

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷



[12] 发明专利申请公开说明书

[21] 申请号 03808897.5

A01H 5/00

A01H 5/10

C12N 1/21

C12N 5/04

C12N 15/31

C12N 15/62

C12N 15/82

[43] 公开日 2005 年 7 月 27 日

[11] 公开号 CN 1646006A

[22] 申请日 2003.2.20 [21] 申请号 03808897.5

[30] 优先权

[32] 2002. 3. 6 [33] US [31] 60/362,250

[86] 国际申请 PCT/US2003/004735 2003.2.20

[87] 国际公布 WO2003/075655 英 2003.9.18

[85] 进入国家阶段日期 2004.10.20

[71] 申请人 辛根塔参与股份公司

地址 瑞士巴塞尔

[72] 发明人 沈志诚 G·W·瓦伦

F·舒特科斯克 V·卡拉米尔

[74] 专利代理机构 中国国际贸易促进委员会专利商
标事务所

代理人 罗菊华

C07K 14/32 C07K 19/00

权利要求书 12 页 说明书 53 页 序列表 42 页

[54] 发明名称 新的 VIP3 毒素和应用的方法

[57] 摘要

本发明公开了具有高度抗广谱鳞翅目昆虫害虫活性的新的 Vip3 毒素。编码 Vip3 毒素的 DNA 可以用于转化各种原核生物和真核生物的生物体以表达 Vip3 毒素。这些重组的生物体可以用于控制各种环境中的鳞翅目昆虫。

I S S N 1 0 0 8 - 4 2 7 4

1、一种分离的核酸分子，其包含编码具有抗昆虫活性的毒素的核苷酸序列，其中所述核苷酸序列：

(a) 具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50℃，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行的，并在 65℃，0.1XSSC，0.1% SDS 中漂洗；或者

(b) 与 (a) 的核苷酸序列同类编码；或

(c) 与 SEQ ID NO:1 具有至少 93% 序列同一性；或

(d) 编码与 SEQ ID NO:2 具有至少 91% 序列同一性的氨基酸序列。

2、如权利要求 1 所述的分离的核酸分子，其中所述核苷酸序列具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50℃，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行的，并在 65℃，0.1XSSC，0.1% SDS 中漂洗。

3、如权利要求 1 所述的分离的核酸分子，其中所述核苷酸序列和具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列的核苷酸序列同类编码，该杂交是在 50℃，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行的，并在 65℃，0.1XSSC，0.1% SDS 中漂洗。

4、如权利要求 1 所述的分离的核酸分子，其包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 75% 序列同一性的核苷酸序列。

5、如权利要求 1 所述的分离的核酸分子，其包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 85% 序列同一性的核苷酸序列。

6、如权利要求 1 所述的分离的核酸分子，其包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 95% 序列同一性的核苷酸序列。

7、如权利要求 1 所述的分离的核酸分子，其包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 99% 序列同一性的核苷酸序列。

8、如权利要求 1 所述的分离的核酸分子，其包含 SEQ ID NO:1 或 SEQ ID NO:3 核苷酸 1981-2367。

9、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列与SEQ ID NO:1具有至少93%序列同一性。

10、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列与SEQ ID NO:1具有至少95%序列同一性。

11、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列与SEQ ID NO:1具有至少99%序列同一性。

12、如权利要求1所述的分离的核酸分子,其包含如SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 11, SEQ ID NO : 32 或 SEQ ID NO: 34 中所述的核苷酸序列。

13、如权利要求1所述的分离的核酸分子,其包含SEQ ID NO: 1的核苷酸序列。

14、如权利要求1所述的分离的核酸分子,其包含SEQ ID NO: 3的核苷酸序列。

15、如权利要求1所述的分离的核酸分子,其包含SEQ ID NO: 11的核苷酸序列。

16、如权利要求1所述的分离的核酸分子,其包含SEQ ID NO: 32的核苷酸序列。

17、如权利要求1所述的分离的核酸分子,其包含SEQ ID NO: 34的核苷酸序列。

18、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列编码与SEQ ID NO: 2具有至少91%同一性的氨基酸序列。

19、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列编码与SEQ ID NO: 2具有至少95%同一性的氨基酸序列。

20、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列编码与SEQ ID NO: 2具有至少99%同一性的氨基酸序列。

21、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列编码SEQ ID NO: 2, SEQ ID NO : 12 或 SEQ ID NO: 33 中所述氨基酸序列。

22、如权利要求1所述的分离的核酸分子,其中所述核苷酸序列编

码 SEQ ID NO: 2 中所述氨基酸序列。

23、如权利要求 1 所述的分离的核酸分子,其中所述核苷酸序列编码 SEQ ID NO: 12 中所述氨基酸序列。

24、如权利要求 1 所述的分离的核酸分子,其中所述核苷酸序列编码 SEQ ID NO: 33 中所述氨基酸序列。

25、如权利要求 1 所述的分离的核酸分子,其包含在苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离物中,该分离物选自命名为 NRRL 保藏号 B-30556 的 C1674; 和命名为 NRRL 保藏号 B-30557 的 C536。

26、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在大肠杆菌(*E.coli*)克隆中,该大肠杆菌(*E.coli*)克隆选自命名为 NRRL 保藏号 B-30553 的 pNOV3910; 命名为 NRRL 保藏号 B-30552 的 pNOV3911; 命名为 NRRL 保藏号 B-30555 的 pNOV3906; 命名为 NRRL 保藏号 B-30554 的 pNOV3905; 和命名为 NRRL 保藏号 B-30551 的 pNOV3912。

27、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在命名为 NRRL 保藏号 B-30553 的大肠杆菌 (*E.coli*) 克隆 pNOV3910 中。

28、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在命名为 NRRL 保藏号 B-30552 的大肠杆菌 (*E.coli*) 克隆 pNOV3911 中。

29、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在命名为 NRRL 保藏号 B-30555 的大肠杆菌 (*E.coli*) 克隆 pNOV3906 中。

30、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在命名为 NRRL 保藏号 B-30554 的大肠杆菌 (*E.coli*) 克隆 pNOV3905 中。

31、如权利要求 1 所述的分离的核酸分子,其包含约 2.4kb 的 DNA 片段,该 DNA 片段包含在命名为 NRRL 保藏号 B-30551 的大肠杆菌 (*E.coli*) 克隆 pNOV3912 中。

32、如权利要求 1 所述的分离的核酸分子，其中所述毒素包含与 SEQ ID NO: 2 氨基酸序列的氨基酸 681-788 具有至少 75% 同一性的氨基酸序列。

33、如权利要求 1 所述的分离的核酸分子，其中所述毒素包含与 SEQ ID NO: 2 氨基酸序列的氨基酸 681-788 具有至少 85% 同一性的氨基酸序列。

34、如权利要求 1 所述的分离的核酸分子，其中所述毒素包含与 SEQ ID NO: 2 氨基酸序列的氨基酸 681-788 具有至少 95% 同一性的氨基酸序列。

35、如权利要求 1 所述的分离的核酸分子，其中所述毒素包含与 SEQ ID NO: 2 氨基酸序列的氨基酸 681-788 具有至少 99% 同一性的氨基酸序列。

36、如权利要求 1 所述的分离的核酸分子，其中所述毒素包含 SEQ ID NO: 2 氨基酸序列的氨基酸 681-788。

37、如权利要求 1 所述的分离的核酸分子，其中所述毒素具有抗鳞翅目昆虫的活性。

38、如权利要求 37 所述的分离的核酸分子，其中所述鳞翅目昆虫选自：欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫(*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾(*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

39、一种包含可操作地连接权利要求 1 所述核酸分子的异源启动子序列的嵌合基因。

40、一种包含权利要求 39 所述嵌合基因的重组载体。

41、一种包含权利要求 39 所述嵌合基因的转基因宿主细胞。

42、如权利要求 41 所述的转基因宿主细胞，其是细菌细胞。

43、如权利要求 41 所述的转基因宿主细胞，其是植物细胞。

44、一种包含权利要求 43 所述转基因植物细胞的转基因植物。

45、如权利要求 44 所述的转基因植物，其中所述植物选自高粱，小麦，向日葵，番茄，菜类作物，棉花，稻，大豆，甜菜，甘蔗，烟草，大麦，油籽油菜和玉米。

46、如权利要求 45 所述的转基因植物，其中所述植物是玉米植物。

47、来源于权利要求 44 所述转基因植物的转基因种子。

48、来源于权利要求 46 所述玉米植物的转基因种子。

49、一种具有抗昆虫活性的分离的毒素，其中所述毒素包含氨基酸序列，该氨基酸序列

a) 与 SEQ ID NO: 2 氨基酸 661-788 具有至少 75% 同一性；或

b) 与 SEQ ID NO: 2 具有至少 91% 同一性；或

c) 是通过含有核苷酸序列的核酸分子表达产生的，该核苷酸序列具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行的，并在 65°C，0.1XSSC，0.1% SDS 中漂洗；或

d) 是通过含有核苷酸序列的核酸分子的表达产生的，该核苷酸序列与 (c) 的核苷酸序列同类编码；或

e) 是通过含有核苷酸序列的核酸分子的表达产生的，该核苷酸序列与 SEQ ID NO:1 具有至少 93% 序列同一性。

50、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 75% 同一性的氨基酸序列。

51、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 85% 同一性的氨基酸序列。

52、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 95% 同一性的氨基酸序列。

53、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 99% 同一性的氨基酸序

列。

54、如权利要求 49 所述的分离的毒素，其中所述毒素包含 SEQ ID NO:2 氨基酸序列的氨基酸 661-788。

55、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 91% 同一性的氨基酸序列。

56、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 95% 同一性的氨基酸序列。

57、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 97% 同一性的氨基酸序列。

58、如权利要求 49 所述的分离的毒素，其中所述毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 99% 同一性的氨基酸序列。

59、如权利要求 49 所述的分离的毒素，其中所述毒素包含 SEQ ID NO: 2, SEQ ID NO: 12 或 SEQ ID NO: 33 中所述氨基酸序列。

60、如权利要求 59 所述的分离的毒素，其中所述毒素包含 SEQ ID NO: 2 中所述氨基酸序列。

61、如权利要求 59 所述的分离的毒素，其中所述毒素包含 SEQ ID NO: 12 中所述氨基酸序列。

62、如权利要求 59 所述的分离的毒素，其中所述毒素包含 SEQ ID NO: 33 中所述氨基酸序列。

63、如权利要求 49 所述的分离的毒素，其中所述毒素是通过含有核苷酸序列的核酸分子表达产生的，该核苷酸序列具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行的，并在 65°C，0.1XSSC，0.1% SDS 中漂洗。

64、如权利要求 49 所述的分离的毒素，其中所述毒素是通过含有核苷酸序列的核酸分子表达产生的，该核苷酸序列和具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列的核苷酸序列同类编码，该杂交是在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行，并在 65°C，0.1XSSC，0.1% SDS 中漂洗。

65、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 75% 同一性。

66、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 85% 同一性。

67、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 95% 同一性。

68、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 99% 同一性。

69、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO:1 或 SEQ ID NO:3 核苷酸 1981-2367 的核酸分子表达产生的。

70、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 具有至少 93% 序列同一性。

71、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 具有至少 95% 序列同一性。

72、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含核苷酸序列的核酸分子表达产生的，该核苷酸序列与 SEQ ID NO:1 具有至少 99% 序列同一性。

73、如权利要求 49 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 11, SEQ ID NO: 32 或 SEQ ID NO: 34 中所述核苷酸的核酸分子表达产生的。

74、如权利要求 73 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 1 的核酸分子表达产生的。

75、如权利要求 73 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 3 的核酸分子表达产生的。

76、如权利要求 73 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 11 的核酸分子表达产生的。

77、如权利要求 73 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 32 的核酸分子表达产生的。

78、如权利要求 73 所述的分离的毒素，其中所述毒素是通过包含 SEQ ID NO: 34 的核酸分子表达产生的。

79、如权利要求 49 所述的分离的毒素，其中所述毒素具有抗鳞翅目昆虫的活性。

80、如权利要求 79 所述的分离的毒素，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

81、如权利要求 49 所述的分离的毒素，其中所述毒素是由选自命名为 NRRL 保藏号 B-30556 的 C1674 和命名为 NRRL 保藏号 B-30557 的 C536 的苏云金芽孢杆菌 (*Bacillus thuringiensis*) 株系产生的。

82、如权利要求 81 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30556 的苏云金芽孢杆菌 (*Bacillus thuringiensis*) 株系 C1674 产生的。

83、如权利要求 81 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30557 的苏云金芽孢杆菌 (*Bacillus thuringiensis*) 株系 C536 产生的。

84、如权利要求 49 所述的分离的毒素，其中所述毒素是由选自命名为 NRRL 保藏号 B-30553 的 pNOV3910；命名为 NRRL 保藏号 B-30552 的 pNOV3911；命名为 NRRL 保藏号 B-30555 的 pNOV3906；命名为 NRRL 保藏号 B-30554 的 pNOV3905；和命名为 NRRL 保藏号

B-30551 的 pNOV3912 的大肠杆菌 (*E.coli*) 克隆产生的。

85、如权利要求 84 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30553 的大肠杆菌 (*E.coli*) 克隆产生的。

86、如权利要求 84 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30552 的大肠杆菌 (*E.coli*) 克隆产生的。

87、如权利要求 84 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30555 的大肠杆菌 (*E.coli*) 克隆产生的。

88、如权利要求 84 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30554 的大肠杆菌 (*E.coli*) 克隆产生的。

89、如权利要求 84 所述的分离的毒素，其中所述毒素是由命名为 NRRL 保藏号 B-30551 的大肠杆菌 (*E.coli*) 克隆产生的。

90、一种含有有效控制昆虫量的权利要求 49 中所述毒素的组合物。

91、一种产生具有抗昆虫活性的毒素的方法，该方法包括：

(a) 获得如权利要求 41 所述的转基因宿主细胞；

(b) 在允许产生毒素的条件下培养所述转基因宿主细胞；和

(c) 回收所述毒素。

92、一种产生抗昆虫转基因植物的方法，包括将权利要求 1 所述的核酸分子引入植物细胞；和从所述植物细胞再生转化的植物，其中所述转化的植物具有昆虫抗性。

93、如权利要求 92 所述的方法，其中所述昆虫是鳞翅目昆虫。

94、如权利要求 93 所述的方法，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾 (*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

95、一种控制昆虫的方法，包括给所述昆虫递送有效量的权利要求 49 所述的毒素。

96、如权利要求 95 所述的方法，其中所述昆虫是鳞翅目昆虫。

97、如权利要求 96 所述的方法，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫(*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾(*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

98、如权利要求 95 所述的方法，其中所述毒素通过口服递送给昆虫。

99、一种保护玉米植物免受至少一种昆虫害虫侵害的方法，包括：在玉米细胞中引入权利要求 1 中所述核酸分子；和再生稳定转化的玉米植物，其中转化的玉米植物产生足够保护玉米植物免受至少一种昆虫害虫侵害量的杀虫毒素。

100、一种具有抗昆虫活性的杂合毒素，其中所述毒素由包含权利要求 1 所述核苷酸序列的核酸分子编码。

101、如权利要求 100 所述的杂合毒素，其中所述昆虫是鳞翅目昆虫。

102、如权利要求 101 所述的杂合毒素，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫(*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾(*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

103、如权利要求 100 所述的杂合毒素，其中所述杂合毒素是由 SEQ ID NO: 10 中所述核苷酸序列编码的。

104、一种包含杀虫有效量权利要求 100 所述杂合毒素的组合物。

105、一种产生具有抗昆虫活性的杂合毒素的方法，包括：

- a) 获得如权利要求 41 所述的转基因宿主细胞；
- b) 在允许产生杂合毒素的条件下培养所述转基因宿主细胞；和
- c) 回收所述杂合毒素。

106、一种产生抗昆虫转基因植物的方法，包括将权利要求 1 所述的核酸分子引入植物细胞；和从所述植物细胞再生转化的植物，其中所述转化植物产生杂合毒素并且具有昆虫抗性。

107、如权利要求 106 所述的方法，其中所述昆虫是鳞翅目昆虫。

108、如权利要求 107 所述的方法，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫(*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾(*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

109、一种控制昆虫的方法，包括给所述昆虫递送有效量权利要求 100 所述的杂合毒素。

110、如权利要求 109 所述的方法，其中所述昆虫是鳞翅目昆虫。

111、如权利要求 110 所述的方法，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫(*Spodoptera frugiperda*)，小地老虎(*Agrotis ipsilon*)，谷实夜蛾(*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

112、如权利要求 109 所述的方法，其中所述杂合毒素通过口服递送给昆虫。

113、一种具有抗昆虫活性的杂合毒素，其包含 Vip3 毒素羧基末端区域和按氨基到羧基的方向与之连接的不同 Vip3 毒素的氨基末端区域，其中所述羧基末端区域包含与 SEQ ID NO: 2 氨基酸 661-788 具有至少 75% 同一性的氨基酸序列；和其中所述氨基末端区域与 SEQ ID NO: 6 氨基酸 1-660 有至少 85% 同一性。

114、如权利要求 113 所述的杂合毒素，其中所述羧基末端区域包含 SEQ ID NO: 2 氨基酸 661-788，且所述氨基末端区域包含 SEQ ID NO: 6 氨基酸 1-660。

115、如权利要求 113 所述的杂合毒素，其中所述杂合毒素包含 SEQ ID NO: 12 氨基酸 1-788。

116、一种控制昆虫的方法，包括给所述昆虫递送有效量权利要求 113 所述的杂合毒素。

117、如权利要求 116 所述的方法，其中所述昆虫是鳞翅目昆虫。

118、如权利要求 117 所述的方法，其中所述鳞翅目昆虫选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾(*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾(*Heliothis virescens*)，贪夜蛾(*Spodoptera exigua*)，棉红铃虫(*Pectinophora gossypiella*)，粉纹夜蛾(*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

119、一种包含编码权利要求 113 所述杂合毒素的核苷酸序列的核酸分子。

120、一种包含核苷酸序列的分离的 *vip3* 核酸分子，所述核苷酸序列

- a) 包含 SEQ ID NO : 8; 或
- b) 编码 SEQ ID NO : 9 中所述氨基酸序列。

新的 VIP3 毒素和应用的方法

技术领域

本发明涉及来源于苏云金芽孢杆菌 (*Bacillus thuringiensis*) 的新的 Vip3 毒素, 其表达产生所述毒素的核酸序列, 和制备该毒素的方法, 及利用该毒素和相应的核酸序列控制昆虫的方法。

背景技术

植物害虫是世界重要农作物损失的主要因素。在美国, 每年单由非哺乳动物害虫包括昆虫的侵染就损失约 80 亿美元。除了农田作物损失外, 昆虫害虫对于蔬菜和水果种植者, 对于观赏花卉的生产者, 和对于家宅园丁也是负担。

主要通过化学杀虫剂的密集应用控制昆虫害虫, 该化学杀虫剂可抑制昆虫生长, 妨碍昆虫进食或繁殖, 或引起死亡。因此可以达到良好的昆虫控制, 但是这些化学药品有时也影响其它有益昆虫。化学杀虫剂广泛应用引起的另一问题是抗药性昆虫品种的出现。通过各种抗药性处理实践已经部分缓解了这种现象, 但是对可替代害虫控制剂存在日益增加的需要。生物害虫控制剂, 如表达杀虫毒素象 δ -内毒素的苏云金芽孢杆菌 (*Bacillus thuringiensis*) 株系也已经施用于农作物植物, 并且具有满意的结果, 提供了化学杀虫剂的可替代物或补充物。已经分离了一些编码这些 δ -内毒素的基因, 表明在异源宿主中它们的表达提供了用于经济重要昆虫害虫控制的另一种工具。特别地, 转基因植物中杀虫毒素如苏云金芽孢杆菌 (*Bacillus thuringiensis*) δ -内毒素的表达提供了免受所针对的昆虫害虫侵害的有效保护, 已经表达这种毒素的转基因植物已经商品化了, 使得农民能够减少化学昆虫控制剂的施用。

现在已经鉴定了其它非内毒素基因和它们编码的蛋白质。专利 5,877,012, 6,107,279, 6,137,033 和 6,291,156, 及 Estruch 等(1996, Proc.

Natl. Acad. Sci. 93: 5389-5394)和 Yu 等(1997, Appl. Environ. Microbiol. 63: 532-536)描述了称为 Vip3 的新一类杀虫蛋白质, 这里并入所有这些文献为参考文献。Vip3 基因编码序列编码芽孢杆菌 (*Bacillus*) 营养生长阶段期间产生和分泌的约 88kDa 蛋白质 (营养杀虫蛋白质, VIP)。Vip3A 蛋白质具有抗广谱鳞翅目害虫, 包括但不限于小地老虎 (BCW, *Agrotis ipsilon*), 草地粘虫 (FAW, *Spodoptera frugiperda*), 烟芽夜蛾 (TBW, *Heliothis virescens*), 和谷实夜蛾 (CEW, *Helicoverpa zea*) 的杀虫活性。最近, 已经发现表达 Vip3A 蛋白质的植物抗半翅目昆虫害虫引起的进食损害。因此 Vip3A 蛋白质显示了独特的杀虫活性谱。其它的公开文献 WO 98/18932, WO 98/33991, WO 98/00546, 和 WO 99/57282 现在也已经鉴定了 Vip3 类蛋白质的同系物。

控制昆虫害虫的化学和生物药剂的连续应用增加了昆虫发展对这种控制方法抗药性的机会。而且, 每种控制剂仅仅可以控制几种特定的昆虫害虫。

因此, 需要开发出能为农民提供经济利益, 并且是环境可接受的新的和有效的害虫控制剂。特别需要的是提供靶向广谱经济重要昆虫害虫的控制剂和有效控制昆虫株系的控制剂, 所述昆虫株系对现存昆虫控制剂是抗药性的或能变得有抗药性, 以及与目前控制剂比较具有增加的效力的控制剂。而且, 期望药剂的施用最小化对环境的负担。

发明内容

本发明通过提供新的基因和毒素致力于满足对新的害虫控制剂的需要, 所述新的基因和毒素不同于美国专利 5,877,012, 6,107,279, 和 6,137,033 和 Estruch 等(1996), 和 Yu 等(1997)及 WO 98/18932, WO 98/33991, WO 99/57282, 和 WO 98/00546 中所公开的基因和毒素。

本发明提供了控制植物害虫的组合物和方法。特别地, 提供了从苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离的新 vip3 核酸序列, 基本与其相同的序列, 所述序列的表达产生了对经济上重要昆虫害虫, 特别是侵染植物的昆虫害虫具有毒性的杀虫毒素。本发明进一步涉及由该核

酸序列表达产生的新的杀虫毒素，含有该杀虫毒素的组合物和制剂，它们能抑制昆虫害虫存活，生长和繁殖的能力，或能限制与昆虫有关的作物植物损害或损失。本发明也涉及例如在制造具有增加杀虫活性的杂合毒素中或在重组基因方法如 DNA 改组中利用该核酸序列的方法。本发明进一步涉及制造毒素的方法和涉及例如在微生物中利用该核酸序列控制昆虫的方法或在转基因植物中利用该核酸序列提供免受昆虫损害保护的方法，并涉及利用杀虫毒素和包含该杀虫毒素的组合物和制剂的方法，例如向昆虫侵染的区域施用杀虫毒素或组合物或制剂或预防性处理昆虫易侵染的区域或植物以提供免受昆虫害虫侵害的保护。

可以利用本领域通常公知的方法改造本发明的核苷酸序列以为了各种目的改变核苷酸序列，所述目的包括但不限于扩大杀虫活性谱或增加抗特异性害虫的特异性活性。通过随机片段化和基因片段和合成的寡核苷酸的 PCR 重装配的 DNA 改组可以用于基因工程改造核苷酸序列。

这里所述新的杀虫毒素是高度抗昆虫活性的。例如，通过该杀虫毒素可以控制大量经济上重要昆虫害虫，如鳞翅目欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，西南玉米秆草螟 (*Diatraea grandiosella*)，小蔗秆草螟 (*Diatraea saccharalis*)，细点突夜蛾 (*Helicoverpa punctigera*) 和棉铃虫 (*Helicoverpa armigera*)。可以单独使用或与其它昆虫控制策略联合使用杀虫毒素以提供具有最小环境影响的最大害虫控制效力。

根据一个方面，本发明提供了包含编码毒素的核苷酸序列的分离的核酸分子，所述毒素具有抗昆虫活性，其中所述核苷酸序列：(a) 具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行，同时在 65°C，0.1XSSC，0.1% SDS 中漂洗；或者 (b) 与 (a) 的核苷酸序列同类编码；或 (c) 包含连续 20 个碱基对的核苷酸部分，其序列与 (a) 或 (b) 核苷酸序列的连续 20 个碱基对核苷酸部分相同；或 (d)

与 SEQ ID NO:1 具有至少 93% 序列同一性; 或(e) 编码与 SEQ ID NO:2 具有至少 91% 序列同一性的氨基酸序列。

在这方面的一个实施方式中, 分离的核酸分子包含具有互补序列的核苷酸序列, 该互补序列与 SEQ ID NO:1 核苷酸 1981-2367 杂交, 该杂交是在 50°C, 7% 十二烷基硫酸钠 (SDS), 0.5M NaPO₄, 1mM EDTA 中进行, 同时在 65°C, 0.1XSSC, 0.1% SDS 中漂洗。

在该方面的另一个实施方式中, 分离的核酸分子包含与具有互补序列的核苷酸序列同类编码的核苷酸序列, 该互补序列与 SEQ ID NO:1 核苷酸 1981-2367 杂交, 该杂交是在 50°C, 7% 十二烷基硫酸钠 (SDS), 0.5M NaPO₄, 1mM EDTA 中进行, 同时在 65°C, 0.1XSSC, 0.1% SDS 中漂洗。

在另一个实施方式中, 分离的核酸分子包含连续 20 个碱基对的核苷酸部分, 其序列与 SEQ ID NO:1 或 SEQ ID NO:3 中所述核苷酸序列核苷酸 1981-2367 的连续 20 个碱基对核苷酸部分相同。

在另一个实施方式中, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 75% 序列同一性的核苷酸序列。优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 85% 序列同一性的核苷酸序列。更优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 95% 序列同一性的核苷酸序列。甚至更优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 99% 序列同一性的核苷酸序列。最优选地, 分离的核酸分子包含 SEQ ID NO:1 或 SEQ ID NO:3 的核苷酸 1981-2367。

在另一个实施方式中, 分离的核酸分子包含与 SEQ ID NO:1 具有至少 93% 序列同一性的核苷酸序列。优选地, 分离的核酸分子包含与 SEQ ID NO:1 具有至少 95% 序列同一性的核苷酸序列。更优选地, 分离的核酸分子包含与 SEQ ID NO:1 具有至少 99% 序列同一性的核苷酸序列。最优选地, 分离的核酸分子包含选自 SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:10, SEQ ID NO:31 和 SEQ ID NO:33 的核苷酸序列的核苷酸 1-2367。

在本发明一个实施方式中，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 75% 同一性。优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 85% 同一性。更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 95% 同一性。甚至更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 99% 同一性。最优选地，分离的核酸分子编码包含 SEQ ID NO:2 的氨基酸 661-788 的毒素。

在另一个实施方式中，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 91% 同一性。优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 95% 同一性。更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 99% 同一性。最优选地，分离的核酸分子编码包含 SEQ ID NO:2 或 SEQ ID NO:12 中所述氨基酸序列的毒素。

在一个实施方式中，分离的核酸分子包含在苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离物中，该分离物选自命名为 NRRL 保藏号 B-30556 的 C1674；和命名为 NRRL 保藏号 B-30557 的 C536。

在另一个实施方式中，分离的核酸分子包含约 2.4kb 的 DNA 片段，该 DNA 片段包含在大肠杆菌 (*E.coli*) 克隆中，该大肠杆菌 (*E.coli*) 克隆选自命名为 NRRL 保藏号 B-30553 的 pNOV3910；命名为 NRRL 保藏号 B-30552 的 pNOV3911；命名为 NRRL 保藏号 B-30555 的 pNOV3906；命名为 NRRL 保藏号 B-30554 的 pNOV3905；和命名为 NRRL 保藏号 B-30551 的 pNOV3912。

根据本发明一个实施方式，分离的核酸分子编码具有抗鳞翅目昆虫活性的毒素。优选地，根据该实施方式，毒素具有抗欧洲玉米螟 (*Ostrinia*

nubilalis), 小菜蛾(*Plutella xylostella*), 草地粘虫(*Spodoptera frugiperda*), 小地老虎(*Agrotis ipsilon*), 谷实夜蛾(*Helicoverpa zea*), 烟芽夜蛾(*Heliothis virescens*), 贪夜蛾(*Spodoptera exigua*), 棉红铃虫(*Pectinophora gossypiella*), 粉纹夜蛾(*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟(*Homoeosoma electellum*) 的活性。

本发明也提供了嵌合基因, 其包含与本发明核酸分子可操作地连接的异源启动子序列。进一步, 本发明提供了包含这种嵌合基因的重组载体。本发明也提供了包含这种嵌合基因的转基因宿主细胞。本发明该方面的转基因宿主细胞可以是动物细胞、动物病毒、植物病毒、细菌细胞、酵母细胞或植物细胞, 优选植物细胞。进一步, 本发明提供了包含这种植物细胞的转基因植物。本发明该方面的转基因植物可以是高粱, 小麦, 向日葵, 番茄, 菜类作物, 棉花, 稻, 大豆, 甜菜, 甘蔗, 烟草, 大麦, 油籽油菜或玉米, 优选玉米和棉花。进一步, 本发明提供了来自转基因植物的种子, 所述转基因植物包括: 高粱, 小麦, 向日葵, 番茄, 菜类作物, 棉花, 稻, 大豆, 甜菜, 甘蔗, 烟草, 大麦, 油籽油菜和玉米。在特别优选的实施方式中, 种子来自于转基因玉米植物或转基因棉花植物。

本发明也提供了进一步包含编码第二杀虫成分的第二核酸序列或核酸序列组的转基因植物。特别优选的第二核酸序列是编码 δ -内毒素的核酸序列, 编码另外营养杀虫蛋白质毒素的核酸序列或编码非蛋白质杀虫成分产生途径的核酸序列。

在另一方面, 本发明提供了通过本发明核酸分子表达产生的毒素。

在优选的实施方式中, 通过包含核苷酸序列核苷酸 1-2367 的核酸分子表达产生该毒素, 所述核苷酸序列选自 SEQ ID NO:1, SEQ ID NO:3 和 SEQ ID NO:10。

在另一个实施方式中, 本发明毒素具有抗鳞翅目昆虫活性, 优选地抗欧洲玉米螟(*Ostrinia nubilalis*), 小菜蛾(*Plutella xylostella*), 草地粘虫(*Spodoptera frugiperda*), 小地老虎(*Agrotis ipsilon*), 谷实夜蛾(*Helicoverpa zea*), 烟芽夜蛾(*Heliothis virescens*), 贪夜蛾(*Spodoptera*

exigua), 棉红铃虫(*Pectinophora gossypiella*), 粉纹夜蛾(*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*) 。

在一个实施方式中, 本发明毒素通过苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离物产生, 该分离物选自命名为 NRRL 保藏号 B-30556 的 C1674; 和命名为 NRRL 保藏号 B-30557 的 C536。

在另一个实施方式中, 毒素通过大肠杆菌 (*E.coli*) 克隆产生, 该大肠杆菌(*E.coli*)克隆选自命名为 NRRL 保藏号 B-30553 的 pNOV3910; 命名为 NRRL 保藏号 B-30552 的 pNOV3911; 命名为 NRRL 保藏号 B-30555 的 pNOV3906; 命名为 NRRL 保藏号 B-30554 的 pNOV3905; 命名为 NRRL 保藏号 B-30551 的 pNOV3912。

在一个实施方式中, 本发明毒素含有氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 75% 同一性。优选地, 该毒素含有氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 85% 同一性。更优选地, 毒素含有氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 95% 同一性。甚至更优选地, 毒素含有氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 99% 同一性。最优选地, 毒素含有 SEQ ID NO:2 的氨基酸 661-788。

在另一个实施方式中, 本发明毒素包含氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 91% 同一性。优选地, 毒素包含氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 95% 同一性。更优选地, 毒素包含氨基酸序列, 所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 99% 同一性。最优选地, 毒素包含 SEQ ID NO:2, SEQ ID NO:11 或 SEQ ID NO:32 中所述氨基酸序列。

本发明也提供了包含有效昆虫控制量的本发明毒素的组合物。

在另一方面, 本发明提供了产生有抗昆虫活性的毒素的方法, 该方法包括 (a) 获得包含嵌合基因的转基因宿主细胞, 所述嵌合基因本身包含与本发明核酸分子可操作地连接的异源启动子序列; 和 (b) 在转

基因细胞中表达核酸分子，这产生了至少一种有抗昆虫活性的毒素。

在进一步方面，本发明提供了生产抗昆虫转基因植物的方法，包括向转基因植物中引入本发明的核酸分子，其中核酸分子可以在转基因植物中以控制昆虫的有效量表达。根据一个实施方式，昆虫是鳞翅目昆虫，优选地选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾 (*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

在进一步方面，本发明提供了控制昆虫的方法，包括递送给昆虫有效量的本发明毒素。根据一个实施方式，昆虫是鳞翅目昆虫，优选地选自欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾 (*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。优选地，口服给昆虫递送毒素。在一个优选的实施方式中，通过转基因植物来口服递送毒素，所述转基因植物包含表达本发明毒素的核酸序列。

本发明也提供了具有抗昆虫活性的杂合毒素，其中由包含本发明核苷酸序列的核酸分子编码杂合毒素。

在一个实施方式中，本发明的杂合毒素具有抗鳞翅目昆虫活性，优选地抗欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾 (*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

在另一个实施方式中，由包含在命名为 NRRL 保藏号 B-30551 的 pNOV3912 大肠杆菌 (*E.coli*) 克隆中的约 2.4kb DNA 片段编码杂合毒素。在优选的实施方式中，由 SEQ ID NO: 10 中所述核苷酸序列编码杂

合毒素。

本发明也提供了包含本发明杀虫有效量杂合毒素的组合物。

在另一方面，本发明提供了生产具有抗昆虫活性的杂合毒素的方法，包括（a）获得包含嵌合基因的转基因宿主细胞，所述嵌合基因本身包含与本发明核酸分子可操作地连接的异源启动子序列；和（b）在转基因细胞中表达此核酸分子，这产生了至少一种有抗昆虫活性的杂合毒素。

在另一方面，本发明提供了生产抗昆虫转基因植物的方法，包括向植物中引入本发明的核酸分子，其中核酸分子编码杂合毒素和其中杂合毒素可以在转基因植物中以控制昆虫的有效量表达。根据一个实施方式，昆虫是鳞翅目昆虫，优选地选自欧洲玉米螟（*Ostrinia nubilalis*），小菜蛾（*Plutella xylostella*），草地粘虫（*Spodoptera frugiperda*），小地老虎（*Agrotis ipsilon*），谷实夜蛾（*Helicoverpa zea*），烟芽夜蛾（*Heliothis virescens*），贪夜蛾（*Spodoptera exigua*），棉红铃虫（*Pectinophora gossypiella*），粉纹夜蛾（*Trichoplusia ni*），*Cochyles hospes* 和向日葵同斑螟（*Homoeosoma electellum*）。

在另一个方面，本发明提供了控制昆虫的方法，包括给昆虫递送有效量的本发明杂合毒素。根据一个实施方式，昆虫是鳞翅目昆虫，优选地选自欧洲玉米螟（*Ostrinia nubilalis*），小菜蛾（*Plutella xylostella*），草地粘虫（*Spodoptera frugiperda*），小地老虎（*Agrotis ipsilon*），谷实夜蛾（*Helicoverpa zea*），烟芽夜蛾（*Heliothis virescens*），贪夜蛾（*Spodoptera exigua*），棉红铃虫（*Pectinophora gossypiella*），粉纹夜蛾（*Trichoplusia ni*），*Cochyles hospes* 和向日葵同斑螟（*Homoeosoma electellum*）。优选地，口服给昆虫递送杂合毒素。在一个优选的实施方式中，通过转基因植物口服递送杂合毒素，所述转基因植物包含表达本发明杂合毒素的核酸序列。

本发明也提供了具有抗昆虫活性的杂合毒素，其包含按氨基到羧基的方向连接到不同 Vip3 毒素的氨基末端区域的 Vip3 毒素羧基末端区域，其中所述羧基末端区域包含与 SEQ ID NO: 2 氨基酸 661-788 有至

少 75% 同一性, 优选至少 85% 同一性, 更优选至少 95% 同一性, 最优选至少 99% 同一性的氨基酸序列; 和其中所述氨基末端区域与 SEQ ID NO: 5 氨基酸 1-660 有至少 75% 同一性, 优选至少 85% 同一性, 更优选至少 95% 同一性, 最优选至少 99% 同一性。在优选的实施方式中, 该羧基末端区域包含 SEQ ID NO: 2 氨基酸 661-788, 并且该氨基末端区域包含 SEQ ID NO: 5 氨基酸 1-660。在更优选的实施方式中, 杂合毒素包含 SEQ ID NO: 11 氨基酸 1-788。

根据本发明该方面, 优选地, 杂合毒素具有抗鳞翅目昆虫活性, 更优选地抗选自欧洲玉米螟 (*Ostrinia nubilalis*), 小菜蛾 (*Plutella xylostella*), 草地粘虫 (*Spodoptera frugiperda*), 小地老虎 (*Agrotis ipsilon*), 谷实夜蛾 (*Helicoverpa zea*), 烟芽夜蛾 (*Heliothis virescens*), 贪夜蛾 (*Spodoptera exigua*), 棉红铃虫 (*Pectinophora gossypiella*), 粉纹夜蛾 (*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*) 的鳞翅目昆虫。

本发明该方面还包括的是包含核苷酸序列的核酸分子, 所述核苷酸序列编码该方面的杂合毒素。

本发明也提供了控制昆虫的方法, 其中包含本发明的杂合毒素的转基因植物进一步包含编码第二杀虫成分的第二核酸序列或核酸序列组。特别优选的第二核酸序列是编码 δ -内毒素的核酸序列, 编码另外营养杀虫蛋白质毒素的核酸序列或编码非蛋白质杀虫成分产生途径的核酸序列。

而且, 本发明另一方面提供了诱变本发明核酸分子的方法, 其中核酸分子已经被切割为期望大小的双链随机片段群, 所述方法包括: (a) 向双链随机片段群加入一种或多种单链或双链寡核苷酸, 其中每种寡核苷酸包含与双链模板多核苷酸同一的区域和异源的区域; (b) 变性由此得到的双链随机片段和寡核苷酸的混合物为单链片段; (c) 在引起单链片段在同一性区域退火以形成成对的退火片段的条件下, 温育由此得到的单链片段群和聚合酶, 同一性区域足以使所述对的一个成员引发另一个成员的复制, 由此形成了诱变的双链多核苷酸; 和 (d) 重复第

二和第三步骤至少两轮进一步的循环，其中在下一轮循环第二步骤中得到的混合物包含从前一轮循环第三步骤得到的诱变的双链多核苷酸，且其中所述下一轮循环形成了进一步诱变的双链多核苷酸。

通过本发明下面的描述和非限制性实施例的学习，本发明的其它方面和优点对于本领域技术人员将变得显而易见。

序列表中序列的简要说明

SEQ ID NO: 1 是天然 *vip3C* 核苷酸序列。

SEQ ID NO: 2 是 SEQ ID NO: 1 编码的氨基酸序列。

SEQ ID NO: 3 是玉米优化的 *vip3C* 核苷酸序列。

SEQ ID NO: 4 是天然 *vip3A (a)* 核苷酸序列。

SEQ ID NO: 5 是 SEQ ID NO: 5 编码的氨基酸序列。

SEQ ID NO: 6 是天然 *vip3B* 核苷酸序列。

SEQ ID NO: 7 是 SEQ ID NO: 7 编码的氨基酸序列。

SEQ ID NO: 8 是天然 *vip3Z* 核苷酸序列。

SEQ ID NO: 9 是 SEQ ID NO: 9 编码的氨基酸序列。

SEQ ID NO: 10 是杂合 *vip3A-C* 核苷酸序列。

SEQ ID NO: 11 是 SEQ ID NO: 11 编码的氨基酸序列。

SEQ ID NO: 12-29 是在本发明实践中有用的引物序列。

SEQ ID NO: 30 是载体 pNOV2149 的核苷酸序列。

SEQ ID NO: 31 是 *vip3C-12168* 核苷酸序列。

SEQ ID NO: 32 是 SEQ ID NO: 32 编码的氨基酸序列。

SEQ ID NO: 33 是玉米优化的 *vip3C-12168* 核苷酸序列。

保藏

按照关于国际承认用于专利程序目的微生物保藏的布达佩斯条约，下列材料保藏在农业研究机构保藏中心 (NRRL)，1815 North University Street, Peoria, Illinois 61604。一旦授予专利权，将不可撤销地解除对保藏材料获得的所有限制。

<u>分离物/克隆</u>	<u>保藏号</u>	<u>保藏日期</u>
苏云金芽孢杆菌 (B. t.) 株系 C1674	NRRL B-30556	2002年2月7日
苏云金芽孢杆菌 (B. t.) 株系 C536	NRRL B-30557	2002年2月7日
大肠杆菌 (E.coli) BL21 (pNOV3906)	NRRL B-30555	2002年2月7日
大肠杆菌 (E.coli) BL21 (pNOV3905)	NRRL B-30554	2002年2月7日
大肠杆菌 (E.coli) DH5 α (pNOV3910)	NRRL B-30553	2002年2月7日
大肠杆菌 (E.coli) DH5 α (pNOV3911)	NRRL B-30552	2002年2月7日
大肠杆菌 (E.coli) DH5 α (pNOV3912)	NRRL B-30551	2002年2月7日

定义

本发明毒素的“活性”意指毒素起口服活性昆虫控制剂的作用，有毒性作用，或能破坏或阻止昆虫进食，这可以或可以不引起昆虫死亡。当递送给昆虫本发明毒素时，结果通常是昆虫死亡，或昆虫不以使昆虫可得到毒素的资源为食。

“相关”/“可操作地连接”指两个物理或功能相关的核酸序列。例如，如果启动子或调节 DNA 序列和编码 RNA 或蛋白质的 DNA 序列可操作地连接或定位以至于调节 DNA 序列将影响编码或结构 DNA 序列的表达水平，那么称启动子或调节 DNA 序列与编码 RNA 或蛋白质的 DNA 序列“相关”。

“嵌合基因”是重组核酸序列，其中启动子或调节核酸序列可操作地连接编码 mRNA 或作为蛋白质表达的核酸序列，或与编码 mRNA 或作为蛋白质表达的核酸序列相关，使得调节核酸序列能调节相关核酸序列的转录或表达。嵌合基因的调节核酸序列不是如自然界中所发现的正常可操作地连接相关核酸序列。

“编码序列”是转录为 RNA 如 mRNA, rRNA, tRNA, snRNA, 有义 RNA 或反义 RNA 的核酸序列。优选地，随后在生物体中翻译 RNA 以产生蛋白质。

“控制”昆虫意指通过毒性作用抑制昆虫害虫存活，生长，进食和/或繁殖的能力，或限制作物植物中昆虫相关的损害或损失。尽管“控制”

昆虫优选意指杀死昆虫，但是“控制”昆虫也可以是或不是指杀死昆虫。

对应于：在本发明上下文中“对应于”意指当不同 Vip3 基因或蛋白质的核酸编码序列或氨基酸序列互相对比时，“对应于”一些计数位置的核酸或氨基酸是与这些位置对比，但不必是在相对于特定 Vip3 各自核酸编码序列或氨基酸序列的这些精确数字位置中的核酸或氨基酸。同样，当特定 Vip3（例如，Vip3Z）的编码或氨基酸序列与参照 Vip3（例如，Vip3C）的编码或氨基酸序列对比时，“对应于”Vip3C 序列一些计数位置的 Vip3Z 序列中核酸或氨基酸是与 Vip3C 序列的这些位置对比，但不必是在 Vip3Z 蛋白质各自核酸编码序列或氨基酸序列的这些精确数字位置中的核酸或氨基酸。

“递送”毒素意指毒素与昆虫接触，引起有毒作用和昆虫的控制。可以用许多公认的方法，例如通过昆虫摄食口服或通过转基因植物表达，制成的蛋白质组合物，可喷射的蛋白质组合物，食饵基质与昆虫接触，或任何其它本领域公认的毒素递送系统递送毒素。

“有效的昆虫控制量”意指通过毒性作用抑制昆虫存活，生长，进食和/或繁殖的能力或限制作物植物中昆虫相关的损害或损失的毒素浓度。尽管“有效的昆虫控制量”优选意指杀死昆虫，但是“有效的昆虫控制量”也可以是或不是指杀死昆虫。

这里所用的“表达盒”意指能指导适合宿主细胞中特定核苷酸序列表达的核酸序列，包含与目的核苷酸序列可操作地连接的启动子，所述目的核苷酸序列可操作地连接终止信号。通常，它也包含核苷酸序列正确翻译所需的序列。包含目的核苷酸序列的表达盒可以是嵌合的，意指至少其成分之一相对于至少它的其它成分之一是异源的。表达盒也可以是天然存在的，但以重组形式获得用于异源表达的表达盒。然而，通常，表达盒相对于宿主是异源的，即，表达盒的特定核酸序列不天然出现在宿主细胞中，必须通过转化事件将其引入宿主细胞或宿主细胞的前体。表达盒中核苷酸序列的表达可以受组成型启动子或诱导型启动子控制，其中仅当宿主细胞暴露于一些特定外部刺激时，所述诱导型启动子才起始转录。如果是多细胞生物体的情况，如植物，启动子也可以是对特定

组织，或器官或发育阶段特异的。

“基因”是位于基因组内的限定区域，除了前述编码核酸序列外，包含其它主要是调节性的核酸序列，所述调节性核酸序列负责编码部分的表达，即转录和翻译控制。基因也可以包含其它 5'和 3'非翻译序列和终止序列。进一步可以存在的元件是，例如内含子。

“目的基因”指任何基因，当转移其到植物中时，其赋予了植物期望的特性如抗生素抗性，病毒抗性，昆虫抗性，疾病抗性或对其它害虫的抗性，除草剂耐受性，改进的营养价值，工业生产过程中改进的性能或改变的繁殖能力。“目的基因”也可以是转移到植物中用于植物中商业价值的酶或代谢物生产的目的基因。

“异源”核酸序列是与其被引入的宿主细胞不天然相关的核酸序列，包含非天然存在的天然存在核酸序列的多拷贝。

“同源”核酸序列是与其被引入的宿主细胞天然相关的核酸序列。

“同源重组”是同源核酸分子间核酸片段的相互交换。

这里所用的“杂合毒素”是人工制造的杀虫毒素，其包含与来源于不同毒素氨基酸区域或片段连接的一种毒素的氨基酸区域或片段。例如，非限制性的，连接来源于 SEQ ID NO:2 氨基酸 661-788 的 Vip3C C-末端区域和来源于 SEQ ID NO:4 氨基酸 1-660 的 Vip3A N-末端区域，产生了具有 SEQ ID NO:11 中所述氨基酸序列的杂合毒素。

“杀虫的”定义为能控制昆虫，优选杀死它们的毒性生物活性。

当核酸序列编码与参照核酸序列编码的多肽有相同氨基酸序列的多肽时，该核酸序列与参照核酸序列是“同类编码”。

“分离的”核酸分子或分离的蛋白质或毒素是人工地与其天然环境分离而存在，因此不是天然产物的核酸分子或蛋白质或毒素。分离的核酸分子或蛋白质或毒素可以以纯化形式存在，或者可以存在于非天然环境如，例如重组宿主细胞或转基因植物中。

天然：指在未转化细胞的基因组中存在的基因。

天然存在：术语“天然存在”用于描述可以在自然界中发现的客体，其与人工产生的客体不同。例如，可以从自然来源分离，并没有在实验

室有意进行人工修饰的、生物体（包括病毒）中存在的蛋白质或核苷酸序列是“天然存在”的。

“核酸分子”或“核酸序列”是可以从任何来源分离的单或双链 DNA 或 RNA 的线性片段。在本发明上下文中，优选地，核酸分子是 DNA 片段。

“植物”是在任何发育阶段的任何植物，特别是种子植物。

“植物细胞”是植物的结构和生理学单位，包含原生质体和细胞壁。植物细胞可以是分离的单个细胞或培养细胞形式，或作为高等有组织的单位如，例如植物组织，植物器官或整个植物的一部分。

“植物细胞培养物”意指各种发育阶段的植物单位如，例如原生质体，细胞培养细胞，植物组织中的细胞，花粉，花粉管，胚珠，胚囊，合子和胚的培养物。

“植物材料”指叶，茎，根，花或花的部分，果实，花粉，卵细胞，合子，种子，插条，细胞或组织培养物，或植物的任何其它部分或产物。

“植物器官”是植物清楚的和明显结构化和分化的部分，如根，茎，叶，花蕾或胚。

这里所用的“植物组织”意指组织成结构和功能单位的一组植物细胞。包括植物中或培养物中植物的任何组织。该术语包括但不限于整个植物，植物器官，植物种子，组织培养物和组织成结构和/或功能单位的任何植物细胞组。该术语与以上列举的或该定义包含的任何特定类型植物组织的联合应用或单独应用不意味排除任何其它类型植物组织。

“启动子”是编码区域上游非翻译的 DNA 序列，其包含 RNA 聚合酶 11 的结合位点，并起始 DNA 的转录。启动子区域也可以包含作为基因表达调节物的其它元件。

“原生质体”是没有细胞壁或仅有部分细胞壁的分离的植物细胞。

“调节元件”指参与控制核苷酸序列表达的序列。调节元件包含可操作地连接目的核苷酸序列的启动子和终止信号。通常它们也包含核苷酸序列正确翻译所需的序列。

“改组的”核酸是通过改组方法，如这里所述的任何改组方法产生的

核酸。通过人工的和可选地循环的方式（物理地或实际上）重组两个或多个核酸（或字符串）产生改组核酸。一般地，在改组方法中利用一步或多步筛选步骤以鉴定目的核酸；可以在任何重组步骤前或后进行该筛选步骤。在一些（但不是所有）改组实施方式中，期望在筛选前进行多轮重组以增加待筛选库的多样性。可选地，可以循环重复重组和筛选的全部过程。根据上下文，改组可以指重组和筛选的全部过程，或可替代地，可以仅指全部过程的重组部分。

基本相同：在两个核酸或蛋白质序列上下文中的短语“基本相同”指当比较和比对以获得最大对应时，如利用下面序列比较算法之一或目测所测定的，具有至少 60%，优选 80%，更优选 90%，甚至更优选 95% 和最优选至少 99% 核苷酸或氨基酸残基同一性的两个或多个序列或亚序列。优选地，基本同一性存在于至少约 50 个残基长度的序列区域，更优选至少约 100 个残基的区域上，最优选地，在至少约 150 残基中的序列基本相同。在特别优选的实施方式中，在编码区整个长度中序列基本相同。而且，基本相同的核酸或蛋白质序列具有基本相同的功能。

为了进行序列比较，通常，一个序列作为参照序列而与检测序列比较。当利用序列比较算法时，将检测和参照序列输入到计算机中，如果必要的话指定亚序列的坐标，并指定序列算法程序的参数。然后，根据选定的程序参数，序列比较算法将计算出检测序列相对于参照序列的百分序列同一性。

例如，通过 Smith & Waterman, *Adv. Appl. Math.* 2: 482 (1981) 的局部同源性算法，通过 Needleman & Wunsch, *J. Mol. Biol.* 48: 443 (1970) 的同源性比对算法，通过 Pearson & Lipman, *Proc. Nat'l. Acad. Sci. USA* 85: 2444 (1988) 的相似性检索方法，通过这些算法的计算机化实施 (Wisconsin Genetics 软件包中 GAP, BESTFIT, FASTA 和 TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI) 或通过目测 (通常参见, Ausubel 等, 下文) 可以进行用于比较的序列的最佳比对。

适于测定百分序列同一性和序列相似性的一个算法例子是 BLAST 算法，在 Altschul 等, *J. Mol. Biol.* 215: 403-410 (1990) 中描述了该算法。

通过国家生物技术信息中心([http://www. Ncbi.nlm. nih. gov/](http://www.Ncbi.nlm.nih.gov/))公众可得到进行 BLAST 分析的软件。该算法包括：通过鉴定出查寻序列中长度为 W 的短字而首先鉴定出高分值序列对 (HSPs)，所述短字在与数据库序列中相同长度的字比对时匹配或满足一些正值阈值记分 T 。 T 称为相邻字记分阈值 (Altschul 等, 1990)。这些最初的邻近字命中作为开始查寻的线索去发现包含它们的更长的 HSPs。然后，这些字命中将沿每个序列的两个方向尽可能远的延伸，直到积累比对分值不再增加。对于核苷酸序列，用参数 M (成对匹配残基的奖励分值;总是大于零) 和 N (错配残基的罚分值;总是小于零) 计算积累分值。对于氨基酸序列，用记分矩阵计算积累分值。当积累比对分值从获得的最大值回落数量 X ，由于一个或更多负分值残基比对积累，积累分值达到或低于零，或两个序列的任一个到达终点时，每个方向的字命中延伸停止。BLAST 算法的参数 W , T 和 X 决定了比对的敏感性和速度。BLASTN 程序 (对于核苷酸序列) 使用字长值 (W) 11, 期望值 (E) 10, 截断值 100, $M=5$, $N=-4$ 和两个链的比较为缺省值。对于氨基酸序列，BLASTP 程序使用字长值 (W) 3, 期望值 (E) 10 和 BLOSUM62 记分矩阵(参见, Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA* 89 : 10915 (1989))为缺省值。

除了计算百分序列同一性外，BLAST 算法也进行两个序列间相似性的统计学分析 (参见, 例如 Karlin & Altschul, *Proc. Nat'l. Acad. Sci. USA* 90: 5873-5787 (1993))。BLAST 算法提供的一个相似性测定是最小和概率 ($P(N)$)，其提供了两个核苷酸或氨基酸序列间偶然出现匹配的概率的指示。例如，如果检测核酸序列与参照核酸序列比较的最小和概率少于约 0.1，更优选少于约 0.01，最优选少于约 0.001，那么认为检测核酸序列与参照序列相似。

两个核酸序列基本相同的另一个指标是两个分子在严格条件下互相杂交。短语“特异性杂交”指当该序列存在于复杂混合物 (例如, 总细胞的) DNA 或 RNA 中时, 在严格条件下, 分子仅与特定核苷酸序列结合, 形成双螺旋或杂交。“基本结合”指探针核酸和靶核酸间互补杂交,

并且包含较少的错配，通过降低杂交介质的严格性可耐受所述的错配，以实现靶核酸序列的期望检测。

在核酸杂交试验如 Southern 和 Northern 杂交上下文中“严格杂交条件”和“严格杂交漂洗条件”是序列依赖性的，并且在不同环境参数下是不同的。较长的序列在较高温度特异性杂交。在 Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, 第 I 部分第 2 章 "Overview of principles of hybridization and the strategy of nucleic acid probe assays" Elsevier, New York 中可以发现核酸杂交的大量指南。一般地，对于在限定离子强度和 pH 下的特定序列，高严格性杂交和漂洗条件选择为低于热熔点 (T_m) 约 5°C 。典型地，在“严格条件”下，探针将与其靶亚序列杂交，而不与其它序列杂交。

T_m 是 (在限定离子强度和 pH 条件下) 50% 靶序列与完全匹配的探针杂交时的温度。对于特定的探针，非常严格的条件选择为等于 T_m 。在 Southern 或 Northern 印迹中在滤膜上有多于 100 个互补残基的互补核酸杂交的一个严格杂交条件的例子是在 42°C ，具有 1mg 肝素的 50% 甲酰胺，过夜进行该杂交。高严格性漂洗条件的例子是 72°C ，0.15M NaCl 约 15 分钟。严格漂洗条件的例子是在 65°C ，0.2 x SSC 漂洗 15 分钟 (参见, Sambrook, 下文, SSC 缓冲液 的描述)。通常，在高严格性漂洗前进行低严格性漂洗以除去背景探针信号。对于例如多于 100 个核苷酸的双螺旋而言，中严格性漂洗的例子是 45°C ，1 x SSC 漂洗 15 分钟。对于例如多于 100 个核苷酸的双螺旋而言，低严格性漂洗的例子是 40°C ，4-6 x SSC 漂洗 15 分钟。对于短探针 (例如，约 10 到 50 个核苷酸)，严格条件通常包括在 pH7.0 到 8.3 的少于约 1.0M Na 离子的盐浓度，通常约 0.01 到 1.0Na 离子浓度 (或其它盐)，通常温度至少是约 30°C 。通过添加去稳定剂如甲酰胺也可以获得严格条件。一般地，在特定杂交测定中，信噪音比就无关探针所观察到的值高 $2\times$ (或更高) 表明特异杂交的检测。在严格条件下不互相杂交的核酸如果它们编码的蛋白质是基本相同的，那么它们仍是基本相同的。例如，当用遗传密码所

允许的最大密码子简并性创造核酸拷贝时，就会出现这种情况。

下面是杂交/漂洗条件设置的例子，所述条件可以用于克隆与本发明参照核苷酸序列基本相同的同源核苷酸序列：参照核苷酸序列与参照核苷酸序列优选在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中杂交，在 50°C，2 X SSC，0.1% SDS 中漂洗，更期望在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中杂交，在 50°C，1 X SSC，0.1% SDS 中漂洗，更期望在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中杂交，在 50°C，0.5 X SSC，0.1% SDS 中漂洗，优选地，在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中杂交，在 50°C，0.1 X SSC，0.1% SDS 中漂洗，更优选地，在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中杂交，在 65°C，0.1 X SSC，0.1% SDS 中漂洗。

两个核酸序列或蛋白质基本相同的另一个指标是第一核酸编码的蛋白质与第二核酸编码的蛋白质免疫交叉反应或特异结合。因此，蛋白质通常与第二蛋白质基本相同，例如，其中两个蛋白质仅仅由于保守性置换而不同。

“合成的”指包含天然序列中不存在的结构特征的核苷酸序列。例如，称更密切地类似双子叶和/或单子叶植物基因 G+C 含量和正常密码子分布的人工序列是合成的。

“转化”是向宿主细胞或生物体中引入异源核酸的过程，特别地，“转化”意指 DNA 分子稳定整合进入目的生物体基因组中。

“转化的/转基因的/重组的”指已经引入异源核酸分子的宿主生物体，如细菌或植物。核酸分子可以稳定地整合进入宿主基因组或者核酸分子也可以作为染色体外分子存在。这种染色体外分子可以是自主复制的。转化的细胞，组织，或植物理解为不仅包含转化过程的最终产物，也包含其转基因子代。“非转化的”，“非转基因的”，或“非重组的”宿主指不含有异源核酸分子的野生型生物体，例如细菌或植物。

“Vip 类蛋白质”包含 Vip3A(a)，Vip3A(b)，Vip3A(c)，Vip3B，Vip3C(a)，Vip3C(b)，Vip3Z 和它们的同系物。这里所用的“同系物”意指

所述蛋白质或多肽与 Vip3 类蛋白质的其它成员具有确定的关系。这种确定的关系包括,但不限于: 1) 在序列水平上,与 Vip3 类蛋白质另一个成员有至少 70%,更优选至少 80%,最优选至少 90% 同一性,同时还保持了杀虫活性的蛋白质, 2) 与免疫识别 Vip3 类蛋白质另一个成员的抗体交叉反应的蛋白质, 3) 与 Vip3 类蛋白质另一个成员的受体交叉反应,并且保留了诱导程序性细胞死亡的能力的蛋白质,和 4) 在序列水平上,与 Vip3 类蛋白质另一个成员的毒性核心区域有至少 70%,更优选至少 80%,最优选至少 90% 同一性,同时还保持了杀虫活性的蛋白质。在 WO 98/18932, WO 98/33991, WO 98/00546 和 WO 99/57282 中已经公开了其它 Vip3 同系物。

用下面的标准缩写的碱基表示核苷酸: 腺嘌呤(A),胞嘧啶(C),胸腺嘧啶(T)和鸟嘌呤(G)。同样地,用下面标准缩写表示氨基酸: 丙氨酸(Ala; A),精氨酸(Arg; R),天冬酰胺(Asn; N),天冬氨酸(Asp; D),半胱氨酸(Cys; C),谷氨酰胺(Gln; Q),谷氨酸(Glu; E),甘氨酸(Gly; G),组氨酸(His; H),异亮氨酸(Ile; I),亮氨酸(Leu; L),赖氨酸(Lys; K),甲硫氨酸(Met; M),苯丙氨酸(Phe; F),脯氨酸(Pro; P),丝氨酸(Ser; S),苏氨酸(Thr; T),色氨酸(Trp; W),酪氨酸(Tyr; Y)和缬氨酸(Val; V)。

本发明的详细描述

本发明涉及核酸序列,其表达产生新的毒素,还涉及制造和应用该毒素控制昆虫害虫。核酸序列来源于芽孢杆菌属(*Bacillus*),一种革兰氏阳性形成孢子的微生物。特别地,提供了作为杀虫剂的新的 Vip3 蛋白质。

为了于本发明的目的,昆虫害虫包括选自例如鞘翅目,双翅目,膜翅目,鳞翅目,食毛目,同翅目,半翅目,直翅目,缨翅目,革翅目,等翅目,虱目,蚤目,毛翅目等,特别是鳞翅目的昆虫。

表 1-7 给出了与主要作物植物有关的害虫列表。这种害虫包含在本发明范围内。

表 1

鳞翅目 (Lepidoptera)	
欧洲玉米螟 (<i>Ostrinia nubilalis</i>)	贪夜蛾 (<i>Spodoptera exigua</i>)
小地老虎 (<i>Agrotis ipsilon</i>)	棉红铃虫 (<i>Pectinophora gossypiella</i>)
谷实夜蛾 (<i>Helicoverpa zea</i>)	稻白螟 (<i>Scirpophaga innotata</i>)
草地粘虫 (<i>Spodoptera frugiperda</i>)	稻纵卷叶野螟 (<i>Cnaphalocrocis medinalis</i>)
西南玉米秆草螟 (<i>Diatraea grandiosella</i>)	七星稻螟 (<i>Chilo plejadellus</i>)
南美玉米苗斑螟 (<i>Elasmopalpus lignosellus</i>)	三点水螟 (<i>Nymphula depunctalis</i>)
小蔗秆草螟 (<i>Diatraea saccharalis</i>)	斜纹贪夜蛾 (<i>Spodoptera litura</i>)
烟芽夜蛾 (<i>Heliothis virescens</i>)	灰翅贪夜蛾 (<i>Spodoptera mauritia</i>)
三化螟 (<i>Scirpophaga incertulas</i>)	向日葵细卷叶蛾 (<i>Cochylis hospes</i>)
轴禾草螟 (<i>Chilo polychrysa</i>)	<i>Pseudaletia unipunctata</i>
东方粘虫 (<i>Mythimna separata</i>)	西部灰地老虎 (<i>Agrotis orthogonia</i>)
斑禾草螟 (<i>Chilo partellus</i>)	<i>Pseudoplusia includens</i>
粒肤脏切夜蛾 (<i>Feltia subterranea</i>)	大豆夜蛾 (<i>Anticarsia gemmatalis</i>)
向日葵同斑螟 (<i>Homoeosoma electellum</i>)	苜蓿绿夜蛾 (<i>Plathypena scabra</i>)

表 2

鞘翅目 (Coleoptera)	
<i>Diabrotica virgifera</i>	蛴螬 (<i>Phyllophaga crinita</i>)
<i>Diabrotica longicornis</i>	金针虫 (<i>Melanotus spp.</i> , <i>Eleodes</i> , <i>Conoderus</i> , 和 <i>Aeolus spp.</i> ,)
黄瓜十一星叶甲 (<i>Diabrotica undecimpunctata</i>)	黑角负泥虫 (<i>Oulema melanopus</i>)
圆头犀金龟(蛴螬) (<i>Cyclocephala borealis</i>)	玉米跳甲 (<i>Chaetocnema pulicaria</i>)
圆头无斑犀金龟(蛴螬) (<i>Cyclocephala immaculata</i>)	黑角负泥虫 (<i>Oulema melanopus</i>)
日本弧丽金龟 (<i>Popillia japonica</i>)	三叶草叶象 (<i>Hypera punctata</i>)
玉米跳甲 (<i>Chaetocnema pulicaria</i>)	棉铃象 (<i>Anthonomus grandis</i>)
玉米尖隐喙象 (<i>Sphenophorus maidis</i>)	葡萄肖叶甲 (<i>Colaspis brunnea</i>)
	稻水象 (<i>Lissorhoptrus oryzophilus</i>)
	米象 (<i>Sitophilus oryzae</i>)
	墨西哥豆瓢虫 (<i>Epilachna varivestis</i>)

表 3

同翅目 (Homoptera)	
玉米缢管蚜 (<i>Rhopalosiphum maidis</i>)	棉花盲蝽象 (<i>Pseudatomoscelis seriatus</i>)
玉米根蚜 (<i>Anuraphis maidiradicis</i>)	结翅粉虱 (<i>Trialeurodes abutilonea</i>)
类甘蔗伪毛蚜 (<i>Sipha flava</i>)	二条黑尾叶蝉 (<i>Nephotettix nigropictus</i>)
麦二叉蚜 (<i>Schizaphis graminum</i>)	桃蚜 (<i>Myzus persicae</i>)
麦长管蚜 (<i>Macrosiphum avenae</i>)	马铃薯小叶蝉 (<i>Empoasca fabae</i>)
棉蚜 (<i>Aphis gossypii</i>)	

表 4

半翅目 (Hemiptera)	
麦长蝽 (<i>Blissus leucopterus leucopterus</i>)	喜绿蝽 (<i>Acrosternum hilare</i>)
美国牧草盲蝽 (<i>Lygus lineolaris</i>)	褐臭蝽 (<i>Euschistus servus</i>)

表 5

直翅目 (Orthoptera)	
赤胫黑蝗 (<i>Melanoplus femurrubrum</i>)	
血黑蝗 (<i>Melanoplus sanguinipes</i>)	
异黑蝗 (<i>Melanoplus differentialis</i>)	

表 6

双翅目 (Diptera)	
灰地种蝇 (<i>Hylemya platura</i>)	美州麦秆蝇 (<i>Meromyza americana</i>)
玉米斑潜蝇 (<i>Agromyza parvicornis</i>)	Wheat Bulb Fly (<i>Hylemya coarctata</i>)
高粱康瘿蚊 (<i>Contarinia sorghicola</i>)	向日葵籽瘿蚊 (<i>Neolasioptera murtfeldtiana</i>)
小麦瘿蚊 (<i>Mayetiola destructor</i>)	
麦红吸浆虫 (<i>Sitodiplosis mosellana</i>)	

表 7

 缨翅目 (Thysanoptera)

玉米黄足蓟马 (*Anaphothrips obscurus*)

Frankliniella fusca

烟蓟马 (*Thrips tabaci*)

大豆蓟马 (*Sericothrips variabilis*)

本发明核酸序列的表达产生了可用于控制鳞翅目昆虫的毒素, 所述鳞翅目昆虫例如, 非限制地包括欧洲玉米螟 (*Ostrinia nubilalis*), 小菜蛾 (*Plutella xylostella*), 草地粘虫 (*Spodoptera frugiperda*), 小地老虎 (*Agrotis ipsilon*), 谷实夜蛾 (*Helicoverpa zea*), 烟芽夜蛾 (*Heliothis virescens*), 贪夜蛾 (*Spodoptera exigua*), 棉红铃虫 (*Pectinophora gossypiella*), 粉纹夜蛾 (*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。

在一个优选的实施方式中, 本发明包括含有核苷酸序列的分离的核酸分子, 所述核苷酸序列 (a) 具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列 (compliment), 该杂交是在 50°C, 7% 十二烷基硫酸钠 (SDS), 0.5M NaPO₄, 1mM EDTA 中进行, 同时在 65°C, 0.1XSSC, 0.1% SDS 中漂洗; 或者 (b) 与 (a) 的核苷酸序列同类编码; 或 (c) 包含连续 20 个碱基对的核苷酸部分, 其序列与 (a) 或 (b) 核苷酸序列的连续 20 个碱基对核苷酸部分相同; 或 (d) 与 SEQ ID NO:1 具有至少 93% 序列同一性; 或 (e) 编码与 SEQ ID NO:2 具有至少 91% 序列同一性的氨基酸序列, 其中分离的核酸分子的表达产生昆虫控制活性。当 SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:10 和 SEQ ID NO:31 在异源宿主中表达时, 它们产生了抗欧洲玉米螟 (*Ostrinia nubilalis*), 小菜蛾 (*Plutella xylostella*), 草地粘虫 (*Spodoptera frugiperda*), 小地老虎 (*Agrotis ipsilon*), 谷实夜蛾 (*Helicoverpa zea*), 烟芽夜蛾 (*Heliothis virescens*), 贪夜蛾 (*Spodoptera exigua*), 棉红铃虫 (*Pectinophora gossypiella*), 粉纹夜蛾 (*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*) 的昆虫控制活性, 表明 SEQ ID NO:1, SEQ

ID NO:3, SEQ ID NO:10 和 SEQ ID NO:31 中所述核苷酸序列对于这种昆虫控制活性是充分的。

在一个实施方式中, 本发明包含核酸分子, 所述核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 75% 序列同一性的核苷酸序列。优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 85% 序列同一性的核苷酸序列。更优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 95% 序列同一性的核苷酸序列。甚至更优选地, 分离的核酸分子包含与 SEQ ID NO:1 核苷酸 1981-2367 具有至少 99% 序列同一性的核苷酸序列。最优选地, 分离的核酸分子包含 SEQ ID NO:1 或 SEQ ID NO:3 的核苷酸 1981-2367。

在另一个实施方式中, 本发明包含核酸分子, 该核酸分子包含与 SEQ ID NO:1 具有至少 93% 序列同一性的核苷酸序列。优选地, 分离的核酸分子包含与 SEQ ID NO:1 具有至少 95% 序列同一性的核苷酸序列。更优选地, 分离的核酸分子包含与 SEQ ID NO:1 具有至少 99% 序列同一性的核苷酸序列。最优选地, 分离的核酸分子包含选自 SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:10, SEQ ID NO:31 和 SEQ ID NO:33 的核苷酸序列的核苷酸 1-2367。

而在另一个实施方式中, 本发明包含核酸分子, 所述核酸分子包含在苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离物中, 该分离物选自命名为 NRRL 保藏号 B-30556 的 C1674; 和命名为 NRRL 保藏号 B-30557 的 C536。在优选的实施方式中, 本发明包含核酸分子, 所述核酸分子包含在大肠杆菌 (*E.coli*) 克隆中, 该大肠杆菌 (*E.coli*) 克隆选自命名为 NRRL 保藏号 B-30553 的 pNOV3910; 命名为 NRRL 保藏号 B-30552 的 pNOV3911; 命名为 NRRL 保藏号 B-30555 的 pNOV3906; 命名为 NRRL 保藏号 B-30554 的 pNOV3905; 和命名为 NRRL 保藏号 B-30551 的 pNOV3912, 它们的表达产生了杀虫毒素。

本发明也包含编码毒素的分离的核酸分子, 所述毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 75% 同一性的氨基酸序列。优选地, 分离的核酸分子编码包含氨基酸序列的毒素, 所述氨基酸

序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 85% 同一性。更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 95% 同一性。甚至更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 99% 同一性。最优选地，分离的核酸分子编码包含 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 的毒素。

在另一个实施方式中，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 91% 同一性。优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有至少 95% 同一性。更优选地，分离的核酸分子编码包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 中所述氨基酸序列具有 99% 同一性。最优选地，分离的核酸分子编码包含 SEQ ID NO:2 或 SEQ ID NO:11 或 SEQ ID NO:32 中所述氨基酸序列的毒素。

本发明也包括包含本发明核酸序列的重组载体。在这种载体中，优选地，该核酸序列包含在表达盒中，该表达盒包含用于在能表达所述核苷酸序列的转基因宿主细胞中表达该核苷酸序列的调节元件。这种调节元件通常包含启动子和终止信号，优选地，也包含使得本发明核酸序列编码的多肽能够有效翻译的元件。包含所述核酸序列的载体通常具有在特定宿主细胞中复制的能力，优选地，其作为染色体外分子，因此可以利用它在宿主细胞中扩增本发明的核酸分子。在一个实施方式中，用于这种载体的宿主细胞是微生物，如细菌，特别是大肠杆菌 (*E.coli*)。在另一个实施方式中，用于这种重组载体的宿主细胞是内寄生菌或体表寄生菌。用于这种载体的一种优选宿主细胞是真核生物细胞，如酵母细胞，植物细胞或昆虫细胞。植物细胞如玉米细胞或棉花细胞是最优选的宿主细胞。在另一个优选实施方式中，这种载体是病毒载体，并且可用于在特定宿主细胞，例如昆虫细胞或植物细胞中复制所述核苷酸序列。重组载体也用于将本发明核苷酸序列转化到转基因宿主细胞内，从而该

核苷酸序列稳定整合进入这种转基因宿主细胞的 DNA 中。在一个实施方式中，这种转基因宿主细胞是原核生物细胞。在优选实施方式中，这种转基因宿主细胞是真核细胞，如酵母细胞，昆虫细胞或植物细胞。在最优选的实施方式中，转基因宿主细胞是植物细胞，如玉米细胞或棉花细胞。

而在另一方面，本发明提供了由本发明核酸分子表达所产生的毒素。

在优选的实施方式中，本发明杀虫毒素包含由本发明核苷酸序列编码的多肽。在进一步优选的实施方式中，通过从选自命名为 NRRL 保藏号 B-30556 的 C1674；和命名为 NRRL 保藏号 B-30557 的 C536 的苏云金芽孢杆菌 (*Bacillus thuringiensis*) 分离物产生该毒素。

在另一个实施方式中，通过从选自命名为 NRRL 保藏号 B-30553 的 pNOV3910；命名为 NRRL 保藏号 B-30552 的 pNOV3911；命名为 NRRL 保藏号 B-30555 的 pNOV3906；命名为 NRRL 保藏号 B-30554 的 pNOV3905；和命名为 NRRL 保藏号 B-30551 的 pNOV3912 的大肠杆菌 (*E.coli*) 克隆产生所述毒素。在优选的实施方式中，通过表达包含 SEQ ID NO:1 核苷酸 1-2367，或 SEQ ID NO:3 核苷酸 1-2367，或 SEQ ID NO:10 核苷酸 1-2367，或 SEQ ID NO:31 核苷酸 1-2367 的核酸分子产生毒素。

本发明包括包含氨基酸序列的毒素，所述氨基酸序列与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 75% 同一性。优选地，毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 85% 同一性的氨基酸序列。更优选地，毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 95% 同一性的氨基酸序列。甚至更优选地，毒素包含与 SEQ ID NO:2 氨基酸序列的氨基酸 661-788 具有至少 99% 同一性的氨基酸序列。最优选地，毒素包含 SEQ ID NO:2 氨基酸 661-788。

在另一优选的实施方式中，本发明毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 91% 同一性的氨基酸序列。优选地，毒素包含与

SEQ ID NO:2 中所述氨基酸序列具有至少 95% 同一性的氨基酸序列。更优选地，毒素包含与 SEQ ID NO:2 中所述氨基酸序列具有至少 99% 同一性的氨基酸序列。最优选地，毒素包含 SEQ ID NO:2, SEQ ID NO:11 或 SEQ ID NO:32 中所述氨基酸序列。

当在生物测定中试验本发明毒素抗昆虫害虫时，本发明毒素有控制昆虫的活性。在另一个优选实施方式中，本发明毒素具有抗鳞翅目昆虫的活性，优选地抗欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾 (*Trichoplusia ni*)，*Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*)。在实施例 6, 8, 9 和 13 中进一步示例了本发明杀虫毒素控制昆虫的性质。

本发明也包括具有抗昆虫活性的杂合毒素，其中所述杂合毒素是由包含如下核苷酸序列的核酸分子编码的，该核苷酸序列：(a) 具有与 SEQ ID NO:1 核苷酸 1981-2367 杂交的互补序列，该杂交是在 50°C，7% 十二烷基硫酸钠 (SDS)，0.5M NaPO₄，1mM EDTA 中进行，在 65°C，0.1XSSC，0.1% SDS 中漂洗；或者 (b) 与 (a) 的核苷酸序列同类编码；或 (c) 包含连续 20 个碱基对的核苷酸部分，其序列与 (a) 或 (b) 核苷酸序列的连续 20 个碱基对核苷酸部分相同，其中所述核酸分子的表达产生了昆虫控制活性。在优选的实施方式中，杂合毒素是由包含在 pNOV3912 中的约 2.4kb DNA 片段编码的，所述 pNOV3912 保藏在大肠杆菌 (*E.coli*) 株系 DH5α 中，命名为 NRRL 保藏号 B-30551，它的表达产生杀虫杂合毒素。这里具体示例的是 SEQ ID NO:10 中所述核苷酸序列编码的杂合毒素。当在异源宿主中表达 SEQ ID NO:10 核酸分子时，它产生了抗欧洲玉米螟 (*Ostrinia nubilalis*)，小菜蛾 (*Plutella xylostella*)，草地粘虫 (*Spodoptera frugiperda*)，小地老虎 (*Agrotis ipsilon*)，谷实夜蛾 (*Helicoverpa zea*)，烟芽夜蛾 (*Heliothis virescens*)，贪夜蛾 (*Spodoptera exigua*)，棉红铃虫 (*Pectinophora gossypiella*)，粉纹夜蛾

(*Trichoplusia ni*), *Cochyles hospes* 和向日葵同斑螟 (*Homoeosoma electellum*) 的控制昆虫活性。在实施例 9 中进一步阐述了本发明示例杂合毒素的控制昆虫特性。

本发明也包括具有抗昆虫活性的杂合毒素, 其包含 Vip3 毒素羧基末端区域和按氨基到羧基的方向与之连接的不同 Vip3 毒素的氨基末端区域, 其中所述羧基末端区域包含与 SEQ ID NO: 2 氨基酸 661-788 有至少 75% 同一性, 优选至少 85% 同一性, 更优选至少 95% 同一性, 最优选至少 99% 同一性的氨基酸序列; 和其中所述氨基末端区域与 SEQ ID NO: 5 氨基酸 1-660 有至少 75% 同一性, 优选至少 85% 同一性, 更优选至少 95% 同一性, 最优选至少 99% 同一性。在优选的实施方式中, 该羧基末端区域包含 SEQ ID NO: 2 氨基酸 661-788, 并且该氨基末端区域包含 SEQ ID NO: 65 氨基酸 1-660。在更优选的实施方式中, 杂合毒素包含 SEQ ID NO: 11 氨基酸 1-788。

在进一步实施方式中, 可以在称作体外重组或 DNA 改组的技术中, 通过掺入随机突变修饰本发明核苷酸序列。在 Stemmer 等, *Nature* 370: 389-391 (1994) 和 U. S. 专利 5,605, 793 (并入本文作为参考) 中描述了该技术。根据本发明原始的核苷酸序列可以产生几百万个核苷酸序列的突变体拷贝, 回收具有改进特性, 诸如增加的杀虫活性, 增加的稳定性, 或不同的特异性或靶昆虫害虫范围的变异体。该方法包括从包含本发明核苷酸序列的双链多核苷酸模板形成诱变的双链多核苷酸, 其中双链多核苷酸模板已经被切割为期望大小的双链随机片段, 并且包括包括如下步骤: 向由此得到的双链随机片段群加入一种或多种单链或双链寡核苷酸, 其中所述寡核苷酸包含与双链模板多核苷酸同一的区域和异源的区域; 将由此得到的双链随机片段和寡核苷酸的混合物变性为单链片段; 在引起所述单链片段在所述同一性区域退火以形成成对的退火片段的条件下, 温育由此得到的单链片段群和聚合酶, 所述同一性区域足以使配对的一个成员引发另一个成员的复制, 由此形成了突变的双链多核苷酸; 和重复第二和第三步骤至少另两轮循环, 其中在下一轮循环第二步骤中得到的混合物包含从前一轮循环第三步骤得到的诱变的双链多核

苷酸，并且所述下一轮循环形成了进一步诱变的双链多核苷酸。在优选的实施方式中，双链随机片段群中单个种类的双链随机片段的浓度小于总 DNA 的 1% 重量。在进一步优选的实施方式中，双链多核苷酸模板包含至少约 100 个种类的多核苷酸。在另一个优选的实施方式中，双链随机片段的大小约是 5bp 到 5kb。在进一步优选的实施方式中，该方法的第四步骤包括重复第二和第三步骤至少 10 个循环。

异源微生物宿主中核苷酸序列的表达

作为生物学昆虫控制剂，所述杀虫毒素可通过在能表达所述核苷酸序列的异源宿主细胞中表达核苷酸序列而产生。在第一个实施方式中，产生了包含本发明核苷酸序列修饰的苏云金芽孢杆菌 (*B. thuringiensis*) 细胞。这种修饰包含现存调节元件的突变或缺失，因此引起了核苷酸序列改变的表达，或者引入了控制核苷酸序列表达的新调节元件。在另一个实施方式中，通过在染色体内插入或通过含有该核苷酸序列的染色体外复制分子的引入，向苏云金芽孢杆菌 (*Bacillus thuringiensis*) 细胞中加入一个或多个核苷酸序列的额外拷贝。

在另一个实施方式中，将至少一个本发明核苷酸序列插入到适合的表达盒中，该表达盒包含启动子和终止信号。核苷酸序列的表达是组成型的，或利用对各种类型刺激物反应以起始转录的诱导型启动子。在优选实施方式中，表达毒素的细胞是微生物，如病毒，细菌或真菌。在优选实施方式中，病毒如杆状病毒在基因组中含有本发明核苷酸序列，感染适合病毒复制和所述核苷酸序列表达的适当真核生物细胞后可表达大量的相应杀虫毒素。由此产生的杀虫毒素用作杀虫剂。可替代地，用基因工程改造成包含该核苷酸序列的杆状病毒体内感染昆虫，通过杀虫毒素的表达或通过病毒感染和杀虫毒素表达的联合杀死昆虫。

细菌细胞也是用于本发明核苷酸序列表达的宿主。在优选实施方式中，利用能在植物组织内生活和繁殖的非致病共生细菌（通常所说的内寄生菌）或能在叶际或根际建群的非致病共生细菌（通常所说的体表寄生菌）。这种细菌包括农杆菌属 (*Agrobacterium*)，产碱杆菌属

(*Alcaligenes*), 固氮根瘤菌属 (*Azospirillum*), 固氮菌属 (*Azotobacter*), 芽孢杆菌属 (*Bacillus*), 棍状杆菌属 (*Clavibacter*), 肠杆菌属 (*Enterobacter*), 欧文氏菌属 (*Erwinia*), 黄杆菌属 (*Flavobacter*), 克氏杆菌属 (*Klebsiella*), 假单胞菌属 (*Pseudomonas*), 根瘤菌属 (*Rhizobium*), 沙雷氏菌属 (*Serratia*), 链霉菌属 (*Streptomyces*) 和黄单胞菌属 (*Xanthomonas*) 的细菌。共生真菌, 如木霉菌属 (*Trichoderma*) 和胶霉属 (*Gliocladium*) 也可能用作本发明核苷酸序列表达的宿主以用于同一目的。

这些基因操作技术对于可供利用的不同宿主是特定的, 而且是本领域已知的。例如, 表达载体 pKK223-3 和 pKK223-2 可以用于在大肠杆菌 (*E.coli*) 中在 *tac* 或 *trc* 启动子后以转录或翻译融合形式表达异源基因。对于编码多 ORFs (开放读框) 的操纵子的表达, 最简单的方法是以转录融合方式在载体, 如 pKK223-3 中插入该操纵子, 使得能够利用异源基因的相关核糖体结合位点。在革兰氏阳性物种如芽孢杆菌 (*Bacillus*) 中超量表达的技术也是本领域已知的, 其也可以在本发明上下文中使用 (Quax 等, In: *Industrial Microorganisms : Basic and Applied Molecular Genetics*, Baltz 等编辑, American Society for Microbiology, Washington (1993))。可替代的过量表达系统依赖于例如酵母载体, 并包括毕赤酵母属 (*Pichia*), 酵母属 (*Saccharomyces*) 和克鲁维酵母菌属 (*Kluyveromyces*) 的应用 (Sreekrishna, In: *Industrial microorganisms : basic and applied molecular genetics*, Baltz, Hegeman 和 Skatrud 编辑, American Society for Microbiology, Washington (1993); Dequin & Barre, *Biotechnology L2*: 173-177 (1994); van den Berg 等, *Biotechnology 8* : 135-139 (1990))。

植物转化

在特别优选实施方式中, 在高等生物体例如植物中表达至少一种本发明的杀虫毒素。在这种情况下, 表达有效量毒素的转基因植物保护其本身免受昆虫害虫的侵害。当昆虫开始以这种转基因植物为食时, 它也

摄食了表达的毒素。这将阻止昆虫进一步咬食植物组织，或者，甚至伤害或杀死昆虫。可将本发明核苷酸序列插入到表达盒中，然后优选地，表达盒稳定整合在所述植物基因组中。在另一个优选实施方式中，核苷酸序列包含在非致病自我复制的病毒中。按照本发明转化的植物可以是单子叶植物或双子叶植物，包括但不限于玉米，小麦，大麦，黑麦，甘薯，豆，豌豆，菊苣，莴苣，甘蓝，花椰菜，花茎甘蓝，芜菁，萝卜，菠菜，芦笋，洋葱，大蒜，胡椒，芹菜，笋瓜，南瓜，大麻，夏南瓜，苹果，梨，温梓，瓜，李子，樱桃，桃子，油桃，杏，草莓，葡萄，木莓，黑莓，菠萝，鳄梨，蕃木瓜，芒果，香蕉，大豆，番茄，高粱，甘蔗，甜菜，向日葵，菜籽油菜，三叶草，烟草，胡萝卜，棉花，苜蓿，稻，马铃薯，茄子，黄瓜，拟南芥属和木本植物如针叶树和落叶树。

一旦已将期望的核苷酸序列转化进入特定植物物种中，可以在该物种中繁殖它或用常规育种技术将它转移进入相同物种的其它品种，特别包括商业品种中。

优选地，在转基因植物中表达本发明的核苷酸序列，由此在转基因植物中引起相应毒素的生物合成。以这种方式，可产生具有增强抗昆虫抗性的转基因植物。为了在转基因植物中表达本发明核苷酸序列，本发明核苷酸序列可能需要修饰和优化。尽管在许多情况下来源于微生物的基因不用修饰就可以在植物中高水平地表达，但是具有在植物中不偏爱的密码子的微生物核苷酸序列可能引起转基因植物中低水平的表达。所有生物体都有特定的密码子使用偏爱性，这是本领域已知的，可以在保持本发明所述核苷酸序列编码的氨基酸的同时改变其密码子以符合植物偏爱性。而且，从有至少约 35%，优选多于约 45%，更优选多于 50%，最优选多于约 60% GC 含量的编码序列可以最好地实现植物中高水平的表达。由于存在可能使信息去稳定的 ATTTA 基序和可以引起不适合聚腺苷酸化的 AATAAA 基序，有低 GC 含量的微生物核苷酸序列在植物中可能低水平地表达。尽管可以在单子叶植物和双子叶植物物种中充分地表达优选的基因序列，但可以修饰序列以适应单子叶植物或双子叶植物的特异密码子偏好和 GC 含量偏好，因为这些偏好已被证明是不

同的(Murray 等, Nucl. Acids Res. 17: 477-498 (1989))。此外, 可筛选核苷酸序列以寻找引起信息截断的非常规剪接位点的存在。利用公开专利申请 EP 0 385 962(Monsanto), EP 0 359 472(Lubrizol) 和 WO 93/07278(Ciba-Geigy)中所述的方法, 用本领域熟知的位点定向诱变技术, PCR 和合成基因构建进行在这些核苷酸序列中需要进行的所有改变, 如上述那些改变。

在本发明一个实施方式中, 根据这里并入作为参考文献的美国专利 5,625,136 中公开的方法可制备合成基因。在该方法中, 利用了玉米优选的密码子, 即最常常编码玉米中那个氨基酸的单密码子。特定氨基酸的玉米优选密码子可以来源于, 例如玉米的已知基因序列。在 Murray 等, Nucleic Acids Research 17: 477-498 (1989)中教导了玉米植物 28 个基因的玉米密码子使用, 这里并入这篇文献的公开内容为参考文献。制备的具有玉米最优密码子的本发明具体示例的合成序列如 SEQ ID NO: 3 和 SEQ ID NO: 33 中所述。

以这种方式可以优化核苷酸序列以便在任何植物中的表达。公认基因序列的所有或任何部分都可以优化或合成。即, 也可以利用合成或部分优化的序列。

为了翻译的有效起始, 可能需要修饰邻近起始甲硫氨酸的序列。例如, 通过包含已知在植物中有效的序列可以修饰它们。Joshi 提出了植物适合的共有序列(NAR 15: 6643-6653 (1987)), Clonetech 提出了进一步的翻译起始子共有序列(1993/1994 目录, 210 页)。这些共有序列适合与本发明核苷酸序列一起使用。向包含所述核苷酸序列, 直到和包含 ATG(同时保持第二个氨基酸没有被修饰)或可替代地直到和包含 ATG 后的 GTC(具有修饰转基因第二个氨基酸的可能性)的构建物中引入该序列。

可将作为其天然序列或作为如上所述优化合成序列的本发明新的 *vip3* 毒素基因可操作地与在植物中表达的各种启动子, 包括组成型, 诱导型, 时序调节, 发育调节, 化学调节, 组织优选和组织特异性启动子相融合以制备重组 DNA 分子, 即嵌合基因。启动子的选择将随着表达

时间和空间需要而变化，而且也取决于靶物种。因此，优选在叶，主茎或茎秆，穗，花序（例如，穗状花序，圆锥花序，穗轴等），根，和/或幼苗中表达本发明核苷酸序列。然而，在许多情况下，寻求抗多于一种类型的昆虫害虫的保护，因此期望在多组织中的表达。尽管证明了来源于双子叶植物的许多启动子在单子叶植物中是可起作用的，反之亦然，但是理想地，选择双子叶植物启动子用于双子叶植物中的表达，单子叶植物的启动子用于单子叶植物中的表达。然而，没有限制所选启动子的起源，只要启动子起作用驱动期望细胞中核苷酸序列的表达就足够了。

优选的组成型启动子包括 CaMV 35S 和 19S 启动子(Fraley 等, 1994 年 10 月 4 日公布的美国专利号 5,352,605)。另外优选的启动子来源于在大多数细胞类型中表达的几种肌动蛋白基因的任何一种。可以容易地修饰 McElroy 等(Mol. Gen. Genet. 231: 150-160 (1991))所述的启动子表达盒以用于新毒素基因的表达，该基因表达盒特别适合在单子叶植物宿主中使用。

另一个优选的组成型启动子来源于泛素，它是已知在许多细胞类型中积累的另一种基因产物。从几种物种，例如向日葵(Binet 等, 1991. Plant Science 79: 87-94)，玉米(Christensen 等, 1989. Plant Molec. Biol. 12 : 619-632)和拟南芥(Norris 等 1993. Plant Molec. Biol. 21: 895-906)克隆了泛素启动子，可用于转基因植物中。已发展了转基因单子叶植物系统中的玉米泛素启动子，并且在专利公布 EP 0 342 926 中公开了其序列和用于单子叶植物转化的载体。泛素启动子适合在转基因植物，特别是单子叶植物中新毒素基因的表达。

可用于在植物，特别是玉米中表达本发明新毒素基因的组织特异性或组织偏爱启动子是在根，木髓，叶或花粉中引导表达的启动子。在 WO 93/07278 中公开了这种启动子，这里整体并入其为参考文献。其它可用在本发明中的组织特异性启动子包括美国专利 6,040,504 中公开的棉花核酮糖二磷酸羧化酶-加氧酶启动子；美国专利 5,604,121 中公开的水稻蔗糖合酶启动子；和 WO 01/73087 中公开的夜香树黄化叶卷曲病

毒启动子，并入所有这些文献为参考文献。在美国专利 5,614,395 中公开了可用于指导植物中新毒素基因表达的化学诱导型启动子，这里整体并入这篇文献为参考文献。

也可以在化学调节型启动子调节下表达本发明核苷酸序列。这能够使得仅仅在用诱导化学试剂处理作物时才合成 Vip3 毒素。在公开的申请 EP 0 332 104 (申请人 Ciba-Geigy) 和美国专利 5,614,395 中详述了基因表达化学诱导的优选技术。化学诱导的优选启动子是烟草 PR-1a 启动子。

一种优选的启动子种类是创伤诱导型启动子。已经描述了在创伤位点及植物病原体感染位点表达的大量启动子。理想地，这种启动子应该仅在感染位点局部有活性，从而杀虫毒素仅在需要合成杀虫毒素以杀死侵入昆虫害虫的细胞中积累。优选的这种启动子包括 Stanford 等, *Mol. Gen. Genet.* 215: 200-208 (1989), Xu 等, *Plant Molec. Biol.* 22: 573-588 (1993), Logemann 等, *Plant Cell* 1: 151-158 (1989), Rohrmeier & Lehle, *Plant Molec. Biol.* 22: 783-792 (1993), Firek 等, *Plant Molec. Biol.* 22: 129-142 (1993), 和 Warner 等, *Plant J.* 3: 191-201 (1993) 所述的启动子。

优选的组织特异性表达模式包括绿色组织特异性，根特异性，茎特异性和花特异性。适于绿色组织中表达的启动子包括许多调节参与光合作用的基因的启动子，这样的许多启动子已经从单子叶植物和双子叶植物克隆出来了。优选的启动子是来源于磷酸烯醇羧化酶基因的玉米 PEPC 启动子 (Hudspeth & Grula, *Plant Molec. Biol.* 12: 579-589 (1989))。根特异性表达的优选启动子是 de Framond (*FEBS* 290: 103-106 (1991); EP 0 452 269 to Ciba-Geigy) 所述的启动子。优选的茎特异性启动子是美国专利 5,625,136 (申请人 Ciba-Geigy) 中所述的启动子，该启动子驱动玉米 *trpA* 基因的表达。

其它优选实施方式是以创伤诱导方式或病原体感染诱导方式表达核苷酸序列的转基因植物。

除了适合启动子的选择外，植物中杀虫毒素表达的构建需要连接在异源核苷酸序列下游的适合的转录终止子。可得到几种这种终止子，并

且它们是本领域已知的(例如来源于 CaMV 的 tml, 来源于 rbcS 的 E9)。任何已知在植物中起作用的可得到的终止子都可以用在本发明中。

也可以向本发明所述表达盒中引入大量其它的序列。这些序列包括已经证明增强表达的序列, 如内含子序列(例如来源于 Adh1 和 bronzel) 和病毒前导序列(例如来源于 TMV, MCMV 和 AMV)。

有可能优选将本发明核苷酸序列的表达靶向到植物中不同细胞定位。在一些情况下, 可能期望定位在细胞溶质中, 而在其它情况下, 可能优选定位在一些亚细胞细胞器中。转基因所编码的酶的亚细胞定位采用了本领域熟知的技术。通常, 操作编码来源于已知靶向细胞器的基因产物的靶肽的 DNA, 使之融合到核苷酸序列的上游。已知用于叶绿体的许多这种靶序列, 并且已知证明了它们在异源构建中的功能。本发明核苷酸序列的表达也可靶向到宿主细胞的内质网或液泡中。实现上述这些目的的技术是本领域熟知的。

可用于植物转化的大量转化载体是植物转化领域技术人员已知的, 并且本发明的核酸分子可以与任何这种载体联合使用。载体的选择将依赖于用于转化的优选转化技术和靶植物物种。对于一些靶物种, 可以优选不同的抗生素或除草剂选择性标记。通常用在转化中的选择性标记包括赋予对卡那霉素和相关抗生素抗性的 *nptII* 基因 (Messing & Vierra, 1982. Gene 19: 259-268; 和 Bevan 等, 1983. Nature 304: 184-187), 赋予对除草剂膦丝菌素抗性的 *bar* 基因 gene, (White 等, 1990. Nucl. Acids Res 18: 1062, 和 Spencer 等, 1990. Theor. Appl. Genet 79: 625-631), 赋予对抗生素潮霉素抗性的 *hph* 基因 (Blochinger & Diggelmann, Mol Cell Biol 4: 2929-2931), 和赋予对 methatrexate 抗性的 *dhfr* 基因 (Bourouis 等, 1983. EMBO J. 2 (7): 1099-1104), 赋予对草甘磷抗性的 EPSPS 基因 (U. S. 专利号: 4,940, 935 和 5,188, 642), 和提供代谢甘露糖能力的甘露糖-6-磷酸异构酶基因(美国专利号 5,767, 378 和 5,994, 629)。然而, 选择性标记的选择对本发明不是至关重要的。

在另一个优选的实施方式中, 将本发明的核苷酸序列直接转化到质体基因组中。质体转化的主要优点是质体通常不需要实质修饰就能表达

细菌基因，并且质体能表达单启动子控制下的多个开放读框。在美国专利 5,451,513、5,545,817 和 5,545,818 中，PCT 申请号 WO 95/16783 中和 McBride 等，(1994) Proc. Nati. Acad. Sci. USA 91,7301-7305 中详尽地描述了质体转化技术。叶绿体转化的基本技术包括例如利用生物轰击或原生质体转化（例如氯化钙或 PEG 介导的转化）将位于目的基因和选择标记侧翼的克隆质体 DNA 的区域一起引入适合的靶组织中。1 到 1.5kb 侧翼区域，称为导向序列，可促进与质体基因组的同源重组，因此允许原质体系特定区域的置换或修饰。最初，可利用在提供了对壮观霉素和/或链霉素抗性的叶绿体 16S rRNA 和 rps12 基因的点突变作为转化的选择标记(Svab, Z., Hajdukiewicz, P.,和 Maliga, P. (1990) Proc. Nati. Acad. Sci. USA 87,8526- 8530; Staub, J. M.,和 Maliga, P. (1992) Plant Cell 4,39-45)。这以约每 100 次靶叶片轰击 1 次的频率产生了稳定的同质转化体。这些标记间存在的克隆位点可以用来产生用于导入外源基因的质体靶向载体(Staub, J. M.,和 Maliga, P. (1993) EMBO J. 12,601-606)。通过用显性选择标记，编码壮观霉素去毒酶（spectinomycin-cletoxifying enzyme）氨基糖苷类-3'-腺苷酰转移酶的细菌 aadA 基因置换隐性 rRNA 或 r-蛋白质抗生素抗性基因可获得转化频率的显著增加(Svab, Z.和 Maliga, P. (1993) Proc. Natl. Acad. Sci. USA 90,913-917)。以前，该标记成功地用于高频率转化绿藻 *Chlamydomonas reinhardtii* 质体基因组(Goldschmidt-Clermont, M. (1991) Nucl. Acids Res. 19: 4083-4089)。其它用于质体转化的选择标记是本领域已知的，并且包含在本发明范围内。通常，转化后需要约 15 到 20 个细胞分裂周期以达到同质状态。通过同源重组将基因插入每个植物细胞中存在的所有几千个环形质体基因组拷贝中的质体表达利用了拷贝数大大高于核表达基因的优势，使得表达水平可以容易地超过总可溶植物蛋白质的 10%。在优选实施方式中，将本发明核苷酸序列插入到质体靶向载体中，并且转化进入期望的植物宿主质体基因组中。获得了对于含有本发明核苷酸序列的质体基因组而言属同质的植物，优选地，该植物具有高水平地表达核苷酸序列的能力。

昆虫控制成分的联合

可以联合 Bt δ -内毒素或其它杀虫成分来使用本发明的杀虫毒素以增加害虫靶范围。而且,本发明杀虫毒素与 Bt δ -内毒素或其它不同特性的杀虫要素的联合使用,对于昆虫抗性的预防和/或处理具有特定的用途。

已经根据苏云金芽孢杆菌 (*Bacillus thuringiensis*) 的各种杀虫结晶蛋白质的活性谱和序列相似性对其进行了分类。Hofte 和 Whiteley, *Microbiol. Rev.* 53: 242-255 (1989) 提出的分类将当时已知的杀虫结晶蛋白质分为 4 大类。通常,通过活性谱定义所述的大类, Cry1 蛋白质具有抗鳞翅目活性, Cry2 蛋白质具有抗鳞翅目和双翅目活性, Cry3 蛋白质具有抗鞘翅目活性, 和 Cry4 蛋白质具有抗双翅目活性。

在每个大类中,根据序列相似性将 δ -内毒素分组。Cry1 蛋白质通常以 130 - 140kDa 前毒素蛋白质产生,蛋白酶解切割该前毒素后产生约 60 - 70kDa 的活性毒素。 δ -内毒素的活性部分存在于全长分子的 NH₂-末端部分。Hofte 和 Whiteley (上文) 将当时已知的 Cry1 蛋白质分为 6 组, 1Aa, 1Ab, 1Ac, 1B, 1C 和 1D。从此,也已鉴定了分类为 Cry1Ea, Cry1Fa, Cry9A, Cry9C 和 Cry9B 及其它类的蛋白质。

苏云金芽孢杆菌 (*Bacillus thuringiensis*) 各 δ -内毒素的杀虫活性谱趋向于非常窄,一种 δ -内毒素仅仅具有抗几种昆虫活性。特异性是参与产生活性毒素蛋白质各个步骤的效率和其随后与昆虫消化道上皮细胞相互作用的能力的结果。在一个优选实施方式中,转基因植物中本发明核酸分子的表达伴随着一种或多种 Bt δ -内毒素的表达。特别优选的 Bt δ -内毒素是在美国专利 5,625,136 中公开的 Bt δ -内毒素,这里并入其为参考文献。

熟知许多苏云金芽孢杆菌 (*Bacillus thuringiensis*) 的 δ -内毒素蛋白质实际上以前毒素形式表达。这些前毒素溶解在昆虫肠道碱性环境中,通过蛋白酶的蛋白酶解转化为有毒的核心片段(Hofte 和 Whiteley, *Microbiol. Rev.* 53: 242-255 (1989))。对于 Cry1 类的 δ -内毒素蛋白质,

有毒的核心片段定位在前毒素 N 末端一半中。在本发明范围内，可以在植物转化载体中利用编码新的毒素蛋白质全长前毒素形式或截短有毒核心片段的基因以提供对宿主植物的杀虫特性。

其它的杀虫成分包括蛋白酶抑制剂（丝氨酸和半胱氨酸类型），凝集素， α -淀粉酶，过氧化物酶和胆固醇氧化酶。本发明中也可利用其它 Vip 基因，如美国专利 5,849,870 中所公开的 *vip1A(a)* 和 *vip2A(a)*，这里并入其为参考文献。

通过基因工程改造植物以含有和表达所有必需的基因可以实现相同转基因植物中该一个以上杀虫成分的共表达。可替代地，可以基因工程改造植物亲本 1 以用于本发明基因的表达。可以基因工程改造第二种植物，亲本 2 以用于补充的控制昆虫成分的表达。通过杂交亲本 1 和亲本 2，可获得表达引入亲本 1 和亲本 2 中所有基因的子代植物。

本发明也包含公开核酸分子的变异体。通过应用熟知的分子生物学技术，例如，如下所概述的 PCR 和杂交技术可以鉴定和/或分离天然存在的变异体序列。

变异体 *vip3* 核苷酸序列包含合成来源的核苷酸序列，如，例如通过利用位点定向诱变产生的核苷酸序列，或通过整个结构域交换产生的核苷酸序列，但是这些核苷酸序列仍然显示了杀虫活性。诱变和核苷酸序列改变的方法是本领域已知的。参见，例如 Kunkel (1985) Proc. Natl. Acad. Sci. USA 82: 488-492; Kunkel 等(1987) Methods in Enzymol. 154: 367-382 ; 美国专利号 4,873,192; Walker 和 Gastra 编辑(1983) Techniques in Molecular Biology (MacMillan Publishing Company, New York)及其中引用的参考文献。通常，本发明的核苷酸序列与其各自的参照 *vip3* 核苷酸序列有至少 80%，优选 85%，90%，95%，直到 98% 或更多的序列同一性，并且有杀虫活性。

变异体 *vip3* 核苷酸序列也包含来源于诱变和重组方法如 DNA 改组的序列。用这种方法，可以将本发明的一种或多种不同的 *vip3* 序列（例如 *vip3C(a)*，*vip3C(b)*，*vip3A-C* 和 *vip3C-12168*，但不限于此）重组在一起，或与其它 *vip3* 或相关序列（例如 *vip3A*(SEQ ID NO: 4)，

vip3B(SEQ ID NO: 6)和 *vip3Z*(SEQ ID NO: 8) , 但不限于此)相重组, 以产生编码具有期望特性的 Vip3 毒素的新的 *vip3* 核酸分子。用这种方式, 从序列相关 *vip3* 多核苷酸群产生重组 *vip3* 多核苷酸文库, 所述序列相关 *vip3* 多核苷酸包含具有基本序列同一性的序列区域, 并且可以在体外或体内同源重组, 对于这种 DNA 改组的策略是本领域已知的。参见, 例如 Stemmer (1994) Proc. Natl. Acad. Sci. USA 91: 10747-10751; Stemmer (1994) Nature 370: 389- 391; Cramer 等, (1997) Nature Biotech. 15: 436-438 ; Moore 等, (1997) J. Mol. Biol. 272: 336-347; Zhang 等, (1997) Proc. Natl. Acad. Sci. USA 94: 4504-4509; Cramer 等, (1998) Nature 391: 288-291; 国际专利申请 WO 99/57128 和美国专利号 5,605,793,5,837,458 和 6,335,179。

可以将这里公开的诱变方法和高处理量的筛选方法联合应用, 以检测宿主细胞中克隆化的诱变的 Vip3 多肽的杀虫活性。可以从宿主细胞回收编码活性 Vip3 多肽(例如, 分泌性的和通过抗体检测的; 或昆虫生物测定中有杀虫活性的)的诱变 DNA 分子, 并用标准工艺方法快速测序。这些方法可以快速测定目的 Vip3 多肽中单个氨基酸残基的重要性, 可以应用于未知结构的多肽。

可对用 DNA 改组方法产生的重组 *vip3* 基因文库进行筛选, 以鉴定显示在用于保护植物抗害虫方面具有改进特性的重组 *vip3* 基因。可利用 DNA 改组获得改进的 *vip3* 害虫抗性基因的特性包括增加的抗靶害虫的效力, 增加的靶害虫范围, 害虫发展抗性可能性降低, 增加的表达水平, 对蛋白酶降解有增加的抗性, 增加的环境稳定性和对宿主植物的毒性降低。通过利用适合的筛选策略, 能同时或依次获得对多个特性进行优化的 *vip3* 基因。

利用 DNA 改组可获得编码毒素的 *vip3* 害虫抗性基因, 所述毒素显示了增强的抗靶害虫效力。一旦完成改组, 即可筛选由此得到的改组 *vip3* 基因文库以鉴定显示增强杀虫活性的 *vip3* 基因。进行该筛选的一个方法是将改组 *vip3* 基因的蛋白质编码区域克隆到表达载体中, 所述表达载体适于在选定宿主细胞如, 例如大肠杆菌 (*E.coli*) 或苏云金芽孢杆

菌 (*Bacillus thuringiensis*) 结晶缺陷株系中表达基因。本领域技术人员将认识到利用这两种表达系统的优点和缺点。例如在产生分泌性 Vip3 蛋白质中将更期望使用苏云金芽孢杆菌 (*Bacillus thuringiensis*)。如果期望, 可以对克隆进行初步筛选, 例如通过免疫测定以鉴定产生正确大小 Vip3 蛋白质的克隆。然后在功能筛选中检测初步筛选中的阳性克隆以鉴定编码有期望增强活性的毒素的改组 *vip3* 基因。

整个昆虫测定法可以用于检测毒性。在这些测定法中, 在昆虫饵料, 例如人工饵料和植物组织上放置改组 *vip3* 基因表达的 Vip3 毒素, 并且使靶昆虫消耗该 Vip3 毒素。可以在生物测定中进一步检测引起靶昆虫生长抑制或死亡率的这些克隆以测定效力。编码具有增强效力的毒素的改组 *vip3* 基因可以鉴定为有降低 EC₅₀ (降低 50% 昆虫生长必需的毒素浓度) 和/或 LC₅₀ (引起 50% 死亡率必需的毒素浓度) 的改组 *vip3* 基因。

体外测定法也可以用于筛选改组 *vip3* 基因文库。这种测定法通常包括应用对 Vip3 毒素敏感的培养的昆虫细胞和/或表达 Vip3 毒素受体的细胞, 所述 Vip3 毒素受体或者是天然的或者是异源基因表达的结果。可以利用其它的体外测定法, 例如测定细胞形态学改变, 用于测定细胞死亡的染料和标记, 或细胞 ATPase 释放的测定。利用培养的昆虫细胞用于 Vip3 毒性测定的一个适合的体外测定例子是 Sf9 (*Spodoptera frugiperda*) 细胞。Sf9 对 Vip3 毒素高度敏感。当混合 Vip3 毒素和 Sf9 细胞时, 细胞膜变得对小分子高度可渗透。当向细胞悬浮液加入染料, 如台盼蓝时, 那些被 Vip3 毒素杀死的细胞染色为蓝色。因此, 可以通过图象分析测定 Vip3 毒素的细胞毒性。

另一个体外测定法包括 Vip3 毒素受体的应用。在美国专利 6,291,156 中公开了一种这种受体, 这里并其为参考文献。可以在接受表面 (例如 96 孔板或硝化纤维素膜, 但不限于此) 固定 Vip3 受体蛋白质, 并且暴露于包含改组 *vip3* 基因的克隆。因此, 可以根据对 Vip3 受体的结合亲和性鉴定编码功能毒素的改组 *vip3* 基因。而且, 利用本领域已知的方法 (参见, 例如 Clem 和 Miller, 1194, Mol. Cel. Biol. 14: 5212-522) 可以将编码 Vip3 受体的基因转化进入非 Vip3 敏感细胞系, 例如

Schneider 2 (S2)果蝇细胞系。然后,转化的 S2 细胞暴露于包含改组 vip3 基因的克隆。因此,可以根据细胞死亡的诱导鉴定编码功能性毒素的改组 vip3 基因。

实施例

参考下面详细的实施例将进一步描述本发明。仅仅为了示例目的提供这些实施例,除了另外说明,这些实施例不意味是限制性的。这里所用的标准重组 DNA 和分子克隆技术也是本领域已知的,并且 Ansel (编辑) *Current Protocols in Molecular Biology*, John Wiley and Sons, Inc. (1994); J. Sambrook, *Molecular Cloning: A Laboratory Manual*, 第三版, Cold Spring Harbor, NY :Cold Spring Harbor laboratory Press(2001); 和 T. J. Silhavy, M. L. Berman,和 L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984)描述了这些技术。

实施例 1: 鉴定含有 Vip3 同源蛋白质的 Bt 分离物

在 PCR 反应中使用三组 PCR 引物以从苏云金芽孢杆菌 (*Bacillus thuringiensis*) (*Bt*)分离物扩增可能的同源 vip3 基因片段,所述 PCR 引物序列是根据 vip3A 基因 (SEQ ID NO:5) 设计的。所用的三组引物是:

1F: 5'-ATGAACAAGAATAATACTAAATTAAGCACAAGAGCC-3'	(SEQ ID NO: 12)
1R: 5'-CTCAACATAGAGGTAATTTTAGGTAGATATACCCG-3'	(SEQ ID NO: 13)
p3: 5'-GATGATGGGGTGTATATGCCGTTAG-3'	(SEQ ID NO: 14)
p4: 5'-AATAAATTGTGAAATTCCTCCGTCC-3'	(SEQ ID NO: 15)
4F: 5'-AGTCAAATGGAGATCAAGGTTGGGGAGATAAC-3'	(SEQ ID NO: 16)
4R: 5'-TTACTTAATAGAGAGATCGTGAAATGTACAATA-3'	(SEQ ID NO: 17)

如果 *Bt* 分离物含有与 vip3A 基因 (SEQ ID NO:4) 相同的基因,那么可以预测有 3 种 PCR 产物。用引物组 1F/1R, p3/p4 和 4F/4R 产生的 PCR 产物大小分别是 377 bp, 344 bp 和 419 bp。对仅产生 1 种或 2 种 PCR 产物,表明它们可能含有与 vip3A 具有一些序列差异的 vip3 基因

的分离物进行进一步序列分析。

实施例 2: 克隆和测序 PCR 产物以证实 Vip3 同源序列

用引物组 1F/1R (SEQ ID NO:12/SEQ ID NO:13) 及下面两种引物对实施例 1 中鉴定的产生 1 种或 2 种 PCR 产物的 Bt 分离物再一次进行 PCR:

p5: 5'-AATGGAGATGAAGCTTGGGGAGAT-3' (SEQ ID NO: 18)

p6: 5'-CGTGGAAATGTACAATAGGACCACC-3' (SEQ ID NO: 19)

然后用本领域的标准方法, 将 PCR 产物克隆到 pCR2.1-Topo (Invitrogen) 载体中, 并且测序。

确定 3 种 Bt 分离物含有同源 vip3 基因, 命名为 vip3C, 其与 vip3A 具有显著序列差异。这 3 种 Bt 分离物被命名为 C536, C1674 和 AB727。

实施例 3: PCR 克隆全长 vip3C 基因

利用从 Bt 株系 C536 或 C1674 分离的总质粒 DNA 做为模板, 通过 PCR 获得 vip3C 基因的 3' 末端。所用的引物是:

Vip3CF4: 5'-GTTTAGAAGATTTTCAAACCATTAC-3' (SEQ ID NO: 20)

T7: 5'-TTAATACGACTCACTATAGGG-3' (SEQ ID NO: 21)

引物 T7 是非基因特异性引物, 其识别 vip3C 基因 3' 的侧翼核苷酸序列。

利用本领域的标准方法克隆和测序 PCR 产物。利用位于 vip3C 3'-和 5'-末端的 2 种引物:

Vip3Cc: 5'-TTTATTTAATAGAAACGTTTTCAAATGATATATG-3' (SEQ ID NO: 22)

Vip3Cn: 5'-CACCATGAACAAGAATAATACTAAATTAAGCACAAAGAG-3' (SEQ ID NO: 23)

通过 PCR 获得最终的全长 vip3C 基因。

获得 2 种全长的 vip3C 基因。来源于 Bt 分离物 C536 的 vip3C 基因

被命名为 *vip3C(a)*，从 C1674 分离的 *vip3C* 基因被命名为 *vip3C(b)*。*vip3C(a)* 和 *vip3C(b)* 在位置 2213 有 1 个核苷酸不同（参见 SEQ ID NO: 1），其中 *vip3C(a)* 在位置 2213 含有核苷酸“a”，因此编码在 SEQ ID NO: 2 位置 738 的氨基酸 Glu，其中 *vip3C(b)* 在位置 2213 含有核苷酸“g”，因此编码在 SEQ ID NO: 2 位置 738 的 Gly。

将 *vip3C(a)* 和 *vip3C(b)* 基因分别克隆到 pET 101/D-Topo 表达载体中，并且命名为 pNOV3911 和 pNOV3910，保藏在大肠杆菌 (*E. coli*) DH5 α 细胞中，并且分别给出保藏号为 NRRL B-30552 和 NRRL B-30553。

实施例 4: 粘粒克隆全长 *vip3Z* 基因

通过在 37°C 用 2mg/ml 溶菌酶处理重悬在 100 mM Tris pH 8, 10 mM EDTA 中的新鲜生长的细胞 30 分钟，从 AB727 分离总 DNA。在 1% SDS, 50mM EDTA, 1M 胍中加入蛋白酶 K 至终浓度 100 μ g/ml，在 55°C 温育。加入等体积的苯酚-氯仿-异戊醇。轻轻混合样品 5 分钟，以 3K 离心。重复该步骤 2 次。然后用 0.7 体积异丙醇混合水相，离心。用 70% 乙醇洗 DNA 沉淀 3 次，在 0.5 X TE 中轻轻重悬浮。在 37°C，在 100 μ l 体积中，用 0.3 单位 *Sau3A*/ μ gDNA 处理 12 μ g DNA。每 2 分钟间隔取样，持续 10 分钟。然后加入 1/10 体积 10 X TE，在 65°C 加热样品 30 分钟以灭活酶。对样品进行电泳以测定 40kb 范围的部分，并且在连接中利用该样品。

利用 *BamHI* 克隆位点，如供应商所述制备 SuperCos 粘粒载体 (Stratagene, La Jolla, CA)。在 5 μ l 体积中，以 2: 1 比率将 100ng/ml 制备的 SuperCos 和以前用 *Sau3A* 消化的 AB727 DNA 在 6°C 连接过夜。如供应商所述，用 Gigapack XL III (Stratagene) 包装连接混合物。如供应商所述，包装的噬菌体感染进入 XL-1MR 大肠杆菌 (*E. coli*) 细胞 (Stratagene)。在有 50 μ g/ml 卡那霉素的 L-琼脂上涂布粘粒文库，在 37°C 温育 16 小时。挑选和培养 200 个菌落用于筛选 *vip3Z* 基因的存在。

利用引物 Vip3ZA: 5'-GGCATTATGGATTTGCCACTGGTATC-3' (SEQ ID NO: 28) 和引物 Vip3ZB: 5'-TCCTTTGATACGCAGGTGT

AATTCAG-3' (SEQ ID NO: 29), 通过 PCR 筛选 200 个粘粒克隆的 *vip3Z* 基因的存在。

一个命名为 5g 的粘粒克隆显示含有编码 Vip3Z 蛋白质 (SEQ ID NO: 9) 的 *vip3Z* 基因 (SEQ ID NO: 8)。

实施例 5: 玉米优化的 *vip3C* 基因的构建

根据美国专利 5,625,136 中公开的方法制备玉米优化的 *vip3C* 基因, 这里并入这篇文献为参考文献。在该方法中, 利用了玉米偏爱的密码子, 即最常编码玉米中该氨基酸的单密码子。特定氨基酸的玉米偏爱密码子可从玉米的已知基因序列推知。在 Murray 等, *Nucleic Acids Research* 17: 477-498 (1989) 中可找到玉米植物 28 个基因的密码子使用。

制备编码 SEQ ID NO: 2 中所述氨基酸序列的合成的 *vip3C* (a) 和 *vip3C* (b) 基因。在 SEQ ID NO: 3 的位置 2213 和 2214, 合成的 *vip3C* (a) 基因分别包含核苷酸“a”和“g”, 编码在 SEQ ID NO: 2 位置 738 的氨基酸 Glu, 而合成的 *vip3C* (b) 基因分别包含核苷酸“g”和“a”, 编码在 SEQ ID NO: 2 位置 738 的氨基酸 Gly。将合成的 *vip3C* (a) 和 *vip3C* (b) 基因分别克隆到 pET101/D-Topo 表达载体中, 并且由此得到的载体命名为 pNOV3905, 保藏在大肠杆菌 (*E. coli*) BL21 细胞中, 并且给出保藏号为 NRRL B-30554, 和 pNOV3906, 保藏在大肠杆菌 (*E. coli*) BL21 细胞中, 并且给出保藏号为 NRRL B-30555。

实施例 6: Vip3C 蛋白质的生物测定

向 50mm 培养皿中倾入小地老虎饵料 (BioServ, Frenchtown, NJ)。冷却该饵料, 并用移液管转移 200 μ l 含有 pNOV3905, pNOV3906, pNOV3910 或 pNOV3911 的大肠杆菌 (*E. coli*) 细胞悬浮液到饵料表面。用细菌环均一地涂抹溶液以使悬浮液覆盖整个饵料的表面。使表面完全干燥。用细尖刷在饵料上放置下表列出的鳞翅目物种的第一龄期幼虫。分开检测每个物种。在幼虫侵染饵料后 3 天和 5 天记录幼虫死亡率及进食和生长抑制的出现。用含有没有表达载体的大肠杆菌 (*E. coli*) 细胞的样品作为阴性对照。为了比较的目的, 也可以在相同生物测定中检测

Vip3A 蛋白质，或对于该实施例，比较 Vip3C 数据和 Vip3A 的已知活性谱。

结果如表 8 中所示。昆虫侵染平板后第 5 天观察到杀虫活性。结果表明 Vip3C (a) (来源于 pNOV3911 和 3905) 和 Vip3C (b) (来源于 pNOV3910 和 3906) 比 Vip3A 毒素有更广谱的活性。检测也表明 Vip3 毒素不具有抗环境有益昆虫 *Danaus plexippus* 的活性。

表 8

检测的昆虫	% 昆虫死亡率		
	Vip3C (a)	Vip3C (b)	Vip3A ^b 的活性谱
小地老虎(<i>Agrotis ipsilon</i>)	100	100	+
谷实夜蛾(<i>Helicoverpa zea</i>)	75 ^a	75 ^a	+
烟芽夜蛾(<i>Heliothis virescens</i>)	80	50	+
贪夜蛾(<i>Spodoptera exigua</i>)	100	100	+
草地粘虫(<i>Spodoptera frugiperda</i>)	70 ^a	70 ^a	+
粉纹夜蛾(<i>Trichoplusia ni</i>)	100	100	+
棉红铃虫(<i>Pectinophora gossypiella</i>)	50 ^a	60 ^a	+
向日葵细卷叶蛾(<i>Cochylis hospes</i>)	90	90	+
向日葵同斑螟 (<i>Homoeosoma electellum</i>)	40 ^a	30 ^a	+
欧洲玉米螟 (<i>Ostrinia nubilalis</i>)	100	100	-
小菜蛾(<i>Plutella xylostella</i>)	100	100	-

^a 观察到存活的昆虫有严重的进食和生长抑制。

^b “+”表示对 Vip3A 敏感的昆虫物种。“-”表示对 Vip3A 几乎没有或没有敏感性的昆虫物种。

实施例 7: 产生含有 vip3C 基因的转基因玉米植物

选择优化的 vip3C 基因 (SEQ ID NO: 3) 用于转化到玉米植物中。将含有 vip3C (a) 序列的表达盒转移到用于农杆菌 (*Agrobacterium*) 介导的玉米转化的适合载体。对于该实施例，表达盒除了包含 vip3C (a)

基因外，还包含本领域已知的玉米泛素启动子和 nos 终止子，及用于转基因系筛选的磷酸甘露糖异构酶 (PMI) 基因(Negrotto 等, (2000) *Plant Cell Reports* 19: (798-803)。由此得到的载体被命名为 pNOV2149 (SEQ ID NO: 30)。

基本如 Negrotto 等, (2000) *Plant Cell Reports* 19: 798-803 所述进行未成熟玉米胚的转化。对于该实施例，所有培养基组分都如上文 Negrotto 等中所述。然而，可以替换为本领域已知的各种培养基组分。

在 28℃，在 YEP(酵母提取物(5g/L)，蛋白胨(10g/L)，NaCl(5g/L)，15g/L 琼脂，pH 6.8)固体培养基上培养包含植物转化质粒的农杆菌 (*Agrobacterium*) 株系 LBA4404(pSBI)2 - 4 天。在补充了 100μM As(Negrotto 等, (2000) *Plant Cell Rep* 19: 798-803)的 LS-inf 培养基中悬浮大约 0.8×10^9 农杆菌 (*Agrobacterium*)。在该培养基上预诱导细菌 30 - 60 分钟。

从 8 - 12 日龄的穗切下 A188 玉米基因型的未成熟胚，放入液体 LS-inf+100μM As 中。用新鲜感染培养基润洗胚一次。然后加入农杆菌 (*Agrobacterium*)溶液，涡旋胚 30 秒，使其和细菌共存 5 分钟。然后，将胚以盾片侧朝上的方向转移到 LSAs 培养基，在黑暗中培养 2 到 3 天。随后，按每个培养皿 20 到 25 个胚的方式，将胚转移到补充了头孢噻肟 (250mg/l) 和硝酸银 (1.6mg/l) 的 LSDc 培养基中，在 28℃，黑暗中培养约 10 天。

将产生胚性愈伤组织的未成熟胚转移到 LSD1M 0.5S 培养基。在该培养基上筛选培养物约 6 周，同时在 3 周时进行继代培养步骤。转移存活的愈伤组织到补充了甘露糖的 Reg1 培养基。在光中 (16 小时光/8 小时黑暗管理) 培养后，然后将绿色组织转移到没有生长调节物的 Reg2 培养基，培育约 1 - 2 周。将小植物转移到含有 Reg3 培养基的 Magenta GA-7 盒(Magenta Corp, Chicago Ill.)，在光下生长。约 2 - 3 周后，通过 PCR 检测植物中 PMI 基因和 *vip3C(a)*基因的存在。PCR 检测为阳性的植物转移到温室，检测对鳞翅目害虫的抗性。

实施例 8: 转基因玉米植物的分析

当从 Magenta GA-7 盒移植植物到土壤中时, 从植物取样。取样包括切下两小片叶片 (ca.2 - 4cm 长), 放置每片叶片在小培养皿中。阴性对照或者是来自于相同转化试验的 *vip3C(a)* 基因呈 PCR 阴性的转基因植物, 或者是在人工气候室中生长的 (与检测植物有相似大小的) 非转基因植物。

通过在每个叶片上放置 10 只第一龄幼虫, 用欧洲玉米螟 (*Ostrinia nubilalis*) 或草地粘虫 (*Spodoptera frugiperda*) 接种每个植物的叶片样品。然后, 严密地密封培养皿。

在接种约 3-4 天后, 收集数据。计算幼虫的百分死亡率, 同时确定叶片的可视损害等级。摄食损害分级为高级, 中级, 低级或没有, 分别给出了数值 3, 2, 1 或 0。

表 9 中所示结果表明, 包含 *vip3C(a)* 基因和表达 *Vip3C(a)* 蛋白质的转基因玉米植物对欧洲玉米螟 (ECB) (*Ostrinia nubilalis*) 和草地粘虫 (FAW) (*Spodoptera frugiperda*) 具有杀虫活性。

表 9

事件	植物号码	ECB 死亡率	损害等级	FAW 死亡率	损害等级
557	12A	80 %	2	100 %	0
557	20B	100 %	1	100 %	0
557	8A	80 %	1	70 %	0
557	11A	100 %	2	100 %	0
557	16B	95 %	2	95 %	0
557	18B	90 %	2	100 %	0
557	14B	100 %	1	100 %	0
556	1A	100 %	1	100 %	0
556	3B	80 %	1	100 %	0
556	4A	95 %	1	100 %	0
556	13A	100 %	1	100 %	0
A188	NEG	0	10	0 %	10

实施例 9: 杂合 *Vip3* 毒素

Vip3C 对欧洲玉米螟 (*Ostrinia nubilalis*) 和小菜蛾 (*Plutella xylostella*) 是有毒性的, 而同源 Vip3 毒素, 例如 Vip3A (a), Vip3A (b) 和 Vip3A (c) 是没有毒的。Vip3C 和 Vip3A 主要不同在于它们各自氨基酸序列的 C-末端区域, 特别是 SEQ ID NO: 2 氨基酸 661 到氨基酸 788 区域。为了证明 Vip3C 的 C-末端区域是 Vip3C 毒素中负责抗欧洲玉米螟 (*Ostrinia nubilalis*) 和小菜蛾 (*Plutella xylostella*) 活性的部分, 构建了一种杂合毒素, 其中包含 SEQ ID NO: 2 氨基酸 661 到氨基酸 788 的 Vip3C C-末端区域, 该区域按从氨基到羧基方向与 SEQ ID NO: 5 氨基酸 1 到氨基酸 660 的 Vip3A N-末端区域连接。命名该杂合毒素为 Vip3A-C。

用下面引物, 利用两步 PCR 构建编码 Vip3A-C 杂合毒素的核酸分子:

Vip3A-N: 5'-CACCATGAACAAGAATAATACTAAATTAAGCACAAGAG-3' (SEQ ID NO: 24)

Vip3A2050: 5'-TAAAGTTATCTCCCAAGCTTCATCTCCA-3' (SEQ ID NO: 25)

Vip3C-C1: 5'-AATGGAGATGAAGCTTGGGGAGAT-3' (SEQ ID NO: 26)

Vip3C-C2: 5'-TTTATTTAATAGAAACGTTTTCAAATGATATATG-3' (SEQ ID NO: 27)

在第一步 PCR 中, 用引物 Vip3A-N (SEQ ID NO: 24) 和 Vip3A2050 (SEQ ID NO: 25) 产生编码 N-末端区域的 *vip3A* 基因 5' 末端约 2.0kb 的片段, 用引物 Vip3C-C1 (SEQ ID NO: 26) 和 Vip3C-C2 (SEQ ID NO: 27) 产生编码 C-末端区域的 *vip3C* 基因 3' 末端约 0.4kb 的片段。在第二步 PCR 中, 联合这两个片段作为引物 Vip3A-N (SEQ ID NO: 24) 和 Vip3C-C2 (SEQ ID NO: 27) 的模板以产生约 2.4kb 杂合 *vip3A-vip3C* 基因, 命名为 *vip3A-C*。

制得杂合 *vip3A-vip3C* (b) 基因, 其序列如 SEQ ID NO: 10 中所述。将杂合 *vip3A-C* 基因克隆到 pET101D (Novagen) 中, 命名由此得到的载体为 pNOV3912, 并且将其转化到大肠杆菌 (*E.coli*) DH5 α 中进行表达。检测该大肠杆菌 (*E.coli*) 克隆 (NRRL B-30551) 抗表 10 中列出的昆虫物种的能力。Vip3C 蛋白质用作比较对照。数据与已知的 Vip3A 活性

谱比较。

表 10 中所示的结果证实，由 SEQ ID NO: 2 氨基酸 661 到氨基酸 788 的 Vip3C C-末端区域足以提供杂合毒素抗欧洲玉米螟 (*Ostrinia nubilalis*) 和小菜蛾 (*Plutella xylostella*) 的活性。

表 10

检测的昆虫	%昆虫死亡率		
	Vip3A-C	Vip3C (b) ^b	Vip3A ^c 的活性谱
小地老虎 (<i>Agrotis ipsilon</i>)	100	100	+
谷实夜蛾 (<i>Helicoverpa zea</i>)	100	75 ^a	+
烟芽夜蛾 (<i>Heliothis virescens</i>)	60	50	+
贪夜蛾 (<i>Spodoptera exigua</i>)	80	100	+
草地粘虫 (<i>Spodoptera frugiperda</i>)	70 ^a	70 ^a	+
粉纹夜蛾 (<i>Trichoplusia ni</i>)	80	100	+
棉红铃虫 (<i>Pectinophora gossypiella</i>)	80	60 ^a	+
向日葵细卷叶蛾 (<i>Cochylis hospes</i>)	100	90	+
向日葵同斑螟 (<i>Homoeosoma electellum</i>)	40 ^a	30 ^a	+
欧洲玉米螟 (<i>Ostrinia nubilalis</i>)	100	100	-
小菜蛾 (<i>Plutella xylostella</i>)	100	100	-

^a 观察到存活的昆虫有严重的进食和生长抑制。

^b 来源于实施例 6 的数据

^c “+”表示对 Vip3A 敏感的昆虫物种。“-”表示对 Vip3A 几乎没有或没有敏感性的昆虫物种。

实施例 10: 通过 DNA 改组的 *vip3* 基因的体外重组

通过 PCR 扩增本发明的 *vip3* 基因之一 (SEQ ID NO: 1, 3, 或 11)。基本如 Stemmer 等, *PNAS* 91: 10747-10751 (1994) 所述, 通过 DNase I 处理消化由此得到的 DNA 片段, 从反应混合物除去 PCR 引物。均如 Stemmer 等 (1994) 所述, 进行没有引物的 PCR 反应, 随后进行有引

物的PCR反应。将由此得到的DNA片段克隆进入pTRC99a(Pharmacia, 目录号: 27-5007-01), 利用 Biorad 基因脉冲发生器和制造商的条件, 通过电穿孔转化进入大肠杆菌 (*E.coli*) 株系 SASX38。在培养基上过夜培养转化的细菌, 筛选杀虫活性。

在相似的反应中, 使包含本文所述 *vip3* 基因之一 (SEQ ID NO: 1,3,5,7,9 或 11 或其突变体) 的 PCR 扩增的 DNA 片段和包含至少一个本文所述的其它 *vip3* 基因的 PCR 扩增的 DNA 片段 (或其突变体) 发生体外重组, 如下所述回收由此得到的具有改进杀虫特性的变异体。

为了增加改组 *vip3* 基因文库的多样性, 利用合成的寡核苷酸改组对一个或多个 *vip3* 基因 (称为初级基因) 进行改组。合成对应于至少一个多样性区域的多种 (例如, 2, 5, 10, 20, 50, 75, 或 100 或更多) 寡核苷酸。可以直接对这些寡核苷酸改组, 或可以与该核酸家族的一种或多种重组。

可以从称为次级基因的其他 *vip3* 基因获取寡核苷酸序列。次级基因与初级基因有一定程度的同源性。有几种选择用于寡核苷酸合成的次级基因部分的方法。例如, 可以随机选择次级基因的部分。DNA 改组方法将选择那些可以引入到改组基因中的寡核苷酸。

只要选定的部分适合合成, 它们可以是任何长度。也可以根据初级和次级基因之间的同源性设计寡核苷酸。一定程度的同源性对于改组期间必需发生在 DNA 片段间的交叉是必要的。同时, 强的异质性是改组基因文库的多样性所希望的。而且, 根据蛋白质序列知识和功能关系, 可以选择用于寡核苷酸合成的特异次级基因部分。

本发明公开了 Vip3 的 C-末端结构域部分地负责 Vip3 毒素的活性谱。当利用 DNA 改组技术, 通过本发明修饰杀虫谱时, 可以选择次级基因核苷酸序列的 C-末端区域作为靶区域用于合成寡核苷酸改组方法中应用的寡核苷酸。

因为 Vip3 蛋白质的杀虫活性至少部分地依赖于 N-末端区域, 可以选择次级基因的 N-末端区域进行寡核苷酸改组, 以增加杀虫活性。

一方面, 用几种根据次级 *vip3A* 基因序列合成的寡核苷酸改组初级

*vip3C(a)*和 *vip3C(b)*基因。*vip3C(a)*和 *vip3C(b)*是高度同源的，但是 *vip3A* 实质上不同于这些基因。因此，期望 *vip3A* 和 *vip3C(a)*和 *vip3C(b)* 一起改组以增加所得改组重组核酸的多样性。选择实质不同于 *vip3C(a)* 和 *vip3C(b)*相应部分的 *Vip3A* 序列部分，并且合成一系列覆盖这些部分的 50-mer 寡核苷酸。用 *vip3C(a)*和 *vip3C(b)*改组这些寡核苷酸。然后从改组的基因文库筛选一定数量的克隆，通过限制酶切作图检查多样性。考虑该多样性比从单独 *vip3C(a)*和 *vip3C(b)*的改组所正常预期的多样性要多。

实施例 11: 高通量筛选杀虫活性

筛选大肠杆菌 (*E.coli*) 或者苏云金芽孢杆菌 (*Bacillus thuringiensis*) 中改组 *vip3* 基因文库的杀虫活性。用 Q-bot (Beckman) 挑选菌落，放在标准 96 孔规格板中的生长培养基中，过夜生长。然后，以 96 孔规格板将每个克隆铺到昆虫饵料表面上，使表面干燥。可选地，向每个孔加入转化细胞集合以增加起始筛选循环中检测的克隆数。例如，每孔筛选 100 个克隆，并且利用 10,000 孔提供了 10^6 克隆筛选。

向每个孔加入几种新生靶昆虫幼虫，例如烟芽夜蛾 (*Heliothis virescens*)，谷实夜蛾 (*Helicoverpa zea*) 或草地粘虫 (*Spodoptera frugiperda*)。用透气膜包裹该板，以使幼虫保留在所放置的孔中。5 天后，评估各孔消耗饵料的量和/或昆虫死亡率。选择显示几乎没有或完全没有消耗饵料和/或观察到高的死亡率的孔中的克隆用于进一步分析。应该发现有增强的抗靶昆虫活性的几个克隆。

实施例 12: 粘粒克隆全长 *vip3C* 基因

通过在 37°C 用 2mg/ml 溶菌酶处理重悬在 100 mM Tris pH 8, 10 mM EDTA 中的新鲜生长的细胞 30 分钟，从 C1674 (NRRL B-30556) 分离总 DNA。在 1% SDS, 50mM EDTA, 1M 脲中加入蛋白酶 K 至终浓度 100µg/ml，在 55°C 温育。加入等体积的苯酚-氯仿-异戊醇。轻轻混匀样品 5 分钟，以 3K 离心。重复该步骤 2 次。然后用 0.7 体积异丙醇混合水相，离心。用 70% 乙醇洗 DNA 沉淀 3 次，在 0.5 X TE 中轻轻重悬

浮。在 37°C, 在 100 μ l 体积中, 用 0.3 单位 *Sau3A*/ μ g DNA 处理 12 μ gDNA。每 2 分钟间隔取样, 持续 10 分钟。然后加入 1/10 体积 10 X TE, 在 65°C 加热样品 30 分钟以灭活酶。对样品进行电泳以测定 40kb 范围的部分, 并且在连接中利用该样品。

利用 *BamHI* 克隆位点, 如供应商所述制备 SuperCos 粘粒载体 (Stratagene, La Jolla, CA)。在 5 μ l 体积中, 以 2: 1 比率将 100ng/ml 制备的 SuperCos 和以前用 *Sau3A* 消化的 C1674 DNA 在 6°C 连接过夜。如供应商所述, 用 Gigapack XL III (Stratagene) 包装连接混合物。如供应商所述, 将包装的噬菌体感染进入 XL-1MR 大肠杆菌 (*E.coli*) 细胞 (Stratagene)。在有 50 μ g/ml 卡那霉素的 L-琼脂上涂布粘粒文库, 在 37°C 温育 16 小时。挑选和培养 200 个菌落用于筛选 *vip3C* 基因的存在。

利用 *vip3C* 特异性引物, 通过 PCR 筛选 200 个粘粒克隆的 *vip3C* 基因的存在。

2 个粘粒克隆显示含有 *vip3C* 编码序列。几轮测序循环后, 证实该序列是 SEQ ID NO: 31 中所述序列。该 *vip3C* 编码序列被命名为 *vip3C-12168* 并且编码 Vip3C-12168 蛋白质 (SEQ ID NO: 32)。

实施例 13: Vip3C-12168 的生物测定

利用实施例 6 中所述的方法, 检测含有表达载体 (pTrcHis; Invitrogen) 的大肠杆菌 (*E.coli*) 细胞的生物活性, 所述表达载体含有 *vip3C-12168* 编码序列。检测的昆虫物种是欧洲玉米螟 (*Ostrinia nubilalis*) (ECB), 草地粘虫 (*Spodoptera frugiperda*) (FAW), 小地老虎 (*Agrotis ipsilon*) (BCW), 烟芽夜蛾 (*Heliothis virescens*) (TBW) 和谷实夜蛾 (*Helicoverpa zea*) (CEW)。在幼虫侵染饵料 7 天后记录幼虫死亡率及进食和生长抑制的出现。含有具有空表达载体 (pTrcHis) 的大肠杆菌 (*E.coli*) 细胞的样品做为阴性对照。为了比较活性谱的目的, 在相同生物测定中也检测表达 δ -内毒素 Cry1Ab 的大肠杆菌 (*E.coli*) 细胞和表达 Vip3A 蛋白质的大肠杆菌 (*E.coli*) 细胞。

结果如表 11 中所示。该数据表明 Vip3C-12168 具有和 Cry1Ab 与

Vip3A 之组合相同的活性谱。

表 11

处理	死亡率 %				
	ECB	FAW	BCW	TBW	CEW
Cry1Ab	100	0	10	0 ^a	8
Vip3A	0	100	100	83 ^b	100
Vip3C-12168	100	100	100	92 ^b	100
PTrcHis (空载体)	0	0	10	0	8

^a 生长抑制; ^b 进食抑制

实施例 14: 玉米优化的 Vip3C-12168

根据实施例 5 中所述的方法设计玉米优化的 *vip3C-12168* 编码序列。玉米优化的 *vip3C-12168* 编码序列的核苷酸序列如 SEQ ID NO: 33 中所述。

在本发明说明书中提到的所有出版物和专利申请都表明了本发明所属领域普通技术人员的技术水平。这里并入所有出版物和专利申请为参考文献, 其有如具体和单独地说明并入参考文献的每篇单独出版物或专利申请一样。

应该理解这里所述的实施例和实施方式仅仅是示例目的, 根据这些, 各种修饰或改变对本领域技术人员将是暗示性的, 并将包括在该申请精神和范围内及所附权利要求范围内。

<110> Syngenta Participations AG
Shen, Zhicheng
Warren, Gregory
Shotkoski, Frank
Kramer, Vance

<120> 新的 Vip3 毒素和应用的方法

<130> 60163PCT

<150> US 60/362250

<151> 2002-03-06

<160> 33

<170> PatentIn version 3.2

<210> 1

<211> 2367

<212> DNA

<213> 苏云金芽孢杆菌

<220>

<221> misc_feature

<222> (1)..(2367)

<223> 天然 Vip3C 编码序列。

第 2213 位的“r”代表核苷酸 g 或 a。

<400> 1
atgaacaaga ataactactaa attaagcaca agagccctac cgagttttat tgattatttt 60
aatggcattt atggatttgc cactggtatc aaagacatta tgaatatgat ttttaaacg 120
gatacaggtg gtaatctaac cttagacgaa atcctaaaga atcagcagtt actaaatgag 180
atctctggtg aattggatgg ggtaaattgg agcttaaattg atcttatcgc acagggaaac 240
ttaaatacag aattatctaa ggaaatctta aaaatcgcaa atgaacagaa tcaagtctta 300
aatgatgtta ataacaaact cgatgcgata aatcagatgc ttcatatata tctacctaaa 360
attacatcta tgtaagtga tgtaatgag caaaattatg cgctaagtct gcaaatagaa 420
tacttaagta agcaattgca agaaatttct gataaattag atattattaa cgtaaattgtt 480
cttattaact ctacacttac tgaaattaca cctgcatatc aacggattaa atatgtgaat 540
gaaaaatttg aagaattaac ttttgctaca gaaaccactt taaaagtaaa aaaggatagc 600
tcgcctgctg atattcttga tgagttaact gaattaactg aactagcgaa aagtgttaca 660
aaaaatgacg ttgatggttt tgaattttac cttatacat tccacgatgt aatggtagga 720
aataatttat tcgggcgttc agctttaaaa actgcttcag aattaattgc taaagaaaat 780
gtgaaaacaa gtggcagtga agtaggaaat gtttataatt tcttaattgt attaacagct 840
ctacaagcaa aagcttttct tactttaaca acatgccgaa aattattagg cttagcaggt 900
attgattata cttctattat gaatgaacat ttaaataagg aaaaagagga atttagagta 960

```

aacatccttc ctacactttc taatactttt tctaatocta attatgcaaa agttaaagga 1020
agtgatgaag atgcaaagat gattgtggaa gctaaaccag gacatgcatt ggttgggttt 1080
gaaatgagca atgattcaat cacagtatta aaagtatatg aggctaagct aaaacaaaat 1140
tatcaagttg ataaggattc cctatcggag gttatztatg gtgatacggg taaattattt 1200
tgtccagatc aatctgaaca aatatattat acaaataaca tagtattccc aaatgaatat 1260
gtaattacta aaattgattt cactaaaaaa atgaaaactt taagatatga ggtaacagcg 1320
aatttttatg attcttctac aggagaaatt gacttaaata agaaaaagt agaatcaagt 1380
gaagcggagt atagaacggt aagtgcta atgatggag tgtatatgcc attaggtgtc 1440
atcagtgaaa ctttttgac tccgataaat gggtttggcc tccaagctga tgaaaattca 1500
agattaatta ctttaacatg taaatcatat ttaagagaac tactgctagc aacagactta 1560
agcaataaag aaactaaatt gatcgtcca ccaagtgggt ttattagcaa tattgtagag 1620
aacgggtcca tagaagagga caatttagag ccgtggaaag caaataataa gaatgcgtat 1680
gtagatcata caggcggagt gaatggaact aaagctttat atgttcataa ggacggagga 1740
ttttcacaat ttattggaga taagttaaaa ccgaaaactg agtatgtaat ccaatatact 1800
gttaaaggaa aaccttctat tcatttaaaa gatgaaaata ctggatatat tcattatgaa 1860
gatacaaata ataatttaaa agattatcaa actattacta aacgttttac tacaggaact 1920
gatttaaagg gagtgtattt aatttttaaaa agtcaaaatg gagatgaagc ttggggagat 1980
aaatttacia ttttagaaat taagcctgag gaggatttat taagcccaga attaattaat 2040
ccgaattctt ggattacgac tccaggggct agcatttcag gaaataaact tttcattaac 2100
ttggggacia atgggacctt tagacaaagt ctttcattaa acagttattc aacttatagt 2160
ataagcttta ctgcatcagg accattta atgtgacgtaa gaaattctag ggragtatta 2220
tttgaacgaa gcaaccttat gtcttcaact agtcatatct ctgggacatt caaaactgaa 2280
tccaataata ccgattata tgtagaactt tcccgtcgtc ctggtggtgg tggatcatata 2340
tcatttgaaa acgtttctat taaataa 2367

```

```

<210> 2
<211> 788
<212> PRT
<213> 苏云金芽孢杆菌

```

```

<220>
<221> MISC_FEATURE
<222> (1)..(788)
<223> Vip3C 毒素

```

738 位的 Xaa 是氨基酸 Glu 或 Gly

```
<400> 2
```

Met Asn Lys Asn Asn Thr Lys Leu Ser Thr Arg Ala Leu Pro Ser Phe

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

Asn Phe Leu Ile Val Leu Thr Ala Leu Gln Ala Lys Ala Phe Leu Thr
 275 280 285

Leu Thr Thr Cys Arg Lys Leu Leu Gly Leu Ala Gly Ile Asp Tyr Thr
 290 295 300

Ser Ile Met Asn Glu His Leu Asn Lys Glu Lys Glu Glu Phe Arg Val
 305 310 315 320

Asn Ile Leu Pro Thr Leu Ser Asn Thr Phe Ser Asn Pro Asn Tyr Ala
 325 330 335

Lys Val Lys Gly Ser Asp Glu Asp Ala Lys Met Ile Val Glu Ala Lys
 340 345 350

Pro Gly His Ala Leu Val Gly Phe Glu Met Ser Asn Asp Ser Ile Thr
 355 360 365

Val Leu Lys Val Tyr Glu Ala Lys Leu Lys Gln Asn Tyr Gln Val Asp
 370 375 380

Lys Asp Ser Leu Ser Glu Val Ile Tyr Gly Asp Thr Asp Lys Leu Phe
 385 390 395 400

Cys Pro Asp Gln Ser Glu Gln Ile Tyr Tyr Thr Asn Asn Ile Val Phe
 405 410 415

Pro Asn Glu Tyr Val Ile Thr Lys Ile Asp Phe Thr Lys Lys Met Lys
 420 425 430

Thr Leu Arg Tyr Glu Val Thr Ala Asn Phe Tyr Asp Ser Ser Thr Gly
 435 440 445

Glu Ile Asp Leu Asn Lys Lys Lys Val Glu Ser Ser Glu Ala Glu Tyr
 450 455 460

Arg Thr Leu Ser Ala Asn Asp Asp Gly Val Tyr Met Pro Leu Gly Val
 465 470 475 480

Ile Ser Glu Thr Phe Leu Thr Pro Ile Asn Gly Phe Gly Leu Gln Ala
 485 490 495

Asp Glu Asn Ser Arg Leu Ile Thr Leu Thr Cys Lys Ser Tyr Leu Arg
 500 505 510

Glu Leu Leu Leu Ala Thr Asp Leu Ser Asn Lys Glu Thr Lys Leu Ile
 515 520 525

Val Pro Pro Ser Gly Phe Ile Ser Asn Ile Val Glu Asn Gly Ser Ile
 530 535 540

Glu Glu Asp Asn Leu Glu Pro Trp Lys Ala Asn Asn Lys Asn Ala Tyr
 545 550 555 560

Val Asp His Thr Gly Gly Val Asn Gly Thr Lys Ala Leu Tyr Val His
 565 570 575

Lys Asp Gly Gly Phe Ser Gln Phe Ile Gly Asp Lys Leu Lys Pro Lys
 580 585 590

Thr Glu Tyr Val Ile Gln Tyr Thr Val Lys Gly Lys Pro Ser Ile His
 595 600 605

Leu Lys Asp Glu Asn Thr Gly Tyr Ile His Tyr Glu Asp Thr Asn Asn
 610 615 620

Asn Leu Lys Asp Tyr Gln Thr Ile Thr Lys Arg Phe Thr Thr Gly Thr
 625 630 635 640

Asp Leu Lys Gly Val Tyr Leu Ile Leu Lys Ser Gln Asn Gly Asp Glu
 645 650 655

Ala Trp Gly Asp Lys Phe Thr Ile Leu Glu Ile Lys Pro Ala Glu Asp
 660 665 670

Leu Leu Ser Pro Glu Leu Ile Asn Pro Asn Ser Trp Ile Thr Thr Pro
 675 680 685

Gly Ala Ser Ile Ser Gly Asn Lys Leu Phe Ile Asn Leu Gly Thr Asn
 690 695 700

Gly Thr Phe Arg Gln Ser Leu Ser Leu Asn Ser Tyr Ser Thr Tyr Ser
 705 710 715 720

Ile Ser Phe Thr Ala Ser Gly Pro Phe Asn Val Thr Val Arg Asn Ser
 725 730 735

Arg Xaa Val Leu Phe Glu Arg Ser Asn Leu Met Ser Ser Thr Ser His
 740 745 750

Ile Ser Gly Thr Phe Lys Thr Glu Ser Asn Asn Thr Gly Leu Tyr Val
 755 760 765

Glu Leu Ser Arg Arg Ser Gly Gly Gly Gly His Ile Ser Phe Glu Asn
 770 775 780

Val Ser Ile Lys
785

<210> 3
<211> 2367
<212> DNA
<213> 人工序列

<220>
<223> 玉米优化的 vip3C 编码序列.

第 2213 和 2214 位的 “r” 代表核苷酸 g 或 a.

<400> 3
atgaacaaga acaacaccaa gctctccacc cgcgcctcc cgctcttcat cgactacttc 60
aacggcatct acggcttcgc caccggcatc aaggacatca tgaacatgat cttcaagacc 120
gacaccggcg gcaacctcac cctcgacgag atcctcaaga accagcagct cctcaacgag 180
atcagcggca agctcgacgg cgtgaacggc tccctcaacg acctcatcgc ccagggcaac 240
ctcaacaccg agctgtccaa ggagatcctc aagatcgcca acgagcagaa ccaggtgctc 300
aacgacgtga acaacaagct cgacgccatc aacaccatgc tccacatcta cctcccgaag 360
atcacctcca tgctctccga cgtgatgaag cagaactacg ccctctccct ccagatcgag 420
tacctctcca agcagctcca ggagatcagc gacaagctcg acatcatcaa cgtgaacgtg 480
ctcatcaact ccacctcac cgagatcacc cgggectacc agcgcacaa gtacgtgaac 540
gagaagttcg aggagctgac cttcgccacc gagaccacc tcaaggtgaa gaaggactcc 600
tccccggccg acatcctcga cgagctgacc gagctgaccg agctggccaa gtccgtgacc 660
aagaacgacg tggacggctt cgagttctac ctcaacacct tccacgacgt gatggtgggc 720
aacaacctct tcggccgctc cgccctcaag accgcctccg agctgatcgc caaggagaac 780
gtgaagacct cggctccga ggtgggcaac gtgtacaact tctcatcgt gtcaccgcc 840
ctgcaggcca aggccttct caccctcac acctgccga agctcctcg cctcgccggc 900
atcgactaca cctccatcat gaacgagcac ctcaacaagg agaaggagga gttccgcgtg 960
aacatcctcc cgaccctctc caacacctc tccaaccga actacgcaa ggtgaagggc 1020
tccgacgagg acgccaagat gatcgtggag gccaagccgg gccacgccct cgtgggcttc 1080
gagatgtcca acgactccat caccgtgctc aaggtgtacg aggccaagct caagcagaac 1140
taccaggtgg acaaggactc cctctccgag gtgatctacg gcgacaccga caagctcttc 1200
tgcccgacc agtccgagca gatatactac accaacaaca tcgtgttccc gaacgagtac 1260
gtgatcacca agatcgactt caccaagaag atgaagacc tccgctacga ggtgaccgcc 1320
aacttctacg actcctccac cggcgagatc gacctcaaca agaagaaggt ggagtccctc 1380
gaggccgagt accgcaccct ctccgccaac gacgacggcg tgtacatgcc gctcggcgtg 1440
atctccgaaa ccttccctcac cccgatcaac ggcttcggcc tccaggccga cgagaactcc 1500


```

cgctcatca ccctcacctg caagtcctac ctccgcgagc tgctcctcgc caccgacctc 1560
tccaacaagg agaccaagct catcgtgccg ccgtccgget tcatctccaa catcgtggag 1620
aacggctcca tcgaggagga caacctcgag ccgtggaagg ccaacaacaa gaacgcctac 1680
gtggaccaca ccggcggcgt gaacggcacc aaggccctct acgtgcacaa ggacggcggc 1740
ttctcccagt tcatcggcga caagctcaag ccgaagaccg agtacgtgat ccagtacacc 1800
gtgaagggca agccgtccat ccacctcaag gacgagaaca ccggctacat ccactacgag 1860
gacaccaaca acaacctcaa ggactaccag accatcacca agcgccttac caccggcacc 1920
gacctcaagg gcgtgtacct catcctcaag tcccagaacg gcgacgaggc ctggggcgac 1980
aagttcacca tccttgagat caagccggcc gaggacctcc tctccccgga gctgatcaac 2040
ccgaactcct ggatcaccac ccggggcgcc tccatctccg gcaacaagct cttcatcaac 2100
ctcggcacca acggcacctt ccgccagtc ctctccctca actcctactc caactactcc 2160
atctccttca ccgcctccgg ccggtcaac gtgaccgtgc gcaactcccg cgrrgtgctc 2220
ttcgagcgct ccaacctcat gtcctccacc tcccacatct ccggcacctt caagaccgag 2280
tccaacaaca ccggcctcta cgtggagctg tcccgcgct ccggcggcgg cggccacatc 2340
tccttcgaga acgtgtccat caagtag 2367

```

```

<210> 4
<211> 2370
<212> DNA
<213> 苏云金芽孢杆菌

```

```

<220>
<221> misc_feature
<222> (1)..(2370)
<223> vip3A(a) 天然编码序列.

```

```

<400> 4
atgaacaaga ataatactaa attaagcaca agagccttac caagttttat tgattatfff 60
aatggcattt atggatttgc cactggatc aaagacatta tgaacatgat ttttaaacg 120
gatacaggtg gtgatctaac cctagacgaa attttaaaga atcagcagtt actaaatgat 180
atftctggta aattggatgg ggtgaatgga agcttaaattg atcttatcgc acagggaaac 240
ttaaatacag aattatctaa ggaaatatta aaaattgcaa atgaacaaaa tcaagtttta 300
aatgatgtta ataacaaact cgatgcgata aatacgatgc ttcgggtata tctacctaaa 360
attacctcta tgttgagtga tgtaatgaaa caaaattatg cgctaagtct gcaaatagaa 420
tacttaagta aacaattgca agagatttct gataagttgg atattattaa tgtaaatgta 480
cttattaact ctacacttac tgaaattaca cctgcgtatc aaaggattaa atatgtgaac 540
gaaaaatttg aggaattaac ttttgctaca gaaactagtt caaaagtaaa aaaggatggc 600
tctcctgcag atattottga tgagttaact gagttaactg aactagcgaa aagtgttaaca 660

```

```

aaaaatgatg tggatggttt tgaattttac cttatacat tccacgatgt aatggtagga      720
aataatttat tcgggcgttc agctttaaaa actgcatcgg aattaattac taaagaaaat      780
gtgaaaacaa gtggcagtgga ggtcggaaat gtttataact tcttaattgt attaacagct      840
ctgcaagcaa aagcttttct tactttaaca acatgccgaa aattattagg cttagcagat      900
attgattata cttctattat gaatgaacat ttaaataagg aaaaagagga atttagagta      960
aacatcctcc ctacactttc taatactttt tctaateccta attatgcaaa agttaaagga     1020
agtgatgaag atgcaaagat gattgtggaa gctaaaccag gacatgcatt gattgggttt     1080
gaaattagta atgattcaat tacagtatta aaagtatatg aggctaagct aaaacaaaat     1140
tatcaagtcg ataaggattc ctfatcggaa gttatttatg gtgatatgga taaattattg     1200
tgcccagatc aatctgaaca aatctattat acaataaca tagtatttcc aatgaatat     1260
gtaattacta aaattgattt cactaaaaaa atgaaaactt taagatatga ggtaacagcg     1320
aatttttatg attcttctac aggagaaatt gacttaaata agaaaaaagt agaatcaagt     1380
gaagcggagt atagaacggt aagtgctaata gatgatgggg tgtatatgcc gttaggtgtc     1440
atcagtgaaa catttttgac tccgattaat gggtttgcc tccaagctga tgaaaattca     1500
agattaatta ctttaacatg taaatcatat ttaagagaac tactgctagc aacagactta     1560
agcaataaag aaactaaatt gatcgtcccg ccaagtgggt ttattagcaa tattgtagag     1620
aacgggtcca tagaagagga caatttagag ccgtggaaag caaataataa gaatgcgtat     1680
gtagatcata caggcggagt gaatggaact aaagctttat atgttcataa ggacggagga     1740
atrtcacaat ttattggaga taagttaaaa ccgaaaactg agtatgtaat ccaatatact     1800
gttaaaggaa aaccttctat tcatttaaaa gatgaaaata ctggatatat tcattatgaa     1860
gatacaaata ataatttaga agattatcaa actattaata aacgttttac tacaggaact     1920
gatttaaagg gagtgtatth aattttaaaa agtcaaaatg gagatgaagc ttggggagat     1980
aactttatta ttttggaat tagtccttct gaaaagttat taagtccaga attaatat     2040
acaaataatt ggacgagtac gggatcaact aatattagcg gtaatacact cactctttat     2100
cagggaggac gagggattct aaaacaaaac cttcaattag atagtttttc aacttataga     2160
gtgtatthtt ctgtgtccgg agatgctaata gtaaggatta gaaattctag ggaagtgtta     2220
tttgaaaaaa gatatatgag cggtgctaaa gatgtttctg aatgttcac tacaaaattt     2280
gagaaagata acttttatat agagctttct caaggaata atttatatgg tggctctatt     2340
gtacatthtt acgatgtctc tattaagtaa                                     2370

```

```

<210> 5
<211> 789
<212> PRT
<213> 苏云金芽孢杆菌

```

```

<220>
<221> MISC_FEATURE
<222> (1)..(789)
<223> Vip3A 毒素

<400> 5

Met Asn Lys Asn Asn Thr Lys Leu Ser Thr Arg Ala Leu Pro Ser Phe
1          5          10          15

Ile Asp Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp
          20          25          30

Ile Met Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Asp Leu Thr Leu
          35          40          45

Asp Glu Ile Leu Lys Asn Gln Gln Leu Leu Asn Asp Ile Ser Gly Lys
          50          55          60

Leu Asp Gly Val Asn Gly Ser Leu Asn Asp Leu Ile Ala Gln Gly Asn
65          70          75          80

Leu Asn Thr Glu Leu Ser Lys Glu Ile Leu Lys Ile Ala Asn Glu Gln
          85          90          95

Asn Gln Val Leu Asn Asp Val Asn Asn Lys Leu Asp Ala Ile Asn Thr
          100          105          110

Met Leu Arg Val Tyr Leu Pro Lys Ile Thr Ser Met Leu Ser Asp Val
          115          120          125

Met Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys
          130          135          140

Gln Leu Gln Glu Ile Ser Asp Lys Leu Asp Ile Ile Asn Val Asn Val
145          150          155          160

Leu Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile
          165          170          175

Lys Tyr Val Asn Glu Lys Phe Glu Glu Leu Thr Phe Ala Thr Glu Thr
          180          185          190

Ser Ser Lys Val Lys Lys Asp Gly Ser Pro Ala Asp Ile Leu Asp Glu
          195          200          205

Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp Val
          210          215          220

Asp Gly Phe Glu Phe Tyr Leu Asn Thr Phe His Asp Val Met Val Gly

```

225		230		235		240									
Asn	Asn	Leu	Phe	Gly	Arg	Ser	Ala	Leu	Lys	Thr	Ala	Ser	Glu	Leu	Ile
				245					250					255	
Thr	Lys	Glu	Asn	Val	Lys	Thr	Ser	Gly	Ser	Glu	Val	Gly	Asn	Val	Tyr
			260					265					270		
Asn	Phe	Leu	Ile	Val	Leu	Thr	Ala	Leu	Gln	Ala	Lys	Ala	Phe	Leu	Thr
		275					280					285			
Leu	Thr	Thr	Cys	Arg	Lys	Leu	Leu	Gly	Leu	Ala	Asp	Ile	Asp	Tyr	Thr
	290					295					300				
Ser	Ile	Met	Asn	Glu	His	Leu	Asn	Lys	Glu	Lys	Glu	Glu	Phe	Arg	Val
305					310					315					320
Asn	Ile	Leu	Pro	Thr	Leu	Ser	Asn	Thr	Phe	Ser	Asn	Pro	Asn	Tyr	Ala
				325					330					335	
Lys	Val	Lys	Gly	Ser	Asp	Glu	Asp	Ala	Lys	Met	Ile	Val	Glu	Ala	Lys
			340					345					350		
Pro	Gly	His	Ala	Leu	Ile	Gly	Phe	Glu	Ile	Ser	Asn	Asp	Ser	Ile	Thr
		355					360					365			
Val	Leu	Lys	Val	Tyr	Glu	Ala	Lys	Leu	Lys	Gln	Asn	Tyr	Gln	Val	Asp
	370					375					380				
Lys	Asp	Ser	Leu	Ser	Glu	Val	Ile	Tyr	Gly	Asp	Met	Asp	Lys	Leu	Leu
385					390					395					400
Cys	Pro	Asp	Gln	Ser	Glu	Gln	Ile	Tyr	Tyr	Thr	Asn	Asn	Ile	Val	Phe
				405					410					415	
Pro	Asn	Glu	Tyr	Val	Ile	Thr	Lys	Ile	Asp	Phe	Thr	Lys	Lys	Met	Lys
			420					425						430	
Thr	Leu	Arg	Tyr	Glu	Val	Thr	Ala	Asn	Phe	Tyr	Asp	Ser	Ser	Thr	Gly
		435					440					445			
Glu	Ile	Asp	Leu	Asn	Lys	Lys	Lys	Val	Glu	Ser	Ser	Glu	Ala	Glu	Tyr
	450					455					460				
Arg	Thr	Leu	Ser	Ala	Asn	Asp	Asp	Gly	Val	Tyr	Met	Pro	Leu	Gly	Val
465					470					475					480
Ile	Ser	Glu	Thr	Phe	Leu	Thr	Pro	Ile	Asn	Gly	Phe	Gly	Leu	Gln	Ala
				485					490					495	

Asp Glu Asn Ser Arg Leu Ile Thr Leu Thr Cys Lys Ser Tyr Leu Arg
 500 505 510

Glu Leu Leu Leu Ala Thr Asp Leu Ser Asn Lys Glu Thr Lys Leu Ile
 515 520 525

Val Pro Pro Ser Gly Phe Ile Ser Asn Ile Val Glu Asn Gly Ser Ile
 530 535 540

Glu Glu Asp Asn Leu Glu Pro Trp Lys Ala Asn Asn Lys Asn Ala Tyr
 545 550 555 560

Val Asp His Thr Gly Gly Val Asn Gly Thr Lys Ala Leu Tyr Val His
 565 570 575

Lys Asp Gly Gly Ile Ser Gln Phe Ile Gly Asp Lys Leu Lys Pro Lys
 580 585 590

Thr Glu Tyr Val Ile Gln Tyr Thr Val Lys Gly Lys Pro Ser Ile His
 595 600 605

Leu Lys Asp Glu Asn Thr Gly Tyr Ile His Tyr Glu Asp Thr Asn Asn
 610 615 620

Asn Leu Glu Asp Tyr Gln Thr Ile Asn Lys Arg Phe Thr Thr Gly Thr
 625 630 635 640

Asp Leu Lys Gly Val Tyr Leu Ile Leu Lys Ser Gln Asn Gly Asp Glu
 645 650 655

Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Ile Ser Pro Ser Glu Lys
 660 665 670

Leu Leu Ser Pro Glu Leu Ile Asn Thr Asn Asn Trp Thr Ser Thr Gly
 675 680 685

Ser Thr Asn Ile Ser Gly Asn Thr Leu Thr Leu Tyr Gln Gly Gly Arg
 690 695 700

Gly Ile Leu Lys Gln Asn Leu Gln Leu Asp Ser Phe Ser Thr Tyr Arg
 705 710 715 720

Val Tyr Phe Ser Val Ser Gly Asp Ala Asn Val Arg Ile Arg Asn Ser
 725 730 735

Arg Glu Val Leu Phe Glu Lys Arg Tyr Met Ser Gly Ala Lys Asp Val
 740 745 750

Ser Glu Met Phe Thr Thr Lys Phe Glu Lys Asp Asn Phe Tyr Ile Glu
755 760 765

Leu Ser Gln Gly Asn Asn Leu Tyr Gly Gly Pro Ile Val His Phe Tyr
770 775 780

Asp Val Ser Ile Lys
785

<210> 6
<211> 2364
<212> DNA
<213> 苏云金芽孢杆菌

<220>
<221> misc_feature
<222> (1)..(2364)
<223> vip3B 天然编码序列.

<400> 6
atgaacaaga ataatactaa attaaacgca agggccttac cgagttttat tgattatfff 60
aatggcattt atggatttgc cactggtatc aaagacatta tgaacatgat ttttaaacg 120
gatacaggtg gaaatctaac cctagacgaa attttaaaaa atcagcagtt attaatgag 180
atctctggta aattggatgg ggtaaatggg agcttaaacg atcttatcgc acagggaaac 240
ttaaatacag aattatctaa ggaaatctta aaaattgcaa atgagcagaa tcaagtctta 300
aatgatgtta ataacaaact taatgcgata aatacaatgc ttcacatata tctacctaaa 360
attacatcta tgttaaataga tgtaatagaaa caaaattatg cactaagtct gcaaatagaa 420
tacctaagta aacaattgca agaaatttcc gacaagttag atgtcattaa cgtgaatgta 480
cttattaact ctacacttac tgaaattaca cctgcgatc aacggatgaa atatgtaaat 540
gaaaaatttg aagatttaac ttttgctaca gaaaccactt taaaagtaaa aaagaatagc 600
tcccctgcag atattcttga tgagtttaact gagtttaactg aactagcgaa aagtgtaaca 660
aaaaatgacg tggatggttt tgaattttac cttaatacat tccacgatgt aatggtagga 720
aacaatttat tggggcgttc agcttttaaaa actgcttcgg aattaatcgc taaagaaaat 780
gtgaaaacaa gtggcagtga ggtaggaaat gtttataatt tcttaattgt attaacagct 840
ctgcaagcaa aagcttttct tactttaaca acatgccgga aattattagg cttagcagat 900
attgattata ctttcattat gaatgaacat ttagataagg aaaaagagga atttagagta 960
aatatccttc ctacactttc taatactttt tctaactcta actatgcaa agctaaagga 1020
agcaatgaag atgcaaagat aattgtggaa gctaaaccag gatatgcttt ggttggattt 1080
gaaatgagca atgattcaat cacagtatta aaagcatatc aggctaagct aaaacaagat 1140
tatcaagttg ataaagattc gttatcagaa attgtctatg gtgatatgga taaattattg 1200

```

tgcccgatc aatctgaaca aatatattat acaaataaca ttgcttttcc caatgaatat 1260
gtaattacta aaattacttt tactaaaaaa atgaatagtt taagatatga ggcaacagct 1320
aatttttatg attcttctac aggggatatt gatctaaata agacaaaagt agaatcaagt 1380
gaagcagagt atagtacgct aagtgctagt actgatggag tctatatgcc gttaggtatt 1440
atcagtghaa catttttgac tccaattaat gggtttgaa tcgtagtcga tghaaaattca 1500
aaattagtaa atttaacatg taaatcatat ttaagagagg tattattagc aacagactta 1560
agtaataaag aaactaaatt gattgtccca cctattgggt ttattagcaa tattgtagaa 1620
aatgggaact tagagggaga aaacttagag ccgtggaaag caaataacaa aatgcggtat 1680
gtagatcata caggcggcgt aatggaact aaagctttat atgttcataa ggatggtgag 1740
ttttcacaat ttattggaga taagttgaaa tcgaaaacag aatatgtaat tcaatatatt 1800
gtaaagggaa aagcttctat tcttttgaaa gatgaaaaaa atggtgattg catttatgaa 1860
gatacaaaata atggtttaga agattttcaa accattacta aaagttttat tacaggaacg 1920
gattcttcag gagttcattt aatatttaat agtcaaaatg gcgatgaagc atttggggaa 1980
aactttacta tttcagaaat taggctttcc gaagatttat taagtccaga attgataaat 2040
tcagatgctt gggttggatc tcagggaact tggatctcag gaaattcact cactattaat 2100
agtaatgtga atggaacttt tcgacaaaac ctttcgtag aaagctattc aacttatagt 2160
atgaacttta atgtgaatgg atttgccaaag gtgacagtaa gaaattcccg tgaagtatta 2220
tttghaaaaa attatccgca gctttcacct aaagatattt ctghaaaatt cacaaactgca 2280
gccaaataata ccgggttgta tgtagagctt tctcgtttta catcgggtgg cgctataaat 2340
ttccggaatt tttcgattaa gtga 2364

```

```

<210> 7
<211> 787
<212> PRT
<213> 苏云金芽孢杆菌

```

```

<220>
<221> MISC_FEATURE
<222> (1)..(787)
<223> Vip3B 毒素

```

```

<400> 7

```

```

Met Asn Lys Asn Asn Thr Lys Leu Asn Ala Arg Ala Leu Pro Ser Phe
1          5          10          15

```

```

Ile Asp Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp
20          25          30

```

```

Ile Met Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Asn Leu Thr Leu
35          40          45

```

Asp Glu Ile Leu Lys Asn Gln Gln Leu Leu Asn Glu Ile Ser Gly Lys
 50 55 60

Leu Asp Gly Val Asn Gly Ser Leu Asn Asp Leu Ile Ala Gln Gly Asn
 65 70 75 80

Leu Asn Thr Glu Leu Ser Lys Glu Ile Leu Lys Ile Ala Asn Glu Gln
 85 90 95

Asn Gln Val Leu Asn Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Thr
 100 105 110

Met Leu His Ile Tyr Leu Pro Lys Ile Thr Ser Met Leu Asn Asp Val
 115 120 125

Met Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys
 130 135 140

Gln Leu Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Val Asn Val
 145 150 155 160

Leu Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Met
 165 170 175

Lys Tyr Val Asn Glu Lys Phe Glu Asp Leu Thr Phe Ala Thr Glu Thr
 180 185 190

Thr Leu Lys Val Lys Lys Asn Ser Ser Pro Ala Asp Ile Leu Asp Glu
 195 200 205

Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp Val
 210 215 220

Asp Gly Phe Glu Phe Tyr Leu Asn Thr Phe His Asp Val Met Val Gly
 225 230 235 240

Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ser Glu Leu Ile
 245 250 255

Ala Lys Glu Asn Val Lys Thr Ser Gly Ser Glu Val Gly Asn Val Tyr
 260 265 270

Asn Phe Leu Ile Val Leu Thr Ala Leu Gln Ala Lys Ala Phe Leu Thr
 275 280 285

Leu Thr Thr Cys Arg Lys Leu Leu Gly Leu Ala Asp Ile Asp Tyr Thr
 290 295 300

Phe Ile Met Asn Glu His Leu Asp Lys Glu Lys Glu Glu Phe Arg Val
 305 310 315 320

Asn Ile Leu Pro Thr Leu Ser Asn Thr Phe Ser Asn Pro Asn Tyr Ala
 325 330 335

Lys Ala Lys Gly Ser Asn Glu Asp Ala Lys Ile Ile Val Glu Ala Lys
 340 345 350

Pro Gly Tyr Ala Leu Val Gly Phe Glu Met Ser Asn Asp Ser Ile Thr
 355 360 365

Val Leu Lys Ala Tyr Gln Ala Lys Leu Lys Gln Asp Tyr Gln Val Asp
 370 375 380

Lys Asp Ser Leu Ser Glu Ile Val Tyr Gly Asp Met Asp Lys Leu Leu
 385 390 395 400

Cys Pro Asp Gln Ser Glu Gln Ile Tyr Tyr Thr Asn Asn Ile Ala Phe
 405 410 415

Pro Asn Glu Tyr Val Ile Thr Lys Ile Thr Phe Thr Lys Lys Met Asn
 420 425 430

Ser Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Ser Ser Thr Gly
 435 440 445

Asp Ile Asp Leu Asn Lys Thr Lys Val Glu Ser Ser Glu Ala Glu Tyr
 450 455 460

Ser Thr Leu Ser Ala Ser Thr Asp Gly Val Tyr Met Pro Leu Gly Ile
 465 470 475 480

Ile Ser Glu Thr Phe Leu Thr Pro Ile Asn Gly Phe Gly Ile Val Val
 485 490 495

Asp Glu Asn Ser Lys Leu Val Asn Leu Thr Cys Lys Ser Tyr Leu Arg
 500 505 510

Glu Val Leu Leu Ala Thr Asp Leu Ser Asn Lys Glu Thr Lys Leu Ile
 515 520 525

Val Pro Pro Ile Gly Phe Ile Ser Asn Ile Val Glu Asn Gly Asn Leu
 530 535 540

Glu Gly Glu Asn Leu Glu Pro Trp Lys Ala Asn Asn Lys Asn Ala Tyr
 545 550 555 560

Val Asp His Thr Gly Gly Val Asn Gly Thr Lys Ala Leu Tyr Val His

565	570	575
Lys Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys Leu Lys Ser Lys 580	585	590
Thr Glu Tyr Val Ile Gln Tyr Ile Val Lys Gly Lys Ala Ser Ile Leu 595	600	605
Leu Lys Asp Glu Lys Asn Gly Asp Cys Ile Tyr Glu Asp Thr Asn Asn 610	615	620
Gly Leu Glu Asp Phe Gln Thr Ile Thr Lys Ser Phe Ile Thr Gly Thr 625	630	635
Asp Ser Ser Gly Val His Leu Ile Phe Asn Ser Gln Asn Gly Asp Glu 645	650	655
Ala Phe Gly Glu Asn Phe Thr Ile Ser Glu Ile Arg Leu Ser Glu Asp 660	665	670
Leu Leu Ser Pro Glu Leu Ile Asn Ser Asp Ala Trp Val Gly Ser Gln 675	680	685
Gly Thr Trp Ile Ser Gly Asn Ser Leu Thr Ile Asn Ser Asn Val Asn 690	695	700
Gly Thr Phe Arg Gln Asn Leu Ser Leu Glu Ser Tyr Ser Thr Tyr Ser 705	710	715
Met Asn Phe Asn Val Asn Gly Phe Ala Lys Val Thr Val Arg Asn Ser 725	730	735
Arg Glu Val Leu Phe Glu Lys Asn Tyr Pro Gln Leu Ser Pro Lys Asp 740	745	750
Ile Ser Glu Lys Phe Thr Thr Ala Ala Asn Asn Thr Gly Leu Tyr Val 755	760	765
Glu Leu Ser Arg Phe Thr Ser Gly Gly Ala Ile Asn Phe Arg Asn Phe 770	775	780
Ser Ile Lys 785		

<210> 8
 <211> 2407
 <212> DNA
 <213> 苏云金芽孢杆菌

<220>
 <221> misc_feature
 <222> (1)..(2406)
 <223> vip3Z 天然编码序列.

<400> 8
 atgaataata ctaagttaaa cgcaagggct ttaccaagtt ttattgatta ttttaatggc 60
 atttatggat ttgccactgg tatcaaagac attatgaaca tgatttttaa aacggataca 120
 ggtggtggta atttaacact agatgaaatt ttaaagaatc aagatttatt aaatcaaact 180
 tcagataaac tcgatggaat taatggagat ttaggtgatc ttattgcaca aggcaattta 240
 aattcagaac taactaagga attattaana attgcgaatg agcagaatct gatgttaaat 300
 aatgttaatg ctcaacttaa ttcaataaat tcaacactta acacctatct gccaaaaatt 360
 acatctatgc taagtgaggt aatgaaaca aactatgat taagtctaca aatagaatth 420
 cttagtgaac aattacaaga aatcagat aaacttgatg ttatcaatth aatgtatta 480
 attaactcta cattgacaga aattacgcct gcatacaac gtattaaata tgtaaatgat 540
 aaatttgatg aattgacttc tactgtggaa aaaaatccga aaattaatca agataatth 600
 actgaagatg ttattgataa ttaactgat ttaactgaac tagcacgaag tgtaacgaga 660
 aatgatatgg atagtthtga atthtatatt aaaactthcc atgatgtgat gataggaaat 720
 aatthattca gtcgttctgc attaaaaaact gcttcagaat taattgctaa ggaaaatata 780
 catactatgg gaagtgaat tggtaatgtc tacactthta tggttgthtt gacttcctta 840
 caagcaaaag cgttcctaac ttaactgca tgccgtaaat tattaggatt aacagatath 900
 gattatacac aaattatgaa tgaaaatth aatagagaaa aagaggaatt tcgcttaaat 960
 attcttccta cactthctaa tgattthttct aatcctaatt atacagaaac tthaggaagt 1020
 gatctttag atcctattgt tacgttagaa gctgatcctg gttatgctth aataggtthtt 1080
 gagattctca atgatccact tccagtatta aaagtatath aggcaaagct aaaaccaaat 1140
 tatcaagtcg acaagagtc gattatggaa aatathtatg gaaatatcca caaactactt 1200
 tgtccaaaac aacgtcacca aaaatathat ataaaagaca ttacaththcc tgaaggtht 1260
 gtaatcacca aaattgthtt tgaaaaaaa ttgaatctat taggatatga agtaacagca 1320
 aatctthtatg acccaththac aggaagtath gattthgata agactathct agaathcatgg 1380
 aaggaagaat gctgtgaaga agaathgctgt gaagaagaat gctgtgaaga agaathgctgt 1440
 gaagaathat ataaaathat agaggcggat actaacggtg thtatatgcc gttgggagta 1500
 attagtgaac cathththaac accaathctat agthththaac taathattga cgaaagaaca 1560
 aagagaathat cththagcggg taaathctat thacgtgaath cththactagc cacagathta 1620
 gthtaataaag atacgaathth aathctthca ccaathggtt tcathaacag thttgtggaa 1680
 aathggaata taacathcgga thathatagag ccttggaag cgaaataath aathgcatath 1740
 gtcgataaga cggatgath ggtgggathth aactctthth atactcataa ggatggggaa 1800

```

ttcttgcaat ttattggagc taagttaaag gctaaaactg agtatatcat tcaatatact 1860
gtaaaagggg gtccggaagt ttatttgaaa aacaataaag gtatctttta tgaggataca 1920
acaaataaat ttgatacggt tcaaactata actaaaaagt tcaattcagg agtagatcca 1980
tccgaaatat atctagtttt taaaaatcaa attggatatg aagcatgggg aaataaattt 2040
attatactag aaatcaagtc atttgaaacc ctaccacaaa tattaaaacc tgaaaattgg 2100
atgccttttg gtaatgctga gattaaagaa gatggaaaaa ttgagatttc aggtaatgga 2160
actatgacgc aaaatattca attagaacag aattccaagt atcatctaag attttctgta 2220
aaaggaaaag ggagagtagc gatacaaact caaagctccc atataaatgt accagctaca 2280
aacgaagagg tttctacaat gattacaact agaaacttat acggtgaagg tatgatatac 2340
ctatttaatg atgacgtgga gaactccaaa gttatttttt cggatgtatc tctagttaaa 2400
gaatagg 2407

```

```

<210> 9
<211> 801
<212> PRT
<213> 苏云金芽孢杆菌

```

```

<220>
<221> MISC_FEATURE
<222> (1)..(801)
<223> Vip3Z 毒素

```

```

<400> 9

```

```

Met Asn Asn Thr Lys Leu Asn Ala Arg Ala Leu Pro Ser Phe Ile Asp
1          5          10          15

```

```

Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp Ile Met
          20          25          30

```

```

Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Gly Asn Leu Thr Leu Asp
          35          40          45

```

```

Glu Ile Leu Lys Asn Gln Asp Leu Leu Asn Gln Ile Ser Asp Lys Leu
          50          55          60

```

```

Asp Gly Ile Asn Gly Asp Leu Gly Asp Leu Ile Ala Gln Gly Asn Leu
65          70          75          80

```

```

Asn Ser Glu Leu Thr Lys Glu Leu Leu Lys Ile Ala Asn Glu Gln Asn
          85          90          95

```

```

Leu Met Leu Asn Asn Val Asn Ala Gln Leu Asn Ser Ile Asn Ser Thr
          100          105          110

```

Leu Asn Thr Tyr Leu Pro Lys Ile Thr Ser Met Leu Ser Glu Val Met
 115 120 125

Lys Gln Asn Tyr Val Leu Ser Leu Gln Ile Glu Phe Leu Ser Glu Gln
 130 135 140

Leu Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu
 145 150 155 160

Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys
 165 170 175

Tyr Val Asn Asp Lys Phe Asp Glu Leu Thr Ser Thr Val Glu Lys Asn
 180 185 190

Pro Lys Ile Asn Gln Asp Asn Phe Thr Glu Asp Val Ile Asp Asn Leu
 195 200 205

Thr Asp Leu Thr Glu Leu Ala Arg Ser Val Thr Arg Asn Asp Met Asp
 210 215 220

Ser Phe Glu Phe Tyr Ile Lys Thr Phe His Asp Val Met Ile Gly Asn
 225 230 235 240

Asn Leu Phe Ser Arg Ser Ala Leu Lys Thr Ala Ser Glu Leu Ile Ala
 245 250 255

Lys Glu Asn Ile His Thr Met Gly Ser Glu Ile Gly Asn Val Tyr Thr
 260 265 270

Phe Met Val Val Leu Thr Ser Leu Gln Ala Lys Ala Phe Leu Thr Leu
 275 280 285

Thr Ala Cys Arg Lys Leu Leu Gly Leu Thr Asp Ile Asp Tyr Thr Gln
 290 295 300

Ile Met Asn Glu Asn Leu Asn Arg Glu Lys Glu Glu Phe Arg Leu Asn
 305 310 315 320

Ile Leu Pro Thr Leu Ser Asn Asp Phe Ser Asn Pro Asn Tyr Thr Glu
 325 330 335

Thr Leu Gly Ser Asp Leu Val Asp Pro Ile Val Thr Leu Glu Ala Asp
 340 345 350

Pro Gly Tyr Ala Leu Ile Gly Phe Glu Ile Leu Asn Asp Pro Leu Pro
 355 360 365

Val Leu Lys Val Tyr Gln Ala Lys Leu Lys Pro Asn Tyr Gln Val Asp

370	375	380
Lys Glu Ser Ile Met	Glu Asn Ile Tyr Gly	Asn Ile His Lys Leu Leu
385	390	395 400
Cys Pro Lys Gln Arg	His Gln Lys Tyr Tyr	Ile Lys Asp Ile Thr Phe
	405	410 415
Pro Glu Gly Tyr Val	Ile Thr Lys Ile Val Phe	Glu Lys Lys Leu Asn
	420	425 430
Leu Leu Gly Tyr Glu	Val Thr Ala Asn Leu	Tyr Asp Pro Phe Thr Gly
	435	440 445
Ser Ile Asp Leu Asn	Lys Thr Ile Leu Glu	Ser Trp Lys Glu Glu Cys
	450	455 460
Cys Glu Glu Glu Cys	Cys Glu Glu Glu Cys	Cys Glu Glu Cys Cys
465	470	475 480
Glu Glu Leu Tyr Lys	Ile Ile Glu Ala Asp	Thr Asn Gly Val Tyr Met
	485	490 495
Pro Leu Gly Val Ile	Ser Glu Thr Phe Leu	Thr Pro Ile Tyr Ser Phe
	500	505 510
Lys Leu Ile Ile Asp	Glu Arg Thr Lys Arg	Ile Ser Leu Ala Gly Lys
	515	520 525
Ser Tyr Leu Arg Glu	Ser Leu Leu Ala Thr	Asp Leu Val Asn Lys Asp
	530	535 540
Thr Asn Leu Ile Pro	Ser Pro Asn Gly Phe	Ile Asn Ser Ile Val Glu
545	550	555 560
Asn Trp Asn Ile Thr	Ser Asp Asn Ile Glu	Pro Trp Lys Ala Asn Asn
	565	570 575
Lys Asn Ala Tyr Val	Asp Lys Thr Asp Asp	Met Val Gly Phe Asn Ser
	580	585 590
Leu Tyr Thr His Lys	Asp Gly Glu Phe Leu	Gln Phe Ile Gly Ala Lys
	595	600 605
Leu Lys Ala Lys Thr	Glu Tyr Ile Ile Gln	Tyr Thr Val Lys Gly Ser
	610	615 620
Pro Glu Val Tyr Leu	Lys Asn Asn Lys Gly	Ile Phe Tyr Glu Asp Thr
625	630	635 640

Thr Asn Lys Phe Asp Thr Phe Gln Thr Ile Thr Lys Lys Phe Asn Ser
645 650 655

Gly Val Asp Pro Ser Glu Ile Tyr Leu Val Phe Lys Asn Gln Ile Gly
660 665 670

Tyr Glu Ala Trp Gly Asn Lys Phe Ile Ile Leu Glu Ile Lys Ser Phe
675 680 685

Glu Thr Leu Pro Gln Ile Leu Lys Pro Glu Asn Trp Met Pro Phe Gly
690 695 700

Asn Ala Glu Ile Lys Glu Asp Gly Lys Ile Glu Ile Ser Gly Asn Gly
705 710 715 720

Thr Met Thr Gln Asn Ile Gln Leu Glu Gln Asn Ser Lys Tyr His Leu
725 730 735

Arg Phe Ser Val Lys Gly Lys Gly Arg Val Ala Ile Gln Thr Gln Ser
740 745 750

Ser His Ile Asn Val Pro Ala Thr Asn Glu Glu Val Ser Thr Met Ile
755 760 765

Thr Thr Arg Asn Leu Tyr Gly Glu Gly Met Ile Tyr Leu Phe Asn Asp
770 775 780

Asp Val Glu Asn Ser Lys Val Ile Phe Ser Asp Val Ser Leu Val Lys
785 790 795 800

Glu

<210> 10
<211> 2367
<212> DNA
<213> 人工序列

<220>
<223> vip3A-C 杂合毒素编码序列.

<400> 10
atgaacaaga ataatactaa attaagcaca agagccttac caagttttat tgattatttt 60
aatggcattt atggatttgc cactggtatc aaagacatta tgaacatgat ttttaaacg 120
gatacaggtg gtgatctaac cctagacgaa attttaaaga atcagcagtt actaaatgat 180
atctctggta aattggatgg ggtgaatgga agcttaaagt atcttatcgc acagggaaac 240
ttaaatacag aattatctaa ggaaatatta aaaattgcaa atgaacaaaa tcaagtttta 300

aatgatgta ataacaaact cgatgcgata aatacgatgc ttcgggtata tctacctaaa	360
attacctcta tgttgagtga tgtaatgaaa caaaattatg cgctaagtct gcaaatagaa	420
tacttaagta aacaattgca agagatttct gataagttgg atattattaa tgtaaattgta	480
cttattaact ctacacttac tgaaattaca cctgcgtatc aaaggattaa atatgtgaac	540
gaaaaatttg aggaattaac ttttgctaca gaaactagtt caaaagtaaa aaaggatggc	600
tctcctgcag atattcttga tgagttaact gagttaactg aactagcgaa aagtgtaaaca	660
aaaaatgatg tggatggttt tgaattttac cttatacat tccacgatgt aatggtagga	720
aataatttat tggggcgctc agcttttaaaa actgcatcgg aattaattac taaagaaaat	780
gtgaaaacaa gtggcagtga ggtcggaaat gtttataact tcttaattgt attaacagct	840
ctgcaagcaa aagcttttct tactttaaca acatgccgaa aattattagg cttagcagat	900
attgattata cttctattat gaatgaacat ttaaataagg aaaaagagga atttagagta	960
aacatcctcc ctacactttc taatactttt tctaactcta attatgcaa agttaaagga	1020
agtgatgaag atgcaaagat gattgtggaa gctaaaccag gacatgcatt gattgggttt	1080
gaaattagta atgattcaat tacagtatta aaagtatatg aggctaagct aaaacaaaat	1140
tatcaagtcg ataaggattc cttatcggaa gttattttatg gtgatatgga taaattattg	1200
tgcccagatc aatctgaaca aatctattat acaaataaca tagtatttcc aatgaatat	1260
gtaattacta aaattgattt cactaaaaaa atgaaaactt taagatatga ggtaacagcg	1320
aatttttatg attcttctac aggagaaatt gacttaaaata agaaaaagt agaatcaagt	1380
gaagcggagt atagaacggt aagtgctaat gatgatgggg tgtatatgcc gttaggtgtc	1440
atcagtgaaa catttttgac tccgattaat gggtttgcc tccaagctga tgaaaattca	1500
agattaatta cttaacatg taaatcatat ttaagagaac tactgctagc aacagactta	1560
agcaataaag aactaaatt gatcgtcccg ccaagtgggt ttattagcaa tattgtagag	1620
aacgggtcca tagaagagga caatttagag ccgtggaaag caaataataa gaatgcgtat	1680
gtagatcata caggcggagt gaatggaact aaagctttat atgttcataa ggacggagga	1740
atttcacaat ttattggaga taagttaaaa ccgaaaactg agtatgtaat ccaatatact	1800
gttaaaggaa aaccttctat tcatttaaaa gatgaaaata ctggatatat tcattatgaa	1860
gatacaaata ataatttaga agattatcaa actattaata aacgttttac tacaggaact	1920
gatttaaagg gagtgtattht aatttttaaaa agtcaaaatg gagatgaagc ttggggagat	1980
aaatttacia ttttagaaat taagcctgcg gaggatttat taagcccaga attaatat	2040
ccgaattctt ggattacgac tccaggggct agcatttcag gaaataaact tttcattaac	2100
ttggggacia atgggacctt tagacaaagt ctttcattaa acagttattc aacttatagt	2160
ataagcttta ctgcatcagg accatttaat gtgacggtaa gaaattctag gggagtatta	2220
tttgaacgaa gcaaccttat gtcttcaact agtcatattht ctgggacatt caaaactgaa	2280

tccaataata ccggattata tgtagaactt tcccgtcgcct ctggtggtgg tggatcatata 2340

tcatttgaaa acgtttctat taaataa 2367

<210> 11
 <211> 788
 <212> PRT
 <213> 人工序列

<220>
 <223> 杂合 Vip3A-C 毒素

<400> 11

Met Asn Lys Asn Asn Thr Lys Leu Ser Thr Arg Ala Leu Pro Ser Phe
 1 5 10 15

Ile Asp Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp
 20 25 30

Ile Met Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Asp Leu Thr Leu
 35 40 45

Asp Glu Ile Leu Lys Asn Gln Gln Leu Leu Asn Asp Ile Ser Gly Lys
 50 55 60

Leu Asp Gly Val Asn Gly Ser Leu Asn Asp Leu Ile Ala Gln Gly Asn
 65 70 75 80

Leu Asn Thr Glu Leu Ser Lys Glu Ile Leu Lys Ile Ala Asn Glu Gln
 85 90 95

Asn Gln Val Leu Asn Asp Val Asn Asn Lys Leu Asp Ala Ile Asn Thr
 100 105 110

Met Leu Arg Val Tyr Leu Pro Lys Ile Thr Ser Met Leu Ser Asp Val
 115 120 125

Met Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys
 130 135 140

Gln Leu Gln Glu Ile Ser Asp Lys Leu Asp Ile Ile Asn Val Asn Val
 145 150 155 160

Leu Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile
 165 170 175

Lys Tyr Val Asn Glu Lys Phe Glu Glu Leu Thr Phe Ala Thr Glu Thr
 180 185 190

Ser Ser Lys Val Lys Lys Asp Gly Ser Pro Ala Asp Ile Leu Asp Glu

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

Arg Thr Leu Ser Ala Asn Asp Asp Gly Val Tyr Met Pro Leu Gly Val
 465 470 475 480
 Ile Ser Glu Thr Phe Leu Thr Pro Ile Asn Gly Phe Gly Leu Gln Ala
 485 490 495
 Asp Glu Asn Ser Arg Leu Ile Thr Leu Thr Cys Lys Ser Tyr Leu Arg
 500 505 510
 Glu Leu Leu Leu Ala Thr Asp Leu Ser Asn Lys Glu Thr Lys Leu Ile
 515 520 525
 Val Pro Pro Ser Gly Phe Ile Ser Asn Ile Val Glu Asn Gly Ser Ile
 530 535 540
 Glu Glu Asp Asn Leu Glu Pro Trp Lys Ala Asn Asn Lys Asn Ala Tyr
 545 550 555 560
 Val Asp His Thr Gly Gly Val Asn Gly Thr Lys Ala Leu Tyr Val His
 565 570 575
 Lys Asp Gly Gly Ile Ser Gln Phe Ile Gly Asp Lys Leu Lys Pro Lys
 580 585 590
 Thr Glu Tyr Val Ile Gln Tyr Thr Val Lys Gly Lys Pro Ser Ile His
 595 600 605
 Leu Lys Asp Glu Asn Thr Gly Tyr Ile His Tyr Glu Asp Thr Asn Asn
 610 615 620
 Asn Leu Glu Asp Tyr Gln Thr Ile Asn Lys Arg Phe Thr Thr Gly Thr
 625 630 635 640
 Asp Leu Lys Gly Val Tyr Leu Ile Leu Lys Ser Gln Asn Gly Asp Glu
 645 650 655
 Ala Trp Gly Asp Lys Phe Thr Ile Leu Glu Ile Lys Pro Ala Glu Asp
 660 665 670
 Leu Leu Ser Pro Glu Leu Ile Asn Pro Asn Ser Trp Ile Thr Thr Pro
 675 680 685
 Gly Ala Ser Ile Ser Gly Asn Lys Leu Phe Ile Asn Leu Gly Thr Asn
 690 695 700
 Gly Thr Phe Arg Gln Ser Leu Ser Leu Asn Ser Tyr Ser Thr Tyr Ser
 705 710 715 720

Ile Ser Phe Thr Ala Ser Gly Pro Phe Asn Val Thr Val Arg Asn Ser
725 730 735

Arg Gly Val Leu Phe Glu Arg Ser Asn Leu Met Ser Ser Thr Ser His
740 745 750

Ile Ser Gly Thr Phe Lys Thr Glu Ser Asn Asn Thr Gly Leu Tyr Val
755 760 765

Glu Leu Ser Arg Arg Ser Gly Gly Gly Gly His Ile Ser Phe Glu Asn
770 775 780

Val Ser Ile Lys
785

<210> 12
<211> 36
<212> DNA
<213> 人工序列

<220>
<223> 1F 正向引物

<400> 12
atgaacaaga ataatactaa attaagcaca agagcc 36

<210> 13
<211> 35
<212> DNA
<213> 人工序列

<220>
<223> 1R 反向引物

<400> 13
ctcaacatag aggtaatttt aggtagatat acccg 35

<210> 14
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物 P3

<400> 14
gatgatgggg tgtatatgcc gttag 25

<210> 15
<211> 25
<212> DNA
<213> 人工序列

<220>
<223> 引物 P4

<400> 15 aataaattgt gaaattcctc cgtcc	25
<210> 16 <211> 33 <212> DNA <213> 人工序列	
<220> <223> 引物 4F	
<400> 16 agtcaaaatg gagatcaagg ttggggagat aac	33
<210> 17 <211> 34 <212> DNA <213> 人工序列	
<220> <223> 引物 4R	
<400> 17 ttacttaata gagagatcgt ggaaatgtac aata	34
<210> 18 <211> 23 <212> DNA <213> 人工序列	
<220> <223> 引物 P5	
<400> 18 aatggagatg aagcttgggg aga	23
<210> 19 <211> 25 <212> DNA <213> 人工序列	
<220> <223> 引物 P6	
<400> 19 cgtggaaatg tacaatagga ccacc	25
<210> 20 <211> 25 <212> DNA <213> 人工序列	
<220> <223> 引物 Vip3CF4	
<400> 20 gtttagaaga ttttcaaacc attac	25
<210> 21	

<211> 21
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 T7

 <400> 21
 ttaatacgac tcactatagg g 21

 <210> 22
 <211> 34
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 Vip3Cc

 <400> 22
 tttatattaat agaaacgttt tcaaatagata tatg 34

 <210> 23
 <211> 38
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 Vip3Cn

 <400> 23
 caccatgaac aagaataata ctaaattaag cacaagag 38

 <210> 24
 <211> 38
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 Vip3A-N

 <400> 24
 caccatgaac aagaataata ctaaattaag cacaagag 38

 <210> 25
 <211> 29
 <212> DNA
 <213> 人工序列

 <220>
 <223> 引物 Vip3A2050

 <400> 25
 taaagttatc tccccagct tcattctcca 29

 <210> 26
 <211> 24
 <212> DNA
 <213> 人工序列

 <220>

<223> Vip3C-C1	
<400> 26	
aatggagatg aagcttgggg agat	24
<210> 27	
<211> 34	
<212> DNA	
<213> 人工序列	
<220>	
<223> Vip3C-C2	
<400> 27	
ttttatttaaat agaaacgttt tcaaatgata tatg	34
<210> 28	
<211> 27	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 Vip3Za	
<400> 28	
ggcatttatg gatttgccac tggatc	27
<210> 29	
<211> 27	
<212> DNA	
<213> 人工序列	
<220>	
<223> 引物 Vip3Zb	
<400> 29	
tcctttgata cgcaggtgta atttcag	27
<210> 30	
<211> 13829	
<212> DNA	
<213> 人工序列	
<220>	
<223> pNOV2149	
<400> 30	
aagcttgcac gcctgcagtg cagcgtgacc cggcgtgcc cctctctaga gataatgagc	60
attgcatgtc taagttataa aaaattacca catatTTTTT ttgtcacact tgtttgaagt	120
gcagtttatc tatctttata catatattta aactttactc tacgaataat ataatctata	180
gtactacaat aatatcagtg ttttagagaa tcatataaat gaacagttag acatggctca	240
aaggacaatt gagtatTTTg acaacaggac tctacagttt tatctTTTTa gtgtgcatgt	300
gttctccttt ttttttgcaa atagcttcac ctatataata cttcatccat tttattagta	360
catccattta gggtttaggg ttaatggttt ttatagacta atTTTTTTtag tacatctatt	420

ttattctatt ttagcctcta aattaagaaa actaaaactc tatttttagtt tttttattta	480
ataatttaga tataaaaatag aataaaaataa agtgactaaa aattaaacaa atacccttta	540
agaaattaaa aaaactaagg aaacatTTTT cttgttttoga gtagataatg ccagcctggt	600
aaacgccgtc gacgagtcta acggacacca accagcgaac cagcagcgtc gcgtcggggc	660
aagcgaagca gacggcacgg catctctgtc gctgcctctg gacccctctc gagagtcccg	720
ctccaccggt ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg agcggcagac	780
gtgagccggc acggcaggcg gcctcctcct cctctcaccg caccggcagc tacgggggat	840
tcctttccca ccgctccttc gctttccctt cctcgcctgc cgtaataaat agacaccccc	900
tccacaccct ctttcccaaa cctcgtggtg ttcggagcgc acacacacac aaccagatct	960
cccccaaatc caccogtcgg cacctccgct tcaaggtagc ccgctcgtcc tcccccccc	1020
cccctctcta ccttctctag atcggcgctc cggtcctatg ttagggcccg gtagttctac	1080
ttctgttcat gtttgtgtta gatccgtgtt tgtgttagat ccgtgctgct agcgttcgta	1140
cacggatgcg acctgtacgt cagacacggt ctgattgcta acttgccagt gtttctcttt	1200
ggggaatcct gggatggctc tagccgttcc gcagacggga tcgatttcat gatttttttt	1260
gtttcgttgc atagggtttg gtttgcctt ttcttttatt tcaatatatg ccgtgcactt	1320
gtttgtcggg tcatcttttc atgctttttt ttgtcttggg tgtgatgatg tggctcgggt	1380
gggcggtcgt tctagatcgg agtagaatc tgtttcaaac tacctgggtg atttattaat	1440
tttgatctg tatgtgtgtg ccatacatat tcatagttac gaattgaaga tgatggatgg	1500
aaatatcgat ctaggatagg tatacatggt gatgcggggt ttactgatgc atatacagag	1560
atgctttttg ttcgcttggg tgtgatgatg tgggtgtggt gggcggtcgt tcattcgttc	1620
tagatcggag tagaatactg tttcaaacta cctgggtgat ttattaattt tggaaactgta	1680
tgtgtgtgtc atacatcttc atagttacga gtttaagatg gatggaaata tcgatctagg	1740
ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcagcat	1800
ctattcatat gctetaacct tgagtaccta tctattataa taaacaagta tgttttataa	1860
ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg tggatttttt	1920
tagccctgcc ttcatacgtt atttatttgc ttggtaactgt ttcttttgtc gatgctcacc	1980
ctgttgtttg gtgttacttc tgcagggatc caccatgaac aagaacaaca ccaagctctc	2040
caccgcgccc ctccgctcct tcatcgacta cttcaacggc atctacggct tcgccaccgg	2100
catcaaggac atcatgaaca tgatcttcaa gaccgacacc ggccggcaacc tcaccctoga	2160
cgagatcctc aagaaccagc agctcctcaa cgagatcagc ggcaagctcg accggcgtgaa	2220
cggtcctctc aacgacctca tcgcccaggg caacctcaac accgagctgt ccaaggagat	2280
cctcaagatc gccaacgagc agaaccaggt gctcaacgac gtgaacaaca agctcgacgc	2340
catcaacacc atgctccaca tctacctccc gaagatcacc tccatgctct ccgacgtgat	2400

gaagcagaac	tacgccctct	ccctccagat	cgagtacctc	tccaagcagc	tccaggagat	2460
cagcgacaag	ctcgacatca	tcaacgtgaa	cgtgctcatc	aactccaccc	tcaccgagat	2520
caccccggcc	taccagcgca	tcaagtacgt	gaacgagaag	ttcgaggagc	tgaccttcgc	2580
caccgagacc	accctcaagg	tgaagaagga	ctcctccccg	gccgacatcc	tcgacgagct	2640
gaccgagctg	accgagctgg	ccaagtcctg	gaccaagaac	gacgtggacg	gcttcgagtt	2700
ctacctcaac	accttccacg	acgtgatggt	gggcaacaac	ctcttcggcc	gctccgcctt	2760
caagaccgcc	tccgagctga	tcgccaagga	gaacgtgaag	acctccggct	ccgaggtggg	2820
caacgtgtac	aacttctctc	tcgtgctcac	cgccctgcag	gccaaggcct	tcctcaccct	2880
caccacctgc	cgcaagctcc	tcggcctcgc	cggcctcgac	tacacctcca	tcatgaacga	2940
gcacctcaac	aaggagaagg	aggagtccg	cgtgaacatc	ctcccgacct	tctccaacac	3000
cttctccaac	ccgaactacg	ccaaggtgaa	gggctccgac	gaggacgcca	agatgatcgt	3060
ggaggccaag	ccgggccacg	ccctcgtggg	cttcgagatg	tccaacgact	ccatcaccgt	3120
gctcaaggtg	tacgaggcca	agctcaagca	gaactaccag	gtggacaagg	actccctctc	3180
cgaggtgatc	tacggcgaca	ccgacaagct	cttctgcccc	gaccagtccg	agcagatata	3240
ctacaccaac	aacatcgtgt	tcccgaacga	gtacgtgatc	accaagatcg	acttcaccaa	3300
gaagatgaag	accctccgct	acgaggtgac	cgccaacttc	tacgactcct	ccaccggcga	3360
gatcgacctc	aacaagaaga	aggtggagtc	ctccgaggcc	gagtaccgca	ccctctccgc	3420
caacgacgac	ggcgtgtaca	tgccgctcgg	cgtgatctcc	gaaaccttcc	tcaccccgat	3480
caacggcttc	ggcctccagg	ccgacgagaa	ctcccgcctc	atcaccttca	cctgcaagtc	3540
ctacctccgc	gagctgctcc	tcgccaccga	cctctccaac	aaggagacca	agctcatcgt	3600
gccgccgtcc	ggcttcatct	ccaacatcgt	ggagaacggc	tccatcgagg	aggacaacct	3660
cgagccgtgg	aaggccaaca	acaagaacgc	ctacgtggac	cacaccggcg	gcgtgaacgg	3720
caccaaggcc	ctctacgtgc	acaaggacgg	cggtctctcc	cagttcatcg	gcgacaagct	3780
caagccgaag	accgagtacg	tgatccagta	caccgtgaag	ggcaagccgt	ccatccacct	3840
caaggacgag	aacaccggct	acatccacta	cgaggacacc	aacaacaacc	tcaaggacta	3900
ccagaccatc	accaagcgct	tcaccaccgg	caccgacctc	aagggcggtg	acctatcctt	3960
caagtcccag	aacggcgacg	aggcctgggg	cgacaagttc	accatccttg	agatcaagcc	4020
ggccgaggac	ctcctctccc	cggagctgat	caacctgaac	tcctggatca	ccacccgggg	4080
cgctccatc	tcgggaaca	agctcttcat	caacctcggc	accaacggca	ccttcgcgca	4140
gtccctctcc	ctcaactcct	actccaacta	ctccatctcc	ttcaccgctt	ccggcccgtt	4200
caacgtgacc	gtgcgcaact	cccgcagggg	gctcttcgag	cgctccaacc	tcatgtctct	4260
cacctcccac	atctccggca	ccttcaagac	cgagtccaac	aacaccggcc	tctacgtgga	4320

gctgtcccg	cgctccggcg	gcgccggcca	catctccttc	gagaacgtgt	ccatcaagta	4380
gatctgagct	ctagatcccc	gaatttcccc	gatcgttcaa	acatttggca	ataaagtttc	4440
ttaagattga	atcctgttgc	cggtcttgcg	atgattatca	tataatttct	gttgaattac	4500
gttaagcatg	taataattaa	catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	4560
attagagtcc	cgcaattata	catttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	4620
taggataaat	tatcgcgcgc	ggtgtcatct	atgttactag	atcgggaatt	gggtaccagc	4680
ttgcatgcct	gcagtgcagc	gtgaccgggt	cgtgcccttc	tctagagata	atgagcattg	4740
catgtctaag	ttataaaaaa	ttaccacata	ttttttttgt	cacacttgtt	tgaagtgcag	4800
tttatctatc	tttatacata	tatttaaact	ttactctacg	aataatataa	tctatagtac	4860
tacaataata	tcagtgtttt	agagaatcat	ataaatgaac	agttagacat	ggtctaaagg	4920
acaattgagt	attttgacaa	caggactcta	cagttttatc	tttttagtgt	gcatgtgttc	4980
tccttttttt	ttgcaaatag	cttcacctat	ataaacttcc	atccatttta	ttagtacatc	5040
catttagggg	ttagggttaa	tggtttttat	agactaattt	ttttagtaca	tctattttat	5100
tctattttag	cctctaaatt	aagaaaacta	aaactctatt	ttagtttttt	tatttaataa	5160
tttagatata	aatagaata	aaataaagtg	actaaaaatt	aaacaaatac	cctttaagaa	5220
attaaaaaaa	ctaaggaaac	atttttcttg	tttcgagtag	ataatgccag	cctgttaaac	5280
gccgtcgcag	agtctaacgg	acaccaacca	gcgaaccagc	agcgtcgcgt	cgggccaagc	5340
gaagcagacg	gcacggcatc	tctgtcgctg	cctctggacc	cctctcgaga	gttccgctcc	5400
accgttggac	ttgctccgct	gtcggcatcc	agaaattgcg	tggcggagcg	gcagacgtga	5460
gccggcacgg	caggcggcct	cctcctcctc	tcacggcacc	ggcagctacg	ggggattcct	5520
ttcccaccgc	tccttcgctt	tccttctctc	gcccgcgcta	ataaatagac	accccctcca	5580
caccctcttt	ccccaacctc	gtgttgttcg	gagcgcacac	acacacaacc	agatctcccc	5640
caaatccacc	cgtcggcacc	tcogcttcaa	ggtacgcgcg	tcgtcctccc	ccccccccc	5700
tctctacctt	ctctagatcg	gogttccggt	ccatggttag	ggcccggtag	ttctacttct	5760
gttcatgttt	gtgttagatc	cgtgtttgtg	ttagatccgt	gctgctagcg	ttcgtacacg	5820
gatgcgacct	gtacgtcaga	cacgttctga	ttgctaactt	gccagtgttt	ctctttgggg	5880
aatcctggga	tggctctagc	cgttccgcag	acgggatcga	tttcatgatt	ttttttgttt	5940
cgttgcatag	ggtttggttt	gcccttttcc	tttatttcaa	tatatgccgt	gcacttgttt	6000
gtcgggtcat	cttttcatgc	ttttttttgt	cttggttgtg	atgatgtggt	ctggttgggc	6060
ggtcgttcta	gatcggagta	gaattctgtt	tcaaactacc	tggtggattt	attaattttg	6120
gatctgtatg	tgtgtgccat	acatattcat	agttacgaat	tgaagatgat	ggatggaaat	6180
atcgatctag	gataggtata	catgttgatg	cgggttttac	tgatgcatat	acagagatgc	6240
tttttgttcc	cttggttgtg	atgatgtggt	gtggttgggc	ggtcgttcat	tcgttctaga	6300

tccgagtaga	atactgtttc	aaactacctg	gtgtatztat	taattttgga	actgtatgtg	6360
tgtgtcatac	atcttcatag	ttacgagttt	aagatggatg	gaaatatcga	tctaggatag	6420
gtatacatgt	tgatgtgggt	tttactgatg	catatacatg	atggcatatg	cagcatctat	6480
tcatatgctc	taaccttgag	tacctatcta	ttataataaa	caagtatggt	ttataattat	6540
tttgatcttg	atataacttg	atgatggcat	atgcagcagc	tatatgtgga	tttttttagc	6600
cctgccttca	tacgctattd	atttgcttgg	tactgtttct	tttgtcgatg	ctcacctgtt	6660
tgtttggtgt	tacttctgca	gggatccccg	atcatgcaaa	aactcattaa	ctcagtgcaa	6720
aactatgcct	ggggcagcaa	aacggcggtg	actgaacttt	atggtatgga	aaatccgtcc	6780
agccagccga	tggccgagct	gtggatgggc	gcacatccga	aaagcagttc	acgagtgca	6840
aatgccgccc	gagatatcgt	ttcactgcgt	gatgtgattg	agagtgataa	atcgactctg	6900
ctcggagagg	ccgttgccaa	acgetttggc	gaactgcctt	tcctgttcaa	agtattatgc	6960
gcagcacagc	cactctccat	tcaggttcat	ccaaacaaac	acaattctga	aatcggtttt	7020
gccaaagaaa	atgccgcagg	tatcccgatg	gatgccgccg	agcgttaacta	taaagatcct	7080
aaccacaagc	eggagctggt	ttttgcgctg	acgcctttcc	ttgcgatgaa	cgcgtttctg	7140
gaattttccg	agattgtctc	cctactccag	ccggtcgcag	gtgcacatcc	ggcgattgct	7200
cactttttac	aacagcctga	tgccgaacgt	ttaagcgaa	tgttcgccag	cctggtgaat	7260
atgcagggtg	aagaaaaatc	ccgcgcgctg	gcgattttaa	aatcggccct	cgatagccag	7320
cagggtgaac	cgtggcaaac	gattcgttta	atctctgaat	tttaccggga	agacagcggg	7380
ctgttctccc	cgctattgct	gaatgtgggt	aaattgaacc	ctggcgaagc	gatgttcctg	7440
ttcgctgaaa	caccgcacgc	ttacctgcaa	ggcgtggcgc	tggaagtgat	ggcaaacctc	7500
gataacgtgc	tgcggtcggg	tctgacgcct	aaatacattg	atattccgga	actggttgcc	7560
aatgtgaaat	tcgaagccaa	accggctaac	cagttgttga	cccagccggg	gaaacaaggt	7620
gcagaactgg	acttcccgat	tccagtggat	gattttgcct	tctcgctgca	tgaccttagt	7680
gataaagaaa	ccaccattag	ccagcagagt	gccgccattd	tgttctgcgt	cgaaggcgat	7740
gcaacgttgt	ggaaaggttc	tcagcagtta	cagcttaaac	cgggtgaatc	agcgtttatt	7800
gcgcgcaacg	aatcaccggg	gactgtcaaa	ggccacggcc	gtttagcgcg	tgtttacaac	7860
aagctgtaag	agcttactga	aaaaattaac	atctcttgct	aagctgggag	ctcgatccgt	7920
cgacctgcag	atcgttcaaa	catttgcaaa	taaagtttct	taagattgaa	tctgttgccc	7980
ggtcttgcca	tgattatcat	ataatttctg	ttgaattacg	ttaagcatgt	aataattaac	8040
atgtaatgca	tgacgttatt	tatgagatgg	gtttttatga	ttagagtccc	gcaattatac	8100
atthaatacg	cgatagaaaa	caaaatatag	cgcgcaaac	aggataaatt	atcgcgcgcg	8160
gtgtcatcta	tgttactaga	tctgctagcc	ctgcaggaaa	tttaccgggt	cccgggcggc	8220

cagcatggcc	gtatccgcaa	tgtgttatta	agttgtctaa	gcgtaattt	gtttacacca	8280
caatataatcc	tgccaccagc	cagccaacag	ctccccgacc	ggcagctcgg	cacaaaatca	8340
ccactcgata	caggcagccc	atcagaatta	attctcatgt	ttgacagctt	atcatcgact	8400
gcacggtgca	ccaatgcttc	tggcgtcagg	cagccatcgg	aagctgtggt	atggctgtgc	8460
aggtcgtaaa	tactgcata	attcgtgtcg	ctcaaggcgc	actcccgttc	tggataatgt	8520
tttttgccg	gacatcataa	cggttctggc	aaatattctg	aatgagctg	ttgacaatta	8580
atcatccggc	tcgtataatg	tgtggaattg	tgagcggata	acaatttcac	acaggaaaca	8640
gacatgagg	gaagcgttga	tcgccgaagt	atcgactcaa	ctatcagagg	tagttggcgt	8700
catcgagcgc	catctcgaac	cgacgttget	ggccgtacat	ttgtacggct	ccgcagtgga	8760
tggcggcctg	aagccacaca	gtgatattga	tttgctggtt	acggtgaccg	taaggcttga	8820
tgaacaacg	cggcgagctt	tgatcaacga	ccttttgaa	acttcggctt	cccctggaga	8880
gagcgagatt	ctccgcgctg	tagaagtcac	cattgttgtg	cacgacgaca	tcattccgtg	8940
gcgttatcca	gctaagcgcg	aactgcaatt	tggagaatgg	cagcgcaatg	acattcttgc	9000
aggtatcttc	gagccagcca	cgatcgacat	tgatctggct	atcttgctga	caaaagcaag	9060
agaacatagc	gttgcttgg	taggtccagc	ggcggaggaa	ctctttgatc	cggttcctga	9120
acaggatcta	tttgaggcgc	taaataaac	cttaacgcta	tggaaactcg	cgcccgactg	9180
ggctggcgat	gagcgaatg	tagtgcttac	gttgctccgc	atcttggtaca	gcgcagtaac	9240
cggcaaaatc	gcgccgaagg	atgtcgctgc	cgactgggca	atggagcgc	tgccggccca	9300
gtatcagccc	gtcactctt	aagctaggca	ggcttatctt	ggacaagaag	atcgcttggc	9360
ctcgcgcgca	gatcagttgg	aagaatttgt	tactacgtg	aaaggcgaga	tcaccaaagt	9420
agtcggcaaa	taaagctcta	gtggatctcc	gtacccccgg	gggatctggc	tcgcggcgga	9480
cgacgcgacgc	cggggcgaga	ccataggcga	tctcctaaat	caatagtagc	tgtaacctcg	9540
aagcgttca	cttgtaacaa	cgattgagaa	tttttgatc	aaaattgaaa	tacttggttc	9600
gcatttttgt	catccgcggt	cagccgcaat	tctgacgaac	tgcccattta	gctggagatg	9660
attgtacatc	cttcacgtga	aaatttctca	agcgtctgta	acaagggttc	agattttaga	9720
ttgaaagggtg	agccgttgaa	acacgttctt	cttgctgatg	acgacgtcgc	tatgcggcat	9780
cttattattg	aataccttac	gatccacgcc	ttcaaagtga	ccgcggtagc	cgacagcacc	9840
cagttcacia	gagtactctc	ttccgcgacg	gtcgatgtcg	tggttgttga	tctaaattta	9900
ggctcgtgaag	atgggctoga	gatcgttcgt	aatctggcgg	caaagtctga	tattccaatc	9960
ataattatca	gtggcgaccg	ccttgaggag	acggataaag	ttgctgcact	cgagctagga	10020
gcaagtgatt	ttatcgctaa	gccgttcagt	atcagagagt	ttctagcacg	cattcggggt	10080
gccttgccgcg	tgccgcccga	cgttgtccgc	tccaaagacc	gacggctttt	ttgttttact	10140
gactggacac	ttaatctcag	gcaacgtcgc	ttgatgtccg	aagctggcgg	tgagggtgaaa	10200

cttacggcag gtgagttcaa tcttctcctc gcgttttttag agaaacccccg cgacgttcta 10260
tcgcgcgagc aactttctcat tgccagtcga gtacgcgacg aggaggttta tgacaggagt 10320
atagatgttc tcattttgag gctgcccgc aaacttgagg cagatccgtc aagccctcaa 10380
ctgataaaaa cagcaagagg tgccggttat ttctttgacg cggacgtgca ggtttcgcac 10440
ggggggacga tggcagcctg agccaattcc cagatccccg aggaatcggc gtgagcggtc 10500
gcaaaccatc cggcccggta caaatcggcg cggcgctggg tgatgacctg gtggagaagt 10560
tgaaggccgc gcaggccgcc cagcggcaac gcatcgaggc agaagcacgc cccggtgaat 10620
cgtggcaagc ggccgctgat cgaatccgca aagaatccccg gcaaccgccg gcagccggtg 10680
cgccgtcgat taggaagccg cccaagggcg acgagcaacc agattttttc gttccgatgc 10740
tctatgacgt gggcaccgc gatagtcga gcatcatgga cgtggccgtt ttccgtctgt 10800
cgaagcgtga ccgacgagct ggcgaggta tccgctacga gcttccagac gggcacgtag 10860
aggtttccgc agggccggcc ggcatggcca gtgtgtggga ttacgacctg gtactgatgg 10920
cggtttccca tctaaccgaa tccatgaacc gataccggga agggaaggga gacaagccccg 10980
gccgcgtgtt ccgtccacac gttgcccgc tactcaagtt ctgccggcga gccgatggcg 11040
gaaagcagaa agacgacctg gtagaaacct gcattcgggtt aaacaccacg cacgttgcca 11100
tgcagcgtac gaagaaggcc aagaacggcc gcctggtgac ggtatccgag ggtgaagcct 11160
tgattagccg ctacaagatc gtaaagagcg aaaccggcg gccggagtac atcgagatcg 11220
agctagctga ttgatgtac cgcgagatca cagaaggcaa gaaccggac gtgctgacgg 11280
ttcaccocga ttactttttg atcgatccccg gcatcggccg ttttctctac cgctggcac 11340
gccgcgccgc aggcaaggca gaagccagat ggttgttcaa gacgatctac gaacgcagtg 11400
gcagcgcgg agagttcaag aagttctgtt tcaccgtgcg caagctgatc gggtaaagt 11460
acctgccgga gtacgatttg aaggaggagg cggggcaggc tggcccgatc ctagtcatgc 11520
gctaccgcaa cctgatcgag ggcaagcat ccgcccgttc ctaatgtacg gagcagatgc 11580
tagggcaaat tgccctagca ggggaaaaag gtcgaaaagg tctctttcct gtggatagca 11640
cgtacattgg gaaccctaaag ccgtacattg ggaaccggaa cccgtacatt gggaaaccctaa 11700
agccgtacat tgggaaccgg tcacacatgt aagtgactga tataaaagag aaaaaaggcg 11760
atthttccgc ctaaaactct ttaaaaactta ttaaaaactct taaaaccgc ctggcctgtg 11820
cataactgtc tggccagcgc acagccgaag agctgcaaaa agcgcctacc ctccggtcgc 11880
tgcgctccct acgccccgc gcttcgcgtc ggccatcgc gcccgctggc cgctcaaaaa 11940
tggctggcct acggccaggc aatctaccag ggcgcggaca agccgcgccg tcgccactcg 12000
accgccggcg ctgaggtctg cctcgtgaag aaggtgttgc tgactcatac caggcctgaa 12060
tcgccccatc atccagccag aaagtgaggg agccacgggt gatgagagct ttgtttagg 12120

tggaccagtt ggtgattttg aacttttgc tggccacgga acggtctgcg ttgtcgggaa 12180
 gatgcgatgat ctgatacctc aactcagcaa aagttcgatt tattcaacaa agccgcccgtc 12240
 ccgtcaagtc agcgtaatgc tctgccagtg ttacaaccaa ttaaccaatt ctgattagaa 12300
 aaactcatcg agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata 12360
 tttttgaaaa agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat 12420
 ggcaagatcc tggatcgggt ctgcgattcc gactcgtcca acatcaatac aacctattaa 12480
 tttcccctcg tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc 12540
 cggtgagaat ggcaaaagct ctgcattaat gaatcggcca acgcgcgggg agaggcgggtt 12600
 tgcgtattgg gcgctcttcc gttcctcgc tcaactgactc gctgcgctcg gtcgttcggc 12660
 tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca gaatcagggg 12720
 ataacgcagc aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg 12780
 ccgcttgct ggcgttttcc cataggctcc gccccctga cgagcatcac aaaaatcgac 12840
 gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg tttcccctg 12900
 gaagctccct cgtgcgctct cctgttccga ccctgcccgt taccggatac ctgtccgcct 12960
 ttctcccttc ggaagcgtg gcgctttctc atagctcacg ctgtaggtat ctgagttcgg 13020
 tgtaggtegt togetccaag ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct 13080
 gcgccttata cggtaactat cgtcttgagt ccaaccgggt aagacacgac ttatcgccac 13140
 tggcagcagc cactggtaac aggattagca gagcagagta tgtaggcgggt gctacagagt 13200
 tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttggt atctgcgctc 13260
 tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc aaacaaacca 13320
 ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga aaaaaggat 13380
 ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac gaaaactcac 13440
 gttaagggat tttggatcatg agattatcaa aaaggatctt cacctagatc cttttgatcc 13500
 ggaattaatt cctgtggttg gcatgcacat acaaatggac gaacggataa accttttcac 13560
 gcccttttaa atatccgatt attctaataa acgctctttt ctcttaggtt taccgcca 13620
 tatatcctgt caaacactga tagtttaaac tgaaggcggg aaacgacaat ctgatcatga 13680
 gcggagaatt aagggagtca cgttatgacc cccgccgatg acgcgggaca agccgtttta 13740
 cgtttggaac tgacagaacc gcaacgctgc aggaattggc cgcagcggcc atttaaatca 13800
 attgggcgcg ccgaattcga gctcggtag 13829

<210> 31
 <211> 2367
 <212> DNA
 <213> 苏云金芽孢杆菌

```

<220>
<221> misc_feature
<222> (1)..(2367)
<223> Vip3C-12168 编码序列

<400> 31
atgaatatga ataatactaa attaaacgca agggccctac cgagttttat tgattatttt      60
aatggcattt atggatttgc cactgggtatc aaagacatta tgaatatgat ttttaaaacg      120
gatacaggtg gtaatctaac cttagacgaa atcctaaaga atcagcagtt actaaatgag      180
atctctggta aattggatgg ggtaaattgg agcttaaattg atcttatcgc acagggaaac      240
ttaaatacag aattatctaa ggaaatctta aaaattgcaa atgaacagaa tcaagtctta      300
aatgatgtta ataacaaact cgatgcgata aatacgatgc ttcatatata tctacctaaa      360
attacatcta tgtaagtga tgtaatgaag caaaattatg cgctaagtct gcaaatagaa      420
tacttaagta agcaattgca agaaatttct gataaattag atattattaa cgtaaattgt      480
cttattaact ctacacttac tgaattaca cctgcatatc aacggattaa atatgtgaat      540
gaaaaatttg aagaattaac ttttgctaca gaaaccactt taaaagtaaa aaaggatagc      600
tcgcctgctg atattcttga tgagttaact gaattaactg aactagcgaa aagtgttaca      660
aaaaatgacg ttgatggttt tgaattttac cttaaatacat tccacgatgt aatggtagga      720
aataatttat tcgggcggtc agcttttaaaa actgcttcag aattaattgc taaagaaaat      780
gtgaaaacaa gtggcagtga agtaggaaat gtttataatt tcttaattgt attaacagct      840
ctacaagcaa aagcttttct tactttaaca acatgccgaa aattattagg cttagcagat      900
attgattata cttctattat gaatgaacat ttaaataagg aaaaagagga atttagagta      960
aacatccttc ctacactttc taatactttt tctaactcta attatgcaa agttaaagga     1020
agtgatgaag atgcaaagat gattgtggaa gctaaaccag gacatgcatt ggttgggttt     1080
gaaatgagca atgattcaat cacagtatta aaagtatatg aggctaagct aaaacaaaat     1140
tatcaagttg ataaggattc cttatcggag gttatttatg gtgatacggg taaattattt     1200
tgtccagatc aatctgaaca aatatattat acaaataaca tagtattccc aatgaatat     1260
gtaattacta aaattgattt cactaaaaaa atgaaaactt taagatatga ggtaacagcg     1320
aatttttatg attcttctac aggagaaatt gacttaataa agaaaaaagt agaatcaagt     1380
gaagcggagt atagaacggt aagtgcta atgatggag tgtatatgcc attaggtgtc     1440
atcagtgaaa catttttgac tccgataaat gggtttggcc tccaagctga tgaaaattca     1500
agattaatta cttaacatg taaatcatat ttaagagaac tactgctagc aacagactta     1560
agcaataaag aaactaaatt gatcgtccca ccaagtgggt ttattagcaa tattgtagag     1620
aacgggtcca tagaagagga caatttagag ccgtggaaag caaataataa gaatgcgtat     1680
gtagatcata caggcggagt gaatggaact aaagctttat atgttcataa ggacggagga     1740
ttttcacaat ttattggaga taagttaaaa ccgaaaactg agtatgtaat ccaatatact     1800

```

```

gttaaaggaa aaccttctat tcatttataaa gatgaaaata ctggatataat tcattatgaa 1860
gatacaaata ataatttataa agattatcaa actattacta aacgttttac tacaggaact 1920
gatttaaagg gagtgtattht aattttataaa agtcaaaatg gagatgaagc ttggggagat 1980
aaatttataa ttttagaaat taagcctgag gaggatttat taagcccaga attaattaat 2040
ccgaattctt ggattacgac tccaggggct agcatttcag gaaataaact tttcattaac 2100
ttggggacaa atgggacctt tagacaaagt ctttcattaa acagttattc aacttatagt 2160
ataagcttta ctgcatcagg accatttaat gtgacggtaa gaaattctag ggaagtatta 2220
tttgaacgaa gcaaccttat gtcttcaact agtcatattht ctgggacatt caaaactgaa 2280
tccaataata ccggattata tgtagaactt tcccgtcgtc ctgggtggtg tggtcatata 2340
tcatttgaaa acgtttctat taaataa 2367

```

```

<210> 32
<211> 788
<212> PRT
<213> 苏云金芽孢杆菌

```

```

<220>
<221> MISC_FEATURE
<222> (1)..(788)
<223> Vip3C-12168 毒素

```

```

<400> 32

```

```

Met Asn Met Asn Asn Thr Lys Leu Asn Ala Arg Ala Leu Pro Ser Phe
1          5          10          15

```

```

Ile Asp Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp
20          25          30

```

```

Ile Met Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Asn Leu Thr Leu
35          40          45

```

```

Asp Glu Ile Leu Lys Asn Gln Gln Leu Leu Asn Glu Ile Ser Gly Lys
50          55          60

```

```

Leu Asp Gly Val Asn Gly Ser Leu Asn Asp Leu Ile Ala Gln Gly Asn
65          70          75          80

```

```

Leu Asn Thr Glu Leu Ser Lys Glu Ile Leu Lys Ile Ala Asn Glu Gln
85          90          95

```

```

Asn Gln Val Leu Asn Asp Val Asn Asn Lys Leu Asp Ala Ile Asn Thr
100         105         110

```

```

Met Leu His Ile Tyr Leu Pro Lys Ile Thr Ser Met Leu Ser Asp Val
115         120         125

```


Met Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys
 130 135 140

Gln Leu Gln Glu Ile Ser Asp Lys Leu Asp Ile Ile Asn Val Asn Val
 145 150 155 160

Leu Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile
 165 170 175

Lys Tyr Val Asn Glu Lys Phe Glu Glu Leu Thr Phe Ala Thr Glu Thr
 180 185 190

Thr Leu Lys Val Lys Lys Asp Ser Ser Pro Ala Asp Ile Leu Asp Glu
 195 200 205

Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp Val
 210 215 220

Asp Gly Phe Glu Phe Tyr Leu Asn Thr Phe His Asp Val Met Val Gly
 225 230 235 240

Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ser Glu Leu Ile
 245 250 255

Ala Lys Glu Asn Val Lys Thr Ser Gly Ser Glu Val Gly Asn Val Tyr
 260 265 270

Asn Phe Leu Ile Val Leu Thr Ala Leu Gln Ala Lys Ala Phe Leu Thr
 275 280 285

Leu Thr Thr Cys Arg Lys Leu Leu Gly Leu Ala Asp Ile Asp Tyr Thr
 290 295 300

Ser Ile Met Asn Glu His Leu Asn Lys Glu Lys Glu Glu Phe Arg Val
 305 310 315 320

Asn Ile Leu Pro Thr Leu Ser Asn Thr Phe Ser Asn Pro Asn Tyr Ala
 325 330 335

Lys Val Lys Gly Ser Asp Glu Asp Ala Lys Met Ile Val Glu Ala Lys
 340 345 350

Pro Gly His Ala Leu Val Gly Phe Glu Met Ser Asn Asp Ser Ile Thr
 355 360 365

Val Leu Lys Val Tyr Glu Ala Lys Leu Lys Gln Asn Tyr Gln Val Asp
 370 375 380

Lys Asp Ser Leu Ser Glu Val Ile Tyr Gly Asp Thr Asp Lys Leu Phe
 385 390 395 400

Cys Pro Asp Gln Ser Glu Gln Ile Tyr Tyr Thr Asn Asn Ile Val Phe
 405 410 415

Pro Asn Glu Tyr Val Ile Thr Lys Ile Asp Phe Thr Lys Lys Met Lys
 420 425 430

Thr Leu Arg Tyr Glu Val Thr Ala Asn Phe Tyr Asp Ser Ser Thr Gly
 435 440 445

Glu Ile Asp Leu Asn Lys Lys Lys Val Glu Ser Ser Glu Ala Glu Tyr
 450 455 460

Arg Thr Leu Ser Ala Asn Asp Asp Gly Val Tyr Met Pro Leu Gly Val
 465 470 475 480

Ile Ser Glu Thr Phe Leu Thr Pro Ile Asn Gly Phe Gly Leu Gln Ala
 485 490 495

Asp Glu Asn Ser Arg Leu Ile Thr Leu Thr Cys Lys Ser Tyr Leu Arg
 500 505 510

Glu Leu Leu Leu Ala Thr Asp Leu Ser Asn Lys Glu Thr Lys Leu Ile
 515 520 525

Val Pro Pro Ser Gly Phe Ile Ser Asn Ile Val Glu Asn Gly Ser Ile
 530 535 540

Glu Glu Asp Asn Leu Glu Pro Trp Lys Ala Asn Asn Lys Asn Ala Tyr
 545 550 555 560

Val Asp His Thr Gly Gly Val Asn Gly Thr Lys Ala Leu Tyr Val His
 565 570 575

Lys Asp Gly Gly Phe Ser Gln Phe Ile Gly Asp Lys Leu Lys Pro Lys
 580 585 590

Thr Glu Tyr Val Ile Gln Tyr Thr Val Lys Gly Lys Pro Ser Ile His
 595 600 605

Leu Lys Asp Glu Asn Thr Gly Tyr Ile His Tyr Glu Asp Thr Asn Asn
 610 615 620

Asn Leu Lys Asp Tyr Gln Thr Ile Thr Lys Arg Phe Thr Thr Gly Thr
 625 630 635 640

Asp Leu Lys Gly Val Tyr Leu Ile Leu Lys Ser Gln Asn Gly Asp Glu
645 650 655

Ala Trp Gly Asp Lys Phe Thr Ile Leu Glu Ile Lys Pro Ala Glu Asp
660 665 670

Leu Leu Ser Pro Glu Leu Ile Asn Pro Asn Ser Trp Ile Thr Thr Pro
675 680 685

Gly Ala Ser Ile Ser Gly Asn Lys Leu Phe Ile Asn Leu Gly Thr Asn
690 695 700

Gly Thr Phe Arg Gln Ser Leu Ser Leu Asn Ser Tyr Ser Thr Tyr Ser
705 710 715 720

Ile Ser Phe Thr Ala Ser Gly Pro Phe Asn Val Thr Val Arg Asn Ser
725 730 735

Arg Glu Val Leu Phe Glu Arg Ser Asn Leu Met Ser Ser Thr Ser His
740 745 750

Ile Ser Gly Thr Phe Lys Thr Glu Ser Asn Asn Thr Gly Leu Tyr Val
755 760 765

Glu Leu Ser Arg Arg Ser Gly Gly Gly Gly His Ile Ser Phe Glu Asn
770 775 780

Val Ser Ile Lys
785

<210> 33
<211> 2367
<212> DNA
<213> 人工序列

<220>
<223> 玉米优化的 Vip3C-12168

<400> 33
atgaacaaga acaacaccaa gctcaacgcc cgcgccctcc cgctcctcat cgactacttc 60
aacggcatct acggcttcgc caccggcatc aaggacatca tgaacatgat cttcaagacc 120
gacaccggcg gcaacctcac cctcgacgag atcctcaaga accagcagct cctcaacgag 180
atcagcggca agctcgacgg cgtgaacggc tcctcaacg acctcatcgc ccagggcaac 240
ctcaacaccg agctgtccaa ggagatcctc aagatcgcca acgagcagaa ccaggtgctc 300
aacgacgtga acaacaagct cgacgccatc aacaccatgc tccacatcta cctcccgaag 360
atcacctcca tgctctccga cgtgatgaag cagaactacg ccctctccct ccagatcgag 420
tacctctcca agcagctcca ggagatcagc gacaagctcg acatcatcaa cgtgaacgtg 480

ctcatcaact ccaccctoac cgagatcacc ccggcctacc agcgcaccaa gtacgtgaac	540
gagaagttag aggagctgac cttegccacc gagaccacc tcaaggtgaa gaaggactcc	600
tccccggccg acatcctcga cgagctgacc gagctgaccg agctggccaa gtccgtgacc	660
aagaacgacg tggacggctt cgagttctac ctcaacacct tccacgacgt gatgggtggc	720
aacaacctet tcggccgctc cgccctcaag accgcctccg agctgatcgc caaggagaac	780
gtgaagacct ccggctccga ggtgggcaac gtgtacaact tcctcatcgt gctcaccgcc	840
ctgcaggcca aggccttctt caccctcacc acctgccga agctcctcgg cctcggccgac	900
atcgactaca cctccatcat gaacgagcac ctcaacaagg agaaggagga gttccgcgtg	960
aacatcctcc cgaccctctc caacaccttc tccaaccgga actacgcca ggtgaagggc	1020
tccgacgagg acgccaagat gatcgtggag gccaaagccg gccacgccct cgtgggcttc	1080
gagatgtcca acgactccat caccgtgctc aagggtgtacg aggccaagct caagcagaac	1140
taccaggtgg acaaggactc cctctccgag gtgatctacg gcgacaccga caagctcttc	1200
tgccccgacc agtccgagca gatatactac accaacaaca tcgtgttccc gaacgagtac	1260
gtgatcacca agatcgactt caccaagaag atgaagacc tccgctacga ggtgaccgcc	1320
aacttctacg actcctccac cggcgagatc gacctcaaca agaagaaggt ggagtctctc	1380
gaggccgagt accgcaccct ctccgccaac gacgacggcg tgtacatgcc gctcggcgtg	1440
atctccgaaa ccttctcac ccgatcaac ggcttcggcc tccaggccga cgagaactcc	1500
cgctcatca cctcaccctg caagtcttac ctccgagagc tgctcctcgc caccgacctc	1560
tccaacaagg agaccaagct catcgtgccg ccgtccggct tcactctcaa catcgtggag	1620
aacggctcca tcgaggagga caacctcgag ccgtggaagg ccaacaacaa gaacgcctac	1680
gtggaccaca ccggcggcgt gaacggcacc aaggccctct acgtgcacaa ggacggcggc	1740
ttctcccagt tcacggcga caagctcaag ccgaagaccg agtacgtgat ccagtacacc	1800
gtgaagggca agccgtccat ccacctcaag gacgagaaca ccggctacat ccaactacgag	1860
gacaccaaca acaacctcaa ggactaccag accatcacca agcgtctcac caccggcacc	1920
gacctcaagg gcgtgtacct catcctcaag tcccagaacg gcgacgaggc ctggggcgac	1980
aagttcacca tccttgagat caagccggcc gaggacctcc tctccccgga gctgatcaac	2040
ccgaactcct ggatcaccac ccggggcgcc tccatctccg gcaacaagct cttcatcaac	2100
ctcggcacca acggcacctt ccggcagtc ctctccctca actcctactc cacctactcc	2160
atctccttca ccgcctccgg ccggttcaac gtgaccgtgc gcaactcccg cgagggtgctc	2220
ttcgagcgct ccaacctcat gtcctccacc tcccacatct ccggcacctt caagaccgag	2280
tccaacaaca ccggccteta cgtggagctg tcccgcgct ccggcggcgg cggccacatc	2340
tccttcgaga acgtgtccat caagtag	2367