



(12) 发明专利申请

(10) 申请公布号 CN 114058631 A

(43) 申请公布日 2022. 02. 18

(21) 申请号 202111327615.1

(22) 申请日 2016.08.25

(30) 优先权数据

62/210,737 2015.08.27 US

(62) 分案原申请数据

201680053628.5 2016.08.25

(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

T·A·奇彻 U·R·凯森帕利

J·L·卢特凯

(74) 专利代理机构 北京市金杜律师事务所

11256

代理人 孟凡宏 谢燕军

(51) Int. Cl.

C12N 15/31 (2006.01)

C07K 14/195 (2006.01)

C12N 15/82 (2006.01)

A01H 5/10 (2018.01)

A01H 5/00 (2018.01)

A01H 6/46 (2018.01)

A01H 6/54 (2018.01)

A01H 6/60 (2018.01)

A01N 37/46 (2006.01)

A01N 63/50 (2020.01)

A01P 7/04 (2006.01)

G01N 33/569 (2006.01)

权利要求书2页 说明书29页

序列表62页

(54) 发明名称

新型昆虫抑制蛋白

(57) 摘要

公开了对鳞翅类害虫的种表现出毒性活性的杀虫蛋白,并且包括但不限于TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL。提供了含有编码一种或多种公开的杀虫蛋白的重组核酸序列的DNA构建体。提供了抗鳞翅类侵袭的转基因植物、植物细胞、种子和植物部分,其含有编码本发明的所述杀虫蛋白的重组核酸序列。还提供了用于检测生物样品中本发明的重组核酸序列或蛋白质的存在的方法,以及使用TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL杀虫蛋白中的任一种来防治鳞翅类种的害虫的方法。

1. 一种重组核酸分子,其包含与编码杀虫蛋白的多核苷酸区段可操作地连接的异源启动子,其中所述杀虫蛋白包含与选自下组的序列具有至少90%氨基酸序列同一性的氨基酸序列:SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18。

2. 如权利要求1所述的重组核酸分子,其中:

- a. 所述重组核酸分子包含用于在植物中表达所述杀虫蛋白的序列;或
- b. 所述重组核酸分子在不可再生的植物细胞中表达以产生杀虫有效量的所述杀虫蛋白;或
- c. 所述重组核酸分子与载体可操作地连接,并且所述载体选自质粒、噬菌粒、杆粒、粘粒和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,其限定为存在于宿主细胞内,其中所述宿主细胞选自细菌细胞和植物细胞。

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以下的细菌属:土壤杆菌属、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,其中所述芽孢杆菌属的种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所述的重组核酸分子,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所述的重组核酸分子,其中所述植物宿主细胞选自苜蓿、香蕉、大麦、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、三叶草、棉花、葫芦、黄瓜、花旗松、茄子、桉树、亚麻、蒜、葡萄、啤酒花、韭菜、莴苣、粟、坚果、燕麦、橄榄、洋葱、棕榈、牧草、豌豆、花生、胡椒、木豆、松树、马铃薯、杨树、南瓜、辐射松、萝卜、油菜籽、稻、根茎、黑麦、红花、高粱、南方松、大豆、菠菜、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香、甘薯、柳枝稷、茶、烟草、番茄、小黑麦、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类昆虫的活性。

9. 如权利要求8所述的重组核酸分子,其中所述鳞翅类昆虫选自:绒毛豆夜蛾、甘蔗螟虫、小玉米茎蛀虫、玉米穗虫、烟芽夜蛾、大豆尺夜蛾、莎草粘虫、草地贪夜蛾、甜菜夜蛾、美国棉铃虫、东方叶虫、棉红铃虫、秋菜夜蛾、西南玉米螟、棉叶波纹夜蛾、菱纹背蛾、斑纹钻夜蛾、烟斜纹夜蛾、西方豆夜蛾和欧洲玉米螟。

10. 一种昆虫抑制组合物,其包含如权利要求1所述的重组核酸分子。

11. 如权利要求10所述的昆虫抑制组合物,其还包含编码至少一种不同于所述杀虫蛋白的其他杀虫剂的核苷酸序列。

12. 如权利要求11所述的昆虫抑制组合物,其中所述至少一种其他杀虫剂选自昆虫抑制蛋白、昆虫抑制dsRNA分子和辅助蛋白。

13. 如权利要求11所述的昆虫抑制组合物,其中所述至少一种其他杀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种的活性。

14. 如权利要求11所述的昆虫抑制组合物,其中所述至少一种其他杀虫蛋白选自Cry1A、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B、Cry1C、Cry1C、Cry1D、Cry1E、Cry1F、

Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry2A、Cry2Ab、Cry2Ae、Cry3、Cry3A、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、IP3、DIG-3、DIG-5、DIG-10、DIG-657和DIG-11蛋白。

15. 一种生产种子的方法,其包括:

- a. 种植包含权利要求1-9任一项所述的重组核酸分子的第一种子;
- b. 由所述第一种子生长出植物;以及
- c. 从所述植物收获种子,其中所述收获的种子包含所述重组核酸分子。

16. 一种用于防治鳞翅类害虫或虫害的方法,所述方法包括使所述害虫与杀虫有效量的杀虫蛋白接触,所述杀虫蛋白包含与选自下组的序列具有至少90%氨基酸序列同一性的氨基酸序:SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18。

17. 一种检测包含蛋白质的样品中杀虫蛋白的存在的方法,其中所述杀虫蛋白包含与选自下组的序列具有至少90%氨基酸序列同一性的氨基酸序:SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16和SEQ ID NO:18;所述方法包括:

- a. 使所述样品与免疫反应性抗体接触;以及
- b. 检测所述杀虫蛋白的存在。

18. 如权利要求17所述的方法,其中所述检测步骤包括ELISA或Western印迹。

新型昆虫抑制蛋白

[0001] 本申请是申请日为2016年8月25日、申请号为201680053628.5、发明名称为“新型昆虫抑制蛋白”的发明专利申请的分案申请。

[0002] 相关申请的引用

[0003] 本申请要求2015年8月27日提交的美国临时申请No.62/210,737的权益,其全部内容通过引用并入本文。

[0004] 序列表的引入

[0005] 包含序计算机可读形式的序列表的名为“38-21_61627US0001SEQLISTING_ST25.txt”的文件创建于2016年8月3日。该文件为94,508字节(在MS-Windows®中测量),通过电子提交同时提交(使用美国专利局EFS-Web提交系统(United States Patent Office EFS-Web filing system)),并且通过引用整体并入本文。

发明领域

[0006] 本发明一般涉及昆虫抑制蛋白的领域。公开了一类新型的对作物植物和种子的农业相关害虫表现出昆虫抑制活性的蛋白质。特别地,所公开的一类蛋白质对作物植物和种子的农业相关害虫,特别是对鳞翅类种的害虫具有杀虫活性。提供了含有编码一种或多种公开的毒素蛋白的重组多核苷酸构建体的植物、植物部分和种子。

[0007] 发明背景

[0008] 提高包括玉米、大豆、甘蔗、稻、小麦、蔬菜和棉花等在内的农业重要植物的作物产量已变得越来越重要。除了需要越来越多的农产品来为增长的人口提供食物、衣物和能量以外,预期气候相关影响和来自增长的人口使用农业实践以外的土地的压力也会减少农业可用耕地的量。这些因素导致对粮食安全的严峻预测,特别是在植物生物技术和农艺实践缺乏重大改进的情况下。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的改进是在有限的耕地面积上扩大作物生产的重要工具。

[0009] 昆虫,特别是鳞翅目和鞘翅目的昆虫被认为是大田作物受损的主要原因,从而降低了侵袭地区的作物产量。对农业有不利影响的鳞翅类害虫的种包括但不限于莎草粘虫(Black armyworm) (*Spodoptera exempta*)、秋菜夜蛾(Black cutworm) (*Agrotis ipsilon*)、玉米穗虫(Corn earworm) (*Helicoverpa zea*)、棉叶波纹夜蛾(Cotton leaf worm) (*Alabama argillacea*)、菱纹背蛾(Diamondback moth) (*Plutella xylostella*)、欧洲玉米螟(European corn borer) (*Ostrinia nubilalis*)、草地贪夜蛾(Fall armyworm) (*Spodoptera frugiperda*)、Cry1Fa1抗性草地贪夜蛾(Cry1Fa1 resistant Fall armyworm) (*Spodoptera frugiperda*)、旧世界棉铃虫(Old World bollworm) (OWB, *Helicoverpa armigera*)、南部夜蛾(Southern armyworm) (*Spodoptera eridania*)、大豆尺夜蛾(Soybean looper) (*Chrysodeixis includens*)、斑纹钻夜蛾(Spotted bollworm) (*Earias vittella*)、西南玉米螟(Southwestern Corn Borer) (*Diatraea grandiosella*)、烟芽夜蛾(Tobacco budworm) (*Heliothis virescens*)、烟斜纹夜蛾(Tobacco cutworm) (*Spodoptera litura*,也称为群毛虫(cluster caterpillar))、西方豆夜蛾(Western bean

cutworm) (*Striacosta albicosta*) 和绒毛豆夜蛾 (Velvetbean caterpillar) (*Anticarsia gemmatalis*)。

[0010] 历史上,依赖于合成化学杀虫剂的广泛应用来作为农业中的害虫防治剂。除了出现的抗性问题的外,对环境和人类健康的担忧激发了生物杀虫剂的研发。这项研究工作导致逐渐发现和使用各种昆虫病原性微生物物种,包括细菌。

[0011] 当昆虫病原性细菌,特别是属于芽孢杆菌属的细菌的潜力被发现并开发为生物害虫防治剂时,生物防治范例发生了转移。细菌苏云金芽孢杆菌 (*Bacillus thuringiensis*, Bt) 的菌株已被用作杀虫蛋白的来源,这是因为发现Bt菌株对特定昆虫具有高毒性。已知Bt菌株产生 δ -内毒素,其在孢子形成开始时和静止生长期期间位于伴孢结晶包涵体内(例如Cry蛋白质),并且也已知其产生分泌的杀虫蛋白。在被易感昆虫摄食后, δ -内毒素以及分泌毒素在中肠上皮表面发挥其作用,破坏细胞膜,导致细胞破裂和死亡。编码杀虫蛋白的基因也在除Bt以外的细菌物种中被鉴定出来,包括其他芽孢杆菌属和多种其他细菌物种,例如侧孢短芽孢杆菌 (*Brevibacillus laterosporus*)、球形赖氨酸芽孢杆菌 (*Lysinibacillus sphaericus*, “Ls”, 以前称为球形芽孢杆菌 (*Bacillus sphaericus*)) 和日本甲虫类芽孢杆菌 (*Paenibacillus popilliae*)。

[0012] 结晶和分泌的可溶性杀虫毒素对它们的宿主是高度特异性的并且作为化学杀虫剂的替代物已经获得全世界的认可。例如,杀虫毒素蛋白已被用于各种农业应用中,以保护农业上重要的植物免受昆虫侵袭,减少对化学杀虫剂应用的需求并提高产量。通过机械方法,例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上,以及通过使用遗传转化技术产生表达杀虫毒素蛋白的转基因植物和种子,杀虫毒素蛋白被用于防治作物植物的农业相关害虫。

[0013] 表达杀虫毒素蛋白的转基因植物的使用已经在全球范围内适用。例如,2012年,有2610万公顷种植了表达Bt毒素的转基因作物 (James, C., Global Status of Commercialized Biotech/GM Crops:2012. ISAAA Brief No.44)。转基因防虫作物的全球使用和这些作物中使用的有限数目的杀虫毒素蛋白为赋予对目前使用的杀虫蛋白的抗性的现有昆虫等位基因产生了选择压力。

[0014] 目标害虫对杀虫毒素蛋白的抗性的发展产生了对发现和开发新形式的杀虫毒素蛋白的持续需求,所述新型杀虫毒素蛋白可用于管理对表达杀虫毒素蛋白的转基因作物的昆虫抗性的增加。新的蛋白毒素具有改善的效力并且对更广泛的易感昆虫种类表现出防治,这将减少可发展抗性等位基因的存活昆虫的数量。另外,在一个植物中使用两种或更多种对相同害虫有毒性并显示不同作用方式的转基因杀虫毒素蛋白,降低了在任何单个目标昆虫物种中的抗性的可能性。

[0015] 因此,发明人在本文中公开了对以下物种表现出杀虫活性的来自日本甲虫类芽孢杆菌的新型蛋白毒素家族,以及类似毒素蛋白、变体蛋白和示例性重组蛋白:鳞翅类的种,特别是莎草粘虫 (*Spodoptera exempta*)、秋菜夜蛾 (*Agrotis ipsilon*)、玉米穗虫 (*Helicoverpa zea*)、棉叶波纹夜蛾 (*Alabama argillacea*)、菱纹背蛾 (*Plutella xylostella*)、欧洲玉米螟 (*Ostrinia nubilalis*)、草地贪夜蛾 (*Spodoptera frugiperda*)、Cry1Fa1抗性草地贪夜蛾 (*Spodoptera frugiperda*)、旧世界棉铃虫 (OWB, *Helicoverpa armigera*)、南部夜蛾 (*Spodoptera eridania*)、大豆尺夜蛾 (*Chrysodeixis includens*)、斑

纹钻夜蛾 (*Earias vittella*)、西南玉米螟 (*Diatraea grandiosella*)、烟芽夜蛾 (*Heliothis virescens*)、烟斜纹夜蛾 (*Spodoptera litura*, 也称为群毛虫)、西方豆夜蛾 (*Striacosta albicosta*) 和绒毛豆夜蛾 (*Anticarsia gemmatalis*)。

发明内容

[0016] 本文公开了一组新型的具有昆虫抑制活性的杀虫蛋白(毒素蛋白),在本文称为TIC6757、TIC7472和TIC7473,属于TIC6757蛋白毒素类别,其对作物植物的一种或多种害虫表现出抑制活性。TIC6757蛋白和TIC6757蛋白毒素类别中的蛋白质可以在制剂中和在植物中(in planta)单独使用或与其他杀虫蛋白和毒性剂组合使用,从而提供了目前农业系统中使用的杀虫蛋白和杀虫剂化学物的替代品。

[0017] 在一个实施方案中,本申请公开了重组核酸分子,其包含与编码杀虫蛋白或其片段的多核苷酸区段可操作地连接的异源启动子片段,其中:(a)所述杀虫蛋白包含SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18的氨基酸序列;或(b)所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18具有至少85%、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;或(c)所述多核苷酸区段与具有SEQ ID NO:3、SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15和SEQ ID NO:17的核苷酸序列的多核苷酸杂交;或(d)编码杀虫蛋白或其片段的所述多核苷酸区段包含与SEQ ID NO:3、SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15或SEQ ID NO:17的核苷酸序列具有至少65%、或70%、或75%、或80%、或85%、或90%、或95%、或98%、或99%、或约100%序列同一性的多核苷酸序列;或(e)所述重组核酸分子与载体可操作地连接,并且所述载体选自质粒、噬菌粒、杆粒、粘粒和细菌或酵母人工染色体。重组核酸分子可以包含用于在植物中表达杀虫蛋白的序列;或在植物细胞中表达以产生杀虫有效量的杀虫蛋白。

[0018] 本申请的另一个实施方案中是包含本申请的重组核酸分子的宿主细胞,其中所述宿主细胞选自细菌细胞和植物细胞。考虑的细菌宿主细胞包括土壤杆菌属(*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,所述芽孢杆菌属的种是蜡样芽孢杆菌(*Bacillus cereus*)或苏云金芽孢杆菌(*Bacillus thuringiensis*),所述短芽孢杆菌属是侧孢短芽孢杆菌,或者埃希氏菌属是大肠杆菌(*Escherichia coli*)。考虑的植物宿主细胞包括双子叶植物细胞和单子叶植物细胞。考虑的植物细胞还包括苜蓿(alfalfa)、香蕉(banana)、大麦(barley)、豆类(bean)、西兰花(broccoli)、卷心菜(cabbage)、芸苔(brassica)、胡萝卜(carrot)、木薯(cassava)、蓖麻(castor)、花椰菜(cauliflower)、芹菜(celery)、鹰嘴豆(chickpea)、白菜(Chinese cabbage)、柑橘(citrus)、椰子(coconut)、咖啡(coffee)、玉米(corn)、三叶草(clover)、棉花(cotton)(棉属的种)、葫芦(cucurbit)、黄瓜(cucumber)、花旗松(Douglas fir)、茄子(eggplant)、桉树(eucalyptus)、亚麻(flax)、蒜(garlic)、葡萄(grape)、啤酒花(hops)、韭

菜(leek)、莴苣(lettuce)、火炬松(Loblolly pine)、粟(millet)、瓜类(melons)、坚果(nut)、燕麦(oat)、橄榄(olive)、洋葱(onion)、观赏植物(ornamental)、棕榈(palm)、牧草(pasture grass)、豌豆(pea)、花生(peanut)、胡椒(pepper)、木豆(pigeonpea)、松树(pine)、马铃薯(potato)、杨树(poplar)、南瓜(pumpkin)、辐射松(Radiata pine)、萝卜(radish)、油菜籽(rapeseed)、稻(rice)、根茎(rootstocks)、黑麦(rye)、红花(safflower)、灌木(shrub)、高粱(sorghum)、南方松(Southern pine)、大豆(soybean)、菠菜(spinach)、倭瓜(squash)、草莓(strawberry)、甜菜(sugar beet)、甘蔗(sugarcane)、向日葵(sunflower)、甜玉米(sweet corn)、枫香(sweet gum)、甘薯(sweet potato)、柳枝稷(switchgrass)、茶(tea)、烟草(tobacco)、番茄(tomato)、小黑麦(triticale)、草坪草(turf grass)、西瓜(watermelon)和小麦(wheat)植物细胞。

[0019] 在另一个实施方案中,杀虫蛋白表现出针对鳞翅类昆虫的活性,所述鳞翅类昆虫包括绒毛豆夜蛾、甘蔗螟虫(Sugarcane borer)、小玉米茎蛀虫(Lesser cornstalk borer)、玉米穗虫、烟芽夜蛾、大豆尺夜蛾、莎草粘虫、南方夜蛾、草地贪夜蛾、甜菜夜蛾(Beet armyworm)、旧世界棉铃虫、东方叶虫(Oriental leafworm)、棉红铃虫(Pink bollworm)、秋菜夜蛾、西南玉米螟、棉叶波纹夜蛾、菱纹背蛾、斑纹钻夜蛾(Spotted bowl worm)、烟斜纹夜蛾、西方豆夜蛾和欧洲玉米螟。

[0020] 本申请还考虑了包含重组核酸分子的植物,所述重组核酸分子包含与编码杀虫蛋白或其片段的多核苷酸区段可操作地连接的异源启动子片段,其中:(a)所述杀虫蛋白包含SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18的氨基酸序列;或(b)所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18具有至少85%、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;或(c)所述多核苷酸区段在严格杂交条件下与SEQ ID NO:3、SEQ ID NO:15或SEQ ID NO:17的核苷酸序列的互补序列杂交;或(d)所述植物表现出可检测量的所述杀虫蛋白。在某些实施方案中,杀虫蛋白包含SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18。在一个实施方案中,所述植物是双子叶植物或单子叶植物。在另一个实施方案中,所述植物进一步选自苜蓿、香蕉、大麦、豆类、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、花旗松、茄子、桉树、亚麻、蒜、葡萄、啤酒花、韭菜、莴苣、火炬松、粟、瓜类、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、杨树、南瓜、辐射松、萝卜、油菜籽、稻、根茎、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、倭瓜、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香、甘薯、柳枝稷、茶、烟草、番茄、小黑麦、草坪草、西瓜和小麦。

[0021] 在另一些实施方案中,公开了包含重组核酸分子的种子。

[0022] 在另一个实施方案中,预期了包含本申请中公开的重组核酸分子的昆虫抑制组合物。所述昆虫抑制组合物可以进一步包含编码至少一种不同于所述杀虫蛋白的其他杀虫剂的核苷酸序列。在某些实施方案中,所述至少一种其他杀虫剂选自昆虫抑制蛋白、昆虫抑制性dsRNA分子和辅助蛋白。还预期了昆虫抑制组合物中的至少一种其他杀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种的活性。在一个实施方案中,昆虫抑制组合物中的至少一种其他杀虫剂选自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和DIG-11蛋白。

[0023] 还预期了包含可检测量的本申请中公开的重组核酸分子的商品。这样的商品包括由谷物处理者装袋的商品玉米、玉米片、玉米饼、玉米面、玉米粉、玉米糖浆、玉米油、玉米青贮、玉米淀粉、玉米谷物等；以及相应的大豆、稻、小麦、高粱、木豆、花生、水果、瓜类和蔬菜商品，适当时包括含有可检测量的本申请的这些多核苷酸和/或多肽的这些商品的汁、浓缩物、果酱、果冻、酱和其他可食用形式；完整或经处理的棉籽、棉籽油、棉绒；经处理用于饲料或食物的种子和植物部分；纤维、纸、生物质；以及燃料产品，例如来自棉籽油或来自轧棉机废物的丸粒的燃料；完整或经处理的大豆种子、大豆油、大豆蛋白、大豆粉、大豆面、大豆片，大豆麸皮、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、包含大豆的奶油、大豆生物质以及使用大豆植物和大豆植物部分生产的燃料产品。

[0024] 本申请还涵盖了生产包含本申请中公开的重组核酸分子的种子的方法。该方法包括种植至少一个包含本申请中公开的重组核酸分子的种子；由所述种子生长出植物；以及从所述植物收获种子，其中所述收获的种子包含本申请中的重组核酸分子。

[0025] 在另一个说明性实施方案中，提供了抗昆虫侵袭的植物，其中所述植物的细胞包含：(a) 重组核酸分子，其编码杀虫有效量的SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18所示的杀虫蛋白；或(b) 杀虫有效量的蛋白质，所述蛋白质包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18具有至少85%、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。

[0026] 在本申请中还公开了用于防治鳞翅类种害虫和防治植物，特别是作物植物的鳞翅类种虫害的方法。在一个实施方案中，该方法包括(a) 使所述害虫与杀虫有效量的SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18所示的杀虫蛋白接触；或(b) 使所述害虫与杀虫有效量的一种或多种杀虫蛋白接触，所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16或SEQ ID NO:18具有至少85%、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。

[0027] 本文还提供了检测包含编码杀虫蛋白或其片段的多核苷酸区段的重组核酸分子的存在的方法，其中：(a) 所述杀虫蛋白包含SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18的氨基酸序列；或(b) 所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID

NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18具有至少65%、或70%、或75%、或80%、或85%、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；或(c)所述多核苷酸区段与具有SEQ ID NO:3、SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15和SEQ ID NO:17的核苷酸序列的多核苷酸杂交。在本发明的一个实施方案中，该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包含编码本文提供的杀虫蛋白或其片段的多核苷酸区段的植物的基因组DNA杂交，并且在这样的杂交条件下不与来自不包含所述区段的其他同基因植物的基因组DNA杂交，其中所述探针与SEQ ID NO:3、SEQ ID NO:1、SEQ ID NO:5、SEQ ID NO:7、SEQ ID NO:9、SEQ ID NO:11、SEQ ID NO:13、SEQ ID NO:15或SEQ ID NO:17或编码杀虫蛋白的序列同源或互补，所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18具有至少65%、或70%、或75%、或80%、或85%、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。该方法可还包括(a)使所述样品和所述探针经受严格杂交条件；以及(b)检测所述探针与所述样品的DNA的杂交。

[0028] 本发明还提供了检测包含蛋白质的样品中杀虫蛋白或其片段的存在的方法，其中所述杀虫蛋白包含SEQ ID NO:2的氨基酸序列；或所述杀虫蛋白包含与SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:6、SEQ ID NO:8、SEQ ID NO:10、SEQ ID NO:12、SEQ ID NO:14、SEQ ID NO:16或SEQ ID NO:18具有至少65%、或70%、或75%、或80%、或85%、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。在一个实施方案中，该方法包括：(a)使所述样品与免疫反应性抗体接触；以及(b)检测蛋白质的存在。在一些实施方案中，检测步骤包括ELISA或Western印迹。

[0029] 序列简述

[0030] SEQ ID NO:1是获自日本甲虫类芽孢杆菌种DSC004343的编码TIC6757杀虫蛋白的核酸序列。

[0031] SEQ ID NO:2是TIC6757杀虫蛋白的氨基酸序列。

[0032] SEQ ID NO:3是设计用于在植物细胞中表达的编码TIC6757PL杀虫蛋白的合成编码序列，其中紧接起始甲硫氨酸密码子插入额外的丙氨酸密码子。

[0033] SEQ ID NO:4是由设计用于在植物细胞中表达的合成编码序列(SEQ ID NO:3)编码的TIC6757PL的氨基酸序列，并且其中紧接起始甲硫氨酸插入额外的丙氨酸。

[0034] SEQ ID NO:5是编码TIC6757_His杀虫蛋白的核酸序列，其中编码组氨酸标签的核酸序列可操作地连接至TIC6757编码序列的5'并且与其同框。

[0035] SEQ ID NO:6是TIC6757_His杀虫蛋白的氨基酸序列。

[0036] SEQ ID NO:7是获自日本甲虫类芽孢杆菌种DSC007648的编码TIC7472杀虫蛋白的核酸序列。

[0037] SEQ ID NO:8是TIC7242杀虫蛋白的氨基酸序列。

[0038] SEQ ID NO:9是编码TIC7472_His杀虫蛋白的核酸序列，其中编码组氨酸标签的核酸序列可操作地连接至TIC7472编码序列的3'并且与其同框。

[0039] SEQ ID NO:10是TIC7472_His杀虫蛋白的氨基酸序列。

[0040] SEQ ID NO:11是从核苷酸位置1-2391的开放阅读框和翻译终止密码子的编码TIC7473杀虫蛋白的核酸序列。

[0041] SEQ ID NO:12是获自日本甲虫类芽孢杆菌种DSC008493的TIC7243杀虫蛋白的氨基酸序列翻译。

[0042] SEQ ID NO:13是编码TIC7473_His杀虫蛋白的重组核酸序列,其中编码组氨酸标签的核酸序列可操作地连接至TIC7472编码序列的3'并且与其同框。

[0043] SEQ ID NO:14是TIC7473_His杀虫蛋白的氨基酸序列翻译。

[0044] SEQ ID NO:15是设计用于在植物细胞中表达的编码TIC7472PL杀虫蛋白的合成编码序列,其中紧接起始甲硫氨酸密码子插入额外的丙氨酸密码子。

[0045] SEQ ID NO:16是由设计用于在植物细胞中表达的合成编码序列(SEQ ID NO:15)编码的TIC7472PL的氨基酸序列,并且其中紧接起始甲硫氨酸插入额外的丙氨酸。

[0046] SEQ ID NO:17是设计用于在植物细胞中表达的编码TIC7473PL杀虫蛋白的合成编码序列,其中紧接起始甲硫氨酸密码子插入额外的丙氨酸密码子。

[0047] SEQ ID NO:18是由设计用于在植物细胞中表达的合成编码序列(SEQ ID NO:17)编码的TIC7473PL的氨基酸序列,并且其中紧接起始甲硫氨酸插入额外的丙氨酸。

[0048] 发明详述

[0049] 农业害虫防治领域中的问题可以表征为需要新的毒素蛋白,所述新的毒素蛋白对目标害虫有效,对目标害虫物种表现出广谱毒性,能够在植物中表达而不导致不希望的农业问题,并且与目前植物中商业使用的毒素相比提供替代作用方式。

[0050] 本文公开了由TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL示例的新型杀虫蛋白,解决了这些需求中的每一个,特别是针对宽范围的鳞翅类害虫,特别是针对莎草粘虫(*Spodoptera exempta*)、秋菜夜蛾(*Agrotis ipsilon*)、玉米穗虫(*Helicoverpa zea*)、棉叶波纹夜蛾(*Alabama argillacea*)、菱纹背蛾(*Plutella xylostella*)、欧洲玉米螟(*Ostrinia nubilalis*)、草地贪夜蛾(*Spodoptera frugiperda*)、Cry1Fa1抗性草地贪夜蛾(*Spodoptera frugiperda*)、旧世界棉铃虫(OWB,*Helicoverpa armigera*)、南部夜蛾(*Spodoptera eridania*)、大豆尺夜蛾(*Chrysodeixis includens*)、斑纹钻夜蛾(*Earias vittella*)、西南玉米螟(*Diatraea grandiosella*)、烟芽夜蛾(*Heliothis virescens*)、烟斜纹夜蛾(*Spodoptera litura*,也称为群毛虫)、西方豆夜蛾(*Striacosta albicosta*)和绒毛豆夜蛾(*Anticarsia gemmatalis*)。

[0051] 本申请提及的TIC6757、“TIC6757蛋白”、“TIC6757蛋白毒素”、“TIC6757毒素蛋白”、“TIC6757杀虫蛋白”、“TIC6757相关毒素”、“TIC6757相关毒素蛋白”、TIC6757PL、“TIC6757PL蛋白”、“TIC6757PL蛋白毒素”、“TIC6757PL毒素蛋白”、“TIC6757PL杀虫蛋白”、“TIC6757PL相关毒素”、“TIC6757PL相关毒素蛋白”、TIC7472、“TIC7472蛋白”、“TIC7472蛋白毒素”、“TIC7472毒素蛋白”、“TIC7472杀虫蛋白”、“TIC7472相关毒素”、“TIC7472相关毒素蛋白”、TIC7472PL、“TIC7472PL蛋白”、“TIC7472PL蛋白毒素”、“TIC7472PL毒素蛋白”、“TIC7472PL杀虫蛋白”、“TIC7472PL相关毒素”、“TIC7472PL相关毒素蛋白”、TIC7473、“TIC7473蛋白”、“TIC7473蛋白毒素”、“TIC7473毒素蛋白”、“TIC7473杀虫蛋白”、“TIC7473相关毒素”、“TIC7473相关毒素蛋白”、TIC7473PL、“TIC7473PL蛋白”、“TIC7473PL蛋白毒素”、“TIC7473PL毒素蛋白”、“TIC7473PL杀虫蛋白”、“TIC7473PL相关毒素”、“TIC7473PL相

关毒素蛋白”等是指包含以下、由以下组成、与以下基本上同源、类似于以下、或衍生自以下的任何新型杀虫蛋白或昆虫抑制蛋白：TIC6757 (SEQ ID NO:2)、TIC6757PL (SEQ ID NO:4)、TIC7472 (SEQ ID NO:8)、TIC7472PL (SEQ ID NO:16)、TIC7473 (SEQ ID NO:12) 或TIC7473PL (SEQ ID NO:18) 和赋予针对鳞翅类昆虫的活性的其杀虫或昆虫抑制片段,或者其组合,包括表现出杀虫或昆虫抑制活性的任何蛋白,如果此类蛋白与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL的比对导致从约85%到100%百分比的任何百分比分数的氨基酸序列同一性。TIC6757和TIC6757PL蛋白质包括蛋白质的质体靶向和非质体靶向形式。

[0052] 在本申请中使用的术语“区段”或“片段”来描述比描述TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的完整氨基酸或核酸序列短连续氨基酸或核酸序列。在以下情况下,表现出昆虫抑制活性的区段或片段也由本申请公开:如果这样的区段或片段与SEQ ID NO:2所示的TIC6757蛋白、SEQ ID NO:4所示的TIC6757PL蛋白、SEQ ID NO:8所示的TIC7472蛋白、SEQ ID NO:16所示的TIC7472PL蛋白、SEQ ID NO:12所示的TIC7473蛋白、或SEQ ID NO:18所示的TIC7473PL蛋白的相应部分的比对导致所述区段或片段与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的相应部分之间从约85%到约100%的任何百分比分数的氨基酸序列同一性。

[0053] 在本申请中,提及的术语“活力”或“活性”、“杀虫活性”或“杀虫”或“杀昆虫活性”、“昆虫抑制”或“杀虫”是指毒性剂如蛋白毒素抑制(抑制生长、摄食、生殖力或生存能力)、遏制(遏制生长、摄食、生殖力或生存能力)、防治(在包含有效量的TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的特定作物上防治虫害、控制害虫摄食活动)或杀死(引起发病、死亡或降低繁殖力)害虫的效力。这些术语旨在包括向害虫提供杀虫有效量的毒性蛋白的结果,其中害虫暴露于毒性蛋白导致发病、死亡、降低繁殖力或发育迟缓。这些术语还包括由于在植物中或植物上提供了杀虫有效量的毒性蛋白,所以从植物、植物组织、植物部分、种子、植物细胞或植物可能生长的特定地理位置驱除害虫。一般而言,杀虫活性是指毒性蛋白有效抑制特定目标昆虫(包括但不限于鳞翅目的昆虫)的生长、发育、存活力、摄食行为、交配行为、生殖力的能力,或者由摄食该蛋白质、蛋白质片段、蛋白质区段或多核苷酸的昆虫造成的不利影响的任何可测量降低。毒性蛋白可以由植物产生,或者可以施用于植物或施用于植物所在位置中的环境。术语“生物活性”、“有效”、“效力”或其变体也是本申请中可交换使用的术语,用于描述本发明的蛋白质对目标害虫的影响。

[0054] 当在目标害虫的饮食中提供时,杀虫有效量的毒性剂在毒性剂接触害虫时表现出杀虫活性。毒性剂可以是本领域已知的杀虫蛋白或一种或多种化学试剂。杀虫或杀昆虫化学试剂和杀虫或杀昆虫蛋白试剂可以单独使用或彼此组合使用。化学试剂包括但不限于靶向特定基因以遏制目标害虫的dsRNA分子、有机氯化物、有机磷酸酯、氨基甲酸酯、拟除虫菊酯、新烟碱类和ryanoids。杀虫或杀昆虫蛋白试剂包括本申请所述的蛋白质毒素,以及其他蛋白质毒性剂,包括靶向鳞翅类昆虫的毒性剂,以及用于防治其他植物害虫的蛋白质毒素,例如用于防治鞘翅类、半翅类和同翅类的种的本领域可获得的Cry和Cyt蛋白。

[0055] 意图是提及害虫,特别是作物植物的害虫时是指作物植物的被TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白毒素类别防治的害虫,特别是那些鳞翅目害虫。然而,在以下情况下提及害虫也可包括植物的鞘翅类、半翅类和同翅类害

虫,以及线虫和真菌:当靶向这些害虫的毒性剂与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白或者与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白具有85%至约100%同一性的蛋白质共定位或一起存在时。

[0056] TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白通过共同功能相关并且对来自鳞翅目昆虫种的害虫(包括成虫、蛹、幼虫和新生虫)表现出杀虫活性。

[0057] 鳞翅目的昆虫包括但不限于夜蛾科(Noctuidae)的粘虫(*armyworm*)、切根虫(*cutworm*)、尺夜蛾(*looper*)和*heliathine*,例如草地贪夜蛾(*Spodoptera frugiperda*)、甜菜夜蛾(*Spodoptera exigua*)、莎草粘虫(*Spodoptera exempta*)、南部夜蛾(*Spodoptera eridania*)、蓓带夜蛾(*bertha armyworm*) (*Mamestra configurata*)、秋菜夜蛾(*Agrotis ipsilon*)、粉纹夜蛾(*cabbage looper*) (*Trichoplusia ni*)、大豆尺夜蛾(*Pseudoplusia includens*)、绒毛豆夜蛾(*Anticarsia gemmatalis*)、绿夜蛾(*green cloverworm*) (*Hypena scabra*)、烟芽夜蛾(*Heliothis virescens*)、*granulate cutworm* (*Agrotis subterranea*)、粘虫(*Pseudaletia unipuncta*)、西方夜蛾(*western cutworm*) (*Agrotis orthogonia*);螟蛾科的蛀虫(*borer*)、潜蛾(*casebearer*)、结网虫(*webworm*)、*coneworm*、菜青虫(*cabbageworm*);和螟蛾科(*Pyrilidae*)的食叶虫(*skeletonizer*),例如欧洲玉米螟(*Ostrinia nubilalis*)、脐橙螟(*navel orangeworm*) (*Amyelois transitella*)、玉米根螟(*corn root webworm*) (*Crambus caliginosellus*)、草地螟(*sod webworm*) (*Herpetogramma licarsisalis*)、向日葵螟(*sunflower moth*) (*Homoeosoma electellum*)、小玉米茎蛀虫(*Elasmopalpus lignosellus*);卷蛾科(*Tortricidae*)的卷叶虫(*Leafroller*)、蛱虫(*budworm*)、种子蠕虫(*seed worms*)和果实蠕虫(*fruit worm*),例如苹果蠹蛾(*codling moth*) (*Cydia pomonella*)、葡萄卷夜蛾(*grape berry moth*) (*Endopiza viteana*)、东方果蛾(*oriental fruit moth*) (*Grapholita molesta*)、向日葵芽蛾(*sunflower bud moth*) (*Suleima helianthana*);以及许多其他在经济上重要的鳞翅目,例如菱纹背蛾(*Plutella xylostella*)、棉红铃虫(*Pectinophora gossypiella*)和舞毒蛾(*gypsy moth*) (*Lymantria dispar*)。鳞翅目的其他害虫包括例如棉叶波纹夜蛾(*Alabama argillacea*)、果树卷夜蛾(*fruit tree leaf roller*) (*Archips argyrospila*)、欧洲卷叶虫(*European leafroller*) (*Archips rosana*)和其他卷蛾属的种(二化螟(*Chilo suppressalis*)、亚洲稻螟(*Asiatic rice borer*)或稻螟(*rice stem borer*))、稻纵卷叶螟(*rice leaf roller*) (*Cnaphalocrocis medinalis*)、玉米根螟(*corn root webworm*) (*Crambus caliginosellus*)、兰草结网虫(*bluegrass webworm*) (*Crambus teterrellus*)、西南玉米螟(*Diatraea grandiosella*)、甘蔗螟虫(*Diatraea saccharalis*)、埃及钻夜蛾(*spiny bollworm*) (*Earias insulana*)、斑纹钻夜蛾(*Earias vittella*)、美国棉铃虫(*American bollworm*) (*Helicoverpa armigera*)、玉米穗虫(*Helicoverpa zea*,也称为大豆豆荚虫(*soybean podworm*)和棉铃虫)、烟芽夜蛾(*Heliothis virescens*)、草地螟(*Herpetogramma licarsisalis*)、西方豆夜蛾(*Striacosta albicosta*)、葡萄蛾(*European grape vine moth*) (*Lobesia botrana*)、柑橘潜叶蛾(*citrus leafminer*) (*Phyllocnistis citrella*)、大菜粉蝶(*large white butterfly*) (*Pieris brassicae*)、菜粉蝶(*small white butterfly*) (*Pieris rapae*,也称为进口菜青虫(*imported cabbageworm*))、甜菜夜蛾(*Spodoptera exigua*)、烟斜纹夜蛾(*Spodoptera litura*,也称为群毛虫)以及番茄斑潜蝇

(tomato leafminer) (*Tuta absoluta*)。

[0058] 本申请中提及的“分离的DNA分子”或等同的术语或短语旨在表示DNA分子是单独存在的或与其他组合物组合存在,但不存在于其天然环境中的分子。例如,对于生物体基因组的DNA中天然发现的核酸元件,例如编码序列、内含子序列、非翻译前导序列、启动子序列、转录终止序列等,只要元件在生物体的基因组中并且在基因组中其天然存在的位置处,就不认为其是“分离的”。然而,只要元件不在生物体的基因组中以及基因组中其天然存在的位置处,就认为这些元件中的每个以及这些元件的子元件在本公开内容的范围内是“分离的”。类似地,编码杀虫蛋白或该蛋白质的任何天然存在的杀虫变体的核苷酸序列可以是分离的核苷酸序列,只要该核苷酸序列不在天然发现编码该蛋白质的序列的细菌的DNA内。对于本公开内容的目的,编码天然存在的杀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。对于本公开内容的目的,任何转基因核苷酸序列,即插入到植物或细菌细胞的基因组中或存在于染色体外载体中的DNA的核苷酸序列将被认为是分离的核苷酸序列,无论其存在于用于转化细胞的质粒或类似结构中,植物或细菌的基因组中,还是以可检测量存在于来自植物或细菌的组织、子代、生物样品或商品中。

[0059] 如本申请中进一步描述的,编码TIC6757 (SEQ ID NO:19) 的开放阅读框 (ORF) 发现于获自日本甲虫类芽孢杆菌菌株DSC004343的DNA中。将编码序列在微生物宿主细胞中克隆并表达以产生用于生物测定的重组蛋白。使用高通量筛选和生物信息学技术筛选微生物序列的编码表现出与TIC6757的相似性的蛋白质的基因。编码TIC7472 (SEQ ID NO:7) 的开放阅读框 (ORF) 发现于获自日本甲虫类芽孢杆菌菌株DSC007648的DNA中。编码TIC7473 (SEQ ID NO:11) 的开放阅读框 (ORF) 发现于获自日本甲虫类芽孢杆菌菌株DSC008493的DNA中。使用来自微生物宿主细胞的蛋白质TIC6757的生物测定证明了针对以下鳞翅类的种的活性:甜菜夜蛾 (*Spodoptera exigua*)、秋菜夜蛾 (*Agrotis ipsilon*)、玉米穗虫 (*Helicoverpa zea*)、棉叶波纹夜蛾 (*Alabama argillacea*)、菱纹背蛾 (*Plutella xylostella*)、欧洲玉米螟 (*Ostrinia nubilalis*)、草地贪夜蛾 (*Spodoptera frugiperda*)、Cry1Fa1抗性草地贪夜蛾 (*Spodoptera frugiperda*)、旧世界棉铃虫 (OWB, *Helicoverpa armigera*)、南部夜蛾 (*Spodoptera eridania*)、大豆尺夜蛾 (*Chrysodeixis includens*)、斑纹钻夜蛾 (*Earias vittella*)、西南玉米螟 (*Diatraea grandiosella*)、烟芽夜蛾 (*Heliothis virescens*)、烟斜纹夜蛾 (*Spodoptera litura*, 也称为群毛虫) 和绒毛豆夜蛾 (*Anticarsia gemmatalis*)。使用来自微生物宿主细胞的蛋白质TIC7472和TIC7473证明了针对以下鳞翅类的种的活性:玉米穗虫 (*Helicoverpa zea*)、草地贪夜蛾 (*Spodoptera frugiperda*)、南部夜蛾 (*Spodoptera eridania*)、大豆尺夜蛾 (*Chrysodeixis includens*) 和西南玉米螟 (*Diatraea grandiosella*)。

[0060] 对于在植物细胞中的表达,TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白可被表达以存在于胞质溶胶中或靶向植物细胞的各种细胞器。例如,将蛋白质靶向叶绿体可导致转基因植物中表达的蛋白质的水平增加,同时防止脱表型 (off-phenotype) 的发生。靶向也可导致转基因事件中害虫抗性效力的增加。靶肽或转运肽是指指导将蛋白质运输至细胞中的特定区域的短 (3-70个氨基酸长) 肽链,所述区域包括细胞核、线粒体、内质网 (ER)、叶绿体、质外体、过氧化物酶体和质膜。在蛋白质被运输后,一些靶肽通过信号肽酶从蛋白质上裂解。为了靶向叶绿体,蛋白质含有为约40-50个氨基酸的转运

肽。关于使用叶绿体转运肽的描述,参见美国专利No.5,188,642和5,728,925。许多叶绿体定位蛋白从核基因表达为前体,并通过叶绿体转运肽(CTP)被靶向叶绿体。此类分离的叶绿体蛋白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、捕光复合蛋白I和蛋白II、硫氧还蛋白F、烯醇丙酮莽草酸磷酸合酶(EPSPS)结合的那些,以及美国专利No.7,193,133中描述的转运肽。已经在体内和体外证明了非叶绿体蛋白质可以通过使用与异源CTP的蛋白融合被靶向叶绿体,并且CTP足以将蛋白质靶向叶绿体。已经表明引入合适的叶绿体转运肽如拟南芥(*Arabidopsis thaliana*) EPSPS CTP (CTP2) (参见Klee等, *Mol. Gen. Genet.* 210:437-442, 1987) 或矮牵牛 (*Petunia hybrida*) EPSPS CTP (CTP4) (参见, della-Cioppa等, *Proc. Natl. Acad. Sci. USA* 83:6873-6877, 1986) 在转基因植物中将异源EPSPS蛋白质序列靶向叶绿体 (参见美国专利No.5,627,061; 5,633,435; 和5,312,910; 以及EP 0218571; EP 189707; EP 508909; 和EP 924299)。为了将TIC6757或TIC6757PL毒素蛋白靶向叶绿体,将编码叶绿体转运肽的序列置于与已设计用于在植物细胞中最佳表达的编码TIC6757或TIC6757PL毒素蛋白的合成编码序列的5'端可操作连接并且同框。

[0061] 预期通过使用TIC6757、TIC7472或TIC7473的氨基酸序列可以产生与TIC6757、TIC7472和TIC7473相关的另外的毒素蛋白序列,以产生具有新型特性的新型蛋白质。可以将TIC6757、TIC7472和TIC7473毒素蛋白进行比对,以将氨基酸序列水平的差异组合到新的氨基酸序列变体中,并对编码变体的重组核酸序列进行适当的改变。

[0062] 本发明进一步预期可通过使用多种本领域已知的基因编辑方法在植物中工程改造TIC6757蛋白毒素类别的改良变体。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 (转录活化因子样效应子核酸酶) 和CRISPR (成簇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用于将在植物细胞内转化的毒素蛋白质编码序列改变为不同的毒素编码序列。特别地,通过这些方法,改变毒素编码序列内的一个或多个密码子以设计新的蛋白质氨基酸序列。或者,将编码序列内的片段替换或缺失,或将额外的DNA片段插入编码序列中,以设计新的毒素编码序列。新的编码序列可以编码具有新特性的毒素蛋白,所述新特性例如针对害虫的增强的活性或谱,以及提供针对其中已对原始昆虫毒素蛋白产生抗性的害虫物种的活性。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的毒素蛋白的完整植物。

[0063] 还预期TIC6757、TIC7472和TIC7473的片段或其蛋白变体可以是截短的形式,其中一个或多个氨基酸从蛋白质的N端、C端、中间或其组合缺失,其中所述片段和变体保留了昆虫抑制活性。这些片段可以是TIC6757、TIC7472和TIC7473的天然存在的或合成的变体,或者衍生的蛋白质变体,但应保留至少TIC6757、TIC7472或TIC7473的昆虫抑制活性。

[0064] 类似于TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白的蛋白质可以使用本领域已知的各种基于计算机的算法来鉴定并相互比较 (参见表1和2)。本申请中报道的氨基酸序列同一性是使用以下这些默认参数的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,亲水间隙:开,亲水残基:GPSNDQERK,残基特异性空位罚分:开 (Thompson等 (1994) *Nucleic Acids Research*, 22:4673-4680)。进一步通过100%乘以 (氨基酸同一性/目标蛋白质的长度) 的乘积来计算氨基酸同一性百分比。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比对获得的那些相似

的结果,并且在本文中被考虑。

[0065] 意图是表现出对鳞翅类昆虫种的昆虫抑制活性的蛋白质在以下情况下与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL相关:如果在例如Clustal W比对中的查询中使用该蛋白质,并且本发明如SEQ ID NO:2、SEQ ID NO:4、SEQ ID NO:8、SEQ ID NO:16、SEQ ID NO:12或SEQ ID NO:18所示的蛋白质在这样的比对中鉴定为命中,其中查询蛋白沿着查询蛋白的长度表现出至少85%至约100%氨基酸同一性,即约85%、86%、87%、88%、89%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%、100%,或者该范围内的任何百分比分数。

[0066] 使用Clustal W算法将示例性蛋白TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL相互比对。产生了每个全长蛋白质的氨基酸序列同一性百分比的成对矩阵,如表1所报告的。

[0067] 表1. 示例性蛋白质TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL的成对矩阵显示。

[0068]

毒素	TIC6757 (SEQ ID NO:2)	TIC6757 PL (SEQ ID NO:4)	TIC7472 (SEQ ID NO:8)	TIC7472PL (SEQ ID NO:16)	TIC7473 (SEQ ID NO:12)	TIC7473PL (SEQ ID NO:18)
TIC6757 (SEQ ID NO:2)	-	99.9(796)	99.7(795)	99.6(794)	99.9(796)	99.7(795)
TIC6757 PL (SEQ ID NO:4)	99.7(796)	-	99.5(794)	99.7(796)	99.6(795)	99.9(797)
TIC7472 (SEQ ID NO:8)	99.7(795)	99.6(794)	-	99.9(796)	99.9(796)	99.7(795)
TIC7472PL (SEQ ID NO:16)	99.5(794)	99.7(796)	99.7(796)	-	99.6(795)	99.9(797)
TIC7473 (SEQ ID NO:12)	99.9(796)	99.7(795)	99.9(796)	99.7(795)	-	99.9(796)
TIC7473PL (SEQ ID NO:18)	99.6(795)	99.9(797)	99.6(795)	99.9(797)	99.7(796)	-

[0069] 表格说明:在成对矩阵中报告了(X)和(Y)之间的Clustal W比对。计算所有对之间的氨基酸同一性百分比,并由每个框中的第一个数字表示。每个框中的第二个数字(括号内)表示对之间相同氨基酸的数目。

[0070] 除了同一性百分比之外,TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473、TIC7473PL和相关蛋白还可以通过一级结构(保守氨基酸基序)、长度(约797个氨基酸)和其他特征来关联。表2中报道了TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白毒素的特征。

[0071] 表2. TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白的选择

的特征。

[0072]	蛋白质	分子量 (单位道尔顿)	氨基酸 酸长度	等电点	PH 7.0 时电荷	强碱性 (-) 氨基酸的 数目	强酸性 氨基酸的 数目	疏水性 氨基酸的 数目	极性 氨基酸的 数目
	TIC6757	90011.21	797	4.4289	-34.5	81	112	391	406
	TIC6757PL	90082.29	798	4.4289	-34.5	81	112	392	406
	TIC7472	90096.28	797	4.4141	-35.5	81	113	390	407
	TIC7472PL	90167.36	798	4.4141	-35.5	81	113	391	407
	TIC7473	90069.25	797	4.4141	-35.5	81	113	390	407
	TIC7473PL	90140.33	798	4.4141	-35.5	81	113	391	407

[0073] 如本申请的实施例中进一步描述的,设计编码TIC6757、TIC6757PL的变体的合成核酸分子序列以用于在植物中使用。设计用于在植物中使用的编码TIC6757PL蛋白的示例性重组核酸分子序列如SEQ ID NO:3所示。相对于TIC6757蛋白,TIC6757PL蛋白具有紧接起始甲硫氨酸之后的额外的丙氨酸氨基酸。认为插入到TIC6757氨基酸序列中的额外的丙氨酸残基提高了植物中蛋白质的表达。同样,编码TIC7472和TIC7473的变体的合成核酸分子序列在本文中分别称为TIC7472PL和TIC7473PL,并且被设计用于在植物中使用。设计用于在植物中使用的编码TIC7472PL和TIC7473PL的示例性合成核酸分子序列分别如SEQ ID NO:15和SEQ ID NO:17所示。相对于TIC7472和TIC7473蛋白,TIC7472PL和TIC7473PL蛋白具有紧接起始甲硫氨酸之后的额外的丙氨酸氨基酸。

[0074] 可以构建包含重组核酸分子序列的表达盒和载体并根据本领域已知的转化方法和技术将其导入玉米、大豆或棉花植物细胞中。例如,土壤杆菌介导的转化描述在以下文献中:美国专利申请公开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No.5,750,871(油菜)、7,026,528(小麦)和6,365,807(稻),以及Arenicibia等.(1998)Transgenic Res.7:213-222(甘蔗),所述文献全部内容通过引用整体并入本文。转化的细胞可以再生成表达TIC6757PL、TIC7472和TIC7473蛋白的转化植物,并且通过使用从转化植物获得的植物叶片在鳞翅类害虫幼虫的存在下进行生物测定来证明了杀虫活性。植物可以来自植物细胞(通过再生)、种子、花粉或分生组织转化技术。用于转化植物的方法是本领域已知的。

[0075] 作为传统转化方法的替代方法,可以通过定点整合将DNA序列(例如转基因、表达盒等)插入或整合到植物或植物细胞基因组内的特定位点或基因座中。因此,本公开内容的重组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个转基因、表达盒或用于插入植物或植物细胞基因组中的其他DNA序列。用于定点整合的此类供体模板可进一步包含在插入序列(例如,待插入植物基因组中的序列、转基因、盒等)的侧翼的一个或两个同源臂。本公开内容的重组DNA构建体可进一步包含编码位点特异性核酸酶和/或进行定点整合的任何任何相关蛋白的表达盒。这些核酸酶表达盒可以与供体模板存在于相同的分子或载体中(顺式)或存在于分开的分子或载体上(反式)。用于定点整合的多种方法是本领域已知的,涉及切割基因组DNA以在期望的基因组位点或基因座处产生双链断裂(DSB)或

缺口的不同蛋白质(或蛋白质和/或引导RNA的复合物)。简而言之,如本领域所理解的那样,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或切口的位点整合到基因组中。供体模板中的同源臂的存在可以促进在修复过程中通过同源重组采用插入序列并靶向到植物基因组中,尽管插入事件可以通过非同源端连接(NHEJ)发生。可以使用的位点特异性核酸酶的实例包括锌指核酸酶、工程或天然大范围核酸酶、TALE核酸内切酶和RNA引导的核酸内切酶(例如Cas9或Cpf1)。对于使用RNA引导的位点特异性核酸酶(例如,Cas9或Cpf1)的方法,重组DNA构建体还将包含编码一个或多个引导RNA的序列,以将核酸酶引导至植物基因组中的期望位点。

[0076] 考虑了编码TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL的重组核酸分子组合物。例如,TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL蛋白可以用重组DNA构建体表达,在所述重组DNA构建体中具有编码该蛋白质的ORF的多核苷酸分子可操作地连接至基因表达元件,例如启动子以及用于构建体所针对的系统中表达所需的任何其他调控元件。非限制性实例包括与用于在植物中表达蛋白质的TIC6757PL、TIC7472PL或TIC7473PL蛋白质编码序列可操作地连接的植物功能性启动子,或与用于在Bt细菌或其他芽孢杆菌属种中表达蛋白质的TIC6757、TIC7472或TIC7473蛋白质编码序列可操作地连接的Bt功能性启动子。其他元件可以与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质编码序列可操作地连接,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含但不限于与诸如SEQ ID NO:3、SEQ ID NO:1、SIQ ID NO:7、SEQ ID NO:11、SEQ ID NO:15和SEQ ID NO:17的多核苷酸可操作地连接的异源启动子,所述多核苷酸编码具有SEQ ID NO:4、SEQ ID NO:2、SEQ ID NO:8、SEQ ID NO:12、SEQ ID NO:16和SEQ ID NO:18所示的氨基酸序列的相应多肽或蛋白质。异源启动子还可以与编码质体靶向的TIC6757PL、TIC7472PL或TIC7473PL或非靶向的TIC6757PL、TIC7472PL或TIC7473PL的合成DNA编码序列可操作地连接。编码本文公开的蛋白质的重组核酸分子的密码子可以用同义密码子替换(本领域已知为沉默替换)。

[0077] 包含TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质编码序列的重组DNA构建体可进一步包含编码以下的DNA区域:一种或多种昆虫抑制剂,其可被配置与编码TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的DNA序列同时表达或共表达;不同于TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的蛋白质;昆虫抑制性dsRNA分子;或辅助蛋白。辅助蛋白包括但不限于辅助因子、酶、结合伴侣或其他试剂,所述其他试剂用于辅助昆虫抑制剂的效力,例如通过辅助其表达,影响其在植物中的稳定性,优化用于低聚反应的自由能,增强其毒性,以及提高其活性的范围来实现。辅助蛋白可以例如促进一种或多种昆虫抑制剂的摄取,或增强毒性剂的毒性作用。

[0078] 可以组装重组DNA构建体,使得所有蛋白质或dsRNA分子从一个启动子表达,或者每个蛋白质或dsRNA分子处于单独的启动子控制下,或其一些组合。本发明的蛋白质可以从多基因表达系统中表达,其中TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL中的一种或多种蛋白质从共同核苷酸区段表达,根据选择的表达系统的类型,所述区段还包含其他开放阅读框和启动子。例如,细菌多基因表达系统可以利用单个启动子

驱动单个操纵子内的多连接/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利用多个-未连接或连接的表达盒,每个表达盒表达不同的蛋白质或其他试剂,例如一种或多种dsRNA分子。

[0079] 可以通过载体例如质粒、杆状病毒、合成染色体、病毒粒子、粘粒、噬菌粒、噬菌体或病毒载体将包含TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质编码序列的重组多核苷酸或重组DNA构建体可以递送到宿主细胞中。此类载体可用于实现TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质编码序列在宿主细胞中的稳定或瞬时表达,或随后编码的多肽的表达。包含TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质编码序列并且被引入宿主细胞的外源重组多核苷酸或重组DNA构建体在本申请中被称为“转基因”。

[0080] 本文提供了转基因细菌、转基因植物细胞、转基因植物和转基因植物部分,其包含表达TIC6757或相关家族毒素蛋白编码序列中的任何一种或多种的重组多核苷酸。术语“细菌细胞”或“细菌”可以包括但不限于土壤杆菌属、芽孢杆菌属、埃希氏菌属、沙门氏菌属(*Salmonella*)、假单胞菌属、短芽孢杆菌属、克雷伯菌属、欧文氏菌属或根瘤菌属细胞。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”还可以包括但不限于苜蓿、香蕉、大麦、豆类、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、花旗松、茄子、桉树、亚麻、蒜、葡萄、啤酒花、韭菜、莴苣、火炬松、粟、瓜类、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、杨树、南瓜、辐射松、萝卜、油菜籽、稻、根茎、黑麦、红花、灌木、高粱、南方松、大豆、菠菜、倭瓜、草莓、甜菜、甘蔗、向日葵、甜玉米、枫香、甘薯、柳枝稷、茶、烟草、番茄、小黑麦、草坪草、西瓜和小麦植物细胞或植物。在某些实施方案中,提供了从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,转基因植物可以从转基因种子获得,通过切割、弯折、研磨或以其他方式使部分与植物分离获得。在某些实施方案中,植物部分可以是种子、莢、叶、花、茎、根或其任何部分,或转基因植物部分的不可再生部分。如本文所用,转基因植物部分的“不可再生”部分是不能诱导形成完整植物或不能诱导形成能够进行有性和/或无性繁殖的完整植物的部分。在某些实施方案中,植物部分的不可再生部分是转基因种子、莢、叶、花、茎或根的一部分。

[0081] 提供了制备包含鳞翅目昆虫抑制量的TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的转基因植物的方法。可以通过将编码本申请中提供的任何蛋白质的重组多核苷酸引入植物细胞并选择由表达鳞翅目昆虫抑制量的蛋白质的所述植物细胞衍生的植物来制备这样的植物。植物可以通过再生从植物细胞、从种子、花粉或分生组织转化技术获得。用于转化植物的方法是本领域已知的。

[0082] 本文还公开了经加工植物产品,其中经加工产品包含可检测量的TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白,其昆虫抑制区段或片段,或其任何区别部分。在某些实施方案中,加工的产品选自植物部分、植物生物质、油、粕、糖、动物饲料、粉、薄片、麸皮、绒、外壳、经加工种子和种子。在某些实施方案中,经加工产品是不可再生的。植物产品可以包括来自自转基因植物或转基因植物部分的商品或其他商业产品,其中所述商品或其他产品可以通过检测编码或包含TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的区别部分的核酸片段或表达的RNA或蛋白质通过商

业追踪。

[0083] 表达TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的植物可通过育种与表达其他毒素蛋白和/或表达其他转基因性状(如抗除草剂基因、赋予产量或抗逆性形状的基因等)的转基因事件交杂,或者这样的性状可在单一载体中组合以使得所有性状都联系起来。

[0084] 如实施例中进行进一步描述的,TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白编码序列以及与TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL具有较大的同一性百分比的序列可使用本领域普通技术人员已知的方法如聚合酶链式反应(PCR)、热扩增和杂交来鉴定。例如,蛋白质TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL可用于产生与相关蛋白质特异性结合的抗体,并且可用于筛选和发现其他密切相关的蛋白质成员。

[0085] 此外,编码TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473和TIC7473PL毒素蛋白的核苷酸序列可以用作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鉴定该类别的其他成员。例如,衍生自如SEQ ID NO:3、SEQ ID NO:15或SEQ ID NO:17所示序列的寡核苷酸可用于确定来自商品的脱氧核糖核酸样品中TIC6757PL、TIC7472PL或TIC7473PL转基因的存在或不存在。考虑到使用寡核苷酸的某些核酸检测方法的灵敏度,预期衍生自如SEQ ID NO:3、SEQ ID NO:15和SEQ ID NO:17所示序列的寡核苷酸可用于检测来自合并来源的商品中的TIC6757PL、TIC7472PL和TIC7473PL转基因,在所述合并来源中,仅一部分商品来自包含任何转基因的转基因植物。进一步认识到,这样的寡核苷酸可用于在SEQ ID NO:3、SEQ ID NO:15和SEQ ID NO:17中的每一个中引入核苷酸序列变异。这种“诱变”寡核苷酸可用于鉴定在转基因植物宿主细胞中显示一定范围的昆虫抑制活性或不同表达的TIC6757PL、TIC7472PL和TIC7473PL氨基酸序列变体。

[0086] 核苷酸序列同源物,例如由在严格杂交条件下与本申请中公开的每个或任何序列杂交的核苷酸序列编码的杀虫蛋白也是本发明的实施方案。本发明还提供了检测与第二核苷酸序列杂交的第一核苷酸序列的方法,其中第一核苷酸序列(或其反向互补序列)编码杀虫蛋白或其杀虫片段并与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由SEQ ID NO:3、SEQ ID NO:1、SIQ ID NO:7、SEQ ID NO:11、SEQ ID NO:15或SEQ ID NO:17所示的任何核苷酸序列。核苷酸编码序列在适当的杂交条件(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任一种其他蛋白质产生的抗血清发生交叉反应。如本文所定义,严格杂交条件包括:至少在42℃下杂交,然后在室温下洗涤两次,每次用2X SSC、0.1% SDS洗涤5分钟,然后在65℃下洗涤两次,每次用0.5X SSC、0.1% SDS洗涤30分钟。在甚至更高的温度下洗涤构成甚至更严格的条件,例如68℃的杂交条件,接着在68℃下在含有0.1% SDS的2xSSC中洗涤。

[0087] 本领域技术人员将认识到,由于遗传密码的冗余性,许多其他序列能够编码此类相关蛋白,并且在其用于在芽孢杆菌属菌株或植物细胞中表达杀虫蛋白的程度上的那些序列是本发明的实施方案,当然也认识到许多这样的冗余编码序列在这些条件下不与编码TIC6757、TIC7472和TIC7473的天然芽孢杆菌属或类芽孢杆菌属(Paenibacillus)序列杂交。本申请考虑了使用这些以及本领域普通技术人员已知的其他鉴定方法来鉴定TIC6757、TIC7472和TIC7473蛋白质编码序列以及与TIC6757、TIC7472和TIC7473蛋白质编码序列具

有较大的同一性百分比的序列。

[0088] 本公开还考虑使用本领域已知的分子方法来工程改造和克隆包含来自杀虫蛋白的蛋白质嵌合体的商业上有用的蛋白质；例如，嵌合体可以由TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的片段组装，以得到额外有用的实施方案，包括TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白的片段以及不同于TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL和相关蛋白的不同蛋白质的片段的组装体。TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白可彼此比对并且与其他芽孢杆菌属、类芽孢杆菌属其他杀虫蛋白（无论其在系统发生上是密切还是远缘相关）进行比对，并且可以鉴定可用于在比对的蛋白质之间进行替换的每种这样的蛋白的区段，导致嵌合蛋白质的构建。可对这样的嵌合蛋白质进行害虫生物测定分析，并且表征与嵌合体中每个这样的片段所来源的亲本蛋白质相比增强的生物活性或扩大的目标害虫谱的存在或不存在。可通过与其他蛋白质交换结构域或区段或通过使用本领域已知的定向进化方法对多肽的杀虫活性可进一步工程改造，已得到针对特定害虫或更广谱害虫的活性。

[0089] 本申请中还公开了用TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL蛋白质防治作物植物的昆虫（特别是鳞翅目）侵袭的方法。这样的方法可包括培育包含昆虫或鳞翅目抑制量的TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的植物。在某些实施方案中，这样的方法可以进一步包括以下任何一项或多项：(i) 将包含或编码TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的任何组合物施用于植物或产生植物的种子；以及(ii) 用编码TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的多核苷酸转化植物或产生植物的植物细胞。一般而言，预期可以在组合物中、在微生物中或在转基因植物中提供TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白，以赋予针对鳞翅类昆虫的昆虫抑制活性。

[0090] 在某些实施方案中，TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的重组核酸分子是昆虫抑制组合物的杀虫活性成分，所述组合物通过以下方式来制备：在适于表达TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的条件下培养被转化来表达TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的重组芽孢杆菌或任何其他重组细菌细胞。这样的组合物可以通过干燥、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此类重组细胞的培养物来制备。这样的过程可以得到芽孢杆菌或其他昆虫病原性细菌细胞提取物、细胞悬液、细胞匀浆、细胞裂解物、细胞上清液、细胞滤液或细胞沉淀。通过获得如此产生的重组多肽，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，并且可以配制用于各种用途，包括作为农业昆虫抑制喷雾产品或作为饮食生物测定中的昆虫抑制制剂。

[0091] 在一个实施方案中，为了降低出现抗性的可能性，包含TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL的昆虫抑制组合物还包含表现出对相同鳞翅类昆虫种的昆虫抑制活性，但不同于TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL毒素蛋白的至少一种额外的多肽。用于这种组合物的可能的额外多肽包括昆虫

抑制蛋白和昆虫抑制dsRNA分子。Baum等描述了使用这种核糖核苷酸序列来防治害虫的一个实例(美国专利公开2006/0021087A1)。用于防治害虫鳞翅类害虫的此类额外多肽可选自昆虫抑制蛋白,例如但不限于:Cry1A(美国专利No.5,880,275)、Cry1Ab、Cry1Ac、Cry1A.105、Cry1Ae、Cry1B(美国专利公开No.10/525,318)、Cry1C(美国专利No.6,033,874)、Cry1D、Cry1Da及其变体、Cry1E、Cry1F和Cry1A/F嵌合体(美国专利No.7,070,982;6,962,705;和6,713,063)、Cry1G、Cry1H、Cry1I、Cry1J、Cry1K、Cry1L、Cry1-型嵌合体,例如但不限于TIC836、TIC860、TIC867、TIC869和TIC1100(国际申请公开W02016/061391(A2))、TIC2160(国际申请公开W02016/061392(A2))、Cry2A、Cry2Ab(美国专利No.7,064,249)、Cry2Ae、Cry4B、Cry6、Cry7、Cry8、Cry9、Cry15、Cry43A、Cry43B、Cry51Aa1、ET66、TIC400、TIC800、TIC834、TIC1415、Vip3A、VIP3Ab、VIP3B、AXMI-001、AXMI-002、AXMI-030、AXMI-035、和AXMI-045(美国专利公开2013-0117884A1)、AXMI-52、AXMI-58、AXMI-88、AXMI-97、AXMI-102、AXMI-112、AXMI-117、AXMI-100(美国专利公开2013-0310543A1)、AXMI-115、AXMI-113、AXMI-005(美国专利公开2013-0104259A1)、AXMI-134(美国专利公开2013-0167264A1)、AXMI-150(美国专利公开2010-0160231A1)、AXMI-184(美国专利公开2010-0004176A1)、AXMI-196、AXMI-204、AXMI-207、AXMI-209(美国专利公开2011-0030096A1)、AXMI-218、AXMI-220(美国专利公开2014-0245491A1)、AXMI-221z、AXMI-222z、AXMI-223z、AXMI-224z、AXMI-225z(美国专利公开2014-0196175A1)、AXMI-238(美国专利公开2014-0033363A1)、AXMI-270(美国专利公开2014-0223598A1)、AXMI-345(美国专利公开2014-0373195A1)、AXMI-335(国际申请公开W02013/134523(A2))、DIG-3(美国专利公开2013-0219570A1)、DIG-5(美国专利公开2010-0317569A1)、DIG-11(美国专利公开2010-0319093A1)、AfIP-1A及其衍生物(美国专利公开2014-0033361A1)、AfIP-1B及其衍生物(美国专利公开2014-0033361A1)、PIP-1A/PIP-1B(美国专利公开2014-0007292A1)、PSEEN3174(美国专利公开2014-0007292A1)、AECFG-592740(美国专利公开2014-0007292A1)、Pput_1063(美国专利公开2014-0007292A1)、DIG-657(国际申请公开W02015/195594A2)、Pput_1064(美国专利公开2014-0007292A1)、GS-135及其衍生物(美国专利公开2012-0233726A1)、GS153及其衍生物(美国专利公开2012-0192310A1)、GS154及其衍生物(美国专利公开2012-0192310A1)、GS155及其衍生物(美国专利公开2012-0192310A1)、美国专利公开2012-0167259A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2012-0047606A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2011-0154536A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2011-0112013A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2010-0192256A1中所述的SEQ ID NO:2和4及其衍生物、美国专利公开2010-0077507A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2010-0077508A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2009-0313721A1中所述的SEQ ID NO:2及其衍生物、美国专利公开2010-0269221A1中所述的SEQ ID NO:2或4及其衍生物、美国专利No.7,772,465(B2)中所述的SEQ ID NO:2及其衍生物、W02014/008054A2中所述的CF161_0085及其衍生物、美国专利公开US2008-0172762A1、US2011-0055968A1和US2012-0117690A1中所述的鳞翅类毒素蛋白及其衍生物;US7510878(B2)中所述的SEQ ID NO:2及其衍生物、美国专利No.7812129(B1)中所述的SEQ ID NO:2及其衍生物;等。

[0092] 在另一些实施方案中,这样的组合物/制剂可以进一步包含对不被本发明的其他

昆虫抑制蛋白抑制的昆虫具有昆虫抑制活性的至少一种额外的多肽,以扩大所获得的昆虫抑制谱。例如,为了防治半翅类害虫,本发明的昆虫抑制蛋白的组合可以与半翅类活性蛋白例如TIC1415(美国专利公开2013-0097735A1)、TIC807(美国专利No.8609936)、TIC834(美国专利公开2013-0269060A1)、AXMI-036(美国专利公开2010-0137216A1)和AXMI-171(美国专利公开2013-0055469A1)一起使用。另外,用于防治鞘翅类害虫的多肽可以选自例如但不限于以下的昆虫抑制蛋白:Cry3Bb(美国专利No.6,501,009)、Cry1C变体、Cry3A变体、Cry3、Cry3B、Cry34/35、5307、AXMI134(美国专利公开2013-0167264A1)、AXMI-184(美国专利公开2010-0004176A1)、AXMI-205(美国专利公开2014-0298538A1)、AXMI-207(美国专利公开2013-0303440A1)、AXMI-218、AXMI-220(美国专利公开20140245491A1)、AXMI-221z、AXMI-223z(美国专利公开2014-0196175A1)、AXMI-279(美国专利公开2014-0223599A1)、AXMI-R1及其变体(美国专利公开2010-0197592A1)、TIC407、TIC417、TIC431、TIC807、TIC853、TIC901、TIC1201、TIC3131、DIG-10(美国专利公开2010-0319092A1)、eHIPs(美国专利申请公开No.2010/0017914)、IP3及其变体(美国专利公开2012-0210462A1),以及 **Hexatoxin-Hv1a**(美国专利申请公开US2014-0366227A1)。

[0093] 用于防治鞘翅类、鳞翅类和半翅类害虫的其他多肽可以在由Neil Crickmore维护的苏云金芽孢杆菌毒素命名网站(在万维网btomenclature.info)上找到。

[0094] 本领域已经记录了昆虫对某些杀虫剂产生抗性的可能性。一种昆虫抗性管理策略是采用表达通过不同的作用方式发挥作用的两种不同昆虫抑制剂的转基因作物。因此,对任一种昆虫抑制剂具有抗性的任何昆虫都可以通过另一种昆虫抑制剂来防治。另一种昆虫抗性管理策略使用没有对目标鳞翅类害虫物种提供保护的植物,以为这些未受保护的植物提供庇护。一个具体事例描述在美国专利No.6,551,962中,所述美国专利的全部内容通过引用并入本文。

[0095] 其他实施方案,例如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、在制剂上喷雾、滴落或擦拭的局部施用杀虫化学剂可直接施用于土壤(土壤浇灌),或施用于表达本文公开的蛋白质的生长中的植物,或配制为施用于包含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。用于种子处理的此类制剂可以与本领域已知的各种粘着剂和增粘剂一起施用。这样的制剂可包含在作用方式上与所公开的蛋白质协同作用的杀虫剂,使得制剂杀虫剂通过不同的作用方式起作用以防治可由所公开的蛋白质防治的相同或类似害虫,或这样的杀虫剂用于防治更宽宿主范围内的害虫或无法由TIC6757、TIC6757PL、TIC7472、TIC7472PL、TIC7473或TIC7473PL杀虫蛋白有效防治的植物害虫物种。

[0096] 前述组合物/制剂可以进一步包含农业上可接受的载体(例如饵剂)、粉剂、粉尘、丸剂、颗粒剂、喷雾剂、乳剂、胶体悬浮剂、水溶液、芽孢杆菌孢子/晶体制剂、种子处理剂、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,或转化以表达一种或多种蛋白质的细菌。取决于重组多肽中固有的昆虫抑制或杀虫抑制水平和待施用于植物或饮食测定的制剂水平,组合物/制剂可包含各种重量的重组多肽,例如0.0001重量%至0.001重量%至0.01重量%至1重量%至99重量%的重组多肽。

[0097] 鉴于前述内容,本领域技术人员应该理解,可以对所公开的特定方面进行改变,并且仍然获得相似或类似的结果,而不偏离本发明的精神和范围。因此,本文公开的特定结构

和功能细节不应解释为限制。应该理解的是,本文引用的每篇参考文献的全部公开内容并入本申请的公开内容内。

实施例

[0098] 实施例1

[0099] TIC6757的发现、克隆和表达

[0100] 对编码三种新型日本甲虫类芽孢杆菌杀虫蛋白的序列进行了鉴定、克隆、序列确证,并在昆虫生物测定中测试。分别从日本甲虫类芽孢杆菌菌株DSC004343、DSC007648和DSC008493分离的杀虫蛋白TIC6757、TIC7472和TIC7473代表了新的Vip3C样蛋白。与TIC6757、TIC7472和TIC7473远缘相关的序列是Vip3Ca2 (83.7%同一性,最接近的已知亲缘关系)、Vip3Aa1 (66.75%同一性) 和Vip3B样蛋白 (60.93%同一性)。TIC6757、TIC7472和TIC7473的与众不同的独特品质表明这些杀虫蛋白可能具有新的作用方式(MOA)。

[0101] 设计聚合酶链式反应(PCR)引物以从分别分离自日本甲虫类芽孢杆菌菌株DSC004343、DSC007648和DSC008493的总基因组DNA扩增TIC6757、TIC7472和TIC7473的编码区的全长拷贝。PCR扩增子也包含每个编码序列的翻译起始密码子和终止密码子。

[0102] 使用本领域已知的方法将每个扩增子克隆到与Bt可表达启动子可操作地连接的两个不同的Bt表达载体中。一个Bt表达载体包含在芽孢杆菌孢子形成期间开启的启动子。另一种表达载体包含非孢子形成启动子。另外,将每个扩增子克隆到用于大肠杆菌(E.coli)中进行蛋白质表达的载体中。为了分离大肠杆菌表达的蛋白质,将组氨酸标签与表达的编码序列可操作地连接以促进蛋白质的柱纯化。以下表3列出了用于细菌表达的编码序列及其各自的蛋白质序列。

[0103] 表3.用于在Bt和大肠杆菌中表达的毒素编码序列和相应的蛋白质序列。

	毒素	DNA 编码序列 SEQ ID NO:	蛋白质 SEQ ID NO:	细菌表达 宿主
[0104]	TIC6757	1	2	<i>Bt</i>
	TIC7472	7	8	<i>Bt</i>
	TIC7473	11	12	<i>Bt</i>
	TIC6757_His	5	6	大肠杆菌
	TIC7472_His	9	10	大肠杆菌
[0105]	TIC7473_His	13	14	大肠杆菌

[0106] 实施例2

[0107] TIC6757、TIC7472和TIC7473在昆虫生物测定中证实了对鳞翅类的活性

[0108] 使杀虫蛋白TIC6757、TIC7472和TIC7473在Bt和大肠杆菌中表达并测定其对鳞翅目、鞘翅目和半翅目各个种的毒性。针对下物种测定来自Bt的每种毒素的制备物:鳞翅类的种甜菜夜蛾(BAW, *Spodoptera exigua*)、秋菜夜蛾(BCW, *Agrotis ipsilon*)、玉米穗虫(CEW, *Helicoverpa zea*)、棉叶波纹夜蛾(CLW, *Alabama argillacea*)、菱纹背蛾(DBM, *Plutella xylostella*)、欧洲玉米螟(ECB, *Ostrinia nubilalis*)、草地贪夜蛾(FAW, *Spodoptera*

frugiperda)、Cry1Fa1抗性草地贪夜蛾(FAWR1, *Spodoptera frugiperda*)、美国棉铃虫(AWB, *Helicoverpa armigera*)、棉红铃虫(PBW, *Pectinophora gossypiella*)、南部夜蛾(SAW, *Spodoptera eridania*)、大豆尺夜蛾(SBL, *Chrysodeixis includens*)、斑纹钻夜蛾(SBW, *Earias vittella*)、西南玉米螟(SWCB, *Diatraea grandiosella*)、烟芽夜蛾(TBW, *Heliothis virescens*)、烟斜纹夜蛾(TCW, *Spodoptera litura*, 也称为群毛虫)和绒毛豆夜蛾(VBW, *Anticarsia gemmatilis*)；鞘翅类的种马铃薯甲虫(Colorado potato beetle (CPB, *Leptinotarsa decemlineata*))、西部玉米根虫(Western Corn Rootworm) (WCB, *Diabrotica virgifera virgifera*)；以及半翅类的种牧草盲蝽(Tarnished plant bug) (TPB, *Lygus lineolaris*)、西部牧草盲蝽(Western tarnished plant bug) (WTP, *Lygus hesperus*)、新热带褐蝽(Neotropical Brown Stink Bug) (NBSB, *Euschistus heros*)和绿蝽(Green Stink Bug) (GSB, *Nezara viridula*)。

[0109] 通过在大肠杆菌或Bt表达宿主中产生蛋白质来评估杀虫蛋白TIC6757、TIC7472和TIC7473的生物活性。在Bt宿主的情况下，使表达TIC6757、TIC7472或TIC7473的Bt菌株生长二十四(24)小时，然后将培养物添加到昆虫饮食中。通过将食用具有来自表达TIC6757、TIC7472或TIC7473的Bt菌株的培养物的饮食的昆虫的生长和发育与食用具有未处理对照培养物的饮食的昆虫进行比较来评估死亡率和发育迟缓。以类似的方式处理表达TIC6757、TIC7472或TIC7473的大肠杆菌菌株并且也提供在昆虫饮食中。对于来自Bt或大肠杆菌制备物或两种制备物的每种蛋白质观察的生物测定活性列于下表4和5中，其中“+”表示活性，“NT”表示未针对该特定害虫测定毒素。

[0110] 表4. TIC6757、TIC7472或TIC7473针对害虫的生物测定活性

毒素	BAW	BCW	CEW	CLW	DBM	ECB	FAW	FAWR1	AWB	PBW	SAW	SBL
TIC6757	+	+	+	+	+	+	+	+	+		+	+
TIC7472	NT	NT	+	NT	NT	NT	+	NT	NT	NT	+	+
TIC7473	NT	NT	+	NT	NT	NT	+	NT	NT	NT	+	+

[0112] 表5. TIC6757、TIC7472或TIC7473针对害虫的生物测定活性

毒素	SBW	SWCB	TBW	TCW	VBC	CPB	WCR	TPB	WTP	NBSB	SGB
TIC6757	+	+	+	+	+						
TIC7472	NT	+		NT	NT		NT			NT	NT
TIC7473	NT	+		NT	NT		NT			NT	NT

[0114] 从上表4和5中可以看出，昆虫毒素TIC6757显示出对许多鳞翅类害虫(BAW、BCW、CEW、CLW、DBM、ECB、FAW、FAWR1、AWB、SAW、SBL、SBW、SWCB、TBW、TCW和VBC)的活性。观察到了TIC7472和TIC7473对于大多数测定的害虫(CEW、FAW、SAW、SBL、SWCB)的活性。

[0115] 实施例3

[0116] 在稳定转化的玉米植物中测定TIC6757PL对鳞翅类害虫的活性

[0117] 使用本领域已知的方法克隆二元植物转化载体，所述二元植物转化载体包含设计为表达质体靶向的和非靶向的TIC6757PL杀虫蛋白二者的转基因盒。将所得载体用于稳定转化玉米植物。从转化体收获组织并用于针对多种鳞翅类害虫的昆虫生物测定。

[0118] 构建用于在植物中表达编码的蛋白质的合成编码序列，将其克隆到二元植物转化载体中并用于转化玉米植物细胞。根据美国专利5,500,365中一般描述的方法来合成合成

序列,以避免某些不合理的问题序列,例如富含ATTTA和A/T的植物聚腺苷酸化序列,同时保留天然类芽孢杆菌属蛋白的氨基酸序列。合成编码序列编码TIC6757PL蛋白质,其相对于TIC6757蛋白质在紧接起始甲硫氨酸之后包含额外的丙氨酸残基。对于质体靶向的蛋白质,将合成TIC6757PL杀虫蛋白质编码序列在框内与叶绿体靶向信号肽编码序列可操作地连接。所得的植物转化载体包含用于表达TIC6757PL杀虫蛋白的第一转基因盒,所述第一转基因盒包含组成型启动子,所述组成型启动子的5'与前导序列可操作地连接,前导序列的5'与内含子可操作地连接,内含子的5'与编码质体靶向的或未靶向的TIC6757PL蛋白的合成编码序列可操作地连接,合成编码序列的5'又与3'UTR可操作地连接;以及用于使用草甘膦选择来选择转化的植物细胞的第二转基因盒。TIC6757PL杀虫蛋白的合成编码序列如SEQ ID NO:3所示,并编码SEQ ID NO:4所示的蛋白质。

[0119] 使用土壤杆菌介导的转化方法,用如上所述的四种不同的二元转化载体转化玉米植物。二元植物转化载体构建体1和3包含编码质体靶向的TIC6757PL蛋白的编码序列,而构建体2和4包含编码非靶向的TIC6757PL蛋白的编码序列。通过本领域已知的方法诱导转化的细胞形成植物。类似于美国专利No.8,344,207中所述的那些,使用植物叶盘进行生物测定。每个叶盘样品上放置一只新孵化的不到一天龄的新生幼虫,并使其进食约4天。使用未转化的玉米植物来获得用作阴性对照的组织。针对秋菜夜蛾(BCW, *Agrotis ipsilon*)、玉米穗虫(CEW, *Helicoverpa zea*)、草地贪夜蛾(FAW, *Spodoptera frugiperda*)和西南玉米螟(SWCB, *Diatraea grandiosella*),评估来自每个二元载体的多重转化 R_0 单拷贝插入事件。

[0120] 如表6所示,表达TIC6757PL的转化的 R_0 植物对于所有四种害虫高度有效(定义为具有小于或等于百分之十七点五的叶损伤和百分之百的死亡率)。高外显率(由“(H)”指示)定义为每种构建体的超过百分之五十的测定事件具有小于或等于百分之十七点五的叶损伤和百分之百的死亡率。低外显率(由“(L)”指示)定义为每种构建体的小于或等于百分之五十的测定事件具有小于或等于百分之十七点五的叶损伤和百分之百的死亡率。

[0121] 表6.具有 $\leq 17.5\%$ 的叶损伤和百分之百的死亡率和外显率的表达TIC6757的事件的数目

		具有≤ 17.5%的叶损伤和 100%死亡率 (外显率)的事件的数目				
[0122]	构建体	事件总数	BCW	CEW	FAW	SWC
	构建体 1	22	17 (H)	18 (H)	18 (H)	11 (L)
	构建体 2	20	14 (H)	14 (H)	14 (H)	4 (L)
	构建体 3	19	17 (H)	17 (H)	17 (H)	17 (H)
	构建体 4	20	16 (H)	16 (H)	15 (H)	7 (L)

[0123] 使来自 R_0 构建体1(质体靶向的)和构建体2(质体未靶向的)的所选 R_0 事件自花授粉,产生 F_1 子代。选择来自每个 R_0 事件的多个杂合 F_1 子代植物用于叶盘生物测定,并针对秋菜夜蛾(BCW, *Agrotis ipsilon*)、玉米穗虫(CEW, *Helicoverpa zea*)、草地贪夜蛾(FAW, *Spodoptera frugiperda*)和西南玉米螟(SWCB, *Diatraea grandiosella*)进行测定。下表7示出了来自每个构建体/事件的每种植物的平均叶损伤百分比和平均死亡率。 F_1 子代植物相对于 R_0 事件来提及。例如,“事件-1_1”是来自事件-1的第一个杂合子 F_1 子代植物,“事件-

1_2”是来自事件-1的第一个杂合子 F_1 子代植物。“N”代表测定中使用的来自每个植物的样品数。从表7和表8可以看出,来自每个 R_0 事件的大多数植物表现出不大于百分之五的叶片损伤和对于BCW、CEW和FAW的百分之百的死亡率。关于SWCB,来自每个 R_0 事件的多个植物在测定中表现出小于10%的叶损害和大于50%的死亡率。

[0124] 表7.来自表达TIC6757PL的所选 R_0 事件的 F_1 子代的平均叶损伤百分比和死亡率。

构建体	事件_植物	N	BCW		CEW	
			平均叶损伤%	平均死亡率	平均叶损伤%	平均死亡率
构建体 1	事件-1_1	3	5.00	100.00	5.00	100.00
构建体 1	事件-1_2	3	5.00	100.00	5.00	100.00
构建体 1	事件-1_3	3	5.00	100.00	5.00	100.00
构建体 1	事件-1_4	3	5.00	100.00	6.65	100.00
构建体 1	事件-2_1	3	5.00	100.00	5.00	100.00
构建体 1	事件-2_2	3	NT	NT	7.50	100.00
构建体 1	事件-2_3	3	NT	NT	8.35	100.00
构建体 2	事件-3_1	3	5.00	100.00	5.00	100.00
构建体 2	事件-3_2	3	5.00	100.00	5.00	100.00
构建体 2	事件-4_1	3	5.00	100.00	5.00	100.00
构建体 2	事件-4_2	3	5.00	100.00	5.00	100.00
构建体 2	事件-4_3	3	6.65	66.67	5.00	100.00
构建体 2	事件-4_4	3	6.65	66.67	5.00	100.00
构建体 2	事件-4_5	3	20.00	33.33	10.00	100.00
构建体 2	事件-5_1	3	5.00	100.00	5.00	100.00
构建体 2	事件-5_2	3	5.00	100.00	5.00	100.00
构建体 2	事件-5_3	3	5.00	100.00	5.00	100.00
无	阴性对照	3	55.00	0.00	55.00	0.00

[0126] 表8.来自表达TIC6757PL的所选 R_0 事件的 F_1 子代的平均叶损伤百分比和死亡率。

构建体	事件_植物	N	FAW		SWCB	
			平均叶损伤%	平均死亡率	平均叶损伤%	平均死亡率
构建体 1	事件-1_1	3	5.00	100.00	6.65	66.67
构建体 1	事件-1_2	3	5.00	100.00	6.65	66.67
构建体 1	事件-1_3	3	5.00	100.00	7.50	50.00
构建体 1	事件-1_4	3	5.00	100.00	8.35	66.67
构建体 1	事件-2_1	3	5.00	100.00	5.00	50.00
构建体 1	事件-2_2	3	5.00	100.00	5.00	50.00
构建体 1	事件-2_3	3	5.00	100.00	6.65	66.67
构建体 2	事件-3_1	3	5.00	100.00	5.00	100.00
构建体 2	事件-3_2	3	5.00	100.00	15.00	50.00
构建体 2	事件-4_1	3	5.00	100.00	12.50	0.00

[0128]

构建体 2	事件-4_2	3	5.00	100.00	40.00	100.00
构建体 2	事件-4_3	3	5.00	100.00	48.35	0.00
构建体 2	事件-4_4	3	5.00	100.00	55.00	0.00
构建体 2	事件-4_5	3	5.00	100.00	55.00	0.00
构建体 2	事件-5_1	3	5.00	100.00	5.00	100.00
构建体 2	事件-5_2	3	5.00	100.00	6.65	66.67
构建体 2	事件-5_3	3	5.00	100.00	8.35	0.00
无	阴性对照	3	55.00	0.00	51.65	0.00

[0129] 使来自构建体3(质体靶向的)和构建体4(质体未靶向的)的所选 R_0 事件自花授粉,产生 F_1 子代。选择来自每个 R_0 事件的杂合 F_1 子代植物用于叶盘生物测定,并针对西部豆夜蛾(WBC, *Striacosta albicosta*)进行测定。表9示出了来自每个 R_0 事件和阴性对照的 F_1 子代植物的平均叶损伤百分比和平均百分比死亡率。“N”代表测定中使用的来自每个植物的样品数。

[0130] 表9. 来自表达TIC6757PL的所选 R_0 事件的 F_1 子代的平均叶损伤百分比和死亡率。

[0131]

构建体	事件	N	平均叶损伤%	平均死亡率
构建体 3	事件-6_1	4	5.00	100.00
构建体 3	事件-7_1	4	5.00	100.00
构建体 3	事件-8_1	4	5.00	100.00
构建体 3	事件-9_1	4	5.00	100.00
构建体 3	事件-10_1	4	5.00	100.00
构建体 3	事件-11_1	4	5.00	100.00
构建体 3	事件-12_1	4	5.00	100.00
构建体 3	事件-13_1	4	5.00	100.00
构建体 3	事件-14_1	4	5.00	100.00
构建体 3	事件-15_1	4	27.50	50.00
构建体 4	事件-16_1	4	5.00	100.00
构建体 4	事件-17_1	4	5.00	100.00
构建体 4	事件-18_1	4	5.00	100.00
阴性对照		4	45.00	0.00

[0132] 从上表9中可以看出,来自针对WBC进行测定的每个 R_0 事件的所有 F_1 子代植物(除了一个以外)表现出不超过百分之五的叶损害和百分之百的死亡率。

[0133] 测定来自用构建体3(质体靶向的)和构建体4(未靶向的)转化的所选杂合子 F_1 子代植物的幼苗对秋菜夜蛾(BCW, *Agrotis ipsilon*)的抗性。将 F_1 子代种子以及未转化的种子(阴性对照)种植在盆中。8天后,当幼苗从土壤中出苗时,每株植物用3只三龄BCW侵袭。在侵袭14天后,检查植物以对被BCW切咬的植物的数量进行计数。在测定中使用用构建体3转化的10个不同 R_0 事件得到的68个 F_1 子代植物和用构建体4转化的4个不同 R_0 事件得到的10个 F_1 子代植物。测定中还使用了15个阴性对照植物。观察植物后,观察到百分之八十的阴性对照被BCW切咬,而百分之零的用构建体3和构建体4转化的 F_1 子代植物表现出切咬。

[0134] 前文表明表达TIC6757PL的转化的玉米植物提供了对鳞翅类害虫的优异抗性,所

述鳞翅类害虫特别是秋菜夜蛾 (*Agrotis ipsilon*)、玉米穗虫 (*Helicoverpa zea*)、草地贪夜蛾 (*Spodoptera frugiperda*)、西南玉米螟 (*Diatraea grandiosella*) 和西部豆西部豆夜蛾 (*Striacosta albicosta*)。

[0135] 实施例4

[0136] 在稳定转化的大豆植物中测定TIC6757PL针对鳞翅类害虫的活性

[0137] 使用本领域已知的方法克隆包含设计为表达质体靶向的和非靶向的TIC6757PL杀虫蛋白二者的转基因盒的二元植物转化载体。将所得载体用于稳定转化大豆植物。从转化体收获组织并用于针对多种鳞翅类害虫的昆虫生物测定。

[0138] 将如上文实施例3中所述设计用于植物表达的合成编码序列克隆到二元植物转化载体中,并用于转化大豆植物细胞。使用本领域已知的方法构建包含质体靶向的和非靶向的TIC6757PL编码序列的二元载体。所得的植物转化载体包含用于表达TIC6757PL杀虫蛋白的第一转基因盒,所述第一转基因盒包含组成型启动子,组成型启动子的5'与前导序列可操作地连接,前导序列的5'与编码质体靶向的或未靶向的TIC6757PL蛋白的合成编码序列可操作地连接,合成编码序列的5'又与3'UTR可操作地连接;以及用于使用壮观霉素选择来选择转化的植物细胞的第二转基因盒。构建体1、3和5包含编码非靶向TIC6757PL杀虫蛋白的编码序列。构建体2、4和6包含编码质体靶向的TIC6757PL蛋白质的编码序列。

[0139] 通过本领域已知的方法诱导转化的大豆细胞形成植物。类似于美国专利No.8,344,207中所述的那些,使用植物叶盘进行生物测定。使用未转化的大豆植物来获得待用作阴性对照的组织。针对南部夜蛾 (SAW, *Spodoptera eridania*)、大豆尺夜蛾 (SBL, *Chrysodeixis includens*) 和大豆豆荚虫 (SPW, *Helicoverpa zea*) 评估来自每个二元载体的多重转化事件。

[0140] 如表10所示,表达TIC6757PL的转化的R₀大豆植物对SAW、SBL和SPW高度有效(定义为具有小于或等于百分之二十的叶损伤)。高外显率(由“(H)”指示)定义为每种构建体的超过百分之五十的测定事件具有小于或等于百分之二十的叶损伤。低外显率(由“(L)”指示)定义为每种构建体的小于或等于百分之五十的测定事件具有小于或等于百分之二十的叶损伤。

[0141] 表10. 具有≤20%的叶损伤和外显率的表达TIC6757PL的事件的数目

		具有≤20%的叶损伤 (外显率) 的事件的数目		
构建体	事件总数	SAW	SBL	SPW
构建体 1	15	14 (H)	14 (H)	12 (H)
构建体 2	15	5 (L)	3 (L)	8 (H)
构建体 3	15	12 (H)	13 (H)	13 (H)
构建体 4	15	15 (H)	15 (H)	15 (H)
构建体 5	15	14 (H)	13 (H)	14 (H)
构建体 6	15	15 (H)	15 (H)	15 (H)

[0143] 使来自构建体3、4、5和6的转化的表达TIC6757PL蛋白毒素的选择的R₀转基因大豆植物自花授粉,并产生R₁种子。使R₁种子萌发产生R₁植物。选择TIC6757PL表达盒纯合的R₁植物用于针对南部夜蛾 (SAW, *Spodoptera eridania*)、大豆尺夜蛾 (SBL, *Chrysodeixis*

includens)、大豆豆荚虫 (SPW, *Helicoverpa zea*) 和绒毛豆夜蛾 (VBW, *Anticarsia gemmatalis*) 的叶盘生物测定。表11和12示出了每个R₁子代植物和阴性对照品种A3555的通过每种昆虫显示的平均叶损伤百分比。表11和12还示出了相对于阴性对照,每个测定的事件的通过每种昆虫显示的标准误差平均值 (SEM) 叶损伤百分比。“N”代表测定中使用的来自每个植物的样品数。“SEM”表示平均损伤百分比的标准误差。

[0144] 表11. 表达TIC6757PL的R₁大豆植物的平均叶损伤百分比

构建体	事件数	植物/事件数	SAW			SBL		
			N	平均损伤%	SEM	N	平均损伤%	SEM
构建体 3	5	6	4	0.37	0.30	4	1.91	0.72
构建体 4	8	6	4	0.31	0.25	4	1.25	0.34
构建体 5	8	6	4	0.02	0.02	4	0.75	0.35
构建体 6	8	6	4	0.76	0.34	4	0.97	0.35
阴性对照	品种 A3555	8	4	87.93	9.74	4	79.44	12.44

[0146] 表12. 表达TIC6757PL的R₁大豆植物的平均叶损伤百分比

构建体	事件数	植物/事件数	SPW			VBC		
			N	平均损伤%	SEM	N	平均损伤%	SEM
构建体 3	5	6	4	16.32	3.83	4	1.89	0.60
构建体 4	8	6	4	2.25	0.30	4	0.96	0.31
构建体 5	8	6	4	2.40	0.50	4	0.51	0.25
构建体 6	8	6	4	3.65	0.53	4	0.71	0.32
阴性对照	品种 A3555	8	4	97.25	1.09	4	88.88	10.30

[0148] 如表11和12中可见,表达TIC6757PL毒素蛋白的R₁大豆植物对SAW、SBL、SPW和VBC提供了优异的抗性。对于SAW,所有四个事件表现出小于百分之一(1%)的叶损害,而阴性对照则具有约百分之八十八(88%)的叶损害。对于SBL,所有四(4)个事件表现出小于百分之二(2)的叶损害,而对照具有约百分之八十(80)的叶损害。对于SPW,四个事件中的三个表现出小于百分之四(4)的叶损害,而对照具有约百分之九十七(97)的叶损害。对于VBC,三个事件表现出小于百分之一(1)的叶损伤,一个事件表现出小于百分之二(2)的叶损伤,而阴性对照具有接近百分之八十九(89)的叶损害。

[0149] 前文表明表达TIC6757PL的转化的大豆植物提供了对鳞翅类害虫的优异抗性,特别是对南部夜蛾 (*Spodoptera eridania*)、大豆尺夜蛾 (*Chrysodeixis includens*)、大豆豆荚虫 (*Helicoverpa zea*) 和绒毛豆夜蛾 (*Anticarsia gemmatalis*)。

[0150] 实施例5

[0151] 在稳定转化的棉花植物中测定针对鳞翅类害虫的TIC6757PL活性

[0152] 使用本领域已知的方法克隆包含设计为表达质体靶向的和非靶向的TIC6757PL杀虫蛋白二者的转基因盒的二元植物转化载体。将所得载体用于稳定转化棉花植物。从转化

体收获组织并用于针对多种鳞翅类害虫的昆虫生物测定。

[0153] 将如上文实施例3中所述设计用于植物表达的合成编码序列克隆到二元植物转化载体中,并用于转化棉花植物细胞。使用本领域已知的方法构建包含质体靶向的和非靶向的TIC6757PL编码序列的二元载体。所得的植物转化载体包含用于表达TIC6757PL杀虫蛋白的第一转基因盒,所述第一转基因盒包含组成型启动子,组成型启动子的5'与前导序列可操作地连接,前导序列的5'与编码质体靶向的或未靶向的TIC6757PL蛋白的合成编码序列可操作地连接,合成编码序列的5'又与3'UTR可操作地连接;以及用于使用壮观霉素选择来选择转化的植物细胞的第二转基因盒。

[0154] 通过本领域已知的方法诱导转化的棉花细胞形成植物。类似于美国专利No.8,344,207中所述的那些,使用植物叶盘进行生物测定。使用未转化的棉花植物来获得待用作阴性对照的组织。针对棉铃虫(Southern armyworm Cotton bollworm)(CBW,*Helicoverpa zea*)、草地贪夜蛾(FAW,*Spodoptera frugiperda*)、大豆尺夜蛾(SBL,*Chrysodeixis includens*)和烟芽夜蛾(TBW,*Heliothis virescens*)评估来自每个二元载体的多重转化事件。

[0155] 如表13所示,表达TIC6757PL的转化的R₀棉花植物对CBW、FAW、SBL和TBW高度有效(定义为具有小于或等于百分之十的叶损伤)。高外显率(由“(H)”指示)定义为每种构建体的超过百分之五十的测定事件具有小于或等于百分之十的叶损伤。低外显率(由“(L)”指示)定义为每种构建体的小于或等于百分之五十的测定事件具有小于或等于百分之十的叶损伤。

[0156] 表13具有≤10%的叶损伤和外显率的表达TIC6757PL的事件的数目

		具有≤10%的叶损伤的事件数/测定的事件数 (外显率)			
构建体		CBW	FAW	SBL	TBW
[0157]	构建体 1	22/25 (H)	21/24 (H)	21/25 (H)	21/25 (H)
	构建体 2	12/15 (H)	6/15 (L)	13/15 (H)	13/15 (H)
	构建体 3	7/13 (H)	8/14 (H)	4/13 (L)	6/14 (L)
	构建体 4	11/14 (H)	8/14 (H)	9/14 (H)	10/14 (H)
	构建体 5	20/25 (H)	19/23 (H)	20/24 (H)	19/23 (H)
	构建体 6	6/7 (H)	7/7 (H)	7/7 (H)	6/7 (H)
	构建体 7	22/25 (H)	22/25 (H)	22/25 (H)	22/25 (H)

[0158] 实施例6

[0159] 在稳定转化的玉米植物中测定针对鳞翅类害虫的TIC7472PL和TIC7473PL活性

[0160] 使用本领域已知的方法克隆包含设计为表达质体靶向的和非靶向的TIC7472PL或TIC7473PL杀虫蛋白二者的转基因盒的二元植物转化载体。将所得载体用于稳定转化玉米植物。从转化体收获组织并用于针对多种鳞翅类害虫的昆虫生物测定。

[0161] 构建合成编码序列用于在植物中表达编码的蛋白质,将其克隆到二元植物转化载体中并用于转化玉米植物细胞。按照美国专利5,500,365中一般描述的方法来合成合成序列,以避免某些不合理的问题序列,例如富含ATTGA和A/T的植物聚腺苷酸化序列,同时保留

天然类芽孢杆菌属蛋白的氨基酸序列。合成编码序列编码TIC7472PL和TIC7473PL蛋白质，其相对于TIC7472和TIC7473蛋白质在紧接甲硫氨酸之后包含额外的丙氨酸残基。对于质体靶向的蛋白质，合成TIC7472PL或TIC7473PL杀虫蛋白质编码序列在框内与叶绿体靶向信号肽编码序列可操作地连接。所得的植物转化载体包含用于表达TIC7472PL或TIC7473PL杀虫蛋白的第一转基因盒，所述第一转基因盒包含组成型启动子，组成型启动子的5'与前导序列可操作地连接，前导序列的5'与内含子可操作地连接，内含子的5'与编码质体靶向的或未靶向的TIC7472PL或TIC7473PL蛋白的合成编码序列可操作地连接，合成编码序列的5'又与3'UTR可操作地连接；以及用于使用草甘膦选择来选择转化的植物细胞的第二转基因盒。TIC7472PL杀虫蛋白的合成编码序列如SEQ ID NO:15所示，并编码SEQ ID NO:16所示的蛋白质。TIC7473PL杀虫蛋白的合成编码序列如SEQ ID NO:17所示，并编码SEQ ID NO:18所示的蛋白质。

[0162] 使用土壤杆菌介导的转化方法，用上述二元转化载体转化玉米植物。通过本领域已知的方法诱导转化的细胞形成植物。类似于美国专利No.8,344,207中所述的那些，使用植物叶盘进行生物测定。使用未转化的玉米植物来获得待用作阴性对照的组织。针对秋菜夜蛾(BCW, *Agrotis ipsilon*)、玉米穗虫(CEW, *Helicoverpa zea*)、草地贪夜蛾(FAW, *Spodoptera frugiperda*)和西南玉米螟(SWCB, *Diatraea grandiosella*)以及其他鳞翅类害虫，评估来自每个二元载体的多重转化事件。

[0163] 观察害虫的由摄取表达TIC7472PL或TIC7473PL的本发明叶盘造成的死亡和发育迟缓，并与来自未转化的玉米植物的叶盘进行比较。

[0164] 实施例7

[0165] 在稳定转化的大豆和棉花植物中测定针对鳞翅类害虫的TIC6757PL活性

[0166] 使用本领域已知的方法克隆包含设计为表达质体靶向的和非靶向的TIC7472PL或TIC7473PL杀虫蛋白二者的转基因盒的二元植物转化载体。将所得载体用于稳定转化大豆和棉花植物。从转化体收获组织并用于针对多种鳞翅类害虫的昆虫生物测定。

[0167] 将如上文实施例6中所述设计用于植物表达的合成编码序列克隆到二元植物转化载体中，并用于转化大豆或棉花植物细胞。使用本领域已知的方法构建包含质体靶向的和非靶向的TIC7472PL或TIC7473PL编码序列的二元载体。所得的植物转化载体包含用于表达TIC7472PL或TIC7473PL杀虫蛋白的第一转基因盒，所述第一转基因盒包含组成型启动子，所述组成型启动子的5'与前导序列可操作地连接，所述前导序列的5'与编码质体靶向的或未靶向的TIC7472PL或TIC7473PL蛋白的合成编码序列可操作地连接，所述合成编码序列的5'又与3'UTR可操作地连接；以及用于使用壮观霉素选择来选择转化的植物细胞的第二转基因盒。构建体1、2和7包含编码非靶向TIC6757PL杀虫蛋白的编码序列。构建体3、4、5和6包含编码靶向的TIC6757PL杀虫蛋白的编码序列。

[0168] 通过本领域已知的方法诱导转化的大豆或棉花细胞形成植物。类似于美国专利No.8,344,207中所述的那些，使用植物叶盘进行生物测定。使用未转化的大豆或棉花植物来获得用作阴性对照的组织。针对南部夜蛾(SAW, *Spodoptera eridania*)、大豆尺夜蛾(SBL, *Chrysodeixis includens*)、大豆豆荚虫(SPW, *Helicoverpa zea*)、草地贪夜蛾(FAW, *Spodoptera frugiperda*)、大豆尺夜蛾(SBL, *Chrysodeixis includens*)、烟芽夜蛾(TBW, *Heliothis virescens*)、棉铃虫(Cotton bollworm) (CBW, *Helicoverpa zea*)和绒毛豆夜蛾

(VBW, *Anticarsia gemmatalis*) 以及其他鳞翅类害虫评估来自每个二元载体的多重转化事件。观察害虫的由摄取表达TIC7472PL或TIC7473PL的本发明叶盘造成的死亡和发育迟缓,并与来自未转化的大豆或棉花植物的叶盘进行比较。

[0169] 根据本公开内容,本文公开和要求保护的所有组合物可以在不需过度实验的情况下制备和实施。尽管已经根据前述说明性实施方案了描述本发明的组合物,但是对于本领域技术人员来说显而易见的是,可以对本文所述的组合物的进行变化、改变、修改和变更,而不脱离本发明的真实的观念、精神和范围。更具体地说,显而易见的是,化学和生理学上相关的某些试剂可以替代本文所述的试剂,而获得相同或相似的结果。对于本领域技术人员显而易见的所有这些类似的替代和修改被认为是在由所附权利要求限定的本发明的精神、范围和观念内。

[0170] 在说明书中引用的所有出版物和公开的专利文件通过引用并入本文,其程度如同每个单独的出版物或专利申请被具体地和单独地指示通过引用并入。

序列表

	<110>	孟山都技术公司 (Monsanto Technology LLC)	
	<120>	新型昆虫抑制蛋白	
	<130>	38-21(61627)-0001	
	<150>	US 62/210, 737	
	<151>	2015-08-27	
	<160>	18	
	<170>	PatentIn 3.5 版	
	<210>	1	
	<211>	2394	
	<212>	DNA	
	<213>	日本甲虫类芽孢杆菌	
	<220>		
[0001]	<221>	misc_feature	
	<222>	(1)..(2394)	
	<223>	衍生自日本甲虫类芽孢杆菌菌株 DSC004343 的编码 TIC6757 的 DNA 序列	
	<400>	1	
		atgaagcaga ataataattt tagtgtaagg gccttaccaa gttttattga tgtttttaat	60
		ggaatttatg gttttgccac tggcattcaa gatattttta acatgatttt tggaacagat	120
		acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct	180
		ggtaaacttg aggggattag tggagaccta agtgagatta ttgcgcaggg aaatttgaat	240
		acagaattag ctaaggaatt gctaaaaatc gctaattgagc agaacaacgt attaactgat	300
		gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca	360
		aatatgttaa gcgatgttat gaaacagaat tatgctctga gtcttcaaat agaatatctc	420
		agtaaacaac tacaggagat atcagataaa cttgatgtta ttaacttaaa tgtactcatt	480
		aactctacac tcacagaaat cactcctgct tatcaacgta ttaaatatgt aaatgaaaaa	540

	tttgatgaat taactcttgc tacagaaaaa actctaagag caaaacaagg tagcgaagac	600
	attattgcta atgatactct tgaaaattta actgagctaa cagaactagc gaaaagtgt	660
	acaaaaaatg acatggatag ttctgagttt tatctccata cattccatga tgtattgatt	720
	ggcaataatt tatttggtcg ttcggcttta aaaacagctg cagaattgat tactaaagac	780
	gagataaaga cgagtgaag tgagatagga aaagtttata gtttcttaat tgtactaact	840
	tgtctacaag caaaagcctt tctcacttta acggcatgcc gaaaattatt gggcttatca	900
	gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	960
	gataacatac ttcctacact gtccaataaa ttttctaacc ctaattatgt aaaaactata	1020
	ggtagtgata attatgcaaa agttatttta gaagctgaac caggatatgc ttaggtgga	1080
	tttgaaatta tcaatgatcg aatcccggta ttaaagcgt ataaagctaa gctaaaacaa	1140
	aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1200
[0002]	ttttgtccaa aaaattctga acaaaaatat tatactaaaa gtctgacatt tcctgatggc	1260
	tatgttatta ctaagattac ctttgaaaaa aagctgaaca acctaagata tgaggcaaca	1320
	gcaaattttt atgaccatc tacaggagat attgatttaa atgagaagca agtggaatct	1380
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgatgg tgtttacatg	1440
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1500
	gacgagaaat cgaaaatctt aactttaaca tgtaaacttt atttacgaga atatttatta	1560
	gaatctgatt taataaataa agagacaagc ctcatgtct cgcctaagt tttatcagt	1620
	aatatcgtag aaaattggaa catagaagcg gataatctag aaccatgggt agcaaa taac	1680
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1740
	ggtgatgggg aattttcaca atttattgga gataaattaa aaccaaatac agattatatt	1800
	attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat	1860

actatgtacg aagatacaaa cggtagttct gaagaatttc aaactatagc tgtaaattat 1920
acttcagaaa ctgatccttc acaaacacat ttagttttta aaagtcaaag tggctatgag 1980
gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggtcca 2040
gaattgataa aatttgatga ttggattagt tttggtacta cttacattag agatgatgta 2100
cttactatcg atccaagtcg tggaggttat ttagacaat ctcttaaatt agacagctat 2160
tcaacttata atttgagctt ttctttttct ggattatggg ctaaggttat tataaaaaat 2220
tcccacggag tagtatgtt tgaaaaagta agtcagcagt cttcatacgt agatattagt 2280
gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt 2340
cgtggtggtt ttgggtcggt ccgtgatttt tctatgaagg aaaagtttga ataa 2394

[0003]

<210> 2
<211> 797
<212> PRT
<213> 日本甲虫类芽孢杆菌

<220>
<221> MISC_FEATURE
<222> (1)..(797)
<223> 衍生自编码 TIC6757 的日本甲虫类芽孢杆菌 DSC004343 编码序列的 TIC6757 的
氨基酸序列

<400> 2

Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe Ile
1 5 10 15

Asp Val Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Gln Asp Ile
20 25 30

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

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

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

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

Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met Leu
100 105 110

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

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

[0004]

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

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

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

Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu Glu
195 200 205

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

Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu Ile

225	230	235	240
Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu Leu			
245	250	255	
Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys Val			
260	265	270	
Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe Leu			
275	280	285	
Thr Leu Thr Ala Cys Arg Lys Leu Leu Gly Leu Ser Asp Ile Asp Tyr			
290	295	300	
Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe Arg			
305	310	315	320
[0005]			
Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser Asn Pro Asn Tyr			
325	330	335	
Val Lys Thr Ile Gly Ser Asp Asn Tyr Ala Lys Val Ile Leu Glu Ala			
340	345	350	
Glu Pro Gly Tyr Ala Leu Val Gly Phe Glu Ile Ile Asn Asp Arg Ile			
355	360	365	
Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln Val			
370	375	380	
Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys Leu			
385	390	395	400
Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu Thr			
405	410	415	

	Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys Leu	
	420	425 430
	Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Pro Ser Thr	
	435	440 445
	Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu Gln	
	450	455 460
	Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr Met	
	465	470 475 480
	Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser Phe	
	485	490 495
[0006]	Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys Lys	
	500	505 510
	Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys Glu	
	515	520 525
	Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val Glu	
	530	535 540
	Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn Asn	
	545	550 555 560
	Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys Ala	
	565	570 575
	Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys	
	580	585 590

	Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly Lys	
	595	600 605
	Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr Glu	
	610	615 620
	Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn Tyr	
	625	630 635 640
	Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser Gln	
		645 650 655
	Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys Lys	
		660 665 670
[0007]	Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp Trp	
	675	680 685
	Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile Asp	
	690	695 700
	Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser Tyr	
	705	710 715 720
	Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys Val	
		725 730 735
	Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser Gln	
	740	745 750
	Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr Thr Thr Ser Asn	
	755	760 765

Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly Phe
770 775 780

Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu
785 790 795

<210> 3

<211> 2397

<212> DNA

<213> 人工

<220>

<223> 衍生自编码 TIC6757 的日本甲虫类芽孢杆菌菌株 DSC004343 的设计用于植物表达的编码 TIC6757PL 的合成 DNA 序列，相对于细菌 TIC6757 氨基酸序列，TIC6757PL 具有插入在 2 位的额外的丙氨酸残基

<400> 3

[0008] atggctaagc agaacaacaa cttctccgtg cgcgcgctcc cgtcgttcat cgacgtcttc 60

aacggcatct acgggttcgc caccggcatc caggacatct tcaacatgat cttcggcacc 120

gacacgggcg acctcacgct ggaggaggtg ctcaagaacc aggaactgct ctacgacatc 180

tcgggcaagc tggagggtcat ctccggcgac ctacgcgaaa ttatcgcgca agggaaacctc 240

aacacggagc tggcgaaaga gctgctcaag atcgccaacg agcagaacaa cgtgctgacg 300

gacgtgaaca acaagctcaa cgcgatcaac tcgatgctcc acatctacct cccaaagatt 360

acgaacatgc tgtccgacgt catgaagcag aactatgcc tcagcctcca gatcgagtac 420

ctctcgaagc agtccagga gatttcggat aagctggacg tcatcaacct gaacgtgctg 480

ataaactcca cgtcacgga gatcactccg gcctaccagc gcatcaagta cgtgaacgag 540

aagttcgacg agctgacact ggcgactgag aagaccctcc gcgccaagca agggctccgag 600

gacatcattg cgaacgacac gctggagaac ttgacggagt tgaccgagct ggccaagagc 660

gtgacgaaga acgacatgga cagcttcgag ttctacctcc acactttcca cgacgtgctg 720

atcggaaca acctcttcgg ccggagcgcg ctcaagaccg cagccgagct gattacgaag 780

	gacgagatca agacctccgg gtccgagatc ggaaaggtgt acagcttcct gatcgtcctc	840
	acgtgcttac aagctaaggc gtccctcacc ctgaccgcct gccgtaagct gttgggcctg	900
	tccgacatcg actacaccaa catcctcaac cagcacctca acgacgagaa gaacgtcttt	960
	cgggacaaca tcctcccgac actgagcaac aagttcagta acccgaacta cgtcaagacg	1020
	attggcagcg acaactacgc caaggttata ctagaggccg agcccgtta cgcctgtgc	1080
	gggttcgaga tcatcaacga ccgcattccc gtctcaagg cgtacaaggc caagctcaag	1140
	cagaactacc aagtggacca ccagagcctc agcgagatcg tgtacctgga catcgacaag	1200
	ttgtttctgcc cgaagaactc ggagcagaag tattacacca agagcctgac cttcccgac	1260
	ggctacgtta ttacaaagat cacgttcgag aagaagctca acaacctccg ttacgaggca	1320
	actgccaaact tctacgacce gtccaccggc gacattgact tgaacgagaa gcaagtcgag	1380
	tccacgttcc tccaggccga ctacatcagc atcaacgtca gcgacgacga cggcgtgtac	1440
[0009]	atgccgctcg gattcatcag cgagacctc ctacagccga tcaactcgtt cgagctggag	1500
	gtggacgaga agtccaagat tctcaccctg acctgcaaga gctacctccg ggaatactc	1560
	ctggagagcg acctcatcaa taaggagact tcgctcatag ctccgccaa cgtcttcac	1620
	tccaacatcg tcgagaactg gaacatcgag gccgacaacc tggagccgtg ggtggcaaac	1680
	aacaagaacg cctacgtgga ctccaccggc gggatcgagg gaagcaaggc cctgttcacc	1740
	caggcgacg gtgagttctc gcagttcacc ggcgacaagg tcaagcccaa cacggactac	1800
	atcatccagt acaccgtcaa gggcaagcca gcgatctacc tcaagaacaa gaacaccggg	1860
	tacacatgt acgaggacac gaacggcagc agcgaggagt tccagaccat cgcggtcaac	1920
	tacacctccg agaccgatcc ctcccagacc caccttgtct tcaagtcca gagcggctac	1980
	gaggcgtggg gcgacaactt catcatcctg gattgcaagg ctttcgagac tcccagggc	2040
	ccggagctta tcaagttcga tgactggatt tcgtttggca ccacctacat ccgggacgac	2100
	gtgctaacca tcgacctgtc gcgcggcggc tacttccgcc aaagcctcaa gctggactcg	2160

tactccacgt acaacctatc tttcagcttc tcgggcctgt gggcgaaggt gattatcaag 2220
aactcgcacg gcgtggtcct gttcgagaag gtctcccagc agtcgtcata cgtggacatc 2280
agcgagtcct tcaccaccac cagcaacaag gagggcttct tcatcgagct gaccggcgac 2340
tcgcgcggcg ggttcggtc cttccgcgac ttctccatga aggagaaatt cgagtga 2397

<210> 4

<211> 798

<212> PRT

<213> 人工

<220>

<223> 由合成 DNA 序列编码的 TIC6757PL 的氨基酸序列，其中相对于细菌 TIC6757 氨基酸序列，在 2 位插入了额外的丙氨酸残基

<400> 4

[0010] Met Ala Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe
1 5 10 15
Ile Asp Val Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Gln Asp
20 25 30
Ile Phe Asn Met Ile Phe Gly Thr Asp Thr Gly Asp Leu Thr Leu Glu
35 40 45
Glu Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu
50 55 60
Glu Gly Ile Ser Gly Asp Leu Ser Glu Ile Ile Ala Gln Gly Asn Leu
65 70 75 80
Asn Thr Glu Leu Ala Lys Glu Leu Leu Lys Ile Ala Asn Glu Gln Asn
85 90 95

	Asn Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met	
	100	105 110
	Leu His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser Asp Val Met	
	115	120 125
	Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys Gln	
	130	135 140
	Leu Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu	
	145	150 155 160
	Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys	
	165	170 175
[0011]	Tyr Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr	
	180	185 190
	Leu Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu	
	195	200 205
	Glu Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn	
	210	215 220
	Asp Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu	
	225	230 235 240
	Ile Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu	
	245	250 255
	Leu Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys	
	260	265 270

Val Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe
275 280 285

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

Tyr Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe
305 310 315 320

Arg Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser Asn Pro Asn
325 330 335

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

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

[0012]

Ile Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln
370 375 380

Val Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys
385 390 395 400

Leu Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu
405 410 415

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

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

Thr Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu

450	455	460	
Gln Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr			
465	470	475	480
Met Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser			
485	490	495	
Phe Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys			
500	505	510	
Lys Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys			
515	520	525	
Glu Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val			
530	535	540	
[0013]			
Glu Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn			
545	550	555	560
Asn Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys			
565	570	575	
Ala Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp			
580	585	590	
Lys Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly			
595	600	605	
Lys Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr			
610	615	620	
Glu Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn			
625	630	635	640

	Tyr Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser	
	645	650 655
	Gln Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys	
	660	665 670
	Lys Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp	
	675	680 685
	Trp Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile	
	690	695 700
	Asp Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser	
	705	710 715 720
[0014]	Tyr Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys	
	725	730 735
	Val Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser	
	740	745 750
	Gln Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr Thr Thr Ser	
	755	760 765
	Asn Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly	
	770	775 780
	Phe Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu	
	785	790 795
	<210> 5	
	<211> 2454	
	<212> DNA	

<213> 人工	
<220>	
<223> 编码组氨酸标记的 TIC6757 蛋白的重组核酸序列	
<400> 5	
atgcatcacc atcaccatca ccatcaccat cacggtaccg agaccgtccg cttccaatcc	60
atgaagcaga ataataattt tagtgtaagg gccttaccaa gttttattga tgttttta	120
ggaatttatg gttttgccac tggcattcaa gatattttta acatgatttt tggaacagat	180
acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct	240
ggtaaaactg aggggattag tggagaccta agtgagatta ttgcgcaggg aaatttgaat	300
acagaattag ctaaggaatt gctaaaaatc gctaattgagc agaacaacgt attaactgat	360
gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca	420
aatatgttaa gcgatgttat gaaacagaat tatgctctga gtcttcaa	480
[0015] agtaacaac tacaggagat atcagataaa cttgatgtta ttaacttaaa tgtactcatt	540
aactctacac tcacagaaat cactcctgct tatcaacgta ttaa	600
tttgatgaat taactcttgc tacagaaaaa actctaagag caaaacaagg tagcgaagac	660
attattgcta atgatactct tgaatttta actgagctaa cagaactagc gaaaagtgt	720
acaaaaaatg acatggatag ttctgagttt tatctccata cattccatga tgtattgatt	780
ggcaataatt tatttggtcg ttcggttta aaaacagctg cagaattgat tactaaagac	840
gagataaaga cgagtgaag tgagatagga aaagtttata gtttctta	900
tgtctacaag caaaagcctt tctcatttta acggcatgcc gaaaattatt gggcttatca	960
gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	1020
gataacatac ttctacact gtccaataaa ttttctaacc ctaattatgt aaaaactata	1080
ggtagtgata attatgcaaa agttatttta gaagctgaac caggatatgc tttagttgga	1140
tttgaaatta tcaatgatcg aatcccggtt ttaaaagcgt ataaagctaa gctaaaacaa	1200

	aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1260
	ttttgtccaa aaaattctga acaaaaatat tataactaaaa gtctgacatt tcctgatggc	1320
	tatgttatta ctaagattac ctttgaaaaa aagctgaaca acctaagata tgaggcaaca	1380
	gcaaattttt atgaccatc tacaggagat attgatttaa atgagaagca agtggaatct	1440
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgatgg tgtttacatg	1500
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1560
	gacgagaaat cgaaaatctt aactttaaca tgtaaactct atttacgaga atatttatta	1620
	gaatctgatt taataaataa agagacaagc ctcatgtct cgcctaagt ttttatcagt	1680
	aatatcgtag aaaattggaa catagaagcg gataatctag aacctgggt agcaaataac	1740
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1800
	ggtgatgggg aattttcaca atttattgga gataaattaa aaccaaatac agattatatt	1860
[0016]	attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat	1920
	actatgtacg aagatacaaa cggtagttct gaagaatttc aaactatagc tgtaaattat	1980
	acttcagaaa ctgaccttc acaaacacat ttagttttta aaagtcaaag tggctatgag	2040
	gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggtcca	2100
	gaattgataa aatttgatga ttggattagt ttgtgacta cttacattag agatgatgta	2160
	cttactatcg atccaagtcg tggaggttat ttagacaat ctcttaaatt agacagctat	2220
	tcaacttata atttgagctt ttctttttct ggattatggg ctaaggttat tataaaaaat	2280
	tcccacggag tagtatgtt tgaaaaagta agtcagcagt cttcatacgt agatattagt	2340
	gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt	2400
	cgtggtggtt ttgggtcggt ccgtgatttt tctatgaagg aaaagtttga ataa	2454

<210> 6

<211> 817
<212> PRT
<213> 人工

<220>
<223> 组氨酸标记的 TIC6757 蛋白的氨基酸序列

<400> 6

Met His His His His His His His His His His Gly Thr Glu Thr Val
1 5 10 15

Arg Phe Gln Ser Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu
20 25 30

Pro Ser Phe Ile Asp Val Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly
35 40 45

[0017] Ile Gln Asp Ile Phe Asn Met Ile Phe Gly Thr Asp Thr Gly Asp Leu
50 55 60

Thr Leu Glu Glu Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser
65 70 75 80

Gly Lys Leu Glu Gly Ile Ser Gly Asp Leu Ser Glu Ile Ile Ala Gln
85 90 95

Gly Asn Leu Asn Thr Glu Leu Ala Lys Glu Leu Leu Lys Ile Ala Asn
100 105 110

Glu Gln Asn Asn Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile
115 120 125

Asn Ser Met Leu His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser
130 135 140

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

Ser Lys Gln Leu Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu
165 170 175

Asn Val Leu Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln
180 185 190

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

Glu Lys Thr Leu Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn
210 215 220

Asp Thr Leu Glu Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val
225 230 235 240

[0018]

Thr Lys Asn Asp Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His
245 250 255

Asp Val Leu Ile Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr
260 265 270

Ala Ala Glu Leu Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu
275 280 285

Ile Gly Lys Val Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala
290 295 300

Lys Ala Phe Leu Thr Leu Thr Ala Cys Arg Lys Leu Leu Gly Leu Ser
305 310 315 320

Asp Ile Asp Tyr Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys

325	330	335
Asn Val Phe Arg Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser		
340	345	350
Asn Pro Asn Tyr Val Lys Thr Ile Gly Ser Asp Asn Tyr Ala Lys Val		
355	360	365
Ile Leu Glu Ala Glu Pro Gly Tyr Ala Leu Val Gly Phe Glu Ile Ile		
370	375	380
Asn Asp Arg Ile Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln		
385	390	395
400		
Asn Tyr Gln Val Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp		
405	410	415
[0019]		
Ile Asp Lys Leu Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr		
420	425	430
Lys Ser Leu Thr Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe		
435	440	445
Glu Lys Lys Leu Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr		
450	455	460
Asp Pro Ser Thr Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser		
465	470	475
480		
Thr Phe Leu Gln Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp		
485	490	495
Gly Val Tyr Met Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro		
500	505	510

	Ile Asn Ser Phe Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr	
	515	520 525
	Leu Thr Cys Lys Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu	
	530	535 540
	Ile Asn Lys Glu Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser	
	545	550 555 560
	Asn Ile Val Glu Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp	
		565 570 575
	Val Ala Asn Asn Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu	
		580 585 590
[0020]	Gly Ser Lys Ala Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe	
	595	600 605
	Ile Gly Asp Lys Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr	
	610	615 620
	Val Lys Gly Lys Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr	
	625	630 635 640
	Thr Met Tyr Glu Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile	
		645 650 655
	Ala Val Asn Tyr Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val	
	660	665 670
	Phe Lys Ser Gln Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile	
	675	680 685

Leu Glu Cys Lys Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys
690 695 700

Phe Asp Asp Trp Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val
705 710 715 720

Leu Thr Ile Asp Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys
725 730 735

Leu Asp Ser Tyr Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu
740 745 750

Trp Ala Lys Val Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu
755 760 765

[0021] Lys Val Ser Gln Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr
770 775 780

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

Arg Gly Gly Phe Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe
805 810 815

Glu

<210> 7
<211> 2394
<212> DNA
<213> 日本甲虫类芽孢杆菌

<220>
<221> misc_feature

<222> (1)..(2394)

<223> 衍生自编码 TIC7472 的日本甲虫类芽孢杆菌菌株 DSC007648 的 DNA 序列

<400> 7

atgaagcaga ataataattt tagtgtaagg gccttaccaa gttttattga tgtttttaat	60
ggaatttatg attttgccac tggcattcaa gatattttta acatgatttt tggaacagat	120
acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct	180
ggtaaacttg aggggattag tggagaccta agtgagatta ttgcgcaggg aaatttgaat	240
acagaattag ctaaggaatt gctaaaaatc gctaatagagc agaacaacgt attaactgat	300
gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca	360
aatatgttaa gcgatgttat gaaacagaat tatgctctga gtcttcaaata agaatatctc	420
agtaaacaac tacaggagat atcagataaa ctgatgtta ttaacttaaa tgtactcatt	480
aactctacac tcacagaaat cactcctgct tatcaacgta ttaaataatgt aaatgaaaaa	540
[0022] tttgatgaat taactcttgc tacagaaaaa actctaagag caaaacaagg tagcgaagac	600
attattgcta atgatactct tgaataatga actgagctaa cagaactagc gaaaagtgtg	660
acaaaaaatg acatggatag tttcgagttt tatctccata cattccatga tgtattgatt	720
ggcaataatt tatttggtcg ttcggcttta aaaacagctg cagaattgat tactaaagac	780
gagataaaga cgagtggaag tgagatagga aaagtttata gtttcttaat tgtactaact	840
tgtctacaag caaaagcctt tctcacttta acggcatgcc gaaaattatt gggcttatca	900
gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	960
gataacatac ttcctacact gtccaataaa ttttctaacc ctaattatgt aaaaactata	1020
ggtagtata attatgcaa agttatttta gaagctgaac caggatatgc tttagttgga	1080
tttgaaatta tcaatgatcg aatcccggta ttaaaagcgt ataaagctaa gctaaaacaa	1140
aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1200
ttttgtccaa aaaattctga acaaaaatat tataactaaa gtctgacatt tcctgatggc	1260

	tatgtttatta ctaagattac ctttgaaaa aagctgaaca acctaagata tgaggcaaca	1320
	gcaaattttt atgacccatc tacaggagat attgatttaa atgagaagca agtggaatct	1380
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgat gg tgtttacatg	1440
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1500
	gacgagaaat cgaaaatctt aactttaaca tgtaaactctt atttacgaga atatttatta	1560
	gaatctgatt taataaataa agagacaagc ctcatgtgct cgcctaagt ttttatcagt	1620
	aatatcgtag aaaattggaa catagaagcg gataatctag aacatgggt agcaaataac	1680
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1740
	ggtgatgggg aattttcaca atttattgga gataaattaa aaccaaatac agattatatt	1800
	attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat	1860
	actatgtacg aagatacaaa cggtagtctt gaagaatttc aaactatagc tgtaaattat	1920
[0023]	acttcagaaa ctgatccttc acaaacacat ttagttttta aaagtcaaag tggctatgag	1980
	gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggtcca	2040
	gaattgataa aatttgatga ttggattagt ttggtacta cttacattag agatgatgta	2100
	cttactatcg atccaagtcg tggaggttat ttagacaat ctcttaaatt agacagctat	2160
	tcaacttata atttgagctt ttctttttct ggattatggg ctaaggttat tataaaaaat	2220
	tcccacggag tagtattgtt tgaaaaagta agtcagcagt cttcatagct agatattaat	2280
	gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt	2340
	cgtggtggtt ttgggtcggt ccgtgatttt tctatgaagg aaaagtttga ataa	2394
	<210> 8	
	<211> 797	
	<212> PRT	
	<213> 日本甲虫类芽孢杆菌	

<220>

<221> MISC_FEATURE

<222> (1)..(797)

<223> 衍生自编码 TIC7472 的日本甲虫类芽孢杆菌菌株 DSC007648 编码序列的
TIC7472 的氨基酸序列

<400> 8

Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe Ile
1 5 10 15

Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp Ile
20 25 30

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

[0024] Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu Glu
50 55 60

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

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

Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met Leu
100 105 110

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

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

	Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu Ile	
	145	150 155 160
	Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys Tyr	
		165 170 175
	Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr Leu	
		180 185 190
	Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu Glu	
		195 200 205
	Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp	
		210 215 220
[0025]	Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu Ile	
	225	230 235 240
	Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu Leu	
		245 250 255
	Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys Val	
		260 265 270
	Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe Leu	
		275 280 285
	Thr Leu Thr Ala Cys Arg Lys Leu Leu Gly Leu Ser Asp Ile Asp Tyr	
		290 295 300
	Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe Arg	
	305	310 315 320

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

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

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

Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln Val
370 375 380

Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys Leu
385 390 395 400

Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu Thr
405 410 415

[0026]

Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys Leu
420 425 430

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

Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu Gln
450 455 460

Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr Met
465 470 475 480

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

Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys Lys

500	505	510
Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys Glu		
515	520	525
Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val Glu		
530	535	540
Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn Asn		
545	550	555 560
Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys Ala		
565	570	575
Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys		
580	585	590
[0027]		
Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly Lys		
595	600	605
Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr Glu		
610	615	620
Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn Tyr		
625	630	635 640
Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser Gln		
645	650	655
Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys Lys		
660	665	670
Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp Trp		
675	680	685

	Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile Asp 690 695 700	
	Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser Tyr 705 710 715 720	
	Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys Val 725 730 735	
	Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser Gln 740 745 750	
	Gln Ser Ser Tyr Val Asp Ile Asn Glu Ser Phe Thr Thr Thr Ser Asn 755 760 765	
[0028]	Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly Phe 770 775 780	
	Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu 785 790 795	
	<210> 9	
	<211> 2418	
	<212> DNA	
	<213> 人工	
	<220>	
	<223> 编码组氨酸标记的 TIC7472 蛋白的重组核酸序列	
	<400> 9	
	atgaagcaga ataataattt tagtgtaagg gccttaccaa gttttattga tgtttttaat	60
	ggaatttatg attttgccac tggcattcaa gatattttta acatgatttt tggaacagat	120
	acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct	180

	ggtaaacttg aggggattag tggagaccta agtgagatta ttgcgcaggg aaatttgaat	240
	acagaattag ctaaggaatt gctaaaaatc gctaattgagc agaacaacgt attaactgat	300
	gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca	360
	aatatgttaa gcgatgttat gaaacagaat tatgctctga gtcttcaaata agaatatctc	420
	agtaaacaac tacaggagat atcagataaa cttgatgtta ttaacttaaa tgtactcatt	480
	aactctacac tcacagaaat cactcctgct tatcaacgta ttaaataatgt aaatgaaaaa	540
	tttgatgaat taactcttgc tacagaaaaa actctaagag caaacaagg tagcgaagac	600
	attattgcta atgatactct tgaaaattta actgagctaa cagaactagc gaaaagtgt	660
	acaaaaaatg acatggatag ttctgagttt tatctccata cattccatga tgtattgatt	720
	ggcaataatt tatttggtcg ttcggttta aaaacagctg cagaattgat tactaaagac	780
	gagataaaga cgagtggaag tgagatagga aaagtttata gtttcttaat tgtactaact	840
[0029]	tgtctacaag caaaagcctt tctcacttta acggcatgcc gaaaattatt gggcttatca	900
	gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	960
	gataacatac ttctacact gtccaataaa ttttctaacc ctaattatgt aaaaactata	1020
	ggtagtata attatgcaaa agttatttta gaagctgaac caggatatgc tttagttgga	1080
	tttgaaatta tcaatgatcg aatcccggta ttaaaagcgt ataaagctaa gctaaaacaa	1140
	aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1200
	ttttgtccaa aaaattctga acaaaaatat tatactaaaa gtctgacatt tctgatggc	1260
	tatgttatta ctaagattac ctttgaaaaa aagctgaaca acctaagata tgaggcaaca	1320
	gcaaattttt atgaccatc tacaggagat attgatttaa atgagaagca agtggaatct	1380
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgatgg tgtttacatg	1440
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1500
	gacgagaaat cgaaaatctt aactttaaca tgtaaactctt atttacgaga atatttatta	1560

	gaatctgatt taataataa agagacaagc ctcatgtgct cgcctaattgt ttttatcagt	1620
	aatatcgtag aaaattggaa catagaagcg gataatctag aacctgggt agcaaataac	1680
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1740
	ggtgatgggg aattttcaca atttattgga gataaattaa aaccaaatac agattatatt	1800
	attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat	1860
	actatgtacg aagatacaaa cggtagttct gaagaatttc aaactatagc tgtaaattat	1920
	acttcagaaa ctgataccttc acaaacacat ttagttttta aaagtcaaag tggctatgag	1980
	gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggcca	2040
	gaattgataa aatttgatga ttggattagt tttggtacta cttacattag agatgatgta	2100
	cttactatcg atccaagtcg tggaggttat tttagacaat ctcttaaatt agacagctat	2160
	tcaacttata atttgagctt ttctttttct ggattatggg ctaaggttat tataaaaaat	2220
[0030]	tcccacggag tagtattgtt tgaaaaagta agtcagcagt cttcatacgt agatattaat	2280
	gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt	2340
	cgtggtggtt ttgggtcggt ccgtgatttt tctatgaagg aaaagtttga acaccacat	2400
	cacgctcacc atcactga	2418

<210> 10

<211> 805

<212> PRT

<213> 人工

<220>

<223> 组氨酸标记的 TIC7472 蛋白的氨基酸序列

<400> 10

Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe Ile

1

5

10

15

	Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp Ile	
	20 25 30	
	Phe Asn Met Ile Phe Gly Thr Asp Thr Gly Asp Leu Thr Leu Glu Glu	
	35 40 45	
	Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu Glu	
	50 55 60	
	Gly Ile Ser Gly Asp Leu Ser Glu Ile Ile Ala Gln Gly Asn Leu Asn	
	65 70 75 80	
	Thr Glu Leu Ala Lys Glu Leu Leu Lys Ile Ala Asn Glu Gln Asn Asn	
	85 90 95	
[0031]	Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met Leu	
	100 105 110	
	His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser Asp Val Met Lys	
	115 120 125	
	Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys Gln Leu	
	130 135 140	
	Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu Ile	
	145 150 155 160	
	Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys Tyr	
	165 170 175	
	Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr Leu	
	180 185 190	

Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu Glu
195 200 205

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

Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu Ile
225 230 235 240

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

Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys Val
260 265 270

Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe Leu
275 280 285

[0032]

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

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

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

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

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

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

370	375	380	
Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys Leu			
385	390	395	400
Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu Thr			
	405	410	415
Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys Leu			
	420	425	430
Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Pro Ser Thr			
	435	440	445
Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu Gln			
	450	455	460
[0033]			
Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr Met			
465	470	475	480
Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser Phe			
	485	490	495
Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys Lys			
	500	505	510
Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys Glu			
	515	520	525
Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val Glu			
	530	535	540
Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn Asn			
545	550	555	560

	Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys Ala	
	565	570 575
	Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys	
	580	585 590
	Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly Lys	
	595	600 605
	Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr Glu	
	610	615 620
	Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn Tyr	
	625	630 635 640
[0034]	Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser Gln	
	645	650 655
	Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys Lys	
	660	665 670
	Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp Trp	
	675	680 685
	Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile Asp	
	690	695 700
	Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser Tyr	
	705	710 715 720
	Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys Val	
	725	730 735

Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser Gln
 740 745 750

Gln Ser Ser Tyr Val Asp Ile Asn Glu Ser Phe Thr Thr Thr Ser Asn
 755 760 765

Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly Phe
 770 775 780

Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu His His His
 785 790 795 800

His Ala His His His
 805

[0035] <210> 11
 <211> 2394
 <212> DNA
 <213> 日本甲虫类芽孢杆菌

<220>
 <221> misc_feature
 <222> (1)..(2394)
 <223> 衍生自编码 TIC7473 的日本甲虫类芽孢杆菌菌株 DSC008493 的 DNA 序列

<400> 11
 atgaagcaga ataataattt tagtgaagg gccttacaa gttttattga tgtttttaat 60
 ggaatttatg attttgccac tggcattcaa gatattttta acatgatttt tggaacagat 120
 acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct 180
 ggtaaacttg aggggattag tggagaccta agtgagatta ttgcgcaggg aaatttgaat 240
 acagaattag ctaaggaatt gctaaaaatc gctaatgagc agaacaacgt attactgat 300
 gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca 360

	aatatgttaa gcatgttat gaaacagaat tatgctctga gtcttcaa at agaatatctc	420
	agtaacaac tacaggagat atcagataaa ctgatgtta ttaacttaaa tgtactcatt	480
	aactctacac tcacagaaat cactcctgct tatcaacgta ttaa atatgt aaatgaaaa	540
	tttgatgaat taactcttgc tacagaaaaa actctaagag caaaacaagg tagcgaagac	600
	attattgcta atgatactct tgaaaattta actgagctaa cagaactagc gaaaagtgt	660
	acaaaaaatg acatggatag ttctgagttt tatctccata cattccatga tgtattgatt	720
	ggcaataatt tatttggctg ttccggcttta aaaacagctg cagaattgat tactaaagac	780
	gagataaaga cgagtgaag tgagatagga aaagtttata gtttctta at tgtactaact	840
	tgtctacaag caaaagcctt tctcacttta acggcatgcc gaaaattatt gggcttatca	900
	gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	960
	gataacatac ttcctacact gtccaataaa ttttctaacc ctaattatgt aaaaactata	1020
[0036]	ggtagtata attatgcaaa agttatttta gaagctgaac caggatatgc tttagttgga	1080
	tttgaaatta tcaatgatcg aatcccgta ttaaagcgt ataaagctaa gctaaaacaa	1140
	aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1200
	ttttgtccaa aaaattctga acaaaaatat tatactaaaa gtctgacatt tcctga tggc	1260
	tatgttatta ctaagattac ctttgaaaaa aagctgaaca acctaagata tgaggcaaca	1320
	gcaaattttt atgaccatc tacaggagat attgatttaa atgagaagca agtggaatct	1380
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgatgg tgtttacatg	1440
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1500
	gacgagaaat cgaaaatctt aactttaaca tgtaa atctt atttacgaga atatttatta	1560
	gaatctgatt taataaataa agagacaagc ctcatgtctc cgccta atgt ttttatcagt	1620
	aatatcgtag aaaattggaa catagaagcg gataatctag aacctgggt agcaataac	1680
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1740


```

ggatgatgggg aattttcaca atttattgga gataaattaa aaccaaatac agattatatt 1800

attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat 1860

actatgtacg aagatacaaaa cggtagttct gaagaatttc aaactatagc tgtaaattat 1920

acttcagaaa ctgatccttc acaaacacat ttagttttta aaagtcaaag tggctatgag 1980

gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggtcca 2040

gaattgataa aatttgatga ttggattagt tttggtacta cttacattag agatgatgta 2100

cttactatcg atccaagtcg tggaggttat tttagacaat ctcttaaatt agacagctat 2160

tcaacttata atttgagctt ttctttttct ggattatggg ctaaggttat tataaaaaat 2220

tcccacggag tagtattgtt tgaaaaagta agtcagcagt cttcatacgt agatattagt 2280

gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt 2340

cgtggtggtt ttgggtcggt ccgtgatatt tctatgaagg aaaagtttga ataa 2394

```

[0037]

```

<210> 12
<211> 797
<212> PRT
<213> 日本甲虫类芽孢杆菌

<220>
<221> MISC_FEATURE
<222> (1)..(797)
<223> 衍生自编码 TIC7473 的日本甲虫类芽孢杆菌菌株 DSC008493 编码序列的
TIC7473 的氨基酸序列

```

<400> 12

```

Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe Ile
1           5           10           15

```

```

Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp Ile
          20           25           30

```

	Phe Asn Met Ile Phe Gly Thr Asp Thr Gly Asp Leu Thr Leu Glu Glu	
	35 40 45	
	Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu Glu	
	50 55 60	
	Gly Ile Ser Gly Asp Leu Ser Glu Ile Ile Ala Gln Gly Asn Leu Asn	
	65 70 75 80	
	Thr Glu Leu Ala Lys Glu Leu Leu Lys Ile Ala Asn Glu Gln Asn Asn	
	85 90 95	
	Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met Leu	
	100 105 110	
[0038]	His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser Asp Val Met Lys	
	115 120 125	
	Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys Gln Leu	
	130 135 140	
	Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu Ile	
	145 150 155 160	
	Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys Tyr	
	165 170 175	
	Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr Leu	
	180 185 190	
	Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu Glu	
	195 200 205	

	Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp	
	210	215 220
	Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu Ile	
	225	230 235 240
	Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu Leu	
	245	250 255
	Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys Val	
	260	265 270
	Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe Leu	
	275	280 285
[0039]	Thr Leu Thr Ala Cys Arg Lys Leu Leu Gly Leu Ser Asp Ile Asp Tyr	
	290	295 300
	Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe Arg	
	305	310 315 320
	Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser Asn Pro Asn Tyr	
	325	330 335
	Val Lys Thr Ile Gly Ser Asp Asn Tyr Ala Lys Val Ile Leu Glu Ala	
	340	345 350
	Glu Pro Gly Tyr Ala Leu Val Gly Phe Glu Ile Ile Asn Asp Arg Ile	
	355	360 365
	Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln Val	
	370	375 380

Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys Leu
385 390 395 400

Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu Thr
405 410 415

Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys Leu
420 425 430

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

Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu Gln
450 455 460

Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr Met
465 470 475 480

[0040]

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

Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys Lys
500 505 510

Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys Glu
515 520 525

Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val Glu
530 535 540

Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn Asn
545 550 555 560

Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys Ala

565	570	575
Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys		
580	585	590
Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly Lys		
595	600	605
Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr Glu		
610	615	620
Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn Tyr		
625	630	635
640		
Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser Gln		
645	650	655
[0041]		
Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys Lys		
660	665	670
Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp Trp		
675	680	685
Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile Asp		
690	695	700
Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser Tyr		
705	710	715
720		
Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys Val		
725	730	735
Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser Gln		
740	745	750

Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr Thr Thr Ser Asn
 755 760 765

Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly Phe
 770 775 780

Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu
 785 790 795

<210> 13
 <211> 2418
 <212> DNA
 <213> 人工

<220>
 <223> 编码组氨酸标记的 TIC7473 蛋白的重组核酸序列

[0042] <400> 13
 atgaagcaga ataataattt tagtgtaagg gccttaccaa gttttattga tgtttttaat 60
 ggaatttatg attttgccac tggcattcaa gatattttta acatgatttt tggaacagat 120
 acaggtgatc taacactaga agaagtttta aaaaatcaag agttacttta tgatatttct 180
 ggtaaacttg aggggattag tggagacctt agtgagatta ttgcgcaggg aaatttgaat 240
 acagaattag ctaaggaatt gctaaaaatc gctaatgagc agaacaacgt attactgat 300
 gttaataaca aactcaatgc gataaattcg atgctccaca tctatcttcc taaaattaca 360
 aatatgttaa gcgatgttat gaaacagaat tatgctctga gtcttcaaat agaatatctc 420
 agtaaacac tacaggagat atcagataaa cttgatgtta ttaacttaaa tgtactcatt 480
 aactctacac tcacagaaat cactcctgct tatcaacgta ttaaataatgt aaatgaaaaa 540
 tttgatgaat taactcttgc tacagaaaaa actctaagag caaaacaagg tagcgaagac 600
 attattgcta atgatactct tgaaaattta actgagctaa cagaactagc gaaaagtgt 660

	acaaaaaatg acatggatag ttctgagttt tatctccata cattccatga tgtattgatt	720
	ggcaataatt tatttggtcg ttccggcttta aaaacagctg cagaattgat tactaaagac	780
	gagataaaga cgagtggaag tgagatagga aaagtttata gtttcttaat tgtactaact	840
	tgtctacaag caaaagcctt tctcacttta acggcatgcc gaaaattatt gggcttatca	900
	gatattgatt atactaatat tctaaatcag catctaaatg atgaaaagaa tgtatttcgt	960
	gataacatac ttcttacct gtccaataaa ttttctaacc ctaattatgt aaaaactata	1020
	ggtagtgata attatgcaaa agttatttta gaagctgaac caggatatgc tttagttgga	1080
	tttgaaatta tcaatgatcg aatcccggta ttaaaagcgt ataaagctaa gctaaaacaa	1140
	aattatcaag ttgatcatca gtcgttatca gagattgttt atttagatat cgataaacta	1200
	ttttgtccaa aaaattctga acaaaaatat tataactaaa gtctgacatt tccatgatggc	1260
	tatgttatta ctaagattac ctttgaaaaa aagctgaaca acctagata tgaggcaaca	1320
[0043]	gcaaattttt atgaccatc tacaggagat attgatttaa atgagaagca agtggaatct	1380
	acttttcttc aagcagatta tatttctata aatgttagtg atgatgatgg tgtttacatg	1440
	ccgttaggcg ttatcagcga aacatttttg tctccaatta atagttttga attagaagtt	1500
	gacgagaaat cgaaaatctt aactttaaca tgtaaactctt atttacgaga atatttatta	1560
	gaatctgatt taataaataa agagacaagc ctcatgtctc cgcctaattg ttttatcagt	1620
	aatatcgtag aaaattggaa catagaagcg gataatctag aacctgggt agcaaataac	1680
	aagaatgcat atgtcgatag tacaggcggc atagagggat ctaaagctct atttactcaa	1740
	ggtgatgggg aattttcaca atttattgga gataaattaa aaccaatac agattatatt	1800
	attcaatata ctgtaaaagg aaaacctgct atttatttaa aaaacaaaaa tactggatat	1860
	actatgtacg aagatacaaa cggtagttct gaagaatttc aaactatagc tgtaaattat	1920
	acttcagaaa ctgatccttc acaaacacat ttagttttta aaagtcaaag tggctatgag	1980
	gcttgggggg acaactttat tattctagaa tgtaaggcat ttgaaactcc agaaggtcca	2040

	gaattgataa aatttgatga ttggattagt tttggtacta cttacattag agatgatgta	2100
	cttactatcg atccaagtcg tggaggttat tttagacaat ctcttaaatt agacagctat	2160
	tcaacttata atttgagctt ttccttttct ggattatggg ctaaggttat tataaaaaat	2220
	tcccacggag tagtattggt tgaaaaagta agtcagcagt cttcatacgt agatattagt	2280
	gaaagtttta ctaccacatc aaataaagaa ggatttttta tagaactaac gggcgatagt	2340
	cgtggtggtt ttgggtcggt ccgtgatttt tctatgaagg aaaagtttga acaccacat	2400
	cacgctcacc atcactga	2418
	<210> 14	
	<211> 805	
	<212> PRT	
	<213> 人工	
	<220>	
[0044]	<223> 组氨酸标记的 TIC7473 蛋白的氨基酸序列	
	<400> 14	
	Met Lys Gln Asn Asn Asn Phe Ser Val Arg Ala Leu Pro Ser Phe Ile	
	1 5 10 15	
	Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp Ile	
	20 25 30	
	Phe Asn Met Ile Phe Gly Thr Asp Thr Gly Asp Leu Thr Leu Glu Glu	
	35 40 45	
	Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu Glu	
	50 55 60	
	Gly Ile Ser Gly Asp Leu Ser Glu Ile Ile Ala Gln Gly Asn Leu Asn	
	65 70 75 80	

	Thr Glu Leu Ala Lys Glu Leu Leu Lys Ile Ala Asn Glu Gln Asn Asn	
	85	90 95
	Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met Leu	
	100	105 110
	His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser Asp Val Met Lys	
	115	120 125
	Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys Gln Leu	
	130	135 140
	Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu Ile	
	145	150 155 160
[0045]	Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys Tyr	
	165	170 175
	Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr Leu	
	180	185 190
	Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu Glu	
	195	200 205
	Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn Asp	
	210	215 220
	Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu Ile	
	225	230 235 240
	Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu Leu	
	245	250 255

Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys Val
260 265 270

Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe Leu
275 280 285

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

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

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

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

[0046]

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

Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln Val
370 375 380

Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys Leu
385 390 395 400

Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu Thr
405 410 415

Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys Leu
420 425 430

Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Pro Ser Thr

435	440	445
Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu Gln		
450	455	460
Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr Met		
465	470	475 480
Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser Phe		
485	490	495
Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys Lys		
500	505	510
Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys Glu		
515	520	525
[0047]		
Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val Glu		
530	535	540
Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn Asn		
545	550	555 560
Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys Ala		
565	570	575
Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp Lys		
580	585	590
Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly Lys		
595	600	605
Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr Glu		
610	615	620

	Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn Tyr	
	625	630 635 640
	Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser Gln	
		645 650 655
	Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys Lys	
		660 665 670
	Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp Trp	
		675 680 685
	Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile Asp	
		690 695 700
[0048]	Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser Tyr	
	705	710 715 720
	Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys Val	
		725 730 735
	Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser Gln	
		740 745 750
	Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr Thr Thr Ser Asn	
		755 760 765
	Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly Phe	
		770 775 780
	Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu His His His	
		785 790 795 800

His Ala His His His
805

<210> 15
<211> 2397
<212> DNA
<213> 人工

<220>

<223> 衍生自编码 TIC7472 的日本甲虫类芽孢杆菌菌株 DSC007648 的设计用于植物表达的编码 TIC7472PL 的合成 DNA 序列, 相对于细菌 TIC7472 氨基酸序列, TIC7472PL 具有插入在 2 位的额外的丙氨酸残基

<400> 15

[0049]

atggctaagc agaacaacaa cttcagcgtg cgggcgctcc cgtecttcat cgacgtcttc	60
aacggcatct acgacttcgc cacgggcac caggacatct tcaacatgat ctttgggacg	120
gataccggcg acctcaccct cgaagaagtc cttagaacc aggaactcct gtacgacatc	180
agcgggaagc tggagggcat ctcaggcgac ctgtcggaga tcatcgccca gggcaacctc	240
aacacggagc tcgccaagga actgcttaag atcgccaacg agcagaacaa cgttctgacc	300
gacgtcaaca acaagctcaa cgcgatcaac tccatgctcc acatctacct gccgaagatc	360
accaacatgc tgagcgacgt catgaagcag aactacgcgc tgtcgctcca gatcgagtat	420
ttgagcaagc agtccagga aatctcggac aagctcgacg tgataaacct gaacgtcctc	480
atcaactcca cgtgaccga gatcacaccc gcctaccagc gcatcaagta cgtcaacgag	540
aagttcgacg agctgacgt ggcgaccgag aagaccctgc gggcgaagca aggcagcgag	600
gacattattg cgaacgacac gcttgagaat ctcacggagc tgactgagct ggcgaagtcc	660
gtgaccaaga acgacatgga ctcgttcgag ttctacctac acacctttca tgacgtcctg	720
atcggaaca acctcttcgg gcgcagcgcg ctcaagaccg ctgcggagct tatcacgaag	780
gacgagatca agacctccgg ctcggagatc gggaaggtgt acagcttcct gatcgtgttg	840

	acgtgcctcc aggccaaggc gttcctcacg ctactgcct gccggaagct cttgggcctc	900
	tcagacatcg actacacgaa catcctcaac cagcacctca acgacgagaa gaacgtcttc	960
	cgcgacaaca tccttcgac gctttcgaat aagttcagca acccgaacta cgtgaagacc	1020
	atcggcagcg ataactacgc gaaggtgata ctggaggcgg agcccggcta cgccctggtc	1080
	ggcttcgaga tcattaacga ccgtatcccg gtcctcaagg cgtacaaggc caagctcaag	1140
	cagaactacc aagtggacca ccagtccttc tccgagatcg tgtacctgga catcgacaag	1200
	ctgttctgcc cgaagaacag cgagcagaag tactacacca agtcgctgac cttcccggac	1260
	gggtacgtca tcacaaagat cactttcgag aagaagctga acaacctgcg ctacgaggcg	1320
	acggcgaaact tctacgacct gagcacgggt gacatcgacc tgaacgagaa gcaagtggag	1380
	tccacgttcc tccaggcgga ctacatctct atcaacgtga gcgacgacga cggcgtgtac	1440
	atgccgtgg gcgtcatctc cgagaccttc ctctctccca tcaactcggt cgagcttgaa	1500
[0050]	gtggacgaga aatcgaagat cctgacgtg acctgcaaga gctacctgcg cgagtacctg	1560
	ctggagtccg acctcatcaa caaggagacc agcctgatcg cgccgcctaa tgtgttcac	1620
	agcaacatcg tggagaactg gaacatcgag gccgacaatt tggaacctg ggtcgccaac	1680
	aacaagaacg cctacgtgga cagcacgggc ggcatcgagg gctccaaggc cctgtttacc	1740
	caggagacg gcgagttcag tcagttcatc ggcgacaagc tcaagcccaa cacggactac	1800
	atcatccagt acaccgtcaa agggaagcct gcgatctacc tcaagaacaa gaacaccgga	1860
	tacacgatgt acgaggacac caacggctcc tcggaggagt tccagaccat cgcggtgaac	1920
	tacacctccg agacggacct gtcccagacg cacctcgtgt tcaagtccca gtcaggctac	1980
	gaagcgtggg gtgacaactt tatcatcctg gagtgcagg cgttcgagac gcccgagggc	2040
	ccggaactca tcaagttcga cgactggatc tcattcggca ccacgtacat ccgggacgac	2100
	gtcctcacca tcgacctgct tcgcggcggc tacttccgcc agtccctcaa gctcgactcg	2160
	tacagcacgt acaacctgtc cttctctttc agcgggctgt gggccaaggt catcatcaag	2220

aactcgcatg gcgtcgtcct cttcgagaag gtgtcccagc agagttccta cgtggacatc 2280
aacgagagct tcacgacgac gtccaacaag gagggattct tcatcgagct gaccggcgac 2340
agtcgcggag gcttcgggag cttcggggac ttctccatga aggagaagtt cgagtag 2397

<210> 16

<211> 798

<212> PRT

<213> 人工

<220>

<223> 由合成 DNA 序列编码的 TIC7472PL 的氨基酸序列，其中相对于细菌 TIC7472 氨基酸序列，在 2 位插入了额外的丙氨酸残基

<400> 16

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

[0051]

Ile Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp
20 25 30

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

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

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

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

Asn Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met

100	105	110
Leu His Ile Tyr Leu Pro Lys Ile Thr Asn Met Leu Ser Asp Val Met		
115	120	125
Lys Gln Asn Tyr Ala Leu Ser Leu Gln Ile Glu Tyr Leu Ser Lys Gln		
130	135	140
Leu Gln Glu Ile Ser Asp Lys Leu Asp Val Ile Asn Leu Asn Val Leu		
145	150	155 160
Ile Asn Ser Thr Leu Thr Glu Ile Thr Pro Ala Tyr Gln Arg Ile Lys		
165	170	175
Tyr Val Asn Glu Lys Phe Asp Glu Leu Thr Leu Ala Thr Glu Lys Thr		
180	185	190
[0052]		
Leu Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu		
195	200	205
Glu Asn Leu Thr Glu Leu Thr Glu Leu Ala Lys Ser Val Thr Lys Asn		
210	215	220
Asp Met Asp Ser Phe Glu Phe Tyr Leu His Thr Phe His Asp Val Leu		
225	230	235 240
Ile Gly Asn Asn Leu Phe Gly Arg Ser Ala Leu Lys Thr Ala Ala Glu		
245	250	255
Leu Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys		
260	265	270
Val Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe		
275	280	285

	Leu Thr Leu Thr Ala Cys Arg Lys Leu Leu Gly Leu Ser Asp Ile Asp	
	290	295 300
	Tyr Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe	
	305	310 315 320
	Arg Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser Asn Pro Asn	
		325 330 335
	Tyr Val Lys Thr Ile Gly Ser Asp Asn Tyr Ala Lys Val Ile Leu Glu	
		340 345 350
	Ala Glu Pro Gly Tyr Ala Leu Val Gly Phe Glu Ile Ile Asn Asp Arg	
		355 360 365
[0053]	Ile Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln	
	370	375 380
	Val Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys	
	385	390 395 400
	Leu Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu	
		405 410 415
	Thr Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys	
		420 425 430
	Leu Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Pro Ser	
		435 440 445
	Thr Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu	
		450 455 460

	Gln Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr	
	465	470 475 480
	Met Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser	
		485 490 495
	Phe Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys	
		500 505 510
	Lys Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys	
		515 520 525
	Glu Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val	
		530 535 540
[0054]	Glu Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn	
	545	550 555 560
	Asn Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys	
		565 570 575
	Ala Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp	
		580 585 590
	Lys Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly	
		595 600 605
	Lys Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr	
		610 615 620
	Glu Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn	
	625	630 635 640

Tyr Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser
645 650 655

Gln Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys
660 665 670

Lys Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp
675 680 685

Trp Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile
690 695 700

Asp Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser
705 710 715 720

[0055]

Tyr Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys
725 730 735

Val Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser
740 745 750

Gln Gln Ser Ser Tyr Val Asp Ile Asn Glu Ser Phe Thr Thr Thr Ser
755 760 765

Asn Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly
770 775 780

Phe Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu
785 790 795

<210> 17
<211> 2397
<212> DNA
<213> 人工

<220>

<223> 衍生自编码 TIC7473 的日本甲虫类芽孢杆菌菌株 DSC008493 的设计用于植物表达的编码 TIC7473PL 的合成 DNA 序列, 相对于细菌 TIC7473 氨基酸序列, TIC7473PL 具有插入在 2 位的额外的丙氨酸残基

<400> 17

atggctaagc agaacaacaa cttcagtgtc cgcgccctgc cctccttcat tgacgtcttc	60
aacggcatct acgacttcgc taccggcatc caggacatct tcaacatgat ctttgaacg	120
gataccggcg acctgacgct ggaggaggtg ctcaagaacc aggagctgct ctacgacatc	180
tccgggaagc tggagggcat ctccggcgac ctgtcggaga tcatcgccca gggcaacctc	240
aacacggagt tagccaaaga gctgctcaag atcgccaacg agcagaacaa cgctctgacc	300
gacgtgaaca acaagctcaa cgccatcaac tccatgctcc acatctacct cccgaagatc	360
accaacatgc tctccgacgt catgaagcag aactacgcc tcagcctcca gatcgagtac	420
ctctccaagc agctacagga gatttccgac aagctggacg tcatcaacct gaacgtcttc	480
atcaattcga cgctcaccga gatcaactcc gcgtaccagc gcatcaagta cgtgaacgag	540
aagttcgacg agctgacgct ggcgactgag aagaccctcc gcgccaacaa gggcagcgag	600
gacatcatcg ctaacgacac gctggagaac ctgacggaac tcacggagct ggccaagtcc	660
gtgaccaaga acgacatgga cagcttcgag ttctacctgc atacgttcca cgacgtgctc	720
atcggcaaca acctgttcgg ccgctccgcc ctcaagacgg ccgccgagct gattacaaag	780
gacgagatca agacgagcgg ctccggagatc ggcaaggtgt acagtttcct gatcgtcctg	840
acctgtctcc aggctaaggc gttcctgacg ctaaccgcct gccggaagct cctgggcctc	900
agcgacatcg actacaccaa catcctgaac cagcacctga acgacgagaa gaacgtcttc	960
cgcgacaaca tcctgcccac actgtcgaac aagttctcaa acccgaacta cgtgaagacc	1020
atcgggagcg acaactacgc caagtgatc ctggaggccg agccgggcta cgcgctggtg	1080
ggcttcgaga tcatcaacga ccgcatcccg gtcctcaagg cgtacaaggc gaagctcaag	1140

[0056]

	cagaactacc aagtggacca ccagagccta tccgagatcg tgtacctgga catcgacaaa	1200
	ctgtttctgcc cgaagaactc cgagcagaag tactacacca agtcgctcac ctccccggac	1260
	ggctacgtca tcaccaagat cacgttcgag aagaagctca acaacctgcg ttacgaggcg	1320
	accgccaact tctacgaccc gtccaccggc gacatcgacc ttaacgagaa gcaagtcgag	1380
	agcaccttcc tccaggccga ctacatctcc atcaacgtct cggacgacga cggcgtgtac	1440
	atgccgctgg gcgtcatctc cgagaccttc ctgagcccgga tcaacagctt cgagctggag	1500
	gtggacgaga agtccaagat cctgacccta acgtgcaaga gctacctcag ggagtacctc	1560
	ctggagtccg acctcatcaa caaggagacg agcctgatcg cgctccaaa cgtcttcac	1620
	agcaacattg tggagaactg gaacatcgag gcggacaacc tagaaccttg ggtggcgaac	1680
	aacaagaacg cctacgtgga cagcaccggc ggcatcgagg gcagcaaagc actgttcact	1740
	cagggtgacg gcgagttctc gcagttcatc ggcgacaagc tcaagccaaa caccgactac	1800
[0057]	atcatccagt acacggtaaa gggcaagcct gctatctacc tcaagaacaa gaacaccggc	1860
	tacacgatgt acgaggacac gaacgggtcc agcgaggagt tccagaccat cgccgtgaac	1920
	tacaccagcg agaccgaccc gtcccagacc cacctcgtgt tcaagtcgca gagcgggtac	1980
	gaggcttggg gagataactt cattatcttg gagtgcagg cgttcgagac gccggaaggc	2040
	ccggagctca tcaagttcga cgactggatc tcgttcggga ccacctacat ccgcgacgac	2100
	gtgctcacca tcgacccgag ccgtggcggc tacttccgcc agtccttgaa actcgactcg	2160
	tactcgacgt acaacctctc gttcagcttc tcgggcctct gggctaaggt catcatcaag	2220
	aactcccacg gcgtcgctct gttcgagaag gtgtcgcagc agagtctgta cgtggacatc	2280
	tcggagtcct tcaccaccac cagcaacaag gagggcttct ttatcgagct cacgggcgac	2340
	tcgcgcggcg gcttcggctc gttccgggac tttagtatga aggagaagtt cgagtag	2397
	<210> 18	
	<211> 798	

<212> PRT

<213> 人工

<220>

<223> 由合成 DNA 序列编码的 TIC7473PL 的氨基酸序列，其中相对于细菌 TIC7473 氨基酸序列，在位置 2 插入了额外的丙氨酸残基。

<400> 18

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

Ile Asp Val Phe Asn Gly Ile Tyr Asp Phe Ala Thr Gly Ile Gln Asp
20 25 30

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

[0058] Glu Val Leu Lys Asn Gln Glu Leu Leu Tyr Asp Ile Ser Gly Lys Leu
50 55 60

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

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

Asn Val Leu Thr Asp Val Asn Asn Lys Leu Asn Ala Ile Asn Ser Met
100 105 110

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

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

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

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

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

Leu Arg Ala Lys Gln Gly Ser Glu Asp Ile Ile Ala Asn Asp Thr Leu
195 200 205

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

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

[0059]

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

Leu Ile Thr Lys Asp Glu Ile Lys Thr Ser Gly Ser Glu Ile Gly Lys
260 265 270

Val Tyr Ser Phe Leu Ile Val Leu Thr Cys Leu Gln Ala Lys Ala Phe
275 280 285

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

Tyr Thr Asn Ile Leu Asn Gln His Leu Asn Asp Glu Lys Asn Val Phe
305 310 315 320

Arg Asp Asn Ile Leu Pro Thr Leu Ser Asn Lys Phe Ser Asn Pro Asn

325	330	335
Tyr Val Lys Thr Ile Gly Ser Asp Asn Tyr Ala Lys Val Ile Leu Glu 340	345	350
Ala Glu Pro Gly Tyr Ala Leu Val Gly Phe Glu Ile Ile Asn Asp Arg 355	360	365
Ile Pro Val Leu Lys Ala Tyr Lys Ala Lys Leu Lys Gln Asn Tyr Gln 370	375	380
Val Asp His Gln Ser Leu Ser Glu Ile Val Tyr Leu Asp Ile Asp Lys 385	390	395 400
Leu Phe Cys Pro Lys Asn Ser Glu Gln Lys Tyr Tyr Thr Lys Ser Leu 405	410	415
[0060]		
Thr Phe Pro Asp Gly Tyr Val Ile Thr Lys Ile Thr Phe Glu Lys Lys 420	425	430
Leu Asn Asn Leu Arg Tyr Glu Ala Thr Ala Asn Phe Tyr Asp Pro Ser 435	440	445
Thr Gly Asp Ile Asp Leu Asn Glu Lys Gln Val Glu Ser Thr Phe Leu 450	455	460
Gln Ala Asp Tyr Ile Ser Ile Asn Val Ser Asp Asp Asp Gly Val Tyr 465	470	475 480
Met Pro Leu Gly Val Ile Ser Glu Thr Phe Leu Ser Pro Ile Asn Ser 485	490	495
Phe Glu Leu Glu Val Asp Glu Lys Ser Lys Ile Leu Thr Leu Thr Cys 500	505	510

	Lys Ser Tyr Leu Arg Glu Tyr Leu Leu Glu Ser Asp Leu Ile Asn Lys	
	515	520 525
	Glu Thr Ser Leu Ile Ala Pro Pro Asn Val Phe Ile Ser Asn Ile Val	
	530	535 540
	Glu Asn Trp Asn Ile Glu Ala Asp Asn Leu Glu Pro Trp Val Ala Asn	
	545	550 555 560
	Asn Lys Asn Ala Tyr Val Asp Ser Thr Gly Gly Ile Glu Gly Ser Lys	
	565	570 575
	Ala Leu Phe Thr Gln Gly Asp Gly Glu Phe Ser Gln Phe Ile Gly Asp	
	580	585 590
[0061]	Lys Leu Lys Pro Asn Thr Asp Tyr Ile Ile Gln Tyr Thr Val Lys Gly	
	595	600 605
	Lys Pro Ala Ile Tyr Leu Lys Asn Lys Asn Thr Gly Tyr Thr Met Tyr	
	610	615 620
	Glu Asp Thr Asn Gly Ser Ser Glu Glu Phe Gln Thr Ile Ala Val Asn	
	625	630 635 640
	Tyr Thr Ser Glu Thr Asp Pro Ser Gln Thr His Leu Val Phe Lys Ser	
	645	650 655
	Gln Ser Gly Tyr Glu Ala Trp Gly Asp Asn Phe Ile Ile Leu Glu Cys	
	660	665 670
	Lys Ala Phe Glu Thr Pro Glu Gly Pro Glu Leu Ile Lys Phe Asp Asp	
	675	680 685

Trp Ile Ser Phe Gly Thr Thr Tyr Ile Arg Asp Asp Val Leu Thr Ile
690 695 700

Asp Pro Ser Arg Gly Gly Tyr Phe Arg Gln Ser Leu Lys Leu Asp Ser
705 710 715 720

Tyr Ser Thr Tyr Asn Leu Ser Phe Ser Phe Ser Gly Leu Trp Ala Lys
725 730 735

[0062] Val Ile Ile Lys Asn Ser His Gly Val Val Leu Phe Glu Lys Val Ser
740 745 750

Gln Gln Ser Ser Tyr Val Asp Ile Ser Glu Ser Phe Thr Thr Thr Ser
755 760 765

Asn Lys Glu Gly Phe Phe Ile Glu Leu Thr Gly Asp Ser Arg Gly Gly
770 775 780

Phe Gly Ser Phe Arg Asp Phe Ser Met Lys Glu Lys Phe Glu
785 790 795