



(12) 发明专利

(10) 授权公告号 CN 110267976 B

(45) 授权公告日 2024.01.26

(21) 申请号 201880010953.2

B·M·塔奇克 T·T·沃恩

(22) 申请日 2018.01.11

M·W·F·F·雷兴贝格

(65) 同一申请的已公布的文献号
申请公布号 CN 110267976 A

(74) 专利代理机构 北京市金杜律师事务所
11256

(43) 申请公布日 2019.09.20

专利代理师 孟凡宏 王月

(30) 优先权数据

(51) Int.Cl.

62/445,313 2017.01.12 US

C07K 14/325 (2006.01)

(85) PCT国际申请进入国家阶段日
2019.08.08

C12N 15/82 (2006.01)

(86) PCT国际申请的申请数据

A01H 5/00 (2018.01)

PCT/US2018/013298 2018.01.11

A01H 6/46 (2018.01)

(87) PCT国际申请的公布数据

A01N 37/46 (2006.01)

W02018/132556 EN 2018.07.19

A01P 7/04 (2006.01)

(73) 专利权人 孟山都技术公司
地址 美国密苏里州

(56) 对比文件

WO 2016061391 A2, 2016.04.21

(72) 发明人 J·A·鲍姆 D·J·伯温

US 2012035135 A1, 2012.02.09

C·A·蔡 D·J·池

US 2015148288 A1, 2015.05.28

W·P·克林顿 C·L·达特

US 2016108427 A1, 2016.04.21

L·英格里什 S·弗拉辛斯基

CN 105348374 A, 2016.02.24

V·M·古佐夫 K·A·贾雷尔

G Osman等.Purification and
characterization of Bacillus
thuringiensis vegetative insecticidal
toxin protein(s).《Lett Appl Microbiol .
》.2013,第57卷(第4期),310-316.

U·R·凯森帕利 T·M·马尔瓦尔

R·M·麦卡罗尔 J·S·米利甘

J·P·摩根斯特恩 D·G·洛克

S·A·萨尔瓦多 T·F·史密斯

C·E·索托 C·M·斯塔尔茨

审查员 蔡放

权利要求书2页 说明书18页

序列表74页 附图3页

(54) 发明名称

对鳞翅目昆虫具有活性的杀害虫毒素蛋白
质

或所述杀昆虫剂在生物样品中的存在性的方法。

(57) 摘要

本文公开了编码展现鳞翅目昆虫抑制活性的
杀昆虫蛋白质的核苷酸序列,含有黑切根虫
(BCW) 毒性蛋白质的嵌合体和BCW毒素杀昆虫剂,
表达所述嵌合体或所述杀昆虫剂的转基因植物,
以及在本文中称为BCW 001、BCW 002、BCW 003的
新型杀昆虫蛋白质,和用于检测所述核苷酸序列

CN 110267976 B

1. 一种包含核苷酸序列的多核苷酸构建体,所述核苷酸序列编码:
SEQ ID NO:6的氨基酸序列的杀昆虫蛋白质;
其中所述核苷酸序列可操作地连接到异源性启动子序列。
2. 一种对黑切根虫鳞翅目种具有毒性的蛋白质,其氨基酸序列如SEQ ID NO:6所示。
3. 根据权利要求1所述的多核苷酸构建体,其中所述杀昆虫蛋白质展现针对选自以下组成的组的鳞翅目种的活性:小地老虎、豆白缘切根虫、玉米穗虫、玉米螟、小蔗秆草螟、西南玉米秆草螟、粉纹夜蛾和大豆夜蛾。
4. 根据权利要求2所述的蛋白质,其中所述蛋白质展现针对选自以下组成的组的鳞翅目种的生物活性:小地老虎、豆白缘切根虫、玉米穗虫、玉米螟、小蔗秆草螟、西南玉米秆草螟、粉纹夜蛾和大豆夜蛾。
5. 一种载体,其包含如权利要求1所述的多核苷酸构建体。
6. 一种宿主细胞,其包含如权利要求1所述的多核苷酸构建体,其中所述宿主细胞选自细菌细胞、酵母细胞和植物细胞组成的组,所述植物细胞是不可再生的。
7. 如权利要求6所述的宿主细胞,其中所述植物细胞选自以下组成的组:芸苔、观赏植物、砧木和灌木的植物细胞。
8. 如权利要求6所述的宿主细胞,其中所述植物细胞选自以下组成的组:苜蓿、香蕉、大麦、菜豆、西兰花、卷心菜、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、大白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、花旗松、茄子、桉树、亚麻、大蒜、葡萄、蛇麻子、韭菜、莴苣、火炬松、粟、甜瓜、坚果、燕麦、橄榄、洋葱、棕榈、牧场草、豌豆、花生、胡椒、木豆、松树、马铃薯、白杨、南瓜、辐射松、萝卜、油菜籽、稻米、红花、高粱、南方松、大豆、菠菜、南瓜小果、草莓、糖用甜菜、甘蔗、向日葵、甜玉米、枫香、甜薯、柳枝黍、茶、烟草、番茄、草皮草、西瓜,以及小麦的植物细胞,所述植物细胞是不可再生的。
9. 如权利要求6所述的宿主细胞,其中所述植物细胞选自黑麦或黑小麦的植物细胞。
10. 一种生物样品,其包含可检测量的如权利要求1所述的多核苷酸构建体,所述生物样品是不可再生的。
11. 一种组合物,其提供杀昆虫有效量的如权利要求2所述的蛋白质以防治鳞翅目害虫物种,以及:
 - (a) 不同于所述蛋白质并且也对相同鳞翅目种具有毒性的剂,其中所述剂选自以下组成的组:具有不同于所述蛋白质的氨基酸序列的多肽、RNA分子,和化学化合物;或
 - (b) 选自以下组成的组的剂:Cry1A、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B、Cry1C、Cry1D、Cry1E、Cry1F、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab、Cry2Ae、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry43A、Cry43B、ET35、ET66、TIC400、TIC800、TIC807、TIC834、TIC853、TIC1415、VIP3A、VIP3Ab、AXMI杀昆虫蛋白、DIG杀昆虫蛋白、eHIP和VIP蛋白。
12. 如权利要求11所述的组合物,其进一步包含额外杀害虫剂,其中所述额外杀害虫剂选自以下组成的组:Cry3A、Cry3B、Cry34、Cry35、Cry51Aa1、ET29、ET33、ET34、ET70、TIC407、TIC417、TIC431、TIC901、TIC1201、TIC3131、DIG-10、Axmi184、Axmi205和AxmiR1。
13. 一种产生包含如权利要求1所述的多核苷酸构建体的种子的方法,所述方法包括:
 - (a) 种植一个或多个包含所述多核苷酸构建体的种子;

- (b) 从所述种子生长出植物;以及
- (c) 从所述植物收集许多种子,其中所述种子包含所述多核苷酸构建体。

对鳞翅目昆虫具有活性的杀害虫毒素蛋白质

[0001] 相关申请的引用

[0002] 本申请要求2017年1月12日提交的美国临时申请号62/445,313的权益,所述美国临时申请以引用的方式整体并入本文。

[0003] 序列表的并入

[0004] 本文提供计算机可读形式的序列列表,其含有于2018年1月11日创建的名为“MONS434W0-sequence listing.txt”的文件。如在MS-Windows®中所测量,该文件为166,134字节,并且以引用的方式整体并入本文。这个序列列表由SEQ ID NO:1-SEQ ID NO:12组成。

技术领域

[0005] 本发明总体涉及昆虫抑制性蛋白质的领域,特别是涉及展现针对植物和种子的农业相关鳞翅目(Lepidopteran)害虫的昆虫抑制活性的蛋白质,所述植物和种子的农业相关鳞翅目害虫特别是诸如黑切根虫(“BCW”,小地老虎(*Agrotis ipsilon*))的鳞翅目害虫。

[0006] 发明背景

[0007] 由苏云金芽孢杆菌(*Bacillus thuringiensis*, Bt)细菌物种产生的昆虫抑制性蛋白质是在本领域中已知的。某些Bt蛋白质可用于通过将农业上可接受的含有一种或多种此类蛋白质的制剂喷雾于植物上,用被配制为含有杀昆虫有效量的此类蛋白质的组合物对种子进行包衣,或通过植物/种子中表达结果有效的一种或多种所述蛋白质来防治作物植物的农业害虫。

[0008] 仅少许Bt蛋白质已被开发来用作转基因性状以实现由农民进行商业使用以防治昆虫害虫。农民依赖于这些蛋白质来提供指定害虫防治谱,并且可继续在叶和土壤施用中依赖于广谱化学物质来防治害虫。某些鳞翅目昆虫,诸如地老虎属(*Agrotis*)种和*Striacosta*种,已被证明特别难以使用当前在用的转基因杀昆虫性状进行防治,所述性状包括Cry1Ab、Cry1Ac、Cry1Fa、Cry2Ab、Cry2Ae、VIP3Aa和已较不频繁使用的各种其他Bt毒素。因此,需要这样的昆虫抑制性蛋白质,所述昆虫抑制性蛋白质展现针对更广泛谱系的昆虫害虫物种的活性,并且用于在毒素中使用以用于克服害虫对包括当前用于害虫管理系统中的毒素的现有杀害虫剂的抗性发展。

[0009] 本申请描述新型蛋白质家族、变体和嵌合毒素蛋白质构建体,所述新型蛋白质家族、变体和嵌合毒素蛋白质构建体各自展现惊人有效的针对鳞翅目,特别是针对诸如黑切根虫之类的地老虎属种害虫的杀昆虫活性。

发明内容

[0010] 显示一组新型昆虫抑制性多肽(毒素蛋白质BCW 001、BCW 002和BCW 003,以及它们的杀害虫片段)会展现针对作物植物的若干鳞翅目害虫,特别是针对黑切根虫物种(地老虎属种)的抑制活性。各蛋白质可单独或彼此组合以及与其他Bt蛋白质和昆虫抑制剂组合用于制剂中以及植物中,由此提供当前用于农业系统中的Bt蛋白质和杀昆虫剂化学物质的

替代物。

[0011] 本发明提供多核苷酸构建体,其以可操作键联方式含有异源性启动子区段,所述异源性启动子区段连接到编码具有Cry1A特征的杀昆虫蛋白质或其杀昆虫活性片段的核苷酸序列,所述杀昆虫蛋白质相对于Cry1A类毒素蛋白质小于全长,并且具有如以SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO 6、SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12阐述的从约位置1至位置607的氨基酸序列。展现此类杀昆虫活性的小于全长多肽应展现与如以SEQ ID NO:2阐述的BCW 001氨基酸序列从约位置1至约位置606或从约位置5至约位置600具有至少约100%、99%、98%、97%、96%、95%、94%、94%、92%、91%或90%同一性。如果将使用全长或显著较大毒素片段,那么同一性百分比应具有较小严格性,并且扩展至与如以SEQ ID NO:2、SEQ ID NO:4和SEQ ID NO:6阐述的全长毒素蛋白质序列具有从约100%至约95%、至约90%、至约85%或甚至80%同一性的同一性百分比,因为当在膳食生物测定中针对黑切根虫幼虫测试时,以及当在植物中在玉米、棉花中以及在表达此类蛋白质的大豆转基因事件中测试时,这些毒素蛋白质展现商业适用水平的生物活性。

[0012] 本发明还提供对黑切根虫鳞翅目种具有毒性的蛋白质,包括具有如以SEQ ID NO:2阐述的从位置256至606的氨基酸序列的蛋白质(BCW 001蛋白),以及具有如以SEQ ID NO:4和SEQ ID NO:6中的任一者阐述的从氨基酸位置257至607的氨基酸序列的蛋白质(在本文中分别称为BCW 002毒素蛋白和BCW 003毒素蛋白)。

[0013] 还观察到所述杀昆虫蛋白质展现针对选自由以下组成的组的鳞翅目种的活性:草地夜蛾(*Spodoptera frugiperda*)、甜菜夜蛾(*Spodoptera exigua*)、斜纹夜蛾(*Spodoptera litura*)、蓓带夜蛾(*Mamestra configurata*)、豆白缘切根虫(*Striacosta albicosta*)、粉纹夜蛾(*Trichoplusia ni*)、大豆夜蛾(*Pseudoplusia includens*)、黎豆夜蛾(*Anticarsia gemmatalis*)、苜蓿绿夜蛾(*Hypena scabra*)、烟芽夜蛾(*Heliothis virescens*)、粒肤地老虎(*Agrotis subterranea*)、一星粘虫(*Pseudaletia unipuncta*)、小地老虎、西方灰地老虎(*Agrotis orthogonia*)、玉米螟(*Ostrinia nubilalis*)、脐橙螟(*Amyelois transitella*)、玉米根草螟(*Crambus caliginosellus*)、稻切叶野螟(*Herpetogramma licarsisalis*)、向日葵螟(*Homoeosoma electellum*)、南美玉米苗斑螟(*Elasmopalpus lignosellus*)、苹果蠹蛾(*Cydia pomonella*)、葡萄小食心虫(*Endopiza viteana*)、梨小食心虫(*Grapholita molesta*)、向日葵芽卷叶蛾(*Suleima helianthana*)、小菜蛾(*Plutella xylostella*)、棉红铃虫(*Pectinophora gossypiella*)、舞毒蛾(*Lymantria dispar*)、棉叶波纹夜蛾(*Alabama argillacea*)、果树黄卷蛾(*Archips argyrospila*)、玫瑰黄卷蛾(*Archips rosana*)、二化螟(*Chilo suppressalis*)、稻纵卷叶螟(*Cnaphalocrocis medinalis*)、玉米根草螟(*Crambus caliginosellus*)、早熟禾草螟(*Crambus teterrellus*)、西南玉米秆草螟(*Diatraea grandiosella*)、小蔗秆草螟(*Diatraea saccharalis*)、埃及钻夜蛾(*Earias insulana*)、翠纹钻夜蛾(*Egrias vittella*)、棉铃虫(*Helicoverpa armigera*)、玉米穗虫(*Helicoverpa zea*)、烟芽夜蛾(*Heliothis virescens*)、稻切叶野螟(*Herpetogramma licarsisalis*)、葡萄花翅小卷蛾(*Lobesia botrana*)、棉红铃虫(*Pectinophora gossypiella*)、柑桔潜叶蛾(*Phyllocnistis citrella*)、欧洲粉蝶(*Pieris brassicae*)、菜粉蝶(*Pieris rapae*)、小菜蛾(*Plutella xylostella*)和番茄斑潜蝇(*Tuta absoluta*)。

[0014] 本发明的蛋白质也可展现针对选自由以下组成的组的鳞翅目种的生物活性:草地

夜蛾、甜菜夜蛾、斜纹夜蛾、蓓带夜蛾、豆白缘切根虫、粉纹夜蛾、大豆夜蛾、黎豆夜蛾、苜蓿绿夜蛾、烟芽夜蛾、粒肤地老虎、一星粘虫、小地老虎、西方灰地老虎、玉米螟、脐橙螟、玉米根草螟、稻切叶野螟、向日葵螟、南美玉米苗斑螟、苹果蠹蛾、葡萄小食心虫、梨小食心虫、向日葵芽卷叶蛾、小菜蛾、棉红铃虫、舞毒蛾、棉叶波纹夜蛾、果树黄卷蛾、玫瑰黄卷蛾、二化螟、稻纵卷叶螟、玉米根草螟、早熟禾草螟、西南玉米秆草螟、小蔗秆草螟、埃及钻夜蛾、翠纹钻夜蛾、棉铃虫、玉米穗虫、烟芽夜蛾、稻切叶野螟、葡萄花翅小卷蛾、棉红铃虫、柑桔潜叶蛾、欧洲粉蝶、菜粉蝶、小菜蛾和番茄斑潜蝇。

[0015] 本发明的蛋白质和本文涵盖的构建体可包括在任何载体中,所述载体包括质粒、粘粒、杆粒、噬菌体介导的载体等。

[0016] 此类载体可用于将本发明的构建体引入许多宿主细胞中,包括引入细菌细胞、酵母细胞和植物细胞中。

[0017] 为酵母细胞的宿主细胞可为酿酒酵母 (*Saccharomyces cerevisiae*) 或裂殖酵母 (*Saccharomyces pombe*) 等。细菌宿主细胞可为许多已知的此类宿主细胞,包括但不限于大肠杆菌 (*E. coli*)、苏云金芽孢杆菌和其他相关芽孢杆菌。植物宿主细胞可从许多植物物种获得,包括但不限于来自以下植物的植物细胞:苜蓿、香蕉、大麦、菜豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、大白菜、柑桔、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、花旗松、茄子、桉树、亚麻、大蒜、葡萄、蛇麻子、韭菜、莴苣、火炬松、粟、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧场草、豌豆、花生、胡椒、木豆、松树、马铃薯、白杨、南瓜、辐射松、萝卜、油菜籽、稻米、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、南瓜小果、草莓、糖用甜菜、甘蔗、向日葵、甜玉米、枫香、甜薯、柳枝黍、茶、烟草、番茄、黑小麦、草皮草、西瓜和小麦植物。

[0018] 可通过以下方式产生转基因植物事件,特别是玉米、棉花和大豆转基因植物品种:将本发明的含有例如以SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11阐述的经适当修饰的多核苷酸序列的构建体引入此类植物细胞的基因组中,以及回收在基因组中包含用于表达至少本发明的蛋白质毒素(即,BCW 001、BCW 002或BCW 003蛋白质毒素)的遗传构建体的可繁殖转基因玉米、大豆或棉花植物。此类转基因植物将已向它们的植物基因组中引入了这样的多核苷酸构建体,所述多核苷酸构建体包含至少异源性启动子区段,所述异源性启动子区段可操作地连接到编码具有选自SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12组成的组的氨基酸序列的BCW 001、BCW 002或BCW 003杀昆虫蛋白或其杀昆虫蛋白质片段的核苷酸序列。

[0019] 种子也被构想为本发明的特征,其中种子从所述转基因植物产生,并且此类种子含有可检测量的引入到植物基因组中的多核苷酸构建体。从各个此类转基因植物产生的花粉、种子、子代植物细胞、植物组织和商品产品将含有可检测量的多核苷酸构建体。

[0020] 含有至少可检测量的编码此类BCW 001、BCW 002或BCW 003蛋白的多核苷酸构建体的任何生物样品都将在本发明的范围内。

[0021] 提供杀昆虫有效量的本发明的BCW 001、002或003蛋白的组合物被构想并且提供来防治鳞翅目害虫物种。此类组合物也可含有不同于BCW毒素蛋白的补充剂。此类药剂将如同BCW毒素蛋白一样也将对相同鳞翅目种具有毒性。补充剂将选自具有不同于BCW蛋白的氨基酸序列的蛋白质或多肽组成的药剂组,并且也可为赋予对靶标昆虫害虫的毒性效应的

RNA分子(诸如dsRNA、miRNA或siRNA)的药剂,并且也可为杀昆虫化合物(诸如除虫菊酯、有机磷酸酯杀虫剂),等等。或者,补充剂可为任何可相容的Cry或相关毒素蛋白质,诸如另一Cry1A、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae(但这些不是优选的,因为这些可能不赋予适当抗性管理性质)、或Cry1B、Cry1C、Cry1D、Cry1E、Cry1F、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab、Cry2Ae、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry43A、Cry43B、ET35、ET66、TIC400、TIC800、TIC807、TIC834、TIC853、TIC1415、VIP3A、VIP3Ab、Axmi杀昆虫蛋白、DIG杀昆虫蛋白、eHIP和VIP蛋白,以及本领域中已知会赋予对黑切根虫物种幼虫或可适用的其他鳞翅目靶标害虫物种的毒性性质的任何毒素蛋白质。

[0022] 此类组合物也可包括未必对相同靶标害虫具有毒性的额外杀虫剂,诸如选自由以下组成的组的额外药剂:Cry1C、Cry3A、Cry3B、Cry34、Cry35、Cry51Aa1、ET29、ET33、ET34、ET70、TIC407、TIC417、TIC431、TIC901、TIC1201、TIC3131、5307、DIG-10、Axmi184、Axmi205和AxmiR1。

[0023] 利用本发明的蛋白质的杀虫性质的用于产生种子的方法也被构思。所述方法包括被设计来表达BCW 001、BCW 002或BCW 003蛋白或展现与所述蛋白的至少约90%同一性的蛋白质的多核苷酸构建体,所述方法包括种植一个或多个含有表达一种或多种本发明的BCW蛋白毒素的多核苷酸的种子,从所述种子生长出植物,以及接着从所述植物收集许多此类种子。所收集种子将含有所述多核苷酸构建体,并且将产生也将对黑切根虫害虫侵扰具有抗性的植物。

[0024] 此类植物可为玉米、棉花、大豆或易感于被证明受本发明的蛋白质防治的鳞翅目害虫物种的任何其他植物。此类植物被构思为先前产生的将受益于本发明的蛋白质的毒性性质的效应的转基因植物。属于该种类的玉米植物包括但不限于选自由以下组成的组的转基因事件:DKB89614-9、MON801、MON802、MON809、MON810、MON863、MON88017、MON89034、事件4114-3、事件5307、DAS59122-7、Bt10、Bt11、Bt176、CBH-351、DKB-83614-9、MIR162、MIR604、TC1507、TC6275、事件676、事件678、事件680、事件98140、DAS40278-9、DKB89790-5、MON21-9、HCEM485、MON832、MON87427、NK603、T14、T25和VC001981-5。属于该种类的转基因植物的大豆植物选自由以下组成的组:MON87751、DAS81419-2、MON87701、A2704-12、A2704-21、A5547-127、A5547-35、CV127、DAS44406-6、DAS68416-4、DP356043、FG72、MON4032、ACS-GM003-1、MON87705、MON87708、MON89788、W62、W98和GFM Cry1A。属于该种类的棉花转基因植物选自由以下组成的组:DAS24236-5、DAS21023-5、事件31707、事件31803、事件31807、事件31808、事件42317、BNLA-601、COT102、COT67B、事件1、GHB119、GK12、MON15985、MLS9124、MON1076、MON531、MON757、T303-3、T304-40、SGK321、事件19-51a、GHB614、LLCotton25、MON88701、MON88702、MON1445、MON1698和MON88913。属于该种类的甘蔗转基因植物包括甘蔗植物转基因事件NXI-1T。将受益于具有编码此类BCW蛋白毒素的构建体的稻米植物在本领域中是已知的,包括选自由以下组成的组的稻米植物转基因事件:LLRICE06、LLRICE601、LLRICE62、GM-A17054和GM-A17054。

[0025] 此外,提供包含可检测量的所公开重组多核苷酸的经加工植物产品。所述经加工产品包括但不限于:植物生物质、油、粗粉、动物饲料、面粉、薄片、糠、棉绒、壳,和经加工种子。

[0026] 还提供制备转基因植物的方法。此类方法包括将重组多核苷酸引入植物细胞中,

以及选择表达昆虫抑制量的由所述重组多核苷酸编码的重组多肽的转基因植物。

[0027] 本发明的其他实施方案、特征和优势将根据以下具体实施方式、实施例和权利要求而显而易见。

附图说明

[0028] 图1示出在生物测定中用BCW 001、BCW 002和BCW 003毒素蛋白测试的鳞翅目昆虫害虫物种的清单。“+”指示相对于缓冲剂对照具有致死性；“-”指示未观察到超过缓冲剂对照的水平显著致死性；“ND”指示未使用可适用的毒素蛋白加以测试。BCW 001展现针对小地老虎(黑切根虫)、豆白缘切根虫(西方豆切根虫)、玉米穗虫(玉米耳虫)、玉米螟(欧洲玉米螟)、西南玉米秆草螟(西南玉米螟)、粉纹夜蛾(卷心菜尺蠖)、大豆夜蛾(大豆尺蠖)的致死性,并且在针对草地夜蛾(秋行军虫)测试时不展现致死性或生长阻碍。未用BCW 001测试小蔗秆草螟(蔗螟)。BCW 002展现针对小地老虎(黑切根虫)、豆白缘切根虫(西方豆切根虫)、玉米穗虫(玉米耳虫)和玉米螟(欧洲玉米螟)的致死性,并且在针对草地夜蛾(秋行军虫)测试时不展现致死性或生长阻碍。未用BCW 002测试小蔗秆草螟(蔗螟)、西南玉米秆草螟(西南玉米螟)、粉纹夜蛾(卷心菜尺蠖)和大豆夜蛾(大豆尺蠖)。BCW 003展现针对小地老虎(黑切根虫)、豆白缘切根虫(西方豆切根虫)、玉米穗虫(玉米耳虫)、玉米螟(欧洲玉米螟)、小蔗秆草螟(蔗螟)、西南玉米秆草螟(西南玉米螟)、粉纹夜蛾(卷心菜尺蠖)和大豆夜蛾(大豆尺蠖)的致死性,并且在针对草地夜蛾(秋行军虫)测试时不展现致死性或生长阻碍。

[0029] 图2示出BCW 001(SEQ ID NO:2,顶行)相对于BCW 002(SEQ ID NO:4,中行)相对于BCW 003(SEQ ID NO:6,底行)的氨基酸序列比对;在各三联行以下的星号代表在三个不同序列中的至少一者中在可适用的氨基酸位置处的差异。

[0030] 序列简述

[0031] SEQ ID NO:1是编码BCW 001鳞翅目昆虫毒性蛋白的天然苏云金芽孢杆菌菌株EG4384核苷酸序列。

[0032] SEQ ID NO:2是从如以SEQ ID NO:1阐述的开放阅读框推导的BCW 001氨基酸序列。

[0033] SEQ ID NO:3是编码嵌合BCW 002鳞翅目昆虫毒性蛋白的人工序列。

[0034] SEQ ID NO:4是从如以SEQ ID NO:3阐述的开放阅读框推导的BCW 002氨基酸序列,其中此类BCW 002蛋白由Cry1Ac的结构域I可操作地连接到BCW 001的结构域II和III(如以SEQ ID NO:2阐述的氨基酸位置258至氨基酸位置606),并且可操作地连接到如以SEQ ID NO:4和SEQ ID NO:6阐述的从氨基酸位置608至1177的Cry1Ac原毒素结构域所组成。SEQ ID NO:4(BCW 002)仅在位置259处不同于SEQ ID NO:6(BCW 003),BCW 002在该位置处含有异亮氨酸,BCW 003如同BCW 001中一样含有苏氨酸。

[0035] SEQ ID NO:5是编码嵌合BCW 003鳞翅目昆虫毒性蛋白的人工序列。

[0036] SEQ ID NO:6是从如以SEQ ID NO:3阐述的开放阅读框推导的BCW 003氨基酸序列,其中所述BCW 003蛋白由Cry1Ac的结构域I可操作地连接到BCW 001的结构域II和III(如以SEQ ID NO:2阐述的氨基酸位置258至氨基酸位置606),并且可操作地连接到如以SEQ ID NO:4和SEQ ID NO:6阐述的从氨基酸位置608至1177的Cry1Ac原毒素结构域所组成。SEQ

ID NO:6 (BCW 003) 仅在位置259处不同于SEQ ID NO:4 (BCW 002), BCW 002在该位置处含有异亮氨酸, BCW 003如同BCW 001中一样含有苏氨酸。

[0037] SEQ ID NO:7是用于在植物中表达的编码BCW 001蛋白的人工序列。

[0038] SEQ ID NO:8是由SEQ ID NO:7获得的BCW 001推导氨基酸序列。

[0039] SEQ ID NO:9是用于在植物中表达的编码BCW 002蛋白的人工序列。

[0040] SEQ ID NO:10是由SEQ ID NO:9获得的BCW 002推导氨基酸序列。

[0041] SEQ ID NO:11是用于在植物中表达的编码BCW 003蛋白的人工序列。

[0042] SEQ ID NO:12是由SEQ ID NO:11获得的BCW 003推导氨基酸序列。

具体实施方式

[0043] 通过将含有杀昆虫蛋白质的制剂喷雾于植物/种子上来防治作物植物的农业害虫的一替代方案是将编码这些蛋白质的多核苷酸插入植物基因组中以达成在植物或植物部分中表达。用这些多核苷酸转化的植物携带昆虫抗性, 所述昆虫抗性为这些表达的蛋白质作为转基因性状提供的。

[0044] 为避免发展出针对当前所用蛋白质的昆虫抗性或规避所述昆虫抗性, 需要具有不同作用模式 (MOA) 以及广泛谱系和功效的新型蛋白质来进行鳞翅目昆虫防治。一种用以解决该需要的方式是对Bt基因组测序以希望发现新型杀昆虫蛋白质。另一方法是将来自各种Bt蛋白质的区段互换以创建具有昆虫抑制性质的新型嵌合Bt蛋白质。由重新拣选本领域中已知的众多天然杀昆虫晶体蛋白质的结构域结构来任意创建的嵌合蛋白具有增强性质的可能性是微小的 (参见例如A Strategy for Shuffling Numerous Bacillus thuringiensis Crystal Protein Domains; J. Economic Entomology, 97 (6): 1805-1813. 2004)。

[0045] 本文公开编码在本文中标识为BCW蛋白的杀昆虫蛋白质的核苷酸序列, 所述杀昆虫蛋白质解决对替代性MOA的需要, 提供针对更广泛谱系的昆虫害虫的活性, 并且起用于延迟或避免抗性发展, 特别是用于防治黑切根虫 (BCW) 害虫的作用。

[0046] 在对苏云金芽孢杆菌菌株EG4384的基因组测序之后, BCW 001作为预测氨基酸序列具有Cry1A型蛋白的特征的开放阅读框而被发现。BCW 001开放阅读框 (ORF) 编码具有1180个氨基酸的蛋白质, 并且所述蛋白质被预测具有Cry1蛋白毒素的许多特征, 包括可鉴定结构域I、II和III结构以及在所预测蛋白质的羧基末端半部处的特征性Cry1A型原毒素结构域。所预测结构域I氨基酸序列 (SEQ ID NO:2的残基1至约258) 展现与Cry1Ac蛋白毒素结构域I的约67%同一性。所预测结构域II氨基酸序列 (如以SEQ ID NO:2阐述的从约259至约残基459的残基) 展现与Cry1Ai2结构域II的完全 (100%) 同一性。所预测结构域III氨基酸序列 (如以SEQ ID NO:2阐述的从约260至约606的残基) 展现与Cry1Ah2中的相应结构域III残基的约63%同一性。BCW 001所预测蛋白的原毒素结构域结构 (如以SEQ ID NO:2阐述的约残基607至1180) 展现与Cry1Aa9中的相应残基的约96%同一性。总的说来, 该所预测全长蛋白质展现与Cry1Ai的约83%氨基酸序列同一性, 并且如以SEQ ID NO:2阐述的从氨基酸位置1至约残基607的所预测毒素区域展现与Cry1Ai1的76%同一性。难以将该新型毒素蛋白质指定为特定Cry1A类别, 因此将向苏云金芽孢杆菌命名委员会提供该序列, 并且所述委员会将确定该蛋白质是否应得到它自身的单独和新型类别。

[0047] BCW 001蛋白在苏云金芽孢杆菌的无晶体菌株中从质粒载体产生,并且针对多种鳞翅目害虫来测试芽孢晶体制备物。对于数据,参见图1第2列。证据指示该蛋白质不具有它从其得出它的起源的任何蛋白质即本领域中已知的Cry1Aa、Cry1Ah或Cry1Ai毒素蛋白的特征。当针对黑切根虫加以测试时,先前技术蛋白质都不展现任何可观活性,然而,当针对黑切根虫加以测试时,该新型蛋白质BCW 001在生物测定中具有毒性,并且当针对一连串其他鳞翅目害虫加以测试时以及在图1中所阐述,惊人地展现毒性性质。

[0048] 具体地,在生物测定中,相较于未处理昆虫膳食对照,BCW 001蛋白展现针对西方豆切根虫(“WBC”,豆白缘切根虫)、玉米耳虫(“CEW”,玉米穗虫)、欧洲玉米螟(“ECB”,玉米螟)、西南玉米螟(“SWC”,西南玉米秆草螟)、大豆尺蠖(“SL”,大豆夜蛾)、卷心菜尺蠖(“CLW”,粉纹夜蛾)以及第1龄期和第3龄期黑切根虫(“BCW”,小地老虎)的活性。

[0049] 如以下进一步所述,使用Cry1Ac的结构域I和BCW 001的结构域II和III产生嵌合毒素蛋白质(即分别如以SEQ ID NO:4和SEQ ID NO:6阐述的BCW 002和BCW 003),并且将这些嵌合毒素蛋白质引入玉米和甘蔗植物中。两种嵌合蛋白均在玉米中展现针对BCW、WBC、CEW和SWC的活性。BCW 003在甘蔗中展现针对SCB的活性。

[0050] 如本文所用的短语“BCW蛋白”是指赋予针对特别是但不限于BCW、WBC和/或SCB的鳞翅目昆虫的活性的包含BCW 001(SEQ ID NO:2)、BCW 002(SEQ ID NO:4)和BCW 003(SEQ ID NO:6),由BCW001(SEQ ID NO:2)、BCW 002(SEQ ID NO:4)和BCW 003(SEQ ID NO:6)组成,与BCW 001(SEQ ID NO:2)、BCW 002(SEQ ID NO:4)和BCW 003(SEQ ID NO:6)大致上同源,或源于BCW 001(SEQ ID NO:2)、BCW 002(SEQ ID NO:4)和BCW 003(SEQ ID NO:6)的任何昆虫抑制性多肽序列的任何新型昆虫抑制性蛋白质,及其昆虫抑制性区段或其组合。编码BCW 001的多核苷酸源于菌株EG4384。BCW 001的核心毒性氨基酸序列对应于如以SEQ ID NO:2阐述的从约位置28至约位置606以及直至位置618的氨基酸,并且BCW 002和003的核心毒性氨基酸序列分别对应于如以SEQ ID NO:4和SEQ ID NO:6阐述的从约位置29至约位置607以及直至约位置619的氨基酸。

[0051] 在一个实施方案中,本文公开的蛋白质由于主要 δ -内毒素结构、长度(约1176-1180个氨基酸)、蛋白质的不具有原毒素的长度(从约600至约619个氨基酸)、毒性核心的长度(约591个氨基酸)、或存在至少一个BCW特异性区段而相关。

[0052] 利用以下这些缺省参数,使用Clustal W算法将示例性蛋白质加以彼此比对,从而产生各对的成对氨基酸数目同一性和成对氨基酸同一性百分比:权重矩阵:blosum;空位开放罚分:10.0;空位延伸罚分:0.05;亲水性空位:开启;亲水性残基:GPSNDQERK;残基特异性空位罚分:开启。Clustal W算法描述于Thompson,J.D.,Higgins,D.G.和Gibson,T.J.(1994)“CLUSTAL W:improving the sensitivity of progressive multiple sequence alignment through sequence weighting,positions-specific gap penalties and weight matrix choice.”Nucleic Acids Research,22:4673-4680中。

[0053] 其他比对算法在本领域中也是可用的,并且提供与使用Clustal W比对获得的结果类似的结果。

[0054] 术语“约”在本文中用于描述这些区段边界可变化1、2、3、4、5、10、20、25、30或35个残基,此取决于亲本蛋白质的序列和它们的彼此比对。为进一步描述各种区段的可变性和构型,表2和表3对BCW 002和003以及其他对黑切根虫具有活性的嵌合体的区段边界进行列

表。

[0055] 术语“片段”在本文中用于描述相比于描述BCW毒性蛋白的完整氨基酸或核酸序列,较短的连续氨基酸或核酸序列。

[0056] 短语“昆虫抑制性”和“杀昆虫性”在本文中可互换使用,并且是指蛋白质、蛋白质片段、蛋白质区段或多核苷酸产生对昆虫活力、生长、昆虫发育、昆虫繁殖、昆虫摄食行为、昆虫交配行为的任何可测量抑制作用和/或导致由摄食该蛋白质、蛋白质片段、蛋白质区段或多核苷酸的昆虫引起的不利作用的任何可测量降低。

[0057] 术语“生物活性”、“活性的”、“活性”、“有效的”、“有效力的”或其变化形式在本文中可互换用于描述本发明的蛋白质对靶标昆虫害虫的作用。

[0058] 作物是在它的生长阶段的某一点收集其产物的自生植物或栽培植物。所述产物的非限制性实例是种子、圆莢、叶、花、茎、根或其任何部分。

[0059] 从包含如本文例示的多核苷酸或表达如本文例示的蛋白质的任何植物组织、细菌、病毒或载体获得生物样品,诸如但不限于种子、圆莢、叶、花、茎、根或其任何部分,并且含有可检测量的所述多核苷酸、蛋白质或两者。

[0060] 短语“可检测量”在本文中用于描述本文公开的蛋白质或多核苷酸的可通过标准分析方法来检测的最小量,所述分析方法诸如但不限于聚合酶链式反应(PCR)和酶联免疫吸附测定(ELISA)技术等。

[0061] 在一个实施方案中,本文所述的毒素蛋白质由于共同功能而相关,并且展现对于鳞翅目昆虫物种的杀昆虫活性。

[0062] BCW 001区段对含有所述区段的Cry1嵌合体给予黑切根虫活性。给予黑切根虫活性的BCW 001区段的实例以SEQ ID NO:2阐述,从约氨基酸位置250至约606,并且更特别是从约氨基酸位置255至约606。含有对应于BCW 001毒素蛋白的结构域II和III的该氨基酸区段的Cry1嵌合体将经常也对嵌合蛋白赋予与防治黑切根虫相关的毒性性质,并且这已在Cry1A、Cry1B、Cry1C、Cry1D、Cry1E和Cry1F骨架内加以测试,其中可适用的毒素构建体已使它的结构域II和III组分被来自BCW 001的该氨基酸范围取代,并且在许多情况下,黑切根虫活性在嵌合构建体中得以惊人地维持(数据未显示)。

[0063] 在本发明的一方面,具体来说,由可适用的BCW毒素蛋白防治的害虫是来自鳞翅目的昆虫害虫,包括成虫、蛹、幼虫和新生体。

[0064] 鳞翅目的昆虫包括但不限于夜蛾科(Noctuidae)中的行军虫、切根虫、尺蠖和棉铃虫(例如秋行军虫(草地夜蛾)、甜菜行军虫(甜菜夜蛾)、披肩行军虫(蓓带夜蛾(Mamestra configurata))、黑切根虫(小地老虎)、卷心菜尺蠖(粉纹夜蛾)、大豆尺蠖(大豆夜蛾(Pseudoplusia includens))、黎豆毛虫(黎豆夜蛾)、绿三叶草虫(苜蓿绿夜蛾(Hypena scabra))、烟草蚜虫(烟芽夜蛾)、颗粒切根虫(粒肤地老虎(Agrotis subterranea))、西方切根虫(西方灰地老虎(Agrotis orthogonia))、行军虫(一星粘虫(Pseudaletia unipuncta));来自螟蛾科(Pyralidae)的螟虫、鞘蛾幼虫、结网虫、锥虫、卷心菜虫和雕叶虫(例如欧洲玉米螟(玉米螟)、脐橙虫(脐橙螟(Amyelois transitella))、玉米根结网虫(玉米根草螟(Crambus caliginosellus))、草地结网虫(稻切叶野螟(Herpetogramma licarsialis))、向日葵蛾(向日葵螟(Homoeosoma electellum))、小玉米秆螟(南美玉米苗斑螟(Elasmopalpus lignosellus));卷蛾科(Tortricidae)中的卷叶虫、蚜虫、种子蠕

虫和果实蠕虫(例如苹果蠹蛾(苹果蠹蛾(*Cydia pomonella*))、葡萄浆果蛾(葡萄小食心虫(*Endopiza viteana*))、东方果蛾(梨小食心虫(*Grapholita molesta*))、向日葵芽蛾(向日葵芽卷叶蛾(*Suleima helianthana*))) ;以及许多其他在经济上重要的鳞翅目昆虫(例如菱纹背蛾(小菜蛾)、粉螟蛉虫(棉红铃虫)、吉卜赛蛾(舞毒蛾(*Lymantria dispar*)))。鳞翅目的其他昆虫害虫包括例如棉叶波纹夜蛾(*Alabama argillacea*) (棉叶蠕虫)、果树黄卷蛾(*Archips argyrospila*) (果树卷叶虫)、玫瑰黄卷蛾(*A. rosana*) (欧洲卷叶虫)和其他黄卷蛾属种、二化螟(*Chilo suppressalis*) (亚洲稻螟或稻茎螟)、稻纵卷叶螟(*Cnaphalocrocis medinalis*) (稻卷叶虫)、玉米根草螟(玉米根结网虫)、早熟禾草螟(*C. teterrellus*) (早熟禾结网虫)、西南玉米秆草螟(西南玉米螟)、小蔗秆草螟(*D. saccharalis*) (甘蔗螟)、埃及钻夜蛾(*Earias insulana*) (多刺螟蛉虫)、翠纹钻夜蛾(*E. vittella*) (斑点螟蛉虫)、棉铃虫(美洲螟蛉虫)、玉米穗虫(*H. zea*) (玉米耳虫或棉花螟蛉虫)、烟芽夜蛾(烟草蚜虫)、稻切叶野螟(草地结网虫)、葡萄花翅小卷蛾(*Lobesia botrana*) (欧洲葡萄藤蛾)、棉红铃虫(粉螟蛉虫)、柑桔潜叶蛾(*Phyllocnistis citrella*) (柑桔潜叶虫)、欧洲粉蝶(*Pieris brassicae*) (大白蝴蝶)、菜粉蝶(*P. rapae*) (外来卷心菜虫或小白蝴蝶)、小菜蛾(菱纹背蛾)、甜菜夜蛾(甜菜行军虫)、斜纹夜蛾(*S. litura*) (烟草切根虫、群集毛虫)、草地夜蛾(*S. frugiperda*) (秋行军虫)和番茄斑潜蝇(*Tuta absoluta*) (番茄潜叶虫)。

[0065] 本文公开的蛋白质也可用于产生特异性结合BCW特异性毒素蛋白的抗体,并且可用于筛选和发现BCW毒素种类的其他成员。

[0066] 在一个实施方案中,编码昆虫抑制性BCW 001相关蛋白的示例性多核苷酸以SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11阐述。编码这些蛋白质的核苷酸序列可作为筛选探针和引物用于使用热扩增或等温扩增和/或杂交方法以及为本领域普通技术人员所知的其他鉴定方法来鉴定所述种类的其他成员。

[0067] 本发明的一方面提供用于发现相关蛋白质的方法,并且所述方法包括对Bt基因组测序,汇编序列数据,鉴定和克隆编码所述杀害虫蛋白质的Bt基因,以及表达和测试新型Bt蛋白质以测定杀害虫活性。本发明的另一方面采用分子方法来工程改造和克隆商业适用蛋白质,其包括来自杀害虫蛋白质种类的蛋白质的嵌合体,例如嵌合体可由BCW毒性蛋白的区段装配以获得额外实施方案。可使所公开蛋白质经受彼此比对以及与其他Bt杀害虫蛋白质的比对,并且可鉴定各所述蛋白质的适用于在所比对蛋白质之间进行取代,从而导致构建嵌合蛋白的区段。可使所述嵌合蛋白经受害虫生物测定分析,并且对相较于嵌合体中各所述区段所源于的亲本蛋白质,生物活性增加和/或靶标害虫谱扩大的存在性进行表征。可通过与其他蛋白质调换结构域或区段来进一步工程改造多肽的杀害虫活性以获得对特定害虫或对更广泛谱系的害虫的活性。

[0068] 在一个实施方案中,本文公开的蛋白质包括本文公开的蛋白质的功能等效片段(N末端或C末端缺失)。

[0069] 本文提供BCW毒性蛋白。在某些实施方案中,BCW 001相关毒素蛋白可被分离,可被提供在组合物、转基因微生物或转基因植物中。在该实施方案中,BCW 002蛋白以及特别是BCW 003蛋白赋予鳞翅目昆虫抑制活性,特别是针对黑切根虫和/或蔗螟的抑制活性。在本申请中提及“经分离蛋白质”或等效术语或短语意指所述蛋白质是单独与其他组成成分组合存在,但不在它的天然环境内的蛋白质。举例来说,只要本发明的毒素蛋白质在这些毒

素蛋白质所天然见于其中的生物体内,那么天然见于所述生物体内的本发明的毒素蛋白质等不被视为是“经分离的”。然而,只要蛋白质不在它所天然见于其中的生物体内,那么在本公开的范围内,这些毒素蛋白质各自将是“经分离的”。

[0070] 如本文所用的“可操作地连接”是指使核酸序列或氨基酸序列接合以使一个序列可对连接的序列提供所需功能或可相容或适用特征。

[0071] 与BCW 001、BCW 002和BCW 003具有生物功能等效性的肽、多肽和蛋白质包括但不限于含有这些BCW毒素蛋白序列中的保守性氨基酸取代的氨基酸序列。在所述氨基酸序列中,序列中的一个或多个氨基酸被另外氨基酸取代,从而导致沉默或保守性氨基酸序列变化。

[0072] 尽管所公开的昆虫抑制性多肽优选包含BCW 001、BCW 002或BCW 003蛋白质序列,但本文也公开该序列的具有与该昆虫抑制性蛋白质的昆虫抑制活性相同或类似的昆虫抑制活性的片段和变体。举例来说,具有至少30、35、38、40、45、50、55、60、65、70、75、100、150、200、500、550个或更多个处于BCW相关毒素蛋白中的氨基酸的连续序列具有昆虫抑制活性。在另一实施方案中,BCW相关毒素蛋白的具有昆虫抑制活性的片段可包含BCW毒素蛋白序列中的氨基酸取代、缺失、插入或添加。

[0073] 在一个实施方案中,昆虫抑制性多肽包含如以SEQ ID NO:2阐述的BCW 001蛋白质序列的从约残基28至约残基618的昆虫抑制性区段。非限制性实例包括以下中的任一者:SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:6或独自或可操作连接于嵌合蛋白中的具有与该特定BCW 001蛋白的昆虫抑制活性相同或类似的昆虫抑制活性的较短片段或变体。在另一实施方案中,具有含至少约38个或更多个处于SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:6中的任一者中的氨基酸的连续氨基酸序列的区段也提供功能性杀昆虫蛋白质。昆虫抑制性BCW001毒性片段也可包含具有含591个氨基酸的区域的至少30、35、38、40、45、50、100、150、200、500、550、555、560、565、570、572、574、580或585个氨基酸残基的区段,所述区域对应于具有SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:6中的任一者的序列的约残基28至约618。

[0074] 在一些实施方案中,成熟BCW 001蛋白(成熟意指是1180个氨基酸的原毒素形式的蛋白质在昆虫害虫消化道中通过蛋白水解而被裂解以释放就在残基607至约残基618的N末端的核心毒素,从而释放或多或少地包含如以SEQ ID NO:2阐述的残基1至残基606或从约5至约618的任何数目的残基的活性毒素区段,只要所释放区段展现对黑切根虫幼虫的毒性性质即可)的片段可为具有昆虫抑制活性的截短形式,其中一个或多个氨基酸从蛋白质的N末端、C末端、中间或其组合缺失。这些片段可为BCW 001的天然存在或合成变体,并且保留BCW毒素蛋白的昆虫抑制活性。在某些实施方案中,成熟BCW 001、BCW 002或BCW 003蛋白的片段展现由它们所源于的起始蛋白质分子具有的杀害虫活性。本文所述的片段或变体可进一步包含本文鉴定的导致蛋白质的杀害虫活性的结构域。具有昆虫抑制活性的截短衍生物是对应于如以SEQ ID NO:2阐述的BCW 001毒素蛋白质序列的从约28至约606或至约618的残基,或对应于如以SEQ ID NO:4或SEQ ID NO:6阐述的BCW毒素蛋白的从约29至约607以及至残基619的残基的BCW毒素蛋白。

[0075] 在另一实施方案中,截短N末端缺失突变形式包括但不限于缺乏来自BCW毒素蛋白的不具有原毒素的毒素部分或毒性核心的N末端和/或C末端的氨基酸残基的BCW 001毒性蛋白。举例来说,可使BCW 001蛋白的对应于SEQ ID NO:2的残基28至618或对应于SEQ ID

NO:4或SEQ ID NO:6的残基29至619的毒性核心的1至6个N末端氨基酸残基缺失。对应于SEQ ID NO:2的残基28至618或SEQ ID NO:4或SEQ ID NO:6的残基29至619的BCW毒素蛋白的截短C末端缺失突变形式包括但不限于缺乏1至6个C末端氨基酸残基的BCW毒素蛋白。在其他实施方案中,具有SEQ ID NO:2的相应残基28至618或SEQ ID NO:4或SEQ ID NO:6的相应残基29至619的BCW毒素蛋白可具有N末端1至6个氨基末端残基截短与C末端1至6个羧基末端残基截短两者。

[0076] 在一些实施方案中,CPR24719蛋白的对不同于CPR24719-1的蛋白质给予黑切根虫活性的个别区段1至6或区段1至6的组合也可展现相同或类似功能。

[0077] 相对于成熟BCW毒素蛋白的具有以SEQ ID NO:2的残基28至618显示或对应于SEQ ID NO:4或SEQ ID NO:6的残基29至619的相应氨基酸序列的相应区段,本文公开的BCW毒素蛋白的片段和变体可具有约62%或更大序列同一性、约65%、70%、75%、80%、85%、90%、92%、95%、96%、97%、98%或更大序列同一性、以及约99%、99.5%、100%氨基酸序列同一性。

[0078] 本发明的一实施方案包括编码BCW毒素蛋白的重组多核苷酸组合物。举例来说,BCW毒素蛋白可用重组DNA构建体表达,其中具有编码蛋白质的开放阅读框的经分离多核苷酸分子可操作地连接到各种元件,诸如启动子和对于在构建体所意图用于的系统中的表达具有功能性的任何其他调控元件。在本申请中提及“经分离DNA分子”或等效术语或短语意指所述DNA分子是单独与其他组成成分组合存在,但不在它的天然环境内的DNA分子。举例来说,只要元件在生物体的基因组内以及在其天然见于其中的基因组内的位置处,那么天然见于所述生物体的基因组的DNA内的核酸元件诸如编码序列、内含子序列、未翻译前导序列、启动子序列、转录终止序列等不被视为是“经分离的”。然而,只要元件不在生物体的基因组内以及不在其所天然见于其中的基因组内的位置处,那么在本公开的范围内,这些元件各自以及这些元件的子部分将是“经分离的”。类似地,编码杀昆虫蛋白质或那个蛋白质的任何天然存在的杀昆虫变体的核苷酸序列将是经分离核苷酸序列,只要所述核苷酸序列不在编码所述蛋白质的所述序列所天然见于的细菌的DNA内即可。出于本公开的目的,编码天然存在的杀昆虫蛋白质的氨基酸序列的合成核苷酸序列将被视为是经分离的。出于本公开的目的,任何转基因核苷酸序列,即插入植物或细菌的细胞的基因组中或存在于染色体外载体中的DNA的核苷酸序列,将被视为是经分离核苷酸序列,无论它存在于用于转化细胞的质粒或类似结构内,植物或细菌的基因组内,抑或以可检测量存在于源于植物或细菌的组织、子代、生物样品或商品产品中。非限制性实例包括植物功能性启动子可操作地连接到BCW毒素蛋白编码序列以在植物中表达蛋白质,或Bt功能性启动子可操作地连接到BCW毒素蛋白编码序列以在Bt中表达蛋白质。可被可操作地连接到BCW毒素蛋白编码序列的其他元件包括但不限于增强子、内含子、前导物、所编码蛋白质固定标签(HIS标签)、所编码亚细胞易位肽(例如质体转运肽、信号肽)、翻译后修饰酶的所编码多肽位点、核糖体结合位点和RNAi靶标位点。

[0079] 如本文所用,“重组DNA分子”是包含在无人干预下将不一起天然存在的DNA分子的组合的DNA分子。举例来说,重组DNA分子可为包含至少两个相对于彼此是异源性的DNA分子的DNA分子、包含与在自然界中存在的DNA序列不同的DNA序列的DNA分子、或已通过遗传转化或基因编辑来并入宿主细胞的DNA中的DNA分子。类似地,“重组蛋白质分子”是包含在

无人为干预下将不一起天然存在的氨基酸的组合的蛋白质分子。举例来说,重组蛋白质分子可为包含至少两个相对于彼此是异源性的氨基酸分子的蛋白质分子、包含与在自然界中存在的氨基酸序列不同的氨基酸序列的蛋白质分子、或由于对宿主细胞进行遗传转化或通过对宿主细胞基因组进行基因编辑而在宿主细胞中表达的蛋白质分子。

[0080] 与此一起提供的示例性重组多核苷酸分子包括但不限于SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11,以及以SEQ ID NO:3和SEQ ID NO:5阐述的核苷酸区段,和编码具有如以SEQ ID NO:2(BCW001)、SEQ ID NO:4(BCW 002)、SEQ ID NO:6(BCW 003)以及SEQ ID NO:8、SEQ ID NO:10和SEQ ID NO:12阐述的氨基酸序列的相应多肽或蛋白质的核苷酸区段中的各者。编码本文公开的蛋白质的重组多核苷酸分子的密码子可被同义密码子取代(也被称为沉默取代)。也提供编码任何本文公开的BCW毒素变异蛋白的重组多核苷酸。

[0081] 包含BCW毒素蛋白编码序列的重组DNA构建体也可进一步包含编码一种或多种昆虫抑制剂的DNA区域,其可被配置来与编码BCW毒素蛋白、不同于BCW毒素蛋白的蛋白质、昆虫抑制性dsRNA分子或杀昆虫化合物的DNA序列并行表达或共表达。杀昆虫化合物的非限制性实例是有机氯化物、有机磷酸酯和氨基甲酸酯、拟除虫菊酯(pyrethroid)、新烟碱类(neonicotinoid)和莱恩碱类(ryanoid)。

[0082] 可装配重组DNA构建体以使所有蛋白质或dsRNA分子都由一个启动子表达,或各蛋白质或dsRNA分子处于单独启动子控制下,或处于其某一组合的情况下。本文公开的蛋白质可由多基因表达系统表达,其中一种或多种本文公开的蛋白质由共同核苷酸区段表达,在所述区段上视所选表达系统的类型而定也含有其他开放阅读框和/或启动子。举例来说,细菌多基因表达系统可利用单一启动子来驱动来自单一操纵元内的多重连接/串联开放阅读框的表达。在另一实例中,植物多基因表达系统可利用多重连接表达盒,各自表达不同蛋白质或其他药剂诸如一种或多种dsRNA分子。在另一实例中,植物多基因表达系统可利用未多重连接表达盒,各自表达不同蛋白质或其他药剂诸如一种或多种dsRNA分子。用于本文所述的重组核酸中的启动子可包含完整启动子序列或其具有启动子活性或基因调控活性的任何变体或片段。

[0083] 包含BCW毒素蛋白编码序列的重组多核苷酸或重组DNA构建体可通过载体例如质粒、杆状病毒、人工染色体、病毒粒子、粘粒、噬菌粒、噬菌体或病毒载体来递送至宿主细胞中。所述载体可用于实现在宿主细胞中稳定或短暂表达BCW毒素蛋白编码序列,或随后表达成多肽。包含BCW毒素蛋白编码序列,并且引入宿主细胞中的外源性重组多核苷酸或重组DNA构建体在本文中被称作“转基因”。

[0084] 本文也提供含有表达任何一种或多种本文提供的BCW毒素蛋白编码序列的任何重组多核苷酸的转基因细菌、转基因植物细胞、转基因植物、真菌和酵母以及转基因植物部分。术语“细菌细胞”或“细菌”可包括但不限于土壤杆菌属(*Agrobacterium*)、芽孢杆菌属(*Bacillus*)、埃希氏菌属(*Escherichia*)、沙门氏菌属(*Salmonella*)、假单胞菌属(*Pseudomonas*)或根瘤菌属(*Rhizobium*)细胞。术语“植物细胞”或“植物”可包括但不限于苜蓿、香蕉、大麦、菜豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、大白菜、柑桔、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、花旗松、茄子、桉树、亚麻、大蒜、葡萄、蛇麻子、韭菜、莴苣、火炬松、粟、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧场草、豌豆、花生、胡椒、木豆、松树、马铃薯、白杨、南瓜、辐射松、萝卜、油菜籽、稻米、砧木、黑麦、红

花、灌木、高粱、南方松、大豆、菠菜、南瓜小果、草莓、糖用甜菜、甘蔗、向日葵、甜玉米、枫香、甜薯、柳枝黍、茶、烟草、番茄、黑小麦、草皮草、西瓜和小麦植物细胞或植物。在某些实施方案中,提供由转基因植物细胞再生的转基因植物和转基因植物部分。在某些实施方案中,转基因植物可通过繁殖、切割、折断、研磨或另外解离来自植物的部分来从转基因种子获得。在某些实施方案中,植物部分可为种子、圆荚、叶、花、茎、根或其任何部分或转基因植物部分的非可再生部分。如在该情形下所用,转基因植物部分的“非可再生”部分是不能被诱导以形成完整植物,或不能被诱导以形成能够有性和/或无性繁殖的完整植物的部分。在某些实施方案中,植物部分的非可再生部分是转基因种子、圆荚、叶、花、茎或根的部分。

[0085] 本文还提供制备包含昆虫或鳞翅目昆虫抑制量的BCW毒素蛋白的转基因植物的方法。所述植物可通过以下方式来制备:将编码任何本文提供的BCW毒素蛋白的重组多核苷酸引入植物细胞中,以及选择源于所述植物细胞的表达昆虫或鳞翅目昆虫抑制量的BCW毒素蛋白的植物。植物可通过再生、播种、授粉或分生组织转化技术来由植物细胞获得。

[0086] 用于对植物进行转化的方法在本领域中是已知的。举例来说,土壤杆菌介导的转化描述于美国专利申请公布2009/0138985A1(大豆)、2008/0280361A1(大豆)、2009/0142837A1(玉米)、2008/0282432(棉花)、2008/0256667(棉花)、2003/0110531(小麦)、2001/0042257A1(糖用甜菜)、美国专利5,750,871(加拿大油菜)、7,026,528(小麦)和6,365,807(稻米)中以及Arencebia等(1998)Transgenic Res.7:213-222(甘蔗)中。

[0087] 本文还提供表达昆虫或鳞翅目昆虫抑制量的BCW毒素蛋白的转基因植物用以防治昆虫或鳞翅目昆虫侵扰的用途。任何以上提及的转基因植物都可用于本文提供的用于保护植物免遭昆虫或鳞翅目昆虫侵扰的方法中。获得表达对鳞翅目昆虫具有活性的蛋白质诸如Cry1A蛋白(美国专利号5,880,275)、Cry1B(美国专利申请号10/525318)、Cry1C(美国专利号6,033,874)、Cry1F、Cry1A/F嵌合体(美国专利号7,070,982;6,962,705和6,713,063)和Cry2Ab蛋白(美国专利号7,064,249)的转基因植物的方法是充分表征的。

[0088] 本文还提供任何以上提及的转基因宿主细胞用以产生BCW毒素蛋白的用途。

[0089] 本发明的额外方面包括用于检测编码BCW毒素蛋白或其特性片段和区段的多核苷酸的抗体和方法、用于鉴定所述蛋白质种类的额外昆虫抑制性成员的方法、用于防治昆虫生长和/或侵扰的制剂和方法、以及用于对植物和其他接受宿主提供所述防治的方法。

[0090] 在某些实施方案中,植物产品可包括源于转基因植物或转基因植物部分的商品或其他贸易产品,其中商品或其他产品可通过检测编码或包含BCW毒素蛋白的特性部分的核苷酸区段或所表达RNA或蛋白质来在贸易全程加以追踪。所述商品或其他贸易产品包括但不限于植物部分、生物质、油、粗粉、糖、动物饲料、面粉、薄片、糠、棉绒、经加工种子和种子。

[0091] 还与此一起提供的是经加工植物产品,其中所述经加工产品包含可检测量的编码BCW毒素蛋白、其区段、其昆虫抑制性片段或其任何特性部分的重组多核苷酸。在某些实施方案中,经加工产品选自由以下组成的组:植物生物质、油、粗粉、动物饲料、面粉、薄片、糠、棉绒、壳和经加工种子。在某些实施方案中,经加工产品是非可再生的。

[0092] 本文还提供防治昆虫的方法。在某些实施方案中,防治作物植物的鳞翅目昆虫侵扰。所述方法可包括使包含昆虫或鳞翅目昆虫抑制量的BCW毒素蛋白的植物生长。在某些实施方案中,所述方法可进一步包括以下中的任何一者或多者:(i)将包含或编码BCW毒素蛋白的任何组合物施加于植物或产生植物的种子;和/或(ii)用编码BCW毒素蛋白的多核苷酸

转化植物或产生植物的植物细胞。在某些实施方案中,植物是经短暂或稳定转化的转基因植物,其包含表达昆虫或鳞翅目昆虫抑制量的BCW毒素蛋白的转基因。在某些实施方案中,植物是已向其施加包含BCW毒素蛋白的组合物的非转基因植物。在所述方法的某些实施方案中,植物是玉米或甘蔗植物。在某些实施方案中,鳞翅目种是小地老虎。在某些实施方案中,鳞翅目种是小蔗秆草螟。在某些实施方案中,鳞翅目种在作物田地中。

[0093] 在植物中或通过某一方法使本文公开的蛋白质富集可包括在表达/产生重组多肽/蛋白质的条件下培养重组Bt细胞。这种方法可包括通过干燥、冻干、均质化、提取、过滤、离心、沉降或浓缩表达/产生所述重组多肽的重组Bt细胞的培养物进行制备。这种方法可产生Bt细胞提取物、细胞混悬液、细胞匀浆、细胞溶解产物、细胞上清液、细胞滤液或细胞集结块。通过获得如此产生的重组多肽/蛋白质,包括重组多肽/蛋白质的组合物可包括细菌细胞、细菌芽孢和伴胞包涵体,并且可被配制以用于各种用途,包括作为农业昆虫抑制性喷雾产品或作为膳食生物测定中的昆虫抑制性制剂。

[0094] 在一个实施方案中,包含所公开重组多肽/蛋白质的昆虫抑制性组合物/制剂可进一步包含至少一种展现针对相同鳞翅目昆虫物种的昆虫抑制活性,但不同于重组多肽的额外多肽,以提供鳞翅目昆虫对BCW毒素蛋白或其他鳞翅目昆虫抑制性组合物的抗性的发生降低。所述多肽选自由以下组成的组:昆虫抑制性蛋白质、昆虫抑制性dsRNA分子和化合物。使用所述核糖核苷酸序列来防治昆虫害虫的一个实例描述于美国专利申请公布号2006/0021087中。其他所述组合物的实例包括但不限于Cry1A(美国专利号5,880,275)、Cry1Ab、Cry1Ac、Cry1Ae、Cry1B(美国专利申请号10/525,318)、Cry1C(美国专利号6,033,874)、Cry1E、Cry1F以及Cry1A/F嵌合体(美国专利号7,070,982;6,962,705;和6,713,063)、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab(美国专利号7,064,249)、Cry2Ae、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry43A、Cry43B、ET35、ET66、TIC400、TIC800、TIC807、TIC834、TIC853和TIC1415。其他非限制性实例是可与本文公开的蛋白质组合的对鳞翅目昆虫具有活性的蛋白质VIP、Axmi和DIG,诸如但不限于Vip3A、VIP3Ab、AXMI-184、AXMI-196、DIG-3、DIG-4、DIG-5和DIG-11。

[0095] 在其他实施方案中,所述组合物可进一步包含至少一种展现对不由另外昆虫抑制性BCW毒素蛋白抑制的昆虫的昆虫抑制活性的额外多肽以使所得昆虫抑制谱扩大。举例来说,对于防治鞘翅目(Coleopteran)害虫,昆虫抑制性BCW毒素蛋白的组合可与诸如但不限于以下的对鞘翅目昆虫具有活性的蛋白质一起使用:Cry1C变体、Cry3A变体、Cry3Bb(美国专利号6,501,009)、Cry34/35、5307、Axmi184、Axmi205、AxmiR1、TIC407、TIC417、TIC431、TIC901、TIC1201、TIC3131、DIG-10和eHIP(美国专利申请公布号2010/0017914)。

[0096] 在本领域中已用文件记载昆虫发展出对某些杀昆虫剂的抗性的可能性。一种昆虫抗性管理策略在于采用表达两种通过不同作用模式来起作用的不同昆虫抑制剂的转基因作物。因此,对任一昆虫抑制剂具有抗性的任何昆虫都可由另一昆虫抑制剂防治。另一昆虫抗性管理策略采用对未针对鳞翅目害虫物种加以保护的植物的使用以产生庇护。一个特定实例描述于美国专利号6,551,962中。

[0097] 本文公开的其他实施方案包括被设计用于防治也由本文公开的蛋白质防治的害虫的待与蛋白质一起用于种子处理的经表面施加的杀害虫化合物,喷洒制剂、滴注制剂或擦拭制剂可直接施加于土壤(土壤浸润),施加于表达本文公开的蛋白质的生长植物,或被

配制以施加于含有一种或多种编码一种或多种所公开蛋白质的转基因的种子。用于种子处理的所述制剂可与本领域中已知的各种粘着剂和增粘剂一起施加。所述制剂可含有在MOA方面与所公开蛋白质协同的杀虫剂,以使配制杀虫剂通过不同MOA来起防治可由所公开蛋白质防治的相同或类似害虫的作用,或所述杀虫剂起防治在更广泛宿主范围内的害虫诸如鳞翅目种或半翅目(Hemipteran)种或不被有效防治的其他植物害虫物种诸如鞘翅目种的作用。

[0098] 以上提及的组合物/制剂可进一步包含农业上可接受的载体,所述组合物/制剂诸如诱饵、粉剂、粉尘剂、丸粒、颗粒剂、喷雾剂、乳液、胶体混悬液、水溶液、芽孢杆菌芽孢/晶体制剂、种子处理剂、被转化以表达一种或多种蛋白质的重组植物细胞/植物组织/种子/植物、或被转化以表达一种或多种蛋白质的细菌。视重组多肽固有的昆虫抑制或杀昆虫抑制水平和待施加于植物或应用于膳食测定的制剂的水平而定,以重量计,组合物/制剂可包括各种量的重组多肽,例如以重量计,0.0001%至0.001%至0.01%至1%至99%的重组多肽。

[0099] 本文公开的蛋白质可组合在用于向植物表面、向土壤进行表面施加的制剂、用于种子处理的制剂、以及具有对鳞翅目种的靶标害虫具有毒性的其他药剂的制剂中。所述药剂包括但不限于:Cry1A蛋白、Cry1B、Cry1C、Cry1F、Cry1A/F嵌合体和Cry2Ab蛋白。

[0100] 实施例

[0101] 以下所公开实施方案仅代表可以各种形式体现的本发明。因此,本文公开的特定结构和功能细节不应解释为具有限制性。应了解本文引用的各参考文献的整个公开内容以引用的方式并入本文。

[0102] 实施例1

[0103] 该实施例教导对毒素蛋白质BCW 001的探索和分析以及对嵌合毒素BCW 002和BCW 003的构建。

[0104] 通过使用芽孢晶体制备物,在膳食生物测定中,Bt菌株EG4384被鉴定为赋予鳞翅目昆虫活性。产生该菌株的基因组的序列,加工原始序列读段,由经加工的读段装配重叠群,鉴定显示与Cry1蛋白的同源性的开放阅读框,并且分析推导的氨基酸序列。鉴定出了如以SEQ ID NO:1阐述的特定开放阅读框,其编码相较于本领域中已知的大多数Cry1蛋白,展现新型氨基酸序列的蛋白质的推导氨基酸序列(BCW 001,SEQ ID NO:2)。从开放阅读框推导的蛋白质因为其长度为1180个氨基酸而具有的新型Cry1型蛋白的全部特征,并且与已知Cry1蛋白的比对指示该蛋白质具有在约600-630个氨基末端氨基酸内的特征性三结构域结构,以及Cry1A型特征性原毒素氨基酸序列结构。编码该预测氨基酸序列的多核苷酸序列含有也具有Cry1特征性的开放阅读框,即在编码毒素的预测结构域I的C末端区域的DNA区段内具有NheI限制位点,并且在编码预测原毒素结构域的N末端部分的DNA区段内具有KpnI限制位点。

[0105] BCW 001毒素与Cry1Ac进行氨基酸序列比较揭示对应于结构域I的氨基酸区段(从约位置1直至约位置258的氨基酸)展现与Cry1Ac内的那个相同区段的仅约67%同一性,对应于结构域II的氨基酸区段(从约位置58直至约位置460的氨基酸)展现与Cry1Ac内的那个相同区段的极低同一性百分比,并且对应于结构域III的氨基酸区段(从约位置460直至约位置607的氨基酸)展现与来自Cry1Ah2的结构域III区段的约63%同一性。

[0106] 将从NheI限制位点直至KpnI限制位点的编码大致上预测结构域II和III的DNA区段切除,并且取代含有编码Cry1Ac的DNA区段的表达载体中Cry1Ac编码序列的相应区段,从而导致开放阅读框由编码Cry1Ac的结构域I的第一区段、编码BCW 001的结构域II和III的第二区段和编码Cry1Ac毒素蛋白的原毒素结构域的第三区段组成,并且以连续顺序从5'至3'使所述第一区段、所述第二区段和所述第三区段同框连接。该嵌合构建体(SEQ ID NO:3)编码在本文中称为BCW 002的嵌合毒素蛋白质(SEQ ID NO:4)。略微改变结构域I与结构域II之间的分界点产生了开放阅读框(SEQ ID NO:5),所述开放阅读框编码在本文中称为BCW 003的嵌合毒素蛋白质(SEQ ID NO:6),所述嵌合毒素蛋白质具有仅在氨基酸位置259处不同于BCW 002的氨基酸序列。BCW 003,如同BCW 001一样,在位置259处含有苏氨酸(T),而BCW 002在那个位置处含有异亮氨酸(I)。BCW 001主要在毒素的结构域I(即氨基酸1-202)内不同于BCW 002和BCW 003,并且如上所陈述,BCW 002和BCW 002实际上是同一的,不同之处是在位置259处的I/T差异。

[0107] 实施例2

[0108] 该实施例教导BCW 001、BCW 002和BCW 003蛋白的有效鳞翅目害虫防治生物活性。

[0109] 将表达BCW 001、BCW 002和BCW 003毒素蛋白的构建体转化至大肠杆菌或可适用的苏云金芽孢杆菌或其他芽孢杆菌中允许在生物测定中测试所表达蛋白质,并且与本领域中已知的对黑切根虫具有毒性的蛋白质诸如Cry1Fa和Cry1Ac进行比较。观察到所得重组菌株表达具有针对鳞翅目害虫的活性的重组蛋白。当针对黑切根虫和玉米耳虫幼虫测试时,生物测定活性是特别强烈的。如以上在本发明的具体实施方式、发明背景和发明内容中所指定,存在极少数已被发现展现任何可观水平的针对黑切根虫的生物活性的毒素蛋白质,因此,本领域中需要鉴定此类蛋白质以在植物中用于保护此类植物免遭黑切根虫侵扰,以及确保存在对黑切根虫具有活性的补充性蛋白质的足够供给,所述补充性蛋白质可用于克服对当前在用的任何所述对黑切根虫具有活性的蛋白质诸如Cry1Fa毒素蛋白的抗性的任何发展。

[0110] 实施例3

[0111] 该实施例教导当BCW 001的结构域II和III取代其他Cry1毒素中的相应结构域时,此类结构域足以将黑切根虫生物活性传递至此类其他Cry1毒素蛋白。

[0112] 鉴定了具有针对鳞翅目昆虫的活性的许多BCW毒素嵌合体,两种嵌合体特别展现针对BCW、WBC和SCB的强烈活性。

[0113] 具有编码Cry1Ab、Cry1Ac和Cry1Ca的核苷酸序列的构建体用于构建含有取代Cry1Ab、Cry1Ac和Cry1Ca的可适用结构域的结构域II和结构域III BCW 001区段的嵌合体,并且在芽孢晶体生物测定中测试所得天然和嵌合蛋白。对在所述膳食生物测定中针对BCW、FAW和CEW的活性进行列表。在测试的实验条件下,Cry1Ac展现针对FAW、BCW和CEW的活性,Cry1Ab展现针对FAW和CEW但不针对BCW的活性,并且Cry1Ca不展现针对FAW、BCW以及CEW的活性。BCW 001展现针对BCW和CEW而非FAW的活性。相较于BCW 003,对于Cry1Ac,Cry1Ac针对BCW的活性是约1/10。当在生物测定中针对FAW、BCW和CEW测试时,Cry1Ab和Cry1Ac、含有BCW 001的结构域II和III的嵌合体展现毒性性质。Cry1Ac针对CEW不具有毒性,并且Cry1Ab针对BCW不具有毒性。构建了Cry1Ca/BCW 001嵌合体,其中Cry1Ca结构域III取代BCW 001的相应结构域,并且所得嵌合毒素对FAW、BCW和CEW展现毒性性质,而Cry1Ca毒素在针对这些害虫

中的任一者测试时都是无效的。

[0114] 实施例4

[0115] 该实施例说明BCW 001、BCW 002和BCW 003在生物测定中针对多种鳞翅目害虫测试时的毒性性质。

[0116] 用于生物测定以及对昆虫的致死性和生长阻碍进行评分的方案在本领域中是已知的,所述方案的实例描述于PCT专利申请公布号W0 2012/139004和美国专利号7,927,598中。

[0117] 图1使各种BCW 001、BCW 002和BCW 003毒素蛋白与在膳食生物测定中根据昆虫物种获得的杀害虫活性相关联。这些毒素蛋白质各自显示针对鳞翅目昆虫的活性。

[0118] 实施例5

[0119] 该实施例教导用于植物中的编码本发明的蛋白质的人工序列的构建、用于植物中的植物载体和构建体的制备、以及表达本发明的蛋白质的植物的产生。

[0120] 根据美国专利5,500,365中所述的方法,设计和合成编码BCW 001蛋白(SEQ ID NO:1)、BCW 002蛋白(SEQ ID NO:3)和BCW 003蛋白(SEQ ID NO:5)的核苷酸序列。设计用于植物表达的这些经工程改造编码区在本文中提供为编码BCW 001的SEQ ID NO:7、编码BCW 002的SEQ ID NO:9和编码BCW 003的SEQ ID NO:11。

[0121] 构建具有有如SEQ ID NO:7、SEQ ID NO:9和SEQ ID NO:11阐述的序列的多种植物表达盒。所述表达盒适用于在植物原生质体中短暂表达或转化植物细胞。关于蛋白质在细胞内的最终放置来设计典型表达盒。以使蛋白质被翻译以及保持在胞质液中的方式设计一组表达盒。设计另一组表达盒以具有与毒素蛋白质邻接的转运肽,以允许靶向细胞的细胞器诸如叶绿体或质体。所有表达盒都被设计在5'末端以启动子起始,所述启动子可包含多个连续连接的启动子元件和增强子元件以加强转基因的表达。启动子序列通常连续继之以一个或多个在启动子的3'的前导序列。在前导序列的3'提供内含子序列以改进转基因的表达。毒素或转运肽的编码序列以及毒素的编码序列位于启动子、前导物和内含子构型的3'。在编码序列的3'提供序列以有助于转录终止,并且提供对所得转录物的多腺苷酸化重要的序列。所有上述元件都经常与被提供用于构建表达盒的额外序列诸如限制核酸内切酶位点或连接非依赖性克隆位点连续排列。

[0122] 对于玉米植物,设计一组用于胞质表达BCW 001,包含墨西哥玉米(Mexicana)泛素1启动子;用于胞质表达BCW 002,包含水稻(Oryza sativa)肌动蛋白(actin)15启动子或35S启动子;以及用于胞质表达BCW 003,包含35S启动子的表达盒。

[0123] 设计另一组用于在玉米植物中靶向表达BCW 002和BCW 003昆虫毒素蛋白的表达盒,其中叶绿体肽编码序列(例如CTP2)同框融合在编码BCW毒素蛋白的DNA区段的5'末端,所述表达盒包含水稻肌动蛋白15启动子或35S启动子以及包含35S启动子的序列。

[0124] 在植物转化载体中构建包含CaMV 35S启动子或可操作地连接到35S启动子的PC1SV.FL启动子的甘蔗植物表达盒。在一些情况下,包括表达靶向叶绿体的Cry2Ab的盒。

[0125] 针对第三龄期BCW、WBC、CEW、SWC和SCB幼虫测试表达本发明的蛋白质的植物。将BCW 001的胞质表达盒以及BCW 002和BCW 003的胞质和质体靶向表达盒克隆,并且用于产生表达这些蛋白质的转基因玉米事件。通过本领域中已知的方法来诱导经转化细胞以形成植物。如美国专利8,344,207中所述,以类似方式进行使用植物叶片的生物测定。基于由昆

虫吞食的叶盘的百分比,在0(0%被食用)至11(大于50%被食用)的标度上对叶损害等级(LDR)指定等级评分。等级评分步幅以5%以递增方式增加。同基因玉米株系用于获得组织作为阴性对照,并且对结果进行评估。相对于未转化对照,质体靶向表达BCW 002和BCW 003昆虫毒素蛋白与胞质表达BCW 002和BCW 003昆虫毒素蛋白两者均使摄食损害降低。针对这些昆虫的叶盘测定的结果与来自以上呈现的实施例的生物测定数据一致。一种包含用于胞质表达的BCW 003盒的构建体导致34个转化事件,这些事件中的25个事件展现完全防治BCW新生体。也针对CEW、SWC和FAW测试将BCW 001和BCW 003表达至胞质液中的玉米植物。表达BCW 003的植物展现100%防治CEW和SWC,并且LDR值在1与2之间的范围内。三个表达BCW 001的转化事件导致植物展现针对CEW和SWC的功效,并且LDR值在1与3之间的范围内。这与先前实施例中呈现的膳食生物测定数据一致。

[0126] 产生表达BCW 003的转基因甘蔗植物,并且在生物测定中针对SCB加以测试。各生物测定包括作为阴性对照的来自野生型甘蔗的叶盘和表达高水平的Cry2Ab的阳性对照。在侵扰之后四(4)天测量昆虫致死性和叶损害。发现来自表达BCW 003的若干转基因甘蔗事件的叶盘以植物中方式防治蔗螟,并且展现类似于阳性对照的低于2的损害等级,以及90%-100%的平均昆虫致死率。

[0127] 相较于仅表达BCW 003的事件,表达BCW 003和Cry2Ab的转基因事件展现更好SCB防治。

序列表

<110> 孟山都技术公司
 <120> 对鳞翅目昆虫具有活性的杀害虫毒素蛋白质
 <130> MONS:434W0
 <150> US 62/445,313
 <151> 2017-01-12
 <160> 12
 <170> PatentIn version 3.5
 <210> 1
 <211> 3543
 <212> DNA
 <213> 苏云金芽孢杆菌(Bacillus thuringiensis)

<220>
 <221> CDS
 <222> (1)..(3540)

[0001] <400> 1
 atg gag gaa aat aat cag aat caa tgc gtc cct tat aat tgt ttg aat 48
 Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15
 aat cct gca atc gaa ata tta gaa gga gac aga ata tca gtt ggt aac 96
 Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn
 20 25 30
 act cca atc gat att tet cta tca ctt gtg gaa ctt ctt att agt gaa 144
 Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu
 35 40 45
 ttt gtc cca ggc ggt gga ata ata aca gga ttg ttg aac ata gta tgg 192
 Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp
 50 55 60
 gga ttt gta ggg cct tcc caa tgg gac gca ttt ctt gct caa gtg gaa 240
 Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu
 65 70 75 80
 cag tta att aac caa agg ata tca gaa gct gta aga aat aca gca att 288
 Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile
 85 90 95
 cag gaa tta gag gga atg gcg cgg gtt tat aga acc tat gct act gct 336

	Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala	
	100 105 110	
	ttt gct gag tgg gaa aga gat cct aat aac aca gat cta aga gaa gca	384
	Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala	
	115 120 125	
	gta cgg aca cag ttt aca gca act gag act tat atc agt gga aga ata	432
	Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile	
	130 135 140	
	tct gtt tta aaa att caa aat ttt gaa gtg cag ctg tta tcg gtg ttt	480
	Ser Val Leu Lys Ile Gln Asn Phe Glu Val Gln Leu Leu Ser Val Phe	
	145 150 155 160	
	gcc caa gct gcc aat tta cat tta tct tta tta aga gac gtt gtg ttt	528
	Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe	
	165 170 175	
	ttt ggg caa aga tgg ggg ttt tca acg aca acc gta aat aat tac tac	576
	Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr	
	180 185 190	
[0002]	aat gat tta aca gaa gag att agt acc tat aca gat tat gca gta cgc	624
	Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg	
	195 200 205	
	tgg tac aat acg gga tta gag cgt gta tgg gga ccg gat tct aga gat	672
	Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp	
	210 215 220	
	tgg gta agg tat aat caa ttt aga aga gag cta aca ctt act gta tta	720
	Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu	
	225 230 235 240	
	gat atc gtt gct cta ttc cca aat tat gat agt cga agg tat cca att	768
	Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile	
	245 250 255	
	cga aca gtt tcc caa tta aca aga gaa att tat acg aac cca gta tta	816
	Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu	
	260 265 270	
	gaa aat ttt gat ggt agt ttt cgt gga atg gct cag aga ata gaa cag	864
	Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln	
	275 280 285	
	aat att agg caa cca cat ctt atg gat atc ctt aat agt ata acc att	912
	Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile	
	290 295 300	

	tat act gat gtg cat aga ggc ttt aat tat tgg tca ggg cat caa ata Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile 305 310 315 320	960
	aca gct tct cct gta ggg ttt tca gga cca gaa ttc gca ttc cct tta Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu 325 330 335	1008
	ttt ggg aat gcg gga aat gca gct cca ccc gta ctt gtc tca tta act Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr 340 345 350	1056
	ggc ttg ggg att ttt aga aca tta tct tca cct tta tat aga aga att Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile 355 360 365	1104
	ata ctt ggt tca ggc cca aat aat cag gaa ctg ttt gtc ctt gat gga Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly 370 375 380	1152
	acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act ata Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile 385 390 395 400	1200
[0003]	tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccg cca cag Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln 405 410 415	1248
	gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt cat Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His 420 425 430	1296
	ggt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga gct Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala 435 440 445	1344
	cca acg ttt tct tgg cag cat cgc agt gct acg aca act aat ata att Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile 450 455 460	1392
	gca gcg gat agt att act caa att cct gct gtt aaa gga cgt tct att Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile 465 470 475 480	1440
	att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc gat Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp 485 490 495	1488
	ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca tac Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr 500 505 510	1536

	ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt gtt Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg Val 515 520 525	1584
	cgt gtt cgt tat gct tct aca tct tca ctc cct gta gat gta gtt ttc Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val Phe 530 535 540	1632
	gga aat att agt cat cct act aca ttc cca gcc act gcc aga tca tta Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser Leu 545 550 555 560	1680
	gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga act Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly Thr 565 570 575	1728
	ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta tct Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu Ser 580 585 590	1776
	act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act gca Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr Ala 595 600 605	1824
[0004]	acc ttt gaa gca gaa tcg gat tta gaa aga gca caa aag gcg gtg aat Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn 610 615 620	1872
	gcg ctg ttt act tct aca aac caa cta ggg ata aaa aca gat gtg acg Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val Thr 625 630 635 640	1920
	gat tat cat att gat caa gtg tcc aat tta gtg gag tgt tta tcg gat Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp 645 650 655	1968
	gaa ttt tat ctg gat gaa aag cga gaa ttg tcc gag aaa gtc aaa cat Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His 660 665 670	2016
	gcg aag cga ctc agt gat gag cga aat tta ctt caa gat cca aac ttc Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe 675 680 685	2064
	agg ggc atc aat aga caa cca gat cgt ggc tgg aga gga agt acg gat Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp 690 695 700	2112
	att acc atc caa gga gga gat gac gta ttc aaa gag aat tac gtc aca Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr	2160

705	710	715	720	
cta cca ggt acc ttt gat gag tgc tat cca acg tat tta tat caa aaa				2208
Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys				
	725	730	735	
ata gat gag tcg aaa tta aaa gcc tat acc cgt tat caa tta aga ggg				2256
Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly				
	740	745	750	
tat atc gag gat agt caa gac tta gaa atc tat tta att cgc tac aat				2304
Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn				
	755	760	765	
gca aaa cat gaa aca gta aat gtg cca ggt acg ggt tcc tta tgg ccg				2352
Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro				
	770	775	780	
ctt tca gcc caa agt cca atc gga aag tgt gga gag ccg aat cga tgc				2400
Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys				
	785	790	800	
gcg cca cac ctt gaa tgg aat cct gat tta cac tgt tcc tgc aga gac				2448
Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp				
	805	810	815	
[0005]				
ggg gaa aaa tgt gct cat cat tct cat cat ttc tcc ttg gac att gat				2496
Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp				
	820	825	830	
gtt gga tgt aca gac tta aat gag gat tta ggt gta tgg gtg ata ttc				2544
Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe				
	835	840	845	
aag att aag acg caa gat ggc cat gca aga cta gga aat cta gag ttt				2592
Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe				
	850	855	860	
ctc gaa gag aaa cca tta gta ggg gaa gca cta gct cgt gtg aaa aga				2640
Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg				
	865	870	875	880
gcg gag aaa aaa tgg aga gac aaa cgc gaa aaa tta caa ttg gaa aca				2688
Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr				
	885	890	895	
aat atc gtt tat aaa gag gca aaa gaa tct gta gat gct tta ttt gta				2736
Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val				
	900	905	910	
aac tct caa tat gat caa tta caa gcg gat acg aac atc gcg atg att				2784

	Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile	
	915 920 925	
	cat gcg gca gat aaa cgt gtt cat aga atc cga gaa gcg tac ctt cca	2832
	His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro	
	930 935 940	
	gag tta tct gtg att ccg ggt gta aat gca gac att tcc gaa gaa tta	2880
	Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu	
	945 950 955 960	
	gaa ggg cgt att ttc act gca ttc tct cta tat gat gcg aga aat gtc	2928
	Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val	
	965 970 975	
	att aaa aat ggc gat ttc aat aat ggc tta tta tgc tgg aac gtg aaa	2976
	Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys	
	980 985 990	
	ggg cat gta gat gta gaa gaa caa aat aac cac cgt tcg gtc ctt gtt	3024
	Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val	
	995 1000 1005	
[0006]	gtt ccg gaa tgg gaa gca gaa gtg tca caa gag gtt cgt gtc tgt	3069
	Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys	
	1010 1015 1020	
	ccg ggg cgt ggc tat atc ctt cgt gtc aca gcg tac aag gag gga	3114
	Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly	
	1025 1030 1035	
	tat gga gaa ggt tgc gta acc att cat gag atc gag aac aat aca	3159
	Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr	
	1040 1045 1050	
	gac gaa ctg aag ttt agc aac tgt gta gaa gag gaa gtc tat cca	3204
	Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro	
	1055 1060 1065	
	aac aac acg gta acg tgt aat gat tat act gca aat caa gaa gaa	3249
	Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu	
	1070 1075 1080	
	tat gag ggt acg tac act tct cgt aat cga gga tat gac gaa gcc	3294
	Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala	
	1085 1090 1095	
	tat gaa agc aat tct tct gta cca gct gag tat gcg tca gtc tat	3339
	Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr	
	1100 1105 1110	

gaa gaa aaa gtg tat aca gat gga cga aga ggg aat cct tgt gaa 3384
 Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu
 1115 1120 1125

tct aac aga gga tat ggg gat tac aca cca cta cca get ggc tat 3429
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr
 1130 1135 1140

gtg aca aaa gaa tta gag tac ttc cca gaa acc gat aag gta tgg 3474
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp
 1145 1150 1155

att gag att gga gaa aca gaa gga aca ttc att gtg gat agt gtg 3519
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val
 1160 1165 1170

gaa tta ctc ctt atg gag gaa taa 3543
 Glu Leu Leu Leu Met Glu Glu
 1175 1180

[0007] <210> 2
 <211> 1180
 <212> PRT
 <213> 苏云金芽孢杆菌(Bacillus thuringiensis)
 <400> 2

Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn
 1 5 10 15

Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn
 20 25 30

Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu
 35 40 45

Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp
 50 55 60

Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu
 65 70 75 80

Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile
 85 90 95

Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala
100 105 110

Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala
115 120 125

Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile
130 135 140

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

Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe
165 170 175

Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr
180 185 190

[0008] Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg
195 200 205

Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp
210 215 220

Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu
225 230 235 240

Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile
245 250 255

Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu
260 265 270

Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln
275 280 285

Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile
290 295 300

Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile
 305 310 315 320

Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu
 325 330 335

Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr
 340 345 350

Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile
 355 360 365

Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly
 370 375 380

Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile
 385 390 395 400

[0009] Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln
 405 410 415

Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His
 420 425 430

Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala
 435 440 445

Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile
 450 455 460

Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile
 465 470 475 480

Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp
 485 490 495

Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr

	500	505	510
	Leu Glu Val Pro Ile Phe Phe	Gln Ser Pro Ser Thr	Asn Tyr Arg Val
	515	520	525
	Arg Val Arg Tyr Ala Ser Thr	Ser Ser Leu Pro Val	Asp Val Val Phe
	530	535	540
	Gly Asn Ile Ser His Pro Thr Thr	Phe Pro Ala Thr	Ala Arg Ser Leu
	545	550	555 560
	Asp Asn Leu Gln Ser Asn Asp Phe	Gly Tyr Ile Asp	Ile Ala Gly Thr
		565	570 575
	Phe Leu Pro Ser Leu Gly Pro Ser	Ile Gly Ile Arg	Pro Met Leu Ser
		580	585 590
[0010]	Thr Ile Asn Leu Ile Val Asp	Arg Phe Glu Phe Ile	Pro Val Thr Ala
	595	600	605
	Thr Phe Glu Ala Glu Ser Asp	Leu Glu Arg Ala	Gln Lys Ala Val Asn
	610	615	620
	Ala Leu Phe Thr Ser Thr	Asn Gln Leu Gly	Ile Lys Thr Asp Val Thr
	625	630	635 640
	Asp Tyr His Ile Asp	Gln Val Ser Asn	Leu Val Glu Cys Leu Ser Asp
		645	650 655
	Glu Phe Tyr Leu Asp	Glu Lys Arg	Glu Leu Ser Glu Lys Val Lys His
	660	665	670
	Ala Lys Arg Leu Ser Asp	Glu Arg Asn Leu Leu	Gln Asp Pro Asn Phe
	675	680	685
	Arg Gly Ile Asn Arg	Gln Pro Asp	Arg Gly Trp Arg Gly Ser Thr Asp
	690	695	700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
 705 710 715 720

Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys
 725 730 735

Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly
 740 745 750

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn
 755 760 765

Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro
 770 775 780

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys
 785 790 795 800

[0011] Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp
 805 810 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
 820 825 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
 835 840 845

Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
 850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
 865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr
 885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
 900 905 910

	Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile	
	915	920 925
	His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro	
	930	935 940
	Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu	
	945	950 955 960
	Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val	
	965	970 975
	Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys	
	980	985 990
	Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val	
	995	1000 1005
[0012]	Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys	
	1010	1015 1020
	Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly	
	1025	1030 1035
	Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr	
	1040	1045 1050
	Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro	
	1055	1060 1065
	Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu	
	1070	1075 1080
	Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala	
	1085	1090 1095
	Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr	
	1100	1105 1110

Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu
 1115 1120 1125

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr
 1130 1135 1140

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp
 1145 1150 1155

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val
 1160 1165 1170

Glu Leu Leu Leu Met Glu Glu
 1175 1180

- <210> 3
- <211> 3534
- <212> DNA
- <213> 人工序列(Artificial Sequence)

[0013]

- <220>
- <223> 编码 BCW 002 的完全合成 DNA

- <220>
- <221> CDS
- <222> (1).. (3531)

<400> 3
 atg gat aac aat ccg aac atc aat gaa tgc att cct tat aat tgt tta 48
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

agt aac cct gaa gta gaa gta tta ggt gga gaa aga ata gaa act ggt 96
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

tac acc cca atc gat att tcc ttg tcg cta acg caa ttt ctt ttg agt 144
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

gaa ttt gtt ccc ggt gct gga ttt gtg tta gga cta gtt gat ata ata 192
 Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile
 50 55 60

	260	265	270	
	tta gaa aat ttt gat ggt agt ttt cgt gga atg gct cag aga ata gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu			864
	275	280	285	
	cag aat att agg caa cca cat ctt atg gat atc ctt aat agt ata acc Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr			912
	290	295	300	
	att tat act gat gtg cat aga ggc ttt aat tat tgg tca ggg cat caa Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln			960
	305	310	315	320
	ata aca gct tct cct gta ggg ttt tca gga cca gaa ttc gca ttc cct Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro			1008
	325	330	335	
	tta ttt ggg aat gcg gga aat gca gct cca ccc gta ctt gtc tca tta Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu			1056
	340	345	350	
	act ggt ttg ggg att ttt aga aca tta tct tca cct tta tat aga aga Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg			1104
	355	360	365	
[0015]	att ata ctt ggt tca ggc cca aat aat cag gaa ctg ttt gtc ctt gat Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp			1152
	370	375	380	
	gga acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr			1200
	385	390	395	400
	ata tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccg cca Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro			1248
	405	410	415	
	cag gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser			1296
	420	425	430	
	cat gtt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg			1344
	435	440	445	
	gct cca acg ttt tct tgg cag cat cgc agt gct acg aca act aat ata Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile			1392
	450	455	460	
	att gca gcg gat agt att act caa att cct gct gtt aaa gga cgt tct			1440

	Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser	
	465	470 475 480
	att att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc	1488
	Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly	
	485	490 495
	gat ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca	1536
	Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala	
	500	505 510
	tac ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt	1584
	Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg	
	515	520 525
	gtt cgt gtt cgt tat gct tct aca tct tca ctc cct gta gat gta gtt	1632
	Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val	
	530	535 540
	ttc gga aat att agt cat cct act aca ttc cca gcc act gcc aga tca	1680
	Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser	
	545	550 555 560
[0016]	tta gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga	1728
	Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly	
	565	570 575
	act ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta	1776
	Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu	
	580	585 590
	tct act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act	1824
	Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr	
	595	600 605
	gca acc ttt gaa gca gaa tcg gat tta gaa aga gca caa aag gcg gtg	1872
	Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val	
	610	615 620
	aat gcg ctg ttt act tct aca aac caa cta ggg ata aaa aca gat gtg	1920
	Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val	
	625	630 635 640
	acg gat tat cat att gat caa gtg tcc aat tta gtg gag tgt tta tcg	1968
	Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser	
	645	650 655
	gat gaa ttt tat ctg gat gaa aag cga gaa ttg tcc gag aaa gtc aaa	2016
	Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys	
	660	665 670

	cat gcg aag cga ctc agt gat gag cga aat tta ctt caa gat cca aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn 675 680 685	2064
	ttc agg ggc atc aat aga caa cca gat cgt ggc tgg aga gga agt acg Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr 690 695 700	2112
	gat att acc atc caa gga gga gat gac gta ttc aaa gag aat tac gtc Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val 705 710 715 720	2160
	aca cta cca ggt acc ttt gat gag tgc tat cca aca tat ttg tat caa Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln 725 730 735	2208
	aaa atc gat gaa tca aaa tta aaa gcc ttt acc cgt tat caa tta aga Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg 740 745 750	2256
	ggg tat atc gaa gat agt caa gac tta gaa atc tat tta att cgc tac Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr 755 760 765	2304
[0017]	aat gca aaa cat gaa aca gta aat gtg cca ggt acg ggt tcc tta tgg Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp 770 775 780	2352
	ccg ctt tca gcc caa agt cca atc gga aag tgt gga gag ccg aat cga Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg 785 790 795 800	2400
	tgc gcg cca cac ctt gaa tgg aat cct gac tta gat tgt tcg tgt agg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg 805 810 815	2448
	gat gga gaa aag tgt gcc cat cat tcg cat cat ttc tcc tta gac att Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile 820 825 830	2496
	gat gta gga tgt aca gac tta aat gag gac cta ggt gta tgg gtg atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845	2544
	ttt aag att aag acg caa gat ggg cac gca aga cta ggg aat cta gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu 850 855 860	2592
	ttt etc gaa gag aaa cca tta gta gga gaa gca cta gct cgt gtg aaa Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys 865 870 875 880	2640

	aga gcg gag aaa aaa tgg aga gac aaa cgt gaa aaa ttg gaa tgg gaa	2688
	Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu	
	885 890 895	
	aca aat atc gtt tat aaa gag gca aaa gaa tct gta gat gct tta ttt	2736
	Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe	
	900 905 910	
	gta aac tct caa tat gat caa tta caa gcg gat acg aat att gcc atg	2784
	Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met	
	915 920 925	
	att cat gcg gca gat aaa cgt gtt cat agc att cga gaa gct tat ctg	2832
	Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu	
	930 935 940	
	cct gag ctg tct gtg att ccg ggt gtc aat gcg gct att ttt gaa gaa	2880
	Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu	
	945 950 955 960	
	tta gaa ggg cgt att ttc act gca ttc tcc cta tat gat gcg aga aat	2928
	Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn	
	965 970 975	
[0018]	gtc att aaa aat ggt gat ttt aat aat ggc tta tcc tgc tgg aac gtg	2976
	Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val	
	980 985 990	
	aaa ggg cat gta gat gta gaa gaa caa aac aac caa cgt tcg gtc ctt	3024
	Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu	
	995 1000 1005	
	gtt gtt ccg gaa tgg gaa gca gaa gtg tca caa gaa gtt cgt gtc	3069
	Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val	
	1010 1015 1020	
	tgt ccg ggt cgt ggc tat atc ctt cgt gtc aca gcg tac aag gag	3114
	Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu	
	1025 1030 1035	
	gga tat gga gaa ggt tgc gta acc att cat gag atc gag aac aat	3159
	Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn	
	1040 1045 1050	
	aca gac gaa ctg aag ttt agc aac tgc gta gaa gag gaa atc tat	3204
	Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr	
	1055 1060 1065	
	cca aat aac acg gta acg tgt aat gat tat act gta aat caa gaa	3249
	Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu	

1070	1075	1080	
gaa tac gga ggt gcg tac act tct cgt aat cga gga tat aac gaa			3294
Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu			
1085	1090	1095	
gct cct tcc gta cca gct gat tat gcg tca gtc tat gaa gaa aaa			3339
Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys			
1100	1105	1110	
tcg tat aca gat gga cga aga gag aat cct tgt gaa ttt aac aga			3384
Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg			
1115	1120	1125	
ggg tat agg gat tac acg cca cta cca gtt ggt tat gtg aca aaa			3429
Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys			
1130	1135	1140	
gaa tta gaa tac ttc cca gaa acc gat aag gta tgg att gag att			3474
Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile			
1145	1150	1155	
gga gaa acg gaa gga aca ttt atc gtg gac agc gtg gaa tta ctc			3519
Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu			
1160	1165	1170	
[0019] ctt atg gag gaa tag			3534
Leu Met Glu Glu			
1175			
<210> 4			
<211> 1177			
<212> PRT			
<213> 人工序列(Artificial Sequence)			
<220>			
<223> 合成构建体			
<400> 4			
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

[0020] Ile Pro Leu Leu 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 Tyr 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 Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro

	245	250	255
	Ile Arg Ile Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260	265	270
	Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu 275	280	285
	Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290	295	300
	Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln 305	310	315
	Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro 325	330	335
[0021]	Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu 340	345	350
	Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg 355	360	365
	Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp 370	375	380
	Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr 385	390	395
	Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro 405	410	415
	Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser 420	425	430
	His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg 435	440	445

Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile
450 455 460

Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser
465 470 475 480

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

Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala
500 505 510

Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg
515 520 525

Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val
530 535 540

[0022] Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser
545 550 555 560

Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly
565 570 575

Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu
580 585 590

Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr
595 600 605

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val
625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
645 650 655

Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
 675 680 685

Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr
 690 695 700

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
 740 745 750

[0023] Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
 755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
 770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
 785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
 805 810 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
 820 825 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
 835 840 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
 850 855 860

	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	
	865					870					875					880	
	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp	Glu	
					885					890					895		
	Thr	Asn	Ile	Val	Tyr	Lys	Glu	Ala	Lys	Glu	Ser	Val	Asp	Ala	Leu	Phe	
				900					905						910		
	Val	Asn	Ser	Gln	Tyr	Asp	Gln	Leu	Gln	Ala	Asp	Thr	Asn	Ile	Ala	Met	
			915					920					925				
	Ile	His	Ala	Ala	Asp	Lys	Arg	Val	His	Ser	Ile	Arg	Glu	Ala	Tyr	Leu	
	930						935						940				
	Pro	Glu	Leu	Ser	Val	Ile	Pro	Gly	Val	Asn	Ala	Ala	Ile	Phe	Glu	Glu	
	945					950					955					960	
[0024]	Leu	Glu	Gly	Arg	Ile	Phe	Thr	Ala	Phe	Ser	Leu	Tyr	Asp	Ala	Arg	Asn	
					965					970						975	
	Val	Ile	Lys	Asn	Gly	Asp	Phe	Asn	Asn	Gly	Leu	Ser	Cys	Trp	Asn	Val	
			980							985					990		
	Lys	Gly	His	Val	Asp	Val	Glu	Glu	Gln	Asn	Asn	Gln	Arg	Ser	Val	Leu	
			995						1000						1005		
	Val	Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val		
	1010						1015							1020			
	Cys	Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu		
	1025						1030							1035			
	Gly	Tyr	Gly	Glu	Gly	Cys	Val	Thr	Ile	His	Glu	Ile	Glu	Asn	Asn		
	1040						1045							1050			
	Thr	Asp	Glu	Leu	Lys	Phe	Ser	Asn	Cys	Val	Glu	Glu	Glu	Ile	Tyr		

1055	1060	1065
Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu		
1070	1075	1080
Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu		
1085	1090	1095
Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys		
1100	1105	1110
Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg		
1115	1120	1125
Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys		
1130	1135	1140
Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile		
1145	1150	1155
Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu		
1160	1165	1170
Leu Met Glu Glu		
1175		

[0025]

<210> 5
 <211> 3534
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 编码 BCW 003 的完全合成序列

<220>
 <221> CDS
 <222> (1)..(3531)

<400> 5
 atg gat aac aat ccg aac atc aat gaa tgc att cct tat aat tgt tta
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu

48

1	5	10	15	
agt aac cct gaa gta gaa gta tta ggt gga gaa aga ata gaa act ggt				96
Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly				
	20	25	30	
tac acc cca atc gat att tcc ttg tcg cta acg caa ttt ctt ttg agt				144
Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser				
	35	40	45	
gaa ttt gtt ccc ggt gct gga ttt gtg tta gga cta gtt gat ata ata				192
Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile				
	50	55	60	
tgg gga att ttt ggt ccc tct caa tgg gac gca ttt ctt gta caa att				240
Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile				
65	70	75	80	
gaa cag tta att aac caa aga ata gaa gaa ttc gct agg aac caa gcc				288
Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala				
	85	90	95	
att tct aga tta gaa gga cta agc aat ctt tat caa att tac gca gaa				336
Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu				
	100	105	110	
[0026] tct ttt aga gag tgg gaa gca gat cct act aat cca gca tta aga gaa				384
Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu				
	115	120	125	
gag atg cgt att caa ttc aat gac atg aac agt gcc ctt aca acc gct				432
Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala				
	130	135	140	
att cct ctt ttg gca gtt caa aat tat caa gtt cct ctt tta tca gta				480
Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val				
145	150	155	160	
tat gtt caa gct gca aat tta cat tta tca gtt ttg aga gat gtt tca				528
Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser				
	165	170	175	
gtg ttt gga caa agg tgg gga ttt gat gcc gcg act atc aat agt cgt				576
Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg				
	180	185	190	
tat aat gat tta act agg ctt att ggc aac tat aca gat tat gct gta				624
Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val				
	195	200	205	
cgc tgg tac aat acg gga tta gaa cgt gta tgg gga ccg gat tct aga				672

	Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg 210 215 220	
	gat tgg gta agg tat aat caa ttt aga aga gaa tta aca cta act gta Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
	tta gat atc gtt gct ctg ttc ccg aat tat gat agt aga aga tat cca Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro 245 250 255	768
	att cga aca gtt tcc caa tta aca aga gaa att tat acg aac cca gta Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val 260 265 270	816
	tta gaa aat ttt gat ggt agt ttt cgt gga atg gct cag aga ata gaa Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu 275 280 285	864
	cag aat att agg caa cca cat ctt atg gat atc ctt aat agt ata acc Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr 290 295 300	912
[0027]	att tat act gat gtg cat aga ggc ttt aat tat tgg tca ggg cat caa Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln 305 310 315 320	960
	ata aca gct tct cct gta ggg ttt tca gga cca gaa ttc gca ttc cct Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro 325 330 335	1008
	tta ttt ggg aat gcg gga aat gca gct cca ccc gta ctt gtc tca tta Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu 340 345 350	1056
	act ggt ttg ggg att ttt aga aca tta tct tca cct tta tat aga aga Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg 355 360 365	1104
	att ata ctt ggt tca ggc cca aat aat cag gaa ctg ttt gtc ctt gat Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp 370 375 380	1152
	gga acg gag ttt tct ttt gcc tcc cta acg acc aac ttg cct tcc act Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr 385 390 395 400	1200
	ata tat aga caa agg ggt aca gtc gat tca cta gat gta ata ccg cca Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro 405 410 415	1248

	cag gat aat agt gta cca cct cgt gcg gga ttt agc cat cga ttg agt Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser	1296
	420 425 430	
	cat gtt aca atg ctg agc caa gca gct gga gca gtt tac acc ttg aga His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg	1344
	435 440 445	
	gct cca acg ttt tct tgg cag cat cgc agt gct acg aca act aat ata Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile	1392
	450 455 460	
	att gca gcg gat agt att act caa att cct gct gtt aaa gga cgt tct Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser	1440
	465 470 475 480	
	att att aat aat ggc acg gta att tca gga cca ggg ttt acc gga ggc Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly	1488
	485 490 495	
	gat ttg gtt aga tta tac aat gct gat ttt aat att aat aat aga gca Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala	1536
	500 505 510	
[0028]	tac ctt gaa gtt ccg ata ttc ttc caa tca ccc tct aca aat tat cgt Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg	1584
	515 520 525	
	gtt cgt gtt cgt tat gct tct aca tct tca ctc cct gta gat gta gtt Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val	1632
	530 535 540	
	ttc gga aat att agt cat cct act aca ttc cca gcc act gcc aga tca Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser	1680
	545 550 555 560	
	tta gat aat cta caa tcc aat gat ttt gga tat att gat att gct gga Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly	1728
	565 570 575	
	act ttc tta cct tca cta ggg cct agt ata ggt atc aga cct atg tta Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu	1776
	580 585 590	
	tct act att aat ttg ata gta gat aga ttt gaa ttt att cca gta act Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr	1824
	595 600 605	
	gca acc ttt gaa gca gaa tcg gat tta gaa aga gca caa aag gcg gtg Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val	1872
	610 615 620	

	820	825	830	
	gat gta gga tgt aca gac tta aat gag gac cta ggt gta tgg gtg atc			2544
	Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile			
	835	840	845	
	ttt aag att aag acg caa gat ggg cac gca aga cta ggg aat cta gag			2592
	Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu			
	850	855	860	
	ttt ctc gaa gag aaa cca tta gta gga gaa gca cta gct cgt gtg aaa			2640
	Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys			
	865	870	875	880
	aga gcg gag aaa aaa tgg aga gac aaa cgt gaa aaa ttg gaa tgg gaa			2688
	Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu			
	885	890	895	
	aca aat atc gtt tat aaa gag gca aaa gaa tct gta gat gct tta ttt			2736
	Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe			
	900	905	910	
	gta aac tct caa tat gat caa tta caa gcg gat acg aat att gcc atg			2784
	Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met			
	915	920	925	
[0030]	att cat gcg gca gat aaa cgt gtt cat agc att cga gaa gct tat ctg			2832
	Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu			
	930	935	940	
	cct gag ctg tct gtg att ccg ggt gtc aat gcg gct att ttt gaa gaa			2880
	Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu			
	945	950	955	960
	tta gaa ggg cgt att ttc act gca ttc tcc cta tat gat gcg aga aat			2928
	Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn			
	965	970	975	
	gtc att aaa aat ggt gat ttt aat aat ggc tta tcc tgc tgg aac gtg			2976
	Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val			
	980	985	990	
	aaa ggg cat gta gat gta gaa gaa caa aac aac caa cgt tcg gtc ctt			3024
	Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu			
	995	1000	1005	
	gtt gtt ccg gaa tgg gaa gca gaa gtg tca caa gaa gtt cgt gtc			3069
	Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val			
	1010	1015	1020	
	tgt ccg ggt cgt ggc tat atc ctt cgt gtc aca gcg tac aag gag			3114

	Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu	
	1025 1030 1035	
	gga tat gga gaa ggt tgc gta acc att cat gag atc gag aac aat	3159
	Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn	
	1040 1045 1050	
	aca gac gaa ctg aag ttt agc aac tgc gta gaa gag gaa atc tat	3204
	Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr	
	1055 1060 1065	
	cca aat aac acg gta acg tgt aat gat tat act gta aat caa gaa	3249
	Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu	
	1070 1075 1080	
	gaa tac gga ggt gcg tac act tct cgt aat cga gga tat aac gaa	3294
	Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu	
	1085 1090 1095	
	gct cct tcc gta cca get gat tat gcg tca gtc tat gaa gaa aaa	3339
	Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys	
	1100 1105 1110	
[0031]	tcg tat aca gat gga cga aga gag aat cct tgt gaa ttt aac aga	3384
	Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg	
	1115 1120 1125	
	ggg tat agg gat tac acg cca cta cca gtt ggt tat gtg aca aaa	3429
	Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys	
	1130 1135 1140	
	gaa tta gaa tac ttc cca gaa acc gat aag gta tgg att gag att	3474
	Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile	
	1145 1150 1155	
	gga gaa acg gaa gga aca ttt atc gtg gac agc gtg gaa tta ctc	3519
	Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu	
	1160 1165 1170	
	ctt atg gag gaa tag	3534
	Leu Met Glu Glu	
	1175	
	<210> 6	
	<211> 1177	
	<212> PRT	
	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 合成构建体	

<400> 6

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

[0032]

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 Leu 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 Tyr 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 Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
 245 250 255

Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
 260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu
 275 280 285

[0033] Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
 290 295 300

Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln
 305 310 315 320

Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro
 325 330 335

Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu
 340 345 350

Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg
 355 360 365

Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp
 370 375 380

Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr
 385 390 395 400

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val
625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
645 650 655

Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
675 680 685

Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr
690 695 700

[0035] Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
705 710 715 720

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
725 730 735

Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg

	805	810	815
	Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile		
	820	825	830
	Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile		
	835	840	845
	Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu		
	850	855	860
	Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys		
	865	870	880
	Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu		
	885	890	895
[0036]	Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe		
	900	905	910
	Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met		
	915	920	925
	Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu		
	930	935	940
	Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu		
	945	950	955
	Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn		
	965	970	975
	Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val		
	980	985	990
	Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu		
	995	1000	1005

	Val	Val	Pro	Glu	Trp	Glu	Ala	Glu	Val	Ser	Gln	Glu	Val	Arg	Val
	1010						1015					1020			
	Cys	Pro	Gly	Arg	Gly	Tyr	Ile	Leu	Arg	Val	Thr	Ala	Tyr	Lys	Glu
	1025						1030					1035			
	Gly	Tyr	Gly	Glu	Gly	Cys	Val	Thr	Ile	His	Glu	Ile	Glu	Asn	Asn
	1040						1045					1050			
	Thr	Asp	Glu	Leu	Lys	Phe	Ser	Asn	Cys	Val	Glu	Glu	Glu	Ile	Tyr
	1055						1060					1065			
	Pro	Asn	Asn	Thr	Val	Thr	Cys	Asn	Asp	Tyr	Thr	Val	Asn	Gln	Glu
	1070						1075					1080			
	Glu	Tyr	Gly	Gly	Ala	Tyr	Thr	Ser	Arg	Asn	Arg	Gly	Tyr	Asn	Glu
	1085						1090					1095			
[0037]	Ala	Pro	Ser	Val	Pro	Ala	Asp	Tyr	Ala	Ser	Val	Tyr	Glu	Glu	Lys
	1100						1105					1110			
	Ser	Tyr	Thr	Asp	Gly	Arg	Arg	Glu	Asn	Pro	Cys	Glu	Phe	Asn	Arg
	1115						1120					1125			
	Gly	Tyr	Arg	Asp	Tyr	Thr	Pro	Leu	Pro	Val	Gly	Tyr	Val	Thr	Lys
	1130						1135					1140			
	Glu	Leu	Glu	Tyr	Phe	Pro	Glu	Thr	Asp	Lys	Val	Trp	Ile	Glu	Ile
	1145						1150					1155			
	Gly	Glu	Thr	Glu	Gly	Thr	Phe	Ile	Val	Asp	Ser	Val	Glu	Leu	Leu
	1160						1165					1170			
	Leu	Met	Glu	Glu											
	1175														
	<210>	7													
	<211>	3543													
	<212>	DNA													

	<213> 人工序列(Artificial Sequence)	
	<220>	
	<223> 编码 BCW 001 的用于植物中的完全合成序列	
	<220>	
	<221> CDS	
	<222> (1)..(3540)	
	<223> BCW 001	
	<400> 7	
	atg gag gag aac aac cag aac cag tgt gtc cca tac aac tgc ctc aac	48
	Met Glu Glu Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Asn	
	1 5 10 15	
	aac cca gcc atc gaa atc ctt gag ggc gac cga att tca gtc ggc aac	96
	Asn Pro Ala Ile Glu Ile Leu Glu Gly Asp Arg Ile Ser Val Gly Asn	
	20 25 30	
	acg ccc atc gac atc tcc ctg agt ctt gtg gaa ctc ctc atc tcg gag	144
	Thr Pro Ile Asp Ile Ser Leu Ser Leu Val Glu Leu Leu Ile Ser Glu	
	35 40 45	
[0038]	ttc gtc cct ggc ggc ggc ata atc acc ggt ctg ctc aac atc gtg tgg	192
	Phe Val Pro Gly Gly Gly Ile Ile Thr Gly Leu Leu Asn Ile Val Trp	
	50 55 60	
	gga ttc gtg ggc cca tcc cag tgg gat gcg ttc ctg gcc caa gtg gag	240
	Gly Phe Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Ala Gln Val Glu	
	65 70 75 80	
	cag ctc atc aac cag agg atc tcc gag gcc gtc cgc aat acc gcg atc	288
	Gln Leu Ile Asn Gln Arg Ile Ser Glu Ala Val Arg Asn Thr Ala Ile	
	85 90 95	
	caa gag ctg gag ggc atg gcc cgc gtg tac cgc acc tac gcc acc gcc	336
	Gln Glu Leu Glu Gly Met Ala Arg Val Tyr Arg Thr Tyr Ala Thr Ala	
	100 105 110	
	ttt gct gaa tgg gag cgc gac ccg aac aac act gac ctg cgc gag gcc	384
	Phe Ala Glu Trp Glu Arg Asp Pro Asn Asn Thr Asp Leu Arg Glu Ala	
	115 120 125	
	gtc cga aca cag ttc acg gcg acc gag acc tac atc agc ggc cgg atc	432
	Val Arg Thr Gln Phe Thr Ala Thr Glu Thr Tyr Ile Ser Gly Arg Ile	
	130 135 140	
	tca gtg ctc aag atc cag aac ttc gag gtg cag ctc cta tcg gtc ttc	480
	Ser Val Leu Lys Ile Gln Asn Phe Glu Val Gln Leu Leu Ser Val Phe	
	145 150 155 160	

	gcc cag gcc gcc aac ttg cac ctg agc ctc ctg cgg gac gtt gtg ttc Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe 165 170 175	528
	ttc ggc cag cgg tgg ggc ttc tct act acg acc gtg aac aac tac tac Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr 180 185 190	576
	aac gac ctg acg gag gaa atc agc acc tac aca gat tac gca gtt cgt Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg 195 200 205	624
	tgg tac aac acc ggc ctt gag cgc gtg tgg ggc ccg gac tcc cgc gat Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp 210 215 220	672
	tgg gtc cgc tac aac cag ttc cgc cgc gag ctg acg ctt aca gtg ctg Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu 225 230 235 240	720
	gac atc gtc gca ctc ttt cct aac tac gac tcc agg cgc tat ccc atc Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile 245 250 255	768
[0039]	agg aca gtg tca cag ctc acc cgc gag atc tac aca aac ccg gtg ctt Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu 260 265 270	816
	gag aac ttc gac ggc agc ttc cgt ggc atg gcg cag cgc att gaa cag Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln 275 280 285	864
	aac atc cgc cag ccg cac ctt atg gac atc ttg aac agt atc act atc Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile 290 295 300	912
	tac acc gac gtc cac aga ggc ttc aac tac tgg agc gga cac cag atc Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile 305 310 315 320	960
	aca gcc agc ccg gta ggc ttc tcg ggt cca gag ttc gcc ttc ccg ctg Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu 325 330 335	1008
	ttt ggg aac gct ggc aat gcc gcg ccg ccc gtg ctg gtc agc ctc act Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr 340 345 350	1056
	ggt ctc ggc atc ttc cgc aca ctt tcc tcg ccg ctg tac agg agg atc Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg Ile	1104

	355	360	365	
	atc ctc ggg tcc ggt ccg aac aac cag gag ctg ttc gtg ctc gac ggg			1152
	Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly			
	370	375	380	
	acc gag ttc agt ttc gcc agc ctc acg acg aac ctc ccg tcc acc atc			1200
	Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile			
	385	390	395	400
	tat cgc cag cgc gga acg gtc gat tcc ctg gat gtt atc cca ccg caa			1248
	Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln			
	405	410	415	
	gac aat tct gtg ccg ccg agg gcc ggg ttc tcc cac ccg ctg tct cac			1296
	Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His			
	420	425	430	
	gtg act atg ctt tca cag gcc gcc gga gcc gtg tac acg ctc cgt gcg			1344
	Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala			
	435	440	445	
	cct act ttc tcc tgg cag cac cgc agc gcg acc acg acc aac atc atc			1392
	Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile			
	450	455	460	
[0040]	gca gca gac tcc atc acc cag atc ccg gcc gtt aag ggc cgc agc atc			1440
	Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile			
	465	470	475	480
	atc aac aac gga act gtc atc agc ggt ccg gcc ttc acg gcc gcc gac			1488
	Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp			
	485	490	495	
	ctg gtc ccg ctc tac aac gca gac ttc aac atc aat aac cgc gct tat			1536
	Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr			
	500	505	510	
	ctt gaa gta cct atc ttc ttc cag agc ccg agc act aac tac ccg gtt			1584
	Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg Val			
	515	520	525	
	cgc gtc cgc tac gcc agc acc tcc agc ctc cct gtg gat gtc gtg ttc			1632
	Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val Phe			
	530	535	540	
	gga aac ata agc cat ccg acc acg ttc cca gcc acg gct agg agc ctg			1680
	Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser Leu			
	545	550	555	560
	gac aac cta cag agc aac gac ttc ggc tac atc gac atc gcg gcc acc			1728

	Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly Thr	
	565 570 575	
	ttt ctg cca agc ctg ggt ccg tct atc ggc atc cgc ccg atg ctg agc	1776
	Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu Ser	
	580 585 590	
	act atc aac cta att gtg gac cgg ttc gag ttt atc ccg gtg acg gca	1824
	Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr Ala	
	595 600 605	
	acg ttc gag gcg gag tct gac ctc gaa agg gca cag aag gcc gtg aac	1872
	Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn	
	610 615 620	
	gcc ctg ttc acg agc acc aac cag ctt ggc att aag act gat gtc acc	1920
	Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val Thr	
	625 630 635 640	
	gac tac cac att gac caa gtc agc aac ctg gtg gag tgc ctc tcg gac	1968
	Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp	
	645 650 655	
	gag ttc tat ctt gat gag aaa cgg gaa cta agc gag aag gtg aag cac	2016
[0041]	Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His	
	660 665 670	
	gca aag cgc ttg agc gac gag cgg aac tta ctc cag gac cct aac ttc	2064
	Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe	
	675 680 685	
	cgt ggg att aac cgc cag ccg gat cgc ggg tgg cgc ggc tca acg gac	2112
	Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp	
	690 695 700	
	atc acc atc cag ggc ggc gat gac gtc ttc aag gag aac tac gtg acc	2160
	Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr	
	705 710 715 720	
	ctc cct ggc acg ttc gac gag tgc tac ccg acg tac ctt tat cag aag	2208
	Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys	
	725 730 735	
	att gac gaa agc aag ctg aaa gcc tac acc cgc tac cag ttg cgc ggc	2256
	Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly	
	740 745 750	
	tac atc gag gac tct caa gac ctg gag atc tac ttg att cga tac aac	2304
	Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn	
	755 760 765	

	gcg aaa cac gag acc gtc aac gtg cgg ggc act ggg agc ctg tgg ccg Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro 770 775 780	2352
	ttg tct gca caa agt ccg atc ggc aag tgc ggc gag cca aac cgg tgc Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys 785 790 795 800	2400
	gct ccg cac ctg gag tgg aac cca gac ctt cat tgc tcc tgt agg gat Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp 805 810 815	2448
	ggc gag aag tgc gct cac cac agc cat cac ttc agc ctc gac att gac Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp 820 825 830	2496
	gtc ggc tgt acc gac ctt aat gag gat ctg ggt gtg tgg gtg atc ttc Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe 835 840 845	2544
	aag atc aag acc cag gac ggt cac gcc cgg ttg ggc aat ctg gag ttc Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe 850 855 860	2592
[0042]	ctg gag gag aag ccg ctg gtt ggc gag gct ctc gcg cgg gtc aag ccg Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg 865 870 875 880	2640
	gcg gag aag aag tgg cgg gac aaa cgc gag aag ctc cag tta gag acg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr 885 890 895	2688
	aac atc gtg tac aag gag gcg aag gaa tcc gtg gac gca cta ttc gtg Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val 900 905 910	2736
	aac agc cag tac gac caa ctc cag gcc gac acc aac atc gcc atg att Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile 915 920 925	2784
	cac gca gcc gac aag agg gtg cac cgc atc cgc gaa gcc tac ctt ccc His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro 930 935 940	2832
	gaa ctt tcg gtc atc cca ggc gtc aac gct gac atc tcg gag gaa ttg Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu 945 950 955 960	2880
	gag ggc aga atc ttc acg gcc ttc tct ttg tac gat gcc agg aac gtc Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val 965 970 975	2928

	atc aag aac ggc gac ttc aac aac ggc ctg ctg tgc tgg aac gtg aag Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys 980 985 990	2976
	ggc cac gtg gac gtc gag gag cag aac aac cac aga tca gtc ctg gtg Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val 995 1000 1005	3024
	gtg ccc gag tgg gaa gcc gaa gtc tca caa gaa gtc cgg gtg tgc Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys 1010 1015 1020	3069
	cct gga cgc ggg tac att ctc cgc gtg acc gcc tac aag gag ggc Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly 1025 1030 1035	3114
	tac ggt gag ggc tgc gtg acc atc cac gag atc gag aac aac acc Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr 1040 1045 1050	3159
	gac gag ctg aaa ttc agt aac tgt gtt gag gag gag gtg tac ccg Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro 1055 1060 1065	3204
[0043]	aac aac acc gtc acc tgc aac gac tac act gcg aac cag gag gaa Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu 1070 1075 1080	3249
	tac gag ggc acg tac acg agc cgc aat cgc ggg tac gac gag gcg Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala 1085 1090 1095	3294
	tac gag agc aac tcc agc gtc ccg gcc gag tac gcc tcc gtg tac Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr 1100 1105 1110	3339
	gag gag aag gtt tac acc gac ggg agg cgt ggc aac ccg tgc gag Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu 1115 1120 1125	3384
	agc aac aga ggc tac ggc gat tac act ccg ctt ccc gct ggc tac Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr 1130 1135 1140	3429
	gtg acg aaa gag ctg gag tac ttc cca gag acc gac aag gtg tgg Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp 1145 1150 1155	3474
	atc gag atc gga gaa acg gag ggc acg ttc ata gtg gac tcc gtt Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val 1160 1165 1170	3519

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

Ala Gln Ala Ala Asn Leu His Leu Ser Leu Leu Arg Asp Val Val Phe
 165 170 175

Phe Gly Gln Arg Trp Gly Phe Ser Thr Thr Thr Val Asn Asn Tyr Tyr
 180 185 190

Asn Asp Leu Thr Glu Glu Ile Ser Thr Tyr Thr Asp Tyr Ala Val Arg
 195 200 205

Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg Asp
 210 215 220

Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val Leu
 225 230 235 240

[0045] Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro Ile
 245 250 255

Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val Leu
 260 265 270

Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu Gln
 275 280 285

Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr Ile
 290 295 300

Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln Ile
 305 310 315 320

Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro Leu
 325 330 335

Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu Thr

	340		345		350
	Gly Leu Gly Ile Phe Arg Thr	Leu Ser Ser Pro Leu Tyr Arg Arg Ile			
	355		360		365
	Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp Gly				
	370		375		380
	Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr Ile				
	385		390		395
	Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro Gln				
		405		410	415
	Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser His				
		420		425	430
[0046]	Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg Ala				
	435		440		445
	Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile Ile				
	450		455		460
	Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser Ile				
	465		470		475
	Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly Asp				
		485		490	495
	Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala Tyr				
		500		505	510
	Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg Val				
	515		520		525
	Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val Phe				
	530		535		540

Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser Leu
 545 550 555 560

Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly Thr
 565 570 575

Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu Ser
 580 585 590

Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr Ala
 595 600 605

Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val Asn
 610 615 620

Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val Thr
 625 630 635 640

[0047] Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser Asp
 645 650 655

Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His
 660 665 670

Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe
 675 680 685

Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp
 690 695 700

Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val Thr
 705 710 715 720

Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln Lys
 725 730 735

Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg Gly
 740 745 750

Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr Asn
 755 760 765

Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp Pro
 770 775 780

Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys
 785 790 795 800

Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg Asp
 805 810 815

Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile Asp
 820 825 830

Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile Phe
 835 840 845

[0048] Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu Phe
 850 855 860

Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys Arg
 865 870 875 880

Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Gln Leu Glu Thr
 885 890 895

Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe Val
 900 905 910

Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met Ile
 915 920 925

His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu Pro
 930 935 940

Glu Leu Ser Val Ile Pro Gly Val Asn Ala Asp Ile Ser Glu Glu Leu
 945 950 955 960

Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn Val
 965 970 975

Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val Lys
 980 985 990

Gly His Val Asp Val Glu Glu Gln Asn Asn His Arg Ser Val Leu Val
 995 1000 1005

Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys
 1010 1015 1020

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly
 1025 1030 1035

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr
 1040 1045 1050

[0049] Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Val Tyr Pro
 1055 1060 1065

Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala Asn Gln Glu Glu
 1070 1075 1080

Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala
 1085 1090 1095

Tyr Glu Ser Asn Ser Ser Val Pro Ala Glu Tyr Ala Ser Val Tyr
 1100 1105 1110

Glu Glu Lys Val Tyr Thr Asp Gly Arg Arg Gly Asn Pro Cys Glu
 1115 1120 1125

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr
 1130 1135 1140

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp

	1145	1150	1155	
	Ile Glu	Ile Gly Glu Thr Glu	Gly Thr Phe Ile Val	Asp Ser Val
	1160	1165	1170	
	Glu Leu	Leu Leu Met Glu Glu		
	1175	1180		
	<210>	9		
	<211>	3534		
	<212>	DNA		
	<213>	人工序列(Artificial Sequence)		
	<220>			
	<223>	用于植物中的编码 CW 002 的完全合成序列		
	<220>			
	<221>	CDS		
	<222>	(1)..(3531)		
	<223>	BCW 002		
[0050]	<400>	9		
	atg gac aac aac ccg aac atc aac gag tgc atc ccc tac aac tgc etc			48
	Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu			
	1	5	10	15
	tcc aac ccg gag gtc gag gtg ctg ggc ggc gaa agg atc gag acc ggc			96
	Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly			
		20	25	30
	tac act ccc atc gac atc agc etc agc ctg acc cag ttc ctg etc tct			144
	Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser			
		35	40	45
	gag ttc gtg ccc ggc gcg ggg ttc gtt etc ggc ctg gtc gac atc atc			192
	Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile			
		50	55	60
	tgg ggc atc ttc ggt ccg agc cag tgg gac gcc ttt etc gtt cag att			240
	Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile			
	65	70	75	80
	gag cag ctg atc aac cag cgc atc gag gag ttc gcc cgc aac cag gcg			288
	Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala			
		85	90	95
	atc tcc cgg ctg gag ggc etc tcc aac ctg tac caa atc tac gcc gag			336

	Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu	
	100 105 110	
	agc ttc cgg gag tgg gaa gcc gat ccg acc aac ccc gct ctc agg gag	384
	Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu	
	115 120 125	
	gag atg cgg att cag ttc aac gac atg aac tcc gct ctc acg act gcc	432
	Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala	
	130 135 140	
	atc cca ctc ctc gct gtg cag aac tac caa gtg ccg ctc ctg tcc gtg	480
	Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val	
	145 150 155 160	
	tac gtg cag gcc gcc aat ctg cac ctc tcc gtc ctc cgg gac gtt agc	528
	Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser	
	165 170 175	
	gtg ttc ggg cag cgc tgg ggc ttc gac gcc gct acc atc aac tcc cgt	576
	Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg	
	180 185 190	
[0051]	tac aac gat ctc act cgc ctc atc ggc aac tac acc gac tat gcc gtg	624
	Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val	
	195 200 205	
	cgc tgg tac aac act ggt ctt gag aga gtc tgg ggc ccg gac agc cgc	672
	Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg	
	210 215 220	
	gac tgg gtg cgc tac aac cag ttc cgg cgc gag ctg acc ctc acc gtg	720
	Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val	
	225 230 235 240	
	ctc gac atc gta gcc ctc ttt ccc aac tac gac tcc cgg cgc tac ccg	768
	Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro	
	245 250 255	
	att cgc atc gtc agc cag ctc acc agg gag atc tac acc aac cct gtg	816
	Ile Arg Ile Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val	
	260 265 270	
	ctg gag aac ttc gac ggc tcc ttt cgc ggg atg gcc caa cgc ata gag	864
	Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu	
	275 280 285	
	cag aac atc cgc caa cct cat ctg atg gac atc ctt aat tct atc acc	912
	Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr	
	290 295 300	

	atc tac act gac gtt cat cgc ggg ttt aac tac tgg tgc ggc cac caa Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln 305 310 315 320	960
	atc act gcg teg ccc gtt ggt ttc tcc ggc ccg gag ttc gcg ttc cct Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro 325 330 335	1008
	ctg ttc gga aac gcg ggc aat gcc gct cca ccc gta ttg gtg agc ctg Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu 340 345 350	1056
	acc ggc ctc ggc atc ttc cgt aca ctg tct agc cct ctg tac aga agg Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg 355 360 365	1104
	atc att ctt ggc agc ggt ccc aat aac cag gaa ctc ttc gtg ttg gac Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp 370 375 380	1152
	ggc acc gag ttc agc ttc gcc agt ctt acg acc aat ttg ccc tcc aca Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr 385 390 395 400	1200
[0052]	atc tat cgc cag cgc ggt act gtg gac tcc ctt gat gtg ata cca cct Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro 405 410 415	1248
	cag gac aac tct gtc cca cct cgc gcc ggt ttc tcc cac cgc ctc agc Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser 420 425 430	1296
	cac gtc act atg ctg agt cag gct gcg gga gcc gtg tac acc ctt cgg His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg 435 440 445	1344
	gct ccg acg ttt agc tgg cag cac agg agc gcg act acc acg aac atc Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile 450 455 460	1392
	att gcg gct gac tcc atc act caa atc cct gcc gtt aag ggt cgc tcc Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser 465 470 475 480	1440
	atc atc aac aat ggg aca gtg atc tcg gga ccg ggc ttc acc ggc ggt Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly 485 490 495	1488
	gac ctg gtg agg ctg tac aac gcg gac ttc aac atc aac aac agg gcg Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala 500 505 510	1536

	tac ctc gaa gtc ccg atc ttc ttc cag tcg ccc agc acg aac tat cgt Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg 515 520 525	1584
	gtc agg gtc cgg tac gcc tca acc tca tcc ctc ccg gtc gat gtg gtc Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val 530 535 540	1632
	ttc ggc aac atc agc cac ccg acc acg ttt ccg gct acc gcc cga tcc Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser 545 550 555 560	1680
	ctg gac aat ctg caa agc aac gat ttc ggc tac att gac att gcc ggg Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly 565 570 575	1728
	acg ttc ctc ccg agc ctc ggc cca tcc atc ggc atc cgg ccc atg ctc Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu 580 585 590	1776
	tcc acc atc aac ctg atc gtg gat cgg ttt gag ttc atc cca gtg aca Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr 595 600 605	1824
[0053]	gcc act ttc gag gct gag tcc gac cta gag cgt gct cag aag gca gtc Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val 610 615 620	1872
	aat gct ctg ttt acc tcc acc aat cag ctc ggc att aag acc gat gtg Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val 625 630 635 640	1920
	acc gat tac cac att gac caa gtc tca aac ctc gtt gag tgc ctc tcg Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser 645 650 655	1968
	gat gag ttc tac ctt gat gag aag agg gag ctt tca gag aaa gtt aag Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys 660 665 670	2016
	cac gct aag aga ctc tcg gac gaa cgc aat ctg ttg caa gat ccc aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn 675 680 685	2064
	ttc aga ggg atc aac cgt cag cca gac cgg gga tgg cgc ggg tcc acg Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr 690 695 700	2112
	gac atc act atc cag ggc ggt gat gac gtc ttc aag gag aac tac gtg Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val	2160

705	710	715	720	
acc ctg ccg ggc	acc ttt gac gaa tgc	tac ccc act tac ctc	tac cag	2208
Thr Leu Pro Gly	Thr Phe Asp Glu Cys	Tyr Pro Thr Tyr Leu	Tyr Gln	
	725	730	735	
aag att gac gag tcc	aag ctc aag gcg ttc	aca cgc tac cag ctc	agg	2256
Lys Ile Asp Glu Ser	Lys Leu Lys Ala Phe	Thr Arg Tyr Gln Leu	Arg	
	740	745	750	
ggt tac atc gag gac	tcc caa gac ctg gaa	atc tac ctg atc cgc	tac	2304
Gly Tyr Ile Glu Asp	Ser Gln Asp Leu Glu	Ile Tyr Leu Ile Arg	Tyr	
	755	760	765	
aac gct aag cac gag	act gtc aac gtg ccc	ggc acc ggc agc ctg	tgg	2352
Asn Ala Lys His Glu	Thr Val Asn Val Pro	Gly Thr Gly Ser Leu	Trp	
	770	775	780	
ccc ttg tcc gct cag	agc cca atc ggc aag	tgc ggc gag ccc aac	cgc	2400
Pro Leu Ser Ala Gln	Ser Pro Ile Gly Lys	Cys Gly Glu Pro Asn	Arg	
	785	790	800	
tgc gcg ccc cac ctg	gaa tgg aac ccc gac	ctc gac tgt agc tgc	cgc	2448
Cys Ala Pro His Leu	Glu Trp Asn Pro Asp	Leu Asp Cys Ser Cys	Arg	
	805	810	815	
[0054]				
gac gga gag aag tgc	gcg cat cac tcc cac	cac ttc agc ctc gac	atc	2496
Asp Gly Glu Lys Cys	Ala His His Ser His	His Phe Ser Leu Asp	Ile	
	820	825	830	
gac gtc ggt tgc acc	gac ctt aac gag gat	ctg ggc gtt tgg gtg	atc	2544
Asp Val Gly Cys Thr	Asp Leu Asn Glu Asp	Leu Gly Val Trp Val	Ile	
	835	840	845	
ttc aag atc aag act	cag gac ggc cac gcc	cgc ctg gga aac ctg	gag	2592
Phe Lys Ile Lys Thr	Gln Asp Gly His Ala	Arg Leu Gly Asn Leu	Glu	
	850	855	860	
ttc ctg gag gag aag	ccc ctc gtt ggc gag	gcc ctg gcc cgc gtc	aag	2640
Phe Leu Glu Glu Lys	Pro Leu Val Gly Glu	Ala Leu Ala Arg Val	Lys	
	865	870	875	880
agg gcc gag aag aaa	tgg cgc gac aag cgc	gag aag ctg gag tgg	gag	2688
Arg Ala Glu Lys Lys	Trp Arg Asp Lys Arg	Glu Lys Leu Glu Trp	Glu	
	885	890	895	
acc aac atc gtg tac	aag gaa gcg aag gag	tca gtt gac gcc ctg	ttc	2736
Thr Asn Ile Val Tyr	Lys Glu Ala Lys Glu	Ser Val Asp Ala Leu	Phe	
	900	905	910	
gtc aac agc cag tac	gac cag ctc cag gca	gac aca aac atc get	atg	2784

	Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met	
	915 920 925	
	atc cat gcg gcc gac aag cgc gtc cac tcc atc cgc gag gcg tac ctg	2832
	Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu	
	930 935 940	
	ccc gag ctg tcc gtc atc ccc ggc gtc aac gcc gcg atc ttt gag gag	2880
	Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu	
	945 950 955 960	
	ctg gag ggc cgc atc ttc acc gcc ttc tcc ctc tac gac gca cgc aac	2928
	Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn	
	965 970 975	
	gtt atc aag aat ggc gac ttc aac aac ggg ctg tcc tgc tgg aat gtc	2976
	Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val	
	980 985 990	
	aag ggc cac gtg gac gtc gag gag cag aac aac cag cgc tca gtc ctg	3024
	Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu	
	995 1000 1005	
[0055]	gtc gtc ccg gag tgg gag gcc gaa gtc agc cag gaa gtc cgc gtc	3069
	Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val	
	1010 1015 1020	
	tgc cct gga cgc ggg tac atc ctg cgc gtc act gcc tac aag gaa	3114
	Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu	
	1025 1030 1035	
	ggc tac gga gag ggc tgc gtc acc atc cat gag atc gaa aac aac	3159
	Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn	
	1040 1045 1050	
	acg gat gag ctt aag ttc agc aac tgt gtt gaa gag gaa atc tac	3204
	Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr	
	1055 1060 1065	
	ccg aac aac acg gtc acc tgc aat gat tac acc gtc aac cag gag	3249
	Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu	
	1070 1075 1080	
	gaa tac ggt gga gct tac acc tcc cgc aac agg ggc tac aac gag	3294
	Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu	
	1085 1090 1095	
	gca ccc tet gtc ccg gcc gac tac gct tca gtc tac gaa gag aag	3339
	Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys	
	1100 1105 1110	

tcg tac acc gac gga cgc aga gag aac ccg tgt gag ttc aac cgc 3384
 Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125

ggc tac cgc gat tac acc ccg ctg cct gtc ggg tac gtc acc aaa 3429
 Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140

gag ctg gaa tac ttc cca gag acc gac aaa gtc tgg att gag atc 3474
 Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155

ggc gag acc gag ggc acg ttc atc gtg gac tcc gtc gaa ctc ctt 3519
 Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170

ctg atg gaa gag tga 3534
 Leu Met Glu Glu
 1175

- <210> 10
- <211> 1177
- <212> PRT
- [0056] <213> 人工序列(Artificial Sequence)

- <220>
- <223> 合成构建体

<400> 10

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 Leu 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

[0057] 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 Tyr 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 Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val
 225 230 235 240

Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro
 245 250 255

Ile Arg Ile Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val
 260 265 270

Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu
 275 280 285

Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr
 290 295 300

Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln
 305 310 315 320

Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro
 325 330 335

Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu
 340 345 350

Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg
 355 360 365

Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp
 370 375 380

[0058] Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr
 385 390 395 400

Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro
 405 410 415

Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser
 420 425 430

His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg
 435 440 445

Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile
 450 455 460

Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser
 465 470 475 480

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

Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala
 500 505 510

Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg
 515 520 525

Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val
 530 535 540

Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser
 545 550 555 560

Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly
 565 570 575

Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu
 580 585 590

[0059] Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr
 595 600 605

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val
 625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
 645 650 655

Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
 675 680 685

Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr

	690		695				700										
	Asp	Ile	Thr	Ile	Gln	Gly	Gly	Asp	Asp	Val	Phe	Lys	Glu	Asn	Tyr	Val	
	705					710					715					720	
	Thr	Leu	Pro	Gly	Thr	Phe	Asp	Glu	Cys	Tyr	Pro	Thr	Tyr	Leu	Tyr	Gln	
					725					730						735	
	Lys	Ile	Asp	Glu	Ser	Lys	Leu	Lys	Ala	Phe	Thr	Arg	Tyr	Gln	Leu	Arg	
				740					745						750		
	Gly	Tyr	Ile	Glu	Asp	Ser	Gln	Asp	Leu	Glu	Ile	Tyr	Leu	Ile	Arg	Tyr	
			755					760							765		
	Asn	Ala	Lys	His	Glu	Thr	Val	Asn	Val	Pro	Gly	Thr	Gly	Ser	Leu	Trp	
	770							775								780	
[0060]	Pro	Leu	Ser	Ala	Gln	Ser	Pro	Ile	Gly	Lys	Cys	Gly	Glu	Pro	Asn	Arg	
	785					790					795					800	
	Cys	Ala	Pro	His	Leu	Glu	Trp	Asn	Pro	Asp	Leu	Asp	Cys	Ser	Cys	Arg	
					805					810						815	
	Asp	Gly	Glu	Lys	Cys	Ala	His	His	Ser	His	His	Phe	Ser	Leu	Asp	Ile	
				820					825						830		
	Asp	Val	Gly	Cys	Thr	Asp	Leu	Asn	Glu	Asp	Leu	Gly	Val	Trp	Val	Ile	
			835					840						845			
	Phe	Lys	Ile	Lys	Thr	Gln	Asp	Gly	His	Ala	Arg	Leu	Gly	Asn	Leu	Glu	
	850							855						860			
	Phe	Leu	Glu	Glu	Lys	Pro	Leu	Val	Gly	Glu	Ala	Leu	Ala	Arg	Val	Lys	
	865					870					875					880	
	Arg	Ala	Glu	Lys	Lys	Trp	Arg	Asp	Lys	Arg	Glu	Lys	Leu	Glu	Trp	Glu	
					885					890						895	

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
 900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
 915 920 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
 930 935 940

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
 945 950 955 960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
 965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990

[0061] Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn
 1040 1045 1050

Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr
 1055 1060 1065

Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu
 1070 1075 1080

Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu
 1085 1090 1095

Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys
 1100 1105 1110

Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg
 1115 1120 1125

Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys
 1130 1135 1140

Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile
 1145 1150 1155

Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu
 1160 1165 1170

Leu Met Glu Glu
 1175

[0062]

<210> 11
 <211> 3534
 <212> DNA
 <213> 人工序列(Artificial Sequence)

<220>
 <223> 用于植物中的编码 BCW 003 的完全合成核苷酸序列

<220>
 <221> CDS
 <222> (1)..(3531)
 <223> BCW 003

<400> 11
 atg gac aac aac ccg aac atc aac gag tgc atc ccc tac aac tgc ctc 48
 Met Asp Asn Asn Pro Asn Ile Asn Glu Cys Ile Pro Tyr Asn Cys Leu
 1 5 10 15

tcc aac ccg gag gtc gag gtg ctg ggc ggc gaa agg atc gag acc ggc 96
 Ser Asn Pro Glu Val Glu Val Leu Gly Gly Glu Arg Ile Glu Thr Gly
 20 25 30

tac act ccc atc gac atc agc ctc agc ctg acc cag ttc ctg ctc tet 144
 Tyr Thr Pro Ile Asp Ile Ser Leu Ser Leu Thr Gln Phe Leu Leu Ser
 35 40 45

	gag ttc gtg ccc ggc gcg ggg ttc gtt ctc ggc ctg gtc gac atc atc Glu Phe Val Pro Gly Ala Gly Phe Val Leu Gly Leu Val Asp Ile Ile 50 55 60	192
	tgg ggc atc ttc ggt ccg agc cag tgg gac gcc ttt ctc gtt cag att Trp Gly Ile Phe Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile 65 70 75 80	240
	gag cag ctg atc aac cag cgc atc gag gag ttc gcc cgc aac cag gcg Glu Gln Leu Ile Asn Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala 85 90 95	288
	atc tcc cgg ctg gag ggc ctc tcc aac ctg tac caa atc tac gcc gag Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Gln Ile Tyr Ala Glu 100 105 110	336
	agc ttc cgg gag tgg gaa gcc gat ccg acc aac ccc gct ctc agg gag Ser Phe Arg Glu Trp Glu Ala Asp Pro Thr Asn Pro Ala Leu Arg Glu 115 120 125	384
	gag atg cgg att cag ttc aac gac atg aac tcc gct ctc acg act gcc Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala 130 135 140	432
[0063]	atc cca ctc ctc gct gtg cag aac tac caa gtg ccg ctc ctg tcc gtg Ile Pro Leu Leu Ala Val Gln Asn Tyr Gln Val Pro Leu Leu Ser Val 145 150 155 160	480
	tac gtg cag gcc gcc aat ctg cac ctc tcc gtc ctc cgg gac gtt agc Tyr Val Gln Ala Ala Asn Leu His Leu Ser Val Leu Arg Asp Val Ser 165 170 175	528
	gtg ttc ggg cag cgc tgg ggc ttc gac gcc gct acc atc aac tcc cgt Val Phe Gly Gln Arg Trp Gly Phe Asp Ala Ala Thr Ile Asn Ser Arg 180 185 190	576
	tac aac gat ctc act cgc ctc atc ggc aac tac acc gac tat gcc gtg Tyr Asn Asp Leu Thr Arg Leu Ile Gly Asn Tyr Thr Asp Tyr Ala Val 195 200 205	624
	cgc tgg tac aac act ggt ctt gag aga gtc tgg ggc ccg gac agc cgc Arg Trp Tyr Asn Thr Gly Leu Glu Arg Val Trp Gly Pro Asp Ser Arg 210 215 220	672
	gac tgg gtg cgc tac aac cag ttc cgg cgc gag ctg acc ctc acc gtg Asp Trp Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val 225 230 235 240	720
	ctc gac atc gta gcc ctc ttt ccc aac tac gac tcc cgg cgc tac ccg Leu Asp Ile Val Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro	768

	245	250	255	
	att cgc acc gtc agc cag ctc acc agg gag atc tac acc aac cct gtg			816
	Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val			
	260	265	270	
	ctg gag aac ttc gac ggc tcc ttt cgc ggg atg gcc caa cgc ata gag			864
	Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu			
	275	280	285	
	cag aac atc cgc caa cct cat ctg atg gac atc ctt aat tct atc acc			912
	Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr			
	290	295	300	
	atc tac act gac gtt cat cgc ggg ttt aac tac tgg tgc ggc cac caa			960
	Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln			
	305	310	315	320
	atc act gcg tgc ccc gtt ggt ttc tcc ggc ccg gag ttc gcg ttc cct			1008
	Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro			
	325	330	335	
	ctg ttc gga aac gcg ggc aat gcc gct cca ccc gta ttg gtg agc ctg			1056
	Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu			
	340	345	350	
[0064]	acc ggc ctc ggc atc ttc cgt aca ctg tct agc cct ctg tac aga agg			1104
	Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg			
	355	360	365	
	atc att ctt ggc agc ggt ccc aat aac cag gaa ctc ttc gtg ttg gac			1152
	Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp			
	370	375	380	
	ggc acc gag ttc agc ttc gcc agt ctt acg acc aat ttg ccc tcc aca			1200
	Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr			
	385	390	395	400
	atc tat cgc cag cgc ggt act gtg gac tcc ctt gat gtg ata cca cct			1248
	Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro			
	405	410	415	
	cag gac aac tct gtc cca cct cgc gcc ggt ttc tcc cac cgc ctc agc			1296
	Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser			
	420	425	430	
	cac gtc act atg ctg agt cag gct gcg gga gcc gtg tac acc ctt cgg			1344
	His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg			
	435	440	445	
	gct ccg acg ttt agc tgg cag cac agg agc gcg act acc acg aac atc			1392

	Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile 450	455	460		
	att gcg gct gac tcc atc act caa atc cct gcc gtt aag ggt cgc tcc Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser 465	470	475	480	1440
	atc atc aac aat ggg aca gtg atc tcg gga ccg ggc ttc acc ggc ggt Ile Ile Asn Asn Gly Thr Val Ile Ser Gly Pro Gly Phe Thr Gly Gly 485	490	495		1488
	gac ctg gtg agg ctg tac aac gcg gac ttc aac atc aac aac agg gcg Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala 500	505	510		1536
	tac ctc gaa gtc ccg atc ttc ttc cag tcg ccc agc acg aac tat cgt Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg 515	520	525		1584
	gtc agg gtc cgg tac gcc tca acc tca tcc ctc ccg gtc gat gtg gtc Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val 530	535	540		1632
[0065]	ttc ggc aac atc agc cac ccg acc acg ttt ccg gct acc gcc cga tcc Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser 545	550	555	560	1680
	ctg gac aat ctg caa agc aac gat ttc ggc tac att gac att gcc ggg Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly 565	570	575		1728
	acg ttc ctc ccg agc ctc ggc cca tcc atc ggc atc cgg ccc atg ctc Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu 580	585	590		1776
	tcc acc atc aac ctg atc gtg gat cgg ttt gag ttc atc cca gtg aca Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr 595	600	605		1824
	gcc act ttc gag gct gag tcc gac cta gag cgt gct cag aag gca gtc Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val 610	615	620		1872
	aat gct ctg ttt acc tcc acc aat cag ctc ggc att aag acc gat gtg Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val 625	630	635	640	1920
	acc gat tac cac att gac caa gtc tca aac ctc gtt gag tgc ctc tcg Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser 645	650	655		1968

	gat gag ttc tac ctt gat gag aag agg gag ctt tca gag aaa gtt aag Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys 660 665 670	2016
	cac gct aag aga ctc tcg gac gaa cgc aat ctg ttg caa gat ccc aac His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn 675 680 685	2064
	ttc aga ggg atc aac cgt cag cca gac cgg gga tgg cgc ggg tcc acg Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr 690 695 700	2112
	gac atc act atc cag ggc ggt gat gac gtc ttc aag gag aac tac gtg Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val 705 710 715 720	2160
	acc ctg ccg ggc acc ttt gac gaa tgc tac ccc act tac ctc tac cag Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln 725 730 735	2208
	aag att gac gag tcc aag ctc aag gcg ttc aca cgc tac cag ctc agg Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg 740 745 750	2256
[0066]	ggt tac atc gag gac tcc caa gac ctg gaa atc tac ctg atc cgc tac Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr 755 760 765	2304
	aac gct aag cac gag act gtc aac gtg ccc ggc acc ggc agc ctg tgg Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp 770 775 780	2352
	ccc ttg tcc gct cag age cca atc ggc aag tgc ggc gag ccc aac cgc Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg 785 790 795 800	2400
	tgc gcg ccc cac ctg gaa tgg aac ccc gac ctc gac tgt agc tgc cgc Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg 805 810 815	2448
	gac gga gag aag tgc gcg cat cac tcc cac cac ttc agc ctc gac atc Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile 820 825 830	2496
	gac gtc ggt tgc acc gac ctt aac gag gat ctg ggc gtt tgg gtg atc Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile 835 840 845	2544
	ttc aag atc aag act cag gac ggc cac gcc cgc ctg gga aac ctg gag Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu 850 855 860	2592

	1055	1060	1065	
	ccg aac aac acg gtc acc tgc aat gat tac acc gtc aac cag gag			3249
	Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu			
	1070	1075	1080	
	gaa tac ggt gga gct tac acc tcc cgc aac agg ggc tac aac gag			3294
	Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu			
	1085	1090	1095	
	gca ccc tct gtc ccg gcc gac tac gct tca gtc tac gaa gag aag			3339
	Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys			
	1100	1105	1110	
	tcg tac acc gac gga cgc aga gag aac ccg tgt gag ttc aac cgc			3384
	Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg			
	1115	1120	1125	
	ggc tac cgc gat tac acc ccg ctg cct gtc ggg tac gtc acc aaa			3429
	Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys			
	1130	1135	1140	
	gag ctg gaa tac ttc cca gag acc gac aaa gtc tgg att gag atc			3474
	Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile			
	1145	1150	1155	
[0068]	ggc gag acc gag ggc acg ttc atc gtg gac tcc gtc gaa ctc ctt			3519
	Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu			
	1160	1165	1170	
	ctg atg gaa gag tga			3534
	Leu Met Glu Glu			
	1175			
	<210> 12			
	<211> 1177			
	<212> PRT			
	<213> 人工序列(Artificial Sequence)			
	<220>			
	<223> 合成构建体			
	<400> 12			
	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

[0069] Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Thr Thr Ala
 130 135 140

Ile Pro Leu Leu 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 Tyr 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 Val Arg Tyr Asn Gln Phe Arg Arg Glu Leu Thr Leu Thr Val

225	230	235	240
Leu Asp Ile Val	Ala Leu Phe Pro Asn Tyr Asp Ser Arg Arg Tyr Pro		
	245	250	255
Ile Arg Thr Val Ser Gln Leu Thr Arg Glu Ile Tyr Thr Asn Pro Val			
	260	265	270
Leu Glu Asn Phe Asp Gly Ser Phe Arg Gly Met Ala Gln Arg Ile Glu			
	275	280	285
Gln Asn Ile Arg Gln Pro His Leu Met Asp Ile Leu Asn Ser Ile Thr			
	290	295	300
Ile Tyr Thr Asp Val His Arg Gly Phe Asn Tyr Trp Ser Gly His Gln			
305	310	315	320
Ile Thr Ala Ser Pro Val Gly Phe Ser Gly Pro Glu Phe Ala Phe Pro			
[0070]	325	330	335
Leu Phe Gly Asn Ala Gly Asn Ala Ala Pro Pro Val Leu Val Ser Leu			
	340	345	350
Thr Gly Leu Gly Ile Phe Arg Thr Leu Ser Ser Pro Leu Tyr Arg Arg			
	355	360	365
Ile Ile Leu Gly Ser Gly Pro Asn Asn Gln Glu Leu Phe Val Leu Asp			
	370	375	380
Gly Thr Glu Phe Ser Phe Ala Ser Leu Thr Thr Asn Leu Pro Ser Thr			
385	390	395	400
Ile Tyr Arg Gln Arg Gly Thr Val Asp Ser Leu Asp Val Ile Pro Pro			
	405	410	415
Gln Asp Asn Ser Val Pro Pro Arg Ala Gly Phe Ser His Arg Leu Ser			
	420	425	430

His Val Thr Met Leu Ser Gln Ala Ala Gly Ala Val Tyr Thr Leu Arg
 435 440 445

Ala Pro Thr Phe Ser Trp Gln His Arg Ser Ala Thr Thr Thr Asn Ile
 450 455 460

Ile Ala Ala Asp Ser Ile Thr Gln Ile Pro Ala Val Lys Gly Arg Ser
 465 470 475 480

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

Asp Leu Val Arg Leu Tyr Asn Ala Asp Phe Asn Ile Asn Asn Arg Ala
 500 505 510

Tyr Leu Glu Val Pro Ile Phe Phe Gln Ser Pro Ser Thr Asn Tyr Arg
 515 520 525

[0071] Val Arg Val Arg Tyr Ala Ser Thr Ser Ser Leu Pro Val Asp Val Val
 530 535 540

Phe Gly Asn Ile Ser His Pro Thr Thr Phe Pro Ala Thr Ala Arg Ser
 545 550 555 560

Leu Asp Asn Leu Gln Ser Asn Asp Phe Gly Tyr Ile Asp Ile Ala Gly
 565 570 575

Thr Phe Leu Pro Ser Leu Gly Pro Ser Ile Gly Ile Arg Pro Met Leu
 580 585 590

Ser Thr Ile Asn Leu Ile Val Asp Arg Phe Glu Phe Ile Pro Val Thr
 595 600 605

Ala Thr Phe Glu Ala Glu Ser Asp Leu Glu Arg Ala Gln Lys Ala Val
 610 615 620

Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Ile Lys Thr Asp Val
 625 630 635 640

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
 645 650 655

Asp Glu Phe Tyr Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
 660 665 670

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
 675 680 685

Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr
 690 695 700

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
 705 710 715 720

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
 725 730 735

[0072] Lys Ile Asp Glu Ser Lys Leu Lys Ala Phe Thr Arg Tyr Gln Leu Arg
 740 745 750

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
 755 760 765

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
 770 775 780

Pro Leu Ser Ala Gln Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg
 785 790 795 800

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu Asp Cys Ser Cys Arg
 805 810 815

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
 820 825 830

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
 835 840 845

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
 850 855 860

Phe Leu Glu Glu Lys Pro Leu Val Gly Glu Ala Leu Ala Arg Val Lys
 865 870 875 880

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Lys Leu Glu Trp Glu
 885 890 895

Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
 900 905 910

Val Asn Ser Gln Tyr Asp Gln Leu Gln Ala Asp Thr Asn Ile Ala Met
 915 920 925

Ile His Ala Ala Asp Lys Arg Val His Ser Ile Arg Glu Ala Tyr Leu
 930 935 940

[0073] Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
 945 950 955 960

Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
 965 970 975

Val Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Ser Cys Trp Asn Val
 980 985 990

Lys Gly His Val Asp Val Glu Glu Gln Asn Asn Gln Arg Ser Val Leu
 995 1000 1005

Val Val Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val
 1010 1015 1020

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu
 1025 1030 1035

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn

1040	1045	1050
Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu Glu Ile Tyr		
1055	1060	1065
Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Val Asn Gln Glu		
1070	1075	1080
Glu Tyr Gly Gly Ala Tyr Thr Ser Arg Asn Arg Gly Tyr Asn Glu		
1085	1090	1095
Ala Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr Glu Glu Lys		
1100	1105	1110
[0074]		
Ser Tyr Thr Asp Gly Arg Arg Glu Asn Pro Cys Glu Phe Asn Arg		
1115	1120	1125
Gly Tyr Arg Asp Tyr Thr Pro Leu Pro Val Gly Tyr Val Thr Lys		
1130	1135	1140
Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp Ile Glu Ile		
1145	1150	1155
Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val Glu Leu Leu		
1160	1165	1170
Leu Met Glu Glu		
1175		

昆虫物种	BCW 001	BCW 002	BCW 003
小地老虎	+	+	+
豆白缘切根虫	+	+	+
玉米穗虫	+	+	+
玉米螟	+	+	+
小蔗秆草螟	ND	ND	+
西南玉米秆草螟	+	ND	+
粉纹夜蛾	+	ND	+
大豆夜蛾	+	ND	+
草地夜蛾	-	-	-

图1

SEQ ID NO:2	MEEN-NQNQCVPYNCLNNAIEILEGDRISVGNTPIDISLSLVELLISEFVPGGGIITGL	59
SEQ ID NO:4	MDNNPNINECIPYNCLSNPEVEVLGGERIETGYTPIDISLSLQFLLSEFVPGAGFVLGL	60
SEQ ID NO:6	MDNNPNINECIPYNCLSNPEVEVLGGERIETGYTPIDISLSLQFLLSEFVPGAGFVLGL * * * * *	60
SEQ ID NO:2	LNIVWGFVGPSQWDAFLAQVEQLINQRISEAVRNTAIQELEGMARVYRXYATAFAEWERD	119
SEQ ID NO:4	VDIIWGFVGPSQWDAFLVQIEQLINQRIEEFARNQAISRLEGLSNLYQIYAESFREWEAD	120
SEQ ID NO:6	VDIIWGFVGPSQWDAFLVQIEQLINQRIEEFARNQAISRLEGLSNLYQIYAESFREWEAD * * * * *	120
SEQ ID NO:2	PNNTDLREAVRTQFTATETYISGRISVLKIQNFVQLLSVFAQAANLHLSLLRDVVFVGGQ	179
SEQ ID NO:4	PTNPALREEMRIQFNDMNSALTTAIPLLAVQNYQVPLLSVYVQAANLHLSVLRDVSVFGQ	180
SEQ ID NO:6	PTNPALREEMRIQFNDMNSALTTAIPLLAVQNYQVPLLSVYVQAANLHLSVLRDVSVFGQ * * * * *	180
SEQ ID NO:2	RWGFSTTTVNNYNDLTEEISTYTDYAVRWYNTGLERVWGPDSRDWVRYNQFRRELTTLV	239
SEQ ID NO:4	RWGFDAATINSRYNDLTRLIGNYTDYAVRWYNTGLERVWGPDSRDWVRYNQFRRELTTLV	240
SEQ ID NO:6	RWGFDAATINSRYNDLTRLIGNYTDYAVRWYNTGLERVWGPDSRDWVRYNQFRRELTTLV * * * * *	240
SEQ ID NO:2	LDIVALFPNYDSRRYPIRTVSQLTREIYTNPVLENFDGSGFRGMAQRIEQNIRQPHLMDIL	299
SEQ ID NO:4	LDIVALFPNYDSRRYPIRTVSQLTREIYTNPVLENFDGSGFRGMAQRIEQNIRQPHLMDIL	300
SEQ ID NO:6	LDIVALFPNYDSRRYPIRTVSQLTREIYTNPVLENFDGSGFRGMAQRIEQNIRQPHLMDIL *	300
SEQ ID NO:2	NSITIIYTDVHRGFNYWSGHQITASPVGFSGPEFAFPLFGNAGNAAPPVLVSLTGLGIFRT	359
SEQ ID NO:4	NSITIIYTDVHRGFNYWSGHQITASPVGFSGPEFAFPLFGNAGNAAPPVLVSLTGLGIFRT	360
SEQ ID NO:6	NSITIIYTDVHRGFNYWSGHQITASPVGFSGPEFAFPLFGNAGNAAPPVLVSLTGLGIFRT	360
SEQ ID NO:2	LSSPLYRRIILGSGPNNQELFVLDGTEFSFASLTNLPSTIYRQRGTVDSDLVIPPQDNS	419
SEQ ID NO:4	LSSPLYRRIILGSGPNNQELFVLDGTEFSFASLTNLPSTIYRQRGTVDSDLVIPPQDNS	420
SEQ ID NO:6	LSSPLYRRIILGSGPNNQELFVLDGTEFSFASLTNLPSTIYRQRGTVDSDLVIPPQDNS	420
SEQ ID NO:2	VPPRAGFSHRLSHVTMLSQAAGAVYTLRAPTFQHSRATTTNIIAADSITQIPAVKGRS	479
SEQ ID NO:4	VPPRAGFSHRLSHVTMLSQAAGAVYTLRAPTFQHSRATTTNIIAADSITQIPAVKGRS	480
SEQ ID NO:6	VPPRAGFSHRLSHVTMLSQAAGAVYTLRAPTFQHSRATTTNIIAADSITQIPAVKGRS	480
SEQ ID NO:2	IINNGTVISGPGFTGGDLVRLYNADFNINNRAYLEVPIFFQSPSTNYRVRVRYASTSSLP	539
SEQ ID NO:4	IINNGTVISGPGFTGGDLVRLYNADFNINNRAYLEVPIFFQSPSTNYRVRVRYASTSSLP	540
SEQ ID NO:6	IINNGTVISGPGFTGGDLVRLYNADFNINNRAYLEVPIFFQSPSTNYRVRVRYASTSSLP	540
SEQ ID NO:2	VDVVFVGNISHPTTFPATARSLDNLQSNDFGYIDIAAGTFPLSLGPSIGIRPMLSTINLIVD	599
SEQ ID NO:4	VDVVFVGNISHPTTFPATARSLDNLQSNDFGYIDIAAGTFPLSLGPSIGIRPMLSTINLIVD	600
SEQ ID NO:6	VDVVFVGNISHPTTFPATARSLDNLQSNDFGYIDIAAGTFPLSLGPSIGIRPMLSTINLIVD	600
SEQ ID NO:2	RFEFIPVTATFEAESDLERAQKAVNALFTSTNQLGIKTDVTDYHIDQVSNLVECLSDEFY	659
SEQ ID NO:4	RFEFIPVTATFEAESDLERAQKAVNALFTSTNQLGIKTDVTDYHIDQVSNLVECLSDEFY	660
SEQ ID NO:6	RFEFIPVTATFEAESDLERAQKAVNALFTSTNQLGIKTDVTDYHIDQVSNLVECLSDEFY	660

图2

SEQ ID NO:2	LDEKRELSEKVKHAKRLSDERNLLQDPNFRGINRQPDRGWRGSTDITIQGGDDVFKENYV	719
SEQ ID NO:4	LDEKRELSEKVKHAKRLSDERNLLQDPNFRGINRQPDRGWRGSTDITIQGGDDVFKENYV	720
SEQ ID NO:6	LDEKRELSEKVKHAKRLSDERNLLQDPNFRGINRQPDRGWRGSTDITIQGGDDVFKENYV	720
SEQ ID NO:2	TLPGTFDECYPTYLYQKIDESKCLKAYTRYQLRGYIEDSQDLEIYLI RYNAKHETVNVPGT	779
SEQ ID NO:4	TLPGTFDECYPTYLYQKIDESKCLKAFTRYQLRGYIEDSQDLEIYLI RYNAKHETVNVPGT	780
SEQ ID NO:6	TLPGTFDECYPTYLYQKIDESKCLKAFTRYQLRGYIEDSQDLEIYLI RYNAKHETVNVPGT	780
	*	
SEQ ID NO:2	GSLWPLSAQSPIGKCGEPNRCAPHLEWNPDLHCSCRDGEKCAHSHHFSLDIDVGCTDLN	839
SEQ ID NO:4	GSLWPLSAQSPIGKCGEPNRCAPHLEWNPDLDCSCRDGEKCAHSHHFSLDIDVGCTDLN	840
SEQ ID NO:6	GSLWPLSAQSPIGKCGEPNRCAPHLEWNPDLDCSCRDGEKCAHSHHFSLDIDVGCTDLN	840
	*	
SEQ ID NO:2	EDLGVWVIFKIKTQDGHARLGNLEFLEEKPLVGEALARVKRAEKKWRDKREKLQLETNIV	899
SEQ ID NO:4	EDLGVWVIFKIKTQDGHARLGNLEFLEEKPLVGEALARVKRAEKKWRDKREKLEWETNIV	900
SEQ ID NO:6	EDLGVWVIFKIKTQDGHARLGNLEFLEEKPLVGEALARVKRAEKKWRDKREKLEWETNIV	900
	**	
SEQ ID NO:2	YKEAKESVDALFVNSQYDQLQADTNIAMIHAADKRVHRIREAYLPELSVIPGVNADISEE	959
SEQ ID NO:4	YKEAKESVDALFVNSQYDQLQADTNIAMIHAADKRVHSIREAYLPELSVIPGVNAIFEE	960
SEQ ID NO:6	YKEAKESVDALFVNSQYDQLQADTNIAMIHAADKRVHSIREAYLPELSVIPGVNAIFEE	960
	* * *	
SEQ ID NO:2	LEGRIFTAFSLYDARNVIKNGDFNNGLLCWNVKGHV DVEEQNNHRSVLVVPEWEAEVSQE	1019
SEQ ID NO:4	LEGRIFTAFSLYDARNVIKNGDFNNGLSCWNVKGHV DVEEQNNQRSVLVPEWEAEVSQE	1020
SEQ ID NO:6	LEGRIFTAFSLYDARNVIKNGDFNNGLSCWNVKGHV DVEEQNNQRSVLVPEWEAEVSQE	1020
	* *	
SEQ ID NO:2	VRVCPGRGYILRVTAYKEGYGEGCVTIHEIENNTDELKFSNCVVEEVYPNNTVTCNDYTA	1079
SEQ ID NO:4	VRVCPGRGYILRVTAYKEGYGEGCVTIHEIENNTDELKFSNCVVEE IYPNNTVTCNDYTV	1080
SEQ ID NO:6	VRVCPGRGYILRVTAYKEGYGEGCVTIHEIENNTDELKFSNCVVEE IYPNNTVTCNDYTV	1080
	* *	
SEQ ID NO:2	NQEEYEGTYTSRNRGYDEAYESNSSVPAEYASVYEEKVYTDGRRGNPCESNRGYDYTPL	1139
SEQ ID NO:4	NQEEYGGAYTSRNRGYNEAP----SVPADYASVYEEKSYTDGRRENPCFNRYRDTPL	1136
SEQ ID NO:6	NQEEYGGAYTSRNRGYNEAP----SVPADYASVYEEKSYTDGRRENPCFNRYRDTPL	1136
	* * * * * * * * *	
SEQ ID NO:2	PAGYVTKELEYFPETDKVWIEIGETEGTFIVDSVELLLMEE	1180
SEQ ID NO:4	PVGYVTKELEYFPETDKVWIEIGETEGTFIVDSVELLLMEE	1177
SEQ ID NO:6	PVGYVTKELEYFPETDKVWIEIGETEGTFIVDSVELLLMEE	1177
	*	

图2(接续)