(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 101932701 A
(43)申请公布日 2010.12.29

(21)申请号 200980103634.7

(22)申请日 2009.01.30

(30)优先权数据
61/025,406 2008.02.01 US

(85)PCT申请进入国家阶段日
2010.07.30

(86)PCT申请的申请数据
PCT/US2009/032515 2009.01.30

(87)PCT申请的公布数据
WO2009/099906 EN 2009.08.13

(71)申请人 阿森尼克斯公司
地址 美国北卡罗来纳州

(72)发明人 V·海因里希斯 L·C·斯考滕 B·范迪博格

(74)专利代理机构 北京市中咨律师事务所 11247
代理人 黄革生 林柏楠

(51)Int.Cl.
C12N 9/10(2006.01)
C12N 15/82(2006.01)

权利要求书 5 页 说明书 23 页 序列表 44 页
附图 0 页

(54)发明名称

GRG31 和 GRG36 EPSP 合酶的定向进化

(57)摘要

本发明提供赋予细菌、植物、植物细胞、组织和种子除草剂抗性或耐受性的组合物和方法。组合物包括编码除草剂抗性或耐受性多肽的多核苷酸，含有那些多核苷酸的载体，以及含有载体的宿主细胞。本发明的核苷酸序列可用在 DNA 构建体或表达盒中用于转化生物体，包括微生物和植物，和在其中表达。组合物还包括转化的细菌、植物、植物细胞、组织和种子。本发明尤其是提供了编码草甘膦抗性或耐受性多肽的分离的多核苷酸，特别是 SEQ ID NO:2 和 4 的多肽变体。另外，本发明涵盖与所述多核苷酸对应的氨基酸序列。尤其是本发明提供了分离的多核苷酸，其含有编码 SEQ ID NO:7-28 中所示氨基酸序列的核苷酸序列、或者 SEQ ID NO:29 或 30 中所示的核苷酸序列。

CN 101932701 A

1. 由编码草甘膦耐受性 EPSP 合酶的分离的核酸分子组成的组合物，其中所述核酸分子选自下组：

a) 编码 SEQ ID NO:2 的变体的核苷酸序列，其中所述变体包含如下一或多个残基：

i) 与 SEQ ID NO:2 氨基酸位置 75 相对应位置上的半胱氨酸残基；

ii) 与 SEQ ID NO:2 氨基酸位置 76 相对应位置上的天冬氨酸残基；

iii) 与 SEQ ID NO:2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基；

iv) 与 SEQ ID NO:2 氨基酸位置 78 相对应位置上的丝氨酸残基；

v) 与 SEQ ID NO:2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基；

vi) 与 SEQ ID NO:2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基；

vii) 与 SEQ ID NO:2 氨基酸位置 86 相对应位置上的丙氨酸残基；

viii) 与 SEQ ID NO:2 氨基酸位置 87 相对应位置上的甘氨酸残基；

ix) 与 SEQ ID NO:2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基；

x) 与 SEQ ID NO:2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基；

xi) 与 SEQ ID NO:2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基；和

xii) 与 SEQ ID NO:2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基；和

b) 编码 SEQ ID NO:4 的变体的核苷酸序列，其中所述变体包含如下一或多个残基：

i) 与 SEQ ID NO:4 氨基酸位置 90 相对应位置上的异亮氨酸残基；

ii) 与 SEQ ID NO:4 氨基酸位置 145 相对应位置上的谷氨酰胺残基；

iii) 与 SEQ ID NO:4 氨基酸位置 350 相对应位置上的苏氨酸残基；和

iv) 与 SEQ ID NO:4 氨基酸位置 410 相对应位置上的甲硫氨酸残基。

2. 权利要求 1 的分离的核酸分子，其中所述核苷酸序列是已经设计用于在植物中表达的合成序列。

3. 含有权利要求 1 的核酸分子的载体。

4. 含有权利要求 3 的载体的宿主细胞。

5. 权利要求 4 的宿主细胞，其为细菌宿主细胞或植物宿主细胞。

6. 含有权利要求 1 的核酸分子的转基因种子。

7. 由分离的多肽组成的组合物，所述多肽包含选自下组的草甘膦耐受性 EPSP 合酶：

a)SEQ ID NO:2 的变体，其中所述变体包含如下一个或多个残基：

i) 与 SEQ ID NO:2 氨基酸位置 75 相对应位置上的半胱氨酸残基；

ii) 与 SEQ ID NO:2 氨基酸位置 76 相对应位置上的天冬氨酸残基；

iii) 与 SEQ ID NO:2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基；

iv) 与 SEQ ID NO:2 氨基酸位置 78 相对应位置上的丝氨酸残基；

v) 与 SEQ ID NO:2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基；

vi) 与 SEQ ID NO :2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基 ;

vii) 与 SEQ ID NO :2 氨基酸位置 86 相对应位置上的丙氨酸残基 ;

viii) 与 SEQ ID NO :2 氨基酸位置 87 相对应位置上的甘氨酸残基 ;

ix) 与 SEQ ID NO :2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基 ;

x) 与 SEQ ID NO :2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基 ;

xi) 与 SEQ ID NO :2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基 ;和

xii) 与 SEQ ID NO :2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基 ;和

b)SEQ ID NO :4 的变体,其中所述变体包含如下一个或多个残基 :

i) 与 SEQ ID NO :4 氨基酸位置 90 相对应位置上的异亮氨酸残基 ;

ii) 与 SEQ ID NO :4 氨基酸位置 145 相对应位置上的谷氨酰胺残基 ;

iii) 与 SEQ ID NO :4 氨基酸位置 350 相对应位置上的苏氨酸残基 ;和

iv) 与 SEQ ID NO :4 氨基酸位置 410 相对应位置上的甲硫氨酸残基。

8. 由具有稳定整合至其基因组中的 DNA 构建体的植物组成的组合物,所述 DNA 构建体含有编码具有草甘膦耐受性活性的蛋白质的核苷酸序列,其中所述核苷酸序列选自下组 :

a) 编码 SEQ ID NO :2 的变体的核苷酸序列,其中所述变体包含如下一个或多个残基 :

i) 与 SEQ ID NO :2 氨基酸位置 75 相对应位置上的半胱氨酸残基 ;

ii) 与 SEQ ID NO :2 氨基酸位置 76 相对应位置上的天冬氨酸残基 ;

iii) 与 SEQ ID NO :2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基 ;

iv) 与 SEQ ID NO :2 氨基酸位置 78 相对应位置上的丝氨酸残基 ;

v) 与 SEQ ID NO :2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基 ;

vi) 与 SEQ ID NO :2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基 ;

vii) 与 SEQ ID NO :2 氨基酸位置 86 相对应位置上的丙氨酸残基 ;

viii) 与 SEQ ID NO :2 氨基酸位置 87 相对应位置上的甘氨酸残基 ;

ix) 与 SEQ ID NO :2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基 ;

x) 与 SEQ ID NO :2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基 ;

xi) 与 SEQ ID NO :2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基 ;和

xii) 与 SEQ ID NO :2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基 ;和

b) 编码 SEQ ID NO :4 的变体的核苷酸序列,其中所述变体包含如下一个或多个残基 :

i) 与 SEQ ID NO :4 氨基酸位置 90 相对应位置上的异亮氨酸残基 ;

ii) 与 SEQ ID NO :4 氨基酸位置 145 相对应位置上的谷氨酰胺残基 ;

iii) 与 SEQ ID NO :4 氨基酸位置 350 相对应位置上的苏氨酸残基 ;和

iv) 与 SEQ ID NO :4 氨基酸位置 410 相对应位置上的甲硫氨酸残基 ;

其中所述核苷酸序列有效连接于驱动编码序列在植物细胞中表达的启动子上。

9. 权利要求 8 的组合物,其中所述植物是植物细胞。

10. 权利要求 9 的组合物,其中所述植物选自下组 :玉蜀黍、高粱、小麦、向日葵、番茄、十字花科植物、胡椒属植物、马铃薯、棉花、稻、大豆、甜菜、甘蔗、烟草、大麦和油菜籽油菜。

11. 权利要求 1-10 中任一项的组合物,其中所述 SEQ ID NO :2 的变体选自下组 :SEQ ID NO :7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25 和 26。

12. 权利要求 1-10 中任一项的组合物,其中所述 SEQ ID NO :4 的变体选自 SEQ ID NO :27 或 28。

13. 产生具有草甘膦耐受性活性的多肽的方法,包括在编码所述多肽的核酸分子表达的条件下培养权利要求 4 的宿主细胞。

14. 赋予植物草甘膦耐受性的方法,所述方法包括用 DNA 构建体转化所述植物,所述构建体含有驱动在植物细胞中表达的启动子,所述启动子有效连接于编码草甘膦耐受性 EPSP 合酶的核苷酸序列,其中所述核苷酸序列选自下组 :

a) 编码 SEQ ID NO :2 的变体的核苷酸序列,其中所述变体包含如下一个或多个残基 :

i) 与 SEQ ID NO :2 氨基酸位置 75 相对应位置上的半胱氨酸残基 ;

ii) 与 SEQ ID NO :2 氨基酸位置 76 相对应位置上的天冬氨酸残基 ;

iii) 与 SEQ ID NO :2 氨基酸位置 77 相对应位置的上天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基 ;

iv) 与 SEQ ID NO :2 氨基酸位置 78 相对应位置上的丝氨酸残基 ;

v) 与 SEQ ID NO :2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基 ;

vi) 与 SEQ ID NO :2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基 ;

vii) 与 SEQ ID NO :2 氨基酸位置 86 相对应位置上的丙氨酸残基 ;

viii) 与 SEQ ID NO :2 氨基酸位置 87 相对应位置上的甘氨酸残基 ;

ix) 与 SEQ ID NO :2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基 ;

x) 与 SEQ ID NO :2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基 ;

xi) 与 SEQ ID NO :2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基 ;和

xii) 与 SEQ ID NO :2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基 ;和

b) 编码 SEQ ID NO :4 的变体的核苷酸序列,其中所述变体包含如下一个或多个残基 :

i) 与 SEQ ID NO :4 氨基酸位置 90 相对应位置上的异亮氨酸残基 ;

ii) 与 SEQ ID NO :4 氨基酸位置 145 相对应位置上的谷氨酰胺残基 ;

iii) 与 SEQ ID NO :4 氨基酸位置 350 相对应位置上的苏氨酸残基 ;和

iv) 与 SEQ ID NO :4 氨基酸位置 410 相对应位置上的甲硫氨酸残基 ;

和再生转化的植物。

15. 增加植物活力或产量的方法，包括：

a) 向所述植物中引入包含核苷酸序列的核酸分子，所述核苷酸序列编码具有高于周围环境温度的最适温度的草甘膦耐受性 EPSP 合酶，其中所述核苷酸序列选自下组：

i) 编码 SEQ ID NO:2 的变体的核苷酸序列，其中所述变体包含如下一个或多个残基：

1) 与 SEQ ID NO:2 氨基酸位置 75 相对应位置上的半胱氨酸残基；

2) 与 SEQ ID NO:2 氨基酸位置 76 相对应位置上的天冬氨酸残基；

3) 与 SEQ ID NO:2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基；

4) 与 SEQ ID NO:2 氨基酸位置 78 相对应位置上的丝氨酸残基；

5) 与 SEQ ID NO:2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基；

6) 与 SEQ ID NO:2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基；

7) 与 SEQ ID NO:2 氨基酸位置 86 相对应位置上的丙氨酸残基；

8) 与 SEQ ID NO:2 氨基酸位置 87 相对应位置上的甘氨酸残基；

9) 与 SEQ ID NO:2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基；

10) 与 SEQ ID NO:2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基；

11) 与 SEQ ID NO:2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基；和

12) 与 SEQ ID NO:2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基；和

ii) 编码 SEQ ID NO:4 的变体的核苷酸序列，其中所述变体包含如下一个或多个残基：

1) 与 SEQ ID NO:4 氨基酸位置 90 相对应位置上的异亮氨酸残基；

2) 与 SEQ ID NO:4 氨基酸位置 145 相对应位置上的谷氨酰胺残基；

3) 与 SEQ ID NO:4 氨基酸位置 350 相对应位置上的苏氨酸残基；和

4) 与 SEQ ID NO:4 氨基酸位置 410 相对应位置上的甲硫氨酸残基；

b) 使所述植物与有效浓度的草甘膦接触；和

c) 在接触所述草甘膦之后，在每天至少连续 2 小时温度高于周围环境温度的条件下培养所述植物至少 4 天，其中所述接触后的天数在植物的生长季内，

其中所述植物的活力或产量与表达最适温度不高于周围环境温度的草甘膦耐受性 EPSP 合酶的植物的活力或产量相比更高。

16. 赋予植物草甘膦抗性的方法，包括：

a) 向所述植物中引入包含核苷酸序列的核酸分子，所述核苷酸序列编码草甘膦耐受性 EPSP 合酶，其中所述核苷酸序列选自下组：

i) 编码 SEQ ID NO:2 的变体的核苷酸序列，其中所述变体包含如下一个或多个残基：

1) 与 SEQ ID NO:2 氨基酸位置 75 相对应位置上的半胱氨酸残基；

2) 与 SEQ ID NO:2 氨基酸位置 76 相对应位置上的天冬氨酸残基；

3) 与 SEQ ID NO:2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基；

4) 与 SEQ ID NO :2 氨基酸位置 78 相对应位置上的丝氨酸残基 ;

5) 与 SEQ ID NO :2 氨基酸位置 81 相对应位置上的丝氨酸或甘氨酸残基 ;

6) 与 SEQ ID NO :2 氨基酸位置 82 相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基 ;

7) 与 SEQ ID NO :2 氨基酸位置 86 相对应位置上的丙氨酸残基 ;

8) 与 SEQ ID NO :2 氨基酸位置 87 相对应位置上的甘氨酸残基 ;

9) 与 SEQ ID NO :2 氨基酸位置 88 相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基 ;

10) 与 SEQ ID NO :2 氨基酸位置 95 相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基 ;

11) 与 SEQ ID NO :2 氨基酸位置 206 相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基 ;和

12) 与 SEQ ID NO :2 氨基酸位置 215 相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基 ;和

ii) 编码 SEQ ID NO :4 的变体的核苷酸序列,其中所述变体包含如下一个或多个残基 :

1) 与 SEQ ID NO :4 氨基酸位置 90 相对应位置上的异亮氨酸残基 ;

2) 与 SEQ ID NO :4 氨基酸位置 145 相对应位置上的谷氨酰胺残基 ;

3) 与 SEQ ID NO :4 氨基酸位置 350 相对应位置上的苏氨酸残基 ;和

4) 与 SEQ ID NO :4 氨基酸位置 410 相对应位置上的甲硫氨酸残基 ;

b) 使所述植物与有效浓度的草甘膦接触 ;和

c) 在接触所述草甘膦之后,在每天至少连续 2 小时温度高于周围环境温度的条件下培养所述植物至少 4 天,其中所述接触后的天数在植物的生长季内。

17. 权利要求 13-16 中任一项的方法,其中所述 SEQ ID NO :2 的变体选自下组 :SEQ ID NO :7、8、9、10、11、12、13、14、15、16、17、18、19、20、21、22、23、24、25 和 26。

18. 权利要求 13-16 中任一项的方法,其中所述 SEQ ID NO :4 的变体选自 SEQ ID NO :27 或 28。

19. 权利要求 15、16、17 或 18 中任一项的方法,其中步骤 (c) 中的温度是约 32℃至约 60℃。

20. 权利要求 19 的方法,其中步骤 (c) 中的温度是约 37℃。

GRG31 和 GRG36EPSP 合酶的定向进化

发明领域

[0001] 本发明涉及植物分子生物学，特别是赋予提高的除草剂草甘膦抗性或耐受性的新 EPSP 合酶多肽。

[0002] 发明背景

[0003] N- 膦酰基甲基甘氨酸通常称为草甘膦，是一种重要的农事化学品。草甘膦抑制将磷酸烯醇丙酮酸 (PEP) 和 3- 磷酸莽草酸 (S3P) 转化成 5- 烯醇丙酮酰 -3- 磷酸莽草酸的酶。对该酶 (5- 烯醇丙酮酰莽草酸 -3- 磷酸合酶 ；本文称为“EPSP 合酶”或“EPSPS”) 的抑制通过关闭莽草酸路径、由此抑制芳族氨基酸生物合成而杀死植物细胞。

[0004] 由于草甘膦类除草剂抑制芳族氨基酸生物合成，故它们不仅杀死植物细胞，而且还对细菌细胞有毒。草甘膦抑制许多细菌 EPSP 合酶，因而对这些细菌有毒。不过，某些细菌 EPSP 合酶对草甘膦具有高耐受性。

[0005] 植物细胞对草甘膦毒性的抗性可以通过转化植物细胞使之表达抗草甘膦的细菌 EPSP 合酶而产生。特别地，来自根癌农杆菌 (Agrobacteriumtumefaciens)CP4 株的该细菌基因已经用来继在植物中表达后赋予植物细胞除草剂抗性。来自鼠伤寒沙门氏菌 (Salmonella typhimurium)CT7 株的突变 EPSP 合酶赋予细菌细胞草甘膦抗性，并赋予植物细胞草甘膦抗性 (美国专利 No.4,535,060 ；4,769,061 ；和 5,094,945)。

[0006] 美国专利 6,040,497 报道了突变玉米 EPSP 合酶，具有 102 位苏氨酸至异亮氨酸取代、以及 106 位脯氨酸至丝氨酸取代 (“TIPS”突变)。这些改变赋予所述玉米酶草甘膦抗性。来自鼠伤寒沙门氏菌 CT7 株的突变 EPSP 合酶赋予细菌细胞草甘膦抗性，而且据报道赋予植物细胞草甘膦抗性 (美国专利 No.4,535,060 ；4,769,061 ；和 5,094,945)。He 等人 ((2001)Biochim etBiophysica Acta 1568 ：1-6) 通过诱变以及重组大肠杆菌和鼠伤寒沙门氏菌 EPSP 合酶基因，开发了草甘膦耐受性增加的 EPSP 合酶，并提出 42 位 (T42M) 和 230 位 (Q230K) 的突变很可能是造成所观察到的抗性的原因。随后的工作 (He 等人 (2003) Biosci.Biotech.Biochem.67 ：1405-1409) 表明，T42M 突变 (苏氨酸突变为甲硫氨酸) 足以提高大肠杆菌和鼠伤寒沙门氏菌酶的耐受性。由于除草剂抗性植物提供许多优势，期望除草剂抗性基因提高的草甘膦抗性活性。

[0007] 其他的诱变方法描述在 2006 年 6 月 13 日递交的、发明名称为“全排列诱变”的美国专利 No.60/813,095 中。

[0008] 发明概述

[0009] 本文提供了赋予抗性或耐受性的组合物和方法。组合物包括抗草甘膦除草剂的 EPSP 合酶，编码此类酶的核酸分子，含有那些核酸分子的载体，以及含有载体的宿主细胞。组合物包括编码除草剂抗性多肽的核酸分子，包括编码 SEQ ID NO ：2 和 4 的多肽变体的那些核酸分子。在多个实施方案中，多肽变体如 SEQ ID NO ：7-28 所示。这些编码序列可用在 DNA 构建体或表达盒中，用于转化生物体，包括微生物和植物，及在其中表达。组合物还包括通过将本发明组合物引入生物体基因组中而呈草甘膦抗性的转化细菌、植物、植物细胞、组织和种子。在生物体是植物的情况下，引入所述序列允许对所述植物施加含草甘膦的除草

剂，以选择性地杀死草甘膦敏感性杂草或其他未转化的植物，而非转化的生物体。这些序列还可以用作为标记，以选择在草甘膦条件下生长的植物细胞。

[0010] 另外还提供了鉴定具有草甘膦抗性活性的 EPSP 合酶的方法。所述方法包括基于本发明结构域的存在来鉴定具有草甘膦抗性的其他 EPSP 合酶序列。

[0011] 发明详述

[0012] 现在将于下文参照附图更充分地描述本发明，附图中显示了一些但非所有的本发明的实施方案。事实上，本发明可以有多种不同形式的实施方案，故而不应阐释为局限于本文所阐释的实施方案；相反，这些实施方案是为了使发明公开满足适用的法律要求而提供的。通篇相似的数值适用于相似的要素。

[0013] 在得益于前述说明以及所附附图中所给出的教导的情况下，本发明所属领域的技术人员自然会想到本文所阐释的本发明的许多变动和其他实施方案。所以，应当理解本发明不局限于所公开的具体实施方案，且变动和其他实施方案旨在落入所附权利要求书的范围之内。虽然本文采用了具体术语，但其仅仅是在一般性和描述性意义上使用的，而不是为了构成限制。

[0014] 本发明涉及调控生物特别是植物或植物细胞的除草剂抗性的组合物和方法。所述方法包括用编码本发明草甘膦抗性基因的核苷酸序列转化生物体。本发明的核苷酸序列可用于制备对除草剂草甘膦显示出增加的耐受性的植物。因而，本发明的“草甘膦抗性”或“草甘膦耐受性”基因旨在指编码 SEQ ID NO:7-28 中任一所示氨基酸序列的核苷酸序列，及其编码草甘膦抗性或耐受性多肽的片段和变体。同样，本发明的“草甘膦抗性”或“草甘膦耐受性”多肽是指具有 SEQ ID NO:7-28 所示氨基酸序列的多肽，及其赋予宿主细胞草甘膦抗性或草甘膦耐受性的片段和变体。

[0015] A. 分离的多核苷酸及其变体和片段

[0016] 在一些实施方案中，本发明包括分离的、重组的或纯化的多核苷酸。“分离的”或“纯化的”多核苷酸或多肽或其生物活性片段，当通过重组技术生产时基本上不含其它细胞材料或培养基，或当通过化学合成时基本上不含化学前体或其他化学品。“生物活性”表示拥有天然多肽的期望生物活性，即，保留除草剂抗性活性。“分离的”多核苷酸可以不含在所述多核苷酸来源生物体的基因组 DNA 中天然位于所述核酸侧翼的序列（即，位于所述核酸 5’和 3’端的序列）（例如，蛋白质编码序列）。出于本发明的目的，“分离的”在用来述及多核苷酸时不包括分离的染色体。例如，在多个实施方案中，分离的草甘膦抗性编码多核苷酸可以含有少于约 5kb、4kb、3kb、2kb、1kb、0.5kb 或 0.1kb 的在所述多核苷酸来源细胞的基因组 DNA 中天然位于所述多核苷酸侧翼的核苷酸序列。

[0017] 本发明的多核苷酸包括编码作为 SEQ ID NO:2 和 4 的变体的草甘膦抗性多肽的那些多核苷酸。特别是，多核苷酸编码 SEQ ID NO:2 的变体，在对应于 SEQ ID NO:2 氨基酸残基 75、77、77、78、81、82、86、87、88 和 95 的位置上具有一个或多个取代，或在对应于 SEQ ID NO:4 氨基酸位置 90、145、350 或 410 的一个或多个位置上具有取代。在一些实施方案中，变体包含如下一个或多个：与 SEQ ID NO:2 氨基酸位置 75 相对应位置上的半胱氨酸残基；与 SEQ ID NO:2 氨基酸位置 76 相对应位置上的天冬氨酸残基；与 SEQ ID NO:2 氨基酸位置 77 相对应位置上的天冬酰胺、丙氨酸、谷氨酰胺或苏氨酸残基；与 SEQ ID NO:2 氨基酸位置 78 相对应位置上的丝氨酸残基；与 SEQ ID NO:2 氨基酸位置 81 相对应位置上的丝氨酸

或甘氨酸残基；与SEQ ID NO:2氨基酸位置82相对应位置上的缬氨酸、苏氨酸、亮氨酸或苯丙氨酸残基；与SEQ ID NO:2氨基酸位置86相对应位置上的丙氨酸残基；与SEQ ID NO:2氨基酸位置87相对应位置上的甘氨酸残基；与SEQ ID NO:2氨基酸位置88相对应位置上的苯丙氨酸、缬氨酸、亮氨酸或组氨酸残基；与SEQ ID NO:2氨基酸位置95相对应位置上的丝氨酸、谷氨酰胺或苏氨酸残基；与SEQ ID NO:2氨基酸位置206相对应位置上的缬氨酸、丝氨酸、精氨酸、谷氨酰胺、谷氨酸或苏氨酸残基；与SEQ ID NO:2氨基酸位置215相对应位置上的赖氨酸、谷氨酰胺、精氨酸、丝氨酸或苏氨酸残基；与SEQ ID NO:4氨基酸位置90相对应位置上的异亮氨酸残基；与SEQ ID NO:4氨基酸位置145相对应位置上的谷氨酰胺残基；与SEQ ID NO:4氨基酸位置350相对应位置上的苏氨酸残基；和与SEQ ID NO:4氨基酸位置410相对应位置上的甲硫氨酸残基。在多个实施方案中，变体选自SEQ ID NO:7-28中任一，以及SEQ ID NO:29或30所示的合成核苷酸序列。

[0018] “草甘膦”表示N-膦酰基甲基甘氨酸（包括其任何盐）的任何除草剂形式以及导致在植物中(in planta)产生草甘膦阴离子的其他形式。“除草剂抗性蛋白质”或由“除草剂抗性编码核酸分子”的表达而产生的蛋白质，包括赋予细胞如下能力的蛋白质，即，与不表达该蛋白质的细胞相比耐受更高浓度除草剂的能力、或与不表达该蛋白质的细胞相比更长时间地耐受一定浓度的除草剂的能力。“草甘膦抗性蛋白质”包括赋予细胞如下能力的蛋白质，即，与不表达该蛋白质的细胞相比耐受更高浓度的草甘膦、或与不表达该蛋白质的细胞相比更长时间地耐受一定浓度的草甘膦的能力。“耐受”或“耐受性”是指，以不易区别于未处理细胞的方式，存活、或者执行基本细胞功能，例如蛋白质合成和呼吸。

[0019] 本发明还考虑本文所述多核苷酸的变体和片段。多核苷酸的“片段”可以编码多肽的生物活性部分，或者其可以是能够在本文别处所公开的方法中用作杂交探针或PCR引物的片段。作为多核苷酸片段的多核苷酸包含至少约15、20、50、75、100、200、300、350、400、450、500、550、600、650、700、750、800、850、900、950、1000、1050、1100、1150、1200、1250、1300、1350、1400、1450、1500、1550、1600、1650、1700、1750、1800、1850、1900、1950个连续核苷酸，或多达本文所公开的全长多核苷酸中存在的核苷酸数，这取决于目的用途（例如，包含SEQ ID NO:1的EPSP合酶多核苷酸）。“连续”核苷酸表示彼此紧邻的核苷酸残基。

[0020] 本发明的多核苷酸片段一般将编码保留全长草甘膦抗性蛋白质的生物活性、即除草剂抗性活性的多肽片段。“保留除草剂抗性活性”是指，片段将具有本文所公开的如SEQ ID NO:2或4的全长草甘膦抗性蛋白质的除草剂抗性活性的至少约30％、至少约50％、至少约70％、至少约80％、85％、90％、95％、100％、110％、125％、150％、175％、200％、250％、至少约300％或更高。测量除草剂抗性活性的方法在本领域公知。参见例如美国专利No.4,535,060和5,188,642，每一篇在此全文并入作为参考。

[0021] 编码本发明多肽的生物活性部分的多核苷酸片段将编码至少约15、25、30、50、75、100、125、150、175、200、250、300、350、400个连续氨基酸，或多达本发明全长多肽中存在的氨基酸总数。

[0022] 本发明还包括如前文所述的变体多核苷酸。多核苷酸的“变体”还包括编码本文所公开的多肽、但是由于遗传密码的简并性而存在保守性差别的那些序列，以及足够相同的那些序列。

[0023] 术语“足够相同”是指，利用比对程序之一，使用标准参数，多肽或多核苷酸序列与

参照序列相比具有至少约 60％或 65％序列同一性，约 70％或 75％序列同一性，约 80％或 85％序列同一性，约 90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％序列同一性。本领域技术人员将会意识到，可以通过考虑密码子简并性、氨基酸相似性、读框定位等等，适当地调整这些数值，以确定两个多核苷酸所编码的多肽的相应同一性。

[0024] 细菌基因常常在邻近开放读框起点处具有多个甲硫氨酸起始密码子。通常在这些起始密码子的一个或多个处的起始翻译会导致产生功能蛋白质。这些起始密码子可以包括 ATG 密码子。不过，细菌例如芽孢杆菌属菌种 (Bacillus sp.) 也识别密码子 GTG 作为起始密码子，在 GTG 密码子处起始翻译的蛋白质含有甲硫氨酸作为第一个氨基酸。此外，不是常常能先验性地确定这些密码子中哪个天然地在细菌中使用。因而，可以理解，这些可选甲硫氨酸密码子之一的使用可能导致产生赋予除草剂抗性的变体。这些除草剂抗性蛋白质涵盖在本发明之中，并且可以用在本发明的方法之中。

[0025] 天然存在的等位变体可以用公知的分子生物学技术来鉴定，例如下文所述的聚合酶链式反应 (PCR) 和杂交技术。变体多核苷酸也包括合成来源的、已经例如利用定点或其他诱变策略生成的、但仍然编码具有期望生物活性的多肽的多核苷酸。

[0026] 技术人员还会理解，可以通过进一步突变本发明的多核苷酸而引入变化，由此引起所编码多肽的氨基酸序列的进一步变化，而不改变多肽的生物活性。因而，分离的变体多核苷酸可以通过在编码如本文所公开的 EPSP 合酶结构域的相应多核苷酸中引入一个或多个另外的核苷酸取代、添加或缺失而产生，从而在编码的多肽中引入一个或多个氨基酸取代、添加或缺失。进一步的突变可通过标准技术引入，例如定点诱变和 PCR 介导的诱变或者基因改组技术。此类变体多核苷酸也涵盖在本发明之中。

[0027] 变体多核苷酸可以通过沿编码序列的全长或部分随机引入突变（例如通过饱和诱变）而制备，可以筛选所得突变体赋予除草剂抗性活性的能力，以鉴定保留活性的突变体。

[0028] 基因改组或有性 PCR 方法（例如，Smith(1994)Nature 370 :324-325 ;美国专利 No.5,837,458 ;5,830,721 ;5,811,238 和 5,733,731，每一篇都在此并入作为参考）可用来进一步修饰或增强本文公开的多核苷酸和多肽（例如，赋予草甘膦抗性的多肽）。基因改组包括随机片段化若干突变体 DNA，接着通过 PCR 将它们重组装成全长分子。多种基因改组方法的实例包括但不限于 :在 DNA 酶处理后组装、交错延伸方法 (STEP) 以及体外随机引发重组。在 DNA 酶介导的方法中，分离自一组阳性突变体的 DNA 区段以 DNaseI 切割成随机片段，并进行多轮 PCR 而不添加引物。随着 PCR 循环的执行，随机片段的长度接近未切割区段的长度，导致不同克隆中的突变在一些所得序列中混合和累积。通过多个选择和改组循环，已经导致了若干酶的功能性增强 (Stemmer(1994)Nature 370 :389-391 ;Stemmer(1994)Proc. Natl.Acad.Sci.USA 91 :10747-10751 ;Crameri 等 (1996)Nat.Biotechnol.14 :315-319 ; Zhang 等 (1997)Proc.Natl.Acad.Sci.USA94 :4504-4509 以 及 Crameri 等 (1997)Nat. Biotechnol.15 :436-438)。可对例如编码本文公开多肽的多核苷酸实施此类方法。

[0029] 生成变体的其他方法包括使表达本文公开蛋白质（或其文库）的细胞接触对蛋白质的活性造成胁迫的特定条件。特定条件可包括（但不限于）温度改变、pH 改变、及底物或抑制剂浓度的改变。蛋白质文库可在蛋白质表达期间（例如在大肠杆菌或其他宿主中）、或在制备蛋白质提取物之后、或在蛋白质纯化之后接触这些条件。

[0030] 已经接触过胁迫条件的蛋白质文库的功能或酶促活性可随之与野生型蛋白质进行比较，以鉴定具有改善性能的蛋白质。活性比较可以作为生长筛选的一部分、或可选地作为定量蛋白质活性的酶测定法的一部分来实施。可鉴定为改善的性能可以包括草甘膦耐受性、动力学常数（包括 Km、Ki、Vmax）、蛋白质稳定性、蛋白质热稳定性、或蛋白质最适温度的变化。

[0031] B. 分离的蛋白质及其变体和片段

[0032] 除草剂抗性多肽也涵盖在本发明之中。基本上不含细胞物质的除草剂抗性多肽包括具有少于约 30％、20％、10％或 5％（以干重计）非除草剂抗性多肽（本文也称为“杂质蛋白质”）的多肽制品。在本发明中，“除草剂抗性蛋白质”表示本文公开的 EPSP 合酶多肽。也提供了其片段、生物活性部分及变体，并且它们可以用来实施本发明的方法。

[0033] “片段”或“生物活性部分”包括这样的多肽片段，其包含编码除草剂抗性蛋白质的氨基酸序列的一部分、且保留除草剂抗性活性。例如，除草剂抗性蛋白质的生物活性部分可以是长度 10、25、50、100 个或更多氨基酸的多肽。此类生物活性部分可通过重组技术制备，并评估其除草剂抗性活性。

[0034] “变体”表示这样的蛋白质或多肽，其具有与 SEQ ID NO：7-28 中任一至少约 60％、65％、约 70％、75％、约 80％、85％、90％、91％、92％、93％、94％、95％、96％、97％、98％或 99％相同的氨基酸序列。本领域技术人员将会意识到，可以考虑密码子简并性、氨基酸相似性、读框定位等等，对这些数值进行适当的调整，以确定两个多核苷酸所编码的多肽的相应同一性。

[0035] 例如，可在一个或多个非必需氨基酸残基处进行保守氨基酸取代。“非必需”氨基酸残基是能够自多肽野生型序列改变而不影响生物活性的残基，而“必需”氨基酸残基是生物活性所必需的。“保守氨基酸取代”是指将氨基酸残基替换为具有相似侧链的氨基酸残基。具有相似侧链的氨基酸残基家族在本领域已经定义。这些家族包括具有碱性侧链（例如，赖氨酸、精氨酸、组氨酸）、酸性侧链（例如，天冬氨酸、谷氨酸）、不带电荷的极性侧链（例如，甘氨酸、天冬酰胺、谷氨酰胺、丝氨酸、苏氨酸、酪氨酸、半胱氨酸）、非极性侧链（例如，丙氨酸、缬氨酸、亮氨酸、异亮氨酸、脯氨酸、苯丙氨酸、甲硫氨酸、色氨酸）、β 分枝侧链（例如，苏氨酸、缬氨酸、异亮氨酸）和芳族侧链（例如，酪氨酸、苯丙氨酸、色氨酸、组氨酸）的氨基酸。可非保守区进行保留功能的氨基酸取代。一般，不对多肽活性所必需的保守氨基酸残基或位于保守基序中的氨基酸残基进行此取代。不过，本领域技术人员将会理解，功能变体可以在保守残基处具有微小的保守或非保守改变。

[0036] 本发明也涵盖本发明多肽或其变体或片段的抗体。产生抗体的方法在本领域公知（参见例如，Harlow 和 Lane(1988)Antibodies：A LaboratoryManual，Cold Spring Harbor Laboratory，Cold Spring Harbor，NY；美国专利 No.4,196,265）。

[0037] C. 多核苷酸构建体

[0038] 可修饰编码本发明 EPSP 合酶多肽的多核苷酸，以获得或增强在植物细胞中的表达。编码本文鉴定的多肽的多核苷酸可以在表达盒中提供，以用于在目的植物中表达。“植物表达盒”包括能够引起多核苷酸在植物细胞中表达的 DNA 构建体，包括重组 DNA 构建体。盒可以，按 5’-3’ 转录方向，包括有效连接于一个或多个目的多核苷酸的在植物中有功能的转录起始区域（即启动子，尤其是异源启动子），和 / 或翻译及转录终止区（即终止区）。

盒可以另外含有至少一个待引入生物体的其他多核苷酸，例如可选择标记基因。可选地，该其他多核苷酸可以在多个表达盒上提供。这样的表达盒提供有多个限制性位点，用来插入多核苷酸，使之处于调控区域的转录调控之下。

[0039] “异源”一般指，多核苷酸或多肽并非细胞内源的、或者对于其所存在于的天然基因组中的位置而言并非内源的，并且已经通过感染、转染、显微注射、电穿孔、微粒轰击(microprojection)等而添加到细胞中。“有效连接”表示两个多核苷酸之间的功能性连接。例如，当启动子有效连接于DNA序列时，启动子序列起始和介导DNA序列的转录。公认，有效连接的多核苷酸可以是或者可以不是连续的，并且当用来述及两个多肽编码区的连接时，所述多肽在相同的读框中表达。

[0040] 启动子可以是在选择的植物细胞、植物部分或植物中显示出转录活性的任何多核苷酸序列。启动子对于植物宿主和/或本发明的DNA序列可以是天然的或同源的、或者外来的或异源的。启动子对于植物宿主是“天然的”或“同源的”，表示启动子见于该启动子所引入的天然植物中。启动子对于本发明的DNA序列是“外来的”或“异源的”，表示启动子对于有效连接的本发明DNA序列并非是天然的或天然存在的启动子。启动子可以是诱导型或组成型的。其可以是天然存在的，可以由多个天然存在启动子的部分组成，或者可以部分或全部是合成的。通过启动子结构的研究提供了启动子的设计指南，例如Harley和Reynolds(1987)Nucleic Acids Res.15:2343-2361。此外，可以优化启动子相对于转录起点的定位。参见例如Roberts等(1979)Proc.Natl.Acad.Sci.USA,76:760-764。许多在植物中使用的适宜启动子在本领域公知。

[0041] 例如，在植物中使用的适宜组成型启动子包括：来自植物病毒的启动子，例如花生退绿条死病花椰菜花叶病毒(PClSV)启动子(美国专利No.5,850,019)；来自花椰菜花叶病毒(CaMV)的35S启动子(Odell等(1985)Nature 313:810-812)；小球藻病毒甲基转移酶基因的启动子(美国专利No.5,563,328)以及来自玄参花叶病毒(FMV)的全长转录物启动子(美国专利No.5,378,619)；来自诸如稻肌动蛋白(McElroy等(1990)Plant Cell 2:163-171)；泛素(Christensen等(1989)Plant Mol.Biol.12:619-632和Christensen等(1992)Plant Mol.Biol.18:675-689)；pEMU(Last等(1991)Theor.Appl.Genet.81:581-588)；MAS(Velten等(1984)EMBO J.3:2723-2730)；玉米H3组蛋白(Lepetit等(1992)Mol.Gen.Genet.231:276-285and Atanassova等(1992)Plant J.2(3):291-300)；欧洲油菜(Brassica napus)ALS3(PCT申请WO 97/41228)等基因的启动子；以及各种农杆菌(Agrobacterium)基因的启动子(参见美国专利No.4,771,002;5,102,796;5,182,200;和5,428,147)。

[0042] 在植物中使用的适宜诱导型启动子包括：来自ACE1系统的响应铜的启动子(Mett等(1993)PNAS 90:4567-4571)；玉米In2基因的响应苯磺酰胺除草剂安全剂的启动子(Hershey等(1991)Mol.Gen.Genetics227:229-237和Gatz等(1994)Mol.Gen.Genetics 243:32-38)；以及来自Tn 10 Tet阻遏物的启动子(Gatz等(1991)Mol.Gen.Genet.227:229-237)。在植物中使用的另一种诱导型启动子是响应植物通常不响应的诱导剂的启动子。例示性的此类诱导型启动子是来自类固醇激素基因的诱导型启动子，其转录活性受糖皮质激素诱导(Schena等(1991)Proc.Natl.Acad.Sci.USA88:10421)；或最近应用的嵌合转录激活物XVE，用于受雌二醇激活的、基于雌激素受体的诱导型植物表达系统(Zuo

等 (2000)Plant J.,24 :265-273)。其他在植物中使用的诱导型启动子在 EP 332104、PCT WO93/21334 和 PCT WO 97/06269 中描述,在此将其整体并入并入作为参考。还可以使用由其他启动子的部分组成的启动子,以及部分或全部合成的启动子。参见例如 Ni 等 (1995) Plant J.7 :661-676 和 PCT WO 95/14098,描述了在植物中使用的此类启动子。

[0043] 启动子可以包括,或者经修饰而包括,一个或多个增强子元件。在一些实施方案中,启动子可以包括多个增强子元件。含有增强子元件的启动子与不含它们的启动子相比可以确保更高水平的转录。在植物中使用的适宜增强子元件包括 PClSV 增强子元件 (美国专利 No.5,850,019)、CaMV35S 增强子元件 (美国专利 No.5,106,739 和 5,164,316) 以及 FMV 增强子元件 (Maiti 等 (1997)Transgenic Res.6 :143-156)。还请参见 PCT WO96/23898。

[0044] 通常此类构建体可以含有 5' 和 3' 非翻译区。此类构建体可以含有“信号序列”或“前导序列”以利于目的肽向某个细胞内结构如叶绿体 (或其他质体)、内质网或高尔基体的共翻译或翻译后转运,或者分泌。例如,可以工程化设计构建体以含有信号肽,以利于肽转运到内质网。“信号序列”是指,已知会或者怀疑会引起跨细胞膜的共翻译或翻译后肽转运的序列。在真核细胞中,这一般涉及分泌到高尔基体中以及由此的一定糖基化。“前导序列”是指,任何在翻译后产生如下氨基酸序列的序列,所述氨基酸序列足以触发肽链共翻译转运到亚细胞器。因而,这包括通过穿运进入内质网、进入液泡、质体 (包括叶绿体)、线粒体等而靶向转运和 / 或糖基化的前导序列。还可以优选工程化设计植物表达盒以含有内含子,以便内含子的 mRNA 加工是表达所必需的。

[0045] “3'非翻译区”表示位于编码序列下游的多核苷酸。多聚腺苷酸化信号序列及其他编码能够影响多聚腺苷酸尾添加到 mRNA 前体 3' 端的调控信号的序列是 3' 非翻译区。“5' 非翻译区”表示位于编码序列上游的多核苷酸。

[0046] 其他上游或下游非翻译元件包括增强子。增强子是发挥作用以增加启动子区表达的多核苷酸。增强子在本领域公知,并且包括但不限于 SV40 增强子区和 35S 增强子元件。

[0047] 终止区可以对于转录起始区而言是天然的,可以对本发明的序列而言是天然的,或者可以来自其他来源。便利的终止区可以来自根癌农杆菌 (A.tumefaciens)Ti 质粒,例如章鱼碱合酶和胭脂碱合酶终止区。参见例如 Guerineau 等 (1991)Mol.Gen.Genet.262 :141-144 ;Proudfoot(1991)Cell64 :671-674 ;Sanfacon 等 (1991)Genes Dev.5 :141-149 ;Mogen 等 (1990)Plant Cell 2 :1261-1272 ;Munroe 等 (1990)Gene 91 :151-158 ;Ballas 等 (1989)Nucleic Acids Res.17 :7891-7903 和 Joshi 等 (1987)NucleicAcidRes.15 :9627-9639。

[0048] 在本发明的一方面,针对给定多肽,例如本发明的多肽,设计合成的 DNA 序列。合成 DNA 序列的开放读框在细胞中的表达导致产生本发明的多肽。合成 DNA 序列可用于简单地去除不想要的限制性核酸内切酶位点以利于 DNA 克隆策略,改变或去除任何潜在的密码子偏好,改变或改善 GC 含量,去除或改变可变读框,和 / 或改变或去除可能存在于天然 DNA 序列之中的内含子 / 外显子剪接识别位点、多聚腺苷酸化位点、Shine-Delgarno 序列、不想要的启动子元件等等。也有可能合成的 DNA 序列可以用来向 DNA 序列中引入其他改进,例如引入内含子序列、创建以与细胞器靶向序列 (例如叶绿体转运肽、质外体 / 液泡靶向肽、或导致所得肽滞留在内质网中的肽序列) 融合的蛋白质形式表达的 DNA 序列。合成的基因还可以利用宿主细胞偏好的密码子合成以提高表达,或者以宿主偏好的密码子使用频

率使用密码子合成。参见例如 Campbell 和 Gowri(1990)Plant Physiol.92 :1-11 ;美国专利 No.6,320,100 ;6,075,185 ;5,380,831 和 5,436,391,美国公开申请 No.20040005600 和 20010003849,以及 Murray 等 (1989)Nucleic Acids Res.17 :477-498,在此并入作为参考。

[0049] 在一个实施方案中,目的多核苷酸靶向叶绿体进行表达。在这种方式中,在目的多核苷酸不直接插入到叶绿体中的情况下,表达盒将另外含有编码转运肽的多核苷酸以指导目的核苷酸进入叶绿体。这样的转运肽在本领域已知。参见例如 Von Heijne 等 (1991)Plant Mol.Biol.Rep.9 :104-126 ;Clark 等 (1989)J.Biol.Chem.264 :17544-17550 ; Della-Cioppa 等 (1987)Plant Physiol.84 :965-968 ;Romer 等 (1993)Biochem.Biophys. Res.Commun.196 :1414-1421 和 Shah 等 (1986)Science 233 :478-481。

[0050] 可以考虑植物细胞核和该细胞器之间密码子使用的差异,对待靶向叶绿体的目的多核苷酸进行优化以在叶绿体中表达。在这种方式中,目的多核苷酸可以利用叶绿体偏好的密码子来合成。参见例如美国专利 No.5,380,831,在此并入作为参考。

[0051] 该植物表达盒可以插入到植物转化载体中。“转化载体”表示允许转化细胞的 DNA 分子。这样的分子可以由一个或多个表达盒组成,并且可以组织成一个以上的载体 DNA 分子。例如,双元载体是植物转化载体,利用两个不连续的 DNA 载体编码转化植物细胞所必需的所有顺式和反式功能 (Hellens 和 Mullineaux(2000)Trends in Plant Science 5 : 446-451)。“载体”是指设计用于在不同宿主细胞之间转移的多核苷酸构建体。“表达载体”是指具有在外来细胞中掺入、整合和表达异源 DNA 序列或片段的能力的载体。

[0052] 植物转化载体包括一个或多个用于实现植物转化的 DNA 载体。例如,本领域常见的做法是利用含有一个以上连续 DNA 区段的植物转化载体。这些载体在本领域往往称为双元载体。双元载体以及伴有辅助质粒的载体最常用于农杆菌介导的转化,其中实现有效转化所需的 DNA 区段的大小和复杂度相当大,故而将功能分割到分开的 DNA 分子上是有利的。双元载体一般包括这样的质粒载体,其含有 T-DNA 转移所需的顺式作用序列 (例如左边界和右边界)、经工程化设计能够在植物细胞中表达的可选择标记、以及“目的多核苷酸”(经工程化设计能够在期望生成转基因植物的植物细胞中表达的多核苷酸)。还存在于该质粒载体上的有细菌复制所需的序列。这些顺式作用序列以允许有效转移到植物细胞中并在其中表达的方式排列。例如,可选择标记序列和目的序列位于左右边界之间。通常第二质粒载体含有介导 T-DNA 从农杆菌转移到植物细胞中的反式作用因子。该质粒往往含有毒力功能 (Vir 基因),允许农杆菌感染植物细胞,并通过边界序列处的切割和 vir 介导的 DNA 转移而转移 DNA,如本领域所理解的那样 (Hellens 和 Mullineaux(2000)Trends in Plant Science,5 :446-451)。若干类型的农杆菌菌株 (例如,LBA4404、GV3101、EHA101、EHA105 等) 可用于植物转化。第二质粒载体对于通过其他方法例如微粒轰击、显微注射、电穿孔、聚乙二醇等向植物中引入多核苷酸而言并不是必需的。

[0053] D. 植物转化

[0054] 本发明的方法包括向植物中引入核苷酸构建体。“引入”表示以构建体获得进入植物细胞内部的方式将核苷酸构建体呈递给植物。本发明的方法不要求使用特定的方法将核苷酸构建体引入到植物中,仅仅是要求核苷酸构建体获得进入植物的至少一个细胞的内部。将核苷酸构建体引入到植物中的方法在本领域已知,包括但不限于稳定转化方法、瞬时转化方法以及病毒介导的方法。

[0055] 一般而言，植物转化方法包括将异源DNA转移到靶植物细胞中（例如，未成熟或成熟的胚、悬浮培养物、未分化的愈伤组织、原生质体等），接着施加最大阈值水平的适宜选择（取决于可选择标记基因，在本案中是“草甘膦”），以从未转化的细胞团中回收转化的植物细胞。外植体一般被转移到新鲜供应的相同培养基中并常规培养。随后，转化的细胞在置于补充有最大阈值水平选择剂（例如“草甘膦”）的再生培养基中之后分化出芽。芽随之转移到选择性生根培养基中，以回收生根的芽或小植株。转基因小植株随之长成成熟植株，并产生可育种子（例如Hiei等(1994)The Plant Journal6:271-282;Ishida等(1996)Nature Biotechnology 14:745-750)。外植体一般转移到新鲜供应的相同培养基中并常规培养。有关生成转基因植物的技术和方法的一般说明见Ayres和Park(1994)Critical Reviews in PlantScience 13:219-239以及Bommineni和Jauhar(1997)Maydica 42:107-120。由于转化的材料含有许多细胞；转化的以及未转化的细胞皆存在于任何一块的靶受试愈伤组织或者组织或细胞群中。杀死未转化的细胞并允许转化的细胞增殖的能力将导致转化的植物培养物。通常，去除未转化细胞的能力对于快速回收转化的植物细胞和成功生成转基因植物是一种限制。可利用分子和生化方法来确认转基因植物基因组中整合的异源目的基因的存在。

[0056] 生成转基因植物可以通过若干方法之一来实施，包括但不限于，通过农杆菌将异源DNA引入到植物细胞之中（农杆菌介导的转化）、用附于粒子上的异源外来DNA轰击植物细胞、以及多种其他非粒子直接介导的方法（例如，Hiei等(1994)The Plant Journal 6:271-282;Ishida等(1996)NatureBiotechnology 14:745-750;Ayres和Park(1994)Critical Reviews in PlantScience 13:219-239;Bommineni和Jauhar(1997)Maydica 42:107-120)以转移DNA。

[0057] 叶绿体转化方法在本领域已知。参见例如Svab等(1990)Proc.Natl.Acad.Sci.USA 87:8526-8530;Svab和Maliga(1993)Proc.Natl.Acad.Sci.USA 90:913-917;Svab和Maliga(1993)EMBO J.12:601-606。该方法依赖于基因枪递送含可选择标记的DNA，以及通过同源重组使DNA靶向该质体基因组。另外，质体转化可通过核编码的、质体定向的RNA聚合酶的组织偏好型表达，通过反式激活沉默的质体携带的转基因而实现。这样的系统已经在McBride等(1994)Proc.Natl.Acad.Sci.USA 91:7301-7305中报道。

[0058] 已经转化的细胞可以按照常规方法培养成植物。参见例如McCormick等(1986)Plant Cell Reports 5:81-84。随之可以生长这些植物，并用相同的转化的株系或不同的株系授粉，鉴定组成型表达期望的表型特征的所得杂种。可以栽培两代或更多代，以确保该期望表型特征的表达得以稳定维持和遗传，然后收获种子以确保已经实现期望表型特征的表达。以这种方式，本发明提供了具有稳定掺入至其基因组中的本发明核苷酸构建体[例如本发明的表达盒]的转化的种子（也称为“转基因种子”）。

[0059] E. 植物转化评价

[0060] 在将异源外来DNA引入到植物细胞中之后，该异源基因在植物基因组中的转化或整合可以通过多种方法来确认，例如分析与整合基因相关的核酸、蛋白质以及代谢物。

[0061] PCR分析是一种在移植入土壤之前的较早阶段筛选存在掺入的基因的转化细胞、组织或芽的快速方法(Sambrook和Russell(2001)MolecularCloning:A Laboratory Manual(Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY))。PCR利用特

异于目的基因或农杆菌载体背景等的寡核苷酸引物进行。

[0062] 植物转化可以通过基因组 DNA 的 Southern 印迹分析来确认 (Sambrook 和 Russell(2001) 同前)。一般而言，从转化株中提取总 DNA，用适宜的限制性酶消化，在琼脂糖凝胶上分级分离，并转至硝基纤维素或尼龙膜上。膜或“斑点”随之可以按照标准技术用例如放射标记的 ^{32}P 靶 DNA 片段探测，以确认引入基因在植物基因组中的整合 (Sambrook 和 Russell，2001，同前)。

[0063] 在 Northern 分析中，按照本领域常规使用的标准方法，从转化株的特定组织中分离 RNA，在甲醛琼脂糖凝胶中分级分离，并印迹到尼龙滤膜上 (Sambrook 和 Russell(2001) 同前)。本发明 grg 序列所编码的 RNA 的表达随之可以利用本领域已知的方法通过使滤膜与源自 GDC 的放射性探针杂交来检验 (Sambrook 和 Russell(2001) 同前)。

[0064] 可以通过标准方法 (Sambrook 和 Russell(2001) 同前)，利用与除草剂抗性蛋白质上存在的一个或多个表位结合的抗体，对转基因植株进行 Western 印迹和生化测定等，以确定除草剂抗性基因所编码的蛋白质的存在。

[0065] 在本发明的一个方面，本文所述的 grg 基因可用作标记来评估细菌或植物细胞的转化。

[0066] F. 植物和植物部分

[0067] “植物”是指，整个植物、植物器官 (例如，叶、茎、根等)、种子、植物细胞、繁殖体、胚及其子代。植物细胞可以是分化的或未分化的 (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)。本发明可以用来将多核苷酸引入到任何植物物种中，包括但不限于单子叶植物和双子叶植物。目的植物的实例包括但不限于 : 玉米 (玉蜀黍)、高粱、小麦、向日葵、番茄、十字花科植物、胡椒属植物、马铃薯、棉花、稻、大豆、甜菜、甘蔗、烟草、大麦、和油菜籽油菜、芸苔属物种 (Brassicasp.)、苜蓿、黑麦、粟、红花、花生、甘薯、木薯、咖啡、椰子、菠萝、柑橘属果树、可可、茶树、香蕉、鳄梨、无花果、番石榴、芒果、橄榄、木瓜、腰果、澳大利亚坚果、杏、燕麦、蔬菜类、观赏植物类以及针叶树类。

[0068] 蔬菜类包括但不限于 : 番茄、莴苣、绿豆、菜豆、豌豆、以及甜瓜属 (Curcumis) 的成员，例如黄瓜、甜瓜和香瓜。观赏植物类包括但不限于 : 杜鹃、绣球属植物、木槿属植物、蔷薇属植物、郁金香属植物、水仙、矮牵牛属植物、康乃馨、大戟属植物和菊属植物。作物植物也有重要意义，包括例如玉蜀黍、高粱、小麦、向日葵、番茄、十字花科植物、胡椒属植物、马铃薯、棉花、稻、大豆、甜菜、甘蔗、烟草、大麦、油菜籽油菜等。

[0069] 本发明适合于任何单子叶植物家族的成员，包括但不限于 : 玉蜀黍、稻、大麦、燕麦、小麦、高粱、黑麦、甘蔗、菠萝、薯蓣、洋葱、香蕉、椰子和枣。

[0070] G. 增加植物产量的方法

[0071] 本发明提供了增加植物产量的方法。所述方法包括向植物或植物细胞中引入含本文所公开的 grg 序列的多核苷酸。如本文所定义的那样，植物“产量”是指植物所产生的生物质的质量和 / 或数量。“生物质”是指任何测量到的植物产品。生物质产出的增加可以是所测量的该植物产品的产量方面的任何改进。增加的植物产量具有若干商业应用。例如，增加植物叶生物质可以增加人或动物所消耗的叶类蔬菜的产量。另外，可以通过增加叶生物质，增加植物来源的药用或工业产品的产出。产量增加可以包括任何统计学显著的增加，包括但不限于 : 至少 1％的增加、至少 3％的增加、至少 5％的增加、至少 10％的增加、至少

20％的增加、至少 30％、至少 50％、至少 70％、至少 100％或更高的增加。

[0072] 在特定方法中，用有效浓度的除草剂处理植物，其中除草剂施用导致增强的植物产量。“有效浓度”表示允许植物产量增加的浓度。目的除草剂的此类有效浓度在本领域普遍已知。可以在出苗前或者出苗后，按照除草剂施用的通常技术对大田施用除草剂，所述大田包含已经因异源表达本发明的 grg 基因而对除草剂具有抗性的作物。

[0073] 本发明还提供了对植物或植物部分赋予除草剂抗性的方法。在此类方法中，将本文公开的 grg 多核苷酸引入到植物中，其中多核苷酸的表达导致草甘膦耐受性或抗性。借助该方法产生的植物能够用有效浓度的除草剂处理，并展示出对该除草剂的增加的耐受性。本申请中除草剂的“有效浓度”是足以使天然不具有除草剂抗性或未曾被赋予除草剂抗性的植物或植物部分的生长减慢或停止的量。

[0074] H. 控制大田中杂草的方法

[0075] 本发明还提供了选择性地控制含植物的大田中的杂草的方法。在一个实施方案中，植物种子或植物由于所述植物种子或植物中插入了本文公开的 grg 多核苷酸而为草甘膦抗性的。在一个特定方法中，用有效浓度的除草剂处理植物，其中除草剂施用导致选择性地控制杂草或其他未转化的植物。“有效浓度”是指，控制杂草或其他未转化植物的生长或扩散、而不会显著影响草甘膦抗性植物或植物种子的浓度。目的除草剂的此有效浓度在本领域普遍已知。可以在出苗前或者出苗后，按照除草剂施用的通常技术，对大田施用除草剂，其中所述大田包含已经被赋予了抗除草剂抗性的植物或植物种子。

[0076] I. 温度谱

[0077] 已经进行了植物中草甘膦代谢的若干研究，并揭示植物不代谢草甘膦或者代谢极其缓慢。草甘膦可以迅速穿透角质层，并经过一个相当长的时间转移到整个植株（在 Kearney 和 Kaufman 编辑 (1988)Herbicides ;Chemistry, Degradation & Mode of Action Marcel Dekker, Inc., New York,3 :1-70，以及 Grossbard 和 Atkinson 编辑 (1985)The Herbicide GlyphosateButterworths,London,25-34 页中综述）。因而，很可能在农业经济意义植物中在接触草甘膦后持续的一段时间内草甘膦耐受性是必要的。在生长季期间温度频繁超过 30℃的情况下，采用在升高温度时维持活性的草甘膦耐受性 EPSP 合酶将是有利的。

[0078] 在本发明的一个实施方案中，EPSP 合酶在高于或低于周围环境温度的温度时呈现出热稳定性。“热稳定性”表示酶在高于或低于周围环境温度的温度时比在该温度并非热稳定的 EPSP 合酶在更长的时间阶段内具有活性。例如，热稳定的 EPSP 合酶在高于或低于周围环境温度的温度时在大于约 1 小时、大于约 2 小时、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 20、约 25 小时或更长的时间内具有酶活。出于本发明的目的，“周围”环境温度为约 30℃。在一些实施方案中，高于周围温度是是指等于或高于约 32℃、约 34℃、约 37℃、约 40℃、约 45℃、约 50℃、约 55℃、约 60℃、约 65℃或更高的温度。低于周围温度是指等于或低于约 28℃、低于约 27℃、约 26℃、约 25℃、约 23℃、约 20℃、约 18℃、约 15℃、约 10℃、或等于或低于约 5℃，或 0℃左右的温度。测定 EPSP 合酶活性的方法在本文别处更详细地讨论。出于本发明的目的，热稳定的 EPSP 合酶在以于该酶最适温度观察到的最大活性水平的约 90％至 100％、约 80％至约 90％、约 70％至约 80％、约 60％至约 70％、或约 50％至约 60％发挥功能时视为有活性。

[0079] 因而，本文提供了增加在高于周围环境温度的温度时的草甘膦耐受性的方法和组合物。在一个实施方案中，所述方法包括向植物中引入编码 SEQ ID NO :7-28 中任一所示草甘膦耐受性 EPSP 合酶的核苷酸序列，并在高于周围环境温度的温度培养植物。在具体实施方案中，在植物生长季，培养温度高于周围环境温度平均至少约 2 小时 / 天、至少约 3 小时 / 天、至少约 4 小时 / 天、至少约 5、约 6、约 7、约 8、约 9、约 10、约 11、约 12、约 14、约 16、约 18、约 20、至少约 22 小时 / 天或高达约 24 小时 / 天。

[0080] 在另一实施方案中，所述方法包括向植物中引入编码 SEQ ID NO :7-28 中任一所示草甘膦耐受性 EPSP 合酶的核苷酸序列，使所述植物与除草有效浓度的草甘膦接触，并在向植物施用草甘膦之后，以至少 1 小时、至少约 2 小时、至少约 3、至少约 4 或更长小时 / 天，在超过周围环境温度的温度培养植物，为期至少 2、3、4、5、6、7、8、9、10、11、12、13、14、15、16、17、18、19、20 或更多天，其中温度超过周围环境温度的天出现在植物生长季期间。

[0081] 提供如下实施例作为举例说明而非限制。

[0082] 实验

[0083] 实施例 1. GRG31 的 Q 环坐标

[0084] 2007 年 1 月 10 日递交的美国专利申请 No. 11/651,752（在此并入作为参考）公开了 Q 环是赋予 EPSP 合酶草甘膦抗性的重要区域。Q 环定义为自对应于 SEQ ID NO :5(GRG1) 氨基酸位置 81 的丙氨酸至对应于 SEQ IDNO :5 氨基酸位置 102 的亮氨酸的区域，其中 Q 环的“核心”定义为 SEQ IDNO :5 的氨基酸位置 84-99。

[0085] 本文对此核心区域中的氨基酸给予位置编号，以简化对此区域每个氨基酸的引述。因而，GRG31Q 环核心 (I-D-I-G-P-A-G-T-A-M-R-F-L-T-A-Y) 的位置对应于 GRG31(SEQ ID NO :2) 的氨基酸 73-88，并在本文指定如下 :

[0086] 表 1. Q 环核心氨基酸的位置坐标的指定

[0087]

GRG1 (SEQ ID NO: 5)中的氨基酸 (单字母代码)	GRG31 (SEQ ID NO: 2)中的氨基酸 (单字母代码)	Q 环核心中指定的位置
I	I73	位置 1
D	D74	位置 2
C	I75	位置 3
G	G76	位置 4
E	P77	位置 5
S	A78	位置 6
G	G79	位置 7
L	T80	位置 8
S	A81	位置 9
I	M82	位置 10
R	R83	位置 11
M	F84	位置 12
F	L85	位置 13
T	T86	位置 14
P	A87	位置 15
I	Y88	位置 16

[0088] 实施例 2. Syngrg31 设计和表达

[0089] 设计并合成编码 GRG31 蛋白(美国专利申请 No. 11/760,570,2007 年 6 月 8 日递交,在此全文并入作为参考)的新基因序列。此序列在本文提供为 SEQ ID NO:6。此开放读框在本文命名为“syngrg31”,通过本领域已知的方法克隆到表达载体 pRSF1b(Invitrogen) 中。

[0090] 实施例 3. GRG31 的定点诱变

[0091] 美国专利申请 No. 11/762,580,2008 年 6 月 13 日递交,在此全文并入作为参考,描述了产生多样性的方法。此方法在本文称为“全排列诱变”(permutational mutagenesis)。利用其他草甘膦抗性 EPSPS 酶中存在的 Q 环多样性设计了 GRG31 的全排列文库。

[0092] 开发了草甘膦生长筛选法。发现含有 GRG31 表达构建体 (pAX1945 ;GRG31 在 pRSF1b 中)的 BL21 * DE3 细胞在含有高达 50mM 草甘膦的琼脂板 (1×M63 基本培养基 ; 0.05mM IPTG ;50ug/ml 卡那霉素)上生长。将文库 1 克隆到 pAX1945 的 MfeI 和 AlwN1 限制性位点之间。

[0093] 克隆 pAX3625 含有 grg31 变体,其具有符合读框的终止密码子,该密码子破坏了

GRG31 的开放读框。因而，pAX3625 并不赋予草甘膦抗性。利用 pAX3625 作为载体进行 GRG31 文库的后续克隆和筛选，因为含 pAX3625 的大肠杆菌细胞不会在含草甘膦的琼脂板上生长，除非 pAX3625 的 Q 环盒已被替换为草甘膦抗性 Q 环。

[0094] 将文库 1 克隆到 pAX3625 中，将 130,000 个克隆铺到含 50mM 草甘膦的板上。1.5 天后，由在草甘膦板上的生长，选择到三十一个 (31) 草甘膦抗性克隆。制备提取物并通过在小量草甘膦浓度 (0、1 和 2mM 草甘膦) 存在下测定活性来评估草甘膦抗性。在此预筛中，6 个克隆鉴定为具有最好活性 :GRG31pm3a11、GRG31pm3a6、GRG31pm3a8、GRG31pm3b8、GRG31pm3b9 和 GRG31pm3d3(分别是 SEQ ID NO :7、8、9、10、11 和 12)。

[0095] 这些克隆共有在 Q 环位置 3 上异亮氨酸到半胱氨酸的取代。有意思的是，GRG1 在此位置上也含有半胱氨酸。对来自这 6 个克隆的提取物测定在一系列 PEP 浓度 (0、20、50 和 100 μ M PEP) 时的活性。克隆 3a11 和 3b8 在低 PEP 浓度时显示出最好活性，并进行了纯化。克隆 3b8 相对于野生型 GRG31 显示出提高的 K_i(116 μ M)。

[0096] 表 2. 克隆 3b8 的氨基酸变化

[0097]

克隆	氨基酸变化
3b8	I75 → C ;P77 → N ;M82 → T ;

[0098] 实施例 4. GRG31 文库 3

[0099] 4 个 Q 环位置 (位置 5、9、10 和 16) 鉴定为能够接受变异，选择之在组合文库中进行随机化。此文库命名为文库 3。

[0100] 文库 3 克隆到 pAX3625 中，对未选择的克隆进行测序，发现文库中已经掺入了所述的多样性。

[0101] 11 万 5 千 (115,000) 个克隆铺到含 50mM 草甘膦的板上，挑取了 384 个草甘膦抗性克隆进行测序。通过这些克隆的 Q 环区域的测序，鉴定到五十三个 (53) 独特的序列。通过测试在 0 和 0.2mM 草甘膦时的活性进行了所述独特克隆的 K_i 评估。根据该评估，鉴定了具有最高草甘膦抗性增加迹象的九个 (9) 克隆。通过比较 0、20 和 50 μ M PEP 时的活性对这 9 个克隆所编码蛋白质的 K_m 进行了评估。通过这些评估鉴定了克隆 L3P2b11(SEQ IDNO : 13) 和 L3P2c11(SEQ ID NO :14) 具有草甘膦抗性和低 PEP 浓度时的活性的最佳组合。还选择了克隆 L3P2d11(SEQ ID NO :15) 进行进一步研究。L3P2c11 含有在 Q 环位置 16 上的单个氨基酸变化 (Y88-F)。

[0102] 表 3. GRG31 变体的氨基酸变化

[0103]

克隆	**氨基酸变化**
L3P2b11	**P77→Q; A81→S; M82→L Y88→F**
L3P2c11	**Y88→F**
L3P2d11	**M82→F; Y88→L**

[0104] 实施例 5.GRG31 文库 4

[0105] 生成了命名为文库 4 的全排列文库 (permutational library)，并克隆到 pAX3625 中。3 万 5 千个 (35,000) 克隆铺到含 50mM 草甘膦的板上，并从筛选的该组克隆中鉴定到二十三个 (23) 独特的克隆。制备来自这 23 个克隆的提取物，并通过测试 0、0.2 和 0.5mM 草甘膦时其编码蛋白质的活性进行 K_i 评估。选择了 9 个克隆，并评估 0、20 和 50 μM PEP 时其编码蛋白质的 K_m。在这些测试中克隆 L4e11(SEQ ID NO:16) 鉴定为具有有利性能的克隆。纯化了 L4e11 蛋白质，并确定了其动力学性能。与本文所述其他 GRG31 变体相比较，L4e11 的动力学示于表 5。

[0106] 实施例 6.GRG31 变体 L5e3 的分离

[0107] 从其中 Q 环位置 2、3、4、5、6、7、9、10、14、15 和 16 诱变的 GRG31 变体文库中分离了能够赋予对 50mM 草甘膦的抗性的独特克隆。对这些变体如之前的文库那样测试 K_m 和 K_i，并对若干克隆进行更详细的动力学表征。一个这样的克隆 L5E3(SEQ ID NO:17) 具有在 GRG31 氨基酸 77 处脯氨酸至苏氨酸的单个氨基酸变化。根据表 1，此氨基酸改变对应于 Q 环位置 5。

[0108] 表 4.GRG31 变体的氨基酸变化

[0109]

克隆	氨基酸变化
L4e11	I75→C;A78→S;Y88→H
L5e3	P77→T

[0110] 实施例 7.GRG31 变体的动力学分析

[0111] 通过如本文所述的酶测定法，表征了 GRG31 变体 L3P2b11、L3P2c11 和 L4e11 以及 L5e3，并与天然 GRG31 酶进行了比较。对于每种酶，确定了在几个草甘膦浓度之每一个处的表观 K_m($K_{m(app)}$)，并利用 $K_{m(app)}$ 对草甘膦浓度的曲线来计算每种酶的 K_i。通过在 37℃孵育酶 16 小时来评估每种酶的热稳定性，然后相对于 4℃孵育的对照酶来定量残留酶活（按 V_{max} 计）。

[0112] 动力学分析揭示，克隆 L3P2b11、L3P2c11 和 L4e11 相对于 GRG31 具有改善的动力学性能，如表 5 所示。

[0113] 表 5. 与其他 GRG31 变体相比较的 L4e11 动力学

[0114]

	GRG31	L3P2b11	L3P2c11	L4e11	L5e3
K_i(μM)	73	983	715	389	507
K_m(μM)	23	11	39.6	12.4	11.4
V_{max}	29.3	5.7	5.3	3.5	3.6

[0115] 表 6. GRG31 变体中 Q 环变化的位置

[0116]

Q 环中的位置	GRG31 中的氨基酸	Pm3b8	L3P2b11	L3P2c11	L3P2d11	L4e11	L5e3
位置 1	I73	-	-	-	-	-	-
位置 2	D74	-	-	-	-	-	-
位置 3	I75	C	-	-	-	C	-
位置 4	G76	-	-	-	-	-	-
位置 5	P77	N	Q	-	-	-	T
位置 6	A78	-	-	-	-	S	-
位置 7	G79	-	-	-	-	-	-
位置 8	T80	-	-	-	-	-	-
位置 9	A81	-	S	-	-	-	-
位置 10	M82	T	L	-	F	-	-
位置 11	R83	-	-	-	-	-	-
位置 12	F84	-	-	-	-	-	-
位置 13	L85	-	-	-	-	-	-
位置 14	T86	-	-	-	-	-	-
位置 15	A87	-	-	-	-	-	-
位置 16	Y88	-	F	F	L	H	-

[0117] 克隆 L3P2b11、L3P2c11 和 L4e11 相对于 GRG31 显示出改善的 K_i。

[0118] 观察表 6，说明每一种动力学性能改善的 GRG31Q 环变体要么具有位置 16 上的突变（优选该位置上突变成苯丙氨酸），要么是位置 5 上的突变。表 6 中的每一种突变都具有大为改进的 K_i，和改进的针对磷酸烯醇丙酮酸 (PEP) 的 K_m。这种观察结果与这些变体中至少部分地因 PEP 结合亲和力提高所致的抗性机制相符。PEP 受草甘膦的竞争性抑制，而提高的 PEP 亲和力将导致存在草甘膦时表观 K_m 的减小，以及由此更高的 K_i。有意思的是，这些克隆中的每一种都相对于 GRG31 具有减小的 V_{max}，与此区域改变能够破坏蛋白质折叠的观察结果相符。不过，针对这些蛋白质观察到的 V_{max} 仍然足以使这些蛋白质赋予大肠杆菌细胞对高草甘膦浓度的抗性。如果期望这些变体的 V_{max} 提高，可实施如本文所提供的诱变和筛选策略（包括 Q 环外的区域），并选择呈现出 V_{max} 提高的变体。此类策略已经证明对类似的 EPSP 合酶有效（参见美国专利申请 No. 11/762,526，2007 年 6 月 13 日递交，在此全文并入作为参考）。可选地，可以产生提高 EPSPS 例如 GRG31 蛋白的折叠和 / 或热稳定性的突变，

然后组合这些突变，并测试具有变体诸如 L3P2b11、L3P2c11、L3P2d11、L4e11 和 / 或 L5e3 的提高的结合特性，但还具有 V_{max} 方面的改进的变体。

[0119] 实施例 8. 开发具有提高的热稳定性的变体

[0120] 为提高 GRG31 的热稳定性，进行 GRG31 的易错 PCR 诱变。4 万 5 千个 (45,000) 克隆铺到 50mM 草甘膦板上，并挑取九十五个 (95) 克隆。制备提取物，并测试增加的热稳定性。2 个克隆 (6b2，SEQ ID NO :18，和 6e7，SEQ ID NO :19) 鉴定为具有增加的热稳定性。序列分析显示，相对于 GRG31，克隆 6b2 含有两个突变 (N95S 和 A206V)，而克隆 6e7 含有 1 个突变 (E215K)。

[0121] 表 7. 具有提高的热稳定性的 GRG31 变体的氨基酸变化

[0122]

克隆	氨基酸变化
6b2	N95 → S，A206 → V
6e7	E215 → K

[0123] 分子建模表明，这三个突变位置位于 GRG31 的表面。

[0124] 为进一步研究这些位置上不同取代的贡献，在对应于 GRG31 氨基酸位置 95、206 和 215 的位置上产生几个点突变。在 37℃测试了点突变体提取物的热稳定性 :

[0125] 表 8.GRG31 变体的热稳定性

[0126]

氨基酸变化	Wt GRG31 活性的百分比 (4 小时 37℃)
N95→S, A206→V	228%
E215→K	213%
N95→S	229%
N95→Q	170%
A206→V	213%
A206→S	121%
E215→Q	276%

[0127] 在分开的实验中，分析了更多的残基变化。

[0128] 表 9.GRG31 变体的热稳定性

[0129]

氨基酸变化	Wt GRG31 活性的百分比 (4 小时 37℃)
N95→S, A206→V	139%
E215→K	152%
N95→E	48%
N95→D	57%
N95→T	118%
A206→R	104%
A206→Q	123%
A206→E	152%
A206→T	157%
E215→R	118%
E215→S	143%
E215→T	111%

[0130] 数据表明，点突变 N95S 和 A206V 两者均有助于增加的稳定性。

[0131] 培养 3 个变体 (6b2、6e7、以及三联 EK(N95S A206V E215K ;SEQ IDNO :20))，纯化其蛋白质，并更详细地分析所述蛋白质的热稳定性。数据（表 10）表明，变体 6b2、6e7 以及三联突变体与野生型 GRG31 相比具有增加的热稳定性。

[0132] 表 10. GRG31 变体的热稳定性

[0133]

	GRG31	6e7	6b2	三联 EK
热稳定性 (4 小时 37℃)	++	+++	+++	+++

[0134] 实施例 9. GRG31 变体的动力学

[0135] 还利用纯化的蛋白质确定了 6e2、6b2 和三联 EK 变体的动力学特性，并测试了蛋白质的溶解性。结果示于表 11。所有三个变体 6e7、6b2 和三联 EK 编码的蛋白质相对于 GRG31 具有针对 PEP 的改进的（更低的）K_m。而且，6e7 还具有提高的 V_{max}。

[0136] 表 11. GRG31 变体的动力学特性

[0137]

	GRG31	6e7	6b2	三联 EK
K_i(μM)	73	111	92	48.2
K_m(μM)	23	3.4	6.5	7.2
V_{max}	29.3	62	34	33

[0138] 克隆 6e7 值得注意，因为其 V_{max} 比 GRG31 提高 2 倍。

[0139] 因而，这些改进的克隆可以适于与在其他动力学性能方面（例如针对 PEP 的 K_m 或就草甘膦而言的 K_i）具有改进的其他变体组合。

[0140] 实施例 10. 组合热稳定性改进与动力学改进

[0141] 设计 GRG31 变体，将 6e7、6b2 和三联 EK 突变中鉴定到的热稳定性改进突变掺入到 L3P2b11 或 L3P2c11 变体中。如表 12 所示，L3P2c11+6e7(SEQ ID NO:21) 编码的蛋白质相对于 GRG31 具有 2 个变化。如表 12 所示，L3P2b11+6e7(SEQ ID NO:22) 编码的蛋白质相对于 GRG31 具有 5 个变化。如表 12 所示，L3P2c11+ 三联 EK(SEQ ID NO:23) 编码的蛋白质相对于 GRG31 具有 4 个变化。如表 12 所示，L3P2b11+ 三联 EK(SEQ ID NO:24) 编码的蛋白质相对于 GRG31 具有 7 个变化。如表 12 所示，L3P2c11+6b2(SEQ ID NO:25) 编码的蛋白质相对于 GRG31 具有 3 个变化。L3P2b11+6b2(SEQ ID NO:26) 编码的蛋白质相对于 GRG31 具有 6 个变化。

[0142] 表 12. GRG31 变体的氨基酸变化

[0143]

克隆	氨基酸变化
L3P2c11 +6e7	**Y88→F; E215→K**
L3P2b11 +6e7	**P77→Q; A81→S; M82→L Y88→F ; E215→K**
L3P2c11+三联 EK	**Y88→F ; N95→S, A206→V; E215→K**
L3P2b11+三联 EK	**P77→Q; A81→S; M82→L Y88→F ; N95→S, A206→V; E215→K**
L3P2c11 +6b2	**Y88→F ; N95→S, A206→V**
L3P2b11 +6b2	**P77→Q; A81→S; M82→L Y88→F ; N95→S, A206→V**

[0144] 表 13. GRG31 变体的动力学

[0145]

	GRG31	L3P2c11 +6e7	L3P2b11 +6e7	L3P2c11 + 三联 EK	L3P2b11 + 三联 EK	L3P2c11 +6b2	L3P2b11 +6b2
K_i (μM)	73	416	1013	298	606	405	628
K_m (μM)	23	8.6	10.3	12.6	12.3	7.9	8
V_{max}	29.3	23.5	0.86	14.4	0.72	23.4	0.59

[0146] 表 13 显示，具有组合突变的变体的性能优于具有任一单组改变的变体。

[0147] 变体 GRG31(L3P2c11+6e7) 和 GRG31(L3P2c11+6b2) 呈现出超过 4 倍改进的对草甘膦的 K_i，大约两倍改进的针对 PEP 的 K_m，以及大约等同于 GRG31 并且相对于 GRG31(L3P2c11) 蛋白约 4-5 倍改进的 V_{max}。

[0148] 实施例 11. GRG31 变体的热稳定性

[0149] 通过测量 37℃的残存活性，分析了若干 GRG31 变体的热稳定性。虽说 L3P2c11+ 三联 EK 显示出等同于 GRG31 的热稳定性，但 L3P2c11+6b2 和 L5E3 两种均显示出提高的稳定性。因而，这两个克隆具有改进的性能：更高的草甘膦 K_i，和 37℃改进的热稳定性。

[0150] 表 14. GRG31 在 37℃的半衰期

[0151]

	GRG31	L5E3	L3P2c11+ 三联 EK	L3P2c11+6b2
37℃的 $t_{1/2}$（小时）	2.7	10.8	2.7	3.5

[0152] 实施例 12. GRG36 的改进

[0153] 将编码 GRG36(SEQ ID NO:4)、优化用于植物表达的基因（称为“syngrg36”，如 2007 年 6 月 27 日递交的美国专利申请 No. 11/769,327 所述）克隆到植物转化质粒 pAG3532 中。该质粒用作为模板利用 Mutazyme II 系统 (Stratagene) 进行易错 PCR，以向 syngrg36 中引入随机突变。模板在易错 PCR 反应中进行 1 ： 50 稀释，且扩增进行 30 个循环。利用此载体中位于 syngrg36 侧翼 (T7 启动子，T7 终止子）的 PCR 引物。PCR 产物以 BamHI 和 sgs I 在 Tango 缓冲液中消化，凝胶纯化，并连接到细菌表达载体 pRSF-1b 载体中，以创建诱变了的 grg36 文库。

[0154] syngrg36DNA 文库转化到大肠杆菌菌株 BL21 * DE3 star(Invitrogen) 中以诱导蛋白质表达。转化后，将单菌落铺到含 25mM 草甘膦的 1×M63 培养基中，以选择保留酶活和生长耐受性的克隆。挑取单菌落，并排列到 384 孔板中，以创建诱变了的文库进行酶活筛选。以这种方式创建 2 个 384 孔板。酶活筛选进行如下：将文库克隆种在含 LB 培养基的 96 孔深孔板中，并培养至 A600 = 0.6。然后添加 IPTG(0.5mM)，并在 20℃过夜孵育深孔板，以诱导蛋白质表达。接下来，利用**POPCULTURE®**试剂 (Novagen) 由各单个培养物制备蛋白质提取物，并如 2007 年 1 月 10 日递交的美国专利申请 No. 11/651,752（在此全文并

入作为参考）所公开的那样测量酶活。通过这种方法，鉴定到 6 个提取物，其拥有的酶活比一同测试的野生型 GRG36 样品的平均值高 2 个标准差。对这些克隆中的每一个进行 DNA 测序，并发现每一个都在 GRG36 中携带突变。选择这 6 个进行进一步表征。

[0155] 6 个改进克隆中的每一个都在 250mL LB 培养液中培养和诱导。诱导后，每一种 GRG36 蛋白变体都通过亲和层析纯化到钴树脂 (Novagen) 上。纯化的蛋白质在 37℃加热 0、4 和 6 小时后测试酶活。一种 GRG36 蛋白质变体，称为 GRG36(ace7)(SEQ ID NO :27)，由 syngrg36(ace7)(SEQ IDNO :29) 编码，发现在 37℃拥有比野生型 GRG36 酶高 7.5 倍的热稳定性。另一种 GRG36 蛋白质变体，称为 GRG36(ace8)(SEQ ID NO :28)，由 syngrg36(ace8)(SEQ ID NO :30) 编码，发现拥有比野生型 GRG36 高 2.5 倍的热稳定性。结果总结于表 15 :

[0156] 表 15.GRG36 酶在 37℃的半衰期

[0157]

酶	37℃的半衰期
GRG36	2.5 小时
GRG36(ace7)	18.8 小时
GRG36(ace8)	7.3 小时

[0158] 另外，进行酶测定试验，以计算底物结合亲和力 (K_m)、草甘膦结合 (K_i) 以及酶速率 (V_{max}) 的动力学常数。简言之，用底物磷酸烯醇式丙酮酸滴定每一种酶，并相对于底物浓度绘制线性酶速率曲线图。通过此曲线拟合 Michaelis-Menton 方程，以计算产生半数最大速率的底物浓度 (K_m)，且用饱和底物时获得的最大速率除以酶浓度以得出 V_{max}。接下来，在不同的草甘膦浓度进行另外的底物滴定，并相对于草甘膦浓度，绘制在每种草甘膦浓度时的表观结合常数 (K_{obs}) 的图，得到直线。K_i 测量为 (-)x- 截距。对于酶 GRG36(ace7)，未进行完整的 K_i 测量，但在 1mM 草甘膦时观察到 GRG36 和 GRG36(ace7) 之间非常相似的活性，提示这两种酶拥有相似的草甘膦结合性能。

[0159] 表 16.GRG36 酶的动力学数值

[0160]

酶	V_{max} (nmol/min/μg)	K_m (μM)	K_i (μM)
GRG36	**8**	**6.3**	**421**
GRG36(ace7)	**1.8**	**9.8**	**未测**
GRG36(ace8)	**6.8**	**11.03**	**435**

[0161] 实施例 13. 克隆 grg 变体到植物表达盒中

[0162] 对于本文所述的每一种 grg 变体，通过 PCR 从全长 DNA 模板扩增开放读框 (ORF)。PCR 期间在 ORF 的每一端添加 Hind III 限制性位点。另外，在紧邻基因起始密码子的

5’处添加核苷酸序列 ACC，以增加翻译效率 (Kozak(1987)NucleicAcids Research 15：8125-8148；Joshi(1987)NucleicAcids Research 15：6643-6653)。PCR 产物利用本领域公知的技术进行克隆和测序，以确保 PCR 期间不引入突变。

[0163] 含 PCR 产物的质粒用 Hind III 消化，并分离含完整 ORF 的片段。此片段克隆到质粒的 Hind III 位点，所述质粒例如 pAX200，一种含有稻肌动蛋白启动子 (McElroy 等 (1991)Molec.Gen.Genet.231：150-160) 和 PinII 终止子 (An 等 (1989)The Plant Cell 1：115-122) 的植物表达载体。然后将来自此中间质粒的启动子 - 基因 - 终止子片段亚克隆到质粒 pSB11(JapanTobacco，Inc.) 中，以形成最终基于 pSB11 的质粒。在一些情况下，可能优选生成备选构建体，其中叶绿体前导序列以与 grg 构建体 N- 端的融合物形式编码。这些基于 pSB11 的质粒一般这样组织，从而含有启动子 - 基因 - 终止子构建体或启动子 - 叶绿体前导序列 - 基因 - 终止子构建体的 DNA 片段，可以通过限制性酶如 KpnI 和 PmeI 的双消化而被切出，并通过气溶胶束注射而转化到植物中。所得基于 pSB11 的克隆的结构通过限制性消化和凝胶电泳以及通过跨各克隆接合点的测序来验证。

[0164] 利用本领域公知的三亲交配法，将质粒转移到还持有质粒 pSB1 的根癌农杆菌菌株 LBA4404(Japan Tobacco，Inc.) 中，并铺到含壮观霉素的培养基板上。基于 pSB11 的质粒克隆携带壮观霉素抗性，但属于窄宿主范围质粒，不能够在农杆菌中复制。当基于 pSB11 的质粒通过同源重组整合到宽宿主范围质粒 pSB1 中时，产生壮观霉素抗性菌落。pSB1 和基于 pSB11 的质粒的共整合产物通过 Southern 杂交来验证。持有所述共整合体的农杆菌菌株用来通过本领域已知的方法转化玉蜀黍，例如像 PureIntro 法 (JapanTobacco)。

[0165] 实施例 14. 通过农杆菌介导的转化法转化植物细胞

[0166] 玉米穗最好在授粉后 8-12 天收集。从穗中分离胚，并优选 0.8-1.5mm 大小的那些胚用于转化。胚盾片面向上放在适宜的孵育培养基上，例如 DN62A5S 培养基 (3.98g/L N6 盐；1ml/L(1000× 贮液)N6 维生素；800mg/LL- 天冬酰胺；100mg/L 肌醇；1.4g/L L- 脯氨酸；100mg/L 酪蛋白氨基酸；50g/L 蔗糖；1ml/L(1mg/ml 贮液)2,4-D)。不过，除 DN62A5S 之外的培养基和盐也适宜，并且是本领域已知的。胚在黑暗中 25℃孵育过夜。不过，本质上并不需要孵育胚过夜。

[0167] 所得的外植体转移到载网 (30-40/ 板) 上，转移到渗透培养基上约 30-45 分钟，然后转移到 beaming plate 上 (参见例如 PCT 公开号 WO/0138514 和美国专利 No.5,240,842)。

[0168] 设计用来在植物细胞中表达本发明 GRG 蛋白质的 DNA 构建体，利用基本上如 PCT 公开号 WO/0138514 中所述的条件，利用气溶胶束加速器加速进入到植物组织中。气溶胶束处理后，胚在渗透培养基上孵育约 30 分钟，然后置于孵育培养基上 25℃黑暗过夜。为避免过度地损伤经气溶胶束处理的外植体，在向恢复培养基转移之前孵育至少 24 小时。然后将胚铺在恢复期培养基上，25℃黑暗约 5 天，然后转移到选择培养基中。外植体在选择培养基中孵育长达 8 周，这取决于所利用的具体选择的性质和特征。选择期之后，将所得的愈伤组织转移到胚成熟培养基中，直至观察到形成成熟的体细胞胚。所得的成熟体细胞胚随之置于低光照条件下，并通过本领域已知的方法起始再生过程。允许所得的芽在生根培养基上生根，并将所得的植株转移到育苗盆中，作为转基因植物繁殖。

[0169] 材料

[0170] DN62A5S 培养基

[0171]

组分	每升	来源
Chu's N6 基础盐混合物 (产品号 C 416)	3.98 g/L	Phytotechnology Labs
Chu's N6 维生素溶液 (产品号 C 149)	1 ml/L (1000×贮液)	Phytotechnology Labs
L-天冬酰胺	800 mg/L	Phytotechnology Labs
肌醇	100 mg/L	Sigma
L-脯氨酸	1.4 g/L	Phytotechnology Labs
酪蛋白氨基酸	100 mg/L	Fisher Scientific
蔗糖	50 g/L	Phytotechnology Labs
2, 4-D (产品号 D-7299)	1 ml/L (1 mg/ml 贮液)	Sigma

[0172] 用 1N KOH/1N KCl 调整溶液 pH 至 pH 5.8，添加 Gelrite(Sigma) 至 3g/L，并高压灭菌。冷却到 50℃之后，添加 2ml/L 的 5mg/ml 硝酸银贮液 (Phytotechnology Labs)。所述配方支持约 20 个板。

[0173] 实施例 15. 通过农杆菌介导的转化法将 EPSP 合酶转化到玉蜀黍植物细胞中

[0174] 穗最好在授粉后 8-12 天收集。从穗中分离胚，并优选 0.8-1.5mm 大小的那些胚用于转化。胚盾片面向上置于适宜的孵育培养基中，并在黑暗中 25℃孵育过夜。

[0175] 不过，本身并不需要过夜孵育胚。胚与含有具有本发明 EPSP 合酶的适当载体的农杆菌菌株接触，进行 Ti 质粒介导的转移约 5-10 分钟，然后铺在共培养培养基上约 3 天 (25℃黑暗中)。共培养后，外植体转移到恢复期培养基中约 5 天 (25℃黑暗中)。外植体在选择培养基中孵育长达 8 周，这取决于所利用的具体选择的性质和特征。选择期之后，将所得的愈伤组织转移到胚成熟培养基中，直至观察到形成成熟的体细胞胚。所得的成熟体细胞胚随之置于低光照条件下，并如本领域已知的那样起始再生过程。允许所得的芽在生根培养基上生根，并将所得的植株转移到育苗盆中，作为转基因植物繁殖。

[0176] 本说明书中提及的所有出版物和专利申请指示了本发明所述领域技术人员的技术水平。所有出版物和专利申请在此并入作为参考，就如同每一篇单独出版物或专利申请明确单独地说明并入作为参考一样。

[0177] 虽然出于澄清理解的目的，已经通过举例说明和实施例的方式对前述发明进行了一些详细的说明，但是显然可以在所附权利要求书的范围内进行某些变化和改动。

[0001]

序列表

<110> 阿森尼克斯公司

<120> GRG31 和 GRG36 EPSP 合酶的定向进化

<130> 45600/367530

<150> 61/025,406
<151> 2008-02-01

<160> 30

<170> FastSEQ for Windows Version 4.0

<210> 1
<211> 1251
<212> DNA
<213> 多食/屎鞘氨醇杆菌 (Sphingobacterium multivorum/faecium)

<220>
<221> CDS
<222> (1)...(1251)

<400> 1

```
atg aat caa caa gtc atc acg ctg acg cat cct tca aaa aaa ata cag   48
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
 1               5                  10                  15

ggt acg gtt caa ctc aca ggt tca aaa tct gag agc aac cgt gct ctt   96
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
            20                  25                  30

att atc cag tca ttg agc aaa gga caa gtt gaa ata gcc aac ctt tct   144
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
        35                  40                  45

gaa gct gca gat acg gta acg tta aac cgt gtg cta caa att gct tca   192
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
    50                  55                  60

gat ccg aaa cca gga ttc aac aca att gac atc ggt cca gcg gga acg   240
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65                  70                  75                  80
```

[0002]

gcc atg cga ttc tta acg gct tac ctc aac ctt gtc aaa gga aat ttt 288
Ala Met Arg Phe Leu Thr Ala Tyr Leu Asn Leu Val Lys Gly Asn Phe
85 90 95

atc ctt aca ggt act gaa cgc atg caa cag cgt cct ata ggt ata tta 336
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110

gtt gat gcc atg aaa gaa att ggt gca gac atc cac tat gac aag aaa 384
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125

gtc gga tac cct cct ttg aaa att gag ggc ggg ctg ttt caa gaa aaa 432
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140

gac cgt gtc aag att aaa ggt aat atc agc agc caa tat ata tca gcc 480
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160

ctc tta tta att gca cct gca tta aaa aaa ggg ctt act tta gag att 528
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

gag ggt gaa tta acc tcc aga cct tat gta tca atg acc ttg gat atg 576
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190

ctg aaa agt gtc ggg att cag cat gaa tgg aaa aac aat gcg att aaa 624
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205

att gcg ccg cag gca ttt gag aaa caa aca ata tat gtc gag cca gat 672
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220

tgg agc gct gct tcc tat tgg tac gct atc gcc gca cta gca gat gca 720
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240

aac gca tcg atc gta ttg ccc gga tta aga aaa aac agc tta cag ggt 768
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255

[0003]

```
gat att gct att ata agc att atg gag cat ttt ggt gta caa tcg agc   816
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
            260                 265                 270

ttt gaa tcg gac gga tta cac tta aat aaa aag gta atc ggt tcg gat   864
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
        275                 280                 285

gta agc tta ttt aac ttt aaa gaa tgt ccc gat ctc gca caa act gta   912
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
    290                 295                 300

gtt gtt gtc gcg gct gcg tta aaa cga gat gta tct ttt acg ggc ttg   960
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305                 310                 315                 320

gag acc tta aaa att aag gag act gac cgt atc gcg gca cta caa aag   1008
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
                325                 330                 335

gaa att gcg aaa ttt gga gcc gag cta att gaa gat ggc gat acc tac   1056
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
            340                 345                 350

cat ctg aaa aca gcg cag gta tat cag cct gaa gag gtt act ttc gat   1104
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
        355                 360                 365

act tac gaa gat cat cgc atg gcg atg gcg ttc gca cca ctg gca tta   1152
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
    370                 375                 380

gtt ttc gac cag att aag att gct gaa cct caa gtt gta gaa aaa tca   1200
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385                 390                 395                 400

tat cct gat ttt tgg aat cat tta cag gcg caa gct ttt gtc att gaa   1248
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
                405                 410                 415

tag                                                               1251
 *
```

[0004]

<210> 2
<211> 416
<212> PRT
<213> Sphingobacterium multivorum/faecium

<400> 2

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

[0005]

305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 3
<211> 1344
<212> DNA
<213> 未知

<220>
<223> 源于分离自土壤的生物

<221> CDS
<222> (1)...(1344)

<400> 3
atg aaa aat caa aat ttt gat gct aaa gcc cgt agc ccg tgg aca ccg 48
Met Lys Asn Gln Asn Phe Asp Ala Lys Ala Arg Ser Pro Trp Thr Pro
1 5 10 15

tta aaa ggt gta aat aag ata agt gtt tca cct agt aaa gga aga ata 96
Leu Lys Gly Val Asn Lys Ile Ser Val Ser Pro Ser Lys Gly Arg Ile
20 25 30

aat gga acc gtt act att cct ggt agc aag agt tta acc aac aga gct 144
Asn Gly Thr Val Thr Ile Pro Gly Ser Lys Ser Leu Thr Asn Arg Ala
35 40 45

tta atc ata agt tct cta gct agt gga aag tca aaa gtg caa ggt att 192
Leu Ile Ile Ser Ser Leu Ala Ser Gly Lys Ser Lys Val Gln Gly Ile
50 55 60

tta aaa agc gat gat tca ttt tgg tgt tta gac tcc ttg aaa aag cta 240
Leu Lys Ser Asp Asp Ser Phe Trp Cys Leu Asp Ser Leu Lys Lys Leu

[0006]

65 70 75 80

gga gta aat gtg aaa att caa gga gat aca gct ttt atc gaa gga aac 288
Gly Val Asn Val Lys Ile Gln Gly Asp Thr Ala Phe Ile Glu Gly Asn
85 90 95

ggt ggt aaa tgg gaa tca ggt gat tta tat att ggt gca gca gga acg 336
Gly Gly Lys Trp Glu Ser Gly Asp Leu Tyr Ile Gly Ala Ala Gly Thr
100 105 110

att gca cga ttt cta cct gga gca tta gca gtt tca ggt aca ggc ata 384
Ile Ala Arg Phe Leu Pro Gly Ala Leu Ala Val Ser Gly Thr Gly Ile
115 120 125

tgg gag tta gaa gcc agt aaa agt atg agt aaa cga cct att tca ccc 432
Trp Glu Leu Glu Ala Ser Lys Ser Met Ser Lys Arg Pro Ile Ser Pro
130 135 140

tta gta gat gct tta aaa gag ctt ggg gct gaa ata aca tat cta agc 480
Leu Val Asp Ala Leu Lys Glu Leu Gly Ala Glu Ile Thr Tyr Leu Ser
145 150 155 160

gat caa ggc tac tat ccg ttg tta gtt aaa gga aaa caa cta aat ggg 528
Asp Gln Gly Tyr Tyr Pro Leu Leu Val Lys Gly Lys Gln Leu Asn Gly
165 170 175

ggc gag gtt gaa ctc tca ggt aga att tct agt cag ttt ata agt ggt 576
Gly Glu Val Glu Leu Ser Gly Arg Ile Ser Ser Gln Phe Ile Ser Gly
180 185 190

cta ttg att gcc tcg cct tat tta aat gat cca atc aag att aat att 624
Leu Leu Ile Ala Ser Pro Tyr Leu Asn Asp Pro Ile Lys Ile Asn Ile
195 200 205

aaa gat cac atc gtt caa cac tca tat gta ctc tta act ttg gaa tta 672
Lys Asp His Ile Val Gln His Ser Tyr Val Leu Leu Thr Leu Glu Leu
210 215 220

atg aaa aag ttt ggt gca aaa gtt aaa tac gat agt agc cta aaa gaa 720
Met Lys Lys Phe Gly Ala Lys Val Lys Tyr Asp Ser Ser Leu Lys Glu
225 230 235 240

ata gtc gtc tat cca tct aag tac act cca caa gat ata aat tta gaa 768
Ile Val Val Tyr Pro Ser Lys Tyr Thr Pro Gln Asp Ile Asn Leu Glu
245 250 255

[0007]

```
gca gat gtt tcc act gca tgt tat ttt ctg gct ctt gct gca gtg acc   816
Ala Asp Val Ser Thr Ala Cys Tyr Phe Leu Ala Leu Ala Ala Val Thr
            260                 265                 270

aac ggt aaa gta caa att gat aat cta act tat gaa aca aaa caa cca   864
Asn Gly Lys Val Gln Ile Asp Asn Leu Thr Tyr Glu Thr Lys Gln Pro
        275                 280                 285

gat ata aaa atg gtt gat atc ctt gaa cgg atg gga tgc aaa gta aca   912
Asp Ile Lys Met Val Asp Ile Leu Glu Arg Met Gly Cys Lys Val Thr
    290                 295                 300

aga ggt tct tca ttc att gaa ata gag gga gtt agt caa tta aaa ggt   960
Arg Gly Ser Ser Phe Ile Glu Ile Glu Gly Val Ser Gln Leu Lys Gly
305                 310                 315                 320

gga ttt gaa atc tct atg agg gaa atg tct gac caa gtg tta act cta  1008
Gly Phe Glu Ile Ser Met Arg Glu Met Ser Asp Gln Val Leu Thr Leu
                325                 330                 335

gca gca att gct cca ttc gca gat gaa cca ata acc ata aaa gac gtt  1056
Ala Ala Ile Ala Pro Phe Ala Asp Glu Pro Ile Thr Ile Lys Asp Val
            340                 345                 350

gaa cat ata cgc cat cac gaa tca aat cga atc agt gta cta gtt gat  1104
Glu His Ile Arg His His Glu Ser Asn Arg Ile Ser Val Leu Val Asp
        355                 360                 365

tca cta tct agg tta gga att ata gta gaa gaa ttt aaa gat gga cta  1152
Ser Leu Ser Arg Leu Gly Ile Ile Val Glu Glu Phe Lys Asp Gly Leu
    370                 375                 380

aaa gtg tat ccg ggt aat ccg aaa gcc act tta cta gat aca cac gat  1200
Lys Val Tyr Pro Gly Asn Pro Lys Ala Thr Leu Leu Asp Thr His Asp
385                 390                 395                 400

gat cat aga gtt gca atg gca tta tca ctt ata ggt tca aga gtt gaa  1248
Asp His Arg Val Ala Met Ala Leu Ser Leu Ile Gly Ser Arg Val Glu
                405                 410                 415

ggt ata caa ata aat gat cca gga tgt gta tct aaa act tgt cct cag  1296
Gly Ile Gln Ile Asn Asp Pro Gly Cys Val Ser Lys Thr Cys Pro Gln
            420                 425                 430
```

[0008]

```
tat ttt gaa tta ttg gaa aaa cta ggt ttg aat ata att aaa cat tga   1344
Tyr Phe Glu Leu Leu Glu Lys Leu Gly Leu Asn Ile Ile Lys His  *
        435                 440                 445
```

<210> 4
<211> 447
<212> PRT
<213> 未知

<220>
<223> 源于分离自土壤的生物

<400> 4

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

[0009]

Ile Val Val Tyr Pro Ser Lys Tyr Thr Pro Gln Asp Ile Asn Leu Glu
245 250 255
Ala Asp Val Ser Thr Ala Cys Tyr Phe Leu Ala Leu Ala Ala Val Thr
260 265 270
Asn Gly Lys Val Gln Ile Asp Asn Leu Thr Tyr Glu Thr Lys Gln Pro
275 280 285
Asp Ile Lys Met Val Asp Ile Leu Glu Arg Met Gly Cys Lys Val Thr
290 295 300
Arg Gly Ser Ser Phe Ile Glu Ile Glu Gly Val Ser Gln Leu Lys Gly
305 310 315 320
Gly Phe Glu Ile Ser Met Arg Glu Met Ser Asp Gln Val Leu Thr Leu
325 330 335
Ala Ala Ile Ala Pro Phe Ala Asp Glu Pro Ile Thr Ile Lys Asp Val
340 345 350
Glu His Ile Arg His His Glu Ser Asn Arg Ile Ser Val Leu Val Asp
355 360 365
Ser Leu Ser Arg Leu Gly Ile Ile Val Glu Glu Phe Lys Asp Gly Leu
370 375 380
Lys Val Tyr Pro Gly Asn Pro Lys Ala Thr Leu Leu Asp Thr His Asp
385 390 395 400
Asp His Arg Val Ala Met Ala Leu Ser Leu Ile Gly Ser Arg Val Glu
405 410 415
Gly Ile Gln Ile Asn Asp Pro Gly Cys Val Ser Lys Thr Cys Pro Gln
420 425 430
Tyr Phe Glu Leu Leu Glu Lys Leu Gly Leu Asn Ile Ile Lys His
435 440 445

<210> 5
<211> 431
<212> PRT
<213> 肠杆菌科(Enterobacteriaceae)

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

[0010]

85 90 95
Thr Pro Ile Val Ala Leu Ser Lys Glu Glu Val Thr Ile Lys Gly Ser
100 105 110
Gly Ser Leu Val Thr Arg Pro Met Asp Phe Phe Asp Glu Ile Leu Pro
115 120 125
His Leu Gly Val Lys Val Lys Ser Asn Gln Gly Lys Leu Pro Leu Val
130 135 140
Ile Gln Gly Pro Leu Lys Pro Ala Asp Val Thr Val Asp Gly Ser Leu
145 150 155 160
Ser Ser Gln Phe Leu Thr Gly Leu Leu Leu Ala Tyr Ala Ala Ala Asp
165 170 175
Ala Ser Asp Val Ala Ile Lys Val Thr Asn Leu Lys Ser Arg Pro Tyr
180 185 190
Ile Asp Leu Thr Leu Asp Val Met Lys Arg Phe Gly Leu Lys Thr Pro
195 200 205
Glu Asn Arg Asn Tyr Glu Glu Phe Tyr Phe Lys Ala Gly Asn Val Tyr
210 215 220
Asp Glu Thr Lys Met Gln Arg Tyr Thr Val Glu Gly Asp Trp Ser Gly
225 230 235 240
Gly Ala Phe Leu Leu Val Ala Gly Ala Ile Ala Gly Pro Ile Thr Val
245 250 255
Arg Gly Leu Asp Ile Ala Ser Thr Gln Ala Asp Lys Ala Ile Val Gln
260 265 270
Ala Leu Met Ser Ala Asn Ala Gly Ile Ala Ile Asp Ala Lys Glu Ile
275 280 285
Lys Leu His Pro Ala Asp Leu Asn Ala Phe Glu Phe Asp Ala Thr Asp
290 295 300
Cys Pro Asp Leu Phe Pro Pro Leu Val Ala Leu Ala Ser Tyr Cys Lys
305 310 315 320
Gly Glu Thr Lys Ile Lys Gly Val Ser Arg Leu Ala His Lys Glu Ser
325 330 335
Asp Arg Gly Leu Thr Leu Gln Asp Glu Phe Gly Lys Met Gly Val Glu
340 345 350
Ile His Leu Glu Gly Asp Leu Met Arg Val Ile Gly Gly Lys Gly Val
355 360 365
Lys Gly Ala Glu Val Ser Ser Arg His Asp His Arg Ile Ala Met Ala
370 375 380
Cys Ala Val Ala Ala Leu Lys Ala Val Gly Glu Thr Thr Ile Glu His
385 390 395 400
Ala Glu Ala Val Asn Lys Ser Tyr Pro Asp Phe Tyr Ser Asp Leu Lys
405 410 415
Gln Leu Gly Gly Val Val Ser Leu Asn His Gln Phe Asn Phe Ser
420 425 430

[0011]

<210> 6
<211> 1251
<212> DNA
<213> 人工序列

<220>
<223> 编码 GRG31 蛋白的合成的核苷酸序列
(syngrg31)

<400> 6
atgaaccagc aggtgatcac cttgacacat ccaagcaaga agatccaagg caccgtgcag 60
ctcaccggca gcaagagcga gagcaacagg gcgctcatca ttcagagctt gagcaagggc 120
caagtggaga tcgccaacct ctcagaagct gctgacaccg tcaccctcaa cagggtgctg 180
caaattgctt cagatccaaa gcctggcttc aacaccatcg acattgggcc ggccggcacc 240
gccatgaggt tcttgacggc ctacctcaac ctggtgaagg gcaacttcat cctcaccggc 300
actgaaagga tgcagcagcg gcccatcggc atcctggtgg acgccatgaa ggagatcggc 360
gccgacatcc actacgacaa gaaggtgggc tacccgccgc taaagattga aggaggcctc 420
ttccaagaga aggaccgcgt caagatcaag ggcaacatca gcagccagta catctcggcg 480
ctgctgctca tcgcgccggc gctgaagaag gggctgacgc tggagattga aggagagctg 540
acatcaaggc cttatgtgag catgacgctg gacatgctga agagcgtcgg catccagcat 600
gaatggaaga acaacgccat caagattgct cctcaagcat ttgagaagca gaccatctat 660
gtggagccag attggagcgc cgcctcatat tggtacgcca tcgccgcgct ggctgatgca 720
aatgcaagca tcgtgctgcc agggctgagg aagaacagcc tccaaggaga catcgccatc 780
atcagcatca tggagcattt tggagttcaa agttcatttg agagcgacgg cctccacctc 840
aacaagaagg tgattggaag tgatgtgagc ctcttcaact tcaaggagtg cccagatctt 900
gctcaaacag tggtggtggt ggcggcggcg ctgaagagag atgtgagctt caccggcctg 960
gagacgctga agatcaagga gaccgacagg atcgccgcgc tacagaagga gatcgccaag 1020
ttcggcgccg agctgattga agatggagac acctaccacc tcaagacggc gcaggtgtac 1080
cagccagaag aagtcacctt cgacacctat gaagatcaca ggatggcaat ggccttcgcg 1140
ccgctagctc tggtgttcga ccagatcaag attgctgaac ctcaagtggt ggagaagagc 1200
tacccggact tctggaacca cctccaagct caagccttcg tcatcgagta g 1251

<210> 7
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (克隆 GRG31pm3a11)

<400> 7
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30

[0012]

Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Gly Asn Ala Gly Thr
65 70 75 80
Ser Val Arg Phe Leu Thr Gly Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser

[0013]

385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 8
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 GRG31pm3a6）

<400> 8
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Gly Asn Ala Gly Thr
65 70 75 80
Ser Met Arg Phe Leu Thr Ala Val Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly

[0014]

245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 9
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 GRG31pm3a8）

<400> 9
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Gly Pro Ala Gly Thr
65 70 75 80
Gly Val Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu

[0015]

100 105 110

Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125

Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140

Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160

Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190

Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205

Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220

Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240

Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255

Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270

Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285

Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300

Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335

Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350

His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365

Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380

Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400

Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 10
<211> 416
<212> PRT
<213> 人工序列

[0016]

<220>

<223> GRG31 变体 (克隆 GRG31pm3b8)

<400> 10

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

[0017]

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 11
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (克隆 GRG31pm3b9)

<400> 11
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Asp Ala Ala Gly Thr
65 70 75 80
Ser Met Arg Phe Leu Ala Ala Tyr Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

[0018]

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

<210> 12
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 GRG31pm3b3）

<400> 12
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30

[0019]

Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Asp Ala Ala Gly Thr
65 70 75 80
Ala Thr Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser

[0020]

385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 13
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 L3P2b11）

<400> 13
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Gln Ala Gly Thr
65 70 75 80
Ser Leu Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly

[0021]

245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 14
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 L3P2c11）

<400> 14
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu

[0022]

100 105 110

Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125

Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140

Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160

Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190

Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205

Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220

Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240

Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255

Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270

Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285

Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300

Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335

Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350

His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365

Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380

Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400

Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 15
<211> 416
<212> PRT
<213> 人工序列

[0023]

<220>

<223> GRG31 变体（克隆 L3P2d11）

<400> 15

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

[0024]

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 16
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 4e11）

<400> 16
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Cys Gly Pro Ser Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala His Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

[0025]

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

<210> 17
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 L5e3）

<400> 17
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30

[0026]

Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Thr Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Tyr Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser

[0027]

385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 18
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 6b2）

<400> 18
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Tyr Leu Asn Leu Val Lys Gly Ser Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Val Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Glu Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly

[0028]

245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 19
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体（克隆 6e7）

<400> 19
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Tyr Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu

[0029]

100 105 110

Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125

Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140

Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160

Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190

Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205

Ile Ala Pro Gln Ala Phe Lys Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220

Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240

Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255

Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270

Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285

Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300

Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335

Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350

His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365

Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380

Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400

Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 20
<211> 416
<212> PRT
<213> 人工序列

[0030]

<220>

<223> GRG31 变体（三联 EK）

<400> 20

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

[0031]

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 21
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (L3P2c11+6e7)

<400> 21
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

[0032]

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

<210> 22
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (L3P2b11+6e)

<400> 22
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30

[0033]

Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Gln Ala Gly Thr
65 70 75 80
Ser Leu Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Asn Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Ala Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Lys Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser

[0034]

385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 23
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (L3P2c11+三联 EK)

<400> 23
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Pro Ala Gly Thr
65 70 75 80
Ala Met Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Ser Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175
Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190
Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Val Ile Lys
195 200 205
Ile Ala Pro Gln Ala Phe Lys Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220
Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240
Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly

[0035]

245 250 255
Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270
Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285
Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300
Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320
Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 24
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (L3P2b11+三联 EK)

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

[0036]

100 105 110

Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125

Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140

Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160

Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

Glu Gly Glu Leu Thr Ser Arg Pro Tyr Val Ser Met Thr Leu Asp Met
180 185 190

Leu Lys Ser Val Gly Ile Gln His Glu Trp Lys Asn Asn Val Ile Lys
195 200 205

Ile Ala Pro Gln Ala Phe Lys Lys Gln Thr Ile Tyr Val Glu Pro Asp
210 215 220

Trp Ser Ala Ala Ser Tyr Trp Tyr Ala Ile Ala Ala Leu Ala Asp Ala
225 230 235 240

Asn Ala Ser Ile Val Leu Pro Gly Leu Arg Lys Asn Ser Leu Gln Gly
245 250 255

Asp Ile Ala Ile Ile Ser Ile Met Glu His Phe Gly Val Gln Ser Ser
260 265 270

Phe Glu Ser Asp Gly Leu His Leu Asn Lys Lys Val Ile Gly Ser Asp
275 280 285

Val Ser Leu Phe Asn Phe Lys Glu Cys Pro Asp Leu Ala Gln Thr Val
290 295 300

Val Val Val Ala Ala Ala Leu Lys Arg Asp Val Ser Phe Thr Gly Leu
305 310 315 320

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335

Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350

His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365

Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380

Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400

Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 25
<211> 416
<212> PRT
<213> 人工序列

[0037]

<220>

<223> GRG31 变体 (L3P2c11+6b2)

<400> 25

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

[0038]

Glu Thr Leu Lys Ile Lys Glu Thr Asp Arg Ile Ala Ala Leu Gln Lys
325 330 335
Glu Ile Ala Lys Phe Gly Ala Glu Leu Ile Glu Asp Gly Asp Thr Tyr
340 345 350
His Leu Lys Thr Ala Gln Val Tyr Gln Pro Glu Glu Val Thr Phe Asp
355 360 365
Thr Tyr Glu Asp His Arg Met Ala Met Ala Phe Ala Pro Leu Ala Leu
370 375 380
Val Phe Asp Gln Ile Lys Ile Ala Glu Pro Gln Val Val Glu Lys Ser
385 390 395 400
Tyr Pro Asp Phe Trp Asn His Leu Gln Ala Gln Ala Phe Val Ile Glu
405 410 415

<210> 26
<211> 416
<212> PRT
<213> 人工序列

<220>
<223> GRG31 变体 (L3P2b11+6b2)

<400> 26
Met Asn Gln Gln Val Ile Thr Leu Thr His Pro Ser Lys Lys Ile Gln
1 5 10 15
Gly Thr Val Gln Leu Thr Gly Ser Lys Ser Glu Ser Asn Arg Ala Leu
20 25 30
Ile Ile Gln Ser Leu Ser Lys Gly Gln Val Glu Ile Ala Asn Leu Ser
35 40 45
Glu Ala Ala Asp Thr Val Thr Leu Asn Arg Val Leu Gln Ile Ala Ser
50 55 60
Asp Pro Lys Pro Gly Phe Asn Thr Ile Asp Ile Gly Gln Ala Gly Thr
65 70 75 80
Ser Leu Arg Phe Leu Thr Ala Phe Leu Asn Leu Val Lys Gly Ser Phe
85 90 95
Ile Leu Thr Gly Thr Glu Arg Met Gln Gln Arg Pro Ile Gly Ile Leu
100 105 110
Val Asp Ala Met Lys Glu Ile Gly Ala Asp Ile His Tyr Asp Lys Lys
115 120 125
Val Gly Tyr Pro Pro Leu Lys Ile Glu Gly Gly Leu Phe Gln Glu Lys
130 135 140
Asp Arg Val Lys Ile Lys Gly Asn Ile Ser Ser Gln Tyr Ile Ser Ala
145 150 155 160
Leu Leu Leu Ile Ala Pro Ala Leu Lys Lys Gly Leu Thr Leu Glu Ile
165 170 175

[0039]

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

<210> 27
<211> 447
<212> PRT
<213> 人工序列

<220>
<223> GRG36 变体 (GRG36(ace7))

<400> 27
Met Lys Asn Gln Asn Phe Asp Ala Lys Ala Arg Ser Pro Trp Thr Pro
1 5 10 15
Leu Lys Gly Val Asn Lys Ile Ser Val Ser Pro Ser Lys Gly Arg Ile
20 25 30

[0040]

Asn Gly Thr Val Thr Ile Pro Gly Ser Lys Ser Leu Thr Asn Arg Ala
35 40 45
Leu Ile Ile Ser Ser Leu Ala Ser Gly Lys Ser Lys Val Gln Gly Ile
50 55 60
Leu Lys Ser Asp Asp Ser Phe Trp Cys Leu Asp Ser Leu Lys Lys Leu
65 70 75 80
Gly Val Asn Val Lys Ile Gln Gly Asp Ile Ala Phe Ile Glu Gly Asn
85 90 95
Gly Gly Lys Trp Glu Ser Gly Asp Leu Tyr Ile Gly Ala Ala Gly Thr
100 105 110
Ile Ala Arg Phe Leu Pro Gly Ala Leu Ala Val Ser Gly Thr Gly Ile
115 120 125
Trp Glu Leu Glu Ala Ser Lys Ser Met Ser Lys Arg Pro Ile Ser Pro
130 135 140
Gln Val Asp Ala Leu Lys Glu Leu Gly Ala Glu Ile Thr Tyr Leu Ser
145 150 155 160
Asp Gln Gly Tyr Tyr Pro Leu Leu Val Lys Gly Lys Gln Leu Asn Gly
165 170 175
Gly Glu Val Glu Leu Ser Gly Arg Ile Ser Ser Gln Phe Ile Ser Gly
180 185 190
Leu Leu Ile Ala Ser Pro Tyr Leu Asn Asp Pro Ile Lys Ile Asn Ile
195 200 205
Lys Asp His Ile Val Gln His Ser Tyr Val Leu Leu Thr Leu Glu Leu
210 215 220
Met Lys Lys Phe Gly Ala Lys Val Lys Tyr Asp Ser Ser Leu Lys Glu
225 230 235 240
Ile Val Val Tyr Pro Ser Lys Tyr Thr Pro Gln Asp Ile Asn Leu Glu
245 250 255
Ala Asp Val Ser Thr Ala Cys Tyr Phe Leu Ala Leu Ala Ala Val Thr
260 265 270
Asn Gly Lys Val Gln Ile Asp Asn Leu Thr Tyr Glu Thr Lys Gln Pro
275 280 285
Asp Ile Lys Met Val Asp Ile Leu Glu Arg Met Gly Cys Lys Val Thr
290 295 300
Arg Gly Ser Ser Phe Ile Glu Ile Glu Gly Val Ser Gln Leu Lys Gly
305 310 315 320
Gly Phe Glu Ile Ser Met Arg Glu Met Ser Asp Gln Val Leu Thr Leu
325 330 335
Ala Ala Ile Ala Pro Phe Ala Asp Glu Pro Ile Thr Ile Lys Asp Val
340 345 350
Glu His Ile Arg His His Glu Ser Asn Arg Ile Ser Val Leu Val Asp
355 360 365
Ser Leu Ser Arg Leu Gly Ile Ile Val Glu Glu Phe Lys Asp Gly Leu
370 375 380
Lys Val Tyr Pro Gly Asn Pro Lys Ala Thr Leu Leu Asp Thr His Asp

[0041]

385 390 395 400
Asp His Arg Val Ala Met Ala Leu Ser Leu Ile Gly Ser Arg Val Glu
405 410 415
Gly Ile Gln Ile Asn Asp Pro Gly Cys Val Ser Lys Thr Cys Pro Gln
420 425 430
Tyr Phe Glu Leu Leu Glu Lys Leu Gly Leu Asn Ile Ile Lys His
435 440 445

<210> 28
<211> 447
<212> PRT
<213> 人工序列

<220>
<223> GRG36 变体 (GRG36(ace8))

<400> 28
Met Lys Asn Gln Asn Phe Asp Ala Lys Ala Arg Ser Pro Trp Thr Pro
1 5 10 15
Leu Lys Gly Val Asn Lys Ile Ser Val Ser Pro Ser Lys Gly Arg Ile
20 25 30
Asn Gly Thr Val Thr Ile Pro Gly Ser Lys Ser Leu Thr Asn Arg Ala
35 40 45
Leu Ile Ile Ser Ser Leu Ala Ser Gly Lys Ser Lys Val Gln Gly Ile
50 55 60
Leu Lys Ser Asp Asp Ser Phe Trp Cys Leu Asp Ser Leu Lys Lys Leu
65 70 75 80
Gly Val Asn Val Lys Ile Gln Gly Asp Thr Ala Phe Ile Glu Gly Asn
85 90 95
Gly Gly Lys Trp Glu Ser Gly Asp Leu Tyr Ile Gly Ala Ala Gly Thr
100 105 110
Ile Ala Arg Phe Leu Pro Gly Ala Leu Ala Val Ser Gly Thr Gly Ile
115 120 125
Trp Glu Leu Glu Ala Ser Lys Ser Met Ser Lys Arg Pro Ile Ser Pro
130 135 140
Leu Val Asp Ala Leu Lys Glu Leu Gly Ala Glu Ile Thr Tyr Leu Ser
145 150 155 160
Asp Gln Gly Tyr Tyr Pro Leu Leu Val Lys Gly Lys Gln Leu Asn Gly
165 170 175
Gly Glu Val Glu Leu Ser Gly Arg Ile Ser Ser Gln Phe Ile Ser Gly
180 185 190
Leu Leu Ile Ala Ser Pro Tyr Leu Asn Asp Pro Ile Lys Ile Asn Ile
195 200 205
Lys Asp His Ile Val Gln His Ser Tyr Val Leu Leu Thr Leu Glu Leu

[0042]

210 215 220
Met Lys Lys Phe Gly Ala Lys Val Lys Tyr Asp Ser Ser Leu Lys Glu
225 230 235 240
Ile Val Val Tyr Pro Ser Lys Tyr Thr Pro Gln Asp Ile Asn Leu Glu
245 250 255
Ala Asp Val Ser Thr Ala Cys Tyr Phe Leu Ala Leu Ala Ala Val Thr
260 265 270
Asn Gly Lys Val Gln Ile Asp Asn Leu Thr Tyr Glu Thr Lys Gln Pro
275 280 285
Asp Ile Lys Met Val Asp Ile Leu Glu Arg Met Gly Cys Lys Val Thr
290 295 300
Arg Gly Ser Ser Phe Ile Glu Ile Glu Gly Val Ser Gln Leu Lys Gly
305 310 315 320
Gly Phe Glu Ile Ser Met Arg Glu Met Ser Asp Gln Val Leu Thr Leu
325 330 335
Ala Ala Ile Ala Pro Phe Ala Asp Glu Pro Ile Thr Ile Thr Asp Val
340 345 350
Glu His Ile Arg His His Glu Ser Asn Arg Ile Ser Val Leu Val Asp
355 360 365
Ser Leu Ser Arg Leu Gly Ile Ile Val Glu Glu Phe Lys Asp Gly Leu
370 375 380
Lys Val Tyr Pro Gly Asn Pro Lys Ala Thr Leu Leu Asp Thr His Asp
385 390 395 400
Asp His Arg Val Ala Met Ala Leu Ser Met Ile Gly Ser Arg Val Glu
405 410 415
Gly Ile Gln Ile Asn Asp Pro Gly Cys Val Ser Lys Thr Cys Pro Gln
420 425 430
Tyr Phe Glu Leu Leu Glu Lys Leu Gly Leu Asn Ile Ile Lys His
435 440 445

<210> 29
<211> 1377
<212> DNA
<213> 人工序列

<220>
<223> 编码 GRG36 变体的合成的序列
(GRG36(ace7))

<400> 29
atggcacatc accaccacca tcacggatcc ggcatgaaga accagaactt cgacgccaag 60
gcaagaagcc cctggacgcc gctcaagggc gtcaacaaga tctccgtctc gccaagcaaa 120
ggaaggatca atggcaccgt caccatccct ggaagcaaga gcttgacaaa tagagctctc 180
atcatcagca gcctagcaag cggcaagagc aaggttcaag gcatcctcaa gagcgacgac 240

[0043]

agcttctggt gcctggattc attgaagaag ctcggcgtca atgtgaagat ccaaggagac 300
atcgccttca ttgaaggaaa tggaggaaaa tgggagagtg gagatctcta catcggcgcc 360
gccggcacca ttgcaagatt tcttcctggc gcgctggcgg tgagcggcac aggcatctgg 420
gagctggagg caagcaagag catgagcaag aggcccatct caccgcaggt ggatgctctc 480
aaggagctcg gcgccgagat cacctacctc tctgatcaag gctactaccc gctgctggtg 540
aagggcaagc agctcaatgg aggagaggtg gagctctctg gaaggatcag cagccagttc 600
atcagcggcc tgctgattgc ttcaccatat ttgaatgatc ccatcaagat caacatcaag 660
gaccacatcg tgcagcacag ctatgtgctg ctgacgctgg agctgatgaa gaagttcggc 720
gccaaggtga agtacgacag cagcttgaag gagatcgtcg tctacccaag caagtacacg 780
ccgcaggaca tcaacctgga agctgatgtt tcaacagcat gctacttcct ggcgctggcg 840
gcggtgacaa atggtaaggt gcaaattgac aacctcacct acgagaccaa gcagccggac 900
atcaagatgg tggacatcct ggagagaatg ggctgcaagg tgacaagagg aagcagcttc 960
atcgagattg aaggagtgag ccagctgaag ggcggcttcg agatctcaat gagggagatg 1020
agcgaccagg tgctgacgct ggcggccatc gcgccatttg ctgatgagcc catcaccatc 1080
aaggatgtgg agcacatccg ccatcatgaa agcaacagga tctccgtgct ggtggacagc 1140
ctctcaaggc tgggcatcat cgtggaggag ttcaaggatg gcctcaaggt gtaccctgga 1200
aatccaaagg cgacgctgct ggacacccat gatgatcacc gcgtggcaat ggcgctgagc 1260
ttgattggat caagggtgga aggcatccag atcaatgatc ctggctgcgt cagcaagacc 1320
tgccctcaat attttgagct gctggagaag ctgggcctca acatcatcaa gcactag 1377

<210> 30
<211> 1377
<212> DNA
<213> 人工序列

<220>
<223> 编码 GRG36 变体的合成的序列
(GRG36(ace8))

<400> 30
atggcacatc accaccacca tcacggatcc ggcatgaaga accagaactt cgacgccaag 60
gcaagaagcc cctggacgcc gctcaagggc gtcaacaaga tctccgtctc gccaagcaaa 120
ggaaggatca atggcaccgt caccatcccc ggaagcaaga gcttgacaaa tagagctctc 180
atcatcagca gcttagcaag cggcaagagc aaggttcaag gcatcctcaa gagcgacgac 240
agcttctggt gcctggattc attgaagaag ctcggcgtca atgtgaagat ccaaggagac 300
accgccttca ttgaaggaaa tggaggaaaa tgggagagtg gagatctcta catcggcgcc 360
gccggcacca ttgcaagatt tcttcctggc gcgctggcgg tgagcggcac cggcatctgg 420
gagctggagg caagcaagag catgagcaag aggcccatct caccgctggt ggatgctctc 480
aaggagctcg gcgccgagat cacctacctc tctgatcaag gctactaccc gctgctggtg 540
aagggcaagc agctcaatgg aggagaggtg gagctctctg gaaggatcag cagccagttc 600
atcagcggcc tgctgattgc ttcaccatat ttgaatgatc ccatcaagat caacatcaag 660
gaccacatcg tgcagcacag ctatgtgctg ctgacgctgg agctgatgaa gaagttcggc 720
gccaaggtga agtacgacag cagcttgaag gagatcgtcg tctacccaag caagtacacg 780
ccgcaggaca tcaacctgga agctgatgtt tcaacagcat gctacttcct ggcgctggcg 840
gcggtgacaa atggcaaggt gcaaattgac aacctcacct acgagaccaa gcagccggac 900

[0044]

atcaagatgg tggacatcct ggagaggatg ggctgcaagg tgacaagagg aagcagcttc 960
atcgagattg aaggagtgag ccagctgaag ggcggcttcg agatctcaat gagggagatg 1020
agcgaccagg tgctgacgct ggcggccatc gcgccatttg ctgatgagcc catcaccatc 1080
acggatgtgg agcacatccg ccatcatgaa agcaacagga tctccgtgct ggtggacagc 1140
ctctcaaggc tgggcatcat cgtggaggag ttcaaggacg gcctcaaggt gtaccctgga 1200
aatccaaagg cgacgctgct ggacacccat gatgatcacc gcgtggcaat ggcgctgagc 1260
atgattggat caagggtgga aggcatccag atcaatgatc ctggctgcgt cagcaagacc 1320
tgccctcaat attttgagct gctggagaag ctgggcctca acatcatcaa gcactag 1377