(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 105017391 A
(43)申请公布日 2015.11.04

(21)申请号 201510388438.6
(22)申请日 2015.06.30
(71)申请人 杭州瑞丰生物科技有限公司
地址 311121 浙江省杭州市余杭区文一西路1500号
(72)发明人 沈志成 张先文 林朝阳 王东芳
(74)专利代理机构 杭州天正专利事务所有限公司 33201
代理人 黄美娟 李世玉
(51)Int.Cl.
C07K 14/325(2006.01)
C12N 15/32(2006.01)
C12N 5/10(2006.01)
C07K 19/00(2006.01)
C12N 15/62(2006.01)
A01N 47/44(2006.01)
A01P 7/04(2006.01)

权利要求书1页 说明书7页
序列表10页

(54)发明名称

抗虫蛋白、抗虫融合蛋白、编码基因、载体及应用

(57)摘要

本发明公开了一种抗虫蛋白、抗虫融合蛋白、编码基因、载体及应用，抗虫蛋白的氨基酸序列与SEQ ID NO:1相比具有90%的同源性；所述抗虫融合蛋白从N端到C端依次含有BT晶体毒素Cry1蛋白和抗虫蛋白Cry2Aj；本发明首次公开了一种具有高杀虫能力的BT晶体毒素Cry2Aj蛋白，并且设计了用一个Cry1Ab晶体毒素与该Cry2Aj晶体毒素融合成的人工蛋白质分子，与原先的Cry1和Cry2晶体毒素物理混合相比，具有如下优点：杀虫活性提高了10倍以上，杀虫谱更广；本发明的杀虫蛋白质可以杀灭棉铃虫、甜菜夜蛾、玉米螟、二化螟等水稻、玉米和棉花的主要鳞翅目害虫；此外，本发明发现的融合蛋白质还可以有效的减缓害虫抗性的发生。

CN 105017391 A

1.一种抗虫蛋白,其特征在于:所述抗虫蛋白的氨基酸序列与 SEQ ID NO:1 相比具有 90%的同源性。

2.如权利要求 1 所述抗虫蛋白,其特征在于:所述抗虫蛋白的氨基酸序列为 SEQ ID NO:1 所示。

3.一种编码权利要求 1 或 2 所述抗虫蛋白的基因,其特征在于所述基因的核苷酸序列为 SEQ ID NO:2 所示。

4.一种由权利要求 2 所述抗虫蛋白编码基因构建的重组载体。

5.一种权利要求 1 所述抗虫蛋白在制备抗虫植物细胞中的应用。

6.一种含权利要求 1 所述抗虫蛋白的抗虫融合蛋白,其特征在于所述抗虫融合蛋白从 N 端到 C 端依次含有 BT 晶体毒素 Cry1 蛋白和抗虫蛋白。

7.如权利要求 6 所述含抗虫蛋白的抗虫融合蛋白,其特征在于所述抗虫融合蛋白的氨基酸序列为 SEQ ID NO:3 所示。

8.一种权利要求 6 所述的抗虫融合蛋白的编码基因,其特征在于:所述编码基因的核苷酸序列为 SEQ ID NO:4 所示。

9.一种由权利要求 8 所述抗虫融合蛋白编码基因构建的重组载体。

10.一种权利要求 6 或 7 所述抗虫融合蛋白在制备抗虫植物细胞中的应用。

抗虫蛋白、抗虫融合蛋白、编码基因、载体及应用

（一）技术领域

[0001] 本发明涉及一种具有高杀虫能力的抗虫基因和抗虫蛋白质、一种具有高杀虫能力的抗虫融合基因和融合蛋白质、以及用该抗虫蛋白或抗虫融合蛋白构建抗虫农作物的具体应用。

（二）背景技术

[0002] 害虫给全球农业生产带来每年约 80 亿美元的损失。水稻害虫种类复杂，自播种、秧苗期至成熟期，由种、苗、根、茎、叶、花穗、谷粒均可受不同害虫危害。国内已知的水稻害虫已达 385 余种，其中最主要的有 40 余种，如三化螟、二化螟和大螟等分布广泛，危害严重。目前，害虫的防治主要是依靠化学农药，化学农药的使用对于农业持续稳定的发展发挥了重要作用。但是，在经过半个多世纪的使用之后，化学农药已经暴露出了严重的问题。例如化学农药在自然界和农作物中的残留以及在食物链中的富积已经造成了严重的环境污染，对人畜安全直接构成威胁；由于化学农药非特异性的作用方式，杀死靶标害虫的同时也对毒害了许多益虫和捕虫鸟类，破坏了生态平衡；并且由于长期使用化学农药，特别是不合理的滥用，许多害虫已逐步对其产生了抗性（林良斌等，1997），常规剂量已不能有效地防治害虫，然而随着用药剂量的加大，害虫的抗性也随之增强，由此形成恶性循环，导致防治成本显著增加，并且带来了更为严重的环境污染、生态失调和人类健康的危害。

[0003] 近年来，随着分子生物学技术特别是植物基因工程的迅速发展，为防治农作物害虫提供了新的途径。抗虫是目前农作物转基因改良领域中应用最广泛的技术之一，国际上的主要农业生物技术公司和种子公司都在从事该方面的研发，并且已经取得了革命性的进展。利用基因工程技术，将抗虫基因导入农作物，使其在农作物体内表达抗虫蛋白并能稳定遗传，形成抗虫转基因农作物新品系。种植转基因抗虫农作物大幅度降低了化学农药的使用量，减少了害虫造成的损失，为农业生产创造了巨大的价值。目前应用于农作物抗虫性状改良的外源基因主要有苏云金芽孢杆菌（Bacillus thuringiensis，Bt）杀虫基因、植物凝集素基因和蛋白酶抑制剂基因等。其中 Bt 杀虫基因是目前世界上应用最为广泛的抗虫基因，Bt 蛋白对鳞翅目、双翅目和鞘翅目昆虫都具有很好的杀虫活性。

[0004] 获得转基因抗虫农作物的关键技术是获得性能优良的杀虫蛋白质。杀虫蛋白质有多种，比较常用的是苏云金芽孢杆菌（简称 Bt）晶体蛋白质，例如 Cry1Ab、Cry1Ac、Cry2Ab 等，它们已被大量运用于转基因抗虫农作物（Crickmore N,Zeigler D,Feitelson J,Schnepf E,Van Rie J,Lereclus D,Baum J,Dean D.(1998)Microbiol.Mol.Biol.Rev.62,807-813）。但是，单个杀虫蛋白质往往杀虫谱比较窄，杀虫活性比较低。同时，长期大量使用单一的杀虫蛋白质还可能引起害虫抗性的发展（Ferre J,Van Rie J.(2002)Annu.Rev.Entomol.47,501-533；Gassmann A,Carrière Y,Tabashnik B.(2009)Annu.Rev.Entomol.54,147-163）。因此，获得具有高杀虫活性的新型蛋白质杀虫毒素，对提高杀虫能力和减缓害虫抗性的发生具有重要的应用价值（Chen,M.,Shelton,A.,Ye,G.(2011)Annu.Rev.Entomol.56,81-101）。

[0005] 相比单个基因的转化，多个基因的聚合使用在转基因作物中具有明显的优势。所谓基因聚合就是指在同种作物品种中转入两个不同的抗虫基因。这两个基因可以是两种不同的Bt基因，也可以是一种Bt基因和一种非Bt抗虫基因，例如Cheng等在转基因水稻中表达Bt融合蛋白基因Cry1AbCry1Ac，对于二化螟防治效果良好(Cheng XY,Sardana R,Kaplan H,Altosaar I.(1998)Proc Natl Acad Sci USA.95,2767-2772)；Bohorova等将融合基因cry1Bcry1Ab转入玉米中来防治多种农业害虫(Bohorova N,Frutos R,Royer M, Estañol P,Pacheco M,Rascon Q,McLean S,Hoisington D.(2011)Thoer Appl Genet.103,817-826)。使用二价或多价抗虫基因相对于单个基因有以下几个方面的优势：杀虫效果更好，经济实用；杀虫谱更广，多个抗虫基因之间杀虫谱可以互相弥补，不同的基因对同种害虫的杀虫效果也存在差异，可以相互补充；可以有效延缓抗性品种的使用寿命，理论上昆虫同时对两种不同的Bt基因同时产生抗性的几率大大低于对单个Bt基因产生抗性的几率，因而能有效延缓昆虫产生抗性的时间。因此，同时使用无交互抗性的Bt毒素对于治理昆虫抗性是一种经济而有效的措施，必须加快发展利用抗虫融合基因创建具有更好抗虫活性的转基因农作物。

[0006] 然而，本领域的技术人员都知道，并不是所有的两个杀虫基因的融合都能够增强杀虫能力的。例如Cry1Ab和Cry1Ca的融合蛋白质的杀虫能力比相同重量的单独Cry1Ab和Cry1Ca的混合物的杀虫能力还要低。另外，即使两个单独的抗虫蛋白质组合使用具有增效作用，但也并不能得出它们的融合蛋白具有高效杀虫能力；这是因为：在蛋白质生物化学上，两个单独蛋白质的物理混和与融合是不同的，不能根据物理混和所具有的增效作用来肯定融合后所获得的增效作用；而且在不同位置融合所得的融合蛋白的活性是大相径庭的。

[0007] 本发明首次公开了一种具有高杀虫能力的BT晶体毒素Cry2蛋白质，并且设计了用一个Cry1晶体毒素与该Cry2晶体毒素融合成的人工蛋白质分子，与原先的Cry1和Cry2晶体毒素物理混合相比，具有如下优点：杀虫活性提高了10倍以上，杀虫谱更广。本发明的杀虫蛋白质可以杀灭棉铃虫、甜菜夜蛾、玉米螟、二化螟等水稻、玉米和棉花的主要鳞翅目害虫。此外，本发明发现的融合蛋白质还可以有效的减缓害虫抗性的发生。

（三）发明内容

[0008] 本发明目的是提供一种具有高杀虫能力的抗虫蛋白及抗虫融合蛋白，以及在抗虫药物及抗虫农作物中的应用。

[0009] 本发明采用的技术方案是：

[0010] 本发明提供一种抗虫蛋白，所述抗虫蛋白(Cry2Aj)的氨基酸序列与SEQ ID NO:1相比具有90％的同源性，更优选所述氨基酸序列为SEQ ID NO:1所示。

[0011] 本发明还涉及所述抗虫蛋白编码基因，所述编码基因的核苷酸序列为SEQ ID NO:2所示。

[0012] 本发明涉及一种由所述抗虫蛋白编码基因构建的重组载体，以及由重组载体转化制备的重组工程菌。

[0013] 本发明提供一种所述抗虫蛋白在制备抗虫药物中的应用，所述虫为水稻、玉米和棉花的主要鳞翅目害虫，如棉铃虫、甜菜夜蛾、玉米螟、二化螟等。

[0014] 本发明提供一种所述抗虫蛋白在制备抗虫植物细胞中的应用，所述植物为水稻、玉米、棉花、小麦、高粱、大豆、油菜、萝卜或甘蓝等农作物。

[0015] 此外，本发明还提供一种含所述抗虫蛋白的抗虫融合蛋白，所述抗虫融合蛋白从N端到C端依次含有BT晶体毒素Cry1蛋白和抗虫蛋白(Cry2Aj)，优选所述抗虫融合蛋白的氨基酸序列为SEQ ID NO:3所示，更优选所述融合蛋白是由SEQ ID NO:1所示蛋白Cry2Aj与SEQ ID NO:5所示蛋白Cry1Ab连接构成，其中蛋白Cry1Ab编码基因的核苷酸序列为SEQ ID NO:6所示，该融合蛋白从N端至C端依次为Cry1Ab晶体毒素和抗虫蛋白Cry2Aj。

[0016] 本发明涉及一种所述的抗虫融合蛋白的编码基因，所述编码基因与SEQ ID NO:4所示核苷酸序列相比具有92%以上的同源性，更优选所述编码基因的核苷酸序列为SEQ ID NO:4所示。抗虫融合基因包括从5'-3'依次含有编码BT晶体毒素Cry1的核苷酸序列和编码抗虫蛋白Cry2Aj的核苷酸序列，且这两个核苷酸序列位于同一个开放阅读框内，其中BT晶体毒素Cry1为Cry1Ab、Cry1Ac或Cry1Aa及其修饰改良基因。

[0017] 本发明提供一种由所述抗虫融合蛋白编码基因构建的重组载体，以及由重组载体转化制备的重组工程菌。

[0018] 本发明涉及一种所述抗虫融合蛋白在制备抗虫药物中的应用，所述虫为水稻、玉米和棉花的主要鳞翅目害虫，如棉铃虫、甜菜夜蛾、玉米螟、二化螟等。

[0019] 本发明涉及一种所述抗虫融合蛋白在制备抗虫植物细胞中的应用，所述植物为水稻、玉米、棉花、小麦、高粱、大豆、油菜、萝卜或甘蓝等农作物。

[0020] 本发明所述抗虫蛋白、抗虫融合蛋白在制备抗虫植物细胞中的应用，具体为：将含抗虫蛋白编码基因或抗虫融合蛋白编码基因的载体转入农杆菌中，然后用农杆菌侵染植物细胞或植物组织，从而获得具有抗虫性能的植物细胞或植物组织。

[0021] 与现有技术相比，本发明的有益效果主要体现在：本发明首次公开了一种具有高杀虫能力的BT晶体毒素Cry2Aj蛋白，并且设计了用一个Cry1Ab晶体毒素与该Cry2Aj晶体毒素融合成的人工蛋白质分子，与原先的Cry1和Cry2晶体毒素物理混合相比，具有如下优点：杀虫活性提高了10倍以上，杀虫谱更广；本发明的杀虫蛋白质可以杀灭棉铃虫、甜菜夜蛾、玉米螟、二化螟等水稻、玉米和棉花的主要鳞翅目害虫。此外，本发明发现的融合蛋白质还可以有效的减缓害虫抗性的发生。

（四）具体实施方式

[0022] 下面结合具体实施例对本发明进行进一步描述，但本发明的保护范围并不仅限于此：

[0023] 本发明实施例所使用的分子生物学和生物化学方法均为已知的技术。在Ausubel编写的John Wiley and Sons公司出版的Current Protocols in Molecular Biology和J.Sambrook等编写Cold Spring Harbor Laboratory Press(2001)出版的Molecular Cloning:ALabortory Manual,3rd ED.等文献均有详细的说明。

[0024] 实施例1、Cry2Aj基因的获得及表达载体构建

[0025] 编码Cry2Aj杀虫蛋白的基因委托上海生工合成，其DNA序列为SEQ ID NO:2。Cry2Aj基因被克隆到表达载体pET28a(Novagen,Cambridge,UK)中限制性内切酶BamHI和SacI的位点之间，得到的载体命名为pET-2Aj，Cry2Aj基因编码的氨基酸序列为SEQ ID

NO:1 所示的蛋白质。

[0026] 说明:SEQ ID NO:2 核酸序列的表达框从第一个起始密码子(ATG)开始,到最后终止密码子(TAA)前结束,长度为 1902bp。第一个 ATG 前的核酸序列 GGATCC 为 BamHI 酶切位点,其后的 ACC 为核糖体结合位点,最后 GAGCTC 为 SacI 酶切位点。

[0027] 实施例 2、Cry1Ab-Cry2Aj 融合基因的获得及表达载体构建

[0028] 编码 Cry1Ab-Cry2Aj 的杀虫蛋白的融合基因由上海生工合成,在 Cry1Ab 和 Cry2Aj 之间是一个连接肽,其氨基酸序列是 PGKGGG。Cry1Ab-Cry2Aj 融合基因的 DNA 序列为 SEQ ID NO:4。Cry1Ab-Cry2Aj 融合基因被克隆到表达载体 pET28a 中限制性内切酶 BamHI 和 SacI 的位点之间,得到的载体命名为 pET-1Ab-2Aj,Cry1Ab-Cry2Aj 融合基因编码一个氨基酸序列为 SEQ ID No:3 的融合蛋白质。

[0029] 因实验需要,委托上海生工合成 Cry1Ab 基因,其 DNA 序列为 SEQ ID NO:6。Cry1Ab 基因被克隆到表达载体 pET28a 中限制性内切酶 BamHI 和 SacI 的位点之间,得到的载体命名为 pET-1Ab,Cry1Ab 基因编码一个氨基酸序列为 SEQ ID NO:5 的蛋白质。

[0030] 说明:SEQ ID NO:4 核酸序列的表达框从第一个起始密码子(ATG)开始,到最后终止密码子(TAA)前结束,长度为 3864bp。第一个 ATG 前的核酸序列 GGATCC 为 BamHI 酶切位点,其后的 ACC 为核糖体结合位点,最后 GAGCTC 为 SacI 酶切位点。

[0031] SEQ ID NO:6 核酸序列的表达框从第一个起始密码子(ATG)开始,到最后终止密码子(TAA)前结束,长度为 1950bp。第一个 ATG 前的核酸序列 GGATCC 为 BamHI 酶切位点,其后的 ACC 为核糖体结合位点,最后 GAGCTC 为 SacI 酶切位点。

[0032] 实施例 3、杀虫蛋白质的制备

[0033] 包含杀虫基因的载体 pET-1Ab、pET-2Aj 和 pET-1Ab-2Aj 分别导入 BL21Star(Escherichia coli)细胞系,在包含 50mg/L 卡那霉素的 LB 固体培养基上,37℃培养选取单克隆。单个菌落接种到 100 毫升 LB 细菌培养液,在 37℃下震动培养至 OD600 = 0.6,然后加 IPTG(Isopropyl-β-D-thiogalactoside)至浓度为 0.5mM,并继续在同样的条件下培养 4 小时。培养液经过 5000g 离心 10 分钟沉淀大肠杆菌细胞,然后弃上清收集沉淀。沉淀中加 30 毫升 pH7.0、20mM Tris-HCl 缓冲液,超声破碎,分别获得杀虫蛋白 Cry1Ab、杀虫蛋白 Cry 2Aj 和杀虫蛋白 Cry 1Ab-Cry 2Aj。这样获得

[0034] 的重组蛋白质用来进行杀虫活性的测定。

[0035] (1)LB 固体培养基组成为:

[0036]

胰化蛋白胨	10g
酵母提取物	5g
NaCl	5g
琼脂	15 g
去离子水	定溶于 1L
	pH 7.0 高压灭菌

[0037] LB 细菌培养液组成为:

[0038]

胰化蛋白胨	10g
酵母提取物	5g
NaCl	5g
去离子水	定溶于 1L
pH 7.0 高压灭菌	

[0039] 实施例 4、在大肠杆菌中表达的杀虫蛋白的杀虫活性测定

[0040] 实施例 3 所获得的杀虫蛋白 Cry1Ab、杀虫蛋白 Cry 2Aj 和杀虫蛋白 Cry1Ab-Cry2Aj 各 100 微升单独或者混合涂在 0.5 平方厘米昆虫人工饲料的表面，饲养新生一龄幼虫用来进行杀虫活性测定。以含有 pET28a 空载体的 BL21Star 为阴性对照，准备方法与实施例 3 类同。饲养 7 天以后统计杀虫率，结果如表 1 所示。

[0041] 表 1、Cry1Ab、Cry2Aj 和 Cry1Ab-Cry2Aj 的杀虫率

[0042]

杀虫蛋白质	棉铃虫	玉米螟
Cry1Ab （100 微升）	0%	100%
Cry2Aj（100 微升）	100%	80%
Cry1Ab+Cry2Aj（各 100 微升）	50%	80%
Cry1Ab-Cry2Aj 融合蛋白（100 微升）	100%	100%
阴性对照（100 微升）	0%	0%

[0043] 所述昆虫人工饲料制备方法为：(1) 将玉米面 1200g、黄豆面 600g、酵母粉 540g 放入牛皮纸袋，再放入烘箱中 120℃ 2 小时。(2) 将步骤 (1) 烘好的原料与 300g 白糖放入一小锅内，在大锅内烧水 6000ml，温水时向小锅内放入大锅一半的温水，用勺子搅拌均匀，同时将 100g 琼脂粉放入大锅内，放入 11g 山梨酸和 11g 对羟基苯甲酸甲酯，熬到开锅，搅拌，加入小锅内的东西，熬，加入 KOH(22.4g+100ml 水)，再加 10％甲醛 90ml(40％甲醛 30ml+90ml 水)，再加乙酸 240ml(180ml 冰乙酸 +320ml 水，然后取 40ml 上述乙酸溶液再 +200ml 水) 放入锅内，搅拌均匀，加入 37.5g 抗坏血酸和 90ml 维生素，搅拌均匀，分装到 4 个白瓷盘中，置于 4 度冰箱备用。

[0044] 结果表明，Cry2Aj 对棉铃虫有 100％的杀虫活性，对玉米螟有 80％的杀虫活性；融合蛋白质 Cry1Ab-Cry2Aj 的杀虫活性比单独的 Cry2Aj 或者 Cry1Ab 明显提高，同时也比 Cry1Ab 和 Cry2Aj 混合的杀虫活性明显提高。

[0045] 实施例 5、农杆菌转化 T-DNA 载体的构建

[0046] 农杆菌转化 T-DNA 载体是基于 pCambia 1300(CAMBIA, Canberra, Australia) 载体而构建的。编码杀虫蛋白质的合成核苷酸序列 Cry2Aj(SEQ ID No :2，5’端被设计上 BamHI 位点，3’端被设计上 SacI 位点，BamHI-SacI 酶切片段) 和 Cry1Ab-Cry2Aj(SEQ ID No :4，5’端被设计上 BamHI 位点，3’ 端被设计上 SacI 位点，BamHI-SacI 酶切片段)，分别与玉米的

pepc 终止子（SEQ ID No:8,5'端被设计上 SacI 位点,3'端被设计上 KpnI 位点,SacI-KpnI 酶切片段）连接,获得包括基因和终止子的 BamHI-KpnI 片段。

[0047] 玉米 ubiquitin-1 启动子从玉米的基因组中通过 PCR 获得,使用的引物分别是：

[0048] ZmUbiF(5'GCGAAGCTTGCATGCCTACAGTGCAGCGTGACCCGGTCGTGC,添加了 HindIII 位点,SEQ ID No:9)。

[0049] ZmUbiR(5'GTGGGATCCTCTAGAGTCGACCTGCAGAAGTAACACCAAACAA CAG,添加了 BamHI 位点,SEQ ID No:10)。

[0050] 玉米ubiquitin-1启动子(SEQ ID No:7)经过HindIII和BamHI酶切后和基因-终止子片段（BamHI-KpnI 片段）共同连接到经过 HindIII 和 KpnI 酶切的 pCambia1300 载体中,获得 T-DNA 载体 pCAM-Cry2Aj 和 pCAM-Cry1Ab-Cry2Aj。由于玉米 ubiquitin-1 启动子在禾本科植物中均有启动子活性,pCAM-Cry2Aj 和 pCAM-Cry1Ab-Cry2Aj 可以作为多种禾本科植物的抗虫转化 T-DNA 载体。

[0051] 实施例 6、转 Cry2Aj 基因及 Cry1Ab-Cry2Aj 融合基因水稻的获得

[0052] 转基因植物的获得方法是采用现有技术（卢雄斌龚祖埙,1998 生命科学 10：125-131;刘凡等,2003 分子植物育种 1:108-115)。选取成熟饱满种子去壳,诱导产生愈伤组织作为转化材料。分别取含 T-DNA 载体 pCAM-Cry2Aj 和 pCAM-Cry1Ab-Cry2Aj 的农杆菌划板,挑单菌落接种,准备转化用农杆菌。将待转化的愈伤组织放入适当浓度的农杆菌液中（含乙酰丁香酮),让农杆菌结合到愈伤组织表面,然后把愈伤组织转移到共培养基中,共培养 2-3 天。用无菌水冲洗转化后的愈伤,转移到含抗生素的筛选培养基上,筛选培养(50ng/mL 潮霉素）两个月（中间继代一次）。把筛选后,生长活力良好的愈伤转移到预分化培养基上培养 20 天左右,然后将预分化好的愈伤组织转移到分化培养基,14 小时光照分化发芽。2-3 周后,把抗性再生植株转移到生根培养基上壮苗生根,最后将再生植株洗去琼脂移植于温室,分别获得含基因 Cry2Aj 和基因 Cry1Ab-Cry2Aj 的植株,作为鉴定材料。

[0053] 实施例 7、转 Cry2Aj 基因及 Cry1Ab-Cry2Aj 融合基因玉米的获得

[0054] 玉米转化方法基本按照 Ishida 的方法 (Yuji Ishida,Hideaki Saito,Shozo Ohta,Yukoh Hiei,Toshihiko Komari&Takashi Kumashiro,1996,High efficiency transformation of maize(Zea mays L.)mediated by Agrobacterium tumefaciens. NatureBiotechnlogy,14:745-750)。取授粉后 8-10 天的 Hi-II 玉米穗。收集所有的未成熟胚（大小为 1.0-1.5mm)。分别将含有 T-DNA 载体 pCAM-Cry2Aj 和 pCAM-Cry1Ab-Cry2Aj 的农杆菌与未成熟胚共培养 2-3 天 (22℃）。转移未成熟胚到愈伤诱导培养基上（含 200mg/L 的 Timentin 杀农杆菌),28℃暗培养 10-14 天。将所有的愈伤转到带有 50ng/mL 潮霉素的筛选培养基上,28℃暗培养 2-3 周。

[0055] 转移所有的组织到新鲜潮霉素的筛选培养基上,28℃暗培养 2-3 周。然后,转移所有筛选后成活的胚性组织到再生培养基上,28℃暗培养 10-14 天,每皿一个株系。转移胚性组织到新鲜的再生培养基上,26℃光照培养 10-14 天。转移所有发育完全的植株到生根培养基上,26℃光照培养直到根发育完全,分别获得含基因 Cry2Aj 和基因 Cry1Ab-Cry2Aj 的植株。

[0056] 实施例 8、转基因农作物抗虫能力的测定

[0057] 利用二化螟和稻纵卷叶螟分别测定了通过实施例 6 获得的 10 个转基因水稻的杀

虫活性。新生二化螟的一龄幼虫取食转 Cry2Aj 基因水稻后，有 4 个株系的致死率是 100％。稻纵卷叶螟的一龄幼虫取食转 Cry2Aj 基因水稻后，有 6 个株系的致死率是 100％；新生二化螟的一龄幼虫取食转 Cry1Ab-Cry2Aj 融合基因水稻后，有 8 个株系的致死率是 100％。稻纵卷叶螟的一龄幼虫取食转 Cry1Ab-Cry2Aj 融合基因水稻后，有 9 个株系的致死率是 100％。

[0058] 利用玉米螟和棉铃虫分别测定了通过实施例 7 获得的 10 个转基因玉米株系的杀虫活性。新生玉米螟一龄幼虫取食 10 个转 Cry2Aj 基因玉米株系后，有 8 个株系 100％死亡，其他 2 个株系对玉米螟杀灭活性在 20％到 50％之间。新生棉铃虫一龄幼虫取食 10 个转 Cry2Aj 基因玉米株系后，10 个株系全部 100％死亡；

[0059] 新生玉米螟一龄幼虫取食 10 个转 Cry1Ab-Cry2Aj 融合基因玉米株系后，10 个株系全部 100％死亡。新生棉铃虫一龄幼虫取食 10 个转 Cry1Ab-Cry2Aj 融合基因玉米株系后，10 个株系全部 100％死亡。

[0060] 最后，还需要注意的是，以上列举的仅是本发明的若干个具体实施例。显然，本发明不限于以上实施例，还可以有许多变形。本领域的普通技术人员能从本发明公开的内容直接导出或联想到的所有变形，均应认为是本发明的保护范围。

[0001]

SEQUENCE LISTING

<110> 杭州瑞丰生物科技有限公司

<120> 抗虫蛋白、抗虫融合蛋白、编码基因、载体及应用
<130>
<160> 10
<170> PatentIn version 3.5
<210> 1
<211> 632
<212> PRT
<213> unknown
<220>
<223> 人工序列
<400> 1
Ala Ser Val Leu Asn Ser Gly Arg Thr Thr Ile Cys Asp Gly His Asn
1 5 10 15
Val Met Ala Gln Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp Thr
20 25 30
Ile Gln Lys Glu Trp Thr Glu Trp Lys Arg Asp Asp His Ser Leu Tyr
35 40 45
Val Ala Pro Val Val Gly Thr Val Ala Ile Phe Leu Leu Lys Lys Leu
50 55 60
Gly Gly Phe Ala Ala Lys Arg Ile Leu Asn Gly Leu Trp Asp Leu Val
65 70 75 80
Phe Pro Asn Asp Asn Thr Lys Leu Met Gln Asp Ile Leu Arg Glu Thr
85 90 95
Glu Lys Phe Leu Asn Gln Arg Leu Asn Ala Asp Thr Leu Ser Arg Val
100 105 110
Asn Ala Glu Leu Glu Gly Leu Gln Lys Asn Val Ala Glu Phe Asn Arg
115 120 125
Gln Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Pro Thr Leu Leu Ser
130 135 140
Ile Thr Ser Pro Val Asn Thr Met Gln Gln Leu Phe Leu Asn Arg Leu
145 150 155 160
Thr Gln Phe Gln Leu Arg Gly Tyr Glu Leu Leu Leu Leu Pro Leu Phe
165 170 175
Ala Gln Ala Ala Asn Leu His Leu Ser Phe Ile Arg Asp Val Ile Leu
180 185 190
Asn Ala Glu Glu Trp Gly Ile Ser Ala Ala Thr Leu Arg Thr Tyr Arg
195 200 205
Gly Tyr Leu Lys Asp Tyr Thr Lys Glu Tyr Ser Asn Tyr Cys Ile Asn
210 215 220
Thr Tyr Gln Thr Ala Phe Lys Lys Leu Asn Thr Pro Leu His Glu Met
225 230 235 240
Leu Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val Ser
245 250 255
Ile Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly Ala
260 265 270
Asn Leu Tyr Ala Ser Gly Ser Gly Pro Gln Gln Thr Gln Ser Phe Thr
275 280 285
Ser Gln Asp Trp Pro Phe Leu Tyr Ser Leu Phe Gln Val Asn Ser Asn
290 295 300
Tyr Val Leu Asn Gly Phe Ser Gly Ala Arg Leu Ser Asn Thr Phe Pro
305 310 315 320
Asn Ile Val Gly Leu Pro Gly Ser Thr Thr Thr His Ala Leu Leu Ala
325 330 335
Ala Arg Val Asn Tyr Ser Gly Gly Ile Ser Ser Gly Asp Ile Gly Ala

[0002]

340 345 350
Ser Pro Phe Asn Gln Asn Phe Asn Cys Ser Thr Phe Leu Pro Pro Leu
355 360 365
Leu Thr Pro Phe Val Arg Ser Trp Leu Asp Ser Gly Ser Asp Arg Glu
370 375 380
Gly Val Ala Thr Val Thr Asn Trp Gln Thr Glu Ser Phe Glu Thr Thr
385 390 395 400
Leu Gly Leu Arg Ser Gly Ala Phe Thr Ala Arg Gly Asn Ser Asn Tyr
405 410 415
Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly Val Pro Leu Val Val
420 425 430
Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr Asn Glu Ile Arg Asn
435 440 445
Ile Ala Ser Pro Ser Gly Thr Pro Gly Gly Ala Arg Ala Tyr Met Val
450 455 460
Ser Val His Asn Arg Lys Asn Asn Ile His Ala Val His Glu Asn Gly
465 470 475 480
Ser Met Ile His Leu Ala Pro Asn Asp Tyr Thr Gly Phe Thr Ile Ser
485 490 495
Pro Ile His Ala Thr Gln Val Asn Asn Gln Thr Arg Thr Phe Ile Ser
500 505 510
Glu Lys Phe Gly Asn Gln Gly Asp Ser Leu Arg Phe Glu Gln Asn Asn
515 520 525
Thr Thr Ala Arg Tyr Thr Leu Arg Gly Asn Gly Asn Ser Tyr Asn Leu
530 535 540
Tyr Leu Arg Val Ser Ser Ile Gly Asn Ser Thr Ile Arg Val Thr Ile
545 550 555 560
Asn Gly Arg Val Tyr Thr Ala Thr Asn Val Asn Thr Thr Thr Asn Asn
565 570 575
Asp Gly Val Asn Asp Asn Gly Ala Arg Phe Ser Asp Ile Asn Ile Gly
580 585 590
Asn Val Val Ala Ser Ser Asn Ser Asp Val Pro Leu Asp Ile Asn Val
595 600 605
Thr Leu Asn Ser Gly Thr Gln Phe Asp Leu Met Asn Ile Met Leu Val
610 615 620
Pro Thr Asn Ile Ser Pro Leu Tyr
625 630
<210> 2
<211> 1917
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 2
ggatccacca tggcttccgt gctcaactcc ggacgcacca ccatctgcga cggccacaac 60
gtgatggccc aggacccgtt ctccttccag cacaagtccc tcgacaccat ccagaaggag 120
tggaccgagt ggaagcgcga cgaccactca ctctacgtgg ctcctgtggt gggtactgtg 180
gccatcttcc tcctcaagaa gctcggaggc ttcgctgcca agcgcatcct caatggcctc 240
tgggacctcg tgttcccaaa cgacaacacc aagctcatgc aggacatcct ccgtgagacc 300
gagaagttcc tcaaccagcg cctcaacgct gacaccctct cccgcgtgaa cgccgaactt 360
gagggcctcc agaagaacgt tgccgagttc aaccgtcagg tggacaactt cctcaacccg 420
aaccgtaacc cgactctcct ttccatcaca tccccggtga acaccatgca gcagctcttc 480
ctcaaccgcc tcacccagtt ccagctccgc ggctacgaac tccttctcct tccactcttc 540
gctcaggccg ccaacctcca cctctccttc atccgcgatg tgatcctcaa cgcagaggag 600
tggggtatct ccgctgccac tctccgcacc taccgtggct acctcaagga ctacaccaag 660
gagtactcca actactgcat caacacctac cagacagcct tcaagaagct caacacccct 720
ctccacgaga tgcttgagtt ccgcacctac atgttcctca acgtgttcga gtacgtgtcc 780
atctggtccc tcttcaagta ccagtccctc ctcgtgtcct ccggcgccaa cctctacgcc 840

[0003]

tccggctccg gcccgcagca gacccagtcc ttcacctccc aggactggcc gttcctctac 900
tccctcttcc aggtgaactc caactacgtg ctcaacggct tctccggcgc ccgcctctcc 960
aacaccttcc cgaacatcgt gggcctcccct ggctccacaa ctacccacgc cctcctcgcc 1020
gctcgcgtga actactccgg tggcatctct tccggcgaca tcggcgcttc cccgttcaac 1080
cagaacttca actgctctac cttcctccct ccgctcctca caccgttcgt gcgctcctgg 1140
ctcgactccg gctccgaccg cgagggcgtg gctaccgtga ccaactggca gaccgagtcc 1200
ttcgagacca ccctcggcct ccgctccggt gccttcaccg cacgcggcaa ctccaactac 1260
ttcccggact acttcatccg caacatctcc ggcgtgccgc tcgtggtgcg caacgaggac 1320
ctccgtcgcc cactccacta caacgagatc cgcaacatcg cctccccgtc cggcacccct 1380
ggaggtgcac gcgcttacat ggtgtccgtg cacaaccgca agaacaacat ccacgctgtg 1440
cacgagaacg gctccatgat ccacctcgct ccaaacgact acaccggctt caccatctcc 1500
ccgatccacg ctaccccaggt gaacaaccag acccgcacct tcatctccga gaagttcggc 1560
aaccagggcg actctctccg cttcgagcag aacaacacca ccgctcgtta caccctccgc 1620
ggtaacggca actcctacaa cctctacctc cgcgtgtctt ccatcggcaa ctccaccatc 1680
cgcgtgacca tcaacggccg tgtgtacacc gccacaaacg tgaacaccac cacaaacaac 1740
gacggcgtga acgacaacgg tgctcgcttc tccgacatca acatcggcaa cgtggtggcc 1800
tcctccaact ccgacgtgcc gctcgacatc aacgtgaccc tcaactctgg cacacagttc 1860
gacctcatga acatcatgct cgtgccgacc aacatctccc cgctctacta agagctc 1917

<210> 3
<211> 1287
<212> PRT
<213> unknown
<220>
<223> 人工序列
<400> 3

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

[0004]

260 265 270
Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile Glu
275 280 285
Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
290 295 300
Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His Gln
305 310 315 320
Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe Pro
325 330 335
Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val Ala
340 345 350
Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr Arg
355 360 365
Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu Asp
370 375 380
Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala Val
385 390 395 400
Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro Gln
405 410 415
Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser His
420 425 430
Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile Ile
435 440 445
Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn Asn
450 455 460
Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser Thr
465 470 475 480
Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr Gly
485 490 495
Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu Arg
500 505 510
Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile Arg
515 520 525
Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly Arg
530 535 540
Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser Asn
545 550 555 560
Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe Asn
565 570 575
Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe Asn
580 585 590
Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala Glu
595 600 605
Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val
610 615 620
Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
625 630 635 640
Thr Asp Tyr His Ile Asp Gln Val Arg Pro Gly Lys Gly Gly Gly Ala
645 650 655
Ser Val Leu Asn Ser Gly Arg Thr Thr Ile Cys Asp Gly His Asn Val
660 665 670
Met Ala Gln Asp Pro Phe Ser Phe Gln His Lys Ser Leu Asp Thr Ile
675 680 685
Gln Lys Glu Trp Thr Glu Trp Lys Arg Asp Asp His Ser Leu Tyr Val
690 695 700
Ala Pro Val Val Gly Thr Val Ala Ile Phe Leu Leu Lys Lys Leu Gly
705 710 715 720
Gly Phe Ala Ala Lys Arg Ile Leu Asn Gly Leu Trp Asp Leu Val Phe

[0005]

725 730 735
Pro Asn Asp Asn Thr Lys Leu Met Gln Asp Ile Leu Arg Glu Thr Glu
740 745 750
Lys Phe Leu Asn Gln Arg Leu Asn Ala Asp Thr Leu Ser Arg Val Asn
755 760 765
Ala Glu Leu Glu Gly Leu Gln Lys Asn Val Ala Glu Phe Asn Arg Gln
770 775 780
Val Asp Asn Phe Leu Asn Pro Asn Arg Asn Pro Thr Leu Leu Ser Ile
785 790 795 800
Thr Ser Pro Val Asn Thr Met Gln Gln Leu Phe Leu Asn Arg Leu Thr
805 810 815
Gln Phe Gln Leu Arg Gly Tyr Glu Leu Leu Leu Leu Pro Leu Phe Ala
820 825 830
Gln Ala Ala Asn Leu His Leu Ser Phe Ile Arg Asp Val Ile Leu Asn
835 840 845
Ala Glu Glu Trp Gly Ile Ser Ala Ala Thr Leu Arg Thr Tyr Arg Gly
850 855 860
Tyr Leu Lys Asp Tyr Thr Lys Glu Tyr Ser Asn Tyr Cys Ile Asn Thr
865 870 875 880
Tyr Gln Thr Ala Phe Lys Lys Leu Asn Thr Pro Leu His Glu Met Leu
885 890 895
Glu Phe Arg Thr Tyr Met Phe Leu Asn Val Phe Glu Tyr Val Ser Ile
900 905 910
Trp Ser Leu Phe Lys Tyr Gln Ser Leu Leu Val Ser Ser Gly Ala Asn
915 920 925
Leu Tyr Ala Ser Gly Ser Gly Pro Gln Gln Thr Gln Ser Phe Thr Ser
930 935 940
Gln Asp Trp Pro Phe Leu Tyr Ser Leu Phe Gln Val Asn Ser Asn Tyr
945 950 955 960
Val Leu Asn Gly Phe Ser Gly Ala Arg Leu Ser Asn Thr Phe Pro Asn
965 970 975
Ile Val Gly Leu Pro Gly Ser Thr Thr Thr His Ala Leu Leu Ala Ala
980 985 990
Arg Val Asn Tyr Ser Gly Gly Ile Ser Ser Gly Asp Ile Gly Ala Ser
995 1000 1005
Pro Phe Asn Gln Asn Phe Asn Cys Ser Thr Phe Leu Pro Pro Leu
1010 1015 1020
Leu Thr Pro Phe Val Arg Ser Trp Leu Asp Ser Gly Ser Asp Arg
1025 1030 1035
Glu Gly Val Ala Thr Val Thr Asn Trp Gln Thr Glu Ser Phe Glu
1040 1045 1050
Thr Thr Leu Gly Leu Arg Ser Gly Ala Phe Thr Ala Arg Gly Asn
1055 1060 1065
Ser Asn Tyr Phe Pro Asp Tyr Phe Ile Arg Asn Ile Ser Gly Val
1070 1075 1080
Pro Leu Val Val Arg Asn Glu Asp Leu Arg Arg Pro Leu His Tyr
1085 1090 1095
Asn Glu Ile Arg Asn Ile Ala Ser Pro Ser Gly Thr Pro Gly Gly
1100 1105 1110
Ala Arg Ala Tyr Met Val Ser Val His Asn Arg Lys Asn Asn Ile
1115 1120 1125
His Ala Val His Glu Asn Gly Ser Met Ile His Leu Ala Pro Asn
1130 1135 1140
Asp Tyr Thr Gly Phe Thr Ile Ser Pro Ile His Ala Thr Gln Val
1145 1150 1155
Asn Asn Gln Thr Arg Thr Phe Ile Ser Glu Lys Phe Gly Asn Gln
1160 1165 1170
Gly Asp Ser Leu Arg Phe Glu Gln Asn Asn Thr Thr Ala Arg Tyr

[0006]

1175 1180 1185
Thr Leu Arg Gly Asn Gly Asn Ser Tyr Asn Leu Tyr Leu Arg Val
1190 1195 1200
Ser Ser Ile Gly Asn Ser Thr Ile Arg Val Thr Ile Asn Gly Arg
1205 1210 1215
Val Tyr Thr Ala Thr Asn Val Asn Thr Thr Thr Asn Asn Asp Gly
1220 1225 1230
Val Asn Asp Asn Gly Ala Arg Phe Ser Asp Ile Asn Ile Gly Asn
1235 1240 1245
Val Val Ala Ser Ser Asn Ser Asp Val Pro Leu Asp Ile Asn Val
1250 1255 1260
Thr Leu Asn Ser Gly Thr Gln Phe Asp Leu Met Asn Ile Met Leu
1265 1270 1275
Val Pro Thr Asn Ile Ser Pro Leu Tyr
1280 1285

<210> 4
<211> 3878
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 4

ggatccacca tggacaacaa ccccaacatc aacgagtgca tcccctacaa ctgcctgagc 60
aaccccgagg tggaggtgct gggcggcgag cgcatcgaga ccggctacac ccccatcgac 120
atcagcctga gcctgaccca gttcctgctg agcgagttcg tgcccggcgc cggcttcgtg 180
ctgggcctgg tggacatcat ctggggcatc ttcggcccca gccagtggga cgccttcctg 240
gtgcagatcg agcagctgat caaccagcgc atcgaggagt tcgcccgcaa ccaggccatc 300
agccgcctgg agggcctgag caacctgtac caaatctacg ccgagagctt ccgcgagtgg 360
gaggccgacc ccaccaaccc cgccctgcgc gaggagatgc gcatccagtt caacgacatg 420
aacagcgccc tgaccaccgc catccccctg ttcgccgtgc agaactacca ggtgcccctg 480
ctgagcgtgt acgtgcaggc cgccaacctg cacctgagcg tgctgcgcga cgtcagcgtg 540
ttcggccagc gctggggctt cgacgccgcc accatcaaca gccgctacaa cgacctgacc 600
cgcctgatcg gcaactacac cgaccacgcc gtgcgctggt acaacaccgg cctggagcgc 660
gtgtggggtc ccgacagccg cgactggatc aggtacaacc agttccgccg cgagctgacc 720
ctgaccgtgc tggacatcgt gagcctgttc cccaactacg acagccgcac ctaccccatc 780
cgcaccgtga gccagctgac ccgcgagatt tacaccaacc ccgtgctgga gaacttcgac 840
ggcagcttcc gcggcagcgc ccagggcatc gagggcagca tccgcagccc ccacctgatg 900
gacatcctga acagatcacc atctacaccg acgcccaccg cggcgagtac tactggagcg 960
gccaccagat catggccagc cccgtcggct tcagcggccc cgagttcacc ttccccctgt 1020
acggcaccat gggcaacgct gcacctcagc agcgcatagt ggcacagctg ggccagggag 1080
tgtaccgcac cctgagcagc accctgtacc gtcgaccttt caacatcggc atcaacaacc 1140
agcagctgag cgtgctggac ggcaccgagt tcgcctacgg caccagcagc aacctgccca 1200
gcgccgtgta ccgcaagagc ggcaccgtgg acagcctgga cgagatcccc cctcagaaca 1260
acaacgtgcc acctcgacag ggcttcagcc accgtctgag ccacgtgagc atgttccgca 1320
gtggcttcag caacagcagc gtgagcatca tccgtgcacc tatgttcagc tggattcacc 1380
gcagtgccga gttcaacaac atcatccccca gcagccagat cacccagatc cccctgacca 1440
agagcaccaa cctgggcagc ggcaccagcg tggtgaaggg ccccggcttc accggcggcg 1500
acatcctgcg ccgcaccagc cccggccaga tcagcaccct gcgcgtgaac atcaccgccc 1560
ccctgagcca gcgctaccgc gtccgcatcc gctacgccag caccaccaac ctgcagttcc 1620
acaccagcat cgacggccgc cccatcaacc agggcaactt cagcgccacc atgagcagcg 1680
gcagcaacct gcagagcggc agcttccgca ccgtgggctt caccaccccc ttcaacttca 1740
gcaacggcag cagcgtgttc accctgagcg cccacgtgtt caacagcggc aacgaggtgt 1800
acatcgaccg catcgagttc gtgcccgccg aggtgacctt cgaggccgag tacgacctgg 1860
agagggctca gaaggccgtg aacgagctgt tcaccagcag caaccagatc ggcctgaaga 1920
ccgacgtgac cgactaccac atcgatcagg tgcgacccgg gaagggtgga ggagcttccg 1980
tgctcaactc cggacgcacc accatctgcg acggccacaa cgtgatggcc caggacccgt 2040
tctccttcca gcacaagtcc ctcgacacca tccagaagga gtggaccgag tggaagcgcg 2100
acgaccactc actctacgtg gctcctgtgg tgggtactgt ggccatcttc ctcctcaaga 2160

[0007]

agctcggagg cttcgctgcc aagcgcatcc tcaatggcct ctgggacctc gtgttcccaa 2220
acgacaacac caagctcatg caggacatcc tccgtgagac cgagaagttc ctcaaccagc 2280
gcctcaacgc tgacacccctc tcccgcgtga acgccgaact tgagggcctc cagaagaacg 2340
ttgccgagtt caaccgtcag gtggacaact tcctcaaccc gaaccgtaac ccgactctcc 2400
ttttccatcac atccccggtg aacaccatgc agcagctctt cctcaaccgc ctcacccagt 2460
tccagctccg cggctacgaa ctccttctcc ttccactctt cgctcaggcc gccaacctcc 2520
acctctcctt catccgcgat gtgatcctca acgcagagga gtggggtatc tccgctgcca 2580
ctctccgcac ctaccgtggc tacctcaagg actacaccaa ggagtactcc aactactgca 2640
tcaacaccta ccagacagcc ttcaagaagc tcaacacccc tctccacgag atgcttgagt 2700
tccgcaccta catgttcctc aacgtgttcg agtacgtgtc catctggtcc ctcttcaagt 2760
accagtccct cctcgtgtcc tccggcgcca acctctacgc ctccggctcc ggcccgcagc 2820
agacccagtc cttcacctcc caggactggc cgttcctcta ctccctcttc caggtgaact 2880
ccaactacgt gctcaacggc ttctccggcg cccgcctctc caacaccttc ccgaacatcg 2940
tgggcctccc tggctccaca actaccccacg ccctcctcgc cgctcgcgtg aactactccg 3000
gtggcatctc ttccggcgac atcggcgctt ccccgttcaa ccagaacttc aactgctcta 3060
ccttcctccc tccgctcctc acaccgttcg tgcgctcctg gctcgactcc ggctccgacc 3120
gcgagggcgt ggctaccgtg accaactggc agaccgagtc cttcgagacc accctcggcc 3180
tccgctccgg tgccttcacc gcacgcggca actccaacta cttcccggac tacttcatcc 3240
gcaacatctc cggcgtgccg ctcgtggtgc gcaacgagga cctccgtcgc ccactccact 3300
acaacgagat ccgcaacatc gcctccccgt ccggcacccc tggaggtgca cgcgcttaca 3360
tggtgtccgt gcacaaccgc aagaacaaca tccacgctgt gcacgagaac ggctccatga 3420
tccacctcgc tccaaacgac tacaccggct tcaccatctc cccgatccac gctacccagg 3480
tgaacaacca gacccgcacc ttcatctccg agaagttcgg caaccagggc gactctctcc 3540
gcttcgagca gaacaacacc accgctcgtt acaccctccg cggtaacggc aactcctaca 3600
acctctacct ccgcgtgtct tccatcggca actccaccat ccgcgtgacc atcaacggcc 3660
gtgtgtacac cgccacaaac gtgaacacca ccacaaacaa cgacggcgtg aacgacaacg 3720
gtgctcgctt ctccgacatc aacatcggca acgtggtggc ctcctccaac tccgacgtgc 3780
cgctcgacat caacgtgacc ctcaactctg gcacacagtt cgacctcatg aacatcatgc 3840
tcgtgccgac caacatctcc ccgctctact aagagctc 3878

<210> 5
<211> 650
<212> PRT
<213> unknown
<220>
<223> 人工序列
<400> 5

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

[0008]

Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg
180 185 190
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp His Ala Val
195 200 205
Arg Trp Tyr Asn Thr Gly Leu Glu Ala Arg Val Trp Gly Pro Asp Ser
210 215 220
Arg Asp Trp Ile Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr
225 230 235 240
Val Leu Asp Ile Val Ser Leu Phe Pro Asn Tyr Asp Ser Arg Thr Tyr
245 250 255
Pro Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro
260 265 270
Val Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Ser Ala Gln Gly Ile
275 280 285
Glu Gly Ser Ile Arg Ser Pro His Leu Met Asp Ile Leu Asn Ser Ile
290 295 300
Thr Ile Tyr Thr Asp Ala His Arg Gly Glu Tyr Tyr Trp Ser Gly His
305 310 315 320
Gln Ile Met Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Thr Phe
325 330 335
Pro Leu Tyr Gly Thr Met Gly Asn Ala Ala Pro Gln Gln Arg Ile Val
340 345 350
Ala Gln Leu Gly Gln Gly Val Tyr Arg Thr Leu Ser Ser Thr Leu Tyr
355 360 365
Arg Arg Pro Phe Asn Ile Gly Ile Asn Asn Gln Gln Leu Ser Val Leu
370 375 380
Asp Gly Thr Glu Phe Ala Tyr Gly Thr Ser Ser Asn Leu Pro Ser Ala
385 390 395 400
Val Tyr Arg Lys Ser Gly Thr Val Asp Ser Leu Asp Glu Ile Pro Pro
405 410 415
Gln Asn Asn Asn Val Pro Pro Arg Gln Gly Phe Ser His Arg Leu Ser
420 425 430
His Val Ser Met Phe Arg Ser Gly Phe Ser Asn Ser Ser Val Ser Ile
435 440 445
Ile Arg Ala Pro Met Phe Ser Trp Ile His Arg Ser Ala Glu Phe Asn
450 455 460
Asn Ile Ile Pro Ser Ser Gln Ile Thr Gln Ile Pro Leu Thr Lys Ser
465 470 475 480
Thr Asn Leu Gly Ser Gly Thr Ser Val Val Lys Gly Pro Gly Phe Thr
485 490 495
Gly Gly Asp Ile Leu Arg Arg Thr Ser Pro Gly Gln Ile Ser Thr Leu
500 505 510
Arg Val Asn Ile Thr Ala Pro Leu Ser Gln Arg Tyr Arg Val Arg Ile
515 520 525
Arg Tyr Ala Ser Thr Thr Asn Leu Gln Phe His Thr Ser Ile Asp Gly
530 535 540
Arg Pro Ile Asn Gln Gly Asn Phe Ser Ala Thr Met Ser Ser Gly Ser
545 550 555 560
Asn Leu Gln Ser Gly Ser Phe Arg Thr Val Gly Phe Thr Thr Pro Phe
565 570 575
Asn Phe Ser Asn Gly Ser Ser Val Phe Thr Leu Ser Ala His Val Phe
580 585 590
Asn Ser Gly Asn Glu Val Tyr Ile Asp Arg Ile Glu Phe Val Pro Ala
595 600 605
Glu Val Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala
610 615 620
Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp
625 630 635 640

[0009]

Val Thr Asp Tyr His Ile Asp Gln Val Arg
645 650

<210> 6
<211> 1965
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 6

ggatccacca tggacaacaa ccccaacatc aacgagtgca tcccctacaa ctgcctgagc 60
aaccccgagg tggaggtgct gggcggcgag cgcatcgaga ccggctacac ccccatcgac 120
atcagcctga gcctgaccca gttcctgctg agcgagttcg tgcccggcgc cggcttcgtg 180
ctgggcctgg tggacatcat ctggggcatc ttcggcccca gccagtggga cgccttcctg 240
gtgcagatcg agcagctgat caaccagcgc atcgaggagt tcgcccgcaa ccaggccatc 300
agccgcctgg agggcctgag caacctgtac caaatctacg ccgagagctt ccgcgagtgg 360
gaggccgacc ccaccaaccc cgccctgcgc gaggagatgc gcatccagtt caacgacatg 420
aacagcgccc tgaccaccgc catccccctg ttcgccgtgc agaactacca ggtgcccctg 480
ctgagcgtgt acgtgcaggc cgccaacctg cacctgagcg tgctgcgcga cgtcagcgtg 540
ttcggccagc gctggggctt cgacgccgcc accatcaaca gccgctacaa cgacctgacc 600
cgcctgatcg gcaactacac cgaccacgcc gtgcgctggt acaacaccgg cctggagcgc 660
gtgtggggtc ccgacagccg cgactggatc aggtacaacc agttccgccg cgagctgacc 720
ctgaccgtgc tggacatcgt gagcctgttc cccaactacg acagccgcac ctaccccatc 780
cgcaccgtga gccagctgac ccgcgagatt tacaccaacc ccgtgctgga gaacttcgac 840
ggcagcttcc gcggcagcgc ccagggcatc gagggcagca tccgcagccc ccacctgatg 900
gacatcctga acagcatcac catctacacc gacgcccacc gcggcgagta ctactggagc 960
ggccaccaga tcatggccag ccccgtcggc ttcagcggcc ccgagttcac cttccccctg 1020
tacggcacca tgggcaacgc tgcacctcag cagcgcatag tggcacagct gggccaggga 1080
gtgtaccgca ccctgagcag caccctgtac cgtcgacctt tcaacatcgg catcaacaac 1140
cagcagctga gcgtgctgga cggcaccgag ttcgcctacg gcaccagcag caacctgccc 1200
agcgccgtgt accgcaagag cggcaccgtg gacagcctgg acgagatccc ccctcagaac 1260
aacaacgtgc cacctcgaca gggcttcagc caccgtctga gccacgtgag catgttccgc 1320
agtggcttca gcaacagcag cgtgagcatc atccgtgcac ctatgttcag ctggattcac 1380
cgcagtgccg agttcaacaa catcatcccc agcagccaga tcacccagat ccccctgacc 1440
aagagcacca acctgggcag cggcaccagc gtggtgaagg gccccggctt caccggcggc 1500
gacatcctgc gccgcaccag ccccggccag atcagcaccc tgcgcgtgaa catcaccgcc 1560
cccctgagcc agcgctaccg cgtccgcatc cgctacgcca gcaccaccaa cctgcagttc 1620
cacaccagca tcgacggccg ccccatcaac cagggcaact tcagcgccac catgagcagc 1680
ggcagcaacc tgcagagcgg cagcttccgc accgtgggct tcaccacccc cttcaacttc 1740
agcaacggca gcagcgtgtt cacccctgagc gcccacgtgt tcaacagcgg caacgaggtg 1800
tacatcgacc gcatcgagtt cgtgcccgcc gaggtgacct tcgaggccga gtacgacctg 1860
gagagggctc agaaggccgt gaacgagctg ttcaccagca gcaaccagat cggcctgaag 1920
accgacgtga ccgactacca catcgatcag gtgcgataag agctc 1965

<210> 7
<211> 2016
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 7

aagcttgcat gcctacagtg cagcgtgacc cggtcgtgcc cctctctaga gataatgagc 60
attgcatgtc taagttataa aaaattacca catatttttt ttgtcacact tgtttgaagt 120
gcagtttatc tatctttata catatattta aactttactc tacgaataat ataatctata 180
gtactacaat aatatcagtg ttttagagaa tcatataaat gaacagttag acatggtcta 240
aaggacaatt gagtattttg acaacaggac tctacagttt tatcttttta gtgtgcatgt 300
gttctccttt ttttttgcaa atagcttcac ctatataata cttcatccat tttattagta 360
catccattta gggtttaggg ttaatggttt ttatagacta atttttttag tacatctatt 420
ttattctatt ttagcctcta aattaagaaa actaaaactc tattttagtt tttttattta 480
ataatttaga tataaaatag aataaaataa agtgactaaa aattaaacaa ataccccttta 540

[0010]

agaaattaaa aaaactaagg aaacattttt cttgtttcga gtagataatg ccagcctgtt 600
aaacgccgtc gacgagtcta acggacacca accagcgaac cagcagcgtc gcgtcgggcc 660
aagcgaagca gacggcacgg catctctgtc gctgcctctg gacccctctc gagagttccg 720
ctccaccgtt ggacttgctc cgctgtcggc atccagaaat tgcgtggcgg agcggcagac 780
gtgagccggc acggcaggcg gcctcctcct cctctcacgg cacggcagct acgggggatt 840
cctttcccac cgctccttcg ctttcccttc ctcgcccgcc gtaataaata gacacccccct 900
ccacaccctc tttccccaac ctcgtgttgt tcggagcgca cacacacaca accagatctc 960
ccccaaatcc acccgtcggc acctccgctt caaggtacgc cgctcgtcct cccccccccc 1020
ccctctctac cttctctaga tcggcgttcc ggtccatggt tagggcccgg tagttctact 1080
tctgttcatg tttgtgttag atccgtgttt gtgttagatc cgtgctgcta gcgttcgtac 1140
acggatgcga cctgtacgtc agacacgttc tgattgctaa cttgccagtg tttctctttg 1200
gggaatcctg ggatggctct agccgttccg cagacgggat cgatttcatg attttttttg 1260
tttcgttgca tagggtttgg tttgcccttt tcctttattt caatatatgc cgtgcacttg 1320
tttgtcgggt catcttttca tgctttttttt tgtcttggtt gtgatgatgt ggtctggttg 1380
ggcggtcgtt ctagatcgga gtagaattct gtttcaaact acctggtgga tttattaatt 1440
ttggatctgt atgtgtgtgc catacatatt catagttacg aattgaagat gatggatgga 1500
aatatcgatc taggataggt atacatgttg atgcgggttt tactgatgca tatacagaga 1560
tgctttttgt tcgcttggtt gtgatgatgt ggtgtggttg ggcggtcgtt cattcgttct 1620
agatcggagt agaatactgt ttcaaactac ctggtgtatt tattaatttt ggaactgtat 1680
gtgtgtgtca tacatcttca tagttacgag tttaagatgg atggaaatat cgatctagga 1740
taggtataca tgttgatgtg ggttttactg atgcatatac atgatggcat atgcagcatc 1800
tattcatatg ctctaacctt gagtacctat ctattataat aaacaagtat gttttataat 1860
tattttgatc ttgatatact tggatgatgg catatgcagc agctatatgt ggatttttttt 1920
agccctgcct tcatacgcta tttatttgct tggtactgtt tcttttgtcg atgctcaccc 1980
tgttgtttgg tgttacttct gcaggtcgac tctaga 2016

<210> 8
<211> 297
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 8

aagagctcac tggctaggcg gcttctcttc actcacctgc agagtgcacc gcaataatca 60
gcttccggat ggtggccgtt ttgtcagttt tggatggaaa tgccgaactg gccagcgtct 120
gttttcccta tgcatatgta atttcctgcc tctttatatt cactcttgtt gtcaagtcca 180
agtggaaaat cttggcatat tatacatatt gtaataataa acatcgtaca atctgcatgc 240
tgttttgtaa taattaatta atatcccagc ccattggatg gacttgttta aggtacc 297

<210> 9
<211> 42
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 9

gcgaagcttg catgcctaca gtgcagcgtg acccggtcgt gc 42

<210> 10
<211> 46
<212> DNA
<213> unknown
<220>
<223> 人工序列
<400> 10

gtgggatcct ctagagtcga cctgcagaag taacaccaaa caacag 46