(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116670155 A
(43)申请公布日 2023.08.29

(21)申请号 202180085702.2
(22)申请日 2021.12.20
(30)优先权数据
63/128,775 2020.12.21 US
(85)PCT国际申请进入国家阶段日
2023.06.19
(86)PCT国际申请的申请数据
PCT/US2021/064286 2021.12.20
(87)PCT国际申请的公布数据
WO2022/140227 EN 2022.06.30
(71)申请人 孟山都技术公司
地址 美国密苏里州
(72)发明人 D·J·伯温 C·A·蔡 A·R·豪 M·S·马伦戈 J·迈耶 J·S·米利甘 B·E·韦纳
(74)专利代理机构 北京市金杜律师事务所 11256
专利代理师 孟凡宏
(51)Int.Cl.
C07K 14/195(2006.01)

权利要求书4页 说明书27页
序列表28页 附图2页

(54)发明名称

新型昆虫抑制性蛋白质

(57)摘要

本发明公开了表现出针对鳞翅目害虫物种的毒性活性的杀害虫性蛋白质,所述蛋白质包括但不限于TIC4029、TIC4029_1和TIC4029_8。提供了DNA构建体,所述DNA构建体含有编码所公开的杀害虫性蛋白质中的一种或多种的重组核酸序列。提供了对鳞翅目侵扰具有耐受性的转基因植物、植物细胞、种子和植物部分,所述转基因植物、植物细胞、种子和植物部分含有编码所述TIC4029类的杀害虫性蛋白质的重组核酸序列。还提供了对于检测生物样品中存在所述重组核酸序列或所述TIC4029类的蛋白质的方法,以及使用TIC4029、TIC4029_1和TIC4029_8杀害虫性蛋白质中的任一种来防治鳞翅目物种害虫的方法。

CN 116670155 A

1.一种重组核酸分子,所述重组核酸分子包含异源性启动子,所述异源性启动子可操作地连接至编码杀害虫性蛋白质或其杀害虫性片段的多核苷酸区段,其中,任选地:

a.所述杀害虫性蛋白质包含SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列;

b.所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有至少87%、或90%、或95%、或98%、或99%、或约100%同一性的氨基酸序列;或者

c.所述多核苷酸区段在严格杂交条件下与具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ IDNO:8或SEQ ID NO:9中所示的核苷酸序列的多核苷酸杂交。

2.如权利要求1所述的重组核酸分子:

a.在植物细胞中表达以产生杀害虫有效量的所述杀害虫性蛋白质或杀害虫性片段;或者

b.与载体可操作地连接,并且所述载体选自由以下组成的组:质粒、噬菌粒、杆粒、粘粒以及细菌或酵母人工染色体。

3.如权利要求1所述的重组核酸分子,所述重组核酸分子存在于宿主细胞中,其中所述宿主细胞选自由细菌细胞和植物细胞组成的组。

4.如权利要求3所述的重组核酸分子,其中所述细菌宿主细胞来自选自由以下组成的组的细菌属:农杆菌属、根瘤菌属、芽孢杆菌属、短芽孢杆菌属、埃希氏菌属、假单胞菌属、克雷伯氏菌属、泛菌属和欧文氏菌属。

5.如权利要求4所述的重组核酸分子,其中所述芽孢杆菌是蜡样芽孢杆菌或苏云金芽孢杆菌,所述短芽孢杆菌属是侧孢短芽孢杆菌,并且所述埃希氏菌属是大肠杆菌。

6.如权利要求2所述的重组核酸,其中所述植物细胞是双子叶植物细胞或单子叶植物细胞。

7.如权利要求6所述的重组核酸,其中所述植物细胞选自由以下组成的组:苜蓿、香蕉、大麦、豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、道格拉斯冷杉、茄子、桉树、亚麻、大蒜、葡萄、啤酒花、韭葱、莴苣、火炬松、小米、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧草、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、水稻、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、西葫芦、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香树、甘薯、柳枝稷、茶、烟草、番茄、黑小麦、草坪草、西瓜和小麦植物细胞。

8.如权利要求1所述的重组核酸分子,其中所述蛋白质表现出针对鳞翅目昆虫的活性。

9.如权利要求8所述的重组核酸分子,其中所述鳞翅目昆虫选自由以下组成的组:豆芽蛾(豆茎蛾)、黑粘虫(考斯夜蛾)、黑地蚕(小地老虎)、玉米穗虫(玉米耳虫)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)、南美豆荚虫(大豆荚虫)、南部行军虫(南部灰翅夜蛾)、大豆尺蠖(大豆尺蠖)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾)、烟草蚜虫(烟芽夜蛾)和天鹅绒豆毛虫(黎豆夜蛾)。

10.一种包含如权利要求1所述的重组核酸分子的植物,或其一部分。

11.如权利要求10所述的植物,其中所述植物是单子叶植物或双子叶植物,或其一部

分。

12.如权利要求10所述的植物，其中所述植物选自由以下组成的组：苜蓿、香蕉、大麦、豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、道格拉斯冷杉、茄子、桉树、亚麻、大蒜、葡萄、啤酒花、韭葱、莴苣、火炬松、小米、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧草、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、水稻、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、西葫芦、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香树、甘薯、柳枝稷、茶、烟草、番茄、黑小麦、草坪草、西瓜和小麦。

13.如权利要求10所述的植物，其中所述其一部分是种子，并且其中所述种子包含所述重组核酸分子。

14.一种昆虫抑制性组合物，所述昆虫抑制性组合物包含如权利要求1所述的重组核酸分子。

15.如权利要求14所述的昆虫抑制性组合物，其还包含编码不同于所述杀害虫性蛋白质或其杀害虫性片段的至少一种其他杀害虫剂的核苷酸序列。

16.如权利要求15所述的昆虫抑制性组合物，其中所述至少一种其他杀害虫剂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、化学分子和辅助蛋白，并且

其中所述至少一种其他杀害虫剂，与所述杀害虫性蛋白质或其杀害虫性片段相比，对相同的害虫具有毒性。

17.如权利要求15所述的昆虫抑制性组合物，其中所述至少一种其他杀害虫剂表现出针对鳞翅目、鞘翅目或半翅目的一种或多种害虫物种的活性。

18.如权利要求15所述的昆虫抑制性组合物，其中所述至少一种其他杀害虫性蛋白质选自由以下组成的组：Cry1A、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B、Cry1C、Cry1C变体、Cry1D、Cry1E、Cry1F、Cry1A/F嵌合体、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab、Cry2Ae、Cry3、Cry3A变体、Cry3B、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry34、Cry35、Cry43A、Cry43B、Cry51Aa1、ET29、ET33、ET34、ET35、ET66、ET70、TIC400、TIC407、TIC417、TIC431、TIC800、TIC807、TIC834、TIC853、TIC900、TIC901、TIC1201、TIC1415、TIC2160、TIC3131、TIC836、TIC860、TIC867、TIC869、TIC1100、VIP3A、VIP3B、VIP3Ab、AXMI-AXMI-、AXMI-88、AXMI-97、AXMI-102、AXMI-112、AXMI-117、AXMI-100、AXMI-115、AXMI-113和AXMI-005、AXMI134、AXMI-150、AXMI-171、AXMI-184、AXMI-196、AXMI-204、AXMI-207、AXMI-209、AXMI-205、AXMI-218、AXMI-220、AXMI-221z、AXMI-222z、AXMI-223z、AXMI-224z和AXMI-225z、AXMI-238、AXMI-270、AXMI-279、AXMI-345、AXMI-335、AXMI-R1及其变体、IP3及其变体、DIG-3、DIG-5、DIG-10、DIG-657、DIG-11蛋白、IPD102Aa及其同源物、IPD110Aa及其同源物、TIC868、Cry1Da1_7、BCW003、TIC1100、TIC867、TIC867_23、TIC6757、TIC7941、IDP072Aa、TIC5290、TIC3668、TIC3669、TIC3670、IPD103及其同源物、PIP-50和PIP-65以及它们的同源物、PIP-83及其同源物以及Cry1B.34。

19.如权利要求14所述的昆虫抑制性组合物，其包含表达杀昆虫有效量的所述杀害虫性蛋白质的植物细胞。

20.一种由如权利要求10所述的植物或其一部分产生的商品性产品，其中所述商品性产品包含可检测的量的所述重组核酸分子、所述杀害虫性蛋白质或其杀害虫性片段。

21.如权利要求20所述的商品性产品，其选自由以下组成的组：通过谷物装卸运输机装袋的商品玉米、玉米片、玉米饼、玉米粉、玉米粕、玉米糖浆、玉米油、玉米青贮料、玉米淀粉、玉米谷物等等，以及对应的大豆、水稻、小麦、高粱、木豆、花生、水果、甜瓜以及蔬菜商品性产品，在适用的情况下包括汁、浓缩物、果酱、胶冻、柑橘酱以及含有可检测量的本申请的此类多核苷酸和/或多肽的此类商品性产品的其他可食用形式，全棉籽或加工棉籽、棉油、皮棉、针对饲料或食品加工的种子和植株部分、纤维、纸张、生物质以及燃料产品，诸如来源于棉油的燃料或来源于轧棉机废料的颗粒燃料，全大豆籽或加工大豆籽、大豆油、大豆蛋白、大豆粕、大豆粉、大豆片、大豆麸、大豆浆、大豆干酪、大豆酒、包含大豆的动物饲料、包含大豆的纸张、包含大豆的乳膏、大豆生物质以及使用大豆植株和大豆植株部分产生的燃料产品。

22.一种产生子代种子的方法，所述子代种子包含如权利要求1所述的重组核酸分子，所述方法包括：

a.种植第一种子，所述第一种子包含所述重组核酸分子；

b.从步骤a.所述的种子培植植株；以及

c.从所述植株收获所述子代种子，其中所述收获的种子包含所述重组核酸分子。

23.一种对昆虫侵扰具有耐受性的植物，其中所述植物的细胞包含如权利要求1所述的重组核酸分子。

24.一种用于防治鳞翅目物种害虫或害虫侵扰的方法，所述方法包括使所述害虫接触杀昆虫有效量的：

a.杀害虫性蛋白质，所述杀害虫性蛋白质包含如SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列；或者

b.一种或多种杀害虫性蛋白质，所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有至少87％、或90％、或95％、或98％、或99％、或约100％同一性的氨基酸序列。

25.一种检测包含植物基因组DNA的样品中存在如权利要求1所述的重组核酸分子的方法，所述方法包括：

a.使所述样品与核酸探针接触，所述核酸探针在严格杂交条件下与所述多核苷酸区段杂交，并且在此类杂交条件下不与来自另一种同基因植物的不包含所述多核苷酸区段的基因组DNA杂交，其中所述探针与SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9中所示的核酸序列；或者编码这样的杀害虫性蛋白质的序列同源或互补，所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有至少87％、或90％、或95％、或98％、或99％、或约100％同一性的氨基酸序列；

b.使所述样品和所述探针经受严格杂交条件；以及

c.检测所述探针与所述重组核酸分子的杂交；

其中检测出所述杂交证实了所述植物样品中存在包含所述多核苷酸片段的植物基因组DNA。

26.一种检测包含蛋白质的样品中存在杀害虫性蛋白质或其片段的方法，其中所述杀害虫性蛋白质包含SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10的氨基酸序列；或者所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有

至少87％、或90％、或95％、或98％、或99％、或约100％同一性的氨基酸序列，所述方法包括：

a.使所述样品与免疫反应性抗体接触，所述免疫反应性抗体对结合所述杀害虫性蛋白质或其片段具有特异性；以及

b.检测所述抗体与所述杀害虫性蛋白质或其片段的所述结合。

27.如权利要求26所述的方法，其中所述检测步骤包括ELISA或蛋白质印迹。

28.一种杀害虫有效量的蛋白质，所述蛋白质包含如SEQ ID NO:2、SEQ ID NO:4和SEQ ID NO:10中所示的氨基酸序列。

新型昆虫抑制性蛋白质

[0001] 相关申请的交叉引用

[0002] 本申请要求2020年12月21日提交的美国临时申请63/128,775的权益,该临时申请以引用的方式整体并入本文。

[0003] 序列表的并入

[0004] 名称为－MONS484WO-sequence_listing.txt”的文件包含DOS操作系统生成的计算机可读形式的序列表,该文件创建于2021年12月9日。该文件为48,197字节(在MS-Windows®中测量),通过电子提交(使用美国专利局EFS-Web提交系统)方式与本申请同时提交,并且以引用的方式整体并入本文。

技术领域

[0005] 本发明整体涉及昆虫抑制性蛋白质领域。公开了一类新型蛋白质,所述蛋白质表现出针对作物植株和种子的农业相关害虫,尤其是鳞翅目昆虫物种的昆虫抑制性活性。提供了植物、植物部分和种子,包括植物和微生物细胞,以及含有编码所公开的毒素蛋白中的一种或多种的重组多核苷酸构建体的载体。

背景技术

[0006] 提高具有重要农业意义的植物(包括玉米、大豆、甘蔗、水稻、小麦、蔬菜和棉花等等)的作物产量变得越来越重要。除了越来越需要农产品来为不断增长的人口提供食物、衣服以及提供能量之外,气候相关的影响以及不断增长的人口对将土地用于农业实践以外的用途带来的压力预计会减少可用于耕作的可耕地的数量。这些因素产生了对粮食安全的严峻预测,尤其是在植物生物技术和农艺实践无重大改进的情况下。鉴于这些压力,技术、农业技术和害虫管理方面的环境可持续性改进是在有限的可用于耕作的可耕地上扩大作物生产的重要工具。

[0007] 昆虫,尤其是鳞翅目昆虫,被认为是损害大田作物,从而降低受侵染区域的作物产量的主要原因。对农业产生负面影响的鳞翅目害虫物种包括但不限于:豆芽蛾(豆茎蛾)、黑粘虫(考斯夜蛾)、黑地蚕(小地老虎)、玉米穗虫(玉米耳虫)、棉叶虫(棉叶波纹夜蛾)、小菜蛾(小菜蛾)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)(包括Cry1Ab、Cry1Ac和Cry1Fa耐受性草地贪夜蛾)、旧大陆棉铃虫(OWB,棉铃虫)、南部行军虫(南部灰翅夜蛾)、大豆尺蠖(大豆尺蠖)、斑点棉铃虫(棉斑实蛾)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾(Rachiplusia nu))、烟草蚜虫(烟芽夜蛾)、烟草地老虎(斜纹夜蛾,也称为集群毛虫)、西部豆地老虎(西部豆角夜蛾)和天鹅绒豆毛虫(黎豆夜蛾)。

[0008] 从历史上看,合成化学杀昆虫剂的密集应用依赖于在农业中作为害虫防治剂。对此类杀昆虫剂的无差别毒性作用的担忧、对环境和人类健康的担忧以及新出现的耐受性问题,都刺激了生物杀害虫剂的研究和开发。这项研究工作的结果是各种昆虫病原性微生物物种(包括细菌)的逐步发现和使用。

[0009] 当昆虫病原性细菌,尤其是属于芽孢杆菌属(Bacillus)的细菌的潜力被发现并且

开发为生物害虫防治剂时,生物防治模式发生了变化。自从分子生物学诞生以来,苏云金芽孢杆菌(Bacillus thuringiensis,Bt)细菌菌株就被用作杀害虫性蛋白质的来源,并且发现某些Bt菌株针对特定昆虫表现出高水平的毒性。已知Bt菌株产生δ-内毒素,所述δ-内毒素在孢子形成开始时和稳定生长期期间定位于伴孢结晶包涵体(例如,Cry蛋白)中,并且还已知Bt菌株产生分泌型杀昆虫性蛋白质。在被易感昆虫摄取时,δ-内毒素以及分泌型毒素可以在中肠上皮细胞表面发挥作用,破坏细胞膜,导致细胞破裂和死亡。编码杀昆虫性蛋白质的基因也在除Bt之外的细菌菌种中得以鉴定,这些细菌菌种包括其他芽孢杆菌和多种另外的细菌菌种,诸如假单胞菌属(Pseudomonas)、沙雷氏菌属(Serratia)、致病杆菌属(Xenorhabdus)和光杆状菌属(Photorhabdus)物种,以及其他芽孢杆菌,包括但不限于侧孢短芽孢杆菌(Brevibacillus laterosporus)、球形赖氨酸芽孢杆菌(Lysinibacillus sphaericus)(—Ls”,此前称为球形芽孢杆菌(Bacillus sphaericus))、日本甲虫类芽孢杆菌(Paenibacillus popilliae)和缓病类芽孢杆菌(Paenibacillus lentimorbus)。此外,杀昆虫性毒素还从多种非细菌来源(包括蕨类植物和其他植物相关物种以及蛛形纲动物毒液)得以鉴定,并且在一些害虫物种的饮食中输送dsRNA已经被鉴定为有效的害虫管理策略。

[0010] 用于本发明的目的的优选的结晶和分泌型可溶性杀昆虫毒素对目标害虫物种具有高度特异性,并且作为化学杀昆虫剂的替代品已经获得全世界的认可。例如,杀昆虫性毒素蛋白已经被用于各种农业应用,以保护具有重要农业意义的植物免受昆虫侵扰、减少对化学杀害虫剂施用的需要并且提高产量。杀昆虫性毒素蛋白可用于通过机械方法来防治作物植物的农业相关害虫,所述机械方法诸如喷洒以将含有各种细菌菌株的微生物制剂分散到植物表面,或者通过使用遗传转化技术来产生表达杀昆虫性毒素蛋白的转基因植株和种子。

[0011] 表达杀昆虫性毒素蛋白的转基因植物的使用已经在全球范围内得到认可。例如,在2016年,种植了2310万公顷表达Bt毒素的转基因作物,并且种植了7540万公顷表达复合有除草剂耐受性状的Bt毒素的转基因作物(ISAAA.2016.Global Status of Commercialized Biotech/GM Crops:2016.ISAAA Brief No.52.ISAAA:Ithaca,NY)。转基因防昆虫作物的全球使用以及可在这些作物中使用的杀昆虫性毒素蛋白的有限数量,已经对现有的昆虫等位基因产生了选择压力,所述等位基因赋予目前使用的杀昆虫性蛋白质以耐受性。

[0012] 目标害虫对杀昆虫性毒素蛋白的耐受性发展产生了对新型杀昆虫性毒素蛋白的发现和开发的持续需求,所述新型杀昆虫性毒素蛋白可用于管理害虫种群,包括具有耐受性等位基因的那些。具有管理功效并且表现出对更广谱的易感昆虫物种的防治的新型蛋白毒素将减少可以产生耐受性等位基因的存活昆虫的数量。此外,对同一种昆虫害虫具有毒性并且显示出不同的作用模式或者两种或更多种不同的毒性作用模式(例如,编码靶向用于抑制的必需基因的dsRNA的转基因与编码肽或蛋白毒素的转基因相结合,这两种转基因均对同一昆虫物种具有毒性)的两种或更多种转基因杀昆虫性毒素蛋白在一种植物中的使用,降低了在任何单一目标昆虫物种中产生耐受性的概率。另外,自限性技术(诸如Oxitec Ltd提供的那些)的使用,当与本发明的蛋白质一起使用时,可以改善赋予表达本发明的蛋白质的转基因作物的性状的持久性(Zhou等人2018.Combining the high-dose/refuge

strategy and self-limiting transgenic insects in resistance management—atest in experimental mesocosms.Evol Appl 11(5):727-738;Alphey等人2009.Combining pest control and resistance management:synergy of engineered insects with Bt crops.Journal of Economic Entomology,102:717-732)。

[0013] 因此,本发明人在本文中公开了一种来自苏云金芽孢杆菌的新型蛋白质和示例性重组蛋白质,它们各自表现出针对目标鳞翅目物种的杀昆虫活性,尤其是针对以下物种的杀昆虫活性:黑粘虫(考斯夜蛾)、黑地蚕(小地老虎)、玉米穗虫(玉米耳虫)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)、南美豆荚虫(大豆荚虫)、南部行军虫(南部灰翅夜蛾)、大豆尺蠖(大豆尺蠖)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾)、烟草蚜虫(烟芽夜蛾)和天鹅绒豆毛虫(黎豆夜蛾)。

发明内容

[0014] 本文公开了一种具有昆虫抑制性活性的新型杀害虫性蛋白质TIC4029及其截短变体,所述蛋白质及其截短变体显示表现出针对作物植物的一种或多种害虫的抑制性活性。TIC4029蛋白毒素类中的TIC4029蛋白和变体蛋白可以单独使用或者与其他杀昆虫性蛋白质和毒剂组合用于制剂和植物体,从而为目前在农业系统中使用的杀昆虫性蛋白质和杀昆虫剂化学物质提供替代品。

[0015] 在一个实施方案中,本申请公开了一种重组核酸分子,所述重组核酸分子包含异源性启动子,所述异源性启动子可操作地连接至编码杀害虫性蛋白质或其杀害虫性片段的多核苷酸区段,其中所述杀害虫性蛋白质包含SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10的氨基酸序列;或者所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有至少87%、或90%、或95%、或98%、或99%、或约100%同一性的氨基酸序列;或者所述多核苷酸区段在严格杂交条件下与具有SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9中所示的核苷酸序列的多核苷酸杂交。所述重组核酸分子可以包含功能是在植物中表达所述杀害虫性蛋白质的序列,并且所述序列在植物细胞中表达时产生杀害虫有效量的杀害虫性蛋白质或其杀害虫性片段。

[0016] 在本申请的另一个实施方案中,所述重组核酸分子存在于细菌或植物宿主细胞内。所设想的细菌宿主细胞包括至少以下属:农杆菌属(Agrobacterium)、根瘤菌属(Rhizobium)、芽孢杆菌属(Bacillus)、短芽孢杆菌属(Brevibacillus)、埃希氏菌属(Escherichia)、假单胞菌属(Pseudomonas)、克雷伯氏菌属(Klebsiella)、泛菌属(Pantoea)和欧文氏菌属(Erwinia)。在某些实施方案中,所述芽孢杆菌属(Bacillus)是蜡样芽孢杆菌(Bacillus cereus)或苏云金芽孢杆菌,所述短芽孢杆菌属(Brevibacillus)是侧孢短芽孢杆菌(Brevibacillus laterosporus),或者所述埃希氏菌属(Escherichia)物种是大肠杆菌(Escherichia coli)。所设想的植物宿主细胞包括双子叶植物细胞和单子叶植物细胞。所设想的植物细胞还包括苜蓿、香蕉、大麦、豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花(Gossypium sp.)、葫芦、黄瓜、道格拉斯冷杉、茄子、桉树、亚麻、大蒜、葡萄、啤酒花、韭葱、莴苣、火炬松、小米、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧草、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、水稻、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、西葫芦(squash)、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香树(sweet gum)、甘薯、柳枝稷、茶、烟草、番茄、黑小麦、草坪草、西瓜和小麦植物细胞。

[0017] 在另一个实施方案中,所述杀害虫性蛋白质表现出针对鳞翅目昆虫的活性,所述鳞翅目昆虫包括至少豆芽蛾(豆茎蛾)、黑粘虫(考斯夜蛾)、黑地蚕(小地老虎(Agrotis ipsilon))、玉米穗虫(玉米耳虫)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)、南美豆荚虫(大豆荚虫)、南部行军虫(南部灰翅夜蛾)、大豆尺蠖(大豆尺蠖)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾)、烟草蚜虫(烟芽夜蛾)和天鹅绒豆毛虫(黎豆夜蛾)。

[0018] 本申请还设想了包含重组核酸分子的植物或植物部分,所述重组核酸分子编码所述TIC4029毒素蛋白类的杀害虫性蛋白质或其片段。设想了双子叶植物和单子叶植物二者。在另一个实施方案中,所述植物还选自由以下组成的组:苜蓿、香蕉、大麦、豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、道格拉斯冷杉、茄子、桉树、亚麻、大蒜、葡萄、啤酒花、韭葱、莴苣、火炬松、小米、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧草、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、水稻、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、西葫芦、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香树、甘薯、柳枝稷、茶、烟草、番茄、黑小麦、草坪草、西瓜和小麦。

[0019] 在某些实施方案中,公开了包含所述重组核酸分子的种子。

[0020] 在又一个实施方案中,设想了一种昆虫抑制性组合物,所述昆虫抑制性组合物包含本申请中公开的重组核酸分子。所述昆虫抑制性组合物可以另外包含编码不同于所述杀害虫性蛋白质的至少一种其他杀害虫剂的核苷酸序列。在某些实施方案中,所述至少一种其他杀害虫剂选自由以下组成的组:昆虫抑制性蛋白质、昆虫抑制性dsRNA分子和辅助蛋白。还设想了,所述昆虫抑制性组合物中的所述至少一种其他杀害虫剂表现出针对鳞翅目(Lepidoptera)、鞘翅目(Coleoptera)或半翅目(Hemiptera)的一种或多种害虫物种的活性。所述昆虫抑制性组合物中的所述至少一种其他杀害虫剂在一个实施方案中选自由以下组成的组:Cry1A、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B、Cry1C、Cry1C变体、Cry1D、Cry1E、Cry1F、Cry1A/F嵌合体、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab、Cry2Ae、Cry3、Cry3A变体、Cry3B、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry34、Cry35、Cry43A、Cry43B、Cry51Aa1、ET29、ET33、ET34、ET35、ET66、ET70、TIC400、TIC407、TIC417、TIC431、TIC800、TIC807、TIC834、TIC853、TIC900、TIC901、TIC1201、TIC1415、TIC2160、TIC3131、TIC836、TIC860、TIC867、TIC869、TIC1100、VIP3A、VIP3B、VIP3Ab、AXMI-AXMI-、AXMI-88、AXMI-97、AXMI-102、AXMI-112、AXMI-117、AXMI-100、AXMI-115、AXMI-113和AXMI-005、AXMI134、AXMI-150、AXMI-171、AXMI-184、AXMI-196、AXMI-204、AXMI-207、AXMI-209、AXMI-205、AXMI-218、AXMI-220、AXMI-221z、AXMI-222z、AXMI-223z、AXMI-224z和AXMI-225z、AXMI-238、AXMI-270、AXMI-279、AXMI-345、AXMI-335、AXMI-R1及其变体、IP3及其变体、DIG-3、DIG-5、DIG-10、DIG-657、DIG-11蛋白、IPD102Aa及其同源物、IPD110Aa及其同源物、TIC868、Cry1Da1_7、BCW003、TIC1100、TIC867、TIC867_23、TIC6757、TIC7941、IDP072Aa、TIC5290、TIC3668、TIC3669、TIC3670、IPD103及其同源物、PIP-50和PIP-65以及它们的同源

物、PIP-83及其同源物以及Cry1B.34。

[0021] 还设想了商品性产品，所述商品性产品包含可检测量的本申请中公开的重组核酸分子和毒素蛋白。此类商品性产品包括通过谷物装卸运输机装袋的商品玉米、玉米片、玉米饼、玉米粉、玉米粕、玉米糖浆、玉米油、玉米青贮料、玉米淀粉、玉米谷物等等，以及对应的大豆、水稻、小麦、高粱、木豆、花生、水果、甜瓜以及蔬菜商品性产品，在适用的情况下包括汁、浓缩物、果酱、胶冻、柑橘酱(marmalade)以及含有可检测量的本申请的此类多核苷酸和/或多肽的此类商品性产品的其他可食用形式，全棉籽或加工棉籽、棉油、皮棉、针对饲料或食品加工的种子和植株部分、纤维、纸张、生物质以及燃料产品，诸如来源于棉油的燃料或来源于轧棉机废料的颗粒燃料，全大豆籽或加工大豆籽、大豆油、大豆蛋白、大豆粕、大豆粉、大豆片、大豆麸、大豆浆、大豆干酪、大豆酒、包含大豆的动物饲料、包含大豆的纸张、包含大豆的乳膏、大豆生物质以及使用大豆植株和大豆植株部分产生的燃料产品。

[0022] 本申请还设想了一种产生种子的方法，所述种子包含所述重组核酸分子和来自所述TIC4029蛋白毒素类的蛋白毒素。所述方法包括种植至少一种包含本申请中公开的重组核酸分子的种子；从所述种子培植植株；以及从所述植株收获种子，其中所述收获的种子包含参考重组核酸分子。

[0023] 在另一个示例性实施方案中，提供了一种对鳞翅目昆虫侵扰具有耐受性的植物，其中所述植物的细胞包含本文公开的重组核酸分子。

[0024] 本申请还公开了用于防治鳞翅目物种害虫和防治植物，尤其是作物植物的鳞翅目物种害虫侵染的方法。所述方法包括，在一个实施方案中，首先使所述害虫与杀昆虫有效量的SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的杀害虫性蛋白质接触；或者使所述害虫与杀昆虫有效量的一种或多种这样的杀害虫性蛋白质接触，所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中所示的氨基酸序列具有至少87％、或90％、或95％、或98％、或99％、或约100％同一性的氨基酸序列。

[0025] 本文还提供了一种检测TIC4029类的重组核酸分子的存在的方法，其中所述方法包括使核酸样品与核酸探针接触，所述核酸探针在严格杂交条件下与来自一种植物的基因组DNA杂交，所述基因组DNA包含编码本文提供的杀害虫性蛋白质或其片段的多核苷酸区段，并且在此类杂交条件下不与来自另一种同基因植物的不包含所述区段的基因组DNA杂交，其中所述探针与SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9，或者编码这样的杀害虫性蛋白质的序列同源或互补，所述杀害虫性蛋白质包含与SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10具有至少87％、或90％、或95％、或98％、或99％、或约100％氨基酸序列同一性的氨基酸序列；使所述样品和探针经受严格杂交条件；以及检测所述探针与所述样品的DNA的杂交。在一些实施方案中，检测所述TIC4029毒素蛋白类的成员的存在的步骤可以包括ELISA或蛋白质印迹。

[0026] 本文还提供了检测来自所述TIC4029类的杀害虫性蛋白质或其片段的存在的方法，其中所述方法包括使样品与向TIC4029类毒素蛋白的结合具有特异性免疫反应性抗体接触；以及检测所述抗体与所述TIC4029类蛋白的结合，从而确认所述蛋白质在所述样品中的存在。在一些实施方案中，所述检测的步骤包括ELISA或蛋白质印迹。

[0027] 本申请还设想了一种用于在表达TIC4029类毒素蛋白的转基因作物田地中防治鳞翅目害虫物种或害虫侵扰的方法，其中所述方法包括培植表达杀昆虫有效量的杀害虫性蛋

白质的作物植物,所述杀害虫性蛋白质具有SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:4、SEQ ID NO:6、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16、SEQ ID NO:18、SEQ ID NO:20、SEQ ID NO:22、SEQ ID NO:24、SEQ ID NO:26、SEQ ID NO:28、SEQ ID NO:30或SEQ ID NO:32中所示的氨基酸序列;或者培植表达杀昆虫有效量的一种或多种杀害虫性蛋白质的作物植物,所述杀害虫性蛋白质包含与SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15、SEQ ID NO:17、SEQ ID NO:19、SEQ ID NO:21、SEQ ID NO:23、SEQ ID NO:25、SEQ ID NO:27、SEQ ID NO:29或SEQ ID NO:31中所示的氨基酸序列具有至少87%、或90%、或95%、或98%、或99%、或99.5%、或约100%同一性的氨基酸序列;以及任选地将各自携带自限性基因的一种或多种转基因鳞翅目害虫物种释放至含有编码本发明的毒素蛋白的基因的作物田地,以防止或延迟所述一种或多种鳞翅目害虫物种对所述毒素蛋白的耐受性的产生。在一个实施方案中,所述作物植物可以是单子叶植物或双子叶植物。在另一个实施方案中,所述单子叶作物植物可以是玉米、小麦、高粱、水稻、黑麦或小米。在又一个实施方案中,所述双子叶作物植物可以是大豆、棉花或芥花。这种包含自限性基因的鳞翅目物种的组合在转基因作物田地的适当附近处释放,将防止或延迟缺乏(或不具有)所述自限性基因的所述鳞翅目害虫物种对本发明的毒素蛋白的耐受性的产生。

附图说明

[0028] 图1由1a和1b两部分组成,描绘了昆虫毒素蛋白TIC4029(SEQ ID NO:2)、TIC4029_1(SEQ ID NO:4)和TIC4029_8(SEQ ID NO:10)的氨基酸序列比对。具体而言,图1a-1b显示了相对于全长TIC4029毒素蛋白的引入到TIC4029_1和TIC4029_8中的截短。TIC4029_1包含原毒素结构域的截短。TIC4029_8代表TIC4029的胰蛋白酶核心,包含原毒素结构域的截短和成熟毒素片段的氨基末端的一部分。

[0029] 序列简要说明

[0030] SEQ ID NO:1是编码从苏云金芽孢杆菌物种EGBS0016获得的TIC4029杀害虫性蛋白质的核酸序列。

[0031] SEQ ID NO:2是TIC4029杀害虫性蛋白质的氨基酸序列。

[0032] SEQ ID NO:3是编码TIC4029_1杀害虫性蛋白质的核酸序列。TIC4029_1是截短,其中编码TIC4029的原毒素结构域的编码序列已经缺失。

[0033] SEQ ID NO:4是TIC4029_1杀害虫性蛋白质的氨基酸序列。

[0034] SEQ ID NO:5是用于在植物细胞中表达的合成的编码序列,编码TIC4029的TIC4029PL-1。

[0035] SEQ ID NO:6是用于在植物细胞中表达的合成的编码序列,编码TIC4029的TIC4029PL-2。

[0036] SEQ ID NO:7是用于在植物细胞中表达的合成的编码序列,编码TIC4029的TIC4029PL-3。

[0037] SEQ ID NO:8是用于在植物细胞中表达的合成的编码序列,编码TIC4029_1的TIC4029_1PL。

[0038] SEQ ID NO:9是用于表达氨基酸序列TIC4029_8的合成的编码序列。TIC4029_8包

含天然TIC4029蛋白的氨基酸28-607，并且通过缺失N-末端氨基酸1-27和天然蛋白质的氨基酸607之后的原毒素结构域的氨基酸来制备。此外，在天然蛋白质的氨基酸位置27处的TIC4029的精氨酸残基被甲硫氨酸残基置换。

[0039] SEQ ID NO:10是TIC4029_8杀害虫性蛋白质的氨基酸序列。

具体实施方式

[0040] 农业害虫防治领域的一个问题可以表征为对新型毒素蛋白的需要，所述新型毒素蛋白是针对目标害虫有效的，表现出针对目标害虫物种的广谱毒性，能够在植物中表达而不引起不期望的农艺问题，并且提供相较于目前在植物中商业化使用的毒素的替代作用模式。

[0041] 以TIC4029和氨基酸序列变体为示例的新型杀害虫性蛋白质是本文公开的并且解决了本领域中的这些问题中的每个，尤其是针对广谱鳞翅目昆虫害虫，具体而言针对豆芽蛾(豆茎蛾)、黑粘虫(考斯夜蛾)、黑地蚕(小地老虎)、玉米穗虫(玉米耳虫)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)、南美豆荚虫(大豆荚虫)、南部行军虫(南部灰翅夜蛾)、大豆尺蠖(大豆尺蠖)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾)、烟草蚜虫(烟芽夜蛾)和天鹅绒豆毛虫(黎豆夜蛾)。

[0042] 在本申请中提及TIC4029、－TIC4029蛋白”、－TIC4029蛋白毒素”、－TIC4029杀害虫性蛋白质”、－TIC4029相关毒素”、－TIC4029相关性毒素”、－TIC4029类”、－TIC4029蛋白毒素类”、－TIC4029毒素蛋白类”等等，是指这样的任何新型杀害虫性蛋白质或昆虫抑制性蛋白质，所述杀害虫性蛋白质或昆虫抑制性蛋白质包含TIC4029(SEQ ID NO:2)的任何杀害虫性蛋白质或昆虫抑制性蛋白质序列、以及截短毒素蛋白、TIC4029_1(SEQ ID NO:4)和TIC4029_8(SEQ ID NO:10)以及它们的杀害虫性或昆虫抑制性区段、或者它们的组合，由其组成，与其基本上同源，与其类似或由其衍生，所述杀害虫性蛋白质或昆虫抑制性蛋白质赋予针对鳞翅目害虫的活性，包括任何表现出杀害虫性或昆虫抑制性活性的蛋白质，条件是这种蛋白质与TIC4029的比对产生从约86％至约100％的任何分数百分比的同一性的氨基酸序列。TIC4029蛋白包括质体靶向和非质体靶向形式的蛋白质二者。

[0043] 本申请中使用术语－区段”或－片段”来描述比描述TIC4029或TIC4029截短变体蛋白的完整氨基酸或核酸序列更短的连续氨基酸或核酸序列。表现出昆虫抑制性活性的区段或片段也在本申请中有所公开，条件是这种区段或片段与SEQ ID NO:2中所示的TIC4029蛋白的相应的区段、SEQ ID NO:4中所示的TIC4029_1蛋白、或SEQ ID NO:10中所示的TIC4029_8蛋白的比对，在TIC4029或TIC4029截短变体蛋白内的氨基酸的区段或片段与相应的区段之间，产生从约87％至约100％的任何分数百分比的氨基酸序列同一性。如本文所述的片段可以包含SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10的至少50、至少100、至少250、至少500、至少600、至少800或至少1000个连续氨基酸。如本文所述的片段可以具有SEQ ID NO:2、SEQ ID NO:4或SEQ ID NO:10中的任一者的杀害虫活性。

[0044] 在本申请中提及术语－活性的”或－活性”、－杀害虫活性”或－杀害虫性”或－杀昆虫活性”、－昆虫抑制性”、－杀害虫有效性”或－杀昆虫剂”是指毒剂诸如蛋白毒素在抑制(抑制生长、摄食、生殖力或生存力)、压制(压制生长、摄食、生殖力或生存力)、防治(防治害虫侵扰、防治害虫对含有有效量的TIC4029或TIC4029截短变体蛋白的特定作物的摄食活

性)或杀死(导致发病、死亡或生殖力降低)害虫方面的功效。这些术语旨在包括向害虫提供杀昆虫有效量的毒性蛋白的结果,其中害虫向毒性蛋白的暴露导致发病、死亡、生殖力降低或发育迟缓。这些术语还包括作为在植物之中或之上提供杀害虫有效量的毒性蛋白的结果,从植物、植物的组织、植物部分、种子、植物细胞,或者从培植植物特定地理位置排斥害虫。一般而言,杀害虫活性是指毒性蛋白有效抑制生长、发育、生存力、摄食行为、交配行为、生殖力的能力,或者由特定目标害虫(包括但不限于鳞翅目昆虫)对该蛋白、蛋白片段、蛋白区段或多核苷酸的昆虫摄食导致的任何可测量的不良反应的减少。毒性蛋白可以由植物产生或者可以施用于植物或植物所处位置内的环境。术语－生物活性”、－有效性”、－有效”或它们的变化形式在本申请中也是可互换使用的术语,用于描述本发明的蛋白质对目标昆虫害虫的作用。

[0045] 当在目标害虫的饮食中提供杀害虫有效量的毒剂时,当毒剂接触害虫时表现出杀害虫活性。毒剂可以是杀害虫性蛋白质或本领域已知的一种或多种化学剂。杀害虫性或杀昆虫性化学剂以及杀害虫性或杀昆虫性蛋白质剂可以单独或彼此组合使用。化学剂包括但不限于靶向特定基因以抑制目标害虫的dsRNA分子、有机氯化物、有机磷酸酯、氨基甲酸酯、拟除虫菊酯、新烟碱和莱恩碱。杀害虫性蛋白质剂或杀昆虫性蛋白质剂包含本申请中所示的蛋白毒素,以及其他蛋白毒剂,包括靶向鳞翅目的那些,以及用于防治其他植物害虫的蛋白毒素,诸如可在本领域中用于防治鞘翅目、半翅目和同翅目物种的Cry、Vip和Cyt蛋白。

[0046] 提及害虫,尤其是作物植物的害虫,意指作物植物的昆虫害虫,尤其是通过TIC4029蛋白毒素类防治的那些鳞翅目昆虫害虫。然而,当靶向这些害虫的毒剂与TIC4029蛋白或TIC4029截短变体蛋白或者与TIC4029蛋白或TIC4029截短变体蛋白具有87％至约100％同一性的蛋白质一起共定位或存在时,提及害虫还可以包括植物的鞘翅目、半翅目和同翅目昆虫害虫,以及线虫和真菌。短语－一起存在”或－共定位”旨在包括目标昆虫害虫已经与TIC4029蛋白毒素类接触,以及相对于目标昆虫害虫而言任何其他毒剂也以杀害虫有效量存在的任何情况。－接触”旨在指存在于目标害虫的饮食中,并且该饮食被目标害虫消耗。

[0047] TIC4029或TIC4029截短变体蛋白因共同功能而相关,并且针对鳞翅目昆虫物种(包括成虫、蛹、幼虫和新生虫)的昆虫害虫表现出杀害虫活性。

[0048] 鳞翅目昆虫包括但不限于:夜蛾科(Noctuidae)的粘虫、地老虎、尺蠖和实夜蛾,例如,秋粘虫(草地贪夜蛾)、豆芽蛾(豆茎蛾)、甜菜夜蛾(甜菜夜蛾(Spodoptera exigua))、黑粘虫(考斯夜蛾)、南部行军虫(南部灰翅夜蛾)、披肩粘虫(披肩粘虫(Mamestra configurata))、黑地蚕(小地老虎)、卷心菜尺蠖(卷心菜尺蠖(Trichoplusia ni))、大豆尺蠖(大豆夜蛾)、向日葵尺蠖(薄荷灰夜蛾)、天鹅绒豆毛虫(黎豆夜蛾)、苜蓿绿夜蛾(苜蓿绿夜蛾(Hypena scabra))、烟草蚜虫(烟芽夜蛾)、颗粒地老虎(粒肤地老虎)、粘虫(粘虫(Pseudaletia unipuncta))、甘蔗螟(小蔗螟)、向日葵尺蠖(薄荷灰夜蛾)、南美豆荚虫(大豆荚虫)、西部地老虎(西部灰地老虎);螟蛾科(Pyralidae)的螟、鞘蛾、结网蠕虫、锥虫、甘蓝虫和雕叶虫,例如,欧洲玉米螟(欧洲玉米螟)、脐橙虫(脐橙虫(Amyelois transitella))、玉米根结网蠕虫(玉米根草螟)、草皮结网蠕虫(水稻切叶野螟)、向日葵蛾(向日葵螟)、小玉米杆螟(南美玉米苗斑螟);卷叶蛾科(Tortricidae)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(葡萄浆果蛾

(Endopiza viteana))、东部果蛾(东方果蛾)、向日葵芽蛾(向日葵食芽蛾);以及很多其他具有重要经济意义的鳞翅目昆虫,例如,小菜蛾(小菜蛾)、粉红棉铃虫(棉红铃虫)和舞毒蛾(舞毒蛾)。其他鳞翅目昆虫害虫包括,例如,棉叶虫、果树卷叶虫(果树黄卷蛾)、欧洲卷叶虫(玫瑰黄卷蛾)和其他黄卷蛾属(Archips)物种(二化螟(Chilo suppressalis),亚洲水稻螟或水稻二化螟)、水稻卷叶虫(稻纵卷叶螟)、玉米根结网蠕虫(玉米根草螟)、早熟禾结网蠕虫(熟禾草螟)、西南玉米螟(西南玉米螟)、甘蔗螟(小蔗螟)、多刺棉铃虫(埃及钻夜蛾)、斑点棉铃虫(棉斑实蛾)、美洲棉铃虫(稻田棉铃虫)、玉米穗虫(玉米耳虫,也称为大豆豆荚虫和棉花棉铃虫)、烟草蚜虫(烟芽夜蛾)、草皮结网蠕虫(水稻切叶野螟)、西部豆地老虎(西部豆角夜蛾)、欧洲葡萄藤蛾(葡萄花翅小卷蛾)、柑橘潜叶蛾(柑橘潜叶蛾(Phyllocnistis citrella))、大白蝶(大菜粉蝶)、小白蝶(菜粉蝶,也称为输入菜粉蝶)、甜菜夜蛾(甜菜夜蛾(Spodoptera exigua))、烟草地老虎(斜纹夜蛾,也称为集群毛虫)和番茄潜叶蛾(番茄潜叶蛾(Tuta absoluta))。

[0049] 在本申请中提及－分离的DNA分子”或等同的术语或短语,旨在意指DNA分子是单独或与其他组合物组合存在的DNA分子,但是不在其天然环境中。例如,在生物体的基因组的DNA中天然存在的核酸元件(诸如编码序列、内含子序列、非翻译前导序列、启动子序列、转录终止序列等),只要所述元件位于生物体的基因组内并且在其天然存在的基因组内的位置处,就不被认为是－分离的”。然而,这些元件中的每者以及这些元件的子部分,在本公开的范围内将是－分离的”,只要该元件不在生物体的基因组内并且不在其天然存在的基因组内的位置处即可。类似地,编码杀害虫性蛋白质或该蛋白质的任何天然存在的杀昆虫性变体的核苷酸序列将是分离的核苷酸序列,只要该核苷酸序列不在编码该蛋白的序列天然存在的细菌的DNA内即可。出于本公开的目的,编码天然存在的杀昆虫蛋白的氨基酸序列的合成核苷酸序列将被认为是分离的。出于本公开的目的,任何转基因核苷酸序列,即,插入植物或细菌细胞的基因组或存在于染色体外载体中的DNA的核苷酸序列,无论存在于用于转化细胞的质粒或类似结构中,存在于植物或细菌的基因组中,还是以可检测的量存在于来源于植物或细菌的组织、子代、生物样品或商品性产品中,都将被视为分离的核苷酸序列。

[0050] 在本申请中提及术语－自限性基因”是指一种或多种限制宿主存活,导致宿主群体减少的基因。这种技术由Oxitech Ltd.提供。携带转基因自限性基因的转基因雄性昆虫被释放并且与野生雌性一起繁殖。结果,子代继承了自限性基因的拷贝。自限性基因通过在昆虫细胞中过量产生蛋白质来破坏昆虫细胞的正常功能,干扰细胞产生发育所需的其他必需蛋白质的能力。该基因通过破坏昆虫的正常发育,阻止昆虫存活到成虫。例如,自限性小菜蛾(小菜蛾)品系OX4319L由Oxitech Ltd开发,携带雄性选择基因,该雄性选择基因利用性别决定基因双性基因(dsx)的序列。该基因表达性别交替剪接,以对雌性特异性自限性基因表达进行工程化,防止雌性后代存活超过幼虫阶段,并允许产生仅雄性的自限性蛾群。在释放后,雄性与雌性害虫交配,导致下一代中的雌性后代数量减少,从而局部抑制小菜蛾群体。为了促进在小菜蛾生产设施内饲养大量供释放的雄性,通过在幼虫饲料中添加四环素或合适的类似物来阻遏OX4319L品系中的雌性特异性dsx的表达。OX4319L还表达荧光蛋白DsRed,以允许有效监测该品系在田地中的存在(Jin等人,2013.Engineered female-specific lethality for control of pest Lepidoptera.ACS Synthetic Biology,2:

160-166)。当该技术应用于含有本发明的毒素基因的植物田地中时,可以延迟或防止针对本发明的毒素基因和蛋白质所防治的害虫物种的耐受性产生,从而提供更大的耐久性的含有本发明的毒素基因和蛋白质的任何植物产品。该技术可以应用于秋粘虫、玉米穗虫、玉米根虫和大豆尺蠖,以及许多其他作物害虫物种。

[0051] 如本申请中所进一步描述,编码TIC4029(SEQ ID NO:1)的开放阅读框(ORF)在从苏云金芽孢杆菌菌株EGBS0016获得的DNA中被发现。编码序列被克隆并在微生物宿主细胞中表达,以产生用于生物测定法的重组蛋白。使用微生物宿主细胞衍生的TIC4029蛋白质进行的生物测定法展示出针对以下鳞翅目物种的活性:黑粘虫(BAW,考斯夜蛾)、黑地蚕(BCW,小地老虎)、玉米穗虫(CEW,玉米耳虫)、欧洲玉米螟(ECB,欧洲玉米螟)、秋粘虫(FAW,草地贪夜蛾)、南部行军虫(SAW,南部灰翅夜蛾)、大豆尺蠖(SBL,大豆尺蠖)、西南玉米螟(SWC,西南玉米螟)、甘蔗螟(SCB,小蔗螟)、向日葵尺蠖(SFL,薄荷灰夜蛾)、烟草蚜虫(TBW,烟芽夜蛾)和天鹅绒豆毛虫(VBC,黎豆夜蛾)。

[0052] 本文公开的TIC4029_1和TIC4029_8是TIC4029的两种不同的氨基酸序列缺失变体。TIC4029_1包含羧基末端原毒素结构域的缺失,即,天然毒素的第603位之后的氨基酸。TIC4029_1在细菌中表达并且针对BAW、SAW和VBC而测定,并且展示出针对全部三种鳞翅目昆虫害虫的活性。TIC4029_8包含天然TIC4029蛋白的氨基酸28-607,并且通过缺失N-末端氨基酸1-27和天然蛋白质的氨基酸607之后的原毒素结构域的氨基酸来制备。此外,在天然蛋白质的氨基酸位置27处的TIC4029的精氨酸残基被甲硫氨酸残基置换。TIC4029_8毒素蛋白代表天然TIC4029毒素蛋白的胰蛋白酶核心。

[0053] 图1提供了TIC4029、TIC4029_1和TIC4029_8的比对,图中展示了TIC4029_1和TIC4029_8相对于TIC4029的截短。产生了被设计用于在植物细胞中表达TIC4029蛋白类毒素的合成的编码序列,包括TIC4029(SEQ ID NO:5-7)、TIC4029_1(SEQ ID NO:8)和TIC4029_8(SEQ ID NO:9)。

[0054] 表达毒素TIC4029(由SEQ ID NO:5编码)、TIC4029_1(由SEQ ID NO:8编码)和TIC4029_8(由SEQ ID NO:9编码)的大豆植物使用叶盘来测定并且展示出抗SBL的功效和SAW的抑制。表达TIC4029(由SEQ ID NO:7编码)和TIC4029_8(由SEQ ID NO:9编码)的玉米植物展示出抗SWC的功效,表达TIC4029_1(由SEQ ID NO:8编码)的玉米植物展示出SWC的抑制。在美国的温网室试验中,表达TIC4029(由SEQ ID NO:5编码)的大豆植物展示出抗SBL和VBC的功效,以及SAW的抑制。在巴西的温网室试验中测试时,表达TIC4029(由SEQ ID NO:5编码)的大豆植物展示出抗SBL、SFL和VBC的功效。

[0055] 对于在植物细胞中的表达,TIC4029(SEQ ID NO:2)、TIC4029_1(SEQ ID NO:4)和TIC4029_8(SEQ ID NO:10)蛋白可以表达为驻留在胞质溶胶中或靶向植物细胞的各种细胞器。例如,使蛋白质靶向叶绿体可以引起转基因植物中所表达的蛋白质的水平增加,同时防止出现异常表型。靶向还可以在转基因事件中引起害虫耐受性功效的增加。靶标肽或转运肽是短(长度为3-70个氨基酸)肽链,它将蛋白质运输到细胞中的特定区域,包括细胞核、线粒体、内质网(ER)、叶绿体、质外体、过氧化物酶体和质膜。在蛋白质运输后,一些靶标肽被信号肽酶从蛋白质切割下来。为了靶向叶绿体,蛋白质含有约40-50个氨基酸的转运肽。关于叶绿体转运肽的使用的描述,参见美国专利5,188,642和5,728,925。很多叶绿体定位蛋白作为前体从核基因表达,并且通过叶绿体转运肽(CTP)来靶向叶绿体。此类分离的CTP的

实例包括但不限于与核酮糖-1,5,-二磷酸羧化酶的小亚基(SSU)、铁氧还蛋白、铁氧还蛋白氧化还原酶、捕光复合物蛋白I和蛋白II、硫氧还蛋白F、烯醇丙酮酰莽草酸磷酸合酶(EPSPS)和美国专利7,193,133中描述的转运肽相关的那些。已经在体内和体外得到证明的是,非叶绿体蛋白可以通过使用与异源性CTP的蛋白质融合体来靶向叶绿体,并且CTP足以使蛋白质靶向叶绿体。合适的叶绿体转运肽,诸如拟南芥(Arabidopsis thaliana)EPSPS CTP(CTP2)(参见,Klee等人,Mol.Gen.Genet.210:437-442,1987)或碧冬茄(Petunia hybrida)EPSPS CTP(CTP4)(参见,della-Cioppa等人,Proc.Natl.Acad.Sci.USA 83:6873-6877,1986)的掺入已经显示出靶向转基因植物中的叶绿体的异源性EPSPS蛋白序列(参见,美国专利5,627,061;5,633,435和5,312,910;以及EP 0218571;EP 189707;EP 508909;以及EP 924299)。为了使TIC4029或截短变体TIC4029毒素蛋白靶向叶绿体,将编码叶绿体转运肽的序列放置于可操作连接的5′处,并且与编码TIC4029或截短变体TIC4029毒素蛋白(已经被设计用于在植物细胞中表达)的合成的编码序列同框。

[0056] 可以设想,使用TIC4029的氨基酸序列来产生与TIC4029相关的另外的毒素蛋白序列,以产生具有新性质的新型蛋白质。可以比对TIC4029毒素蛋白以将氨基酸序列水平的差异组合成新型氨基酸序列变体,并且对编码变体的重组核酸序列进行适当的改变。

[0057] 可以设想,TIC4029蛋白毒素类的改进的变体可以通过使用本领域已知的各种基因编辑方法在植物体中进行工程化。用于基因组编辑的此类技术包括但不限于ZFN(锌指核酸酶)、大范围核酸酶、TALEN(转录激活因子样效应物核酸酶)和CRISPR(成簇的规律间隔短回文重复序列)/Cas(CRISPR相关)系统。这些基因组编辑方法可以用于将植物细胞内转化的毒素蛋白编码序列改变为不同的毒素编码序列。具体而言,通过这些方法,毒素编码序列中的一个或多个密码子被改变,以对新型蛋白质氨基酸序列进行工程化。或者,使编码序列中的片段置换或缺失,或者将另外的DNA片段插入编码序列,以对新型毒素编码序列进行工程化。新型编码序列可以编码具有新性质(诸如增加的针对昆虫害虫的活性或活性谱)的毒素蛋白,以及提供其中已经针对原始昆虫毒素蛋白产生耐受性的针对昆虫害虫物种的活性。包含基因编辑的毒素编码序列的植物细胞可以通过本领域已知的方法使用,以产生表达新型毒素蛋白的完整植物。

[0058] 还可以设想,TIC4029的片段或它们的蛋白变体可以是截短形式,其中一个或多个氨基酸从蛋白质的N-末端、C-末端、中间或它们的组合缺失,其中片段和变体保持昆虫抑制性活性。这些片段可以是TIC4029的天然存在或合成变体或者衍生的蛋白变体,但是应保持至少TIC4029的昆虫抑制性活性。TIC4029的截短变体蛋白的实例包括TIC4029_1(SEQ ID NO:4)和TIC4029_8(SEQ ID NO:10)。

[0059] 可以使用本领域已知的各种基于计算机的算法来鉴定和彼此比较类似于TIC4029蛋白的蛋白质(参见表1)。本申请中报告的氨基酸序列同一性是使用以下默认参数进行Clustal W比对的结果: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(Thompson等人(1994)Nucleic Acids Research,22:4673-4680)。氨基酸同一性百分比通过100%乘以(氨基酸同一性/主题蛋白质的长度)的乘积进一步计算。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比对获得的那些结果类似的结果,并且在本文中设想到。

[0060] 如果在查询中,例如,在Clustal W比对中使用蛋白质,则意味着表现出针对鳞翅

目昆虫物种的昆虫抑制性活性的蛋白质与TIC4029有关，SEQ ID NO:2中所示的本发明的蛋白质被鉴定为这种比对的命中，其中查询蛋白沿着查询蛋白的长度表现出至少87％至约100％的氨基酸同一性，即约87％、88％、89％、90％、91％、92％、93％、94％、95％、96％、97％、98％、99％、100％或该范围内的任何分数百分比。

[0061] 除了同一性百分比之外，TIC4029和TIC4029的截短变体还可以通过一级结构(保守氨基酸基序)、长度和其他特征来进行关联。表1中报告了TIC4029蛋白毒素类的特征。

[0062] 表1.TIC4029和截短变异蛋白毒素的选定特征。

[0063]

蛋白质	分子量(单位：道尔顿)	氨基酸长度	等电点	pH 7.0时的电荷	疏水性氨基酸的数量	极性氨基酸的数量	强碱性(-)氨基酸的数量	强酸性氨基酸的数量
TIC4029	133674.75	1177	4.8540	-32.0	593	584	139	159
TIC4029_1	68156.27	603	6.7263	1.0	317	286	60	55
TIC4029_8	65610.58	581	8.5413	6.0	307	274	60	50

[0064] 如本申请的实施例中所进一步描述，编码TIC4029、TIC4029_1和TIC4029_8的合成核酸分子序列被设计用于植物，它们分别由SEQ ID NO:5-7(TIC4029)、SEQ ID NO:8(TIC4029_1)和SEQ ID NO:9(TIC4029_8)编码。

[0065] 可以根据本领域已知的转化方法和技术将含有重组核酸分子序列的表达盒和载体构建和引入植株中，具体而言构建和引入玉米、大豆或棉花植物细胞中。例如，农杆菌介导的转化在美国专利申请公布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(甜菜)、美国专利5,750,871(芥花)、7,026,528(小麦)和6,365,807(水稻)，以及Arencibia等人(1998)Transgenic Res.7:213-222(甘蔗)中有所描述，这些专利和文献均以引用的方式整体并入本文。经转化的细胞可以再生为表达TIC4029、TIC4029_1和TIC4029_8的经转化的植物，并且通过使用从经转化的植物获得的植物叶盘在存在鳞翅目害虫幼虫的情况下进行的生物测定法来展示杀害虫活性。植物可以通过再生、种子、花粉或分生组织转化技术来从植物细胞衍生。转化植物的方法是本领域已知的。

[0066] 作为传统转化方法的替代，可将DNA序列(诸如转基因，表达盒等)经由定点整合插入或整合到植物或植物细胞的基因组内的特定位点或基因座中。因此，本公开的重组DNA构建体和分子可包含供体模板序列，所述供体模板序列包含至少一种转基因、表达盒或其他DNA序列，以用于插入到植物或植物细胞的基因组中。这种用于定点整合的供体模板还可包含侧接插入序列(即，待被插入到植物基因组中的序列、转基因、盒等)的一个或两个同源臂。本公开的重组DNA构建体还可包含表达盒，所述表达盒编码位点特异性核酸酶和/或任何相关蛋白质以进行定点整合。这些核酸酶表达盒可与供体模板存在于同一分子或载体中(顺式)，或存在于单独的分子或载体中(反式)。用于定点整合的若干方法在本领域中是已知的，这些方法涉及切割基因组DNA以在所需的基因组位点或基因座处产生双链断裂(DSB)或切口的不同的蛋白质(或蛋白质复合物和/或指导RNA)。简而言之，如本领域所理解，在修复由核酸酶引入的DSB或切口的过程中，供体模板DNA可以在DSB或切口的位点处整合到基因组中。同源臂在供体模板中的存在可以在修复过程中通过同源重组来促进插入序列的采用和向植物基因组的靶向，虽然插入事件可以通过非同源末端连接(NHEJ)来发生。可使用

的位点特异性核酸酶的实例包括锌指核酸酶、工程化的或天然的大范围核酸酶、TALE-核酸内切酶和RNA指导的核酸内切酶(例如,Cas9或Cpf1)。对于使用RNA指导的位点特异性核酸酶(例如,Cas9或Cpf1)的方法,所述重组DNA构建体还将包含编码一个或多个指导RNA以将核酸酶引导至植物基因组内的所需位点的序列。

[0067] 编码细菌表达的TIC4029、TIC4029_1和TIC4029_8以及植物表达的TIC4029、TIC4029_1和TIC4029_8蛋白的重组核酸分子组合物可以用重组DNA构建体来表达,其中具有编码蛋白质的ORF的多核苷酸分子可操作地连接至基因表达元件,诸如启动子和在构建体所期望的系统中表达所需的任何其他调控元件。非限制性实例包括植物功能性启动子(可操作地连接至用于在植物中表达蛋白质的TIC4029、TIC4029_1或TIC4029_8蛋白编码序列)或者Bt功能性启动子(可操作地连接至用于在Bt细菌或其他芽孢杆菌属物种中表达蛋白质的TIC4029、TIC4029_1或TIC4029_8蛋白编码序列)。其他元件可以可操作地连接至TIC4029、TIC4029_1或TIC4029_8蛋白编码序列,包括但不限于增强子、内含子、非翻译前导序列、编码的蛋白质固定标签(HIS-标签)、易位肽(即,质体转运肽、信号肽)、用于翻译后修饰酶的多肽序列、核糖体结合位点和RNAi靶位点。本文提供的示例性重组多核苷酸分子包括但不限于可操作地连接至多核苷酸的异源性启动子,诸如SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9,它们分别编码具有SEQ ID NO:2(由SEQ ID NO:1和SEQ ID NO:5-7编码)、SEQ ID NO:4(由SEQ ID NO:3和SEQ ID NO:8编码)和SEQ ID NO:10(由SEQ ID NO:9编码)中所示的氨基酸序列的多肽或蛋白质。异源性启动子也可以可操作地连接至编码质体靶向的TIC4029、TIC4029_1或TIC4029_8的合成的DNA编码序列。编码本文公开的蛋白质的重组核酸分子的密码子可以被同义密码子置换(在本领域中称为沉默置换)。

[0068] 包含TIC4029、TIC4029_1或TIC4029_8蛋白编码序列的重组DNA构建体还可以包含编码一种或多种昆虫抑制剂的DNA区域,所述DNA区域可以被构造为与编码TIC4029、TIC4029_1或TIC4029_8蛋白的DNA序列、昆虫抑制性dsRNA分子、或辅助蛋白同时表达或共表达。辅助蛋白包括但不限于辅因子、酶、结合配偶体或其他有助于昆虫抑制剂的有效性的剂,例如,通过帮助其表达、影响其在植物中的稳定性、优化寡聚化反应的自由能、增加其毒性以及增加其活性谱。例如,辅助蛋白可以促进一种或多种昆虫抑制剂的摄取,或者增强毒剂的毒性作用。

[0069] 可以组装重组DNA构建体,以使得所有蛋白质或dsRNA分子都从一个启动子表达,或者每个蛋白质或dsRNA分子处于单独的启动子控制或其某种组合之下。本发明的蛋白质可以从多基因表达系统表达,其中TIC4029、TIC4029_1或TIC4029_8的一种或多种蛋白质从共同核苷酸区段表达,所述共同核苷酸区段还含有其他开放阅读框和启动子,这取决于所选的表达系统类型。例如,细菌多基因表达系统可以利用单个启动子从单个操纵子驱动多重连锁/串联开放阅读框的表达(即,多顺反子表达)。在另一个实例中,植物多基因表达系统可以利用多重非连锁或连锁表达盒,每个盒表达不同的蛋白质或其他剂(诸如一种或多种dsRNA分子)。

[0070] 可通过载体将包含TIC4029、TIC4029_1或TIC4029_8蛋白编码序列的重组多核苷酸或重组DNA构建体递送至宿主细胞,所述载体例如质粒、杆状病毒、合成染色体、病毒颗粒、粘粒、噬菌粒、噬菌体或病毒载体。此类载体可以用于实现TIC4029、TIC4029_1或

TIC4029_8蛋白编码序列在宿主细胞中的稳定或瞬时表达，或者编码多肽的后续表达。包含TIC4029、TIC4029_1或TIC4029_8蛋白编码序列并且被引入宿主细胞的外源性重组多核苷酸或重组DNA构建体在本申请中称为－转基因”。

[0071] 本文提供了含有重组多核苷酸的转基因细菌、转基因植物细胞、转基因植物和转基因植物部分，所述重组多核苷酸表达TIC4029、TIC4029_1或TIC4029_8中的任何一者或多者，或者相关的家族毒素蛋白编码序列。术语－细菌细胞”或－细菌”可以包括但不限于农杆菌、芽孢杆菌、埃希氏菌、沙门氏菌、假单胞菌、短芽孢杆菌、克雷伯氏菌、欧文氏菌或根瘤菌细胞。术语－植物细胞”或－植物”可以包括但不限于双子叶植物或单子叶植物。术语－植物细胞”或－植物”可以另外包括但不限于：苜蓿、香蕉、大麦、豆、西兰花、卷心菜、芸苔、胡萝卜、木薯、蓖麻、花椰菜、芹菜、鹰嘴豆、白菜、柑橘、椰子、咖啡、玉米、三叶草、棉花、葫芦、黄瓜、道格拉斯冷杉、茄子、桉树、亚麻、大蒜、葡萄、啤酒花、韭葱、莴苣、火炬松、小米、甜瓜、坚果、燕麦、橄榄、洋葱、观赏植物、棕榈、牧草、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、水稻、砧木、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、西葫芦、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香树、甘薯、柳枝稷、茶、烟草、番茄、黑小麦、草坪草、西瓜和小麦植物细胞或植物。在某些实施方案中，提供了转基因植物和从转基因植物细胞再生的转基因植物部分。在某些实施方案中，转基因植物可以通过切割、折断、研磨或者从植物中分离一部分而从转基因种子获得。在某些实施方案中，植物部分可以是种子、棉铃、叶、花、茎、根或它们的任何部分，或者转基因植物部分的不可再生部分。如该语境中所用，转基因植物部分的－不可再生”部分是不能被诱导形成整株植物或者不能被诱导形成能够进行有性和/或无性繁殖的整株植物的部分。在某些实施方案中，植物部分的不可再生部分是转基因种子、棉铃、叶、花、茎或根的一部分。

[0072] 提供了制备包含昆虫、鳞翅目抑制量的TIC4029、TIC4029_1或TIC4029_8蛋白的转基因植物的方法。此类植物可以通过以下步骤来制备：将编码本申请中提供的蛋白质中的任一者的重组多核苷酸引入植物细胞中，以及选择来源于表达昆虫、鳞翅目抑制量的蛋白质的所述植物细胞的植物。植物可以通过再生、种子、花粉或分生组织转化技术来从植物细胞衍生。转化植物的方法是本领域已知的。

[0073] 本文还公开了经加工的植物产品，其中所述经加工的产品包含可检测的量的TIC4029、TIC4029_1或TIC4029_8蛋白、它们的昆虫抑制性区段或片段、或者它们的任何区分部分。在某些实施方案中，经加工的产品选自由以下组成的组：植物部分、植物生物质、油、粕、糖、动物饲料、粉、片、麸、皮棉、壳、经加工的种子和种子。在某些实施方案中，经加工的产品是不可再生的。植物产品可以包括来源于转基因植物或转基因植物部分的商品或其他商业产品，其中商品或其他产品可以通过检测编码或包含TIC4029、TIC4029_1或TIC4029_8蛋白的区分部分的核苷酸区段或表达的RNA或蛋白质通过商业来追踪。

[0074] 表达TIC4029、TIC4029_1或TIC4029_8蛋白的植物可以通过用转基因事件繁育来杂交，所述转基因事件表达其他毒素蛋白和/或表达其他转基因性状，诸如除草剂耐受性基因、赋予产量或胁迫耐受性性状的基因等等，或者此类性状可以被组合在单一堆叠载体中，以使得所述性状完全连锁。

[0075] 如实施例中所进一步描述，TIC4029、TIC4029_1或TIC4029_8蛋白编码序列以及与TIC4029、TIC4029_1或TIC4029_8具有大量同一性百分比的序列可以使用本领域的普通技

术人员已知的方法来鉴定,所述方法诸如聚合酶链式反应(PCR)、热扩增和杂交。例如,蛋白质TIC4029、TIC4029_1或TIC4029_8可以用于产生与相关蛋白质特异性结合的抗体,并且可以用于筛选和寻找其他密切相关的蛋白质成员。

[0076] 此外,编码TIC4029、TIC4029_1或TIC4029_8毒素蛋白的核苷酸序列可以用作探针和引物,所述探针和引物用于筛选以使用热循环或等温扩增和杂交方法来鉴定类别中的其他成员。例如,来源于如SEQ ID NO:5-9中所示的序列的寡核苷酸可以用于确定来源于商品性产品的脱氧核糖核酸样品中是否存在TIC4029、TIC4029_1或TIC4029_8转基因。鉴于采用寡核苷酸的某些核酸检测方法的灵敏度,预期来源于如SEQ ID NO:5-9中所示的序列的寡核苷酸可以用于检测来源于合并来源的商品性产品中的TIC4029、TIC4029_1或TIC4029_8转基因,其中只有一小部分商品性产品来源于含有转基因中的任一个的转基因植物。还认识到,此类寡核苷酸可以用于在SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8和SEQ ID NO:9中的每者中引入核苷酸序列变异。此类－诱变”寡核苷酸可用于鉴定TIC4029、TIC4029_1和TIC4029_8氨基酸序列变体,所述变体在转基因植物宿主细胞中表现出一定范围的昆虫抑制性活性或不同表达。

[0077] 核苷酸序列同源物也是本发明的一个实施方案,所述核苷酸序列同源物例如由核苷酸序列编码的杀昆虫性蛋白质,所述核苷酸序列在严格杂交条件下与本申请中公开的序列中的每者或任一者杂交。本发明还提供了一种用于检测第一核苷酸序列的方法,所述第一核苷酸序列与第二核苷酸序列杂交,其中所述第一核苷酸序列(或其反向互补序列)编码杀害虫性蛋白质或其杀害虫性片段,并且与所述第二核苷酸序列杂交。在这种情况下,第二核苷酸序列可以是在严格杂交条件下如SEQ ID NO:1、SEQ ID NO:3、SEQ ID NO:5、SEQ ID NO:6、SEQ ID NO:7、SEQ ID NO:8或SEQ ID NO:9所提供的核苷酸序列中的任一者。核苷酸编码序列在适当的杂交条件(诸如严格杂交条件)下彼此杂交,并且由这些核苷酸序列编码的蛋白质与针对其他蛋白质中的任一种产生的抗血清发生交叉反应。如本文所定义的严格杂交条件包括至少在42℃处杂交,然后在室温处用2X SSC、0.1％ SDS洗涤两次,每次五分钟,然后在65℃处,在0.5X SSC、0.1％ SDS中洗涤两次,每次三十分钟。在更高的温度下洗涤构成更严格的条件,例如,杂交条件68℃,然后在68℃处,在含有0.1％ SDS的2xSSC中洗涤。

[0078] 本领域的技术人员将认识到,由于遗传密码的冗余性,很多其他序列能够编码此类相关蛋白,并且那些序列是本发明的实施方案,在一定程度上发挥作用以在芽孢杆菌菌株或植物细胞中表达杀害虫性蛋白质,当然还认识到,很多此类冗余性编码序列在这些条件下不与编码TIC4029、TIC4029_1和TIC4029_8变体的天然芽孢杆菌序列杂交。本申请设想了使用本领域的普通技术人员已知的这些和其他鉴定方法来鉴定TIC4029、TIC4029_1和TIC4029_8蛋白编码序列以及与TIC4029、TIC4029_1和TIC4029_8蛋白编码序列具有大量同一性百分比的序列。

[0079] 本公开还设想了使用本领域已知的分子方法来工程化和克隆商业化使用的蛋白质,所述蛋白质包含来自杀害虫性蛋白质的蛋白质嵌合体;例如,嵌合体可以从TIC4029、TIC4029_1或TIC4029_8蛋白的区段组装,衍生出另外有用的实施方案,包括将TIC4029、TIC4029_1或TIC4029_8蛋白的区段与不同于TIC4029、TIC4029_1或TIC4029_8蛋白和相关蛋白的多种蛋白质的区段一起组装。TIC4029、TIC4029_1或TIC4029_8蛋白可以进行彼此比

对以及与其他芽孢杆菌、类芽孢杆菌或其他杀害虫性蛋白质比对(无论它们在系统发生上的关系密切还是疏远),并且可以鉴定每种此类蛋白质的区段,以用于比对的蛋白质之间的置换,从而产生嵌合蛋白的构建体。此类嵌合蛋白可以进行害虫生物测定分析,并且与衍生出嵌合体中的每个此类片段的亲本蛋白相比,对是否存在增加的生物活性或扩大的目标害虫谱进行表征。通过与其他蛋白质交换结构域或区段或者通过使用本领域已知的定向进化方法,可以对多肽的杀害虫活性进行进一步工程化,以使其对特定的害虫或更广谱的害虫具有活性。

[0080] 本申请公开了用TIC4029、TIC4029_1或TIC4029_8蛋白来防治昆虫,尤其是鳞翅目对作物植物的侵扰的方法。此类方法可以包括培植包含昆虫或鳞翅目抑制量的TIC4029、TIC4029_1或TIC4029_8毒素蛋白的植物。在某些实施方案中,此类方法可以进一步包括以下步骤中的任一者或多者:(i)将包含或编码TIC4029、TIC4029_1或TIC4029_8毒素蛋白的任何组合物施加于植株或产生植株的种子;以及(ii)用编码TIC4029、TIC4029_1或TIC4029_8毒素蛋白的多核苷酸转化植株或产生植株的植物细胞。通常,设想可以在组合物中、在微生物中或在转基因植物中提供TIC4029、TIC4029_1或TIC4029_8毒素蛋白,以赋予针对鳞翅目昆虫的昆虫抑制性活性。

[0081] 在某些实施方案中,编码TIC4029、TIC4029_1或TIC4029_8毒素蛋白的重组核酸分子是通过培养重组芽孢杆菌或任何其他重组细菌细胞来制备的昆虫抑制性组合物的杀昆虫活性成分,所述重组芽孢杆菌或任何其他重组细菌细胞被转化为在适合表达TIC4029、TIC4029_1或TIC4029_8毒素蛋白的条件下表达TIC4029、TIC4029_1或TIC4029_8毒素蛋白。可以通过干燥、冻干、匀浆、提取、过滤、离心、沉降或浓缩表达/产生所述重组多肽的此类重组细胞的培养物来制备这种组合物。这种过程可以产生芽孢杆菌或其他昆虫病原性细菌细胞提取物、细胞悬浮液、细胞匀浆物、细胞裂解物、细胞上清液、细胞滤液或细胞沉淀。通过获得这样产生的重组多肽,包含重组多肽的组合物可以包括细菌细胞、细菌孢子和伴孢包涵体,并且可以被配制用于各种用途,包括作为农业昆虫抑制性喷洒产品或作为饮食生物测定法中的昆虫抑制性制剂。

[0082] 在一个实施方案中,为了降低耐受性产生的可能性,包含TIC4029、TIC4029_1或TIC4029_8蛋白的昆虫抑制性组合物可以另外包含至少一种另外的多肽,所述多肽表现出针对同一鳞翅目昆虫物种的昆虫抑制性活性,但是不同于TIC4029、TIC4029_1或TIC4029_8毒素蛋白。这种组合物的可能的另外的多肽包括昆虫抑制性蛋白质和昆虫抑制性dsRNA分子。使用此类核糖核苷酸序列来防治昆虫害虫的一个实例在Baum等人(美国专利公布2006/0021087A1)中有所描述。此类用于防治鳞翅目害虫的另外的多肽可以选自由以下组成的组:昆虫抑制性蛋白质,诸如但不限于:Cry1A(美国专利5,880,275)、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B(美国专利公布10/525,318)、Cry1C(美国专利6,033,874)、Cry1D、Cry1Da以及它们的变体、Cry1E、Cry1F和Cry1A/F嵌合体(美国专利7,070,982、6,962,705和6,713,063)、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry1型嵌合体(诸如但不限于TIC836、TIC860、TIC867、TIC869和TIC1100(国际申请公布WO2016/061391(A2))、TIC2160(国际申请公布WO2016/061392(A2)))、Cry2A、Cry2Ab(美国专利7,064,249)、Cry2Ae、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry43A、Cry43B、Cr y51Aa1、ET66、TIC400、TIC800、TIC834、TIC1415、Vip3A、VIP 3Ab、VIP3B、AXMI-001、AXMI-002、AXMI-030、AXMI-035

和AXMI-045(美国专利公布2013-0117884A1)、AXMI-52、AXMI-58、AXMI-88、AXMI-97、AXMI-102、AXMI-112、AXMI-117、AXMI-100(美国专利公布2013-0310543A1)、AXMI-115、AXMI-113、AX MI-005(美国专利公布2013-0104259A1)、AXMI-134(美国专利公布2013-0167264A1)、AXMI-150(美国专利公布2010-0160231A1)、AXMI-184(美国专利公布2010-0004176A1)、AXMI-196、AXMI-204、AXMI-207、AXMI-209(美国专利公布2011-0030096A1)、AX MI-218、AXMI-220(美国专利公布2014-0245491A1)、AXMI-221z、AXMI-222z、AXMI-223z、AXMI-224z、AXMI-225z(美国专利公布2014-0196175A1)、AXMI-238(美国专利公布2014-0033363A1)、AXMI-270(美国专利公布2014-0223598A1)、AXMI-345(美国专利公布2014-0373195A1)、AXMI-335(国际申请公布WO2013/134523(A2))、DIG-3(美国专利公布2013-0219570A1)、DIG-5(美国专利公布2010-0317569A1)、DIG-11(美国专利公布2010-0319093A1)、AfIP-1A及其衍生物(美国专利公布2014-0033361A1)、AfIP-1B及其衍生物(美国专利公布2014-0033361A1)、PIP-1APIP-1B(美国专利公布2014-0007292A1)、PSEEN3174(美国专利公布2014-0007292A1)、AECFG-592740(美国专利公布2014-0007292A1)、Pput_1063(美国专利公布2014-0007292A1)、DIG-657(国际申请公布WO2015/195594A2)、Pput_1064(美国专利公布2014-0007292A1)、GS-135及其衍生物(美国专利公布2012-0233726A1)、GS153及其衍生物(美国专利公布2012-0192310A1)、GS154及其衍生物(美国专利公布2012-0192310A1)、GS155及其衍生物(美国专利公布2012-0192310A1)、如美国专利公布2012-0167259A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2012-0047606A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2011-0154536A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2011-0112013A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2010-0192256A1所述的SEQ ID NO:2和SEQ ID NO:4及其衍生物、如美国专利公布2010-0077507A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2010-0077508A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2009-0313721A1所述的SEQ ID NO:2及其衍生物、如美国专利公布2010-0269221A1所述的SEQ ID NO:2或SEQ ID NO:4及其衍生物、如美国专利7,772,465(B2)所述的SEQ ID NO:2及其衍生物、如WO2014/008054A2所述的CF161_0085及其衍生物、如美国专利公布US2008-0172762A1、US2011-0055968A1和US2012-0117690A1所述的鳞翅目毒性蛋白及其衍生物;如US7510878(B2)所述的SE Q ID NO:2或SEQ ID NO:4及其衍生物、如美国专利7812129(B1)所述的SEQ ID NO:2或SEQ ID NO:4及其衍生物;IPD110Aa和同源物(国际申请公布WO2019/178038A2);TIC868(美国专利US10233217)、Cry1Da1_7(美国专利US10059959)、BCW003(美国专利US10703782)、TIC1100(美国专利US10494408)、TIC867(美国专利US10669317)、TIC867_23(美国专利US10611806)、TIC6757(美国专利US10155960)、TIC7941(美国专利公布2020-0229445A1)、对鳞翅目物种具有毒性的蕨类植物毒素,诸如美国专利10,227,608所公开的那些,等等。

[0083] 在其他实施方案中,此类组合物/制剂可以另外包含至少一种另外的多肽,所述多肽表现出针对不被本发明的另一种昆虫抑制性蛋白质抑制的昆虫的昆虫抑制性活性,以扩大所获得的昆虫抑制谱。例如,为了防治半翅目害虫,本发明的昆虫抑制性蛋白质的组合可以与半翅目活性蛋白诸如TIC1415(美国专利公布2013-0097735A1)、TIC807(美国专利8609936)、TIC834(美国专利公布2013-0269060A1)、AXMI-036(美国专利公布2010-0137216A1)和AXMI-171(美国专利公布2013-0055469A1)一起使用。此外,用于防治鞘翅目

害虫的多肽可以选自由以下组成的组:昆虫抑制性蛋白质,诸如但不限于:Cr y3Bb(美国专利6,501,009)、Cry1C变体、Cry3A变体、Cry3、Cry3B、Cry34/35、5307、AXMI134(美国专利公布2013-0167264A1)、AXMI-184(美国专利公布2010-0004176A1)、AXMI-205(美国专利公布2014-0298538A1)、AXMI-207(美国专利公布2013-0303440A1)、AXMI-218、AXMI-220(美国专利公布20140245491A1)、AXMI-221z、AXMI-223z(美国专利公布2014-0196175A1)、AXMI-279(美国专利公布2014-0223599A1)、AXMI-R1及其变体(美国专利公布2010-0197592A1)、TIC407、TIC417、TIC431、TIC807、TIC853、TI C901、TIC1201、TIC3131、DIG-10(美国专利公布2010-0319092A1)、eHIP(美国专利申请公布2010/0017914)、IP3及其变体(美国专利公布2012-0210462A1)、假单胞菌毒素IDP072Aa(美国专利申请公布2014/055128)和ω-己毒素-Hv1a(美国专利申请公布US2014-0366227A1)。

[0084] 用于防治鞘翅目、鳞翅目和半翅目昆虫害虫的其他多肽,可以与TIC4029家族的昆虫抑制性蛋白质组合,可见于由Neil Crickmore维护的苏云金芽孢杆菌毒素命名法网站上(万维网网址btnomenclature.info)。广泛地,设想了本领域的技术人员已知的任何昆虫抑制性蛋白质均可以与TIC4029家族的蛋白质组合用于植物体中(通过育种或分子堆叠组合)或者在组合物或制剂中用作生物杀害虫剂或生物杀害虫剂的组合。

[0085] 昆虫对某些杀昆虫剂或杀昆虫性蛋白质产生耐受性的可能性在本领域中已有所记载。一种昆虫耐受性管理策略是采用表达两种不同的昆虫抑制剂的转基因作物,这两种昆虫抑制剂通过不同的作用模式发挥作用。因此,对昆虫抑制剂中的任一种具有耐受性的任何昆虫都可以通过另一种昆虫抑制剂来防治。另一种昆虫耐受性管理策略采用不针对目标鳞翅目害虫物种进行保护的植物的使用,从而为此类未受保护的植物提供保护设施。美国专利6,551,962中描述了一个特定实施例,该专利以引用的方式整体并入。

[0086] 其他实施方案(诸如局部施用的杀害虫性化学物质,该化学物质被设计用于防治另外可用本文公开的蛋白质防治的害虫,与蛋白质一起用于种子处理、喷洒、滴落或擦拭制剂)可以直接施用于土壤(土壤灌溉剂),施用于表达本文公开的蛋白质的生长植物,或配制用于含有编码一种或多种公开的蛋白质的一种或多种转基因的种子。此类用于种子处理的制剂可以与本领域已知的各种粘合剂和增粘剂一起施用。此类制剂可以含有在作用模式上与所公开的蛋白质协同作用的杀害虫剂,以使得杀昆虫剂制剂通过不同的作用模式发挥作用,以防治可以用所公开的蛋白质防治的相同的或相似的害虫,或者此类杀昆虫剂作用于防治更广泛宿主范围内的害虫或TIC4029、TIC4029_1和TIC4029_8杀害虫性蛋白质无法有效防治的植物害虫物种。

[0087] 前述组合物/制剂可以另外包含农业可接受的载剂,诸如诱饵、粉末、粉剂、沉淀、颗粒、喷洒剂、乳液、胶体悬浮液、水溶液、芽孢杆菌孢子/晶体制备物、种子处理剂、被转化为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、或者被转化为表达一种或多种蛋白质的细菌。根据重组多肽中固有的昆虫抑制性或杀虫抑制剂的水平以及将施用于植物或饮食测定法的制剂水平,组合物/制剂可以包含以重量计各种量的重组多肽,例如,以重量计0.0001%至0.001%至0.01%至1%至99%的重组多肽。

[0088] 综上所述,本领域的技术人员应当理解,可以对所公开的具体方面进行修改并且仍然获得相同或相似的结果,而不脱离本发明的精神和范围。因此,本文公开的特定结构和功能细节不应被解释为限制性的。应当理解,本文引用的每篇参考文献的全部公开内容均

并入本申请的公开内容中。

[0089] 实施例

[0090] 实施例1

[0091] TIC4029和TIC4029变体TIC4029_1和TIC4029_8的发现、克隆和表达

[0092] 对编码新型苏云金芽孢杆菌杀害虫性蛋白质的序列进行鉴定、克隆、序列确认，并且在昆虫生物测定法中进行测试。杀害虫性蛋白质TIC4029从Bt物种EGBS0016分离，代表新型Cry1Ca相关蛋白。Bt菌株EGBS0016最初被鉴定为含有Bt或Bt样细菌菌株的孢子形成、晶体和质粒。DNA从EGBS0016分离并测序。然后对组装的序列进行生物信息学分析。TIC4029蛋白通过pfam分析鉴定为内毒素结构域的命中以及与已知的Cry1Ca毒素的同一性。全长TIC4029蛋白氨基酸序列表现出与TIC1425(美国专利10,626,151)具有86％同一性。昆虫毒素TIC1425展示出针对棉叶虫(棉叶波纹夜蛾)、欧洲玉米螟(欧洲玉米螟)、秋粘虫(草地贪夜蛾)、甘蔗螟(小蔗螟)和西南玉米螟(西南玉米螟)的活性。聚合酶链式反应(PCR)引物被设计为从Bt物种EGBS0016分离的总基因组DNA扩增TIC4029的编码区的全长拷贝。PCR扩增子还包含编码序列的翻译起始和终止密码子。

[0093] 使用本领域已知的方法将TIC4029编码序列克隆到与Bt表达启动子可操作连接的Bt表达载体。将孢子和可溶性蛋白制备物用于生物测定法。

[0094] 此外，还生成了TIC4029的两个截短变体TIC4029_1和TIC4029_8。TIC4029_1包含原毒素结构域的缺失并且由相对于TIC4029的氨基酸1-603组成(参见图1a-1b)。合成TIC4029_1的编码序列并将该编码序列克隆到与Bt表达启动子可操作连接的细菌表达载体中。将孢子和可溶性蛋白制备物用于生物测定法。

[0095] TIC4029_8包含TIC4029的N-末端氨基酸片段和原毒素结构域的缺失，并且由TIC4029的氨基酸28-607组成(参见图1a-1b)。此外，随后引入ATG密码子，置换TIC4029的第27位氨基酸的精氨酸残基，以启动TIC4029_8蛋白的翻译。TIC4029_8氨基酸序列代表TIC4029的胰蛋白酶核心。如实施例4和5所述在植物体中测试编码TIC4029_8的序列的活性。

[0096] 实施例2

[0097] TIC4029和TIC4029_1在昆虫生物测定法中表现出鳞翅目毒性活性

[0098] 杀害虫性蛋白质TIC4029和TIC4029_1在Bt中表达并测定这两种蛋白质针对各种鳞翅目物种的毒性。还测定了TIC4029对鞘翅目、半翅目和双翅目的各种物种的毒性。

[0099] 测定了TIC4029针对以下物种的毒性：鳞翅目昆虫物种：黑地蚕(BCW，小地老虎)、玉米穗虫(CEW，玉米耳虫，也称为大豆豆荚虫)、欧洲玉米螟(ECB，欧洲玉米螟)、秋粘虫(FAW，草地贪夜蛾)、南部行军虫(SAW，南部灰翅夜蛾)、大豆尺蠖(SBL，大豆尺蠖)、西南玉米螟(SWC，西南玉米螟)、甘蔗螟(SCB，小蔗螟)、向日葵尺蠖(SFL，薄荷灰夜蛾)、烟草蛀虫(TBW，烟芽夜蛾)和天鹅绒豆毛虫(VBC，黎豆夜蛾)；鞘翅目物种：科罗拉多马铃薯甲虫(CPB，科罗拉多金花虫(Leptinotarsa decemlineata))和西部玉米根虫(WCR，美洲玉米根虫(Diabrotica virgifera))；半翅目物种：新热带棕纹椿象(NBSB，英雄美洲蝽(Euschistus heros))和西部牧草盲蝽(WTP，豆荚草盲蝽(Lygus hesperus))；以及双翅目物种：埃及伊蚊(YFM，埃及伊蚊(Aedes aegypti))。使用含有TIC4029的微生物宿主细胞来源的蛋白质样品进行的生物测定法展示出针对鳞翅目物种BCW、CEW、ECB、FAW、SAW、SBL、SWC、SCB、SFL、TBW和

VBC的活性。还观察到针对双翅目物种YFM的活性。

[0100] 细菌编码的TIC4029_1表现出针对SBL和VBC的毒性活性。此外，TIC4029和TIC4029_1表现出针对黑粘虫(BAW，考斯夜蛾)的毒性活性。

[0101] 实施例3

[0102] 用于植物表达的编码TIC4029、TIC4029_1或TIC4029_8的合成序列。

[0103] 合成的编码序列被设计为编码TIC4029、TIC4029_1或TIC4029_8，以在植物中表达这些毒素中的每种。

[0104] 合成序列是根据美国专利5,500,365中一般描述的方法合成的，以避免某些有害的问题序列，诸如ATTTA和富含A/T的植物多腺苷酸化序列，同时基本上保留天然蛋白质的氨基酸序列。TIC4029_PL-1(SEQ ID NO:5)、TIC4029_PL-2(SEQ ID NO:6)和TIC4029_PL-3(SEQ ID NO:7)各自是编码TIC4029的不同的编码序列构建体。一种合成的编码序列针对TIC4029_1(SEQ ID NO:8)而设计，一种合成的编码序列针对TIC4029_8(SEQ ID NO:9)而设计。

[0105] 这些编码序列使用本领域已知的技术转移到植物转化载体中。用于转化大豆植物的转化载体包含用于表达适用的杀害虫性蛋白质的第一转基因盒，包含可操作地连接至前导序列的植物功能性组成型启动子、毒素编码序列、非翻译区(UTR)以及植物功能性转录终止和多腺苷酸化序列。第二转基因盒也包含在载体中，所述第二转基因盒表达选择性标记，在这种情况下是壮观霉素耐受性。用于转化玉米植物的转化载体包含用于表达适用的杀害虫性蛋白质的第一转基因盒，并且包含可操作地连接至前导序列的植物功能性组成型启动子、内含子序列、非翻译区(UTR)(包含转录和翻译终止序列)和植物多腺苷酸化序列，并且包含编码用于选择经转化的植物细胞的草甘膦耐受性的第二转基因盒。

[0106] 实施例4

[0107] TIC4029、TIC4029_1和TIC4029_8在稳定转化的大豆植物中表现出鳞翅目毒性活性

[0108] 使用本领域已知的方法来构建包含转基因盒的二元植物转化载体，所述转基因盒被设计为表达杀害虫性蛋白质TIC4029、TIC4029_1和TIC4029_8。所得的载体被用于使用编码这些毒素蛋白的表达盒来稳定转化大豆植物。如本文所述，将表达这些毒素蛋白的植物组织用于针对各种鳞翅目昆虫物种的昆虫生物测定法。

[0109] 使用农杆菌介导的转化方法利用二元转化载体来转化大豆植物细胞。通过本领域已知的方法来诱导经转化的细胞形成植株。使用从经转化的植物获得的植物叶盘的生物测定法的进行方式类似于美国专利8,344,207中描述的方法。将一只不到一日龄的新孵化的新生幼虫放置于每个叶盘样品上，并允许摄食大约四天。将未经转化的大豆植物用作阴性对照。将从每个二元载体产生的多个R_0单拷贝插入事件针对南部行军虫(SAW，南部灰翅夜蛾)、大豆尺蠖(SBL，大豆尺蠖)、大豆豆荚虫(SPW，玉米耳虫)和天鹅绒豆毛虫(VBC，黎豆夜蛾)进行评估。根据生物测定法中每个事件的叶片损伤百分比以及共有最低损伤百分比范围(外显率)的事件的百分比，将范围为0至3的功效评分分数(功效分数)分配给每个事件，如表2所示。评分为3表示基于摄食该特定样品的毒性的最大有效影响。评分为2表示幼虫喂养的样品表现出生长抑制、蜕皮或导致发病以及通常死亡。

[0110] 表2.表达毒素的大豆植物的功效评分分数、叶片损伤百分比和外显率标准。

[0111]

功效分数	叶片损伤百分比	外显率
0	>50%	≥80%
1	<50%	≥20%

[0112]

2	<30%	≥20%
3	≤10%	≥50%

[0113] 表3展示了表达TIC4029、TIC4029_1和TIC4029_8的R_0大豆植物的功效分数。括号中的数字表示共有最低百分比范围的损伤的事件数/所测定的事件总数，—NT”表示未测试。对于TIC4029和TIC4029_1，使用多个构建体来转化植物，每个构建体包含不同的表达元件。

[0114] 表3.表达TIC4029、TIC4029_1和TIC4029_8的R_0大豆植物的功效分数。

[0115]

毒素	构建体	编码序列	SEQ ID NO:	SAW	SBL	SPW	VBC
			功效评分分数				
TIC4029	1	TIC4029PL-1	5	3 (17/20)	3 (19/20)	0 (18/20)	2 (7/20)
TIC4029	1	TIC4029PL-1	5	2 (14/15)	3 (14/15)	0 (15/15)	0 (15/15)
TIC4029	1	TIC4029PL-1	5	3 (12/14)	3 (13/14)	0 (14/14)	NT
TIC4029	2	TIC4029PL-1	5	2 (8/13)	3 (12/13)	0 (12/13)	NT
TIC4029	3	TIC4029PL-1	5	3 (12/14)	3 (13/14)	0 (14/14)	NT
TIC4029	4	TIC4029PL-1	5	0 (19/19)	2 (13/19)	0 (19/19)	NT
TIC4029	5	TIC4029PL-1	5	0 (12/14)	3 (12/14)	0 (14/14)	NT
TIC4029	6	TIC4029PL-1	5	2 (13/20)	3 (18/20)	0 (20/20)	NT
TIC4029	7	TIC4029PL-1	5	3 (18/20)	3 (20/20)	0 (20/20)	2 (5/20)
TIC4029	8	TIC4029PL-1	5	2 (14/20)	3 (19/20)	0 (20/20)	NT
TIC4029_1	1	TIC4029_1PL	8	3 (8/14)	3 (11/14)	0 (14/14)	2 (8/14)
TIC4029_1	1	TIC4029_1PL	8	3 (17/20)	3 (19/20)	0 (20/20)	3 (19/20)

[0116]

毒素	构建体	编码序列	SEQ ID NO:	SAW	SBL	SPW	VBC
			功效评分分数				
TIC4029_1	1	TIC4029_1PL	8	3 (19/20)	3 (20/20)	0 (20/20)	NT
TIC4029_1	2	TIC4029_1PL	8	NT	3 (15/16)	0 (16/16)	NT
TIC4029_1	3	TIC4029_1PL	8	2 (5/12)	3 (12/12)	0 (12/12)	NT
TIC4029_1	4	TIC4029_1PL	8	2 (4/16)	3 (13/16)	0 (16/16)	NT
TIC4029_1	5	TIC4029_1PL	8	1 (6/10)	3 (8/10)	0 (10/10)	NT
TIC4029_1	6	TIC4029_1PL	8	3 (16/18)	3 (17/18)	0 (18/18)	NT
TIC4029_1	7	TIC4029_1PL	8	3 (14/17)	3 (17/17)	0 (17/17)	2 (10/17)
TIC4029_8	1	TIC4029_8PL	9	3 (14/16)	3 (16/16)	0 (16/16)	NT
TIC4029_8	1	TIC4029_8PL	9	3 (13/14)	3 (11/14)	NT	3 (12/14)

[0117] 从表3中可以看出，R_0用于在稳定转化的大豆中表达TIC4029的10个构建体中有9个对SBL有效。用于在稳定转化的大豆中表达TIC4029_1的全部9种构建体和用于在稳定转化的大豆中表达TIC4029_8的两种构建体均对SBL有效。用于表达TIC4029PL-1的四种构建体对SAW有效，而另外4种构建体提供SAW的抑制。

[0118] 使选定的表达TIC4029、TIC4029_1和TIC4029_8的R_0大豆植物自花授粉并产生R_1种子。将R_1种子用于培植R_1植株。选择杀害虫性蛋白质表达盒纯合的R_1植物来进行针对SAW、SBL、SPW和VBC的叶盘生物测定法。表4显示了表达TIC4029、TIC4029_1和TIC4029_8的R_1纯合植物的功效评分分数。

[0119] 表4.表达TIC4029和TIC4029_1氨基酸变体的R_1大豆植物的功效评分分数。

[0120]

毒素	构建体	编码序列	SEQ ID NO:	SAW	SBL	SPW	VBC
TIC4029	构建体-1	TIC4029PL-1	5	3 (10/10)	3 (10/10)	0 (10/10)	1 (8/10)
TIC4029	构建体-1	TIC4029PL-1	5	3 (8/11)	NT	NT	NT
TIC4029	构建体-2	TIC4029PL-1	5	3 (5/7)	NT	NT	1 (5/7)
TIC4029	构建体-3	TIC4029PL-1	5	3 (8/8)	NT	NT	1 (6/8)
TIC4029	构建体-4	TIC4029PL-1	5	0 (7/7)	3 (6/7)	NT	0 (7/7)
TIC4029	构建体-5	TIC4029PL-1	5	3 (4/7)	NT	NT	0 (7/7)
TIC4029	构建体-6	TIC4029PL-1	5	3 (5/9)	3 (7/9)	NT	0 (9/9)
TIC4029	构建体-7	TIC4029PL-1	5	3 (10/10)	3 (10/10)	NT	0 (6/10)
TIC4029	构建体-8	TIC4029PL-1	5	3 (7/9)	3 (7/9)	NT	0 (8/9)
TIC4029_1	构建体-1	TIC4029_1PL	8	3 (8/10)	NT	NT	3 (8/10)
TIC4029_1	构建体-1	TIC4029_1PL	8	3 (11/11)	NT	NT	NT
TIC4029_1	构建体-2	TIC4029_1PL	8	3 (7/8)	NT	NT	3 (5/8)
TIC4029_1	构建体-3	TIC4029_1PL	8	3 (9/9)	NT	NT	3 (9/9)
TIC4029_1	构建体-4	TIC4029_1PL	8	3 (6/9)	NT	NT	0 (7/9)
TIC4029_1	构建体-5	TIC4029_1PL	8	3 (3/5)	NT	NT	1 (3/5)
TIC4029_1	构建体-6	TIC4029_1PL	8	3 (11/11)	NT	NT	3 (9/11)
TIC4029_1	构建体-7	TIC4029_1PL	8	3 (7/8)	NT	NT	3 (7/8)
TIC4029_8	构建体-1	TIC4029_8PL	10	3 (8/9)	NT	NT	NT

[0121] 从表4中可以看出，来源于表达TIC4029的5个选定的构建体的R_1纯合植物对SBL有效。来源于表达TIC4029的8个选定的构建体的R_1纯合植物对SAW有效。来源于表达TIC4029_1的8个选定的构建体和表达TIC4029_8的1个选定的构建体的R_1纯合植物对SAW有效。

[0122] 在一项试验中，使用如上文所述的叶盘测定法对用表达TIC4029的构建体-1和构建体-3转化产生的R_1纯合植物进行针对豆芽蛾(豆茎蛾)的测定，产生50％至100％的死亡率。TIC4029对豆芽蛾(豆茎蛾)具有活性。对用表达TIC4029的构建体-3转化产生R_1纯合植

物进行针对Cry1Ac耐受性向日葵尺蠖(SFL,薄荷灰夜蛾)的测定,该纯合植物对Cry1Ac耐受性SFL非常有效。

[0123] 实施例5

[0124] TIC4029、TIC4029_1和TIC4029_8在稳定转化的玉米植物中展示出针对西南玉米螟的活性

[0125] 使用本领域已知的方法来克隆包含转基因盒的二元植物转化载体,所述转基因盒被设计为表达TIC4029、TIC4029_1和TIC4029_8杀害虫性蛋白质。将所得的载体用于稳定转化玉米植物。从转化子收获组织,并将组织用于针对各种鳞翅目昆虫物种的昆虫生物测定法。

[0126] 如实施例3中所述,使用农杆菌介导的转化方法利用二元转化载体来转化玉米植物。通过本领域已知的方法来诱导经转化的细胞形成植株。使用植物叶盘的生物测定法的进行方式类似于美国专利8,344,207中描述的那些。将一只不到一日龄的新孵化的新生幼虫放置于每个叶盘样品上,并允许摄食大约四天。将未经转化的玉米植物用于获得组织,以用作阴性对照。将来自二元载体的多个转化R_0单拷贝插入事件针对西南玉米螟(SWC,西南玉米螟)进行评估。根据生物测定法中每个事件的叶片损伤百分比以及共有最低损伤百分比范围(外显率)的事件的百分比,将范围为0至3的功效评分分数分配给每个事件,如实施例3的表2所示。

[0127] 下表5显示了表达TIC4029、TIC4029_1和TIC4029_8的R_0稳定转化的玉米植物的功效分数。括号中的数字表示共有最低百分比范围的损伤的事件数/所测定的事件总数。对于TIC4029,使用两种构建体来转化植物,每种构建体包含编码TIC4029的不同编码序列,但是包含相同的表达元件。

[0128] 表5.表达TIC4029、TIC4029_1和TIC4029_8的R_0玉米植物针对西南玉米螟的功效分数。

[0129]

毒素	构建体	编码序列	SEQ ID NO:	SWC
TIC4029	构建体-1	TIC4029PL-1	5	1(2/3)
TIC4029	构建体-1	TIC4029PL-1	5	0(37/40)
TIC4029	构建体-1	TIC4029PL-1	5	2(11/31)
TIC4029	构建体-2	TIC4029PL-2	6	3(20/25)
TIC4029_1	构建体-1	TIC4029_1PL	8	2(12/40)
TIC4029_1	构建体-1	TIC4029_1PL	8	3(26/40)
TIC4029_8	构建体-1	TIC4029_8PL	10	3(13/18)

[0130] 从表5中可以看出,用构建体-2转化的表达TIC4029的R_0单拷贝玉米植物对SWC有效,所述构建体-2包含TIC4029编码序列TIC4029-2(SEQ ID NO:6)和TIC4029_1和TIC4029_8。来自编码序列TIC4029PL-1的TIC4029的表达水平远低于针对TIC4029PL-2观察到的表达水平,这可能是来源于包含TIC4029PL-1编码序列的构建体的功效分数较低的原因。使选定的R_0单拷贝植物与未经转化的优良玉米品种杂交,该单拷贝植物使用编码序列TIC4029PL-2利用表达TIC4029的构建体-2来转化。将来源于该杂交的杂合F_1事件针对SWC进行测定。这些杂合F_1事件对SWC有效。

[0131] 实施例6

[0132] TIC4029对大豆尺蠖、向日葵尺蠖和天鹅绒豆毛虫有效并且在温网室试验中抑制南部行军虫

[0133] 在美国和阿根廷的温网室试验中，对表达TIC4029的大豆植物进行针对选定的昆虫害虫物种的保护测定。

[0134] 在美国的2019年生长季节期间，在多个位置的温网室试验中，将表达TIC4029的大豆植物针对南部行军虫(SAW，南部灰翅夜蛾)、大豆尺蠖(SBL，大豆尺蠖)、大豆豆荚虫(SPW，玉米耳虫)和天鹅绒豆毛虫(VBC，黎豆夜蛾)进行测定。温网室试验在伊利诺伊州泽西维尔(Jerseyville,IL)针对SAW和SBL进行，在田纳西州尤宁城(Union City,TN)针对SBL进行，在伊利诺伊州沃特曼(Waterman,IL)针对VBC和SPW进行。使用随机完整时钟设计来评估事件。每个事件田地均种植成一个六(6)英尺的行，每英尺有大约八(8)粒种子。每个事件有三(3)个重复；因此，每个事件均在温网室中以三(3)个单独的田地表示，这些田地随机定位于温网室中。将未经转化的事件作为阴性对照，将其田地随机分配到温网室内的位置。

[0135] SPW和VBC的侵扰使用成虫来完成。昆虫在田纳西州尤宁城的昆虫饲养场的成虫羽化笼中饲养成蛹，并维持在气候控制的培养箱中。将昆虫运输至伊利诺伊州沃特曼和泽西维尔，在温网室中释放。在温网室中，每次释放使用大约一千二百(1,200)至两千(2,000)只成虫。对于SPW，从大豆发育的R1至R2阶段，每周在温网室释放成虫一次。对于VBC，在V4至R3的发育阶段之间，每两周在温网室中释放成虫一次。在温网室中，每次释放大约一千二百(1,200)至两千(2,000)只成虫。成虫需要持续接触百分之十(10％)的蔗糖溶液才能正常生存和生殖。向塑料食品容器中添加脱脂棉，然后将糖溶液倒入容器中，使棉花完全浸透。每天补充糖溶液直到成虫活动平息，这通常在成虫最终释放后两周左右。

[0136] 对于SAW侵扰，使用直接卵侵扰，因为这种昆虫不会优先或均匀地在大豆上产卵。每次侵扰使用大约二十五万(250,000)至三十二万(320,000)个卵，并且从R1至R3发育阶段每两周施用一次。通过以下步骤将含有相同数量SAW卵的纸片附着在植物上：将纸片折叠在上层树冠中的坚固叶柄上，并且将纸片牢固地装订在一起。将一(1)张纸片放置于田地前端一(1)英尺内的植物上，将第二张纸片放置于田地中间的植物上，将第三张纸片放置于田地后端一(1)英尺内的植物上。

[0137] 在植物发育的不同阶段评估落叶百分比。对于SAW，在伊利诺伊州泽西维尔(at Jerseyville,IL)，在R2.8、R4.1、R4.8和R6.0发育阶段评估落叶百分比。对于SBL，在田纳西州尤宁城(Union City,TN)，在R2.0、R3.1、R4.2和R5.5发育阶段确定落叶百分比，并且在伊利诺伊州泽西维尔(Jerseyville,IL)，在R5.4和R5.8发育阶段确定落叶百分比。对于VBC，在伊利诺伊州沃特曼(Waterman,IL)，在R3.9、R5.0和R5.4发育阶段评估落叶百分比。对于SPW，在伊利诺伊州沃特曼(Waterman,IL)，在R4.1、R4.7、R5.4和R5.8发育阶段评估落叶百分比。最大落叶百分比来源于在每种昆虫的不同发育阶段观察到的最高落叶百分比。下表6显示了表达TIC4029的植物针对SAW、SBL和VBC的平均最大落叶百分比。SPW的平均最大落叶百分比与阴性对照相似，未在表6中提供。

[0138] 表6.在美国温网室试验中表达TIC4029的大豆植物的平均最大落叶百分比。

[0139]

	SAW		SBL		VBC	
位置	阴性	TIC4029	阴性	TIC4029	阴性	TIC4029
Jerseyville, IL	56.5	12.5	25.0	0.0		
Union City, TN			71.5	0.2		
Waterman, IL					50.3	0.8

[0140] 从表6中可以看出,表达TIC4029的植物可有效防治SBL和VBC;并且展示出对SAW的抑制。

[0141] 在2019-2020年生长季节期间,还在阿根廷的两个位置弗兰路易斯(Fran Luis,BA)和佩尔加米诺(Pergamino,BA)对表达TIC4029的大豆植物进行温网室试验。温网室试验的进行方式与2017年在美国进行的试验相似。温网室中的每个田地包括种植成两(2)米的行,每行有四十二(42)粒种子。每个事件以随机位于温网室内的三(3)个代表性样品表示。温网室试验针对鳞翅目昆虫害虫大豆尺蠖(SBL,大豆尺蠖)、天鹅绒豆毛虫(VBC,黎豆夜蛾)、南美豆荚虫(SAPW,大豆荚虫)、向日葵尺蠖(SFL,薄荷灰夜蛾)和黑粘虫(BAW,考斯夜蛾)进行。

[0142] 在植物发育的不同阶段评估落叶百分比。对于SBL,在弗兰路易斯(Fran Luis,BA),在R5.0,R5.5和R6.0发育阶段评估落叶百分比,并且在佩尔加米诺(Pergamino,BA),在R4.0、R5.1和R5.6发育阶段评估落叶百分比。对于VBC,在弗兰路易斯(Fran Luis,BA),在R5.5,R6.0和R6.5发育阶段评估落叶百分比,并且在佩尔加米诺(Pergamino,BA),在R5.0、R5.6和R6.0发育阶段评估落叶百分比。对于SFL,在弗兰路易斯(Fran Luis,BA),在R5.0、R5.3、R5.5和R6.0发育阶段评估落叶百分比,并且在佩尔加米诺(Pergamino,BA),在R3.0、R4.0、R5.2和R6.2发育阶段评估落叶百分比。对于SAPW,在弗兰路易斯(Fran Luis,BA),在R4.4、R5.1、R5.5和R6.0发育阶段评估落叶百分比,并且在佩尔加米诺(Pergamino,BA),在R3.0、R4.0、R5.1和R6.2发育阶段评估落叶百分比。对于BAW,在弗兰路易斯(Fran Luis,BA),在R3.0、R5.0、R5.5和R6.0发育阶段评估落叶百分比,并且在佩尔加米诺(Pergamino,BA),在R5.1、R5.4和R6.0发育阶段评估落叶百分比。对于每个位置的昆虫害虫中的每种,如本实施例中的上文所述测定最大落叶百分比。下表7显示了表达TIC4029的植物的平均最大落叶百分比。SAPW和BAW的最大落叶百分比与对照相同,未包括在表7中。

[0143] 表7.在阿根廷温网室试验中表达TIC4029的大豆植物的平均最大落叶百分比。

[0144]

昆虫	转基因	Fran Luis, BA, ARG	Pergamino, BA, ARG
SBL	阴性	82.5	62.0
	TIC4029	3.1	5.2
VBC	阴性	32.0	88.0
	TIC4029	4.5	6.0

[0145]

SFL	阴性	47.2	37.6
	TIC4029	5.5	4.3

[0146] 从表7中可以看出，表达TIC4029的大豆植物对SBL、VBC和SFL有效。

[0147] TIC4029对SBL、VBC和SFL有效；并且提供SAW的抑制。

[0148] 根据本公开，无需过度实验即可制备和执行本文公开和要求保护的所有组合物。虽然本发明的组合物已经根据前述示例性实施方案进行了描述，但是对于本领域的技术人员而言显而易见的是，可以对本文所述的组合物进行变异、变化、修改和改变，而不脱离本发明的真实概念、精神和范围。更具体而言，将会明白，在化学和生理学上均相关的某些剂可以替代本文所述的剂，同时将实现相同或相似的结果。对本领域技术人员来说显而易见的所有这些类似的替换和修改均被认为是在所附权利要求所限定的本发明的精神、范围和概念之内。

[0149] 本说明书中引用的所有出版物和公布的专利文档均以引用的方式并入本文，如同每个单独的出版物或专利申请均被具体地和单独地指示为以引用的方式并入的程度。

序列表

<110> 孟山都技术有限责任公司(Monsanto Technology LLC)

<120> 新型昆虫抑制性蛋白质

<130> MONS:484WO

<150> US 63/128,775
<151> 2020-12-21

<160> 10

<170> PatentIn version 3.5

<210> 1
<211> 3534
<212> DNA
<213> 苏云金芽孢杆菌(Bacillus thuringienses)

<220>
<221> misc_feature
<222> (1)..(3534)
<223> 编码从苏云金芽孢杆菌(Bacillus thuringiensis)物种EGBS0016获得的TIC4029杀虫性蛋白质的核酸序列。

[0001]

<400> 1
atgaataata atattgaaaa ccaatgcgta ccttacaatt gtttaagtaa tcctgaagaa 60

gtacttttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca 120

cttgttcaac ttctggtatc taactttgta ccaggggggag gatttttagt tggattaata 180

gattttgtat ggggaatagt aggcccttct caatgggatg catttctagt gcaaattgaa 240

caattaattc agcaaagaat agaagcatat gctagggctg cagcaatttc taatttagaa 300

ggaataggaa acaatttcaa tatatatgtg gaagcatttc aagaatggga agaagatcct 360

aataatccag cgacaagaaa tagagtagtt gatcgctttc gtatacttga tgggctactt 420

gaaagggaca ttccttcgtt tcgaatttct ggatttgaag tccccctttt atccgtttat 480

actcaagcgg ccaatctgca tctagctata ttgagggatt ctgtaatttt tggggaaaga 540

tggggattaa caacgacaaa tgtcaatgaa aactataata gacaaattag gcatattaat 600

gaatatgccg atcactgtgc aaatacgtac aatcggggac tcaataattt accgaaatct 660

acgtatcaag attggataac atataatcga ttacggagag acttaacatt gactgtatta 720

gatatcgtca atttctttcc aaactatgac aataggagat atccaattca aacagttggt 780

caactaacaa gggaagttta tacggaccca ttaattaatt ttaatccaca gttacagtct 840

gtagctcaat tacctacttt taacgttatg gaaagcagcg caattagaaa tcctcattta 900

tttgatgtat tgaataatct tacaattttt acagactggt ttagtgttgg acgcaacttt 960

tattggggag gccatcgagt aatatctagc cgtataggag gtggtaacat aacatctcct 1020

atatatggaa gagaggcgaa tcaggagcct ccaagatctt ttacttttaa tggaccggtt 1080

tttaggactc tatcaaatcc tactttaaga tcattacagc aaccttggcc agcgccacca 1140

tttaatttac gtggtgttga aggagtagaa ttttctacac ctacaaatag ctttacgtat 1200

cgaggaagag gtacggttga ttctttaact gagttaccgc ctcaggataa tagtgtgcca 1260

cctcgcgaag gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaaca 1320

ccttttttaa caacaggtgt agtattttct tggacgcatc gtagtgccga tcaaaatata 1380

atctatccaa atagaattac ccaaatacca ttggtaaaag catctgatct cccttcaggt 1440

[0002] accactgttg ttagaggacc agggtttaca ggtggggata ttcttcgaag aacaagtact 1500

ggtggattcg gcacgataag agtaaatgtt aatggaacac taacacaaag atatcgtata 1560

ggatttcgct atgcttcaac agtagatttt gatttctttg tagtacgtgg agggactact 1620

gtaaataatt ttagattccc acgtacaatg aacagtggag aggaattaag atacggatcc 1680

tttgagacaa ggtctttcac tactcctttt acatttactc aaattcagga tacaattcga 1740

acgtctattc aaggtcttag cggtaatggt gaagtgtatc ttgacagaat cgagatcatt 1800

ccagttactg caaccttaaa ggcagaatat gatttagaaa gagcgcagaa ggtggtgggt 1860

gccctgttta cttccacaaa ccaactagcg ctaaaaacaa atattacgga ttatcatatt 1920

gatcaagttt ccaatttagt ggattgttta tccgatgaat tttgtctgga tgaaaagcga 1980

gaattgtccg agaaagtcaa acatgcgaag cgactcagtg atgagcggaa tttactccaa 2040

gatccaaact ttagaggcat caatagacaa ccagaccgtg gctggagagg aagtacggat 2100

attaccatcc aaggaggaaa tgacgtattc aaagaaaatt acgtcacact accaggtacc 2160

tttgatgagt gttatccaac gtatttgtat caaaaaatag atgaatcaaa attaaaatct 2220

gatacccgtt atcaattaag agggtatatt gaagatagtc aagacttaga aatctattta 2280

attcgctaca atgcaaaaca cgaaacagta aatgttcccg gtacgagttt cttgtggccg 2340

ctttcggtcg aaagtcctat tgggaagtgc ggagaaccga atcgttgcgc accacacatt 2400

gaatggaatc ctgatctaga atgttcctgt agagacggag aaaaatgtgc acatcattcc 2460

catcatttct ccctagacat tgatgttgga tgtacagact taaatgagga cttaggtgta 2520

tgggtgatat tcaagattaa gacgcaagat ggccatgcaa gactagggaa tttagagttt 2580

ctcgaagaga aaccgttatt aggagaagcg ttagctcgtg cgaaaagagc ggagaaaaaa 2640

tggagagaca aacgcgaaac attgcaattg gaaacaaata ttgtttataa agaggcaaaa 2700

gaatctgtag atgctttatt tgtgaactct caatatgata gattacaagc ggataccgac 2760

atcgcgatga ttcatgcggc agataaacgc gttcatcgaa ttcgagaagc atatcttcca 2820

gagttatctg taattccggg tgtaaatgcg ggtatttttg aagaattaga gggacgtatt 2880

ttcacagcat actctctata tgatgcgaga aatgtcatta aaaatggcga tttcaataat 2940

[0003] ggcttatcat gctggaacgt gaaagggcat gtagatatag aagaacaaaa caaccaccgt 3000

tcggttcttg ttgtcccgga atgggaagca gaagtgtcac aagaagttcg tgtctgtcca 3060

ggacgtggtt atatccttcg agtcacagcg tacaaagagg gatatggaga aggttgtgta 3120

accattcatg aaattgaaga tcatacagat gaactaaagt ttagcaactg tgtagaagag 3180

gacgcatatc caggtaatac ggtggcatgt gataattatc ccgcgaatca agaagaagga 3240

tgtacagagt tacgtcattc cggtaatcgc ggatatgatg aaacctatgt gaattctact 3300

tcctcatcca ctgattacac agcggtctat aaggaagaat catatacggg tgaacagaga 3360

tatgattctt gtgaatctaa cagaggatat gggaattata cgccactacc agctggttat 3420

gtgacaaaag agttagagta cttcccagaa accgataagg tatggattga gatcggagaa 3480

acggaaggaa cattcatcgt agacagcgtg gaattactcc ttatggaaga atag 3534

<210> 2
<211> 1177
<212> PRT
<213> 苏云金芽孢杆菌(Bacillus thuringienses)

<220>
<221> MISC_FEATURE
<222> (1)..(1177)
<223> TIC4029 杀虫性蛋白质的氨基酸序列。

<400> 2

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

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

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Leu Leu Val Ser Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
50 55 60

[0004] Gly Ile Val Gly Pro Ser Gln Trp Asp Ala Phe Leu Val Gln Ile Glu
65 70 75 80

Gln Leu Ile Gln Gln Arg Ile Glu Ala Tyr Ala Arg Ala Ala Ala Ile
85 90 95

Ser Asn Leu Glu Gly Ile Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
100 105 110

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

Val Val Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
130 135 140

Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145 150 155 160

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

Phe Gly Glu Arg Trp Gly Leu Thr Thr Thr Asn Val Asn Glu Asn Tyr
180 185 190

Asn Arg Gln Ile Arg His Ile Asn Glu Tyr Ala Asp His Cys Ala Asn
195 200 205

Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
210 215 220

Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
225 230 235 240

Asp Ile Val Asn Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
245 250 255

Gln Thr Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu Ile
260 265 270

[0005]

Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
275 280 285

Val Met Glu Ser Ser Ala Ile Arg Asn Pro His Leu Phe Asp Val Leu
290 295 300

Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
305 310 315 320

Tyr Trp Gly Gly His Arg Val Ile Ser Ser Arg Ile Gly Gly Gly Asn
325 330 335

Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
340 345 350

Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
355 360 365

Leu Arg Ser Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
370 375 380

Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr Tyr
385 390 395 400

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

Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
420 425 430

Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val Val
435 440 445

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

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

[0006]

Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg
485 490 495

Arg Thr Ser Thr Gly Gly Phe Gly Thr Ile Arg Val Asn Val Asn Gly
500 505 510

Thr Leu Thr Gln Arg Tyr Arg Ile Gly Phe Arg Tyr Ala Ser Thr Val
515 520 525

Asp Phe Asp Phe Phe Val Val Arg Gly Gly Thr Thr Val Asn Asn Phe
530 535 540

Arg Phe Pro Arg Thr Met Asn Ser Gly Glu Glu Leu Arg Tyr Gly Ser
545 550 555 560

Phe Glu Thr Arg Ser Phe Thr Thr Pro Phe Thr Phe Thr Gln Ile Gln
565 570 575

Asp Thr Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn Gly Glu Val
580 585 590

Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Thr Ala Thr Leu Lys Ala
595 600 605

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

Ser Thr Asn Gln Leu Ala Leu Lys Thr Asn Ile Thr Asp Tyr His Ile
625 630 635 640

Asp Gln Val Ser Asn Leu Val Asp Cys Leu Ser Asp Glu Phe Cys Leu
645 650 655

Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys His Ala Lys Arg Leu
660 665 670

Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn Phe Arg Gly Ile Asn
675 680 685

[0007] Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr Asp Ile Thr Ile Gln
690 695 700

Gly Gly Asn Asp Val Phe Lys Glu Asn Tyr Val Thr Leu Pro Gly Thr
705 710 715 720

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

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

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

Thr Val Asn Val Pro Gly Thr Ser Phe Leu Trp Pro Leu Ser Val Glu
770 775 780

Ser Pro Ile Gly Lys Cys Gly Glu Pro Asn Arg Cys Ala Pro His Ile

785 790 795 800

Glu Trp Asn Pro Asp Leu Glu Cys Ser Cys Arg Asp Gly Glu Lys Cys
805 810 815

Ala His His Ser His His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr
820 825 830

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

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

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

Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu Thr Asn Ile Val Tyr
[0008] 885 890 895

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

Asp Arg Leu Gln Ala Asp Thr Asp Ile Ala Met Ile His Ala Ala Asp
915 920 925

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

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

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

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

Ile Glu Glu Gln Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp
995 1000 1005

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

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

Cys Val Thr Ile His Glu Ile Glu Asp His Thr Asp Glu Leu Lys
1040 1045 1050

Phe Ser Asn Cys Val Glu Glu Asp Ala Tyr Pro Gly Asn Thr Val
1055 1060 1065

Ala Cys Asp Asn Tyr Pro Ala Asn Gln Glu Glu Gly Cys Thr Glu
1070 1075 1080

[0009]

Leu Arg His Ser Gly Asn Arg Gly Tyr Asp Glu Thr Tyr Val Asn
1085 1090 1095

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

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

Gly Tyr Gly Asn Tyr Thr Pro Leu Pro Ala 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> 3
<211> 1812
<212> DNA
<213> 人工(Artificial)

<220>
<223> 编码包含截短的 TIC4029_1 杀虫性蛋白质的核酸序列，
其中编码 TIC4029 的原毒素结构域的编码序列
已经缺失。

<400> 3

[0010]

atgaataata atattgaaaa ccaatgcgta ccttacaatt gtttaagtaa tcctgaagaa 60

gtacttttgg atggagaacg gatatcaact ggtaattcat caattgatat ttctctgtca 120

cttgttcaac ttctggtatc taactttgta ccaggggggag gatttttagt tggattaata 180

gattttgtat ggggaatagt aggcccttct caatgggatg catttctagt gcaaattgaa 240

caattaattc agcaaagaat agaagcatat gctagggctg cagcaatttc taatttagaa 300

ggaataggaa acaatttcaa tatatatgtg gaagcatttc aagaatggga agaagatcct 360

aataatccag cgacaagaaa tagagtagtt gatcgctttc gtatacttga tgggctactt 420

gaaagggaca ttccttcgtt tcgaatttct ggatttgaag tccccctttt atccgtttat 480

actcaagcgg ccaatctgca tctagctata ttgagggatt ctgtaatttt tggggaaaga 540

tggggattaa caacgacaaa tgtcaatgaa aactataata gacaaattag gcatattaat 600

gaatatgccg atcactgtgc aaatacgtac aatcggggac tcaataattt accgaaatct 660

acgtatcaag attggataac atataatcga ttacggagag acttaacatt gactgtatta 720

gatatcgtca atttctttcc aaactatgac aataggagat atccaattca aacagttggt 780

caactaacaa gggaagttta tacggaccca ttaattaatt ttaatccaca gttacagtct 840

gtagctcaat tacctacttt taacgttatg gaaagcagcg caattagaaa tcctcattta 900

tttgatgtat tgaataatct tacaattttt acagactggt ttagtgttgg acgcaacttt 960

tattggggag gccatcgagt aatatctagc cgtataggag gtggtaacat aacatctcct 1020

atatatggaa gagaggcgaa tcaggagcct ccaagatctt ttacttttaa tggaccggtt 1080

tttaggactc tatcaaatcc tactttaaga tcattacagc aaccttggcc agcgccacca 1140

tttaatttac gtggtgttga aggagtagaa ttttctacac ctacaaatag ctttacgtat 1200

cgaggaagag gtacggttga ttctttaact gagttaccgc ctcaggataa tagtgtgcca 1260

cctcgcgaag gatatagtca tcgtttatgt catgcaactt ttgttcaaag atctggaaca 1320

ccttttttaa caacaggtgt agtattttct tggacgcatc gtagtgccga tcaaaatata 1380

atctatccaa atagaattac ccaaatacca ttggtaaaag catctgatct cccttcaggt 1440

accactgttg ttagaggacc agggtttaca ggtggggata ttcttcgaag aacaagtact 1500

ggtggattcg gcacgataag agtaaatgtt aatggaacac taacacaaag atatcgtata 1560

ggatttcgct atgcttcaac agtagatttt gatttctttg tagtacgtgg agggactact 1620

gtaaataatt ttagattccc acgtacaatg aacagtggag aggaattaag atacggatcc 1680

tttgagacaa ggtctttcac tactcctttt acatttactc aaattcagga tacaattcga 1740

acgtctattc aaggtcttag cggtaatggt gaagtgtatc ttgacagaat cgagatcatt 1800

ccagttactt ag 1812

[0011]

<210> 4
<211> 603
<212> PRT
<213> 人工(Artificial)

<220>
<223> 包含原毒素结构域的截短的 TIC4029_1
杀虫性蛋白质的氨基酸序列。

<400> 4

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

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

Ser Ser Ile Asp Ile Ser Leu Ser Leu Val Gln Leu Leu Val Ser Asn
35 40 45

Phe Val Pro Gly Gly Gly Phe Leu Val Gly Leu Ile Asp Phe Val Trp
50 55 60

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

Gln Leu Ile Gln Gln Arg Ile Glu Ala Tyr Ala Arg Ala Ala Ala Ile
85 90 95

Ser Asn Leu Glu Gly Ile Gly Asn Asn Phe Asn Ile Tyr Val Glu Ala
100 105 110

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

Val Val Asp Arg Phe Arg Ile Leu Asp Gly Leu Leu Glu Arg Asp Ile
130 135 140

Pro Ser Phe Arg Ile Ser Gly Phe Glu Val Pro Leu Leu Ser Val Tyr
145 150 155 160

[0012]

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

Phe Gly Glu Arg Trp Gly Leu Thr Thr Thr Asn Val Asn Glu Asn Tyr
180 185 190

Asn Arg Gln Ile Arg His Ile Asn Glu Tyr Ala Asp His Cys Ala Asn
195 200 205

Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro Lys Ser Thr Tyr Gln Asp
210 215 220

Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp Leu Thr Leu Thr Val Leu
225 230 235 240

Asp Ile Val Asn Phe Phe Pro Asn Tyr Asp Asn Arg Arg Tyr Pro Ile
245 250 255

Gln Thr Val Gly Gln Leu Thr Arg Glu Val Tyr Thr Asp Pro Leu Ile
260 265 270

Asn Phe Asn Pro Gln Leu Gln Ser Val Ala Gln Leu Pro Thr Phe Asn
275 280 285

Val Met Glu Ser Ser Ala Ile Arg Asn Pro His Leu Phe Asp Val Leu
290 295 300

Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe Ser Val Gly Arg Asn Phe
305 310 315 320

Tyr Trp Gly Gly His Arg Val Ile Ser Ser Arg Ile Gly Gly Gly Asn
325 330 335

Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala Asn Gln Glu Pro Pro Arg
340 345 350

Ser Phe Thr Phe Asn Gly Pro Val Phe Arg Thr Leu Ser Asn Pro Thr
355 360 365

[0013]

Leu Arg Ser Leu Gln Gln Pro Trp Pro Ala Pro Pro Phe Asn Leu Arg
370 375 380

Gly Val Glu Gly Val Glu Phe Ser Thr Pro Thr Asn Ser Phe Thr Tyr
385 390 395 400

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

Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser His Arg Leu Cys His Ala
420 425 430

Thr Phe Val Gln Arg Ser Gly Thr Pro Phe Leu Thr Thr Gly Val Val
435 440 445

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

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

Thr Thr Val Val Arg Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Arg
485 490 495

Arg Thr Ser Thr Gly Gly Phe Gly Thr Ile Arg Val Asn Val Asn Gly
500 505 510

Thr Leu Thr Gln Arg Tyr Arg Ile Gly Phe Arg Tyr Ala Ser Thr Val
515 520 525

Asp Phe Asp Phe Phe Val Val Arg Gly Gly Thr Thr Val Asn Asn Phe
530 535 540

Arg Phe Pro Arg Thr Met Asn Ser Gly Glu Glu Leu Arg Tyr Gly Ser
545 550 555 560

Phe Glu Thr Arg Ser Phe Thr Thr Pro Phe Thr Phe Thr Gln Ile Gln
565 570 575

[0014] Asp Thr Ile Arg Thr Ser Ile Gln Gly Leu Ser Gly Asn Gly Glu Val
580 585 590

Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val Thr
595 600

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

<220>
<223> 用于在植物细胞中表达的合成编码序列，
编码 TIC4029 的 TIC4029PL-1。

<400> 5
atgaacaaca acattgagaa ccagtgcgtg ccctacaact gcctgagcaa ccctgaggag 60

gtgctcctag acggtgagcg catctccacc ggtaactcta gcatcgacat ctccctgagc 120

ctagtgcagc tcctagtgtc taacttcgtg cccggcggtg gcttcctcgt gggcctcatc 180

gacttcgtgt ggggcatcgt gggcccctcc cagtgggacg ccttcctcgt gcagatcgag 240

cagctcatcc agcaacgcat cgaggcctac gcccgcgccg ctgccatcag caacctcgag 300

ggcatcggca acaatttcaa catctacgtg gaggccttcc aggagtggga ggaagacccc 360

aacaatcccg ccacccgcaa ccgcgtggtc gaccgcttcc gcatcctcga cggcctcctg 420

gagcgcgaca tcccctcctt ccgcatctcc ggcttcgagg tgcccctcct gtccgtgtac 480

acccaggccg ctaacctcca cctcgccatc ctccgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactaccaa cgtgaacgag aactacaacc gccagatccg ccacatcaac 600

gagtacgccg accactgcgc caacacctac aaccgcggcc tcaacaatct ccccaagtcc 660

acctaccagg actggatcac ctacaaccgc ctccgcaggg acctcaccct caccgtgctc 720

gacatcgtga acttctttcc caactacgac aaccgcaggt accccatcca gaccgtgggc 780

cagctcaccc gcgaggtgta caccgacccc ctcatcaact tcaaccccca gctccagtcc 840

gtggcccagc tccccacctt caacgtgatg gagtccagcg ccatccgcaa cccccacctc 900

ttcgacgtgc tcaacaatct caccatcttc accgactggt tctccgtggg ccgcaacttc 960

[0015] tactggggcg gtcaccgcgt gatctccagc cgcatcggcg gtggcaacat cacctccccc 1020

atctacggcc gcgaggccaa ccaggagcct ccgcgctcct tcaccttcaa cggccctgtg 1080

ttccgcaccc tctccaaccc taccctcagg tccctccagc aaccttggcc tgcccctccg 1140

ttcaacctca ggggcgtgga gggcgtggag ttctccaccc ctaccaactc cttcacctac 1200

aggggcaggg gcaccgtgga ctccctcacc gagctgcctc cgcaggacaa ctccgtgcct 1260

ccgagggagg gctacagcca caggctgtgc cacgccacct tcgtgcagag gagcggcacc 1320

cctttcctga ccactggcgt cgttttcagc tggacccaca ggagcgccga ccagaacatc 1380

atttacccta acaggatcac ccagatccct ctggtcaagg ccagcgacct gcctagcggc 1440

accactgtcg ttaggggccc tggcttcacc ggcggtgaca tcctgaggcg gaccagcacc 1500

ggcggtttcg gcaccatcag ggtcaacgtc aacggcaccc tgactcagag gtacaggatc 1560

ggcttcaggt acgccagcac tgtcgacttc gacttctttg tcgttagggg cggtactact 1620

gtcaacaatt tcaggttccc gaggactatg aacagcggcg aggaactgag gtacggtagc 1680

ttcgagacta ggtctttcac tacgccgttc actttcactc agatccagga cactatccgg 1740

actagtatcc agggtctgtc tggtaacggt gaggtctacc tggaccggat cgagatcatt 1800

ccggtcactg ctactctgaa ggctgagtac gacctggagc gggctcagaa ggtcgttggt 1860

gctctgttca cttctactaa ccagctggct ctgaagacta acatcactga ctaccacatc 1920

gaccaggtct ctaacctggt cgactgcctg tctgacgagt tctgcctgga cgagaagcgg 1980

gagctgtctg agaaggtcaa gcacgctaag cggctgtctg acgagcggaa cctgcttcag 2040

gacccgaact tccggggtat caaccggcag ccggaccggg gttggcgggg tagtactgac 2100

atcactattc agggtgggaa cgacgtcttt aaggagaact acgtcacgct gccgggtacg 2160

tttgacgagt gctacccgac gtacctgtac cagaagattg acgagtctaa gctgaagtct 2220

gacacgcgtt accagctgcg tggttacatt gaggatagtc aggatcttga gatttacctt 2280

attcgttaca acgctaagca cgagacggtc aacgtcccgg gtacgtcttt tctttggccg 2340

ctttctgtcg agtcgccgat tgggaagtgc ggggagccaa accgttgcgc tccacacatt 2400

gagtggaacc cagatcttga gtgctcgtgc cgtgatgggg agaagtgcgc tcaccatagt 2460

caccatttta gtcttgatat tgatgtcggg tgcacggatc ttaatgagga tcttggggtc 2520

[0016]

tgggtcattt ttaagattaa gacgcaagat gggcacgcta gacttgggaa tcttgagttt 2580

cttgaggaaa agccactttt gggggaggct cttgctagag ctaagcgtgc ggagaagaaa 2640

tggcgtgata agcgtgagac gcttcaactt gagacgaata ttgtctacaa ggaggcgaag 2700

gagtcggtcg atgcgctttt tgtcaatagt caatacgata gacttcaagc ggatacggat 2760

attgcgatga ttcacgcggc agataagcgt gtccatagaa ttcgtgaggc gtaccttcca 2820

gagttgtcgg tcattccagg ggttaatgcg gggatttttg aggaattgga ggggcgtatt 2880

tttacggcgt actcgttgta cgatgcgaga aatgttatta agaatgggga ttttaataat 2940

gggttgtcgt gctggaatgt taaggggcat gttgatatag aggaacaaaa taatcataga 3000

tcggttttgg ttgtaccaga gtgggaggca gaggtttcgc aagaggttag agtttgccca 3060

gggagaggat acatattgag agttacggca tataaggagg gatatggaga gggatgtgtt 3120

acgatacatg agatagagga tcatacagat gagttgaaat ttagtaattg tgttgaggaa 3180

gatgcatatc caggaaatac agttgcatgt gataattatc cagcaaatca agaggaagga 3240

tgtacagagt tgcgacattc gggaaatcga ggatatgatg aaacatatgt taattcaaca 3300

tcatcatcaa cagattatac agcagtttat aaagaagaat catatacagg agaacaacga 3360

tatgattcat gtgaatcaaa tcgaggatat ggaaattata caccactacc agcaggatat 3420

gttacaaaag aactagaata ttttccagaa acagataaag tatggataga aataggagaa 3480

acagaaggaa catttatagt agattcagta gaattattat tgatggaaga atga 3534

<210> 6
<211> 3534
<212> DNA
<213> 人工(Artificial)

<220>
<223> 用于在植物细胞中表达的合成编码序列，
编码 TIC4029 的 TIC4029PL-2。

<400> 6
atgaacaaca acatcgagaa ccagtgcgtc ccctacaact gcctgtcgaa cccagaggag 60

gtgctgctag acggcgagcg tatctcaacg ggcaacagct ccatcgacat ctccctttcg 120

ctcgtgcagc tcttggtgtc caacttcgtg cccggcggtg gattcctggt ggggctgatc 180

[0017]

gacttcgtgt ggggcattgt cggccctagc cagtgggatg ctttcctcgt tcagatcgaa 240

caactgatac agcagcggat cgaagcctac gcccgcgccg ccgcgattag caacttagag 300

gggatcggca acaacttcaa catctatgtg gaggcgttcc aggagtggga ggaagaccca 360

aacaaccccg cgacccgcaa tcgcgtagtt gatcgtttcc gcatcctgga cggccttctg 420

gagcgagaca ttccttcgtt tcgcatcagc gggttcgaag tgccgctgct ttcagtctat 480

acccaggcgg cgaacctgca ccttgcaatc ctcagggact cggtgatctt cggcgagcgt 540

tggggcttga cgactacgaa cgtcaatgag aattacaata gacagatccg ccacatcaac 600

gagtacgccg accactgcgc gaacacctac aaccgtgggc ttaacaacct ccctaagagc 660

acgtaccaag actggataac atacaaccgg ctccgtcgtg acctcacact tacggtcttg 720

gacattgtga acttcttccc caactacgac aaccgccgtt acccgatcca gacggtcggc 780

cagctcactc gggaggtcta taccgacccg ctcatcaact tcaacccgca gttgcagagc 840

gtcgcccagc tccccacgtt taacgtgatg gaaagctcgg ccatacggaa cccgcaccta 900

ttcgacgtgc tgaacaacct cacgatcttc accgactggt tcagtgtcgg gcgcaacttc 960

tactggggcg gacaccgcgt tatctccagc cggatcggcg gaggcaacat taccagcccg 1020

atctacggcc gcgaggccaa ccaggagcca ccgcgctcgt tcactttcaa cggccccgtc 1080

tttcggaccc ttagcaatcc aacactgcgt tcactccaac aaccctggcc tgcgccacct 1140

ttcaacctca gaggcgttga gggtgttgag ttctcaaccc cgacaaattc gttcacctat 1200

cgcgggcgcg gaaccgtgga cagtctcacc gagctgccgc cacaggacaa tagtgttcct 1260

ccacgggagg gatattccca ccgcttgtgc catgcgacct tcgtgcagag gtccgggaca 1320

ccgttcctga ccaccggcgt cgtattctcg tggacccacc gctctgctga ccagaacatt 1380

atctacccga accgtatcac gcagatacca ctggttaagg cttccgacct cccctccggg 1440

acaaccgtcg tgcgcgggcc cggattcaca ggcggcgaca tccttcggcg cacctcgacg 1500

ggtggcttcg gcaccatcag ggtcaacgtg aacggcactc tgacgcagcg ctatcgcatc 1560

ggtttccgtt acgcctcgac ggtggacttc gatttctttg tcgtgcgtgg cggaaccacg 1620

gtcaacaatt tccgtttccc acgcaccatg aactctggtg aggaactgcg ttacggcagc 1680

[0018] ttcgaaactc ggtccttcac aaccccattc accttcaccc aaatccagga caccatccga 1740

accagtatcc agggcttgag cgggaacggc gaagtgtacc tggaccgcat cgagatcatc 1800

ccggtgactg cgaccttgaa agcggagtac gacctggagc gggcgcagaa ggtggtgggt 1860

gcgctgttca cctccacaaa tcagctcgca ttgaaaacga acatcactga ctaccatatc 1920

gaccaagtga gcaaccttgt ggactgcctg agcgacgagt tctgtcttga cgagaaacgt 1980

gagctgtccg aaaaggttaa gcacgccaag aggctgagcg acgagcgcaa tctcctgcaa 2040

gacccgaact ttcgagggat caaccggcag ccggatcgag gctggcgtgg ctccacggac 2100

attaccatcc agggaggcaa cgacgtcttc aaggagaact acgtcaccct gccggggacg 2160

ttcgacgagt gctatccaac gtacctctat cagaaaatcg acgagagcaa actcaagagt 2220

gacacacgct accagctcag agggtacatc gaggacagcc aagacttaga gatctatctt 2280

attcggtaca acgccaagca cgagaccgtg aacgttcccg gtacgtcctt cctctggccg 2340

ctgagtgtgg agagccccat cgggaagtgc ggagagccaa accgttgcgc tccccacatc 2400

gagtggaacc ctgacctgga atgctcgtgc cgggacgggg agaagtgcgc gcatcattct 2460

caccactttt cccttgatat tgacgtcggg tgcaccgacc tcaacgagga tctgggcgta 2520

tgggtgatat tcaagattaa gacccaggac ggccatgctc ggctcgggaa cctagagttc 2580

ctggaggaga agccgttact gggcgaggcc ctggccagag ccaagcgcgc ggagaagaag 2640

tggcgtgaca agcgggagac cttacagctt gagaccaaca tcgtgtacaa ggaggcaaag 2700

gaaagcgttg acgcactctt cgtgaactct cagtacgaca gactacaggc cgacaccgac 2760

atcgccatga ttcacgctgc cgacaagcgg gtgcacagaa tacgggaagc ctatctgccc 2820

gagctgagtg tgatccctgg ggtgaacgct ggcatcttcg aggagctgga gggccgaatc 2880

ttcaccgcgt actccctcta cgacgcaagg aacgtcatca agaacggcga tttcaacaac 2940

ggccttagct gctggaacgt gaagggtcac gtggatattg aggagcagaa caaccatcgc 3000

tcggtgctcg tggtgccgga gtgggaggcg gaggtgagcc aggaggtccg ggtgtgccca 3060

gggcgcgggt atatccttcg ggtgacggcc tacaaggagg gatacggcga ggggtgcgtg 3120

acgatccatg agattgaaga ccacaccgac gagttaaagt tcagtaattg cgtcgaggag 3180

[0019] gatgcgtacc ctgggaacac ggtggcttgc gacaactacc cggcaaacca ggaggagggg 3240

tgcacagagc tgcgccacag cgggaaccgg ggctacgacg agacttatgt caactctact 3300

agcagctcga cggactacac cgctgtgtac aaggaggagt cctacactgg cgagcagcgc 3360

tacgacagtt gcgagtcgaa ccggggttac gggaattaca cgccgctccc ggcgggctac 3420

gtcactaagg agctggagta tttcccagag accgataagg tctggataga gatcggcgag 3480

acagagggca cgttcattgt cgattcggtg gagctgctcc tgatggagga gtga 3534

<210> 7
<211> 3534
<212> DNA
<213> 人工(Artificial)

<220>
<223> 用于在植物细胞中表达的合成编码序列，
编码 TIC4029 的 TIC4029PL-3。

<400> 7
atgaacaaca acatcgagaa tcagtgcgtg ccctacaact gcctaagcaa ccctgaggag 60

gtgctcctag acggcgaacg catctccacc gggaactcta gcatcgacat ctctctgagc 120

ctagtgcaac tcctagtgtc gaacttcgtg cctggcggtg gcttcctcgt tggcttgatc 180

gacttcgtgt ggggtatcgt gggcccctcc cagtgggacg ccttcctcgt tcaaatcgag 240

cagctcatac aacaacgcat cgaggcttac gcccgcgccg ctgccatcag taacctagag 300

ggcatcggca acaacttcaa catctacgtg gaagccttcc aagagtggga ggaagaccca 360

aacaaccccg ccacccgcaa ccgcgtggtg gaccgctttc gaatcctcga cggcctcctg 420

gagcgcgaca tcccctcctt ccggatctcc ggtttcgagg tgcccctcct gtccgtttac 480

acccaggccg ctaacttgca cctcgccatc ctccgcgact ccgtgatctt cggggagcgc 540

tggggcctca ccactaccaa cgtgaacgag aactacaacc gtcagatccg ccacatcaac 600

gagtacgccg accactgcgc caacacctac aatcgcggcc tcaacaatct ccccaagtcc 660

acctaccaag actggatcac ctacaaccgc ttgagaaggg atctcactct caccgtgctc 720

gacatcgtga acttctttcc caactatgat aaccgtcgtt atcccattca aaccgtgggc 780

cagctcaccc gcgaagtcta taccgaccca ctcattaact tcaaccctca gctccaatcc 840

gttgcccagc tccccacctt caacgtgatg gagtcctcgg caatccgcaa ccctcactta 900

[0020]

ttcgacgtgc ttaacaattt gaccatcttc accgactggt tctccgtggg ccgcaacttc 960

tactggggcg gtcaccgcgt gatctctagc cgcatcggag gtggcaacat cacctcccca 1020

atctacggcc gcgaggctaa tcaggagcct ccgcgttcct ttaccttcaa cggccctgtg 1080

ttccgcaccc tttccaaccc tacccttaga tcactccaac aaccctggcc cgcaccaccg 1140

ttcaacttac gcggcgtgga gggcgttgag ttctccaccc ctaccaactc ctttacctat 1200

cgcggtagag gcaccgtgga ttccctcact gaactgcctc cgcaagacaa ttccgtgcct 1260

ccgagggagg gctactctca ccgtctgtgc catgccacct tcgtgcaaag gagcggcacc 1320

cctttcctga ccactggcgt cgtgttctcc tggacccacc gatcagccga ccaaaacatc 1380

atatacccaa accgcatcac ccagatccct ctggtcaagg ccagcgacct gcctagcggc 1440

accactgtcg ttcgaggccc tgggttcacc ggcggtgaca tcctgaggcg gaccagcacc 1500

ggcggtttcg gcaccataag ggtcaacgtc aacggcaccc tgactcaaag ataccgtatc 1560

ggctttcgtt acgccagcac agtcgatttc gacttcttcg tcgtaagggg cggaactact 1620

gtcaacaatt tccgcttccc gcgtactatg aactctggcg aggaactgag atacgggagt 1680

ttcgagacaa ggagcttcac tacgccgttt actttcactc aaattcaaga cactatccgg 1740

accagcatac aaggtttaag tggaaacgga gaagtctatc tggaccggat cgagatcatt 1800

cctgtcactg ctactctaaa agctgagtac gacctggagc gggctcagaa ggtcgttggt 1860

gctctgttca cttctactaa ccagttggct ctgaaaacaa acatcactga ttaccacatt 1920

gaccaagtat ccaatctggt tgactgcctg tctgatgagt tctgcctgga cgagaaacgg 1980

gagctgtctg agaaggtcaa gcacgctaag cggctgtctg acgaacggaa cctgcttcaa 2040

gacccgaact tccggggtat caatcgacag ccagatcggg gttggcgggg aagcactgac 2100

atcactatac aaggagggaa tgatgtcttc aaggaaaact acgtcacgct gccgggaacg 2160

tttgacgagt gctacccgac gtacctttat cagaaaatcg acgagtctaa gctgaaatct 2220

gacactcgtt accagttgcg tggttacatt gaggatagcc aggatcttga aatctacctt 2280

attcgttaca acgctaagca cgagacggtc aacgttcctg gaacgtcatt cttgtggccg 2340

ctttctgtcg agtcgccgat tgggaagtgc ggggagccca accgttgcgc tccacacatt 2400

[0021] gagtggaacc ctgaccttga gtgctcgtgc cgtgatgggg agaagtgcgc tcaccatagt 2460

caccactttt cacttgacat tgatgtcggg tgcacggatc ttaatgagga tcttggggtc 2520

tgggtcatat tcaagattaa gacgcaagat gggcacgcta gactgggaaa cttggagttc 2580

cttgaggaaa agccactgct aggggaagct cttgctagag ctaagcgggc ggagaagaag 2640

tggcgtgata agcgtgagac acttcaactt gagacgaata tcgtttacaa ggaggcgaag 2700

gagtcagtcg atgcgctctt cgtgaatagt caatacgacc gacttcaagc ggataccgat 2760

attgcgatga ttcacgcggc tgataagcgt gtccatcgta ttcgtgaagc gtaccttccg 2820

gagttgtcgg tcattcctgg ggttaatgcg gggatattcg aggaattgga ggggaggata 2880

ttcacggcgt attcgcttta cgatgcgaga aatgttatta agaatgggga tttcaataac 2940

gggttgtcgt gctggaatgt gaaagggcat gttgacattg aggagcaaaa caaccatcgt 3000

agtgttcttg tcgtacctga gtgggaggcc gaggtttcgc aagaggtgcg ggtttgccca 3060

gggcgaggct acatcttgag agttacggcg tacaaagagg gatacgggga gggatgtgtt 3120

acgattcatg agatagagga tcatactgat gagcttaagt tctcaaactg tgttgaggaa 3180

gatgcttacc caggaaatac cgttgcctgt gataactatc ctgcaaacca ggaggaagga 3240

tgcaccgagt tgcgacattc gggaaatcga ggatatgatg agacttatgt taattcaaca 3300

tcatcatcaa ctgattacac tgctgtgtac aaggaggaat catatacggg agaacaacgt 3360

tatgattcat gcgagtcgaa tcgaggatat ggaaactata caccactacc cgctggatat 3420

gtgacaaagg aactagagta tttcccagaa actgataagg tgtggataga gattggtgaa 3480

accgaaggaa cattcatcgt tgactctgtg gaactcttgc ttatggagga atga 3534

<210> 8
<211> 1812
<212> DNA
<213> 人工(Artificial)

<220>
<223> 用于在植物细胞中表达的合成编码序列，
编码 TIC4029_1 的 TIC4029_1PL。

<400> 8
atgaacaaca acattgagaa ccagtgcgtg ccctacaact gcctgagcaa ccctgaggag 60

[0022]

gtgctcctag acggtgagcg catctccacc ggtaactcta gcatcgacat ctccctgagc 120

ctagtgcagc tcctagtgtc taacttcgtg cccggcggtg gcttcctcgt gggcctcatc 180

gacttcgtgt ggggcatcgt gggcccctcc cagtgggacg ccttcctcgt gcagatcgag 240

cagctcatcc agcaacgcat cgaggcctac gcccgcgccg ctgccatcag caacctcgag 300

ggcatcggca acaatttcaa catctacgtg gaggccttcc aggagtggga ggaagacccc 360

aacaatcccg ccacccgcaa ccgcgtggtc gaccgcttcc gcatcctcga cggcctcctg 420

gagcgcgaca tcccctcctt ccgcatctcc ggcttcgagg tgcccctcct gtccgtgtac 480

acccaggccg ctaacctcca cctcgccatc ctccgcgact ccgtgatctt cggcgagcgc 540

tggggcctca ccactaccaa cgtgaacgag aactacaacc gccagatccg ccacatcaac 600

gagtacgccg accactgcgc caacacctac aaccgcggcc tcaacaatct ccccaagtcc 660

acctaccagg actggatcac ctacaaccgc ctccgcaggg acctcaccct caccgtgctc 720

gacatcgtga acttctttcc caactacgac aaccgcaggt accccatcca gaccgtgggc 780

cagctcaccc gcgaggtgta caccgacccc ctcatcaact tcaaccccca gctccagtcc 840

gtggcccagc tccccacctt caacgtgatg gagtccagcg ccatccgcaa cccccacctc 900

ttcgacgtgc tcaacaatct caccatcttc accgactggt tctccgtggg ccgcaacttc 960

tactggggcg gtcaccgcgt gatctccagc cgcatcggcg gtggcaacat cacctccccc 1020

atctacggcc gcgaggccaa ccaggagcct ccgcgctcct tcaccttcaa cggccctgtg 1080

ttccgcaccc tctccaaccc taccctcagg tccctccagc aaccttggcc tgcccctccg 1140

ttcaacctca ggggcgtgga gggcgtggag ttctccaccc ctaccaactc cttcacctac 1200

aggggcaggg gcaccgtgga ctccctcacc gagctgcctc cgcaggacaa ctccgtgcct 1260

ccgagggagg gctacagcca caggctgtgc cacgccacct tcgtgcagag gagcggcacc 1320

cctttcctga ccactggcgt cgttttcagc tggacccaca ggagcgccga ccagaacatc 1380

atttacccta acaggatcac ccagatccct ctggtcaagg ccagcgacct gcctagcggc 1440

accactgtcg ttaggggccc tggcttcacc ggcggtgaca tcctgaggcg gaccagcacc 1500

ggcggtttcg gcaccatcag ggtcaacgtc aacggcaccc tgactcagag gtacaggatc 1560

ggcttcaggt acgccagcac tgtcgacttc gacttctttg tcgttagggg cggtactact 1620

[0023]

gtcaacaatt tcaggttccc gaggactatg aacagcggcg aggaactgag gtacggtagc 1680

ttcgagacta ggtctttcac tacgccgttc actttcactc agatccagga cactatccgg 1740

actagtatcc agggtctgtc tggtaacggt gaggtctacc tggaccggat cgagatcatt 1800

ccggtcactt ga 1812

<210> 9
<211> 1746
<212> DNA
<213> 人工(Artificial)

<220>
<223> 用于在植物细胞中表达的合成编码序列，
编码 TIC4029_8 的 TIC4029_8PL，该序列包含在氨基末端处的截短和原毒素结构域的截短。

<400> 9
atgatctcca ccggtaactc tagcatcgac atctccctga gcctagtgca gctcctagtg 60

tctaacttcg tgcccggcgg tggcttcctc gtgggcctca tcgacttcgt gtggggcatc 120

gtgggccccc cccagtggga cgccttcctc gtgcagatcg agcagctcat ccagcaacgc 180

atcgaggcct acgcccgcgc cgctgccatc agcaacctcg agggcatcgg caacaatttc 240

aacatctacg tggaggcctt ccaggagtgg gaggaagacc ccaacaatcc cgccacccgc 300

aaccgcgtgg tcgaccgctt ccgcatcctc gacggcctcc tggagcgcga catcccctcc 360

ttccgcatct ccggcttcga ggtgcccctc ctgtccgtgt acacccaggc cgctaacctc 420

cacctcgcca tcctccgcga ctccgtgatc ttcggcgagc gctggggcct caccactacc 480

aacgtgaacg agaactacaa ccgccagatc cgccacatca acgagtacgc cgaccactgc 540

gccaacacct acaaccgcgg cctcaacaat ctccccaagt ccacctacca ggactggatc 600

acctacaacc gcctccgcag ggacctcacc ctcaccgtgc tcgacatcgt gaacttcttt 660

cccaactacg acaaccgcag gtaccccatc cagaccgtgg gccagctcac ccgcgaggtg 720

tacaccgacc ccctcatcaa cttcaacccc cagctccagt ccgtggccca gctccccacc 780

ttcaacgtga tggagtccag cgccatccgc aacccccacc tcttcgacgt gctcaacaat 840

ctcaccatct tcaccgactg gttctccgtg ggccgcaact tctactgggg cggtcaccgc 900

[0024] gtgatctcca gccgcatcgg cggtggcaac atcacctccc ccatctacgg ccgcgaggcc 960

aaccaggagc ctccgcgctc cttcaccttc aacggccctg tgttccgcac cctctccaac 1020

cctaccctca ggtccctcca gcaaccttgg cctgcccctc cgttcaacct caggggcgtg 1080

gagggcgtgg agttctccac ccctaccaac tccttcacct acaggggcag gggcaccgtg 1140

gactccctca ccgagctgcc tccgcaggac aactccgtgc ctccgaggga gggctacagc 1200

cacaggctgt gccacgccac cttcgtgcag aggagcggca ccccctttcct gaccactggc 1260

gtcgttttca gctggaccca caggagcgcc gaccagaaca tcatttaccc taacaggatc 1320

acccagatcc ctctggtcaa ggccagcgac ctgcctagcg gcaccactgt cgttaggggc 1380

cctggcttca ccggcggtga catcctgagg cggaccagca ccggcggttt cggcaccatc 1440

agggtcaacg tcaacggcac cctgactcag aggtacagga tcggcttcag gtacgccagc 1500

actgtcgact tcgacttctt tgtcgttagg ggcggtacta ctgtcaacaa tttcaggttc 1560

ccgaggacta tgaacagcgg cgaggaactg aggtacggta gcttcgagac taggtctttc 1620

actacgccgt tcactttcac tcagatccag gacactatcc ggactagtat ccagggtctg 1680

tctggtaacg gtgaggtcta cctggaccgg atcgagatca ttccggtcac tgctactctg 1740

aagtga 1746

<210> 10
<211> 581
<212> PRT
<213> 人工(Artificial)

<220>
<223> TIC4029_8 杀虫性蛋白质的氨基酸序列
该序列包含在氨基末端处的截短和
原毒素结构域的截短。

<400> 10

Met Ile Ser Thr Gly Asn Ser Ser Ile Asp Ile Ser Leu Ser Leu Val
1 5 10 15

Gln Leu Leu Val Ser Asn Phe Val Pro Gly Gly Gly Phe Leu Val Gly
20 25 30

[0025] Leu Ile Asp Phe Val Trp Gly Ile Val Gly Pro Ser Gln Trp Asp Ala
35 40 45

Phe Leu Val Gln Ile Glu Gln Leu Ile Gln Gln Arg Ile Glu Ala Tyr
50 55 60

Ala Arg Ala Ala Ala Ile Ser Asn Leu Glu Gly Ile Gly Asn Asn Phe
65 70 75 80

Asn Ile Tyr Val Glu Ala Phe Gln Glu Trp Glu Glu Asp Pro Asn Asn
85 90 95

Pro Ala Thr Arg Asn Arg Val Val Asp Arg Phe Arg Ile Leu Asp Gly
100 105 110

Leu Leu Glu Arg Asp Ile Pro Ser Phe Arg Ile Ser Gly Phe Glu Val
115 120 125

Pro Leu Leu Ser Val Tyr Thr Gln Ala Ala Asn Leu His Leu Ala Ile
130 135 140

Leu Arg Asp Ser Val Ile Phe Gly Glu Arg Trp Gly Leu Thr Thr Thr
145 150 155 160

Asn Val Asn Glu Asn Tyr Asn Arg Gln Ile Arg His Ile Asn Glu Tyr
165 170 175

Ala Asp His Cys Ala Asn Thr Tyr Asn Arg Gly Leu Asn Asn Leu Pro
180 185 190

Lys Ser Thr Tyr Gln Asp Trp Ile Thr Tyr Asn Arg Leu Arg Arg Asp
195 200 205

Leu Thr Leu Thr Val Leu Asp Ile Val Asn Phe Phe Pro Asn Tyr Asp
210 215 220

Asn Arg Arg Tyr Pro Ile Gln Thr Val Gly Gln Leu Thr Arg Glu Val
225 230 235 240

[0026] Tyr Thr Asp Pro Leu Ile Asn Phe Asn Pro Gln Leu Gln Ser Val Ala
245 250 255

Gln Leu Pro Thr Phe Asn Val Met Glu Ser Ser Ala Ile Arg Asn Pro
260 265 270

His Leu Phe Asp Val Leu Asn Asn Leu Thr Ile Phe Thr Asp Trp Phe
275 280 285

Ser Val Gly Arg Asn Phe Tyr Trp Gly Gly His Arg Val Ile Ser Ser
290 295 300

Arg Ile Gly Gly Gly Asn Ile Thr Ser Pro Ile Tyr Gly Arg Glu Ala
305 310 315 320

Asn Gln Glu Pro Pro Arg Ser Phe Thr Phe Asn Gly Pro Val Phe Arg
325 330 335

Thr Leu Ser Asn Pro Thr Leu Arg Ser Leu Gln Gln Pro Trp Pro Ala

340 345 350

Pro Pro Phe Asn Leu Arg Gly Val Glu Gly Val Glu Phe Ser Thr Pro
355 360 365

Thr Asn Ser Phe Thr Tyr Arg Gly Arg Gly Thr Val Asp Ser Leu Thr
370 375 380

Glu Leu Pro Pro Gln Asp Asn Ser Val Pro Pro Arg Glu Gly Tyr Ser
385 390 395 400

His Arg Leu Cys His Ala Thr Phe Val Gln Arg Ser Gly Thr Pro Phe
405 410 415

Leu Thr Thr Gly Val Val Phe Ser Trp Thr His Arg Ser Ala Asp Gln
420 425 430

Asn Ile Ile Tyr Pro Asn Arg Ile Thr Gln Ile Pro Leu Val Lys Ala
[0027] 435 440 445

Ser Asp Leu Pro Ser Gly Thr Thr Val Val Arg Gly Pro Gly Phe Thr
450 455 460

Gly Gly Asp Ile Leu Arg Arg Thr Ser Thr Gly Gly Phe Gly Thr Ile
465 470 475 480

Arg Val Asn Val Asn Gly Thr Leu Thr Gln Arg Tyr Arg Ile Gly Phe
485 490 495

Arg Tyr Ala Ser Thr Val Asp Phe Asp Phe Phe Val Val Arg Gly Gly
500 505 510

Thr Thr Val Asn Asn Phe Arg Phe Pro Arg Thr Met Asn Ser Gly Glu
515 520 525

Glu Leu Arg Tyr Gly Ser Phe Glu Thr Arg Ser Phe Thr Thr Pro Phe
530 535 540

Thr Phe Thr Gln Ile Gln Asp Thr Ile Arg Thr Ser Ile Gln Gly Leu
545 550 555 560

[0028]

Ser Gly Asn Gly Glu Val Tyr Leu Asp Arg Ile Glu Ile Ile Pro Val
565 570 575

Thr Ala Thr Leu Lys
580

```
TIC4029         MNNNIENQCVPYNCLSNPEEVLLDGERISTGNSSIDISLSLVQLLVSNFVPGGGFLVGLI
TIC4029_1       MNNNIENQCVPYNCLSNPEEVLLDGERISTGNSSIDISLSLVQLLVSNFVPGGGFLVGLI
TIC4029_8       -------------------------MISTGNSSIDISLSLVQLLVSNFVPGGGFLVGLI
                                          ********************************

TIC4029         DFVWGIVGPSQWDAFLVQIEQLIQQRIEAYARAAAISNLEGIGNNFNIYVEAFQEWEEDP
TIC4029_1       DFVWGIVGPSQWDAFLVQIEQLIQQRIEAYARAAAISNLEGIGNNFNIYVEAFQEWEEDP
TIC4029_8       DFVWGIVGPSQWDAFLVQIEQLIQQRIEAYARAAAISNLEGIGNNFNIYVEAFQEWEEDP
                ************************************************************

TIC4029         NNPATRNRVVDRFRILDGLLERDIPSFRISGFEVPLLSVYTQAANLHLAILRDSVIFGER
TIC4029_1       NNPATRNRVVDRFRILDGLLERDIPSFRISGFEVPLLSVYTQAANLHLAILRDSVIFGER
TIC4029_8       NNPATRNRVVDRFRILDGLLERDIPSFRISGFEVPLLSVYTQAANLHLAILRDSVIFGER
                ************************************************************

TIC4029         WGLTTTNVNENYNRQIRHINEYADHCANTYNRGLNNLPKSTYQDWITYNRLRRDLTLTVL
TIC4029_1       WGLTTTNVNENYNRQIRHINEYADHCANTYNRGLNNLPKSTYQDWITYNRLRRDLTLTVL
TIC4029_8       WGLTTTNVNENYNRQIRHINEYADHCANTYNRGLNNLPKSTYQDWITYNRLRRDLTLTVL
                ************************************************************

TIC4029         DIVNFFPNYDNRRYPIQTVGQLTREVYTDPLINFNPQLQSVAQLPTFNVMESSAIRNPHL
TIC4029_1       DIVNFFPNYDNRRYPIQTVGQLTREVYTDPLINFNPQLQSVAQLPTFNVMESSAIRNPHL
TIC4029_8       DIVNFFPNYDNRRYPIQTVGQLTREVYTDPLINFNPQLQSVAQLPTFNVMESSAIRNPHL
                ************************************************************

TIC4029         FDVLNNLTIFTDWFSVGRNFYWGGHRVISSRIGGGNITSPIYGREANQEPPRSFTFNGPV
TIC4029_1       FDVLNNLTIFTDWFSVGRNFYWGGHRVISSRIGGGNITSPIYGREANQEPPRSFTFNGPV
TIC4029_8       FDVLNNLTIFTDWFSVGRNFYWGGHRVISSRIGGGNITSPIYGREANQEPPRSFTFNGPV
                ************************************************************

TIC4029         FRTLSNPTLRSLQQPWPAPPFNLRGVEGVEFSTPTNSFTYRGRGTVDSLTELPPQDNSVP
TIC4029_1       FRTLSNPTLRSLQQPWPAPPFNLRGVEGVEFSTPTNSFTYRGRGTVDSLTELPPQDNSVP
TIC4029_8       FRTLSNPTLRSLQQPWPAPPFNLRGVEGVEFSTPTNSFTYRGRGTVDSLTELPPQDNSVP
                ************************************************************

TIC4029         PREGYSHRLCHATFVQRSGTPFLTTGVVFSWTHRSADQNIIYPNRITQIPLVKASDLPSG
TIC4029_1       PREGYSHRLCHATFVQRSGTPFLTTGVVFSWTHRSADQNIIYPNRITQIPLVKASDLPSG
TIC4029_8       PREGYSHRLCHATFVQRSGTPFLTTGVVFSWTHRSADQNIIYPNRITQIPLVKASDLPSG
                ************************************************************

TIC4029         TTVVRGPGFTGGDILRRTSTGGFGTIRVNVNGTLTQRYRIGFRYASTVDFDFFVVRGGTT
TIC4029_1       TTVVRGPGFTGGDILRRTSTGGFGTIRVNVNGTLTQRYRIGFRYASTVDFDFFVVRGGTT
TIC4029_8       TTVVRGPGFTGGDILRRTSTGGFGTIRVNVNGTLTQRYRIGFRYASTVDFDFFVVRGGTT
                ************************************************************

TIC4029         VNNFRFPRTMNSGEELRYGSFETRSFTTPFTFTQIQDTIRTSIQGLSGNGEVYLDRIEII
TIC4029_1       VNNFRFPRTMNSGEELRYGSFETRSFTTPFTFTQIQDTIRTSIQGLSGNGEVYLDRIEII
TIC4029_8       VNNFRFPRTMNSGEELRYGSFETRSFTTPFTFTQIQDTIRTSIQGLSGNGEVYLDRIEII
                ************************************************************

TIC4029         PVTATLKAEYDLERAQKVVGALFTSTNQLALKTNITDYHIDQVSNLVDCLSDEFCLDEKR
TIC4029_1       PVT-------------------------------------------------------
TIC4029_8       PVTATLK---------------------------------------------------
                ***
```

图1a

```
TIC4029         ELSEKVKHAKRLSDERNLLQDPNFRGINRQPDRGWRGSTDITIQGGNDVFKENYVTLPGT
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         FDECYPTYLYQKIDESKLKSDTRYQLRGYIEDSQDLEIYLIRYNAKHETVNVPGTSFLWP
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         LSVESPIGKCGEPNRCAPHIEWNPDLECSCRDGEKCAHHSHHFSLDIDVGCTDLNEDLGV
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         WVIFKIKTQDGHARLGNLEFLEEKPLLGEALARAKRAEKKWRDKRETLQLETNIVYKEAK
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         ESVDALFVNSQYDRLQADTDIAMIHAADKRVHRIREAYLPELSVIPGVNAGIFEELEGRI
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         FTAYSLYDARNVIKNGDFNNGLSCWNVKGHVDIEEQNNHRSVLVVPEWEAEVSQEVRVCP
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         GRGYILRVTAYKEGYGEGCVTIHEIEDHTDELKFSNCVEEDAYPGNTVACDNYPANQEEG
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         CTELRHSGNRGYDETYVNSTSSSTDYTAVYKEESYTGEQRYDSCESNRGYGNYTPLPAGY
TIC4029_1       ------------------------------------------------------------
TIC4029_8       ------------------------------------------------------------

TIC4029         VTKELEYFPETDKVWIEIGETEGTFIVDSVELLLMEE
TIC4029_1       -------------------------------------
TIC4029_8       -------------------------------------
```

图1b