案中,该疾病为基因疾病、癌症、代谢疾病或溶酶体贮积病。

[0173] X. 药物组合物及治疗方法

[0174] 提供药物组合物,该药物组合物包含:本发明公开的RGN多肽及其变体或片段以及编码该RGN多肽及其变体或片段的多核苷酸、本发明公开的gRNA或编码该gRNA的多核苷酸、本发明公开的系统、或包含该RGN多肽或RGN编码多核苷酸、gRNA或gRNA编码多核苷酸、或该RGN系统中任一个的细胞及药学上可接受的载剂。

[0175] 药物组合物为被运用于防止、降低程度、治愈或治疗目标病况或疾病的组合物,该组合物包含活性成分(即,RGN多肽、RGN编码多核苷酸、gRNA、gRNA编码多核苷酸、RGN系统、或包含这些中任一个的细胞)及药学上可接受的载剂。

[0176] 如本文中使用的,“药学上可接受的载剂”是指不对生物体引起显著刺激且不消除该活性成分(即,RGN多肽、RGN编码多核苷酸、gRNA、gRNA编码多核苷酸、RGN系统、或包含这些中任一个的细胞)的活性及特性的材料。载剂必须具有足够高的纯度及足够低的毒性,以使它们合适施用于正被治疗的个体。该载剂可为惰性的,其也可具有医药效益。在一些实施方案中,药学上可接受的载剂包含合适对人或其他脊椎动物施用的一种或多种兼容固体或液体填充剂、稀释剂或封装物质。在一些实施方案中,药学上可接受的载剂不是天然存在的。在一些实施方案中,未发现该药学上可接受的载剂在自然界中与该活性成分在一起。

[0177] 本发明公开的方法中使用的药学组合物可用提供合适转移、递送、耐受性等等的合适载剂、赋形剂和其他药剂配制。众多恰当制剂为本领域技术人员所知。参见,例如,Remington, The Science and Practice of Pharmacy (21st ed. 2005)。合适的制剂例如包括:粉剂、糊剂、膏剂、胶冻剂、蜡、油类、脂质、含脂质(阳离子或阴离子)的囊泡(诸如, LIPOFECTIN囊泡)、脂质纳米颗粒、DNA缀合物、无水吸收糊剂、水包油及油包水乳液、乳液卡

波蜡(emulsions carbowax)(各种分子量的聚乙二醇)、半固体凝胶和含有卡波蜡的半固体混合物。经口或肠胃外使用的药学组合物可被制备为适合供给某个剂量的活性成分的单位剂量的剂型。这些单位剂量的剂型例如包括片剂、丸剂、胶囊、注射剂(安瓿)、栓剂等。

[0178] 在其中包括或以本发明公开的RGN、gRNA、RGN系统或编码该RGN、gRNA、RGN系统的多核苷酸修饰的细胞被施用个体的一些实施方案中,细胞与药学上可接受的载剂一起作为悬浮剂被施用。本领域技术人员将认识到将被使用于细胞组合物中的药学上可接受的载剂将不包括实质上干扰将被递送至该个体的细胞的存活率的量的缓冲液、化合物、冷冻保存药剂、保存剂、或其他药剂。包括细胞的制剂可包含例如允许细胞膜保持完整性的渗透压缓冲液,且视情况地,在施用时保持细胞存活率或增强植入的营养剂。此类制剂及悬浮剂为本领域技术人员所知,和/或使用常规实验可适用于与本文公开的细胞一起使用。

[0179] 细胞组合物也可被乳化为或呈现为脂质体组合物,前提条件为该乳化程序不会不利地影响细胞存活率。该细胞及其他活性成分可以与药学上可接受且与该活性成分兼容的赋形剂和以适合在本文公开的治疗方法中使用的量混合。

[0180] 包含在细胞组合物中的额外药剂可包含其内的组成的药学上可接受的盐。药学上可接受的盐包含与诸如例如盐酸或磷酸的无机酸、或与诸如乙酸、酒石酸、苦杏仁酸等等的有机酸形成的酸加成盐(与多肽的自由氨基基团形成)。与自由羧基基团形成的盐也可衍生自诸如例如钠、钾、铵、钙或铁的氢氧化物的无机碱和诸如异丙胺、三甲胺、2-乙胺乙醇、组氨酸、普鲁卡因等等的有机碱。

[0181] 生理学上可耐受且药学上可接受的载剂为本领域中所知。示例性液体载剂为无菌水溶液,其除了活性成分及水外不含有其他材料,或其含有诸如在生理pH值、生理盐水或二者的磷酸钠的缓冲液(诸如,磷酸盐缓冲盐水)。更进一步地,水性载剂可含有一种以上的缓冲盐以及诸如氯化钠及氯化钾、右旋糖、聚乙二醇及其他溶质的盐。液体组合物也可含有除了水及排除水的液相。这种额外液相的实例为甘油、诸如棉花籽油的植物油和水油乳液。在特定病症或病况的治疗中有效的细胞组合物中使用的活性化合物的量可依赖于该病症或病况的本质,且可通过标准临床技术决定。

[0182] 本发明公开的RGN多肽、指导RNA、RGN系统、或编码该RGN多肽、指导RNA、RGN系统的多核苷酸可依赖于施用的特定方式及剂型以诸如载剂、溶剂、稳定剂、佐剂、稀释剂等等的药学上可接受的赋形剂配制。在一些实施方案中,这些药学组合物被配制以达成生理学上相容的pH,且依赖于制剂及施用途径,其pH范围为约3的pH至约11的pH、约pH3至约pH7。在一些实施方案中,可将该pH调整至约pH5.0至约pH8的范围。在一些实施方案中,组合物可包括本文描述的治疗有效量的至少一个化合物,连同一种或多种药学上可接受的赋形剂。在一些实施方案中,组合物包括本文描述的化合物的组合,或包含在治疗或防止细菌生长中有用的第二活性成分(例如而不限于,抗菌或抗微生物药剂),或包含本公开内容的试剂的组合。

[0183] 例如,适合的赋形剂包含载剂分子,该载剂分子包含大的缓慢代谢的大分子,诸如,蛋白质、多糖、聚乳酸、聚乙醇酸、聚合的氨基酸、氨基酸共聚物和无活性的病毒颗粒。其他示例性赋形剂可包含:抗氧化剂(例如而不限于,抗坏血酸)、螯合剂(例如而不限于,EDTA)、碳水化合物(例如而不限于,糊精、羟烷基纤维素和羟烷基甲基纤维素)、硬脂酸、液体(例如而不限于,油、水、盐水、甘油及乙醇)、润湿或乳化剂、pH缓冲物质等等。

[0184] 在一些实施方案中,在单位剂量或多剂量容器(例如,密封安瓿及小瓶)中提供制

剂,且可被储存于冻干(冷冻干燥)条件下,需要在立即使用之前,添加无菌液体载剂(例如,盐水、注射用水、半液体泡沫或凝胶)。实时注射溶液及悬浮剂可自前面描述种类的无菌粉剂、颗粒及片剂制备。在一些实施方案中,活性成分溶解于缓冲液体溶液中,其在单位剂量或多剂量容器中被冷冻,且之后被解冻用于注射或被保持/稳定在冷冻下直到使用。

[0185] 治疗剂可包含在控释系统中。为延长药物的作用,常常期望自皮下、鞘内或肌肉内注射减缓药物的吸收。这可通过使用具水难溶性的结晶或非结晶材料的液体悬浮剂来完成。然后,药物的吸收速率依赖于其溶解速率,而其溶解速率又可依赖于结晶大小及结晶的方式。作为另一种选择,通过该药物溶解或悬浮于油运载体中完成肠胃外施用药物的延迟吸收。在一些实施方案中,长期缓释植入剂的使用特别适合用于慢性病况的治疗。长期缓释植入剂是本领域技术人员众所周知的。

[0186] 本文提供治疗有需要的个体的疾病的方法。该方法包括以有效量对有需要的个体施用本发明公开的RGN多肽或其活性变体或片段或编码该RGN多肽或其活性变体或片段的多核苷酸、本发明公开的gRNA或编码该gRNA的多核苷酸、本发明公开的RGN系统、或通过这些组合物中任一个修饰的或包括这些组合物中任一个的细胞。

[0187] 在一些实施方案中,治疗包括通过施用本发明公开的RGN多肽、gRNA、或RGN系统、或编码该RGN多肽、gRNA、或RGN系统的多核苷酸的体内基因编辑。在一些实施方案中,治疗包括离体基因编辑,其细胞是以本发明公开的RGN多肽、gRNA、或RGN系统、或编码该RGN多肽、gRNA、或RGN系统的多核苷酸离体基因修饰,且然后将经修饰细胞施用至个体。在一些实施方案中,经基因修饰细胞源于之后被施用该修饰细胞的该个体,且所移植细胞在本文中被称为自体的。在一些实施方案中,经基因修饰细胞源自与被施用该经修饰细胞的个体(即,受者)属相同物种的不同个体(即,供体),且该所移植细胞在本文中被称作异体的。在本文描述的一些实例中,该细胞在对有需要的个体施用之前可在培养中被扩增。

[0188] 在一些实施方案中,将要用本发明公开的组合物治疗的疾病为可用免疫疗法(诸如,用嵌合抗原受体(CAR) T细胞)治疗的疾病。此类疾病包括但不限于癌症。

[0189] 在一些实施方案中,将要用本发明公开的组合物治疗的疾病与被突变的序列(即,该序列为对于该疾病或病症是因果的或为对于与该疾病或病症关联的症状是因果的)关联,以便治疗该疾病或病症或与该疾病或病症关联的症状的降低。在一些实施方案中,将要用本发明公开的组合物治疗的疾病与因果突变关联。如本文中使用的,“因果突变”是指对个体中疾病或病症的严重程度或出现有贡献的基因组中的特殊核苷酸、多个核苷酸或核苷酸序列。因果突变的校正引起由疾病或病症导致的至少一个症状改良。在一些实施方案中,因果突变与本文公开的RGN所识别的PAM位点相邻。该因果突变可用本发明公开的RGN或包括本发明公开的RGN及碱基编辑的多肽(即,碱基编辑器)的融合多肽来校正。与因果突变关联的疾病的非限制性实例包含:囊肿纤维化、赫尔勒综合征、弗里德赖希共济失调、亨廷顿病和镰形细胞疾病。疾病关联基因及突变的额外非限制性实例可自在万维网上找到的约翰·霍普金斯大学(马里兰州巴尔的摩)麦库席克-内森斯遗传医学研究所(McKusick-Nathans Institute of Genetic Medicine)和美国国家医学图书馆(马里兰州贝塞斯达)(National Library of Medicine(Bethesda,Md.))的国家生物技术信息中心(National Center for Biotechnology Information)找到。

[0190] 在一些实施方案中,本文提供的方法使用于将去活化点突变引入将与疾病或病症

关联的基因产物编码的基因或等位基因内。例如,在一些实施方案中,本文提供运用本发明公开的组合物将去活化点突变引入致癌基因(例如在增生性疾病的治疗中)内的方法。在一些实施方案中,去活化突变可在编码序列中产生提前终止密码子,这导致经截短基因产物(例如缺少全长蛋白的功能的经截短蛋白)的表达。在一些实施方案中,本文提供的方法的目的为通过基因组编辑来恢复功能不良性基因的功能。本发明公开的RGN多肽及包含该RGN多肽的系统可对例如通过在人细胞培养中校正疾病关联突变的体外基于基因编辑的人疗法有效。

[0191] 如本文中使用的,“治疗(treatment)”或“治疗(treating)”、“缓和”或“改善”可被互换地使用。这些术语是指用于获得有益或期望结果的措施,该有益结果或期望结果包括但不限于治疗性益处和/或预防性益处。通过治疗性益处意味着对治疗中的一种或多种疾病、病况或症状之中的任何治疗相关改良或对治疗中的一种或多种疾病、病况或症状的作用。对于预防性益处,该组合物可被施用处于特殊疾病、病况或症状的发展风险中的个体,或被施用报告疾病的一种或多种生理症状的个体,即使该疾病、病症或症状可能尚未呈现征兆。

[0192] 术语“有效量”或“治疗有效量”是指足以实现有益或期望结果的药剂量。治疗有效量可取决于如下的一种或多种而变化:被治疗的个体及疾病病况、该个体的体重及年龄、疾病病况的严重性、施用方式等等,本领域技术人员可容易地对此做出决定。特定剂量可取决于如下一种或多种而变化:所选的特殊药剂、将要遵循的给药方案、是否与其他化合物组合地施用、施用时机和输送其的递送系统。

[0193] 术语“施用”是指通过导致所引入的活性成分至少部分地定位于期望位点(诸如,受伤或修复的位点)处的方法或途径,将活性成分置于个体内,使得产生期望效应。在其中细胞被施用的那些实施方案中,可通过导致递送至个体中的期望位置的任何恰当途径施用该细胞,其中至少一部分所移植细胞或该细胞的组成维持存活。施用个体后的细胞的存活期可短至几小时(例如,二十四小时)、至几天、至长至若干年、甚或患者的一生,即,长期植入。例如,在本文描述的一些方面中,光受体细胞或视网膜前体细胞的有效量是通过身体的施用途径(诸如腹膜内或静脉内途径)施用的。

[0194] 在一些实施方案中,施用包括通过病毒递送的施用。在一些实施方案中,施用包括通过电穿孔的施用。在一些实施方案中,施用包括通过纳米颗粒递送的施用。在一些实施方案中,施用包括通过脂质体递送的施用。施用的任何有效途径均可用于施用本文描述的药学组合物的有效量。在一些实施方案中,施用包括通过选自以下组成的组的方法的施用:静脉内地、皮下地、肌肉内地、口服地、经直肠地、通过气溶胶、肠胃外地、经眼地、经肺地、经皮地、经阴道地、经耳地、经鼻地和通过局部施用、或其任意组合。在一些实施方案中,对于细胞的递送,使用通过注射或灌注的施用。

[0195] 如本文中使用的,术语“个体(subject)”是指期望对其进行诊断、治疗(treatment)或治疗(therapy)的任何个体(individual)。在一些实施方案中,个体为动物。在一些实施方案中,个体为哺乳动物。在一些实施方案中,个体为人。

[0196] 治疗功效可由熟练的临床医师决定。然而,如果疾病或病症的症候或症状的任何一个或全部是以有益方式被改变(例如,至少减少10%)、或其他临床上接受的疾病的症状或标记被改良或改善,则该治疗被认为“有效治疗”。还可通过个体未发生如通过住院评估

的恶化、或不需要医学介入(例如,疾病的进展被停止或至少减慢)来测量功效。本领域技术人员知晓测量这些指标的方法。治疗包括:(1)抑制疾病,例如,停止或减慢症状的进展;或(2)减缓疾病,例如,引起症状消退;和(3)预防或降低症状发展的可能性。

[0197] A. 使用碱基编辑修饰因果突变

[0198] 在一些实施方案中,本发明的RGN被使用于使用碱基编辑来修饰因果突变。可使用取决于本发明的RGN-碱基编辑器融合蛋白的措施而校正的基因遗传性疾病的实例是赫尔勒综合征。赫尔勒综合征(也称为MPS-1)为 α -L-艾杜糖苷酶(IDUA)缺乏而导致在分子水平上由溶酶体中硫酸皮肤素和硫酸乙酰肝素的累积表征的溶酶体贮积症的结果。此疾病一般是由编码 α -L-艾杜糖苷酶(IDUA)基因中的突变引起的遗传性基因病症。常见的IDUA突变是W402X和Q70X,两者都是无义突变导致翻译的提前终止。通过精确的基因组编辑(PGE)措施很好地解决了这种突变,由于单核苷酸的回复突变(例如,通过碱基编辑措施)将恢复野生型编码序列并导致由基因座的内源性调节机制控制的蛋白质表达。另外,由于已知杂合子是无症状的,所以靶向这些突变中的一个的PGE疗法对于大部分患有此疾病的患者是有用的,因为突变的等位基因中只有一个需被校正(Bunge等人(1994) Hum. Mol. Genet. 3 (6) : 861-866,通过引用并入本文)。

[0199] 目前对赫尔勒综合征的目前治疗包括酶替代疗法及骨髓移植(Vellodi等人,((1997)) Arch. Dis. Child. 76 (2) :92-99;Peters等人,((1998)) Blood 91 (7) :2601-2608,通过引用并入本文)。尽管酶替代疗法对赫尔勒综合征患者的生存和生活质量已具有显著效应,但该措施需要昂贵且耗时的每周灌注。额外措施包含在表达载体上IDUA基因的递送或将该基因插入高度表达的基因座内,诸如,血清白蛋白的基因座内(第9,956,247号美国专利,通过引用并入本文)。然而,这些措施不能将原始IDUA基因座恢复至正确的编码序列。基因组编辑策略可具有很多优点,最显著的是基因表达的调节将受健康个体中存在的天然机制的控制。另外,使用碱基编辑不一定引起双链DNA断裂,这可能由肿瘤阻遏机制的破坏引起大规模染色体重排、细胞死亡或致癌性。一般策略可涉及使用本发明的RGN碱基编辑器融合蛋白来靶向并校正人基因组中的某些疾病引起的突变。应理解,还可寻求类似措施来靶向通过碱基编辑可校正的疾病。还应进一步理解,也可使用本发明的RGN来部署以靶向其他物种(特别是一般家庭宠物或家畜)中疾病引起的突变的类似措施。常见家庭宠物和家畜包含狗、猫、马、猪、牛、羊、鸡、驴、蛇、雪貂和包括鲑鱼的鱼和虾。

[0200] B. 通过靶向缺失修饰因果突变

[0201] 本发明的RGN在因果突变更复杂的人治疗措施中也有用。例如,诸如弗里德赖希共济失调和亨廷顿病的一些疾病是在基因特定区域处的三个核苷酸基序的重复序列中显著增加的结果,这会影响表达蛋白起作用或被表达的能力。弗里德赖希共济失调(FRDA)为一种导致脊髓中神经组织渐进性退化的常染色体隐性遗传病。线粒体中共济蛋白(FXN)蛋白质的降低水平引起细胞水平的氧化损伤及铁缺乏。降低的FXN表达已被联系至体细胞和生殖系FXN基因的内含子1内的GAA三联体扩增。在FRDA患者中,GAA重复序列往往由超过70个、有时甚至超过1000个(最常见的是600-900个)三联体组成,而未受影响的个体具有约40个或更少的重复序列(Pandolfo等人,(2012) Handbook of Clinical Neurology 103:275-294;Campuzano等人,(1996) Science 271:1423-1427;Pandolfo(2002) Adv. Exp. Med. Biol. 516:99-118;全部通过引用并入本文)。

[0202] 引起弗里德赖希共济失调 (FRDA) 的三核苷酸重复序列的扩增发生在 FXN 基因内定义的基因座中, 被称为 FRDA 不稳定区域。RNA 指导核酸酶 (RGN) 可用于切除 FRDA 患者细胞中的不稳定区域。此措施需要: 1) 可被程序化以靶向人基因组中的等位基因的 RGN 及指导 RNA 序列; 以及 2) 对于 RGN 及指导序列的递送措施。特别是当除了功能性表达盒所需的其他基因元件之外, 还要考虑到 SpCas9 基因及指导 RNA 的长度时, 用于基因组编辑的许多核酸酶 (诸如, 常用的来自化脓性链球菌 (*S. pyogenes*) 的 Cas9 核酸酶 (SpCas9)) 太大而无法包装到腺相关病毒 (AAV) 载体内。这使得使用 SpCas9 的措施更加困难。

[0203] 本发明的某些 RNA 指导核酸酶极适合随指导 RNA 一起被包装到 AAV 载体内。包装两个指导 RNA 可能需要第二个载体, 但该措施与诸如 SpCas9 的较大核酸酶所需要的相较仍然是有利的, 这可能需要在两个载体之间割裂蛋白质序列。本发明涵盖使用本发明的 RGN 的策略, 其中移除了基因组不稳定区域。这种策略适用于具有类似基因基础的其他疾病及病症, 诸如亨廷顿病。使用本发明的 RGN 的类似策略也可适用于具有农艺或经济重要性的非人动物 (包含: 狗、猫、马、猪、牛、羊、鸡、驴、蛇、雪貂、以及包含鲑鱼的鱼和虾) 的类似疾病及病症。

[0204] C. 通过靶向诱变修饰因果突变

[0205] 本发明的 RGN 也可引入可导致有益效应的破坏性突变。编码血红蛋白的基因 (特别是 β 球蛋白链 (HBB 基因)) 中的基因缺陷可造成很多称为血红蛋白病变的疾病 (包括镰状细胞贫血和地中海贫血)。

[0206] 在成年人中, 血红蛋白为包括两个 α 样球蛋白链及两个 β 样球蛋白链及 4 个血红素基团的异四聚体。在成人中, $\alpha_2\beta_2$ 四聚体被称为血红蛋白 A (HbA) 或成人血红蛋白。通常情况下, α 及 β 球蛋白链以大约 1:1 的比例合成, 并且此比例就血红蛋白和红细胞 (RBC) 稳定而言似乎是关键的。在发育中的胎儿中, 产生了不同形式的血红蛋白 (胎儿血红蛋白 (HbF)), 其对氧的结合亲和力高于血红蛋白 A, 使得氧可通过母亲的血流递送到婴儿的系统。胎儿血红蛋白也含有两个 α 球蛋白链, 但是代替成人 β 球蛋白链, 其具有两个胎儿 γ 球蛋白链 (即, 胎儿血红蛋白为 $\alpha_2\gamma_2$)。自 γ 球蛋白的产生转换至 β 球蛋白的产生的调节相当复杂, 且主要涉及 γ 球蛋白转录的下调与 β 球蛋白转录的同时上调。在妊娠约 30 周时, 胎儿中 γ 球蛋白的合成开始下降, 而 β 球蛋白的产生增加。到约 10 个月龄时, 新生儿的血红蛋白几乎都是 $\alpha_2\beta_2$, 但是一些 HbF 持续到成年期 (约为总血红蛋白的 1-3%)。在大多数具有血红蛋白病变的患者中, 编码 γ 球蛋白的基因仍然存在, 但由于如上所述在临近分娩时发生正常的基因阻遏, 表达相对较低。

[0207] 镰状细胞疾病是由 β 球蛋白基因 (HBB) 中的 V6E 突变 (DNA 水平的 GAG 至 GTG) 引起的, 其中所得的血红蛋白被称为“血红蛋白 S”或“HbS”。在低氧条件下, HbS 分子聚集并形成纤维状沉淀物。这些聚集物引起 RBC 异常或“变成镰状细胞”, 导致细胞的灵活性丧失。变成镰状细胞的 RBC 不再能够挤入毛细血管床, 且可能导致镰状细胞患者的血管阻塞性危机。除此之外, 变成镰状细胞的 RBC 比正常 RBC 更脆弱, 且倾向于溶血, 最终引起患者贫血。

[0208] 镰状细胞患者的治疗及管理是终生的课题, 其涉及抗生素治疗、疼痛管理及急性发作期间的输液。一种措施为使用羟基脲, 其通过增加 γ 球蛋白的产生来部分地发挥其效应。然而, 长期羟基脲疗法的长期副作用仍然是未知的, 而且治疗产生不想要的副作用且可能在患者之间具有不同的功效。尽管镰状细胞治疗的功效有所增加, 但患者的预期寿命仍

然仅在50岁至60岁的中期至晚期,并且关联的疾病发病对患者的生活质量具有深远的影响。

[0209] 地中海贫血(α 地中海贫血及 β 地中海贫血)也是与血红蛋白有关的疾病,并且通常情况下涉及减少的球蛋白链表达。这可以通过基因调节区域中的突变或从导致减少的表达或减少的水平或功能性球蛋白蛋白质的球蛋白编码序列中的突变发生。地中海贫血的治疗通常涉及血液输注及祛铁疗法。如果适当的供体可被鉴定,则骨髓移植也被使用于治疗有严重地中海贫血的人,但这种手术可能具有重大风险。

[0210] 已经提出用于治疗镰状细胞病(SCD)及 β 地中海贫血的一种措施为增加 γ 球蛋白的表达,使得HbF在功能上取代异常的成人血红蛋白。如上所提及,用羟基脲治疗SCD患者由于其在增加 γ 球蛋白表达上的效应而被认为是部分成功的(DeSimone(1982)Proc Nat'l Acad Sci USA 79(14):4428-31;Ley等人(1982)N.Engl.J.Medicine,307:1469-1475;Ley等人(1983)Blood 62:370-380;Cons tantoulakis等人(1988)Blood 72(6):1961-1967,均通过引用并入本文)。增加HbF的表达涉及鉴定其产物在 γ 球蛋白表达的调节中起作用的基因。一种这样的基因为BCL11A。BCL11A编码在成人红细胞前体细胞中表达的锌指蛋白,且其表达的下调引起 γ 球蛋白表达增加(Sankaran等人(2008)Science 322:1839,通过引用并入本文)。已经提出使用靶向至BCL11A基因的抑制性RNA(例如,第2011/0182867号美国专利公开,通过引用并入本文),但此技术具有若干潜在的缺点,包含可能无法实现完全敲减,这种RNA的递送可能是有问题的,且RNA必须连续存在且终生需要多次治疗。

[0211] 本发明的RGN可用于靶向BCL11A增强子区域,以破坏BCL11A的表达,从而增加 γ 球蛋白表达。这种靶向的破坏可通过非同源末端连接(NHEJ)来实现,由此本发明的RGN靶向至BCL11A增强子区域内的特殊序列,使双链断裂,且细胞的机械修复该断裂,通常同时引入有害突变。类似于针对其他疾病目标所描述的,本发明的RGN由于其相对小的尺寸,使得能够将RGN及其指导RNA的表达盒包装至用于体内递送的单一AAV载体中,从而具有优于其他已知RGN的优点。使用本发明的RGN的类似策略还可应用于人和农艺或经济重要性的非人动物中的类似疾病和病症。

[0212] XI. 包括多核苷酸基因修饰的细胞

[0213] 本文提供了包括已使用通过如本文所描述的RGN、crRNA和/或tracrRNA介导的过程修饰的所关注的目标序列的细胞和生物体。在这些实施方案中的一些实施方案中,RGN包括SEQ ID NO:1至109的氨基酸序列中任一个、或其活性变体或片段。在各种实施方案中,指导RNA包括CRISPR重复序列,该CRISPR重复序列包括SEQ ID NO:110至119、139、141、143、146和201至309的核苷酸序列中任一个、或其活性变体或片段。在具体实施方案中,指导RNA包括tracrRNA,该tracrRNA包括SEQ ID NO:120至128、140、142、145、147和148的核苷酸序列中任一个、或其活性变体或片段。该系统的指导RNA可为单指导RNA或双指导RNA。

[0214] 经修饰的细胞可为真核细胞(例如,哺乳动物、植物、昆虫细胞)或原核细胞。还提供包括至少一种核苷酸序列的细胞器和胚胎,该核苷酸序列已通过利用如本文所描述的RGN、crRNA和/或tracrRNA的过程进行了修饰。经基因修饰的细胞、生物体、细胞器和胚胎对于经修饰的核苷酸序列可为杂合的或纯合的。

[0215] 该细胞、生物体、细胞器或胚胎的染色体修饰可导致改变的表达(上调或下调)、灭活、或改变的蛋白质产物或整合的序列的表达。在其中染色体修饰导致基因灭活或非功能

性蛋白质产物表达的那些实施方案中,经基因修饰的细胞、生物体、细胞器或胚胎被称为“敲除”。该敲除表型可为缺失突变(即,至少一个核苷酸的缺失)、插入突变(即,至少一个核苷酸的插入)或无义突变(即,至少一个核苷酸的置换,使得终止密码子被引入)的结果。

[0216] 在一些实施方案中,细胞、生物体、细胞器或胚胎的染色体修饰可产生“敲入”,该“敲入”由编码蛋白质的核苷酸序列的染色体整合导致。在这些实施方案中的一些实施方案中,编码序列被整合至该染色体,使得编码野生型蛋白质的染色体序列被灭活,但表达出外源引入的蛋白质。

[0217] 在一些实施方案中,染色体修饰导致变体蛋白质产物的产生。表达的变体蛋白质产物可具有至少一个氨基酸置换和/或至少一个氨基酸的添加或缺失。当与野生型蛋白质比较时,由改变的染色体序列所编码的变体蛋白质产物可呈现出经修饰的特征或活性,包括但不限于改变的酶活性或基质特异性。

[0218] 在一些实施方案中,染色体修饰可导致改变的蛋白质表达模式。作为非限制性实例,控制蛋白质产物表达的调节区域中的染色体改变可导致蛋白质产物过度表达或下调或改变的组织或时间表达模式。

[0219] 已经被修饰的细胞可按照传统方式生长成生物体,诸如,植物。参见,例如,McCormick等人(1986)Plant Cell Reports 5:81-84。然后,可使这些植物生长,且用相同修饰品系或不同品系授粉,且所得杂合体具有基因修饰。本发明提供基因修饰种子。再生植物的后代、变体及突变体也包含于本发明的范围内,前提条件是这些部分包括基因修饰。还提供了保持了基因修饰的经加工的植物产品或副产物,例如,包含豆粕。

[0220] 本文提供的方法可用于修饰任何植物物种,包括但不限于单子叶植物及双子叶植物。所关注的植物的实例包括但不限于:玉蜀黍(玉米)、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦和油菜、芸苔属物种、苜蓿、黑麦、小米、红花、花生、甘薯、木薯、咖啡、椰子、菠萝、柑橘树、可可、茶、香蕉、鳄梨、无花果、番石榴、芒果、橄榄、木瓜、腰果、澳洲胡桃、杏仁、燕麦、蔬菜、观赏植物以及针叶树。

[0221] 蔬菜包括但不限于:西红柿、莴苣、绿豆、青豆、豌豆和诸如黄瓜、哈密瓜及香瓜的黄瓜属的成员。观赏植物包括但不限于:杜鹃花、绣球花、芙蓉、玫瑰、郁金香、水仙、矮牵牛、康乃馨、猩猩木及菊花。优选地,本发明的植物为农作物(例如,玉米、高粱、小麦、向日葵、西红柿、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、甜菜、甘蔗、烟草、大麦、油菜等)。

[0222] 本文提供的方法还可用于基因修饰任何原核物种,包括但不限于:古菌及细菌(例如,芽孢杆菌属、克雷伯氏菌属、链霉菌属、根瘤菌属、埃希氏菌属、假单胞菌属、沙门氏菌属、志贺氏杆菌属、弧菌属、耶尔森菌属、支原体菌属、农杆菌属、乳酸乳杆菌属)。

[0223] 本文提供的方法可用于基因修饰任何真核物种或来自于其的细胞,包括但不限于:动物(例如,哺乳动物、昆虫、鱼类、鸟类和爬行动物)、真菌、变形虫、藻类和酵母。在一些实施方案中,本发明公开的方法修饰的细胞包含造血源的细胞,诸如,免疫系统的细胞(即,免疫细胞),包括但不限于:B细胞、T细胞、自然杀伤(NK)细胞、包含多能干细胞及经诱导的多能干细胞的干细胞、嵌合抗原受体T(CAR-T)细胞、单核细胞、巨噬细胞和树突细胞。

[0224] 可将经修饰的细胞可被引入生物体内。在自体细胞移植的情况中,这些细胞可源自同一个生物体(例如,人),其中该细胞以离体措施被修饰。作为另一种选择,在异体细胞移植的情况中,该细胞源自同一个物种中的另一个生物体(例如,另一个人)。

[0225] XII. 用于检测目标DNA或切割单链DNA群体的试剂盒和方法

[0226] 本发明公开的RGN,特别是APG09624及APG05405(如SEQ ID NO:2及4所示)一旦通过目标DNA的缺失而被活化则可不加区分地切割未靶向的单链DNA(ssDNA)。因此,本文提供用于检测样本中的目标DNA(双链或单链)的组合物及方法。

[0227] 检测DNA分子的目标DNA的方法包括:使样本与RGN(或编码其的多核苷酸)、能够与DNA分子中的RGN及目标DNA杂交的指导RNA(或编码其的多核苷酸)和与该指导RNA杂交的检测单链DNA(检测ssDNA)接触,随后,测量通过该RGN切割该ssDNA产生的可检测信号,从而检测该DNA分子的目标DNA序列。在一些实施方案中,方法可包含在与RGN及指导RNA接触之前或同时进行的样本中的核酸分子的扩增步骤。在这些实施方案中的一些实施方案中,为增加检测方法的灵敏度,该指导RNA将与之杂交的特异性序列可被扩增。

[0228] 在其中样本与编码RGN多肽的多核苷酸和/或编码该指导RNA的多核苷酸接触的那些实施方案中,样本包含完整细胞且该多核苷酸被引入细胞内,其中它们然后被表达。在这些实施方案中的一些实施方案中,至少一个多核苷酸还包含启动子,该启动子可操作地连接至编码该RGN多肽和/或指导RNA的核苷酸序列。

[0229] 在一些实施方案中,期望目标可作为RNA(诸如,例如诸如冠状病毒的RNA病毒的基因组或基因组的局部)存在。在一些实施方案中,冠状病毒可为SARS样冠状病毒。在进一步的实施方案中,冠状病毒可为SARS-CoV-2、SARS-CoV或诸如bat-SL-CoVZC45(寄存编号:MG772933)的蝙蝠SARS样冠状病毒。在其中该目标作为RNA存在的实施方案中,目标可被逆转录为可通过RGN有效靶向的DNA分子。逆转录之后可为诸如本领域中已知的涉及热循环的RT-PCR方法的扩增步骤,或可为诸如RT-LAMP(逆转录-环介导等温扩增)的等温方法(Notomi等人Nucleic Acids Res 28:E63, (2000))。

[0230] 该核酸扩增可在该样本与该RGN、指导RNA及检测ssDNA接触之前发生,或扩增可与该接触步骤同时地发生。

[0231] 在某些实施方案中,该方法涉及使样本与RGN及一个以上的指导RNA接触。为扩增该可检测信号且引起对那个DNA分子的检测,每一个均能够与RGN杂交的指导RNA可与单DNA分子的唯一目标序列结合。

[0232] 这些组合物及方法涉及不与该指导RNA杂交且为非目标ssDNA的检测ssDNA的使用。在一些实施方案中,检测ssDNA包含可检测标记,该可检测标记在切割该检测ssDNA后提供可检测信号。非限制性实例为包含荧光团/淬灭剂对的检测ssDNA,其中该荧光团当该检测ssDNA为完整的(即,未被切割)时因其信号通过紧密相邻的淬灭剂的存在被阻遏而不发荧光。检测ssDNA的切割导致淬灭剂的移除,且然后该荧光标记可被检测。荧光标记或荧光染料的非限制性实例包括Cy5、荧光素(例如,FAM、6FAM、5(6)FAM、FITC)、Cy3、Alexa Fluor®染料和得克萨斯红。淬灭剂的非限制性实例包括Iowa Black®FQ、Iowa Black®RQ、Qx1淬灭剂、ATTO淬灭剂和QSY染料。在一些实施方案中,检测ssDNA包含第二淬灭剂,诸如,比如ZEN™、TAO™和黑洞淬灭剂(Black Hole Quencher®)的内部淬灭剂,其可降低背景而增加信号检测。

[0233] 在其他实施方案中,检测ssDNA包含可检测标记,该可检测标记在切割检测ssDNA之前提供可检测信号,而ssDNA的切割抑制或防止该信号的检测。这种情况的非限制性实例为包含荧光共振能量转移(FRET)对的检测ssDNA。FRET为通过其能量的无辐射转移自第一

(供体)荧光团的激发态至紧密相邻的第二(受体)荧光团发生的过程。该供体荧光团的发射光谱与受体荧光团的激发光谱重叠。因此,当检测ssDNA为完整的(即,未被切割)时,受体荧光团将发荧光,而当检测ssDNA被切割时,因为供体与受体荧光团不再彼此紧密相邻,所以受体荧光团不再发荧光。FRET供体及受体荧光团为本领域中所知,且包括但不限于青色荧光蛋白(CFP)/绿色荧光蛋白(GFP)、Cy3/Cy5和GFP/黄色荧光蛋白(YFP)。

[0234] 在一些实施方案中,检测ssDNA具有的长度为约2个核苷酸至约30个核苷酸,包括但不限于:约2、约3、约4、约5、约6、约7、约8、约9、约10、约11、约12、约13、约14、约15、约16、约17、约18、约19、约20、约21、约22、约23、约24、约25个核苷酸、约26个核苷酸、约27个核苷酸、约28个核苷酸、约29个核苷酸和约30个核苷酸。

[0235] 其中使用包含检测ssDNA的这些组合物及方法可检测目标DNA的样本包括包含或据信包含核酸(例如,DNR或RNA分子)的任何样本。该样本可来源于包含纯化核酸的合成组合或诸如呼吸拭子(例如,鼻咽拭子)提取物、细胞裂解物、患者样本、细胞、组织、唾液、血液、血清、血浆、尿液、抽吸物、活检样本、脑脊液或生物体(例如,细菌、病毒)的生物样本的任何来源。

[0236] 使样本与RGN、指导RNA和检测ssDNA接触可包括在体外、离体或在体内接触。在一些实施方案中,检测ssDNA和/或RGN和/或指导RNA被固定在例如横向流动装置上,其中该样本接触经固定化检测ssDNA和/或RGN和/或指导RNA。在一些实施方案中,针对检测ssDNA上的抗原部分的抗体以自完整检测ssDNA允许分化经切割的检测ssDNA的方式被固定在例如横向流动装置上。还提供了装置(例如,横向流动装置、微流体装置),诸如,第W0 2020/028729号国际专利公开所描述的装置,该国际专利公开的全部内容通过引用并入本文,该装置包含经固定化检测ssDNA。可在对该装置添加该样本之前、同时或之后,对该样本添加该RGN及指导RNA,而当该目标DNA存在于该样本内时,该RGN将切割该目标DNA以及该检测ssDNA,引起可检测信号的增加或减少,该可检测信号可被测量,以检测该目标DNA序列的存在。作为另一种选择,该RGN和/或指导RNA固定在该装置(例如,横向流动装置、微流体装置)上,且将该样本及检测ssDNA添加至该装置。可在对该装置添加该样本之前、期间或之后,对该样本添加该检测ssDNA。另一种替代装置(例如,横向流动装置、微流体装置)包含针对该检测ssDNA上的抗原部分的经固定化抗体,且对该装置添加该样本、检测ssDNA、RGN及指导RNA。在一些实施方案中,方法还可包含确定样本中存在的目标DNA的量。可将测试样本中的可检测信号的测量值与参考测量值(例如,参考样本的或其包含已知量的目标DNA的系列的测量值)进行比较。

[0237] 该组合物及方法的应用的非限制性实例包括单核苷酸多态性(SNP)检测、癌症筛查、细菌感染的检测、抗生素抗性的检测和病毒感染的检测。

[0238] 使用本领域已知的任何适合方法可测量通过RGN切割该ssDNA产生的可检测信号,该方法包括但不限于测量荧光信号、凝胶上的带的视觉分析、比色变化以及电信号的存在或不存在。

[0239] 本发明提供用于检测样本中的DNA分子的目标DNA的试剂盒,其中该试剂盒包含:本发明的RGN多肽(或包含编码该RGN多肽的核苷酸序列的多核苷酸)、能够与该RGN及DNA分子中的目标DNA序列杂交的指导RNA(或包含编码该指导RNA的核苷酸序列的多核苷酸)和与该指导RNA杂交的检测ssDNA。在要检测的目标为RNA的那些实施方案中,试剂盒还可包含

逆转录酶。在其中使用核酸扩增的那些实施方案中,包含该RGN及指导RNA(或编码其的多核苷酸)及检测ssDNA的试剂盒还可包含核酸扩增试剂(例如,DNA聚合酶、多核苷酸、缓冲液)。在其中该试剂盒包含编码RGN多肽的多核苷酸和/或编码该指导RNA的多核苷酸的那些实施方案中,多核苷酸被引入细胞内,其中它们然后被表达。在这些实施方案中的一些实施方案中,至少一个多核苷酸还包含启动子,该启动子可操作地连接至编码该RGN多肽和/或指导RNA的核苷酸序列。在某些实施方案中,试剂盒包含一个以上的指导RNA(或编码一个以上的指导RNA的多核苷酸),每一个指导RNA都能够与该RGN杂交。为扩增该可检测信号且引起对那个DNA分子的检测,该指导RNA可与单DNA分子的唯一目标序列结合。该试剂盒的组件可个别地或组合地被提供,且可以被提供在任何适合的容器中,例如小瓶、瓶子或管子。在一些实施方案中,试剂盒包括一种或多种语言的说明。在一些实施方案中,试剂盒包含一种或更多种试剂,其在利用本文描述的一种或多种组件的过程中使用。试剂可被提供于任何适合的容器中。例如,试剂盒可以提供一种或多种反应或储存缓冲液。试剂可为以可用于特别测定的形式、或以在使用前需要添加一种或多种其他组成的形式(例如,以浓缩或冷冻干燥形式)被提供。缓冲液可以是任何缓冲液,包括但不限于碳酸钠缓冲液、碳酸氢钠缓冲液、硼酸盐缓冲液、Tris缓冲液、MOPS缓冲液、HEPES缓冲液及其组合。在一些实施方案中,缓冲剂为碱性的。在一些实施方案中,缓冲液具有约7至约10的pH。

[0240] 本文还提供了通过接触核酸群体切割单链DNA的方法,其中该群体包括DNA分子的目标DNA序列及多个非目标ssDNA,该非目标ssDNA具有RGN和能够与该RGN及该目标DNA序列杂交的指导RNA。在这些实施方案中的一些实施方案中,核酸群体在细胞裂解物内。在这些实施方案中的一些实施方案中,非目标ssDNA非该细胞所原有,且在这些实施方案中的一些实施方案中,非目标ssDNA为病毒DNA。在特别实施方案中,目标DNA序列为病毒序列。可在体外、体内、或离体进行该方法。例如,可在体内进行该方法,其中对个体施用RGN多肽及指导RNA或一种或多种包含编码该RGN多肽和/或该指导RNA的核苷酸序列的多核苷酸,且病毒目标DNA序列通过该RGN的结合及切割可导致所感染的细胞内的非目标病毒ssDNA的切割。

[0241] 冠词“a”和“an”在本文中用于指该冠词的语法对象中的一种或一种以上(即,至少一种)的语法对象。作为实例,“多肽”表示一种或多种多肽。

[0242] 本说明书中提及的所有出版物及专利申请表示本公开内容所属领域中具有通常知识者的水平。所有出版物及专利申请通过引用并入本文,如同每一个单独出版物或专利申请被具体且单独地指示通过引用而被并入本文。

[0243] 虽然为了清楚理解的目的已经作为示例及实例一定程度上详细地描述了前述发明,但显然可在所附实施方案的范围内实践某些改变及修改。

[0244] 非限制性实施方案包括:

[0245] 1. 一种核酸分子,包含编码RNA指导核酸酶(RGN)多肽的多核苷酸,其中该多核苷酸包含编码RGN多肽的核苷酸序列,该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列;

[0246] 其中当与能够与该目标DNA序列杂交的指导RNA(gRNA)结合时,该RGN多肽能够以RNA指导的序列特异性方式结合DNA分子的目标DNA序列,且

[0247] 其中编码RGN多肽的该多核苷酸可操作地连接至与该多核苷酸异源的启动子。

[0248] 2. 如实施方案1的核酸分子,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具

有至少95%序列同一性的氨基酸序列。

[0249] 3.如实施方案1的核酸分子,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。

[0250] 4.如实施方案1-3中任一实施方案的核酸分子,其中该目标DNA序列位于为单链的该DNA分子的区域内。

[0251] 5.如实施方案4的核酸分子,其中该RGN多肽能够在结合时切割该目标DNA序列。

[0252] 6.如实施方案1-3中任一实施方案的核酸分子,其中该目标DNA序列位于为双链的该DNA分子的区域内。

[0253] 7.如实施方案6的核酸分子,其中该RGN多肽能够在结合时切割该目标DNA序列。

[0254] 8.如实施方案7的核酸分子,其中通过该RGN多肽的切割产生双链断裂。

[0255] 9.如实施方案7的核酸分子,其中通过该RGN多肽的切割产生单链断裂。

[0256] 10.如实施方案1-9中任一实施方案的核酸分子,其中该RGN多肽与碱基编辑多肽可操作地融合。

[0257] 11.如实施方案10的核酸分子,其中该碱基编辑多肽为脱氨酶。

[0258] 12.如实施方案1-11中任一实施方案的核酸分子,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

[0259] 13.如实施方案1-12中任一实施方案的核酸分子,其中该RGN多肽包含一种或多种核定位信号。

[0260] 14.如实施方案1-13中任一实施方案的核酸分子,其中该RGN多肽针对在真核细胞中的表达而被密码子优化。

[0261] 15.一种包含如实施方案1-14中任一实施方案的核酸分子的载体。

[0262] 16.如实施方案15的载体,还包含至少一个核苷酸序列,该至少一个核苷酸序列编码选自以下组成的组的RGN辅助蛋白:

[0263] a)与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;

[0264] b)与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0265] c)与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;

[0266] d)与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0267] e)与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列;和

[0268] f)与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0269] 17.如实施方案16的载体,其中该RGN辅助蛋白选自以下组成的组:

[0270] a)与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0271] b)与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助

蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0272] c) 与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0273] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0274] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0275] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0276] 18. 如实施方案16的载体,其中该RGN辅助蛋白选自以下组成的组:

[0277] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0278] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0279] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0280] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0281] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0282] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0283] 19. 如实施方案15-18中任一实施方案的载体,还包含至少一个核苷酸序列,该至少一个核苷酸序列编码能够与该目标DNA序列杂交的该gRNA。

[0284] 20. 如实施方案19的载体,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列,且该gRNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

[0285] 21. 如实施方案19的载体,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列,且该gRNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

[0286] 22. 如实施方案19的载体,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列,且该gRNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

[0287] 23. 如实施方案19的载体,其中该gRNA包含tracrRNA。

[0288] 24. 如实施方案23的载体,其中该tracrRNA选自以下组成的组:

[0289] a) 与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

[0290] b) 与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中该gRNA还包含

CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

[0291] c)与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

[0292] d)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

[0293] e)与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

[0294] f)与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

[0295] g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0296] h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

[0297] i)与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0298] 25.如实施方案23的载体,其中该tracrRNA选自以下组成的组:

[0299] a)与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

[0300] b)与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

[0301] c)与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

[0302] d)与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

[0303] e)与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

[0304] f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

[0305] g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0306] h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

[0307] i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0308] 26. 如实施方案23的载体,其中该tracrRNA选自以下组成的组:

[0309] a) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

[0310] b) 与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

[0311] c) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

[0312] d) 与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

[0313] e) 与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

[0314] f) 与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

[0315] g) 与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0316] h) 与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

[0317] i) 与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中该gRNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中

该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0318] 27.如实施方案23-26中任一实施方案的载体,其中该gRNA为单指导RNA。

[0319] 28.如实施方案23-26中任一实施方案的载体,其中该gRNA为双指导RNA。

[0320] 29.一种包含如实施方案1-14中任一实施方案的核酸分子或如实施方案15-28中任一实施方案的载体的细胞。

[0321] 30.一种制备RGN多肽的方法,包含在该RGN多肽被表达的条件下,培养实施方案29的细胞。

[0322] 31.一种制备RGN多肽的方法,包含将异源核酸分子引入细胞内,该异源核酸分子包含编码RNA指导核酸酶(RGN)多肽的核苷酸序列,该RNA指导核酸酶(RGN)多肽包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列;

[0323] 其中当与能够与该目标DNA序列杂交的指导RNA(gRNA)结合时,该RGN多肽以RNA指导的序列特异性方式结合DNA分子的目标DNA序列;

[0324] 且在该RGN多肽被表达的条件下,培养该细胞。

[0325] 32.如实施方案31的方法,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有至少95%序列同一性的氨基酸序列。

[0326] 33.如实施方案31的方法,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有100%序列同一性的氨基酸序列。

[0327] 34.如实施方案30-33中任一实施方案的方法,还包括纯化该RGN多肽。

[0328] 35.如实施方案30-33中任一实施方案的方法,其中该细胞还表达一种或多种指导RNA,该一种或多种指导RNA与该RGN多肽结合,以形成RGN核糖核蛋白复合物。

[0329] 36.如实施方案35的方法,还包括纯化该RGN核糖核蛋白复合物。

[0330] 37.一种分离的RNA指导核酸酶(RGN)多肽,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列;和

[0331] 其中当与能够与该目标DNA序列杂交的指导RNA(gRNA)结合时,该RGN多肽能够以RNA指导的序列特异性方式结合DNA分子的目标DNA序列。

[0332] 38.如实施方案37的分离的RGN多肽,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有至少95%序列同一性的氨基酸序列。

[0333] 39.如实施方案37的分离的RGN多肽,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有100%序列同一性的氨基酸序列。

[0334] 40.如实施方案37-39中任一实施方案的分离的RGN多肽,其中该目标DNA序列位于为单链的该DNA分子的区域内。

[0335] 41.如实施方案40的分离的RGN多肽,其中该RGN多肽能够在结合时切割该目标DNA序列。

[0336] 42.如实施方案37-39中任一实施方案的分离的RGN多肽,其中该目标DNA序列位于为双链的该DNA分子的区域内。

[0337] 43.如实施方案42的分离的RGN多肽,其中该RGN多肽能够在结合时切割该目标DNA序列。

[0338] 44.如实施方案43的分离的RGN多肽,其中通过该RGN多肽的切割产生双链断裂。

[0339] 45.如实施方案43的分离的RGN多肽,其中通过该RGN多肽的切割产生单链断裂。

[0340] 46.如实施方案37-45中任一实施方案的分离的RGN多肽,其中该RGN多肽与碱基编辑多肽可操作地融合。

[0341] 47.如实施方案46的分离的RGN多肽,其中该碱基编辑多肽为脱氨酶。

[0342] 48.如实施方案37-47中任一实施方案的分离的RGN多肽,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

[0343] 49.如实施方案37-48中任一实施方案的分离的RGN多肽,其中该RGN多肽包含一种或多种核定位信号。

[0344] 50.一种包含编码CRISPR RNA(crRNA)的多核苷酸的核酸分子,其中该crRNA包含间隔序列及CRISPR重复序列,其中该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少90%序列同一性的核苷酸序列;

[0345] 其中指导RNA包含:

[0346] a) 该crRNA;或

[0347] b) 该crRNA及能够与该crRNA的该CRISPR重复序列杂交的反式活化的CRISPR RNA(tracrRNA);

[0348] 当该指导RNA与RNA指导核酸酶(RGN)多肽结合时,通过该crRNA的间隔序列,能够以序列特异性方式与DNA分子的目标DNA序列杂交,且

[0349] 其中编码crRNA的该多核苷酸被可操作地连接至与该多核苷酸异源的启动子。

[0350] 51.如实施方案50的核酸分子,其中该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少95%序列同一性的核苷酸序列。

[0351] 52.如实施方案50的核酸分子,其中该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有100%序列同一性的核苷酸序列。

[0352] 53.一种包含实施方案50-52中任一个的核酸分子的载体。

[0353] 54.如实施方案53的载体,其中该载体还包含编码该tracrRNA的多核苷酸。

[0354] 55.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:110具有至少90%序列同一性,且该tracrRNA包含与SEQ ID NO:120具有至少90%序列同一性的核苷酸序列。

[0355] 56.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:110具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:120具有至少95%序列同一性的核苷酸序列。

[0356] 57.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:110具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:120具有100%序列同一性的核苷酸序列。

[0357] 58.如实施方案55-57中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列。

[0358] 59.如实施方案58的载体,其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列。

[0359] 60.如实施方案58的载体,其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列。

[0360] 61.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:111具有至少90%序列同一性,且该tracrRNA包含与SEQ ID NO:121具有至少90%序列同一性的核苷酸序列。

[0361] 62.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:111具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:121具有至少95%序列同一性的核苷酸序列。

[0362] 63.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:111具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:121具有100%序列同一性的核苷酸序列。

[0363] 64.如实施方案61-63中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列。

[0364] 65.如实施方案64的载体,其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列。

[0365] 66.如实施方案64的载体,其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列。

[0366] 67.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:112具有至少90%序列同一性,且该tracrRNA包含与SEQ ID NO:122具有至少90%序列同一性的核苷酸序列。

[0367] 68.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:112具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:122具有至少95%序列同一性的核苷酸序列。

[0368] 69.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:112具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:122具有100%序列同一性的核苷酸序列。

[0369] 70.如实施方案67-69中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列。

[0370] 71.如实施方案70的载体,其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列。

[0371] 72.如实施方案70的载体,其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列。

[0372] 73.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:113具有至少90%序列同一性,且该tracrRNA包含与SEQ ID NO:123具有至少90%序列同一性的核苷酸序列。

[0373] 74.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:113具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:123具有至少95%序列同一性的核苷酸序列。

[0374] 75.如实施方案54的载体,其中该CRISPR重复序列与SEQ ID NO:113具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:123具有100%序列同一性的核苷酸序列。

[0375] 76.如实施方案73-75中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列。

[0376] 77.如实施方案76的载体,其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列。

[0377] 78.如实施方案76的载体,其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列。

[0378] 79.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:114具有至少90%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:124具有至少90%序列同一性的核苷酸序列。

[0379] 80.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:114具有至少95%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:124具有至少95%序列同一性的核苷酸序列。

[0380] 81.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:114具有

100%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:124具有100%序列同一性的核苷酸序列。

[0381] 82.如实施方案79-81中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列。

[0382] 83.如实施方案82的载体,其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列。

[0383] 84.如实施方案82的载体,其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列。

[0384] 85.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:115具有至少90%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:125具有至少90%序列同一性的核苷酸序列。

[0385] 86.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:115具有至少95%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:125具有至少95%序列同一性的核苷酸序列。

[0386] 87.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:115具有100%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:125具有100%序列同一性的核苷酸序列。

[0387] 88.如实施方案85-87中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列。

[0388] 89.如实施方案88的载体,其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列。

[0389] 90.如实施方案88的载体,其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列。

[0390] 91.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:117具有至少90%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:126具有至少90%序列同一性的核苷酸序列。

[0391] 92.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:117具有至少95%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:126具有至少95%序列同一性的核苷酸序列。

[0392] 93.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:117具有100%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:126具有100%序列同一性的核苷酸序列。

[0393] 94.如实施方案91-93中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列。

[0394] 95.如实施方案94的载体,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列。

[0395] 96.如实施方案94的载体,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列。

[0396] 97.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:118具有至少

90%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:127具有至少90%序列同一性的核苷酸序列。

[0397] 98.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:118具有至少95%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:127具有至少95%序列同一性的核苷酸序列。

[0398] 99.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:118具有100%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:127具有100%序列同一性的核苷酸序列。

[0399] 100.如实施方案97-99中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列。

[0400] 101.如实施方案100的载体,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列。

[0401] 102.如实施方案100的载体,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列。

[0402] 103.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:119具有至少90%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:128具有至少90%序列同一性的核苷酸序列。

[0403] 104.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:119具有至少95%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:128具有至少95%序列同一性的核苷酸序列。

[0404] 105.如实施方案54的载体,其中该CRISPR重复序列包含与SEQ ID NO:119具有100%序列同一性的核苷酸序列,且该tracrRNA包含与SEQ ID NO:128具有100%序列同一性的核苷酸序列。

[0405] 106.如实施方案103-105中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0406] 107.如实施方案106的载体,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0407] 108.如实施方案106的载体,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0408] 109.如实施方案54-108中任一实施方案的载体,其中编码该crRNA的该多核苷酸及编码该tracrRNA的该多核苷酸可操作地连接至同一个启动子,且被编码为单指导RNA。

[0409] 110.如实施方案54-109中任一实施方案的载体,其中编码该crRNA的该多核苷酸及编码该tracrRNA的该多核苷酸可操作地连接至各别启动子。

[0410] 111.一种包含编码反式活化的CRISPR RNA(tracrRNA)的多核苷酸的核酸分子,该反式活化的CRISPR RNA(tracrRNA)包含与SEQ ID NO:120至128中任一个具有至少90%序列同一性的核苷酸序列;

[0411] 其中指导RNA包含:

[0412] a) 该tracrRNA;和

[0413] b) crRNA,该crRNA包含间隔序列及CRISPR重复序列,其中该tracrRNA能够与该

crRNA的该CRISPR重复序列杂交；

[0414] 当该指导RNA与RNA指导核酸酶(RGN)多肽结合时,通过该crRNA的间隔序列,能够以序列特异性方式与目标DNA序列杂交,且

[0415] 其中编码tracrRNA的该多核苷酸被可操作地连接至与该多核苷酸异源的启动子。

[0416] 112.如实施方案111的核酸分子,其中该tracrRNA包含与SEQ ID NO:120至128中任一个具有至少95%序列同一性的核苷酸序列。

[0417] 113.如实施方案111的核酸分子,其中该tracrRNA包含与SEQ ID NO:120至128中任一个具有100%序列同一性的核苷酸序列。

[0418] 114.一种包含实施方案111-113中任一个的核酸分子的载体。

[0419] 115.如实施方案114的载体,其中该载体还包含编码该crRNA的多核苷酸。

[0420] 116.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:120具有至少90%序列同一性的核苷酸序列。

[0421] 117.如实施方案116的载体,其中该CRISPR重复序列与SEQ ID NO:110具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:120具有至少95%序列同一性的核苷酸序列。

[0422] 118.如实施方案116的载体,其中该CRISPR重复序列与SEQ ID NO:110具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:120具有100%序列同一性的核苷酸序列。

[0423] 119.如实施方案116-118中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列。

[0424] 120.如实施方案119的载体,其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列。

[0425] 121.如实施方案119的载体,其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列。

[0426] 122.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:121具有至少90%序列同一性的核苷酸序列。

[0427] 123.如实施方案122的载体,其中该CRISPR重复序列与SEQ ID NO:111具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:121具有至少95%序列同一性的核苷酸序列。

[0428] 124.如实施方案122的载体,其中该CRISPR重复序列与SEQ ID NO:111具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:121具有100%序列同一性的核苷酸序列。

[0429] 125.如实施方案122-124中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列。

[0430] 126.如实施方案125的载体,其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列。

[0431] 127.如实施方案125的载体,其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列。

[0432] 128.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:112具有至少90%序

列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:122具有至少90%序列同一性的核苷酸序列。

[0433] 129.如实施方案128的载体,其中该CRISPR重复序列与SEQ ID NO:112具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:122具有至少95%序列同一性的核苷酸序列。

[0434] 130.如实施方案128的载体,其中该CRISPR重复序列与SEQ ID NO:112具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:122具有100%序列同一性的核苷酸序列。

[0435] 131.如实施方案128-130中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列。

[0436] 132.如实施方案131的载体,其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列。

[0437] 133.如实施方案131的载体,其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列。

[0438] 134.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:123具有至少90%序列同一性的核苷酸序列。

[0439] 135.如实施方案134的载体,其中该CRISPR重复序列与SEQ ID NO:113具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:123具有至少95%序列同一性的核苷酸序列。

[0440] 136.如实施方案134的载体,其中该CRISPR重复序列与SEQ ID NO:113具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:123具有100%序列同一性的核苷酸序列。

[0441] 137.如实施方案134-136中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列。

[0442] 138.如实施方案137的载体,其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列。

[0443] 139.如实施方案137的载体,其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列。

[0444] 140.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:124具有至少90%序列同一性的核苷酸序列。

[0445] 141.如实施方案140的载体,其中该CRISPR重复序列与SEQ ID NO:114具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:124具有至少95%序列同一性的核苷酸序列。

[0446] 142.如实施方案140的载体,其中该CRISPR重复序列与SEQ ID NO:114具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:124具有100%序列同一性的核苷酸序列。

[0447] 143.如实施方案140-142中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列。

[0448] 144.如实施方案143的载体,其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列。

[0449] 145.如实施方案143的载体,其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列。

[0450] 146.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:125具有至少90%序列同一性的核苷酸序列。

[0451] 147.如实施方案146的载体,其中该CRISPR重复序列与SEQ ID NO:115具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:125具有至少95%序列同一性的核苷酸序列。

[0452] 148.如实施方案146的载体,其中该CRISPR重复序列与SEQ ID NO:115具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:125具有100%序列同一性的核苷酸序列。

[0453] 149.如实施方案146-148中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列。

[0454] 150.如实施方案149的载体,其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列。

[0455] 151.如实施方案149的载体,其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列。

[0456] 152.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:126具有至少90%序列同一性的核苷酸序列。

[0457] 153.如实施方案152的载体,其中该CRISPR重复序列与SEQ ID NO:117具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:126具有至少95%序列同一性的核苷酸序列。

[0458] 154.如实施方案152的载体,其中该CRISPR重复序列与SEQ ID NO:117具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:126具有100%序列同一性的核苷酸序列。

[0459] 155.如实施方案152-154中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0460] 156.如实施方案155的载体,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0461] 157.如实施方案155的载体,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0462] 158.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:127具有至少90%序列同一性的核苷酸序列。

[0463] 159.如实施方案158的载体,其中该CRISPR重复序列与SEQ ID NO:118具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:127具有至少95%序列同一性的核苷酸序列。

[0464] 160.如实施方案158的载体,其中该CRISPR重复序列与SEQ ID NO:118具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:127具有100%序列同一性的核苷酸序列。

[0465] 161.如实施方案158-160中任一实施方案的载体,其中该载体还包含编码该RGN多

肽的多核苷酸,该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列。

[0466] 162.如实施方案161的载体,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列。

[0467] 163.如实施方案161的载体,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列。

[0468] 164.如实施方案115的载体,其中该crRNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且该tracrRNA包含与SEQ ID NO:128具有至少90%序列同一性的核苷酸序列。

[0469] 165.如实施方案164的载体,其中该CRISPR重复序列与SEQ ID NO:119具有至少95%序列同一性,且该tracrRNA包含与SEQ ID NO:128具有至少95%序列同一性的核苷酸序列。

[0470] 166.如实施方案164的载体,其中该CRISPR重复序列与SEQ ID NO:119具有100%序列同一性,且该tracrRNA包含与SEQ ID NO:128具有100%序列同一性的核苷酸序列。

[0471] 167.如实施方案164-166中任一实施方案的载体,其中该载体还包含编码该RGN多肽的多核苷酸,该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0472] 168.如实施方案167的载体,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0473] 169.如实施方案167的载体,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0474] 170.如实施方案115-169中任一实施方案的载体,其中编码该crRNA的该多核苷酸及编码该tracrRNA的该多核苷酸可操作地连接至同一个启动子,且被编码为单指导RNA。

[0475] 171.如实施方案115-169中任一实施方案的载体,其中编码该crRNA的该多核苷酸及编码该tracrRNA的该多核苷酸可操作地连接至各别启动子。

[0476] 172.一种用于结合DNA分子的目标DNA序列的系统,该系统包含:

[0477] a) 能够与该目标DNA序列杂交的一种或多种指导RNA,或包含编码该一种或多种指导RNA(gRNA)的一种或多种核苷酸序列的一种或多种多核苷酸;和

[0478] b) 包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽,或包含编码该RGN多肽的核苷酸序列的多核苷酸;

[0479] 其中编码该一种或多种指导RNA的至少一个该核苷酸序列及编码该RGN多肽的该核苷酸序列可操作地连接至与该核苷酸序列异源的启动子;且

[0480] 其中该一种或多种指导RNA能够与该RGN多肽形成复合物,以便将该RGN多肽导向至与该DNA分子的该目标DNA序列结合。

[0481] 173.一种用于结合DNA分子的目标DNA序列的系统,该系统包含:

[0482] a) 能够与该目标DNA序列杂交的一种或多种指导RNA,或包含编码该一种或多种指导RNA(gRNA)的一种或多种核苷酸序列的一种或多种多核苷酸;和

[0483] b) 包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽;

[0484] 其中该一种或多种指导RNA能够与该目标DNA序列杂交,且

[0485] 其中该一种或多种指导RNA能够与该RGN多肽形成复合物,以便将该RGN多肽导向

至与该DNA分子的该目标DNA序列结合。

[0486] 174.如实施方案173的系统,其中编码该一种或多种指导RNA的至少一个该核苷酸序列可操作地连接至与该核苷酸序列异源的启动子。

[0487] 175.如实施方案172-174中任一实施方案的系统,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列。

[0488] 176.如实施方案172-174中任一实施方案的系统,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。

[0489] 177.如实施方案172-176中任一实施方案的系统,其中未发现该RGN多肽与该一种或多种指导RNA在自然界中彼此复合。

[0490] 178.如实施方案172-177中任一实施方案的系统,其中该目标DNA序列为真核目标DNA序列。

[0491] 179.如实施方案172-178中任一实施方案的系统,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

[0492] 180.如实施方案172-178中任一实施方案的系统,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

[0493] 181.如实施方案172-178中任一实施方案的系统,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

[0494] 182.如实施方案172-181中任一实施方案的系统,其中该一种或多种指导RNA包含tracrRNA。

[0495] 183.如实施方案182的系统,其中该tracrRNA选自以下组成的组:

[0496] a)与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

[0497] b)与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

[0498] c)与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

[0499] d)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

[0500] e) 与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

[0501] f) 与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

[0502] g) 与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0503] h) 与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

[0504] i) 与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0505] 184.如实施方案182的系统,其中该tracrRNA选自以下组成的组:

[0506] a) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

[0507] b) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

[0508] c) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

[0509] d) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

[0510] e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的

CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

[0511] f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

[0512] g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0513] h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

[0514] i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0515] 185.如实施方案182的系统,其中该tracrRNA选自以下组成的组:

[0516] a) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

[0517] b) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

[0518] c) 与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

[0519] d) 与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

[0520] e) 与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

[0521] f) 与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

[0522] g) 与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0523] h) 与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

[0524] i) 与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0525] 186.如实施方案182-185中任一实施方案的系统,其中该一种或多种指导RNA为单指导RNA (sgRNA)。

[0526] 187.如实施方案182-185中任一实施方案的系统,其中该一种或多种指导RNA为双指导RNA。

[0527] 188.如实施方案172-187中任一实施方案的系统,其中该系统还包含至少一个RGN辅助蛋白或包含编码其的核苷酸序列的多核苷酸,其选自以下组成的组:

[0528] a) 与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;

[0529] b) 与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0530] c) 与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;

[0531] d) 与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0532] e) 与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列;和

[0533] f) 与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0534] 189.如实施方案188的系统,其中该至少一个RGN辅助蛋白选自以下组成的组:

[0535] a) 与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0536] b) 与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0537] c) 与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0538] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0539] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0540] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0541] 190.如实施方案188的系统,其中该至少一个RGN辅助蛋白选自以下组成的组:

[0542] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0543] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0544] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0545] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0546] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0547] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0548] 191.如实施方案172-190中任一实施方案的系统,其中该目标DNA序列位于细胞内。

[0549] 192.如实施方案191的系统,其中该细胞为真核细胞。

[0550] 193.如实施方案192的系统,其中该真核细胞为植物细胞。

[0551] 194.如实施方案192的系统,其中该真核细胞为哺乳动物细胞。

[0552] 195.如实施方案194的系统,其中该哺乳动物细胞为人细胞。

[0553] 196.如实施方案195的系统,其中该人细胞为免疫细胞。

[0554] 197.如实施方案196的系统,其中该人细胞为干细胞。

[0555] 198.如实施方案197的系统,其中该干细胞为经诱导的多能干细胞。

[0556] 199.如实施方案192的系统,其中该真核细胞为昆虫细胞。

[0557] 200.如实施方案191的系统,其中该细胞为原核细胞。

[0558] 201.如实施方案172-200中任一实施方案的系统,其中该目标DNA序列位于为单链的该DNA分子的区域内。

[0559] 202.如实施方案201的系统,其中当被转录时,该一种或多种指导RNA能够与该目标DNA序列杂交,且该指导RNA能够与该RGN多肽形成复合物,以导向该目标DNA序列的切割。

[0560] 203.如实施方案172-200中任一实施方案的系统,其中该目标DNA序列位于为双链的该DNA分子的区域内。

[0561] 204.如实施方案203的系统,其中当被转录时,该一种或多种指导RNA能够与该目

标DNA序列杂交,且该指导RNA能够与该RGN多肽形成复合物,以导向该目标DNA序列的切割。

[0562] 205.如实施方案204的系统,其中该RGN多肽能够产生双链断裂。

[0563] 206.如实施方案204的系统,其中该RGN多肽能够产生单链断裂。

[0564] 207.如实施方案172-206中任一实施方案的系统,其中该RGN多肽可操作地连接至碱基编辑多肽。

[0565] 208.如实施方案207的系统,其中该碱基编辑多肽为脱氨酶。

[0566] 209.如实施方案172-208中任一实施方案的系统,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

[0567] 210.如实施方案172-209中任一实施方案的系统,其中该RGN多肽包含一种或多种核定位信号。

[0568] 211.如实施方案172-210中任一实施方案的系统,其中该RGN多肽针对在真核细胞中的表达而被密码子优化。

[0569] 212.如实施方案172-211中任一实施方案的系统,其中包含编码一种或多种指导RNA的核苷酸序列的多核苷酸及包含编码RGN多肽的核苷酸序列的多核苷酸位于一个载体上。

[0570] 213.如实施方案172-212中任一实施方案的系统,其中该系统还包含一种或多种供体多核苷酸或包含编码该一种或多种供体多核苷酸的一种或多种核苷酸序列的一种或多种多核苷酸。

[0571] 214.一种药物组合物,包含实施方案1-14、50-52和111-113中任一实施方案的核酸分子、实施方案15-28、53-110和114-171中任一实施方案的载体、实施方案29的细胞、实施方案37-49中任一实施方案的分离的RGN多肽、或实施方案172-213中任一实施方案的系统及药学上可接受的载剂。

[0572] 215.一种结合DNA分子的目标DNA序列的方法,包含将根据实施方案172-213中任一实施方案的系统递送至该目标DNA序列或包含该目标DNA序列的细胞。

[0573] 216.如实施方案215的方法,其中该RGN多肽或该指导RNA还包含可检测标记,从而允许对该目标DNA序列的检测。

[0574] 217.如实施方案215的方法,其中该指导RNA或该RGN多肽还包含表达调控子,从而调控该目标DNA序列的或受该目标DNA序列的转录控制的基因的表达。

[0575] 218.一种切割DNA分子的目标DNA序列的方法,包含将根据实施方案172-213中任一实施方案的系统递送至该目标DNA序列或包含该目标DNA序列的细胞。

[0576] 219.如实施方案218的方法,其中该经修饰的目标DNA序列包含异源DNA在该目标DNA序列内的插入。

[0577] 220.如实施方案218的方法,其中该经修饰的目标DNA序列包含至少一个核苷酸自该目标DNA序列的缺失。

[0578] 221.如实施方案218的方法,其中该经修饰的目标DNA序列包含该目标DNA序列中的至少一个核苷酸的突变。

[0579] 222.一种用于结合DNA分子的目标DNA序列的方法,该方法包括:

[0580] a) 在适合形成RGN核糖核苷酸复合物的条件下,通过组合以下,以在体外组装RNA指导核酸酶(RGN)核糖核苷酸复合物:

- [0581] i) 能够与该目标DNA序列杂交的一种或多种指导RNA;和
- [0582] ii) RGN多肽,该RGN多肽包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列;和
- [0583] b) 使该目标DNA序列或包含该目标DNA序列的细胞与该在体外组装的RGN核糖核苷酸复合物接触;
- [0584] 其中该一种或多种指导RNA与该目标DNA序列杂交,从而将该RGN多肽导向至与该目标DNA序列结合。
- [0585] 223.如实施方案222的方法,其中该目标DNA序列位于为单链的该DNA分子的区域内。
- [0586] 224.如实施方案222的方法,其中该目标DNA序列位于为双链的该DNA分子的区域内。
- [0587] 225.如实施方案222-224中任一实施方案的方法,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。
- [0588] 226.如实施方案222-225中任一实施方案的方法,其中该RGN多肽或该指导RNA还包含可检测标记,从而允许对该目标DNA序列的检测。
- [0589] 227.如实施方案222-225中任一实施方案的方法,其中该指导RNA或该RGN多肽还包含表达调控子,从而允许对该目标DNA序列的表达的调控。
- [0590] 228.一种用于切割和/或修饰DNA分子的目标DNA序列的方法,包含使该DNA分子与以下接触:
- [0591] a) RNA指导核酸酶(RGN)多肽,其中该RGN包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列;和
- [0592] b) 能够将(a)的RGN靶向至该目标DNA序列的一种或多种指导RNA;
- [0593] 其中该一种或多种指导RNA与该目标DNA序列杂交,从而将该RGN多肽导向至与该目标DNA序列结合,并且发生该目标DNA序列的切割和/或修饰。
- [0594] 229.如实施方案228的方法,其中该目标DNA序列位于为单链的该DNA分子的区域内。
- [0595] 230.如实施方案228的方法,其中该目标DNA序列位于为双链的该DNA分子的区域内。
- [0596] 231.如实施方案230的方法,其中通过该RGN多肽的切割产生双链断裂。
- [0597] 232.如实施方案230的方法,其中通过该RGN多肽的切割产生单链断裂。
- [0598] 233.如实施方案228-232中任一实施方案的方法,其中该RGN多肽可操作地连接至碱基编辑多肽。
- [0599] 234.如实施方案233的方法,其中该碱基编辑多肽包括脱氨酶。
- [0600] 235.如实施方案228-234中任一实施方案的方法,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。
- [0601] 236.如实施方案228-235中任一实施方案的方法,其中该经修饰的目标DNA序列包含异源DNA在该目标DNA序列内的插入。
- [0602] 237.如实施方案228-235中任一实施方案的方法,其中该经修饰的目标DNA序列包含至少一个核苷酸自该目标DNA序列的缺失。

[0603] 238.如实施方案228-235中任一实施方案的方法,其中该经修饰的目标DNA序列包含该目标DNA序列中的至少一个核苷酸的突变。

[0604] 239.如实施方案222-238中任一实施方案的方法,其中该RGN包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列。

[0605] 240.如实施方案222-238中任一实施方案的方法,其中该RGN包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。

[0606] 241.如实施方案222-240中任一实施方案的方法,其中未发现该RGN多肽与该一种或多种指导RNA在自然界中彼此复合。

[0607] 242.如实施方案222-241中任一实施方案的方法,其中该目标DNA序列为真核目标DNA序列。

[0608] 243.如实施方案222-242中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

[0609] 244.如实施方案222-242中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

[0610] 245.如实施方案222-242中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

[0611] 246.如实施方案222-242中任一实施方案的方法,其中该一种或多种指导RNA包含tracrRNA。

[0612] 247.如实施方案246的方法,其中该tracrRNA选自以下组成的组:

[0613] a) 与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

[0614] b) 与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

[0615] c) 与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

[0616] d) 与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

[0617] e) 与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中该一种或多种指

导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

[0618] f) 与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

[0619] g) 与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0620] h) 与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

[0621] i) 与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0622] 248.如实施方案246的方法,其中该tracrRNA选自以下组成的组:

[0623] a) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

[0624] b) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

[0625] c) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

[0626] d) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

[0627] e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸

序列;

[0628] f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

[0629] g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0630] h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

[0631] i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0632] 249.如实施方案246的方法,其中该tracrRNA选自以下组成的组:

[0633] a) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

[0634] b) 与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

[0635] c) 与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

[0636] d) 与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

[0637] e) 与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

[0638] f) 与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中该一种或多种指导

RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

[0639] g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0640] h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

[0641] i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0642] 250.如实施方案246-249中任一实施方案的方法,其中该一种或多种指导RNA为单指导RNA (sgRNA)。

[0643] 251.如实施方案246-249中任一实施方案的方法,其中该一种或多种指导RNA为双指导RNA。

[0644] 252.如实施方案222-251中任一实施方案的方法,其中该方法还包含使DNA分子与选自以下组成的组的一种或多种RGN辅助蛋白接触:

[0645] a)与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;

[0646] b)与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0647] c)与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;

[0648] d)与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0649] e)与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列;和

[0650] f)与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0651] 253.如实施方案252的方法,其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0652] a)与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0653] b)与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0654] c)与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助

蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0655] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0656] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0657] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0658] 254. 如实施方案252的方法,其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0659] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0660] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0661] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0662] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0663] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0664] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0665] 255. 如实施方案215-254中任一实施方案的方法,其中该目标DNA序列位于细胞内。

[0666] 256. 如实施方案255的方法,其中该细胞为真核细胞。

[0667] 257. 如实施方案256的方法,其中该真核细胞为植物细胞。

[0668] 258. 如实施方案256的方法,其中该真核细胞为哺乳动物细胞。

[0669] 259. 如实施方案258的方法,其中该哺乳动物细胞为人细胞。

[0670] 260. 如实施方案259的方法,其中该人细胞为免疫细胞。

[0671] 261. 如实施方案260的方法,其中该人细胞为干细胞。

[0672] 262. 如实施方案261的方法,其中该干细胞为经诱导的多能干细胞。

[0673] 263. 如实施方案256的方法,其中该真核细胞为昆虫细胞。

[0674] 264. 如实施方案255的方法,其中该细胞为原核细胞。

[0675] 265. 一种根据实施方案228-254中任一实施方案的方法包括经修饰的目标DNA序列的细胞。

[0676] 266. 如实施方案265的细胞,其中该细胞为真核细胞。

[0677] 267. 如实施方案266的细胞,其中该真核细胞为植物细胞。

[0678] 268. 一种包含实施方案267的细胞的植物。

[0679] 269. 一种包含实施方案267的细胞的种子。

[0680] 270. 如实施方案266的细胞,其中该真核细胞为哺乳动物细胞。

[0681] 271. 如实施方案270的细胞,其中该哺乳动物细胞为人细胞。

- [0682] 272.如实施方案271的细胞,其中该人细胞为免疫细胞。
- [0683] 273.如实施方案272的细胞,其中该人细胞为干细胞。
- [0684] 274.如实施方案273的细胞,其中该干细胞为经诱导的多能干细胞。
- [0685] 275.如实施方案266的细胞,其中该真核细胞为昆虫细胞。
- [0686] 276.如实施方案265的细胞,其中该细胞为原核细胞。
- [0687] 277.一种包含实施方案266及270-274中任一实施方案的细胞的药物组合物及药学上可接受的载剂。
- [0688] 278.一种用于检测样本中的DNA分子的目标DNA序列的试剂盒,该试剂盒包含:
- [0689] a) 包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽或包含编码该RGN多肽的核苷酸序列的多核苷酸,其中当与能够与该目标DNA序列杂交的指导RNA结合时,该RGN多肽能够以RNA指导的序列特异性方式结合及切割DNA分子的该目标DNA序列;
- [0690] b) 该指导RNA或包含编码该指导RNA的核苷酸序列的多核苷酸;和
- [0691] c) 不与指导RNA杂交的检测单链DNA(ssDNA)。
- [0692] 279.如实施方案278的试剂盒,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列。
- [0693] 280.如实施方案278的试剂盒,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。
- [0694] 281.如实施方案278-280中任一实施方案的试剂盒,其中编码该指导RNA的至少一个该核苷酸序列及编码该RGN多肽的该核苷酸序列可操作地连接至与该核苷酸序列异源的启动子。
- [0695] 282.如实施方案278-281中任一实施方案的试剂盒,其中未发现该RGN多肽与该一种或多种指导RNA在自然界中彼此复合。
- [0696] 283.如实施方案278-282中任一实施方案的试剂盒,其中该目标DNA序列为真核目标DNA序列。
- [0697] 284.如实施方案278-283中任一实施方案的试剂盒,其中该检测ssDNA包含荧光团/淬灭剂对。
- [0698] 285.如实施方案278-283中任一实施方案的试剂盒,其中该检测ssDNA包含荧光共振能量转移(FRET)对。
- [0699] 286.如实施方案278-285中任一实施方案的试剂盒,其中该试剂盒还包含选自以下组成的组的至少一个RGN辅助蛋白:
- [0700] a) 与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;
- [0701] b) 与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;
- [0702] c) 与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;
- [0703] d) 与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0704] e) 与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列;和

[0705] f) 与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0706] 287.如实施方案286的试剂盒,其中该至少一个RGN辅助蛋白选自以下组成的组:

[0707] a) 与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0708] b) 与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0709] c) 与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0710] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0711] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0712] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0713] 288.如实施方案286的试剂盒,其中该至少一个RGN辅助蛋白选自以下组成的组:

[0714] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0715] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0716] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0717] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0718] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0719] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0720] 289.如实施方案278-288中任一实施方案的试剂盒,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列,且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

[0721] 290.如实施方案278-288中任一实施方案的试剂盒,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列,且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

[0722] 291.如实施方案278-288中任一实施方案的试剂盒,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列,且该一种或多种指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

[0723] 292.如实施方案278-288中任一实施方案的试剂盒,其中该一种或多种指导RNA包含tracrRNA。

[0724] 293.如实施方案292的试剂盒,其中该tracrRNA选自以下组成的组:

[0725] a)与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

[0726] b)与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

[0727] c)与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

[0728] d)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

[0729] e)与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

[0730] f)与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

[0731] g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0732] h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

[0733] i)与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0734] 294.如实施方案292的试剂盒,其中该tracrRNA选自以下组成的组:

[0735] a) 与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

[0736] b) 与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

[0737] c) 与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

[0738] d) 与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

[0739] e) 与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

[0740] f) 与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

[0741] g) 与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0742] h) 与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

[0743] i) 与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0744] 295.如实施方案292的试剂盒,其中该tracrRNA选自以下组成的组:

[0745] a) 与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中该一种或多种指导

RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

[0746] b)与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

[0747] c)与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

[0748] d)与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

[0749] e)与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

[0750] f)与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

[0751] g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0752] h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

[0753] i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中该一种或多种指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0754] 296.如实施方案292-295中任一实施方案的试剂盒,其中该一种或多种指导RNA为单指导RNA(sgRNA)。

[0755] 297.如实施方案292-295中任一实施方案的试剂盒,其中该一种或多种指导RNA为双指导RNA。

- [0756] 298.如实施方案278-297中任一实施方案的试剂盒,其中该目标DNA序列位于为单链的该DNA分子的区域内。
- [0757] 299.如实施方案278-297中任一实施方案的试剂盒,其中该目标DNA序列位于为双链的该DNA分子的区域内。
- [0758] 300.如实施方案299的试剂盒,其中通过该RGN多肽的切割产生双链断裂。
- [0759] 301.如实施方案299的试剂盒,其中通过该RGN多肽的切割产生单链断裂。
- [0760] 302.如实施方案278-301中任一实施方案的试剂盒,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。
- [0761] 303.一种检测样本中的DNA分子的目标DNA序列的方法,该方法包括:
- [0762] a) 使样本接触:
- [0763] i) 包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶(RGN)多肽,其中当与能够与该目标DNA序列杂交的指导RNA结合时,该RGN多肽能够以RNA指导的序列特异性方式结合及切割DNA分子的该目标DNA序列;
- [0764] ii) 该指导RNA;和
- [0765] iii) 不与指导RNA杂交的检测单链DNA(ssDNA);和
- [0766] b) 测量通过RGN切割检测ssDNA产生的可检测信号,从而测量目标DNA序列。
- [0767] 304.如实施方案303的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列。
- [0768] 305.如实施方案303的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。
- [0769] 306.如实施方案303-305中任一实施方案的方法,其中该样本包含来自细胞裂解物的DNA分子。
- [0770] 307.如实施方案303-305中任一实施方案的方法,其中该样本包含细胞。
- [0771] 308.如实施方案307的方法,其中该细胞为真核细胞。
- [0772] 309.如实施方案303-305中任一实施方案的方法,其中包含目标DNA序列的DNA分子通过存在于包含RNA的样本中的RNA模板分子的逆转录产生。
- [0773] 310.如实施方案309的方法,其中该RNA模板分子为RNA病毒。
- [0774] 311.如实施方案310的方法,其中该RNA病毒为冠状病毒。
- [0775] 312.如实施方案311的方法,其中该冠状病毒为蝙蝠SARS样冠状病毒、SARS-CoV或SARS-CoV-2。
- [0776] 313.如实施方案309-312中任一实施方案的方法,其中包括RNA的样本为自包含细胞的样本取得的。
- [0777] 314.如实施方案303-313中任一实施方案的方法,其中该检测ssDNA包含荧光团/淬灭剂对。
- [0778] 315.如实施方案303-313中任一实施方案的方法,其中该检测ssDNA包括荧光共振能量转移(FRET)对。
- [0779] 316.如实施方案303-315中任一实施方案的方法,其中该方法还包含在步骤a)的接触之前或与步骤a)的接触一起扩增样本中的核酸。
- [0780] 317.如实施方案303-316中任一实施方案的方法,其中该方法还包含使样本与选

自以下组成的组的一种或多种RGN辅助蛋白接触:

[0781] a) 与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;

[0782] b) 与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0783] c) 与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;

[0784] d) 与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0785] e) 与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列;和

[0786] f) 与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0787] 318. 如实施方案317的方法,其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0788] a) 与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0789] b) 与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0790] c) 与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0791] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0792] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0793] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0794] 319. 如实施方案317的方法,其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0795] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0796] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0797] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0798] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0799] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0800] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0801] 320. 一种切割单链DNA (ssDNA) 的方法, 该方法包括使核酸群体与以下接触, 其中该群体包含DNA分子, 该DNA分子包含目标DNA序列及多个非目标ssDNA:

[0802] a) 包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列的RNA指导核酸酶 (RGN) 多肽, 其中当与能够与该目标DNA序列杂交的指导RNA结合时, 该RGN多肽能够以RNA指导的序列特异性方式结合及切割该目标DNA序列; 以及

[0803] b) 该指导RNA;

[0804] 其中该RGN多肽切割该多个的非目标ssDNA。

[0805] 321. 如实施方案320的方法, 其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列。

[0806] 322. 如实施方案320的方法, 其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列。

[0807] 323. 如实施方案320-322中任一实施方案的方法, 其中该核酸群体在细胞裂解物内。

[0808] 324. 如实施方案320-323中任一实施方案的方法, 其中包含目标DNA序列的DNA分子通过RNA模板分子的逆转录产生。

[0809] 325. 如实施方案320-324中任一实施方案的方法, 其中该方法还包含使群体与选自以下组成的组的一种或多种RGN辅助蛋白接触:

[0810] a) 与SEQ ID NO:178-181中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列;

[0811] b) 与SEQ ID NO:182-184中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0812] c) 与SEQ ID NO:185-187中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;

[0813] d) 与SEQ ID NO:191具有至少90%序列同一性的RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:14具有至少90%序列同一性的氨基酸序列;

[0814] e) 与SEQ ID NO:192具有至少90%序列同一性的RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:15具有至少90%序列同一性的氨基酸序列; 和

[0815] f) 与SEQ ID NO:188-190中任一个具有至少90%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0816] 326. 如实施方案325的方法, 其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0817] a) 与SEQ ID NO:178-181中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列;

[0818] b) 与SEQ ID NO:182-184中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0819] c) 与SEQ ID NO:185-187中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;

[0820] d) 与SEQ ID NO:191具有至少95%序列同一性的RGN辅助蛋白, 其中该RGN多肽包含与SEQ ID NO:14具有至少95%序列同一性的氨基酸序列;

[0821] e) 与SEQ ID NO:192具有至少95%序列同一性的RGN辅助蛋白, 其中该RGN多肽包

含与SEQ ID NO:15具有至少95%序列同一性的氨基酸序列;和

[0822] f) 与SEQ ID NO:188-190中任一个具有至少95%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0823] 327.如实施方案325的方法,其中该一种或多种RGN辅助蛋白选自以下组成的组:

[0824] a) 与SEQ ID NO:178-181中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列;

[0825] b) 与SEQ ID NO:182-184中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0826] c) 与SEQ ID NO:185-187中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;

[0827] d) 与SEQ ID NO:191具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:14具有100%序列同一性的氨基酸序列;

[0828] e) 与SEQ ID NO:192具有100%序列同一性的RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:15具有100%序列同一性的氨基酸序列;和

[0829] f) 与SEQ ID NO:188-190中任一个具有100%序列同一性的至少一个RGN辅助蛋白,其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0830] 328.如实施方案303-327中任一实施方案的方法,其中未发现该RGN多肽与该指导RNA在自然界中彼此复合。

[0831] 329.如实施方案303-328中任一实施方案的方法,其中该目标DNA序列为真核目标DNA序列。

[0832] 330.如实施方案303-329中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有至少90%序列同一性的氨基酸序列且该指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少90%序列同一性的CRISPR重复序列。

[0833] 331.如实施方案303-329中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有至少95%序列同一性的氨基酸序列且该指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有至少95%序列同一性的CRISPR重复序列。

[0834] 332.如实施方案303-329中任一实施方案的方法,其中该RGN多肽包含与SEQ ID NO:11具有100%序列同一性的氨基酸序列且该指导RNA包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:116具有100%序列同一性的CRISPR重复序列。

[0835] 333.如实施方案303-329中任一实施方案的方法,其中该指导RNA包含tracrRNA。

[0836] 334.如实施方案333的方法,其中该tracrRNA选自以下组成的组:

[0837] a) 与SEQ ID NO:120具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少90%序列同一性的氨基酸序列;

[0838] b) 与SEQ ID NO:121具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少90%序列同一性的氨基酸序列;

[0839] c) 与SEQ ID NO:122具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少90%序列同一性的CRISPR重复

序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少90%序列同一性的氨基酸序列;

[0840] d)与SEQ ID NO:123具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少90%序列同一性的氨基酸序列;

[0841] e)与SEQ ID NO:124具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少90%序列同一性的氨基酸序列;

[0842] f)与SEQ ID NO:125具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少90%序列同一性的氨基酸序列;

[0843] g)与SEQ ID NO:126具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少90%序列同一性的氨基酸序列;

[0844] h)与SEQ ID NO:127具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少90%序列同一性的氨基酸序列;和

[0845] i)与SEQ ID NO:128具有至少90%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少90%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少90%序列同一性的氨基酸序列。

[0846] 335.如实施方案333的方法,其中该tracrRNA选自以下组成的组:

[0847] a)与SEQ ID NO:120具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有至少95%序列同一性的氨基酸序列;

[0848] b)与SEQ ID NO:121具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有至少95%序列同一性的氨基酸序列;

[0849] c)与SEQ ID NO:122具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有至少95%序列同一性的氨基酸序列;

[0850] d)与SEQ ID NO:123具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有至少95%序列同一性的氨基酸序列;

[0851] e)与SEQ ID NO:124具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有至少95%序列同一性的氨基酸序列;

[0852] f)与SEQ ID NO:125具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有至少95%序列同一性的氨基酸序列;

[0853] g)与SEQ ID NO:126具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含

CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有至少95%序列同一性的氨基酸序列;

[0854] h)与SEQ ID NO:127具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有至少95%序列同一性的氨基酸序列;和

[0855] i)与SEQ ID NO:128具有至少95%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有至少95%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有至少95%序列同一性的氨基酸序列。

[0856] 336.如实施方案333的方法,其中该tracrRNA选自以下组成的组:

[0857] a)与SEQ ID NO:120具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:110具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:1具有100%序列同一性的氨基酸序列;

[0858] b)与SEQ ID NO:121具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:111具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:2具有100%序列同一性的氨基酸序列;

[0859] c)与SEQ ID NO:122具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:112具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:3具有100%序列同一性的氨基酸序列;

[0860] d)与SEQ ID NO:123具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:113具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:4具有100%序列同一性的氨基酸序列;

[0861] e)与SEQ ID NO:124具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:114具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:5具有100%序列同一性的氨基酸序列;

[0862] f)与SEQ ID NO:125具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:115具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:6具有100%序列同一性的氨基酸序列;

[0863] g)与SEQ ID NO:126具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:117具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:12具有100%序列同一性的氨基酸序列;

[0864] h)与SEQ ID NO:127具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:118具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:13具有100%序列同一性的氨基酸序列;和

[0865] i)与SEQ ID NO:128具有100%序列同一性的tracrRNA,其中该指导RNA还包含CRISPR RNA,该CRISPR RNA包含与SEQ ID NO:119具有100%序列同一性的CRISPR重复序列,且其中该RGN多肽包含与SEQ ID NO:16具有100%序列同一性的氨基酸序列。

[0866] 337.如实施方案333-336中任一实施方案的方法,其中该指导RNA为单指导RNA(sgRNA)。

[0867] 338.如实施方案333-336中任一实施方案的方法,其中该指导RNA为双指导RNA。

[0868] 339.如实施方案303-338中任一实施方案的方法,其中该目标DNA序列位于为单链的该DNA分子的区域内。

[0869] 340.如实施方案303-338中任一实施方案的方法,其中该目标DNA序列位于为双链的该DNA分子的区域内。

[0870] 341.如实施方案340的方法,其中该目标DNA序列通过该RGN多肽的切割产生双链断裂。

[0871] 342.如实施方案340的方法,其中该目标DNA序列通过该RGN多肽的切割产生单链断裂。

[0872] 343.如实施方案303-342中任一实施方案的方法,其中该目标DNA序列位于与前间隔序列邻近基序(PAM)相邻的位置。

[0873] 344.一种利用针对基因遗传性疾病的因果突变的校正来产生基因修饰细胞的方法,该方法包括将以下引入该细胞内:

[0874] a) RNA指导核酸酶(RGN)多肽或编码该RGN多肽的多核苷酸,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列,其中编码该RGN多肽的该多核苷酸可操作地连接至启动子,以赋能该RGN多肽在该细胞中的表达;和

[0875] b) 指导RNA(gRNA)或编码该gRNA的多核苷酸,其中该gRNA包含CRISPR重复序列,该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少90%序列同一性的核苷酸序列,其中编码该gRNA的该多核苷酸可操作地连接至启动子,以赋能该gRNA在该细胞中的表达,

[0876] 由此该RGN和gRNA靶向该因果突变的基因组位置且修饰该基因组序列以除去该因果突变。

[0877] 345.如实施方案344的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列,且该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少95%序列同一性的核苷酸序列。

[0878] 346.如实施方案344的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少100%序列同一性的氨基酸序列,且该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少100%序列同一性的核苷酸序列。

[0879] 347.如实施方案344-346中任一实施方案的方法,其中该RGN可操作地连接至碱基编辑多肽。

[0880] 348.如实施方案347的方法,其中该碱基编辑多肽为脱氨酶。

[0881] 349.如实施方案344-348中任一实施方案的方法,其中该细胞为动物细胞。

[0882] 350.如实施方案344-348中任一实施方案的方法,其中该细胞为哺乳动物细胞。

[0883] 351.如实施方案349的方法,其中该细胞为从狗、猫、小鼠、大鼠、兔、马、牛、猪或人取得的。

[0884] 352.如实施方案349的方法,其中该基因遗传性疾病是单核苷酸多态性引起的。

[0885] 353.如实施方案351的方法,其中该基因遗传性疾病为赫尔勒综合征。

[0886] 354.如实施方案353的方法,其中该gRNA还包含靶向靠近该因果单核苷酸多态性的区域的间隔序列。

[0887] 355.一种利用引起疾病的不稳定基因组区域中的缺失来产生基因修饰细胞的方法

法,该方法包括将以下引入该细胞内:

[0888] a) RNA指导核酸酶 (RGN) 多肽或编码该RGN多肽的多核苷酸,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少90%序列同一性的氨基酸序列,其中编码该RGN多肽的该多核苷酸可操作地连接至启动子,以赋能该RGN多肽在该细胞中的表达;以及

[0889] b) 第指导RNA (gRNA) 或编码该第一gRNA的多核苷酸,其中该第一gRNA包含CRISPR重复序列,该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少90%序列同一性的核苷酸序列,其中编码该第一gRNA的该多核苷酸可操作地连接至启动子,以赋能该第一gRNA在该细胞中的表达,且进一步地其中该第一gRNA包含靶向该不稳定基因组区域的5'侧翼的间隔序列;和

[0890] c) 第二指导RNA (gRNA) 或编码该第二gRNA的多核苷酸,其中该第二gRNA包含CRISPR重复序列,该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少90%序列同一性的核苷酸序列,其中编码该第二gRNA的该多核苷酸可操作地连接至启动子,以赋能该第二gRNA在该细胞中的表达,并且进一步地其中该第二gRNA包含靶向该不稳定基因组区域的3'侧翼的间隔序列;

[0891] 由此该RGN及该第一及第二gRNA靶向该不稳定基因组区域,并且该不稳定基因组区域的至少一部分被除去。

[0892] 356. 如实施方案355的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列,且该第一gRNA及该第二gRNA的该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少95%序列同一性的核苷酸序列。

[0893] 357. 如实施方案355的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列,且该第一gRNA及该第二gRNA的该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有100%序列同一性的核苷酸序列。

[0894] 358. 如实施方案355-357中任一实施方案的方法,其中该细胞为动物细胞。

[0895] 359. 如实施方案355-357中任一实施方案的方法,其中该细胞为哺乳动物细胞。

[0896] 360. 如实施方案359的方法,其中该细胞是从狗、猫、小鼠、大鼠、兔、马、牛、猪或人取得的。

[0897] 361. 如实施方案358的方法,其中该基因遗传性疾病为弗里德赖希共济失调或亨廷顿病。

[0898] 362. 如实施方案358的方法,其中该第一gRNA的间隔序列还靶向该不稳定基因组区域内的或靠近该不稳定基因组区域的区域。

[0899] 363. 如实施方案362的方法,其中该第二gRNA的间隔序列还靶向该不稳定基因组区域内的或靠近该不稳定基因组区域的区域。

[0900] 364. 一种用于产生具有降低的BCL11A mRNA和蛋白质表达的基因修饰的哺乳动物造血祖细胞的方法,该方法包括将以下引入分离的人造血祖细胞内:

[0901] a) RNA指导核酸酶 (RGN) 多肽或编码该RGN多肽的多核苷酸,其中该RGN多肽包含与SEQ ID NO:1-109中任一个具有至少90%序列同一性的氨基酸序列,其中编码该RGN多肽的该多核苷酸可操作地连接至启动子,以赋能该RGN多肽在该细胞中的表达;和

[0902] b) 指导RNA (gRNA) 或编码该gRNA的多核苷酸,其中该gRNA包含CRISPR重复序列,该CRISPR重复序列包括与SEQ ID NO:110-119中任一个具有至少90%序列同一性的核苷酸序

列,其中编码该gRNA的该多核苷酸可操作地连接至启动子,以赋能该gRNA在该细胞中的表达,

[0903] 由此该RGN和gRNA在细胞中表达并在BCL11A增强子区域处切割,导致该人造血祖细胞的基因修饰且降低BCL11A的mRNA和/或蛋白质表达。

[0904] 365.如实施方案364的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有至少95%序列同一性的氨基酸序列,且该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有至少95%序列同一性的核苷酸序列。

[0905] 366.如实施方案364的方法,其中该RGN多肽包含与SEQ ID NO:1至109中任一个具有100%序列同一性的氨基酸序列,且该CRISPR重复序列包含与SEQ ID NO:110至119中任一个具有100%序列同一性的核苷酸序列。

[0906] 367.如实施方案364-366中任一实施方案的方法,其中该gRNA还包含靶向位于该BCL11A增强子区域内的或靠近该BCL11A增强子区域的区域的间隔序列。

[0907] 368.如实施方案344-367中任一实施方案的方法,其中该指导RNA、第指导RNA和第二指导RNA包含tracrRNA。

[0908] 369.如实施方案368的方法,其中该tracrRNA包含与SEQ ID NO:120至128中任一个具有至少90%序列同一性的核苷酸序列。

[0909] 370.如实施方案368的方法,其中该tracrRNA包含与SEQ ID NO:120至128中任一个具有至少95%序列同一性的核苷酸序列。

[0910] 371.如实施方案368的方法,其中该tracrRNA包含与SEQ ID NO:120至128中任一个具有100%序列同一性的核苷酸序列。

[0911] 372.一种治疗疾病的方法,该方法包括对需要治疗的个体施用有效量的实施方案214或277的药物组合物。

[0912] 373.如实施方案372的方法,其中该疾病与因果突变关联,且该有效量的该药物组合物校正该基因突变。

[0913] 374.实施方案1-14、50-52和111-113中任一个的核酸分子、实施方案15-28、53-110和114-171中任一个的载体、实施方案29、266和270-274中任一个的细胞、实施方案37-49中任一个的分离的RGN多肽、或实施方案172-213中任一个的系统对于治疗个体的疾病的用途。

[0914] 375.如实施方案374的用途,其中该疾病与因果突变关联,且该治疗包含校正该因果突变。

[0915] 376.实施方案1-14、50-52和111-113中任一个的核酸分子、实施方案15-28、53-110和114-171中任一个的载体、实施方案29、266和270-274中任一个的细胞、实施方案37-49中任一个的分离的RGN多肽、或实施方案172-213中任一个的系统对于制造有用于治疗个体的药品的用途。

[0916] 377.如实施方案376的用途,其中该疾病与因果突变关联,且有效量的该药品校正该因果突变。

[0917] 提供以下实施例为示例而非限制。

[0918] 实验

[0919] 实施例1.RNA指导核酸酶的鉴定

[0920] 在所关注的基因组中鉴定与转座酶具序列相似性的CRISPR关联序列。CRISPR重复序列在被设置为2的阵列中用最少数目的重复序列通过minCED(探勘环境数据集(Environmental Datasets)中的CRISPR重复序列)被鉴定。仅考虑对与同一个重叠群(contig)上的重复序列共定位的推定RNA指导核酸酶做进一步调查。使用重叠群上重复序列与推定cas基因之间距离的若干渐增的严格截断(100kb、50kb、20kb、10kb)。选择5kb的最终过滤器。

[0921] 在该文献中,活性系统中的CRISPR重复序列的长度介于27与47个核苷酸之间。该特征被用于过滤且移除非CRISPR重复特征。为对阵列扩展提供适宜的化学性质,CRISPR阵列中的间隔序列的一部分获得需要重复序列的第一核苷酸为G。作为此步骤的一部分,共有重复序列以及重复序列-间隔序列阵列的取向被预测。包括此过滤器以优先考虑可能的功能性RGN。阵列中所需重复序列的最小数目被增加至3。

[0922] 一些蛋白质(主要是DNA结合蛋白质)在其一级结构中具有重复氨基酸,其可能被错误地检测为CRISPR基因座。自身重复特征发生于蛋白质内部的推定RGN被弃去。仅考虑对基因间重复序列做进一步分析。

[0923] 为确定同源物的簇,蛋白质及共有重复序列被排列且基于它们的谱系被分类为簇。重复序列及蛋白质在谱系上往往共簇,对同源物的簇的概念提供支持。蛋白质簇信息也显示于表1中。使用cdhit,以95%的同一性,使蛋白质成簇。

[0924] 品系关系(relatedness)可通过比较它们的间隔序列内容跟踪。如果系统具有相同的重复序列及类似的蛋白质序列,则称它们有关系,且它们的间隔序列内容可提供有关于它们的共享历史的信息。相同的簇中的同源物往往共享保留的祖间隔序列。趋异品系(Divergent strain)彼此具有整体唯一的间隔序列内容。克隆分离物将完全共享完全相同的间隔序列内容。自身重复序列被保留但自身间隔序列内容在基因组之间不同的系统可促进更可能为活性的且这些被优先考虑。

[0925] 鉴定了109种不同的CRISPR关联的RNA指导核酸酶(RGN)并描述于下表1中。表1提供每一个RGN的名称、其氨基酸序列、其被取得所自的来源和经处理的crRNA及tracrRNA序列。

[0926] 对围绕每一个推定核酸酶的基因座搜索CRISPR免疫可能需要的潜在辅助基因。若干推定核酸酶出现在具潜在辅助基因的操纵子结构中,但没有基因座含有cas1、cas2、cas4或其他已知cas基因的同源物(图1)。另外,若干系统含有对齐该核酸酶的末端且在CRISPR重复序列的上游缺少期望的前导序列的重复序列,这表明CRISPR RNA表达的一种新机制。

[0927] 表1:RGN系统信息

[0928]

RGN ID	SEQ ID NO.	源	蛋白质簇	crRNA 重复序列 (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO.)	共有重复序列 (SEQ ID NO.)
APG07339	1	芽孢杆菌属	4	110	120	129	201
APG09624	2	芽孢杆菌属	4	111	121	130	202
APG03003	3	芽孢杆菌属	4	112	122		203
APG05405	4	芽孢杆菌属	4	113	123	131	204
APG09777	5	芽孢杆菌属	4	114	124	132	205
APG05680	6	芽孢杆菌属	16	115	125	310	206
APG02119	7	芽孢杆菌属	4				207
APG03285	8	芽孢杆菌属	15				208
APG04998	9	芽孢杆菌属	11				209
APG07078	10	芽孢杆菌属	13				210
APG06369	11	戈登氏菌属	0	116			211
APG03847	12	微球菌属	2	117	126	133	212
APG05625	13	微球菌属	2	118	127	134	213
APG03759	14	Paeniglutamicibacter 菌属	5				214
APG05123	15	链霉菌属	3				215
APG03524	16	微球菌属	2	119	128	135	216
APG05361	17	芽孢杆菌属	1				217
APG04303	18	芽孢杆菌属	14				218
APG04291	19	芽孢杆菌属	14				219
APG01006	20	芽孢杆菌属	14				220
APG06547	21	芽孢杆菌属	22				221
APG01699	22	芽孢杆菌属	14				222
APG06155	23	芽孢杆菌属	14				223
APG09116	24	芽孢杆菌属	4				224
APG09403	25	芽孢杆菌属	23				225

[0929]

RGN ID	SEQ ID NO.	源	蛋白质簇	crRNA 重复序列 (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO.)	共有重复序列 (SEQ ID NO.)
APG08954	26	芽孢杆菌属	14				226
APG02589	27	芽孢杆菌属	14				227
APG04061	28	芽孢杆菌属	14				228
APG08773	29	芽孢杆菌属	12				229
APG02836	30	芽孢杆菌属	14				230
APG09123	31	芽孢杆菌属	14				231
APG04288	32	芽孢杆菌属	14				232
APG06873	33	芽孢杆菌属	12				233
APG02381	34	芽孢杆菌属	14				234
APG06947	35	芽孢杆菌属	27				ND
APG04677	36	芽孢杆菌属	14				236
APG07253	37	芽孢杆菌属	14				237
APG08319	38	芽孢杆菌属	24				ND
APG04362	39	芽孢杆菌属	14				239
APG00992	40	芽孢杆菌属	14				240
APG04193	41	芽孢杆菌属	17				241
APG08201	42	芽孢杆菌属	17				242
APG01031	43	芽孢杆菌属	14				243
APG06773	44	芽孢杆菌属	14				244
APG08945	45	芽孢杆菌属	14				245
APG03214	46	芽孢杆菌属	4				246
APG09942	47	芽孢杆菌属	4				247
APG01836	48	芽孢杆菌属	14				248
APG06336	49	芽孢杆菌属	14				249
APG08839	50	芽孢杆菌属	4				250
APG02684	51	芽孢杆菌属	20				ND
APG05281	52	芽孢杆菌属	14				252
APG01046	53	芽孢杆菌属	14				253
APG01240	54	芽孢杆菌属	14				254
APG05981	55	芽孢杆菌属	14				255
APG04054	56	芽孢杆菌属	14				256
APG07301	57	芽孢杆菌属	14				257
APG07284	58	芽孢杆菌属	14				258
APG06812	59	芽孢杆菌属	14				259
APG08143	60	芽孢杆菌属	14				260
APG08031	61	芽孢杆菌属	14				261
APG03966	62	芽孢杆菌属	14				262
APG05371	63	芽孢杆菌属	14				263
APG04324	64	芽孢杆菌属	14				264

[0930]

RGN ID	SEQ ID NO.	源	蛋白质簇	crRNA 重复序列 (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO.)	共有重复序列 (SEQ ID NO.)
APG01233	65	芽孢杆菌属	14				265
APG00823	66	芽孢杆菌属	14				266
APG06704	67	芽孢杆菌属	14				267
APG02228	68	芽孢杆菌属	14				268
APG08636	69	芽孢杆菌属	6				269
APG02665	70	芽孢杆菌属	6				270
APG01832	71	芽孢杆菌属	6				271
APG03625	72	芽孢杆菌属	14				272
APG07479	73	芽孢杆菌属	14				273
APG02608	74	芽孢杆菌属	14				274
APG04337	75	芽孢杆菌属	12				275
APG01431	76	芽孢杆菌属	14				276
APG05423	77	芽孢杆菌属	14				277
APG01452	78	芽孢杆菌属	14				278
APG05729	79	芽孢杆菌属	19				279
APG01946	80	芽孢杆菌属	14				280
APG02414	81	芽孢杆菌属	14				281
APG01839	82	芽孢杆菌属	14				282
APG00752	83	芽孢杆菌属	14				283
APG02156	84	芽孢杆菌属	14				284
APG08789	85	芽孢杆菌属	10				285
APG07736	86	芽孢杆菌属	26				286
APG01573	87	芽孢杆菌属	14				287
APG07722	88	芽孢杆菌属	18				288
APG08071	89	芽孢杆菌属	8				289
APG01280	90	芽孢杆菌属	14				290
APG07455	91	芽孢杆菌属	14				291
APG05150	92	芽孢杆菌属	7				292
APG09405	93	芽孢杆菌属	14				293
APG09583	94	芽孢杆菌属	4				294
APG09909	95	芽孢杆菌属	14				295
APG07075	96	芽孢杆菌属	14				296
APG05892	97	芽孢杆菌属	14				297
APG09648	98	芽孢杆菌属	25				298
APG02311	99	芽孢杆菌属	21				299
APG03906	100	芽孢杆菌属	14				300
APG01953	101	芽孢杆菌属	9				301
APG00903	102	芽孢杆菌属	14				302
APG01543	103	芽孢杆菌属	22				303

	RGN ID	SEQ ID NO.	源	蛋白质簇	crRNA 重复序列 (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	sgRNA (SEQ ID NO.)	共有重复序列 (SEQ ID NO.)
[0931]	APG07261	104	芽孢杆菌属	15				ND
	APG05635	105	芽孢杆菌属	13				305
	APG02962	106	芽孢杆菌属	15				306
	APG07448	107	芽孢杆菌属	15				307
	APG05378	108	芽孢杆菌属	16				308
	APG01852	109	芽孢杆菌属	16				309

[0932] ND=未定

[0933] 实施例2:蛋白质分析

[0934] 通过在interpro数据库中搜索域,预测核酸酶域。使用建立于割裂RuvC域上的hmm核酸酶域概况从已知Cas蛋白预测割裂RuvC核酸酶域。所预测的核酸酶残基可见于表2中 (ND=未定)。

[0935] 表2:核酸酶残基

	RGN ID	RuvC-I	RuvC-II	RuvC-III
[0936]	APG07339	D297	E395	D477
	APG09624	D297	E395	D477
	APG03003	D297	E395	D477
	APG05405	D297	E395	D477
	APG09777	D297	E395	D477
	APG05680	D232	E331	D407
	APG02119	D297	E395	D477
	APG03285	D232	E331	D407
	APG04998	D281	E379	D456
	APG07078	D233	E332	D408
	APG06369	ND	ND	ND
	APG03847	D282	E382	D463
	APG05625	D282	E382	D463
	APG03759	D257	E357	D438
	APG05123	D248	E356	D437
	APG03524	D282	E382	D463
	APG05361	D244	D362	D443
	APG04303	D233	E332	D408
	APG04291	D233	E332	D408
	APG01006	D233	E332	D408

[0937]

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG06547	D183	E277	D359
APG01699	D233	E326	D402
APG06155	D233	E332	D408
APG09116	D297	E395	D477
APG09403	D183	E277	D359
APG08954	D233	E326	D402
APG02589	D236	E329	D405
APG04061	D233	E332	D408
APG08773	D267	E363	D440
APG02836	D233	E332	D408
APG09123	D233	E332	D408
APG04288	D233	E332	D408
APG06873	D267	E363	D440
APG02381	D233	E332	D408
APG06947	D2	E95	D184
APG04677	D233	E326	D402
APG07253	D233	E332	D408
APG08319	D183	E267	D350
APG04362	D233	E332	D408
APG00992	D233	E332	D408
APG04193	D232	E331	D407
APG08201	D232	E331	D407
APG01031	D226	E325	D401
APG06773	D233	E332	D408
APG08945	D233	E332	D408
APG03214	D297	E395	D477
APG09942	D297	E395	D477
APG01836	D233	E332	D408
APG06336	D233	E332	D408
APG08839	D297	E395	D477
APG02684	D221	E305	D387
APG05281	D34	E133	D209
APG01046	D233	E332	D408
APG01240	D233	E326	D402
APG05981	D233	E332	D408
APG04054	D233	E332	D408
APG07301	D233	E332	D408
APG07284	D233	E326	D402
APG06812	D233	E326	D402
APG08143	D233	E332	D408
APG08031	D233	E332	D408
APG03966	D233	E332	D408

[0938]

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG05371	D233	E332	D408
APG04324	D233	E326	D402
APG01233	D233	E332	D408
APG00823	D233	E332	D408
APG06704	D233	E332	D408
APG02228	D233	E332	D408
APG08636	D267	E365	D442
APG02665	D267	E365	D442
APG01832	D267	E365	D442
APG03625	D233	E332	D408
APG07479	D233	E332	D408
APG02608	D233	E332	D408
APG04337	D267	E365	D442
APG01431	D233	E332	D408
APG05423	D233	E332	D408
APG01452	D233	E332	D408
APG05729	D231	E330	D409
APG01946	D233	E332	D408
APG02414	D226	E325	D401
APG01839	D236	E329	D405
APG00752	D226	E325	D401
APG02156	D236	E329	D405
APG08789	D280	E378	D455
APG07736	D52	E145	D221
APG01573	D226	E319	D395
APG07722	D232	E331	D407
APG08071	D281	E379	D456
APG01280	D233	E332	D408
APG07455	D233	E332	D408
APG05150	D267	E365	D442
APG09405	D233	E326	D402
APG09583	D297	E395	D477
APG09909	D233	E332	D408
APG07075	D233	E326	D402
APG05892	D226	E319	D395
APG09648	D86	E170	D253
APG02311	D221	E305	D387
APG03906	D233	E332	D408
APG01953	D281	E379	D456
APG00903	D233	E332	D408
APG01543	D183	E277	D359
APG07261	D232	E331	D407

[0939]

RGN ID	RuvC-I	RuvC-II	RuvC-III
APG05635	D233	E332	D408
APG02962	D232	E331	D407
APG07448	D232	E331	D407
APG05378	D232	E331	D407
APG01852	D232	E331	D407

[0940] 实施例3:指导RNA预测及确认

[0941] 使天然表达了调查中的RNA指导核酸酶系统的细菌培养物生长至对数中期(约0.600的OD600)、成粒状物并快速冷冻。使用mirVANA miRNA分离试剂盒(Life Technologies, Carlsbad, CA)从粒状物中分离RNA,并使用NEBNext Small RNA Library Prep试剂盒(NEB, Beverly, MA)从分离的RNA制备测序文库。将文库制备物在6%聚丙烯酰胺凝胶上分级分离,以捕获小于200nt的RNA物质,从而分别检测crRNA和tracrRNA。由服务提供商(MoGene, St. Louis, MO)在Next Seq 500(高输出试剂盒)上进行深度测序(75bp配对末端)。使用Cutadapt对读数进行质量修整,并使用Bowtie2将读数映射到参考基因组。用python编写定制RNAseq管线(pipeline)来检测crRNA和tracrRNA转录物。通过天然重复序列间隔序列阵列的序列覆盖来确定经处理的crRNA边界。使用允许的BLASTn参数鉴定tracrRNA的抗重复序列部分。通过鉴定含有抗重复序列的转录物, RNA测序深度确认经处理的tracrRNA的边界。使用RNAfold(RNA折叠软件)进行的二级结构预测来进行RNA的手动管理(manual curation)。sgRNA盒是通过DNA合成而被制备的,并且一般如下设计(5'→3') : 经处理的tracrRNA,在其3'末端可操作地连接至4bp非互补接头(AAAG; SEQ ID NO:136)、在其3'末端可操作地连接至crRNA的经处理的重复序列部分,在其3'末端可操作地连接至20-30bp间隔序列。也可以使用其他4bp非互补接头。

[0942] 对于体外测定,通过用TranscriptAid T7高产率转录试剂盒(ThermoFisher)的sgRNA盒的体外转录来合成sgRNA及一些tracrRNA。以合成方式产生crRNA及一些tracrRNA。

[0943] 对于蛋白质表达及纯化,含有与C端His 10标签融合的推定RGN的质粒被构建且被转化到大肠杆菌的BL21(DE3)菌株中。使用补加了卡那霉素的Magic Media自诱导培养基进行表达。裂解以及澄清后,通过固定化金属亲和层析法纯化蛋白质。使用Heparin层析法,进行APG05405的进一步纯化。

[0944] 使用在tracrRNA的上游具T7启动子的dsDNA模板,通过体外转录(IVT)产生tracrRNA(SEQ ID NO:140、145和147)的较长的tracrRNA。用于IVT的模板通过PCR自合成gBlock模板(Integrated DNA Technologies)扩增。以合成方式产生较短的tracrRNA及crRNA。

[0945] 通过差示扫描荧光测定确认RNA结合(Niesen, F.H., H. Berglund, and M. Vedadi. 2007. Nat. Protoc. 2:2212-2221)。通过在Annealing Buffer(Synthego, 60mM KCl 6mM HEPES pH 7)中使过量crRNA与tracrRNA混合来产生双RNA复合物。在磷酸盐缓冲盐水(PBS, Thermo Fisher)中,以0.5μM效应子蛋白及1μM指导RNA的最终浓度,孵育候选效应子蛋白及指导RNA(或双RNA复合物或sgRNA)。在室温下,将它们孵育20分钟,且然后与已在PBS中被稀释的Sypro Orange染料溶液以1:1混合。解链曲线通过测量为温度的函数的荧光强度(FI)来获得,且计算该解链曲线的一次导数(dFI/dT)。为推定RNP的温度的函数的

dFI/dT的曲线图相对于原始核酸酶的位移指示RNA结合,且被用于评估推定指导RNA组合。在所鉴定的原始指导组合中,对于RGN APG09624及APG05405,仅观察到全长推定tracrRNA及crRNA诱发dFI/dT与温度的函数显著位移。对于其他蛋白质/RNA组合观察到此函数的小位移,且相较于以前对功能性RNP形成观察到的位移,该小位移的幅度较小,且不被解释为功能性复合物的形成的指示。第1峰是指对于给定样本与所观察到的最高峰关联的温度。如果第二峰被观察到,则其被指示在第2峰列中。关于复合物的形成的数据的解释被指示于表3的“结合?”列中。“是”指示结合被观察到。“N/A”指示没有足够的资料可用于确定结合是否发生。

[0946] 表3:RGN与指导RNA的结合

RGN	crRNA (SEQ ID NO.)	tracrRNA (SEQ ID NO.)	第1峰 (°C)	第2峰 (°C)	结合?
APG03003	无	无	38	未观察到	N/A
APG03003	139	140	38	未观察到	N/A
APG03003	141	142	37	未观察到	N/A
APG09777	无	无	39	未观察到	N/A
APG09777	139	140	40	未观察到	N/A
APG09777	141	142	39	未观察到	N/A
APG07339	无	无	37	未观察到	N/A
APG07339	143	144	37	未观察到	N/A
APG07339	143	145	38	未观察到	N/A
APG09624	无	无	38	未观察到	N/A
APG09624	143	144	37	未观察到	N/A
APG09624	143	145	39	44	是
APG05405	无	无	36	未观察到	N/A
APG05405	143	144	37	未观察到	N/A
APG05405	143	145	43	未观察到	是
APG09777	无	无	38	未观察到	N/A
APG09777	146	148	39	未观察到	N/A
APG09777	146	147	39	未观察到	N/A

[0948] 实施例4:导向ssDNA目标切割

[0949] 经纯化的APG09624、APG05405及催化灭活的APG05405(如SEQ ID NO:173所示的dAPG05405)与单指导RNA(sgRNA)Gsg.2(如SEQ ID NO:194所示的)在Cutsmart缓冲液(New England Biolabs B7204S)中在200nM核酸酶及400nM sgRNA的最终浓度下一起孵育20分钟。然后,以100nM核酸酶的最终浓度将它们添加至以5'Cy5示踪的ssDNA溶液中,该以5'Cy5示踪的ssDNA在本文中被称为以10nM位于1.5X Cutsmart缓冲液(New England Biolabs B7204S)中的LE111(如SEQ ID NO:195所示)或LE113(如SEQ ID NO:196所示)。

[0950] 样本通过分别以0.1mg/mL及45mM的最终浓度添加RNase及EDTA而被淬灭,且在以下时间点置于冰上:0、40、80和120分钟。淬灭全部样本后,将它们在50℃下孵育30分钟,然后,在95℃下孵育5分钟。将五分之一体积的上样缓冲液(1x TBE、12%Ficoll、7M尿素)添加至每一个反应物中,且在95℃下孵育15分钟,且在15%TBE-尿素丙烯酰胺凝胶(Bio-Rad 3450092)上分析5μl的每一个反应物。

[0951] 随时间、核酸酶及指导RNA的变化的切割产物的定量显示于下面的表4中。LE111 (SEQ ID NO:195) 的序列用作阴性对照,而LE113 (SEQ ID NO:196) 的序列承载被装载于核酸酶上的sgRNA的目标序列。

[0952] 表4:目标寡核苷酸的切割

[0953]

核酸酶	目标寡核苷酸	时间(分钟)	切割%
dAPG05405	LE111	0	0
dAPG05405	LE111	0	0
dAPG05405	LE111	40	0
dAPG05405	LE111	40	0
dAPG05405	LE111	80	0
dAPG05405	LE111	80	0
dAPG05405	LE111	120	0
dAPG05405	LE111	120	0
dAPG05405	LE113	0	0
dAPG05405	LE113	0	0
dAPG05405	LE113	40	0
dAPG05405	LE113	40	0
dAPG05405	LE113	80	0
dAPG05405	LE113	80	0
dAPG05405	LE113	120	0
dAPG05405	LE113	120	0
APG09624	LE111	0	0
APG09624	LE111	0	0
APG09624	LE111	40	0
APG09624	LE111	40	0
APG09624	LE111	80	24
APG09624	LE111	80	24
APG09624	LE111	120	30
APG09624	LE111	120	26
APG09624	LE113	0	0
APG09624	LE113	0	0
APG09624	LE113	40	59
APG09624	LE113	40	65
APG09624	LE113	80	80

[0954]

核酸酶	目标寡核苷酸	时间(分钟)	切割%
APG09624	LE113	80	84
APG09624	LE113	120	92
APG09624	LE113	120	93
APG05405	LE111	0	0
APG05405	LE111	0	0
APG05405	LE111	40	0
APG05405	LE111	40	0
APG05405	LE111	80	0
APG05405	LE111	80	0
APG05405	LE111	120	0
APG05405	LE111	120	0
APG05405	LE113	0	0
APG05405	LE113	0	0
APG05405	LE113	40	58
APG05405	LE113	40	49
APG05405	LE113	80	80
APG05405	LE113	80	85
APG05405	LE113	120	73
APG05405	LE113	120	88

[0955] 凝胶分析揭示主要位于样本中通过sgRNA靶向且具有催化活性核酸酶的切割产物依时间的形成,其证明这些蛋白质的RNA指导核酸酶活性,且提供关键催化残基被正确定义的证据。当APG09624 RNP与LE111一起孵育时,观察到一些非特异性活性,这可能是由于该批次的dAPG05405及APG05405是通过额外层析法步骤纯化的,且因此,由于未完全去除由表达宿主取得的核酸酶,该批次的APG09624可具有一定水平的背景活性。

[0956] 实施例5:可程序化的DNA结合及基因活化

[0957] RGN基因活化及gRNA哺乳动物表达质粒的构建

[0958] 哺乳动物表达的核酸酶构建体被合成。在巨细胞病毒(CMV)启动子(SEQ ID NO: 152)的控制下具N端SV40(SEQ ID NO:149)及C端核质素NLS序列(分别是,SEQ ID NO:149及150)、N端3xFLAG标签(SEQ ID NO:151)和C或N端VPR活化域(SEQ ID NO:154;Chavez等人2015,Nature Methods,12(4):326-328)的人密码子优化的APG05405被产生且被引入哺乳动物表达载体内。使用APG05405的推测具催化活性的及被催化灭活的(“dAPG5405”)两种形式。编码单gRNA的指导RNA表达构建体的每一个也都在人RNA聚合酶III U6启动子(SEQ ID NO:153)的控制下被产生。核酸酶构建体指示于下面的表5中。

[0959] 表5:RGN构建体

[0960]

核酸酶构建体	SEQ ID NO.
APG05405	155
dAPG05405	156
APG05405-VPR	157

dAPG05405-VPR	158
VPR-APG05405	159
VPR-dAPG05405	160

[0961] 哺乳动物细胞的转染及表达

[0962] 转染前一天,将 1×10^4 个HEK293T细胞(Sigma)铺板于杜氏改良伊格尔培养基(DMEM)加10% (vol/vol) 胎牛血清(Gibco) 及1%青霉素-链霉素(Gibco) 的96孔培养皿中。当细胞达到50-60%汇合时的第二天,按照制造商的说明,使用每孔0.3 μ L的Lipofectamine 3000(Thermo Scientific)共转染100ng的RGN表达质粒加100ng的单gRNA表达质粒。生长48小时后,使用Cells-to-Ct One Step试剂盒(ThermoFisher)来收获总RNA。

[0963] 目标基因表达的Taqman测定

[0964] 挑选正常情况下在HEK细胞中具有低表达但在CRISPR活化时可被诱导的内源基因。出于此目的,挑选RHOF2及CD2。TaqMan基因表达测定是作为标准化对照通过对RHOF2及CD2使用FAM示踪探针且对ACTB使用VIC示踪探针(全部探针都来自ThermoFisher)进行的。TaqMan测定是按照制造商的说明在BioRad CFX96 Real Time热循环仪中在Cells-to-CTTMOne Step试剂盒(Thermofisher)中进行的。在不存在gRNA的类似实验中测量背景。使用 $2^{-\Delta\Delta C_t}$ 方法(Livak等人2001,Methods,25(4):402-8)计算基因表达相对于背景的倍数变化,将表达针对ACTB转录物水平进行标准化。

[0965] 表6:目标基因表达的指导RNA

[0966]

核酸酶	基因	指导 (SEQ ID NO)
APG05405-VPR	RHOXF2	167
APG05405-VPR	RHOXF2	168
APG05405-VPR	RHOXF2	169
APG05405-VPR	RHOXF2	170
APG05405-VPR	RHOXF2	171
APG05405-VPR	RHOXF2	172
dAPG05405-VPR	RHOXF2	167
dAPG05405-VPR	RHOXF2	168
dAPG05405-VPR	RHOXF2	169
dAPG05405-VPR	RHOXF2	170
dAPG05405-VPR	RHOXF2	171
dAPG05405-VPR	RHOXF2	172
VPR-APG05405	RHOXF2	167
VPR-APG05405	RHOXF2	168
VPR-APG05405	RHOXF2	169
VPR-APG05405	RHOXF2	170
VPR-APG05405	RHOXF2	171
VPR-APG05405	RHOXF2	172
VPR-dAPG05405	RHOXF2	167
VPR-dAPG05405	RHOXF2	168
VPR-dAPG05405	RHOXF2	169
VPR-dAPG05405	RHOXF2	170
VPR-dAPG05405	RHOXF2	171
VPR-dAPG05405	RHOXF2	172
APG05405-VPR	CD2	161
APG05405-VPR	CD2	162
APG05405-VPR	CD2	163
APG05405-VPR	CD2	164
APG05405-VPR	CD2	165
APG05405-VPR	CD2	166
dAPG05405-VPR	CD2	161
dAPG05405-VPR	CD2	162
dAPG05405-VPR	CD2	163
dAPG05405-VPR	CD2	164
dAPG05405-VPR	CD2	165
dAPG05405-VPR	CD2	166
VPR-APG05405	CD2	161
VPR-APG05405	CD2	162
VPR-APG05405	CD2	163
VPR-APG05405	CD2	164
VPR-APG05405	CD2	165
VPR-APG05405	CD2	166

	核酸酶	基因	指导 (SEQ ID NO)
[0967]	VPR-dAPG05405	CD2	161
	VPR-dAPG05405	CD2	162
	VPR-dAPG05405	CD2	163
	VPR-dAPG05405	CD2	164
	VPR-dAPG05405	CD2	165
	VPR-dAPG05405	CD2	166

[0968] 实施例6:可程序化DNA结合及碱基编辑

[0969] 寡核苷酸及PCR

[0970] 下面描述的全部PCR反应都是在包括0.5 μ M的每一种引物的20 μ L反应物中使用10 μ L的2X Master Mix Phusion High-Fidelity DNA聚合酶(Thermo Scientific)进行的。首先,使用PCR#1引物,使用以下程序:98 $^{\circ}$ C,1分钟;[98 $^{\circ}$ C,10秒;62 $^{\circ}$ C,15秒;72 $^{\circ}$ C,5分钟]的30个循环;72 $^{\circ}$ C,5分钟;12 $^{\circ}$ C,永远,来扩增涵盖每一个目标基因的大基因组区域。然后,使用对每一个指导特异的引物(PCR#2引物),使用以下程序:98 $^{\circ}$ C,1分钟;[98 $^{\circ}$ C,10秒;67 $^{\circ}$ C,15秒;72 $^{\circ}$ C,30秒]的35个循环;72 $^{\circ}$ C,5分钟;12 $^{\circ}$ C,永远,来进一步扩增1微升的此PCR反应物。PCR#2的引物包括用于Illumina测序的Nextera Read 1和Read 2转座酶衔接蛋白突出部序列。

[0971] RGN碱基编辑及gRNA哺乳动物表达质粒的构建

[0972] 哺乳动物表达的核酸酶构建体被合成。在巨细胞病毒(CMV)启动子(SEQ ID NO:152)的控制下具N端SV40(SEQ ID NO:149)及C端核质素NLS序列(分别为SEQ ID NO:149及150)、N端3xFLAG标签(SEQ ID NO:151)和N端脱氨酶(例如,hAPOBEC3A;SEQ ID NO:177)的人密码子优化的APG05405被产生且被引入哺乳动物表达载体内。使用催化去活性形式的APG05405(“dAPG5405”)。编码单gRNA的指导RNA表达构建体的每一个也都在人RNA聚合酶III U6启动子(SEQ ID NO:153)的控制下被产生。

[0973] 哺乳动物细胞的转染及表达

[0974] 转染前一天,将1 $\times 10^5$ 个HEK293T细胞(Sigma)铺板于杜氏改良伊格尔培养基(DMEM)加10% (vol/vol)胎牛血清(Gibco)及1%青霉素-链霉素(Gibco)中的24孔培养皿中。当细胞达到50-60%汇合时,按照制造商的说明,使用每孔1.5 μ L的Lipofectamine 3000(Thermo Scientific)共转染500ng的APG05405表达质粒加500ng的单gRNA表达质粒。生长48小时后,按照制造商的说明,使用基因组DNA分离试剂盒(Machery-Nagel)来收获总基因组DNA。

[0975] 下一代测序

[0976] 来自含有Illumina突出端序列的PCR#2的产物遵循Illumina 16S宏基因组测序文库操作流程进行文库制备。深度测序由服务提供商(MOGene)在Illumina Mi-Seq平台上进行。通常情况下,每扩增子产生200,000个250bp的配对末端读数(2 $\times$ 100,000个读数)。使用CRISPResso(Pinello等人,2016,Nature Biotech 34:695-697)分析读数以计算碱基编辑率。手动管理输出比对,以确认碱基编辑窗口以及鉴定插入或缺失。横跨该目标的每一个位置都被分析,以确定编辑率及在每一个位置发生的特异性核苷酸变化。

[0977] 实施例7:反式ssDNA切割

[0978] 7.1确定反式DNA切割的测定条件

[0979] 在Cutsmart缓冲液(New England Biolabs B7204S)中,将经纯化的APG05405与单指导RNA (sgRNA) 以或者50nM的核酸酶及100nM的sgRNA或者200nM的核酸酶及400nM的sgRNA的最终浓度一起孵育10分钟。将这些RNP溶液添加至在1.5X Cutsmart缓冲液(New England Biolabs B7204S)中ssDNA(目标或错配阴性对照ssDNA)的最终浓度为10nM且报导子ssDNA探针的最终浓度为250nM的溶液。报导子探针(如SEQ ID NO:197及198分别所示的TB0125及TB0089)在5'末端含有荧光染料(56-FAM用于TB0125且Cy5用于TB0089),在3'末端含有淬灭剂(3IABkFQ用于TB0125且3IABRQSp用于TB0089),并且视情况含有内部淬灭剂(内部淬灭剂ZEN仅存在于TB0125上)。报导子探针的切割导致荧光染料去淬灭,且由此导致荧光信号增大。为监测荧光强度,在酶标仪(CLARIOstar Plus)中在30℃下在Corning的小体积384孔微孔板中孵育10μl的每一个反应物。

[0980] 为对此测定确定适合的参数,探测很多条件。为确定是否有淬灭的探针设计或荧光团特性的效应,在每一个反应物中作为混合物包含两个这种报导子。在任何给定的反应物中,它们彼此在相同的浓度。在所有情况中,对照或目标ssDNA的浓度(SEQ ID NO:199及200分别所示的LE201或LE205)为10nM。RNP名称示意了如下面的表7所指示的核酸酶及目标。

[0981] 表7.核糖核蛋白复合物

[0982]	核酸酶	sgRNA (SEQ ID NO)	预期目标	RNP名称
	APG05405	Gsg.1 (193)	LE201	APG05405.1
	APG05405	Gsg.2 (194)	LE205	APG05405.2

[0983] 结果显示于下面的表8中。

[0984] 表8.反式DNA切割测定结果

[0985]

				斜率（任意单位/分钟）	
RNP	[RNP] (nM)	[报导子] (nM)	目标	Cy5通道	FAM通道
APG05405.1	25 nM	0	LE201	-0.60	-0.60
APG05405.1	25 nM	50	LE201	72.60	4.20
APG05405.1	25 nM	250	LE201	198.60	-1.20
APG05405.1	25 nM	500	LE201	294.00	-3.00
APG05405.1	25 nM	750	LE201	428.40	-16.80
APG05405.1	25 nM	1000	LE201	509.40	-19.80
APG05405.1	25 nM	0	LE205	-0.48	-0.42
APG05405.1	25 nM	50	LE205	6.60	-1.80
APG05405.1	25 nM	250	LE205	57.00	0.60
APG05405.1	25 nM	500	LE205	185.40	-6.60
APG05405.1	25 nM	750	LE205	318.60	-14.40
APG05405.1	25 nM	1000	LE205	483.60	3.60
APG05405.2	25 nM	0	LE201	0.00	-2.40
APG05405.2	25 nM	50	LE201	12.00	-3.00
APG05405.2	25 nM	250	LE201	81.00	3.60
APG05405.2	25 nM	500	LE201	232.20	7.80
APG05405.2	25 nM	750	LE201	423.60	7.80
APG05405.2	25 nM	1000	LE201	684.60	9.60
APG05405.2	25 nM	0	LE205	-0.60	-2.40
APG05405.2	25 nM	50	LE205	173.40	44.40
APG05405.2	25 nM	250	LE205	262.80	49.20
APG05405.2	25 nM	500	LE205	381.00	44.40
APG05405.2	25 nM	750	LE205	530.40	31.80
APG05405.2	25 nM	1000	LE205	678.00	21.00
APG05405.1	100 nM	0	LE201	0.36	-3.00

[0986]

				斜率（任意单位/分钟）	
RNP	[RNP] (nM)	[报导子] (nM)	目标	Cy5通道	FAM通道
APG05405.1	100 nM	50	LE201	97.20	5.40
APG05405.1	100 nM	250	LE201	429.00	19.20
APG05405.1	100 nM	500	LE201	651.00	26.40
APG05405.1	100 nM	750	LE201	1176.00	29.40
APG05405.1	100 nM	1000	LE201	1601.40	29.40
APG05405.1	100 nM	0	LE205	-1.80	-0.36
APG05405.1	100 nM	50	LE205	24.00	0.24
APG05405.1	100 nM	250	LE205	142.20	3.60
APG05405.1	100 nM	500	LE205	331.80	5.40
APG05405.1	100 nM	750	LE205	590.40	10.80
APG05405.1	100 nM	1000	LE205	952.80	9.00
APG05405.2	100 nM	0	LE201	-0.36	-3.00
APG05405.2	100 nM	50	LE201	30.60	-4.20
APG05405.2	100 nM	250	LE201	213.60	4.80
APG05405.2	100 nM	500	LE201	499.20	6.00
APG05405.2	100 nM	750	LE201	1128.00	-1.80
APG05405.2	100 nM	1000	LE201	1682.40	7.20
APG05405.2	100 nM	0	LE205	0.06	-1.80
APG05405.2	100 nM	50	LE205	389.40	72.60
APG05405.2	100 nM	250	LE205	1123.80	197.40
APG05405.2	100 nM	500	LE205	1510.20	235.20
APG05405.2	100 nM	750	LE205	2048.40	256.20
APG05405.2	100 nM	1000	LE205	2393.40	234.00

[0987] 由此实验得出的结论是，相较于25nM的RNP浓度，100nM浓度的RNP一般导致较高的报导子探针切割率。一般而言，最多到250nM的报导子浓度，报导子寡核苷酸浓度越高，则报导子切割率越高，几乎没有从报导子浓度的进一步升高观察到什么益处。显然，对于TB0089报导子（在Cy5通道中检测的），尤其是在报导子浓度高于250nM时，存在干扰目标分化的显著较高位准的背景活性。因此，得出结论，报导子浓度高于250nM没有益处，且一般而言，双淬灭TB0125探针（在FAM通道检测的）更适合未来的实验，因为在宽报导子浓度范围内，其对背景活性提供较高比率的特异性。

[0988] 7.2 APG09624反式DNA切割及纯化对非特异性活性的效应

[0989] 经纯化的APG05405及APG09624与单指导RNA (sgRNA) 如以下所示一起在37℃下在1X Cutsmart缓冲液 (New England Biolabs B7204S) 中以200nM的核酸酶及400nM的sgRNA的最终浓度孵育10分钟。

[0990] 表9. 核糖核蛋白复合物

[0991]

核酸酶	核酸酶批次	sgRNA	预期目标	RNP 名称
APG09624	N/A	Gsg.2	LE113	APG09624.2
APG05405	A	Gsg.2	LE113	APG05405.2A
APG05405	B	Gsg.2	LE113	APG05405.2B

[0992] 然后，将这些RNP溶液添加至在1.5X Cutsmart缓冲液 (New England Biolabs

B7204S) 中 ssDNA (目标或错配阴性对照 ssDNA) 的最终浓度为 10nM 且报导子探针 (TP0003; 如 SEQ ID NO:314 所示) 的最终浓度为 250nM 的溶液中。报导子探针在 5' 末端含有荧光染料, 且在 3' 末端含有淬灭剂。报导子探针的切割导致荧光染料去淬灭, 且由此导致荧光信号增大。为监测荧光强度, 在酶标仪 (CLARIOstar Plus) 中在 37℃ 下在 Corning 的小体积 384 孔微孔板中孵育 10μl 的每一个反应物。

[0993] 相对于阴性对照, 与目标序列一起孵育导致以时间为函数的荧光强度实质升高。切割率被适当地概括为荧光对时间函数的线性部分的斜率, 如表 10 所显示。

[0994] 表 10. 反式 DNA 切割测定结果

RNP	ssDNA 目标序列	斜率 (任意单位/分钟)
APG09624.2	LE111	182
APG09624.2	LE113	561
APG05405.2A	LE111	197
APG05405.2A	LE113	1625
APG05405.2B	LE111	73
APG05405.2B	LE113	1054

[0996] 这些数据显示通过 APG05405 及 APG09624 二者形成的 RNP 引起目标序列的分化。

[0997] 7.3 通过 PCR 产物的反式 DNA 切割活化

[0998] 含有目标的简并核苷酸 5' 的寡核苷酸被 PCR 扩增, 以产生目标序列。目标 dsPCR2 及 dsPCR3 在指导 RNA 编码的目标的 5' 侧含有分别如 SEQ ID NO:311 及 312 所示的目标序列 ACTACAACAGCCACAACGTCTATATCATGG (dsPCR2) 及 TGG AATGGGA ACTAAAGTAATGG (dsPCR3), 其分别含有 8bp 及 5bp 的简并区域。用恰当的引物, 寡核苷酸对发生退火且被 PCR 扩增。除此之外, ssDNA 目标被包括于该实验中 - 寡核苷酸含有上面描述的目标序列的反向互补序列 CCATGATATAGACGTTGTGGCTGTTGTAGT (LE205; SEQ ID NO:200) 及 CCATTACTTTAGTTCCCATTC CA (LE501; SEQ ID NO:174)。

[0999] RNP 溶液是通过孵育在 1X NEBuffer 2 (New England Biolabs) 中分别为 0.5μM 及 1μM 的核酸酶及 sgRNA 形成的, 且在室温下被孵育 20 分钟。

[1000] 表 11. 核糖核蛋白复合物

sgRNA	预期目标	RNP 名称
Gsg.3 (SEQ ID NO:175)	dsPCR3, LE501	APG05405.3
Gsg.2 (SEQ ID NO:194)	dsPCR2, LE205	APG05405.2

[1002] 在 1.5X NEBuffer 2 中与具有 5' TEX615 标记及 3' Iowa Black FQ 淬灭剂的 1.5μM 报导子及 100nM 的分别的 PCR 产物或 ssDNA 寡核苷酸 LE501 或 LE205 (分别如 SEQ ID NO:174 及 200 所示; LE501 包含 5' FAM 荧光团) 进行切割反应。报导子探针的切割导致荧光染料去淬灭, 且由此导致荧光信号升高。为监测荧光强度, 在酶标仪 (CLARIOstar Plus) 中在 37℃ 下在 Corning 的小体积 384 孔微孔板中孵育 10μl 的每一个反应物。动力学分析的结果显示于表 12 中。

[1003] 表 12. 反式 DNA 切割测定的结果

[1004]	目标	RNP	斜率 (任意单位/分钟)
	dsPCR3	APG05405.3	780
		APG05405.2	102
	dsPCR2	APG05405.3	78
		APG05405.2	768
	LE501	APG05405.3	624
		APG05405.2	120
	LE205	APG05405.3	90
		APG05405.2	1434

[1005] 这些结果证实ssDNA的非序列特异性切割的目标序列特异性活化。dsDNA PCR产物及目标ssDNA寡核苷酸二者均能够诱导此活性。

[1006] 7.4通过经诱导的非特异性ssDNA切割的PAM确定

[1007] 如果给定系统需要用于DNA结合、修饰或切割的PAM,则含有DNA目标(包括PAM文库)的三元复合物与RNP的分离及由其回收的DNA的测序可用于鉴定PAM序列。该复合物可通过很多方法捕获,诸如,免疫下拉(immuno-pulldown)、固定化金属亲和树脂(诸如,Ni-NTA琼脂糖)捕获、或通过尺寸排阻层析法的分离。

[1008] 作为另一种选择,具与固定目标相邻的不同的PAM序列的DNA片段的平行文库被产生。含有推定核酸酶及适合指导RNA(靶向固定目标的单指导或双RNA指导)的RNP与文库中的每一个片段一起孵育,且使用DNA片段的电泳迁移率位移、尺寸排阻液体层析法、或使用对任一组成亲和的固体支持物的共沉淀而被评估。

[1009] 7.5使用平行质粒DNA文库的PAM确定

[1010] 产生含有前面为8bp简并序列(如SEQ ID NO:176所示的NNNNNNNN)的目标序列(ACTACAACAGCCACAACGTCTATATCATGG(如SEQ ID NO:313所示的))的质粒文库。在转化为感受态细胞且铺板于具选择性培养基的琼脂板上时,自此反应物的转化得到的每一个菌落都原则上与克隆质粒DNA序列对应,且由此来自来源于单菌落的培养物的质粒DNA的制剂为自原始文库取样的唯一质粒制剂。自取样96个菌落获得质粒制剂。对这些制剂个别地进行Sanger测序,以验证它们的PAM序列。

[1011] 在室温下,在1X Cutsmart缓冲液(New England Biolabs B7204S)中,将经纯化的APG05405与单指导RNA(sgRNA)Gsg.2(如SEQ ID NO:194所示的)以200nM的核酸酶及400nM的sgRNA的最终浓度一起孵育20分钟。

[1012] 将这些RNP溶液以100nM的最终浓度添加至在1.5X Cutsmart缓冲液(New England Biolabs B7204S)中质粒DNA及包含荧光团及淬灭剂的ssDNA报导子链为50nM的溶液中。为监测以时间为函数的荧光强度,在酶标仪(CLARIOstar Plus)中在37℃下在Corning的小体积384孔微孔板中孵育10μl的每一个反应物。每一个孔对应于具特异性PAM序列的个别消化反应物。在数据收集完成时,将确定荧光的增加率。通过分析荧光增加率高的孔对应的序列,建立共有PAM序列。如果尚无定论,则可产生且评估额外文库。

[1013] 实施例8:使用ssDNA切割作为诊断

[1014] 由于这些核酸酶在存在目标DNA序列的情况下产生光学可检测信号的能力,它们有希望将功用实施于检测基因疾病的诊断装置或诸如细菌、病毒或真菌的传染性疾病的药

剂内。

[1015] 诊断程序可包括对待测试的样本中的核酸的分离或扩增。在不对核酸进行任何分离或纯化的情况下,使用一些样本也可以是适合的,因为它们可以相当高的量存在于样本中,在没有扩增(诸如PCR)或在没有干扰检测或信号产生的材料的情况下,足以被检测。

[1016] 然后,如其他实施例中所描述而形成的RNP和报导子(诸如,前面的实施例中使用的荧光团及淬灭剂修饰的ssDNA寡核苷酸,或当被切割时产生可见的或以其他方式容易检测的信号的其他种类的ssDNA基质)一起被暴露于样本(或前面的段落中描述的经处理的样本)中。如果使用荧光团-淬灭剂缀合的DNA寡核苷酸(与前面描述的实施例中相同),则使用在前面的实施例中描述的荧光计可将其检测。为简化检测,可进行终点测定,代替前面描述的动力学测定,意味着该测定可相对于阳性及阴性对照进行固定时间,且在此经过时间结束时读出。

[1017] 这些试剂还可被整合于横向流动测试装置内,这样允许利用非常小的仪器对给定致病剂或特异性核酸序列(诸如,个体中的患病等位基因)进行检测。在此测定中,ssDNA报导子会与多种适合抗体或亲和试剂捕获的诸如荧光素、生物素和/或地高辛的分子缀合。

[1018] 实施例8.1-COVID19诊断测定

[1019] 根据标准做法,在通用运输培养基(UTM)中使用鼻咽拭子,自患者收集样本,且提取RNA。类似于Broughton等人2020(Nat.Biotechnol.38:870-874)使用逆转录-环介导等温扩增(RT-LAMP)来扩增基因材料。在一些实施方案中,通过RT-LAMP产生的单链DNA(ssDNA)足可用于通过本文公开的RGN进行PAM独立的检测。在一些实施方案中,使用通过仅位于目标链上的硫代磷酸脂引物的扩增,RT-LAMP产生ssDNA,这允许非目标链的T7核酸外切酶消化。

[1020] 类似于Broughton等人2020,作为对样本收集及制备的质量控制检查,用恰当的引物进行RT-LAMP扩增,以扩增SARS-CoV2基因组的N基因及E基因以及人RNase P。两个LAMP内部引物(常称为FIP或BIP)中一个会含有硫代磷酸脂基团。在用T7核酸外切酶处理所完成的PCR反应物时,自硫代磷酸脂引物延伸的ssDNA为存在于溶液中的主要物质。关于特异性及有效活化,使用上面描述的荧光测定来评估针对此系列的指导。关于特异性,为确保没有交叉反应性,可针对其他冠状病毒中的同源基因测试该检测方案,诸如,HCoV-OC43、HCoV-HKU1、HCoV-229E、HCoV-NL63、MERS-CoV和/或SARS-CoV。通过利用含有FAM及生物素的寡核苷酸,可将该测定转换为横向流动测定。

序列表

<110> 生命编辑制药股份有限公司

<120> RNA指导核酸酶及其活性片段与变体以及使用方法

<130> L103438 1180TW (000076.0)

<140>

<141>

<150> 63/058,169

<151> 2020-07-29

<150> 62/955,014

<151> 2019-12-30

<160> 314

<170> PatentIn版本3.5

<210> 1

<211> 506

<212> PRT

<213> 芽孢杆菌属

<400> 1

```

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

```

				165						170					175				
Leu	Ser	His	Phe	Asn	Asn	Asp	Met	Lys	Thr	Gly	Val	Leu	Arg	Gly	Asp				
				180						185					190				
Arg	Ser	Leu	Arg	Thr	Tyr	Lys	Asn	Ser	His	Pro	Leu	Leu	Val	Arg	Ala				
				195						200					205				
Arg	Ser	Met	Lys	Val	Tyr	Glu	Glu	Asn	Gly	Asp	Tyr	Phe	Ile	Lys	Trp				
				210						215					220				
Val	Lys	Gly	Ile	Val	Phe	Lys	Ile	Val	Ile	Ser	Ala	Gly	Ser	Lys	Gln				
225										230					235				240
Lys	Ala	Asn	Ile	Gly	Glu	Leu	Lys	Ser	Val	Leu	Ile	Asn	Ile	Leu	Asn				
				245						250					255				
Gly	His	Tyr	Lys	Val	Cys	Asp	Ser	Ser	Ile	Ser	Leu	Asn	Lys	Asp	Leu				
				260						265					270				
Ile	Leu	Asn	Leu	Ser	Leu	Asn	Ile	Pro	Val	Ser	Lys	Glu	Asn	Val	Phe				
				275						280					285				
Val	Pro	Gly	Arg	Val	Val	Gly	Val	Asp	Leu	Gly	Leu	Lys	Ile	Pro	Ala				
				290						295					300				
Tyr	Val	Ser	Leu	Asn	Asp	Thr	Pro	Tyr	Ile	Lys	Lys	Gly	Ile	Gly	Asn				
305										310					315				320
Ile	Asp	Asp	Phe	Leu	Arg	Val	Arg	Thr	Gln	Leu	Gln	Ser	Gln	Arg	Lys				
				325						330					335				
Arg	Leu	Gln	Lys	Thr	Leu	Glu	Cys	Thr	Ser	Gly	Gly	Lys	Gly	Arg	Ser				
				340						345					350				
Lys	Lys	Leu	Lys	Gly	Leu	Asp	Arg	Leu	Lys	Ala	Lys	Glu	Lys	Asn	Phe				
				355						360					365				
Val	Asn	Thr	Tyr	Asn	His	Phe	Leu	Ser	Lys	Lys	Ile	Ile	Gln	Phe	Ala				
				370						375					380				
Val	Lys	Asn	Asn	Ala	Gly	Val	Ile	His	Leu	Glu	Glu	Leu	Gln	Phe	Asp				
385										390					395				400
Lys	Leu	Lys	His	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu				
				405						410					415				
Gln	Thr	Met	Ile	Glu	Tyr	Lys	Ala	Glu	Arg	Glu	Gly	Ile	Glu	Val	Lys				
				420						425					430				
Tyr	Val	Asp	Ala	Ser	Tyr	Thr	Ser	Gln	Thr	Cys	Ser	Lys	Cys	Gly	His				
				435						440					445				
Tyr	Glu	Glu	Gly	Gln	Arg	Val	Leu	Gln	Asp	Thr	Phe	Thr	Cys	Lys	Asn				
				450						455					460				
Lys	Glu	Cys	Lys	Gly	Tyr	Val	His	Lys	Val	Asn	Ala	Asp	Phe	Asn	Ala				
465										470					475				480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
 485 490 495
 Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
 500 505
 <210> 2
 <211> 506
 <212> PRT
 <213> 芽孢杆菌属
 <400> 2
 Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
 1 5 10 15
 Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
 20 25 30
 Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
 35 40 45
 Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala
 50 55 60
 Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
 65 70 75 80
 Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
 85 90 95
 Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
 100 105 110
 Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
 115 120 125
 Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
 130 135 140
 Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
 145 150 155 160
 Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
 165 170 175
 Leu Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
 180 185 190
 Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala
 195 200 205
 Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp
 210 215 220
 Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln
 225 230 235 240

Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn		
245	250	255
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu		
260	265	270
Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe		
275	280	285
Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala		
290	295	300
Tyr Val Ser Leu Asn Asp Asn Pro Tyr Ile Lys Lys Gly Ile Gly Asn		
305	310	315
Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys		
325	330	335
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser		
340	345	350
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe		
355	360	365
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala		
370	375	380
Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp		
385	390	395
Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
405	410	415
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys		
420	425	430
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His		
435	440	445
Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn		
450	455	460
Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala		
465	470	475
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met		
485	490	495
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala		
500	505	

<210> 3

<211> 506

<212> PRT

<213> 芽孢杆菌属

<400> 3

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

305	310	315	320
Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys			
	325	330	335
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser			
	340	345	350
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe			
	355	360	365
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala			
	370	375	380
Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp			
385	390	395	400
Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu			
	405	410	415
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys			
	420	425	430
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His			
	435	440	445
Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn			
	450	455	460
Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala			
465	470	475	480
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Asn Glu Met			
	485	490	495
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala			
	500	505	
<210> 4			
<211> 506			
<212> PRT			
<213> 芽孢杆菌属			
<400> 4			
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser			
1	5	10	15
Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn			
	20	25	30
Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn			
	35	40	45
Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala			
	50	55	60
Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr			

65	70	75	80
Lys Tyr Arg Glu Val	Ala Thr Asn Lys Tyr	Gln Asp Tyr Leu	Lys Val
85	90	95	
Lys Glu Lys Val Asn Asp Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg			
100	105	110	
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr			
115	120	125	
Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys			
130	135	140	
Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser			
145	150	155	160
Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val			
165	170	175	
Ile Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp			
180	185	190	
Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala			
195	200	205	
Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp			
210	215	220	
Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln			
225	230	235	240
Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn			
245	250	255	
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu			
260	265	270	
Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe			
275	280	285	
Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala			
290	295	300	
Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn			
305	310	315	320
Ile Asp Asp Phe Leu Lys Val Arg Thr Gln Leu Gln Ser Gln Arg Lys			
325	330	335	
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Asn			
340	345	350	
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe			
355	360	365	
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala			
370	375	380	

Val	Lys	Asn	Asn	Ala	Gly	Val	Ile	His	Leu	Glu	Glu	Leu	Gln	Phe	Asp
385					390				395					400	
Lys	Leu	Lys	His	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu
			405					410						415	
Gln	Thr	Met	Ile	Glu	Tyr	Lys	Ala	Glu	Arg	Glu	Gly	Ile	Glu	Val	Lys
		420					425						430		
Tyr	Val	Asp	Ala	Ser	Tyr	Thr	Ser	Gln	Thr	Cys	Ser	Lys	Cys	Gly	His
	435						440					445			
Tyr	Glu	Glu	Gly	Gln	Arg	Val	Leu	Gln	Asp	Thr	Phe	Thr	Cys	Lys	Asn
	450					455					460				
Lys	Glu	Cys	Lys	Gly	Tyr	Val	His	Lys	Val	Asn	Ala	Asp	Phe	Asn	Ala
465				470					475					480	
Ser	Gln	Asn	Ile	Ala	Lys	Ser	Thr	Asp	Ile	Ile	Arg	Cys	Thr	Glu	Met
			485					490						495	
Ala	Lys	Asn	Asn	Asp	Ile	Glu	Lys	Asn	Ala						
		500					505								
<210> 5															
<211> 506															
<212> PRT															
<213> 芽孢杆菌属															
<400> 5															
Met	Ser	Thr	Pro	Leu	Gln	Gln	Pro	His	Gln	Lys	Ser	Lys	Lys	Thr	Ser
1			5					10						15	
Gln	Met	Ile	Thr	Thr	Arg	Lys	Phe	Lys	Leu	Ala	Ile	Val	Ser	Asp	Asn
		20					25					30			
Arg	Asn	Glu	Ala	Tyr	Ser	Phe	Ile	Arg	Asn	Glu	Ile	Arg	Asn	Gln	Asn
	35					40					45				
Lys	Ala	Leu	Asn	Ala	Ala	Tyr	Asn	His	Leu	Tyr	Phe	Glu	His	Ile	Ala
	50				55					60					
Thr	Glu	Lys	Leu	Lys	His	Ser	Asp	Glu	Glu	Tyr	Gln	Lys	His	Leu	Thr
65				70					75					80	
Lys	Tyr	Arg	Glu	Val	Ser	Thr	Asn	Lys	Tyr	Gln	Asp	Tyr	Leu	Lys	Val
			85					90					95		
Lys	Glu	Lys	Val	Asn	Ala	Ser	Lys	Asp	Asp	Glu	Lys	Leu	Gln	Lys	Arg
	100						105					110			
Val	Asp	Lys	Ala	Arg	Glu	Ala	Tyr	Asn	Lys	Ala	Gln	Glu	Lys	Val	Tyr
	115						120				125				
Lys	Ile	Glu	Lys	Glu	Phe	Asn	Lys	Lys	Ser	Met	Glu	Thr	Tyr	Gln	Lys
	130					135				140					

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

450	455	460
Lys Glu Cys Lys Gly Tyr	Ala His Lys Val	Asn Ala Asp Phe Asn Ala
465	470	475
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile	Ile Arg Cys Thr Glu Met	
	485	490
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala		495
	500	505
<210> 6		
<211> 450		
<212> PRT		
<213> 芽孢杆菌属		
<400> 6		
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val		
1	5	10
Asn Val Glu Trp Lys Ala Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr		
	20	25
Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe		
	35	40
Asp Asn Leu Ser Leu Asn His Phe Lys Glu Thr Gly Glu Tyr Pro Ser		
	50	55
Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile		
65	70	75
Tyr Asn Gln Leu Arg Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met		
	85	90
Ser Thr Thr Ile Gln Lys Ser Leu Lys Asn Trp Asn Ser Arg Lys Lys		
	100	105
Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu		
	115	120
Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Ser		
	130	135
Gly Asp Tyr Phe Ala Asn Val Ser Leu Phe Ser Ser Lys Phe Ile Lys		
145	150	155
Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Val Lys Leu Ser Thr Arg		
	165	170
Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr		
	180	185
Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Asn Lys Trp Tyr		
	195	200
Leu Ser Ile Thr Tyr Lys Ala Thr Lys Lys Glu Gly His Lys Phe Asp		

210	215	220
Glu Glu Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr		
225	230	235
Phe Ala Tyr Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu		240
	245	250
Ile Glu Ala Phe Arg Lys Lys Ile Glu Tyr Arg Arg Ile Ser Leu Leu		255
	260	265
Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys		270
	275	280
Lys Arg Ile Lys Pro Ile Glu Val Leu Asn Asp Lys Ile Thr Lys Phe		285
	290	295
Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys		300
305	310	315
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Gln Gly Ile		320
	325	330
Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln		335
	340	345
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys		350
	355	360
Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile		365
	370	375
His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys		380
385	390	395
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile		400
	405	410
Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys		415
	420	425
Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn		430
	435	440
		445
Ile Asn		
450		
<210> 7		
<211> 506		
<212> PRT		
<213> 芽孢杆菌属		
<400> 7		
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser		
1	5	10
Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn		15

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

Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser		
340	345	350
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe		
355	360	365
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala		
370	375	380
Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp		
385	390	395
Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
405	410	415
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys		
420	425	430
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His		
435	440	445
Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn		
450	455	460
Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala		
465	470	475
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met		
485	490	495
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala		
500	505	

<210> 8

<211> 450

<212> PRT

<213> 芽孢杆菌属

<400> 8

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

Ser Thr Thr	Ile Gln Lys Thr	Leu Lys Asn Trp	Asn Ser Arg	Lys Lys
100		105		110
Glu Ile Trp	Arg Gly Glu Met	Ser Ile Pro Ser	Phe Arg Ser	Asp Leu
115		120		125
Pro Ile Asp	Ile His Gly Asn	Ser Ile Gln Leu	Ile Lys Glu	Lys Gly
130		135		140
Gly Asp Tyr	Ile Ala Ser Val	Ser Leu Phe Ser	Ser Lys Phe	Ile Lys
145		150		155
Glu Asn Asp	Leu Pro Asn Gly	Lys Ile Leu Met	Lys Leu Ser	Thr Arg
	165		170	175
Lys Gln Asn	Ser Met Lys Val	Ile Leu Asp	Arg Leu Ile	Asn Ser Thr
	180		185	190
Tyr Ser Lys	Gly Ala Cys Met	Leu His Lys	His Lys Lys	Lys Trp Tyr
	195		200	205
Leu Ser Ile	Thr Tyr Lys Ser	Asn Ile Lys	Glu Glu Leu	Lys Phe Asp
	210		215	220
Glu Asp Leu	Ile Met Gly Ile	Asp Met Gly	Lys Ile Asn	Val Leu Tyr
	225		230	235
Phe Ala Phe	Asn Lys Ser Leu	Val Arg Gly	Ala Ile Ser	Gly Glu Glu
	245		250	255
Ile Glu Ala	Phe Arg Lys Lys	Ile Glu His	Arg Arg Ile	Ser Leu Leu
	260		265	270
Arg Gln Gly	Lys Tyr Cys Ser	Gly Asn Arg	Ile Gly Lys	Gly Arg Glu
	275		280	285
Lys Arg Ile	Lys Pro Ile Asp	Val Leu Asn	Asp Lys Val	Ala Lys Phe
	290		295	300
Arg Asn Ala	Thr Asn His Lys	Tyr Ala Asn	Tyr Ile Val	Gln Gln Cys
	305		310	315
Leu Lys Tyr	Asn Cys Gly Thr	Ile Gln Leu	Glu Asp Leu	Lys Gly Ile
	325		330	335
Ser Lys Glu	Gln Thr Phe Leu	Lys Asn Trp	Thr Tyr Phe	Asp Leu Gln
	340		345	350
Glu Lys Ile	Lys Asn Gln Ala	Asn Gln Tyr	Gly Val Lys	Val Val Lys
	355		360	365
Ile Asp Pro	Ser Tyr Thr Ser	Gln Arg Cys	Ser Glu Cys	Gly Tyr Ile
	370		375	380
His Lys Asp	Asn Arg Gln Asp	Gln Ser Thr	Phe Glu Cys	Leu Gln Cys
	385		390	395
Ser Phe Lys	Ala His Ala Asp	Tyr Asn Ala	Ala Lys Asn	Ile Ser Val

	405		410		415
Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys					
	420		425		430
Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn					
	435		440		445
Ile Asn					
450					
<210> 9					
<211> 486					
<212> PRT					
<213> 芽孢杆菌属					
<400> 9					
Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ala					
1	5		10		15
Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys					
	20		25		30
Ala Leu Asn Val Ala Leu Asn His Leu Tyr Phe Asn Ser Ile Ala Arg					
	35		40		45
Gln Lys Ile Leu Leu Ala Asp Glu Ala Tyr Gln Gln Lys Leu Lys Asp					
	50		55		60
Ala Ile Asn Ser Gln Glu Lys Ser Phe Asn Thr Leu Lys Glu Leu Glu					
65	70		75		80
Arg Lys Gln Cys Ile Glu Glu Asp Phe Glu Lys Lys Gln Thr Leu Lys					
	85		90		95
Glu Arg Leu Ile Lys Ala Lys Asn Ala Tyr Glu Lys Ala Lys Glu Lys					
	100		105		110
Val Ser His Leu Arg Lys Ser Arg Ser Lys Asp Ser Phe Gln Glu Tyr					
	115		120		125
Lys Asn Ile Ile Gly Gln Val Glu Gln Thr His Leu Arg Asp Ile Ile					
	130		135		140
Ser Ser Gln Phe Asn Leu His Ser Asp Thr Lys Asp Arg Leu Thr Met					
145	150		155		160
Ile Ala Asn Gln Asp Phe Lys Asn Asp Ile Ala Glu Val Leu Ser Gly					
	165		170		175
Asn Cys Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg					
	180		185		190
Gly Arg Asn Thr Val Leu Tyr Lys Glu Gly Asn Glu Phe Phe Ile Lys					
	195		200		205
Trp Ile Lys Gly Ile Val Phe Lys Cys Ile Leu Gly Val Lys Asn Gln					

210	215	220
Asn Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Val Leu Ala Gly Asn		
225	230	235
Tyr Lys Ile Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu		240
245	250	255
Ile Met Asn Leu Thr Leu Asn Met Pro Asp Lys Asp Glu Asn Arg Lys		
260	265	270
Val Pro Gly Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala		
275	280	285
Tyr Phe Ala Val Asn Asp Ala Pro Tyr Ile Arg Lys Ala Leu Gly Lys		
290	295	300
Ile Glu Asp Phe Leu Lys Val Arg Thr Ser Ile Gln Ser Gln Lys Arg		
305	310	315
Ser Leu Glu Arg Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys		
325	330	335
Lys Lys Leu Arg Ala Leu Asp Gln Phe Lys Gly Lys Glu Lys Arg Tyr		
340	345	350
Val Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala		
355	360	365
Ile Gln Tyr Gly Val Glu Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys		
370	375	380
Glu Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
385	390	395
Gln Gln Phe Ile Glu Tyr Lys Ala Lys Arg Glu Gly Ile Leu Ile Lys		
405	410	415
Tyr Val Asp Pro Phe Asn Thr Ser Gln Thr Cys Ser Lys Cys Asp His		
420	425	430
Tyr Glu Gly Gly Gln Arg Glu Lys Gln Ala Asn Phe Leu Cys Lys Ser		
435	440	445
Cys Gly Phe Glu Glu Asn Ala Asp Phe Asn Ala Ala Arg Asn Ile Ala		
450	455	460
Lys Ser Lys Asn Tyr Ile Thr Arg Lys Glu Glu Ser Glu Tyr Tyr Lys		
465	470	475
Arg Asn Asn Glu Ile Ala		480
485		

<210> 10

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 10

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

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Phe Ile Val Lys Gln			
305	310	315	320
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Ile Glu His Leu Lys Gly			
	325	330	335
Ile Ser Lys Asp Asp Lys Val Leu Lys Asp Trp Thr Tyr Tyr Asp Leu			
	340	345	350
Gln Glu Lys Ile Lys Lys Gln Ala Gln Ala Tyr Gly Ile Glu Val Ile			
	355	360	365
Thr Ile Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
	370	375	380
Ile Ser Asn Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln			
385	390	395	400
Cys Glu Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
	405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
	420	425	430
Arg Leu His Ser Lys Lys Tyr Met Glu Asp His Ile Glu Glu Leu Gly			
	435	440	445
Tyr Ser Gly			
450			
<210> 11			
<211> 599			
<212> PRT			
<213> 戈登氏菌属			
<400> 11			
Met Ala Ile Thr Val His Thr Met Gly Ala His Tyr Arg Trp Glu Ile			
1	5	10	15
Pro Glu Gln Leu Arg Ala Gln Leu Trp Leu Ala His Asn Leu Arg Glu			
	20	25	30
Asp Leu Val Thr Leu Gln His Glu Tyr Glu Ala Arg Thr Lys Glu Val			
	35	40	45
Trp Ser Ser Phe Pro Asp Val Ala Ala Thr Glu Asp Arg Leu Thr Ala			
	50	55	60
Ala Glu Leu Glu Ala Glu Ala Leu Ala Glu Lys Val Lys Ser Glu Arg			
65	70	75	80
Ile Arg Gln Arg Thr Lys Arg Val Thr Gly Pro Val Ala Asp Gln Leu			
	85	90	95
Ala Ala Ala Arg Lys Thr Ala Lys Glu Ala Arg Ile Ala Arg Arg Ala			
	100	105	110

Ala Ile Ala Ala Val Arg Asp Thr Ala Lys Ala Gln Leu Asn Glu Leu		
115	120	125
Ser Asp Glu Leu Lys Ala Ala His Lys Arg Leu Tyr Ala Glu Tyr Cys		
130	135	140
Gln Ser Gly His Leu Tyr Trp Ala Thr Phe Asn Ala Thr Leu Asp His		
145	150	155
His Lys Thr Ala Val Lys Arg Met Gln Gln Leu Arg Ala Ala Gly Arg		
165	170	175
Pro Ala Gln Met Arg His His Arg Phe Asp Gly Thr Gly Thr Ile Ala		
180	185	190
Val Gln Leu Gln Arg Gln Ser Gly Gln Pro Gln Arg Thr Pro Ala Leu		
195	200	205
Ile Gly Asp Pro Glu Gly Lys Tyr Arg Asn Val Phe Tyr Ser Pro Trp		
210	215	220
Val Asp Pro Asp Gln Trp Asp Thr Met Thr Arg Ser Glu Gln Arg Lys		
225	230	235
Ala Gly Arg Val Thr Val Thr Met Arg Cys Gly Ser Thr Asp Asp Gly		
245	250	255
Pro Ala Trp Ile Ala Val Pro Val Gln Gln His Arg Met Leu Pro Ala		
260	265	270
Asp Ala Asp Ile Thr Gly Ala Gln Leu Thr Val Thr Arg Lys Gly Ser		
275	280	285
Ala Thr His Val Lys Ile Ala Val Thr Ala Arg Val Gly Glu Pro Glu		
290	295	300
Pro Ile Thr Asp Gly Pro Thr Val Ala Val His Leu Gly Trp Arg Asp		
305	310	315
Thr Asp Thr Gly Thr Val Val Ala His Trp Arg Ala Ser Glu Pro Leu		
325	330	335
Asp Val Pro Phe Asp Met Arg Asp Ile Met Pro Thr Asp Pro Gly Gly		
340	345	350
Arg Thr Gly Thr Val Val Ile Pro Thr Arg Ile Val Asp Arg Ile Glu		
355	360	365
Ser Ala Ala Gly Ile Ala Ser Gln Arg Gly Asp Ala Gln Asn Glu Met		
370	375	380
Lys Arg Gln Leu Val Glu Trp Leu Thr Glu Val Gly Pro Gln Pro His		
385	390	395
Pro Thr Arg Asp Gly Glu Glu Ile Ser Ala Ala Asp Ala Ala Arg Trp		
405	410	415
Arg Asn Pro Gly Arg Phe Ala Ala Leu Ala Met Ala Trp Arg Asp Asn		

420	425	430
Pro Pro Asn Arg Gly Glu Gln Ile	Ala Tyr Val Leu Glu Ser Trp Arg	
435	440	445
Ala Thr Asp Lys Gly Leu Trp Asn Arg Gln Glu Gly Gly Arg Arg Lys		
450	455	460
Ala Leu Gly His Arg Thr Asp Ile Tyr Arg Gln Val Ala Ala Leu Ile		
465	470	475
Ala Asp Gln Ala Gly Val Val Val Val Asp Asp Thr Ser Ile Ala Asp		
485	490	495
Ile Ala Ala Arg Pro Thr Glu Leu Pro Asn Glu Val Glu Ala Arg Ile		
500	505	510
Ala Arg Arg Arg Gly Ser Ala Ala Pro Gly Glu Leu Arg Glu Ala Val		
515	520	525
Arg Ser Ala Ala Val Arg Asp Gly Val Gly Val Glu Val Val Ala His		
530	535	540
Lys Gly Leu Ser Arg Thr His Ala Ala Cys Gly His Glu Asn Pro Ala		
545	550	555
Asp Asp Arg Tyr Met Thr Val Gly Val Val Cys Asp Gly Cys Gly Arg		
565	570	575
Thr Tyr Asp Gln Asp Leu Ser Ala Thr Leu Leu Met Leu Gln Arg Ala		
580	585	590
Thr Ser Thr Ala Ala Ala Thr		
595		
<210> 12		
<211> 553		
<212> PRT		
<213> 微球菌菌属		
<400> 12		
Met Thr Thr Thr Thr Gly Thr Asp Leu Val Pro Arg Leu Arg Ala Phe		
1	5	10
Lys His Arg Leu Asp Pro Asn Pro Ala Gln Ala Thr Leu Leu Ala Gln		
20	25	30
Tyr Ala Gly Ala Ala Arg Val Ala Tyr Asn Met Leu Thr Ala His Asn		
35	40	45
Arg Ala Ala Leu Ala Ala Ser Ala Ala Arg Arg Thr Glu Leu Ala Glu		
50	55	60
Thr Gly Leu Ala Gly Pro Glu Leu Ala Ala Arg Met Lys Ala Glu Arg		
65	70	75
Ala Ala Asp Pro Thr Leu Arg Val Ala Ser Tyr Gln Ser Tyr Ser Thr		

				85					90					95			
Thr	His	Leu	Thr	Pro	Leu	Ile	Arg	Arg	His	Arg	Glu	Ala	Ala	Ala	Ala		
				100					105					110			
Ile	Ala	Ala	Gly	Ala	Asp	Pro	Ala	Glu	Ala	Trp	Thr	Asp	Glu	Arg	Tyr		
				115					120					125			
Ala	Glu	Pro	Trp	Met	His	Thr	Val	Pro	Arg	Arg	Val	Leu	Val	Ser	Gly		
				130					135					140			
Leu	Gln	Asn	Ala	Ala	Lys	Ala	Thr	Glu	Asn	Trp	Met	Ala	Ser	Ala	Ser		
145					150					155				160			
Gly	Thr	Arg	Ala	Gly	Ala	Arg	Val	Gly	Leu	Pro	Arg	Phe	Lys	Lys	Lys		
				165					170					175			
Gly	Arg	Ser	Arg	Asp	Ser	Phe	Thr	Ile	Pro	Ala	Pro	Glu	Val	Ile	Gly		
				180					185					190			
Ala	Ala	Gly	Thr	Pro	Tyr	Lys	Arg	Gly	Glu	Pro	Arg	Arg	Gly	Val	Ile		
				195					200					205			
Thr	Asp	His	Arg	His	Leu	Arg	Leu	Ala	Ser	Leu	Gly	Thr	Ile	Arg	Thr		
				210					215					220			
Tyr	Asp	Lys	Thr	Ser	Arg	Leu	Val	Arg	Ala	Cys	Arg	Arg	Gly	Ala	Gln		
225					230					235				240			
Ile	Arg	Ser	Met	Thr	Ile	Ser	Gln	Ala	Gly	Gly	Arg	Trp	Tyr	Ala	Ser		
				245					250					255			
Ile	Leu	Val	Ala	Asp	Pro	Thr	Pro	Ile	Arg	Thr	Gly	Pro	Ser	Arg	Arg		
				260					265					270			
Gln	Arg	Ala	Asn	Asp	Ala	Val	Gly	Val	Asp	Leu	Gly	Val	Lys	His	Leu		
				275					280					285			
Ala	Ala	Leu	Ser	Thr	Gly	Glu	Val	Ile	Asp	Asn	Gly	Arg	Pro	Gly	Ala		
				290					295					300			
Arg	Gln	Ala	Ala	Arg	Leu	Thr	Arg	Leu	Gln	Arg	Ala	Tyr	Ala	Arg	Thr		
305					310					315				320			
Gln	Pro	Gly	Ser	Asn	Arg	Arg	Glu	Arg	Val	Arg	Arg	Gln	Ile	Ala	Ala		
				325					330					335			
Leu	His	His	Gly	Ile	Ala	Leu	Arg	Arg	Ala	Gly	Leu	Leu	His	Gln	Val		
				340					345					350			
Ser	Thr	Arg	Leu	Ala	Met	Asp	Phe	Ala	Val	Val	Ala	Leu	Glu	Asp	Leu		
				355					360					365			
Asn	Val	Ala	Gly	Met	Thr	Arg	Ser	Ala	Arg	Gly	Thr	Leu	Glu	Ala	Pro		
				370					375					380			
Gly	Arg	Asn	Val	Ala	Ala	Lys	Ser	Gly	Leu	Asn	Arg	Ala	Ile	Leu	Asp		
385					390					395				400			

Ala Gly Leu Gly Met Leu Arg Arg Gln Leu Asp Tyr Lys Thr Ser Trp		
405	410	415
Ala Gly Ser Gln Val Lys Met Ile Asp Arg Phe Ala Pro Ser Ser Lys		
420	425	430
Ala Cys Ser Arg Cys Gly Thr Val Lys Ser Thr Leu Ser Leu Ala Glu		
435	440	445
Arg Thr Phe Glu Cys Glu Ala Cys His Leu Val Ile Asp Arg Asp Val		
450	455	460
Asn Ala Ala Ile Asn Ile Arg Ala Trp Ala Val Gln Glu Glu Arg Gly		
465	470	475
Ala Gly Val Gly Leu Ala Arg Gly Arg Arg Glu Ser Arg Asn Gly Arg		
485	490	495
Gly Ala Ala Val Ser Gly Pro Pro Ser Gly Gly Ala Ala Gly Gln Gly		
500	505	510
Arg Gly Ser Val Lys Pro Ala Pro Gln Gly Val Gly Met Ser Ser Arg		
515	520	525
Ala Thr Gly Trp Ser Ser Gln Pro Pro Ser Thr Glu Gly Glu Ser Ala		
530	535	540
Glu Arg Gly Ala Ser Ala Leu Ala Arg		
545	550	
<210> 13		
<211> 553		
<212> PRT		
<213> 微球菌菌属		
<400> 13		
Met Thr Thr Thr Thr Gly Thr Asp Leu Ala Pro Arg Leu Arg Ala Phe		
1	5	10
Lys His Arg Leu Asp Pro Asn Pro Ala Gln Ala Thr Leu Leu Ala Gln		
20	25	30
Tyr Ala Gly Ala Ala Arg Val Ala Tyr Asn Met Leu Thr Ala His Asn		
35	40	45
Arg Ala Ala Leu Ala Ala Gly Ala Ala Arg Arg Thr Glu Leu Ala Glu		
50	55	60
Thr Gly Leu Ala Gly Pro Glu Leu Ala Ala Arg Met Lys Ala Glu Arg		
65	70	75
Ala Ala Asp Pro Thr Leu Arg Val Ala Ser Tyr Gln Ser Tyr Ser Thr		
85	90	95
Ala His Leu Thr Pro Leu Ile Arg Arg His Arg Glu Ala Ala Ala Ala		
100	105	110

Ile Ala Ala Gly Ala Asp Pro Ala Glu Ala Trp Thr Asp Glu Arg Tyr		
115	120	125
Ala Glu Pro Trp Met His Thr Val Pro Arg Arg Val Leu Val Ser Gly		
130	135	140
Leu Gln Asn Ala Ala Lys Ala Thr Glu Asn Trp Met Ala Ser Ala Ser		
145	150	155
Gly Thr Arg Ala Gly Ala Arg Val Gly Leu Pro Arg Phe Lys Lys Lys		
	165	170
Gly Arg Ser Arg Asp Ser Phe Thr Ile Pro Ala Pro Glu Val Ile Gly		
	180	185
Ala Ala Gly Thr Pro Tyr Lys Arg Gly Glu Pro Arg Arg Gly Val Ile		
	195	200
Thr Asp His Arg His Leu Arg Leu Ala Ser Leu Gly Thr Ile Arg Thr		
	210	215
Tyr Asp Lys Thr Ser Arg Leu Val Arg Ala Cys Arg Arg Gly Ala Gln		
225	230	235
Ile Arg Ser Met Thr Ile Ser Gln Ala Gly Gly Arg Trp Tyr Ala Ser		
	245	250
Ile Leu Val Ala Asp Pro Thr Pro Ile Arg Thr Gly Pro Ser Arg Arg		
	260	265
Gln Arg Ala Asn Gly Ala Val Gly Val Asp Leu Gly Val Lys His Leu		
	275	280
Ala Ala Leu Ser Thr Gly Glu Val Ile Asp Asn Gly Arg Pro Gly Ala		
	290	295
Arg Gln Ala Ala Arg Leu Thr Arg Leu Gln Arg Ala His Ala Arg Thr		
305	310	315
Gln Pro Gly Ser Asn Arg Arg Glu Arg Val Arg Arg Gln Ile Ala Ala		
	325	330
Leu Gln His Gly Ile Ala Leu Arg Arg Ala Gly Leu Leu His Gln Val		
	340	345
Ser Thr Arg Leu Ala Thr Asp Phe Ala Val Val Ala Leu Glu Asp Leu		
	355	360
Asn Val Ala Gly Met Thr Arg Ser Ala Arg Gly Thr Leu Glu Ala Pro		
	370	375
Gly Arg Asn Val Ala Ala Lys Ser Gly Leu Asn Arg Ala Ile Leu Asp		
385	390	395
Ala Gly Leu Gly Met Leu Arg Arg Gln Leu Asp Tyr Lys Thr Ser Trp		
	405	410
Ala Gly Ser Gln Val Lys Met Ile Asp Arg Phe Ala Pro Ser Ser Lys		

420	425	430
Ala Cys Ser Arg Cys Gly Thr Val Lys Ser Thr Leu Ser Leu Ala Glu		
435	440	445
Arg Thr Phe Glu Cys Glu Ala Cys His Leu Val Ile Asp Arg Asp Val		
450	455	460
Asn Ala Ala Ile Asn Ile Arg Ala Trp Ala Val Gln Glu Glu Arg Gly		
465	470	475
Ala Gly Val Glu Leu Ala Arg Gly Arg Arg Glu Ser Arg Asn Gly Arg		
485	490	495
Gly Ala Ala Val Ser Gly Pro Pro Ser Gly Gly Ala Ala Arg Gln Gly		
500	505	510
Arg Gly Ser Val Lys Pro Ala Pro Gln Gly Val Gly Met Ser Ser Arg		
515	520	525
Ala Thr Gly Trp Ser Ser Gln Pro Pro Ser Thr Glu Gly Glu Ser Ala		
530	535	540
Glu Arg Gly Ala Ser Ala Leu Ala Arg		
545	550	
<210> 14		
<211> 505		
<212> PRT		
<213> Paeniglutamicibacter菌属		
<400> 14		
Met Ser Val Gln Ala Ala Met Ala Thr Thr Val Ile His Arg Ala Tyr		
1	5	10
Arg Leu Thr Leu Asp Pro Thr Pro Gln Gln Ala Gln Lys Leu Ser Gln		
20	25	30
Trp Ala Gly Ala Ala Arg Ala Met Tyr Asn His Ala Ile Ala Ala Lys		
35	40	45
Gln Ala Ser His Arg Ser Trp Leu Gln Glu Val Ala Phe Ala Thr Tyr		
50	55	60
Glu Gln Glu Leu Thr Glu Glu Gln Ala Arg Lys Ser Ile Lys Val Pro		
65	70	75
Ile Pro Thr Ala Tyr Gly Phe Asn Ala Trp Leu Thr Glu Thr Arg Asn		
85	90	95
Thr His His Asp Ala Ala Glu Lys Gly Leu Leu Leu Pro Gly Arg Asp		
100	105	110
Gly Arg Glu His Glu Pro Trp Leu His Ala Val Asn Arg Ser Ala Leu		
115	120	125
Met Gly Ala Met Arg His Ala Asp Asp Ala Trp Thr Asn Trp Ile Asp		

130	135	140
Ser Leu Thr Gly Thr Arg Ala Gly Arg Lys Ile Gly Tyr Pro Arg Phe		
145	150	155
Lys Lys Arg Gly Val Ala Arg Asp Ser Phe Thr Ile Thr His Asp Arg		
165	170	175
Lys Ser Pro Gly Ile Arg Leu Ala Thr Thr Arg Arg Leu Arg Ile Pro		
180	185	190
Thr Phe Gly Glu Ile Arg Ile His Asp His Ala Lys Arg Leu His Arg		
195	200	205
Lys Leu His Thr Gly Thr Val Glu Val Thr Ser Val Thr Val Ser Arg		
210	215	220
His Gly Pro Arg Trp Tyr Ala Ser Leu Thr Val Glu Glu Thr Ile Pro		
225	230	235
Thr Pro Arg Leu Ser Lys Arg Lys Arg Ala Ala Gly Ile Ile Gly Val		
245	250	255
Asp Leu Gly Val Lys Ile Thr Ala Ala Leu Ser Asn Gly Asp Leu Ile		
260	265	270
Pro Asn Pro Arg Val Lys Ala Ser His Ala Lys Lys Leu Ala Arg Leu		
275	280	285
Gln Lys Ala Leu Ala Lys Ser Gln Lys Gly Ser Arg Asn Arg Ala Gln		
290	295	300
Leu Val Gln Lys Ile Gly Cys Leu Thr His Leu Glu Ala Arg Gln Arg		
305	310	315
Glu Gly His Ala His Asn Leu Ala Asn Arg Leu Val His Thr Trp Ala		
325	330	335
Ile Ile Gly Ile Glu Asp Leu Asn Val Ala Gly Met Thr Arg Ser Ser		
340	345	350
Arg Gly Thr Ile Glu Lys Pro Gly Lys Asn Val Arg Ala Lys Ala Gly		
355	360	365
Leu Asn Arg Ser Ile Leu Asp Val Ala Pro Ala Gln Ile Arg His Leu		
370	375	380
Leu Asp Tyr Lys Thr Ala Trp Ser Gly Thr Gln Leu Val Val Ile Asp		
385	390	395
Arg Trp Ala Pro Thr Ser Lys Lys Cys Ser Thr Cys Gly Ala Val Lys		
405	410	415
Ala Lys Leu Thr Leu Ala Glu Arg Thr Phe Glu Cys Glu Ala Cys Gly		
420	425	430
Leu Val Leu Asp Arg Asp Ile Asn Ala Ala Arg Asn Ile Ala Ala Leu		
435	440	445

Ala Ala Val Ala Pro Ser Thr Glu Glu Thr Gln Asn Ala Arg Arg Ala
 450 455 460
 Ala Pro Arg Lys Pro Val Pro Ser Thr Val Lys Gln Arg Ala Ala Met
 465 470 475 480
 Lys Arg Glu Asp Pro Pro Gly Ser Ser Pro Pro Ser Asn Gly Arg Thr
 485 490 495
 Phe His Thr Val Leu Val Gln Val Ser
 500 505
 <210> 15
 <211> 515
 <212> PRT
 <213> 链霉菌属
 <400> 15
 Met Glu Arg Glu Val Leu Arg Ala Phe Lys Phe Ala Leu Asp Pro Thr
 1 5 10 15
 Pro Ala Gln Ala Glu Ala Leu Ala Arg His Ala Gly Ala Ala Arg Trp
 20 25 30
 Ala Phe Asn Tyr Ala Leu Ala Val Lys Val Ser Ala His Gln Arg Trp
 35 40 45
 Arg Ala Glu Val Ala Gly Leu Val Ala Gln Gly Ile Glu Glu Ala Glu
 50 55 60
 Ala Arg Arg Arg Val Lys Val Pro Val Pro Ser Lys Pro Gln Ile Gln
 65 70 75 80
 Lys Arg Leu Asn Glu Val Lys Gly Asp Ser Arg Ile Asp Gly Arg Leu
 85 90 95
 Pro Glu Gly Thr Phe Gly Pro Glu Arg Pro Cys Pro Trp Trp Tyr Glu
 100 105 110
 Val Asn Thr Tyr Ala Phe Gln Ser Ala Phe Ile Asp Ala Asp Arg Ala
 115 120 125
 Trp Lys Asn Trp Leu Asp Ser Leu Arg Gly Val Arg Ala Gly Arg Lys
 130 135 140
 Val Gly Tyr Pro Arg Phe Lys Lys Lys Gly Arg Ser Arg Asp Ser Phe
 145 150 155 160
 Arg Leu His His Asp Val Lys Arg Pro Gly Ile Arg Leu Ala Thr Tyr
 165 170 175
 Arg Arg Leu Arg Leu Pro Thr Ile Gly Glu Val Arg Leu His Asp Ser
 180 185 190
 Gly Lys Arg Leu Gly Arg Leu Val Asp Arg Ser Leu Ala Val Val Gln
 195 200 205

Ser Val Thr Val Ser Arg Ala Gly His Arg Trp Tyr Ala Ser Val Leu			
210	215	220	
Cys Lys Val Thr Met Thr Val Pro Asp Gln Pro Ser Arg Arg Gln Arg			
225	230	235	240
Glu Arg Gly Thr Ile Gly Val Asp Leu Gly Val Lys Thr Leu Ala Ala			
	245	250	255
Leu Ser Gln Pro Leu Asp Pro Thr Asp Pro Asp Ser Asp Leu Ile Ala			
	260	265	270
Asn Pro Arg His Leu Ala Arg Ala Gln Gln Arg Leu Leu Lys Ala Gln			
	275	280	285
Arg Ala Leu Ser Arg Thr Glu Lys Gly Ser Arg Arg Arg Asp Arg Ala			
	290	295	300
Arg Arg Lys Val Ala Arg Leu His His Glu Val Ala Leu Arg Arg Glu			
305	310	315	320
Ser Ala Leu His Ala Val Thr Lys Arg Leu Ala Thr Ala Phe Ala Val			
	325	330	335
Val Ala Val Glu Asp Leu His Val Ala Gly Met Thr Ala Ser Ala Arg			
	340	345	350
Gly Thr Leu Glu Lys Pro Gly Arg Arg Val Arg Gln Lys Ala Gly Leu			
	355	360	365
Asn Arg Ala Val Leu Asp Ala Ala Pro Gly Glu Phe Arg Arg Gln Leu			
	370	375	380
Thr Tyr Lys Thr Ser Trp Tyr Gly Ser Lys Leu Ala Val Cys Asp Arg			
385	390	395	400
Trp Phe Pro Ser Ser Lys Thr Cys Ser Ala Cys Gly Trp Gln Asn Pro			
	405	410	415
His Leu Thr Leu Thr Asp Arg Thr Phe His Cys Pro Asp Cys Gly Leu			
	420	425	430
Thr Ile Asp Arg Asp Leu Asn Ala Ala Arg Asn Ile Ala Arg His Ala			
	435	440	445
Thr Val Ala Asp Ala Pro Pro Val Ala Pro Gly Arg Gly Glu Thr Gln			
	450	455	460
Asn Ala Arg Arg Ala Pro Val Arg Pro Asp Asp Arg Lys Ala Ala Arg			
465	470	475	480
His Gly Ala Met Lys Arg Glu Asp Thr Arg Pro Leu Gly Gln Val Pro			
	485	490	495
Pro Gln Arg Ser Asn Pro Leu Ala Ser Pro Pro Thr Gln Lys Gln Ala			
	500	505	510
Thr Leu Phe			

515

<210> 16

<211> 553

<212> PRT

<213> 微球菌菌属

<400> 16

```

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

```


260	265	270
Gln Arg Ala Asn Gly Ala Val	Gly Val Asp Leu Gly Val	Lys His Leu
275	280	285
Ala Ala Leu Ser Thr Gly Glu Val	Ile Asp Asn Gly Arg Pro	Gly Ala
290	295	300
Arg Gln Ala Ala Arg Leu Ala	Arg Leu Gln Arg Ala Tyr	Ala Arg Thr
305	310	315
Gln Pro Gly Ser Asn Arg Arg	Glu Arg Val Arg Arg Gln	Ile Ala Ala
325	330	335
Leu His His Gly Ile Ala Leu	Arg Arg Ala Gly Leu Leu	His Gln Val
340	345	350
Ser Thr Arg Leu Ala Thr Asp	Phe Ala Val Val Ala Leu	Glu Asp Leu
355	360	365
Asn Val Ala Gly Met Thr Arg	Ser Ala Arg Gly Thr Leu	Glu Ala Pro
370	375	380
Gly Arg Asn Val Ala Ala Lys	Ser Gly Leu Asn Arg Ala	Ile Leu Asp
385	390	395
Ala Gly Leu Gly Met Leu Arg	Arg Gln Leu Asp Tyr Lys	Thr Ser Trp
405	410	415
Ala Gly Ser Gln Val Lys Met	Ile Asp Arg Phe Ala Pro	Ser Ser Lys
420	425	430
Ala Cys Ser Arg Cys Gly Thr	Val Lys Ser Thr Leu Ser	Leu Ala Glu
435	440	445
Arg Thr Phe Glu Cys Glu Ala	Cys His Leu Val Ile Asp	Arg Asp Val
450	455	460
Asn Ala Ala Ile Asn Ile Arg	Ala Trp Ala Val Gln Glu	Glu Arg Gly
465	470	475
Ala Gly Val Glu Leu Ala Arg	Gly Arg Arg Glu Ser Arg	Asn Gly Arg
485	490	495
Gly Ala Ala Val Ser Gly Pro	Pro Ser Gly Gly Ala Ala	Gly Gln Gly
500	505	510
Arg Gly Ser Val Lys Pro Ala	Pro Gln Gly Val Gly Met	Ser Ser Arg
515	520	525
Ala Thr Gly Trp Ser Ser Gln	Pro Pro Ser Thr Glu Gly	Glu Ser Ala
530	535	540
Glu Arg Gly Ala Ser Ala Leu	Ala Arg	
545	550	
<210> 17		
<211> 557		

<212> PRT

<213> 芽孢杆菌属

<400> 17

```

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

```

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

<210> 18
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 18

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

305	310	315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly			
	325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu			
	340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Phe His Gly Ile Glu Val Ile			
	355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
	370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln			
385	390	395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala			
	405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
	420	425	430
Lys Leu His Ser Lys Lys Cys Ile Glu Glu Tyr Ile Glu Glu Leu Gly			
	435	440	445
Tyr Leu Asp			
450			
<210> 19			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 19			
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr			
	20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe			
	35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr			
	50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile			
65	70	75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met			
	85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn			
	100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile			

115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430

Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 20
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 20
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Ile Lys Leu Ala Thr
 165 170 175
 Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
 180 185 190
 Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
 195 200 205
 Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
 210 215 220
 Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
 225 230 235 240

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

355	360	365
Glu Gly Leu Arg Thr Leu Ala Leu Val		
370	375	
<210> 22		
<211> 445		
<212> PRT		
<213> 芽孢杆菌属		
<400> 22		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		15
20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		160
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		240

	245		250		255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr					
	260		265		270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro					
	275		280		285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn					
	290		295		300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys					
305		310		315	320
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg					
	325		330		335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn					
	340		345		350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr					
	355		360		365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg					
	370		375		380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His					
385		390		395	400
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn					
	405		410		415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys					
	420		425		430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp					
	435		440		445

<210> 23

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 23

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

65	70	75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met			
	85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn			
	100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile			
	115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Lys Lys Glu Lys Asn			
	130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys			
145	150	155	160
Glu Asn Ser Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr			
	165	170	175
Lys Lys Gln Lys Ser Met Lys Ile Ile Leu Asp Arg Leu Met Asn Gln			
	180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp			
	195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Asn Phe			
210	215	220	
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val			
225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Lys Ala Phe Asn Glu Lys Ile Arg Gln Arg Arg Ile Asn Leu			
	260	265	270
Leu Lys Gln Ser Asn Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg			
	275	280	285
Ala Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys			
	290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln			
305	310	315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Ile Glu Leu Leu Lys Gly			
	325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu			
	340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Val His Gly Ile Glu Val Ile			
	355	360	365
Lys Val Ala Ser Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
370	375	380	

Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 24
 <211> 506
 <212> PRT
 <213> 芽孢杆菌属
 <400> 24
 Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
 1 5 10 15
 Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
 20 25 30
 Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
 35 40 45
 Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala
 50 55 60
 Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr
 65 70 75 80
 Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala
 85 90 95
 Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
 100 105 110
 Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
 115 120 125
 Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
 130 135 140
 Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
 145 150 155 160
 Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
 165 170 175
 Leu Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
 180 185 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala		
195	200	205
Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp		
210	215	220
Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln		
225	230	235
Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn		
	245	250
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu		
	260	265
Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe		
	275	280
Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala		
	290	295
Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn		
305	310	315
Ile Asp Asp Phe Leu Arg Val Arg Thr Gln Leu Gln Ser Gln Arg Lys		
	325	330
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser		
	340	345
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe		
	355	360
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala		
	370	375
Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp		
385	390	395
Lys Leu Lys His Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
	405	410
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys		
	420	425
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His		
	435	440
Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn		
	450	455
Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala		
465	470	475
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met		
	485	490
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala		495

500	505
<210> 25	
<211> 377	
<212> PRT	
<213> 芽孢杆菌属	
<400> 25	
Met Leu Val Asn Lys Ala Tyr Lys Phe Arg Leu Tyr Pro Thr Lys Glu	
1 5 10 15	
Gln Lys Thr Leu Ile Ala Lys Thr Ile Gly Cys Ser Arg Phe Val Phe	
20 25 30	
Asn His Phe Leu Gly Gln Trp Asn Asp Thr Tyr Lys Glu Thr Gly Lys	
35 40 45	
Gly Leu Thr Tyr Asn Ser Ser Ser Ala Glu Leu Thr Lys Leu Lys Lys	
50 55 60	
Glu Leu Val Arg Leu Lys Glu Val Asp Ser Ile Ala Leu Gln Ser Ser	
65 70 75 80	
Leu Lys Asn Leu Ala Asp Ser Tyr Ser Arg Phe Phe Lys Lys Gln Asn	
85 90 95	
Asn Ala Pro Arg Phe Lys Ser Lys Arg Asn Arg Val Gln Ser Tyr Thr	
100 105 110	
Thr Lys Glu Thr Asn Gly Asn Ile Ala Val Val Gly Asn Lys Met Lys	
115 120 125	
Leu Pro Lys Leu Gly Leu Val Arg Phe Ala Lys Ser Arg Glu Val His	
130 135 140	
Gly Arg Val Leu Asn Ala Thr Val Arg Arg Thr Pro Ser Gly Lys Tyr	
145 150 155 160	
Phe Val Ser Ile Leu Ala Glu Val Asp Val Leu Pro Met Glu Lys Ala	
165 170 175	
Glu Ser Ser Ile Gly Ile Asp Leu Gly Ile Thr Asp Phe Ala Ile Phe	
180 185 190	
Ser Asp Gly Arg Met Ile Asp Asn Asn Lys Phe Thr Ala Lys Met Glu	
195 200 205	
Lys Lys Leu Lys Arg Glu Gln Arg Lys Leu Ser Arg Arg Ala Leu His	
210 215 220	
Ala Lys Gln Asn Gly Ile Asn Leu Leu Asp Ala Lys Asn Tyr Gln Lys	
225 230 235 240	
Gln Lys Arg Lys Val Ala Arg Leu His Glu Arg Val Leu Asn Gln Arg	
245 250 255	
Asp Asp Phe Leu Asn Lys Leu Ser Thr Glu Ile Ile Lys Asn His Asp	

260	265	270
Leu Ile Cys Ile Glu Asn Leu Asn Thr Lys Gly Met Leu Arg Asn His		
275	280	285
Lys Leu Ala Lys Ser Ile Ser Asp Val Ser Trp Ser Ala Phe Val Ser		
290	295	300
Lys Leu Glu Tyr Lys Ala Thr Trp Tyr Gly Lys Ala Ile Met Lys Val		
305	310	315
Ser Lys Trp Phe Pro Ser Ser Gln Ile Cys Ser Asp Cys Gly His Gln		
325	330	335
Asp Cys Lys Lys Ser Leu Ala Ile Arg Glu Trp Thr Cys Pro Ile Cys		
340	345	350
His Gln His His Asp Arg Asp Ile Asn Ala Ser Lys Asn Ile Leu Ala		
355	360	365
Glu Gly Leu Arg Thr Leu Ala Leu Ala		
370	375	
<210> 26		
<211> 445		
<212> PRT		
<213> 芽孢杆菌属		
<400> 26		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		

145	150	155	160
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr			
	165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln			
	180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp			
	195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe			
	210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val			
225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr			
	260	265	270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro			
	275	280	285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn			
	290	295	300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys			
305	310	315	320
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg			
	325	330	335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn			
	340	345	350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr			
	355	360	365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg			
	370	375	380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His			
385	390	395	400
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn			
	405	410	415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys			
	420	425	430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp			
	435	440	445
<210> 27			
<211> 448			

<212> PRT

<213> 芽孢杆菌属

<400> 27

```

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

```

Leu	Gln	Pro	Ile	Asp	Val	Leu	Ser	Asn	Lys	Ile	Ala	Lys	Phe	Arg	Asn
290						295					300				
Ser	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Val	Lys	Gln	Cys	Leu	Lys
305						310				315					320
His	Asn	Cys	Gly	Arg	Ile	Gln	Met	Glu	Leu	Leu	Lys	Gly	Ile	Ser	Lys
				325					330					335	
Asn	Asp	Arg	Ile	Leu	Lys	Asp	Trp	Thr	Tyr	Phe	Asp	Leu	Gln	Glu	Lys
			340						345					350	
Ile	Lys	Asn	Gln	Ala	Glu	Ile	His	Gly	Ile	Glu	Val	Ile	Lys	Val	Ala
		355					360						365		
Pro	Ala	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Gln	Cys	Gly	Tyr	Ile	Cys	Lys
		370				375					380				
Glu	Asn	Arg	Cys	Thr	Gln	Ala	Thr	Phe	Glu	Cys	Lys	Gln	Cys	Gly	Tyr
385					390					395					400
Lys	Thr	His	Ala	Asp	Tyr	Asn	Ala	Ala	Lys	Asn	Ile	Ser	Thr	Tyr	Asp
				405						410				415	
Ile	Glu	Asn	Ile	Ile	Asn	Lys	Gln	Leu	Ala	Val	Gln	Ser	Lys	Leu	His
			420						425				430		
Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly	Tyr	Leu	Asp
		435						440					445		
<210> 28															
<211> 451															
<212> PRT															
<213> 芽孢杆菌属															
<400> 28															
Met	Gly	Val	Thr	Ile	Lys	Ile	Met	Lys	Tyr	Gln	Ile	Leu	Cys	Pro	Met
1				5					10					15	
Asn	Val	Asp	Trp	Thr	Ile	Phe	Glu	Lys	His	Leu	Arg	Asn	Leu	Thr	Tyr
		20						25					30		
Gln	Val	Arg	Thr	Ile	Ser	Asn	Arg	Thr	Ile	Gln	Gln	Leu	Trp	Glu	Phe
		35					40						45		
Asp	Ala	Leu	Ser	Phe	Asp	Tyr	Phe	Lys	Glu	Arg	Gly	Thr	Tyr	Pro	Thr
		50				55					60				
Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile
65					70					75					80
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met
				85					90					95	
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn
			100						105					110	

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		

			420				425				430				
Lys	Leu	His	Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly
			435				440				445				
Tyr	Leu	Asp													
			450												
<210>			29												
<211>			471												
<212>			PRT												
<213>			芽孢杆菌属												
<400>			29												
Met	Ile	Ile	Ala	Arg	Lys	Ile	Lys	Leu	Ile	Ile	Ile	Gly	Glu	Asp	Arg
1				5				10				15			
Asp	Thr	Gln	Tyr	Lys	Phe	Ile	Arg	Glu	Glu	Arg	Tyr	Lys	Gln	Asn	Lys
			20				25				30				
Ala	Leu	Asn	Val	Ala	Met	Asn	His	Leu	Tyr	Phe	Leu	His	Val	Ala	Lys
			35				40				45				
Glu	Lys	Ile	Arg	Leu	Leu	Asp	Asn	Lys	Phe	Leu	Gln	Asp	Glu	Lys	Lys
			50				55				60				
Leu	Gln	Glu	Gly	Ile	Lys	Lys	Leu	Tyr	Ala	Glu	Lys	Lys	Val	Ile	Lys
65				70				75				80			
Asp	Gly	Lys	Lys	Arg	Asn	Glu	Leu	Glu	Lys	Lys	Ile	Glu	Lys	Gln	Thr
			85				90				95				
Asn	Glu	Leu	Lys	Lys	Leu	Arg	Ser	Lys	Gly	Asn	Lys	Glu	Ala	Asp	Lys
			100				105				110				
Ile	Leu	Gln	Glu	Ala	Ile	Lys	Ile	Asn	Leu	Ser	Ser	Thr	Thr	Arg	Glu
			115				120				125				
Val	Ile	Ser	Lys	Gln	Phe	Asp	Leu	Ile	Ser	Asp	Thr	Lys	Asp	Arg	Ile
			130				135				140				
Thr	Gln	Lys	Val	Tyr	Gln	Asp	Phe	Lys	Ser	Asp	Leu	Lys	Asn	Gly	Leu
145				150				155				160			
Leu	Ser	Gly	Glu	Arg	Val	Leu	Arg	Thr	Tyr	Lys	Lys	Asn	Asn	Pro	Leu
			165				170				175				
Leu	Ile	Arg	Gly	Arg	Ala	Leu	Asn	Phe	Tyr	Arg	Glu	Gly	Lys	Asp	Val
			180				185				190				
Met	Ile	Lys	Trp	Phe	Gly	Gly	Ile	Ile	Phe	Lys	Cys	Met	Leu	Gly	Gln
			195				200				205				
His	Lys	Asn	Asn	Ala	Gln	Glu	Leu	Lys	Ala	Thr	Leu	Asn	Lys	Val	Leu
			210				215				220				
Glu	Gly	Ser	Tyr	Lys	Val	Cys	Asp	Ser	Ser	Ile	Ser	Val	Gly	Lys	Glu

225	230	235	240
Leu Ile Leu Asn Ile Ser Leu Asp Ile Gly Glu Val Asn Ser Asn Val			
	245	250	255
Ser Cys Lys Lys Gly Arg Val Leu Gly Val Asp Leu Gly Met Lys Val			
	260	265	270
Pro Ala Tyr Met Ser Ile Asn Asp Lys Pro Tyr Ile Arg Lys Ser Leu			
	275	280	285
Gly Ser Leu Asp Asp Phe Leu Arg Ile Arg Val Gln Met Gln Lys Arg			
	290	295	300
Arg Arg Asn Leu His Lys Thr Leu Val Ser Val Lys Gly Gly Lys Gly			
305	310	315	320
Arg Glu Lys Lys Leu Gln Ala Leu Asp Arg Leu Lys Glu Lys Asn Phe			
	325	330	335
Ala Thr Thr Tyr Asn His Phe Leu Ser Tyr Asn Ile Val Lys Phe Ala			
	340	345	350
Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala Leu Ala			
	355	360	365
Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr Gln Leu			
	370	375	380
Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp Val Lys			
385	390	395	400
Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys Arg Asn			
	405	410	415
Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys Lys Ser			
	420	425	430
Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn Ile Ala			
	435	440	445
His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr Phe Lys			
	450	455	460
Lys Leu Gln His Thr Thr Glu			
465	470		
<210> 30			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<220>			
<221> MOD_RES			
<222> (160) .. (160)			
<223> 任何氨基酸			

<400> 30

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

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln			
305	310	315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly			
	325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu			
	340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile			
	355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
	370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln			
385	390	395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
	405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
	420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly			
	435	440	445
Tyr Leu Asp			
450			
<210> 31			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 31			
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr			
	20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe			
	35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr			
	50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile			
65	70	75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met			
	85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn			
	100	105	110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		

			420				425				430				
Lys	Leu	His	Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly
			435				440				445				
Tyr	Leu	Asp													
			450												
<210>			32												
<211>			451												
<212>			PRT												
<213>			芽孢杆菌属												
<400>			32												
Met	Gly	Val	Thr	Ile	Lys	Ile	Met	Lys	Tyr	Gln	Ile	Leu	Cys	Pro	Met
1				5				10				15			
Asn	Val	Asp	Trp	Thr	Ile	Phe	Glu	Lys	His	Leu	Arg	Asn	Leu	Thr	Tyr
			20				25				30				
Gln	Val	Arg	Thr	Ile	Ser	Asn	Arg	Thr	Ile	Gln	Gln	Leu	Trp	Glu	Phe
			35				40				45				
Asp	Ala	Leu	Ser	Phe	Asp	Tyr	Phe	Lys	Glu	Arg	Gly	Thr	Tyr	Pro	Thr
			50				55				60				
Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile
65				70				75				80			
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met
			85				90				95				
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn
			100				105				110				
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile
			115				120				125				
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Ile	Lys	Glu	Lys	Asn
			130				135				140				
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys
145				150				155				160			
Glu	Asn	Gly	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Gly	Thr
			165				170				175				
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln
			180				185				190				
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp
			195				200				205				
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Lys	Phe
			210				215				220				
Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Gly	Ile	Asn	Thr	Val

225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu			
	260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg			
	275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys			
	290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln			
305	310	315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly			
	325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu			
	340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile			
	355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
	370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln			
385	390	395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
	405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
	420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly			
	435	440	445
Tyr Leu Asp			
450			
<210> 33			
<211> 471			
<212> PRT			
<213> 芽孢杆菌属			
<400> 33			
Met Ile Ile Ala Arg Lys Ile Lys Leu Ile Ile Ile Gly Glu Asp Arg			
1	5	10	15
Asp Thr Gln Tyr Lys Phe Ile Arg Glu Glu Arg Tyr Lys Gln Asn Lys			
	20	25	30
Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Leu His Val Ala Lys			

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

Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala Leu Ala
 355 360 365
 Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr Gln Leu
 370 375 380
 Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp Val Lys
 385 390 395 400
 Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys Arg Asn
 405 410 415
 Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys Lys Ser
 420 425 430
 Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn Ile Ala
 435 440 445
 His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr Phe Lys
 450 455 460
 Lys Leu Gln His Thr Thr Glu
 465 470
 <210> 34
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 34
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Thr Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn
 130 135 140

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

450
 <210> 35
 <211> 196
 <212> PRT
 <213> 芽孢杆菌属
 <400> 35
 Met Asp Val Gly Leu Lys Glu Leu Phe Val Ala Ser Asn Gly Met Lys
 1 5 10 15
 Glu Arg Asn Ile Asn Lys Asp Ala Lys Val Lys Lys Leu Leu Lys Arg
 20 25 30
 Lys Lys Ser Ala Gln Arg Asp Met Ser Arg Arg Phe Lys Asn Gly Val
 35 40 45
 Lys Ile Gln Ser Ala Gly Tyr Glu Lys Ala Lys Ala Glu His Leu Arg
 50 55 60
 Leu Ser Arg Lys Ile Thr Asn Ile Arg Asn Asn His Ile His Gln Ala
 65 70 75 80
 Thr Ala Thr Leu Val Lys Thr Lys Pro Met Arg Ile Val Val Glu Asp
 85 90 95
 Leu Ser Ile Ser Asn Leu Leu Lys Asn Lys Lys Leu Ser Lys Ala Leu
 100 105 110
 Ser Phe Gln Lys Leu Asn Leu Phe Phe Gln Cys Leu Ser Tyr Lys Cys
 115 120 125
 Glu Lys Tyr Gly Ile Glu Tyr Val Lys Ala Asp Lys Trp Val Ala Ser
 130 135 140
 Ser Lys Ile Cys Ser Cys Cys Gly Val Lys Tyr Asp His Ser Val Gln
 145 150 155 160
 Ser Glu Gly Gln Trp Ser Leu Lys Ile Arg Glu Trp Arg Cys Val Arg
 165 170 175
 Cys Asn Ser His His Asp Arg Asp Val Asn Ala Ala Ile Asn Leu Ser
 180 185 190
 Arg Trp Val Lys
 195
 <210> 36
 <211> 445
 <212> PRT
 <213> 芽孢杆菌属
 <400> 36
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
 1 5 10 15

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

	325		330		335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn					
	340		345		350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr					
	355		360		365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg					
	370		375		380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His					
385		390		395	400
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn					
	405		410		415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys					
	420		425		430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp					
	435		440		445
<210> 37					
<211> 451					
<212> PRT					
<213> 芽孢杆菌属					
<400> 37					
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met					
1	5		10		15
Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr					
	20		25		30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe					
	35		40		45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr					
50		55		60	
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile					
65		70		75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met					
	85		90		95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn					
	100		105		110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile					
	115		120		125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn					
	130		135		140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys					

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

<210> 38

<211> 362

<212> PRT

<213> 芽孢杆菌属

<400> 38

```

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

```

Asn Met Gln Lys Asn His Lys Leu Thr Arg Ala Ile Ser Glu Val Ser		
275	280	285
Trp Ser Lys Phe Arg Glu Tyr Leu Ala Tyr Lys Thr Ala Trp Lys Gly		
290	295	300
Arg Asp Leu Ile Val Ala Pro Lys Asn Tyr Ala Ser Ser Gln Leu Cys		
305	310	315
Ser Cys Cys Gly Tyr Lys Asn Lys Glu Val Lys Asn Leu Asn Leu Arg		
	325	330
Glu Trp Thr Cys Pro Glu Cys Asn Ser His His Asn Arg Asp Ile Asn		
	340	345
Ala Ser Ile Asn Leu Leu Lys Leu Ala Met		
	355	360
<210> 39		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 39		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
	20	25
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
	35	40
Asp Val Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
	50	55
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
	100	105
Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile		
	115	120
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
	130	135
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
	165	170
		175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Asn Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 40		
<211> 451		
<212> PRT		

<213> 芽孢杆菌属

<400> 40

```

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

```

290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		320
	325	330
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		335
	340	345
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		350
	355	360
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		365
	370	375
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		380
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		400
	405	410
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		415
	420	425
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		430
	435	440
		445
Tyr Leu Asp		
450		
<210> 41		
<211> 450		
<212> PRT		
<213> 芽孢杆菌属		
<400> 41		
Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val		
1	5	10
Asn Ile Glu Trp Lys Thr Phe Glu Ile Tyr Leu Arg Thr Leu Ser Tyr		
	20	25
His Phe Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Tyr		30
	35	40
Asp Asn Gln Ser Leu Lys His Phe Lys Asp Thr Gly Gln Tyr Pro Ser		45
	50	55
Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile		60
65	70	75
Tyr Asp Gln Leu Lys Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Thr Ile Arg Thr Trp Asn Ser Arg Lys Lys		95

100	105	110
Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu		
115	120	125
Pro Ile Asp Ile His Gly Asn Ser Ile Gln Ile Ile Lys Glu Lys Ser		
130	135	140
Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys		
145	150	155
Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Val Lys Leu Ser Thr Arg		
165	170	175
Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Ile Ile Asp Ser Thr		
180	185	190
Tyr Ala Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr		
195	200	205
Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp		
210	215	220
Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr		
225	230	235
Phe Ala Phe Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu		
245	250	255
Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu		
260	265	270
Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu		
275	280	285
Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe		
290	295	300
Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys		
305	310	315
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile		
325	330	335
Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln		
340	345	350
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys		
355	360	365
Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile		
370	375	380
His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln Gln Cys		
385	390	395
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Leu		
405	410	415

Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
 420 425 430
 Leu Asn Gln Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Lys Asn
 435 440 445
 Ile Asn
 450
 <210> 42
 <211> 450
 <212> PRT
 <213> 芽孢杆菌属
 <400> 42
 Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
 1 5 10 15
 Asn Ile Glu Trp Lys Thr Phe Glu Ile Tyr Leu Arg Thr Leu Ser Tyr
 20 25 30
 His Phe Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Tyr
 35 40 45
 Asp Asn Gln Ser Leu Lys His Phe Lys Asp Thr Gly Gln Tyr Pro Ser
 50 55 60
 Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
 65 70 75 80
 Tyr Asp Gln Leu Lys Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Thr Ile Arg Thr Trp Asn Ser Arg Lys Lys
 100 105 110
 Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu
 115 120 125
 Pro Ile Asp Ile His Gly Asn Ser Ile Gln Ile Ile Lys Glu Lys Ser
 130 135 140
 Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
 145 150 155 160
 Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Val Lys Leu Ser Thr Arg
 165 170 175
 Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Ile Ile Asp Ser Thr
 180 185 190
 Tyr Ala Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr
 195 200 205
 Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp
 210 215 220

Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr			
225	230	235	240
Phe Ala Phe Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu			
	245	250	255
Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu			
	260	265	270
Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu			
	275	280	285
Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe			
	290	295	300
Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys			
305	310	315	320
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile			
	325	330	335
Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln			
	340	345	350
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys			
	355	360	365
Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile			
	370	375	380
His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln Gln Cys			
385	390	395	400
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Leu			
	405	410	415
Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys			
	420	425	430
Leu Asn Gln Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Lys Asn			
	435	440	445
Ile Asn			
450			
<210> 43			
<211> 444			
<212> PRT			
<213> 芽孢杆菌属			
<400> 43			
Met Lys Tyr Gln Ile Leu Cys Pro Met Asn Val Asp Trp Thr Ile Phe			
1	5	10	15
Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn			
	20	25	30

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

340	345	350
Ala Glu Ile Tyr Gly Ile Glu Val	Ile Lys Val Val Pro	Ala Tyr Thr
355	360	365
Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys		
370	375	380
Thr Gln Ala Met Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala		
385	390	395
Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile		
405	410	415
Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys		
420	425	430
Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp		
435	440	
<210> 44		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 44		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg His Leu Thr Tyr		
20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35	40	45
Asp Thr Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		

	165		170		175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln					
	180		185		190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp					
	195		200		205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe					
	210		215		220
Asp Lys Glu Leu Val Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val					
225		230		235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp					
	245		250		255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu					
	260		265		270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg					
	275		280		285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys					
	290		295		300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln					
305		310		315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly					
	325		330		335
Ile Ser Lys Asn Asp Lys Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu					
	340		345		350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile					
	355		360		365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr					
	370		375		380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln					
385		390		395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser					
	405		410		415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser					
	420		425		430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly					
	435		440		445
Tyr Leu Asp					
	450				
<210> 45					
<211> 451					

<212> PRT

<213> 芽孢杆菌属

<400> 45

```

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

```

Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
 290 295 300
 Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln
 305 310 315 320
 Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
 325 330 335
 Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
 340 345 350
 Gln Glu Lys Ile Gln Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
 355 360 365
 Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
 370 375 380
 Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 46
 <211> 506
 <212> PRT
 <213> 芽孢杆菌属
 <400> 46
 Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
 1 5 10 15
 Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
 20 25 30
 Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
 35 40 45
 Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala
 50 55 60
 Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
 65 70 75 80
 Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
 85 90 95

Lys	Glu	Lys	Val	Asn	Asp	Ser	Lys	Asp	Asp	Glu	Lys	Leu	Gln	Lys	Arg
100				105				110							
Val	Asp	Lys	Ala	Arg	Glu	Ala	Tyr	Asn	Lys	Ala	Gln	Glu	Lys	Val	Tyr
115				120				125							
Lys	Ile	Glu	Lys	Glu	Phe	Asn	Lys	Lys	Ser	Met	Glu	Thr	Tyr	Gln	Lys
130				135				140							
Val	Val	Gly	Leu	Ser	Lys	Gln	Thr	Arg	Ile	Gly	Lys	Leu	Leu	Lys	Ser
145				150				155				160			
Gln	Phe	Thr	Leu	His	Tyr	Asp	Thr	Glu	Asp	Arg	Ile	Thr	Ser	Thr	Val
165				170				175							
Ile	Ser	His	Phe	Asn	Asn	Asp	Met	Lys	Thr	Gly	Val	Leu	Arg	Gly	Asp
180				185				190							
Arg	Ser	Leu	Arg	Thr	Tyr	Lys	Asn	Ser	His	Pro	Leu	Leu	Val	Arg	Ala
195				200				205							
Arg	Ser	Met	Lys	Val	Tyr	Glu	Glu	Asn	Gly	Asp	Tyr	Phe	Ile	Lys	Trp
210				215				220							
Val	Lys	Gly	Ile	Val	Phe	Lys	Ile	Val	Ile	Ser	Ala	Gly	Ser	Lys	Gln
225				230				235				240			
Lys	Ala	Asn	Ile	Gly	Glu	Leu	Lys	Ser	Val	Leu	Ile	Asn	Ile	Leu	Asn
245				250				255							
Gly	His	Tyr	Lys	Val	Cys	Asp	Ser	Ser	Ile	Ser	Leu	Asn	Lys	Asp	Leu
260				265				270							
Ile	Leu	Asn	Leu	Ser	Leu	Asn	Ile	Pro	Val	Ser	Lys	Glu	Asn	Val	Phe
275				280				285							
Val	Pro	Gly	Arg	Val	Val	Gly	Val	Asp	Leu	Gly	Leu	Lys	Ile	Pro	Ala
290				295				300							
Tyr	Val	Ser	Leu	Asn	Asp	Thr	Pro	Tyr	Ile	Lys	Lys	Gly	Ile	Gly	Asn
305				310				315				320			
Ile	Asp	Asp	Phe	Leu	Lys	Val	Arg	Thr	Gln	Leu	Gln	Ser	Gln	Arg	Lys
325				330				335							
Arg	Leu	Gln	Lys	Thr	Leu	Glu	Cys	Thr	Ser	Gly	Gly	Lys	Gly	Arg	Asn
340				345				350							
Lys	Lys	Leu	Lys	Gly	Leu	Asp	Arg	Leu	Lys	Ala	Lys	Glu	Lys	Asn	Phe
355				360				365							
Val	Asn	Thr	Tyr	Asn	His	Phe	Leu	Ser	Lys	Lys	Ile	Ile	Gln	Phe	Ala
370				375				380							
Val	Lys	Asn	Asn	Ala	Gly	Val	Ile	His	Leu	Glu	Glu	Leu	Gln	Phe	Asp
385				390				395				400			
Lys	Leu	Lys	His	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu

	405		410		415
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys					
	420		425		430
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His					
	435		440		445
Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn					
	450		455		460
Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Asp Phe Asn Ala					
465		470		475	480
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met					
	485		490		495
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala					
	500		505		
<210> 47					
<211> 506					
<212> PRT					
<213> 芽孢杆菌属					
<400> 47					
Met Ser Thr Pro Leu Gln Gln Ala His Gln Lys Ser Lys Lys Thr Ser					
1	5		10		15
Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn					
	20		25		30
Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn					
	35		40		45
Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala					
	50		55		60
Thr Glu Lys Leu Lys His Ser Asp Ala Glu Tyr Gln Lys His Leu Thr					
65		70		75	80
Lys Tyr Arg Glu Val Ala Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Ala					
	85		90		95
Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Glu Lys Arg					
	100		105		110
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr					
	115		120		125
Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Lys Glu Thr Tyr Gln Lys					
	130		135		140
Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser					
145		150		155	160
Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val					

				165						170					175				
Leu	Ser	His	Phe	Asn	Asn	Asp	Met	Lys	Thr	Gly	Val	Leu	Arg	Gly	Asp				
				180						185					190				
Arg	Ser	Leu	Arg	Thr	Tyr	Lys	Asn	Ser	His	Pro	Leu	Leu	Val	Arg	Ala				
				195						200					205				
Arg	Ser	Met	Lys	Val	Tyr	Glu	Glu	Asn	Gly	Asp	Tyr	Phe	Ile	Lys	Trp				
				210						215					220				
Val	Lys	Gly	Ile	Val	Phe	Lys	Ile	Val	Ile	Ser	Ala	Gly	Ser	Lys	Gln				
225										230					235				240
Lys	Ala	Asn	Ile	Gly	Glu	Leu	Lys	Ser	Val	Leu	Ile	Asn	Ile	Leu	Asn				
				245						250					255				
Gly	His	Tyr	Lys	Val	Cys	Asp	Ser	Ser	Ile	Ser	Leu	Asn	Lys	Asp	Leu				
				260						265					270				
Ile	Leu	Asn	Leu	Ser	Leu	Asn	Ile	Pro	Val	Ser	Lys	Glu	Asn	Val	Phe				
				275						280					285				
Val	Pro	Gly	Arg	Val	Val	Gly	Val	Asp	Leu	Gly	Leu	Lys	Ile	Pro	Ala				
				290						295					300				
Tyr	Val	Ser	Leu	Asn	Asp	Thr	Pro	Tyr	Ile	Lys	Lys	Gly	Ile	Gly	Asn				
305										310					315				320
Ile	Asp	Asp	Phe	Leu	Lys	Val	Arg	Thr	Gln	Leu	Gln	Asn	Gln	Arg	Lys				
				325						330					335				
Arg	Leu	Gln	Lys	Thr	Leu	Glu	Cys	Thr	Ser	Gly	Gly	Lys	Gly	Arg	Ser				
				340						345					350				
Lys	Lys	Leu	Lys	Gly	Leu	Asp	Arg	Leu	Lys	Ala	Lys	Glu	Lys	Asn	Phe				
				355						360					365				
Val	Asn	Thr	Tyr	Asn	His	Phe	Leu	Ser	Lys	Lys	Ile	Ile	Gln	Phe	Ala				
				370						375					380				
Val	Lys	Asn	Asn	Ala	Gly	Val	Ile	His	Leu	Glu	Glu	Leu	Gln	Phe	Asp				
385										390					395				400
Lys	Leu	Lys	His	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu				
				405						410					415				
Gln	Thr	Met	Ile	Glu	Tyr	Lys	Ala	Glu	Arg	Glu	Gly	Ile	Lys	Val	Lys				
				420						425					430				
Tyr	Val	Asp	Ala	Ser	Tyr	Thr	Ser	Gln	Thr	Cys	Ser	Lys	Cys	Asp	His				
				435						440					445				
Tyr	Glu	Lys	Gly	Gln	Arg	Val	Leu	Gln	Asp	Thr	Phe	Thr	Cys	Lys	Asn				
				450						455					460				
Lys	Glu	Cys	Lys	Gly	Tyr	Asp	His	Lys	Val	Asn	Ala	Asp	Phe	Asn	Ala				
465										470					475				480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
 485 490 495
 Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
 500 505
 <210> 48
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 48
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
 165 170 175
 Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
 180 185 190
 Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
 195 200 205
 Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
 210 215 220
 Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
 225 230 235 240

Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
245	250	255	
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu			
260	265	270	
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg			
275	280	285	
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys			
290	295	300	
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln			
305	310	315	320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly			
325	330	335	
Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu			
340	345	350	
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile			
355	360	365	
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
370	375	380	
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln			
385	390	395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
405	410	415	
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
420	425	430	
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly			
435	440	445	
Tyr Leu Asp			
450			
<210> 49			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 49			
Met Arg Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr			
20	25	30	
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe			
35	40	45	

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

355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		400
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 50		
<211> 506		
<212> PRT		
<213> 芽孢杆菌属		
<400> 50		
Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser		
1	5	10
Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn		
20	25	30
Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn		
35	40	45
Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala		
50	55	60
Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr		
65	70	75
Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val		
85	90	95
Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg		
100	105	110
Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr		
115	120	125
Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys		
130	135	140
Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser		
145	150	155
Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val		160

				165						170					175				
Leu	Ser	His	Phe	Asn	Asn	Asp	Met	Lys	Thr	Gly	Val	Leu	Arg	Gly	Asp				
				180						185					190				
Arg	Ser	Leu	Arg	Thr	Tyr	Lys	Asn	Ser	His	Pro	Leu	Leu	Val	Arg	Ala				
				195						200					205				
Arg	Ser	Met	Lys	Val	Tyr	Glu	Glu	Asn	Gly	Asp	Tyr	Phe	Ile	Lys	Trp				
				210						215					220				
Val	Lys	Gly	Ile	Val	Phe	Lys	Ile	Val	Ile	Ser	Ala	Gly	Ser	Lys	Gln				
225										230					235				240
Lys	Ala	Asn	Ile	Gly	Glu	Leu	Lys	Ser	Val	Leu	Ile	Asn	Ile	Leu	Asn				
				245						250					255				
Gly	His	Tyr	Lys	Val	Cys	Asp	Ser	Ser	Ile	Ser	Leu	Asn	Lys	Asp	Leu				
				260						265					270				
Ile	Leu	Asn	Leu	Ser	Leu	Asn	Ile	Pro	Val	Ser	Lys	Glu	Asn	Val	Phe				
				275						280					285				
Val	Pro	Gly	Arg	Val	Val	Gly	Val	Asp	Leu	Gly	Leu	Lys	Ile	Pro	Ala				
				290						295					300				
Tyr	Val	Ser	Leu	Asn	Asp	Thr	Pro	Tyr	Ile	Lys	Lys	Gly	Ile	Gly	Asn				
305										310					315				320
Ile	Asp	Asp	Phe	Leu	Asn	Val	Arg	Thr	Gln	Leu	Gln	Asn	Gln	Arg	Lys				
				325						330					335				
Arg	Leu	Gln	Lys	Thr	Leu	Glu	Cys	Thr	Ser	Gly	Gly	Lys	Gly	Arg	Ser				
				340						345					350				
Lys	Lys	Leu	Lys	Gly	Leu	Asp	Arg	Leu	Lys	Ala	Lys	Glu	Lys	Asn	Phe				
				355						360					365				
Val	Asn	Thr	Tyr	Asn	His	Phe	Leu	Ser	Lys	Lys	Ile	Ile	Gln	Phe	Ala				
				370						375					380				
Val	Lys	Asn	Asn	Ala	Gly	Val	Ile	His	Leu	Glu	Glu	Leu	Gln	Phe	Asp				
385										390					395				400
Lys	Leu	Lys	Asn	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu				
				405						410					415				
Gln	Thr	Met	Ile	Glu	Tyr	Lys	Ala	Glu	Arg	Glu	Gly	Ile	Glu	Val	Lys				
				420						425					430				
Tyr	Val	Asp	Ala	Ser	Tyr	Thr	Ser	Gln	Thr	Cys	Ser	Lys	Cys	Gly	His				
				435						440					445				
Tyr	Glu	Lys	Gly	Gln	Arg	Val	Leu	Gln	Asp	Thr	Phe	Thr	Cys	Lys	Asn				
				450						455					460				
Lys	Glu	Cys	Lys	Gly	Tyr	Ala	His	Lys	Val	Asn	Ala	Asp	Phe	Asn	Ala				
465										470					475				480

Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
 485 490 495
 Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
 500 505
 <210> 51
 <211> 405
 <212> PRT
 <213> 芽孢杆菌属
 <400> 51
 Met Ala Arg Lys Thr Val Glu Gln Arg Asn Lys Glu Leu Ala Lys Glu
 1 5 10 15
 Asn Lys Val Leu Arg His Tyr Gly Leu Lys Leu Arg Ala Leu Pro Thr
 20 25 30
 Pro Phe Gln Glu Glu Lys Ile Ala Lys Thr Ile Gly Ser Ala Arg Phe
 35 40 45
 Ser Tyr Asn Phe Tyr Leu Asn Glu Lys Ile Glu Ile Tyr Lys Leu Thr
 50 55 60
 Asn Leu Thr Leu Asp Tyr Ser Ser Phe Lys Lys Ser Phe Asn Gly Leu
 65 70 75 80
 Lys Gln His Pro Ala Phe Asp Trp Leu Lys Glu Val Asp Lys Phe Ser
 85 90 95
 Leu Glu Ser Ala Leu Glu Gln Val Asp Asp Ala Phe Glu Arg Phe Phe
 100 105 110
 Lys Gly Gln Thr Lys Phe Pro Lys Phe Lys Ser Lys His Lys Thr Lys
 115 120 125
 Gln Ser Tyr Thr Thr Lys Glu Thr Asn Gly Asn Ile Ala Phe Asp Leu
 130 135 140
 Glu Asn Gln Val Val Lys Leu Pro Lys Met Gly Lys Val Pro Val Lys
 145 150 155 160
 Leu Ser Lys Lys His Leu Thr Met Phe Gln Lys Asn Glu Phe Thr Gly
 165 170 175
 Lys Ile Lys Ser Ala Thr Val Thr Arg His Ser Ser Gly Gln Tyr Tyr
 180 185 190
 Ile Ser Leu Lys Cys Glu Glu Ile Ile Leu Leu Glu Glu Lys Ile Asp
 195 200 205
 Val Thr Thr Ile Pro Thr Asp Gly Ile Ile Gly Cys Asp Leu Gly Leu
 210 215 220
 Thr Tyr Phe Leu Ile Asp Ser Asn Gly Gln Lys Ile Glu Asn Pro Arg
 225 230 235 240

Tyr Leu Lys Glu Asn Leu Lys Lys Leu Ala Lys Leu Gln Arg Gly Leu
 245 250 255
 Lys His Lys Lys Ile Gly Ser Ser Asn Phe Gln Lys Leu Lys Gln Lys
 260 265 270
 Ile Ala Lys Leu His Leu His Ile Ser Asn Met Arg Lys Asp Phe Leu
 275 280 285
 His Lys Val Ser Arg Lys Leu Val Asn Glu Asn Gln Val Ile Ile Leu
 290 295 300
 Glu Asp Leu Asn Val Lys Gly Met Ile Lys Asn Pro Lys Leu Ala Arg
 305 310 315 320
 Ser Ile Ala Asp Val Gly Trp Gly Met Phe Lys Thr Phe Val Ser Tyr
 325 330 335
 Lys Ala Asn Trp Ala Asn Lys Ile Leu Ile Leu Ile Asn Arg Phe Phe
 340 345 350
 Pro Ser Ser Lys Gln Cys Asn Gly Cys Lys Glu Lys Asn Thr Leu Leu
 355 360 365
 Ser Leu Ser Asp Arg Leu Trp Met Cys Pro Ser Cys Gly Thr His His
 370 375 380
 Asp Arg Asp Asp Asn Ala Ala Leu Asn Ile Lys Glu Glu Gly Ile Arg
 385 390 395 400
 Leu Leu Leu Asn Ala
 405

<210> 52

<211> 252

<212> PRT

<213> 芽孢杆菌属

<400> 52

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

Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His		
100	105	110
Lys Tyr Ala Asn Tyr Ile Ile Lys Gln Cys Leu Lys His Asn Cys Gly		
115	120	125
Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Lys Val		
130	135	140
Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Gln Asn Gln		
145	150	155
Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr Thr		
165	170	175
Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys		
180	185	190
Thr Gln Ala Val Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala		
195	200	205
Asp Tyr Asn Ala Ala Lys Asn Ile Ala Thr Tyr Asp Ile Glu Asn Ile		
210	215	220
Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys		
225	230	235
Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp		
245	250	

<210> 53

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 53

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

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
165	170	175
Gln Arg Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Glu Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Gln Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		

			420				425				430				
Lys	Leu	His	Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly
			435				440				445				
Tyr	Leu	Asp													
			450												
<210>			54												
<211>			445												
<212>			PRT												
<213>			芽孢杆菌属												
<400>			54												
Met	Gly	Val	Thr	Ile	Lys	Ile	Met	Lys	Tyr	Gln	Ile	Leu	Cys	Pro	Leu
1				5				10				15			
Asn	Val	Asp	Trp	Thr	Ile	Phe	Glu	Lys	His	Leu	Arg	Asn	Leu	Thr	Tyr
			20				25				30				
Gln	Val	Arg	Thr	Ile	Ser	Asn	Arg	Thr	Ile	Gln	Gln	Leu	Trp	Glu	Phe
			35				40				45				
Asp	Ala	Leu	Ser	Phe	Asp	Tyr	Phe	Lys	Glu	Arg	Gly	Thr	Tyr	Pro	Thr
			50				55				60				
Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile
65				70				75				80			
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met
			85				90				95				
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn
			100				105				110				
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile
			115				120				125				
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Thr	Lys	Glu	Lys	Asn
			130				135				140				
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys
145				150				155				160			
Glu	Asn	Asp	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Ala	Thr
			165				170				175				
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln
			180				185				190				
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp
			195				200				205				
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Lys	Phe
			210				215				220				
Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Leu	Gly	Gly	Ile	Asn	Thr	Val

225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr			
	260	265	270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro			
	275	280	285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn			
	290	295	300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys			
305	310	315	320
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg			
	325	330	335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn			
	340	345	350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr			
	355	360	365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg			
	370	375	380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His			
385	390	395	400
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn			
	405	410	415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys			
	420	425	430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp			
	435	440	445
<210> 55			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 55			
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15
Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr			
	20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe			
	35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr			

50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		95
	100	105
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		110
	115	120
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		125
	130	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		160
	165	170
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		175
	180	185
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		190
	195	200
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		205
	210	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		240
	245	250
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		255
	260	265
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		270
	275	280
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		285
	290	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		320
	325	330
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		335
	340	345
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		350
	355	360
		365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
 370 375 380
 Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 56
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 56
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Arg Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Ile Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Arg Leu Ala Thr
 165 170 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Met Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Leu Ile Thr Tyr Lys Phe Asn Ala Val Lys Glu Asn Ile Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Leu Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Tyr Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 57		
<211> 451		
<212> PRT		

<213> 芽孢杆菌属

<400> 57

```

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

```

290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		320
	325	330
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		335
	340	345
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		350
	355	360
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		365
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		400
	405	410
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		415
	420	425
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		430
435	440	445
Tyr Leu Asp		
450		
<210> 58		
<211> 445		
<212> PRT		
<213> 芽孢杆菌属		
<400> 58		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
	20	25
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		30
	35	40
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		45
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		95

	100		105		110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile					
115			120		125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn					
130			135		140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys					
145			150		155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr					
	165		170		175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln					
180			185		190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp					
195			200		205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe					
210			215		220
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val					
225			230		235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp					
	245		250		255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr					
260			265		270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro					
275			280		285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn					
290			295		300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys					
305			310		315
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg					
	325		330		335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn					
340			345		350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr					
355			360		365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg					
370			375		380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His					
385			390		395
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn					
	405		410		415

Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
 420 425 430
 Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
 435 440 445
 <210> 59
 <211> 445
 <212> PRT
 <213> 芽孢杆菌属
 <400> 59
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
 165 170 175
 Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln
 180 185 190
 Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
 195 200 205
 Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
 210 215 220
 Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val
 225 230 235 240

Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr		
260	265	270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro		
275	280	285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn		
290	295	300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys		
305	310	315
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg		
325	330	335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn		
340	345	350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr		
355	360	365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg		
370	375	380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His		
385	390	395
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn		
405	410	415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys		
420	425	430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp		
435	440	445

<210> 60

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 60

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

Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile
65					70					75					80
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met
				85						90					95
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn
			100						105					110	
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile
			115						120					125	
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Thr	Lys	Glu	Lys	Asn
			130				135							140	
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys
145					150					155					160
Glu	Asn	Gly	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Ala	Thr
				165						170					175
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln
				180						185					190
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp
			195							200					205
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Lys	Phe
			210							215					220
Asp	Lys	Asp	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Gly	Ile	Asn	Thr	Val
225					230					235					240
Tyr	Phe	Ala	Phe	Asn	Glu	Gly	Phe	Ile	Arg	Ser	Asn	Ile	Lys	Ser	Asp
				245						250					255
Glu	Ile	Lys	Ala	Phe	Asn	Glu	Arg	Ile	Arg	Gln	Arg	Arg	Ile	Asn	Leu
			260							265					270
Leu	Lys	Gln	Ser	Lys	Tyr	Cys	Ser	Asn	Ser	Arg	Thr	Gly	Lys	Gly	Arg
			275							280					285
Glu	Lys	Arg	Leu	Gln	Pro	Ile	Asp	Val	Leu	Ser	Asn	Lys	Ile	Ala	Lys
			290							295					300
Phe	Arg	Asn	Ser	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Ile	Lys	Gln
305					310					315					320
Cys	Leu	Lys	His	Asn	Cys	Gly	Arg	Ile	Gln	Met	Glu	Leu	Leu	Lys	Gly
				325						330					335
Ile	Ser	Lys	Asn	Asp	Lys	Val	Leu	Lys	Asp	Trp	Thr	Tyr	Phe	Asp	Leu
			340							345					350
Gln	Glu	Lys	Ile	Gln	Asn	Gln	Thr	Glu	Ile	His	Gly	Ile	Glu	Val	Ile
			355							360					365
Lys	Val	Ala	Pro	Ala	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Gln	Cys	Gly	Tyr

370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		400
	405	410
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		415
	420	425
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		430
	435	440
Tyr Leu Asp		445
450		
<210> 61		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 61		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		15
	20	25
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		30
	35	40
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Ile Gly Thr Tyr Pro Thr		45
	50	55
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		60
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		95
	100	105
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		110
	115	120
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Lys Lys Glu Lys Asn		125
	130	135
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		140
145	150	155
Glu Asn Ser Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		160
	165	170
Lys Lys Gln Lys Ser Met Lys Ile Ile Leu Asp Arg Leu Met Asn Gln		175

180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met	Ile His Lys Tyr Lys Asn Lys Trp	
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe	Asn Ala Ile Lys Glu Asn Asn Phe	
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile	Asp Met Gly Gly Ile Asn Thr Val	
225	230	235
240	245	250
Tyr Phe Ala Phe Asn Glu Gly Phe	Ile Arg Ser Asn Ile Lys Ser Asp	
255	260	265
Glu Ile Lys Ala Phe Asn Glu Lys	Ile Arg Gln Arg Arg Ile Asn Leu	
270	275	280
Leu Lys Gln Ser Asn Tyr Cys Ser	Asn Ser Arg Thr Gly Lys Gly Arg	
285	290	295
Ala Lys Arg Leu Gln Pro Ile Asp	Val Leu Ser Asn Lys Ile Ala Lys	
300	305	310
Phe Arg Asn Ser Thr Asn His Lys	Tyr Ala Asn Tyr Ile Ile Lys Gln	
315	320	325
Cys Leu Lys His Asn Cys Gly Arg	Ile Gln Ile Glu Leu Leu Lys Gly	
330	335	340
Ile Ser Lys Asn Asp Lys Val Leu	Lys Asp Trp Thr Tyr Phe Asp Leu	
345	350	355
Gln Glu Lys Ile Lys Asn Gln Ala	Glu Val His Gly Ile Glu Val Ile	
360	365	370
Lys Val Ala Ser Ala Tyr Thr Ser	Gln Arg Cys Ser Gln Cys Gly Tyr	
375	380	385
Ile Cys Lys Glu Asn Arg Cys Thr	Gln Ala Val Phe Glu Cys Lys Gln	
390	395	400
Cys Gly Tyr Lys Thr His Ala Asp	Tyr Asn Ala Ala Lys Asn Ile Ala	
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile	Asn Lys Gln Leu Ala Val Gln Ser	
420	425	430
Lys Leu His Ser Lys Lys Cys Met	Glu Glu Tyr Ile Glu Glu Leu Gly	
435	440	445
Tyr Leu Asp		
450		
<210> 62		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		

<400> 62

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

Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
 305 310 315 320
 Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
 325 330 335
 Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
 340 345 350
 Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
 355 360 365
 Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
 370 375 380
 Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445

Tyr Leu Asp

450

<210> 63

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 63

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

Glu Ile Arg Lys Gly Glu Ile Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Ile Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Asn Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		

			420				425					430				
Lys	Leu	His	Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly	
			435				440					445				
Tyr	Leu	Asp														
			450													
<210>			64													
<211>			445													
<212>			PRT													
<213>			芽孢杆菌属													
<400>			64													
Met	Gly	Val	Thr	Ile	Lys	Ile	Met	Lys	Tyr	Gln	Ile	Leu	Cys	Pro	Leu	
1				5				10					15			
Asn	Val	Asp	Trp	Thr	Ile	Phe	Glu	Lys	His	Leu	Arg	Asn	Leu	Thr	Tyr	
			20				25					30				
Gln	Val	Arg	Thr	Ile	Ser	Asn	Arg	Thr	Ile	Gln	Gln	Leu	Trp	Glu	Phe	
			35				40					45				
Asp	Ala	Leu	Ser	Phe	Asp	Tyr	Phe	Lys	Glu	Arg	Gly	Thr	Tyr	Pro	Thr	
			50				55					60				
Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile	
65				70				75					80			
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met	
			85				90					95				
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn	
			100				105					110				
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile	
			115				120					125				
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Thr	Lys	Glu	Lys	Asn	
			130				135					140				
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys	
145				150				155					160			
Glu	Asn	Asp	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Ala	Thr	
			165				170					175				
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln	
			180				185					190				
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp	
			195				200					205				
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Lys	Phe	
			210				215					220				
Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Leu	Gly	Gly	Ile	Asn	Thr	Val	

225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr			
	260	265	270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro			
	275	280	285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn			
	290	295	300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys			
305	310	315	320
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg			
	325	330	335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn			
	340	345	350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr			
	355	360	365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg			
	370	375	380
Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His			
385	390	395	400
Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn			
	405	410	415
Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys			
	420	425	430
Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp			
	435	440	445
<210> 65			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 65			
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr			
	20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe			
	35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr			

50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		95
	100	105
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		110
	115	120
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		125
	130	135
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		140
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		160
	165	170
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		175
	180	185
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		190
	195	200
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		205
	210	215
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		220
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		240
	245	250
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		255
	260	265
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		270
	275	280
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		285
	290	295
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		300
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		320
	325	330
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		335
	340	345
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		350
	355	360
		365

Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
 370 375 380
 Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 66
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 66
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Ala Ile Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Lys Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Ser Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
 165 170 175

Lys	Lys	Gln	Lys	Ser	Met	Lys	Ile	Ile	Leu	Asp	Arg	Leu	Met	Asn	Gln
180				185				190							
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp
195				200				205							
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Asn	Phe
210				215				220							
Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Gly	Ile	Asn	Thr	Val
225				230				235				240			
Tyr	Phe	Ala	Phe	Asn	Glu	Gly	Phe	Ile	Arg	Ser	Asn	Ile	Lys	Ser	Asp
245				250				255							
Glu	Ile	Lys	Ala	Phe	Asn	Glu	Lys	Ile	Arg	Gln	Arg	Arg	Ile	Asn	Leu
260				265				270							
Leu	Lys	Gln	Ser	Asn	Tyr	Cys	Ser	Asn	Ser	Arg	Thr	Gly	Lys	Gly	Arg
275				280				285							
Ala	Lys	Arg	Leu	Gln	Pro	Ile	Asp	Val	Leu	Ser	Asn	Lys	Ile	Ala	Lys
290				295				300							
Phe	Arg	Asn	Ser	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Ile	Lys	Gln
305				310				315				320			
Cys	Leu	Lys	His	Asn	Cys	Gly	Arg	Ile	Gln	Ile	Glu	Leu	Leu	Lys	Gly
325				330				335							
Ile	Ser	Lys	Asn	Asp	Lys	Val	Leu	Lys	Asp	Trp	Thr	Tyr	Phe	Asp	Leu
340				345				350							
Gln	Glu	Lys	Ile	Lys	Asn	Gln	Ala	Glu	Val	His	Gly	Ile	Glu	Val	Ile
355				360				365							
Lys	Val	Ala	Ser	Ala	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Gln	Cys	Gly	Tyr
370				375				380							
Ile	Cys	Lys	Glu	Asn	Arg	Cys	Thr	Gln	Ala	Val	Phe	Glu	Cys	Lys	Gln
385				390				395				400			
Cys	Gly	Tyr	Lys	Thr	His	Ala	Asp	Tyr	Asn	Ala	Ala	Lys	Asn	Ile	Ala
405				410				415							
Thr	Tyr	Asp	Ile	Glu	Asn	Ile	Ile	Asn	Lys	Gln	Leu	Ala	Val	Gln	Ser
420				425				430							
Lys	Leu	His	Ser	Lys	Lys	Cys	Met	Glu	Glu	Tyr	Ile	Glu	Glu	Leu	Gly
435				440				445							
Tyr	Leu	Asp													
450															
<210> 67															
<211> 451															
<212> PRT															

<213> 芽孢杆菌属

<400> 67

```

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

```

290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315 320
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395 400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Ala Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Arg		
435	440	445
Tyr Leu Asp		
450		
<210> 68		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 68		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1 5 10 15		
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
20 25 30		
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35 40 45		
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Ser Pro Thr		
50 55 60		
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65 70 75 80		
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
85 90 95		
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		

100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Met Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Leu Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Lys Gly Leu Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Phe His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Ile Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 69
 <211> 496
 <212> PRT
 <213> 芽孢杆菌属
 <400> 69
 Met Ile Leu Thr Arg Lys Val Lys Leu Val Ile Val Ser Asp Asn Arg
 1 5 10 15
 Asp Glu Gly Tyr Lys Leu Ile Arg Asn Glu Ile Arg Glu Gln His Lys
 20 25 30
 Ala Leu Asn Leu Ala Tyr Asn His Leu Tyr Phe Glu His Asn Ala Ile
 35 40 45
 Gln Ile Leu Lys Gln Asn Asp Glu Asp Tyr Lys Gln Lys Arg Asn Lys
 50 55 60
 Leu Gln Glu Leu Ile Asn Lys Lys Tyr Glu Glu His Gln Lys Ala Lys
 65 70 75 80
 Asn Leu Glu Arg Lys Glu Ala Leu Arg Glu Ala Tyr Asn Asn Lys Lys
 85 90 95
 Gln Glu Leu Tyr Lys Phe Glu Arg Glu Cys Asn Glu Glu Ala Arg Lys
 100 105 110
 Ala Tyr Gln Gln Val Val Gly Phe Thr Gln Gln Thr Arg Val Arg Asn
 115 120 125
 Leu Ile Asn Arg Glu Cys Asn Leu Met Ser Asp Thr Lys Asp Gly Ile
 130 135 140
 Thr Ser Lys Val Thr Gln Asp Tyr Lys Asn Asp Cys Lys Ala Gly Leu
 145 150 155 160
 Leu Ile Gly Lys Arg Ser Leu Arg Asn Tyr Lys Lys Asp Asn Pro Leu
 165 170 175
 Leu Val Arg Gly Arg Ser Leu Lys Phe Tyr Lys Glu Asp Gly Asp Tyr
 180 185 190
 Phe Ile Lys Trp Asn Lys Gly Thr Val Phe Lys Cys Ile Leu His Ile
 195 200 205
 Arg Lys Lys Asn Val Ala Glu Leu Gln Ser Val Leu Glu Asn Val Leu
 210 215 220

Leu Gly Ala Tyr Lys Ile Cys Asp Ser Ser Ile Gly Phe Asn Asn Lys
225 230 235 240
Asp Met Ile Leu Asn Leu Ser Leu Asn Ile Pro Asp Lys Glu Thr Tyr
245 250 255
Asp Tyr Ile Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile
260 265 270
Pro Ala Tyr Val Ser Leu Ser Asp Lys Val Tyr Val Arg Lys Gly Ile
275 280 285
Gly Gly Ile Asp Asp Phe Leu Arg Val Arg Thr Gln Met Gln Lys Arg
290 295 300
Arg Arg Gln Leu Gln Glu Ser Leu Ala Ala Val Lys Gly Gly Lys Gly
305 310 315 320
Arg Glu Lys Lys Leu Lys Ala Leu Asp His Leu Lys Gly Lys Glu Ala
325 330 335
Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr
340 345 350
Phe Ala Val Lys Asn Gln Ala Gly Gln Ile Asn Met Glu Phe Leu Glu
355 360 365
Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr
370 375 380
Gln Leu Gln Met Met Val Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile
385 390 395 400
Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys
405 410 415
Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Lys Gln Glu Lys Phe Ile Cys
420 425 430
Lys Ser Cys Ala Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn
435 440 445
Ile Ala Lys Ser Ala Arg Tyr Ile Ser Asp Ser Thr Glu Arg Glu Tyr
450 455 460
His Lys Lys Lys Gln Glu Asp Leu Lys Glu Ile Leu Gly Glu Asn Asp
465 470 475 480
Ile Ile Asn Glu Gln Leu Ser Leu Phe Asp Asn His Asp Asp Ile Ala
485 490 495
<210> 70
<211> 496
<212> PRT
<213> 芽孢杆菌属
<400> 70

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

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

<210> 71

<211> 496

<212> PRT

<213> 芽孢杆菌属

<400> 71

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

				85					90					95			
Gln	Glu	Leu	Tyr	Lys	Phe	Glu	Arg	Glu	Cys	Asn	Glu	Glu	Ala	Arg	Lys		
				100					105					110			
Ala	Tyr	Gln	Gln	Val	Val	Gly	Phe	Thr	Gln	Gln	Thr	Arg	Val	Arg	Asn		
				115					120					125			
Leu	Ile	Asn	Arg	Glu	Cys	Asn	Leu	Met	Ser	Asp	Thr	Lys	Asp	Gly	Ile		
				130					135					140			
Thr	Ser	Lys	Val	Thr	Gln	Asp	Tyr	Lys	Asn	Asp	Cys	Lys	Ala	Gly	Leu		
145						150				155				160			
Leu	Ile	Gly	Lys	Arg	Ser	Leu	Arg	Asn	Tyr	Lys	Lys	Asp	Asn	Pro	Leu		
				165					170					175			
Leu	Val	Arg	Gly	Arg	Ser	Leu	Lys	Phe	Tyr	Lys	Glu	Asp	Gly	Asp	Tyr		
				180					185					190			
Phe	Ile	Lys	Trp	Asn	Lys	Gly	Thr	Val	Phe	Lys	Cys	Ile	Leu	His	Ile		
				195					200					205			
Arg	Lys	Lys	Asn	Val	Ala	Glu	Leu	Gln	Ser	Val	Leu	Glu	Asn	Val	Leu		
				210					215					220			
Leu	Gly	Ala	Tyr	Lys	Ile	Cys	Asp	Ser	Ser	Ile	Gly	Phe	Asn	Asn	Lys		
225						230				235				240			
Asp	Met	Ile	Leu	Asn	Leu	Ser	Leu	Asn	Ile	Pro	Asp	Lys	Glu	Thr	Tyr		
				245					250					255			
Asp	Tyr	Ile	Pro	Gly	Arg	Val	Val	Gly	Val	Asp	Leu	Gly	Leu	Lys	Ile		
				260					265					270			
Pro	Ala	Tyr	Val	Ser	Leu	Ser	Asp	Lys	Val	Tyr	Val	Arg	Lys	Gly	Ile		
				275					280					285			
Gly	Gly	Ile	Asp	Asp	Phe	Leu	Arg	Val	Arg	Thr	Gln	Met	Gln	Lys	Arg		
				290					295					300			
Arg	Arg	Gln	Leu	Gln	Glu	Ser	Leu	Ala	Ala	Val	Lys	Gly	Gly	Lys	Gly		
305						310				315				320			
Arg	Glu	Lys	Lys	Leu	Lys	Ala	Leu	Asp	His	Leu	Lys	Gly	Lys	Glu	Ala		
				325					330					335			
Asn	Phe	Ala	Lys	Thr	Tyr	Asn	His	Phe	Leu	Ser	Thr	Gln	Ile	Val	Thr		
				340					345					350			
Phe	Ala	Val	Lys	Asn	Gln	Ala	Gly	Gln	Ile	Asn	Met	Glu	Phe	Leu	Glu		
				355					360					365			
Phe	Asp	Lys	Met	Lys	Asn	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr		
				370					375					380			
Gln	Leu	Gln	Met	Met	Val	Glu	Tyr	Lys	Ala	Lys	Arg	Glu	Gly	Ile	Ile		
385						390				395				400			

Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys		
	405	410 415
Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Lys Gln Glu Lys Phe Ile Cys		
	420	425 430
Lys Ser Cys Ala Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn		
	435	440 445
Ile Ala Lys Ser Ala Arg Tyr Ile Ser Asp Ser Thr Glu Arg Glu Tyr		
	450	455 460
His Lys Lys Lys Gln Glu Asp Leu Lys Glu Ile Leu Gly Glu Asn Asp		
465	470	475 480
Ile Ile Asn Glu Gln Leu Ser Leu Phe Asp Asn His Asp Asp Ile Ala		
	485	490 495
<210> 72		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 72		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1	5	10 15
Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr		
	20	25 30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
	35	40 45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
	50	55 60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75 80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
	85	90 95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
	100	105 110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
	115	120 125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		
	130	135 140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155 160
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		
	165	170 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile His Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 73		
<211> 451		
<212> PRT		

<213> 芽孢杆菌属

<400> 73

```

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

```

290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315 320
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Glu		
385	390	395 400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 74		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 74		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1 5 10 15		
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
20 25 30		
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35 40 45		
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50 55 60		
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65 70 75 80		
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
85 90 95		
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		

100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415

Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 75
 <211> 473
 <212> PRT
 <213> 芽孢杆菌属
 <400> 75
 Met Ile Ile Ala Arg Lys Ile Lys Leu Ile Ile Ile Gly Glu Asp Arg
 1 5 10 15
 Asp Thr Gln Tyr Lys Phe Ile Arg Glu Glu Arg Tyr Lys Gln Asn Lys
 20 25 30
 Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Leu His Val Ala Lys
 35 40 45
 Glu Lys Ile Arg Leu Leu Asp Asn Lys Phe Leu Gln Asp Glu Lys Lys
 50 55 60
 Leu Gln Glu Gly Ile Lys Lys Leu Tyr Ala Glu Lys Lys Val Ile Lys
 65 70 75 80
 Asp Gly Lys Lys Arg Asn Glu Leu Glu Lys Lys Ile Glu Lys Gln Thr
 85 90 95
 Asn Glu Leu Lys Lys Leu Arg Ser Lys Gly Asn Lys Glu Ala Asp Lys
 100 105 110
 Ile Leu Gln Glu Ala Ile Lys Ile Asn Leu Ser Ser Thr Thr Arg Glu
 115 120 125
 Val Ile Ser Lys Gln Phe Asp Leu Ile Ser Asp Thr Lys Asp Arg Ile
 130 135 140
 Thr Gln Lys Val Tyr Gln Asp Phe Lys Ser Asp Leu Lys Asn Gly Leu
 145 150 155 160
 Leu Ser Gly Glu Arg Val Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu
 165 170 175
 Leu Ile Arg Gly Arg Ala Leu Asn Phe Tyr Arg Glu Gly Lys Asp Val
 180 185 190
 Met Ile Lys Trp Phe Gly Gly Ile Ile Phe Lys Cys Met Leu Gly Gln
 195 200 205
 His Lys Asn Asn Ala Gln Glu Leu Lys Ala Thr Leu Asn Lys Val Leu
 210 215 220

Glu Gly Ser Tyr Lys Val Cys Asp Ser Ser Ile Ser Val Gly Lys Glu			
225	230	235	240
Leu Ile Leu Asn Ile Ser Leu Asp Ile Gly Glu Val Asn Ser Asn Val			
	245	250	255
Ser Cys Lys Lys Gly Arg Val Leu Gly Val Asp Leu Gly Met Lys Val			
	260	265	270
Pro Ala Tyr Met Ser Ile Asn Asp Lys Pro Tyr Ile Arg Lys Ser Leu			
	275	280	285
Gly Ser Leu Asp Asp Phe Leu Arg Ile Arg Val Gln Met Gln Lys Arg			
	290	295	300
Arg Arg Asn Leu His Lys Thr Leu Val Ser Val Lys Gly Gly Lys Gly			
305	310	315	320
Arg Glu Lys Lys Leu Gln Ala Leu Asp Arg Leu Lys Glu Lys Glu Lys			
	325	330	335
Asn Phe Ala Thr Thr Tyr Asn His Phe Leu Ser Tyr Asn Ile Val Lys			
	340	345	350
Phe Ala Lys Asp Asn Leu Ala Glu Gln Ile Asn Met Glu Phe Leu Ala			
	355	360	365
Leu Ala Gly Glu Asp Lys Asn Ile Ile Leu Arg Asn Trp Ser Tyr Tyr			
	370	375	380
Gln Leu Gln Gln Phe Val Glu Asp Lys Ala Lys Arg Glu Gly Ile Asp			
385	390	395	400
Val Lys Tyr Val Asp Pro Tyr Arg Thr Ser Gln Met Cys Ser Lys Cys			
	405	410	415
Arg Asn Tyr Glu Pro Gly Gln Arg Glu Ser Gln Glu Lys Phe Ile Cys			
	420	425	430
Lys Ser Cys His Leu Glu Ile Asn Ala Asp Tyr Asn Ala Ser Gln Asn			
	435	440	445
Ile Ala His Ser Thr Lys Tyr Ile Thr Asn Lys Asn Gln Ser Glu Tyr			
	450	455	460
Phe Lys Lys Leu Gln His Thr Thr Glu			
465	470		
<210> 76			
<211> 451			
<212> PRT			
<213> 芽孢杆菌属			
<400> 76			
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met			
1	5	10	15

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

	325		330		335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu					
	340		345		350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile					
	355		360		365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr					
	370		375		380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln					
385		390		395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser					
	405		410		415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser					
	420		425		430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly					
	435		440		445
Tyr Leu Asp					
450					
<210> 77					
<211> 451					
<212> PRT					
<213> 芽孢杆菌属					
<400> 77					
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met					
1	5		10		15
Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr					
	20		25		30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe					
	35		40		45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr					
50		55		60	
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile					
65		70		75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met					
	85		90		95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn					
	100		105		110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile					
	115		120		125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn					

130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val		
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445

Tyr Leu Asp

450

<210> 78

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 78

Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met

1 5 10 15

Asn Val Asp Trp Thr Ile Phe Glu Lys Gln Leu Arg Asn Leu Thr Tyr

20 25 30

Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe

35 40 45

Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr

50 55 60

Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile

65 70 75 80

Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met

85 90 95

Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn

100 105 110

Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile

115 120 125

Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn

130 135 140

Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys

145 150 155 160

Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr

165 170 175

Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln

180 185 190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp

195 200 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe

210 215 220

Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val

225 230 235 240

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

245 250 255

Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu		
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln		
305	310	315
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile		
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 79		
<211> 449		
<212> PRT		
<213> 芽孢杆菌属		
<400> 79		
Met Asp Lys Lys Ala Ser Lys Val Met Lys Tyr Glu Ile Ile Lys Pro		
1	5	10
Val Asp Val Asp Trp Asp Val Phe Gly Ser Val Leu Arg Glu Leu Gln		
20	25	30
Phe Glu Ser Lys Asn Leu Met Asn Lys Thr Ile Gln Leu Cys Trp Glu		
35	40	45
Trp Gln Gly Phe Ser Ser Asp Tyr Lys Lys Glu Asn Gly Glu Tyr Pro		
50	55	60

Lys	Leu	Glu	Asn	His	Thr	Lys	Tyr	Lys	Ala	Ile	Thr	Gly	Phe	Val	Tyr
65					70				75					80	
Asp	Arg	Met	Lys	Thr	Gln	Tyr	Asn	Lys	Phe	Asn	Thr	Gly	Asn	Met	Ser
			85						90					95	
Val	Ser	Ile	Lys	Ser	Ala	Thr	Asp	Lys	Trp	Lys	Ser	Asp	Leu	Lys	Asp
			100						105					110	
Ile	Leu	Arg	Gly	Glu	Lys	Ser	Ile	Pro	Ser	Tyr	Lys	Lys	Asn	Val	Pro
			115						120					125	
Ile	Asp	Ile	His	Gly	Asn	Ser	Ile	Lys	Ile	Lys	Glu	Val	Asn	Lys	Asn
			130						135					140	
Gly	Ala	Ile	Leu	Gln	Leu	Ser	Leu	Ile	Ser	Ser	Thr	Tyr	Lys	Lys	Glu
145					150					155					160
Leu	Gly	Ile	Lys	Asn	Gly	Phe	Phe	Asn	Val	Leu	Ile	Lys	Ile	Gly	Asp
					165				170						175
Asn	Ser	Gln	His	Glu	Ile	Val	Lys	Arg	Leu	Phe	Glu	Gly	Asp	Tyr	Lys
			180						185					190	
Ile	Ser	Ala	Ser	Lys	Ile	Val	Lys	His	Lys	Tyr	Lys	Asn	Lys	Trp	Phe
			195						200					205	
Leu	Asn	Leu	Thr	Tyr	Ser	Phe	Phe	Leu	Glu	Lys	Arg	Glu	Leu	Asn	Pro
			210						215					220	
Asp	Asn	Ile	Met	Gly	Ile	Asp	Val	Gly	Val	Val	Asn	Ala	Leu	Tyr	Met
225					230					235					240
Ala	Phe	Asn	Glu	Ser	Leu	Ser	Arg	Tyr	Ser	Ile	Glu	Gly	Gly	Glu	Ile
					245					250					255
Thr	Lys	Phe	Arg	Lys	Gly	Val	Glu	Ala	Arg	Arg	Lys	Ser	Leu	Leu	Arg
			260						265					270	
Gln	Gly	Lys	Tyr	Cys	Gly	Glu	Gly	Arg	Lys	Gly	Arg	Gly	Arg	Ala	Thr
			275						280					285	
Arg	Ile	Lys	Pro	Ile	Glu	Lys	Leu	Ser	Gln	Arg	Val	Asp	Asn	Phe	Lys
			290						295					300	
Asp	Ser	Cys	Asn	His	Lys	Tyr	Ser	Lys	Tyr	Val	Ile	Asp	Met	Ala	Leu
305					310					315					320
Lys	His	Asn	Cys	Gly	Thr	Ile	Gln	Met	Glu	Asp	Leu	Thr	Gly	Ile	Ala
					325					330					335
Glu	Gly	Glu	Lys	Lys	Ser	Ser	Phe	Leu	Gly	Asn	Trp	Thr	Tyr	Phe	Asp
			340						345					350	
Leu	Gln	Glu	Lys	Ile	Thr	Tyr	Lys	Ala	Lys	Glu	His	Gly	Ile	Lys	Val
			355						360					365	
Val	Lys	Ile	Lys	Pro	Lys	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Lys	Cys	Gly

370	375	380
Phe Ile Ser Lys Asp Ser Arg Pro Asp Gln Ala Thr Phe Glu Cys Ile		
385	390	395
Lys Cys Asn Phe Lys Thr Ser Ala Asp Tyr Asn Ala Ala Arg Asn Ile		400
405	410	415
Ala Met Lys Asp Ile Glu Lys Ile Ile Ala Glu Gln Leu Lys Val Gln		
420	425	430
Glu Lys Ala Lys Lys Leu Thr Lys Lys Gln Leu Ala Asn Ile Leu Asp		
435	440	445
Asp		
<210> 80		
<211> 451		
<212> PRT		
<213> 芽孢杆菌属		
<400> 80		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		
85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		
100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		
115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Ile Lys Glu Lys Asn		
130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		
145	150	155
Glu Asn Gly Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Gly Thr		
165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		
180	185	190

Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
195 200 205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe
210 215 220
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val
225 230 235 240
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp
245 250 255
Glu Ile Lys Met Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu
260 265 270
Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg
275 280 285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys
290 295 300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln
305 310 315 320
Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly
325 330 335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu
340 345 350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile Tyr Gly Ile Glu Val Ile
355 360 365
Lys Val Val Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
370 375 380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln
385 390 395 400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
405 410 415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Glu Gln Leu Ala Val Gln Ser
420 425 430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
435 440 445
Tyr Leu Asp
450
<210> 81
<211> 444
<212> PRT
<213> 芽孢杆菌属
<400> 81

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

305	310	315	320
Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Lys Val			
	325	330	335
Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln			
	340	345	350
Ala Glu Ile Tyr Gly Ile Glu Val Ile Lys Val Val Pro Ala Tyr Thr			
	355	360	365
Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys			
	370	375	380
Thr Gln Ala Met Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala			
385	390	395	400
Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile			
	405	410	415
Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys			
	420	425	430
Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp			
	435	440	
<210> 82			
<211> 448			
<212> PRT			
<213> 芽孢杆菌属			
<400> 82			
Met Thr Phe Leu Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu			
1	5	10	15
Cys Pro Leu Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn			
	20	25	30
Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu			
	35	40	45
Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr			
	50	55	60
Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp			
65	70	75	80
Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys			
	85	90	95
Gly Asn Met Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser			
	100	105	110
Arg Arg Asn Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg			
	115	120	125
Asn Arg Ile Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys			

130	135	140
Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp		
145	150	155
Phe His Lys Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys		
165	170	175
Leu Ala Thr Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu		
180	185	190
Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys		
195	200	205
Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu		
210	215	220
Asn Lys Phe Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile		
225	230	235
Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile		
245	250	255
Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln		
260	265	270
Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg		
275	280	285
Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn		
290	295	300
Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys		
305	310	315
His Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys		
325	330	335
Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys		
340	345	350
Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala		
355	360	365
Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys		
370	375	380
Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr		
385	390	395
Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp		
405	410	415
Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His		
420	425	430
Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp		
435	440	445

<210> 83

<211> 444

<212> PRT

<213> 芽孢杆菌属

<400> 83

```

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

```

Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile
 275 280 285
 Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn His
 290 295 300
 Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys Gly
 305 310 315 320
 Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile
 325 330 335
 Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln
 340 345 350
 Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr Thr
 355 360 365
 Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys
 370 375 380
 Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His Ala
 385 390 395 400
 Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile
 405 410 415
 Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys Cys
 420 425 430
 Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
 435 440
 <210> 84
 <211> 448
 <212> PRT
 <213> 芽孢杆菌属
 <400> 84
 Met Thr Phe Leu Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu
 1 5 10 15
 Cys Pro Leu Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn
 20 25 30
 Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu
 35 40 45
 Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr
 50 55 60
 Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp
 65 70 75 80
 Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys
 85 90 95

Gly	Asn	Met	Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	
Arg	Arg	Asn	Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	
Asn	Arg	Ile	Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Thr	Lys	
Glu	Lys	Asn	Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	
Phe	His	Lys	Glu	Asn	Asp	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	
Leu	Ala	Thr	Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	
Ile	Asn	Gln	Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	
Asn	Lys	Trp	Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	
Asn	Lys	Phe	Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Leu	Gly	Gly	Ile	
Asn	Thr	Val	Tyr	Ser	Ala	Phe	Asn	Glu	Gly	Phe	Ile	Arg	Ser	Asn	Ile	
Lys	Ser	Asp	Glu	Ile	Ile	Arg	Gln	Arg	Arg	Ile	Asn	Leu	Leu	Lys	Gln	
Ser	Lys	Tyr	Cys	Ser	Asn	Ser	Arg	Thr	Gly	Lys	Gly	Arg	Thr	Lys	Arg	
Leu	Gln	Pro	Ile	Asp	Val	Leu	Ser	Asn	Lys	Ile	Ala	Lys	Phe	Arg	Asn	
Ser	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Val	Lys	Gln	Cys	Leu	Lys	
His	Asn	Cys	Gly	Arg	Ile	Gln	Met	Glu	Leu	Leu	Lys	Gly	Ile	Ser	Lys	
Asn	Asp	Arg	Ile	Leu	Lys	Asp	Trp	Thr	Tyr	Phe	Asp	Leu	Gln	Glu	Lys	
Ile	Lys	Asn	Gln	Ala	Glu	Ile	His	Gly	Ile	Glu	Val	Ile	Lys	Val	Ala	
Pro	Ala	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Gln	Cys	Gly	Tyr	Ile	Cys	Lys	
Glu	Asn	Arg	Cys	Thr	Gln	Ala	Thr	Phe	Glu	Cys	Lys	Gln	Cys	Gly	Tyr	
Lys	Thr	His	Ala	Asp	Tyr	Asn	Ala	Ala	Lys	Asn	Ile	Ser	Thr	Tyr	Asp	

	405		410		415
Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His					
	420		425		430
Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp					
	435		440		445
<210> 85					
<211> 486					
<212> PRT					
<213> 芽孢杆菌属					
<400> 85					
Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ser					
1	5		10		15
Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys					
	20		25		30
Ala Leu Asn Val Ser Met Asn His Leu Tyr Phe Asn Ser Ile Ala Arg					
	35		40		45
Gln Lys Ile Leu Leu Ala Asp Glu Ala Tyr Gln Gln Lys Leu Asn Asp					
	50		55		60
Ala Ile Asn Ser Gln Glu Lys Ser Phe Ala Ala Leu Lys Glu Ile Glu					
65	70		75		80
Gln Lys Gln Cys Ile Glu Asp Ser Glu Lys Lys Gln Val Leu Lys Glu					
	85		90		95
Arg Leu Ile Lys Thr Lys Asn Ala Tyr Glu Lys Ala Lys Glu Lys Val					
	100		105		110
Asn Asn Leu Arg Lys Ser Arg Ser Lys Asp Ser Phe Gln Glu Tyr Lys					
	115		120		125
Asn Ile Ile Gly Gln Val Glu Gln Thr His Leu Arg Asp Ile Ile Ser					
	130		135		140
Ser Gln Phe Asn Leu His Ser Asp Thr Lys Asp Arg Leu Thr Met Ile					
145	150		155		160
Thr Asn Gln Asp Phe Lys Asn Asp Ile Ala Glu Val Leu Ser Gly Asp					
	165		170		175
Arg Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg Ser					
	180		185		190
Arg Asn Thr Val Leu Tyr Lys Glu Gly Asn Glu Phe Phe Ile Lys Trp					
	195		200		205
Ile Lys Gly Ile Val Phe Lys Cys Ile Leu Gly Val Lys Asn Gln Asn					
	210		215		220
Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Val Leu Ala Gly Ile Tyr					

225	230	235	240
Lys Ile Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu Ile			
	245	250	255
Leu Asn Leu Thr Leu Asp Met Pro Asp Lys Ser Glu His Arg Lys Val			
260	265	270	
Pro Glu Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala Tyr			
275	280	285	
Phe Ala Val Asn Asp Val Ser Tyr Ile Arg Gln Ala Ile Gly Lys Ile			
290	295	300	
Glu Asp Phe Leu Lys Val Arg Thr Ser Ile Gln Ser Gln Lys Arg Ser			
305	310	315	320
Leu Glu His Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys Lys			
325	330	335	
Lys Leu Lys Ala Leu Gly Gln Phe Lys Glu Lys Glu Lys Ser Tyr Ile			
340	345	350	
Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala Ile			
355	360	365	
Gln Tyr Gly Val Val Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys Glu			
370	375	380	
Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu Gln			
385	390	395	400
Gln Phe Ile Glu Tyr Lys Ala Glu Arg Ala Gly Ile Leu Val Lys Tyr			
405	410	415	
Val Asp Pro Phe His Thr Ser Gln Thr Cys Ser Lys Cys Gly His Tyr			
420	425	430	
Glu Asp Gly Gln Arg Glu Lys Gln Asp Thr Phe Cys Cys Lys Ser Cys			
435	440	445	
Gly Phe Thr Glu Asn Ala Asp Tyr Asn Ala Ala Arg Asn Ile Ala Ala			
450	455	460	
Ser Thr Arg Tyr Ile Thr Asn Lys Glu Glu Ser Glu Tyr Tyr Arg Lys			
465	470	475	480
Asn Asn Asn Glu Ile Ala			
485			

<210> 86

<211> 264

<212> PRT

<213> 芽孢杆菌属

<400> 86

Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly

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

<210> 87

<211> 438

<212> PRT

<213> 芽孢杆菌属

<400> 87

Met Lys Tyr Gln Ile Leu Cys Pro Leu Asn Val Asp Trp Thr Ile Phe

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

Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr
 325 330 335
 Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile
 340 345 350
 Glu Val Ile Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln
 355 360 365
 Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu
 370 375 380
 Cys Lys Gln Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys
 385 390 395 400
 Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala
 405 410 415
 Val Gln Ser Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu
 420 425 430
 Glu Leu Gly Tyr Leu Asp
 435
 <210> 88
 <211> 450
 <212> PRT
 <213> 芽孢杆菌属
 <400> 88
 Met Ser Thr Val Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
 1 5 10 15
 Asn Ile Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Leu Trp Glu Phe
 35 40 45
 Asp Asn Gln Ser Leu Asn His Phe Lys Glu Asn Gly Val Tyr Pro Ser
 50 55 60
 Thr Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
 65 70 75 80
 Tyr Asp Gln Leu Lys Phe Glu Tyr Leu Asp Ile Asn Lys Ala Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Thr Ile Lys Thr Trp Asn Ser Arg Lys Lys
 100 105 110
 Glu Ile Trp Ser Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Asn Leu
 115 120 125
 Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Thr Lys Glu Lys Ser
 130 135 140

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

450
 <210> 89
 <211> 487
 <212> PRT
 <213> 芽孢杆菌属
 <400> 89
 Met Ile Thr Ser Arg Lys Ile Arg Leu Ser Ile Val Ser Asp Asn Ser
 1 5 10 15
 Thr Glu Ala Tyr Asn Phe Ile Arg Lys Glu Met Arg Glu Gln Asn Lys
 20 25 30
 Ala Leu Asn Val Ala Met Asn His Leu Tyr Phe Asn Tyr Val Ala Arg
 35 40 45
 Gln Lys Ile Ser Phe Ala Asp Lys Ala Tyr Gln His Lys Leu Asn Lys
 50 55 60
 Ala Ile Glu Ala Gln Gln Lys Ser Phe Ile Leu Leu Lys Glu Leu Glu
 65 70 75 80
 Gln Lys Gln Ile Val Glu Thr Asp Ser Ser Lys Lys Ser Ile Leu Lys
 85 90 95
 Glu Arg Leu Ile Lys Ser Lys Thr Thr Tyr Glu Lys Thr Lys Glu Lys
 100 105 110
 Leu Ser Ser Leu Arg Lys Ala Arg Asn Lys Glu Leu Phe Gln Glu Tyr
 115 120 125
 Asn Asn Met Ile Gly Gln Leu Glu Asp Thr His Leu Arg Asp Ile Val
 130 135 140
 Ser Ser Gln Phe Asn Leu Leu Ser Asp Thr Lys Asp Arg Leu Thr Lys
 145 150 155 160
 Ile Ala Tyr Gln Asp Phe Lys Asn Asp Ile Thr Glu Val Leu Ser Gly
 165 170 175
 Asn Arg Ser Leu Arg Thr Tyr Lys Lys Asn Asn Pro Leu Tyr Ile Arg
 180 185 190
 Gly Arg Thr Leu Asn Leu Phe Lys Glu Ser Glu Glu Phe Phe Ile Lys
 195 200 205
 Trp Thr Lys Lys Ile Val Phe Lys Cys Val Leu Gly Val Lys Tyr Gln
 210 215 220
 Asn Lys Thr Glu Leu Tyr Lys Thr Leu Glu Cys Ile Leu Thr Gly Glu
 225 230 235 240
 Tyr Glu Leu Cys Asp Ser Ser Met Asn Phe Asn Gln Gln Asn Lys Leu
 245 250 255
 Ile Leu Asn Leu Ala Leu Asn Ile Pro Glu Lys Ser Glu Asn Arg Lys

260	265	270
Val Pro Gly Arg Ile Ala Gly Ile Asp Leu Gly Leu Lys Ile Pro Ala		
275	280	285
Tyr Phe Ala Val Asn Asp Val Pro Tyr Ile Arg Lys Ala Leu Gly Lys		
290	295	300
Ile Glu Asp Phe Leu Lys Val Arg Thr Asn Ile Gln Ser Gln Lys Arg		
305	310	315
Ser Leu Gln Arg Ala Leu Gln Ser Ser Lys Gly Gly Lys Gly Arg Lys		
325	330	335
Lys Lys Leu Lys Ala Leu Asp Gln Phe Lys Glu Lys Glu Lys Asn Tyr		
340	345	350
Ile Thr Thr Tyr Asn His Phe Ile Ser Lys Lys Ile Ile Ser Leu Ala		
355	360	365
Val Gln Tyr Gly Val Glu Gln Ile Asn Leu Glu Leu Leu Thr Leu Lys		
370	375	380
Glu Thr Gln Lys Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
385	390	395
Gln Gln Phe Ile Glu Tyr Lys Ala Asn Arg Glu Glu Ile Leu Ile Lys		
405	410	415
Tyr Val Asp Pro Phe His Thr Ser Gln Thr Cys Ser Lys Cys Gly His		
420	425	430
Tyr Glu Asp Gly Gln Arg Glu Lys Gln Asp Thr Phe Cys Cys Lys Ser		
435	440	445
Cys Gly Phe Ile Asp Asn Ala Asp Tyr Asn Ala Ala Arg Asn Ile Ala		
450	455	460
Ala Ser Thr Arg Tyr Ile Thr Asn Lys Glu Glu Ser Glu Tyr Tyr Arg		
465	470	475
Lys Asn Asn Asn Glu Ile Ala		
485		

<210> 90

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 90

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

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

Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile
 355 360 365
 Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr
 370 375 380
 Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln
 385 390 395 400
 Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser
 405 410 415
 Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser
 420 425 430
 Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly
 435 440 445
 Tyr Leu Asp
 450
 <210> 91
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 91
 Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met
 1 5 10 15
 Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn
 100 105 110
 Glu Ile Arg Lys Gly Lys Met Ser Ile Pro Ser Phe Arg Asn Arg Ile
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160

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

<210> 92

<211> 496

<212> PRT

<213> 芽孢杆菌属

<400> 92

```

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

```

275	280	285
Gly Ser Ile Asp Asp Phe Leu Arg Val Arg Thr Gln Met Gln Lys Arg		
290	295	300
Arg Arg Gln Leu Gln Glu Ser Leu Ser Ala Val Lys Gly Gly Lys Gly		
305	310	315
Arg Lys Lys Lys Leu Lys Ala Leu Glu His Leu Lys Glu Lys Glu Ala		
325	330	335
Asn Phe Ala Lys Thr Tyr Asn His Phe Leu Ser Thr Gln Ile Val Thr		
340	345	350
Phe Ala Val Lys Asn His Ala Gly Gln Ile Asn Met Glu Phe Leu Glu		
355	360	365
Phe Asp Lys Met Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr		
370	375	380
Gln Leu Gln Met Met Ile Glu Tyr Lys Ala Lys Arg Glu Gly Ile Ile		
385	390	395
Ile Lys Tyr Val Asp Ala Tyr Leu Thr Ser Gln Thr Cys Ser Lys Cys		
405	410	415
Asp Tyr Tyr Glu Glu Gly Gln Arg Glu Thr Gln Glu Arg Phe Met Cys		
420	425	430
Lys Ser Cys Gly Phe Glu Val Asn Ala Asp Tyr Asn Ala Ser Gln Asn		
435	440	445
Ile Ala Lys Ser Thr Arg Tyr Ile Ser Asp Ser Thr Glu Ser Glu Tyr		
450	455	460
His Lys Lys Lys Gln Glu Ala Leu Lys Gly Ile Leu Gly Glu Asn Asp		
465	470	475
Thr Ile Asn Glu Gln Leu Ser Leu Phe Asn Asn Cys Asp Asp Ile Ala		
485	490	495

<210> 93

<211> 445

<212> PRT

<213> 芽孢杆菌属

<400> 93

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

50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		
65	70	75
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met		80
	85	90
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn		95
	100	105
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile		110
	115	120
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn		125
	130	135
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys		140
145	150	155
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr		160
	165	170
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln		175
	180	185
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp		190
	195	200
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe		205
	210	215
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val		220
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp		240
	245	250
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr		255
	260	265
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro		270
	275	280
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn		285
	290	295
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys		300
305	310	315
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg		320
	325	330
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn		335
	340	345
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr		350
	355	360
		365

Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg
 370 375 380
 Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
 385 390 395 400
 Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
 405 410 415
 Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
 420 425 430
 Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
 435 440 445
 <210> 94
 <211> 506
 <212> PRT
 <213> 芽孢杆菌属
 <400> 94
 Met Ser Thr Pro Leu Gln Gln Pro His Gln Lys Ser Lys Lys Thr Ser
 1 5 10 15
 Gln Met Ile Thr Thr Arg Lys Phe Lys Leu Ala Ile Val Ser Asp Asn
 20 25 30
 Arg Asn Glu Ala Tyr Ser Phe Ile Arg Asn Glu Ile Arg Asn Gln Asn
 35 40 45
 Lys Ala Leu Asn Ala Ala Tyr Asn His Leu Tyr Phe Glu His Ile Ala
 50 55 60
 Thr Glu Lys Leu Lys His Ser Asp Glu Glu Tyr Gln Lys His Leu Thr
 65 70 75 80
 Lys Tyr Arg Glu Val Ser Thr Asn Lys Tyr Gln Asp Tyr Leu Lys Val
 85 90 95
 Lys Glu Lys Val Asn Ala Ser Lys Asp Asp Glu Lys Leu Gln Lys Arg
 100 105 110
 Val Asp Lys Ala Arg Glu Ala Tyr Asn Lys Ala Gln Glu Lys Val Tyr
 115 120 125
 Lys Ile Glu Lys Glu Phe Asn Lys Lys Ser Met Glu Thr Tyr Gln Lys
 130 135 140
 Val Val Gly Leu Ser Lys Gln Thr Arg Ile Gly Lys Leu Leu Lys Ser
 145 150 155 160
 Gln Phe Thr Leu His Tyr Asp Thr Glu Asp Arg Ile Thr Ser Thr Val
 165 170 175
 Leu Ser His Phe Asn Asn Asp Met Lys Thr Gly Val Leu Arg Gly Asp
 180 185 190

Arg Ser Leu Arg Thr Tyr Lys Asn Ser His Pro Leu Leu Val Arg Ala		
195	200	205
Arg Ser Met Lys Val Tyr Glu Glu Asn Gly Asp Tyr Phe Ile Lys Trp		
210	215	220
Val Lys Gly Ile Val Phe Lys Ile Val Ile Ser Ala Gly Ser Lys Gln		
225	230	235
Lys Ala Asn Ile Gly Glu Leu Lys Ser Val Leu Ile Asn Ile Leu Asn		
	245	250
Gly His Tyr Lys Val Cys Asp Ser Ser Ile Ser Leu Asn Lys Asp Leu		
	260	265
Ile Leu Asn Leu Ser Leu Asn Ile Pro Val Ser Lys Glu Asn Val Phe		
	275	280
Val Pro Gly Arg Val Val Gly Val Asp Leu Gly Leu Lys Ile Pro Ala		
	290	295
Tyr Val Ser Leu Asn Asp Thr Pro Tyr Ile Lys Lys Gly Ile Gly Asn		
305	310	315
Ile Asp Asp Phe Leu Asn Val Arg Thr Gln Leu Gln Asn Gln Arg Lys		
	325	330
Arg Leu Gln Lys Thr Leu Glu Cys Thr Ser Gly Gly Lys Gly Arg Ser		
	340	345
Lys Lys Leu Lys Gly Leu Asp Arg Leu Lys Ala Lys Glu Lys Asn Phe		
	355	360
Val Asn Thr Tyr Asn His Phe Leu Ser Lys Lys Ile Ile Gln Phe Ala		
	370	375
Val Lys Asn Asn Ala Gly Val Ile His Leu Glu Glu Leu Gln Phe Asp		
385	390	395
Lys Leu Lys Asn Lys Ser Leu Leu Arg Asn Trp Ser Tyr Tyr Gln Leu		
	405	410
Gln Thr Met Ile Glu Tyr Lys Ala Glu Arg Glu Gly Ile Glu Val Lys		
	420	425
Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His		
	435	440
Tyr Glu Lys Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn		
	450	455
Lys Glu Cys Lys Gly Tyr Ala His Lys Val Asn Ala Asp Phe Asn Ala		
465	470	475
Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met		
	485	490
Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala		495

500	505
<210> 95	
<211> 451	
<212> PRT	
<213> 芽孢杆菌属	
<400> 95	
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Met	
1 5 10 15	
Asn Val Asp Trp Ile Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr	
20 25 30	
Gln Ile Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe	
35 40 45	
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr	
50 55 60	
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile	
65 70 75 80	
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met	
85 90 95	
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn	
100 105 110	
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile	
115 120 125	
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn	
130 135 140	
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys	
145 150 155 160	
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr	
165 170 175	
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln	
180 185 190	
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp	
195 200 205	
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Lys Lys Phe	
210 215 220	
Asp Lys Glu Leu Ile Met Gly Ile Asp Met Gly Gly Ile Asn Thr Val	
225 230 235 240	
Tyr Phe Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp	
245 250 255	
Glu Ile Lys Glu Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu	

260	265	270
Leu Asn Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg		
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys Ile Ala Lys		
290	295	300
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile Ile Lys Gln		
305	310	315
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Met Glu Leu Leu Lys Gly		
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu Lys Asp Trp Thr Tyr Phe Asp Leu		
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile Glu Val Ile		
355	360	365
Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr		
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Met Phe Glu Cys Lys Gln		
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ala		
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Ala Gln Ser		
420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly		
435	440	445
Tyr Leu Asp		
450		
<210> 96		
<211> 445		
<212> PRT		
<213> 芽孢杆菌属		
<400> 96		
Met Gly Val Thr Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Leu		
1	5	10
Asn Val Asp Trp Thr Ile Phe Glu Lys His Leu Arg Asn Leu Thr Tyr		
20	25	30
Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe		
35	40	45
Asp Ala Leu Ser Phe Asp Tyr Phe Lys Glu Arg Gly Thr Tyr Pro Thr		
50	55	60
Val Gln Asp Leu Tyr Gly Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile		

65	70	75	80
Tyr His Thr Leu Gln Ser Lys Tyr Pro Asp Ile His Lys Gly Asn Met			
	85	90	95
Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Asn			
	100	105	110
Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Arg Ile			
	115	120	125
Pro Ile Asp Leu His Asn Asn Ser Val Asp Ile Thr Lys Glu Lys Asn			
	130	135	140
Gly Asp Tyr Ile Ala Gly Ile Ser Leu Phe Ser Arg Asp Phe His Lys			
145	150	155	160
Glu Asn Asp Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr			
	165	170	175
Gln Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Gln			
	180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp			
	195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ala Ile Lys Glu Asn Lys Phe			
210	215	220	
Asp Lys Glu Leu Ile Met Gly Ile Asp Leu Gly Gly Ile Asn Thr Val			
225	230	235	240
Tyr Ser Ala Phe Asn Glu Gly Phe Ile Arg Ser Asn Ile Lys Ser Asp			
	245	250	255
Glu Ile Ile Arg Gln Arg Arg Ile Asn Leu Leu Lys Gln Ser Lys Tyr			
	260	265	270
Cys Ser Asn Ser Arg Thr Gly Lys Gly Arg Thr Lys Arg Leu Gln Pro			
	275	280	285
Ile Asp Val Leu Ser Asn Lys Ile Ala Lys Phe Arg Asn Ser Thr Asn			
	290	295	300
His Lys Tyr Ala Asn Tyr Ile Val Lys Gln Cys Leu Lys His Asn Cys			
305	310	315	320
Gly Arg Ile Gln Met Glu Leu Leu Lys Gly Ile Ser Lys Asn Asp Arg			
	325	330	335
Ile Leu Lys Asp Trp Thr Tyr Phe Asp Leu Gln Glu Lys Ile Lys Asn			
	340	345	350
Gln Ala Glu Ile His Gly Ile Glu Val Ile Lys Val Ala Pro Ala Tyr			
	355	360	365
Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr Ile Cys Lys Glu Asn Arg			
370	375	380	

Cys Thr Gln Ala Thr Phe Glu Cys Lys Gln Cys Gly Tyr Lys Thr His
 385 390 395 400
 Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Thr Tyr Asp Ile Glu Asn
 405 410 415
 Ile Ile Asn Lys Gln Leu Ala Val Gln Ser Lys Leu His Ser Lys Lys
 420 425 430
 Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly Tyr Leu Asp
 435 440 445
 <210> 97
 <211> 438
 <212> PRT
 <213> 芽孢杆菌属
 <400> 97
 Met Lys Tyr Gln Ile Leu Cys Pro Leu Asn Val Asp Trp Thr Ile Phe
 1 5 10 15
 Glu Lys His Leu Arg Asn Leu Thr Tyr Gln Val Arg Thr Ile Ser Asn
 20 25 30
 Arg Thr Ile Gln Gln Leu Trp Glu Phe Asp Ala Leu Ser Phe Asp Tyr
 35 40 45
 Phe Lys Glu Arg Gly Thr Tyr Pro Thr Val Gln Asp Leu Tyr Gly Cys
 50 55 60
 Thr Gln Lys Lys Ile Asp Gly Tyr Ile Tyr His Thr Leu Gln Ser Lys
 65 70 75 80
 Tyr Pro Asp Ile His Lys Gly Asn Met Ser Thr Thr Leu Gln Lys Ile
 85 90 95
 Ile Lys Thr Trp Lys Ser Arg Arg Asn Glu Ile Arg Lys Gly Glu Met
 100 105 110
 Ser Ile Pro Ser Phe Arg Asn Arg Ile Pro Ile Asp Leu His Asn Asn
 115 120 125
 Ser Val Asp Ile Thr Lys Glu Lys Asn Gly Asp Tyr Ile Ala Gly Ile
 130 135 140
 Ser Leu Phe Ser Arg Asp Phe His Lys Glu Asn Asp Asp Val Pro Lys
 145 150 155 160
 Gly Lys Ile Phe Val Lys Leu Ala Thr Gln Lys Gln Lys Ser Met Lys
 165 170 175
 Val Ile Leu Asp Arg Leu Ile Asn Gln Thr Tyr Ser Lys Gly Ala Cys
 180 185 190
 Met Ile His Lys Tyr Lys Asn Lys Trp Tyr Leu Ser Ile Thr Tyr Lys
 195 200 205

Phe Asn Ala Ile Lys Glu Asn Lys Phe Asp Lys Glu Leu Ile Met Gly			
210	215	220	
Ile Asp Leu Gly Gly Ile Asn Thr Val Tyr Ser Ala Phe Asn Glu Gly			
225	230	235	240
Phe Ile Arg Ser Asn Ile Lys Ser Asp Glu Ile Ile Arg Gln Arg Arg			
	245	250	255
Ile Asn Leu Leu Lys Gln Ser Lys Tyr Cys Ser Asn Ser Arg Thr Gly			
	260	265	270
Lys Gly Arg Thr Lys Arg Leu Gln Pro Ile Asp Val Leu Ser Asn Lys			
	275	280	285
Ile Ala Lys Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Tyr Ile			
	290	295	300
Val Lys Gln Cys Leu Lys His Asn Cys Gly Arg Ile Gln Met Glu Leu			
305	310	315	320
Leu Lys Gly Ile Ser Lys Asn Asp Arg Ile Leu Lys Asp Trp Thr Tyr			
	325	330	335
Phe Asp Leu Gln Glu Lys Ile Lys Asn Gln Ala Glu Ile His Gly Ile			
	340	345	350
Glu Val Ile Lys Val Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln			
	355	360	365
Cys Gly Tyr Ile Cys Lys Glu Asn Arg Cys Thr Gln Ala Thr Phe Glu			
	370	375	380
Cys Lys Gln Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys			
385	390	395	400
Asn Ile Ser Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala			
	405	410	415
Val Gln Ser Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu			
	420	425	430
Glu Leu Gly Tyr Leu Asp			
435			
<210> 98			
<211> 276			
<212> PRT			
<213> 芽孢杆菌属			
<400> 98			
Met Pro Gln Phe Lys Ser Lys Lys Asn Lys Ile Lys Ser Tyr Thr Thr			
1	5	10	15
Lys Cys Thr Asn His Asn Ile Ala Ile Val Gly His Thr Ile Lys Leu			
	20	25	30

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

<210> 99
 <211> 405
 <212> PRT
 <213> 芽孢杆菌属
 <400> 99

Met	Ala	Arg	Thr	Thr	Val	Lys	Lys	Arg	Asn	Glu	Glu	Leu	Ala	Lys	Glu
1					5				10					15	

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

				325					330					335		
Lys	Ala	Asn	Trp	Ala	Asn	Lys	Leu	Leu	Ile	Leu	Ile	Asp	Arg	Phe	Phe	
			340					345					350			
Pro	Ser	Ser	Lys	Gln	Cys	Asn	Gly	Cys	Lys	Glu	Thr	Asn	Pro	Leu	Leu	
		355					360					365				
Ser	Leu	Ser	Asp	Arg	Val	Trp	Met	Cys	Pro	Ser	Cys	Gly	Thr	His	His	
	370					375					380					
Asp	Arg	Asp	Asp	Asn	Ala	Ala	Phe	Asn	Ile	Lys	Glu	Glu	Gly	Ile	Arg	
385					390					395					400	
Leu	Leu	Leu	Asn	Ala												
			405													
<210> 100																
<211> 451																
<212> PRT																
<213> 芽孢杆菌属																
<400> 100																
Met	Gly	Val	Thr	Ile	Lys	Ile	Met	Lys	Tyr	Gln	Ile	Leu	Cys	Pro	Met	
1				5					10					15		
Asn	Val	Asp	Trp	Thr	Ile	Phe	Glu	Lys	Gln	Leu	Arg	Asn	Leu	Thr	Tyr	
			20					25					30			
Gln	Val	Arg	Thr	Ile	Ser	Asn	Arg	Thr	Ile	Gln	Gln	Leu	Trp	Glu	Phe	
	35						40					45				
Asp	Ala	Leu	Ser	Phe	Asp	Tyr	Phe	Lys	Glu	Arg	Gly	Thr	Tyr	Pro	Thr	
	50					55					60					
Val	Gln	Asp	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Lys	Ile	Asp	Gly	Tyr	Ile	
65				70					75					80		
Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met	
			85					90					95			
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn	
			100					105					110			
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile	
	115						120					125				
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Ile	Lys	Glu	Lys	Asn	
	130					135					140					
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys	
145				150					155					160		
Glu	Asn	Gly	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Gly	Thr	
			165					170					175			
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln	

180	185	190
Thr Tyr Ser Lys Gly Ala Cys Met	Ile His Lys Tyr Lys Asn Lys Trp	
195	200	205
Tyr Leu Ser Ile Thr Tyr Lys Phe	Asn Ala Ile Lys Glu Asn Lys Phe	
210	215	220
Asp Lys Glu Leu Ile Met Gly Ile	Asp Met Gly Gly Ile Asn Thr Val	
225	230	235
Tyr Phe Ala Phe Asn Glu Gly Phe	Ile Arg Ser Asn Ile Lys Ser Asp	
245	250	255
Glu Ile Lys Met Phe Asn Glu Arg	Ile Arg Gln Arg Arg Ile Asn Leu	
260	265	270
Leu Lys Gln Ser Lys Tyr Cys Ser	Asn Ser Arg Thr Gly Lys Gly Arg	
275	280	285
Thr Lys Arg Leu Gln Pro Ile Asp	Val Leu Ser Asn Lys Ile Ala Lys	
290	295	300
Phe Arg Asn Ser Thr Asn His Lys	Tyr Ala Asn Tyr Ile Val Lys Gln	
305	310	315
Cys Leu Lys His Asn Cys Gly Arg	Ile Gln Met Glu Leu Leu Lys Gly	
325	330	335
Ile Ser Lys Asn Asp Lys Val Leu	Lys Asp Trp Thr Tyr Phe Asp Leu	
340	345	350
Gln Glu Lys Ile Lys Asn Gln Ala	Glu Ile Tyr Gly Ile Glu Val Ile	
355	360	365
Lys Val Val Pro Ala Tyr Thr Ser	Gln Arg Cys Ser Gln Cys Gly Tyr	
370	375	380
Ile Cys Lys Glu Asn Arg Cys Thr	Gln Ala Met Phe Glu Cys Lys Gln	
385	390	395
Cys Gly Tyr Lys Thr His Ala Asp	Tyr Asn Ala Ala Lys Asn Ile Ser	
405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile	Asn Lys Gln Leu Ala Val Gln Ser	
420	425	430
Lys Leu His Ser Lys Lys Cys Met	Glu Glu Tyr Ile Glu Glu Leu Gly	
435	440	445
Tyr Leu Asp		
450		
<210> 101		
<211> 487		
<212> PRT		
<213> 芽孢杆菌属		

<400> 101

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

Ile	Glu	Asp	Phe	Leu	Lys	Val	Arg	Thr	Asn	Ile	His	Ser	Gln	Gln	Arg
305					310					315					320
Ser	Leu	Gln	Arg	Ala	Leu	Gln	Ser	Ser	Lys	Gly	Gly	Lys	Gly	Arg	Lys
				325					330					335	
Lys	Lys	Leu	Lys	Ala	Leu	Asp	Gln	Phe	Lys	Glu	Lys	Glu	Lys	Asn	Tyr
				340				345						350	
Ile	Thr	Thr	Tyr	Asn	His	Phe	Ile	Ser	Lys	Lys	Ile	Ile	Thr	Leu	Ala
				355				360						365	
Ile	Arg	Tyr	Gly	Val	Glu	Gln	Ile	Asn	Leu	Glu	Leu	Leu	Thr	Leu	Lys
				370				375						380	
Asp	Thr	Gln	Lys	Lys	Ser	Leu	Leu	Arg	Asn	Trp	Ser	Tyr	Tyr	Gln	Leu
385					390					395					400
Gln	Gln	Phe	Ile	Glu	Tyr	Lys	Ala	Asn	Arg	Glu	Gly	Ile	Leu	Ile	Lys
				405						410					415
Tyr	Val	Asp	Pro	Phe	His	Thr	Ser	Gln	Thr	Cys	Ser	Lys	Cys	Ser	His
				420						425					430
Tyr	Glu	Asp	Gly	Gln	Arg	Glu	Lys	Gln	Asp	Thr	Phe	Cys	Cys	Lys	Ser
				435						440					445
Cys	Gly	Phe	Lys	Asp	Asn	Ala	Asp	Tyr	Asn	Ala	Ala	Arg	Asn	Ile	Ala
				450						455					460
Ala	Ser	Thr	Lys	Tyr	Ile	Ile	Lys	Lys	Glu	Glu	Ser	Glu	Tyr	Tyr	Arg
465					470					475					480
Gln	Asn	Asn	Asn	Glu	Ile	Ala									
															485

<210> 102

<211> 451

<212> PRT

<213> 芽孢杆菌属

<400> 102

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

Tyr	His	Thr	Leu	Gln	Ser	Lys	Tyr	Pro	Asp	Ile	His	Lys	Gly	Asn	Met	
85					90					95						
Ser	Thr	Thr	Leu	Gln	Lys	Ile	Ile	Lys	Thr	Trp	Lys	Ser	Arg	Arg	Asn	
100					105					110						
Glu	Ile	Arg	Lys	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Asn	Arg	Ile	
115					120					125						
Pro	Ile	Asp	Leu	His	Asn	Asn	Ser	Val	Asp	Ile	Ile	Lys	Glu	Lys	Asn	
130					135					140						
Gly	Asp	Tyr	Ile	Ala	Gly	Ile	Ser	Leu	Phe	Ser	Arg	Asp	Phe	His	Lys	
145					150					155					160	
Glu	Asn	Gly	Asp	Val	Pro	Lys	Gly	Lys	Ile	Phe	Val	Lys	Leu	Gly	Thr	
165					170					175						
Gln	Lys	Gln	Lys	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Gln	
180					185					190						
Thr	Tyr	Ser	Lys	Gly	Ala	Cys	Met	Ile	His	Lys	Tyr	Lys	Asn	Lys	Trp	
195					200					205						
Tyr	Leu	Ser	Ile	Thr	Tyr	Lys	Phe	Asn	Ala	Ile	Lys	Glu	Asn	Lys	Phe	
210					215					220						
Asp	Lys	Glu	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Gly	Ile	Asn	Thr	Val	
225					230					235					240	
Tyr	Phe	Ala	Phe	Asn	Glu	Gly	Phe	Ile	Arg	Ser	Asn	Ile	Lys	Ser	Asp	
245					250					255						
Glu	Ile	Lys	Met	Phe	Asn	Glu	Arg	Ile	Arg	Gln	Arg	Arg	Ile	Asn	Leu	
260					265					270						
Leu	Lys	Gln	Ser	Lys	Tyr	Cys	Ser	Asn	Ser	Arg	Thr	Gly	Lys	Gly	Arg	
275					280					285						
Thr	Lys	Arg	Leu	Gln	Pro	Ile	Asp	Val	Leu	Ser	Asn	Lys	Ile	Ala	Lys	
290					295					300						
Phe	Arg	Asn	Ser	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Val	Lys	Gln	
305					310					315					320	
Cys	Leu	Lys	His	Asn	Cys	Gly	Arg	Ile	Gln	Met	Glu	Leu	Leu	Lys	Gly	
325					330					335						
Ile	Ser	Lys	Asn	Asp	Lys	Val	Leu	Lys	Asp	Trp	Thr	Tyr	Phe	Asp	Leu	
340					345					350						
Gln	Glu	Lys	Ile	Lys	Asn	Gln	Ala	Glu	Ile	Tyr	Gly	Ile	Glu	Val	Ile	
355					360					365						
Lys	Val	Val	Pro	Ala	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Gln	Cys	Gly	Tyr	
370					375					380						
Ile	Cys	Lys	Glu	Asn	Arg	Cys	Thr	Gln	Ala	Met	Phe	Glu	Cys	Lys	Gln	

385	390	395	400
Cys Gly Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
	405	410	415
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
	420	425	430
Lys Leu His Ser Lys Lys Cys Met Glu Glu Tyr Ile Glu Glu Leu Gly			
	435	440	445
Tyr Leu Asp			
450			
<210> 103			
<211> 377			
<212> PRT			
<213> 芽孢杆菌属			
<400> 103			
Met Leu Val Asn Lys Ala Tyr Lys Phe Arg Ile Tyr Pro Asn Lys Glu			
1	5	10	15
Gln Glu Ile Leu Ile Ala Lys Thr Ile Gly Cys Ser Arg Phe Val Phe			
	20	25	30
Asn His Phe Leu Gly Met Trp Asn Asp Thr Tyr Lys Glu Thr Gly Lys			
	35	40	45
Gly Leu Thr Tyr Asn Ser Cys Ser Ala Gln Leu Pro Gln Leu Lys Ile			
	50	55	60
Glu Leu Glu Trp Leu Lys Glu Val Asp Ser Ile Ala Ile Gln Ser Ala			
65	70	75	80
Leu Lys Asn Leu Val Asp Ala Tyr Asn Arg Phe Phe Lys Lys Gln Asn			
	85	90	95
Asp Lys Pro Arg Phe Lys Ser Lys Lys Asn Asp Val Gln Ser Tyr Lys			
	100	105	110
Thr Lys His Thr Asn Gly Asn Ile Ala Ile Val Asn Asn Lys Ile Lys			
	115	120	125
Leu Pro Lys Leu Gly Phe Val Thr Phe Ala Lys Ser Arg Glu Val Asp			
	130	135	140
Gly Arg Ile Met Asn Ala Thr Val Arg Arg Asn Ser Ser Gly Lys Tyr			
145	150	155	160
Phe Val Ala Ile Leu Thr Glu Val Glu Ile Gln Pro Leu Lys Lys Ala			
	165	170	175
Asp Ser Ala Ile Gly Ile Asp Leu Gly Ile Thr Asp Phe Ala Ile Leu			
	180	185	190
Ser Asp Gly His Lys Ile Asp Asn Asn Lys Phe Thr Ser Lys Met Glu			

195	200	205
Lys Lys Leu Lys Arg Glu Gln Arg Lys Leu Ser Lys Arg Ala Leu Leu		
210	215	220
Ala Lys Asn Lys Gly Ile His Leu Leu Asp Ala Gln Asn Tyr Gln Lys		
225	230	235
Gln Lys Cys Lys Val Ala Arg Leu His Glu Arg Val Ile Asn Gln Arg		
245	250	255
Asp Asp Phe Leu Asn Lys Leu Ser Thr Glu Ile Ile Lys Asn His Asp		
260	265	270
Ile Ile Cys Ile Glu Asp Leu Asn Thr Lys Gly Met Leu Arg Asn His		
275	280	285
Lys Leu Ala Lys Ser Ile Ser Asp Val Ser Trp Ser Ala Phe Val Ser		
290	295	300
Lys Leu Glu Tyr Lys Ala Thr Trp Tyr Gly Lys Thr Ile Val Lys Val		
305	310	315
Ser Arg Trp Phe Pro Ser Ser Gln Ile Cys Ser Asp Cys Gly His His		
325	330	335
Asp Gly Lys Lys Ser Leu Glu Ile Arg Gly Trp Thr Cys Pro Ile Cys		
340	345	350
His Ala Asn His Asp Arg Asp Phe Asn Ala Ser Lys Asn Ile Leu Ala		
355	360	365
Glu Gly Leu Arg Thr Leu Ala Leu Val		
370	375	
<210> 104		
<211> 450		
<212> PRT		
<213> 芽孢杆菌属		
<400> 104		
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val		
1	5	10
Asn Val Glu Trp Lys Val Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr		
20	25	30
Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe		
35	40	45
Asp Asn Leu Ser Leu Lys His Phe Lys Glu Thr Gly Glu Tyr Pro Ser		
50	55	60
Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile		
65	70	75
Tyr Asp Gln Leu Lys Glu Glu Tyr Lys Asp Ile Asn Lys Ala Asn Met		

				85					90					95			
Ser	Thr	Thr	Ile	Gln	Lys	Thr	Leu	Lys	Asn	Trp	Asn	Ser	Arg	Lys	Lys		
				100					105					110			
Glu	Ile	Trp	Arg	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Ser	Asp	Leu		
				115					120					125			
Pro	Ile	Asp	Ile	His	Gly	Asn	Ser	Ile	Gln	Leu	Ile	Lys	Glu	Lys	Gly		
				130					135					140			
Gly	Asp	Tyr	Ile	Ala	Ser	Val	Ser	Leu	Phe	Ser	Ser	Lys	Phe	Ile	Lys		
145						150				155					160		
Glu	Asn	Asp	Leu	Pro	Asn	Gly	Lys	Ile	Leu	Met	Lys	Leu	Ser	Thr	Arg		
				165					170						175		
Lys	Gln	Asn	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Ser	Thr		
				180					185					190			
Tyr	Ser	Lys	Gly	Ala	Cys	Met	Leu	His	Lys	His	Lys	Lys	Lys	Trp	Tyr		
				195					200					205			
Leu	Ser	Ile	Thr	Tyr	Lys	Ser	Asn	Ile	Lys	Glu	Glu	Leu	Lys	Phe	Asp		
				210					215					220			
Glu	Asp	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Lys	Ile	Asn	Val	Leu	Tyr		
225						230				235					240		
Phe	Ala	Phe	Asn	Lys	Ser	Leu	Val	Arg	Gly	Ala	Ile	Ser	Gly	Glu	Glu		
				245					250						255		
Ile	Glu	Ala	Phe	Arg	Lys	Lys	Ile	Glu	His	Arg	Arg	Ile	Ser	Leu	Leu		
				260					265					270			
Arg	Gln	Gly	Lys	Tyr	Cys	Ser	Gly	Asn	Arg	Ile	Gly	Lys	Gly	Arg	Glu		
				275					280					285			
Lys	Arg	Ile	Lys	Pro	Ile	Asp	Val	Leu	Asn	Asp	Lys	Val	Ala	Lys	Phe		
				290					295					300			
Arg	Asn	Ala	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Val	Gln	Gln	Cys		
305						310				315					320		
Leu	Lys	Tyr	Asn	Cys	Gly	Thr	Ile	Gln	Leu	Glu	Asp	Leu	Lys	Gly	Ile		
				325					330						335		
Ser	Lys	Glu	Gln	Thr	Phe	Leu	Lys	Asn	Trp	Thr	Tyr	Phe	Asp	Leu	Gln		
				340					345					350			
Glu	Lys	Ile	Lys	Asn	Gln	Ala	Asn	Gln	Tyr	Gly	Val	Lys	Val	Val	Lys		
				355					360					365			
Ile	Asp	Pro	Ser	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Glu	Cys	Gly	Tyr	Ile		
				370					375					380			
His	Lys	Asp	Asn	Arg	Gln	Asp	Gln	Ser	Thr	Phe	Glu	Cys	Leu	Gln	Cys		
385						390				395					400		

Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val
 405 410 415
 Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys
 420 425 430
 Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn
 435 440 445
 Ile Asn
 450
 <210> 105
 <211> 451
 <212> PRT
 <213> 芽孢杆菌属
 <400> 105
 Met Ser Val Met Ile Lys Ile Met Lys Tyr Gln Ile Leu Cys Pro Thr
 1 5 10 15
 Asn Val Asp Trp Ser Thr Phe Glu Lys Asn Leu Arg Asp Leu Thr Tyr
 20 25 30
 Gln Val Arg Thr Ile Ser Asn Arg Thr Ile Gln Gln Leu Trp Glu Phe
 35 40 45
 Asp Ala Leu Ser Tyr Asp Tyr Phe Lys Glu Thr Gly Thr Ser Pro Thr
 50 55 60
 Val Gln Asp Leu Tyr Lys Cys Thr Gln Lys Lys Ile Asp Gly Tyr Ile
 65 70 75 80
 Tyr His Val Leu Gln Ser Lys Tyr Pro Asp Val His Lys Gly Asn Met
 85 90 95
 Ser Thr Thr Leu Gln Lys Ile Ile Lys Thr Trp Lys Ser Arg Arg Ser
 100 105 110
 Glu Ile Arg Lys Gly Glu Met Ser Ile Pro Ser Phe Arg Asn Gln Leu
 115 120 125
 Pro Ile Asp Leu His Asn Asn Ser Val Glu Val Thr Lys Glu Lys Asn
 130 135 140
 Gly Asp Tyr Ile Ala Gly Leu Ser Leu Phe Ser Arg Asp Phe His Lys
 145 150 155 160
 Glu Asn Ser Asp Val Pro Lys Gly Lys Ile Phe Val Lys Leu Ala Thr
 165 170 175
 Lys Lys Gln Lys Ser Met Lys Val Ile Leu Asp Arg Leu Ile Ser Gly
 180 185 190
 Thr Tyr Ser Lys Gly Ala Cys Met Ile His Lys Tyr Lys Asn Lys Trp
 195 200 205

Tyr Leu Ser Ile Thr Tyr Lys Phe Asn Ser Ile Thr Glu Asn Lys Phe			
210	215	220	
Asp Glu Asn Leu Ile Met Gly Ile Asp Met Gly Gly Val Asn Thr Val			
225	230	235	240
Tyr Phe Ala Phe Asn Glu Gly Phe Val Arg Ser Asn Ile Arg Ser Asp			
245	250	255	
Glu Ile Arg Ala Phe Asn Glu Arg Ile Arg Gln Arg Arg Ile Asn Leu			
260	265	270	
Leu Asn Gln Ser Lys Tyr Cys Ser Asp Ser Arg Thr Gly Lys Gly Arg			
275	280	285	
Ala Lys Arg Leu Gln Pro Ile Asp Val Ile Ser Asn Lys Ile Ala Gln			
290	295	300	
Phe Arg Asn Ser Thr Asn His Lys Tyr Ala Asn Phe Ile Val Lys Gln			
305	310	315	320
Cys Leu Lys Tyr Asn Cys Gly Arg Ile Gln Ile Glu His Leu Lys Gly			
325	330	335	
Ile Ser Lys Asp Asp Lys Val Leu Lys Asp Trp Thr Tyr Tyr Asp Leu			
340	345	350	
Gln Glu Lys Ile Lys Lys Gln Ala Gln Ala Tyr Gly Ile Glu Val Ile			
355	360	365	
Thr Ile Ala Pro Ala Tyr Thr Ser Gln Arg Cys Ser Gln Cys Gly Tyr			
370	375	380	
Ile Ser Asn Glu Asn Arg Cys Thr Gln Ala Val Phe Glu Cys Lys Gln			
385	390	395	400
Cys Glu Tyr Lys Thr His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser			
405	410	415	
Thr Tyr Asp Ile Glu Asn Ile Ile Asn Lys Gln Leu Ala Val Gln Ser			
420	425	430	
Arg Leu His Ser Lys Lys Tyr Met Glu Asp His Ile Glu Glu Leu Gly			
435	440	445	
Tyr Ser Gly			
450			
<210> 106			
<211> 450			
<212> PRT			
<213> 芽孢杆菌属			
<400> 106			
Met Ser Leu Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val			
1	5	10	15

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

	325							330							335						
Ser	Lys	Glu	Gln	Thr	Phe	Leu	Lys	Asn	Trp	Thr	Tyr	Phe	Asp	Leu	Gln						
	340							345							350						
Glu	Lys	Ile	Lys	Asn	Gln	Ala	Asn	Gln	Tyr	Gly	Val	Lys	Val	Val	Lys						
	355							360							365						
Ile	Asp	Pro	Ser	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Glu	Cys	Gly	Tyr	Ile						
	370							375							380						
His	Lys	Asp	Asn	Arg	Gln	Asp	Gln	Ser	Thr	Phe	Glu	Cys	Leu	Gln	Cys						
385	390							395							400						
Ser	Phe	Lys	Ala	His	Ala	Asp	Tyr	Asn	Ala	Ala	Lys	Asn	Ile	Ser	Val						
	405							410							415						
Tyr	Thr	Ile	Glu	Lys	Val	Ile	Gln	Lys	Gln	Leu	Glu	Leu	Gln	Glu	Lys						
	420							425							430						
Leu	Asn	Leu	Thr	Lys	Tyr	Lys	Glu	Gln	Tyr	Ile	Glu	Gln	Ile	Lys	Asn						
	435							440							445						
Ile	Asn																				
	450																				
<210>	107																				
<211>	450																				
<212>	PRT																				
<213>	芽孢杆菌属																				
<400>	107																				
Met	Ser	Leu	Ala	Val	Lys	Val	Met	Lys	Tyr	Gln	Ile	Val	Cys	Pro	Val						
1				5					10					15							
Asn	Val	Glu	Trp	Lys	Val	Phe	Glu	Thr	Tyr	Leu	Arg	Thr	Leu	Ser	Tyr						
	20							25							30						
Gln	Ser	Arg	Thr	Ile	Gly	Asn	Arg	Thr	Ile	Gln	Lys	Ile	Trp	Glu	Phe						
	35							40							45						
Asp	Asn	Leu	Ser	Leu	Lys	His	Phe	Lys	Glu	Thr	Gly	Glu	Tyr	Pro	Ser						
	50							55							60						
Ala	Gln	Gln	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Thr	Ile	Ser	Gly	Tyr	Ile						
65	70							75							80						
Tyr	Asp	Gln	Leu	Lys	Glu	Glu	Tyr	Lys	Asp	Ile	Asn	Lys	Ala	Asn	Met						
	85							90							95						
Ser	Thr	Thr	Ile	Gln	Lys	Thr	Leu	Lys	Asn	Trp	Asn	Ser	Arg	Lys	Lys						
	100							105							110						
Glu	Ile	Trp	Arg	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Ser	Asp	Leu						
	115							120							125						
Pro	Ile	Asp	Ile	His	Gly	Asn	Ser	Ile	Gln	Leu	Ile	Lys	Glu	Lys	Gly						

130	135	140
Gly Asp Tyr Ile Ala Ser Val Ser Leu Phe Ser Ser Lys Phe Ile Lys		
145	150	155
Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Met Lys Leu Ser Thr Arg		
165	170	175
Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr		
180	185	190
Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Lys Lys Trp Tyr		
195	200	205
Leu Ser Ile Thr Tyr Lys Ser Asn Ile Lys Glu Glu Leu Lys Phe Asp		
210	215	220
Glu Asp Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr		
225	230	235
Phe Ala Phe Asn Lys Ser Leu Val Arg Gly Ala Ile Ser Gly Glu Glu		
245	250	255
Ile Glu Ala Phe Arg Lys Lys Ile Glu His Arg Arg Ile Ser Leu Leu		
260	265	270
Arg Gln Gly Lys Tyr Cys Ser Gly Asn Arg Ile Gly Lys Gly Arg Glu		
275	280	285
Lys Arg Ile Lys Pro Ile Asp Val Leu Asn Asp Lys Val Ala Lys Phe		
290	295	300
Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys		
305	310	315
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Lys Gly Ile		
325	330	335
Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln		
340	345	350
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Val Lys Val Val Lys		
355	360	365
Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile		
370	375	380
His Lys Asp Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Leu Gln Cys		
385	390	395
Ser Phe Lys Ala His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Val		
405	410	415
Tyr Thr Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys		
420	425	430
Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Ile Lys Asn		
435	440	445

Ile Asn
 450
 <210> 108
 <211> 450
 <212> PRT
 <213> 芽孢杆菌属
 <400> 108
 Met Ser Ile Ala Val Lys Val Met Lys Tyr Gln Ile Val Cys Pro Val
 1 5 10 15
 Asn Val Glu Trp Lys Ala Phe Glu Thr Tyr Leu Arg Thr Leu Ser Tyr
 20 25 30
 Gln Ser Arg Thr Ile Gly Asn Arg Thr Ile Gln Lys Ile Trp Glu Phe
 35 40 45
 Asp Asn Leu Ser Leu Asn His Phe Lys Glu Thr Gly Glu Tyr Pro Ser
 50 55 60
 Ala Gln Gln Leu Tyr Gly Cys Thr Gln Lys Thr Ile Ser Gly Tyr Ile
 65 70 75 80
 Tyr Asn Gln Leu Arg Glu Glu Tyr Gln Asp Ile Asn Lys Ala Asn Met
 85 90 95
 Ser Thr Thr Ile Gln Lys Ser Leu Lys Asn Trp Asn Ser Arg Lys Lys
 100 105 110
 Glu Ile Trp Arg Gly Glu Met Ser Ile Pro Ser Phe Arg Ser Asp Leu
 115 120 125
 Pro Ile Asp Ile His Gly Asn Ser Ile Gln Leu Ile Lys Glu Lys Ser
 130 135 140
 Gly Asp Tyr Phe Ala Asn Val Ser Leu Phe Ser Ser Lys Phe Ile Lys
 145 150 155 160
 Glu Asn Asp Leu Pro Asn Gly Lys Ile Leu Val Lys Leu Ser Thr Arg
 165 170 175
 Lys Gln Asn Ser Met Lys Val Ile Leu Asp Arg Leu Ile Asn Ser Thr
 180 185 190
 Tyr Ser Lys Gly Ala Cys Met Leu His Lys His Lys Asn Lys Trp Tyr
 195 200 205
 Leu Ser Ile Thr Tyr Lys Ala Thr Lys Lys Glu Gly His Lys Phe Asp
 210 215 220
 Glu Glu Leu Ile Met Gly Ile Asp Met Gly Lys Ile Asn Val Leu Tyr
 225 230 235 240
 Phe Ala Tyr Asn Lys Gly Leu Val Arg Gly Ala Ile Ser Gly Glu Glu
 245 250 255

Ile Glu Ala Phe Arg Lys Lys Ile Glu Tyr Arg Arg Ile Ser Leu Leu		
260	265	270
Arg Gln Gly Lys Tyr Cys Ser Glu Asn Arg Ile Gly Lys Gly Arg Lys		
275	280	285
Lys Arg Ile Lys Pro Ile Glu Val Leu Asn Asp Lys Ile Thr Lys Phe		
290	295	300
Arg Asn Ala Thr Asn His Lys Tyr Ala Asn Tyr Ile Val Gln Gln Cys		
305	310	315
Leu Lys Tyr Asn Cys Gly Thr Ile Gln Leu Glu Asp Leu Gln Gly Ile		
325	330	335
Ser Lys Glu Gln Thr Phe Leu Lys Asn Trp Thr Tyr Phe Asp Leu Gln		
340	345	350
Glu Lys Ile Lys Asn Gln Ala Asn Gln Tyr Gly Ile Lys Val Val Lys		
355	360	365
Ile Asp Pro Ser Tyr Thr Ser Gln Arg Cys Ser Glu Cys Gly Tyr Ile		
370	375	380
His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys		
385	390	395
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile		
405	410	415
Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys		
420	425	430
Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn		
435	440	445

Ile Asn
450

<210> 109

<211> 450

<212> PRT

<213> 芽孢杆菌属

<400> 109

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

Ala	Gln	Gln	Leu	Tyr	Gly	Cys	Thr	Gln	Lys	Thr	Ile	Ser	Gly	Tyr	Ile
65					70					75					80
Tyr	Asn	Gln	Leu	Arg	Glu	Glu	Tyr	Gln	Asp	Ile	Asn	Lys	Ala	Asn	Met
			85						90					95	
Ser	Thr	Thr	Ile	Gln	Lys	Ser	Leu	Lys	Asn	Trp	Asn	Ser	Arg	Lys	Lys
			100						105					110	
Glu	Ile	Trp	Arg	Gly	Glu	Met	Ser	Ile	Pro	Ser	Phe	Arg	Ser	Asp	Leu
			115						120					125	
Pro	Ile	Asp	Ile	His	Gly	Asn	Ser	Ile	Gln	Leu	Ile	Lys	Glu	Lys	Ser
			130						135					140	
Gly	Asp	Tyr	Phe	Ala	Asn	Val	Ser	Leu	Phe	Ser	Ser	Lys	Phe	Ile	Lys
145					150					155					160
Glu	Asn	Asp	Leu	Pro	Asn	Gly	Lys	Ile	Leu	Val	Lys	Leu	Ser	Thr	Arg
			165						170					175	
Lys	Gln	Asn	Ser	Met	Lys	Val	Ile	Leu	Asp	Arg	Leu	Ile	Asn	Ser	Thr
			180						185					190	
Tyr	Ser	Lys	Gly	Ala	Cys	Met	Leu	His	Lys	His	Lys	Asn	Lys	Trp	Tyr
			195						200					205	
Leu	Ser	Ile	Thr	Tyr	Lys	Ala	Thr	Lys	Lys	Glu	Gly	His	Lys	Phe	Asp
			210						215					220	
Glu	Glu	Leu	Ile	Met	Gly	Ile	Asp	Met	Gly	Lys	Ile	Asn	Val	Leu	Tyr
225					230					235					240
Phe	Ala	Tyr	Asn	Lys	Gly	Leu	Val	Arg	Gly	Ala	Ile	Ser	Gly	Glu	Glu
			245						250					255	
Ile	Glu	Ala	Phe	Arg	Lys	Lys	Ile	Glu	Tyr	Arg	Arg	Ile	Ser	Leu	Leu
			260						265					270	
Arg	Gln	Gly	Lys	Tyr	Cys	Ser	Glu	Asn	Arg	Ile	Gly	Lys	Gly	Arg	Lys
			275						280					285	
Lys	Arg	Ile	Lys	Pro	Ile	Glu	Val	Leu	Asn	Asp	Lys	Ile	Thr	Lys	Phe
			290						295					300	
Arg	Asn	Ala	Thr	Asn	His	Lys	Tyr	Ala	Asn	Tyr	Ile	Val	Gln	Gln	Cys
305					310					315					320
Leu	Lys	Tyr	Asn	Cys	Gly	Thr	Ile	Gln	Leu	Glu	Asp	Leu	Gln	Gly	Ile
			325						330					335	
Ser	Lys	Glu	Gln	Thr	Phe	Leu	Lys	Asn	Trp	Thr	Tyr	Phe	Asp	Leu	Gln
			340						345					350	
Glu	Lys	Ile	Lys	Asn	Gln	Ala	Asn	Gln	Tyr	Gly	Ile	Lys	Val	Val	Lys
			355						360					365	
Ile	Asp	Pro	Ser	Tyr	Thr	Ser	Gln	Arg	Cys	Ser	Glu	Cys	Gly	Tyr	Ile

370	375	380
His Lys Asn Asn Arg Gln Asp Gln Ser Thr Phe Glu Cys Gln His Cys		
385	390	395
Ser Phe Lys Val His Ala Asp Tyr Asn Ala Ala Lys Asn Ile Ser Ile		400
405	410	415
Tyr Asn Ile Glu Lys Val Ile Gln Lys Gln Leu Glu Leu Gln Glu Lys		
420	425	430
Leu Asn Leu Thr Lys Tyr Lys Glu Gln Tyr Ile Glu Gln Met Glu Asn		
435	440	445
Ile Asn		
450		
<210> 110		
<211> 11		
<212> RNA		
<213> 芽孢杆菌属		
<400> 110		
aggauugaaa u		11
<210> 111		
<211> 11		
<212> RNA		
<213> 芽孢杆菌属		
<400> 111		
aggauugaaa u		11
<210> 112		
<211> 11		
<212> RNA		
<213> 芽孢杆菌属		
<400> 112		
uggauugaaa u		11
<210> 113		
<211> 11		
<212> RNA		
<213> 芽孢杆菌属		
<400> 113		
aggauugaaa u		11
<210> 114		
<211> 11		
<212> RNA		
<213> 芽孢杆菌属		

<400> 114	
uggauugaaa u	11
<210> 115	
<211> 18	
<212> RNA	
<213> 芽孢杆菌属	
<400> 115	
caauagaugu auugaaaau	18
<210> 116	
<211> 37	
<212> RNA	
<213> 戈登氏菌属	
<400> 116	
gucgugagua cgacgcccuc acgcuggcgu ugcgacc	37
<210> 117	
<211> 13	
<212> RNA	
<213> 微球菌菌属	
<400> 117	
cagggaugag ccc	13
<210> 118	
<211> 13	
<212> RNA	
<213> 微球菌菌属	
<400> 118	
cagggaugag ccc	13
<210> 119	
<211> 13	
<212> RNA	
<213> 微球菌菌属	
<400> 119	
cagggaugag ccc	13
<210> 120	
<211> 201	
<212> RNA	
<213> 芽孢杆菌属	
<400> 120	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauguug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120

gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucuaa cucgauacau u	201
<210> 121	
<211> 201	
<212> RNA	
<213> 芽孢杆菌属	
<400> 121	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauguug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucuaa cucgauacau u	201
<210> 122	
<211> 201	
<212> RNA	
<213> 芽孢杆菌属	
<400> 122	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauauug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucuaa cucgauacau u	201
<210> 123	
<211> 201	
<212> RNA	
<213> 芽孢杆菌属	
<400> 123	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauguug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucuaa cucgauacau u	201
<210> 124	
<211> 201	
<212> RNA	
<213> 芽孢杆菌属	
<400> 124	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauauug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucuaa cucgauacau u	201
<210> 125	

<211> 124	
<212> RNA	
<213> 芽孢杆菌属	
<400> 125	
acggucuuug uguuacaccc uucaaagcau ccaggaugcu uauguaucaa uguuuuuuuu	60
acauucgaua cuaccuuuug gagcguuuac acuugguaaa uucaccuugu auuacuucuc	120
cauu	124
<210> 126	
<211> 97	
<212> RNA	
<213> 微球菌属	
<400> 126	
uucugggcgc caggcagagc gaagcccucc cggcgcgag gggggauggc caggagggu	60
ucuaagucgc aggcccugcg ggaugggcuc acguucg	97
<210> 127	
<211> 96	
<212> RNA	
<213> 微球菌属	
<400> 127	
guuuacacag gucgaacgug agcccauccc guagggccug cgacuagaa gccuccugg	60
ccaucacccc uccgcgccgg gagggcuucg cucugc	96
<210> 128	
<211> 97	
<212> RNA	
<213> 微球菌属	
<400> 128	
uucugggcgc caggcagagc gaagcccucc cggcgcgag gggggauggc caggagggu	60
ucuaagucgc aggcccugcg ggaugggcuc acguucg	97
<210> 129	
<211> 221	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 129	
aaacagcuag aauguaacuu aaaguagguc aauguuuuuu uucgauguug cauuuguuuu	60
ggacaagugg auuuuuacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucaa aaagauccua cauggaugug aggauugaaa u	221

<210> 130	
<211> 221	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 130	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauguug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucua aaagauccua cauggaugug aggauugaaa u	221
<210> 131	
<211> 221	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 131	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauguug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucua aaagauccua cauggaugug aggauugaaa u	221
<210> 132	
<211> 216	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 132	
aaacagcuag aauguaacuu aaaguagguc aauguuuuaa uucgauauug caauuuguuu	60
ggacaagugg auuaaaacgu uccuugaaaa ucauauaaag cagccaguuu acgggcuugg	120
gcgaauuugc guccaaaggg ugaggccagg uguaaguaag aaccuacaaa agcacucacc	180
aaagggucua cucgauacau uaaaguggau ugaaau	216
<210> 133	
<211> 114	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	

<400> 133	
uucugggcg c caggcagagc gaagcccucc cggcgcgag gggggauggc caggagggcu	60
ucuaagucgc agggccugcg ggaugggcuc acguucgaaa gcagggauga gccc	114
<210> 134	
<211> 113	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 134	
guuuacacag gucgaacgug agcccauccc guagggccug cgacuuagaa gcccuccugg	60
ccaucggggg uccgcgcccgg gagggcuucg cucugcaaag cagggacgag ccc	113
<210> 135	
<211> 114	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 135	
uucugggcg c caggcagagc gaagcccucc cggcgcgag gggggauggc caggagggcu	60
ucuaagucgc agggccugcg ggaugggcuc acguucgaaa gcagggauga gccc	114
<210> 136	
<211> 4	
<212> PRT	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成胜肽	
<400> 136	
Ala Ala Ala Gly	
1	
<210> 137	
<211> 82	
<212> PRT	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多肽	
<400> 137	
Thr Asn Leu Ser Asp His Glu Lys Glu Thr Gly Lys Gln Leu Val Ile	
1	5 10 15

Gln Glu Ser Ile Leu Met Leu Pro Glu Glu Val Glu Glu Val Ile Gly
 20 25 30
 Asn Lys Pro Glu Ser Asp Ile Leu Val His Thr Ala Tyr Asp Glu Ser
 35 40 45
 Thr Asp Glu Asn Val Met Leu Leu Thr Ser Asp Ala Pro Glu Tyr Lys
 50 55 60
 Pro Trp Ala Leu Val Ile Gln Asp Ser Asn Gly Glu Asn Lys Ile Lys
 65 70 75 80
 Met Leu
 <210> 138
 <211> 1493
 <212> PRT
 <213> 智人
 <400> 138
 Met Pro Asn Glu Gly Ile Pro His Ser Ser Gln Thr Gln Glu Gln Asp
 1 5 10 15
 Cys Leu Gln Ser Gln Pro Val Ser Asn Asn Glu Glu Met Ala Ile Lys
 20 25 30
 Gln Glu Ser Gly Gly Asp Gly Glu Val Glu Glu Tyr Leu Ser Phe Arg
 35 40 45
 Ser Val Gly Asp Gly Leu Ser Thr Ser Ala Val Gly Cys Ala Ser Ala
 50 55 60
 Ala Pro Arg Arg Gly Pro Ala Leu Leu His Ile Asp Arg His Gln Ile
 65 70 75 80
 Gln Ala Val Glu Pro Ser Ala Gln Ala Leu Glu Leu Gln Gly Leu Gly
 85 90 95
 Val Asp Val Tyr Asp Gln Asp Val Leu Glu Gln Gly Val Leu Gln Gln
 100 105 110
 Val Asp Asn Ala Ile His Glu Ala Ser Arg Ala Ser Gln Leu Val Asp
 115 120 125
 Val Glu Lys Glu Tyr Arg Ser Val Leu Asp Asp Leu Thr Ser Cys Thr
 130 135 140
 Thr Ser Leu Arg Gln Ile Asn Lys Ile Ile Glu Gln Leu Ser Pro Gln
 145 150 155 160
 Ala Ala Thr Ser Arg Asp Ile Asn Arg Lys Leu Asp Ser Val Lys Arg
 165 170 175
 Gln Lys Tyr Asn Lys Glu Gln Gln Leu Lys Lys Ile Thr Ala Lys Gln
 180 185 190
 Lys His Leu Gln Ala Ile Leu Gly Gly Ala Glu Val Lys Ile Glu Leu

195	200	205
Asp His Ala Ser Leu Glu Glu Asp Ala Glu Pro Gly Pro Ser Ser Leu		
210	215	220
Gly Ser Met Leu Met Pro Val Gln Glu Thr Ala Trp Glu Glu Leu Ile		
225	230	235
Arg Thr Gly Gln Met Thr Pro Phe Gly Thr Gln Ile Pro Gln Lys Gln		
245	250	255
Glu Lys Lys Pro Arg Lys Ile Met Leu Asn Glu Ala Ser Gly Phe Glu		
260	265	270
Lys Tyr Leu Ala Asp Gln Ala Lys Leu Ser Phe Glu Arg Lys Lys Gln		
275	280	285
Gly Cys Asn Lys Arg Ala Ala Arg Lys Ala Pro Ala Pro Val Thr Pro		
290	295	300
Pro Ala Pro Val Gln Asn Lys Asn Lys Pro Asn Lys Lys Ala Arg Val		
305	310	315
Leu Ser Lys Lys Glu Glu Arg Leu Lys Lys His Ile Lys Lys Leu Gln		
325	330	335
Lys Arg Ala Leu Gln Phe Gln Gly Lys Val Gly Leu Pro Lys Ala Arg		
340	345	350
Arg Pro Trp Glu Ser Asp Met Arg Pro Glu Ala Glu Gly Asp Ser Glu		
355	360	365
Gly Glu Glu Ser Glu Tyr Phe Pro Thr Glu Glu Glu Glu Glu Glu		
370	375	380
Asp Asp Glu Val Glu Gly Ala Glu Ala Asp Leu Ser Gly Asp Gly Thr		
385	390	395
Asp Tyr Glu Leu Lys Pro Leu Pro Lys Gly Gly Lys Arg Gln Lys Lys		
405	410	415
Val Pro Val Gln Glu Ile Asp Asp Asp Phe Phe Pro Ser Ser Gly Glu		
420	425	430
Glu Ala Glu Ala Ala Ser Val Gly Glu Gly Gly Gly Gly Arg Lys		
435	440	445
Val Gly Arg Tyr Arg Asp Asp Gly Asp Glu Asp Tyr Tyr Lys Gln Arg		
450	455	460
Leu Arg Arg Trp Asn Lys Leu Arg Leu Gln Asp Lys Glu Lys Arg Leu		
465	470	475
Lys Leu Glu Asp Asp Ser Glu Glu Ser Asp Ala Glu Phe Asp Glu Gly		
485	490	495
Phe Lys Val Pro Gly Phe Leu Phe Lys Lys Leu Phe Lys Tyr Gln Gln		
500	505	510

Thr Gly Val Arg Trp Leu Trp Glu Leu His Cys Gln Gln Ala Gly Gly		
515	520	525
Ile Leu Gly Asp Glu Met Gly Leu Gly Lys Thr Ile Gln Ile Ile Ala		
530	535	540
Phe Leu Ala Gly Leu Ser Tyr Ser Lys Ile Arg Thr Arg Gly Ser Asn		
545	550	555 560
Tyr Arg Phe Glu Gly Leu Gly Pro Thr Val Ile Val Cys Pro Thr Thr		
565	570	575
Val Met His Gln Trp Val Lys Glu Phe His Thr Trp Trp Pro Pro Phe		
580	585	590
Arg Val Ala Ile Leu His Glu Thr Gly Ser Tyr Thr His Lys Lys Glu		
595	600	605
Lys Leu Ile Arg Asp Val Ala His Cys His Gly Ile Leu Ile Thr Ser		
610	615	620
Tyr Ser Tyr Ile Arg Leu Met Gln Asp Asp Ile Ser Arg Tyr Asp Trp		
625	630	635 640
His Tyr Val Ile Leu Asp Glu Gly His Lys Ile Arg Asn Pro Asn Ala		
645	650	655
Ala Val Thr Leu Ala Cys Lys Gln Phe Arg Thr Pro His Arg Ile Ile		
660	665	670
Leu Ser Gly Ser Pro Met Gln Asn Asn Leu Arg Glu Leu Trp Ser Leu		
675	680	685
Phe Asp Phe Ile Phe Pro Gly Lys Leu Gly Thr Leu Pro Val Phe Met		
690	695	700
Glu Gln Phe Ser Val Pro Ile Thr Met Gly Gly Tyr Ser Asn Ala Ser		
705	710	715 720
Pro Val Gln Val Lys Thr Ala Tyr Lys Cys Ala Cys Val Leu Arg Asp		
725	730	735
Thr Ile Asn Pro Tyr Leu Leu Arg Arg Met Lys Ser Asp Val Lys Met		
740	745	750
Ser Leu Ser Leu Pro Asp Lys Asn Glu Gln Val Leu Phe Cys Arg Leu		
755	760	765
Thr Asp Glu Gln His Lys Val Tyr Gln Asn Phe Val Asp Ser Lys Glu		
770	775	780
Val Tyr Arg Ile Leu Asn Gly Glu Met Gln Ile Phe Ser Gly Leu Ile		
785	790	795 800
Ala Leu Arg Lys Ile Cys Asn His Pro Asp Leu Phe Ser Gly Gly Pro		
805	810	815
Lys Asn Leu Lys Gly Leu Pro Asp Asp Glu Leu Glu Glu Asp Gln Phe		

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

Cys Ser Asn Ser Ser Gly Thr Gly Lys Thr Ser Met Pro Ser Gly		
1130	1135	1140
Asp Glu Ser Ile Asp Glu Lys Leu Gly Leu Ser Tyr Lys Arg Glu		
1145	1150	1155
Arg Pro Ser Gln Ala Gln Thr Glu Ala Phe Trp Glu Asn Lys Gln		
1160	1165	1170
Met Glu Asn Asn Phe Tyr Lys His Lys Ser Lys Thr Lys His His		
1175	1180	1185
Ser Val Ala Glu Glu Glu Thr Leu Glu Lys His Leu Arg Pro Lys		
1190	1195	1200
Gln Lys Pro Lys Asn Ser Lys His Cys Arg Asp Ala Lys Phe Glu		
1205	1210	1215
Gly Thr Arg Ile Pro His Leu Val Lys Lys Arg Arg Tyr Gln Lys		
1220	1225	1230
Gln Asp Ser Glu Asn Lys Ser Glu Ala Lys Glu Gln Ser Asn Asp		
1235	1240	1245
Asp Tyr Val Leu Glu Lys Leu Phe Lys Lys Ser Val Gly Val His		
1250	1255	1260
Ser Val Met Lys His Asp Ala Ile Met Asp Gly Ala Ser Pro Asp		
1265	1270	1275
Tyr Val Leu Val Glu Ala Glu Ala Asn Arg Val Ala Gln Asp Ala		
1280	1285	1290
Leu Lys Ala Leu Arg Leu Ser Arg Gln Arg Cys Leu Gly Ala Val		
1295	1300	1305
Ser Gly Val Pro Thr Trp Thr Gly His Arg Gly Ile Ser Gly Ala		
1310	1315	1320
Pro Ala Gly Lys Lys Ser Arg Phe Gly Lys Lys Arg Asn Ser Asn		
1325	1330	1335
Phe Ser Val Gln His Pro Ser Ser Thr Ser Pro Thr Glu Lys Cys		
1340	1345	1350
Gln Asp Gly Ile Met Lys Lys Glu Gly Lys Asp Asn Val Pro Glu		
1355	1360	1365
His Phe Ser Gly Arg Ala Glu Asp Ala Asp Ser Ser Ser Gly Pro		
1370	1375	1380
Leu Ala Ser Ser Ser Leu Leu Ala Lys Met Arg Ala Arg Asn His		
1385	1390	1395
Leu Ile Leu Pro Glu Arg Leu Glu Ser Glu Ser Gly His Leu Gln		
1400	1405	1410
Glu Ala Ser Ala Leu Leu Pro Thr Thr Glu His Asp Asp Leu Leu		

1415	1420	1425	
Val Glu Met Arg Asn Phe	Ile Ala Phe Gln Ala	His Thr Asp Gly	
1430	1435	1440	
Gln Ala Ser Thr Arg Glu	Ile Leu Gln Glu Phe	Glu Ser Lys Leu	
1445	1450	1455	
Ser Ala Ser Gln Ser Cys	Val Phe Arg Glu Leu	Leu Arg Asn Leu	
1460	1465	1470	
Cys Thr Phe His Arg Thr	Ser Gly Gly Glu Gly	Ile Trp Lys Leu	
1475	1480	1485	
Lys Pro Glu Tyr Cys			
1490			
<210> 139			
<211> 55			
<212> DNA			
<213> 人工序列			
<220>			
<223> 人工序列的描述：合成寡核苷酸			
<400> 139			
gtcacatcct acatggatgt gtggattgaa attggaatgg gaactaaagt aatgg			55
<210> 140			
<211> 201			
<212> DNA			
<213> 人工序列			
<220>			
<223> 人工序列的描述：合成多核苷酸			
<400> 140			
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatattg caatttgttt			60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg			120
gcgaattttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc			180
aaagggtcaa ctcgatacat t			201
<210> 141			
<211> 53			
<212> DNA			
<213> 人工序列			
<220>			
<223> 人工序列的描述：合成寡核苷酸			
<400> 141			
gtttaaacat aacaatagat gtattgaaat tggaatggga actaaagtaa tgg			53
<210> 142			

<211> 124	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 142	
acgggtctttg tgttacaccc ttcaaagcat ccaggatgct tatgtatcaa tgtttttttta	60
acattcgata ctaccttatg gagcgtttac acttggtaaa ttcaccttgt attacttctc	120
catt	124
<210> 143	
<211> 55	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 143	
gtcacatcct acatggatgt gaggattgaa attggaatgg gaactaaagt aatgg	55
<210> 144	
<211> 70	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 144	
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt	60
ggacaagtgg	70
<210> 145	
<211> 201	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 145	
aaacagctag aatgtaactt aaagtaggtc aatgtttaaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa ctcgatacat t	201
<210> 146	
<211> 55	

<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 146	
gtcacatcct acatggatgt gtggattgaa attggaatgg gaactaaagt aatgg	55
<210> 147	
<211> 201	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 147	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatattg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa ctcgatacat t	201
<210> 148	
<211> 101	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<400> 148	
cagccagttt acgggcttgg gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag	60
aacctacaaa agcactcacc aaaggggtcaa ctcgatacat t	101
<210> 149	
<211> 21	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 149	
cctaagaaga aaagaaaggt g	21
<210> 150	
<211> 48	
<212> DNA	
<213> 人工序列	
<220>	

<223> 人工序列的描述：合成寡核苷酸	
<400> 150	
aaaagacctg ccgctacaaa gaaggccggc caggccaaga aaaagaag	48
<210> 151	
<211> 66	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成寡核苷酸	
<400> 151	
gactacaagg accacgacgg cgactacaaa gatcacgata tcgactacaa ggacgacgat	60
gataag	66
<210> 152	
<211> 203	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 152	
tgatgcggtt ttggcagtac atcaatgggc gtggatagcg gtttgactca cggggatttc	60
caagtctcca ccccatcgac gtcaatggga gtttgttttg gcaccaaagt caacgggact	120
ttccaaaatg tcgtaacaac tccgccccat tgacgcaaat gggcggtagg cgtgtacggt	180
gggaggtcta tataagcaga gct	203
<210> 153	
<211> 318	
<212> DNA	
<213> 智人	
<400> 153	
tgtacaaaaa agcaggcttt aaaggaacca attcagtcga ctggatccgg taccaaggctc	60
gggcaggaag agggcctatt tcccatgatt cttcatatt tgcatatacg atacaaggct	120
gttagagaga taattagaat taatttgact gtaaacacaa agatattagt acaaaatacg	180
tgacgtagaa agtaataatt tcttgggtag tttgcagttt taaaattatg ttttaaaatg	240
gactatcata tgcttaccgt aacttgaaag tatttcgatt tcttggcttt atatatcttg	300
tggaaggac gaaacacc	318
<210> 154	
<211> 1644	
<212> DNA	
<213> 人工序列	
<220>	

<223> 人工序列的描述：合成多核苷酸

<400> 154

```

gacgcattgg acgattttga tctggatatg ctgggaagtg acgccctcga tgattttgac      60
cttgacatgc ttggttcgga tgcccttgat gactttgacc tcgacatgct cggcagtgac      120
gcccttgatg atttcgacct ggacatgctg attaactcta gaagttccgg atctggtagc      180
cagtacctgc ccgacaccga cgaccggcac cggatcgagg aaaagcggaa gcggacctac      240
gagacattca agagcatcat gaagaagtcc cccttcagcg gccccaccga ccctagacct      300
ccacctagaa gaatcgccgt gcccagcaga tccagcgcca gcgtgccaaa acctgcccc      360
cagccttacc ccttcaccag cagcctgagc accatcaact acgacgagtt ccctaccatg      420
gtgttcccca gcggccagat ctctcaggcc tctgctctgg ctccagcccc tcctcaggtg      480
ctgcctcagg ctctctgtcc tgcaccagct ccagccatgg tgtctgcaact ggctcaggca      540
ccagcacccg tgctgtgtct ggctcctgga cctccacagg ctgtggctcc accagccct      600
aaacctacac aggccggcga gggcacactg tctgaagctc tgctgcagct gcagttcgac      660
gacgaggatc tgggagccct gctgggaaac agcaccgatc ctgccgtgtt caccgacctg      720
gccagcgtgg acaacagcga gttccagcag ctgctgaacc agggcatccc tgtggccct      780
cacaccaccg agcccatgct gatggaatac cccgaggcca tcaccggct cgtgacaggc      840
gctcagaggc ctctgatcc agtcctgcc cctctgggag caccaggcct gcctaattga      900
ctgctgtctg gcgacgagga cttcagctct atcgccgata tggatttctc agccttgctg      960
ggctctggca gcggcagccg ggattccagg gaagggatgt ttttgccgaa gcctgaggcc      1020
ggctccgcta ttagtgacgt gtttgagggc cgcgagggtg gccagccaaa acgaatccgg      1080
ccatttcata ctccaggaag tccatgggcc aaccgcccac tccccgccag cctcgacca      1140
acaccaaccg gtccagtaca tgagccagtc gggtcactga ccccggcacc agtcctcag      1200
ccactagatc cagcggccgc agtgactccc gaggccagtc acctgttgga agatcccgat      1260
gaagagacga gccaggctgt caaagccctt cgggagatgg ccgatactgt gattccccag      1320
aaggaagagg ctgcaatctg tggccaaatg gacctttccc atccgcccc aaggggcat      1380
ctggatgagc tgacaaccac acttgagtcc atgaccgagg atctgaacct ggactcaccc      1440
ctgaccccg aattgaacga gattctggat accttctga acgacgagtg cctcttgcat      1500
gccatgcata tcagcacagg actgtccatc ttgacacat ctctgttttc cggaggatct      1560
agcggaggct cctctggctc tgagacacct ggcacaagcg agagcgcaac acctgaaagc      1620
agcgggggca gcagcggggg gtca      1644

```

<210> 155

<211> 1518

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 155

```

atgagcacc ctctgcagca gcctcaccag aaatccaaga agaccagcca gatgataaca      60
acacgcaagt tcaagctcgc catcgttagc gacaaccgga acgaggccta cagcttcac      120

```

cggaacgaga	tcagaaacca	gaacaaggcc	ctgaacgccg	cttataacca	cctgtacttc	180
gagcacatcg	caaccgagaa	actgaagcac	agcgatgagg	aataccagaa	gcacctgaca	240
aagtacagag	aggtggccac	caacaagtac	caggactacc	tgaaggtgaa	agaaaaagtg	300
aacgacagca	aggacgacga	gaagctgcag	aagcgggtgg	ataaggctag	agaggcctac	360
aacaaggccc	aggagaaaagt	gtacaagatc	gagaaggaat	tcaacaaaaa	gagcatggaa	420
acctaccaga	aggtggtcgg	cctgagcaaa	cagacaagaa	tcggaaagct	gctcaagagc	480
cagttcacc	tgactacga	caccgaagat	agaatcacca	gcacagtgat	cagccacttt	540
aacaatgata	tgaagacagg	agtgtctgaga	ggcgatagaa	gcctgagaac	atacaagaac	600
agccatcctc	tgctggttag	agccagatct	atgaaggtgt	acgaggaaaa	cggcgactac	660
ttcatcaagt	gggtcaaggg	gatcgtgttt	aagatcgtga	tcagcgccgg	cagcaagcag	720
aaagctaata	tcggcgagct	gaaatctgtg	cttatcaaca	tcctgaatgg	ccactataag	780
gtgtgcgaca	gcagcatcag	cctgaataaa	gacctgatcc	tgaacctgtc	tctgaacatc	840
cctgtgtcta	aggagaacgt	gttcgtgccc	ggcagagtgg	tgggcgtgga	cctgggcctg	900
aagatcccag	cctatgtgtc	cctgaacgac	acaccctaca	tcaagaaggg	catcggaac	960
atcgacgatt	tcctgaaggt	gcggacccaa	ctgcagagcc	agagaaagag	actgcaaaag	1020
accctggaat	gtacctccgg	cggcaagggc	agaaacaaga	agctgaaggg	actggacaga	1080
ctgaaggcca	aggaaaaaaa	cttcgtgaac	acctacaacc	acttcctgag	caagaaaatc	1140
atccagtttg	ccgtgaagaa	caacgccggc	gtgatccacc	tggaagagct	gcagttcgac	1200
aagctgaaac	ataagtccct	gctgcggaac	tggtcctact	accaactgca	gacaatgac	1260
gagtacaagg	ctgaacggga	aggcatcgag	gtgaagtacg	tggaagccag	ctacaccagc	1320
cagacctgta	gcaagtgcgg	ccactacgag	gaaggccagc	gggtgctgca	agacaccttc	1380
acctgcaaga	acaaagagtg	caagggttac	gtgcacaagg	ttaacgccga	cttcaacgcc	1440
tctcagaata	tcgccaagtc	taccgacatc	atccggtgca	ccgagatggc	caagaacaac	1500
gatatcgaga	agaatgcc					1518

<210> 156

<211> 1518

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 156

atgagcacc	ctctgcagca	gcctcaccag	aaatccaaga	agaccagcca	gatgataaca	60
acacgcaagt	tcaagctcgc	catcgtttagc	gacaaccgga	acgaggccta	cagcttcac	120
cggaacgaga	tcagaaacca	gaacaaggcc	ctgaacgccg	cttataacca	cctgtacttc	180
gagcacatcg	caaccgagaa	actgaagcac	agcgatgagg	aataccagaa	gcacctgaca	240
aagtacagag	aggtggccac	caacaagtac	caggactacc	tgaaggtgaa	agaaaaagtg	300
aacgacagca	aggacgacga	gaagctgcag	aagcgggtgg	ataaggctag	agaggcctac	360
aacaaggccc	aggagaaaagt	gtacaagatc	gagaaggaat	tcaacaaaaa	gagcatggaa	420
acctaccaga	aggtggtcgg	cctgagcaaa	cagacaagaa	tcggaaagct	gctcaagagc	480

cagttcaccc	tgcactacga	caccgaagat	agaatcacca	gcacagtgat	cagccacttt	540
aacaatgata	tgaagacagg	agtgtctgaga	ggcgatagaa	gcctgagaac	atacaagaac	600
agccatcctc	tgctggtttag	agccagatct	atgaaggtgt	acgaggaaaa	cggcgactac	660
ttcatcaagt	gggtcaaggg	gatcgtgttt	aagatcgtga	tcagcgccgg	cagcaagcag	720
aaagctaata	tcggcgagct	gaaatctgtg	cttatcaaca	tcctgaatgg	ccactataag	780
gtgtgcgaca	gcagcatcag	cctgaataaa	gacctgatcc	tgaacctgtc	tctgaacatc	840
cctgtgtcta	aggagaacgt	gttcgtgccc	ggcagagtgg	tgggcgtggc	cctgggcctg	900
aagatcccag	cctatgtgtc	cctgaacgac	acaccctaca	tcaagaaggg	catcggaac	960
atcgacgatt	tcctgaaggt	gcggacccaa	ctgcagagcc	agagaaagag	actgcaaaag	1020
accctggaat	gtacctccgg	cggcaagggc	agaaacaaga	agctgaaggg	actggacaga	1080
ctgaaggcca	aggaaaaaaa	cttcgtgaac	acctacaacc	acttcctgag	caagaaaatc	1140
atccagtttg	ccgtgaagaa	caacgccggc	gtgateccacc	tggccgagct	gcagttcgac	1200
aagctgaaac	ataagtccct	gctgcggaac	tggtcctact	accaactgca	gacaatgate	1260
gagtacaagg	ctgaacggga	aggcatcgag	gtgaagtacg	tggacgccag	ctacaccagc	1320
cagacctgta	gcaagtgcgg	ccactacgag	gaaggccagc	gggtgctgca	agacaccttc	1380
acctgcaaga	acaaagagtg	caagggttac	gtgcacaagg	ttaacgccgc	cttcaacgcc	1440
tctcagaata	tcgccaagtc	taccgacatc	atccggtgca	ccgagatggc	caagaacaac	1500
gatatcgaga	agaatgcc					1518

<210> 157

<211> 7494

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述: 合成多核苷酸

<400> 157

gcgatcgcg	ctcccgacat	cttggaccat	tagctccaca	ggtatcttct	tcctctagt	60
ggtcataaca	gcagcttcag	ctacctctca	attcaaaaa	cccctcaaga	ccggtttaga	120
ggccccaagg	ggttatgcta	tcaatcgttg	cgttacacac	acaaaaaacc	aacacacatc	180
catcttcgat	ggatagcgat	tttattatct	aactgctgat	cgagtgtagc	cagatctagt	240
aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	300
cggtaaatgg	cccgccctggc	tgaccgcca	acgacccccg	cccattgacg	tcaataatga	360
cgtatgttcc	catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	420
tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaat	acgcccccta	480
ttgacgtcaa	tgacggtaaa	tggcccgcc	ggcattatgc	ccagtacatg	accttatggg	540
actttcctac	ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	ctgatgcggt	600
tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	660
acccatttga	cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	720
gtcgtaaaca	ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	780
atataagcag	agctggttta	gtgaaccgtc	agatcagatc	tttgtcgatc	ctaccatcca	840

ctcgacacac	ccgccagcgg	ccgcgccacc	atggccccta	agaagaaaag	aaaggtggac	900
tacaaggacc	acgacggcga	ctacaaagat	cacgatatcg	actacaagga	cgacgatgat	960
aagatgagca	ccccctctgca	gcagcctcac	cagaaatcca	agaagaccag	ccagatgata	1020
acaacacgca	agttcaagct	cgccatcggt	agcgacaacc	ggaacgaggc	ctacagcttc	1080
atccggaacg	agatcagaaa	ccagaacaag	gccctgaacg	ccgcttataa	ccacctgtac	1140
ttcgagcaca	tcgcaaccga	gaaactgaag	cacagcgatg	aggaatacca	gaagcacctg	1200
acaaagtaca	gagaggtggc	caccaacaag	taccaggact	acctgaaggt	gaaagaaaaa	1260
gtgaacgaca	gcaaggacga	cgagaagctg	cagaagcggg	tggataaggc	tagagaggcc	1320
tacaacaagg	cccaggagaa	agtgtacaag	atcgagaagg	aattcaacaa	aaagagcatg	1380
gaaacctacc	agaaggtggt	cggcctgagc	aaacagacaa	gaatcggaaa	gctgctcaag	1440
agccagttca	ccctgcacta	cgacaccgaa	gatagaatca	ccagcacagt	gatcagccac	1500
tttaacaatg	atatgaagac	aggagtgtctg	agaggcgata	gaagcctgag	aacatacaag	1560
aacagccatc	ctctgtctgt	tagagccaga	tctatgaagg	tgtacgagga	aaacggcgac	1620
tacttcatca	agtgggtcaa	ggggatcgctg	tttaagatcg	tgatcagcgc	cggcagcaag	1680
cagaaagcta	atatcggcga	gctgaaatct	gtgcttatca	acatcctgaa	tggccactat	1740
aaggtgtgcg	acagcagcat	cagcctgaat	aaagacctga	tcctgaacct	gtctctgaac	1800
atccctgtgt	ctaaggagaa	cgtgttcgtg	cccggcagag	tgggtggcgt	ggacctgggc	1860
ctgaagatcc	cagcctatgt	gtccctgaac	gacacaccct	acatcaagaa	gggcacgcgc	1920
aacatcgacg	atttctgtaa	ggtgcggacc	caactgcaga	gccagagaaa	gagactgcaa	1980
aagaccttgg	aatgtacctc	cggcggcaag	ggcagaaaaca	agaagctgaa	gggactggac	2040
agactgaagg	ccaaggaaaa	aaacttcgtg	aacacctaca	accacttcct	gagcaagaaa	2100
atcatccagt	ttgccgtgaa	gaacaacgcc	ggcgtgatcc	acctggaaga	gctgcagttc	2160
gacaagctga	aacataagtc	cctgctgcgg	aactggctct	actaccaact	gcagacaatg	2220
atcgagtaca	aggctgaacg	ggaaggcatc	gaggtgaagt	acgtggacgc	cagctacacc	2280
agccagacct	gtagcaagtg	cggccactac	gaggaaggcc	agcgggtgct	gcaagacacc	2340
ttcacctgca	agaacaaaga	gtgcaagggc	tacgtgcaca	aggttaacgc	cgacttcaac	2400
gcctctcaga	atatcgccaa	gtctaccgac	atcatccggt	gcaccgagat	ggccaagaac	2460
aacgatatcg	agaagaatgc	caaaagacct	gccgctacaa	agaaggccgg	ccaggccaag	2520
aaaaagaagg	gtcgacttga	cgcgttgata	tcaacaagtt	tgtacaaaaa	agcaggctac	2580
aaagaggcca	gcggttccgg	acgggctgac	gcattggacg	attttgatct	ggatatgctg	2640
ggaagtgacg	ccctcgatga	ttttgacctt	gacatgcttg	gttcggatgc	ccttgatgac	2700
tttgacctcg	acatgctcgg	cagtgcgcgc	cttgatgatt	tcgacctgga	catgctgatt	2760
aactctagaa	gttccggatc	tggtagccag	tacctgcccc	acaccgacga	ccggcaccgg	2820
atcgaggaaa	agcgggaagcg	gacctacgag	acattcaaga	gcacatgaa	gaagtccccc	2880
ttcagcggcc	ccaccgaccc	tagacctcca	cctagaagaa	tcgccgtgcc	cagcagatcc	2940
agcgcacgcg	tgccaaaacc	tgccccccag	ccttaccctt	tcaccagcag	cctgagcacc	3000
atcaactacg	acgagttccc	taccatggtg	ttccccagcg	gccagatctc	tcaggcctct	3060
gctctggctc	cagccccctc	tcaggtgctg	cctcaggctc	ctgctcctgc	accagctcca	3120
gccatggtgt	ctgcactggc	tcaggcacca	gcacccgtgc	ctgtgctggc	tcctggacct	3180

ccacaggctg	tggctccacc	agcccctaaa	cctacacagg	ccggcgaggg	cacactgtct	3240
gaagctctgc	tgcagctgca	gttcgacgac	gaggatctgg	gagccctgct	gggaaacagc	3300
accgatcctg	ccgtgttcac	cgacctggcc	agcgtggaca	acagcgagtt	ccagcagctg	3360
ctgaaccagg	gcatccctgt	ggcccctcac	accaccgagc	ccatgctgat	ggaatacccc	3420
gaggccatca	cccggctcgt	gacaggcgct	cagaggcctc	ctgatccagc	tcctgcccct	3480
ctgggagcac	caggcctgcc	taatggactg	ctgtctggcg	acgaggactt	cagctctatc	3540
gccgatatgg	atttctcagc	cttgctgggc	tctggcagcg	gcagccggga	ttccagggaa	3600
gggatgtttt	tgccgaagcc	tgaggccggc	tccgctatta	gtgacgtgtt	tgagggccgc	3660
gaggtgtgcc	agccaaaacg	aatccggcca	tttcatcctc	caggaagtcc	atgggccaac	3720
cgcccactcc	ccgccagcct	cgcaccaaca	ccaaccggtc	cagtacatga	gccagtcggg	3780
tactgaccc	cggcaccagt	ccctcagcca	ctagatccag	cgcccgcagt	gactcccag	3840
gccagtcacc	tgttgaaga	tcccgatgaa	gagacgagcc	aggctgtcaa	agcccttcgg	3900
gagatggccg	atactgtgat	tccccagaag	gaagaggctg	caatctgttg	ccaaatggac	3960
ctttcccatc	cgcccccaag	gggccatctg	gatgagctga	caaccacact	tgagtccatg	4020
accgaggatc	tgaacctgga	ctcaccctg	accccggaat	tgaacgagat	tctggatacc	4080
ttcctgaacg	acgagtgcct	cttgcctgcc	atgcatatca	gcacaggact	gtccatcttc	4140
gacacatctc	tgttttaaat	aggatccgca	ggcctctgct	agcttgactg	actgagatac	4200
agcgtacctt	cagctcacag	acatgataag	atacattgat	gagtttggac	aaaccacaac	4260
tagaatgcag	tgaaaaaaat	gctttatttg	tgaattttgt	gatgctattg	ctttattttgt	4320
aaccattata	agctgcaata	aacaagttaa	caacaacaat	tgcattcatt	ttatgtttca	4380
ggttcagggg	gaggtgtggg	aggtttttta	aagcaagtaa	aacctctaca	aatgtggtat	4440
tggcccatct	ctatcgggat	cgtagcataa	ccccttgggg	cctctaaacg	ggtcttgagg	4500
ggttttttgt	gcccctcggg	ccggattgct	atctaccggc	attggcgag	aaaaaatgc	4560
ctgatgcgac	gctgcgcgtc	ttatactccc	acatatgcca	gattcagcaa	cggatacggc	4620
ttccccaact	tgcccacttc	catactgttc	ctccttacca	gaaatttatc	cttaaggctg	4680
tcagctatcc	tgcaggcgat	ctctcgattt	cgatcaagac	attcctttta	tggtcttttc	4740
tggacaccac	taggggtcag	aagtagttca	tcaaactttc	ttccctccct	aatctcattg	4800
gttaccttgg	gctatcgaaa	cttaattaac	cagtcaagtc	agctacttgg	cgagatcgac	4860
ttgtctgggt	ttcgactacg	ctcagaattg	cgtcagtcaa	gttcgatctg	gtccttgcta	4920
ttgcaccgt	tctccgatta	cgagtttcat	ttaaattcatg	tgagcaaaaag	gccagcaaaa	4980
ggccaggaac	cgtaaaaaag	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	5040
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aaccgcagag	gactataaag	5100
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	5160
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	5220
ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	5280
ccccgttcag	cccgaaccgt	gcgccttctc	cggtaactat	cgtcttgagt	ccaaccgggt	5340
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	5400
tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	5460
agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	5520

ttgatccggc	aaacaaacca	ccgctggtag	cgggtggtttt	tttgtttgca	agcagcagat	5580
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	5640
tcagtggaac	gaaaactcac	gttaagggat	tttggtcatg	agattatcaa	aaaggatctt	5700
cacctagatc	cttttaaat	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	5760
aacttgggtct	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	5820
atttcgttca	tccatagtgt	catttaaat	tccgaactct	ccaaggccct	cgtcggaaaa	5880
tcttcaaacc	tttcgtccga	tccatcttgc	aggctacctc	tcgaacgaac	tatcgcaagt	5940
ctcttggccg	gccttgcgcc	ttggctattg	cttggcagcg	cctatcgcca	ggtattactc	6000
caatcccga	tatccgagat	cgggatcacc	cgagagaagt	tcaacctaca	tcctcaatcc	6060
cgatctatcc	gagatccgag	gaatatcgaa	atcggggcgc	gcctgggtgta	ccgagaacga	6120
tcctctcagt	gcgagtctcg	acgatccata	tcgttgcttg	gcagtcagcc	agtcggaatc	6180
cagcttggga	cccaggaagt	ccaatcgta	gatattgtac	tcaagcctgg	tcacggcagc	6240
gtaccgatct	gtttaaacct	agatattgat	agtctgatcg	gtcaacgtat	aatcgagtcc	6300
tagcttttgc	aaacatctat	caagagacag	gatcagcagg	aggctttcgc	atgagtattc	6360
aacattttccg	tgtcgccctt	attccctttt	ttgcggcatt	ttgccttctc	gtttttgctc	6420
accagaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	gttggtgctg	cgagtgggtt	6480
acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	ttttcgcccc	gaagaacgct	6540
ttccaatgat	gagcactttt	aaagtcttgc	tatgtggcgc	ggtattatcc	cgtattgacg	6600
ccgggcaaga	gcaactcggt	cgccgcatac	actattctca	gaatgacttg	gttgagtatt	6660
caccagtcac	agaaaagcat	cttacggatg	gcatgacagt	aagagaatta	tgcagtgtctg	6720
ccataaccat	gagtataaac	actgcggcca	acttacttct	gacaacgatt	ggaggaccga	6780
aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	aactcgcctt	gatcggtggg	6840
aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	cctgtagcaa	6900
tggcaacaac	cttgcgtaaa	ctattaactg	gcgaactact	tactctagct	tcccggcaac	6960
agttgataga	ctggatggag	gcggataaag	ttgcaggacc	acttctgcgc	tcggcccttc	7020
cggctggctg	gtttattgct	gataaatctg	gagccggtga	gcgtgggtct	cgcggtatca	7080
ttgcagcact	ggggccagat	ggtaaagcct	cccgtatcgt	agttatctac	acgacgggga	7140
gtcaggcaac	tatggatgaa	cgaaatagac	agatcgctga	gataggtgcc	tactgatta	7200
agcattggta	accgattcta	ggtgcattgg	cgcagaaaaa	aatgcctgat	gcgacgtgc	7260
gcgtcttata	ctcccacata	tgccagattc	agcaacggat	acggcttccc	caacttgccc	7320
acttccatac	gtgtcctcct	taccagaaat	ttatccttaa	gatcccgaat	cgtttaaaact	7380
cgactctggc	tctatcgaat	ctccgtcggt	tcgagcttac	gcgaacagcc	gtggcgctca	7440
tttgctcgtc	gggcatcgaa	tctcgtcagc	tatcgtcagc	ttaccttttt	ggca	7494

<210> 158

<211> 7494

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 158

gcgatcgcg	ctcccacat	cttggacat	tagctccaca	ggtatcttct	tcctctagt	60
ggtcataaca	gcagcttcag	ctacctctca	attcaaaaaa	cccctcaaga	cccgtttaga	120
ggccccaagg	ggttatgcta	tcaatcggtg	cgttacacac	acaaaaaacc	aacacacatc	180
catcttcgat	ggatagcgat	tttattatct	aactgctgat	cgagtgtagc	cagatctagt	240
aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	acataactta	300
cggtaaatgg	cccgccctggc	tgaccgcccc	acgacccccg	cccattgacg	tcaataatga	360
cgtatgttcc	catagtaacg	ccaataggga	ctttccattg	acgtcaatgg	gtggagtatt	420
tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaaagt	acgcccccta	480
ttgacgtcaa	tgacggtaaa	tggcccgccct	ggcattatgc	ccagtacatg	accttatggg	540
actttccctac	ttggcagtac	atctacgtat	tagtcacgc	tattaccatg	ctgatgcggt	600
tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	ccaagtctcc	660
accccatgta	cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	tttccaaaat	720
gtcgtaacaa	ctccgcccc	ttgacgcaaa	tgggcggtag	gcgtgtacgg	tgggaggtct	780
atataagcag	agctggttta	gtgaaccgtc	agatcagatc	tttgtcgatc	ctaccatcca	840
ctcgacacac	ccgccagcgg	ccgcgccacc	atggccccta	agaagaaaag	aaaggtggac	900
tacaaggacc	acgacggcga	ctacaaagat	cacgatatcg	actacaagga	cgacgatgat	960
aagatgagca	cccctctgca	gcagcctcac	cagaaatcca	agaagaccag	ccagatgata	1020
acaacacgca	agttcaagct	cgccatcggt	agcgacaacc	ggaacgaggc	ctacagcttc	1080
atccggaacg	agatcagaaa	ccagaacaag	gccctgaacg	ccgcttataa	ccacctgtac	1140
ttcgagcaca	tcgcaaccga	gaaactgaag	cacagcgatg	aggaatacca	gaagcacctg	1200
acaaagtaca	gagaggtggc	caccaacaag	taccaggact	acctgaaggt	gaaagaaaaa	1260
gtgaacgaca	gcaaggacga	cgagaagctg	cagaagcggg	tggataaggc	tagagaggcc	1320
tacaacaagg	cccaggagaa	agtgtacaag	atcgagaagg	aattcaacaa	aaagagcatg	1380
gaaacctacc	agaaggtggt	cggcctgagc	aaacagacaa	gaatcgga	gctgctcaag	1440
agccagttca	ccctgcacta	cgacaccgaa	gatagaatca	ccagcacagt	gatcagccac	1500
tttaacaatg	atatgaagac	aggagtgtg	agaggcgata	gaagcctgag	aacatacaag	1560
aacagccatc	ctctgctggt	tagagccaga	tctatgaagg	tgtacgagga	aaacggcgac	1620
tacttcatca	agtgggtcaa	ggggatcgtg	tttaagatcg	tgatcagcgc	cggcagcaag	1680
cagaaagcta	atatcggcga	gctgaaatct	gtgcttatca	acatcctgaa	tggccactat	1740
aaggtgtgcg	acagcagcat	cagcctgaat	aaagacctga	tcctgaacct	gtctctgaac	1800
atccctgtgt	ctaaggagaa	cgtgttcgtg	ccggcgagag	tgggtggcgt	ggccctgggc	1860
ctgaagatcc	cagcctatgt	gtccctgaac	gacacacct	acatcaagaa	gggcatcggc	1920
aacatcgacg	atttcctgaa	ggtgcggacc	caactgcaga	gccagagaaa	gagactgcaa	1980
aagaccctgg	aatgtacctc	cggcggcaag	ggcagaaaca	agaagctgaa	gggactggac	2040
agactgaagg	ccaaggaaaa	aaacttcgtg	aacacctaca	accacttcct	gagcaagaaa	2100
atcatccagt	ttgccgtgaa	gaacaacgcc	ggcgtgatcc	acctggccga	gctgcagttc	2160
gacaagctga	aacataagtc	cctgctgcgg	aactggtcct	actaccaact	gcagacaatg	2220
atcgagtaca	aggctgaacg	ggaaggcatc	gaggtgaagt	acgtggacgc	cagctacacc	2280

agccagacct	gtagcaagtg	cggccactac	gaggaaggcc	agcgggtgct	gcaagacacc	2340
ttcacctgca	agaacaaaga	gtgcaagggc	tacgtgcaca	aggttaacgc	cgccttcaac	2400
gcctctcaga	atatcgccaa	gtctaccgac	atcatccggt	gcaccgagat	ggccaagaac	2460
aacgatatcg	agaagaatgc	caaaagacct	gccgctacaa	agaaggccgg	ccaggccaag	2520
aaaaagaagg	gtcgacttga	cgcgttgata	tcaacaagtt	tgtacaaaaa	agcaggctac	2580
aaagaggcca	gcggttccgg	acgggctgac	gcattggacg	atthttgatct	ggatatgctg	2640
ggaagtgacg	ccctcgatga	ttttgacctt	gacatgcttg	gttcggatgc	ccttgatgac	2700
tttgacctcg	acatgctcgg	cagtgcgccc	cttgatgatt	tcgacctgga	catgctgatt	2760
aactctagaa	gttccggatc	tggtagccag	tacctgcccg	acaccgacga	ccggcaccgg	2820
atcgaggaaa	agcgggaagcg	gacctacgag	acattcaaga	gcatcatgaa	gaagtcccc	2880
ttcagcggcc	ccaccgaccc	tagacctcca	cctagaagaa	tcgccgtgcc	cagcagatcc	2940
agcgcacgcg	tgccaaaacc	tgccccccag	cettaccctt	tcaccagcag	cctgagcacc	3000
atcaactacg	acgagttccc	taccatgggtg	ttccccagcg	gccagatctc	tcaggcctct	3060
gctctggctc	cagccccctc	tcagggtgctg	cctcaggctc	ctgctctctgc	accagctcca	3120
gccatgggtg	ctgcactggc	tcaggcacca	gcacccgtgc	ctgtgctggc	tcctggacct	3180
ccacaggctg	tggtccacc	agccccataa	cctacacagg	ccggcgaggg	cacactgtct	3240
gaagctctgc	tgcagctgca	gttcgacgac	gaggatctgg	gagccctgct	gggaaacagc	3300
accgatcctg	ccgtgttcac	cgacctggcc	agcgtggaca	acagcgagtt	ccagcagctg	3360
ctgaaccagg	gcatccctgt	ggccccctac	accaccgagc	ccatgctgat	ggaatacccc	3420
gaggccatca	cccggctcgt	gacaggcgct	cagaggcctc	ctgatccagc	tcctgcccc	3480
ctgggagcac	caggcctgcc	taatggactg	ctgtctggcg	acgaggactt	cagctctatc	3540
gccgatatgg	atttctcagc	cttgctgggc	tctggcagcg	gcagccggga	ttccagggaa	3600
gggatgtttt	tgccgaagcc	tgaggccggc	tccgctatta	gtgacgtgtt	tgagggccgc	3660
gaggtgtgcc	agccaaaacg	aatccggcca	tttcatctc	caggaagtcc	atgggccaac	3720
cgccccactc	ccgccagcct	cgcaccaaca	ccaaccggtc	cagtacatga	gccagtcggg	3780
tactgaccc	cggcaccagt	ccctcagcca	ctagatccag	cgccccgagt	gactcccag	3840
gccagtcacc	tgttggaaga	tcccgatgaa	gagacgagcc	aggctgtcaa	agcccttcgg	3900
gagatggccg	atactgtgat	tccccagaag	gaagaggctg	caatctgtgg	ccaaatggac	3960
ctttcccatc	cgcccccaag	gggccatctg	gatgagctga	caaccacact	tgagtccatg	4020
accgaggatc	tgaacctgga	ctcacccctg	accccggaat	tgaacgagat	tctggatacc	4080
ttcctgaacg	acgagtgcct	cttgcatgcc	atgcatatca	gcacaggact	gtccatcttc	4140
gacacatctc	tgttttaaat	aggatccgca	ggcctctgct	agcttgactg	actgagatac	4200
agcgtacctt	cagctcacag	acatgataag	atacattgat	gagtttggac	aaaccacaac	4260
tagaatgcag	tgaaaaaaat	gctttatttg	tgaattttgt	gatgctattg	ctttattttgt	4320
aaccattata	agctgcaata	aacaagttaa	caacaacaat	tgcattcatt	ttatgtttca	4380
ggttcagggg	gaggtgtggg	aggtttttta	aagcaagtaa	aacctctaca	aatgtggtat	4440
tggcccatct	ctatcggtat	cgtagcataa	ccccttgggg	cctctaaacg	ggtcttgagg	4500
ggttttttgt	gccccctcgg	ccggattgct	atctaccggc	attggcgag	aaaaaatgc	4560
ctgatgcgac	gctgcgcgtc	ttatactccc	acatatgcc	gattcagcaa	cggatacggc	4620

ttccccaact	tgcccacttc	catacgtgtc	ctccttacca	gaaattttatc	cttaaggtcg	4680
tcagctatcc	tgcaggcgat	ctctcgattt	cgatcaagac	attcctttta	tggtcttttc	4740
tggacaccac	taggggtcag	aagtagttca	tcaaactttc	ttccctccct	aatctcattg	4800
gttaccttgg	gctatcgaaa	cttaattaac	cagtcaagtc	agctacttgg	cgagatcgac	4860
ttgtctgggt	ttcgactacg	ctcagaattg	cgtcagtcaa	gttcgatctg	gtccttgcta	4920
ttgcacccgt	tctccgatta	cgagtttcat	ttaaatacatg	tgagcaaaaag	gccagcaaaa	4980
ggccaggaac	cgtaaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	5040
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	5100
ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	ccctgccgct	5160
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgctttctc	atagctcacg	5220
ctgtaggtat	ctcagttcgg	tgtaggtcgt	tcgctccaag	ctgggctgtg	tgcacgaacc	5280
ccccgttcag	cccgaaccgt	gcgccttate	cggtaaactat	cgtcttgagt	ccaacccggt	5340
aagacacgac	ttatcgccac	tggcagcagc	cactggtaac	aggattagca	gagcgaggta	5400
tgtaggcggt	gctacagagt	tcttgaagtg	gtggcctaac	tacggctaca	ctagaagaac	5460
agtatttggg	atctgcgctc	tgctgaagcc	agttaccttc	ggaaaaagag	ttggtagctc	5520
ttgatccggc	aaacaaacca	ccgctggtag	cggtggtttt	tttgtttgca	agcagcagat	5580
tacgcgcaga	aaaaaaggat	ctcaagaaga	tcctttgatc	ttttctacgg	ggtctgacgc	5640
tcagtggaac	gaaaactcac	gttaagggat	tttggtcattg	agattatcaa	aaaggatctt	5700
cacctagatc	cttttaaaatt	aaaaatgaag	ttttaaatca	atctaaagta	tatatgagta	5760
aacttgggtc	gacagttacc	aatgcttaat	cagtgaggca	cctatctcag	cgatctgtct	5820
atttcgttca	tccatagttg	catttaaaatt	tccgaactct	ccaaggccct	cgtcggaaaa	5880
tcttcaaacc	tttcgtccga	tccatcttgc	aggctacctc	tcgaacgaac	tatcgcaagt	5940
ctcttggccg	gccttgcgcc	ttggctattg	cttggcagcg	cctatcgcca	ggtattactc	6000
caatccccgaa	tatccgagat	cgggatcacc	cgagagaagt	tcaacctaca	tcctcaatcc	6060
cgatctatcc	gagatccgag	gaatatcgaa	atcggggcgc	gcctggtgta	ccgagaacga	6120
tcctctcagt	gcgagtctcg	acgatccata	tcgttgcttg	gcagtcagcc	agtcggaatc	6180
cagcttggga	cccaggaagt	ccaatcgtca	gatattgtac	tcaagcctgg	tcacggcagc	6240
gtaccgatct	gtttaaacct	agatattgat	agtctgatcg	gtcaacgtat	aatcgagtcc	6300
tagcttttgc	aaacatctat	caagagacag	gatcagcagg	aggctttcgc	atgagtattc	6360
aacattttccg	tgtcgccctt	attccctttt	ttgcggcatt	ttgccttcct	gtttttgctc	6420
acccagaaac	gctggtgaaa	gtaaaagatg	ctgaagatca	gttgggtgcg	cgagtggggt	6480
acatcgaact	ggatctcaac	agcggtaaga	tccttgagag	ttttcgcccc	gaagaacgct	6540
ttccaatgat	gagcactttt	aaagttctgc	tatgtggcgc	ggtattatcc	cgtattgacg	6600
ccgggcaaga	gcaactcggt	cgccgcatac	actattctca	gaatgacttg	gttgagtatt	6660
caccagtcac	agaaaagcat	cttacggatg	gcatgacagt	aagagaatta	tgcagtgctg	6720
ccataaccat	gagtgataac	actgcggcca	acttacttct	gacaacgatt	ggaggaccga	6780
aggagctaac	cgcttttttg	cacaacatgg	gggatcatgt	aactcgcctt	gatcgttggg	6840
aaccggagct	gaatgaagcc	ataccaaacg	acgagcgtga	caccacgatg	cctgtagcaa	6900
tggcaacaac	cttgcgtaaa	ctattaactg	gcgaactact	tactctagct	tcccggcaac	6960

```

agttgataga ctggatggag gcggataaag ttgcaggacc acttctgcgc tcggcccttc 7020
cggctggctg gtttattgct gataaatctg gagccggtga gcgtgggtct cgcggtatca 7080
ttgcagcact ggggccagat ggtaagccct cccgtatcgt agttatctac acgacgggga 7140
gtcaggcaac tatggatgaa cgaaatagac agatcgctga gataggtgcc tcaactgatta 7200
agcattggta accgattcta ggtgcattgg cgcagaaaaa aatgcctgat gcgacgctgc 7260
gcgtcttata ctcccacata tgccagattc agcaacggat acggcttccc caacttgccc 7320
acttccatac gtgtcctcct taccagaaat ttatccttaa gatcccgaat cgtttaaact 7380
cgactctggc tctatcgaat ctccgtcgtt tcgagcttac gcgaacagcc gtggcgctca 7440
tttgctcgtc gggcatcgaa tctcgtcagc tatcgtcagc ttaccttttt ggca 7494

```

<210> 159

<211> 7512

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 159

```

gcgatcgcg gcgtccacat cttggaccat tagctccaca ggtatcttct tccctctagt 60
ggtcataaca gcagcttcag ctacctctca attcaaaaaa cccctcaaga cccgtttaga 120
ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180
catcttcgat ggatagcgat ttattatct aactgctgat cgagtgtagc cagatctagt 240
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300
cggtaaattg cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga 360
cgtatgttcc catagtaacg ccaataggga ctttcattg acgtcaatgg gtggagtatt 420
tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt acgcccccta 480
ttgacgtcaa tgacggtaaa tggccgcct ggcattatgc ccagtacatg acctatggg 540
actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg ctgatgcggt 600
tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt ccaagtctcc 660
acccattga cgtcaatggg agtttgttt ggcacaaaaa tcaacgggac tttccaaaat 720
gtcgtaaaca ctccgcccc ttgacgcaaa tgggcggtag gcgtgtacgg tgggaggtct 780
atataagcag agctggttta gtgaaccgtc agatcagatc tttgtcgatc ctaccatcca 840
ctcgacacac ccgccagcgg ccgcgccacc atggccccta agaagaaaag aaaggtggac 900
tacaaggacc acgacggcga ctacaaagat cagcatatcg actacaagga cgacgatgat 960
aaggacgcat tggacgattt tgatctggat atgctgggaa gtgacgccct cgatgatttt 1020
gaccttgaca tgcttggttc ggatgccctt gatgactttg acctcgacat gctcggcagt 1080
gacgcccttg atgatttcca cctggacatg ctgattaact ctagaagttc cggatctggt 1140
agccagtacc tgcccagcac cgacgaccgg caccggatcg aggaaaagcg gaagcggacc 1200
tacgagacat tcaagagcat catgaagaag tcccccttca gcggccccac cgaccctaga 1260
cctccacctt gaagaatcgc cgtgcccagc agatccagcg ccagcgtgcc aaaacctgcc 1320
ccccagcctt accccttcac cagcagcctg agcaccatca actacgacga gttccctacc 1380

```

atggtgttcc	ccagcggcca	gatctctcag	gcctctgctc	tggtccagc	ccctcctcag	1440
gtgctgcctc	aggctcctgc	tcctgcacca	gctccagcca	tggtgtctgc	actggctcag	1500
gcaccagcac	ccgtgcctgt	gctggctcct	ggacctccac	aggctgtggc	tccaccagcc	1560
cctaaaccta	cacaggccgg	cgagggcaca	ctgtctgaag	ctctgctgca	gctgcagttc	1620
gacgacgagg	atctgggagc	cctgctggga	aacagcaccc	atcctgccgt	gttcaccgac	1680
ctggccagcg	tggacaacag	cgagttccag	cagctgctga	accagggcat	ccctgtggcc	1740
cctcacacca	ccgagcccat	gctgatggaa	taccccgagg	ccatcacccg	gctcgtgaca	1800
ggcgctcaga	ggcctcctga	tccagctcct	gcccctctgg	gagcaccagg	cctgcctaata	1860
ggactgctgt	ctggcgacga	ggacttcagc	tctatcgccg	atatggattt	ctcagccttg	1920
ctgggctctg	gcagcggcag	ccgggattcc	agggaaggga	tgtttttgcc	gaagcctgag	1980
gccggtccg	ctattagtga	cgtgtttgag	ggccgcgagg	tgtgccagcc	aaaacgaatc	2040
cggccatttc	atcctccagg	aagtccatgg	gccaacgcc	cactccccgc	cagcctcgca	2100
ccaacaccaa	ccggtccagt	acatgagcca	gtcgggtcac	tgaccccgcc	accagtcctt	2160
cagccactag	atccagcgcc	cgcagtgact	cccaggcca	gtcacctgtt	ggaagatccc	2220
gatgaagaga	cgagccaggc	tgtcaaagcc	cttcgggaga	tggccgatac	tgtgattccc	2280
cagaaggaag	aggctgcaat	ctgtggccaa	atggaccttt	cccatccgcc	ccaaggggc	2340
catctggatg	agctgacaac	cacacttgag	tccatgaccg	aggatctgaa	cctggactca	2400
cccctgaccc	cggaaattgaa	cgagattctg	gataccttcc	tgaacgacga	gtgcctcttg	2460
catgccatgc	atatcagcac	aggactgtcc	atcttcgaca	catctctgtt	ttccggagga	2520
tctagcggag	gctcctctgg	ctctgagaca	cctggcacia	gcgagagcgc	aacacctgaa	2580
agcagcgggg	gcagcagcgg	gggtcaatg	agcaccctc	tgcagcagcc	tcaccagaaa	2640
tccaagaaga	ccagccagat	gataacaaca	cgcaagtcca	agctcgccat	cgtagcgac	2700
aaccggaacg	aggcctacag	cttcatccgg	aacgagatca	gaaaccagaa	caaggccctg	2760
aacgccgctt	ataaccacct	gtacttcgag	cacatcgcaa	ccgagaaact	gaagcacagc	2820
gatgaggaat	accagaagca	cctgacaaa	tacagagagg	tggccaccaa	caagtaccag	2880
gactacctga	aggtgaaaga	aaaagtgaac	gacagcaagg	acgacgagaa	gctgcagaag	2940
cgggtggata	aggctagaga	ggcctacaac	aaggcccagg	agaaagtgtg	caagatcgag	3000
aaggaattca	acaaaaagag	catggaaacc	taccagaagg	tggtcggcct	gagcaaacag	3060
acaagaatcg	gaaaagctgct	caagagccag	ttcacctgc	actacgacac	cgaagataga	3120
atcaccagca	cagtgatcag	ccactttaac	aatgatatga	agacaggagt	gctgagaggc	3180
gatagaagcc	tgagaacata	caagaacagc	catcctctgc	tggttagagc	cagatctatg	3240
aaggtgtacg	aggaaaacgg	cgactacttc	atcaagtggg	tcaaggggat	cgtgtttaag	3300
atcgtgatca	gcgccggcag	caagcagaaa	gctaatatcg	gcgagctgaa	atctgtgctt	3360
atcaacatcc	tgaatggcca	ctataaggtg	tgcgacagca	gcacagcct	gaataaagac	3420
ctgatcctga	acctgtctct	gaacatccct	gtgtctaagg	agaacgtgtt	cgtgcccggc	3480
agagtgggtg	gcgtggacct	gggcctgaag	atcccagcct	atgtgtccct	gaacgacaca	3540
ccctacatca	agaagggcag	cggcaacatc	gacgatttcc	tgaaggtgcg	gacccaactg	3600
cagagccaga	gaaagagact	gcaaaagacc	ctggaatgta	cctccggcgg	caagggcaga	3660
aacaagaagc	tgaagggact	ggacagactg	aaggccaagg	aaaaaaactt	cgtgaacacc	3720

tacaaccact	tcctgagcaa	gaaaatcatc	cagtttgccg	tgaagaacaa	cgccggcgtg	3780
atccacctgg	aagagctgca	gttcgacaag	ctgaaacata	agtccctgct	gcggaactgg	3840
tcctactacc	aactgcagac	aatgatcgag	tacaaggctg	aacgggaagg	catcgaggtg	3900
aagtacgtgg	acgccagcta	caccagccag	acctgtagca	agtgcggcca	ctacgaggaa	3960
ggccagcggg	tgctgcaaga	caccttcacc	tgcaagaaca	aagagtgcaa	gggctacgtg	4020
cacaaggtta	acgccgactt	caacgcctct	cagaatatcg	ccaagtctac	cgacatcatc	4080
cggtgcaccg	agatggccaa	gaacaacgat	atcgagaaga	atgccaaaag	acctgccgct	4140
acaaagaagg	ccggccaggc	caagaaaaag	aagtaatgag	gatccgcagg	cctctgctag	4200
cttgactgac	tgagatacag	cgtaccttca	gtcacagac	atgataagat	acattgatga	4260
gtttggacaa	accacaacta	gaatgcagtg	aaaaaatgc	tttatttgtg	aaatttgtga	4320
tgctattgct	ttatttgtaa	ccattataag	ctgcaataaa	caagttaaca	acaacaattg	4380
cattcatttt	atgtttcagg	ttcaggggga	ggtgtgggag	gttttttaaa	gcaagtaaaa	4440
cctctacaaa	tgtggtattg	gccccatctt	atcggtatcg	tagcataacc	ccttggggcc	4500
tctaaacggg	tcttgagggg	ttttttgtgc	ccctcgggcc	ggattgctat	ctaccggcat	4560
tggcgcagaa	aaaaatgcct	gatgcgacgc	tgcgctctt	atactccac	atatgccaga	4620
ttcagcaacg	gatacggtt	ccccaaactg	cccacttcca	tacgtgtcct	ccttaccaga	4680
aatttatcct	taaggtcgtc	agctatcctg	caggcgatct	ctcgatttcg	atcaagacat	4740
tcctttaatg	gtcttttctg	gacaccacta	ggggtcagaa	gtagtccatc	aaactttctt	4800
ccctccctaa	tctcattggg	taccttgggc	tatcgaaact	taattaacca	gtcaagtcag	4860
ctacttggcg	agatcgactt	gtctgggttt	cgactacgct	cagaattgcg	tcagtcaagt	4920
tcgatctggg	ccttgctatt	gcaccggtt	tccgattacg	agtttcattt	aatcatgtg	4980
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	5040
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	5100
cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	5160
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	5220
gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcggtt	gctccaagct	5280
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	5340
tcttgagtcc	aacccggtaa	gacacgactt	atgccactg	gcagcagcca	ctggtaacag	5400
gatttagcaga	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	5460
cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	5520
aaaaagagtt	gtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	5580
tgtttgcaag	cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	5640
ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggtcatgag	5700
attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	5760
ctaaagtata	tatgagtaaa	cttggctctga	cagttaccaa	tgcttaatca	gtgaggcacc	5820
tatctcagcg	atctgtctat	ttcgttcatc	catagttgca	tttaaatttc	cgaactctcc	5880
aaggccctcg	tcgaaaaatc	ttcaaacctt	tcgtccgac	catcttgacg	gctacctctc	5940
gaacgaacta	tcgcaagtct	cttgcccggc	cttgccgctt	ggctattgct	tggcagcgcc	6000
tatcgccagg	tattactcca	atcccgaata	tccgagatcg	ggatcacccg	agagaagttc	6060

```

aacctacatc ctcaatcccc atctatccga gatccgagga atatcgaaat cggggcgcgcg 6120
ctgggtgtacc gagaacgata cttctcagtg gagtctcgac gatccatata gttgcttggc 6180
agtcagccag tcggaatcca gcttgggacc caggaagtcc aatcgtcaga tattgtactc 6240
aagcctggtc acggcagcgt accgatctgt ttaaacctag atattgatag tctgatcggc 6300
caacgtataa tcgagtccta gcttttgcaa acatctatca agagacagga tcagcaggag 6360
gctttcgcat gagtattcaa catttccgtg tcgcccttat tccctttttt gcggcatttt 6420
gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct gaagatcagt 6480
tgggtgcgcg agtgggttac atcgaaactg atctcaacag cggtaagatc cttgagagtt 6540
ttcgccccga agaacgcttt ccaatgatga gcacttttaa agttctgcta tgtggcgcg 6600
tattatcccc tattgacgcc gggcaagagc aactcggtcg ccgcatacac tattctcaga 6660
atgacttggc tgagtattca ccagtcacag aaaagcatct tacggatggc atgacagtaa 6720
gagaattatg cagtgtgcc ataacctga gtgataacac tgcggccaac ttacttctga 6780
caacgattgg aggaccgaag gagtaaccg cttttttgca caacatgggg gatcatgtaa 6840
ctcgccctga tcgttgggaa ccggagctga atgaagccat accaaacgac gagcgtgaca 6900
ccacgatgcc ttagcaatg gcaacaacct tgcgtaaact attactggc gaactactta 6960
ctctagcttc ccggcaacag ttgatagact ggatggaggc ggataaagtt gcaggaccac 7020
ttctgcgctc ggcccttcg gctggctggc ttattgctga taaatctgga gccggtgagc 7080
gtgggtctcg cggatcatt gcagcactgg ggccagatgg taagccctcc cgtatcgtag 7140
ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag atcgtgaga 7200
taggtgcctc actgattaag cattggtaac cgattctagg tgcattggcg cagaaaaaaa 7260
tgcttgatgc gacgctgcgc gtcttatact cccacatatg ccagattcag caacggatac 7320
ggcttcccca acttgccac ttccatacgt gtctcctta ccagaaattt atccttaaga 7380
tcccgaatcg tttaaactcg actctggctc tatcgaatct ccgtcgtttc gagcttacgc 7440
gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc tcgtcagcta tcgtcagctt 7500
accttttttg ca 7512

```

<210> 160

<211> 7512

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 160

```

gcgatcgcgg ctcccgacat cttggaccat tagctccaca ggtatcttct tccctctagt 60
ggtcataaca gcagcttcag ctacctetca attcaaaaaa cccctcaaga cccgtttaga 120
ggccccaagg ggttatgcta tcaatcgttg cgttacacac acaaaaaacc aacacacatc 180
catcttcgat ggatagcgat tttattatct aactgctgat cgagtgtagc cagatctagt 240
aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt acataactta 300
cggtaaatgg cccgcctggc tgaccgcca acgacccccg cccattgacg tcaataatga 360
cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg gtggagtatt 420

```


tacggtaa	ac	tgcccact	tg	gcagta	catc	aagtgt	atca	tatgcc	aa	acgcccc	ccta	480
ttgacgt	caa	tgacgg	taaa	tggccc	gcct	ggcatt	atgc	ccagta	catg	accttat	ggg	540
actttc	tac	ttggcag	tac	atctac	gtat	tagtca	tcgc	tattacc	atg	ctgatgc	ggt	600
tttggc	agta	catcaat	ggg	cgtggat	agc	ggtttg	actc	acgggg	attt	ccaagt	ctcc	660
accccc	attga	cgtcaat	ggg	agtttg	tttt	ggcacca	aaaa	tcaacg	gggac	tttccaa	aat	720
gtcgtaa	caa	ctccgc	cccc	ttgacg	caaa	tggg	cggtag	gcgtgt	acgg	tgggag	gtct	780
atataag	cag	agctgg	ttta	gtgaacc	gtc	agatca	gatac	tttgtc	gatac	ctaccat	ccta	840
ctcgaca	cac	ccgccag	cgg	ccgcgc	cacc	atggccc	ccta	agaagaa	aaag	aaaggt	ggac	900
tacaagg	acc	acgacg	gcga	ctacaa	agat	cacgata	tatcg	actaca	agga	cgacga	tgat	960
aaggac	gcat	tggacg	attt	tgatct	ggat	atgctg	ggga	gtgacg	ccct	cgatga	tttt	1020
gaccttg	aca	tgcttg	gttc	ggatgc	ccct	gatgac	tttg	acctgc	acat	gctcgg	cagt	1080
gacgcc	cttg	atgatt	tcga	cctggac	atg	ctgatta	act	ctagaag	ttc	cggatct	ggt	1140
agccagt	tacc	tgcccga	cac	cgacga	ccgg	caccgga	tgc	aggaaa	agcg	gaagcg	gacc	1200
tacgaga	cat	tcaagag	cat	catgaag	aag	tcccc	cttca	gcggccc	ccac	cgacc	ctaga	1260
cctccac	cta	gaagaat	tcgc	cgtgccc	cagc	agatcc	cagcg	ccagcgt	gcc	aaaacct	gcc	1320
ccccagc	cctt	acccctt	cac	cagcagc	cctg	agcacc	atca	actacg	acga	gttcc	ctacc	1380
atgggtg	ttcc	ccagcgg	cca	gatctct	cag	gcctct	gctc	tggctcc	cagc	ccctc	ctcag	1440
gtgctgc	cctc	aggctc	ctgc	tcctgc	acca	gctccag	cca	tgggtgt	ctgc	actgg	ctcag	1500
gcaccag	cac	ccgtgc	ctgt	gctgg	ctcct	ggacct	ccac	aggctgt	ggc	tccacc	agcc	1560
cctaaac	ccta	cacagg	ccgg	cgaggg	caca	ctgtct	gaag	ctctgt	ctga	gctgc	agttc	1620
gacgacg	agg	atctggg	agc	cctgctg	gga	aacagca	ccg	atcctg	ccgt	gttca	ccgac	1680
ctggcc	cagcg	tggaca	aacag	cgagtt	ccag	cagctgt	ga	accagg	gcat	ccctgt	ggcc	1740
cctcac	acca	ccgagccc	at	gctgat	ggaa	tacccc	gagg	ccatcac	ccg	gctcgt	gaca	1800
ggcgct	caga	ggcctc	ctga	tccag	ctcct	gcccct	ctgg	gagcacc	agg	cctgc	cta	1860
ggactg	ctgt	ctggcg	acga	ggactt	cagc	tctatc	gccg	atatgg	attt	ctcagc	cttg	1920
ctggg	ctctg	gcagcgg	cag	ccggg	attcc	aggga	aggga	tgtttt	tgcc	gaagc	ctgag	1980
gccgg	ctccg	ctattag	tga	cgtgtt	ttag	ggccg	cagagg	tgtgcc	agcc	aaaac	gaatc	2040
cggcc	atttc	atcctcc	cagg	aagtcc	atgg	gccaa	accgcc	cactccc	cg	cagc	ctcga	2100
ccaac	accaa	ccggtcc	ag	acatga	gcca	gtcggg	gtcac	tgacccc	ggc	accagt	ccct	2160
cagcc	actag	atccag	cgc	cgag	tga	ctc	aggcca	gtcac	ctgtt	ggaaga	tccc	2220
gatga	agaga	cgagcc	aggc	tgtcaa	agcc	cttcggg	gaga	tggccg	atac	tgtga	tccc	2280
cagaag	gaag	aggctg	caat	ctgtgg	ccaa	atggac	cttt	cccatc	cgcc	cccaag	gggc	2340
catctg	gatg	agctga	caac	cacact	tgag	tccatg	accg	aggat	ctgaa	cctgg	actca	2400
cccctg	accc	cggaa	ttgaa	cgagatt	ctg	gatac	cttc	tgaacg	acga	gtgc	ctctg	2460
catgcc	atgc	atatcag	cac	aggact	gtcc	atcttc	gaca	catct	ctgtt	ttccg	gagga	2520
tctagc	ggag	gtcct	ctg	ctctga	gaca	cctggc	acaa	gcgag	agcg	aacac	ctgaa	2580
agcagc	gggg	gcagcag	cgg	ggggt	caatg	agcacc	ctc	tgcag	cagcc	tcacc	agaaa	2640
tccaaga	aga	ccagcc	agat	gataa	caaca	cgcaag	ttca	agctc	gcat	cgttag	cgcac	2700
aaccgga	acg	aggc	ctacag	cttcat	ccg	aacga	gatca	gaaacc	agaa	caagg	ccctg	2760

aacgccgctt	ataaccacct	gtacttcgag	cacatcgcaa	ccgagaaact	gaagcacagc	2820
gatgaggaat	accagaagca	cctgacaaaag	tacagagagg	tggccaccaa	caagtaccag	2880
gactacctga	agggtgaaaga	aaaagtgaac	gacagcaagg	acgacgagaa	gctgcagaag	2940
cgggttgata	aggctagaga	ggcctacaac	aaggcccagg	agaaagtgtg	caagatcgag	3000
aaggaattca	acaaaaagag	catggaaacc	taccagaagg	tggtcggcct	gagcaaacag	3060
acaagaatcg	gaaagctgct	caagagccag	ttcacctgc	actacgacac	cgaagataga	3120
atcaccagca	cagtgatcag	ccactttaac	aatgatatga	agacaggagt	gctgagaggc	3180
gatagaagcc	tgagaacata	caagaacagc	catcctctgc	tggttagagc	cagatctatg	3240
aaggtgtacg	aggaaaaacgg	cgactacttc	atcaagtggg	tcaaggggat	cgtgtttaag	3300
atcgtgatca	gcgccggcag	caagcagaaa	gctaatatcg	gcgagctgaa	atctgtgctt	3360
atcaacatcc	tgaatggcca	ctataaggtg	tgcgacagca	gcatcagcct	gaataaagac	3420
ctgatcctga	acctgtctct	gaacatccct	gtgtctaagg	agaacgtgtt	cgtgcccggc	3480
agagtgggtg	gcgtggccct	gggctgaag	atcccagcct	atgtgtccct	gaacgacaca	3540
ccctacatca	agaagggcag	cggcaacatc	gacgatttcc	tgaaggtgcg	gacccaactg	3600
cagagccaga	gaaagagact	gcaaaagacc	ctggaatgta	cctccggcgg	caagggcaga	3660
aacaagaagc	tgaagggact	ggacagactg	aaggccaagg	aaaaaaactt	cgtgaacacc	3720
tacaaccact	tcctgagcaa	gaaaatcatc	cagtttgccg	tgaagaacaa	cgccggcgtg	3780
atccacctgg	ccgagctgca	gttcgacaag	ctgaaacata	agtccctgct	gcggaactgg	3840
tcctactacc	aactgcagac	aatgatcgag	tacaaggctg	aacgggaagg	catcgagggtg	3900
aagtacgtgg	acgccagcta	caccagccag	acctgtagca	agtgcggcca	ctacgaggaa	3960
ggccagcggg	tgctgcaaga	caccttcacc	tgcaagaaca	aagagtgcaa	gggctacgtg	4020
cacaaggtta	acgccgcctt	caacgcctct	cagaatatcg	ccaagtctac	cgacatcatc	4080
cgggtgcaccg	agatggccaa	gaacaacgat	atcgagaaga	atgccaaaag	acctgccgct	4140
acaaagaagg	ccggccaggc	caagaaaaag	aagtaatgag	gatccgcagg	cctctgctag	4200
cttgactgac	tgagatacag	cgtaccttca	gctcacagac	atgataagat	acattgatga	4260
gtttggacaa	accacaacta	gaatgcagtg	aaaaaaatgc	tttattttgtg	aaattttgtga	4320
tgctattgct	ttattttgtaa	ccattataag	ctgcaataaa	caagttaaca	acaacaattg	4380
cattcatttt	atgtttcagg	ttcaggggga	ggtgtgggag	gttttttaaa	gcaagtaaaa	4440
cctctacaaa	tgtggtattg	gccccatctct	atcggtatcg	tagcataacc	ccttggggcc	4500
tctaaacggg	tcttgagggg	ttttttgtgc	ccctcgggcc	ggattgctat	ctaccggcat	4560
tggcgcagaa	aaaaatgcct	gatgcgacgc	tgcgcttctt	atactcccac	atatgccaga	4620
ttcagcaacg	gatacggtt	ccccaaactg	cccacttcca	tacgtgtcct	ccttaccaga	4680
aattttatcct	taaggctcgtc	agctatcctg	caggcgatct	ctcgatttcg	atcaagacat	4740
tcctttaatg	gtcttttctg	gacaccacta	ggggtcagaa	gtagttcatc	aaactttctt	4800
ccctccctaa	tctcattggg	taccttgggc	tatcgaaact	taattaacca	gtcaagtcag	4860
ctacttggcg	agatcgactt	gtctgggttt	cgactacgct	cagaattgcg	tcagtcaagt	4920
tcgatctggg	ccttgctatt	gcacccgttc	tccgattacg	agtttcattt	aaatcatgtg	4980
agcaaaaagg	cagcaaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	5040
taggctccgc	ccccctgacg	agcatcacia	aaatcgacgc	tcaagtcaga	ggtggcgaaa	5100

cccgacagga	ctataaagat	accaggcggt	tccccctgga	agctccctcg	tgcgctctcc	5160
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	5220
gctttctcat	agctcacgct	gtaggtatct	cagttcggtg	taggtcggtc	gctccaagct	5280
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	5340
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	5400
gattagcaga	gcgaggtatg	taggcgggtg	tacagagttc	ttgaagtggg	ggcctaacta	5460
cggctacact	agaagaacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	5520
aaaaagagtt	ggtagctctt	gatccggcaa	aaaaccacc	gctggtagcg	gtggtttttt	5580
tgtttgcaag	cagcagatta	cgcgagaaa	aaaaggatct	caagaagatc	ctttgatctt	5640
ttctacgggg	tctgacgctc	agtggaaacg	aaactcacgt	taagggattt	tggtcatgag	5700
attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aatgaagtt	ttaaatcaat	5760
ctaaagtata	tatgagtaaa	cttggctctg	cagttacca	tgcttaatca	gtgaggcacc	5820
tatctcagcg	atctgtctat	ttcgttcatc	catagttgca	tttaaatttc	cgaactctcc	5880
aaggccctcg	tcgaaaaatc	ttcaaaccct	tcgtccgac	catcttgag	gctacctctc	5940
gaacgaacta	tcgcaagtct	cttggccggc	cttgcgcctt	ggctattgct	tggcagcgcc	6000
tatcgccagg	tattactcca	atcccgaata	tccgagatcg	ggatcacccg	agagaagttc	6060
aacctacatc	ctcaatcccc	atctatccga	gatccgagga	atatcgaaat	cggggcgcg	6120
ctgggtgtacc	gagaacgatc	ctctcagtgc	gagtctcgac	gatccatata	gttgcttggc	6180
agtcagccag	tcggaatcca	gcttgggacc	caggaagtcc	aatcgtcaga	tattgtactc	6240
aagcctgggtc	acggcagcgt	accgatctgt	ttaaaccatg	atattgatag	tctgatcggt	6300
caacgtataa	tcgagtccta	gcttttgcaa	acatctatca	agagacagga	tcagcaggag	6360
gctttcgcat	gagtattcaa	catttccgtg	tcgcccttat	tccctttttt	gcggcatttt	6420
gccttcctgt	ttttgctcac	ccagaaacgc	tggtgaaagt	aaaagatgct	gaagatcagt	6480
tgggtgcg	agtggtgtac	atcgaactgg	atctcaacag	cggtaagatc	cttgagagtt	6540
ttcgccccga	agaacgcttt	ccaatgatga	gcacttttaa	agttctgcta	tgtggcgcg	6600
tattatcccc	tattgacgcc	gggcaagagc	aactcggtcg	ccgcatacac	tattctcaga	6660
atgacttggt	tgagtattca	ccagtcacag	aaaagcatct	tacggatggc	atgacagtaa	6720
gagaattatg	cagtgtctgc	ataaccatga	gtgataacac	tgcggccaac	ttacttctga	6780
caacgattgg	aggaccgaag	gagctaaccg	cttttttgca	caacatgggg	gatcatgtaa	6840
ctcgccctga	tcgttgggaa	ccggagctga	atgaagccat	accaaacgac	gagcgtgaca	6900
ccacgatgcc	tgtagcaatg	gcaacaacct	tgcgtaaact	attaactggc	gaactactta	6960
ctctagcttc	ccggcaacag	ttgatagact	ggatggaggc	ggataaagtt	gcaggaccac	7020
ttctgcgctc	ggcccttcg	gctggctggg	ttattgctga	taaatctgga	gccggtgagc	7080
gtgggtctcg	cggtatcatt	gcagcaactg	ggccagatgg	taagccctcc	cgtatcgtag	7140
ttatctacac	gacggggagt	caggcaacta	tggtatgaac	aatagacag	atcgctgaga	7200
taggtgcctc	actgattaag	cattggtaac	cgattctagg	tgcattggcg	cagaaaaaaa	7260
tgctgatgc	gacgctgcgc	gtcttatact	cccacatatg	ccagattcag	caacggatac	7320
ggcttcccca	acttgccac	ttccatacgt	gtctctctta	ccagaaattt	atccttaaga	7380
tcccgaatcg	tttaaactcg	actctggctc	tatcgaatct	ccgtcgtttc	gagcttacgc	7440

gaacagccgt ggcgctcatt tgctcgtcgg gcatcgaatc tcgtcagcta tcgtcagctt	7500
accttttttg ca	7512
<210> 161	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 161	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa taagctctcg ggggtgtgga	239
<210> 162	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 162	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tactgtaaaa gatgtaaag	239
<210> 163	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 163	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tatgttttat gttactgta	239
<210> 164	
<211> 239	
<212> DNA	
<213> 人工序列	

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 164

aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tacagtaaca taaaacata	239

<210> 165

<211> 239

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 165

aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tggcaaagga gcacatcag	239

<210> 166

<211> 239

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 166

aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tgccaaagca gatgtgttt	239

<210> 167

<211> 239

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 167

aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180

aaaggggtcaa aaagatccta catggatgtg aggattgaaa tgtcctgtag acaggcatg	239
<210> 168	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 168	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tggtggtgtg gggggggag	239
<210> 169	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 169	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tccacactca cgcatgcct	239
<210> 170	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 170	
aaacagctag aatgtaactt aaagtaggtc aatgttttaa ttcgatgttg caatttgttt	60
ggacaagtgg attaaaacgt tccttgaaaa tcatataaag cagccagttt acgggcttgg	120
gcgaatttgc gtccaaaggg tgaggccagg tgtaagtaag aacctacaaa agcactcacc	180
aaaggggtcaa aaagatccta catggatgtg aggattgaaa tcagcacagc agcttcccc	239
<210> 171	
<211> 239	
<212> DNA	
<213> 人工序列	
<220>	

115	120	125
Lys Ile Glu Lys Glu Phe Asn	Lys Lys Ser Met	Glu Thr Tyr Gln Lys
130	135	140
Val Val Gly Leu Ser Lys Gln	Thr Arg Ile Gly	Lys Leu Leu Lys Ser
145	150	155
Gln Phe Thr Leu His Tyr Asp	Thr Glu Asp Arg	Ile Thr Ser Thr Val
165	170	175
Ile Ser His Phe Asn Asn Asp	Met Lys Thr Gly	Val Leu Arg Gly Asp
180	185	190
Arg Ser Leu Arg Thr Tyr Lys	Asn Ser His Pro	Leu Leu Val Arg Ala
195	200	205
Arg Ser Met Lys Val Tyr Glu	Glu Asn Gly Asp	Tyr Phe Ile Lys Trp
210	215	220
Val Lys Gly Ile Val Phe Lys	Ile Val Ile Ser	Ala Gly Ser Lys Gln
225	230	235
Lys Ala Asn Ile Gly Glu Leu	Lys Ser Val Leu	Ile Asn Ile Leu Asn
245	250	255
Gly His Tyr Lys Val Cys Asp	Ser Ser Ile Ser	Leu Asn Lys Asp Leu
260	265	270
Ile Leu Asn Leu Ser Leu Asn	Ile Pro Val Ser	Lys Glu Asn Val Phe
275	280	285
Val Pro Gly Arg Val Val Gly	Val Ala Leu Gly	Leu Lys Ile Pro Ala
290	295	300
Tyr Val Ser Leu Asn Asp Thr	Pro Tyr Ile Lys	Lys Gly Ile Gly Asn
305	310	315
Ile Asp Asp Phe Leu Lys Val	Arg Thr Gln Leu	Gln Ser Gln Arg Lys
325	330	335
Arg Leu Gln Lys Thr Leu Glu	Cys Thr Ser Gly	Gly Lys Gly Arg Asn
340	345	350
Lys Lys Leu Lys Gly Leu Asp	Arg Leu Lys Ala	Lys Glu Lys Asn Phe
355	360	365
Val Asn Thr Tyr Asn His Phe	Leu Ser Lys Lys	Ile Ile Gln Phe Ala
370	375	380
Val Lys Asn Asn Ala Gly Val	Ile His Leu Ala	Glu Leu Gln Phe Asp
385	390	395
Lys Leu Lys His Lys Ser Leu	Leu Arg Asn Trp	Ser Tyr Tyr Gln Leu
405	410	415
Gln Thr Met Ile Glu Tyr Lys	Ala Glu Arg Glu	Gly Ile Glu Val Lys
420	425	430

Tyr Val Asp Ala Ser Tyr Thr Ser Gln Thr Cys Ser Lys Cys Gly His
 435 440 445
 Tyr Glu Glu Gly Gln Arg Val Leu Gln Asp Thr Phe Thr Cys Lys Asn
 450 455 460
 Lys Glu Cys Lys Gly Tyr Val His Lys Val Asn Ala Ala Phe Asn Ala
 465 470 475 480
 Ser Gln Asn Ile Ala Lys Ser Thr Asp Ile Ile Arg Cys Thr Glu Met
 485 490 495
 Ala Lys Asn Asn Asp Ile Glu Lys Asn Ala
 500 505

<210> 174

<211> 56

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<400> 174

ctacgaattc gtttgccatt actttagttc ccattccaaa aaccgatcc atatgg

56

<210> 175

<211> 242

<212> RNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 175

gggaaacagc uagaauuaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug

60

uuuggacaag uggauuuaaa cguuccuuga aaaucauuaa aagcagccag uuuacgggcu

120

ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc

180

accaaagggg caaaaagauc cuacauggau gugaggauug aaauuggaau gggaacuaaa

240

gu

242

<210> 176

<211> 8

<212> DNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成寡核苷酸

<220>

<221> 经修饰_碱基

<222> (1) .. (8)

<223> a, c, t, g, 未知或其他

<400> 176

nnnnnnnn

8

<210> 177

<211> 199

<212> PRT

<213> 智人

<400> 177

Met	Glu	Ala	Ser	Pro	Ala	Ser	Gly	Pro	Arg	His	Leu	Met	Asp	Pro	His
1				5				10					15		
Ile	Phe	Thr	Ser	Asn	Phe	Asn	Asn	Gly	Ile	Gly	Arg	His	Lys	Thr	Tyr
			20					25					30		
Leu	Cys	Tyr	Glu	Val	Glu	Arg	Leu	Asp	Asn	Gly	Thr	Ser	Val	Lys	Met
		35					40					45			
Asp	Gln	His	Arg	Gly	Phe	Leu	His	Asn	Gln	Ala	Lys	Asn	Leu	Leu	Cys
		50				55				60					
Gly	Phe	Tyr	Gly	Arg	His	Ala	Glu	Leu	Arg	Phe	Leu	Asp	Leu	Val	Pro
65				70					75					80	
Ser	Leu	Gln	Leu	Asp	Pro	Ala	Gln	Ile	Tyr	Arg	Val	Thr	Trp	Phe	Ile
			85					90					95		
Ser	Trp	Ser	Pro	Cys	Phe	Ser	Trp	Gly	Cys	Ala	Gly	Glu	Val	Arg	Ala
			100					105					110		
Phe	Leu	Gln	Glu	Asn	Thr	His	Val	Arg	Leu	Arg	Ile	Phe	Ala	Ala	Arg
		115					120					125			
Ile	Tyr	Asp	Tyr	Asp	Pro	Leu	Tyr	Lys	Glu	Ala	Leu	Gln	Met	Leu	Arg
		130				135					140				
Asp	Ala	Gly	Ala	Gln	Val	Ser	Ile	Met	Thr	Tyr	Asp	Glu	Phe	Lys	His
145				150					155					160	
Cys	Trp	Asp	Thr	Phe	Val	Asp	His	Gln	Gly	Cys	Pro	Phe	Gln	Pro	Trp
			165					170					175		
Asp	Gly	Leu	Asp	Glu	His	Ser	Gln	Ala	Leu	Ser	Gly	Arg	Leu	Arg	Ala
		180					185					190			
Ile	Leu	Gln	Asn	Gln	Gly	Asn									
		195													

<210> 178

<211> 173

<212> PRT

<213> 戈登氏菌属

<400> 178

Met	Ser	Gly	Met	Lys	Ile	Arg	Ala	Tyr	Ala	Pro	Ala	Gly	Gly	Pro	Ser
1				5					10					15	
Leu	Ala	Glu	Val	Ala	Leu	Ser	Asn	Val	Thr	Ala	Leu	Cys	Asn	Arg	Trp
			20					25					30		
Leu	Asp	Gln	Pro	Ala	Ser	Glu	Val	Trp	Arg	Glu	Gln	Ala	Ala	Asn	Glu
		35					40					45			
Ile	Leu	Ala	Ala	Met	Asp	Thr	Asp	Pro	Val	Ser	Asp	Leu	Gly	Ile	Val
		50				55					60				
Gly	Leu	Asn	Val	Lys	Asn	Gly	Glu	Ala	Thr	Leu	Asp	Val	Ala	Ser	Cys
65				70						75				80	
Gly	Pro	Ala	Ser	Gln	Ala	Thr	Leu	Leu	Ala	Val	Val	Ala	Ala	Ile	Gly
				85					90					95	
Asp	Leu	Leu	His	Thr	Glu	Asn	Pro	Pro	Asn	Tyr	Val	Glu	Phe	Glu	Val
			100					105					110		
Ile	Gln	Gly	Thr	Arg	Ser	Tyr	Ile	Ala	His	Val	Arg	Thr	Gly	Asp	Arg
		115					120					125			
Pro	Ser	Pro	His	Thr	Leu	Arg	Arg	Gln	Ala	Glu	Gln	Arg	Ile	Thr	Ala
		130				135					140				
Ala	Leu	Asp	Val	Leu	Asp	Asp	His	Ala	Asp	Asp	Pro	Thr	Glu	Val	Ile
145				150						155				160	
Ala	Arg	Ala	Arg	Ala	Val	Leu	Arg	Glu	Gln	Ala	Glu	Gln			
				165					170						

<210> 179

<211> 135

<212> PRT

<213> 戈登氏菌属

<400> 179

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

His Ser Gly Gly Arg Leu Ala Val Thr Asp Ala His Ser Ile Ala Ala		
100	105	110
Val Ala Leu Arg Met Ala Ala Ala Val Thr Glu Ala Arg Ala Pro Ala		
115	120	125
Ala Arg Ser Arg His Ser Ile		
130	135	
<210> 180		
<211> 191		
<212> PRT		
<213> 戈登氏菌属		
<400> 180		
Met Thr Asp Pro Thr Ala Ser Ala Gly Thr His Pro Ala Ala Leu Ala		
1	5	10
Ala Val Ala Val Ala Asn Ser Thr Ala Ala Thr Val Gln Ala Ile Cys		
20	25	30
Val Asp Gly Val Arg Arg His Met His Val Arg Thr His Phe Glu Thr		
35	40	45
Ala Leu Pro His Arg Thr Gln His Ala Gly Arg Arg Phe Arg Ile Val		
50	55	60
Ala Glu His Asp Asp Arg Val Ala Ile Glu Phe Thr Asp Arg Gly Asp		
65	70	75
Arg Met Leu Ala Tyr Pro Glu Glu Ile Thr Ala Asp Thr Tyr Gly Tyr		
85	90	95
Pro Glu Gly Ala His Leu Ile Arg Ala Met Tyr Ala Glu Ala Ala Glu		
100	105	110
Ser Leu Ser His Arg Ile Phe Met Leu Glu Phe Thr Arg Glu Trp Trp		
115	120	125
Leu Glu Leu Ser Glu Gly Arg Gly Gly Pro Leu Asp Thr Ala Pro Gln		
130	135	140
Ser Val Val Leu Leu Gly Ser Arg Met Val Arg Ala Ala Asp Gly Gly		
145	150	155
Arg Ala Val Arg Leu Tyr Arg Arg Ala Ala Asp Gly Ala Pro Asp Val		
165	170	175
Gly Met Leu Arg Ser Ala Ala Arg Asp Phe Val Gly Cys Leu His		
180	185	190
<210> 181		
<211> 234		
<212> PRT		
<213> 戈登氏菌属		

<400> 181

```

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

```

<210> 182

<211> 99

<212> PRT

<213> 微球菌菌属

<400> 182

```

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

```

20	25	30
Pro Asp Ala Val Met Val Ser Lys Ile Gln Gly Ala Gly Trp Thr Leu		
35	40	45
Asp Ser Trp Glu Asp Ala Ala Asn Asn Glu Leu His Thr Arg Arg Glu		
50	55	60
Pro Leu Met Lys Ala Ile Thr Arg Leu Arg Leu Glu Glu Glu Ala Leu		
65	70	75
Arg Gln Glu Tyr Gly Val Thr Asp Gln Asp Glu Pro Ala Asp Gln Lys		
85	90	95
Ala Glu Gly		
<210> 183		
<211> 149		
<212> PRT		
<213> 微球菌菌属		
<400> 183		
Met Asn Ala Ile Glu Leu Lys Cys Arg Arg Glu Ala Leu Gly Leu Ser		
1	5	10
Arg Asp Ala Leu Ala Asp Thr Leu Gly Val Ala Glu Lys Ser Ile Thr		
20	25	30
Arg Trp Glu Phe Gly Lys Asn Pro Pro Arg Asp Trp Ser Trp Ile Asp		
35	40	45
Ser Ala Met Thr Arg Leu Glu Asp Tyr Arg Glu Arg Leu Val Glu Glu		
50	55	60
Leu Ile Ala Glu Thr Leu Arg Val His Glu Gln Ala Gly Ala Ala Leu		
65	70	75
Met Leu Thr Tyr Ala Ser Arg Gly Ser Phe Tyr Arg Trp Leu Pro Glu		
85	90	95
Met Glu Glu Pro Asp Met Gly Asp Gly Ile Glu Val Ile Pro Val Glu		
100	105	110
Leu His Arg Ala Ala Thr Ala Glu Ala Ala Arg Arg Leu Arg Val Ser		
115	120	125
His Gly Leu His Ala Val Ile Glu Ala Val Pro Thr Pro Gln Glu Gly		
130	135	140
Glu Gly Ala Gly Ser		
145		
<210> 184		
<211> 245		
<212> PRT		
<213> 微球菌菌属		

<400> 184

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

<210> 185

<211> 99

<212> PRT

<213> 微球菌菌属

<400> 185

Met Thr Thr Phe Tyr Tyr Ser Val Asn Glu Met Ala Arg Gln Ala Gly

1	5	10	15
Phe Thr Ser Pro Ser Thr Trp Thr Arg Ile Lys Asn Asp Leu Arg Ala			
20	25	30	
Pro Asp Ala Val Met Val Ser Lys Ile Gln Gly Ala Gly Trp Thr Leu			
35	40	45	
Asp Ser Trp Glu Asp Ala Ala Asn Asn Glu Leu His Thr Arg Arg Glu			
50	55	60	
Pro Leu Met Lys Ala Ile Thr Arg Leu Arg Leu Glu Glu Glu Ala Leu			
65	70	75	80
Arg Gln Glu Tyr Gly Val Thr Asp Gln Asp Glu Pro Ala Asp Gln Lys			
85	90	95	
Ala Glu Gly			
<210> 186			
<211> 123			
<212> PRT			
<213> 微球菌菌属			
<220>			
<221> MOD_RES			
<222> (107) .. (107)			
<223> 任何氨基酸			
<400> 186			
Met Asn Ala Ile Glu Leu Lys Cys Arg Arg Glu Ala Leu Gly Leu Ser			
1	5	10	15
Arg Asp Ala Leu Ala Asp Thr Leu Gly Val Ala Glu Lys Ser Ile Thr			
20	25	30	
Arg Trp Glu Phe Gly Lys Asn Pro Pro Arg Asp Trp Ser Trp Ile Asp			
35	40	45	
Ser Ala Met Thr Arg Leu Glu Asp Tyr Arg Glu Arg Leu Val Glu Glu			
50	55	60	
Leu Ile Ala Glu Thr Leu Arg Val His Glu Gln Ala Gly Ala Ala Leu			
65	70	75	80
Met Leu Thr Tyr Ala Ser Arg Gly Ser Phe Tyr Arg Trp Leu Pro Glu			
85	90	95	
Met Glu Glu Pro Asp Met Gly Asp Gly Ile Xaa Val Ile Glu Ala Val			
100	105	110	
Pro Thr Pro Gln Glu Arg Glu Gly Ala Gly Ser			
115	120		
<210> 187			
<211> 245			

<212> PRT

<213> 微球菌菌属

<400> 187

```

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

```

<210> 188

<211> 99

<212> PRT

<213> 微球菌菌属

<400> 188

```

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

```

Ala Glu Gly

<210> 189

<211> 149

<212> PRT

<213> 微球菌菌属

<400> 189

```

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

```

<210> 190

<211> 245

<212> PRT

<213> 微球菌菌属

<400> 190

```

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

```

<210> 191

<211> 206

<212> PRT

<213> Paeniglutamicibacter菌属

<400> 191

```

Met Leu Gly Gln Gly Phe Asp Gln Gly Leu Glu Phe Gly Phe Val Leu
1           5           10           15
Gly Lys Gly Leu Val Gln Gln Leu Leu Ala Cys Pro Val Gln Gly His
           20           25           30
Arg Val Val Val Gly Leu Ala Asp Ile Asp Ala Asp Glu His Val Asp
           35           40           45
Gly Phe Val Val Leu Asp His Cys Ala Pro His Ser Trp Gly His Arg
           50           55           60
Arg Val Ala Ser Asp Glu Thr Val Leu Ala Cys Pro Ala Ala Pro Arg
65           70           75           80
Leu Gly Ile His Val Thr Thr Asp Asn Trp Thr Gly Ala Val Pro Ser
           85           90           95
Arg Ser Gly Pro Tyr Leu Arg Tyr Pro Arg Cys Leu Ser Val Pro Val
           100          105          110
Thr Ala Pro Pro Gly Ser Trp Met Thr Gly Gly Val Asn His Ala Gly
           115          120          125
Thr Asp Arg Pro Ala Ile Pro Ile Leu Ala His Gly Ala Arg Leu Arg
           130          135          140
Asp Tyr Lys Lys Gly Asn Gly Gly Gly Ala Asp Ala Pro Leu Asn Ala
145          150          155          160
Ala Ser Cys Gly Gly Gln Trp Arg Met Ala Cys Asp Arg Thr Ser Val
           165          170          175
Ala Cys Gly Arg Gln Arg Pro Leu Ser Ser Gly Ser Arg Arg Leu Pro
           180          185          190
Arg His Trp Ser Leu Pro His Pro Gly Cys Thr Phe Ile Pro
           195          200          205

```

<210> 192

<211> 106

<212> PRT

<213> 链霉菌属

<220>

<221> MOD_RES

<222> (65) .. (66)

<223> 任何氨基酸

<400> 192

```

Met Gly Ser Ser Ala Val Arg Ala Leu Arg Val Leu Pro Pro Gly Ala

```

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

<210> 193

<211> 244

<212> RNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 193

```

gggaaacagc uagaau guaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug      60
uuuggacaag uggauuaaaa cguuccuuga aaaucauaua aagcagccag uuucacgggcu      120
ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc      180
accaaagggg caaaaagauc cuacauggau gugaggauug aaau guacga gauauaggaa      240
gcug                                             244

```

<210> 194

<211> 242

<212> RNA

<213> 人工序列

<220>

<223> 人工序列的描述：合成多核苷酸

<400> 194

```

gggaaacagc uagaau guaa cuuaaaguag gucaauguuu aaauucgaug uugcaauuug      60
uuuggacaag uggauuaaaa cguuccuuga aaaucauaua aagcagccag uuucacgggcu      120
ugggcgaauu ugcguccaaa gggugaggcc agguguaagu aagaaccuac aaaagcacuc      180
accaaagggg caaaaagauc cuacauggau gugaggauug aaauacuaca acagccacaa      240
cg                                             242

```

<210> 195

<211> 50

<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 195	
aggagcggac agcagcttcc tatactctcg acaggaccgt caacatggtt	50
<210> 196	
<211> 50	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 196	
agccatgata tagacgttgt ggctgttgta gtaggaccgt caacatggtt	50
<210> 197	
<211> 20	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 197	
agcggtagga aaggaatccc	20
<210> 198	
<211> 22	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 198	
ttcgattccg gagaggagc ct	22
<210> 199	
<211> 50	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 199	
aggagcggac agcagcttcc tatactctcg acagggaact caacatggtt	50
<210> 200	

<211> 50	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 200	
agccatgata tagacgttgt ggctgttgta gtagggaact caacatggtt	50
<210> 201	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 201	
gtcacatcct acatggatgt gaggattgaa at	32
<210> 202	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 202	
gtcacatcct acatggatgt gaggattgaa at	32
<210> 203	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 203	
gtcacatcct acatgggtgt gtggattgaa at	32
<210> 204	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 204	
gtcacatcct acatggatgt gaggattgaa at	32
<210> 205	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 205	
gtcacaccct acatgggtgt aaggattgaa at	32
<210> 206	
<211> 30	

<212> DNA	
<213> 芽孢杆菌属	
<400> 206	
gttwaaacat aacaatagat gtwttgaaat	30
<210> 207	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 207	
gtcacatcct acatgggtgt gtggattgaa at	32
<210> 208	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 208	
gttaaaacaa aacaatagat gtattgaaat	30
<210> 209	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 209	
gtttaaactt aacaatagat gtattgaaat	30
<210> 210	
<211> 31	
<212> DNA	
<213> 芽孢杆菌属	
<400> 210	
gtttactatt aacattggat gtatttaaatt	31
<210> 211	
<211> 37	
<212> DNA	
<213> 戈登氏菌属	
<400> 211	
gtcgtgagta cgacgccctc acgctggcgt tgcgacc	37
<210> 212	
<211> 29	
<212> DNA	
<213> 微球菌属	
<400> 212	

ggaagccccg cgcacgcagg gatgagccc	29
<210> 213	
<211> 29	
<212> DNA	
<213> 微球菌菌属	
<400> 213	
ggaagccccg cgcacgcagg gatgagccc	29
<210> 214	
<211> 29	
<212> DNA	
<213> Paeniglutamicibacter菌属	
<400> 214	
gtccttccccg ggcgcgcgcg gaatggctc	29
<210> 215	
<211> 29	
<212> DNA	
<213> 链霉菌属	
<400> 215	
gtacggccca cacgcgcagg ggacgaccg	29
<210> 216	
<211> 29	
<212> DNA	
<213> 微球菌菌属	
<400> 216	
ggaagccccg cgcacgcagg gatgagccc	29
<210> 217	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 217	
gtcggccccg cgcacgcgga gatgagccc	29
<210> 218	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 218	
gtttgacact aacataagat gtattttaa	30
<210> 219	
<211> 30	

<212> DNA	
<213> 芽孢杆菌属	
<400> 219	
gtttaacact aacataagat gtattttaa	30
<210> 220	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 220	
gtttaacatt aacataagat gtattttaa	30
<210> 221	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 221	
gtcgcacctt atataggtgc gtggattgaa at	32
<210> 222	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 222	
gtttastact aacataagat gtattttaa	30
<210> 223	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 223	
gtttaatact aacataagat gtattttaa	30
<210> 224	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 224	
gtcacatcct acatggatgt gaggattgaa at	32
<210> 225	
<211> 22	
<212> DNA	
<213> 芽孢杆菌属	
<400> 225	

tactcatgaa cacctcatgw gt	22
<210> 226	
<211> 33	
<212> DNA	
<213> 芽孢杆菌属	
<400> 226	
gtttactayt aacataagat gtattttaa at ggt	33
<210> 227	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 227	
gtttastayt aacataagat gtattttaa at	30
<210> 228	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 228	
gtttagtact aacataagat gtattttaa at	30
<210> 229	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 229	
gttttgaata aactatgtag aatgtgaat	29
<210> 230	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 230	
gtttaacact aacataagat gtattttaa at	30
<210> 231	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 231	
gtttaacact aacataagat gtattttaa at	30
<210> 232	
<211> 30	

<212> DNA	
<213> 芽孢杆菌属	
<400> 232	
gtttaacact aacataagat gtattttaa	30
<210> 233	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 233	
gttttgaata aactatgtag aatgtgaat	29
<210> 234	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 234	
gtttaacact aacataagat gtattttaa	30
<210> 235	
<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<221> 经修饰_碱基	
<222> (1) .. (19)	
<223> a, c, t, g, 未知或其他	
<400> 235	
nnnnnnnnnn nnnnnnnnn	19
<210> 236	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 236	
gttttagtact aacataagat gtattttaa	30
<210> 237	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 237	
gtttaacact aacataagat gtattttaa	30
<210> 238	

<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<221> 经修饰_碱基	
<222> (1) .. (19)	
<223> a, c, t, g, 未知或其他	
<400> 238	
nnnnnnnnnn nnnnnnnnn	19
<210> 239	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 239	
gtttartact aacataagat gtattttaa	30
<210> 240	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 240	
gtttaacact aacataagat gtattttaa	30
<210> 241	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 241	
gttaaaacat aacaatagat gtattgaa	30
<210> 242	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 242	
gttaaaacat aacaatagat gtattgaa	30
<210> 243	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 243	
gtttaacact aacataagat gtattttaa	30

<210> 244	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 244	
gtttracact aacataagat gtattttaa	30
<210> 245	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 245	
gtttaacatt aacataagat gtattttaa	30
<210> 246	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 246	
gtcacatcct acatggatgt gaggattgaa at	32
<210> 247	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	
<400> 247	
gtcacatcct acatggatgt gtggattgaa at	32
<210> 248	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 248	
gtttagtact aacataagat gtattttaa	30
<210> 249	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 249	
gtttaacatt aacataagat gtattttaa	30
<210> 250	
<211> 33	
<212> DNA	

<213> 芽孢杆菌属	
<400> 250	
gtcacatcct acatggatgt gaggattgaa att	33
<210> 251	
<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<221> 经修饰_碱基	
<222> (1) .. (19)	
<223> a, c, t, g, 未知或其他	
<400> 251	
nnnnnnnnnnn nnnnnnnnnn	19
<210> 252	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 252	
gtttaacatt aacataagat gtattttaa	30
<210> 253	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 253	
gtttaacatt aacataagat gtattttaa	30
<210> 254	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 254	
gtttastayt aacataagat gtattttaa	30
<210> 255	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 255	
gtttaacact aacataagat gtattttaa	30
<210> 256	
<211> 30	

<212> DNA	
<213> 芽孢杆菌属	
<220>	
<221> 经修饰_碱基	
<222> (7) .. (7)	
<223> a, c, t, g, 未知或其他	
<400> 256	
gkttganact aayataagat gtattttaa	30
<210> 257	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 257	
gtttaacact aacataagat gtattttaa	30
<210> 258	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 258	
gtttactayt aacataagat gtattttaa	30
<210> 259	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 259	
gtttastayt aacataagat gtattttaa	30
<210> 260	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 260	
gtttaacact aacataagat gtattttaa	30
<210> 261	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 261	
gtttaatact aacataagat gtattttaa	30
<210> 262	

<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 262	
gtttaacact aacataagat gtattttaa	30
<210> 263	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 263	
gtttartact aacataagat gtattttaa	30
<210> 264	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 264	
gtttagtayt aacataagat gtattttaa	30
<210> 265	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 265	
gtttaacact aacataagat gtattttaa	30
<210> 266	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 266	
gtttaatact aacataagat gtattttaa	30
<210> 267	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 267	
gtttascatt aacataagat gtattttaa	30
<210> 268	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	

<400> 268	
gtttgacact aacataagat gtattttaa	30
<210> 269	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 269	
rtttgatatc aactatgtgg aatgtaa	29
<210> 270	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 270	
rtttgatatc aactatgtgg aatgtaa	29
<210> 271	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 271	
rtttgatatc aactatgtgg aatgtaa	29
<210> 272	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 272	
gtttaacact aacataagat gtattttaa	30
<210> 273	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 273	
gtttartact aacataagat gtattttaa	30
<210> 274	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 274	
gtttagtact aacataagat gtattttaa	30
<210> 275	

<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 275	
gttttgaata aactatgtag aatgtgaat	29
<210> 276	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 276	
gtttaacact aacataagat gtattttaa	30
<210> 277	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 277	
gtttaacact aacataagat gtattttaa	30
<210> 278	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 278	
gtttaacact aacataagat gtattttaa	30
<210> 279	
<211> 34	
<212> DNA	
<213> 芽孢杆菌属	
<400> 279	
agtaacaata cgtatagcta cgtaaattgt agga	34
<210> 280	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 280	
gtttaacact aacataagat gtattttaa	30
<210> 281	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	

<400> 281	
gtttaacact aacataagat gtattttaa	30
<210> 282	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 282	
gtttagtact aacataagat gtattttaa	30
<210> 283	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 283	
gtttastact aacataagat gtattttaa	30
<210> 284	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 284	
gtttagtact aacataagat gtattttaa	30
<210> 285	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 285	
gttttaataca caatatgtat tgttaga	29
<210> 286	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 286	
gtttagyact aacataagat gtattttaa	30
<210> 287	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 287	
gtttastact aacataagat gtattttaa	30
<210> 288	

<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 288	
gtttaaacat aacaatagat gtattgaaat	30
<210> 289	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 289	
rttttaaata caatatgtat tgttagaat	29
<210> 290	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 290	
gtttastayt aacataagat gtatttaaata	30
<210> 291	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 291	
gtttastayt aacataagat gtatttaaata	30
<210> 292	
<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 292	
gtttgatatc aactatatgg aatgtaaata	29
<210> 293	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 293	
gtttastayt aacataagat gtatttaaata	30
<210> 294	
<211> 32	
<212> DNA	
<213> 芽孢杆菌属	

<400> 294	
gtcacatcct acatgggtgt gaggattgaa at	32
<210> 295	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 295	
gtttaahatt aacataagat gtattttaa	30
<210> 296	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 296	
gtttastayt aacataagat gtattttaa	30
<210> 297	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 297	
gtttastayt aacataagat gtattttaa	30
<210> 298	
<211> 21	
<212> DNA	
<213> 芽孢杆菌属	
<400> 298	
ggtcattgtg tcattacatt g	21
<210> 299	
<211> 18	
<212> DNA	
<213> 芽孢杆菌属	
<400> 299	
atccccaaaa cgtaggt	18
<210> 300	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 300	
gtttaacact aacataagat gtattttaa	30
<210> 301	

<211> 29	
<212> DNA	
<213> 芽孢杆菌属	
<400> 301	
gtcttaatta caatatgtat tgtagaat	29
<210> 302	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 302	
gtttaacact aacataagat gtatttaa	30
<210> 303	
<211> 19	
<212> DNA	
<213> 芽孢杆菌属	
<220>	
<221> 经修饰_碱基	
<222> (5) .. (5)	
<223> a, c, t, g, 未知或其他	
<400> 303	
gatanaaagt gtttcata	19
<210> 304	
<211> 19	
<212> DNA	
<213> 人工序列	
<220>	
<221> 经修饰_碱基	
<222> (1) .. (19)	
<223> a, c, t, g, 未知或其他	
<400> 304	
nnnnnnnnnn nnnnnnnnn	19
<210> 305	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 305	
gtttactrtt aacawtggat gtatttaa	30
<210> 306	
<211> 30	

<212> DNA	
<213> 芽孢杆菌属	
<400> 306	
gttaaaacaa aacaatagat gtattgaaat	30
<210> 307	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 307	
gttaaaacaw aacaatagat gtattgaaat	30
<210> 308	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 308	
gttwaacat aacaatagat gtwttgaaat	30
<210> 309	
<211> 30	
<212> DNA	
<213> 芽孢杆菌属	
<400> 309	
gttwaacat aacaatagat gtwttgaaat	30
<210> 310	
<211> 146	
<212> RNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<400> 310	
acggucuuug uguuacaccc uucaaagcau ccaggauucu uauguaucaa uguuuuuuuu	60
acauucgaua cuaccuuuug gagcguuuac acuugguaaa uucaccuugu auuacuucuc	120
cauuuaagca auagauguau ugaaau	146
<210> 311	
<211> 143	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述：合成多核苷酸	
<220>	

<221> 经修饰_碱基	
<222> (52) .. (59)	
<223> a, c, t, g, 未知或其他	
<400> 311	
tcgtcggcag cgtcagatgt gtataagaga cagaacgacg gccagtgaat tnnnnnnnna	60
ctacaacagc cacaacgtct atatcatggg atcctctaga gtcgacctgc tgtctcttat	120
acacatctcc gagcccacga gac	143
<210> 312	
<211> 148	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成多核苷酸	
<220>	
<221> 经修饰_碱基	
<222> (52) .. (56)	
<223> a, c, t, g, 未知或其他	
<400> 312	
tcgtcggcag cgtcagatgt gtataagaga cagaacgacg gccagtgaat tnnnnntgga	60
atgggaacta aagtaatggc aaacgaattc gtaggatcct ctagagtcga cctgctgtct	120
cttatacaca tctccgagcc cagagac	148
<210> 313	
<211> 30	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 313	
actacaacag ccacaacgtc tatatcatgg	30
<210> 314	
<211> 18	
<212> DNA	
<213> 人工序列	
<220>	
<223> 人工序列的描述: 合成寡核苷酸	
<400> 314	
acacctgtaa tcccagca	18

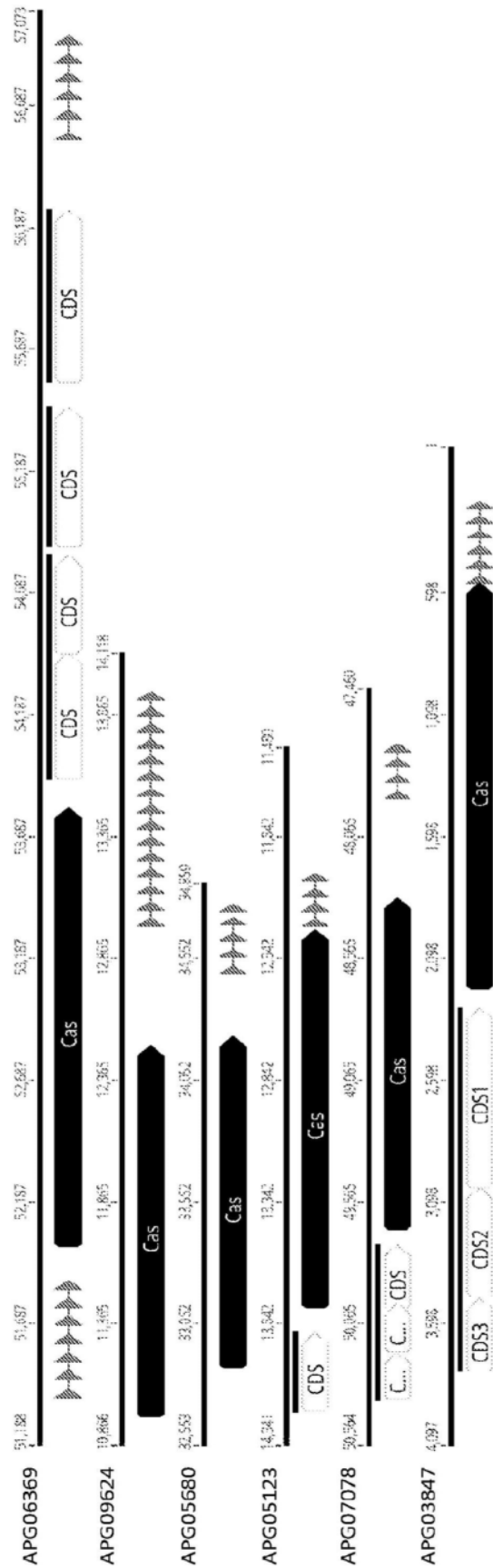


图1