



(12)发明专利申请

(10)申请公布号 CN 107109418 A

(43)申请公布日 2017.08.29

(21)申请号 201580067474.0

(22)申请日 2015.12.03

(30)优先权数据

62/090,899 2014.12.12 US

(85)PCT国际申请进入国家阶段日

2017.06.12

(86)PCT国际申请的申请数据

PCT/US2015/063610 2015.12.03

(87)PCT国际申请的公布数据

W02016/094159 EN 2016.06.16

(71)申请人 先正达参股股份有限公司

地址 瑞士巴塞尔

(72)发明人 M·R·布拉姆莱特 K·赛吉恩

V·C·克拉默 M·S·罗斯

(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 11038

代理人 傅宇昌

(51)Int.Cl.

G12N 15/32(2006.01)

A01H 5/10(2006.01)

C07K 14/325(2006.01)

权利要求书4页 说明书40页
序列表53页

(54)发明名称

用于控制植物有害生物的组合物和方法

(57)摘要

披露了从苏云金芽孢杆菌中分离的新颖的杀昆虫蛋白,这些杀昆虫蛋白针对鳞翅目昆虫有害生物是有活性的。编码这些杀昆虫蛋白的DNA可以用于转化各种原核和真核生物以表达这些杀昆虫蛋白。这些重组生物可以用于控制在不同环境中的鳞翅目昆虫。

1. 一种嵌合基因,该嵌合基因包含可操作地连接至核酸分子的异源启动子,该核酸分子包含编码对至少黑色地老虎(小地老虎)有毒的蛋白质的核苷酸序列,其中该核苷酸序列(a)与SEQ ID NO:1-4中任一项具有至少80%到至少99%序列一致性;或者(b)编码如下蛋白质,该蛋白质包含与SEQ ID NO:13-20中任一项具有至少80%到至少99%序列一致性的氨基酸序列;或者(c)是(a)或(b)的合成序列,该合成序列已经进行密码子优化用于在转基因生物中表达。

2. 如权利要求1所述的嵌合基因,其中该异源启动子是植物可表达型启动子。

3. 如权利要求2所述的嵌合基因,其中该植物可表达型启动子选自下组,该组由以下各项组成:泛素、cnp、玉米TrpA、噬菌体T3基因9'5'UTR、玉米蔗糖合成酶1、玉米醇脱氢酶1、玉米捕光复合物、玉米热休克蛋白、豌豆小亚基RuBP羧化酶、Ti质粒甘露碱合酶、Ti质粒胭脂碱合酶、矮牵牛花查尔酮异构酶、大豆富甘氨酸蛋白1、马铃薯糖蛋白、凝集素、CaMV 35S以及S-E9小亚基RuBP羧化酶启动子。

4. 如权利要求1所述的嵌合基因,其中该蛋白质另外对欧洲玉米蛀虫(欧洲玉米螟)、秋黏虫(草地贪夜蛾)、玉米穗蛾(玉米穗虫)、甘蔗螟(小蔗螟),绒毛豆毛虫(黎豆夜蛾)、大豆夜蛾(大豆尺蠖),西南玉米蛀虫(西南玉米螟)、西部豆切根虫(西部豆夜蛾)、烟夜蛾(烟芽夜蛾)、亚洲玉米蛀虫(亚洲玉米螟)、棉螟蛉(棉铃虫)、条纹蛀茎虫(二化螟)、粉蛀茎虫(非洲大螟)或水稻卷叶螟(稻纵卷叶螟)是有毒的。

5. 如权利要求1所述的嵌合基因,其中该核苷酸序列编码如下蛋白质,该蛋白质包含与SEQ ID NO:13具有至少80%到至少99%一致性的氨基酸序列。

6. 如权利要求1所述的嵌合基因,其中该核苷酸序列编码如下蛋白质,该蛋白质包含与SEQ ID NO:14具有至少99%一致性的氨基酸序列。

7. 如权利要求1所述的嵌合基因,其中该核苷酸序列编码如下蛋白质,该蛋白质包含与SEQ ID NO:15具有至少80%到至少99%一致性的氨基酸序列。

8. 如权利要求1所述的嵌合基因,其中该核苷酸序列编码如下蛋白质,该蛋白质包含与SEQ ID NO:16具有至少80%到至少99%一致性的氨基酸序列。

9. 如权利要求1所述的嵌合基因,其中该合成序列包含SEQ ID NO:5-12中的任一项。

10. 如权利要求1所述的嵌合基因,其中该合成序列编码如下蛋白质,该蛋白质包含SEQ ID NO:13-20中的任一项。

11. 如权利要求1所述的嵌合基因,其中该转基因生物是细菌或植物。

12. 一种合成核酸分子,该合成核酸分子包含编码对至少黑色地老虎(小地老虎)有活性的蛋白质的核苷酸序列,其中该核苷酸序列(a)与SEQ ID NO:5-12中任一项具有至少80%到至少99%序列一致性;或者(b)编码与SEQ ID NO:13-20中任一项具有至少80%到至少99%序列一致性的氨基酸序列。

13. 如权利要求12所述的合成核酸分子,其中该核苷酸序列与SEQ ID NO:5具有至少80%一致性。

14. 如权利要求12所述的合成核酸分子,其中该核苷酸序列与SEQ ID NO:6具有至少80%一致性。

15. 如权利要求12所述的合成核酸分子,其中该核苷酸序列与SEQ ID NO:7具有至少80%一致性。

16. 如权利要求12所述的合成核酸分子,其中该核苷酸序列与SEQ ID NO:8具有至少80%一致性。

17. 一种多核苷酸,该多核苷酸进行至少一个密码子优化用于在转基因生物中表达,其编码对于至少黑色地老虎(小地老虎)和玉米穗蛾(玉米穗虫)有毒的蛋白质,其中该蛋白质包含与SEQ ID NO:13具有至少80%一致性的氨基酸序列,并且在对应于SEQ ID NO:13的氨基酸位置40-44的位置处的氨基酸序列是NLNSC。

18. 一种对至少黑色地老虎(小地老虎)有毒的分离的蛋白质,其中该蛋白质包含与(a)由SEQ ID NO:13-20中任一项代表的氨基酸序列具有至少95%序列一致性的氨基酸序列;或者(b)由SEQ ID NO:5-12中任一项代表的核苷酸序列编码的氨基酸序列具有至少95%序列一致性的氨基酸序列。

19. 如权利要求18所述的蛋白质,其中该氨基酸序列进一步包含对应于SEQ ID NO:6的氨基酸342-354的区域中的氨基酸置换。

20. 如权利要求19所述的蛋白质,其中该氨基酸序列进一步包含在SEQ ID NO:6的位置342、343、344、345、346、347、348、349、350、351、352、353或354处的氨基酸置换。

21. 如权利要求20所述的蛋白质,其中该氨基酸序列包含在对应于SEQ ID NO:6的氨基酸350、351和354的氨基酸位置处的氨基酸置换。

22. 如权利要求21所述的蛋白质,其中该氨基酸序列包含在SEQ ID NO:6的氨基酸位置350、351和354处的氨基酸置换。

23. 如权利要求21所述的蛋白质,其中对应于位置350的氨基酸被异亮氨酸(I)置换,对应于位置351的氨基酸被谷氨酰胺(Q)置换,且对应于位置354的氨基酸被丝氨酸(S)置换。

24. 如权利要求22所述的蛋白质,其中位置350处的亮氨酸(L)被异亮氨酸(I)置换,位置351处的天冬酰胺(N)被谷氨酰胺(Q)置换,且位置354处的苏氨酸(T)被丝氨酸(S)置换。

25. 如权利要求19-24中任一项所述的蛋白质,其中该置换产生具有抗原区域的突变型蛋白,该抗原区域允许该突变型蛋白在蛋白质检测分析中区别于该蛋白质。

26. 如权利要求25所述的蛋白质,其中该蛋白质包含SEQ ID NO:6,且该经突变蛋白包含SEQ ID NO:18。

27. 如权利要求18所述的蛋白质,其中该氨基酸序列包含SEQ ID NO:13-20中的任一项。

28. 如权利要求18所述的蛋白质,其中该氨基酸序列是由SEQ ID NO:5-12中任一项代表的核苷酸序列编码的。

29. 一种杀昆虫组合物,该杀昆虫组合物包含如权利要求18所述的蛋白质和农业上可接受的载体。

30. 如权利要求29所述的组合物,其中该农业上可接受的载体选自下组,该组由以下各项组成:粉剂、尘剂、丸剂、颗粒剂、喷雾剂、乳剂、胶体以及溶液。

31. 如权利要求29所述的组合物,其中该农业上可接受的载体是转基因植物。

32. 如权利要求29所述的组合物,其中该组合物包含转基因细菌细胞,该转基因细菌细胞包含该蛋白质。

33. 如权利要求29所述的组合物,其中该组合物通过脱水、冷冻干燥、均化、萃取、过滤、离心、沉降或浓缩苏云金芽孢杆菌细胞的培养物来制备。

34. 如权利要求29所述的组合物,该组合物包含按重量计从约1%至约99%的杀昆虫蛋白。

35. 如权利要求29-34中任一项所述的组合物,其中该组合物包含如下蛋白质,该蛋白质具有由SEQ ID NO:13-20中任一项代表的氨基酸序列。

36. 一种重组载体,该重组载体包含如权利要求1所述的嵌合基因。

37. 一种转基因非人类宿主细胞,该转基因非人类宿主细胞包含如权利要求1所述的嵌合基因或如权利要求36所述的重组载体。

38. 如权利要求37所述的转基因宿主细胞,该转基因宿主细胞是细菌细胞或植物细胞。

39. 如权利要求38所述的转基因细菌细胞,其中该细菌细胞是芽孢杆菌属、梭菌属、致病杆菌属、发光杆菌属、巴斯德氏芽菌属、埃希氏菌属、假单胞菌属、欧文氏菌属、沙雷氏菌属、克雷伯菌属、沙门氏菌属、巴氏杆菌属、黄单胞菌属、链霉菌属、根瘤菌属、红假单胞菌属、嗜甲基菌属、农杆菌属、醋杆菌属、乳杆菌属、节杆菌属、固氮菌属、明串珠菌属或产碱杆菌属的。

40. 如权利要求39所述的转基因芽孢杆菌属细胞,其中该芽孢杆菌属细胞是苏云金芽孢杆菌细胞。

41. 如权利要求38所述的转基因植物细胞,其中该植物细胞是双子叶植物细胞或单子叶植物细胞。

42. 如权利要求41所述的双子叶植物细胞,其中该双子叶植物细胞选自下组,该组由以下各项组成:大豆细胞、向日葵细胞、番茄细胞、芸苔属作物细胞、棉花细胞、糖用甜菜细胞、烟草细胞。

43. 如权利要求41所述的单子叶植物细胞,其中该单子叶植物细胞选自下组,该组由以下各项组成:大麦细胞、玉蜀黍细胞、燕麦细胞、水稻细胞、高粱细胞、甘蔗细胞以及小麦细胞。

44. 生长在自然光照下的如权利要求41所述的多个植物细胞。

45. 如权利要求44所述的多个植物细胞,其中该多个植物细胞并列以形成质外体。

46. 一种转基因植物,该转基因植物包含如权利要求41所述的植物细胞。

47. 一种转基因植物,该转基因植物包含如权利要求42所述的双子叶植物细胞。

48. 一种转基因植物,该转基因植物包含如权利要求43所述的单子叶植物细胞。

49. 如权利要求48所述的转基因植物,该转基因植物是玉蜀黍植物。

50. 一种转基因植物,该转基因植物包含如权利要求1所述的嵌合基因。

51. 如权利要求50所述的转基因植物,其中该植物选自下组,该组由以下各项组成:大麦、芸苔属作物、棉花、玉蜀黍、水稻、高粱、大豆、糖用甜菜、甘蔗、向日葵、烟草、番茄以及小麦。

52. 如权利要求51所述的转基因植物,其中该植物是玉蜀黍植物。

53. 如权利要求46-52中任一项所述的转基因植物的转基因种子。

54. 一种衍生自如权利要求41-45中任一项所述的转基因植物细胞的收获产物,其中该收获产物包含该蛋白质。

55. 一种衍生自如权利要求54所述的收获产物的加工产物,其中该加工产物选自下组,该组由以下各项组成:细粉、粗粉、油以及淀粉或从其衍生的产物。

56. 一种衍生自如权利要求46-52中任一项所述的转基因植物的收获产物,其中该收获产物包含该蛋白质。

57. 一种衍生自如权利要求56所述的收获产物的加工产物,其中该加工产物选自下组,该组由以下各项组成:细粉、粗粉、油以及淀粉或从其衍生的产物。

58. 来自自如权利要求53所述的转基因种子或自如权利要求46-54中任一项所述的转基因植物的提取物,其中该提取物包含自如权利要求1所述的嵌合基因或蛋白质。

59. 一种产生具有针对至少黑色地老虎(小地老虎)的杀昆虫活性的蛋白质的方法,该方法包括:在自如权利要求37所述的宿主细胞产生该蛋白质的条件下培养该宿主细胞。

60. 自如权利要求59所述的方法,其中该宿主细胞是植物细胞。

61. 自如权利要求60所述的方法,其中该植物细胞是玉蜀黍细胞。

62. 自如权利要求59所述的方法,其中这些条件包括自然光照和使这些细胞在植物基质内生长。

63. 自如权利要求59所述的方法,其中该蛋白质对选自下组的至少一种另外的昆虫具有杀昆虫活性,该组由以下各项组成:欧洲玉米蛀虫(欧洲玉米螟)、秋黏虫(草地贪夜蛾)、玉米穗蛾(玉米穗虫)、甘蔗螟(小蔗螟),绒毛豆毛虫(黎豆夜蛾)、大豆夜蛾(大豆尺蠖),西南玉米蛀虫(西南玉米螟)、西部豆切根虫(西部豆夜蛾)、烟夜蛾(烟芽夜蛾)、亚洲玉米蛀虫(亚洲玉米螟)、棉螟蛉(棉铃虫)、条纹蛀茎虫(二化螟)、粉蛀茎虫(非洲大螟)或水稻卷叶螟(稻纵卷叶螟)及其任何组合。

64. 自如权利要求59所述的方法,其中该核苷酸序列进行密码子优化用于在植物中表达。

65. 自如权利要求59所述的方法,其中该嵌合基因包含SEQ ID NO:1-4中的任一项。

66. 自如权利要求59所述的方法,其中该嵌合基因包含SEQ ID NO:5-8中的任一项。

67. 自如权利要求65所述的方法,其中该蛋白质包含SEQ ID NO:13-16中任一项的氨基酸序列。

68. 自如权利要求66所述的方法,其中该蛋白质包含SEQ ID NO:13-20中任一项的氨基酸序列。

69. 一种产生抗昆虫转基因植物的方法,该方法包括:向植物中引入自如权利要求1所述的嵌合基因,其中该杀昆虫蛋白被表达于该植物中,由此赋予该植物对至少黑色地老虎(小地老虎)的植物抗性,并且产生抗昆虫转基因植物。

70. 自如权利要求69所述的方法,其中通过转化该植物实现该引入步骤。

71. 自如权利要求69所述的方法,其中通过使包含该嵌合基因的第一植物与不同的第二植物杂交来实现该引入步骤。

72. 一种控制至少黑色地老虎(小地老虎)昆虫的方法,该方法包括向这些昆虫递送有效量的自如权利要求1所述的蛋白质。

用于控制植物有害生物的组合和方法

[0001] 电子提交的序列表参考

[0002] 该序列表的官方副本是通过EFS-Web以ASCII格式的序列表以2014年12月5日生成的名为“80668-US-L-ORG-NAT-1_SeqList.txt”的文件进行电子提交的,并且该序列表的大小为135千字节并且与本说明书同时提交。包含在该ASCII格式文件中的序列表是本说明书的一部分,并且通过引用以其全文结合在此。

技术领域

[0003] 本发明涉及杀有害生物蛋白和编码它们的核酸分子,连同用于控制植物有害生物的组合和方法。

背景技术

[0004] 苏云金芽孢杆菌(Bt)是一种革兰氏阳性的孢子形成的土壤细菌,其特征在于它产生晶体包含体的能力,这些晶体包含体对于某些目和种的植物有害生物(包括昆虫)是特异地有毒性的,但是对于植物和其他非目标生物是无害的。出于这个原因,包含苏云金芽孢杆菌菌株或它们的杀昆虫蛋白的组合体可以用作环境上可接受的杀昆虫剂以控制农业昆虫有害生物或多种人或动物疾病的昆虫载体。

[0005] 来自苏云金芽孢杆菌的晶体(Cry)蛋白主要针对鳞翅目的、双翅目的、以及鞘翅目的幼虫具有有力的杀昆虫活性。这些蛋白质还已经显示针对以下目的有害生物的活性:膜翅目、同翅目、毛虱目、食毛目、以及壁虱目,连同其他的无脊椎动物目,例如线虫动物门、扁形动物门、以及肉足鞭毛亚门(Feitelson,J.1993.The Bacillus Thuringiensis family tree.In Advanced Engineered Pesticides.Marcel Dekker,Inc.,New York,N.Y.[Feitelson,J,1993,在前沿的工程化的杀有害生物剂中的苏云金芽孢杆菌家族树,马塞尔德克尔公司,纽约,纽约州)。这些蛋白质最初主要基于它们的杀虫活性而被分类为CryI至CryVI。主要的类别是鳞翅目特异性(I)、鳞翅目和双翅目特异性(II)、鞘翅目特异性(III)、双翅目特异性(IV)、以及线虫特异性(V)和(VI)。这些蛋白质进一步被分类为子族,在各个家族内的更高相关的蛋白质指定了区分的字母,例如CryIA、CryIB、CryIC等。在各个区分内的甚至更紧密相关的蛋白质被给定名称,例如CryIC(a)、CryIC(b)等。术语“Cry毒素”以及“ δ -内毒素”与术语“Cry蛋白”已可互换地使用。对于Cry蛋白和基因的当前新命名法基于氨基酸序列同源性而不是昆虫靶标特异性(Crickmore et al.(1998) Microbiol.Mol.Biol.Rev.62:807-813[Crickmore等人(1998)微生物分子生物学评论,62:807-813])。在这个更可接受的分类中,每种毒素被指定唯一的名称,该名称合并了初级等级(阿拉伯数字)、二级等级(大写字母)、三级等级(小写字母)、以及四级等级(另一个阿拉伯数字)。在当前分类中,在初级等级中罗马数字已经换为阿拉伯数字。例如,在旧命名法下的“CryIA(a)”现在在当前命名法下是“Cry1Aa”。

[0006] Cry蛋白是在Bt的孢子形成阶段期间以晶态形式积聚为原毒素的球状蛋白质分子。在被有害生物摄取后,这些晶体典型地被溶解以释放原毒素,原毒素大小的范围可以

为,例如,对于鳞翅目活性的Cry蛋白为从130-140kDa,并且对于鞘翅目活性的Cry蛋白为60-80kDa。原毒素经由靶有害生物中的肠蛋白酶转变为成熟的毒素片段(大约60-70kDa N末端区域)。这些蛋白质中许多对于特异性目标昆虫是相当有毒性的,但是对于植物和其他非目标生物是无害的。

[0007] Cry蛋白通常具有五个保守序列结构域、以及三个保守结构性结构域(参见例如, de Maagd et al. (2001) Trends Genetics 17:193-199 [de Maagd等人(2001)遗传学趋势, 17:193-199])。第一保守结构性结构域(称作结构域I)典型地由七个 α 螺旋组成并且参与膜插入以及孔形成。结构域II典型地由三个安排为希腊钥匙构型的 β 片层组成,并且结构域III典型地由两个处于“果冻卷”(‘jelly-roll’)构造的反平行的 β 片层组成(de Maagd et al., 2001, supra [de Maagd等人, 2001, 同上])。结构域II和III参与受体识别和结合,并且因此被考虑为毒素特异性的决定物。

[0008] 众多商业上有价值的植物(包括普通的农作物)易受植物有害生物(包括昆虫和线虫有害生物)的攻击的影响,导致作物产量和品质的实质性降低。例如,植物有害生物是在全世界重要农作物损失中的主要因素。由于非哺乳动物有害生物(包括昆虫)的侵染,仅在美国每年就损失大约80亿美元。除了农作物中的损失之外,昆虫有害生物对于菜农和果农,对于观赏性花卉的生产商,以及对于家庭花匠也是负担。

[0009] 昆虫有害生物主要是通过密集使用化学杀有害生物剂来控制,这些化学杀有害生物剂通过抑制昆虫成长、预防昆虫摄食或繁殖、或者导致死亡而有效。生物性有害生物控制剂,例如表达杀有害生物毒素(如Cry蛋白)的苏云金芽孢杆菌(*Bacillus thuringiensis*)菌株,也已经应用至作物植物中,产生了令人满意的结果,提供化学杀有害生物剂的替代物或补充物。已经分离了编码这些Cry蛋白的一些的基因并且它们在异源性宿主(例如转基因植物)中的表达已经显示出提供了另一种用于控制经济上重要的昆虫有害生物的手段。

[0010] 因此可以达到良好的昆虫控制,但是某些化学品有时也能影响非目标有益昆虫,并且某些生物制剂具有非常窄的活性谱。此外,某些化学和生物控制方法的继续使用增加了昆虫有害生物对此类控制措施产生抗性的机会。通过各种抗性管理实践已部分地缓和了这种状况,但仍需要找到新型并有效的有害生物控制剂,这些有害生物控制剂为农民提供经济利益并且是环境可接受的。特别需要的是靶向更广谱的经济上重要的昆虫有害生物并有效控制昆虫品系的控制剂,这些昆虫品系对现有的昆虫控制剂是有抗性或可以变得有抗性。

发明内容

[0011] 鉴于这些需求,本发明的目的是通过提供可以用来控制多种植物有害生物的新颖基因和杀有害生物蛋白来提供新的有害生物控制剂。

[0012] 本发明提供了用于赋予细菌、植物、植物细胞、组织以及种子的杀有害生物活性的组合物以及方法。具体地,提供了包含从苏云金芽孢杆菌(Bt)中分离的编码Cry蛋白的新颖多核苷酸的嵌合基因以及基本上与其一致的序列,这些序列的表达导致具有对经济上重要的昆虫有害生物(特别是侵染植物的昆虫有害生物)的毒性的蛋白质。本发明进一步涉及由这些核酸序列的表达生成的新颖的Cry1蛋白,并且涉及含有这些Cry蛋白的组合物以及配制品,它们通过抑制昆虫有害生物的生存、生长以及繁殖或者限制昆虫相关的对作物植物

的损害或损失的能力对昆虫是有毒的。本发明的Cry蛋白包括天然Cry蛋白以及具有一个或多个氨基酸置换、添加或缺失的突变型Cry蛋白。突变型Cry蛋白的实例包括但不限于被突变为具有比其天然Cry蛋白对应物更宽的活性谱的那些,或经突变以引入表位来产生从天然蛋白中差异性识别出经突变的蛋白的抗体的那些。本发明的新颖Cry蛋白对昆虫有害生物是高活性的。例如,本发明的Cry蛋白可以用于控制一种或多种经济上重要的昆虫有害生物,例如黑色地老虎(black cutworm)(小地老虎(*Agrotis ipsilon*))、欧洲玉米蛀虫(European corn borer)(欧洲玉米螟(*Ostrinia nubilalis*))、秋黏虫(fall armyworm)(草地贪夜蛾(*Spodoptera frugiperda*))、玉米穗蛾(corn earworm)(玉米穗虫(*Helicoverpa zea*))、甘蔗螟(sugarcane borer)(小蔗螟(*Diatraea saccharalis*))、绒毛豆毛虫(velvetbean caterpillar)(黎豆夜蛾(*Anticarsia gemmatalis*))、大豆夜蛾(soybean looper)(大豆尺蠖(*Chrysodeixis includens*))、西南玉米蛀虫(southwest corn borer)(西南玉米螟(*Diatraea grandiosella*))、西部豆切根虫(western bean cutworm)(西部豆夜蛾(*Richia albicosta*))、烟夜蛾(tobacco budworm)(烟芽夜蛾(*Heliothis virescens*))、亚洲玉米蛀虫(Asian corn borer)(亚洲玉米螟(*Ostrinia furnacalis*))、棉螟蛉(cotton bollworm)(棉铃虫(*Helicoverpa armigera*))、条纹蛀茎虫(striped stem borer)(二化螟(*Chilo suppressalis*))、粉蛀茎虫(pink stem borer)(非洲大螟(*Sesamia calamistis*))、水稻卷叶螟(rice leaf folder)(稻纵卷叶螟(*Cnaphalocrocis medinalis*))等。

[0013] 本发明还提供了合成的多核苷酸,其编码本发明的Cry蛋白,并已经进行一个或多个密码子优化用于在转基因生物如细菌和植物中表达。

[0014] 本发明进一步涉及表达盒和重组载体,其包含编码本发明的Cry蛋白的多核苷酸。本发明还提供了包含嵌合基因、或表达盒或重组载体的经转化的细菌、植物、植物细胞、组织、以及种子,该嵌合基因或表达盒或重组载体包含编码本发明的Cry蛋白的多核苷酸。

[0015] 本发明还涉及使用这些多核苷酸的方法,例如在DNA构建体或嵌合基因或表达盒或重组载体中用于在生物(包括微生物和植物)中进行转化和表达。核苷酸或氨基酸序列可以是已经被设计用于在生物(包括但不限于微生物或植物)中表达的合成序列,或者已经被设计用于在制备杂交的、具有增强的杀有害生物活性的毒素的合成序列。本发明进一步涉及制备这些Cry蛋白的方法以及例如在微生物中使用这些核酸序列以控制昆虫或在转基因植物中使用这些核酸序列以赋予对抗昆虫损害的保护的方法,并且涉及使用这些Cry蛋白、以及包含这些Cry蛋白的组合物以及配制品的方法,例如将Cry蛋白或组合物或配制品施用至昆虫侵染的区域,或施用它们以预防性地处理易受昆虫影响的区域或植物从而赋予针对昆虫有害生物的保护。核苷酸或氨基酸序列可以是已经被设计用于在生物(包括但不限于微生物或植物)中表达的合成序列。

[0016] 本发明的组合物以及方法对于对昆虫有毒的生物(确切地为细菌和植物)的产生是有用的。这些生物以及从它们衍生的组合物用于农业目的是所希望的。本发明的组合物对于产生改变的或改进的具有杀有害生物活性的Cry蛋白,或对于检测在产物或生物中的Cry蛋白或核酸的存在也是有用的。

[0017] 参考以下详细说明书和权利要求书,本发明的这些和其他特征、方面、以及优点将变得更好理解。

- [0018] 序列列表中的序列简述
- [0019] SEQ ID NO:1代表对BT-0044蛋白进行编码的核苷酸序列。
- [0020] SEQ ID NO:2代表对BT-0051蛋白进行编码的核苷酸序列。
- [0021] SEQ ID NO:3代表对BT-0068蛋白进行编码的核苷酸序列。
- [0022] SEQ ID NO:4代表对BT-0128蛋白进行编码的核苷酸序列。
- [0023] SEQ ID NO:5代表对BT-0044蛋白进行编码的密码子优化序列。
- [0024] SEQ ID NO:6代表对BT-0051蛋白进行编码的密码子优化序列。
- [0025] SEQ ID NO:7代表对BT-0068蛋白进行编码的密码子优化序列。
- [0026] SEQ ID NO:8代表对BT-0128蛋白进行编码的密码子优化序列。
- [0027] SEQ ID NO:9代表对突变型BT-0044蛋白进行编码的核苷酸序列。
- [0028] SEQ ID NO:10代表对突变型BT-0051蛋白进行编码的核苷酸序列。
- [0029] SEQ ID NO:11代表对突变型BT-0068蛋白进行编码的核苷酸序列。
- [0030] SEQ ID NO:12代表对突变型BT-0128蛋白进行编码的核苷酸序列。
- [0031] SEQ ID NO:13代表BT-0044蛋白的氨基酸序列。
- [0032] SEQ ID NO:14代表BT-0051蛋白的氨基酸序列。
- [0033] SEQ ID NO:15代表BT-0068蛋白的氨基酸序列。
- [0034] SEQ ID NO:16代表BT-0128蛋白的氨基酸序列。
- [0035] SEQ ID NO:17代表突变型BT-0044蛋白的氨基酸序列。
- [0036] SEQ ID NO:18代表突变型BT-0051蛋白的氨基酸序列。
- [0037] SEQ ID NO:19代表突变型BT-0068蛋白的氨基酸序列。
- [0038] SEQ ID NO:20代表突变型BT-0128蛋白的氨基酸序列。
- [0039] SEQ ID NO:21-26代表在本发明中有用的引物。

具体实施方式

[0040] 本说明不旨在是本发明以其而实施的所有不同方式,或可以加入本发明中的所有特征的详细目录。例如,关于一个实施例所说明的特征可以结合入其他实施例中,并且关于具体实施例所说明的特征可以从那个实施例删除。因此,本发明考虑了,在本发明的一些实施例中,可以排除或省略在此陈述的任何特征或特征的组合。另外,鉴于本披露内容,对在此建议的不同实施例的众多变体以及附加对于本领域技术人员是显而易见的,这不脱离本发明。因此,以下说明旨在阐述本发明的一些具体实施例,并且并没有穷尽地叙述其所有排列、组合和变化。

[0041] 除非另外定义,否则所有在此使用的技术和科学术语具有与本发明所属领域的普通技术人员通常所理解的相同的意思。在此的本发明的说明中使用的术语仅仅是出于描述具体实施例的目的并且不旨在限制本发明。还应当理解的是在此使用的术语仅仅是出于描述具体实施例的目的并且不旨在限制本发明的范围。

[0042] 定义

[0043] 当在此和所附权利要求书中使用时,单数形式“一个/种(a)”、“和(and)”和“该(the)”包括复数指代物,除非上下文另外明确地指示。因此,例如,提及“一种植物”是提及一种或多种植物并且包括本领域技术人员已知的其等效物等等。如在此使用的,词语“或

(or)”意指具体清单的任何一个成员并且还包含这个清单的成员的任何一个组合(即,还包括“和”)。

[0044] 术语“约”在此用于意指大约、大致、约或在...左右。当术语“约”结合数值范围来使用时,它通过将边界延伸至高于以及低于所阐明的数值来限定这个范围。通常,术语“约”在此用于将数值限定至以20%的变化,优选10%上下(更高或更低)地高于以及低于规定值。关于温度,术语“约”意指 $\pm 1^{\circ}\text{C}$,优选 $\pm 0.5^{\circ}\text{C}$ 。当术语“约”被用于本发明的上下文中(例如与温度或分子量值组合)时,精确值(即,无“约”)是优选的。

[0045] 本发明的毒性Cry蛋白的“活性”意指该毒性蛋白作为口服活性的昆虫控制剂发挥作用,具有毒性作用、或者能够干扰或阻止昆虫摄食,这可能引起或者可能不引起昆虫的死亡。当本发明的毒性蛋白被递送至昆虫时,这种结果典型地是该昆虫的死亡,或者该昆虫不使该毒性蛋白可供该昆虫利用的来源为食。

[0046] 如在此所使用的,术语“扩增的”意指使用至少一种核酸分子作为模板,构建核苷酸分子的多个拷贝或与该核酸分子互补的多个拷贝。扩增系统包括聚合酶链式反应(PCR)系统、连接酶链式反应(LCR)系统、基于核酸序列的扩增(NASBA,Cangene公司,密西索加,安大略省)、 $Q-\beta$ 复制酶系统、基于转录的扩增系统(TAS)以及链置换扩增(SDA)。参见例如Diagnostic Molecular Microbiology:Principles and Applications,PERSING et al., Ed.,American Society for Microbiology,Washington,D.C.(1993)[诊断分子微生物学:原理与应用,PERSING等人编著,美国微生物学会,华盛顿哥伦比亚特区(1993)]。扩增产物被称为“扩增子”。

[0047] 如在此使用的术语“嵌合构建体”或“嵌合基因”或“嵌合多核苷酸”或“嵌合核酸”(或类似术语)是指如下构建体或分子,该构建体或分子包含被组装进单个核酸分子中的不同来源的两个或更多个多核苷酸。术语“嵌合构建体”、“嵌合基因”、“嵌合多核苷酸”或“嵌合核酸”是指如下任何构建体或分子,其含有但不限于(1)多核苷酸(例如,DNA),包括在自然界中没有被发现在一起的调节多核苷酸和编码多核苷酸(即,构建体中的至少一个多核苷酸相对于它的其他多核苷酸中的至少一个是异源的),或(2)编码不是天然毗邻的蛋白质部分的多核苷酸,或(3)不是天然毗邻的启动子部分。另外,嵌合构建体、嵌合基因、嵌合多核苷酸或嵌合核酸可以包括衍生自不同来源的调节多核苷酸和编码多核苷酸,或包括衍生自相同来源、但以与在自然界中所发现的不同的方式进行安排的调节多核苷酸和编码多核苷酸。在本发明的一些实施例中,嵌合构建体、嵌合基因、嵌合多核苷酸或嵌合核酸包含表达盒,该表达盒包含在调节多核苷酸的控制下、具体地在植物或细菌中具有功能性的调节多核苷酸的控制下的本发明的多核苷酸。

[0048] “编码序列”是转录成RNA(如mRNA、rRNA、tRNA、snRNA、正义RNA或反义RNA)的核酸序列。优选地,RNA进而在生物中被翻译以产生蛋白质。

[0049] 如在此使用的,“密码子优化的”序列意指重组的、转基因的、或合成的多核苷酸的核苷酸序列,其中这些密码子被选择以反映宿主细胞或生物可以具有的特定的密码子偏好性。这典型地是以这样一种方式来完成,该方式是为了保持由密码子优化的核苷酸序列所编码的多肽的氨基酸序列。在某些实施例中,重组DNA构建体的DNA序列包括已经针对该构建体有待在其中进行表达的细胞(例如,动物、植物、或真菌细胞)进行了密码子优化的序列。例如,有待在植物细胞中表达的构建体可以使它的全部或部分序列(例如,第一基因抑

制元件或基因表达元件)进行密码子优化用于在植物中表达。参见例如美国专利号6,121,014,通过引用结合在此。

[0050] “控制”昆虫意指通过毒性作用抑制昆虫有害生物存活、生长、摄食、和/或繁殖的能力,或者限制与昆虫有关的作物植物损害或损失,或者保护在昆虫有害生物存在的条件下生长时的作物的产量潜力。“控制”昆虫可以是或可以不是意指杀死昆虫,尽管它优选意指杀死昆虫。

[0051] 术语“包含(comprises和/或comprising)”当在本说明书中使用,指明所列举特征、整体、步骤、操作、元件、和/或部件的存在,但是不排除一种或多种其他特征、整体、步骤、操作、元件、部件、和/或其组的存在或添加。

[0052] 如在此使用的,连接词“基本上由...组成”(以及语法变体)意指,权利要求书的范围有待被解读为涵盖权利要求书中所列举的指定材料或步骤以及不实质上改变所要求的发明的一个或多个基本和新颖特征的那些。因此,当用于本发明的权利要求中时,术语“基本上由...组成”并不旨在被解释为等同于“包含(comprising)”。

[0053] 在本发明的上下文中,“对应于(corresponding to或corresponds to)”意指当变体Cry蛋白的氨基酸序列与彼此比对时,“对应于”在该变体或同系物蛋白中某些枚举的位置的氨基酸是与参考蛋白中的这些位置比对的那些,但在相对于本发明的特定参考氨基酸序列而言的这些精确的数字位置中是不必要的。例如,如果SEQ ID NO:13是参考序列并且与SEQ ID NO:15比对的话,SEQ ID NO:15的Asn4“对应于”SEQ ID NO:13的Asn6。

[0054] “递送(deliver)”组合物或毒性蛋白意指该组合物或毒性蛋白与昆虫接触,产生毒性作用和对该昆虫的控制。该组合物或毒性蛋白可以按照许多公认的方式进行递送,例如,通过昆虫摄取经口或通过经由转基因植物表达、一种或多种配制的蛋白质组合物、一种或多种可喷洒的蛋白质组合物、饵基、或任何其他领域公认的蛋白质递送系统与昆虫接触。

[0055] 术语“结构域”是指沿着进化相关蛋白的序列的比对在特定位置处保守的一组氨基酸。虽然其他位置上的氨基酸可在同系物之间有所不同,但是在特定位置处高度保守的氨基酸指示在蛋白质的结构、稳定性或功能中很可能是必需的氨基酸。通过其在蛋白质同系物家族的经比对序列中的高度保守性进行鉴别,其可用作鉴别物(identifier),用来确定所讨论的任何多肽是否属于先前鉴别的多肽家族。

[0056] “有效的昆虫控制量”意指毒性蛋白的浓度,它通过毒性作用抑制昆虫存活、生长、摄食和/或繁殖的能力,或者限制与昆虫有关的作物植物损害或损失,或者保护在昆虫有害生物存在的条件下生长时的作物的产量潜力。“有效的昆虫控制量”可以是或可以不是意指杀死昆虫,尽管它优选意指杀死昆虫。

[0057] 如在此使用的“表达盒”意指能够在适当的宿主细胞中指导至少一种感兴趣的多核苷酸的表达的核酸序列,包含可操作地连接至该感兴趣的多核苷酸的启动子,该多核苷酸可操作地连接至终止信号。“表达盒”还典型地包含正确翻译感兴趣的多核苷酸所需的另外的多核苷酸。该表达盒还可以包含在感兴趣的多核苷酸的直接表达中不是必需的但是由于用于从表达载体去除该表达盒的方便限制位点而存在的其他多核苷酸。包含一个或多个感兴趣的核苷酸序列的表达盒可以是嵌合的,意味着它的组分中的至少一种相对于它的其他组分中的至少一种是异源的。该表达盒还可以是天然存在的但已经是以对于异源表达有

用的重组形式而获得的表达盒。然而,典型地,该表达盒相对于该宿主是异源的,即在该表达盒中的感兴趣的多核苷酸不是天然存在于该宿主细胞中的,并且必须已经通过转化过程或育种过程引入到该宿主细胞或该宿主细胞的祖先中。该表达盒中的一个或多个感兴趣的多核苷酸的表达通常是在启动子的控制下。在多细胞生物(如植物)的情况下,启动子还可能对于特定组织、或器官、或者发育阶段是特异性的或优先的。当被转化进植物中时,表达盒或其片段也可被称为“插入的多核苷酸”或者“插入多核苷酸”。

[0058] “基因”在此定义为由多核苷酸组成的遗传单位,该遗传单位占据染色体或质粒上特定位置并且含有用于生物中的具体特征或性状的遗传指令。

[0059] “肠蛋白酶”是在昆虫的消化道中天然发现的蛋白酶。这种蛋白酶通常参与被摄取的蛋白质的消化。

[0060] 当提及基因或核酸使用时,术语“异源”是指编码因子的基因不是在其天然环境中(即,已经通过人工改变)。例如,异源基因可以包括自一个物种引入到另一个物种的基因。异源基因还可以包括对生物来说是天然的基因,该基因已经以一些方式(例如,突变;以多个拷贝添加;连接至非天然启动子或增强子多核苷酸等)被改变。异源基因可以进一步包含植物基因多核苷酸,其包含植物基因的cDNA形式;这些cDNA可以以正义方向(以产生mRNA)或反义方向(以产生反义RNA转录本,其与mRNA转录本是互补的)被表达。在本发明的一个方面中,异源基因区别于内源性植物基因在于该异源基因多核苷酸被典型地连接至包含调节元件如启动子的多核苷酸上,未发现这些多核苷酸与由该异源基因编码的蛋白质的基因或与该染色体中的植物基因多核苷酸天然相关联,或者与在自然界中未发现的染色体的部分(例如,在基因座中表达的基因,其中该基因未正常表达)相关联。另外,“异源”多核苷酸是指不与将多核苷酸引入其中的宿主细胞天然地相关联的多核苷酸,包括天然存在的多核苷酸的非天然存在的多拷贝。

[0061] “同源重组”是在相同的多核苷酸的区域中成对染色体的两个DNA分子或染色单体之间的DNA片段的交换(“交叉”)。“重组事件”在此被理解为意指减数分裂交叉。

[0062] 当核酸序列编码了多肽(该多肽与由参考核酸序列所编码的多肽具有相同的氨基酸序列)时,这种核苷酸序列与这种参考核酸序列“同类编码”。

[0063] 术语“分离的”核酸分子、多核苷酸或毒素是不再存在于其天然环境中的核酸分子、多核苷酸或毒性蛋白。本发明的分离的核酸分子、多核苷酸或毒素可以按照纯化的形式存在,或者可以存在于重组宿主中,例如转基因细菌细胞或转基因植物中。

[0064] “核酸分子”是可以从任何来源中分离的单链或双链DNA或RNA。在本发明的上下文中,核酸分子优选地是DNA区段。

[0065] “可操作地连接”是指在单一核酸片段上多核苷酸的关联,这样使得一者的功能影响另一者的功能。例如,当启动子能够影响编码多核苷酸或功能RNA的表达时(即,该编码多核苷酸或功能RNA处于该启动子的转录控制之下),则该启动子与该编码多核苷酸或功能RNA是可操作地连接的。正义方向或者反义方向的编码多核苷酸能够与调节多核苷酸可操作地连接。

[0066] 如在此使用的“杀有害生物”、“杀昆虫”等是指本发明的Cry蛋白控制有害生物的能力或者可以控制如在此所定义的有害生物Cry蛋白的量。因此,杀有害生物Cry蛋白可以杀死或抑制有害生物(例如,昆虫有害生物)存活、生长、摄食、和/或繁殖的能力。

[0067] “植物”是在发育的任何阶段的任何植物,特别是种子植物。

[0068] “植物细胞”是植物的结构和生理单位,包含原生质体和细胞壁。植物细胞可以处于分离的单细胞或培养细胞的形式,或者是作为较高级的组织单位(例如,植物组织、植物器官、或整株植物)的一部分。

[0069] “植物细胞培养物”意指植物单元(例如像,原生质体、细胞培养物细胞、植物组织中的细胞、花粉、花粉管、胚珠、胚囊、接合子以及处于不同发育阶段的胚)的培养物。

[0070] “植物材料”是指叶、茎、根、花或花的部分、果实、花粉、卵细胞、接合子、种子、切条、细胞或组织培养物、或植物的任何其他部分或产物。

[0071] “植物器官”是植物的独特而明显的已结构化并且分化的部分,如根、茎、叶、花蕾或胚。

[0072] 如在此使用的“植物组织”意指组织化成结构和功能单元的一组植物细胞。包括植物中或培养物中的任何植物组织。这个术语包括但不限于全株植物、植物器官、植物种子、组织培养物以及被组织化成结构和/或功能单元的任何植物细胞群组。这个术语与如以上列出的或由该定义以其他方式涵盖的任何具体类型的植物组织的联合应用或单独应用并不旨在排除任何其他类型的植物组织。

[0073] “多核苷酸”是指由共价键合于链中的许多核苷酸单体构成的聚合物。此类“多核苷酸”包括DNA、RNA、经修饰的寡核苷酸(例如,包含对于生物RNA或DNA不典型的碱基的寡核苷酸,如2'-O-甲基化寡核苷酸)等。在一些实施例中,核酸或多核苷酸可以是单链的、双链的、多链的或其组合。除非另外指示,否则本发明的具体核酸或多核苷酸任选地包含或编码除明确指示的任何多核苷酸之外的互补多核苷酸。

[0074] “感兴趣的多核苷酸”是指如下任何多核苷酸,当其转移至生物(例如,植物)中时赋予该生物所希望的特征,如抗生素抗性、病毒抗性、昆虫抗性、疾病抗性、或对其他有害生物的抗性、除草剂耐受性、改进的营养价值、工业过程中改进的性能、商业上有价值的酶或代谢物的生产、或者改变的繁殖能力。

[0075] 术语“启动子”是指如下多核苷酸,通常在它的编码多核苷酸的上游(5'),它通过提供对正确转录所需的RNA聚合酶以及其他因子的识别来控制该编码多核苷酸的表达。

[0076] “原生质体”是分离的植物细胞,没有细胞壁或仅具有部分的细胞壁。

[0077] 如在此使用的,术语“重组”是指核酸分子(例如,DNA或RNA)和/或蛋白质和/或生物的如下形式,该形式通常不会在自然界中发现并且正因为如此通过人类干预来产生。如在此使用的,“重组核酸分子”是包含多核苷酸组合的核酸分子,这些多核苷酸不会天然地一起存在并且是人类干预的结果,例如,由至少两种彼此异源的多核苷酸的组合组成的核酸分子,和/或人工合成的并且包含偏离通常存在于自然界中的多核苷酸的多核苷酸的核酸分子,和/或包含人工掺入至宿主细胞的基因组DNA中的转基因和该宿主细胞基因组的相关侧翼DNA的核酸分子。重组核酸分子的实例是由将转基因插入至植物的基因组DNA中产生的DNA分子,其可以最终导致该生物中的重组RNA和/或蛋白质分子的表达。如在此使用的,“重组植物”是通常不会在自然界中存在的植物,是人类干预的结果,并且含有掺入至其基因组中的转基因和/或异源核酸分子。由于此类基因组改变,重组植物明显不同于相关的野生型植物。

[0078] “调节元件”是指参与控制核苷酸序列的表达的序列。调节元件包含可操作地连接

至感兴趣的核苷酸序列的启动子以及终止信号。它们还典型地涵盖正确翻译该核苷酸序列所需的序列。

[0079] 在两个核酸或蛋白质序列的背景下,术语“一致的”或“基本上一致的”是指当针对最大对应性进行比较和比对时具有至少60%、优选80%、更优选90%、甚至更优选95%、并且最优选至少99%核苷酸或氨基酸残基一致性的两个或更多个序列或子序列,如使用以下序列比较算法之一或通过目测检查所测量的。优选地,基本上的一致性存在于这些序列的整个长度为至少约50个残基的区域中,更优选地在整个至少约100个残基的区域中,并且最优选地这些序列在至少约150个残基中是基本上一致的。在尤其优选的实施例中,这些序列在编码区的全部长度上基本上是一致的。此外,基本上一致的核酸或蛋白质序列基本上执行相同的功能。

[0080] 对于序列比较,典型地,一个序列充当与测试序列进行比较的参考序列。当使用序列比较算法时,将测试序列和参考序列输入到计算机中(若有必要,指定子序列坐标),并且指定序列算法程序的参数。然后,该序列比较算法基于所指定的程序参数来计算这个或这些测试序列相对于该参考序列的序列一致性百分比。

[0081] 用于比较的序列的最佳比对可以按照以下方式进行,例如通过Smith&Waterman, *Adv. Appl. Math.* 2:482 (1981) [Smith和Waterman, *应用数学进展*, 2:482 (1981)]的局部同源性算法、通过Needleman&Wunsch, *J. Mol. Biol.* 48:443 (1970) [Needleman和Wunsch, *分子生物学杂志*, 48:443 (1970)]的同源比对算法、通过Pearson&Lipman, *Proc. Nat'l. Acad. Sci. USA* 85:2444 (1988) [Pearson和Lipman, *美国国家科学院院刊*, 85:2444 (1988)]的相似性方法的搜索,通过这些算法的计算机化实施(威斯康星州遗传学分析软件包(Wisconsin Genetics Software Package),遗传学计算机组(Genetics Computer Group),科学街575号(575 Science Dr.),麦迪逊,威斯康星州中的GAP、BESTFIT、FASTA、和TFASTA),或通过目测检查(总体上参见Ausubel et al., *infra* [Ausubel等人,下文])。

[0082] 适合于确定序列一致性百分比以及序列相似性的算法的一个实例是BLAST算法,它描述于以下文献中:Altschul et al., *J. Mol. Biol.* 215:403-410 (1990) [Altschul等人, *分子生物学杂志*, 215:403-410 (1990)]。执行BLAST分析的软件是通过国家生物技术信息中心(the National Center for Biotechnology Information, 美国国家医学图书馆(U.S. National Library of Medicine),洛克维尔大道8600号(8600 Rockville Pike),贝塞斯达,马里兰州20894美国)可供公众使用的。这种算法涉及首先通过识别查询序列中具有长度W的短字码而识别得分高的序列对(HSP),这些得分高的序列对当与数据库序列中具有相同长度的字码(word)进行比对时匹配或满足一些正值阈值的得分T。T被称为邻近字码得分阈(Altschul et al., 1990 [Altschul等人, 1990])。这些初始的邻近字码命中充当种子用于起始搜索以发现含有它们的较长的HSP。然后,将这些字码命中在两个方向上沿着每个序列延伸直到累积的比对得分可以得到增加。对于核苷酸序列,使用参数M(对于一对匹配残基的奖赏得分;总是>0)和N(对于错配残基的罚分;总是<0)来计算累积的得分。对于氨基酸序列,使用评分矩阵来计算累积得分。当累积的比对得分从它的最大达到值降低了数量X;由于累积一个或多个负得分的残基比对使累积得分趋于0或0以下;或者到达任一序列的末端时,停止这些字码命中在每个方向上的延伸。BLAST算法的参数W、T、以及X决定了该比对的灵敏度与速度。BLASTN程序(对核苷酸序列来说)使用字长(W)为11、期望值(E)为10、

截止值 (cutoff) 为100、 $M=5$ 、 $N=-4$ 、以及两条链的比较作为默认值。对于氨基酸序列, BLASTP程序使用字长 (W) 为3、期望值 (E) 为10、以及BLOSUM62评分矩阵作为默认值 (参见 Henikoff&Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989) [Henikoff和Henikoff, 美国国家科学院院刊, 89:10915 (1989)])。

[0083] 除计算序列一致性百分比之外, BLAST算法还执行两个序列之间的相似性统计分析 (参见例如, Karlin&Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993) [Karlin和Altschul, 美国国家科学院院刊, 90:5873-5787 (1993)])。由BLAST算法提供的相似性的一种量度是最小概率总和 (P(N)), 它提供了在两个核苷酸或氨基酸序列之间会偶然发生匹配的概率的指示。例如, 若在测试核酸序列与参考核酸序列的比较中最小概率总和小于约0.1、更优选地小于约0.01、并且最优选地小于约0.001, 则该测试核酸序列被认为是与该参考序列相类似的。

[0084] 两个核酸序列基本上一致的另一指示是这两个分子在严格条件下彼此杂交。短语“特异性杂交”是指分子在严格条件下仅与特定的核苷酸序列结合、双链化或杂交, 这是在该序列存在于复合混合物 (例如, 总细胞的) DNA或RNA中时进行的。“实质上结合”是指在探针核酸与靶核酸之间的互补杂交, 并且涵盖少量错配, 这些错配可以通过降低杂交介质的严格来容纳, 以实现靶核酸序列的所希望的检测。

[0085] 在核酸杂交实验 (如DNA杂交和RNA杂交) 的背景下, “严格杂交条件”和“严格杂交洗涤条件”是序列依赖性的, 并且在不同的环境参数下是不同的。较长的序列特定地在较高的温度下杂交。对核酸杂交的广泛指导见于以下文献中: Tijssen (1993) *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes part I chapter 2"Overview of principles of hybridization and the strategy of nucleic acid probe assays"* Elsevier, New York [Tijssen (1993) 生物化学和分子生物学实验室技术-使用核酸探针的杂交第2章第I部分“杂交原理和核酸探针检验策略综述”, 爱思唯尔, 纽约]。通常, 高严格杂交和洗涤条件在限定的离子强度和pH下被选定为比特定序列的热熔点 (T_m) 低约5°C。典型地, 在“严格条件”下, 探针将会与它的靶子序列进行杂交, 但不会与其他序列杂交。

[0086] T_m 是50%的靶序列与完全匹配的探针进行杂交时的温度 (在限定的离子强度和pH下)。极严格条件被选定为等于具体探针的 T_m 。对于互补核酸 (它们在DNA或RNA印迹中在过滤器上具有超过100个互补的残基) 的杂交的严格杂交条件的实例是在42°C下、具有1mg肝素的50%甲酰胺、将杂交进行过夜。高严格洗涤条件的实例是0.15M NaCl在72°C下持续约15分钟。严格洗涤条件的实例是0.2x SSC洗涤在65°C下持续15分钟 (参见, Sambrook, 下文, 对于SSC缓冲剂的说明)。通常, 高严格洗涤之前会先进行低严格洗涤, 以去除背景探针信号。对于例如超过100个核苷酸的双链体的示例性中等严格洗涤是1x SSC在45°C下持续15分钟。对于例如超过100个核苷酸的双链体的低严格洗涤的实例是4-6x SSC在40°C下持续15分钟。对于短探针 (例如, 约10至50个核苷酸), 严格条件典型地涉及小于约1.0M的Na离子的盐浓度, 典型地在pH 7.0至8.3下约0.01至1.0M的Na离子浓度 (或其他盐), 并且温度典型地是至少约30°C。通过加入去稳定剂例如甲酰胺也可以实现严格条件。一般而言, 相比于不相关的探针, 在特定的杂交测定中观察到高出2倍 (或更高) 的信噪比就表明检测到特异性杂交。如果在严格条件下彼此不杂交的核酸所编码的蛋白质是基本上一致的, 则它们仍然是

基本上一致的。例如,当使用遗传密码允许的最大程度的密码子简并而生成核酸的拷贝时,则发生这种情况。

[0087] 以下是杂交/洗涤条件组的实例,这些条件可被用来克隆与本发明的参考核苷酸序列基本上一致的同源核苷酸序列:参考核苷酸序列在以下条件下优选地与该参考核苷酸序列杂交:在7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA中在50℃下,并且在2x SSC、0.1%SDS中在50℃下洗涤;更令人希望的是在7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA中在50℃下,并且在1x SSC、0.1%SDS中在50℃下洗涤;仍更令人希望的是在7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA中在50℃下,并且在0.5x SSC、0.1%SDS中在50℃下洗涤;优选地在7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA中在50℃下,并且在0.1x SSC、0.1%SDS中在50℃下洗涤;更优选地在7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA中在50℃下,并且在0.1x SSC、0.1%SDS中在65℃下洗涤。

[0088] 两个核酸序列或蛋白质基本上一致的另一个指示是由第一核酸编码的蛋白质与由第二核酸编码的蛋白质进行免疫性交联反应或与其特异性结合。因此,蛋白质典型地是与第二蛋白质基本上一致的,例如其中这两种蛋白质仅区别于保守性置换。

[0089] “合成的”是指如下核苷酸序列,该核苷酸序列包含在天然序列中不存在的碱基和/或结构特征。例如,编码本发明的Cry蛋白的人工序列(其更类似于双子叶植物和/或单子叶植物基因的G+C含量和正常密码子分布)被表述为合成的。

[0090] “转化”是用于将异源核酸引入到宿主细胞或生物的方法。具体地,“转化”意指DNA分子稳定整合到感兴趣生物的基因组中。

[0091] “转化的/转基因的/重组的”是指引入了异源核酸分子的宿主生物,例如细菌或植物。该核酸分子可以被稳定地整合到宿主的基因组中,或者该核酸分子还可以作为染色体外分子存在。这样的染色体外分子能够自主复制。转化的细胞、组织或植物应当理解为不仅涵盖转化过程的终产物,而且涵盖其转基因子代。“非转化的”、“非转基因的”、或“非重组的”宿主是指不含有该异源的核酸分子的野生型生物,例如细菌或植物。

[0092] 通过以下标准缩写由核苷酸的碱基来指示这些核苷酸:腺嘌呤(A)、胞嘧啶(C)、胸腺嘧啶(T)、以及鸟嘌呤(G)。氨基酸同样是由以下标准缩写来指示: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)、以及缬氨酸(Val;V)。

[0093] 本发明提供了用于控制有害的植物有害生物的组合物以及方法。具体地,本发明涉及对植物有害生物有毒的Cry蛋白,并且涉及包含编码这些Cry蛋白的核苷酸序列的多核苷酸,并且涉及制备和使用这些多核苷酸和Cry蛋白控制植物有害生物的方法。

[0094] 因此,在一些实施例中,提供了如下嵌合基因,该嵌合基因包含可操作地连接至多核苷酸的异源启动子,该多核苷酸包含编码对于至少黑色地老虎(小地老虎)有毒的蛋白质的核苷酸序列,其中该核苷酸序列(a)与SEQ ID NO:1-4中任一项具有至少80%(例如80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%)到至少99%(99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、

99.8%、99.9%) 序列一致性;或者 (b) 编码包含氨基酸序列的蛋白质,该氨基酸序列与SEQ ID NO:13-16中任一项具有至少80% (例如80%、81%、82%、83%、84%、85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%) 到至少99% (99%、99.1%、99.2%、99.3%、99.4%、99.5%、99.6%、99.7%、99.8%、99.9%) 序列一致性;或者 (c) 是 (a) 或 (b) 的合成序列,该合成序列已经进行密码子优化用于在转基因生物中表达。

[0095] 在其他实施例中,该异源启动子是植物可表达型启动子。例如但不限于,该植物可表达型启动子可以选自下组,该组由以下各项组成:泛素、cnp、玉米TrpA、噬菌体T3基因9 5'UTR、玉米蔗糖合成酶1、玉米醇脱氢酶1、玉米捕光复合物、玉米热休克蛋白、豌豆小亚基RuBP羧化酶、Ti质粒甘露碱合酶、Ti质粒胭脂碱合酶、矮牵牛花查尔酮异构酶、大豆富甘氨酸蛋白1、马铃薯糖蛋白(Potato patatin)、凝集素、CaMV 35S以及S-E9小亚基RuBP羧化酶启动子。

[0096] 在另外的实施例中,由嵌合基因编码的蛋白质另外对选自下组的一个或多个昆虫种类是有毒的,该组由以下各项组成:欧洲玉米蛀虫(欧洲玉米螟)、秋黏虫(草地贪夜蛾)、玉米穗蛾(玉米穗虫)、甘蔗螟(小蔗螟)、绒毛豆毛虫(黎豆夜蛾)、大豆夜蛾(大豆尺蠖)、西南玉米蛀虫(西南玉米螟)、西部豆切根虫(西部豆夜蛾)、烟夜蛾(烟芽夜蛾)、亚洲玉米蛀虫(亚洲玉米螟)、棉蚜蛉(棉铃虫)、条纹蛀茎虫(二化螟)、粉蛀茎虫(非洲大螟)以及水稻卷叶螟(稻纵卷叶螟)。

[0097] 在另外的实施例中,该多核苷酸包含如下核苷酸序列,该核苷酸序列与SEQ ID NO:1具有至少80%到至少99%序列一致性,或者与SEQ ID NO:2具有至少80%到至少99%序列一致性,或者与SEQ ID NO:3具有至少80%到至少99%序列一致性,或者与SEQ ID NO:4具有至少80%到至少99%序列一致性。

[0098] 在其他实施例中,该多核苷酸包含编码如下蛋白质的核苷酸序列,该蛋白质包含与SEQ ID NO:13-16中任一项具有至少80%到至少99%序列一致性的氨基酸序列。

[0099] 在仍其他实施例中,该氨基酸序列与SEQ ID NO:13具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0100] 在另外的实施例中,该氨基酸序列与SEQ ID NO:14具有至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0101] 在仍另外的实施例中,该氨基酸序列与SEQ ID NO:15具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0102] 在其他实施例中,该氨基酸序列与SEQ ID NO:16具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0103] 在一些实施例中,本发明的嵌合基因包含如下多核苷酸,该多核苷酸包含如下核苷酸序列的合成序列,该核苷酸序列与SEQ ID NO:5-12中任一项具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%一致性,其中该合成序列已经进行密码子优化用于在转基因生物中表达。在其他实施例中,本发明的嵌合基因包含如下核酸分子,该核酸分子包含编码如下蛋白质的核苷酸序列的合成序列,该蛋白质包含如下氨基酸序列,该氨基酸序列与SEQ ID NO:13-20中任一项具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性,其中该合成序列已经进行密码子优化用于在转基因生物中表达。在另外的实施例中,该转基因生物是转基因细菌或转基因植物。

[0104] 在一些实施例中,本发明提供了如下合成多核苷酸,该合成多核苷酸包含以下项、基本上由以下项组成、或由以下项组成:编码对于至少黑色地老虎(小地老虎)有活性的蛋白质的核苷酸序列,其中该核苷酸序列与SEQ ID NO:5-12中任一项具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0105] 在其他实施例中,本发明提供了如下合成多核苷酸,该合成多核苷酸包含以下项、基本上由以下项组成、或由以下项组成:编码对于至少黑色地老虎(小地老虎)有活性的蛋白质的核苷酸序列,其中该核苷酸序列编码如下氨基酸序列,该氨基酸序列与SEQ ID NO:13-20中任一项具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0106] 在一些实施例中,本发明提供了如下合成多核苷酸,该合成多核苷酸包含以下项、

基本上由以下项组成、或由以下项组成：如下核苷酸序列，该核苷酸序列进行至少一个密码子优化用于在转基因生物中表达并编码对于至少黑色地老虎(小地老虎)和玉米穗蛾(玉米穗虫)有毒的蛋白质，其中该蛋白质包含与SEQ ID NO:13具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性的氨基酸序列，并且在对应于SEQ ID NO:13的氨基酸位置40-44的位置处的氨基酸序列是NLNSC。在另外的实施例中，该多核苷酸包含以下项、基本上由以下项组成、或由以下项组成：SEQ ID NO:5或SEQ ID NO:9。在另外的实施例中，该氨基酸序列包含以下项、基本上由以下项组成、或由以下项组成：SEQ ID NO:13或SEQ ID NO:17。

[0107] 根据一些实施例，本发明提供了对于至少黑色地老虎(小地老虎)有毒的分离的蛋白质，其中该蛋白质包含以下项、基本上由以下项组成、或由以下项组成：(a)与SEQ ID NO:13-20中任一项代表的氨基酸序列具有至少80%序列一致性到至少99%序列一致性的氨基酸序列；或者(b)由如下核苷酸序列编码的氨基酸序列，该核苷酸序列与由SEQ ID NO:5-12中任一项代表的核苷酸序列具有至少80%序列一致性到至少99%序列一致性。

[0108] 在其他实施例中，该分离的蛋白质包含以下项、基本上由以下项组成、或由以下项组成：与SEQ ID NO:13-16中任一项具有至少80%到至少99%序列一致性的氨基酸序列。在仍其他实施例中，该氨基酸序列与SEQ ID NO:13具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0109] 在其他实施例中，该氨基酸序列与SEQ ID NO:14具有至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0110] 在另外的实施例中，该氨基酸序列与SEQ ID NO:15具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0111] 在仍另外的实施例中，该氨基酸序列与SEQ ID NO:16具有至少80%、或至少81%、或至少82%、或至少83%、或至少84%、或至少85%、或至少86%、或至少87%、或至少88%、或至少89%、或至少90%、或至少91%、或至少92%、或至少94%、或至少94%、或至少95%、或至少96%、或至少97%、或至少98%、或至少99%、或至少99.1%、或至少99.2%、或至少99.3%、或至少99.4%、或至少99.5%、或至少99.6%、或至少99.7%、或至少99.8%、或至少99.9%序列一致性。

[0112] 在一些实施例中,该氨基酸序列包含以下项、基本上由以下项组成、或由以下项组成:SEQ ID NO:13-20中任一项。

[0113] 本发明的响应于通过天然的或突变型BT-0044、BT-0051、BT-0068和BT-0128等或相关蛋白的免疫激发而产生的抗体可以使用生产多克隆抗血清的标准免疫学技术来生产,并且如果需要的话,无限增殖经免疫宿主的抗体产生细胞用作单克隆抗体生产源。用于生产任何感兴趣物质的抗体的技术是熟知的,例如如在以下文献中:Harlow and Lane (1988) [Harlow和Lane (1988)]和Goding (1986) [Goding (1986)]。本发明涵盖杀昆虫蛋白,其与针对本发明的杀昆虫Cry蛋白中的一种或多种产生的抗体交叉反应。

[0114] 本发明中生产的这些抗体在用于确定生物样品中天然或突变型BT-0044、BT-0051、BT-0068和BT-0128或相关蛋白的量或存在的免疫测定中也是有用的。此类测定在质量控制生产含有本发明的毒性蛋白中的一种或多种或相关毒性蛋白的组合物中也是有用的。此外,这些抗体可以用来评估本发明的蛋白中的一种或多种或相关蛋白的重组生产的效力,连同针对编码本发明的蛋白中的一种或多种的核苷酸序列或相关蛋白编码序列的存在而筛选表达文库的效力。抗体还作为亲和配体用于纯化和/或分离本发明的蛋白中的任何一种或多种和相关蛋白是有用的。本发明的蛋白和含有相关抗原表位的蛋白可以通过在优选的宿主细胞中过度表达编码全部或部分本发明的蛋白或相关蛋白的序列的全长或部分长度来获得。

[0115] 应当认识到,可以通过不同的方法来改变编码本发明的Cry蛋白的DNA序列,并且这些改变可以产生编码如下蛋白质的DNA序列,这些蛋白质具有不同于由本发明的天然Cry蛋白所编码的氨基酸序列。这种蛋白质可以按照不同的方式进行改变,包括SEQ ID NO:13-16中任一项的一个或多个氨基酸的氨基酸置换、缺失、截短、以及插入,包括多达约2、约3、约4、约5、约6、约7、约8、约9、约10、约15、约20、约25、约30、约35、约40、约45、约50、约55、约60、约65、约70、约75、约80、约85、约90、约100、约105、约110、约115、约120、约125、约130、约135、约140、约145、约150、约155个、或更多个氨基酸的置换、缺失或插入。用于这样的操作的方法在本领域中通常是已知的。例如,通过在编码该蛋白质的多核苷酸中的突变可以制备天然Cry蛋白的氨基酸序列变体。这还可以通过几种诱变形式之一和/或在定向进化中来完成。在一些方面中,在该氨基酸序列中所编码的改变将实质上不影响该蛋白质的功能。此类变体将具有所希望的杀昆虫活性。在本发明的一个实施例中,由SEQ ID NO:1-4代表的核苷酸序列被改变,以在编码的蛋白质中引入氨基酸置换。在一些实施例中,所得到的突变型蛋白是由合成的突变型多核苷酸编码的,该多核苷酸包含由SEQ ID NO:9-12中任一项代表的核苷酸序列。在其他实施例中,这些突变型蛋白包含以下项、基本上由以下项组成、或由以下项组成:由SEQ ID NO:17-20中任一项代表的氨基酸序列。

[0116] 应当理解的是可以通过使用此类技术改善杀昆虫蛋白对本发明的这些组合物赋予杀昆虫活性的能力。例如,可以在如下宿主细胞中表达Cry蛋白,这些宿主细胞在DNA复制过程中显示出高比率的碱基错掺入,如XL-1 Red (Stratagene公司,拉荷亚,加利福尼亚州)。在此类菌株中繁殖之后,可以分离出DNA(例如通过制备质粒DNA,或通过由PCR进行扩增并且将得到的PCR片段克隆到载体中),在非诱变菌株中培养这些Cry蛋白突变体,并且鉴定具有杀昆虫活性的经突变的基因,例如通过进行对杀昆虫活性进行测试的测定。通常,在摄食测定中混合并使用了该蛋白质。参见例如Marrone et al. (1985) J. of Economic

Entomology 78:290-293[Marrone等人(1985)经济昆虫学杂志,78:290-293]。此类测定可以包括使植物与一种或多种有害生物接触,并且确定该植物存活和/或引起这些有害生物死亡的能力。导致毒性提高的突变的实例见于Schnepf et al.(1998) Microbiol.Mol.Biol.Rev.62:775-806[Schnepf等人(1998)微生物分子生物学综述,62:775-806]。

[0117] 可替代地,可以在氨基或羧基的末端上对本发明的氨基酸序列进行改变,而实质上不影响活性。这可以包括通过现代分子方法所引入的插入、缺失、或改变,这些方法是如PCR,包括PCR扩增,这些PCR扩增借助于将编码氨基酸的序列包含到在PCR扩增中所使用的寡核苷酸之中而改变或延长该蛋白质编码序列。可替代地,所加入的蛋白质序列可以包括完整的蛋白质编码序列,例如在本领域内通常用于产生蛋白质融合物的那些序列。此类融合蛋白常常用于(1)增加感兴趣的蛋白质的表达;(2)引入结合结构域、酶活性、或表位以促进蛋白质纯化、蛋白质检测、或本领域已知的其他实验用途;(3)将蛋白质的分泌或翻译靶向亚细胞器,例如革兰氏阴性菌的壁膜间隙,或真核细胞的内质网,后者常常导致蛋白质的糖基化。

[0118] 本发明的Cry蛋白还可以被突变以引入表位来产生识别该经突变蛋白的抗体。因此,在一些实施例中,本发明提供了经突变的Cry蛋白,其中在天然Cry蛋白中的氨基酸置换产生了具有抗原区域的突变型Cry蛋白,该抗原区域允许该突变型Cry蛋白在蛋白质检测分析中区别于该天然Cry蛋白。在其他实施例中,本发明提供了本发明的经突变Cry蛋白,其中氨基酸序列包含对应于SEQ ID NO:6的氨基酸342-354的区域中的氨基酸置换。在其他实施例中,该氨基酸序列包含在SEQ ID NO:6的位置342、343、344、345、346、347、348、349、350、351、352、353或354处的氨基酸置换。在仍其他实施例中,该氨基酸序列包含在对应于SEQ ID NO:6的氨基酸350、351和354的氨基酸位置处的氨基酸置换。在另外的实施例中,该氨基酸序列包含在对应于SEQ ID NO:6的氨基酸350、351和354的氨基酸位置处的氨基酸置换。在仍另外的实施例中,对应于位置350的氨基酸被异亮氨酸(I)置换,对应于位置351的氨基酸被谷氨酰胺(Q)置换,且对应于位置354的氨基酸被丝氨酸(S)置换。在其他实施例中,SEQ ID NO:6的位置350处的亮氨酸(L)被异亮氨酸(I)置换,在SEQ ID NO:6的位置351处的天冬酰胺(N)被谷氨酰胺(Q)置换,且在SEQ ID NO:6的位置354处的苏氨酸(T)被丝氨酸(S)置换。在其他实施例中,该天然Cry蛋白包含由SEQ ID NO:13-16中任一项代表的氨基酸序列。在仍其他实施例中,该天然Cry蛋白包含由SEQ ID NO:6代表的氨基酸序列,并且该突变型蛋白包含由SEQ ID NO:18代表的氨基酸序列。

[0119] 在一些实施例中,本发明提供了特异性识别本发明的突变型Cry蛋白的表位的抗体,其中该表位包含在对应于SEQ ID NO:6的氨基酸342-354的氨基酸中具有一处或多处置换的氨基酸序列。在其他实施例中,该表位包含在SEQ ID NO:6的氨基酸342-354中具有一处或多处置换的氨基酸序列。在仍其他实施例中,该表位包含SEQ ID NO:18的氨基酸342-354。

[0120] 在一些实施例中,本发明提供了制备如下抗体的方法,所述抗体从衍生出经突变的Cry蛋白的天然Cry蛋白中差异性识别出经突变的Cry蛋白,该方法包括以下步骤:在天然Cry蛋白的抗原环中置换氨基酸;并且产生特异性识别茎突变Cry蛋白的经突变抗原环但不识别该天然Cry蛋白的抗体。在一个实施例中,该抗原环在天然Cry蛋白的结构域I的外部

的非保守区中被鉴定出。在另一个实施例中,该抗原环不是参与Cry蛋白的昆虫肠受体识别或参与Cry蛋白的蛋白酶活化的环。在另一个实施例中,该抗原环包含对应于SEQ ID NO:6的氨基酸341-354的氨基酸序列。在又一个实施例中,该抗原环包含SEQ ID NO:6的氨基酸342-354。

[0121] 本发明的变体核苷酸和氨基酸序列还涵盖了由诱变和引起重组的程序(如DNA改组)所衍生的序列。使用此类程序,可以将一个或多个不同的毒性蛋白编码区用来创造出新的具有所希望特性的毒性蛋白。用这种方式,从相关序列多核苷酸的群体产生重组多核苷酸文库,这些相关序列多核苷酸包含如下序列区域,这些序列区域具有基本的序列一致性并且可以在体外或体内进行同源重组。例如,使用这种方法,可以将编码感兴趣的结构域的序列基序在本发明的杀有害生物基因与其他已知的杀有害生物基因之间进行改组,以获得编码如下蛋白质的新基因,该蛋白质具有改进的感兴趣的特性,例如增加的杀昆虫活性。用于此种DNA改组的策略在本领域中是已知的。参见例如,Stemmer (1994) Proc.Natl.Acad.Sci.USA 91:10747-10751 [Stemmer (1994) 美国国家科学院院刊,91:10747-10751]; Stemmer (1994) Nature 370:389-391 [Stemmer (1994) 自然,370:389-391]; Cramer et al. (1997) Nature Biotech.15:436-438 [Cramer等人 (1997) 自然生物技术,15:436-438]; Moore et al. (1997) J.Mol.Biol.272:336-347 [Moore等人 (1997) 分子生物学杂志,272:336-347]; Zhang et al. (1997) Proc.Natl.Acad.Sci.USA 94:4504-4509 [Zhang等人 (1997) 美国国家科学院院刊,94:4504-4509]; Cramer et al. (1998) Nature 391:288-291 [Cramer等人 (1998) 自然,391:288-291]; 以及美国专利号5,605,793和5,837,458。

[0122] 结构域交换或改组是用于产生本发明的经改变的Cry蛋白的另一种机制。可以在Cry蛋白之间交换结构域,从而产生具有改进的杀有害生物活性或目标谱的杂合或嵌合毒性蛋白。用于产生重组蛋白和测试它们的杀有害生物活性的方法在本领域是熟知的(参见例如,Naimov et al. (2001) Appl.Environ.Microbiol.67:5328-5330 [Naimov等人 (2001) 应用与环境微生物学,67:5328-5330]; de Maagd et al. (1996) Appl.Environ.Microbiol.62:1537-1543 [de Maagd等人 (1996) 应用与环境微生物学,62:1537-1543]; Ge et al. (1991) J.Biol.Chem.266:17954-17958 [Ge等人 (1996) 生物化学杂志,266:17954-17958]; Schnepf et al. (1990) J.Biol.Chem.265:20923-20930 [Schnepf等人 (1990) 生物化学杂志,265:20923-20930]; Rang et al. (1999) Appl.Environ.Microbiol.65:2918-2925 [Rang等人 (1999) 应用与环境微生物学,65:2918-2925])。

[0123] 在一些实施例中,本发明提供了如下重组载体,该重组载体包含本发明的多核苷酸、核酸分子、表达盒或嵌合基因。在其他实施例中,该载体被进一步限定为质粒、粘粒、噬菌粒、人工染色体、噬菌体或病毒载体。用于在植物和其他生物的转化中使用的某些载体在本领域是已知的。

[0124] 因此,本发明的一些实施例针对被设计成表达本发明的多核苷酸和核酸分子的表达盒。如在此使用的,“表达盒”是指如下核酸分子,该核酸分子具有至少一个操作性地连接到感兴趣的核苷酸序列上的控制序列。以这种方式,例如,可操作地连接至待表达的核苷酸序列的植物启动子可以在表达盒中提供,用于在植物、植物部分和/或植物细胞中进行表

达。

[0125] 包含感兴趣的核苷酸序列的表达盒可以是嵌合的,意味着它的组分中的至少一种相对于它的其他组分中的至少一种是异源的。表达盒还可以是天然存在的但已经是以对于异源表达有用的重组形式而获得的表达盒。然而,典型地,该表达盒相对于该宿主是异源的,即该表达盒的特定核酸序列不是天然存在于该宿主细胞中的,并且必须已经通过转化事件引入到该宿主细胞或该宿主细胞的祖先中。

[0126] 除可操作地连接至本发明的核苷酸序列的启动子之外,本发明的表达盒还可以包括其他调节序列。如在此使用的,“调节序列”意指位于编码序列的上游(5'非编码序列)、内部或下游(3'非编码序列)并且影响相关编码序列的转录、RNA加工或稳定性、或翻译的核苷酸序列。调节序列包括但不限于增强子、内含子、翻译前导序列、终止信号、以及多腺苷酸化信号序列。

[0127] 在一些实施例中,本发明的表达盒还可以包括对其他所希望的性状进行编码的核苷酸序列。此类核苷酸序列可以与核苷酸序列的任何组合叠加,以产生具有所希望的表型的植物、植物部分或植物细胞。叠加的组合可以通过任何方法来产生,包括但不限于,通过任何常规的方法学的杂交育种植物或通过遗传转化(即,分子叠加)。如果是通过遗传转化这些植物来进行叠加的,感兴趣的核苷酸序列可以在任何时间并且以任何次序进行组合。例如,包含一种或多种所希望的性状的转基因植物可以用作通过后续转化而引入另外的性状的靶标。另外的核苷酸序列可以在共转化方案中与由表达盒的任何组合提供的本发明的核苷酸序列、核酸分子、核酸构建体、和/或组合物同时引入。例如,如果将引入两个核苷酸序列,则它们可以掺入在分开的盒(反式)中或可以掺入在相同的盒(顺式)上。多核苷酸的表达可以通过相同的启动子或通过不同的启动子来驱动。应进一步认识到多核苷酸可以使用位点特异性重组系统在所希望的基因组位置处叠加。参见例如,国际专利申请公开号WO 99/25821;WO 99/25854;WO 99/25840;WO 99/25855以及WO 99/25853。

[0128] 表达盒还可以包括一种或多种多肽的编码序列,这一种或多种多肽用于主要受益者是种子公司、栽培者或谷物加工者的农艺性状。感兴趣的多肽可以由感兴趣的核苷酸序列编码的任何多肽。适合用于在植物中产生的感兴趣的多肽的非限制性实例包括产生农艺学重要性状的那些多肽,这些性状是如除草剂抗性(有时也称为“除草剂耐受性”)、病毒抗性、细菌病原体抗性、昆虫抗性、线虫抗性、和/或真菌抗性。参见例如,美国专利号5,569,823;5,304,730;5,495,071;6,329,504;以及6,337,431。多肽还可以是提高植物活力或产量(包括允许植物在不同的温度、土壤条件以及日光和沉淀水平下生长的性状)的多肽,或是允许对展现感兴趣性状(例如,选择性标记、种皮颜色等)的植物进行鉴定的多肽。不同的感兴趣的多肽,以及用于将这些多肽引入植物的方法描述于例如,美国专利号4,761,373、4,769,061、4,810,648、4,940,835、4,975,374、5,013,659、5,162,602、5,276,268、5,304,730、5,495,071、5,554,798、5,561,236、5,569,823、5,767,366、5,879,903、5,928,937、6,084,155、6,329,504以及6,337,431;以及美国专利公开号2001/0016956中。还参见,万维网上的lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/。

[0129] 赋予对抑制生长点或分生组织的除草剂(例如咪唑啉酮或磺酰脲)的抗性/耐受性的多核苷酸也可以适用于本发明的一些实施例中。对于突变型ALS和AHAS酶在这一分类号中的示例性多核苷酸如描述于例如,美国专利号5,767,366和5,928,937中。美国专利号4,

761,373和5,013,659针对抵抗不同的咪唑啉酮或磺酰脲除草剂的植物。美国专利号4,975,374涉及含有如下核酸的植物细胞和植物,所述核酸编码突变型谷氨酰胺合成酶(GS),所述突变型谷氨酰胺合成酶抵抗已知抑制GS的除草剂(例如,草胺膦和甲硫氨酸磺基脲(methionine sulfoximine))的抑制作用。美国专利号5,162,602披露了抵抗环己二酮和芳氧苯氧丙酸除草剂的抑制作用的植物。该抗性由改变的乙酰辅酶A羧化酶(ACCase)赋予。

[0130] 赋予对草甘膦抗性的由核苷酸序列编码的多肽也适用于本发明。参见例如,美国专利号4,940,835和美国专利号4,769,061。美国专利号5,554,798披露了抗草甘膦的转基因玉蜀黍植物,所述抗性由改变的5-烯醇丙酮莽草酸-3-磷酸(EPSP)合酶基因赋予。

[0131] 编码对磷酰基化合物(例如草胺膦或草胺膦、以及吡啶氧丙酸或苯氧丙酸以及环己酮)的抗性的多核苷酸也是适合的。参见欧洲专利申请号0 242246。还参见美国专利号5,879,903、5,276,268和5,561,236。

[0132] 其他合适的多核苷酸包括编码对抑制光合作用的除草剂(例如三嗪和苯基氰(腈水解酶))的抗性的那些,参见美国专利号4,810,648。编码用于除草剂抗性的另外的合适多核苷酸包括编码对2,2-二氯丙酸、烯禾啶、吡氟氯禾灵、咪唑啉酮除草剂、磺酰脲除草剂、三唑并嘧啶除草剂、均三嗪除草剂以及溴草腈的抗性的那些。同样适合的是赋予对原卟啉原氧化酶的抗性或者提供增强的对植物疾病的抗性、增强的对不利环境条件(非生物胁迫)的耐受性(这些条件包括但不限于干旱、极冷、极热、或极端的土壤盐度或极端的酸度或碱度)、以及在植物构造或发育中的改变(包括发育时间方面的变化)的多核苷酸。参见例如,美国专利公开号2001/0016956和美国专利号6,084,155。

[0133] 另外的合适的多核苷酸包括对杀有害生物(例如杀昆虫)多肽进行编码的那些。这些多肽可以按足以控制例如昆虫有害生物的量(即,昆虫控制量)进行生产。应认识到在植物中对控制昆虫或其他有害生物必要的杀有害生物多肽的生产量可以变化,这取决于栽培品种、有害生物的类型、环境因素等。有用于另外的昆虫或有害生物抗性的多核苷酸例如包括如下核苷酸序列,它们编码芽孢杆菌属(*Bacillus*)生物中鉴定到的毒素。已经克隆了包含编码来自几个亚种的苏云金芽孢杆菌(*Bt*)杀昆虫蛋白的核苷酸序列的多核苷酸,并且已经发现这些重组克隆对鳞翅目、双翅目和鞘翅目昆虫幼虫是有毒的。此类*Bt*杀昆虫蛋白的实例包括以下Cry蛋白,例如Cry1Aa、Cry1Ab、Cry1Ac、Cry1B、Cry1C、Cry1D、Cry1Ea、Cry1Fa、Cry3A、Cry9A、Cry9B、Cry9C等,连同营养期杀昆虫蛋白例如Vip1、Vip2、Vip3等。*Bt*来源的蛋白质的完整清单可以在万维网在苏塞克斯大学(University of Sussex)维护的苏云金芽孢杆菌毒素命名法数据库中找到(还参见,Crickmore et al.(1998) *Microbiol.Mol.Biol.Rev.*62:807-813[Crickmore等人(1998)微生物分子生物学综述,62:807-813])。

[0134] 适合在植物中产生的多肽进一步包括改进或通过其他方式有助于收获的植物和/或植物部分转化成为商业上有用的产品(包括例如增加的或改变的碳水化合物含量和/或分布、改进的发酵特性、增加的油含量、增加的蛋白含量、改进的消化率、以及增加的营养成分含量(例如,增加的植物甾醇含量、增加的生育酚含量、增加的甾烷醇含量和/或增加的维生素含量))的那些。感兴趣的多肽还包括例如在收获的作物中导致或促成不需要的成分,例如植酸、或降解糖的酶类的含量降低的那些。“导致(resulting in)”或“促成(contributing to)”是指这种感兴趣的多肽可以直接或间接地促成感兴趣的性状的存在

(例如,通过异源纤维素酶的使用来增加纤维素降解)。

[0135] 在一个实施例中,多肽有助于改进的食品或饲料的可消化性。木聚糖酶是半纤维素分解酶,这些酶改善了植物细胞壁的分解,这导致动物更好地利用这些植物营养素。这导致了改进的生长率和饲料转化。同样,可以减小含有木聚糖的饲料的粘度。在植物细胞内异源产生木聚糖酶也可以促进木质纤维素转化成工业加工中的可发酵糖。

[0136] 来自真菌和细菌微生物的多种木聚糖酶已经得到鉴别和表征(参见例如,美国专利号5,437,992;Coughlin et al.(1993)“Proceedings of the Second TRICEL Symposium on *Trichoderma reesei* Cellulases and Other Hydrolases”Espoo [Coughlin等人(1993)“里氏木霉纤维素酶和其他水解酶的第二TRICEL研讨会论文集”,埃斯波];Souminen and Reinikainen,eds.(1993)Foundation for Biotechnical and Industrial Fermentation Research 8:125-135[Souminen和Reinikainen编著(1993)生物技术和工业发酵研究基金会,8:125-135];美国专利公开号2005/0208178;以及PCT公开号W003/16654)。具体地说,在里氏木霉(*T.reesei*)中已经鉴别出三种特异性木聚糖酶(XYL-I、XYL-II、和XYL-III)(Tenkanen et al.(1992)Enzyme Microb.Technol.14:566 [Tenkanen等人(1992)酶与微生物技术,14:566];Torronen et al.(1992)Bio/Technology 10:1461 [Torronen等人(1992)生物/技术,10:1461];以及Xu et al.(1998)Appl.Microbiol.Biotechnol.49:718[Xu等人(1998)应用微生物与生物技术,49:718])。

[0137] 在另一个实施例中,对于本发明有用的多肽可以是多糖降解酶。产生这样的酶的本发明的植物对于产生例如用于生物加工的发酵原料会是有用的。在一些实施例中、可用于发酵过程的酶包括 α 淀粉酶、蛋白酶、支链淀粉酶、异淀粉酶、纤维素酶、半纤维素酶、木聚糖酶、环糊精糖基转移酶、脂肪酶、植酸酶、漆酶、氧化酶、酯酶、角质酶、颗粒淀粉水解酶以及其他葡糖淀粉酶。

[0138] 多糖降解酶包括:淀粉降解酶,例如 α -淀粉酶(EC 3.2.1.1)、葡糖醛酸酶(E.C.3.2.1.131);外-1,4- α -D葡聚糖酶,例如淀粉糖化酶和葡糖淀粉酶(EC 3.2.1.3)、 β -淀粉酶(EC 3.2.1.2)、 α -糖苷酶(EC 3.2.1.20)以及其他外-淀粉酶;淀粉脱支酶,例如a) 异淀粉酶(EC 3.2.1.68)、支链淀粉酶(EC 3.2.1.41)等;b) 纤维素酶,例如外-1,4-3-纤维二糖水解酶(EC3.2.1.91)、外-1,3- β -D-葡聚糖酶(EC 3.2.1.39)、 β -糖苷酶(EC3.2.1.21);c) L-阿拉伯糖酶,例如内-1,5- α -L-阿拉伯糖酶(EC3.2.1.99)、 α -阿拉伯糖苷酶(EC 3.2.1.55)等;d) 半乳聚糖酶,例如内-1,4- β -D-半乳聚糖酶(EC 3.2.1.89)、内-1,3- β -D-半乳聚糖酶(EC3.2.1.90)、 α -半乳糖苷酶(EC 3.2.1.22)、 β -半乳糖苷酶(EC 3.2.1.23)等;e) 甘露聚糖酶,例如内-1,4- β -D-甘露聚糖酶(EC 3.2.1.78)、 β -甘露糖苷酶(EC 3.2.1.25)、 α -甘露糖苷酶(EC 3.2.1.24)等;f) 木聚糖酶,例如内-1,4- β -木聚糖酶(EC 3.2.1.8)、 β -D-木糖苷酶(EC 3.2.1.37)、1,3- β -D-木聚糖酶等;以及g) 其他酶,例如 α -L-岩藻糖苷酶(EC 3.2.1.51)、 α -L-鼠李糖苷酶(EC 3.2.1.40)、果聚糖酶(EC 3.2.1.65)、菊粉酶(EC3.2.1.7)等。在一个实施例中, α -淀粉酶是描述于美国专利号8,093,453中的合成 α -淀粉酶Amy797E,将该专利通过引用以其全文结合在此。

[0139] 可以与本发明一起使用的另外的酶包括蛋白酶,如真菌和细菌蛋白酶。真菌蛋白酶包括但不限于从曲霉属(*Aspergillus*)、木霉属(*Trichoderma*)、毛霉属(*Mucor*)和根霉属(*Rhizopus*),如黑曲霉(*A.niger*)、泡盛曲霉(*A.awamori*)、米曲霉(*A.oryzae*)和米黑毛霉

(M.miehei)获得的那些。在一些实施例中,本发明的多肽可以是纤维二糖水解酶(CBH)(EC 3.2.1.91)。在一个实施例中,该纤维二糖水解酶可以是CBH1或CBH2。

[0140] 与本发明一起使用的其他酶包括但不限于半纤维素酶,如甘露聚糖酶和阿拉伯呋喃糖苷酶(EC 3.2.1.55);木质素酶;脂肪酶(例如,E.C.3.1.1.3)、葡糖氧化酶、果胶酶、木聚糖酶、转葡糖苷酶、 α 1.6葡糖苷酶(例如,E.C.3.2.1.20);酯酶,如阿魏酸酯酶(EC 3.1.1.73)和乙酰基木聚糖酯酶(EC 3.1.1.72);以及角质酶(例如E.C.3.1.1.74)。

[0141] 在一些实施例中,本发明提供了如下转基因非人类宿主细胞,该细胞包含本发明的多核苷酸、核酸分子、嵌合基因、表达盒或重组载体。转基因非人类宿主细胞可以包括但不限于植物细胞、酵母细胞、细菌细胞或昆虫细胞。因此,在一些实施例中,本发明提供了选自以下属的细菌细胞:芽孢杆菌属(*Bacillus*)、短芽孢杆菌属(*Brevibacillus*)、梭菌属(*Clostridium*)、致病杆菌属(*Xenorhabdus*)、发光杆菌属(*Photorhabdus*)、巴斯德氏芽菌属(*Pasteuria*)、埃希氏菌属(*Escherichia*)、假单胞菌属(*Pseudomonas*)、欧文氏菌属(*Erwinia*)、沙雷氏菌属(*Serratia*)、克雷伯菌属(*Klebsiella*)、沙门氏菌属(*Salmonella*)、巴氏杆菌属(*Pasteurella*)、黄单胞菌属(*Xanthomonas*)、链霉菌属(*Streptomyces*)、根瘤菌属(*Rhizobium*)、红假单胞菌属(*Rhodopseudomonas*)、嗜甲基菌属(*Methylophilus*)、农杆菌属(*Agrobacterium*)、醋杆菌属(*Acetobacter*)、乳杆菌属(*Lactobacillus*)、节杆菌属(*Arthrobacter*)、固氮菌属(*Azotobacter*)、明串珠菌属(*Leuconostoc*)或产碱杆菌属(*Alcaligenes*)。因此,例如,作为生物昆虫控制剂,本发明的Cry蛋白可以通过在细菌细胞中表达编码本发明的Cry蛋白的嵌合基因而产生。例如,在一个实施例中,提供了包含本发明的嵌合基因的苏云金芽孢杆菌细胞。

[0142] 在另外的实施例中,本发明提供了作为双子叶植物细胞或单子叶植物细胞的植物细胞。在另外的实施例中,该双子叶植物细胞选自下组,该组由以下各项组成:大豆细胞、向日葵细胞、番茄细胞、芸苔属作物细胞、棉花细胞、糖用甜菜细胞以及烟草细胞。在另外的实施例中,该单子叶植物细胞选自下组,该组由以下各项组成:大麦细胞、玉蜀黍细胞、燕麦细胞、水稻细胞、高粱细胞、甘蔗细胞以及小麦细胞。在一些实施例中,本发明提供了多个双子叶植物细胞或单子叶植物细胞,这些细胞表达由本发明的嵌合基因编码的本发明的毒性蛋白。在其他实施例中,将该多个细胞并列以形成质外体并且使其在自然光照中生长。

[0143] 在本发明的另一个实施例中,在高等生物(例如,植物)中表达本发明的毒性蛋白。在这种情况下,表达有效量的毒性蛋白的转基因植物保护自身免受植物有害生物如昆虫有害生物的伤害。在昆虫开始以这种转基因植物为食时,它也摄取了这种已表达的毒素。这可以妨碍昆虫进一步咬食植物组织或者甚至可以伤害或杀死昆虫。本发明的多核苷酸被插入表达盒中,然后该表达盒被稳定地整合到植物的基因组中。在另一个实施例中,该多核苷酸被包括在非致病性自我复制病毒中。根据本发明转化的植物可以是单子叶植物或双子叶植物,并且包括但不限于玉米(玉蜀黍)、大豆、水稻、小麦、大麦、黑麦、燕麦、高粱、粟、向日葵、红花、糖用甜菜、棉花、甘蔗、油菜、苜蓿、烟草、花生、蔬菜(包括甘薯、豆类、豌豆、菊苣、莴苣、甘蓝、花椰菜、西兰花、芜菁、胡萝卜、茄子、黄瓜、萝卜、菠菜、马铃薯、番茄、芦笋、洋葱、大蒜、瓜类、胡椒、芹菜、南瓜、西葫芦、绿皮西葫芦)、水果(包括苹果、梨、榲桲、李、樱桃、桃、蜜桃、杏、草莓、葡萄、覆盆子、黑莓、菠萝、鳄梨、番木瓜、芒果、香蕉)和特种植物如拟南芥以及木本植物如针叶树和落叶树。优选地,本发明的植物是作物植物,如玉蜀黍、高粱、小麦、向

日葵、番茄、十字花科植物、胡椒、马铃薯、棉花、水稻、大豆、糖用甜菜、甘蔗、烟草、大麦、油菜等。

[0144] 一旦所希望的多核苷酸已经被转化进特定的植物种类中,便可以使用传统的育种技术将其在该种类中繁殖或将其转移到相同种类的其他品种中,特别是包括商业品种。

[0145] 在转基因植物中表达本发明的多核苷酸,由此导致在这些转基因植物中相应Cry蛋白的生物合成。以此方式,产生在存在昆虫压力下具有增强的产量保护的转基因植物。用于它们在转基因植物中的表达,本发明的核苷酸序列可能需要修饰和优化。尽管在许多情况下,来自微生物有机体的基因能够在植物中高水平表达而无需修饰,在转基因植物中的低表达可能是由于微生物核苷酸序列的缘故,这些序列具有在植物中并不优选的密码子。在本领域中已知,活生物具有特定的密码子使用偏好,而且在本发明中所描述的这些核苷酸序列的密码子可以被改变以符合植物偏好,同时维持由其编码的氨基酸。此外,在植物(例如玉米植物)中高表达最好是由如下编码序列实现的,这些编码序列具有至少约35%、或至少约45%、或至少约50%、或至少约60%的GC含量。具有低GC含量的微生物核苷酸序列在植物中也许表达欠佳,这是由于存在着可能使信息不稳定的ATTTA基序,以及可导致不恰当的多腺苷酸化的AATAAA基序。尽管某些基因序列可以在单子叶植物和双子叶植物种类两者中充分表达,但是可以对序列进行修饰以便迎合单子叶植物或双子叶植物的特定密码子偏好以及GC含量偏好,因为这些偏好已经被证明是不同的(Murray et al. Nucl. Acids Res. 17:477-498 (1989) [Murray等人,核酸研究,17:477-498 (1989)])。此外,针对不正常剪接位点的存在来对这些核苷酸序列进行筛选,这些位点可能导致信息平截(message truncation)。使用描述于例如美国专利号5,625,136、5,500,365和6,013,523中的方法,使用熟知的定点诱变、PCR以及合成基因构建技术对在这些核苷酸序列之内所有需要做出的变化(如以上所描述的那些)进行改变。

[0146] 在一些实施例中,本发明提供了根据披露于美国专利号5,625,136中的程序制备的合成基因,将该专利通过引用结合在此。在这个操作中,使用了玉蜀黍偏好的密码子,即最频繁地编码玉蜀黍中的氨基酸的单一密码子。针对特定的氨基酸的玉蜀黍偏好的密码子可源自例如来自玉蜀黍的已知基因序列。例如,针对来自玉蜀黍植物的28个基因的玉蜀黍密码子使用发现于以下文献中:Murray et al., Nucleic Acids Research 17:477-498 (1989) [Murray等人,核酸研究,17:477-498 (1989)],将其披露内容通过引用结合在此。本发明的确切示例的用玉蜀黍优化密码子制备的合成序列由SEQ ID NO:13-20中的任一项代表。以此方式,这些核苷酸序列可以进行优化用于在任何植物中表达。应认识到,核苷酸序列的全部或任何部分可以是优化的或合成的。也就是说,多核苷酸可以包含作为部分天然序列和部分合成优化序列的核苷酸序列。

[0147] 为了有效的翻译起始,可能需要修饰与起始甲硫氨酸相邻的序列。例如,它们可以通过包含已知在植物中有效的序列而被修饰。Joshi已经提出了针对植物的适当的共有序列(NAR 15:6643-6653 (1987)),并且Clonetech提出了另一种共有翻译起始子(1993/1994目录,第210页)。这些共有序列适于与本发明的核苷酸序列一起使用。将这些序列掺入至包含核苷酸序列的构建体中,达到ATG并且包括ATG(同时保持不修饰第二氨基酸),或者可替代地达到ATG后的GTC并且包括ATG后的GTC(具有修饰该转基因的第二氨基酸的可能性)。

[0148] 本发明的新颖cry蛋白编码序列(作为它们的天然序列或作为如上所述的合成序

列)可以可操作地融合至用于在植物中表达的多种启动子(包括组成型、诱导型、时序性调节的、发育调节的、化学调节的、组织优选的以及组织特异性启动子)以制备重组DNA分子(即,嵌合基因)。启动子的选择将取决于表达的时间和空间需要而变化,并且还取决于目标种类而变化。因此,本发明的核苷酸序列在叶、柄(stalk)或茎(stem)、穗、花序(例如穗状花序、圆锥花序、穗轴等)、根、和/或籽苗中的表达是优选的。然而在许多情况下,寻求针对多于一种类型昆虫有害生物的保护,并且因此在多个组织中的表达是令人希望的。尽管已经显示来自双子叶植物的很多启动子在单子叶植物中是可操作的并且反之亦然,但理想的是选择双子叶植物启动子用于在双子叶植物中表达,并且选择单子叶植物启动子用于在单子叶植物中表达。然而,对所选择的启动子的起源并没有限制,足够的是它们在驱动核苷酸序列在所希望的细胞中的表达中是操作性的。

[0149] 在本发明中有用的组成型启动子的实例包括CaMV 35S和19S启动子(Fraley等人,美国专利号5,352,605,通过引用结合在此)。此外,启动子是从肌动蛋白基因的任一种衍生的,这些肌动蛋白基因在大多数细胞类型中被表达。由McElroy等人(Mol.Gen.Genet.231:150-160(1991)[分子遗传学与普通遗传学,231:150-160(1991)])描述的启动子表达盒可以易于被修饰用于新颖的毒素基因的表达并且特别适用于在单子叶植物宿主中使用。又另一种组成型启动子是从泛素衍生的,泛素是在许多细胞类型中积聚的已知的另一种基因产物。泛素启动子已经从一些物种中克隆用于在转基因植物中使用,例如在向日葵(Binet et al.,1991.Plant Science 79:87-94[Binet等人,1991,植物科学,79:87-94])、玉蜀黍(Christensen et al.,1989.Plant Molec.Biol.12:619-632[Christensen等人,1989,植物分子生物学,12:619-632])、以及拟南芥属(Norris et al.1993.Plant Molec.Biol.21:895-906[Norris等人,1993,植物分子生物学,21:895-906])中使用。玉蜀黍泛素启动子已经在转基因单子叶植物系统中得到发展,并且它的序列以及构建用于单子叶植物转化的载体披露于专利公开EP 0 342 926中。泛素启动子适用于新颖的毒素基因在转基因植物(尤其是单子叶植物)中的表达。

[0150] 对于在植物(特别是玉蜀黍)中表达本发明的新颖cry蛋白编码序列有用的组织特异性或组织优先启动子是指导在根、髓、叶或花粉中的表达的那些。此类启动子披露于美国专利号5625136中,通过引用以其全文结合在此。在本发明中有用的其他组织特异性启动子包括棉花二磷酸核酮糖羧化酶(rubisco)启动子,披露于美国专利号6,040,504中;水稻蔗糖合成酶启动子,披露于美国专利号5,604,121中;以及夜香树黄化曲叶病毒启动子,披露于美国专利号7,166,770中,所有这些专利都通过引用以其全文结合。对于指导新颖毒素基因在植物中的表达有用的化学诱导型启动子披露于美国专利号5,614,395中(通过引用以其全文结合在此)。

[0151] 本发明的核苷酸序列也可以在被化学调节的启动子的调节下进行表达。这使得本发明的Cry蛋白能够仅在用诱导化学品对作物植物进行处理时被合成。用于基因表达的化学诱导的此类技术的实例详述于公开申请EP 0 332104和美国专利号5,614,395中。在一个实施例中,该化学调节的启动子是烟草PR-1a启动子。

[0152] 另一类在本发明中有用的启动子是创伤可诱导的启动子。已经描述了数量众多的在创伤部位并且还在植物病原菌感染的部位表达的启动子。理想的是,这样的启动子在昆虫入侵的部位应该仅有局部活性,并且以此方式这些杀昆虫蛋白仅在需要合成这些杀昆虫

蛋白的细胞中积聚以杀死入侵的昆虫有害生物。这类启动子的实例包括由以下文献所描述的那些:Stanford et al.*Mol.Gen.Genet.*215:200-208(1989) [Stanford等人,分子遗传学与普通遗传学,215:200-208(1989)];Xu et al.*Plant Molec.Biol.*22:573-588(1993) [Xu等人,植物分子生物学,22:573-588(1993)];Logemann et al.*Plant Cell* 1:151-158(1989) [Logemann等人,植物细胞,1:151-158(1989)];Rohrmeier&Lehle,*Plant Molec.Biol.*22:783-792(1993) [Rohrmeier和Lehle,植物分子生物,22:783-792(1993)];Firek et al.*Plant Molec.Biol.*22:129-142(1993) [Firek等人,植物分子生物学,22:129-142(1993)]以及Warner et al.*Plant J.*3:191-201(1993) [Warner等人,植物杂志,3:191-201(1993)]。

[0153] 导致在本发明中有用的组织特异性表达模式的启动子的非限制性实例包括绿色组织特异性的、根特异性的、茎特异性的和/或花特异性的。适用于在绿色组织中表达的启动子包括调节参与光合作用的基因的许多启动子,并且这些中的许多已经从单子叶植物和双子叶植物两者中得以克隆。一种这样的启动子是来自磷酸烯醇羧化酶基因的玉米PEPC启动子(Hudspeth&Gruha,*Plant Molec.Biol.*12:579-589(1989) [Hudspeth和Gruha,植物分子生物学,12:579-589(1989)])。另一种用于根特异性表达的启动子是由以下文献描述的启动子:de Framond(*FEBS* 290:103-106(1991) [de Framond(*FEBS* 290:103-106(1991))]或美国专利号5,466,785。另一种在本发明中有用的启动子是描述于美国专利号5,625,136中的茎特异性启动子,它天然地驱动玉蜀黍trpA基因的表达。

[0154] 除了选择适合的启动子之外,用于在植物中表达杀昆虫毒素的构建体还需要适当的可操作地连接在异源核苷酸序列下游的转录终止子。一些此类的终止子是可获得的并且在本领域中是已知的(例如来自CaMV的tml,来自rbcS的E9)。任何已知在植物中发挥功能的可供使用的终止子均可以在本发明的上下文下使用。

[0155] 可以将数量众多的其他序列掺入本发明所描述的表达盒中。这些序列包括已经显示出增强表达的序列,如内含子序列(例如,来自Adh1和bronze1)以及病毒的前导序列(例如,来自TMV、MCMV、和AMV)。

[0156] 本发明的核苷酸序列在植物中针对不同的细胞定位的靶向表达可能是更优选的。在一些情况下,在胞质溶胶中的定位可能是令人希望的,而在其他情况下,在某个亚细胞器中的定位可能是优选的。用于靶向例如植物中的基因产物的任何机构都可以用于实践本发明,并且已知此类机构存在于植物中并且已经相当详细地表征了控制这些机构的功能的序列。已经表征了导致将基因产物靶向其他细胞区室的序列。氨基末端序列可以负责将感兴趣的蛋白质靶向任何细胞区室,如植物的液泡、线粒体、过氧化物酶体、蛋白体、内质网、叶绿体、淀粉颗粒、淀粉体、质外体或细胞壁(例如Unger et.al.*Plant Molec.Biol.*13:411-418(1989) [Unger等人,植物分子生物学,13:411-418(1989)];Rogers et.al.(1985)*Proc.Natl.Acad.Sci.USA* 82:6512-651[Rogers等人(1985)美国国家科学院院刊,82:6512-651];美国专利号7,102,057;WO 2005/096704,将其全部通过引用而特此结合)。任选地,信号序列可以是来自waxy的N-末端信号序列、来自 γ -玉米蛋白的N-末端信号序列、淀粉结合结构域、C-末端淀粉结合结构域、将成熟蛋白引入叶绿体的叶绿体靶向序列(Comai et.al.(1988)*J.Biol.Chem.*263:15104-15109[Comai等人(1988)生物化学杂志,263:15104-15109];van den Broeck,et.al.(1985)*Nature* 313:358-363[van den Broeck等人

(1985)自然,313:358-363];美国专利号5,639,949)或来自糊粉细胞的分泌信号序列(Koehler&Ho,Plant Cell 2:769-783(1990)[Koehler和Ho,植物细胞,2:769-783(1990)])。另外,与羧基末端序列结合的氨基末端序列负责基因产物的液泡靶向(Shinshi et.al.(1990)Plant Molec.Biol.14:357-368[Shinshi等人(1990)植物分子生物学,14:357-368])。在一个实施例中,所选择的信号序列包括已知的切割位点,并且构建的融合体考虑了在一个或多个切割位点之后的需要切割的任何氨基酸。在一些情况下,这个要求可以通过在切割位点与转基因ATG之间添加小数目氨基酸,或可替代地置换转基因序列内的一些氨基酸来满足。这些构建技术在本领域是熟知的并且同样适用于任何细胞区室。

[0157] 应认识到,用于细胞靶向的上述机构不仅可以与其同源启动子结合使用,还可以与异源启动子结合使用,从而在启动子的转录调节下实现特定的细胞靶向目标,该启动子具有不同于自其衍生靶向信号的启动子的表达谱。

[0158] 植物转化

[0159] 用于转化植物的程序在本领域中是熟知且常规的并且普遍描述于文献中。用于植物转化的方法的非限制性实例包括通过以下各项的转化:细菌介导的核酸递送(例如,经由农杆菌)、病毒介导的核酸递送、碳化硅或核酸须晶介导的核酸递送、脂质体介导的核酸递送、微注射、微粒轰击、磷酸钙介导的转化、环糊精介导的转化、电穿孔、纳米粒子介导的转化、超声处理、渗入、PEG介导的核酸吸收、以及使得核酸引入到植物细胞中的任何其他电学的、化学的、物理的(机械的)和/或生物的机制,包括其任何组合。对于本领域已知的不同植物转化方法的一般指导包括以下文献:Miki等人(“Procedures for Introducing Foreign DNA into Plants”in Methods in Plant Molecular Biology and Biotechnology, Glick,B.R.and Thompson,J.E.,Eds.(CRC Press,Inc.,Boca Raton,1993),pages 67-88 [在Glick,B.R.和Thompson,J.E.编辑的植物分子生物学与生物技术方法中的“用于将外来DNA引入植物中的程序”(CRC出版公司,波卡拉顿,1993),第67-88页])和Rakowoczy-Trojanowska(Cell.Mol.Biol.Lett.7:849-858(2002)[细胞与分子生物学快报,7:849-858(2002)])。

[0160] 对于农杆菌介导的转化,二元载体或携带至少一个T-DNA边界序列的载体是适合的,而对于直接基因转移(例如,微粒轰击等),任何载体都是适合的,并且仅含有感兴趣的构建体的线性DNA可以是优选的。在直接基因转移的情况下,可以使用以单个DNA种类的转化或共转化(Schocher et al.,Biotechnology 4:1093-1096(1986)[Schocher等人,生物技术,4:1093-1096(1986)])。对于直接基因转移以及农杆菌介导的转移二者,转化通常(但不是必需的)用如下选择性标记进行,该选择性标记可以是正向选择(磷甘露糖异构酶),提供对抗生素(卡那霉素、潮霉素或甲氨蝶呤)或除草剂(草甘膦或草丁膦)的抗性。然而,选择性标记的选择对于本发明并不是至关重要的。

[0161] 农杆菌介导的转化是用于转化植物(特别是双子叶植物)的常用方法,这是由于其高转化效率并且由于其在许多不同物种中的广泛实用性。农杆菌介导的转化典型地涉及将携带感兴趣的外来DNA的二元载体转移至适当的农杆菌菌株,这可能取决于由宿主农杆菌菌株或者在共同存在的Ti质粒上或染色体地携带的vir基因的互补体(Uknes et al.(1993)Plant Cell 5:159-169[Uknes等人(1993)植物细胞,5:159-169])。将该重组二元载体转移至农杆菌可以使用携带该重组二元载体的大肠杆菌、辅助大肠杆菌菌株(该辅助菌

株携带能够将该重组二元载体移动到目标农杆菌菌株中的质粒)通过三亲本交配程序实现。可替代地,可以通过核酸转化将该重组二元载体转移至农杆菌中(Höfgen&Willmitzer (1988)Nucleic Acids Res.16:9877[Höfgen和Willmitzer(1988)核酸研究,16:9877])。

[0162] 通过重组农杆菌进行的植物转化通常涉及该农杆菌与来自该植物的外植体的共培养,并且遵循本领域熟知的方法。在携带位于这些二元质粒T-DNA边界之间的抗生素或除草剂抗性标记的选择培养基上对转化的组织进行再生。

[0163] 如先前所讨论的,另一种用于转化植物、植物部分和植物细胞的方法涉及在植物组织和细胞上推进惰性或生物活性的粒子。参见例如,美国专利号4,945,050、5,036,006和5,100,792。通常,这种方法涉及在有效于穿透该细胞的外表面并提供掺入在其内部中的条件下在植物细胞处推进惰性或生物活性的粒子。当使用惰性粒子时,可以通过用含有感兴趣的核酸的载体包被这些粒子而将该载体引入该细胞中。可替代地,一个或多个细胞可以被该载体围绕使得该载体通过该粒子的激发而被带入该细胞中。也可以将生物活性的粒子(例如,干酵母细胞、干细菌或噬菌体,各自含有一种或多种被试图引入的核酸)推进到植物组织中。

[0164] 在另一个实施例中,本发明的多核苷酸可以被直接转化进质体基因组中。质体转化的主要优点在于质体通常能够表达细菌基因而无需实质性的修饰,而且质体能够在单一启动子的控制下表达多个开放阅读框。在以下文献中广泛描述了质体转化技术:美国专利号5,451,513、5,545,817和5,545,818;PCT申请号W0 95/16783;以及McBride et al. (1994)Proc.Natl.Acad.Sci.USA 91,7301-7305[McBride等人(1994)美国国家科学院院刊,91,7301-7305]。基本的叶绿体转化技术涉及将位于选择性标记侧翼的经克隆的质体DNA区连同感兴趣的基因一起引入适合的靶组织中,这是例如使用生物射弹(biolytic)或原生质体转化(例如,氯化钙或PEG介导的转化)来进行的。这些1至1.5kb的侧翼区(被命名为靶向序列)促进了与质体基因组的同源重组,并且因而允许置换或修饰原质体(plastome)的特定区域。最初,可以将叶绿体16S rRNA和rps12基因(赋予针对大观霉素和/或链霉素的抗性)的点突变用作供转化用的选择性标记(Svab,Z.,Hajdukiewicz,P.,and Maliga,P. (1990)Proc.Natl.Acad.Sci.USA 87,8526-8530[Svab,Z.,Hajdukiewicz,P.和Maliga,P. (1990)美国国家科学院院刊,87,8526-8530];Staub,J.M.,and Maliga,P. (1992)Plant Cell 4,39-45[Staub,J.M.和Maliga,P. (1992)植物细胞,4,39-45])。在这些标记之间克隆位点的存在允许建立质体靶向载体用于外来基因的引入(Staub,J.M.,and Maliga,P. (1993)EMBO J.12,601-606[Staub,J.M.和Maliga,P. (1993)欧洲分子生物学杂志,12,601-606])。转化效率的实质性增加可以通过用显性的选择性标记(对大观霉素解毒酶氨基糖苷-3'-腺苷转移酶进行编码的细菌aadA基因)置换隐性的rRNA或r蛋白抗生素抗性基因而获得(Svab,Z.,and Maliga,P. (1993)Proc.Natl.Acad.Sci.USA 90,913-917[Svab,Z.和Maliga,P. (1993)美国国家科学院院刊,90,913-917])。先前,这种标记已经被成功地用于莱茵衣藻(*Chlamydomonas reinhardtii*)这种绿藻的质体基因组的高频率转化(Goldschmidt-Clermont,M. (1991)Nucl.Acids Res.19:4083-4089[Goldschmidt-Clermont,M. (1991)核酸研究,19:4083-4089])。有用于质体转化的其他选择性标记在本领域是已知的,并且被包括在本发明的范围之内。典型地,转化之后需要大约15-20个细胞分裂循环以便达到同质状态。质体表达(其中基因通过同源重组被插入到在每个植物细胞中

存在的所有数千个环状质体基因组的拷贝中)利用了超过核表达的基因的庞大的拷贝数目的优点,以便允许能够很容易超过总的可溶性植物蛋白的10%的表达水平。在一个实施例中,可以将本发明的多核苷酸插入质体靶向载体中并转化进所希望的植物宿主的质体基因组中。因此,可以获得与含有本发明的核苷酸序列的质体基因组同型的植物,这些植物能够高表达该多核苷酸。

[0165] 选择转化的转基因植物、植物细胞和/或植物组织培养物的方法在本领域中是常规的,并且可以用于在此提供的本发明的方法中。例如,本发明的重组载体还可以包括包含用于选择性标记的核苷酸序列的表达盒,该选择性标记可以用于选择转化的植物、植物部分和/或植物细胞。如在此使用的,“选择性标记(selectable marker)”意指如下核苷酸序列,当该核苷酸序列表达时向表达该标记的植物、植物部分和/或植物细胞赋予不同的表型,并且因此允许此类转化的植物、植物部分和/或植物细胞与不具有该标记的那些区别开来。这样的核苷酸序列可以编码可选择的或可筛选的标记,这取决于该标记是否赋予可以通过化学手段而被选择的性状,如通过使用选择剂(例如,抗生素、除草剂等),或者取决于该标记是否仅是人们可以通过观察或测试而鉴别的性状,如通过筛选(例如,R基因座性状)。当然,适合的选择性标记的许多实例在本领域中是已知的并且可以用于在此描述的表达盒中。

[0166] 选择性标记的实例包括但不限于编码neo或nptII的核苷酸序列,它赋予对卡那霉素、G418等的抗性(Potrykus et al. (1985) Mol. Gen. Genet. 199:183-188 [Potrykus等人(1985) 分子遗传学与普通遗传学, 199:183-188]); 编码bar的核苷酸序列,它赋予对草丁膦的抗性; 编码改变的5-烯醇丙酮莽草酸-3-磷酸(EPSP) 合酶的核苷酸序列,它赋予对草甘膦的抗性(Hinchee et al. (1988) Biotech. 6:915-922 [Hinchee等人(1988) 生物技术, 6:915-922]); 编码腈水解酶(如来自臭鼻克雷伯菌(*Klebsiella ozaenae*)的bxn)的核苷酸序列,它赋予对溴草腈的抗性(Stalker et al. (1988) Science 242:419-423 [Stalker等人(1988) 科学, 242:419-423]); 编码改变的乙酰乳酸合酶(ALS)的核苷酸序列,它赋予对咪唑啉酮、磺酰脲或其他ALS-抑制化学品的抗性(欧洲专利申请号154204); 编码甲氨蝶呤-抗性的二氢叶酸还原酶(DHFR)的核苷酸序列(Thillet et al. (1988) J. Biol. Chem. 263:12500-12508 [Thillet等人(1988) 生物化学杂志, 263:12500-12508]); 编码茅草枯脱卤素酶的核苷酸序列,它赋予对茅草枯的抗性; 编码甘露糖-6-磷酸异构酶(也称为磷酸甘露糖异构酶(PMI))的核苷酸序列,它赋予代谢甘露糖的能力(美国专利号5,767,378和5,994,629); 编码改变的邻氨基苯甲酸盐合酶的核苷酸序列,它赋予对5-甲基色氨酸的抗性; 和/或编码hph的核苷酸序列,它赋予对潮霉素的抗性。本领域技术人员能够选择用于在本发明的表达盒中使用的适合的选择性标记。

[0167] 另外的选择性标记包括但不限于编码 β -葡糖醛酸酶的核苷酸序列或编码对于多种显色底物已知的酶的uidA (GUS); 编码在植物组织中对花色苷色素(红色)进行调节的产物的R基因座核苷酸序列(Dellaporta et al., “Molecular cloning of the maize R-nj allele by transposon-tagging with Ac” 263-282 In: Chromosome Structure and Function: Impact of New Concepts, 18th Stadler Genetics Symposium (Gustafson & Appels eds., Plenum Press 1988) [Dellaporta等人, “染色体结构与功能: 新概念的影响中的通过Ac转座子标签技术对玉蜀黍R-nj等位基因的分子克隆”, 第18届斯特德莱遗传学专

题讨论会 (Gustafson和Appels编著,Plenum出版社,1988)] ; 编码 β -内酰胺酶的核苷酸序列,对于 β -内酰胺酶而言多种显色底物是已知的(例如,PADAC,一种显色头孢菌素)(Sutcliffe(1978)Proc.Natl.Acad.Sci.USA 75:3737-3741[Sutcliffe(1978)美国国家科学院院刊,75:3737-3741]); 编码xylE的核苷酸序列,xylE编码儿茶酚双加氧酶(Zukowsky et al.(1983)Proc.Natl.Acad.Sci.USA 80:1101-1105[Zukowsky等人(1983)美国国家科学院院刊,80:1101-1105]); 编码酪氨酸酶的核苷酸序列,酪氨酸酶能够氧化酪氨酸成为DOPA和多巴醌,其进而缩合形成黑色素(Katz et al.(1983)J.Gen.Microbiol.129:2703-2714[Katz等人(1983)普通微生物学杂志,129:2703-2714]); 编码 β -半乳糖苷酶的核苷酸序列,对于 β -半乳糖苷酶而言存在显色底物; 编码荧光素酶(lux)的核苷酸序列,荧光素酶允许生物发光检测(Ow et al.(1986)Science 234:856-859[Ow等人(1986)科学,234:856-859]); 编码水母发光蛋白的核苷酸序列,水母发光蛋白可以在钙敏感的生物发光检测中采用(Prasher et al.(1985)Biochem.Biophys.Res.Comm.126:1259-1268[Prasher等人(1985)生物化学与生物物理学研究通讯,126:1259-1268]); 或编码绿色荧光蛋白的核苷酸序列(Niedz et al.(1995)Plant Cell Reports 14:403-406[Niedz等人(1995)植物细胞报道,14:403-406])。本领域技术人员能够选择用于在本发明的表达盒中使用的适合的选择性标记。

[0168] 此外,如本领域中所熟知的,完整的转基因植物可以使用多种已知技术中的任何技术从转化的植物细胞、植物组织培养物和/或培养的原生质体再生而来。例如在以下文献中描述了从植物细胞、植物组织培养物和/或培养的原生质体进行的植物再生:Evans et al.(Handbook of Plant Cell Cultures,Vol.1,MacMilan Publishing Co.New York (1983)) [Evans等人(植物细胞培养手册,第1卷,麦克米兰出版公司,纽约(1983))]; 和 Vasil I.R.(ed.)(Cell Culture and Somatic Cell Genetics of Plants,Acad.Press, Orlando,Vol.I(1984),and Vol.II(1986)) [Vasil I.R.(编著)(植物的细胞培养和体细胞遗传学,学术出版社,奥兰多,第I卷(1984)和第II卷(1986))]

[0169] 另外,工程化进以上所述的本发明的转基因种子和植物、植物部分和/或植物细胞中的遗传特性可以通过有性生殖或营养生长来传递,并且因此可以在子代植物中维持并传代。通常,维持和传代利用了被开发以适合特定目的(如收获、播种或耕作)的已知农业方法。

[0170] 因此,可以按本领域熟知的任何数目的方法(如上所述的)将多核苷酸引入该植物、植物部分和/或植物细胞中。因此,没有依赖于将一种或多种多核苷酸引入植物中的具体方法,相反可以使用允许将该一种或多种多核苷酸稳定地整合到该植物的基因组中的任何方法。在有待引入一种以上多核苷酸的情况下,这些对应的多核苷酸可以作为单一核酸分子的一部分、或者作为分开的核酸分子而进行组装,并且可以位于相同的或不同的核酸分子上。因此,可以在单一转化事件中、在分开的转化事件中、或者例如作为育种方案的一部分在植物中,将这些多核苷酸引入感兴趣的细胞中。

[0171] 本发明的另外的实施例包括从本发明的转基因植物和/或其部分产生的收获产物以及从所述收获产物产生的加工产物。收获产物可以是如在此描述的全株或任何植物部分。因此,在一些实施例中,收获产物的非限制性实例包括种子、果实、花或其部分(例如,花药、柱头等)、叶、茎等。在其他实施例中,加工产物包括但不限于从收获的本发明的种子或

其他植物部分产生的细粉、粗粉、油、淀粉、谷物等,其中所述种子或其他植物部分包含本发明的核酸分子/多核苷酸/核苷酸序列。

[0172] 在其他实施例中,本发明提供了来自本发明的转基因种子和/或转基因植物的提取物,其中该提取物包含本发明的核酸分子、多核苷酸、核苷酸序列或毒性蛋白。可以根据本领域熟知的程序制备来自植物或植物部分的提取物(参见,de la Torre et al.,Food, Agric. Environ. 2(1):84-89(2004) [de la Torre等人,食品农业与环境,2(1):84-89(2004)]; Guidet, Nucleic Acids Res. 22(9):1772-1773(1994) [Guidet, 核酸研究,22(9):1772-1773(1994)]; Lipton et al., Food Agric. Immun. 12:153-164(2000) [Lipton等人,食品农业通讯,12:153-164(2000)])。

[0173] 杀昆虫组合物

[0174] 在一些实施例中,本发明提供了杀昆虫组合物,该组合物包含农业上可接受的载体中的本发明的Cry蛋白。如在此使用的,“农业上可接受的载体”可以包括与该活性组分结合以有助于它施用于植物或其部分的天然或合成、有机或无机材料。农业上可接受的载体的实例包括但不限于粉剂、尘剂、丸剂、颗粒剂、喷雾剂、乳剂、胶体以及溶液。农业上可接受的载体进一步包括但不限于可用于农业配制品中的惰性组分、分散剂、表面活性剂、佐剂、增粘剂、粘着剂、粘合剂或其组合。此类组合物可以按使杀昆虫蛋白或其他有害生物控制剂与这些有害生物接触的任何方式施用。因此,可以将这些组合物施用于植物或植物部分的表面,包括种子、叶、花、茎、块茎、根等。另一种农业上可接受的载体可以是转基因植物或植物部分。

[0175] 在另外的实施例中,该杀昆虫组合物包含本发明的转基因细菌细胞,其中该细菌细胞包含本发明的嵌合基因。例如,这样的杀昆虫组合物可以通过脱水、冷冻干燥、均化、萃取、过滤、离心、沉降或浓缩包含本发明的多核苷酸的苏云金芽孢杆菌细胞的培养物来制备。在另外的实施例中,该组合物包含按重量计从约1%至约99%的本发明的Cry蛋白。

[0176] 本发明的Cry蛋白可以与其他有害生物控制剂组合使用,以增加有害生物目标范围或用于预防和/或管理昆虫抗性。因此,在一些实施例中,本发明提供了控制一种或多种植物有害生物的组合,其中该组合物包含本发明的第一Cry蛋白和不同于该第一Cry蛋白的第二有害生物控制剂。在其他实施例中,该组合物是用于局部施用至植物的配制品。在仍其他实施例中,该组合物是转基因植物。在另外的实施例中,该组合物是局部施用至转基因植物的配制品的组合。在一个实施例中,当该转基因植物包含第二有害生物控制剂时,该配制品包含本发明的第一Cry蛋白。在另一个实施例中,当该转基因植物包含本发明的第一Cry蛋白时,该配制品包括第二有害生物控制剂。

[0177] 在一些实施例中,该第二有害生物控制剂可以是选自下组的试剂,该组由以下各项组成:化学杀有害生物剂、苏云金芽孢杆菌(Bt)杀昆虫蛋白、致病杆菌属杀昆虫蛋白、发光杆菌属杀昆虫蛋白、侧孢短芽孢杆菌(*Brevibacillus laterosporus*)杀昆虫蛋白、球形芽孢杆菌(*Bacillus sphaericus*)杀昆虫蛋白、蛋白酶抑制剂(丝氨酸和半胱氨酸类型两者)、凝集素、 α -淀粉酶、过氧化物酶以及胆固醇氧化酶。

[0178] 在其他实施例中,该第二有害生物控制剂是选自下组的化学杀有害生物剂,该组由以下各项组成:拟除虫菊酯、氨基甲酸酯、新烟碱、神经元钠通道阻断剂、杀昆虫大环内酯、 γ -氨基丁酸(GABA)拮抗剂、杀昆虫脲以及保幼激素模拟物。在另一个实施例中,该化学

杀有害生物剂选自下组,该组由以下各项组成:阿巴美丁、乙酰甲胺磷、啶虫脒、磺胺螨酯(amidoflumet) (S-1955)、除虫菌素(avermectin)、印楝素、甲基谷硫磷、联苯菊酯、联苯腈酯(binfenazate)、噻嗪酮、克百威、溴虫腈、定虫隆、毒死蜱、甲基毒死蜱、环虫酰肼、噻虫胺、氟氯氰菊酯、 β -氟氯氰菊酯、三氯氟氰菊酯、 λ -三氯氟氰菊酯、氯氰菊酯、灭蝇胺、溴氰菊酯、杀螨隆、二嗪磷、除虫脲、乐果、苯虫醚、甲氨基阿维菌素、硫丹、高氰戊菊酯、乙虫腈、苯硫威(fenothicarb)、苯氧威、甲氰菊酯、唑螨酯、氰戊菊酯、氟虫腈、氟啶虫酰胺、氟氰戊菊酯、 τ -氟胺氰菊酯、啉虫胺(UR-50701)、氟虫脲、地虫硫磷、氯虫酰肼、氟铃脲、吡虫啉、茚虫威、异柳磷、虱螨脲、马拉硫磷、聚乙醛、甲胺磷、杀扑磷、灭多威、烯虫酯、甲氧氯、久效磷、甲氧虫酰肼、噻虫啉(nithiazin)、双苯氟脲、多氟脲(XDE-007)、杀线威、对硫磷、甲基对硫磷、氯菊酯、甲拌磷、伏杀磷、亚胺硫磷、磷胺、抗蚜威、丙溴磷、吡蚜酮、啶虫丙醚、蚊蝇醚、鱼藤酮、多杀菌素、螺甲螨酯(spiromesifin) (BSN 2060)、硫丙磷、虫酰肼、伏虫隆、七氟菊酯、特丁硫磷、杀虫畏、噻虫啉、噻虫嗪、硫双威、杀虫双(thiosultap-sodium)、四溴菊酯、敌百虫和杀铃脲、涕灭威、杀线威、苯线磷、双甲脒、灭螨猛、乙酯杀螨醇、三环锡、三氯杀螨醇、除螨灵、依杀螨、啉啉醚、苯丁锡、甲氰菊酯、唑螨酯、噻嗪酮、克螨特、啉啉灵以及吡啉胺。在另一个实施例中,该化学杀有害生物剂选自下组,该组由以下各项组成:氯氰菊酯、三氯氟氰菊酯、氟氯氰菊酯和 β -氟氯氰菊酯、高氰戊菊酯、氰戊菊酯、四溴菊酯、苯硫威、灭多威、杀线威、硫双威、噻虫胺、吡虫啉、噻虫啉、茚虫威、多杀菌素、阿巴美丁、除虫菌素(avermectin)、苯虫醚、硫丹、乙虫腈、氟虫腈、氟虫脲、杀铃脲、苯虫醚、蚊蝇醚、吡蚜酮以及双甲脒。

[0179] 在另外的实施例中,该第二有害生物控制剂可以是任何数目的苏云金芽孢杆菌杀昆虫蛋白中的一种或多种,包括但不限于Cry蛋白、营养期杀昆虫蛋白(VIP)以及任何前述杀昆虫蛋白的杀昆虫嵌合体。在其他实施例中,该第二有害生物控制剂是选自下组的Cry蛋白,该组由以下各项组成:Cry1Aa、Cry1Ab、Cry1Ac、Cry1Ad、Cry1Ae、Cry1Af、Cry1Ag、Cry1Ah、Cry1Ai、Cry1Aj、Cry1Ba、Cry1Bb、Cry1Bc、Cry1Bd、Cry1Be、Cry1Bf、Cry1Bg、Cry1Bh、Cry1Bi、Cry1Ca、Cry1Cb、Cry1Da、Cry1Db、Cry1Dc、Cry1Dd、Cry1Ea、Cry1Eb、Cry1Fa、Cry1Fb、Cry1Ga、Cry1Gb、Cry1Gc、Cry1Ha、Cry1Hb、Cry1Hc、Cry1Ia、Cry1Ib、Cry1Ic、Cry1Id、Cry1Ie、Cry1If、Cry1Ig、Cry1Ja、Cry1Jb、Cry1Jc、Cry1Jd、Cry1Ka、Cry1La、Cry1Ma、Cry1Na、Cry1Nb、Cry2Aa、Cry2Ab、Cry2Ac、Cry2Ad、Cry2Ae、Cry2Af、Cry2Ag、Cry2Ah、Cry2Ai、Cry2Aj、Cry2Ak、Cry2Al、Cry2Ba、Cry3Aa、Cry3Ba、Cry3Bb、Cry3Ca、Cry4Aa、Cry4Ba、Cry4Ca、Cry4Cb、Cry4Cc、Cry5Aa、Cry5Ab、Cry5Ac、Cry5Ad、Cry5Ba、Cry5Ca、Cry5Da、Cry5Ea、Cry6Aa、Cry6Ba、Cry7Aa、Cry7Ab、Cry7Ac、Cry7Ba、Cry7Bb、Cry7Ca、Cry7Cb、Cry7Da、Cry7Ea、Cry7Fa、Cry7Fb、Cry7Ga、Cry7Gb、Cry7Gc、Cry7Gd、Cry7Ha、Cry7Ia、Cry7Ja、Cry7Ka、Cry7Kb、Cry7La、Cry8Aa、Cry8Ab、Cry8Ac、Cry8Ad、Cry8Ba、Cry8Bb、Cry8Bc、Cry8Ca、Cry8Da、Cry8Db、Cry8Ea、Cry8Fa、Cry8Ga、Cry8Ha、Cry8Ia、Cry8Ib、Cry8Ja、Cry8Ka、Cry8Kb、Cry8La、Cry8Ma、Cry8Na、Cry8Pa、Cry8Qa、Cry8Ra、Cry8Sa、Cry8Ta、Cry9Aa、Cry9Ba、Cry9Bb、Cry9Ca、Cry9Da、Cry9Db、Cry9Dc、Cry9Ea、Cry9Eb、Cry9Ec、Cry9Ed、Cry9Ee、Cry9Fa、Cry9Ga、Cry10Aa、Cry11Aa、Cry11Ba、Cry11Bb、Cry12Aa、Cry13Aa、Cry14Aa、Cry14Ab、Cry15Aa、Cry16Aa、Cry17Aa、Cry18Aa、Cry18Ba、Cry18Ca、Cry19Aa、Cry19Ba、Cry19Ca、Cry20Aa、Cry20Ba、Cry21Aa、Cry21Ba、Cry21Ca、Cry21Da、Cry21Ea、Cry21Fa、Cry21Ga、Cry21Ha、Cry22Aa、Cry22Ab、Cry22Ba、Cry22Bb、Cry23Aa、Cry24Aa、Cry24Ba、Cry24Ca、Cry25Aa、Cry26Aa、Cry27Aa、Cry28Aa、Cry29Aa、

Cry29Ba、Cry30Aa、Cry30Ba、Cry30Ca、Cry30Da、Cry30Db、Cry30Ea、Cry30Fa、Cry30Ga、Cry31Aa、Cry31Ab、Cry31Ac、Cry31Ad、Cry32Aa、Cry32Ab、Cry32Ba、Cry32Ca、Cry32Cb、Cry32Da、Cry32Ea、Cry32Eb、Cry32Fa、Cry32Ga、Cry32Ha、Cry32Hb、Cry32Ia、Cry32Ja、Cry32Ka、Cry32La、Cry32Ma、Cry32Mb、Cry32Na、Cry32Oa、Cry32Pa、Cry32Qa、Cry32Ra、Cry32Sa、Cry32Ta、Cry32Ua、Cry33Aa、Cry34Aa、Cry34Ab、Cry34Ac、Cry34Ba、Cry35Aa、Cry35Ab、Cry35Ac、Cry35Ba、Cry36Aa、Cry37Aa、Cry38Aa、Cry39Aa、Cry40Aa、Cry40Ba、Cry40Ca、Cry40Da、Cry41Aa、Cry41Ab、Cry41Ba、Cry42Aa、Cry43Aa、Cry43Ba、Cry43Ca、Cry43Cb、Cry43Cc、Cry44Aa、Cry45Aa、Cry46Aa、Cry46Ab、Cry47Aa、Cry48Aa、Cry48Ab、Cry49Aa、Cry49Ab、Cry50Aa、Cry50Ba、Cry51Aa、Cry52Aa、Cry52Ba、Cry53Aa、Cry53Ab、Cry54Aa、Cry54Ab、Cry54Ba、Cry55Aa、Cry56Aa、Cry57Aa、Cry57Ab、Cry58Aa、Cry59Aa、Cry59Ba、Cry60Aa、Cry60Ba、Cry61Aa、Cry62Aa、Cry63Aa、Cry64Aa、Cry65Aa、Cry66Aa、Cry67Aa、Cry68Aa、Cry69Aa、Cry69Ab、Cry70Aa、Cry70Ba、Cry70Bb、Cry71Aa、Cry72Aa以及Cry73Aa。

[0180] 在另外的实施例中,该第二有害生物控制剂是选自下组的Vip3营养期杀昆虫蛋白,该组由以下各项组成:Vip3Aa1、Vip3Aa2、Vip3Aa3、Vip3Aa4、Vip3Aa5、Vip3Aa6、Vip3Aa7、Vip3Aa8、Vip3Aa9、Vip3Aa10、Vip3Aa11、Vip3Aa12、Vip3Aa13、Vip3Aa14、Vip3Aa15、Vip3Aa16、Vip3Aa17、Vip3Aa18、Vip3Aa19、Vip3Aa20、Vip3Aa21、Vip3Aa22、Vip3Aa23、Vip3Aa24、Vip3Aa25、Vip3Aa26、Vip3Aa27、Vip3Aa28、Vip3Aa29、Vip3Aa30、Vip3Aa31、Vip3Aa32、Vip3Aa33、Vip3Aa34、Vip3Aa35、Vip3Aa36、Vip3Aa37、Vip3Aa38、Vip3Aa39、Vip3Aa40、Vip3Aa41、Vip3Aa42、Vip3Aa43、Vip3Aa44、Vip3Ab1、Vip3Ab2、Vip3Ac1、Vip3Ad1、Vip3Ad2、Vip3Ae1、Vip3Af1、Vip3Af2、Vip3Af3、Vip3Ag1、Vip3Ag2、Vip3Ag3HM117633、Vip3Ag4、Vip3Ag5、Vip3Ah1、Vip3Ba1、Vip3Ba2、Vip3Bb1、Vip3Bb2以及Vip3Bb3。

[0181] 在另外的实施例中,在转基因植物中共表达本发明的第一Cry蛋白和该第二有害生物控制剂。可以通过将植物遗传工程化以含有并表达所有的必需基因来实现一种以上杀有害生物成分在同一个转基因植物中的共表达。可替代地,可以将植物(亲本1)遗传工程化,用于本发明的Cry蛋白的表达。可以将第二植物(亲本2)遗传工程化,用于第二有害生物控制剂的表达。通过将亲本1与亲本2杂交,获得了表达被引入至亲本1和亲本2中的所有基因的子代植物。

[0182] 在另外的实施例中,提供了产生对于至少黑色地老虎(小地老虎)有毒的蛋白质的方法,该方法包括培养包含本发明的多核苷酸或嵌合基因或核酸分子或重组载体的转基因非人类宿主细胞,在该宿主产生对于至少黑色地老虎(小地老虎)有毒的蛋白质的条件下。在一些实施例中,该转基因非人类宿主细胞是植物细胞。在一个实施例中,该植物细胞是玉蜀黍细胞。在其他实施例中,植物细胞或玉蜀黍细胞在其下生长的条件包括自然光照。在其他实施例中,该转基因非人类宿主细胞是细菌细胞。在仍其他实施例中,该转基因非人类宿主细胞是酵母细胞。

[0183] 在其他实施例中,所产生的蛋白质对至少一种另外的昆虫有杀昆虫活性,其中该另外的昆虫选自下组,该组由以下各项组成:欧洲玉米蛀虫(欧洲玉米螟)、秋黏虫(草地贪夜蛾)、玉米穗蛾(玉米穗虫)、甘蔗螟(小蔗螟)、绒毛豆毛虫(黎豆夜蛾)、大豆夜蛾(大豆尺

蠖),西南玉米蛀虫(西南玉米螟)、西部豆切根虫(西部豆夜蛾)、烟夜蛾(烟芽夜蛾)、亚洲玉米蛀虫(亚洲玉米螟)、棉螟蛉(棉铃虫)、条纹蛀茎虫(二化螟)、粉蛀茎虫(非洲大螟)或水稻卷叶螟(稻纵卷叶螟)及其任何组合。

[0184] 在其他实施例中,该嵌合基因包含SEQ ID NO:1-4中的任一项。在仍其他实施例中,所产生的蛋白质包含SEQ ID NO:13-16中任一项的氨基酸序列。

[0185] 在一些实施例中,该嵌合基因包含密码子优化用于在植物中表达的核苷酸序列。在其他实施例中,该嵌合基因包含SEQ ID NO:5-12中的任一项。在另外的实施例中,所产生的蛋白质包含SEQ ID NO:13-20中任一项的氨基酸序列。

[0186] 在另外的实施例中,本发明提供了产生抗有害生物(例如,抗昆虫)转基因植物的方法,该方法包括向植物中引入包含编码本发明的Cry蛋白的核苷酸序列的本发明的多核苷酸、嵌合基因、重组载体、表达盒或核酸分子,其中该核苷酸序列被表达于该植物中,由此赋予该植物对至少欧洲玉米蛀虫的抗性,并且产生抗有害生物(例如,抗昆虫)转基因植物。在一些实施例中,与缺乏本发明的多核苷酸、嵌合基因、重组载体、表达盒或核酸分子的对照植物相比,抗有害生物转基因植物对至少黑色地老虎(小地老虎)有抗性。在一些实施例中,通过转化植物实现引入。在其他实施例中,通过使包含本发明的嵌合基因、重组载体、表达盒或核酸分子的第一植物与不同的第二植物杂交来实现引入。

[0187] 在一些实施例中,对至少黑色地老虎(小地老虎)有抗性的本发明的转基因植物还对至少一种另外的昆虫有抗性,其中该另外的昆虫包括但不限于欧洲玉米蛀虫(欧洲玉米螟)、秋黏虫(草地贪夜蛾)、玉米穗蛾(玉米穗虫)、甘蔗螟(小蔗螟),绒毛豆毛虫(黎豆夜蛾)、大豆夜蛾(大豆尺蠖),西南玉米蛀虫(西南玉米螟)、西部豆切根虫(西部豆夜蛾)、烟夜蛾(烟芽夜蛾)、亚洲玉米蛀虫(亚洲玉米螟)、棉螟蛉(棉铃虫)、条纹蛀茎虫(二化螟)、粉蛀茎虫(非洲大螟)或水稻卷叶螟(稻纵卷叶螟)及其任何组合。

[0188] 在另外的实施例中,提供了控制至少黑色地老虎(小地老虎)昆虫的方法,该方法包括向这些昆虫递送有效量的本发明的Cry蛋白。为了有效,该Cry蛋白首先被昆虫经口摄取。然而,该Cry蛋白可以按许多公认的方式被递送至该昆虫。用于将蛋白质经口递送至昆虫的方式包括但不限于将该蛋白质提供于(1)转基因植物中,其中该昆虫取食(摄取)该转基因植物的一个或多个部分,由此摄取在该转基因植物中表达的多肽;(2)一种或多种配制的蛋白质组合物中,它们可以被施用至或掺入例如昆虫生长介质中;(3)一种或多种蛋白质组合物中,它们可以被施用至表面,例如喷雾在植物部分的表面,然后当该昆虫取食喷雾的一个或多个植物部分时组合物被该昆虫摄取;(4)饵基;(5)经注射进该昆虫;或(6)任何其他本领域公认的蛋白质递送系统。因此,可以使用经口递送至昆虫的任何方法来递送本发明的毒性Cry蛋白。在一些具体实施例中,将本发明的Cry蛋白经口递送至昆虫,其中该昆虫摄取转基因植物的一个或多个部分。

[0189] 在其他实施例中,将本发明的Cry蛋白经口递送至昆虫,其中该昆虫摄取用包含本发明的Cry蛋白的组合物喷雾的植物的一个或多个部分。可以使用本领域技术人员已知的用于将化合物、组合物、配制品等施用于植物表面的任何方法将本发明的组合物递送至植物表面。递送至或接触植物或其部分的一些非限制性实例包括喷雾、撒粉、喷洒、分散、下雾、雾化、撒播、浸泡、土壤注入、土壤掺入、浸透(例如,根、土壤处理)、浸渍、灌注、涂覆、叶或茎渗透、侧施或种子处理等及其组合。用于使植物或其部分与一种或多种化合物、一种或

多种组合物或一种或多种配制品接触的这些和其他程序是本领域技术人员熟知的。

[0190] 在一些实施例中,本发明涵盖为农民提供控制鳞翅目昆虫有害生物的手段的方法,该方法包括向该农民供应或出售植物材料如种子,该植物材料包含能够表达本发明的Cry蛋白的多核苷酸、嵌合基因、表达盒或重组载体,如上所述的。

[0191] 本发明的实施例可以通过参考以下实例而被更好地理解。前述的和以下的本发明的实施例以及各种实施例的描述不是旨在限制权利要求书,而是对其具有说明性。因此,应理解的是权利要求书不旨在限制这些实例的具体细节。本领域技术人员应理解的是本发明的其他实施例可以在不偏离本披露的精神和范围的情况下进行实践,本披露的范围是由所附权利要求书限定的。实例

[0192] 实例1. 活性Bt菌株的鉴别

[0193] 将来自当前收集物中存在的孢子的苏云金芽孢杆菌分离株在T3+盘尼西林琼脂平板上进行培养和维持。在28°C下使每种分离株需氧地生长在24孔深块中约10天直至孢子形成,将其通过用考马斯蓝/乙酸染色并且用显微镜可视化来验证。孢子形成之后,将可溶的和不可溶的部分都针对感兴趣的鳞翅目种类的活性进行测试。在表面污染生物测定中测试各部分,在该生物测定中将各部分覆盖到多种类人工饲料上。针对至少四种鳞翅目种类(包括玉米穗虫(玉米穗蛾)、小地老虎(黑色地老虎)、欧洲玉米螟(欧洲玉米蛀虫)和草地贪夜蛾(秋黏虫)筛选每种分离物,其中样本量为12只新生幼虫。每个测定的持续时间为在室温下约7天;针对死亡率以及幼虫生长抑制对这些平板评分。将在相对于阴性对照观察到的死亡率增加30%时认为是有效的。基于初始昆虫测试,选择菌株C0633、C2080、M0262和M1455用于进一步分析。

[0194] 实例2. Bt基因的分离和测序

[0195] Fosmid基因组文库构建:对于在实例1中鉴别的一些Bt菌株,使用Park等人(FEMS Microbiol. Lett. 284:28-34 (2008) [FEMS微生物学简讯, 284:28-34 (2008)])描述的fosmid文库方法来分离编码推定活性蛋白质的基因。使用CopyControl™ Fosmid Library Production Kit (Epicentre公司, 麦迪逊, 威斯康星州)根据制造商方案构建fosmid文库。简言之,将来自每个Bt菌株的纯化DNA(约0.5μg)经酶处理以对平端进行修端(end repair),并且然后连接到fosmid载体pCC1FOS(Epicentre公司)中。在体外包装进λ噬菌体和感染大肠杆菌(E. coli) EPI1300-T1®后,将细菌细胞铺板在含有12.5μg/ml氯霉素的Luria-Bertani (LB)上。在选择菌落之前,将各板在约37°C下孵育24h。将经转染的大肠杆菌菌落转移到含有150μl含氯霉素的LB培养基的96孔平板中,并在37°C下孵育24h。

[0196] 菌落杂交筛选:以300cfu/100x 15mm L-琼脂加15μg/ml氯霉素平板的密度铺板fosmid文库。总计铺板3000板fosmid。使用Immobilon-Ny+87mm过滤圆片(EMD密理博公司(Millipore), 比勒利卡, 马萨诸塞州)进行膜杂交。如下完成菌落转移:将滤器置于各板上约5min,然后使用镊子,将滤器从琼脂表面转移并将菌落向上放置在使用0.5M NaOH浸泡5min的Whatman滤纸上。然后将菌落滤器放置于2X SSC中浸泡5分钟的Whatman滤纸上。用设定为2000x 100μJ的UV Stratalinker®(Stratagene公司, 拉荷亚, 加利福尼亚州)将DNA固定在膜上。然后将滤器在Whatman滤纸上进行空气干燥。如供应商所述的,将滤器预杂交,并在65°C下在250mM NaPO4 (pH 7.0)、7% SDS、1% BSA中杂交。将杂交滤器在65°C下在2X SSC、0.5% SDS中洗涤30min,随后在65°C下在0.2X SSC、0.2% SDS中洗涤30min。在-80°C下用增感屏将

滤器暴露于X射线胶片(Kodak[®] BIOMAX[®] XAR, 飞世尔科技公司(Fisher Scientific), 匹兹堡, 宾夕法尼亚州)过夜。将阳性菌落影印(patch)到L琼脂(加有15 μ g/ml氯霉素)上。

[0197] 杂交探针:设计PCR引物以从指定为C0633的Bt菌株的基因组DNA扩增cry9B样基因的720bp片段。引物对包括具有序列AAACATGAACCGAAATAATCAAAATG (SEQ ID NO:21)的指定为0AR2613a的正向引物和具有序列ATCCGTCCCTTGTGCGTGTA (SEQ ID NO:22)的指定为0AR2615a的反向引物。PCR反应在以下循环条件下运行:[94 $^{\circ}$ C, 5min], 12x[94 $^{\circ}$ C, 30sec, 57 $^{\circ}$ C至51 $^{\circ}$ C, 每次循环下降0.5 $^{\circ}$ C, 30sec, 72 $^{\circ}$ C 2.5min], 以及35x[94 $^{\circ}$ C, 30sec, 52 $^{\circ}$ C, 30sec, 72 $^{\circ}$ C, 2.5min]。该反应含有1X One Taq[®]缓冲液(新英格兰生物实验室(New England Biolabs), 贝弗利, 马萨诸塞州)、200 μ M dNTP、80ng DNA、2.5U One Taq[®] DNA聚合酶、50ng各引物和无菌蒸馏水至50 μ l总反应。

[0198] 在含有溴化乙锭的1%琼脂糖TAE凝胶上分离得到的扩增子。在UV光下观察扩增子,并从凝胶中切出。使用凝胶提取试剂盒如供应商(凯杰公司(Qiagen), 巴伦西亚, 加利福尼亚州)所述的分离DNA。使用Rediprime II随机引物标记系统(通用电气医疗集团(GE Healthcare), 匹兹堡, 宾夕法尼亚州),用EasyTide(α -32P)dCTP 3000Ci/mmol(珀金埃尔默公司(Perkin Elmer), 沃尔瑟姆, 马萨诸塞州)标记探针。使用Micro Bio-Spin30色谱柱(伯乐公司(BioRad), 赫拉克勒斯(Hercules), 加利福尼亚州)除去未掺入的核苷酸。将探针在添加至杂交溶液之前于95 $^{\circ}$ C下加热5min。

[0199] Bt基因测序:遵循制造商的说明书(凯杰公司)制备2-4个独立克隆的DNA制剂。根据制造商的说明书,使用BigDye[™] Terminator Kit(应用生物系统公司(Applied Biosystems)),福斯特城,加利福尼亚州)进行用针对感兴趣的预测核苷酸序列的两条链设计的引物的测序反应。将反应产物在ABI373或ABI377测序仪上进行电泳。使用Phred/Phrap/Consed软件包(华盛顿大学)对所有测序数据进行分析,其误差率在共有序列水平下等于或小于 10^{-4} 。将序列用程序Sequencher[™](4.7版,基因密码公司(Gene Codes Corp.),安阿伯,密歇根州)进行组装。每个基因被测序到4X覆盖度。

[0200] 实例3. Bt基因克隆和合成

[0201] 设计Cry9特异性引物对以促进cry9型基因的鉴别和克隆。将引物对设计成伴随加入PmeI限制性位点与cry9型基因的5'末端杂交,并伴随加入AscI限制性位点与3'末端杂交。用于扩增5'末端的引物对包括具有序列GTTTAAACATGAATCGAAATAATCAAAATG (SEQ ID NO:23)的正向引物和具有序列GGCGCGCCCTACTCTTGTGTTTCAATAAA (SEQ ID NO:24)的反向引物。用于扩增3'末端的引物对包括具有序列GTTTAAACATGAATCAAAATAAACACGGA (SEQ ID NO:25)的正向引物和具有序列GGCGCGCCTTACTGTTGGGTTTCCATGAACT (SEQ ID NO:26)的反向引物。插入的限制性位点在相应的引物中加下划线。使用以下循环条件进行PCR反应:[94 $^{\circ}$ C, 5min]和30x[94 $^{\circ}$ C, 30sec, 45 $^{\circ}$ C, 30sec, 72 $^{\circ}$ C, 3.5min]。该反应含有1X OneTaq缓冲液、200 μ M dNTP、80ng DNA、2.5U OneTaq DNA聚合酶(新英格兰生物实验室)、50 μ g各引物和无菌蒸馏水至50 μ l总反应。

[0202] 如供应商(生命技术公司(Life Technologies))所述的,将所得扩增子克隆到TOP0 pCR 4.0载体中。如供应商(新英格兰生物实验室)所述的,将分离的质粒DNA用PmeI和

AscI消化。

[0203] 将PmeI/AscI片段克隆到设计用于在大肠杆菌和苏云金芽孢杆菌中表达的指定为pCIB5634`的穿梭载体中。将pCIB5634`载体用PmeI和AscI消化。将消化的载体和基因片段通过在基于1%琼脂糖Tris乙酸盐EDTA缓冲液的凝胶上跑胶而纯化。从凝胶中切出各片段,并使用QIAGEN凝胶提取试剂盒如供应商所述的进行清理。使用来自新英格兰生物实验室的连接试剂盒如供应商所述的将各片段连接在一起。如供应商所述的,将连接反应转化到TOP10细胞(生命技术公司)中,并将其铺板在含有100mg/ml氨苄青霉素的L-琼脂上。从单菌落中分离质粒DNA,并将鉴别的克隆再次测序到2X覆盖度以证实序列正确。

[0204] 选择用于重组产生但未直接克隆出基因组DNA的一些Bt基因被提交给第三销售方用于全基因合成。将这些合成的Bt基因亚克隆到上述穿梭载体中,用于随后的表达和测试进一步的生物活性。

[0205] 实例4.基因组组装与分析

[0206] 使用全基因组测序方法鉴别本发明的一些Bt基因。简言之,使用Covaris S2超声波装置(Covaris公司,沃本,马萨诸塞州)剪切芽孢杆菌属DNA,其中将程序DNA_400bp设为工作循环:10%;强度:4;循环/脉冲:200。将DNA用NEBNext[®] Ultra[™]End Repair/dA-加尾模块(新英格兰生物实验室公司(New England Biolabs,Inc.),伊普斯维奇,马萨诸塞州)处理。如供应商(新英格兰生物实验室公司,伊普斯维奇,马萨诸塞州)所述的,使用NEB Quick Ligation[™]连接生物科技(Biooscience)有索引的1-57个适配子(1-27巴西,28-57美国、英国和瑞士)。如供应商(贝克曼库尔特公司(Beckman Coulter,Inc.),印第安纳波利斯,印第安纳州)所述的,使用Agencourt AMPure XP珠粒清理连接物。

[0207] 将文库如下进行大小分级:将50uL样品与45uL 75%珠粒混合物(25%AMPure珠粒加75%NaCl/PEG溶液TekNova目录号P4136)混合。搅拌混合物并置于磁性支架上。将所得上清液转移至新孔中并且添加45uL 50%珠粒混合物(50%AMPure珠粒加50%NaCl/PEG溶液TekNova目录号P4136)。搅拌此混合物并置于磁性支架上。除去所得上清液并且用80%乙醇洗涤珠粒。添加25uL的洗脱缓冲液(EB)并且将混合物置于磁性支架上。除去所得最终上清液并置于1.5mL管中。这个方法产生了525个DNA碱基对(bp)(插入物加适配子)大小范围内的文库。

[0208] 使用KAPA Biosystem HiFi Hot Start(KAPA生物系统公司(Kapa Biosystems, Inc.),威尔明顿,马萨诸塞州)使用以下循环条件扩增大小确定的DNA文库:[98°C,45s];12x[98°C,15s,60°C,30s,72°C,30s];[72°C,1min]。每个反应含有:5uL DNA文库、1uL生物科技通用引物(25uM)、18uL无菌水、1uL生物科技有索引的引物(25uM)、25uL2X KAPA HiFi聚合酶。

[0209] 使用高灵敏度芯片在Agilent 2100Bioanalyzer(安捷伦科技公司(Agilent Technologies),圣克拉拉,加利福尼亚州)上跑文库,以确定文库大小范围和平均插入物大小。使用标准的制造商测序方案(亿明达公司(Illumina,Inc.),圣地亚哥,加利福尼亚州)在HiSeq 2500测序系统上针对配对末端(PE)测序(100个循环/读数;12-24个文库/泳道)处理所有文库。

[0210] 开发用以鉴别和表征可能的毒性基因的芽孢杆菌属计算分析工具来优先化引导物,用于进一步实验室测试。

[0211] 上述基因组组装与分析以及基因组文库分析在苏云金芽孢杆菌菌株中鉴别出四个Cry9样基因,它们对至少黑色地老虎(小地老虎)有毒性。这些Cry9样基因和蛋白质的鉴别特征示于表1中。

[0212] 表1.苏云金芽孢杆菌菌株中鉴别的Cry9样基因。

菌株	基因/蛋白质名称	分子量(kD)	核苷酸SEQ ID NO:	氨基酸SEQ ID NO:
[0213] C2080	BT0044	127.4	1	13
C0633	BT0051	129.6	2	14
M0262	BT0068	132.3	3	15
M1455	BT0128	132.7	4	16

[0214] 实例5. BT0044、BT0051、BT0068和BT0128与已知Bt Cry蛋白的同源性

[0215] 利用本发明蛋白质的氨基酸序列搜索蛋白质数据库揭示它们与已知的杀昆虫蛋白同源。使用BLAST算法将本发明蛋白质的氨基酸序列与由NCBI维护的非冗余(nr)数据库进行比较,揭示以下蛋白质与本发明的序列具有一批(block)最强的氨基酸一致性(表2)。

[0216] 表2.本发明的Cry蛋白与已知Cry蛋白的一致性百分比。

[0217]

	一致性百分比								
	Cry9Aa1	Cry9Ba1	Cry9Bb1	Cry9Ca1	Cry9Da1	Cry9Db1	Cry9Ea1	Cry9Fa1	Cry9Ga1
0044	73	56	50	52	53	52	52	50	47
0051	56	62	61	98	69	68	70	64	35
0068	54	70	77	69	66	68	69	62	36
0128	60	76	71	69	66	67	68	62	35

[0218] 实例6.重组宿主细胞中的Bt蛋白表达

[0219] 芽孢杆菌属表达。将感兴趣的基因通过上述含有适当的Cry蛋白启动子和红霉素抗性标记的pCIB5634`表达载体在不具有可观察到的鞘翅目或鳞翅目活性的无晶芽孢杆菌菌株中进行表达。通过电穿孔将构建体转化到宿主菌株中,并且随后在含有红霉素的琼脂平板上进行选择。使这些重组菌株在T3培养基中于28℃下生长4-5天至孢子形成阶段。收获细胞沉淀并且在溶解于含有2mM DTT的高pH碳酸盐缓冲液(50mM)中之前反复洗涤。

[0220] 大肠杆菌表达。使用pET28a或pET29a载体(EMD密理博公司)在各个大肠杆菌菌株中表达感兴趣的基因。通过电穿孔转化构建体,并且随后在含有卡那霉素的琼脂平板上进行选择。使这些重组菌株生长并且使用IPTG诱导在28℃下诱导表达。将细胞再悬浮于含有2mM DTT的高pH碳酸盐缓冲液(50mM)中并且然后使用Microfluidics LV-1匀浆器打碎。

[0221] 表达分析。然后通过离心澄清所得的细胞裂解液(来自任一宿主),并通过SDS-PAGE和电泳图谱(BioRad Experion)分析样品的纯度。经由Bradford或Thermo660测定确定总蛋白浓度。然后在生物测定中测试经纯化的Cry蛋白。

[0222] 实例7. Cry蛋白在生物测定中的活性

[0223] 使用本领域公认的人工饲料生物测定法针对以下昆虫有害生物种类中的一种或多种测试实例6中产生的蛋白质:秋黏虫(FAW;草地贪夜蛾)、玉米穗蛾(CEW;玉米穗虫)、欧洲玉米蛀虫(ECB;欧洲玉米螟)、黑色地老虎(BCW;小地老虎)、甘蔗螟(SCB;小蔗螟)、绒毛豆毛虫(VBC;黎豆夜蛾)、大豆夜蛾(SBL;大豆尺夜蛾(Pseudoplusia includens))、西南玉米

蛀虫 (SWCB; 西南玉米螟)、西部豆切根虫 (WBCW; 西部豆夜蛾)、烟夜蛾 (TBW; 烟芽夜蛾)、亚洲玉米蛀虫 (ACB; 亚洲玉米螟)、棉螟蛉 (CBW; 棉铃虫)、条纹蛀茎虫 (SSB; 二化螟)、粉蛀茎虫 (PSB; 水稻大螟 (*Sesamia inferens*)) 和水稻卷叶螟 (RLF; 稻纵卷叶螟)。

[0224] 向24孔平板中的人工昆虫饲料 (Bioserv公司, 弗伦奇敦, 新泽西州) 的表面施加等量的溶液中的蛋白质。在饲料表面干燥之后, 向每个孔中添加有待测试的昆虫种类的幼虫。将这些平板密封并且保持在就温度、光照以及相对湿度而言的环境实验室条件下。阳性对照组由暴露于非常具活性且广谱的野生型芽孢杆菌属菌株的幼虫组成。阴性对照组由暴露于仅用缓冲溶液处理的昆虫饲料的幼虫和未处理的昆虫饲料 (即只有饲料) 上的幼虫组成。约120小时之后评估死亡率并且相对于对照评分。

[0225] 结果示于表3中, 其中“-”意指与对照相比无活性, “+/-”意指与对照相比0-10%的活性 (此类别还包括具有强烈幼虫生长抑制的0%死亡率), “+”意指与对照相比10%-25%的活性, “++”意指与对照相比25%-75%的活性, 并且“+++”意指与对照相比75%-100%的活性。

[0226] 表3. Cry蛋白的生物测定结果。

昆虫	BT 蛋白			
	0044	0051	0068	0128
FAW	-	-	-	-
CEW	+	-	-	+
ECB	-	+++	-	+/-
BCW	+	+++	+++	+++
SCB	+/-	+++	-	+/-
VBC		+++	+++	+++
SBL		-	+++	+++
SWCB	-	+++	+	++
WBCW	-	-		
TBW			+++	+++
ACB		+++		
CBW		+/-		
SSB		+		
PSB		+		
RLF		+++		

[0227] 实例8. Cry蛋白在模拟胃液测定中的结局

[0229] 某些Cry蛋白已经在植物中表达, 并且来自这样的植物的种子每年出售给农民用于控制各种昆虫有害生物。使这样的自我保护的杀有害生物产品经历各种监管机构的审查和注册, 包括例如美国环保署 (EPA)。

[0230] 膳食暴露是人类可以暴露于转基因植物中表达的Cry蛋白的主要途径。哺乳动物急性经口毒性和蛋白质消化率是EPA对人体健康风险评估的终点。Cry蛋白安全性的进一步科学证据在于, 它们已显示使用模拟胃液在体外迅速降解。用代表性Cry1、Cry2和Cry3蛋白进行的七次体外测定的结果确立这些蛋白质通常在30秒内迅速降解。这些结果支持更广泛的结论, 即Cry蛋白的这些组的成员 (共享显著的氨基酸序列一致性) 在人类摄入后可能迅速降解。另一方面考虑是Cry蛋白是否可能引起过敏反应。所证实的Cry蛋白在体外迅速降解应使这种发生的可能性降至最低。相比之下, 食物过敏原通常保留在体外胃肠道模型中,

而无过敏史的常见食物蛋白在模拟胃液中迅速降解 (Metcalf et al.1996[Metcalf等人,1996])。

[0231] 通过对模拟胃液 (SGF) 中蛋白质消化率的分析可以获得对蛋白质潜在过敏原性的另外的见解。在迄今为止已经测试的转基因植物中表达的几乎所有Cry蛋白都被快速消化,并且因此已被确定为不会引起过敏。然而,发现在称为Starlink的转基因玉米产品中发现的Cry9C蛋白对SGF部分稳定。虽然Starlink Cry9C对动物无毒性,但部分消化率和部分加工稳定性的特性使得EPA很难绝对排除Starlink Cry9C蛋白可成为食物过敏原的可能性,最终导致开发Starlink的公司从美国市场召回产品。

[0232] 目前,不存在用于确定新颖蛋白的致敏可能性的确定性测试。因此,EPA使用证据权衡法,其中考虑以下因素:性状来源;与已知过敏原的氨基酸序列比较;以及该蛋白质的生物化学性质,包括模拟胃液 (SGF) 中的体外消化率和糖基化。

[0233] 模拟胃液 (SGF) 测定在代表哺乳动物上消化道的严格控制的条件下测量测试蛋白的体外消化率。简言之,在37°C下经一小时的时间段,将细菌产生的测试Cry蛋白(在0.5-5mg/ml的浓度下)以10单位的胃蛋白酶活性/ μg 测试蛋白的比率暴露于胃蛋白酶(来自猪胃粘膜,溶解于2mg/ml NaCl中,pH 1.2)。将样品在1、2、5、10、30和60分钟时取出,并通过加入预热的(95°C-2分钟)终止缓冲液(65%0.5M碳酸氢钠(pH 11)、35%Tricine加样缓冲液)立即淬灭,以使胃蛋白酶立即失活,并返回加热再持续5分钟。一旦测定完成,便通过SDS-PAGE在10%-20%Tris-Tricine凝胶(肽可见低至1kDa)上检查时间点样品和对照(只有测试蛋白,只有胃蛋白酶)以跟踪由胃蛋白酶进行的消化的动力学和水平。

[0234] SGF测定的结果表明本发明的所有Cry蛋白质都非常迅速地降解。这些结果提供了如下证据,即尽管本发明的Cry蛋白与Cry9蛋白家族有关,但是与某些公开的结果相比,例如Starlink中的Cry9C,它们对SGF测定的响应是非常不同的,这表明在蛋白质中的关键胃蛋白酶切割位点具有显著的结构差异。这些结果进一步表明本发明的Cry蛋白的致敏可能性是极小的。

[0235] 实例9.BT-0051的诱变

[0236] 预测蛋白质中的抗原区域有助于合成可引发与完整蛋白质有反应性的抗体且区分紧密相关的蛋白质的肽的合理方法。对于这个实例,天然BT-0051的氨基酸序列(SEQ ID NO:6)被叠加到Cry8Ea1蛋白(登录号3EB7;worldwide web.rcsb.org/pdb/的Protein Databank;还参见Berman et al.,2000.Nuc.Acids Res.28:235-242[Berman等人,2000,核酸研究,28:235-242])的晶体结构上,并且使用载体NTI 8.0(赛默飞世尔科技公司(Thermo Fisher Scientific,Inc.),沃尔瑟姆,马萨诸塞州;还参见Welling et al.1985.FEBS Lett.188:215-218[Welling等人,1985,欧洲生化学会联合会快报,188:215-218])将预测的抗性区域映射到该结构上。选择适合的诱变区域由选取结构域I外的非保守区中的环结构域组成。已知参与Cry蛋白受体识别的环如预测参与蛋白酶活化的任何残基一样从选择中被淘汰。这留下了一个由SEQ ID NO:6的342-354代表的诱变区域。基于以下预期选取变化L350I、N351Q和T354S(SEQ ID NO:18),即它们将导致相对于天然BT-0051的最小结构变化或功能变化。这样的变化产生允许突变型BT-0051(mBT-0051;SEQ ID NO:18)与天然BT-0051(SEQ ID NO:14)和其他相关Cry9蛋白区分开的抗原区域。

[0237] 实例10.针对植物表达的基因定向

[0238] 在植物中表达之前,在自动化基因合成平台(金斯瑞公司(Genscript, Inc.),皮斯卡塔韦,新泽西州)上合成包含编码Bt Cry蛋白BT-0044、BT-0051、BT-0068和BT-0128中每种核苷酸序列(分别为SEQ ID NO:5-8)的合成多核苷酸以及包含编码突变型Bt Cry蛋白mBT-0044、mBT-0051、mBT-0068和mBT-0128中每种核苷酸序列(分别为SEQ ID NO:17-20)的合成多核苷酸。用于这个实例,制备包含可操作地连接至Cry蛋白编码序列(该编码序列可操作地连接至NOS终止子)的玉蜀黍泛素启动子(Ubi1)的第一表达盒,并且制备包含可操作地连接至磷酸甘露糖异构酶(PMI)编码序列(该编码序列可操作地连接至NOS终止子)的Ubi1启动子的第二表达盒。PMI的表达允许在甘露糖上正向选择转基因植物。将两个表达盒克隆进适于农杆菌介导的玉蜀黍转化的载体中。

[0239] 实例11.Cry蛋白在植物中的表达

[0240] 未成熟的玉蜀黍胚的转化基本上如在以下文献中描述的来进行:Negrotto et al.,2000,Plant Cell Reports 19:798-803[Negrotto等人,2000,植物细胞报告,19:798-803]内。简言之,使包含描述于实例12中的载体的农杆菌菌株LBA4404(pSB1)在28°C下在YEP(酵母提取物(5g/L)、蛋白胨(10g/L)、NaCl(5g/L)、15g/l琼脂,pH 6.8)固体培养基上生长2-4天。将大约 0.8×10^9 个农杆菌细胞悬浮于补充有 $100 \mu\text{M}$ As的LS-inf培养基中。在这个培养基中对细菌预诱导大约30-60分钟。

[0241] 将来自自交玉蜀黍系的未成熟胚从8-12天大的穗中切除到液体LS-inf+ $100 \mu\text{M}$ As中。用新鲜的感染培养基漂洗这些胚。然后添加农杆菌溶液,并且将这些胚涡旋30秒并且允许其与细菌一起沉降5分钟。然后将这些胚盾片向上地转移到LSA培养基中,并且在暗处培养两到三天。随后,将每皮氏板(petri plate)大约20与25个之间的胚转移至补充有头孢噻肟(250mg/l)和硝酸银(1.6mg/l)的LSDc培养基中,并且在大约28°C下在黑暗中培养10天。

[0242] 将产生胚性愈伤组织的未成熟胚转移至LSD1M0.5S培养基中。在这种培养基上对培养物进行持续大约6周的选择,在约3周时进行传代培养步骤。将存活的愈伤组织转移至补充有甘露糖的Reg1培养基中。在光照中(16小时光照/8小时黑暗方案)培养之后,然后将绿色组织转移至没有生长调节剂的Reg2培养基中并且孵育约1-2周。将这些小植株转移至含有Reg3培养基的Magenta GA-7盒(Magenta公司,芝加哥,伊利诺伊州)中并使其在光照中生长。约2-3周之后,通过PCR测试植物的PMI基因和Bt cry基因的存在。将来自PCR测定的阳性植物转移至温室用于进一步评估。

[0243] 在叶切除生物测定中,针对拷贝数(通过Taqman分析确定)、蛋白质表达水平(通过ELISA确定)和对抗感兴趣的昆虫种类的效应对转基因植物进行评估。确切地说,从单拷贝事件(V3-V4阶段)中切取叶组织并且用新生幼虫侵染,然后在室温下孵育5天。叶圆片生物测定的样品量取决于所测试的昆虫种类(欧洲玉米蛀虫(ECB), $n=10$;玉米穗蛾(CEW), $n=3$,黑色地老虎(BCW), $n=5$)而变化。在大约第3天和第5天读取评估组织损伤和死亡率的读数;样品相对于阴性对照的损伤使用以下量表进行评定:“+”: $<5\%$ 组织损伤,所有幼虫死亡;“+/-”: $5\%-20\%$ 组织损伤,所有幼虫死亡;或“-”: $>20\%$ 组织损伤,一些幼虫活着和/或进展到2龄。

[0244] 转基因植物组织生物测定的结果证实当在转基因植物中表达时本发明的Cry蛋白对于目标昆虫是有毒的。例如,在用本发明的嵌合基因稳定转化的玉蜀黍中表达的mBT-0051对于至少黑色地老虎(小地老虎)连同亚洲玉米蛀虫(亚洲玉米螟)、条纹蛀茎虫(二化

螟) 和棉螟蛉 (棉铃虫) 是有活性的。

序列表

<110> Syngenta Participations AG

Bramlett, Matthew Richard

Seguin, Katherine

Kramer, Vance Cary

Rose, Mark Scott

<120> 用于控制植物有害生物的组合物和方法

<130> 80668-WO-REG-ORG-P-1

<160> 26

<170> PatentIn version 3.5

<210> 1

<211> 3444

<212> DNA

<213> 苏云金芽孢杆菌

<400> 1

```

atggatttag acggtataaa aactgaaact gagactgaaa ttgtaaattg ttccgaaagt 60
agtatcgatc catcaagcgt gtcttatgcg ggaaataaca gctattcttc cgctttgaat 120
ctcaattctt gtcaaaacag agggattgca cagtgggtta atacgcttgg aggtgcaatc 180
ggtcaggctg tatccatagg aacatccatc atttccttgc ttgcggcgcc tacgcttact 240
ggaagtattt cgtagcttt taatcttata aggagaatgg ggacaggcag taatggaagc 300
tctatttctg acttgatcaat atgtgactta ctatccataa ttaatttacg tgtaagtcaa 360
gctgtattga acgacgggat tgcagatttt aacggctcag tggctgtata tgatctctat 420
ttgcatgctt tacgcagttg gaacaataac ctaaatgctg ctaccgcgga ggaacttcgc 480
actcgtttct gtattgcaga ttccgaattc gaacgtatct taacgcgggg gtccttgaca 540
catggtgggt cattagctag acaagatgct caagtgttac tgttaccttc ttttghaaat 600
gctgcctatc ttcatctact tatattaagg gatgctagca gatatggggc tagctggggc 660
ttgtttaata cgacaccaca tatcaattat ccagtaagat tacaacaact tattggttct 720
tatacccatt actgcacaca ttggtataat caaggtttaa atgaaatcag acaacgaggc 780
aatacggctg tcaattgggt ggaatttcat agatacagaa gagatatgac attaatggta 840
ctagatgttg tgcattatt ttccagcgtt gatactataa ggtatccgaa cgcaaccgtt 900
gtccaattaa gcagaaccgt ttatacagac ccgattgggt ttgtaaactg tggaagcggc 960
aacagattaa gctggtttga ttggcgggat caagctaatt tttcaacgct agaaagtgaa 1020
atgccaacce cctcgtctcc actttctttg aatcatatga gtatatttac gggteccctt 1080
actttacctg tctctcctaa taccataga gccagggtat ggtatggcaa tcaaaatag 1140
ttcacaacgg gtagtcaaaa ttcaggtcaa acaacaaact ctattcaaaa catttcgggt 1200
ttagaatat ttagaataga ttctcaagcc tgtaactctaa acaataatc gtatggcgtg 1260
aaccgagctg aattttttca tggcgetagt cagggtctcc aaagatctgt ttatcaaggc 1320
tatattagac aaagtggatt ggacaaccgc gtagttatga atcttcaaag ctttttacct 1380

```

ggcgaaaatt cagcgacacc aaccgcacaa gattatacgc atatattaag taatcctggt 1440
 aatataagag gaggacttcg acaaatagta gctgatcgtc gttcttctgt agtcgtttat 1500
 ggttggacac acaaaagttt gagtcgacgt agtttagttg caccagatca aattactcaa 1560
 gtacctgctg ttaaagcaag tccctcatcc cattgtacca tcattgcagg acctggattt 1620
 acgggcgggg atctcgtaag tctgcaacca aatggacaac tcgttatacc gtttcaggta 1680
 tcggcgccag aaacaaatta tcatattcga atatggtatg tttctacgtc cgactgttcc 1740
 ataaatacaa tatgtaatga tgagacccat ttaagtacgt tgccttccac aacctcatca 1800
 cttgaaaatt tacaatgtaa ccatttgcatt tattttaacg tgggcacttt caaacctacg 1860
 atagatagta aactaacgct tgtaaataca agtccaaatg caaatattat catcgacaaa 1920
 attgaattta ttcccgtaga tacggcccaa caacaaaatg aggatctaga agcagcaaaa 1980
 aaagcgggtg cgagcttggt tacacgcaca agggacggat tacaagtaaa tgtgaaagat 2040
 tatcaagtcg atcaageggt aaatttagtg tcatgcttat cagatgaaca atatgggtat 2100
 gacaaaaaga tgttattgga agcgggtacgt gcggcaaac gacttagccg agaacgcaac 2160
 ttacttcagg acccagattt taatacaate aatagtacag aagaaaatgg atggaaagca 2220
 agtaacggcg ttactattag tgagggcggt ccattctata aaggccgtgc aattcagcta 2280
 gcaagtgcac gagaaaatta cccaacatac atctatcaaa aagtagatgc atcggagtta 2340
 aagccgtata cacgttatag actggatggg ttcgtgaaga gtagtcaaga tttagaaatt 2400
 gatctcattc accatcataa agtccatctt gtgaaaaatg taccagataa tttagtatct 2460
 gatacttacc cagatgattc ttgtagtgga atcaatcgtat gtcaggaaca acagatggta 2520
 aatgcgcaac tggaaacaga gcatcatcat ccgatggatt gctgtgaagc agctcaaaca 2580
 catgagtttt ctctctatat tgatacaggg gatttaaatt cgagtgtaga ccagggaatc 2640
 tgggcgatct ttaaagttcg aacaaccgat ggttatgcga cgttaggaaa tcttgaattg 2700
 gtagaggtcg gaccgttacc gggatgaatc ttagaacgtg acaaaggga taatacaaaa 2760
 tggagtgcag agctaggaag aaagcgtgca gaaacagatc gcgtgtatca agatgccaaa 2820
 caatccatca atcatttatt tgtggattat caagatcaac aattaaatcc agaaataggg 2880
 atggcagata ttatggacgc tcaaaatctt gtcgcatcaa tttcagatgt atatagcgat 2940
 gccgtactgc aatccctgg aattaactat gagatttaca cagagctgtc caatcgctta 3000
 caacaagcat cgtatctgta tacgtctcga aatgcggtgc aaaatgggga ctttaacaac 3060
 gggctagata gctggaatgc aacagcgggt gcatcggtac aacaggatgg caatacgcatt 3120
 ttcttagttc tttctcattg ggatgcacaa gtttctcaac aatttagagt gcagccgaat 3180
 tgtaaatatg tattacgtgt aacagcagag aaagtaggcg gcggagacgg atacgtgact 3240
 atccgggatg atgctcatca tacagaaacg cttacattta atgcatgtga ttatgatata 3300
 aatggcacgt acgtgactga taatacgtat ctaacaaaag aagtggattt ccatccggag 3360
 acacaacaca tgtgggtaga ggtaaataaa acagaaggtg catttcatat agatagtatt 3420
 gaattcgttg aaacagaaaa gtaa 3444

<210> 2

<211> 3474

<212> DNA

<213> 苏云金芽孢杆菌

<400> 2

atgaatcgaa ataatacaaaa tgaatatgaa attattgatg cccccattg tgggtgtcca 60
tcagatgacg atgtgaagta tcctttggca agtgacccaa atgcagcgtt acaaaatag 120
aactataaag attacttaca aatgacagat gaggactaca ctgattctta tataaatcct 180
agtttatcta ttagtggtag agatgcagtt cagactgcgc ttactgttgt tgggagaata 240
ctcggggcct taggtgttcc gttttctgga caaatagta gtttttatca attcctttta 300
aatacactgt ggccagttaa tgatacagct atatgggaag ctttcatgcg acaggtggag 360
gaacttgtca atcaacaaat aacagaattt gcaagaaatc aggcacttgc aagattgcaa 420
ggattaggag attccttttaa tgtatatcaa cgttcccttc aaaattggtt ggctgatcga 480
aatgatacac gaaatttaag tgttgttcgt gctcaattta tagctttaga ctttgatttt 540
gttaatgcta ttccattgtt tgcagtaaat ggacagcagg ttccattact gtcagtatat 600
gcacaagctg tgaatttaca tttgttatta ttaaaagat catctctttt tggagaagga 660
tggggattca cacaggggga aatttccaca tattatgacc gteaattgga actaacgct 720
aggtacacta attactgtga aacttggtat aatacaggtt tagatcgttt aagaggaaca 780
aatactgaaa gttgggttaag atatcatcaa ttccgtagag aatgacttt agtggtatta 840
gatgttgtgg cgctatttcc atattatgat gtacgacttt atccaacggg atcaaacca 900
cagcttacac gtgaggata tacagatccg attgtattta atccaccagc taatgttga 960
ctttgccgac gttgggttac taatccctat aatacttttt ctgagctcga aaatgccttc 1020
attcgccac cacatctttt tgataggctg aatagcttaa caatcagcag taatcgattt 1080
ccagtttcat ctaattttat ggattattgg tcaggacata cgttacgccg tagttatctg 1140
aacgattcag cagtacaaga agatagttat ggccetaata caaccacaag agcaacaatt 1200
aatcccggag ttgatggaac aaaccgcata gagtcaacgg cagtagattt tcgttctgca 1260
ttgataggta tatatggcgt gaatagagct tctttgtcc caggaggctt gtttaatggt 1320
acgacttctc ctgctaattg aggatgtaga gatctctatg atacaaatga tgaattacca 1380
ccagatgaaa gtaccggaag ttcaacccat agactatctc atgttacctt ttttagcttt 1440
caaactaatc aggctggatc tatagctaat gcaggaagtg tacctactta tgtttggacc 1500
cgctcgtgatg tggaccttaa taatacgatt accccaaata gaattacaca attaccattg 1560
gtaaaggcat ctgcacctgt ttcgggtact acggtcttaa aaggtccagg atttacagga 1620
gggggtatac tccgaagaac aactaatggc acatttggaa cgtaagagt aacggttaat 1680
tcaccattaa cacaacaata tcgctaaga gttcgtttt cctcaacagg aaatttcagt 1740
ataaggttac tccgtggagg ggtttctatc ggtgatgta gattaggag cacaatgaac 1800
agagggcagg aactaactta cgaatccttt ttcacaagag agtttactac tactggctccg 1860
ttcaatccgc cttttacatt tacacaaget caagagatte taacagtga tgcagaaggt 1920
gttagcaccg gtggtgaata ttatatagat agaattgaaa ttgtccctgt gaatccggca 1980
cgagaagcgg aagaggattt agaagcggcg aagaaagcgg tggcgagctt gtttacacgt 2040
acaagagatg gattacaggt aaatgtgaca gattaccaag tggatcgagc ggcaaattta 2100
gtgtcatgct tatcagatga acaatattcg catgataaaa agatgttatt ggaagccgta 2160
cgcgcagcaa aacgcctcag ccgcgaacgc aacttacttc aagatccaga ttttaataca 2220
atcaatagta cagaagaaaa tggctggaag gcaagtaac gtgttactat tagcgagggc 2280

ggtccattct ttaaaggtcg tgcacttcag ttagcaagcg caagagaaaa ttatccaaca 2340
 tacatttate aaaaagtaga tgcacgggtg ttaaagcett atacacgcta tagactagat 2400
 ggatttgtga agagtagtca agatttagaa attgatctca tccacatca taaagtccat 2460
 cttgtaaaaa atgtaccaga taatttagta tctgatactt actcagatgg ttcttgacgc 2520
 ggaatcaacc gttgtgatga acagcagcag gtagatatgc agctagatgc ggagcatcat 2580
 ccaatggatt gctgtgaagc ggctcaaca catgagtttt cttcctatat taatacaggg 2640
 gatctaaatg caagtgtaga tcagggcatt tgggttgtat taaaagtctg aacaacagat 2700
 gggatgacga cgtaggaaa tcttgaattg gtagaggttg ggccattatc gggatgaatct 2760
 ctagaacgcg aacaaagaga taatgcgaaa tggaatgcag agctaggaag aaagcgtgca 2820
 gaaacagatc gcgtgtatct agctgcgaaa caagcaatta atcatctatt tgtagactat 2880
 caagatcaac aattaaatcc agaaattggg ctacgggaaa taaatgaagc ttcaaactct 2940
 gtgaagtcaa ttctgggtgt atatagtgat aactattac agattctctg aattaactac 3000
 gaaatttaca cagagttate cgategatta caacaagcat cgtatctgta tacgtctcga 3060
 aatgccgtgc aaaaatggaga ctttaacagt ggtctagata gttggaatgc aacaacagat 3120
 gcacgggttc agcaagatgg cagtacacat ttcttagttc ttctgcattg ggatgcacaa 3180
 gtttcccaac aatgagagt aaatttgaat tgtaagtatg ttttacgtgt aacagcaaaa 3240
 aaagtaggag gcggagatgg atacgtcaca atccgagatg gcgctcatca ccaagaaact 3300
 cttacattta atgcatgta ctacgatgta aatggtacgt atgtcaatga caattcgtac 3360
 ataacaaaag aagtggatt ctaccagag acaaaacata tgtgggtaga ggtgagtga 3420
 tccgaagggt cattctatat agacagtatt gaggttattg aaacacaaga gtag 3474

<210> 3

<211> 3522

<212> DNA

<213> 苏云金芽孢杆菌

<400> 3

atgaatcgaa ataatcaagg tgaatatgaa attattgacg cttccacttg tggttgttcg 60
 tcagatgatg ttgttcaata tcctttggca agagatccga atgctgcatt ccaaaatag 120
 aattataaag attatttgaa aatgtctgac ggagactacg tcgattctta tataaaccca 180
 ggcttatcta ttggtcgtag agatgtgacc ctaactggag ttggtattgt tgcgctaata 240
 gtagggactt taggtggtec agttgggggt atagtaactg gcttgatttc ctctctttta 300
 ggattattgt ggccaagtaa tgataatgat gtatgggaag cttttatggc acaaatagaa 360
 gagctaattg aacaaaggat agcagatcaa gtagtaagga atgcactcga taacttaact 420
 ggattgcgcg attattataa tcaataccta ttagcattgg aggagtggca ggaaaggccg 480
 aacgctgtaa gatctacctt agtttttaat agatttgaag cctgcattc tcaactttgta 540
 actagtatgc caagctttgg tagtggccct ggaagtgaag ggtatgcggt acaattgctg 600
 acagtttatg cacaagcggc aaatctgeat ttgttattat taagagatgc tgacatttat 660
 ggggcaagggt ggggacttcg tgaatctcag attgatttat attttaatga gctacaaaat 720
 cgtactcgag attataccaa tcattgtgta actgcgtaca ataatgggtt agaggagata 780
 cgaggaacaa gccctgcaag ttggttgagg taccatcaat tccgtagaga gacaacacta 840

atagcattgg atttagtggc gatattccca tattacaacg tacgagaata tccaattggg 900
 gtaaatcctc agcttacacg tgatgtatat acagatccaa taggggttac tttcagaaga 960
 gaagattggg aacaggagt agaatgcaga ccatgggtaa atactcctta catgagcttt 1020
 tcggatcttg aaaatgcaat aattcgtcca ccacatctat ttgaaacatt acgtaattta 1080
 acaattcata caggtcgata taacctagta ggaggggcca gatttattga aggatgggtc 1140
 ggacattctg taacaaatac tcgcttgggt aattcaacag tatttacaag taattatggt 1200
 tctttgccac ctcgttttca agtttttaat tttactaatt ttgatgttta ccaattaat 1260
 acgagagcag attctacagg tacctttaga atccctggat ttgcagttac aagggcccaa 1320
 ttcattccgg gtgggactta ttcagtagct caccgagatc caggggcatg tcaacaagat 1380
 tatgattcaa ttgaagagtt accaagtcta gaccggatg aacctattaa tagaagtat 1440
 agtcatagat tatcgcatgt taccctttat aaatatactc tctcagatac agattatgga 1500
 gttatcaatt atacagatta tggaagtatg cctgcatatg tctggacaca tcgcatgtg 1560
 gacctacta acacgattac tgcagataga attacacaac tcccattagt aaaggcatct 1620
 aactacctg cgggtactac tgtggtaaaa ggcccaggat ttacaggagg agatatactc 1680
 cgaagaacaa ctaatggaac atttgggaca ttacatgtaa gggttaattc accattaaca 1740
 caacaatata gcctaagagt tcgttttgcc tcaacaggaa atttcagtat aagggtactc 1800
 cgtggaggga cttctatcgg tgatgctaga tttgggagca caatgaacag aggacaggaa 1860
 ctaacttacg aatcctttgt cacaagagag tttactacta ctggtccgtt caatccgct 1920
 tttacattta cacaaactca agaaattcta acagtgaatg cagaaggtgt tagcaccggt 1980
 ggtgaatatt atatagatag tattgagatt gttcctgtaa atccgacgcg agaggcggaa 2040
 gaggatctag aagcagcgaa gaaagcggtg gcgagcttgt ttacacgtac aaggacggga 2100
 ttacaagtaa atgtgacaga ttaccaagtg gatcgagcgg caaatttagt gttatgctta 2160
 tcagatgaac aatatgcgca tgataaaaag atgttattgg aagccgtacg cgcagccaaa 2220
 cgactcagcc gcgagcgtaa cttgcttcaa gatccagatt tcaatgaaat aaatagtacg 2280
 gaagatagtg gttggaagac aagtaacggc attatcatta gtgagggtgg tccattcttt 2340
 aaaggtcgtg cccttcagct agcaagcgca cgtgaaaatt acccaacata catctatcaa 2400
 aaggtagact catcaatgtt aaaaccttat acacgatata aactagatgg atttgtgcaa 2460
 agtagtcaag atttagaaat tgaactcatt caccatcata aagtccacct cgtgaaaaat 2520
 gtaccagata atttagtact tgatacttac ccagatggtt cttgcaacgg aattaaccga 2580
 tgtgaggaac aacagatggt gaattcgcaa ctagaaacag aacatcatcc aatggattgc 2640
 tgtgaagcat cccaaacaca tgagttttct tectatattc atacaggtga cctaaatgca 2700
 agttagatc aaggcatttg ggttgatttg aagattcgga caacagatgg ttctgcgacg 2760
 ttaggaaatc ttgaattggt agaggttggt ccattatcgg gtgaatctct agaacgtgaa 2820
 caaagagata atgcgaaatg gaatgcagag ttaggaagga agcgtgcaga agcagatcgc 2880
 gtgtatcaag gtgcgaaaca agcaattaac catctatttg tagactatca agatcaacaa 2940
 ttaaatccag aagtgggct agcagaaatt agtgaagctc gaaatcttat cgaatcaatt 3000
 tcagatgtat attgcatgc agtactgcga attcctggaa ttaactacga gatgtataca 3060
 gagttatcta atcgtctaca acaagcagcg tatctgtata cgtctcgaaa tgccgtgcaa 3120
 aacggggact ttaacagcgg tttagatagt tggaatgcaa caactgatgc gacggttcag 3180

caggatggca atatgtatth cttagttctt tcccattggg atgcacaagt ttctcaacaa 3240
 ttttagagtac agccgaattg taaatatgtg ttacgtgtga cagcgaagaa agtagggaac 3300
 ggagatggat atgttacgat ccaagatggc gctcatcacc gagaaacact tacattcaat 3360
 gcatgtgact acgatgtaaa tggtagcat gtaaatagaca attcgtatat tacaaaagaa 3420
 ttggagttct atccaaagac agaacatatg tgggtagagg taagtgaac agaaggtacc 3480
 ttctatatag acagcattga gctaattgaa acacaagagt ag 3522

<210> 4

<211> 3540

<212> DNA

<213> 苏云金芽孢杆菌

<400> 4

atgggaggaa aaagtatgaa tcgaaataat caaggtgaat atgaaattat tgacgcttcc 60
 acttgtgggt gtctgcaga tgatgttgtt caatatectt tggcaagaga tccgaatgct 120
 gcattccaaa atatgaatta taaagattat ttgaaaatgt ctgacggaga ctacgtcgat 180
 tcttatataa acccaggctt atctatttgt cgtagagatg tgaccctaac tggagttggg 240
 attgttgcgc taatagtagg gactttagggt ggtccagttg ggggtatagt aactggcttg 300
 atttcctctc ttttaggatt attgtggcca agtaatgata atgatgtatg ggaagctttt 360
 atggcacaaa tagaagagct aattgaacaa aggatagcag atcaagtagt aaggaatgca 420
 ctcgataact taactggatt gcgcgattat tataatcaat acctattage attggaggag 480
 tggcaggaaa ggccgaacgc tgtaagatct acctagttt ttaatagatt tgaaccctg 540
 cattctcact ttgtaactag tatgccaagc tttggtagtg gccctggaag tgaaggtat 600
 gcggtacaat tgctgacagt ttatgcacaa gcggcaaatc tgcatttgtt attattaaga 660
 gatgctgaca tttatggggc aaggtgggga cttcgtgaat ctgagattga tttatatttt 720
 aatgagctac aaaatcgtac tcgagattat accaatcatt gtgtaactgc gtacaataat 780
 gggtagagg agatacgagg aacaagcctt gcaagttggg tgaggtacca tcaattccgt 840
 agagagacaa cactaatagc attggattta gtggcgatat tccatatta caacgtacga 900
 gaatatccaa ttggggtaaa tcctcagctt acacgtgatg tatatacaga tccaataggg 960
 gttactttca gaagagaaga ttgggaaca ggagtagaat gcagaccatg ggtaaatact 1020
 ccttacatga gcttttcgga tcttgaaaat gcaataatc gtccaccaca tctatttgaa 1080
 acattacgta atttaacaat tcatacaggt cgatataacc tagtaggagg ggcgagattt 1140
 attgaaggat gggtcggaca ttctgtaaca aatactcgtt tgggtaattc aacagtattt 1200
 acaagtaatt atggttcttt gccacctcgt tttcaagttt ttaattttac taattttgat 1260
 gtttaccaaaa ttaatacagag agcagattct acaggtacct ttagaatccc tggatttgea 1320
 gttacaaggg cccaattcat tccgggtggg acttattcag tagctcaccg agatccaggg 1380
 gcatgtcaac aagattatga ttcaattgaa gagttacca gtctagaccg ggatgaacct 1440
 attaatagaa gttatagtea tagattatcg catgttacc tttataaata tactctctca 1500
 gatacagatt atggagttat caattataca gattatggaa gtatgcctgc atatgtctgg 1560
 acacatcgcg atgtggacct tactaacacg attactgcag atagaattac acaactccca 1620
 ttagtaaagg catctaacct acctgcgggt actactgtgg taaaaggccc aggatttaca 1680

ggaggagata tactccgaag aacaactaat ggaacatttg ggacattaca tgtaagggtt 1740
 aattcaccat taacacaaca atatgccta agagttegtt ttgcctcaac aggaaatttc 1800
 agtataaggg tactccgtgg agggacttct atcggtgatg ctagatttgg gagcacaatg 1860
 aacagaggac aggaactaac ttacgaatcc tttgtcacia gagagtttac tactactggt 1920
 ccgttcaatc cgccttttac atttacacia actcaagaaa ttctaacagt gaatgcagaa 1980
 ggtgttagca ccggtggtga atattatata gatagtattg agattgttcc tgtaaaccg 2040
 acgcgagagg cggaagagga tctagaagca gcgaagaaag cgggtggcgag cttgtttaca 2100
 cgtacaaggg acggattaca agtaaatgtg acagattatc aagtcgatca agcggcaaat 2160
 ttagtgtcat gcttatcaga tgaacaatat gggtatgaca aaaagatgtt attggaagcg 2220
 gtacgcgcgg caaaacgcct cagccgagaa cgtaacttac ttcaagatcc agattttaat 2280
 acaatcaata gtacagaaga aaatggatgg aaagcaagta acggcgttac tattagttag 2340
 ggcggtccat tctataaagg ccgtgcactt cagctagcaa gtgcacgaga aaattatcca 2400
 acatacattt atcaaaaagt agatgcatcg gagttaaac cttatacacg atatagacta 2460
 gatgggttcg tgaagagtag tcaagattta gaaattgatc tcattacca tcataaagtc 2520
 catcttgtga aaaaatgtacc agataattta gtatctgata cttaccaga tgattcttgt 2580
 agtggaatca atcgatgtca ggaacaacag atggtaaatg cgcaactgga aacagagcat 2640
 catcatccga tggattgctg tgaagcagct caaacacatg agttttcttc ctatattgat 2700
 acaggggatt taaattcgag tntagaccag ggaatctggg cgatctttaa agttcgaaca 2760
 accgatggtt atgcgacgtt aggaaatctt gaattggtag aggtcggacc gttatcgggt 2820
 gaatctttag aacgtgaaca aagggataat acaaaatgga gtgcagagct aggaagaaag 2880
 cgtgcagaaa cagatcgcgt gtatcaagat gccaaacaat ccatcaatca tttatttgtg 2940
 gattatcaag atcaacaatt aaatccagaa atagggatgg cagatattat ggacgctcaa 3000
 aatcttgtcg catcaatttc agatgtatat agcgatgccg tactgcaaat ccctggaatt 3060
 aactatgaga ttacacaga gctgtccaat cgettacaac aagcatcgta tctgtatacg 3120
 tctcgaaatg ccggtgcaaaa tggggacttt aacaacgggc tagatagctg gaatgcaaca 3180
 gcgggtgcat ccgtacaaca ggatggcaat acgcatttct tagttcttcc tcattgggat 3240
 gcacaagttt ctcaacaatt tagagtgcag ccgaattgta aatatgtatt acgtgtaaca 3300
 gcagagaaaag taggcggcgg agacggatac gtgactatcc gggatggtgc tcatcataca 3360
 gaaacgctta catttaatgc atgtgattat gatataaatg gcacgtacgt gactgataat 3420
 acgtatctaa caaaagaagt gatattctat tcacatacag aacacatgtg ggtagaggta 3480
 aatgaaacag aaggtgcatt tcatatagat agtattgaat tcgttgaaac agaaaagtaa 3540
 <210> 5
 <211> 3444
 <212> DNA
 <213> 人工序列
 <220>
 <223> 合成Cry基因
 <400> 5
 atggacctgg atgggaataa gacagagaca gagaccgaga ttgtgaatgg gagcgagagc 60

agcattgacc cgagcagcgt ttcgtacgct gggaacaata gctactccag cgccctgaac 120
 ctcaattcgt gccagaatag gggcatcgct cagtgggtta acacgctggg cggggctatt 180
 gggcaggccg tgagcatcgg cacatctatc atttactcc tggccgcgcc gacactcact 240
 gggctctatt cactggcctt caatctcatc aggaggatgg ggaccggctc caacggctcg 300
 tctatttccg acctgagcat ctgcatctc ctgagcatca ttaacctgcg ggtttcgcag 360
 gctgtgctca acgacgggat cgctgatttc aatggctccg ttgctgtgta cgacctgtac 420
 ctccacgccc tgcgcagctg gaacaataac cctaacgctg ctactgctga ggagctgagg 480
 acccgcttca ggatcgccga ttcggagttc gagaggattc tgacgagggg ctcgctcaca 540
 catggcggct ccctcgcccg ccaggacgct caggtcctcc tgcctccgctc ctctgttaac 600
 gcggcttacc tgcacctgct cactctccgc gatgcttccg gctacggggc ctcttggggc 660
 ctcttcaaca ccacgccga tatcaattac cccgtgagge tgcagcagct cattggcagc 720
 tacacgcaact actgcacaca ttggtacaac caggggctga atgagatccg gcagcgcggc 780
 aacactgccg tgaattgget cgagtccac cgtaccgcc gcgacatgac gctgatggctc 840
 ctcgatgtgg tctcgtgtt ctctgccctc gacacgatcc gctaccgaa cgctacagtt 900
 gtgcagctca gccgcaactg ctacaccgat ccgattgget tcgttaaccg cgggtcagge 960
 aataggctgt cctggttcca ctggaggaac caggcgaatt tctctactct cgagtcagag 1020
 atgccgacc cctcatcccc actgagctc aaccacatgt cgatcttcc tgggctctctg 1080
 accctcccag tgtcccctaa cacgcatagg gcccggtct ggtacggcaa ccagaatag 1140
 ttcaaacctg ggtcacagaa ctccggccag accacgaact ctattcagaa tatctcagge 1200
 ctggagattt tccgcatcga ctctcaggcg tgcaatctca ataacaattc atacggcgtg 1260
 aacagggcgg agttcttcca cggggctagc cagggctcgc agcggctctgt ctaccagga 1320
 tacatccgcc agagcggcct ggacaaccct gtcgttatga atctgcagtc tttctccca 1380
 ggcgagaact cagccacccc tacggcgcag gattacacce acattctgtc caaccgggtt 1440
 aatatcaggg gcgggctcag gcagattgtg gccgacagge gctcctccgt ggtcgtttac 1500
 ggctggacgc acaagtccct gagcaggagg tcaactcgtg ctccagacca gatcaccag 1560
 gtcccagccg ttaaggcgtc ccttcttca cattgcaact tcattgccgg cccaggcttc 1620
 accggcgggg acctggtgtc gctccagccc aacggccagc tcgtcatccc gttccaggtt 1680
 tctgcgcccg agacgaacta ccacattcgc atctgctacg tctcgacgtc tgattgcagc 1740
 attaacacaa tctgcaatga cgagacgeat ctgtccacac tcccagcac aacttccagc 1800
 ctggagaacc tccagtcaa tcacctgeat tacttcaacg tgggeacttt caagccaacc 1860
 atcgactcga agctgacgct cgtaacaca tctctaacg ctaacatcat tatcgacaag 1920
 atcgagttca tcccgttga taccgccag cagcagaacg aggacctga ggccgcgaag 1980
 aaggctgtcg cctccctgtt cacacgcaact agggacggcc tccaggtaa tgttaaggac 2040
 taccagtggt atcaggtgc caactggtc tcatgctct cgcagagca gtacggctac 2100
 gataagaaga tgctgctcga ggccgtgagg gctgctaaga ggctgagcag ggagaggaa 2160
 ctgctccagg accccgattt caacacaatc aactcgaccg aggagaacgg gtggaaggcg 2220
 tcaaatggcg tcaccatctc cgagggcggg ccattctaca agggcagggc tattcagctc 2280
 gcgtctgctc gggagaacta ccccacatac atctaccaga aggtgatgc ctccgagctg 2340
 aagccataca cccgctaccg cctcgacggc ttcgtcaagt cgtctcagga cctggagatt 2400

gatctcatcc accatcaciaa ggtgcacctg gtcaagaacg ttccggaciaa tctcgtgagc 2460
gatacgtacc ccgacgattc atgctccgga atcaacaggt gccaggagca gcagatggtc 2520
aacgcgcagc tggagaccga gcacacacat ccgatggact gctgcgaggc tgctcagacg 2580
cacgagttct catcctacat cgacacaggg gatctgaaca gctcggtcga tcagggcatt 2640
tgggccatct tcaaggtag gaccacggac gggtagccta ccctcggcaa cctggagctg 2700
gtggaggctc ggccactgag cggcgagtcg ctcgagaggg agcagaggga caacactaag 2760
tggtcgctg agctgggccg caagagggt gagaccgacc gcgtctacca ggatgccaag 2820
cagagcatca atcacctggt cgttgactac caggatcagc agctcaacc cgagattggc 2880
atggcggaca tcatggatgc tcagaacctg gtggccagca tctcggacgt gtacagcgat 2940
gcggctctcc agattccagg aatcaactac gagatctaca cggagctgtc gaacaggctc 3000
cagcaggcct cctacctgta cacaagccgg aacgcggtcc agaatgggga cttcaacaat 3060
ggcctcgatt catggaatgc tacggctggg gcttccgtgc agcaggatgg caacacacac 3120
ttcctggtcc tctcccattg ggacgcgcag gttagccagc agttccgctg gcagccgaac 3180
tgcaagtatg tgctgagggt cactgctgag aaggttggcg ggggcgacgg ctacgtgacc 3240
atcaggagc atgcgcacca taccgagacg ctgacattca acgcttgcga ctacgacatc 3300
aacggcacct acgtgacaga caacaactac ctaaccaagg aggtggtctt ccaccggag 3360
actcagcata tgtgggttga ggtgaacgag accgagggcg cttccacat agactccatc 3420
gagttcgtcg agaccgagaa gtga 3444

<210> 6

<211> 3474

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 6

atgaacagga acaaccagaa cgagtacgag attattgacg cccccattg cggctgcccc 60
tccgacgacg atgtgaagta cccactggct agcgacceca acgctgctct gcagaacatg 120
aattacaagg attacctcca gatgaccgac gaggattaca cggactcgta catcaacceca 180
tccctcagca tttcgggcag ggacgctgtc cagacagccc tgactgtggt cggccgcatac 240
ctcggggcgc tgggcgttcc cttctcagc cagattgtgt cttctacca gttctctctg 300
aataccctct ggccagtga aacacagggc atctgggagg ctttcatgcg ccagggtggag 360
gagctggtea atcagcagat tacggagttc gccaggaacc aggtctctgc gcggctgcag 420
ggcctcgggg actccttcaa tgtctaccag aggagcctgc agaactggct cgcgaccgc 480
aacgatacca ggaatctctc cgttgtgcgc gccagttca tcgcgctcga cctggatttc 540
gtgaatgcca ttctctgtt cgtgtgaac ggccagcagg tcccgtctct gtccgtttac 600
gctcaggccg tgaacctgca tctctctctc ctgaaggatg cttegtctct cggcgagggg 660
tggggcttca cacagggcga gatctctact tactacgacc gccagctcga gctgacagcg 720
aggtacacta attactgcga gacctggtac aacacggggc tggacaggct caggggaacc 780
aacacggagt cctggctccg ctaccaccag ttccgcaggg agatgactct ggtcgttctc 840

gatgtggtcg cctgttccc atactacgac gtccgctct acccaaccgg ctccaaccct 900
cagctgacaa gggagggtga cactgaccct atcgtcttca acccaccagc taatgtgggg 960
ctctgcaggc gctggggaac caaccctac aatacgttca gcgagctgga gaacgcgttc 1020
atccggccac ctcatctgtt cगतcgcctc aactctctca ccatttccag caataggttc 1080
cctgtctcgt ctaacttcat ggactactgg tctggccaca cgctgaggcg gagctacctc 1140
aacgattcgg ctgtgcagga ggactcctac ggctcatca ccacgacacg ggccaccatt 1200
aaccgggggg tcgatggcac caaccggatc gagtcgacgg cgggtggactt ccgctctgct 1260
ctcatcggga ttacggcgt taacagggtt tccttcgtgc caggcgggct gttcaatggc 1320
actaccagcc cagctaaccg cggtgcagg gacctgtacg ataccaacga cgagctgcca 1380
ccagacgagt ccacaggctc atccactcat cgctctctgc acgteacatt cttctctttc 1440
cagactaate aggccgggtc aatcgcgaac gctggctcgg tteccacctc cgtgtggacg 1500
cgcagggacg tcgatctgaa caacacgac actccgaacc gcattacgca gctcccactg 1560
gtgaaggctt ctgtctcagt ctcaggeacg acagttctga aggggcccgg cttaccgggc 1620
gggggcatcc tccggcgcac taccaatggg accttcggca cgctgagggt gaccgtcaac 1680
agcccactga cgcagcagta caggtccgc gtgaggttcg cttctacggg caatttctca 1740
atcaggctcc tgaggggggg cgtgagcatt ggggacgtca ggctgggctc gacaatgaac 1800
cggggccagg agctgacata cgagagcttc ttcactcgcg agttcacgac aactggccca 1860
ttcaatccac ctttcaactt cacgcaggcc caggagatcc tcacagttaa cgctgagggc 1920
gtgtcgactg ggggcgagta ctacattgat aggatcgaga ttgttccagt gaaccagct 1980
agggaggctg aggaggacct ggaggctgcc aagaaggctg tggccagcct gttcacacgc 2040
actagggacg gcctccaggt caatgttacc gattaccagg tcgacagggc ggctaacctg 2100
gtttcatgcc tctccgatga gcagtactcc cacgacaaga agatgctctt ggaggccgtc 2160
cgggctgcta agcgcctgtc acgggagcgc aacctcctgc aggacctga tttcaacacg 2220
atcaactcca ctgaggagaa tgggtggaag gccagcaacg gcgtgacat ttcggagggg 2280
ggcccgttct tcaagggccg cgcgctccag ctggctagcg ctagggagaa ctaccctacg 2340
tacatctacc agaaggtcga tgcgtcgggt ctgaagccgt acacacgcta ccgctcgcac 2400
ggcttcgtga agtcctccca ggatctggag atcgacctca ttcacatca caaggtccat 2460
ctggttaaga acgtgcccga caatctctgc tccgatacct acagcgacgg gtctctgcagc 2520
ggaatcaacc gctgcgatga gcagcagcag gtggatatgc agctcgacgc cgagcatcac 2580
ccaatggact gctgcgaggc tgcccagacc cagagttct cttctacat caatacgggg 2640
gatctgaacg cctccgttga ccagggcatt tgggttgtgc tcaaagttag gaccacggac 2700
gggtacgcta ccctgggcaa cctcgagctg gtggaggctc ggccgctgag cggcgagtcg 2760
ctcgagaggg agcagagggga taacgetaag tggaatgctg agctgggcag gaagagggt 2820
gagaccgaca ggtctacct ggetgetaag caggcatca atcactctt cgtggattac 2880
caggaccagc agctgaacct tgagatcggc ctcgctgaga ttaacgagc ctctaatctg 2940
gtcaagtcca tctctggggt ttaactcagat actctctgc agateccggg aattaactac 3000
gagatttaca ccgagctgtc cgaccggtc cagcaggctt cctacctca cacgagccgc 3060
aacgccgtgc agaattggga tttcaactcg ggctggact cttggaacgc gacaactgat 3120
gcttctgtcc agcaggacgg ctcaacctat ttctctgtgc tgtcacactg ggacgctcag 3180

gtgtcccagc agatgagggt caacctgaat tgcaagtacg tctcagggt tacggcgaag 3240
 aaggctcgggg gcggggatgg ctacgtcaca atcagggacg gcgcgcatca ccaggagacc 3300
 ctcacgttca atgcttgcca ctacgatgtc aacggcaca acgttaacga caattcctac 3360
 atcactaagg aggtcgtttt ctaccccgag accaagcaca tgtgggttga ggtgtctgag 3420
 tcggagggct cgttctacat tgatagcatt gagttcattg agacgcagga gtga 3474

<210> 7

<211> 3522

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 7

atgaaccgga acaaccaggg cgagtacgag attattgatg cctccacttg eggctgctcc 60
 tcagatgatg tcgtccagta cccaactcget cgcgaccga acgctgcttt ccagaacatg 120
 aattacaagg actacctgaa gatgtctgac ggcgattacg tcgattcata cattaacca 180
 ggctgtcca tcgggaggag ggacgtcacg ctcacaggcg tcgggatcgt tgetctcatt 240
 gtgggcaccc tgggcggccc agttggcggg attgtgacgg gcctgatctc cagcctctctg 300
 gggctcctgt ggccaagcaa cgacaatgat gtgtgggagg cttcatggc gcagatcgag 360
 gagctgattg agcagaggat cgtgaccag gtggtccgga acgccctgga caatctcacc 420
 ggctgaggg attactacaa ccagtacctc ctggctctcg aggagtggca ggagaggccc 480
 aatgccgtga ggtctacgt ggtcttcaac cggttcgaga cgtccatte aacttctgtg 540
 acatcaatgc catccttcgg cagcgggcct ggcagcgagc gctacgcggt tcagctcctg 600
 accgtgtacg ctcaggctgc caacctgcac ctctgctcc tgagggacgc tgatatctac 660
 ggcgctcggg gggggctcag ggagtcccag atcgacctct acttcaacga gctgcagaat 720
 cggacgcgcg attacacaaa ccattgcgtc acagcctaca acaatggcct ggaggagatc 780
 agggggactt cgccagcttc ttggctgcgc taccaccagt tccggcgcga gaccacgctc 840
 attgccctcg acctggtggc gatcttcca tactacaatg tcagggagta cccaattggc 900
 gttaaccctc agctcacgcg ggacgtgtac acagatccga tcggcgtcac gttcaggcgg 960
 gaggactggg agacaggcgt cgagtgcagg ccgtgggtta ataccata catgtctttc 1020
 tcagatctgg agaacgcat cattagccg cccatctct tcgagacgct ccggaatctg 1080
 acgattcaca caggcaggta caacctggtc ggcggggcga ggttcatcga gggctgggtc 1140
 gggcattccg ttactaatac caggtgggc aacagcactg tgttaccag caattacggg 1200
 tcgctccac ctcggttcca ggtgttcaac ttcacgaatt tcgacgtcta ccagatcaac 1260
 acacgggccg attcgacggg cacattccgc attccggggt tcgcggtcac tagggctcag 1320
 ttcatccccg gcgggacctc ctccgtgget caccgcgacc caggccttg ccagcaggac 1380
 tacgattcaa ttgaggaget gccctccctg gaccagatg agcctataca ccggtctac 1440
 agccatcgc tctcacacgt caccctgtac aagtacactc tctccgacac cgattacggc 1500
 gtgatcaatt acaccgacta cgggagcatg ccagcttacg tgtggacgca tcgacgctc 1560
 gatctgacta acaccattac ggcggatagg atcacgcagc tcccgtggt gaaggcttcg 1620

aactccccg ccggcacaac tgttgaag gggccccgct tcaccggcgg ggacatcctg 1680
 aggaggacca cgaatggcac gttcgggaca ctccacgtga gggtaacag cccactgacc 1740
 cagcagtaca ggctccgggt ccgcttcgct tcgacgggca acttctctat tagggtgctg 1800
 aggggccccga catctatcgg cgacgctcgc ttccgggtcaa ctatgaacag gggccaggag 1860
 ctgacttacg agtccttcgt gacccgcgag ttcaacaact ccggccccgtt caatcccc 1920
 ttcacattca ctcagacca ggagatcctg actgtcaacg ctgaggcgtt ttcgaccggc 1980
 ggggagtact acatcgactc tattgagatc gttccagtga acccaaccag ggaggctgag 2040
 gaggatctcg aggctgctaa gaaggccgtc gcgagcctgt tcacaggagc acgggacggc 2100
 ctccaggctca atgttacgga ctaccagggt gataggcgtg ctaacctcgt gctgtgcctc 2160
 tccgacgagc agtacgcccga cgataagaag atgctcctgg aggcgggtgag ggctgctaag 2220
 aggctgagca gggagaggaa cctcctgcag gacctgatt tcaacgagat caattctact 2280
 gaggactcag gctggaagac cagcaacggg atcattatct cggagggcgg gccgttcttc 2340
 aagggccggg ccctgcagct cgttccgctc cgcgagaact acctaccta catctaccag 2400
 aaggtggact cgtctatgct gaagccgtac acgaggtaca agctcgacgg cttcgtgcag 2460
 tcattcccagg atctcgagat tgagctgate caccatcaca aggtgcacct cgtcaagaac 2520
 gttccagaca atctggctct cgacacctac cctgatggct cgtgcaacgg aatcaaccgc 2580
 tgcgaggagc agcagatggt gaactctcag ctggagacgg agcatcacc tatggactgc 2640
 tgcgaggcct cacagactca tgagttcagc tcgtacatcc acaccggcga cctcaacgcg 2700
 tctgtcgate aggggatttg ggtcgttctg aagatcagga cgacagacgg ctcggctacc 2760
 ctccgggaacc tggagctggt ggaggtcggc ccctgtcag gggagtcct cagaggggag 2820
 cagagggaca acgccaagtg gaatgctgag ctgggccgga agcgcgctga ggctgatcgc 2880
 gtgtaccagg gcgctaagca ggccatcaat cactctctcg tcactacca ggatcagcag 2940
 ctgaaccctg aggttggcct cgcggagatc agcagagctc ggaacctgat tgagctgatc 3000
 tctgacgtgt actgcgatgc cgtcctccgc attccgggaa tcaactacga gatgtacacg 3060
 gagctgtcca acaggctgca gcaggctgct tacctgtaca caagccgcaa cgcggtgcag 3120
 aatggcgact tcaactccgg gctcगतagc tggaatgcta ctaccgacgc caccgttcag 3180
 caggatggca acatgtactt cctggtgctc agccactggg acgcccagggt ttcgagcag 3240
 ttccgctgac agccaaattg caagtatgtg ctgagggtca cagcgaagaa ggtcgggaac 3300
 ggcgacggct acgtgactat ccaggatggc gcgcatcacc gcgagactct gaccttcaat 3360
 gcttgcgact acgatgttaa cggcacgcat gtgaacgaca attcctacat taaaaggag 3420
 ctggagttct acccgaagac tgagcacatg tgggttgagg tgagcgagac tgagggcacc 3480
 ttctacatag attcगतcga gctgattgag acccaggagt ga 3522

<210> 8

<211> 3540

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 8

atggggggga agtctatgaa caggaacaac cagggcgagt acgagattat tgatgcctcc 60
acatgcgggt gctccagcga cgacgtggtc cagtaccac tcgctcgcga ccctaacgct 120
gctttccaga acatgaatta caaggactac ctgaagatgt ccgacggcga ttacgtggat 180
agctacatta acccaggcct ctcgatcggg aggagggacg tcaactctgac cggggttggc 240
atcgtggcgc tgattgttgg cacactcggc gggcctgtgg gcgggattgt cactggcctc 300
atctccagcc tcctggggct cctgtggcca tccaacgaca atgatgtctg ggaggcgttc 360
atggctcaga tcgaggagct gattgagcag cgcacgcgg accaggtggt caggaacgct 420
ctcgacaatc tgaccggcct cagggattac tacaaccagt acctcctggc tctcgaggag 480
tggcaggaga ggccaaatgc cgtgcgctcc acgctcgttt tcaaccgctt cgagaccctg 540
cacagccatt tcgtgacgag catgccgtcg ttcgggtctg gccccgggtc ggagcgcctac 600
gctgtgcagc tcctgaccgt ctacgccag gctgccaacc tccacctct gctcctgcgc 660
gacgctgata tctacggcgc caggtggggg ctcagggaga gccagatcga cctgtacttc 720
aacgagctgc agaatcggac acgcgattac actaaccact gcgtaccgc ctacaacaat 780
ggcctcgagg agatcagggg gacgtacca gettctggc tccgctacca ccagttccgg 840
agggagacca cgctgattgc gctcgacctg gtggetatct tcccctacta caatgtgcgc 900
gagtaccgga ttggcgtcaa cccccagctg accagggacg ttacaccga cccgatcggc 960
gtgacattca ggccgggagga ctgggagact ggcgtggagt gcaggccgtg ggtcaatacc 1020
ccatacatgt ctttctcaga cctcgagaac gccatcatta ggccgcccc cctgttcgag 1080
acgctgagga atctcaccat tcatacgggc aggtacaacc tggtcggcgg ggcgcgcttc 1140
atcgagggct gggttgggca ctcagtacg aacacaaggc tcggcaatc cacagtgttc 1200
acttccaact acggcagcct gccacctcg ttccaggttt tcaacttac aaatttcgac 1260
gtgtaccaga tcaacactag ggccgattcg actggcacct tccgattcc agggttcgcc 1320
gttaccgcg cgcagttcat ccctggcggg acgtactccg tggtcaccg cgaccgggc 1380
gcttgccagc aggactacga tagcattgag gagctgcct cgctcgacc agatgagcct 1440
atcaacaggt cctacagcca ccggctgtct catgtcacc tctacaagta caccctgtca 1500
gacacggatt acggcgtgat caattacac gactacgggt ccatgccagc ttacgtttgg 1560
acgcaccggg acgtggatct cacgaacaca attactgcc accgcatcac acagctccca 1620
ctggtgaagg ccagcactct gcctgcgggc acaactgtt tgaaggccc tgggttcacc 1680
ggcggggaca tcctcaggag gaccacgaat ggcacctcg ggacgctgca tgtccgcgtt 1740
aactccccgc tcacacagca gtacagctg cgggtgcgct tcgcttcgac tggcaacttc 1800
tctattcgcg tcctcagggg cgggacctc atcggcgacg ctaggttcgg gagcacgatg 1860
aacaggggcc aggagctgac atacagctc ttcgtcacta gggagtac aactaccggc 1920
ccgttcaatc cgcccttac cttcacgcag acacaggaga ttctaccgt taacgctgag 1980
ggcgtgagca cgggcgggga gtactacatc gactcgatcg agattgtgcc agtcaacca 2040
accagggagg ctgaggagga tctggagget gctaagaagg ccgtggcgag cctcttact 2100
aggaccggg acggcctgca ggttaatgtg acggactacc aggtcgatca ggccgcgaac 2160
ctggttagct gcctctcgga cgagcagtac ggctacgata agaagatgct cctggaggcc 2220
gtccgcgctg ctaagagget ctcgagggag aggaacctc tgcaggacc cgatttcaac 2280
acaattaatt ctactgagga gaacggctgg aaggccteta atggggtgac catctcagag 2340

ggccggccat tctacaaggg cagggcgctc cagctggctt cagctcggga gaactacccc 2400
 acctacatct accagaaggt cgacgcctcc gagctgaage catacacgcg ctaccgcctg 2460
 gatggcttcg tgaagtcgtc tcaggacctg gagatcgatc tcattcacca tcacaaggtc 2520
 cacctcgta agaacgtgcc ggacaatctg gtctccgata cctaccccga cgattcgtgc 2580
 tctggaatca acaggtgcc ggagcagcag atgggtgaacg cccagctcga gacggagcat 2640
 caccatccta tggactgctg cgaggcggct cagacctatg agttctcatc ctacatcgac 2700
 acgggcgac tcaacagctc ggtcgaccag gggatctggg cgattttcaa ggttaggacg 2760
 acagatggct acgctaccct ggggaatctc gagctggctg aggttggccc cctctctggg 2820
 gagtcaactg agagggagca gagggacaac acaaagtggc ctgctgagct gggcaggaag 2880
 cgggctgaga ctgaccgcgt ctaccaggat gccaaagcagt ccatcaatca cctcttcgtg 2940
 gactaccagg atcagcagct gaacctgag attggcatgg ctgacatcat ggatgcccg 3000
 aacctcgtcg cgtcaatctc cgacgtctac agcgatgcgg ttctgcagat cccgggcatt 3060
 aattacgaga tctacacaga gctgtcgaa aggetccage aggcgtcata cctctacag 3120
 tcccggaacg ctgtgcagaa tggcgacttc aacaatgggc tggattcgtg gaatgcgaca 3180
 gctggcgct ctgtgcagca ggacgggaa actcaactec tegtctgtc tcattgggat 3240
 gcccaggctt cacagcagtt ccgggttcag ccgaactgca agtatgtgct gcgcgttacc 3300
 gctgagaaag tgggcggggg cgacggetac gtcacgatcc gcgatggggc tcaccatacg 3360
 gagacactca ctttcaacgc ctgcgactac gatataatg gcacatacgt tactgacaac 3420
 acctacctga cgaaggaggt catcttctac tcccacacag agcatatgtg ggtggaggtc 3480
 aacgagactg agggcgctt ccacatcgac agcattgagt tcgtggagac cgagaagtga 3540

<210> 9

<211> 3444

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 9

atggacctgg atgggaataa gacagagaca gagaccgaga ttgtgaatgg gagcgagagc 60
 agcattgacc cgagcagcgt ttcgtacgct gggaacaata gctactccag cgccctgaac 120
 ctcaattcgt gccagaatag gggcatcgct cagtgggtta acacgctggg cggggctatt 180
 gggcaggccg tgagcatcgg cacatctatc atttactec tggccgcgcc gacactcact 240
 gggctatatt cactggcctt caatctcatc aggaggatgg ggaccggctc caacggctcg 300
 tctatttccg acctgagcat ctgcgatctc ctgagcatca ttaacctgcg ggtttcgcag 360
 gctgtgctca acgacgggat cgetgatttc aatggetccg ttgctgtgta cgacctgtac 420
 ctccacgccc tgcgcagctg gaacaataac cetaacgctg ctactgctga ggagctgagg 480
 acccgcttca ggatcgccga ttcggagttc gagaggatc tgacgagggg ctcgctcaca 540
 catggcggtt ccctcgeccg ccaggacgct caggtctctc tgctcccgtc ctctgttaac 600
 gcggttacc tgcacctget catctctcgc gatgettcgc gctacggggc ctcttggggc 660
 ctcttcaaca ccacgccgca tatcaattac cccgtgagge tgcagcagct cattggcagc 720

tacacgcaact actgcacaca ttggtacaac caggggctga atgagatccg gcagcgcggc 780
 aacactgccg tgaattggct cgagttccac cgctaccgcc gcgacatgac gctgatggtc 840
 ctcgatgtgg tctcgtctgt ctctgccctc gacacgatcc gctaccgaa cgctacagtt 900
 gtgcagctca gccgcaactgt ctacaccgat ccgattggct tcgttaaccg cgggtcaggc 960
 aataggctgt cctggttcga ctggaggaac caggcgaatt tctctactct cgagtcagag 1020
 atgccgacc cctcatcccc actgagcctc aaccacatgt cgatcttcac tgggcctctg 1080
 accctcccag tgtcccctaa cacgcatagg gcccggtct ggtacggcaa ccagaatatg 1140
 ttcacaactg ggtcacagaa ctccggccag accacgaact ctattcagaa tatctcaggc 1200
 ctggagattt tccgcatcga ctctcaggcg tgcaatctca ataacaattc atacggcgtg 1260
 aacagggcgg agttcttcca cggggctagc cagggctcgc agcggctctgt ctaccagga 1320
 tacatccgcc agagcggcct ggacaaccct gtcgttatga atctgcagtc tttctccca 1380
 ggcgagaact cagccacccc tacggcgcag gattacacc acattctgtc caaccgggtt 1440
 aatatcaggg gcgggctcag gcagattgtg gccgacagge gctctctcgt ggtcgtttac 1500
 ggctggacgc acaagtccct gagcaggagg tcaactcgtg ctccagacca gatcaccag 1560
 gtcccagccg ttaaggcgtc cccttctca cattgcaact tcattgccgg cccaggett 1620
 accggcgggg acctggtgtc gctccagccc aacggccagc tcgtcatccc gttccaggtt 1680
 tctgcgccc agacgaacta ccacattcgc atctgctacg tctcgacgtc tgattgcagc 1740
 attaacacaa tctgcaatga cgagacgcat ctgtccacac tcccagcac aacttccagc 1800
 ctggagaacc tccagtcaa tcacctgcat tacttcaacg tgggcacttt caagccaacc 1860
 atcgactcga agctgacgct cgtaacaca tctcctaacg ctaacatcat tatcgacaag 1920
 atcgagttca tcccgtgga taccgcccag cagcagaacg aggacctcga ggccgcgaag 1980
 aaggctgtcg cctccctgtt cacacgcaact agggacggcc tccaggtcaa tgttaaggac 2040
 taccaggtgg atcaggctgc caacctggtc tcatgctct cgcagagca gtacggctac 2100
 gataagaaga tgctgctcga ggccgtgagg gctgctaaga ggctgagcag ggagaggaac 2160
 ctgctccagg accccgattt caacacaatc aactcgaccg aggagaacgg gtggaaggcg 2220
 tcaaatggcg tcaccatctc cgagggcggg ccattctaca agggcagggc tattcagctc 2280
 gcgtctgctc gggagaacta ccccacatac atctaccaga aggtggatgc ctccgagctg 2340
 aagccataca cccgctaccg cctcgacggc ttctgtaagt cgtctcagga cctggagatt 2400
 gatctcatcc accatcaciaa ggtgcacctg gtcaagaacg ttccggacia tctcgtgagc 2460
 gatacgtacc ccgacgattc atgctccgga atcaacaggt gccaggagca gcagatggtc 2520
 aacgcgcagc tggagaccga gcatcaccat ccgatggact gctgcgagge tgctcagacg 2580
 cacgagttct catctacat cgacacaggg gatctgaaca gctcggtcga tcagggcatt 2640
 tgggcatct tcaaggttag gaccacggac gggtagceta cctcggcaa cctggagctg 2700
 gtggaggtcg ggccaactgag cggcgagtcg ctcgagaggg agcagagggg caacactaag 2760
 tggtcgctg agctgggccc caagagggt gagaccgacc gcgtctacca ggatgccaa 2820
 cagagcatca atcaactgtt cgttgactac caggatcage agctcaacc cgagattggc 2880
 atggcggaca tcatggatgc tcagaacctg gtggccagca tctcggacgt gtacagcgat 2940
 gcggctctcc agattccagg aatcaactac gagatctaca cggagctgtc gaacaggctc 3000
 cagcaggcct cctacctgta cacaagccgg aacgcggtcc agaattggga cttcaacaat 3060

ggctcgatt catggaatgc tacggctggg gcttccgtgc agcaggatgg caacacacac 3120
 ttctggtcc tctccattg ggacgcgag gtagccage agttccgct gcagccgaac 3180
 tgcaagtatg tgctgagggt cactgctgag aaggttggcg gggcgacgg ctacgtgacc 3240
 atcagggacg atgcgacca taccgagacg ctgacattca acgcttgca ctacgacatc 3300
 aacggcacct acgtgacaga caacacttac atcaccaagg agtggtctt ccaccggag 3360
 actcagcata tgtgggttga ggtgaacgag accgagggcg cttccacct tgactccatc 3420
 gagttcgtcg agaccgagaa gtga 3444

<210> 10

<211> 3474

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 10

atgaacagga acaaccagaa cgagtacgag attattgacg cccccattg cggtgcccc 60
 tccgacgacg atgtgaagta cccactgget agcgacceca acgctgctct gcagaacatg 120
 aattacaagg attacctcca gatgaccgac gaggattaca cggactcgta catcaacca 180
 tccctcagca tttcgggcag ggacgctgtc cagacagccc tgactgtggt cggccgcatc 240
 ctcgggggcg tgggcgttcc cttctcaggc cagatttgtt cttctacca gttctctctg 300
 aataccctct ggccagtga cgacacggcg atctgggagg ctttcatgcg ccaggtggag 360
 gagctggtca atcagcagat tacggagttc gccaggaacc aggtctctgc gcggtgcag 420
 ggctcggggg actccttcaa tgtctaccag aggagcctgc agaactggct cgccgaccgc 480
 aacgatacca ggaatctctc cgttgtgcgc gccagttca tcgcgctcga cctggatttc 540
 gtgaatgcca ttctctgtt cgctgtgaac ggccagcagg tcccgtctct gtccgtttac 600
 gctcaggccg tgaacctgca tctctgctc ctgaaggatg cttegtcttt cggcgagggg 660
 tggggcttca cacagggcga gatctctact cactacgacc gccagctcga gctgacagcg 720
 aggtacacta attactgca gacctggtac aacacggggc tggacaggct caggggaacc 780
 aacacggagt cctggctccg ctaccaccag ttccgcaggg agatgactct ggtcgttctc 840
 gatgtggtcg ccctgttccc atactacgac gtccgctct acccaaccgg ctccaacct 900
 cagctgacaa gggagggtga cactgacct ategtctta accaccagc taatgtgggg 960
 ctctgcaggc gctggggaac caaccgtac aatacgtta gcgagctgga gaacgcgttc 1020
 atccggccac ctcatctgtt cgatcgcatc cagtctctct caatttcag caataggttc 1080
 cctgtctctg ctaacttcat ggactactgg tetggccaca cgctgaggcg gagctacctc 1140
 aacgattcgg ctgtgcagga ggactcctac ggcctcatca ccacgacag gccaccatt 1200
 aaccgggggg tcgatggcac caaccggatc gactcgacgg cgggtgactt ccgctctgct 1260
 ctcatcggga tttacggcgt taacaggget tccttcgtgc caggcgggct gttcaatggc 1320
 actaccagcc cagctaacgg cgggtgcagg gacctgtacg ataccaacga cgagctgcca 1380
 ccagacgagt ccacaggetc atccaactcat cgcctctcgc acgtcacatt cttctcttct 1440
 cagactaatc aggccgggtc aatcgcgaac gctggtctcc ttcccaccta cgtgtggacg 1500

cgcagggacg tcgatctgaa caacacgata actccgaacc gcattacgea gctcccactg 1560
 gtgaaggctt ctgctccagt ctccaggcac acagttctga aggggcccgg cttcaccggc 1620
 gggggcatcc tccggcgcac taccaatggg accttcggca cgctgagggt gaccgtcaac 1680
 agcccactga cgcagcagta caggctccgc gtgaggttcg cttctacggg caatttctca 1740
 atcaggctcc tgaggggggg cgtgagcatt ggggacgtca ggctgggctc gacaatgaac 1800
 cggggccagg agctgacata cgagagcttc ttactcgcg agttcacgac aactggccca 1860
 ttcaatccac ctttcacctt cacgcaggcc caggagatcc tcacagttaa cgctgagggc 1920
 gtgtcgactg ggggcgagta ctacattgat aggatcgaga ttgttccagt gaaccagct 1980
 agggaggctg aggaggacct ggaggctgcc aagaaggctg tggccagcct gttcacacgc 2040
 actagggacg gcctccaggt caatgttacc gattaccagg tcgacagggc ggctaacctg 2100
 gtttcatgcc tctccgatga gcagtactcc cagcacaaga agatgctcct ggaggccgtc 2160
 cgggctgcta agcgcctgtc acgggagcgc aacctctgc aggacctga tttcaacacg 2220
 atcaactcca ctgaggagaa tgggtggaag gccagcaacg gcgtgacct ttcggagggg 2280
 ggcccgttct tcaagggccg cgcgctccag ctggctagcg ctagggagaa ctaccctacg 2340
 tacatctacc agaaggtcga tgcgtcggtt ctgaagccgt acacacgcta ccgctcgac 2400
 ggcttcgtga agtccctcca ggatctggag atcgacctc ttcaccatca caaggtccat 2460
 ctggttaaga acgtgcccga caatctcgtc tccgatacct acagcgacgg gtctctgcagc 2520
 ggaatcaacc gctgcgatga gcagcagcag gtggatatgc agctcgacgc cgagcatcac 2580
 ccaatggact gctgcgagge tgcccagacc cagagttct cttctacat caatacgggg 2640
 gatctgaacg cctccgttga ccagggcatt tgggttgtgc tcaaagttag gaccacggac 2700
 gggtagccta ccctgggcaa cctcgagctg gtggaggtcg ggccgctgag cggcgagtcg 2760
 ctcgagaggg agcagagggga taacgctaag tggaatgctg agctgggcag gaagagggt 2820
 gagaccgaca gggctctacct ggctgctaag caggcgatca atcacctctt cgtggattac 2880
 caggaccagc agctgaacct tgagatcggc ctcgctgaga ttaacgaggc ctctaactctg 2940
 gtcaagtcca tctctgggggt ttactcagat actctctgc agatcccggg aattaactac 3000
 gagatttaca ccgagctgtc cgaccggctc cagcaggctt cctacctta cacgagccgc 3060
 aacgccgtgc agaatgggga tttcaactcg ggctggact cttggaacgc gacaactgat 3120
 gcttctgtcc agcaggacgg ctcaacccat ttctctgtc tgtcacactg ggacgctcag 3180
 gtgtcccagc agatgagggt caacctgaat tgcaagtac tcctcagggt tacggcgaag 3240
 aaggctcggg gcggggatgg ctacgtcaca atcaggacg gcgcgatca ccaggagacc 3300
 ctcacgttca atgcttgcga ctacgatgtc aacggccat acgttaacga caattcttac 3360
 atcaactaagg aggtcgtttt ctaccccgag accaagcaca tgtgggttga ggtgtctgag 3420
 tcggagggtc cgttctacat tgatagcatt gaggttcatt agacgcagga gtga 3474

<210> 11

<211> 3522

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 11

atgaaccgga acaaccaggg cgagtacgag attattgatg cctccacttg cggctgctcc 60
tcagatgatg tcgtccagta cccactcget cgcgaccega acgctgcttt ccagaacatg 120
aattacaagg actacctgaa gatgtctgac ggcgattacg tcgattcata cattaacca 180
ggcctgtcca tcgggaggag ggacgtcacg ctcacaggcg tcgggatcgt tgctctcatt 240
gtgggcaccc tgggcggccc agttggcggg attgtgacgg gcctgatctc cagcctcctg 300
gggctcctgt ggccaagcaa cgacaatgat gtgtgggagg cttcatggc gcagatcgag 360
gagctgattg agcagaggat cgctgaccag gtggtcggga acgccctgga caatctcacc 420
ggcctgaggg attactacaa ccagtacctc ctggctctcg aggagtggca ggagaggccc 480
aatgccgtga ggtctacgct ggtcttcaac cggttcgaga cgetccattc acacttcgtg 540
acatcaatgc catccttcgg cagcgggcct ggcagcgagc gctacgcggt tcagctcctg 600
accgtgtacg ctcaggetgc caacctgcac ctctgctec tgagggacgc tgatatctac 660
ggcgtcgggt gggggctcag ggagtcccag atcgacctc acttcaacga gctgcagaat 720
cggacgcgcg attacacaaa ccattgcgtc acagcctaca acaatggcct ggaggagatc 780
agggggactt cgccagcttc ttggetgcgc taccaccagt tccggcgcga gaccacgctc 840
attgccctcg acctggtgge gatettccca tactacaatg tcagggagta cccaattggc 900
gttaaccctc agctcacgcg ggacgtgtac acagatccga tcggcgtcac gttcaggcgg 960
gaggactggg agacaggcgt cgagtgcagg ccgtgggtta atacccata catgtctttc 1020
tcagatctgg agaacgccat cattaggccg ccccatctct tcgagacgct ccggaatctg 1080
acgattcaca caggcaggta caacctggtc ggcggggcga ggttcatcga gggctgggtc 1140
gggcattccg ttactaatac caggetgggc aacagcactg tgttcaccag caattacggg 1200
tcgctccac ctcggtcca ggtgttcaac ttcacgaatt tcgacgteta ccagatcaac 1260
acacgggccg attcgacggg cacattccgc attccggggt tcgcggtcac tagggctcag 1320
ttcatccccg gcgggaccta ctccgtggct caccgcgacc caggcgcttg ccagcaggac 1380
tacgattcaa ttgaggagct gccctccctg gaccagatg agcctatcaa ccggtcctac 1440
agccatcgcc tctcacacgt caccctgtac aagtacactc tctccgacac cgattacggc 1500
gtgatcaatt acaccgacta cgggagcatg ccagcttacg tgtggacgca tcgcgacgctc 1560
gatctgacta acaccattac ggcggatagg atcacgcagc tcccgtggt gaaggcttcg 1620
acactccccg ccggcacaac tgttgtgaag gggcccggct tcaccggcgg ggacatcctg 1680
aggaggacca cgaatggcac gttcgggaca ctccacgtga gggteaacag cccactgacc 1740
cagcagtaca ggctccgggt ccgcttcgct tcgacgggea acttctctat tagggtgctg 1800
aggggcggga catctatcgg cgacgctcgc ttcgggtcaa ctatgaacag gggccaggag 1860
ctgacttacg agtccttcgt gaccgcgag ttcacaacta ccggcccgtt caatecgccc 1920
ttcacattca ctcagacca ggagatcctg actgtcaacg ctgagggcgt ttcgaccggc 1980
ggggagtact acatcgactc tattgagatc gttccagtga acccaaccag ggaggctgag 2040
gaggatctcg aggetgctaa gaaggccgct gcgagcctgt tcacaggac acgggacggc 2100
ctccaggtea atgttacgga ctaccaggtt gatagggtg ctaacctcgt gctgtgctc 2160
tccgacgagc agtacgcca cgataagaag atgctcctgg aggcggtgag ggctgctaag 2220
aggctgagca gggagaggaa cctcctgcag gaccctgatt tcaacgagat caattctact 2280

gaggactcag gctggaagac cagcaacggg atcattatct cggagggecg gccgttcttc 2340
 aagggccggg ccctgcagct cgcttccgct cgcgagaact accctaccta catctaccag 2400
 aaggtggact cgtctatgct gaagccgtac acgaggtaca agctcgacgg cttcgtgcag 2460
 tcatcccagg atctcgagat tgagctgac caccatcaca aggtgcacct cgtcaagaac 2520
 gttccagaca atctggctct cgacacctac cctgatggct cgtgcaacgg aatcaaccgc 2580
 tgcgaggagc agcagatggt gaactctcag ctggagacgg agcatcacc tatggactgc 2640
 tgcgaggcct cacagactca tgagttcagc tcgtacatcc acaccggcga cctcaaccgc 2700
 tctgtcgatc aggggatttg ggtcgttctg aagatcagga cgacagacgg ctcggctacc 2760
 ctcgggaacc tggagctggt ggaggtcggc ccctgtcag gggagtcct cgagagggag 2820
 cagagggaca acgccaagtg gaatgctgag ctgggccgga agcgcgctga ggctgatcgc 2880
 gtgtaccagg gcgctaagca ggccatcaat caectcttcg tcgactacca ggatcagcag 2940
 ctgaaccctg aggttggcct cgcggagatc agcagagctc ggaacctgat tgagtcgac 3000
 tctgacgtgt actgcgatgc cgtctctcgc attccgggaa tcaactacga gatgtacacg 3060
 gagctgtcca acaggetgca gcaggetget tactgtaca caagccgca cgcggtgcag 3120
 aatggcgact tcaactccgg gctcgatagc tggaatgcta ctaccgacgc caccgttcag 3180
 caggatggca acatgtactt cctggtgctc agccactggg acgcccaggt ttgcagcag 3240
 ttcccgctgc agccaaattg caagtatgtg ctgagggtca cagcgaagaa ggtcgggaac 3300
 ggcgacggct acgtgactat ccaggatggc gcgcatacc gcgagactct gaccttcaat 3360
 gcttgcgact acgatgttaa cggcacgcat gtgaacgaca attctacct cacaaaggag 3420
 ctggagttct acccgaagac tgagcacatg tgggttgagg tgagcgagac tgagggcacc 3480
 ttctaccttg attcgatcga gctgattgag acccaggagt ga 3522

<210> 12

<211> 3540

<212> DNA

<213> 人工序列

<220>

<223> 合成Cry基因

<400> 12

atggggggga agtctatgaa caggaacaac cagggecagat acgagattat tgatgcctcc 60
 acatgcgggt gctccagcga cgacgtggc cagtaccac tcgctcgcga ccctaaccgt 120
 gctttccaga acatgaatta caaggactac ctgaagatgt ccgacggcga ttacgtggat 180
 agctacatta acccaggcct ctcgatcggg aggagggacg tcaactctgac cggggttggc 240
 atcgtggcgc tgattgttgg cacactcggc gggcctgtgg gcgggattgt cactggcctc 300
 atctccagcc tcctggggct cctgtggcca tccaacgaca atgatgtctg ggaggcgttc 360
 atggetcaga tcgaggagct gattgagcag cgcatacgg accaggtggt caggaacgct 420
 ctcgacaatc tgaccggcct cagggattac tacaaccagt acctctggc tctcgaggag 480
 tggcaggaga ggccaaatgc cgtgcgetcc acgctcgttt tcaaccgctt cgagaccctg 540
 cacagccatt tcgtgacgag catgccgtcg ttccgggtctg gccccgggct ggagcgtac 600
 gctgtgcage tcctgaccgt ctacgcccag getgccaacc tccacctct getctcgcgc 660

gacgctgata tctacggcgc caggtggggg ctcaggaga gccagatcga cctgtacttc 720
aacgagctgc agaatcggac acgcgattac actaaccact gcgtcaccgc ctacaacaat 780
ggcctcgagg agatcagggg gacgtcacca gcttcctggc tccgctacca ccagttccgg 840
agggagacca cgctgattgc gctcgacctg gtggctatct tcccctacta caatgtgcgc 900
gagtaccgga ttggcgtcaa cccccagctg accagggacg ttacaccga cccgatcggc 960
gtgacattca ggcgggagga ctgggagact ggctgggagt gcaggccgtg ggtcaatacc 1020
ccatacatgt ctttctcaga cctcgagaac gccatcatta ggccgccccca cctgttcgag 1080
acgctgagga atctcaccat tcatacgggc aggtacaacc tggctggcgg ggcgcgcttc 1140
atcgagggct gggttgggca ctcagtacg aacacaaggc tcggcaattc cacagtgttc 1200
acttccaact acggcagcct gccacctcgg ttccaggttt teaacttac aaatttcgac 1260
gtgtaccaga tcaaacactag ggccgattcg actggcaect tccgattcc agggttcgcc 1320
gttaccgcg cgagttcat ccttggcggg acgtactcgg tggtcaccg cgaccggggc 1380
gcttgccagc aggactacga tagcattgag gagctgcct cgctcgacc agatgagcct 1440
atcaacaggt cctacagcca ccggetgtct catgtcacc tctacaagta caccctgtca 1500
gacacggatt acggcgtgat caattacacc gactacgggt ccatgccagc ttacgtttgg 1560
acgcaccggg acgtggatct cacgaacaca attactgcc accgcatcac acagctccca 1620
ctgggtgaagg ccagcactct gctcgcgggc acaactgttg tgaagggcc tgggttcacc 1680
ggcggggaca tcctcaggag gaccacgaat ggcacctcg ggacgctgca tgtccgcgtt 1740
aactccccgc tcacacagca gtacaggctg cgggtgcgct tcgcttcgac tggcaacttc 1800
tctattcgcg tcctcagggg cgggacctcc atcggcgac ctaggttcgg gagcacgatg 1860
aacaggggcc aggagctgac atacagctc ttctcacta gggagtac aactaccggc 1920
ccgttcaate cgccttcac cttcacgcag acacaggaga ttctaccgt taacgctgag 1980
ggcgtgagca cgggcgggga gtactacatc gactcgatcg agattgtgcc agtcaacca 2040
accagggagg ctgaggagga tctggaggct gctaagaagg ccgtggcgag cctcttact 2100
aggaccggg acggcctgca ggttaatgtg acggactacc aggtcgatca ggccgcaac 2160
ctggttagct gcctctcgga cgagcagtac ggctacgata agaagatgct cctggaggcc 2220
gtccgcgctg ctaagaggct ctcgagggag aggaacctc tgcaggacc cgatttcaac 2280
acaattaatt ctactgagga gaacggctgg aaggcctcta atggggtgac catctcagag 2340
ggcgggcat tctacaaggg cagggcgtc cagctggctt cagctcgga gaactaccc 2400
acctacatct accagaaggt cgacgctcc gagctgaagc catacagcg ctaccgctg 2460
gatggcttcg tgaagtcgtc tcaggacctg gagatcgatc teattacca tcacaaggct 2520
cacctcgtta agaactgccc ggacaatctg gtctccgata cctacccga cgattcgtgc 2580
tctggaatca acaggtgcca ggagcagcag atggtgaacg cccagctcga gacggagcat 2640
caccatccta tggactgctg cgaggcggtt cagacctatg agttctcctc ctacatcgac 2700
acgggcgac tcaacagctc ggtcgaccag gggatctggg cgattttcaa ggttaggacg 2760
acagatggct acgctaccct ggggaatctc gagctggctg aggttgccc cctctctggg 2820
gagtcactgg agagggagca gagggacaac acaaagtgtt ctgctgagct gggcaggaag 2880
cgggctgaga ctgaccgct ctaccaggat gccaaagcgt ccatcaatca cctcttcgtg 2940
gactaccagg atcagcagct gaaccctgag attggcatgg ctgacatcat ggatgccag 3000

aacctcgtcg cgtcaatctc cgacgtctac agcgatgcgg ttctgcagat cccgggcatt 3060
aattacgaga tctacacaga gctgtcgaac aggetccage aggcgtcata cctctacacg 3120
tcccggaacg ctgtgcagaa tggcgacttc aacaatgggc tggattcgtg gaatgcgaca 3180
gctggcgcct ctgtgcagca ggacgggaac actcacttcc tcgtcctgtc tcattgggat 3240
gcccaggctc cacagcagtt ccgggttcag ccgaactgca agtatgtgct gcgcgttacc 3300
gctgagaaag tgggcggggg cgacggctac gtcacgatcc gcgatggggc tcaccatacg 3360
gagacactca ctttcaacgc ctgcgactac gatatcaatg gcacatacgt tactgacaac 3420
acctacctga cgaaggaggt catcttctac tcccacacag agcatatgtg ggtggaggtc 3480
aacgagactg agggcgcctt ccacctcgac agccttgagt tcgtggagac cgagaagtga 3540

<210> 13

<211> 1147

<212> PRT

<213> 苏云金芽孢杆菌

<400> 13

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

195	200	205
Leu Arg Asp Ala Ser Arg Tyr Gly Ala Ser Trp Gly Leu Phe Asn Thr		
210	215	220
Thr Pro His Ile Asn Tyr Pro Val Arg Leu Gln Gln Leu Ile Gly Ser		
225	230	235
Tyr Thr His Tyr Cys Thr His Trp Tyr Asn Gln Gly Leu Asn Glu Ile		
245	250	255
Arg Gln Arg Gly Asn Thr Ala Val Asn Trp Leu Glu Phe His Arg Tyr		
260	265	270
Arg Arg Asp Met Thr Leu Met Val Leu Asp Val Val Ser Leu Phe Ser		
275	280	285
Ala Leu Asp Thr Ile Arg Tyr Pro Asn Ala Thr Val Val Gln Leu Ser		
290	295	300
Arg Thr Val Tyr Thr Asp Pro Ile Gly Phe Val Asn Arg Gly Ser Gly		
305	310	315
Asn Arg Leu Ser Trp Phe Asp Trp Arg Asn Gln Ala Asn Phe Ser Thr		
325	330	335
Leu Glu Ser Glu Met Pro Thr Pro Ser Ser Pro Leu Ser Leu Asn His		
340	345	350
Met Ser Ile Phe Thr Gly Pro Leu Thr Leu Pro Val Ser Pro Asn Thr		
355	360	365
His Arg Ala Arg Val Trp Tyr Gly Asn Gln Asn Met Phe Thr Thr Gly		
370	375	380
Ser Gln Asn Ser Gly Gln Thr Thr Asn Ser Ile Gln Asn Ile Ser Gly		
385	390	395
Leu Glu Ile Phe Arg Ile Asp Ser Gln Ala Cys Asn Leu Asn Asn Asn		
405	410	415
Ser Tyr Gly Val Asn Arg Ala Glu Phe Phe His Gly Ala Ser Gln Gly		
420	425	430
Ser Gln Arg Ser Val Tyr Gln Gly Tyr Ile Arg Gln Ser Gly Leu Asp		
435	440	445
Asn Pro Val Val Met Asn Leu Gln Ser Phe Leu Pro Gly Glu Asn Ser		
450	455	460
Ala Thr Pro Thr Ala Gln Asp Tyr Thr His Ile Leu Ser Asn Pro Val		
465	470	475
Asn Ile Arg Gly Gly Leu Arg Gln Ile Val Ala Asp Arg Arg Ser Ser		
485	490	495
Val Val Val Tyr Gly Trp Thr His Lys Ser Leu Ser Arg Arg Ser Leu		
500	505	510

Val Ala Pro Asp Gln Ile Thr Gln Val Pro Ala Val Lys Ala Ser Pro			
515	520	525	
Ser Ser His Cys Thr Ile Ile Ala Gly Pro Gly Phe Thr Gly Gly Asp			
530	535	540	
Leu Val Ser Leu Gln Pro Asn Gly Gln Leu Val Ile Pro Phe Gln Val			
545	550	555	560
Ser Ala Pro Glu Thr Asn Tyr His Ile Arg Ile Cys Tyr Val Ser Thr			
	565	570	575
Ser Asp Cys Ser Ile Asn Thr Ile Cys Asn Asp Glu Thr His Leu Ser			
	580	585	590
Thr Leu Pro Ser Thr Thr Ser Ser Leu Glu Asn Leu Gln Cys Asn His			
	595	600	605
Leu His Tyr Phe Asn Val Gly Thr Phe Lys Pro Thr Ile Asp Ser Lys			
	610	615	620
Leu Thr Leu Val Asn Thr Ser Pro Asn Ala Asn Ile Ile Ile Asp Lys			
625	630	635	640
Ile Glu Phe Ile Pro Val Asp Thr Ala Gln Gln Gln Asn Glu Asp Leu			
	645	650	655
Glu Ala Ala Lys Lys Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp			
	660	665	670
Gly Leu Gln Val Asn Val Lys Asp Tyr Gln Val Asp Gln Ala Ala Asn			
	675	680	685
Leu Val Ser Cys Leu Ser Asp Glu Gln Tyr Gly Tyr Asp Lys Lys Met			
	690	695	700
Leu Leu Glu Ala Val Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn			
705	710	715	720
Leu Leu Gln Asp Pro Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn			
	725	730	735
Gly Trp Lys Ala Ser Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe			
	740	745	750
Tyr Lys Gly Arg Ala Ile Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro			
	755	760	765
Thr Tyr Ile Tyr Gln Lys Val Asp Ala Ser Glu Leu Lys Pro Tyr Thr			
	770	775	780
Arg Tyr Arg Leu Asp Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile			
785	790	795	800
Asp Leu Ile His His His Lys Val His Leu Val Lys Asn Val Pro Asp			
	805	810	815
Asn Leu Val Ser Asp Thr Tyr Pro Asp Asp Ser Cys Ser Gly Ile Asn			

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

Asn Glu Thr Glu Gly Ala Phe His Ile Asp Ser Ile Glu Phe Val
 1130 1135 1140
 Glu Thr Glu Lys
 1145
 <210> 14
 <211> 1157
 <212> PRT
 <213> 苏云金芽孢杆菌
 <400> 14
 Met Asn Arg Asn Asn Gln Asn Glu Tyr Glu Ile Ile Asp Ala Pro His
 1 5 10 15
 Cys Gly Cys Pro Ser Asp Asp Asp Val Lys Tyr Pro Leu Ala Ser Asp
 20 25 30
 Pro Asn Ala Ala Leu Gln Asn Met Asn Tyr Lys Asp Tyr Leu Gln Met
 35 40 45
 Thr Asp Glu Asp Tyr Thr Asp Ser Tyr Ile Asn Pro Ser Leu Ser Ile
 50 55 60
 Ser Gly Arg Asp Ala Val Gln Thr Ala Leu Thr Val Val Gly Arg Ile
 65 70 75 80
 Leu Gly Ala Leu Gly Val Pro Phe Ser Gly Gln Ile Val Ser Phe Tyr
 85 90 95
 Gln Phe Leu Leu Asn Thr Leu Trp Pro Val Asn Asp Thr Ala Ile Trp
 100 105 110
 Glu Ala Phe Met Arg Gln Val Glu Glu Leu Val Asn Gln Gln Ile Thr
 115 120 125
 Glu Phe Ala Arg Asn Gln Ala Leu Ala Arg Leu Gln Gly Leu Gly Asp
 130 135 140
 Ser Phe Asn Val Tyr Gln Arg Ser Leu Gln Asn Trp Leu Ala Asp Arg
 145 150 155 160
 Asn Asp Thr Arg Asn Leu Ser Val Val Arg Ala Gln Phe Ile Ala Leu
 165 170 175
 Asp Leu Asp Phe Val Asn Ala Ile Pro Leu Phe Ala Val Asn Gly Gln
 180 185 190
 Gln Val Pro Leu Leu Ser Val Tyr Ala Gln Ala Val Asn Leu His Leu
 195 200 205
 Leu Leu Leu Lys Asp Ala Ser Leu Phe Gly Glu Gly Trp Gly Phe Thr
 210 215 220
 Gln Gly Glu Ile Ser Thr Tyr Tyr Asp Arg Gln Leu Glu Leu Thr Ala
 225 230 235 240

Arg Tyr Thr Asn Tyr Cys Glu Thr Trp Tyr Asn Thr Gly Leu Asp Arg
 245 250 255
 Leu Arg Gly Thr Asn Thr Glu Ser Trp Leu Arg Tyr His Gln Phe Arg
 260 265 270
 Arg Glu Met Thr Leu Val Val Leu Asp Val Val Ala Leu Phe Pro Tyr
 275 280 285
 Tyr Asp Val Arg Leu Tyr Pro Thr Gly Ser Asn Pro Gln Leu Thr Arg
 290 295 300
 Glu Val Tyr Thr Asp Pro Ile Val Phe Asn Pro Pro Ala Asn Val Gly
 305 310 315 320
 Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu
 325 330 335
 Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Leu Asn Ser
 340 345 350
 Leu Thr Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp
 355 360 365
 Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala
 370 375 380
 Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile
 385 390 395 400
 Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp
 405 410 415
 Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe
 420 425 430
 Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly
 435 440 445
 Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser
 450 455 460
 Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe
 465 470 475 480
 Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr
 485 490 495
 Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro
 500 505 510
 Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser
 515 520 525
 Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu
 530 535 540
 Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn

545	550	555	560
Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr			
	565	570	575
Gly Asn Phe Ser Ile Arg Leu Leu Arg Gly Gly Val Ser Ile Gly Asp			
	580	585	590
Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu			
	595	600	605
Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro			
610	615	620	
Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly			
625	630	635	640
Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Arg Ile Glu Ile Val Pro			
	645	650	655
Val Asn Pro Ala Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys			
	660	665	670
Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn			
	675	680	685
Val Thr Asp Tyr Gln Val Asp Arg Ala Ala Asn Leu Val Ser Cys Leu			
	690	695	700
Ser Asp Glu Gln Tyr Ser His Asp Lys Lys Met Leu Leu Glu Ala Val			
705	710	715	720
Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro			
	725	730	735
Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn Gly Trp Lys Ala Ser			
	740	745	750
Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala			
	755	760	765
Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln			
	770	775	780
Lys Val Asp Ala Ser Val Leu Lys Pro Tyr Thr Arg Tyr Arg Leu Asp			
785	790	795	800
Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile Asp Leu Ile His His			
	805	810	815
His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Ser Asp			
	820	825	830
Thr Tyr Ser Asp Gly Ser Cys Ser Gly Ile Asn Arg Cys Asp Glu Gln			
	835	840	845
Gln Gln Val Asp Met Gln Leu Asp Ala Glu His His Pro Met Asp Cys			
	850	855	860

Cys Glu Ala Ala Gln Thr His Glu Phe Ser Ser Tyr Ile Asn Thr Gly			
865	870	875	880
Asp Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Val			
	885	890	895
Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly Asn Leu Glu Leu Val Glu			
	900	905	910
Val Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn			
	915	920	925
Ala Lys Trp Asn Ala Glu Leu Gly Arg Lys Arg Ala Glu Thr Asp Arg			
	930	935	940
Val Tyr Leu Ala Ala Lys Gln Ala Ile Asn His Leu Phe Val Asp Tyr			
945	950	955	960
Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly Leu Ala Glu Ile Asn Glu			
	965	970	975
Ala Ser Asn Leu Val Lys Ser Ile Ser Gly Val Tyr Ser Asp Thr Leu			
	980	985	990
Leu Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asp			
	995	1000	1005
Arg Leu Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val			
	1010	1015	1020
Gln Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr			
	1025	1030	1035
Thr Asp Ala Ser Val Gln Gln Asp Gly Ser Thr His Phe Leu Val			
	1040	1045	1050
Leu Ser His Trp Asp Ala Gln Val Ser Gln Gln Met Arg Val Asn			
	1055	1060	1065
Leu Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Lys Lys Val Gly			
	1070	1075	1080
Gly Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Gln			
	1085	1090	1095
Glu Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Val Asn Gly Thr			
	1100	1105	1110
Tyr Val Asn Asp Asn Ser Tyr Ile Thr Lys Glu Val Val Phe Tyr			
	1115	1120	1125
Pro Glu Thr Lys His Met Trp Val Glu Val Ser Glu Ser Glu Gly			
	1130	1135	1140
Ser Phe Tyr Ile Asp Ser Ile Glu Phe Ile Glu Thr Gln Glu			
	1145	1150	1155

<210> 15

<211> 1173

<212> PRT

<213> 苏云金芽孢杆菌

<400> 15

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

275	280	285
Phe Pro Tyr Tyr Asn Val Arg	Glu Tyr Pro Ile Gly Val Asn Pro Gln	
290	295	300
Leu Thr Arg Asp Val Tyr Thr Asp Pro Ile Gly Val Thr Phe Arg Arg		
305	310	315
Glu Asp Trp Glu Thr Gly Val Glu Cys Arg Pro Trp Val Asn Thr Pro		
	325	330
Tyr Met Ser Phe Ser Asp Leu Glu Asn Ala Ile Ile Arg Pro Pro His		
	340	350
Leu Phe Glu Thr Leu Arg Asn Leu Thr Ile His Thr Gly Arg Tyr Asn		
	355	360
Leu Val Gly Gly Ala Arg Phe Ile Glu Gly Trp Val Gly His Ser Val		
	370	380
Thr Asn Thr Arg Leu Gly Asn Ser Thr Val Phe Thr Ser Asn Tyr Gly		
385	390	395
Ser Leu Pro Pro Arg Phe Gln Val Phe Asn Phe Thr Asn Phe Asp Val		
	405	410
Tyr Gln Ile Asn Thr Arg Ala Asp Ser Thr Gly Thr Phe Arg Ile Pro		
	420	425
Gly Phe Ala Val Thr Arg Ala Gln Phe Ile Pro Gly Gly Thr Tyr Ser		
	435	440
Val Ala His Arg Asp Pro Gly Ala Cys Gln Gln Asp Tyr Asp Ser Ile		
	450	455
Glu Glu Leu Pro Ser Leu Asp Pro Asp Glu Pro Ile Asn Arg Ser Tyr		
465	470	475
Ser His Arg Leu Ser His Val Thr Leu Tyr Lys Tyr Thr Leu Ser Asp		
	485	490
Thr Asp Tyr Gly Val Ile Asn Tyr Thr Asp Tyr Gly Ser Met Pro Ala		
	500	505
Tyr Val Trp Thr His Arg Asp Val Asp Leu Thr Asn Thr Ile Thr Ala		
	515	520
Asp Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Thr Leu Pro Ala		
	530	535
Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu		
545	550	555
Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu His Val Arg Val Asn		
	565	570
Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr		
	580	585
		590

Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly Thr Ser Ile Gly Asp
 595 600 605
 Ala Arg Phe Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu
 610 615 620
 Ser Phe Val Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro
 625 630 635 640
 Phe Thr Phe Thr Gln Thr Gln Glu Ile Leu Thr Val Asn Ala Glu Gly
 645 650 655
 Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Ser Ile Glu Ile Val Pro
 660 665 670
 Val Asn Pro Thr Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys
 675 680 685
 Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn
 690 695 700
 Val Thr Asp Tyr Gln Val Asp Arg Ala Ala Asn Leu Val Leu Cys Leu
 705 710 715 720
 Ser Asp Glu Gln Tyr Ala His Asp Lys Lys Met Leu Leu Glu Ala Val
 725 730 735
 Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro
 740 745 750
 Asp Phe Asn Glu Ile Asn Ser Thr Glu Asp Ser Gly Trp Lys Thr Ser
 755 760 765
 Asn Gly Ile Ile Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala
 770 775 780
 Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln
 785 790 795 800
 Lys Val Asp Ser Ser Met Leu Lys Pro Tyr Thr Arg Tyr Lys Leu Asp
 805 810 815
 Gly Phe Val Gln Ser Ser Gln Asp Leu Glu Ile Glu Leu Ile His His
 820 825 830
 His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Leu Asp
 835 840 845
 Thr Tyr Pro Asp Gly Ser Cys Asn Gly Ile Asn Arg Cys Glu Glu Gln
 850 855 860
 Gln Met Val Asn Ser Gln Leu Glu Thr Glu His His Pro Met Asp Cys
 865 870 875 880
 Cys Glu Ala Ser Gln Thr His Glu Phe Ser Ser Tyr Ile His Thr Gly
 885 890 895
 Asp Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Ile

900	905	910
Arg Thr Thr Asp Gly Ser Ala Thr	Leu Gly Asn Leu Glu Leu Val Glu	
915	920	925
Val Gly Pro Leu Ser Gly Glu Ser	Leu Glu Arg Glu Gln Arg Asp Asn	
930	935	940
Ala Lys Trp Asn Ala Glu Leu Gly	Arg Lys Arg Ala Glu Ala Asp Arg	
945	950	955
Val Tyr Gln Gly Ala Lys Gln Ala	Ile Asn His Leu Phe Val Asp Tyr	
965	970	975
Gln Asp Gln Gln Leu Asn Pro Glu	Val Gly Leu Ala Glu Ile Ser Glu	
980	985	990
Ala Arg Asn Leu Ile Glu Ser Ile	Ser Asp Val Tyr Cys Asp Ala Val	
995	1000	1005
Leu Arg Ile Pro Gly Ile Asn Tyr	Glu Met Tyr Thr Glu Leu Ser	
1010	1015	1020
Asn Arg Leu Gln Gln Ala Ala Tyr	Leu Tyr Thr Ser Arg Asn Ala	
1025	1030	1035
Val Gln Asn Gly Asp Phe Asn Ser	Gly Leu Asp Ser Trp Asn Ala	
1040	1045	1050
Thr Thr Asp Ala Thr Val Gln Gln	Asp Gly Asn Met Tyr Phe Leu	
1055	1060	1065
Val Leu Ser His Trp Asp Ala Gln	Val Ser Gln Gln Phe Arg Val	
1070	1075	1080
Gln Pro Asn Cys Lys Tyr Val Leu	Arg Val Thr Ala Lys Lys Val	
1085	1090	1095
Gly Asn Gly Asp Gly Tyr Val Thr	Ile Gln Asp Gly Ala His His	
1100	1105	1110
Arg Glu Thr Leu Thr Phe Asn Ala	Cys Asp Tyr Asp Val Asn Gly	
1115	1120	1125
Thr His Val Asn Asp Asn Ser Tyr	Ile Thr Lys Glu Leu Glu Phe	
1130	1135	1140
Tyr Pro Lys Thr Glu His Met Trp	Val Glu Val Ser Glu Thr Glu	
1145	1150	1155
Gly Thr Phe Tyr Ile Asp Ser Ile	Glu Leu Ile Glu Thr Gln Glu	
1160	1165	1170
<210> 16		
<211> 1179		
<212> PRT		
<213> 苏云金芽孢杆菌		

<400> 16
 Met Gly Gly Lys Ser Met Asn Arg Asn Asn Gln Gly Glu Tyr Glu Ile
 1 5 10 15
 Ile Asp Ala Ser Thr Cys Gly Cys Ser Ser Asp Asp Val Val Gln Tyr
 20 25 30
 Pro Leu Ala Arg Asp Pro Asn Ala Ala Phe Gln Asn Met Asn Tyr Lys
 35 40 45
 Asp Tyr Leu Lys Met Ser Asp Gly Asp Tyr Val Asp Ser Tyr Ile Asn
 50 55 60
 Pro Gly Leu Ser Ile Gly Arg Arg Asp Val Thr Leu Thr Gly Val Gly
 65 70 75 80
 Ile Val Ala Leu Ile Val Gly Thr Leu Gly Gly Pro Val Gly Gly Ile
 85 90 95
 Val Thr Gly Leu Ile Ser Ser Leu Leu Gly Leu Leu Trp Pro Ser Asn
 100 105 110
 Asp Asn Asp Val Trp Glu Ala Phe Met Ala Gln Ile Glu Glu Leu Ile
 115 120 125
 Glu Gln Arg Ile Ala Asp Gln Val Val Arg Asn Ala Leu Asp Asn Leu
 130 135 140
 Thr Gly Leu Arg Asp Tyr Tyr Asn Gln Tyr Leu Leu Ala Leu Glu Glu
 145 150 155 160
 Trp Gln Glu Arg Pro Asn Ala Val Arg Ser Thr Leu Val Phe Asn Arg
 165 170 175
 Phe Glu Thr Leu His Ser His Phe Val Thr Ser Met Pro Ser Phe Gly
 180 185 190
 Ser Gly Pro Gly Ser Glu Arg Tyr Ala Val Gln Leu Leu Thr Val Tyr
 195 200 205
 Ala Gln Ala Ala Asn Leu His Leu Leu Leu Leu Arg Asp Ala Asp Ile
 210 215 220
 Tyr Gly Ala Arg Trp Gly Leu Arg Glu Ser Gln Ile Asp Leu Tyr Phe
 225 230 235 240
 Asn Glu Leu Gln Asn Arg Thr Arg Asp Tyr Thr Asn His Cys Val Thr
 245 250 255
 Ala Tyr Asn Asn Gly Leu Glu Glu Ile Arg Gly Thr Ser Pro Ala Ser
 260 265 270
 Trp Leu Arg Tyr His Gln Phe Arg Arg Glu Thr Thr Leu Ile Ala Leu
 275 280 285
 Asp Leu Val Ala Ile Phe Pro Tyr Tyr Asn Val Arg Glu Tyr Pro Ile
 290 295 300

Gly Val Asn Pro Gln Leu Thr Arg Asp Val Tyr Thr Asp Pro Ile Gly
 305 310 315 320
 Val Thr Phe Arg Arg Glu Asp Trp Glu Thr Gly Val Glu Cys Arg Pro
 325 330 335
 Trp Val Asn Thr Pro Tyr Met Ser Phe Ser Asp Leu Glu Asn Ala Ile
 340 345 350
 Ile Arg Pro Pro His Leu Phe Glu Thr Leu Arg Asn Leu Thr Ile His
 355 360 365
 Thr Gly Arg Tyr Asn Leu Val Gly Gly Ala Arg Phe Ile Glu Gly Trp
 370 375 380
 Val Gly His Ser Val Thr Asn Thr Arg Leu Gly Asn Ser Thr Val Phe
 385 390 395 400
 Thr Ser Asn Tyr Gly Ser Leu Pro Pro Arg Phe Gln Val Phe Asn Phe
 405 410 415
 Thr Asn Phe Asp Val Tyr Gln Ile Asn Thr Arg Ala Asp Ser Thr Gly
 420 425 430
 Thr Phe Arg Ile Pro Gly Phe Ala Val Thr Arg Ala Gln Phe Ile Pro
 435 440 445
 Gly Gly Thr Tyr Ser Val Ala His Arg Asp Pro Gly Ala Cys Gln Gln
 450 455 460
 Asp Tyr Asp Ser Ile Glu Glu Leu Pro Ser Leu Asp Pro Asp Glu Pro
 465 470 475 480
 Ile Asn Arg Ser Tyr Ser His Arg Leu Ser His Val Thr Leu Tyr Lys
 485 490 495
 Tyr Thr Leu Ser Asp Thr Asp Tyr Gly Val Ile Asn Tyr Thr Asp Tyr
 500 505 510
 Gly Ser Met Pro Ala Tyr Val Trp Thr His Arg Asp Val Asp Leu Thr
 515 520 525
 Asn Thr Ile Thr Ala Asp Arg Ile Thr Gln Leu Pro Leu Val Lys Ala
 530 535 540
 Ser Thr Leu Pro Ala Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr
 545 550 555 560
 Gly Gly Asp Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu
 565 570 575
 His Val Arg Val Asn Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val
 580 585 590
 Arg Phe Ala Ser Thr Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly
 595 600 605
 Thr Ser Ile Gly Asp Ala Arg Phe Gly Ser Thr Met Asn Arg Gly Gln

610	615	620																
Glu	Leu	Thr	Tyr	Glu	Ser	Phe	Val	Thr	Arg	Glu	Phe	Thr	Thr	Thr	Gly			
625	630	635	640															
Pro	Phe	Asn	Pro	Pro	Phe	Thr	Phe	Thr	Gln	Thr	Gln	Glu	Ile	Leu	Thr			
	645	650	655															
Val	Asn	Ala	Glu	Gly	Val	Ser	Thr	Gly	Gly	Glu	Tyr	Tyr	Ile	Asp	Ser			
	660	665	670															
Ile	Glu	Ile	Val	Pro	Val	Asn	Pro	Thr	Arg	Glu	Ala	Glu	Glu	Asp	Leu			
	675	680	685															
Glu	Ala	Ala	Lys	Lys	Ala	Val	Ala	Ser	Leu	Phe	Thr	Arg	Thr	Arg	Asp			
	690	695	700															
Gly	Leu	Gln	Val	Asn	Val	Thr	Asp	Tyr	Gln	Val	Asp	Gln	Ala	Ala	Asn			
705	710	715	720															
Leu	Val	Ser	Cys	Leu	Ser	Asp	Glu	Gln	Tyr	Gly	Tyr	Asp	Lys	Lys	Met			
	725	730	735															
Leu	Leu	Glu	Ala	Val	Arg	Ala	Ala	Lys	Arg	Leu	Ser	Arg	Glu	Arg	Asn			
	740	745	750															
Leu	Leu	Gln	Asp	Pro	Asp	Phe	Asn	Thr	Ile	Asn	Ser	Thr	Glu	Glu	Asn			
	755	760	765															
Gly	Trp	Lys	Ala	Ser	Asn	Gly	Val	Thr	Ile	Ser	Glu	Gly	Gly	Pro	Phe			
	770	775	780															
Tyr	Lys	Gly	Arg	Ala	Leu	Gln	Leu	Ala	Ser	Ala	Arg	Glu	Asn	Tyr	Pro			
785	790	795	800															
Thr	Tyr	Ile	Tyr	Gln	Lys	Val	Asp	Ala	Ser	Glu	Leu	Lys	Pro	Tyr	Thr			
	805	810	815															
Arg	Tyr	Arg	Leu	Asp	Gly	Phe	Val	Lys	Ser	Ser	Gln	Asp	Leu	Glu	Ile			
	820	825	830															
Asp	Leu	Ile	His	His	His	Lys	Val	His	Leu	Val	Lys	Asn	Val	Pro	Asp			
	835	840	845															
Asn	Leu	Val	Ser	Asp	Thr	Tyr	Pro	Asp	Asp	Ser	Cys	Ser	Gly	Ile	Asn			
	850	855	860															
Arg	Cys	Gln	Glu	Gln	Gln	Met	Val	Asn	Ala	Gln	Leu	Glu	Thr	Glu	His			
865	870	875	880															
His	His	Pro	Met	Asp	Cys	Cys	Glu	Ala	Ala	Gln	Thr	His	Glu	Phe	Ser			
	885	890	895															
Ser	Tyr	Ile	Asp	Thr	Gly	Asp	Leu	Asn	Ser	Ser	Val	Asp	Gln	Gly	Ile			
	900	905	910															
Trp	Ala	Ile	Phe	Lys	Val	Arg	Thr	Thr	Asp	Gly	Tyr	Ala	Thr	Leu	Gly			
	915	920	925															

Asn Leu Glu Leu Val Glu Val Gly Pro Leu Ser Gly Glu Ser Leu Glu
 930 935 940
 Arg Glu Gln Arg Asp Asn Thr Lys Trp Ser Ala Glu Leu Gly Arg Lys
 945 950 955 960
 Arg Ala Glu Thr Asp Arg Val Tyr Gln Asp Ala Lys Gln Ser Ile Asn
 965 970 975
 His Leu Phe Val Asp Tyr Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly
 980 985 990
 Met Ala Asp Ile Met Asp Ala Gln Asn Leu Val Ala Ser Ile Ser Asp
 995 1000 1005
 Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly Ile Asn Tyr Glu
 1010 1015 1020
 Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln Ala Ser Tyr Leu
 1025 1030 1035
 Tyr Thr Ser Arg Asn Ala Val Gln Asn Gly Asp Phe Asn Asn Gly
 1040 1045 1050
 Leu Asp Ser Trp Asn Ala Thr Ala Gly Ala Ser Val Gln Gln Asp
 1055 1060 1065
 Gly Asn Thr His Phe Leu Val Leu Ser His Trp Asp Ala Gln Val
 1070 1075 1080
 Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys Tyr Val Leu Arg
 1085 1090 1095
 Val Thr Ala Glu Lys Val Gly Gly Gly Asp Gly Tyr Val Thr Ile
 1100 1105 1110
 Arg Asp Gly Ala His His Thr Glu Thr Leu Thr Phe Asn Ala Cys
 1115 1120 1125
 Asp Tyr Asp Ile Asn Gly Thr Tyr Val Thr Asp Asn Thr Tyr Leu
 1130 1135 1140
 Thr Lys Glu Val Ile Phe Tyr Ser His Thr Glu His Met Trp Val
 1145 1150 1155
 Glu Val Asn Glu Thr Glu Gly Ala Phe His Ile Asp Ser Ile Glu
 1160 1165 1170
 Phe Val Glu Thr Glu Lys
 1175
 <210> 17
 <211> 1147
 <212> PRT
 <213> 人工序列
 <220>

<223> 突变型BT-0044

<400> 17

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

290	295	300
Arg Thr Val Tyr Thr Asp Pro Ile Gly Phe Val Asn Arg Gly Ser Gly		
305	310	315
Asn Arg Leu Ser Trp Phe Asp Trp Arg Asn Gln Ala Asn Phe Ser Thr		
	325	330
Leu Glu Ser Glu Met Pro Thr Pro Ser Ser Pro Leu Ser Leu Asn His		
	340	345
Met Ser Ile Phe Thr Gly Pro Leu Thr Leu Pro Val Ser Pro Asn Thr		
	355	360
His Arg Ala Arg Val Trp Tyr Gly Asn Gln Asn Met Phe Thr Thr Gly		
	370	380
Ser Gln Asn Ser Gly Gln Thr Thr Asn Ser Ile Gln Asn Ile Ser Gly		
385	390	395
Leu Glu Ile Phe Arg Ile Asp Ser Gln Ala Cys Asn Leu Asn Asn Asn		
	405	410
Ser Tyr Gly Val Asn Arg Ala Glu Phe Phe His Gly Ala Ser Gln Gly		
	420	425
Ser Gln Arg Ser Val Tyr Gln Gly Tyr Ile Arg Gln Ser Gly Leu Asp		
	435	440
Asn Pro Val Val Met Asn Leu Gln Ser Phe Leu Pro Gly Glu Asn Ser		
	450	455
Ala Thr Pro Thr Ala Gln Asp Tyr Thr His Ile Leu Ser Asn Pro Val		
465	470	475
Asn Ile Arg Gly Gly Leu Arg Gln Ile Val Ala Asp Arg Arg Ser Ser		
	485	490
Val Val Val Tyr Gly Trp Thr His Lys Ser Leu Ser Arg Arg Ser Leu		
	500	505
Val Ala Pro Asp Gln Ile Thr Gln Val Pro Ala Val Lys Ala Ser Pro		
	515	520
Ser Ser His Cys Thr Ile Ile Ala Gly Pro Gly Phe Thr Gly Gly Asp		
	530	535
Leu Val Ser Leu Gln Pro Asn Gly Gln Leu Val Ile Pro Phe Gln Val		
545	550	555
Ser Ala Pro Glu Thr Asn Tyr His Ile Arg Ile Cys Tyr Val Ser Thr		
	565	570
Ser Asp Cys Ser Ile Asn Thr Ile Cys Asn Asp Glu Thr His Leu Ser		
	580	585
Thr Leu Pro Ser Thr Thr Ser Ser Leu Glu Asn Leu Gln Cys Asn His		
	595	600
		605

Leu His Tyr Phe Asn Val Gly Thr Phe Lys Pro Thr Ile Asp Ser Lys
 610 615 620
 Leu Thr Leu Val Asn Thr Ser Pro Asn Ala Asn Ile Ile Ile Asp Lys
 625 630 635 640
 Ile Glu Phe Ile Pro Val Asp Thr Ala Gln Gln Gln Asn Glu Asp Leu
 645 650 655
 Glu Ala Ala Lys Lys Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp
 660 665 670
 Gly Leu Gln Val Asn Val Lys Asp Tyr Gln Val Asp Gln Ala Ala Asn
 675 680 685
 Leu Val Ser Cys Leu Ser Asp Glu Gln Tyr Gly Tyr Asp Lys Lys Met
 690 695 700
 Leu Leu Glu Ala Val Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn
 705 710 715 720
 Leu Leu Gln Asp Pro Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn
 725 730 735
 Gly Trp Lys Ala Ser Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe
 740 745 750
 Tyr Lys Gly Arg Ala Ile Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro
 755 760 765
 Thr Tyr Ile Tyr Gln Lys Val Asp Ala Ser Glu Leu Lys Pro Tyr Thr
 770 775 780
 Arg Tyr Arg Leu Asp Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile
 785 790 795 800
 Asp Leu Ile His His His Lys Val His Leu Val Lys Asn Val Pro Asp
 805 810 815
 Asn Leu Val Ser Asp Thr Tyr Pro Asp Asp Ser Cys Ser Gly Ile Asn
 820 825 830
 Arg Cys Gln Glu Gln Gln Met Val Asn Ala Gln Leu Glu Thr Glu His
 835 840 845
 His His Pro Met Asp Cys Cys Glu Ala Ala Gln Thr His Glu Phe Ser
 850 855 860
 Ser Tyr Ile Asp Thr Gly Asp Leu Asn Ser Ser Val Asp Gln Gly Ile
 865 870 875 880
 Trp Ala Ile Phe Lys Val Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly
 885 890 895
 Asn Leu Glu Leu Val Glu Val Gly Pro Leu Ser Gly Glu Ser Leu Glu
 900 905 910
 Arg Glu Gln Arg Asp Asn Thr Lys Trp Ser Ala Glu Leu Gly Arg Lys

915	920	925
Arg Ala Glu Thr Asp Arg	Val Tyr Gln Asp Ala	Lys Gln Ser Ile Asn
930	935	940
His Leu Phe Val Asp Tyr	Gln Asp Gln Gln Leu	Asn Pro Glu Ile Gly
945	950	955
Met Ala Asp Ile Met Asp	Ala Gln Asn Leu Val	Ala Ser Ile Ser Asp
	965	970
Val Tyr Ser Asp Ala Val	Leu Gln Ile Pro Gly	Ile Asn Tyr Glu Ile
	980	985
Tyr Thr Glu Leu Ser Asn	Arg Leu Gln Gln Ala	Ser Tyr Leu Tyr Thr
995	1000	1005
Ser Arg Asn Ala Val Gln	Asn Gly Asp Phe Asn	Asn Gly Leu Asp
1010	1015	1020
Ser Trp Asn Ala Thr Ala	Gly Ala Ser Val Gln	Gln Asp Gly Asn
1025	1030	1035
Thr His Phe Leu Val Leu	Ser His Trp Asp Ala	Gln Val Ser Gln
1040	1045	1050
Gln Phe Arg Val Gln Pro	Asn Cys Lys Tyr Val	Leu Arg Val Thr
1055	1060	1065
Ala Glu Lys Val Gly Gly	Gly Asp Gly Tyr Val	Thr Ile Arg Asp
1070	1075	1080
Asp Ala His His Thr Glu	Thr Leu Thr Phe Asn	Ala Cys Asp Tyr
1085	1090	1095
Asp Ile Asn Gly Thr Tyr	Val Thr Asp Asn Thr	Tyr Ile Thr Lys
1100	1105	1110
Glu Val Val Phe His Pro	Glu Thr Gln His Met	Trp Val Glu Val
1115	1120	1125
Asn Glu Thr Glu Gly Ala	Phe His Leu Asp Ser	Ile Glu Phe Val
1130	1135	1140
Glu Thr Glu Lys		
1145		
<210> 18		
<211> 1157		
<212> PRT		
<213> 人工序列		
<220>		
<223> 突变型BT-0051		
<400> 18		
Met Asn Arg Asn Asn Gln	Asn Glu Tyr Glu Ile	Ile Asp Ala Pro His

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

Leu Cys Arg Arg Trp Gly Thr Asn Pro Tyr Asn Thr Phe Ser Glu Leu
 325 330 335
 Glu Asn Ala Phe Ile Arg Pro Pro His Leu Phe Asp Arg Ile Gln Ser
 340 345 350
 Leu Ser Ile Ser Ser Asn Arg Phe Pro Val Ser Ser Asn Phe Met Asp
 355 360 365
 Tyr Trp Ser Gly His Thr Leu Arg Arg Ser Tyr Leu Asn Asp Ser Ala
 370 375 380
 Val Gln Glu Asp Ser Tyr Gly Leu Ile Thr Thr Thr Arg Ala Thr Ile
 385 390 395 400
 Asn Pro Gly Val Asp Gly Thr Asn Arg Ile Glu Ser Thr Ala Val Asp
 405 410 415
 Phe Arg Ser Ala Leu Ile Gly Ile Tyr Gly Val Asn Arg Ala Ser Phe
 420 425 430
 Val Pro Gly Gly Leu Phe Asn Gly Thr Thr Ser Pro Ala Asn Gly Gly
 435 440 445
 Cys Arg Asp Leu Tyr Asp Thr Asn Asp Glu Leu Pro Pro Asp Glu Ser
 450 455 460
 Thr Gly Ser Ser Thr His Arg Leu Ser His Val Thr Phe Phe Ser Phe
 465 470 475 480
 Gln Thr Asn Gln Ala Gly Ser Ile Ala Asn Ala Gly Ser Val Pro Thr
 485 490 495
 Tyr Val Trp Thr Arg Arg Asp Val Asp Leu Asn Asn Thr Ile Thr Pro
 500 505 510
 Asn Arg Ile Thr Gln Leu Pro Leu Val Lys Ala Ser Ala Pro Val Ser
 515 520 525
 Gly Thr Thr Val Leu Lys Gly Pro Gly Phe Thr Gly Gly Gly Ile Leu
 530 535 540
 Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu Arg Val Thr Val Asn
 545 550 555 560
 Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val Arg Phe Ala Ser Thr
 565 570 575
 Gly Asn Phe Ser Ile Arg Leu Leu Arg Gly Gly Val Ser Ile Gly Asp
 580 585 590
 Val Arg Leu Gly Ser Thr Met Asn Arg Gly Gln Glu Leu Thr Tyr Glu
 595 600 605
 Ser Phe Phe Thr Arg Glu Phe Thr Thr Thr Gly Pro Phe Asn Pro Pro
 610 615 620
 Phe Thr Phe Thr Gln Ala Gln Glu Ile Leu Thr Val Asn Ala Glu Gly

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

Val Tyr Leu Ala Ala Lys Gln Ala Ile Asn His Leu Phe Val Asp Tyr			
945	950	955	960
Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly Leu Ala Glu Ile Asn Glu			
	965	970	975
Ala Ser Asn Leu Val Lys Ser Ile Ser Gly Val Tyr Ser Asp Thr Leu			
	980	985	990
Leu Gln Ile Pro Gly Ile Asn Tyr Glu Ile Tyr Thr Glu Leu Ser Asp			
	995	1000	1005
Arg Leu Gln Gln Ala Ser Tyr Leu Tyr Thr Ser Arg Asn Ala Val			
	1010	1015	1020
Gln Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala Thr			
	1025	1030	1035
Thr Asp Ala Ser Val Gln Gln Asp Gly Ser Thr His Phe Leu Val			
	1040	1045	1050
Leu Ser His Trp Asp Ala Gln Val Ser Gln Gln Met Arg Val Asn			
	1055	1060	1065
Leu Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Lys Lys Val Gly			
	1070	1075	1080
Gly Gly Asp Gly Tyr Val Thr Ile Arg Asp Gly Ala His His Gln			
	1085	1090	1095
Glu Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Val Asn Gly Thr			
	1100	1105	1110
Tyr Val Asn Asp Asn Ser Tyr Ile Thr Lys Glu Val Val Phe Tyr			
	1115	1120	1125
Pro Glu Thr Lys His Met Trp Val Glu Val Ser Glu Ser Glu Gly			
	1130	1135	1140
Ser Phe Tyr Ile Asp Ser Ile Glu Phe Ile Glu Thr Gln Glu			
	1145	1150	1155
<210> 19			
<211> 1173			
<212> PRT			
<213> 人工序列			
<220>			
<223> 突变型BT-0068			
<400> 19			
Met Asn Arg Asn Asn Gln Gly Glu Tyr Glu Ile Ile Asp Ala Ser Thr			
1	5	10	15
Cys Gly Cys Ser Ser Asp Asp Val Val Gln Tyr Pro Leu Ala Arg Asp			
	20	25	30

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

340	345	350
Leu Phe Glu Thr Leu Arg Asn	Leu Thr Ile His Thr Gly Arg Tyr Asn	
355	360	365
Leu Val Gly Gly Ala Arg Phe	Ile Glu Gly Trp Val Gly His Ser Val	
370	375	380
Thr Asn Thr Arg Leu Gly Asn Ser	Thr Val Phe Thr Ser Asn Tyr Gly	
385	390	395
Ser Leu Pro Pro Arg Phe Gln Val	Phe Asn Phe Thr Asn Phe Asp Val	
405	410	415
Tyr Gln Ile Asn Thr Arg Ala Asp	Ser Thr Gly Thr Phe Arg Ile Pro	
420	425	430
Gly Phe Ala Val Thr Arg Ala Gln	Phe Ile Pro Gly Gly Thr Tyr Ser	
435	440	445
Val Ala His Arg Asp Pro Gly Ala	Cys Gln Gln Asp Tyr Asp Ser Ile	
450	455	460
Glu Glu Leu Pro Ser Leu Asp Pro	Asp Glu Pro Ile Asn Arg Ser Tyr	
465	470	475
Ser His Arg Leu Ser His Val Thr	Leu Tyr Lys Tyr Thr Leu Ser Asp	
485	490	495
Thr Asp Tyr Gly Val Ile Asn Tyr	Thr Asp Tyr Gly Ser Met Pro Ala	
500	505	510
Tyr Val Trp Thr His Arg Asp Val	Asp Leu Thr Asn Thr Ile Thr Ala	
515	520	525
Asp Arg Ile Thr Gln Leu Pro Leu	Val Lys Ala Ser Thr Leu Pro Ala	
530	535	540
Gly Thr Thr Val Val Lys Gly Pro	Gly Phe Thr Gly Gly Asp Ile Leu	
545	550	555
Arg Arg Thr Thr Asn Gly Thr Phe	Gly Thr Leu His Val Arg Val Asn	
565	570	575
Ser Pro Leu Thr Gln Gln Tyr Arg	Leu Arg Val Arg Phe Ala Ser Thr	
580	585	590
Gly Asn Phe Ser Ile Arg Val Leu	Arg Gly Gly Thr Ser Ile Gly Asp	
595	600	605
Ala Arg Phe Gly Ser Thr Met Asn	Arg Gly Gln Glu Leu Thr Tyr Glu	
610	615	620
Ser Phe Val Thr Arg Glu Phe Thr	Thr Thr Thr Gly Pro Phe Asn Pro Pro	
625	630	635
Phe Thr Phe Thr Gln Thr Gln Glu	Ile Leu Thr Val Asn Ala Glu Gly	
645	650	655

Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Ser Ile Glu Ile Val Pro
 660 665 670
 Val Asn Pro Thr Arg Glu Ala Glu Glu Asp Leu Glu Ala Ala Lys Lys
 675 680 685
 Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp Gly Leu Gln Val Asn
 690 695 700
 Val Thr Asp Tyr Gln Val Asp Arg Ala Ala Asn Leu Val Leu Cys Leu
 705 710 715 720
 Ser Asp Glu Gln Tyr Ala His Asp Lys Lys Met Leu Leu Glu Ala Val
 725 730 735
 Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn Leu Leu Gln Asp Pro
 740 745 750
 Asp Phe Asn Glu Ile Asn Ser Thr Glu Asp Ser Gly Trp Lys Thr Ser
 755 760 765
 Asn Gly Ile Ile Ile Ser Glu Gly Gly Pro Phe Phe Lys Gly Arg Ala
 770 775 780
 Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro Thr Tyr Ile Tyr Gln
 785 790 795 800
 Lys Val Asp Ser Ser Met Leu Lys Pro Tyr Thr Arg Tyr Lys Leu Asp
 805 810 815
 Gly Phe Val Gln Ser Ser Gln Asp Leu Glu Ile Glu Leu Ile His His
 820 825 830
 His Lys Val His Leu Val Lys Asn Val Pro Asp Asn Leu Val Leu Asp
 835 840 845
 Thr Tyr Pro Asp Gly Ser Cys Asn Gly Ile Asn Arg Cys Glu Glu Gln
 850 855 860
 Gln Met Val Asn Ser Gln Leu Glu Thr Glu His His Pro Met Asp Cys
 865 870 875 880
 Cys Glu Ala Ser Gln Thr His Glu Phe Ser Ser Tyr Ile His Thr Gly
 885 890 895
 Asp Leu Asn Ala Ser Val Asp Gln Gly Ile Trp Val Val Leu Lys Ile
 900 905 910
 Arg Thr Thr Asp Gly Ser Ala Thr Leu Gly Asn Leu Glu Leu Val Glu
 915 920 925
 Val Gly Pro Leu Ser Gly Glu Ser Leu Glu Arg Glu Gln Arg Asp Asn
 930 935 940
 Ala Lys Trp Asn Ala Glu Leu Gly Arg Lys Arg Ala Glu Ala Asp Arg
 945 950 955 960
 Val Tyr Gln Gly Ala Lys Gln Ala Ile Asn His Leu Phe Val Asp Tyr

	965	970	975
Gln Asp Gln Gln Leu Asn Pro Glu Val Gly Leu Ala Glu Ile Ser Glu			
	980	985	990
Ala Arg Asn Leu Ile Glu Ser Ile Ser Asp Val Tyr Cys Asp Ala Val			
	995	1000	1005
Leu Arg Ile Pro Gly Ile Asn Tyr Glu Met Tyr Thr Glu Leu Ser			
	1010	1015	1020
Asn Arg Leu Gln Gln Ala Ala Tyr Leu Tyr Thr Ser Arg Asn Ala			
	1025	1030	1035
Val Gln Asn Gly Asp Phe Asn Ser Gly Leu Asp Ser Trp Asn Ala			
	1040	1045	1050
Thr Thr Asp Ala Thr Val Gln Gln Asp Gly Asn Met Tyr Phe Leu			
	1055	1060	1065
Val Leu Ser His Trp Asp Ala Gln Val Ser Gln Gln Phe Arg Val			
	1070	1075	1080
Gln Pro Asn Cys Lys Tyr Val Leu Arg Val Thr Ala Lys Lys Val			
	1085	1090	1095
Gly Asn Gly Asp Gly Tyr Val Thr Ile Gln Asp Gly Ala His His			
	1100	1105	1110
Arg Glu Thr Leu Thr Phe Asn Ala Cys Asp Tyr Asp Val Asn Gly			
	1115	1120	1125
Thr His Val Asn Asp Asn Ser Tyr Ile Thr Lys Glu Leu Glu Phe			
	1130	1135	1140
Tyr Pro Lys Thr Glu His Met Trp Val Glu Val Ser Glu Thr Glu			
	1145	1150	1155
Gly Thr Phe Tyr Ile Asp Ser Ile Glu Leu Ile Glu Thr Gln Glu			
	1160	1165	1170
<210> 20			
<211> 1179			
<212> PRT			
<213> 人工序列			
<220>			
<223> 突变型BT-0128			
<400> 20			
Met Gly Gly Lys Ser Met Asn Arg Asn Asn Gln Gly Glu Tyr Glu Ile			
1	5	10	15
Ile Asp Ala Ser Thr Cys Gly Cys Ser Ser Asp Asp Val Val Gln Tyr			
	20	25	30
Pro Leu Ala Arg Asp Pro Asn Ala Ala Phe Gln Asn Met Asn Tyr Lys			

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

Ile Arg Pro Pro His Leu Phe Glu Thr Leu Arg Asn Leu Thr Ile His	355	360	365
Thr Gly Arg Tyr Asn Leu Val Gly Gly Ala Arg Phe Ile Glu Gly Trp	370	375	380
Val Gly His Ser Val Thr Asn Thr Arg Leu Gly Asn Ser Thr Val Phe	385	390	395
Thr Ser Asn Tyr Gly Ser Leu Pro Pro Arg Phe Gln Val Phe Asn Phe	405	410	415
Thr Asn Phe Asp Val Tyr Gln Ile Asn Thr Arg Ala Asp Ser Thr Gly	420	425	430
Thr Phe Arg Ile Pro Gly Phe Ala Val Thr Arg Ala Gln Phe Ile Pro	435	440	445
Gly Gly Thr Tyr Ser Val Ala His Arg Asp Pro Gly Ala Cys Gln Gln	450	455	460
Asp Tyr Asp Ser Ile Glu Glu Leu Pro Ser Leu Asp Pro Asp Glu Pro	465	470	475
Ile Asn Arg Ser Tyr Ser His Arg Leu Ser His Val Thr Leu Tyr Lys	485	490	495
Tyr Thr Leu Ser Asp Thr Asp Tyr Gly Val Ile Asn Tyr Thr Asp Tyr	500	505	510
Gly Ser Met Pro Ala Tyr Val Trp Thr His Arg Asp Val Asp Leu Thr	515	520	525
Asn Thr Ile Thr Ala Asp Arg Ile Thr Gln Leu Pro Leu Val Lys Ala	530	535	540
Ser Thr Leu Pro Ala Gly Thr Thr Val Val Lys Gly Pro Gly Phe Thr	545	550	555
Gly Gly Asp Ile Leu Arg Arg Thr Thr Asn Gly Thr Phe Gly Thr Leu	565	570	575
His Val Arg Val Asn Ser Pro Leu Thr Gln Gln Tyr Arg Leu Arg Val	580	585	590
Arg Phe Ala Ser Thr Gly Asn Phe Ser Ile Arg Val Leu Arg Gly Gly	595	600	605
Thr Ser Ile Gly Asp Ala Arg Phe Gly Ser Thr Met Asn Arg Gly Gln	610	615	620
Glu Leu Thr Tyr Glu Ser Phe Val Thr Arg Glu Phe Thr Thr Thr Gly	625	630	635
Pro Phe Asn Pro Pro Phe Thr Phe Thr Gln Thr Gln Glu Ile Leu Thr	645	650	655
Val Asn Ala Glu Gly Val Ser Thr Gly Gly Glu Tyr Tyr Ile Asp Ser			

660	665	670
Ile Glu Ile Val Pro Val Asn Pro Thr Arg Glu Ala Glu Glu Asp Leu		
675	680	685
Glu Ala Ala Lys Lys Ala Val Ala Ser Leu Phe Thr Arg Thr Arg Asp		
690	695	700
Gly Leu Gln Val Asn Val Thr Asp Tyr Gln Val Asp Gln Ala Ala Asn		
705	710	715
Leu Val Ser Cys Leu Ser Asp Glu Gln Tyr Gly Tyr Asp Lys Lys Met		
725	730	735
Leu Leu Glu Ala Val Arg Ala Ala Lys Arg Leu Ser Arg Glu Arg Asn		
740	745	750
Leu Leu Gln Asp Pro Asp Phe Asn Thr Ile Asn Ser Thr Glu Glu Asn		
755	760	765
Gly Trp Lys Ala Ser Asn Gly Val Thr Ile Ser Glu Gly Gly Pro Phe		
770	775	780
Tyr Lys Gly Arg Ala Leu Gln Leu Ala Ser Ala Arg Glu Asn Tyr Pro		
785	790	795
Thr Tyr Ile Tyr Gln Lys Val Asp Ala Ser Glu Leu Lys Pro Tyr Thr		
805	810	815
Arg Tyr Arg Leu Asp Gly Phe Val Lys Ser Ser Gln Asp Leu Glu Ile		
820	825	830
Asp Leu Ile His His His Lys Val His Leu Val Lys Asn Val Pro Asp		
835	840	845
Asn Leu Val Ser Asp Thr Tyr Pro Asp Asp Ser Cys Ser Gly Ile Asn		
850	855	860
Arg Cys Gln Glu Gln Gln Met Val Asn Ala Gln Leu Glu Thr Glu His		
865	870	875
His His Pro Met Asp Cys Cys Glu Ala Ala Gln Thr His Glu Phe Ser		
885	890	895
Ser Tyr Ile Asp Thr Gly Asp Leu Asn Ser Ser Val Asp Gln Gly Ile		
900	905	910
Trp Ala Ile Phe Lys Val Arg Thr Thr Asp Gly Tyr Ala Thr Leu Gly		
915	920	925
Asn Leu Glu Leu Val Glu Val Gly Pro Leu Ser Gly Glu Ser Leu Glu		
930	935	940
Arg Glu Gln Arg Asp Asn Thr Lys Trp Ser Ala Glu Leu Gly Arg Lys		
945	950	955
Arg Ala Glu Thr Asp Arg Val Tyr Gln Asp Ala Lys Gln Ser Ile Asn		
965	970	975

His Leu Phe Val Asp Tyr Gln Asp Gln Gln Leu Asn Pro Glu Ile Gly		
980	985	990
Met Ala Asp Ile Met Asp Ala Gln Asn Leu Val Ala Ser Ile Ser Asp		
995	1000	1005
Val Tyr Ser Asp Ala Val Leu Gln Ile Pro Gly Ile Asn Tyr Glu		
1010	1015	1020
Ile Tyr Thr Glu Leu Ser Asn Arg Leu Gln Gln Ala Ser Tyr Leu		
1025	1030	1035
Tyr Thr Ser Arg Asn Ala Val Gln Asn Gly Asp Phe Asn Asn Gly		
1040	1045	1050
Leu Asp Ser Trp Asn Ala Thr Ala Gly Ala Ser Val Gln Gln Asp		
1055	1060	1065
Gly Asn Thr His Phe Leu Val Leu Ser His Trp Asp Ala Gln Val		
1070	1075	1080
Ser Gln Gln Phe Arg Val Gln Pro Asn Cys Lys Tyr Val Leu Arg		
1085	1090	1095
Val Thr Ala Glu Lys Val Gly Gly Gly Asp Gly Tyr Val Thr Ile		
1100	1105	1110
Arg Asp Gly Ala His His Thr Glu Thr Leu Thr Phe Asn Ala Cys		
1115	1120	1125
Asp Tyr Asp Ile Asn Gly Thr Tyr Val Thr Asp Asn Thr Tyr Leu		
1130	1135	1140
Thr Lys Glu Val Ile Phe Tyr Ser His Thr Glu His Met Trp Val		
1145	1150	1155
Glu Val Asn Glu Thr Glu Gly Ala Phe His Leu Asp Ser Leu Glu		
1160	1165	1170
Phe Val Glu Thr Glu Lys		
1175		
<210> 21		
<211> 26		
<212> DNA		
<213> 人工序列		
<220>		
<223> 0AR2613a正向引物		
<400> 21		
aaacatgaac cgaaataatc aaaatg 26		
<210> 22		
<211> 22		
<212> DNA		

<213> 人工序列
<220>
<223> 0AR2615a反向引物
<400> 22
atccgtccct tgtgcgtgta aa 22
<210> 23
<211> 30
<212> DNA
<213> 人工序列
<220>
<223> 0AR2611a-F正向引物
<400> 23
gtttaaacad gaatcgaaat aatcaaaatg 30
<210> 24
<211> 29
<212> DNA
<213> 人工序列
<220>
<223> 0AR2612a-R反向引物
<400> 24
ggcgcgccct actcttgtgt ttcaataaa 29
<210> 25
<211> 29
<212> DNA
<213> 人工序列
<220>
<223> 0AR2768-F正向引物
<400> 25
gtttaaacad gaatcaaaat aaacacgga 29
<210> 26
<211> 31
<212> DNA
<213> 人工序列
<220>
<223> 0AR2769-R反向引物
<400> 26
ggcgcgcctt actgttgggt ttccatgaac t 31