



(12) 发明专利申请

(10) 申请公布号 CN 105368799 A

(43) 申请公布日 2016. 03. 02

(21) 申请号 201510740118. 2

C12N 15/82(2006. 01)

(22) 申请日 2009. 04. 10

A01H 5/00(2006. 01)

(30) 优先权数据

08154481. 9 2008. 04. 14 EP

61/124, 082 2008. 04. 14 US

(62) 分案原申请数据

200980113723. X 2009. 04. 10

(71) 申请人 拜耳作物科学公司

地址 比利时迭戈姆

申请人 拜耳农科股份公司

(72) 发明人 M·布什 K·菲希尔 B·兰伯

A·萨利安德

(74) 专利代理机构 北京北翔知识产权代理有限

公司 11285

代理人 张广育 姜建成

(51) Int. Cl.

C12N 9/02(2006. 01)

C12N 15/53(2006. 01)

C12N 15/62(2006. 01)

C07K 19/00(2006. 01)

权利要求书1页 说明书27页

序列表31页 附图4页

(54) 发明名称

新的突变羟基苯基丙酮酸双加氧酶, DNA 序列
和耐受 HPPD 抑制剂除草剂的植物分离

(57) 摘要

本发明涉及编码突变的羟基苯基丙酮酸双加氧酶 (HPPD) 的核酸序列、包含该序列作为编码序列的嵌合基因以及其在获得耐受 HPPD 抑制剂除草剂的植物中的应用。

1. 突变的羟基苯基丙酮酸双加氧酶 (HPPD), 其保留催化对羟基苯基丙酮酸 (HPP) 转化为尿黑酸的性质, 且对三酮 HPPD 抑制剂的敏感性低于原始的未突变 HPPD, 其特征在于其在 SEQ ID NO:2 所示假单胞菌 HPPD 的氨基酸序列第 336 位的氨基酸甘氨酸上包含选自下组的突变: Gly336His、Gly336Met、Gly336Phe 和 Gly336Cys。

2. 如权利要求 1 所述的突变 HPPD, 其特征在于, 所述突变 HPPD 包含第二突变。

3. 如权利要求 2 所述的突变 HPPD, 其特征在于, 所述第二突变氨基酸选自以下氨基酸: SEQ ID NO:2 所示假单胞菌 HPPD 序列的 Pro215、Gly298、Gly332、Phe333、Gly334 和 Asn337。

4. 如权利要求 1-3 中任一项所述的突变 HPPD, 所述三酮 HPPD 抑制剂选自: 特波三酮、甲基磺草酮和磺草酮。

5. 如权利要求 4 所述的突变 HPPD, 所述三酮 HPPD 抑制剂是特波三酮。

6. 编码如权利要求 1-5 中任一项所述的突变 HPPD 的核酸序列。

7. 在 5' 和任选在 3' 位置包含编码序列以及异源调控元件的嵌合基因, 其能在宿主生物体中发挥作用, 其特征在于所述编码序列包含至少一个如权利要求 6 所述的核酸序列。

8. 如权利要求 7 所述的嵌合基因, 其特征在于, 其在编码突变 HPPD 的核酸序列的 5' 位置包含编码植物转运肽的核酸序列, 其中该序列位于启动子区域和编码突变 HPPD 的序列之间, 从而允许转运肽 / 突变 HPPD 融合蛋白的表达。

9. 转运肽 / 突变 HPPD 融合蛋白, 其中所述突变 HPPD 如权利要求 1-5 中任一项定义。

10. 用于转化宿主生物体的克隆和 / 或表达载体, 其特征在于其包含至少一个如权利要求 7-8 中任一项所述的嵌合基因。

11. 获得耐受三酮 HPPD 抑制剂的植物的方法, 其特征在于用如权利要求 7-8 中任一项所述的嵌合基因转化所述植物。

12. 如权利要求 11 所述的获得耐受三酮 HPPD 抑制剂的植物的方法, 其特征在于, 用在该植物中发挥作用从而允许 PDH(预苯酸脱氢酶) 酶过表达的第二基因同时或相继进一步转化所述植物。

13. 在包含至少含有如权利要求 6 所述核酸序列或如权利要求 7-8 中任一项所述嵌合基因的种子或植物的区域或田地中控制杂草的方法, 该方法包括对所述区域或田地施用对所述杂草有毒剂量的三酮 HPPD 抑制剂除草剂, 而不显著影响包含如权利要求 6 所述核酸序列或权利要求 7-8 中任一项所述嵌合基因的种子或植物。

14. 获得油或粗碾谷物的方法, 包括种植至少包含如权利要求 6 所述核酸序列或如权利要求 7-8 中任一项所述嵌合基因的转化植物, 任选用三酮 HPPD 抑制剂除草剂处理该植物, 收获谷物, 碾磨谷物以制备粗碾谷物并任选提取油。

15. 如权利要求 11-14 任一项所述的方法, 其中所述三酮 HPPD 抑制剂选自: 特波三酮、甲基磺草酮和磺草酮。

16. 如权利要求 15 所述的方法, 其中所述三酮 HPPD 抑制剂为特波三酮。

17. 在 SEQ ID NO:2 所示假单胞菌 HPPD 的氨基酸序列第 336 位的氨基酸甘氨酸上发生突变的 HPPD 在赋予植物耐受三酮 HPPD 抑制剂中的应用, 其特征在于所述 HPPD 突变选自: Gly336His、Gly336Met、Gly336Phe 和 Gly336Cys。

18. 如权利要求 17 所述突变 HPPD 的应用, 其特征在于, 所述 HPPD 抑制剂选自特波三酮、甲基磺草酮和磺草酮。

新的突变羟基苯基丙酮酸双加氧酶, DNA 序列和耐受 HPPD 抑制剂除草剂的植物分离

[0001] 本申请为分案申请,其原申请的申请日为2009年4月10日,申请号为200980113723.X,名称为“新的突变羟基苯基丙酮酸双加氧酶, DNA 序列和耐受 HPPD 抑制剂除草剂的植物分离”。本发明涉及编码突变的羟基苯基丙酮酸双加氧酶 (hydroxyphenylpyruvate dioxygenase) (HPPD) 的核酸序列、包含该序列作为编码序列的嵌合基因以及其在获得耐受 HPPD 抑制剂除草剂的植物中的用途。

[0002] 羟基苯基丙酮酸双加氧酶 (HPPD ;EC 1.13.11.27) 是一种酶,其催化酪氨酸降解产物 - 对羟基苯基丙酮酸 (HPP) 转化为植物中生育酚和质体醌的前体 - 尿黑酸 (HG) 的反应 (Crouch N.P 等,1997 ;Fritze 等,2004)。生育酚起到膜相关抗氧化剂的作用。质体醌首先是 PSII 和细胞色素 b6/f 复合物之间的电子载体,然后是参与类胡萝卜素生物合成的八氢番茄红素去饱和酶的氧化还原辅因子。

[0003] 大多数植物通过预苯氨酸 (arrogenate) 合成酪氨酸 (Abou-Zeid 等 1995 ;Bonner 等,1995 ;Byng 等,1981 ;Connely 和 Conn 1986 ;Gaines 等,1982)。在这些植物中,HPP 仅来自于酪氨酸降解。另一方面,在如酵母酿酒酵母 (*Sacharomyces cerevisiae*) 或细菌大肠杆菌 (*Escherichia coli*) 等生物体中,HPP 是酪氨酸前体,由预苯酸脱氢酶 (此后称为 PDH) 的作用合成,该酶将预苯酸转化为 HPP (Lingens 等,1967 ;Sampathkumar 和 Morrisson 1982)。因此在这些生物体中,HPP 的产生直接与芳香族氨基酸的生物合成途径 (莽草酸途径) 而非酪氨酸降解途径相关。

[0004] 抑制 HPPD 导致光合作用的解偶联、辅助捕光色素缺乏,最重要的是,由于缺乏通常由类胡萝卜素提供的光保护作用,紫外线辐射和活性氧中间体导致叶绿素破坏 (Norris 等 1995)。对光合作用活性组织进行光漂白会导致生长抑制和植物死亡。

[0005] 一些抑制 HPPD 和与所述酶特异性结合以抑制 HPP 转化为尿黑酸的分子已证实是非常有效的选择性除草剂。

[0006] 市场上最易买到的 HPPD 抑制剂除草剂属于以下四个化学家族之一:

[0007] 1) 三酮类,例如,磺草酮 (sulcotrione) (即 2-[2-氯-4-(甲磺酰基)苯甲酰基]-1,3-环己二酮)、甲基磺草酮 (mesotrione) (即 2-[4-(甲磺酰基)-2-硝基苯甲酰基]-1,3-环己二酮)、特波三酮 (tembotrione) (即 2-[2-氯-4-(甲磺酰基)-3-[(2,2,2-三氟乙氧基)甲基]苯甲酰基]-1,3-环己二酮);

[0008] 2) 二酮腈 (diketonitrile) 类,例如,2-氰基-3-环丙基-1-(2-甲磺酰基-4-三氟甲基苯基)-丙烷-1,3-二酮和 2-氰基-1-[4-(甲磺酰基)-2-三氟甲基苯基]-3-(1-甲基环丙基)丙烷-1,3-二酮;

[0009] 3) 异噁唑类,例如,异噁氟草 (isoxaflutole) (即 5-环丙基-4-异噁唑基 [2-(甲磺酰基)-4-(三氟甲基)苯基]甲酮)。在植物中,异噁氟草在 DKN 中快速代谢,后者是一种表现出 HPPD 抑制剂性质的二酮腈化合物;以及

[0010] 4) 吡唑特 (pyrazolate) 类,例如,苯吡唑草酮 (topramezone) (即 [3-(4,5-二氢-3-异噁唑基)-2-甲基-4-(甲磺酰基)苯基] (5-羟基-1-甲基-1H-吡唑-4-基) 甲

酮)和吡拉塞佛托(pyrasulfotole)(5-羟基-1,3-二甲基吡唑-4-基(2-甲磺酰-4-三氟甲基苯基)甲酮)。

[0011] 这些 HPPD 抑制除草剂可用于表现出代谢耐受性的作物如玉米 (*Zea mays*) 中的草和 / 或阔叶杂草,这些除草剂在其中可快速降解 (Schulz 等,1993;Mitchell 等,2001;Garcia 等,2000;Pallett 等,2001)。为了拓展这些 HPPD 抑制除草剂的范围,已作出努力以赋予植物,尤其是不具有或具有未充分发挥的代谢耐受性的植物对它们的农业水平耐受性。

[0012] 除试图绕过 HPPD 介导的尿黑酸产生 (US 6,812,010) 外,已进行了该敏感酶的过表达以在植物中产生大量对除草剂充足的靶酶 (WO96/38567)。HPPD 的过表达引起了对异噁氟草 (IFT) 的二酮腈衍生物 (DKN) 更好的萌发前耐受性,但该耐受性不足以作为萌发后处理的耐受性 (Matringe 等,2005)。

[0013] 另一种策略是对 HPPD 进行突变以获得相对突变前的天然 HPPD 对 HPPD 抑制剂敏感性较低的靶酶,而同时保留其催化 HPPD 转化为尿黑酸的性质。

[0014] 该策略已成功应用于生产耐受两种属于二酮腈家族的 HPPD 抑制除草剂 (WO 99/24585)2-氰基-3-环丙基-1-(2-甲磺酰基-4-三氟甲基苯基)-丙烷-1,3-二酮和 2-氰基-1-[4-甲磺酰基-2-三氟甲基苯基]-3-(1-甲基环丙基)丙烷-1,3-二酮的植物 (EP496630)。鉴定 Pro215Leu、Gly336Glu、Gly336I1、特别是 Gly336Trp (突变氨基酸的位置根据 SEQ ID NO:2 的假单胞菌 (*Pseudomonas*) HPPD 标出) 是引起对二酮腈除草剂萌发前处理耐受性增强但并不造成酶活性改变的突变。

[0015] 最近,已证实将假单胞菌 HPPD 基因导入烟草和大豆的质体基因组远比核转化有效,赋予对异噁氟草的萌发后使用相同的耐受性 (Dufourmantel 等,2007)。

[0016] 在 WO 04/024928 中,发明人设法通过增强 HPP 前体进入植物细胞的流量来增强植物细胞中异戊烯苯醌的生物合成 (例如,质体醌和生育酚的合成)。这通过过表达 PDH 酶将所述前体的合成与“莽草酸”途径进行连接而完成。他们还注意到,用编码 PDH 酶的基因转化植物有可能增强所述植物对 HPPD 抑制剂的耐受性。

[0017] 尽管对二酮腈除草剂表现出耐受性的植物的开发已经获得了这些成功,仍然有必要开发和 / 或改进对 HPPD 抑制剂的耐受性系统,特别是对属于三酮类 (例如,磺草酮、甲基磺草酮和特波三酮) 和吡唑特类 (例如,苯吡唑草酮和吡拉塞佛托) 的 HPPD 抑制剂。

[0018] 因此,本发明涉及新的突变 HPPD 酶,其保留催化对羟基苯基丙酮酸 (HPP) 转化为尿黑酸的性质,且对 HPPD 抑制剂的敏感性低于原始的未突变 HPPD,其特征在于它们在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位 (天然 HPPD 中为氨基酸甘氨酸) 包含一个选自以下突变的突变:Gly336Arg、Gly336His、Gly336Met、Gly336Phe、Gly336Asn、Gly336Cys、Gly336Val、Gly336Trp、Gly336Glu 和 Gly336Asp。

[0019] 在一个具体的实施方式中,SEQ ID NO:2 所示假单胞菌 HPPD 在第 336 位的突变选自以下突变:Gly336Arg、Gly336His、Gly336Met、Gly336Phe、Gly336Asn、Gly336Cys 和 Gly336Val,条件是突变的 HPPD 不是双重突变体 Gly334Ala-Gly336Arg (位置根据 SEQ ID NO:2 所示假单胞菌 HPPD 给出)。

[0020] 在一个更具体的实施方式中,SEQ ID NO:2 所示假单胞菌 HPPD 在第 336 位的突变选自以下突变:Gly336His、Gly336Met、Gly336Cys 和 Gly336Phe。

[0021] 在另一个具体的实施方式中, HPPD 酶来自植物, 特别是来自拟南芥 (*Arabidopsis thaliana*), SEQ ID NO:4 所示拟南芥 HPPD 的氨基酸序列在第 422 位 (即 SEQ ID NO:2 的假单胞菌 HPPD 的第 336 位) 的甘氨酸包含一个选自以下突变的突变: Gly336Arg、Gly336His、Gly336Met、Gly336Phe、Gly336Asn、Gly336Cys、Gly336Val、Gly336Trp、Gly336Glu 和 Gly336Asp。

[0022] 在一个更具体的实施方式中, SEQ ID NO:4 所示拟南芥 HPPD 在第 422 位 (即 SEQ ID NO:2 的假单胞菌 HPPD 的第 336 位) 的突变选自以下突变: Gly336His、Gly336Asn、Gly336Cys 和 Gly336Val, 且突变的 HPPD 来自植物, 特别是来自拟南芥。应注意, SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位是 SEQ ID NO:4 所示拟南芥 HPPD 的第 422 位。

[0023] 在一个具体的实施方式中, 本发明的突变 HPPD 对异噁唑类、二酮腈类、三酮类、吡唑特类的 HPPD 抑制剂除草剂的敏感性低于原始的未突变 HPPD。

[0024] 在一个具体的实施方式中, 本发明的突变 HPPD 对选自以下的 HPPD 抑制剂除草剂的敏感性低于原始的未突变 HPPD: 异噁氟草、特波三酮、甲基磺草酮、磺草酮、吡拉塞佛托、苯吡唑草酮、2-氰基-3-环丙基-1-(2-SO₂CH₃-4-CF₃苯基)-丙烷-1,3-二酮和 2-氰基-3-环丙基-1-(2-SO₂CH₃-4-2,3Cl₂苯基)-丙烷-1,3-二酮。

[0025] 在另一个具体的实施方式中, 本发明的突变 HPPD 对以下 HPPD 抑制剂的敏感性低于原始的未突变 HPPD: 三酮类 (称为三酮 HPPD 抑制剂), 如特波三酮、磺草酮和甲基磺草酮, 特别是特波三酮, 或吡唑特类 (称为吡唑特 HPPD 抑制剂), 如吡拉塞佛托和苯吡唑草酮。

[0026] 在一个更具体的实施方式中, 本发明的突变 HPPD 对选自特波三酮、磺草酮和甲基磺草酮, 特别是特波三酮的三酮 HPPD 抑制剂的敏感性低于原始的未突变 HPPD。

[0027] 在另一个具体的实施方式中, 除 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位的氨基酸甘氨酸上的第一突变外, 本发明的突变 HPPD 还包含第二突变。

[0028] 在一个更具体的实施方式中, 第二突变氨基酸选自以下 SEQ ID NO:2 所示假单胞菌 HPPD 序列的氨基酸: Pro215、Gly298、Gly332、Phe333、Gly334 和 Asn337。

[0029] 本发明还提供突变的 HPPD 酶, 其保留催化对羟基苯基丙酮酸 (HPP) 转化为尿黑酸的性质, 且对以下 HPPD 抑制剂的敏感性低于原始的未突变 HPPD: 三酮类, 如特波三酮、磺草酮和甲基磺草酮, 或吡唑特类, 如吡拉塞佛托和苯吡唑草酮, 其特征在于它们在 SEQ ID NO:2 所示假单胞菌 HPPD 第 336 位的氨基酸甘氨酸包含突变, 以及提供使用这样的酶使植物对这些 HPPD 抑制剂具有耐受性, 即, 将三酮类或吡唑特类除草剂施用于表达这样突变体酶的植物的, 而植物通过在其基因组中包含编码在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位突变的某些 HPPD 酶的基因而对这样的三酮类或吡唑特类 HPPD 抑制剂具有耐受性的过程。

[0030] 在本发明一个具体的实施方式中, 突变的 HPPD 酶对三酮类如特波三酮、磺草酮和甲基磺草酮的 HPPD 抑制剂的敏感性低于原始的未突变 HPPD, 突变根据选自以下突变的突变, 发生在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位: Gly336Arg、Gly336Asp、Gly336Glu、Gly336His、Gly336Met、Gly336Phe、Gly336Trp、Gly336Asn、Gly336Cys 和 Gly336Val。

[0031] 在本发明一个具体的实施方式中, 突变的 HPPD 酶对三酮类如特波三酮、磺草酮和甲基磺草酮的 HPPD 抑制剂的敏感性低于原始的未突变 HPPD, 突变根据选自以下突变的突变, 发生在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位: Gly336His、Gly336Met、Gly336Phe

和 Gly336Cys。

[0032] 现有技术已对若干 HPPD 及其初级序列进行了描述,特别是以下 HPPD:如假单胞菌等细菌 (Rüetschi 等, Eur. J. Biochem., 205, 459-466, 1992, WO 96/38567)、如拟南芥 (WO 96/38567, Genebank AF047834)、胡萝卜 (WO 96/38567, Genebank 87257)、燕麦 (*Avena sativa*) (WO 02/046387)、小麦 (WO 02/046387)、臂形草 (*Brachiaria platyphylla*) (WO 02/046387)、蒺藜草 (*Cenchrus echinatus*) (WO 02/046387)、刚性黑麦草 (*Lolium rigidum*) (WO 02/046387)、苇状羊茅 (*Festuca arundinacea*) (WO 02/046387)、大狗尾草 (*Setaria faberi*) (WO 02/046387)、牛筋草 (*Eleusine indica*) (WO 02/046387) 和高粱 (WO 02/046387) 等植物、球孢子菌 (*Coccicoides*) (Genebank COITRP) 或如小鼠或猪等哺乳动物的。显示的参考文献中公开的相应序列通过引用纳入本文。

[0033] 通过比对这些已知序列,采用本领域的惯例手段,例如, Thompson, J. D. 等描述的方法 (CLUSTAL W: “通过序列加权、位点特异性空位罚分和加权矩阵选择来提高渐进的多序列比对的敏感性” (improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice.) *Nucleic Acids Research*, 22 ;4673-4680, 1994), 进入这些通过互联网可获得的序列比对的计算机程序,技术人员能够确定与参比序列的序列同源性并找到关键氨基酸或确定共有区域。

[0034] 以本发明为例,参比序列是假单胞菌序列,除非特别说明,对具体氨基酸位置的定义和说明所有均根据 SEQ ID NO:2 所示初级假单胞菌 HPPD 序列作出。附图 1 显示现有技术中描述的若干 HPPD 序列的比对;这些序列的比对以假单胞菌 HPPD 序列作为参比序列,包括以下 HPPD 序列:阿维链霉菌 (*Streptomyces avermitilis*) (Genebank SAV11864)、胡萝卜 (*Daucus carota*) (Genebank DCU 87257)、拟南芥 (Genebank AF047834)、玉米、大麦 (*Hordeum vulgare*) (Genebank HVAJ693)、小麦叶枯病菌 (*Mycosphaerella graminicola*) (Genebank AF038152)、粗球孢子菌 (*Coccicoides immitis*) (Genebank COITRP) 和小鼠 (*Mus musculus*) (Genebank MU54HD)。该图提供假单胞菌序列的氨基酸编号,还有这些序列共有的氨基酸,其中这些氨基酸由星号表示。根据这样的比对,从假单胞菌氨基酸的位置和性质的定义不难鉴定另一个 HPPD 序列中的相应氨基酸的位置。图 1 显示可通过比对不同植物、哺乳动物和细菌来源的序列做到这点,表明技术人员熟知的该比对方法通用于任何其它序列。专利申请 WO 97/49816 中也描述了不同 HPPD 序列的比对。

[0035] 在 W099/24585 中,对假单胞菌 HPPD 单体三级结构的分析显示 HPPD 的 C 末端部分存在,其是酶的活性位点所在,通过保证酶稳定性及其寡聚化(假单胞菌 HPPD 是四聚体,植物 HPPD 是二聚体)的连接肽与其 N 末端部分连接。该结构由研究晶体 X 光衍射的常规方法获得。连接肽可能确定酶 C 末端部分的 N 末端,其中以假单胞菌为例,所述连接肽位于第 145 位和第 157 位的氨基酸之间(参见图 1)。在附图 1 所示的序列比对中显示的所有序列中均会注意到假单胞菌序列中位于第 161 位和第 162 位的两个氨基酸 (D = Asp161, H = His162)。根据假单胞菌 HPPD,由此可能确定连接肽位于氨基酸 Asp161 上游约 5 到 15 个氨基酸之间。

[0036] 根据本发明,应理解“突变的 HPPD”为 HPPD 初级序列的至少一个氨基酸被另一个氨基酸取代。表述“突变氨基酸”以下用于表示被另一个氨基酸取代的氨基酸,从而表示蛋

白质初级序列中的突变位点。

[0037] 根据本发明,突变发生在 SEQ ID NO:2 所示假单胞菌序列第 336 位的氨基酸甘氨酸上,其几乎是所有已鉴定的 HPPD 序列共有的。在至今已知的 240 个 HPPD 序列中,有 238 个在第 336 位包含甘氨酸,仅有聚球菌属 (*Synechococcus* sp.)JA-3-3Ab(Acc-No Q2JX04) 和聚球菌属 JA-2-3B'a(2-13)(Acc-No Q2JPN8) 的 HPPD 序列在该位置上是丙氨酸。Gly336 是发现于大多数 HPPD 序列的共有序列“Gly-Phe-Gly-X-Gly-Asn-Phe”的一部分,其中在各种来源的 HPPD 中,X 可以是 20 种氨基酸中的任何一种,这可能通过序列比对方法鉴定任何来源 HPPD 中的 Gly336 而没有任何困难。

[0038] 作为例子,假单胞菌序列的 Gly336 是 SEQ ID NO:4 所示拟南芥序列的 Gly422(见图 1),但本文提到的 Gly 是 SEQ ID NO:2 所示假单胞菌序列的参比位置第 336 位的 Gly(除非特别说明),即使本发明的突变可发生在任何有用的 HPPD 酶中而并非必然在假单胞菌 HPPD 中。

[0039] HPPD 的酶活性可通过任何方法测定,从而能测定 HPP 或 O₂底物量的减少、或测定任何来源于酶反应的产物即尿黑酸或 CO₂的积累。特别是,HPPD 活性可通过 Garcia 等(1997) 或 Garcia 等(1999) 中描述的方法测定,二者通过引用纳入本文。

[0040] 根据本发明,三酮类的 HPPD 抑制剂(或三酮 HPPD 抑制剂)指具有三酮骨架的 HPPD 抑制剂。这样的三酮 HPPD 抑制剂的一个例子是磺草酮(即 2-[2-氯-4-(甲磺酰基)苯甲酰基]-1,3-环己二酮)、甲基磺草酮(即 2-[4-(甲磺酰基)-2-硝基苯甲酰基]-1,3-环己二酮)和特波三酮(即 2-[2-氯-4-(甲磺酰基)-3-[2,2,2-三-氟乙氧基甲基]苯甲酰基]-1,3-环己二酮)。

[0041] 根据本发明,吡唑特类的 HPPD 抑制剂(或吡唑特 HPPD 抑制剂)指具有吡唑基的 HPPD 抑制剂。这样的吡唑特 HPPD 抑制剂的一个例子是苯吡唑草酮(即 [3-(4,5-二氢-3-异噁唑基)-2-甲基-4-(甲磺酰基)苯基](5-羟基-1-甲基-1H-吡唑-4-基)甲酮)和吡拉塞佛托[(5-羟基-1,3-二甲基吡唑-4-基(2-甲磺酰-4-三氟甲基)甲酮)。

[0042] 在本发明的另一个实施方式中,除 Gly336 的突变外,HPPD 在第二氨基酸位置也发生了突变。该第二突变的存在可进一步增强对与第一突变赋予耐受性的同一 HPPD 抑制剂除草剂的耐受性,或可赋予对另一 HPPD 抑制剂除草剂的耐受性。WO 99/24585 描述了这些突变赋予对 HPPD 抑制剂特别是对二酮腈类和异噁唑类的耐受性的例子。

[0043] 在本发明的一个具体的实施方式中,第二突变的氨基酸选自以下 SEQ ID NO:2 所示假单胞菌序列的参比氨基酸:Pro215、Gly332、Phe333、Gly334 和 Asn337,以及假单胞菌序列中的 Gly298(最后这个氨基酸在其它 HPPD 中没有对应物,见图 1)。

[0044] 在本发明的一个实施方式中,第二突变的氨基酸是 SEQ ID NO:2 所示假单胞菌序列的 Pro215,该突变特别是 Pro215Leu。

[0045] 本发明也涉及一种核酸序列,特别是分离的 DNA,其编码如上所述的突变的 HPPD。

[0046] 本发明也涉及一种编码突变的 HPPD 酶的核酸序列,其中突变的 HPPD 酶保留催化对羟基苯基丙酮酸(HPP)转化为尿黑酸的性质,且对以下 HPPD 抑制剂的敏感性低于原始的未突变 HPPD:三酮类,如特波三酮、磺草酮和甲基磺草酮,或吡唑特类,如吡拉塞佛托和苯吡唑草酮,其特征在于其在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位氨基酸甘氨酸包含突变。

[0047] 在一个更具体的实施方式中,本发明的核酸序列编码突变的 HPPD 酶,其对三酮类的 HPPD 抑制剂如特波三酮、磺草酮和甲基磺草酮的敏感性低于原始的未突变 HPPD,且其中 HPPD 根据选自以下的突变,在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位发生突变:Gly336Arg、Gly336Asp、Gly336Glu、Gly336His、Gly336Met、Gly336Phe、Gly336Trp、Gly336Asn、Gly336Cys 和 Gly336Val。

[0048] 在一个更为具体的实施方式中,本发明的核酸序列编码突变的 HPPD 酶,其对三酮类的 HPPD 抑制剂如特波三酮、磺草酮和甲基磺草酮的敏感性低于原始的未突变 HPPD,且其中 HPPD 根据选自以下的突变,在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位发生突变:Gly336His、Gly336Met、Gly336Phe 和 Gly336Cys。

[0049] 根据本发明,“核酸序列”应理解为 DNA 或 RNA 型的核苷酸序列,优选 DNA 型,特别是双链,无论是天然或合成来源,特别是其中编码本发明突变 HPPD 的密码子已根据要在其中表达的宿主生物体进行了优化(例如,通过采用相比原始宿主在该宿主生物体或该宿主生物体所属组的密码子使用表中更优选或最优选的密码子进行密码子取代),这些优化方法为本领域技术人员所熟知。

[0050] 本文使用的“分离的 DNA”指非天然产生或不再存在于其原始存在的天然环境中的 DNA,例如,嵌合基因中与其它调控元件相关的 DNA 编码序列、转移到另一个宿主细胞如植物细胞中的 DNA、或相比任何已知天然产生的 DNA 具有不同核苷酸序列的人工合成的 DNA。

[0051] 编码根据本发明将发生突变的原始未突变 HPPD 的序列可来自任何来源。其特别可来自细菌来源。有利例子是假单胞菌属的细菌,例如荧光假单胞菌 (*Pseudomonas fluorescens*),或集胞藻 (*Synechocystis*) 属的蓝藻 (cyanobacteria)。所述序列也可来自植物来源,特别是来自双子叶植物、伞状花科植物或单子叶植物。有利例子是植物,如烟草、拟南芥、胡萝卜、玉米、小麦、大麦、燕麦、小麦、臂形草、蒺藜草、刚性黑麦草、苇状羊茅、大狗尾草、牛筋草和高粱。之前引用的参考文献中描述了所述编码序列以及分离和克隆所述编码序列的方法,这些参考文献的内容通过引用纳入本文。

[0052] 在本发明一个具体的实施方式中,所述 HPPD 来自细菌来源,特别是来自假单胞菌属,更特别的是来自荧光假单胞菌,或来自植物来源,特别是来自拟南芥。

[0053] 为本发明目的而突变的 HPPD 可为任何天然产生的 HPPD,或其任何活性片段或其任何变体,其中一些氨基酸(1-10 个氨基酸)为克隆目的而被取代、添加或缺失,以形成转运肽融合等,其保留 HPPD 活性,即催化对羟基苯基丙酮酸转化为尿黑酸的性质。

[0054] 根据本发明,所述 HPPD 可为嵌合 HPPD。术语“嵌合 HPPD”是表示包含来自各种 HPPD 的元件的 HPPD。专利申请 WO 99/24586 中特别描述了这样的嵌合 HPPD。

[0055] 所述突变可发生在编码原始的未突变 HPPD 的核酸序列中,可采用任何适当的手段在所述序列中以对应于用于取代的氨基酸的密码子取代编码突变氨基酸的密码子,其中所述密码子在文献中有广泛描述,且为本领域技术人员熟知。

[0056] 若干分子生物学方法可用于获得该突变。

[0057] 一种制备本发明突变核酸序列和相应蛋白质的优选方法,包括在编码一个或多个预先选择的氨基酸的密码子上实施定点突变,包括 SEQ ID NO:2 所示假单胞菌 HPPD 序列的参比位置 Gly336 的密码子。获得这些定点突变的方法为本领域技术人员熟知,在文献中有广泛描述(特别是在《定点突变:一种实用方法》(Directed Mutagenesis:A Practical

Approach), 1991, M. J. McPHERSON 编, IRL 出版社), 或为可能运用商业试剂盒的方法(例如法玛西亚公司(PHARMACIA)的 U. S. E. 突变试剂盒)。定点突变后,通过采用适当的筛选辅助方法,选择包含对 HPPD 抑制剂具有较低敏感性的突变 HPPD 的细胞是有用的。一种易于实施的筛选方法是测定完全抑制原始的未突变 HPPD 且对表达该未突变 HPPD 的细胞致死的 HPPD 抑制剂的剂量,并对突变细胞施加该预定剂量,然后分离耐受该致死剂量的突变细胞,接着分离并克隆编码该突变 HPPD 的基因。基于本发明一个具体的实施方式和广泛流行的解决方法,即对三酮或吡唑特 HPPD 抑制剂具有较低敏感性的 HPPD,可如上所述进行筛选,采用三酮或吡唑特 HPPD 抑制剂,特别是选自特波三酮、甲基磺草酮、吡拉塞佛托、苯吡唑草酮和磺草酮的 HPPD 抑制剂。

[0058] 基于本发明的另一个实施方式,即除 SEQ ID NO:2 所示假单胞菌 HPPD 序列的第 336 位的参比氨基酸上的第一突变外,在第二氨基酸上又发生突变的 HPPD,可通过与第一突变同时或相继进行定点突变获得第二突变。

[0059] 作为上述定点突变的替代方法,可采用与适当筛选辅助方法相关的随机突变方法(如 EMS 或辐射处理)获得第二突变。这样的突变方法为本领域技术人员熟知,在文献中有充分描述(特别是 Sambrook 等,1989)。可如上所述实施筛选方法。

[0060] 本文中通用的术语“包含”某个序列 X 的 DNA 或蛋白质指至少包括或包含序列 X 的 DNA 或蛋白质,以使其它核苷酸或氨基酸序列可包括于 5' (或 N 末端) 和 / 或 3' (或 C 末端) 端,例如,选择性标记蛋白质(的核苷酸序列)、转运肽(的核苷酸序列)和 / 或 5' 前导序列或 3' 尾随序列。类似地,本申请全文和权利要求书中使用的术语“包含”应理解为意指包括所述的整数或步骤或一组整数或步骤,而不排除任何其它整数或步骤或其它组整数或步骤。

[0061] 因此本发明也涉及一种制备编码本发明突变 HPPD 的核酸序列的方法,所述方法如上定义。

[0062] 本发明也涉及编码本发明突变 HPPD 的核酸作为标记基因或可能赋予植物对 HPPD 抑制剂除草剂耐受性的编码序列在转化植物方法中的用途,以及 HPPD 抑制剂在包含编码本发明突变 HPPD 的核酸序列的植物上的用途。在本发明的一个实施方式中,在这样的用途中,HPPD 抑制剂为三酮类或吡唑特类,优选特波三酮、甲基磺草酮和磺草酮。当然应理解,该序列也可与其它(另一)基因标记和 / 或编码一个或多个具有有用农业性质的蛋白质的序列结合使用。

[0063] 在编码赋予转化植物有用农艺性质的蛋白质的基因中有编码赋予对某些除草剂耐受性、赋予对某些昆虫耐受性、赋予对某些疾病耐受性等的蛋白质的 DNA 序列。专利申请 W091/02071 和 W095/06128 特别描述了这样的基因。在编码赋予转化植物细胞和植物对某些除草剂耐受性的 DNA 序列中有赋予对草丁膦除草剂耐受性的 bar 基因、编码赋予对具有 EPSPS 靶的除草剂如草甘膦及其盐的耐受性的合适 EPSPS 的基因(US 4,535,060、US 4,769,061、US 5,094,945、US 4,940,835、US 5,188,642、US 4,971,908、US 5,145,783、US 5,310,667、US 5,312,910、US 5,627,061、US 5,633,435)、编码草甘膦氧化还原酶的基因(US 5,463,175)。

[0064] 在编码赋予对具有 EPSPS 靶的除草剂耐受性的合适 EPSPS 的 DNA 序列中,特别有编码植物 EPSPS,特别是玉米 EPSPS 的基因,其具有 102 和 106 两个突变,在专利申请 FR

2736926 中有描述,此后称为 EPSPS 双突变体,或编码从农杆菌分离的 EPSPS 的基因,由美国专利 5,633,435 的序列 ID No. 2 和序列 ID No. 3 描述,此后称为 CP4。

[0065] 在编码 EPSPS、更特别是编码以上基因的 DNA 序列的情况中,编码这些酶的序列之前最好存在编码转运肽的序列,特别是美国专利 5,510,471 或 5,633,448 描述的编码“优化转运肽”的序列。

[0066] 在编码赋予对昆虫耐受性的新性质的感兴趣蛋白质的 DNA 序列中,更特别有文献中广泛描述且为本领域技术人员熟知的 Bt 蛋白质。还有从细菌如发光杆菌 (WO 97/17432 和 WO 98/08932) 中提取的蛋白质。

[0067] 本发明也涉及一种嵌合基因(或表达盒),在 5' 和 / 或 3' 位置,至少在 5' 位置包含编码序列以及异源调控元件,它们能够在宿主生物体特别是植物细胞或植物中发挥作用,其中编码序列包含至少一个如前定义的编码突变 HPPD 的核酸序列。

[0068] 因此本发明涉及一种嵌合基因(或表达盒),在 5' 和 / 或 3' 位置,至少在 5' 位置包含编码序列以及异源调控元件,它们能够在宿主生物体特别是植物细胞或植物中发挥作用,其中编码序列包含至少一个如前定义的核酸序列。

[0069] 在一个具体的实施方式中,本发明涉及一种如前所述的嵌合基因,其中宿主生物体选自细菌、酵母、毕赤酵母 (*Pichia*)、真菌、杆状病毒、植物细胞和植物。

[0070] 在另一个具体的实施方式中,本发明涉及一种如前所述的嵌合基因,其中该嵌合基因在编码突变 HPPD 核酸序列的 5' 位置包含编码植物转运肽的核酸序列,其中该序列位于启动子区域和编码突变 HPPD 的序列之间,以允许转运肽 / 突变 HPPD 融合蛋白的表达。

[0071] 作为植物细胞和植物中的启动子调控序列,可使用植物中天然表达的基因的任何启动子序列,特别是特别在植物叶子中表达的启动子,例如,细菌、病毒或植物来源的“组成型”启动子,或“光依赖型”启动子,如植物核酮糖二羧化酶 / 加氧酶 (RuBisCO) 小亚基基因的启动子,或任何合适的已知可使用的启动子。在植物来源的启动子中,可提到如专利申请 EP 0 507 698 所述的组蛋白启动子,或水稻肌动蛋白启动子 (US 5,641,876)。在植物病毒基因的启动子中,有花椰菜花叶病毒 (CAMV 19S 或 35S) 的启动子,或环病毒启动子 (AU 689 311)。

[0072] 也可使用植物特定区域或组织的特异性调控启动子序列,如种子特异性启动子 (Dat1a, R. 等, 1997),特别是 napin 启动子 (EP 255 378)、菜豆蛋白 (phaseolin) 启动子、麦谷蛋白 (glutenin) 启动子、向日葵蛋白 (helianthinin) 启动子 (WO 92/17580)、白蛋白启动子 (WO 98/45460)、油质蛋白 (oleosin) 启动子 (WO 98/45461)、SAT1 启动子或 SAT3 启动子 (PCT/US98/06978)。

[0073] 诱导型启动子也宜选自苯丙氨酸解氨酶 (PAL)、HMG-CoA 还原酶 (HMG)、几丁质酶、葡聚糖酶、蛋白酶抑制剂 (PI)、PR1 家族基因、胭脂碱合成酶 (nos) 和 vspB 启动子 (US 5 670 349, 表 3)、HMG2 启动子 (US 5 670 349)、苹果 β -半乳糖苷酶 (ABG1) 启动子或苹果氨基环丙烷羧酸合成酶 (ACC 合成酶) 启动子 (WO 98/45445)。

[0074] 根据本发明,可联用启动子与位于启动子和编码序列之间的其它调控序列,如转录激活剂(“增强子”),例如专利申请 WO 87/07644 描述的烟草花叶病毒 (TMV) 的翻译激活剂,或 Carrington 和 Freed 1990 描述的烟草蚀纹病毒 (TEV),例如,或内含子,如玉米的 *adh1* 内含子或水稻肌动蛋白的内含子 1。

[0075] 作为调控终止子或多聚腺苷酸化序列,可使用细菌来源的任何相应序列,如例如根癌农杆菌 (*Agrobacterium tumefaciens*) 的 nos 终止子,病毒来源的相应序列,如 CaMV 35S 终止子,或植物来源的相应序列,如例如专利申请 EP 0 633 317 描述的组蛋白终止子。

[0076] “宿主生物体”应理解为可为产生突变 HPPD 而将本发明嵌合基因引入其中的任何单细胞或多细胞生物。这些生物,特别是细菌例如大肠杆菌、酵母特别是酿酒酵母属 (*Saccharomyces*) 或克鲁维酵母属 (*Kluyveromyces*)、毕赤酵母属、真菌特别是曲霉 (*Aspergillus*)、杆状病毒或优选植物细胞和植物。

[0077] 根据本发明,“植物细胞”应理解为来自或发现于植物的任何细胞,其能够形成以下或为其一部分:未分化组织如愈伤组织、分化组织如胚胎、植物的组成部分、植物或种子。

[0078] 根据本发明,“植物”应理解为能够进行光合作用的任何分化的多细胞生物,特别是单子叶或双子叶生物,更特别地为用于或非用于动物或人营养的栽培植物,如玉米、小麦、芸苔属 (*Brassica*) 植物如甘蓝型油菜 (*Brassica napus*) 或芥菜型油菜 (*Brassica juncea*)、大豆、水稻、甘蔗、甜菜根、烟草、棉花、蔬菜植物如黄瓜、韭菜、胡萝卜、番茄、莴苣、胡椒、甜瓜、西瓜等。

[0079] 在一个实施方式中,本发明涉及植物的转化。在植物中天然表达的基因的任何启动子序列,或在植物中天然表达的基因的任何杂合或组合的启动子元件,包括农杆菌或植物病毒启动子,或适合控制除草剂耐受性基因转录的任何启动子均可用作本发明植物中的启动子调控序列。以上描述了这样合适的启动子的例子。

[0080] 根据本发明,也可能联用启动子调控序列与其它位于启动子和编码序列之间的调控序列,如内含子序列,或转录激活剂(增强子)。以上描述了这样合适的调控序列的例子。

[0081] 细菌来源的任何相应序列,如根癌农杆菌的 nos 终止子,或植物来源的任何相应序列,如专利申请 EP 0 633 317 描述的组蛋白终止子,均可用作转录终止(和多聚腺苷酸化)调控元件。

[0082] 在本发明一个具体的实施方式中,在编码突变 HPPD 的核酸序列的 5' 采用编码转运肽的核酸序列,其中该转运肽序列位于启动子区域和编码突变 HPPD 的序列之间,以允许转运肽/突变 HPPD 融合蛋白的表达,其中所述突变 HPPD 如前所述。转运肽可能引导突变的 HPPD 进入质体,更具体说是叶绿体,突变 HPPD 进入质体时在转运肽和突变 HPPD 之间切割融合蛋白。所述转运肽可为单个肽,如 EPSPS 转运肽(美国专利 5,188,642 描述的)或植物核酮糖二羧化酶/加氧酶小亚基(RuBisCO ssu)的转运肽,适当时包括成熟 RuBisCO ssu 的 N 末端部分的几个氨基酸(EP 189 707),或可为若干转运肽的融合如包含与具有质体位置的成熟蛋白质 N 末端序列的一部分融合的第一植物转运肽的转运肽,其中该部分又与如专利 EP 508 909 描述的第二植物转运肽融合,更具体地说,包含与玉米 RuBisCO ssu N 末端的 22 个氨基酸融合的向日葵 RuBisCO ssu 的转运肽的优化转运肽又与专利 EP 508 909 描述其编码序列的玉米 RuBisCO ssu 的转运肽融合。

[0083] 本发明也涉及所述转运肽/突变 HPPD 融合蛋白和编码这样融合蛋白的核酸或植物可表达的嵌合基因,其中该融合蛋白的两个元件如上所述。

[0084] 本发明也涉及转化宿主生物体的克隆和/或表达载体,该载体包含至少一个如上所述的嵌合基因。除以上嵌合基因外,该载体还包含至少一个复制起点。该载体可为通过引入本发明嵌合基因已被转化的质粒、粘粒、噬菌体或病毒。这样的转化载体依赖于转化的宿

主生物体,为本领域技术人员熟知,在文献中有广泛描述。使用的转化载体特别是转化植物细胞或植物的载体可为病毒,可用于转化发育的植物,还包含其本身的复制和表达元件。根据本发明,转化植物细胞或植物的载体优选为质粒,如卸甲的(disarmed)农杆菌Ti质粒。

[0085] 本发明也涉及所述宿主生物体,特别是植物细胞或植物,其是转化的,且包含如上定义的含有编码突变HPPD序列的嵌合基因,以及本发明的植物用于田地生产作物并收获植物产品例如大豆或玉米谷物的用途,其中在一个实施方式中,所述用途包括将HPPD抑制剂除草剂用于这样的植物以控制杂草。在本发明的一个实施方式中,在这样的使用中,所述HPPD抑制剂为三酮类或吡唑特类,优选特波三酮、甲基磺草酮或磺草酮,特别是特波三酮。

[0086] 因此,本发明涉及一种宿主生物体,特别是植物细胞或植物,其特征在于包含至少一个如上所述的嵌合基因,或至少一个如上所述的核酸序列。

[0087] 在一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码突变HPPD酶的核酸序列,其中突变的HPPD酶保留催化对羟基苯基丙酮酸(HPP)转化为尿黑酸的性质,且对HPPD抑制剂的敏感性低于原始的未突变HPPD,其特征在于其在SEQ ID NO:2所示假单胞菌HPPD的第336位(天然HPPD中为氨基酸甘氨酸)包含选自以下的突变:Gly336Arg、Gly336His、Gly336Met、Gly336Phe、Gly336Asn、Gly336Cys和Gly336Val,条件是突变的HPPD不是双重突变体Gly334Ala-Gly336Arg(位置根据SEQ ID NO:2所示假单胞菌HPPD给出)。

[0088] 在一个更具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码如上所述的突变HPPD的核酸序列,其中SEQ ID NO:2所示假单胞菌HPPD的第336位的突变选自以下突变:Gly336His、Gly336Met、Gly336Cys和Gly336Phe,特别是Gly336His。

[0089] 在另一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码突变HPPD的核酸序列,其中突变的HPPD保留催化对羟基苯基丙酮酸(HPP)转化为尿黑酸的性质,且对HPPD抑制剂的敏感性低于原始的未突变HPPD,其中所述HPPD酶来自植物,特别是来自拟南芥,并且在SEQ ID NO:4所示拟南芥HPPD的氨基酸序列第422位(即SEQ ID NO:2所示假单胞菌HPPD的氨基酸序列的第336位)的甘氨酸上包含一个选自以下的突变:Gly336Arg、Gly336His、Gly336Met、Gly336Phe、Gly336Asn、Gly336Cys、Gly336Val、Gly336Trp、Gly336Glu和Gly336Asp。

[0090] 在一个更具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变HPPD的核酸序列,其中SEQ ID NO:2所示假单胞菌HPPD的第336位的突变选自以下突变:Gly336His、Gly336Asn、Gly336Cys和Gly336Val,所述突变HPPD是植物来源,特别是拟南芥。应注意,SEQ ID NO:2所示假单胞菌HPPD的第336位是SEQ ID NO:4所示拟南芥HPPD的第422位。

[0091] 在一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变HPPD的核酸序列,其中本发明的突变HPPD对异噁唑类、二酮腈类、三酮类或吡唑特类的HPPD抑制剂的敏感性低于原始的未突变HPPD。

[0092] 在一个更具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变HPPD的核酸序列,其中所述突变HPPD对选自以下的HPPD抑制剂的敏感性低于原始的未突变HPPD:异噁氟草、特波三酮、甲基磺草酮、磺草酮、吡拉塞佛托、苯

吡唑草酮、2-氰基-3-环丙基-1-(2-SO₂CH₃-4-CF₃苯基)-丙烷-1,3-二酮和2-氰基-3-环丙基-1-(2-SO₂CH₃-4-2,3Cl₂苯基)-丙烷-1,3-二酮。

[0093] 在另一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变 HPPD 的核酸序列,其中所述突变 HPPD 对选自以下的 HPPD 抑制剂的敏感性低于原始的未突变 HPPD:三酮类,如特波三酮、磺草酮和甲基磺草酮,特别是特波三酮,或吡唑特类,如吡拉塞佛托和苯吡唑草酮。

[0094] 在一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变 HPPD 的核酸序列,其中所述突变 HPPD 对选自特波三酮、磺草酮或甲基磺草酮,特别是特波三酮的三酮 HPPD 抑制剂的敏感性较低。

[0095] 在另一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变 HPPD 的核酸序列,其中除 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位氨基酸甘氨酸上的第一突变外,本发明的突变 HPPD 还包含第二突变。

[0096] 在一个更具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码上述突变 HPPD 的核酸序列,其中所述第二突变氨基酸选自以下氨基酸:SEQ ID NO:2 所示假单胞菌 HPPD 序列的 Pro215、Gly298、Gly332、Phe333、Gly334 和 Asn337。

[0097] 本发明还涉及一种植物细胞或植物,其特征在于包含至少一个编码突变 HPPD 酶的核酸序列,其中突变的 HPPD 酶保留催化对羟基苯基丙酮酸 (HPP) 转化为尿黑酸的性质,且对选自以下 HPPD 抑制剂的敏感性低于原始的未突变 HPPD:三酮类,如特波三酮、磺草酮和甲基磺草酮,或吡唑特类,如吡拉塞佛托和苯吡唑草酮,其特征在于 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位氨基酸甘氨酸包含突变。

[0098] 在一个更具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个编码突变 HPPD 酶的核酸序列,其中突变的 HPPD 酶对三酮类或吡唑特类的 HPPD 抑制剂的敏感性低于原始的未突变 HPPD,根据选自以下的突变, HPPD 在 SEQ ID NO:2 所示假单胞菌 HPPD 的第 336 位发生突变:Gly336Arg、Gly336Asp、Gly336Glu、Gly336His、Gly336Met、Gly336Phe、Gly336Trp、Gly336Asn、Gly336Cys 和 Gly336Val。

[0099] 在另一个具体的实施方式中,本发明涉及一种植物细胞或植物,其特征在于包含至少一个上述核酸序列,还有在植物中发挥作用的基因,从而允许 PDH(预苯酸脱氢酶)酶过表达。

[0100] 本发明也涉及包含转化细胞的植物,特别是从转化细胞再生的植物。再生可通过适当的方法获得,例如以上参考文献所述,其中所述方法依赖于物种的性质。可引用以下专利和专利申请,特别是关于转化植物细胞和植物再生的方法:US 4,459,355、US 4,536,475、US 5,464,763、US 5,177,010、US 5,187,073、EP 267,159、EP 604662、EP 672752、US 4,945,050、US 5,036,006、US 5,100,792、US 5,371,014、US 5,478,744、US 5,179,022、US 5,565,346、US 5,484,956、US 5,508,468、US 5,538,877、US 5,554,798、US 5,489,520、US 5,510,318、US 5,204,253、US 5,405,765、EP 442 174、EP 486 233、EP 486 234、EP 539 563、EP 674 725、WO 91/02071 和 WO 95/06128。

[0101] 本发明也涉及转化植物或其部分,其通过栽培和/或杂交以上再生植物得到,还涉及转化植物的种子。

[0102] 本发明也涉及从本发明的植物、其部分或种子获得的终产物如粗碾谷物或油。

[0103] 根据本发明可获得的转化植物可为单子叶型,如谷物、甘蔗、水稻和玉米,或为双子叶型,如烟草、大豆、芸苔属植物如油菜、棉花、甜菜根、三叶草等。

[0104] 本发明涉及转化宿主生物体特别是植物细胞或植物的方法,所述方法将上述至少一个核酸序列或一个嵌合基因整合入这样的生物体中,其中可通过任何适当的已知手段获得转化,所述手段在专门的文献中有充足描述,特别是在本发明引用的参考文献中,更具体地说是采用本发明的载体。

[0105] 一系列方法包括用 DNA 序列附着的粒子轰击细胞、原生质体或组织。另一系列方法包括采用插入到根癌农杆菌 Ti 质粒或毛根农杆菌 Ri 质粒的嵌合基因作为转移入植物的手段。可采用其它方法,如微注射或电穿孔或用 PEG 引导沉淀。技术人员可选择任何适当的方法转化所选宿主生物体,特别是植物细胞或植物。例如,大豆转化的技术在通过引用纳入本文的 EP 1186666 的实施例 1-3 中有广泛描述。对于水稻,可进行农杆菌介导的转化 (Hiei 等, 1994, 和 Hiei 等, 1997, 通过引用纳入本文)、电穿孔 (美国专利 5, 641, 664 和美国专利 5, 679, 558, 通过引用纳入本文) 或轰击 (Christou 等, 1991, 通过引用纳入本文)。通过引用纳入本文的 WO 92/09696 中描述了单子叶植物特别是水稻转化的合适技术。对于棉花,已描述了农杆菌介导的转化 (Gould J.H. 和 Magallanes-Cedeno M., 1998 和 Zapata C., 1999, 通过引用纳入本文)、聚凝胺 (polybrene) 和 / 或处理介导的转化 (Sawahe1 W. A., 2001, 通过引用纳入本文)。

[0106] 在本发明一个具体的实施方式中,突变的 HPPD 靶向叶绿体。可通过整合上述编码转运肽 / 突变 HPPD 融合蛋白的核酸序列来进行。

[0107] 或者,可采用转化叶绿体基因组将突变的 HPPD 直接在叶绿体中表达。一个合适的方法包括用 DNA 包被的粒子轰击树叶部分,并通过同源重组整合编码本发明蛋白质的引入基因。合适的载体和选择系统为本领域技术人员已知。可用于这种整合入烟草品系叶绿体基因组的手段和方法的一个例子在 WO 06/108830 中给出,其内容通过引用纳入本文。在采用叶绿体基因组转化将多肽直接靶向入叶绿体时,通常不需要转运肽序列。

[0108] 本发明也涉及获得耐受 HPPD 抑制剂的植物的方法,其特征在于所述植物用上述嵌合基因进行转化。

[0109] 因此,本发明也涉及一种获得耐受 HPPD 抑制剂的植物的方法,其特征在于所述植物用在 5' 且任选地在 3' 位置包含编码序列以及异源调控元件的嵌合基因进行转化,所述编码序列以及异源调控元件能够在宿主生物体中发挥作用,其特征在于所述编码序列包含至少一个如上所述的核酸序列。

[0110] 在本发明一个具体的实施方式中,在该方法中,所述 HPPD 抑制剂是三酮或吡唑特除草剂,优选特波三酮、甲基磺草酮或磺草酮,特别是特波三酮。

[0111] 在另一个实施方式中,本发明涉及获得上述耐受 HPPD 抑制剂的植物的方法,其特征在于所述植物还同时或相继用在该植物中发挥作用从而允许 PDH (预苯酸脱氢酶) 酶过表达的基因进行转化。

[0112] 本发明也涉及借助 HPPD 抑制剂特别是上述除草剂为植物特别是作物选择性除草的方法,该方法的特征在于在作物播种前、作物萌发前或作物萌发后施用将该除草剂施用于根据本发明进行转化的植物。

[0113] 在本发明一个具体的实施方式中,在该方法中,所述 HPPD 抑制剂是三酮或吡唑特

除草剂, 优选特波三酮、甲基磺草酮或磺草酮, 特别是特波三酮。

[0114] 本发明也涉及在包含如本专利申请之前所述的转化种子的区域或田地里控制杂草的方法, 该方法包括对所述田地区域施用对所述杂草有毒剂量的 HPPD 抑制剂除草剂, 而不显著影响包含如本专利申请之前所述的核酸序列或嵌合基因的种子或植物。

[0115] 在本发明一个具体的实施方式中, 在该方法中, 所述 HPPD 抑制剂是三酮或吡唑特除草剂, 优选特波三酮、甲基磺草酮或磺草酮, 特别是特波三酮。

[0116] 本发明也涉及一种栽培用本发明嵌合基因转化的植物的方法, 该方法包括将包含本发明嵌合基因的种子种植在适合栽培所述植物的田地区域里, 如果存在杂草, 将对杂草有毒剂量的、以上述 HPPD 为靶标的除草剂施用于所述田地的所述区域, 而不显著影响所述转化种子或所述转化植物, 然后在栽培植物或植物部分达到所需成熟阶段时收获所述栽培植物或植物部分, 适当时从收获的植物中分离种子。

[0117] 在本发明一个具体的实施方式中, 在该方法中, 所述 HPPD 抑制剂是三酮或吡唑特除草剂, 优选特波三酮、甲基磺草酮或磺草酮, 特别是特波三酮。

[0118] 在以上方法中, 可根据本发明, 在作物播种前、作物萌发前或作物萌发后施用以所述 HPPD 为靶标的除草剂。

[0119] 本发明也涉及一种获得油, 特别是大豆油或粗碾谷物的过程, 包括在田地里种植表达本发明突变 HPPD 的作物, 特别是大豆作物, 任选用 HPPD 抑制剂除草剂处理这样的作物, 收获谷物, 碾磨谷物以制备粗碾谷物并提取油。包含本发明嵌合基因的整个、破碎或碾碎的植物种子或谷物也是本发明的一部分。

[0120] 因此, 本发明涉及获得油或粗碾谷物的方法, 包括种植上述的转化植物, 任用地用 HPPD 抑制剂除草剂处理这样的植物, 收获谷物, 碾磨谷物以制备粗碾谷物并提取油。

[0121] 在具体的实施方式中, 本发明的以上方法包括选自以下的 HPPD 抑制剂除草剂: 异噁氟草、特波三酮、甲基磺草酮、吡拉塞佛托、磺草酮、苯吡唑草酮、2-氰基-3-环丙基-1-(2-SO₂CH₃-4-CF₃苯基)-丙烷-1, 3-二酮和 2-氰基-3-环丙基-1-(2-SO₂CH₃-4-2, 3Cl₂苯基)丙烷-1, 3-二酮。

[0122] 在其它具体的实施方式中, 本发明的以上方法包括选自以下的 HPPD 抑制剂除草剂: 三酮类, 如特波三酮、磺草酮和甲基磺草酮, 或吡唑特类, 如吡拉塞佛托和苯吡唑草酮, 特别选自特波三酮、磺草酮和甲基磺草酮, 更特别地为特波三酮。

[0123] 在本发明的含义中, “除草剂” 应理解为本身具有除草剂活性的物质或与改变其功效的添加剂如例如增强其活性 (增效剂) 或限制其活性 (安全剂) 的试剂结合的物质。当然应理解, 就其在实践中的应用来说, 以上除草剂以本身已知的方式与农业化学中惯常使用的配方佐剂结合。

[0124] 在根据本发明转化的植物包含一个或多个其它对其它除草剂耐受性的基因 (如例如, 编码赋予植物对草甘膦除草剂耐受性的突变或未突变 EPSPS 的基因, 或赋予对草丁膦除草剂耐受性的 pat 或 bar 基因) 时, 或在转化植物对另一个除草剂具有天然敏感性 (如磺酰脲耐受性) 时, 本发明的方法可包括结合所述除草剂或除草剂的组合, 例如草甘膦和 / 或草丁膦和 / 或磺酰脲除草剂, 同时或按顺序交错施用 HPPD 抑制剂。

[0125] 本发明也涉及基于对上述 HPPD 抑制剂除草剂的选择, 编码本发明突变 HPPD 的嵌合基因作为某种植物转化期间标记基因的用途。

[0126] 本发明也涉及获得耐受三酮或吡唑特 HPPD 抑制剂的植物的方法,其特征在于所述植物用在植物中表达 HPPD 的嵌合基因进行转化,所述 HPPD 在 SEQ ID NO:2 所示假单胞菌 HPPD 氨基酸序列的第 336 位的氨基酸甘氨酸发生突变。

[0127] 在一个具体的实施方式中,本发明涉及所述获得耐受三酮或吡唑特 HPPD 抑制剂的植物的方法,其特征在于所述 HPPD 突变选自 Gly336Arg、Gly336Asp、Gly336Glu、Gly336His、Gly336Met、Gly336Phe、Gly336trp、Gly336Asn、Gly336Cys 和 Gly336Val。

[0128] 在另一个具体的实施方式中,本发明涉及所述获得耐受选自特波三酮、甲基磺草酮和磺草酮的三酮 HPPD 抑制剂的植物的方法。

[0129] 在另一个具体的实施方式中,本发明涉及所述获得耐受三酮或吡唑特 HPPD 抑制剂的植物的方法,其特征在于所述植物还同时或相继用在该植物中发挥作用从而允许 PDH(预苯酸脱氢酶)酶过表达的基因进行转化。

[0130] 本发明也涉及在控制区域或田地中杂草的方法,该方法包括在该区域或田地种植根据上述方法获得的耐受三酮或吡唑特 HPPD 抑制剂的转化植物或来源于其的转化种子,施用对杂草有毒剂量的所述三酮或吡唑特 HPPD 抑制剂,而不显著影响所述转化种子或所述转化植物。

[0131] 本发明也涉及获得油或粗碾谷物的方法,包括种植根据上述方法获得的耐受三酮或吡唑特 HPPD 抑制剂的转化植物或来源于这样植物的转化种子,任选用三酮或吡唑特 HPPD 抑制剂处理这样的植物或种子,收获谷物,碾磨谷物以制备粗碾谷物并提取油。

[0132] 本发明也涉及在 SEQ ID NO:2 所示假单胞菌 HPPD 氨基酸序列的第 336 位的氨基酸甘氨酸发生突变的 HPPD 以使植物耐受三酮或吡唑特 HPPD 抑制剂的用途。

[0133] 本发明也涉及上述突变 HPPD 的用途,其特征在于所述 HPPD 突变选自:Gly336Arg、Gly336Asp、Gly336Glu、Gly336His、Gly336Met、Gly336Phe、Gly336trp、Gly336Asn、Gly336Cys 和 Gly336Val。

[0134] 本发明也涉及上述突变 HPPD 的用途,其特征在于所述 HPPD 抑制剂是选自特波三酮、甲基磺草酮或磺草酮的三酮 HPPD 抑制剂。

[0135] 本发明也涉及宿主生物体,特别是植物细胞或植物,其包含含有编码本发明突变 HPPD 的序列的嵌合基因,还包含在该宿主生物体中发挥作用从而允许预苯酸脱氢酶(本文缩写为 PDH)过表达的基因。

[0136] 在表述“在植物中发挥作用从而允许 PDH 酶过表达的基因”中,术语“PDH”应解释为表现出将预苯酸转化为 HPP 的 PDH 活性的任何天然或突变的 PDH 酶。特别地,所述 PDH 酶可来源于任何类型的生物。可通过任何可能测定预苯酸底物量的减少或测定来自酶反应的产物即 HPP 或辅因子 NADH 或 NADPH 之一的积累的方法鉴定具有 PDH 活性的酶。具体地说,可通过实施例 4 中描述的方法测定 PDH 活性。

[0137] 文献中描述了很多编码 PDH 酶的基因,其序列可在网址 <http://www.ncbi.nlm.nih.gov/entrez/> 进行鉴定。特别熟知的是编码以下的 PDH 酶的基因:如 Mannhaupt 等(1989)中所述的酿酒酵母(登录号 S46037)、杆状菌属(Bacillus)细菌特别是如 Henner 等(1986)中所述的枯草杆菌(B. subtilis)(登录号 P20692)、埃希氏菌属(Escherichia)细菌特别是如 Hudson 等(1984)中所述的大肠杆菌(登录号 KMECTD)、欧文氏菌属(Erwinia)的细菌特别是如 Xia 等(1992)中所述的草生欧文氏菌(E. herbicola)(登录号 S29934)。

[0138] 本发明还涉及获得耐受 HPPD 抑制剂的宿主生物体特别是植物细胞或植物的方法,通过在这样的生物体中整合入上述至少一个核酸序列或一个嵌合基因,再同时或相继在该宿主生物体中发挥作用从而允许 PDH(预苯酸脱氢酶) 酶过表达的基因对其进行转化。

[0139] 在一个具体的实施方式中,本发明涉及获得宿主生物体特别是植物细胞或植物的方法,其对三酮或吡唑特 HPPD 抑制剂,特别是特波三酮、甲基磺草酮或磺草酮具有耐受性。

[0140] 可用于获得用允许 HPPD 酶过表达的基因和允许 PDH 酶过表达的基因二者转化的宿主生物体特别是植物细胞或植物的手段和方法在 WO 04/024928 中有广泛描述,其内容通过引用纳入本文。

[0141] 本说明书中引用的任何在前出版物(或来源于其的信息)或任何已知事项不是也不应看作承认、认可或是任何形式的暗示该在前出版物(或信息)或已知事项形成本发明领域公知常识的一部分。

[0142] 附图

[0143] 图 1 :阿维链霉菌、胡萝卜、拟南芥、玉米、大麦、小麦叶枯病菌、粗球孢子菌、小鼠和荧光假单胞菌的 HPPD 序列比对。氨基酸根据假单胞菌序列编号,星号表示这些序列共有的氨基酸。

[0144] 序列表

[0145] SEQ ID NO 1 :编码荧光假单胞菌 HPPD 的核酸序列

[0146] SEQ ID NO 2 :荧光假单胞菌 HPPD 的氨基酸序列

[0147] SEQ ID NO 3 :编码拟南芥 HPPD 的核酸序列

[0148] SEQ ID NO 4 :拟南芥 HPPD 的氨基酸序列

[0149] SEQ ID NO 5 :编码小鼠 HPPD 的核酸序列

[0150] SEQ ID NO 6 :小鼠 HPPD 的氨基酸序列

[0151] SEQ ID NO 7 :编码粗球孢子菌 HPPD 的核酸序列

[0152] SEQ ID NO 8 :粗球孢子菌 HPPD 的氨基酸序列

[0153] SEQ ID NO 9 :编码小麦叶枯病菌 HPPD 的核酸序列

[0154] SEQ ID NO 10 :小麦叶枯病菌 HPPD 的氨基酸序列

[0155] SEQ ID NO 11 :编码大麦 HPPD 的核酸序列

[0156] SEQ ID NO 12 :大麦 HPPD 的氨基酸序列

[0157] SEQ ID NO 13 :编码玉米 HPPD 的核酸序列

[0158] SEQ ID NO 14 :玉米 HPPD 的氨基酸序列

[0159] SEQ ID NO 15 :编码胡萝卜 HPPD 的核酸序列

[0160] SEQ ID NO 16 :胡萝卜 HPPD 的氨基酸序列

[0161] SEQ ID NO 17 :编码阿维链霉菌 HPPD 的核酸序列

[0162] SEQ ID NO 18 :阿维链霉菌 HPPD 的氨基酸序列

[0163] SEQ ID NO 19 :引物序列 kerfi001

[0164] SEQ ID NO 20 :引物序列 kerfi002

[0165] SEQ ID NO 21 :引物序列 kerfi003

[0166] SEQ ID NO 22 :引物序列 kerfi004

- [0167] SEQ ID NO 23 :引物序列 kerfi007
[0168] SEQ ID NO 24 :引物序列 kerfi008
[0169] SEQ ID NO 25 :引物序列 kerfi011
[0170] SEQ ID NO 26 :引物序列 kerfi012
[0171] SEQ ID NO 27 :引物序列 kerfi014
[0172] SEQ ID NO 28 :引物序列 kerfi016
[0173] SEQ ID NO 29 :引物序列 kerfi019
[0174] SEQ ID NO 30 :引物序列 kerfi020
[0175] SEQ ID NO 31 :引物序列 kerfi015
[0176] SEQ ID NO 32 :引物序列 kerfi018

实施例

[0177] 在以下实施例的辅助下将会更好的理解本发明的各方面。以下在这些实施例中描述的所有方法和操作通过举例的方式给出,对应于从可达到相同或相似结果的不同方法中作出的选择。该选择对结果的质量没有影响,因此技术人员可使用任何合适的方法以达到相同或相似的结果。DNA 片段操作的大多数方法在《现代分子生物学实验技术》(“Current Protocols in Molecular Biology”),第 1 卷和第 2 卷, Ausubel F.M. 等,格林出版公司(Greene Publishing Associates)和 WI 公司(Wiley Interscience)(1989)、《分子克隆》(Molecular cloning), T. Maniatis, E. F. Fritsch, J. Sambrook, 1982 或 Sambrook J. 和 Russell D., 2001,《分子克隆 :实验手册》(Molecular Cloning:a laboratory manual)(第三版)中有描述。

[0178] 实施例 1 :突变 HPPD 的制备

[0179] 概述

[0180] 首先将拟南芥 AtHPPD 编码序列 (1335bp) (Genebank AF047834 ;WO 96/38567) 克隆入表达载体 pQE-30 (QIAGEN) 的限制性位点 BamHI 和 HindIII 之间。

[0181] 首先将荧光假单胞菌 PfHPPD 编码序列 (1174bp) (Rüetschi 等, Eur. J. Biochem., 205, 459-466, 1992, WO 96/38567) 克隆入提供起始密码子的表达载体 pKK233-2(法玛西亚公司)独特的 NcoI 位点。

[0182] 载体 pQE-30-AtHPPD 和 pKK233-2-PfHPPD 用于 PCR 介导的 NcoI 限制性位点和编码 N 末端 His₆- 标签的序列连接于 AtHPPD 和 PfHPPD 的 5' 端,以及 XbaI 限制性位点连接于 AtHPPD 和 PfHPPD 的 3' 段。

[0183] 从琼脂糖凝胶中分离 AtHPPD 基因的 PCR 产物,用限制酶 NcoI 和 XbaI 切割,用 MinElute™PCR 纯化试剂盒(恰根公司(Qiagen))纯化,克隆入用相同限制酶切割的 pSE420(RI)NX 载体。

[0184] 对于 PfHPPD 基因,从琼脂糖凝胶中分离 PCR 产物,克隆入 pCR[®] 2.1-TOPO[®] 载体。用限制酶 NcoI 和 XbaI 将其从该载体上切离,从琼脂糖凝胶中分离,并克隆入用相同限制酶切割的 pSE420(RI)NX 载体。

[0185] 然后对 pSE420(RI)NX-AtHPPD 和 -PfHPPD 二者进行 PCR 介导的定点突变,以在两个基因的相应位点改变指定的密码子。所述密码子分别在 WT PfHPPD 中编码 Gly336 和在

WT AtHPPD 中编码 Gly422。

[0186] 采用焦磷酸测序(Pyrosequencing[®])技术分析编码序列中的突变密码子。

[0187] PCR 介导的编码 N 末端 His₆- 标签的序列与 NcoI 和 XbaI 限制性位点的连接：

[0188] 在 96 孔 PCR 板的 24 孔上分别进行各基因 (AtHPPD 和 PfHPPD) 的 PCR 反应。由于该反应的正向和反向引物在大小上有 18 (AtHPPD) 和 22bp (PfHPPD) 的差异, 因此实施 40.9°C -64.5°C 的退火温度梯度, 各孔置于该范围内的另一个退火温度下。在引物与单链模板进行第一次退火时, 在新链中产生了 5' 突出端, 直到合成出其互补链, 由第一引物的 5' 区域形成的该突出端成为模板的一部分。由此编码序列在两端延伸, 在两端引入编码 N 末端 His₆- 标签的序列和限制性位点。

[0189] 该反应混合物包含 500ng pQE-30-AtHPPD (1 μL 来自质粒大量抽提) 或 1 μg pKK233-2-PfHPPD DNA (0.75 μL 来自质粒大量抽提)、用于 AtHPPD 的 kerfi001 和 kerfi002 分别 1 μl、或用于 PfHPPD 的 kerfi003 和 kerfi004 分别 1 μl (所有引物溶液均具有 10pmol*μL⁻¹ 的浓度)、25 μl HotStarTaq Master Mix (恰根公司) 和分子生物学级超纯水 (HyPure[™] Molecular Biology Grade Water) 至终体积 50 μL。PCR 程序设置如下：

[0190] 1.95°C 15 分钟

[0191] 2.94°C 30 秒

[0192] 40.9°C -60.4°C 30 秒

[0193] 72°C 3 分钟

[0194] 步骤 2 重复 20 次。

[0195] 3.72°C 10 分钟

[0196]

引物名称	引物序列
kerfi001	5'-CCATGGCTCATCACCATCACCATCACCAAAACGCCGCCG TTTCAG-3'
kerfi002	5'-TCTAGATCATCCCACTAACTGTTTGGC-3'
kerfi003	5'-CCATGGCTCATCACCATCACCATCACGCAGATCTATACG AAAACCCAATGG-3'
kerfi004	5'-TCTAGATTAATCGGCGGTCAATACACCAC-3'

[0197] 对 PCR 反应进行琼脂糖凝胶电泳, 所有均产生对应于约 1500bp (AtHPPD) 或 1100bp (PfHPPD) 片段的清晰条带。将所述条带从凝胶中切离, 用 QIAquick[®] 凝胶回收试剂盒 (Gel Extraction Kit) (恰根公司) 纯化 DNA。

[0198] 克隆入 pCR[®]2.1-TOPO[®] 载体 (英杰公司 (Invitrogen))

[0199] pCR[®]2.1-TOPO[®] 载体 (3931bp) 用于展现出 3' - 腺苷 (A) 突出端的 Taq 聚合酶 - 扩增 PCR 产物的一步法克隆。所述载体又经线性化在其末端展现出单个 3' - 胸苷 (T) 突出端。拓扑异构酶 I 与用来将载体共价连接到 PCR 产物的这些 3' - 胸苷共价连接。就选

择携带载体的细菌细胞来说,可使用氨苄青霉素或卡那霉素。所述载体在其多克隆位点内包含 XbaI 限制性位点,在 KanR 基因内包含 NcoI 限制性位点。

[0200] 从各凝胶提取获得的 DNA 溶液分别用于 TOPO TA 克隆。转化大肠杆菌 TOP10 细胞后,各反应产生出三个白色菌落 (A1-A3, P1-P3),用于接种到 5mL LB/amp 培养基上。

[0201] 为确定这些菌落的载体是否携带正确的插入片段,采用 QIAprep[®] 质粒小量抽提试剂盒 (Spin Miniprep Kit) (怡根公司) 从 4mL pCR[®]2.1-TOPO[®]-AtHPPD 培养物 A1-A3 和 -PfHPPD 培养物 P1-P3 制备质粒 DNA。将从这些质粒制备物获得的 DNA 溶液用 HindIII 和 XhoI 进行限制性消化,然后置于 1% 琼脂糖凝胶上分析。HindIII 和 XhoI 各自分别在 pCR[®]2.1-TOPO[®]-AtHPPD/-PfHPPD 载体中包含单个限制性位点。限制性消化克隆 A1 的 DNA 产生了表示 1461bp 片段 (AtHPPD 编码序列) 和 3831bp 载体片段的预期条带;限制性消化 P3 在琼脂糖凝胶上产生了表示 1206bp 片段 (PfHPPD 编码序列) 和 3831bp 载体片段的预期条带。

[0202] 从采用 QIAfilter[™] 大量抽提试剂盒 (Maxi Kit) (怡根公司) 的质粒大量抽提和随后的从 100mL A1 (AtHPPD) 或 P3 (PfHPPD) 液体 LB/amp 培养基的 NaAc/EtOH 沉淀获得的 DNA 用来测定 pCR[®]2.1-TOPO[®] 载体中各个插入 HPPD 基因的 DNA 序列。采用 EMWG 股份有限公司 (Eurofins MWG GmbH) 的引物 M13 uni(-21) 和 M13 rev(-29) 进行 DNA 测序。测序证实了 pCR[®]2.1-TOPO[®] 载体中正确的 AtHPPD 和 PfHPPD 的 DNA 序列,包括编码序列两端的限制性位点。

[0203] 克隆入 pSE420 (RI) NX

[0204] 克隆和表达载体 pSE420 (RI) NX (5261bp) 基于英杰公司的质粒 pSE420。该载体的修饰包括添加卡那霉素耐受性基因和去除大部分超级连接 (superlinker) 区域 (多克隆位点)。

[0205] 该质粒包含 trp-lac (trc) 启动子和在每个大肠杆菌宿主菌株中提供 lac 阻遏物的 lacI^q 基因。所述 lac 阻遏物与 lac 操纵基因 (lacO) 结合,限制靶基因的表达;可通过异丙基-β-D- 硫代半乳糖苷 (IPTG) 诱导减轻该抑制作用。

[0206] 将基因 AtHPPD 和 PfHPPD 克隆入载体 pSE420 (RI) NX 中限制性位点 NcoI 和 XbaI 之间。

[0207] 基于 PCR 的定点突变:

[0208] 通过质粒小量抽提从大肠杆菌 TOP10 液体培养物分离模板 DNA (pSE420 (RI) NX-AtHPPD 和 pSE420 (RI) NX-PfHPPD)。将从这些小量抽提物获得的 DNA 溶液稀释到 0.05 μg*μL⁻¹ 的浓度。

[0209] 基于 PCR 的定点突变需要与相同 DNA 区域互补的两个化学合成的 DNA 引物 (正向和反向引物),两个引物各与双链 DNA 模板的一条链互补。这些引物在其中心包含所需的突变,覆盖约 20-30 个模板核苷酸的区域,包括突变位点和两侧各 10-15 个碱基。所述突变位点覆盖在引物中独立变化的三个核苷酸,以在所选位点获得各可能的密码子。

[0210] 在循环 PCR 突变中,通过从整合入生成链的引物 3'OH 端开始的滚环复制对质粒模板进行完全复制。各新 DNA 分子携带一个或多个包含于引物中的改变的核苷酸。采用高保

真 DNA 聚合酶以降低出现另外不需要的突变的可能性。

[0211] 将寡核苷酸引物对 kerfi007/kerfi008 (AtHPPD) 和 kerfi011/kerfi012 (PfHPPD) 溶解于水中至 $10\text{pmol} \cdot \mu\text{L}^{-1}$ 的浓度。对于突变 PCR 反应, 采用 50ng 来自 pSE420 (RI) NX-AtHPPD 或 pSE420 (RI) NX-PfHPPD 小量抽提的模板质粒, 稀释至 $0.05 \mu\text{g} \cdot \mu\text{L}^{-1}$ 的浓度。反应混合物的组成如下:

[0212] $1 \mu\text{L}$ 模板质粒 ($0.05 \mu\text{g} \cdot \mu\text{L}^{-1}$)

[0213] $1.5 \mu\text{L}$ 引物 kerfi007 (或 kerfi011) ($10\text{pmol} \cdot \mu\text{L}^{-1}$)

[0214] $1.5 \mu\text{L}$ 引物 kerfi008 (或 kerfi012) ($10\text{pmol} \cdot \mu\text{L}^{-1}$)

[0215] $5 \mu\text{L}$ 10x 反应缓冲液

[0216] $1 \mu\text{L}$ dNTP mix

[0217] $40 \mu\text{L}$ HyPure™ 分子生物学级用水

[0218] $1 \mu\text{L}$ *PfuUltra*® 高保真 DNA 聚合酶 ($2.5\text{U} \cdot \mu\text{L}^{-1}$)

[0219] PCR 程序与 AtHPPD 和 PfHPPD 突变的程序相同, 延伸时间设为 7 分钟, 假设 1 分钟复制 1kb 质粒 DNA。

[0220]

1. 95°C 30秒
2. 95°C 30秒
- 55°C 30秒
- 68°C 7分钟

[0221] 步骤 2 重复 18 次。

[0222] PCR 反应后, 将反应置于冰上冷却至室温。

[0223]

引物名称	引物序列
kerfi007	5' -GGTGGTTTTGGCAAANNAATTTCTCTGAGCTC-3'
kerfi008	5' -GAGCTCAGAGAAATTNNNTTGGCCAAAACCACC-3'
kerfi011	5' -CAGCGCCTGAAGTTNNNCTCGCCAAAACCCATC-3'
kerfi012	5' -GATGGGTTTGGCGAGNNNACTTCAAGGCGCTG-3'

[0224] PCR 反应后, 采用 DpnI 限制性内切酶对突变质粒进行选择。仅有 dam- 甲基化 DNA 被限制性酶 DpnI 降解, 其限制性位点 $\text{G}^{\text{Me6}}\text{ATC}$ 相对充足。细菌产生的模板质粒发生了甲基化, 因此降解。然而 PCR- 扩增的 DNA 仍保持原样。

[0225] 将 $1 \mu\text{L}$ DpnI 限制性酶 ($10\text{U} \cdot \mu\text{L}^{-1}$) 加入到 PCR 反应中, 通过上下抽吸混合溶液。离心 ($13,200\text{rpm}$) 1 分钟后, 将反应置于 37°C 温育 1 小时。

[0226] 突变质粒在各引物的 5' 端包含交错的缺口, 可直接转化入感受态细胞。

[0227] 为浓缩突变质粒, 进行 NaAc/EtOH 沉淀, 在 $10 \mu\text{L}$ HyPure™ 分子生物学级用水中重悬 DNA。稍后用 $3 \mu\text{L}$ 这些质粒溶液转化电感受态大肠杆菌 K-12 MG1655 细胞, 在 AtHPPD

的情况下,用 1 μ L 转化电转感受态大肠杆菌 TOP10 细胞。

[0228] 对于 AtHPPD 来说,一共获得 62 个大肠杆菌 K-12 MG1655 克隆,对其进行培养用于随后在 Costar[®] 96 孔 2mL 深孔板中分析突变密码子。为获得更多数量的克隆,将大肠杆菌 TOP10 用作克隆突变质粒的替代宿主。用突变质粒转化大肠杆菌 TOP10 细胞产生了几百个克隆。

[0229] 对于 PfHPPD,一共获得 252 个大肠杆菌 K-12 MG1655 克隆,对其进行培养用于所述对 AtHPPD 质粒转化克隆的分析。

[0230] 实施例 2 : Pyrosequencing[®] 反应证实点突变

[0231] 采用 Pyrosequencing[®] 技术通过测定一段短的限定 DNA 的核酸序列证实点突变。首先进行 PCR 反应扩增包含待测序段的短 DNA 片段。PCR 扩增模板需为单链,与生物素分子在其 5' 端共价结合。生物素用于将模板非共价附着于与交联琼脂糖固定相连接的链霉亲和素。

[0232] 生物素化 DNA 片段的扩增:在 96 孔 PCR 板上进行 PCR 反应。反应混合物包含 1 μ L 正向引物溶液(用于 AtHPPD 的 kerfi016,用于 PfHPPD 的 kerfi020 ;10pmol* μ L⁻¹)、1 μ L 反向引物溶液(包含在其 5' 端的生物素修饰;用于 AtHPPD 的 kerfi019,用于 PfHPPD 的 kerfi014 ;10pmol* μ L⁻¹)、2 μ L 在深孔板中培养克隆的液体细菌培养物、25 μ L HotStarTaq[®] Master Mix 和 21 μ L HyPure[™] 分子生物学级用水。

[0233] 用于 AtHPPD 和 PfHPPD 的 PCR 程序在退火温度方面存在差异,分别设为 55 $^{\circ}$ C 和 60 $^{\circ}$ C。

[0234] 1. 95 $^{\circ}$ C 15 分钟

[0235] 2. 94 $^{\circ}$ C 30 秒

[0236] 55 $^{\circ}$ C /60 $^{\circ}$ C 30 秒

[0237] 72 $^{\circ}$ C 30 秒

[0238] 步骤 2 重复 32 次。

[0239] 3. 72 $^{\circ}$ C 10 分钟

[0240]

引物名称	引物序列
kerfi014 (5'bio)	5'-GATCTTCTCGGAAACCCTGATG-3'
kerfi016	5'-GGGATTCTTGTAGACAGAGATG-3'
kerfi019 (5'bio)	5'- CCCACTAACTGTTTGGCTTC-3'
kerfi020	5'- GGCGGTCAATACACCACGAC-3'

[0241] Pyrosequencing[®] 反应:在 96 孔板上进行 Pyrosequencing[®] 反应(拜尔塔格公司(Biotage))。对于每 45 μ L PCR 反应,加入 40 μ L 结合缓冲液(10mM Tris-HCl ;2M NaCl ;

1mM EDTA ;0.1%吐温 20)、3 μ L 链霉亲和素琼脂糖珠（组成所有权 -GE 医疗集团生物科学部门 (GE Healthcare BioScience AB)) 和 12 μ L ddH₂O。在 96 孔板中将这些混合物摇动 10 分钟。

[0242] 然后采用“真空预装工具”，将各溶液拉穿过连接于小金属管的小过滤器，而此时与生物素化 PCR 产物结合的链霉亲和素珠通过抽吸保留在过滤器上。根据该原理，然后将过滤器浸于 70%乙醇 5 秒钟，以洗涤 DNA 并去除引物、dNTP 和其它 PCR 反应组分。采用 0.2M NaOH 重复该步骤，以使 dsDNA 变性，仅使生物素化 DNA 链与链霉亲和素珠结合。在洗涤缓冲液中对 DNA 进行最终洗涤后，将所述“真空预装工具”置于每孔包含 40 μ L 退火缓冲液和 0.1 μ L Pyrosequencing[®] 引物溶液 (100pmol* μ L* ;用于 AtHPPD 的 kerfi018/ 用于 PfHPPD 的 kerfi015) 的 PSQ[™]96 孔板上。然后关闭真空，将各过滤器浸入其相应的孔中，以溶解过滤器保留的 DNA。然后将孔板置于 80℃温育 2 分钟，以拆分 DNA 模板内最终形成的二级结构。在溶液冷却至室温时，Pyrosequencing[®] 引物与其模板上的结合位点发生杂交。

[0243] 将 Pyrosequencing[®] 反应的剩余组分 (620 μ L 酶混合物、620 μ L 底物混合物和 130 μ L 各 dNTP 溶液) 填充入筒的各个孔中。然后将该筒和 PSQ[™]板置于 PyroMark[™]ID 内。

[0244] Pyrosequencing[®] 设备在测序反应由加入第一 dNTP 启动前自动将酶和底物加入到反应混合物中。为测定引物下游的 DNA 序列，进行 SQA 运行程序。预先定义向反应混合物中加入核苷酸的顺序。可采用 PyroMark[™]ID 软件翻译进入 DNA 序列的 Pyrogram[®] 曲线。

[0245] 结果：

[0246] AtHPPD 的 PCR 扩增片段大小为 239bp，生物素附着于非编码链；PfHPPD 片段包含 142bp，生物素附着于编码链。

[0247] AtHPPD 中的突变密码子位于 kerfi018 引物序列下游三个碱基。测序的第一组三个碱基是腺嘌呤，后面是突变的密码子。AtHPPD 片段的编码链由 DNA 聚合酶合成，因此该序列可直接翻译为氨基酸序列。

[0248] 对 438 个 AtHPPD 菌落进行筛选得到 146 个突变基因、181 个野生型基因（第 422 位的密码子 GGC）和 111 个失败的测序反应或模糊结果。

[0249] 因此在大肠杆菌 K-12 MG1655 或大肠杆菌 TOP10 中转化突变质粒产生突变克隆在所有例子的 33%中取得了成功。可获得编码除赖氨酸外所有氨基酸的密码子。包含谷氨酸、组氨酸、异亮氨酸、苏氨酸、色氨酸和酪氨酸的密码子的基因存在于大肠杆菌 TOP10 克隆中，由其制备 DNA 并转化入大肠杆菌 K-12 MG1655 细胞中。如果可能，可考虑大肠杆菌 K-12 中的密码子使用选择同义密码子。使用频率低于 10%的密码子不予选择，选择最多的密码子使用频率高于 35%（密码子使用数据库；大肠杆菌 K-12 :<http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=83333>）。

[0250] 从引物 kerfi015 开始，PfHPPD 片段的非编码链由 DNA 聚合酶合成，因此该核苷酸序列在可翻译为氨基酸序列前需要翻译为反向互补物。所述突变密码子紧接在引物之后，因此由反应中测序的第一组三个碱基表示。

[0251] 对 252 个 PfHPPD 菌落进行筛选得到 119 个突变基因、73 个未改变基因（第 336 位的密码子 TGG）和 60 个失败的测序反应或模糊结果。

[0252] 因此在大肠杆菌 K-12 MG1655 细胞中转化突变质粒产生突变克隆在所有例子的 47% 中取得了成功。可获得编码除丙氨酸外所有氨基酸的密码子。如果可能,可考虑如上所述 AtHPPD 密码子情况的大肠杆菌 K-12 中的密码子使用选择同义密码子。

[0253]

引物名称	引物序列
kerfi015	5' -GACTCGAACAGCGCCTTGAAGTT-3'
kerfi018	5' -GGATGTGGTGGTTTTGGC-3'

[0254] 实施例 3 :HPPD 活性试验

[0255] HPPD 从 4-HPP 和 O_2 产生尿黑酸和 CO_2 。该酶在有或没有抑制剂存在下与其底物 4-HPP 温育。L-抗坏血酸作为还原剂存在以将活性中心铁保持为亚铁形式,过氧化氢酶存在以降解有毒的 H_2O_2 。在温育 1 小时后,加入 2,4-二硝基苯肼 (DNP) 停止反应。DNP 与碱性 pH 下呈琥珀褐色的试验混合物中的剩余 4-HPP 分子形成腓衍生物。未消耗的 4-HPP 量在 405nm 进行光度测定。

[0256] 对于抑制剂储备溶液的制备,将特波三酮 ($M_w = 440.82$) 和 DKN ($M_w = 359.3$) 溶解于 DMSO 至 10mM 的浓度。首先将该储备溶液在 25% DMSO 中稀释 20 倍至 0.5mM 的浓度。用 ddH_2O 进一步稀释以获得用于试验的抑制剂溶液 (5 μM 、10 μM 和 20 μM)。各抑制剂溶液分别占试验混合物体积的一半,即,其活性浓度再次降低 2 倍。这导致产生 2.5 μM 、5 μM 和 10 μM 的抑制剂浓度。2% DMSO 溶液提供未抑制反应中一半的检测混合物以标准化 DMSO 可能的抑制作用。

[0257] 试验设计用于基于单体的 HPPD 浓度为 444nM 和 4-HPP 浓度为 500 μM 。这对应于 100 μL 试验混合物中的 44.4pmol HPPD 和 50nmol 4-HPP,导致底物相对于酶约过量 1000 倍。AtHPPD 亚基的计算的理论分子量为 49.515kD,导致每份试验混合物 2.2 μg HPPD。PfHPPD 亚基的计算的理论分子量为 41.205kD,导致每份试验混合物 1.8 μg HPPD。该酶溶液提供 1/4 的试验混合物体积,因此通过用 50mM TRIS 缓冲液将 AtHPPD 溶液稀释至 88 $\mu g \cdot mL^{-1}$ 制备酶储备溶液;将 PfHPPD 溶液稀释至 72 $\mu g \cdot mL^{-1}$ 。

[0258] 抑制剂浓度 (2.5 μM 、5 μM 和 10 μM) 提供相比酶量的 5 倍、10 倍和 20 倍的过量抑制剂。制备提供 1/4 试验混合物的缓冲液 / 底物溶液。2.5mL 缓冲液 / 底物溶液包含 1mL 1M TRIS 缓冲液、500 μL 10mM 4-HPP 溶液、500 μL 200mM L-抗坏血酸溶液、13 μL 过氧化氢酶溶液和 487 μL ddH_2O 。

[0259] 在 Greiner F 底 96 孔微板上进行试验,一式三份进行所有反应。每板进行六重对照,包含 25 μL 50mM TRIS 替代 HPPD 溶液 (对应于 4-HPP 的 0% 消耗) 或含 500 μL 1M TRIS 替代 500 μL 10mM 4-HPP (对应于 HPP 的 100% 消耗) 的缓冲液 / 底物溶液。向 50 μL 各抑制剂溶液或 50 μL 2% DMSO 和 25 μL 缓冲液 / 底物溶液的混合物中加入 25 μL HPPD 溶液启动反应。允许反应在室温下进行 1 小时。停止反应,通过加入 50 μL 0.04% DNP/3.8N HCl 溶液诱导 4-HPP 的显色。15 分钟后加入 100 μL 5N KOH,引起腓衍生物的颜色转移。紧接着用 BMG FLUOstar Galaxy 微板分析仪在 405nm 进行光度测定,获得的数据用于分析有和没有抑制剂存在时的 HPPD 活性。

[0260] 结果：

[0261] 在 HPPD 活性试验中检测 SEQ ID N04 所示拟南芥 HPPD 的氨基酸序列第 422 位发生突变的 AtHPPD 突变体（即 Gly422Ala、-Arg、-Asn、-Asp、-Cys、-Glu、-His、-Leu、-Met、-Phe、-Pro、-Ser、-Tyr 和 -Val）和野生型（WT）酶（应注意，SEQ ID N04 所示拟南芥 HPPD 的氨基酸序列的 Gly422 对应于 SEQ ID N0:2 所示假单胞菌参考序列的 Gly336）。所有酶均具有活性，但仅突变体 Gly422Ala、-Asn、-Asp、-Cys、-His、-Met、-Phe、-Tyr 和 -Val 的活性在野生型酶的范围之内或之上（ $\geq 70\%$ ）。野生型酶在 $2.5 \mu\text{M}$ 特波三酮存在下保持其 35% 的活性；仅突变体 Gly422Asn、-Cys、-His 和 -Val 保持较高活性，分别是 39% 、 44% 、 51% 和 43% 。活性在更高的特波三酮浓度下进一步降低。仅突变体 Gly422His 在 5 和 $10 \mu\text{M}$ 特波三酮存在下表现出约 40% 的残余活性，其它所有酶在这些抑制剂浓度下均表现出可与野生型酶相当的活性，分别在约 20% 和 10% （表 1）。

[0262] 测试 PfHPPD 突变体 Gly336Arg、-Asp、-Gln、-Glu、-His、-Leu、-Lys、-Met、-Phe、-Thr、-Trp 和 -Pro 和野生型酶。除未抑制活性处于 70% 野生型活性以下的 Gly336Pro 突变体外，Gly336 突变体的活性在野生型酶的范围之内或之上（ $\geq 75\%$ ）。野生型酶在 $2.5 \mu\text{M}$ 特波三酮存在下仅保持了其 5% 的活性，而突变体 Gly336Asp、-Arg、-Gln、-Glu、-His、-Met、-Phe 和 -Trp 保持了高于 14% 的活性。最高的残余活性来自 Gly336His (26%) 和 Gly336Phe (33%)。有趣的是，Gly336His 突变体在 5 和 $10 \mu\text{M}$ 特波三酮存在下分别表现出 13% 和 11.2% 的残余活性，而 Gly336Phe 的活性分别降低至 12.4% 和 2.5% 。Gly336Met 突变体在这些抑制剂浓度下分别表现出 7% 和 10% 的残余活性，而野生型酶的活性降低至 0 （表 1）。

[0263] 表 1：有和没有特波三酮存在时 PfHPPD 和 AtHPPD 突变体的相对活性（百分比）；通过将未抑制酶活性设为 100% 对活性进行标准化

[0264] 荧光假单胞菌 HPPD

[0265]

Gly336 突 变体	特波三酮浓度 (μM)			
	0	2,5	5	10
Arg	100	14	7	2
Asp	100	18	9	0
Gln	100	14	0	0
Glu	100	15	7	0
Gly	100	5	0	0
His	100	26	13	11
Leu	100	4	0	0
Lys	100	6	0	0
Met	100	16	7	10
Phe	100	33	12	3
Pro	100	5	4	0
Thr	100	8	2	2
Trp	100	21	7	0

[0266] 拟南芥 HPPD

[0267]

Gly422 突 变体*	特波三酮浓度 (μM)			
	0	2,5	5	10
Ala	100	25	21	15
Arg	100	17	1	1
Asn	100	39	26	15
Asp	100	20	7	10
Cys	100	44	27	19
Glu	100	24	24	0
Gly	100	35	21	12
His	100	50	31	40
Leu	100	31	23	14
Met	100	18	13	12
Phe	100	30	16	11
Pro	100	0	0	0
Ser	100	18	4	0
Tyr	100	26	11	0
Val	100	43	22	14

[0268] *SEQ ID N04 所示拟南芥 HPPD 的氨基酸序列第 422 位 gly 处的突变 (对应于 SEQ ID N02 所示假单胞菌 HPPD 的氨基酸序列的 Gly336)

[0269] 实施例 4 :PDH 活性试验

[0270] 在 25°C 下,在包含 50mM tris-HCl、pH 8.6、300 μM 预苯酸和 1mM NAD 或 NADP 总体积为 200 μl 的溶液中,通过在 340nm 用分光光度法监测 NADH 或 NADPH 的形成对预苯酸脱氢酶活性进行测定。

[0271] 实施例 3 :构建嵌合基因以便评价烟草中未突变和突变的 PfHPPD。

[0272] A) 所述嵌合基因的构建 :

[0273] 为制备在 PBD6 型烟草植物中表达的 HPPD (野生型或突变体) 的构建体采用的载体指定为 pRP-RD224。起初设计该载体用于仅通过取代 KpnI 和 BstEII 位点之间该载体的截短 HPPD 基因来克隆所有的假单胞菌 HPPD 突变体。从二元载体 pBI121 (克隆技术公司 (Clontech)) 对其进行构建在 W0 99/24585 中有广泛描述。

[0274] 因此克隆 pRP-RD224 具有以下结构 :

[0275] RB/Nos 启动子 /NPTII/Nos 终止子 / 双组蛋白启动子 /tev/otp/ 截短的 HPPD/Nos 终止子 /LB

[0276] 其中“截短的 HPPD”指截短了约 500 个碱基对的编码 PfHPPD 的序列,以随后促进

筛选整合了突变 HPPD 的转化菌落 (W099/24585)。

[0277] pRP-RD224 突变体 :用 KpnI 和 BstEII 消化携带突变和未突变 HPPD 的载体 DNA,纯化,然后连接入已用 KpnI 和 BstEII 消化并纯化的载体 pRP-RD224。通过用 KpnI 和 BstEII 消化来选择对整合了突变 HPPD 基因的转化体的插入体大小。所得克隆指定为 pRP-RD224,向其加入发生于 HPPD 上的突变类型 ;以该方式创建了以下克隆 :pRP RD224 Pf(未突变酶)、pRP RD224 PfH336(在第 336 位具有组氨酸的酶)、pRP RD224 PfM336(在第 336 位具有甲硫氨酸的酶)和 pRP RD224 PfF336(在第 336 位具有苯丙氨酸的酶)。

[0278] 实施例 4 :过表达 PDH 嵌合基因的构建

[0279] 构建过表达 PDH 的嵌合基因包括在转录方向上装配如专利申请 EP 0 507 698 所述的“双组蛋白”启动子 (PdH4)、Carrington 和 Freed(1990) 中描述的烟草蚀纹病毒翻译增强子 (TEV) 序列、如专利申请 EP 0 508 909 所述的编码优化转运肽 (OTP) 的序列、Mannhaupt 等 (1989) 中描述的酵母 PDH 基因的编码部分和 Bevan 等 (1983) 中描述的胭脂碱合成酶基因的 nos 终止子。然后将该装配物克隆入包含卡那霉素耐受性基因 (NPTII) 的双元载体 pRD 224,以提供载体 pRD 224-PDH。

[0280] 然后该双元载体用于转化农杆菌菌株 EHA 105 以提供农杆菌菌株 EHA 105-pRD 224-PDH。该农杆菌菌株用于转化用如实施例 3 所述的嵌合基因转化的烟草植物。

[0281] 用卡那霉素选择转化植物。

[0282] 引用的参考文献 :

[0283] Abou-Zeid 等,1995,Applied Env Microb 41:1298-1302

[0284] Ausubel F.M. 等,《现代分子生物学实验技术》(“Current Protocols in Molecular Biology”)第 1 和 2 卷,格林出版公司 (Greene Publishing Associates) 和 WI 公司 (Wiley Interscience) (1989)

[0285] Bevan 等,1983,Nucleic Acids Res. 11(2),369-385

[0286] Bonner 等,1995,Plant Cells Physiol. 36,1013-1022

[0287] Byng 等,1981,Phytochemistry 6:1289-1292

[0288] Carrington 和 Freed,1990 ;J. Virol. 64:1590-1597

[0289] Christou 等,1991,Biotechnology 9:957

[0290] Connely 和 Conn,1986,Z. Naturforsch 41c:69-78

[0291] Crouch N. P. 等,1997,Tetrahedron, 53, 20, 6993-7010

[0292] Datla, R. 等,1997,Biotechnology Ann. Rev. 3, 269-296

[0293] Gaines 等,1982,Plants 156:233-240

[0294] Henner 等,1986,Gene 49(1)147-152

[0295] Hiei 等,1994,Plant J 6:271-282

[0296] Hiei 等,1997,Plant Mol Biol. 35:205-21

[0297] Hudson 等,1984,J. Mol. Biol. 180(4), 1023-1051

[0298] Fritze 等,2004,Plant Physiology 134:1388-1400

[0299] Garcia 等,1997,Biochem. J. 325, 761-769

[0300] Garcia 等,1999,Plant Physiol. 119, 1507-1516

[0301] Gould J. H. 和 Magallanes-Cedeno M.,1998,Plant Molecular Biology

reporter, 16:1-10

[0302] Horsch 等, 1985, Science 227:1229-1231

[0303] Lingens 等, 1967, European J. Biochem 1:363-374

[0304] Maniatis T., Fritsch E. F., 刊于《分子克隆》(Molecular cloning), Sambrook, 1982.

[0305] Mannhaupt 等, 1989, Gene 85, 303-311

[0306] Matringe 等, 2005, Pest Management Science 61:269-276

[0307] Mitchell 等, 2001, Pest Management Science 57:120-128

[0308] Pallett 等, 2001, Pest Management Science 57:133-142

[0309] Sambrook 等, 1989, 《分子克隆:实验手册》(Molecular cloning:a laboratory manual), 第二版, 冷泉港实验室出版社, 冷泉港, 纽约 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York)

[0310] Sambrook J. 和 Russell D., 2001, 《分子克隆:实验手册》(Molecular Cloning:A laboratory Manual) (第三版), ISBN 978-087969577-4 CSHL 出版社

[0311] Sampathkumar 和 Morrisson, 1982, Bioch Biophys Acta 701:204-211

[0312] Sawahel W. A., 2001, Plant Molecular Biology reporter, 19:377a-377f

[0313] Schulz 等, 1993, FEBS Letters 318:162-166

[0314] Xia 等, 1992, J. Gen. Microbiol. 138(7), 1309-1316

[0315] Zapata C., 1999, theoretical Applied Genetics, 98(2):1432-2242

[0001]

序列表

<110> 拜耳农科股份公司(Bayer Cropscience AG)
拜耳生物科学股份有限公司(Bayer Bioscience NV)

<120> 新的突变羟基苯基丙酮酸双加氧酶, DNA序列和耐受HPPD抑制剂除草剂的植物分离

<130> BCS 08-4008

<160> 32

<170> PatentIn 版本3.4

<210> 1
<211> 1077
<212> DNA
<213> 荧光假单胞菌 (Pseudomonas fluorescens)

<220>
<221> CDS
<222> (1)..(1077)

<400> 1

atg gca gat cta tac gaa aac cca atg ggc ctg atg ggc ttt gaa ttc	48
Met Ala Asp Leu Tyr Glu Asn Pro Met Gly Leu Met Gly Phe Glu Phe	
1 5 10 15	
atc gaa ttc gcg tgc ccg acg ccg ggt acc ctg gag ccg atc ttc gag	96
Ile Glu Phe Ala Ser Pro Thr Pro Gly Thr Leu Glu Pro Ile Phe Glu	
20 25 30	
atc atg ggc ttc acc aaa gtc gcg acc cac cgt tcc aag aac gtg cac	144
Ile Met Gly Phe Thr Lys Val Ala Thr His Arg Ser Lys Asn Val His	
35 40 45	
ctg tac cgc cag ggc gag atc aac ctg atc ctc aac aac gag ccc aac	192
Leu Tyr Arg Gln Gly Glu Ile Asn Leu Ile Leu Asn Asn Glu Pro Asn	
50 55 60	
agc atc gcc tcc tac ttt gcg gcc gaa cac ggc ccg tgc gtg tgc gcc	240
Ser Ile Ala Ser Tyr Phe Ala Ala Glu His Gly Pro Ser Val Cys Gly	
65 70 75 80	
atg gcg ttc cgc gtg aag gac tgc caa aag gcc tac aac cgc gcc ctg	288
Met Ala Phe Arg Val Lys Asp Ser Gln Lys Ala Tyr Asn Arg Ala Leu	
85 90 95	
gaa ctc ggc gcc cag ccg atc cat att gac acc ggg ccg atg gaa ttg	336
Glu Leu Gly Ala Gln Pro Ile His Ile Asp Thr Gly Pro Met Glu Leu	
100 105 110	
aac ctg ccg gcg atc aag ggc atc ggc ggc gcg ccg ttg tac ctg atc	384
Asn Leu Pro Ala Ile Lys Gly Ile Gly Gly Ala Pro Leu Tyr Leu Ile	
115 120 125	
gac cgt ttc ggc gaa ggc agc tgc atc tac gac atc gac ttc gtg tac	432
Asp Arg Phe Gly Glu Gly Ser Ser Ile Tyr Asp Ile Asp Phe Val Tyr	
130 135 140	
ctc gaa ggt gtg gag cgc aat ccg gtc ggt gea ggt ctc aaa gtc atc	480
Leu Glu Gly Val Glu Arg Asn Pro Val Gly Ala Gly Leu Lys Val Ile	
145 150 155 160	
gac cac ctg acc cac aac gtc tat cgc ggc cgc atg gtc tac tgg gcc	528
Asp His Leu Thr His Asn Val Tyr Arg Gly Arg Met Val Tyr Trp Ala	

[0002]

	165	170	175	
aac ttc tac gag aaa ttg ttc aac ttc cgt gaa ggc egt tac ttc gat				576
Asn Phe Tyr Glu Lys Leu Phe Asn Phe Arg Glu Ala Arg Tyr Phe Asp	180	185	190	
atc aag ggc gag tac acc ggc ctg act tcc aag gcc atg agt ggc cgg				624
Ile Lys Gly Glu Tyr Thr Gly Leu Thr Ser Lys Ala Met Ser Ala Pro	195	200	205	
gac ggc atg atc cgc atc cgg ctg aac gaa gag tcg tcc aag ggc ggc				672
Asp Gly Met Ile Arg Ile Pro Leu Asn Glu Glu Ser Ser Lys Gly Ala	210	215	220	
ggg cag atc gaa gag ttc ctg atg cag ttc aac ggc gaa ggc atc cag				720
Gly Gln Ile Glu Glu Phe Leu Met Gln Phe Asn Gly Glu Gly Ile Gln	225	230	235	240
cac gtg gcg ttc ctc acc gac gac ctg gtc aag acc tgg gac gcg ttg				768
His Val Ala Phe Leu Thr Asp Asp Leu Val Lys Thr Trp Asp Ala Leu	245	250	255	
aag aaa atc ggc atg cgc ttc atg acc ggc cgg cea gac act tat tac				816
Lys Lys Ile Gly Met Arg Phe Met Thr Ala Pro Pro Asp Thr Tyr Tyr	260	265	270	
gaa atg ctc gaa ggc cgc ctg cct gac cac ggc gag ccg gtg gat caa				864
Glu Met Leu Glu Gly Arg Leu Pro Asp His Gly Glu Pro Val Asp Gln	275	280	285	
ctg cag gca cgc ggt atc ctg ctg gac gga tct tcc gtg gaa ggc gac				912
Leu Gln Ala Arg Gly Ile Leu Leu Asp Gly Ser Ser Val Glu Gly Asp	290	295	300	
aaa cgc ctg ctg ctg cag atc ttc tcg gaa acc ctg atg ggc ccg gtg				960
Lys Arg Leu Leu Leu Gln Ile Phe Ser Glu Thr Leu Met Gly Pro Val	305	310	315	320
ttc ttc gaa ttc atc cag cgc aag ggc gac gat ggg ttt ggc gag ggg				1008
Phe Phe Glu Phe Ile Gln Arg Lys Gly Asp Asp Gly Phe Gly Glu Gly	325	330	335	
aac ttc aag gcg ctg ttc gag tcc atc gaa cgt gac cag gtg cgt cgt				1056
Asn Phe Lys Ala Leu Phe Glu Ser Ile Glu Arg Asp Gln Val Arg Arg	340	345	350	
ggt gta ttg acc gcc gat taa				1077
Gly Val Leu Thr Ala Asp	355			

<210> 2
 <211> 358
 <212> PRT
 <213> 荧光假单胞菌 (Pseudomonas fluorescens)

<400> 2

Met Ala Asp Leu Tyr Glu Asn Pro Met Gly Leu Met Gly Phe Glu Phe	1	5	10	15
Ile Glu Phe Ala Ser Pro Thr Pro Gly Thr Leu Glu Pro Ile Phe Glu	20	25	30	
Ile Met Gly Phe Thr Lys Val Ala Thr His Arg Ser Lys Asn Val His	35	40	45	

[0003]

Leu Tyr Arg Gln Gly Glu Ile Asn Leu Ile Leu Asn Asn Glu Pro Asn
 50 55 60
 Ser Ile Ala Ser Tyr Phe Ala Ala Glu His Gly Pro Ser Val Cys Gly
 65 70 75 80
 Met Ala Phe Arg Val Lys Asp Ser Gln Lys Ala Tyr Asn Arg Ala Leu
 85 90 95
 Glu Leu Gly Ala Gln Pro Ile His Ile Asp Thr Gly Pro Met Glu Leu
 100 105 110
 Asn Leu Pro Ala Ile Lys Gly Ile Gly Gly Ala Pro Leu Tyr Leu Ile
 115 120 125
 Asp Arg Phe Gly Glu Gly Ser Ser Ile Tyr Asp Ile Asp Phe Val Tyr
 130 135 140
 Leu Glu Gly Val Glu Arg Asn Pro Val Gly Ala Gly Leu Lys Val Ile
 145 150 155 160
 Asp His Leu Thr His Asn Val Tyr Arg Gly Arg Met Val Tyr Trp Ala
 165 170 175
 Asn Phe Tyr Glu Lys Leu Phe Asn Phe Arg Glu Ala Arg Tyr Phe Asp
 180 185 190
 Ile Lys Gly Glu Tyr Thr Gly Leu Thr Ser Lys Ala Met Ser Ala Pro
 195 200 205
 Asp Gly Met Ile Arg Ile Pro Leu Asn Glu Glu Ser Ser Lys Gly Ala
 210 215 220
 Gly Gln Ile Glu Glu Phe Leu Met Gln Phe Asn Gly Glu Gly Ile Gln
 225 230 235 240
 His Val Ala Phe Leu Thr Asp Asp Leu Val Lys Thr Trp Asp Ala Leu
 245 250 255
 Lys Lys Ile Gly Met Arg Phe Met Thr Ala Pro Pro Asp Thr Tyr Tyr
 260 265 270
 Glu Met Leu Glu Gly Arg Leu Pro Asp His Gly Glu Pro Val Asp Gln
 275 280 285
 Leu Gln Ala Arg Gly Ile Leu Leu Asp Gly Ser Ser Val Glu Gly Asp
 290 295 300
 Lys Arg Leu Leu Leu Gln Ile Phe Ser Glu Thr Leu Met Gly Pro Val
 305 310 315 320
 Phe Phe Glu Phe Ile Gln Arg Lys Gly Asp Asp Gly Phe Gly Glu Gly
 325 330 335
 Asn Phe Lys Ala Leu Phe Glu Ser Ile Glu Arg Asp Gln Val Arg Arg
 340 345 350
 Gly Val Leu Thr Ala Asp
 355

<210> 3
 <211> 1338
 <212> DNA
 <213> 拟南芥 (*Arabidopsis thaliana*)

[0004]

<220>
 <221> CDS
 <222> (1)..(1338)

<400> 3

atg ggc cac caa aac gcc gcc gtt tca gag aat caa aac cat gat gac	48
Met Gly His Gln Asn Ala Ala Val Ser Glu Asn Gln Asn His Asp Asp	
1 5 10 15	
ggc get gcg tgc tgc cgg gga ttc aag ctc gtc gga ttt tcc aag ttc	96
Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val Gly Phe Ser Lys Phe	
20 25 30	
gta aga aag aat cca aag tct gat aaa ttc aag gtt aag cgc ttc cat	144
Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys Val Lys Arg Phe His	
35 40 45	
cac atc gag ttc tgg tgc gcc gac gca acc aac gtc gct cgt cgc ttc	192
His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Val Ala Arg Arg Phe	
50 55 60	
tcc tgg ggt ctg ggg atg aga ttc tcc gcc aaa tcc gat ctt tcc acc	240
Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys Ser Asp Leu Ser Thr	
65 70 75 80	
gga aac atg gtt cac gcc tct tac cta ctc acc tcc ggt gac ctc cga	288
Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr Ser Gly Asp Leu Arg	
85 90 95	
ttc ctt ttc act get cct tac tct ccg tct ctc tcc gcc gga gag att	336
Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu Ser Ala Gly Glu Ile	
100 105 110	
aaa ccg aca acc aca gct tct atc cca agt ttc gat cac gcc tct tgt	384
Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe Asp His Gly Ser Cys	
115 120 125	
egt tcc ttc ttc tct tca cat ggt ctc ggt gtt aga gcc gtt gcg att	432
Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val Arg Ala Val Ala Ile	
130 135 140	
gaa gta gaa gac gca gag tca gct ttc tcc atc agt gta get aat gcc	480
Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile Ser Val Ala Asn Gly	
145 150 155 160	
get att cct tgc tgc cct cct atc gtc ctc aat gaa gca gtt acg atc	528
Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn Glu Ala Val Thr Ile	
165 170 175	
get gag gtt aaa cta tac gcc gat gtt gtt ctc cga tat gtt agt tac	576
Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr	
180 185 190	
aaa gca gaa gat acc gaa aaa tcc gaa ttc ttg cca ggg ttc gag cgt	624
Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu Pro Gly Phe Glu Arg	
195 200 205	
gta gag gat gcg tgc tgc ttc cca ttg gat tat ggt atc cgg cgg ctt	672
Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr Gly Ile Arg Arg Leu	
210 215 220	
gac cac gcc gtg gga aac gtt cct gag ctt ggt ccg gct tta act tat	720
Asp His Ala Val Gly Asn Val Pro Glu Leu Gly Pro Ala Leu Thr Tyr	
225 230 235 240	

[0005]

gta gcg ggg ttc act ggt ttt cac caa ttc gca gag ttc aca gca gac Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala Glu Phe Thr Ala Asp 245 250 255	768
gac gtt gga acc gcc gag agc ggt tta aat tca gcg gtc ctg gct agc Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser Ala Val Leu Ala Ser 260 265 270	816
aat gat gaa atg gtt ctt cta ccg att aac gag cca gtg cac gga aca Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu Pro Val His Gly Thr 275 280 285	864
aag agg aag agt cag att cag acg tat ttg gaa cat aac gaa ggc gca Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu His Asn Glu Gly Ala 290 295 300	912
ggg cta caa cat ctg gct ctg atg agt gaa gac ata ttc agg acc ctg Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp Ile Phe Arg Thr Leu 305 310 315 320	960
aga gag atg agg aag agg agc agt att gga gga ttc gac ttc atg cct Arg Glu Met Arg Lys Arg Ser Ser Ile Gly Gly Phe Asp Phe Met Pro 325 330 335	1008
tct cct ccg cct act tac tac cag aat ctc aag aaa cgg gtc ggc gac Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys Lys Arg Val Gly Asp 340 345 350	1056
gtg ctc agc gat gat cag atc aag gag tgt gag gaa tta ggg att ctt Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu Glu Leu Gly Ile Leu 355 360 365	1104
gta gac aga gat gat caa ggg acg ttg ctt caa ate ttc aca aaa cca Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln Ile Phe Thr Lys Pro 370 375 380	1152
cta ggt gac agg ccg acg ata ttt ata gag ata ate cag aga gta gga Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile Ile Gln Arg Val Gly 385 390 395 400	1200
tgc atg atg aaa gat gag gaa ggg aag gct tac cag agt gga gga tgt Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr Gln Ser Gly Gly Cys 405 410 415	1248
ggt ggt ttt ggc aaa ggc aat ttc tct gag ctc ttc aag tcc att gaa Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile Glu 420 425 430	1296
gaa tac gaa aag act ctt gaa gcc aaa cag tta gtg gga tga Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu Val Gly 435 440 445	1338
<210> 4 <211> 445 <212> PRT <213> 拟南芥 (<i>Arabidopsis thaliana</i>)	
<400> 4 Met Gly His Gln Asn Ala Ala Val Ser Glu Asn Gln Asn His Asp Asp 1 5 10 15 Gly Ala Ala Ser Ser Pro Gly Phe Lys Leu Val Gly Phe Ser Lys Phe 20 25 30	

[0006]

Val Arg Lys Asn Pro Lys Ser Asp Lys Phe Lys Val Lys Arg Phe His
 35 40 45
 His Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Val Ala Arg Arg Phe
 50 55 60
 Ser Trp Gly Leu Gly Met Arg Phe Ser Ala Lys Ser Asp Leu Ser Thr
 65 70 75 80
 Gly Asn Met Val His Ala Ser Tyr Leu Leu Thr Ser Gly Asp Leu Arg
 85 90 95
 Phe Leu Phe Thr Ala Pro Tyr Ser Pro Ser Leu Ser Ala Gly Glu Ile
 100 105 110
 Lys Pro Thr Thr Thr Ala Ser Ile Pro Ser Phe Asp His Gly Ser Cys
 115 120 125
 Arg Ser Phe Phe Ser Ser His Gly Leu Gly Val Arg Ala Val Ala Ile
 130 135 140
 Glu Val Glu Asp Ala Glu Ser Ala Phe Ser Ile Ser Val Ala Asn Gly
 145 150 155 160
 Ala Ile Pro Ser Ser Pro Pro Ile Val Leu Asn Glu Ala Val Thr Ile
 165 170 175
 Ala Glu Val Lys Leu Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr
 180 185 190
 Lys Ala Glu Asp Thr Glu Lys Ser Glu Phe Leu Pro Gly Phe Glu Arg
 195 200 205
 Val Glu Asp Ala Ser Ser Phe Pro Leu Asp Tyr Gly Ile Arg Arg Leu
 210 215 220
 Asp His Ala Val Gly Asn Val Pro Glu Leu Gly Pro Ala Leu Thr Tyr
 225 230 235 240
 Val Ala Gly Phe Thr Gly Phe His Gln Phe Ala Glu Phe Thr Ala Asp
 245 250 255
 Asp Val Gly Thr Ala Glu Ser Gly Leu Asn Ser Ala Val Leu Ala Ser
 260 265 270
 Asn Asp Glu Met Val Leu Leu Pro Ile Asn Glu Pro Val His Gly Thr
 275 280 285
 Lys Arg Lys Ser Gln Ile Gln Thr Tyr Leu Glu His Asn Glu Gly Ala
 290 295 300
 Gly Leu Gln His Leu Ala Leu Met Ser Glu Asp Ile Phe Arg Thr Leu
 305 310 315 320
 Arg Glu Met Arg Lys Arg Ser Ser Ile Gly Gly Phe Asp Phe Met Pro
 325 330 335
 Ser Pro Pro Pro Thr Tyr Tyr Gln Asn Leu Lys Lys Arg Val Gly Asp
 340 345 350
 Val Leu Ser Asp Asp Gln Ile Lys Glu Cys Glu Glu Leu Gly Ile Leu
 355 360 365
 Val Asp Arg Asp Asp Gln Gly Thr Leu Leu Gln Ile Phe Thr Lys Pro

[0007]

370	375	380	
Leu Gly Asp Arg Pro Thr Ile Phe Ile Glu Ile Ile Gln Arg Val Gly			
385	390	395	400
Cys Met Met Lys Asp Glu Glu Gly Lys Ala Tyr Gln Ser Gly Gly Cys			
	405	410	415
Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile Glu			
	420	425	430
Glu Tyr Glu Lys Thr Leu Glu Ala Lys Gln Leu Val Gly			
	435	440	445
<210> 5			
<211> 1182			
<212> DNA			
<213> 小鼠 (Mus musculus)			
<220>			
<221> CDS			
<222> (1)..(1182)			
<400> 5			
atg aca acc tac aac aac aaa gga cca aag cct gag aga ggc cgg ttc			48
Met Thr Thr Tyr Asn Asn Lys Gly Pro Lys Pro Glu Arg Gly Arg Phe			
1	5	10	15
ctc cat ttc cac tcg gtg acc ttc tgg gtt ggc aat gcc aag cag gct			96
Leu His Phe His Ser Val Thr Phe Trp Val Gly Asn Ala Lys Gln Ala			
	20	25	30
gct tcc ttc tac tgc aac aag atg ggc ttt gaa cct ctg gcc tac agg			144
Ala Ser Phe Tyr Cys Asn Lys Met Gly Phe Glu Pro Leu Ala Tyr Arg			
	35	40	45
ggc cta gag act ggc tcc cgg gag gta gtc agc cac gtc atc aag caa			192
Gly Leu Glu Thr Gly Ser Arg Glu Val Val Ser His Val Ile Lys Gln			
	50	55	60
ggg aaa att gtg ttt gtt ctc tgc tct gct etc aat ccc tgg aac aaa			240
Gly Lys Ile Val Phe Val Leu Cys Ser Ala Leu Asn Pro Trp Asn Lys			
65	70	75	80
gag atg ggc gac cac ttg gtg aag cat ggc gac ggg gtg aaa gac atc			288
Glu Met Gly Asp His Leu Val Lys His Gly Asp Gly Val Lys Asp Ile			
	85	90	95
gca ttc gag gtg gaa gac tgc gac cac att gtg cag aaa gct cga gaa			336
Ala Phe Glu Val Glu Asp Cys Asp His Ile Val Gln Lys Ala Arg Glu			
	100	105	110
egg ggc gcc aaa att gtg cgg gag cca tgg gtg gag caa gac aaa ttt			384
Arg Gly Ala Lys Ile Val Arg Glu Pro Trp Val Glu Gln Asp Lys Phe			
	115	120	125
ggg aag gtg aag ttt gct gtg ctg cag acg tat gga gat acc aca cac			432
Gly Lys Val Lys Phe Ala Val Leu Gln Thr Tyr Gly Asp Thr Thr His			
	130	135	140
acc ctg gtg gag aag atc aac tac act ggc cgt ttc tta cct gga ttc			480
Thr Leu Val Glu Lys Ile Asn Tyr Thr Gly Arg Phe Leu Pro Gly Phe			
145	150	155	160
gag gcc cca aca tac aag gat acc ctg ctt cca aaa cta ccc aga tgt			528

[0008]

Glu Ala Pro Thr Tyr Lys Asp Thr Leu Leu Pro Lys Leu Pro Arg Cys	
165	170
175	
aac ctt gag atc att gac cac att gta ggc aac caa ccc gac caa gaa	576
Asn Leu Glu Ile Ile Asp His Ile Val Gly Asn Gln Pro Asp Gln Glu	
180	185
190	
atg cag tct gcc tca gaa tgg tac ctg aaa aac ctg cag ttc cac cgg	624
Met Gln Ser Ala Ser Glu Trp Tyr Leu Lys Asn Leu Gln Phe His Arg	
195	200
205	
ttc tgg tcc gtg gac gac acg cag gtg cac acg gag tac agc tct ctg	672
Phe Trp Ser Val Asp Asp Thr Gln Val His Thr Glu Tyr Ser Ser Leu	
210	215
220	
cgc tcc att gtg gtg acc aac tac gag gaa tcc atc aaa atg ccc atc	720
Arg Ser Ile Val Val Thr Asn Tyr Glu Glu Ser Ile Lys Met Pro Ile	
225	230
235	240
aac gag cca gct ccg ggc agg aag aag tct cag atc cag gaa tat gtg	768
Asn Glu Pro Ala Pro Gly Arg Lys Lys Ser Gln Ile Gln Glu Tyr Val	
245	250
255	
gac tat aat ggg ggt gct ggg gtc cag cac atc gct ctc aag acg gaa	816
Asp Tyr Asn Gly Gly Ala Gly Val Gln His Ile Ala Leu Lys Thr Glu	
260	265
270	
gac atc atc aca gca atc cgc cac ttg agg gag cga ggc acg gag ttc	864
Asp Ile Ile Thr Ala Ile Arg His Leu Arg Glu Arg Gly Thr Glu Phe	
275	280
285	
ttg gcc gcc cca tct tct tac tac aaa ctg ctt cgg gag aat ctc aag	912
Leu Ala Ala Pro Ser Ser Tyr Tyr Lys Leu Leu Arg Glu Asn Leu Lys	
290	295
300	
tca gcc aag atc cag gtg aaa gag agc atg gac gtc ctg gag gag ctg	960
Ser Ala Lys Ile Gln Val Lys Glu Ser Met Asp Val Leu Glu Glu Leu	
305	310
315	320
cat atc cta gtc gac tat gac gag aaa ggc tac ctc cta cag atc ttc	1008
His Ile Leu Val Asp Tyr Asp Glu Lys Gly Tyr Leu Leu Gln Ile Phe	
325	330
335	
acc aag ccc atg cag gac cgg ccc aca ctc ttc ctg gaa gtc att caa	1056
Thr Lys Pro Met Gln Asp Arg Pro Thr Leu Phe Leu Glu Val Ile Gln	
340	345
350	
cgt cac aac cac cag ggc ttt gga gcg ggc aac ttc aac tct ctg ttc	1104
Arg His Asn His Gln Gly Phe Gly Ala Gly Asn Phe Asn Ser Leu Phe	
355	360
365	
aag gcg ttc gag gag gag caa gcc cta cgg ggc aac ctc act gac ctg	1152
Lys Ala Phe Glu Glu Glu Gln Ala Leu Arg Gly Asn Leu Thr Asp Leu	
370	375
380	
gag ccc aat ggt gtg agg tct gga atg taa	1182
Glu Pro Asn Gly Val Arg Ser Gly Met	
385	390
<210> 6	
<211> 393	
<212> PRT	
<213> 小鼠 (Mus musculus)	
<400> 6	

[0009]

Met Thr Thr Tyr Asn Asn Lys Gly Pro Lys Pro Glu Arg Gly Arg Phe
 1 5 10 15
 Leu His Phe His Ser Val Thr Phe Trp Val Gly Asn Ala Lys Gln Ala
 20 25 30
 Ala Ser Phe Tyr Cys Asn Lys Met Gly Phe Glu Pro Leu Ala Tyr Arg
 35 40 45
 Gly Leu Glu Thr Gly Ser Arg Glu Val Val Ser His Val Ile Lys Gln
 50 55 60
 Gly Lys Ile Val Phe Val Leu Cys Ser Ala Leu Asn Pro Trp Asn Lys
 65 70 75 80
 Glu Met Gly Asp His Leu Val Lys His Gly Asp Gly Val Lys Asp Ile
 85 90 95
 Ala Phe Glu Val Glu Asp Cys Asp His Ile Val Gln Lys Ala Arg Glu
 100 105 110
 Arg Gly Ala Lys Ile Val Arg Glu Pro Trp Val Glu Gln Asp Lys Phe
 115 120 125
 Gly Lys Val Lys Phe Ala Val Leu Gln Thr Tyr Gly Asp Thr Thr His
 130 135 140
 Thr Leu Val Glu Lys Ile Asn Tyr Thr Gly Arg Phe Leu Pro Gly Phe
 145 150 155 160
 Glu Ala Pro Thr Tyr Lys Asp Thr Leu Leu Pro Lys Leu Pro Arg Cys
 165 170 175
 Asn Leu Glu Ile Ile Asp His Ile Val Gly Asn Gln Pro Asp Gln Glu
 180 185 190
 Met Gln Ser Ala Ser Glu Trp Tyr Leu Lys Asn Leu Gln Phe His Arg
 195 200 205
 Phe Trp Ser Val Asp Asp Thr Gln Val His Thr Glu Tyr Ser Ser Leu
 210 215 220
 Arg Ser Ile Val Val Thr Asn Tyr Glu Glu Ser Ile Lys Met Pro Ile
 225 230 235 240
 Asn Glu Pro Ala Pro Gly Arg Lys Lys Ser Gln Ile Gln Glu Tyr Val
 245 250 255
 Asp Tyr Asn Gly Gly Ala Gly Val Gln His Ile Ala Leu Lys Thr Glu
 260 265 270
 Asp Ile Ile Thr Ala Ile Arg His Leu Arg Glu Arg Gly Thr Glu Phe
 275 280 285
 Leu Ala Ala Pro Ser Ser Tyr Tyr Lys Leu Leu Arg Glu Asn Leu Lys
 290 295 300
 Ser Ala Lys Ile Gln Val Lys Glu Ser Met Asp Val Leu Glu Glu Leu
 305 310 315 320
 His Ile Leu Val Asp Tyr Asp Glu Lys Gly Tyr Leu Leu Gln Ile Phe
 325 330 335
 Thr Lys Pro Met Gln Asp Arg Pro Thr Leu Phe Leu Glu Val Ile Gln

[0010]

	340		345		350	
Arg His Asn His Gln Gly Phe Gly Ala Gly Asn Phe Asn Ser Leu Phe						
	355		360		365	
Lys Ala Phe Glu Glu Glu Gln Ala Leu Arg Gly Asn Leu Thr Asp Leu						
	370		375		380	
Glu Pro Asn Gly Val Arg Ser Gly Met						
	385		390			
<210>	7					
<211>	1200					
<212>	DNA					
<213>	粗球孢子菌 (Coccidioides immitis)					
<220>						
<221>	CDS					
<222>	(1).. (1200)					
<400>	7					
atg gca cca gcc gct gac tcc ccg acg ctt caa ccc gcc cag ccc tct						48
Met Ala Pro Ala Ala Asp Ser Pro Thr Leu Gln Pro Ala Gln Pro Ser						
1	5		10		15	
gat ctc aat cag tat aga gga tac gac cac gtc cac tgg tat gtc gga						96
Asp Leu Asn Gln Tyr Arg Gly Tyr Asp His Val His Trp Tyr Val Gly						
	20		25		30	
aac get aag cag gcc get acc tac tat gtc act cgc atg ggt ttc gag						144
Asn Ala Lys Gln Ala Ala Thr Tyr Tyr Val Thr Arg Met Gly Phe Glu						
	35		40		45	
aga gta gcc tat cgc gga ttg gag act ggc tcc aaa gcg gtg gcc tcg						192
Arg Val Ala Tyr Arg Gly Leu Glu Thr Gly Ser Lys Ala Val Ala Ser						
	50		55		60	
cat gtt gtg cga aac gga aac atc acc ttc atc ttg act tcg ccc ctt						240
His Val Val Arg Asn Gly Asn Ile Thr Phe Ile Leu Thr Ser Pro Leu						
	65		70		75	80
cga tcc gtt gag cag gct tct cgt ttc ccc gag gac gag gct ctc ctg						288
Arg Ser Val Glu Gln Ala Ser Arg Phe Pro Glu Asp Glu Ala Leu Leu						
	85		90		95	
aag gag atc cac gcc cat ctc gag aga cac ggc gat ggt gtc aag gac						336
Lys Glu Ile His Ala His Leu Glu Arg His Gly Asp Gly Val Lys Asp						
	100		105		110	
gtc gcc ttc gag gtc gac tgc gta gag tct gtc ttc tcg get gcc gtt						384
Val Ala Phe Glu Val Asp Cys Val Glu Ser Val Phe Ser Ala Ala Val						
	115		120		125	
agg aac ggt gct gag gtt gtt tcc gat gtc aga acg gtt gaa gat gag						432
Arg Asn Gly Ala Glu Val Val Ser Asp Val Arg Thr Val Glu Asp Glu						
	130		135		140	
gat ggc cag atc aag atg gcg acc atc cga act tat ggc gag acc act						480
Asp Gly Gln Ile Lys Met Ala Thr Ile Arg Thr Tyr Gly Glu Thr Thr						
	145		150		155	160
caac acc ctc atc gaa aga tcc ggc tac agg ggc gga ttc atg ccg gga						528
His Thr Leu Ile Glu Arg Ser Gly Tyr Arg Gly Gly Phe Met Pro Gly						
	165		170		175	

[0011]

tac cgg atg gag agc aat gcc gac gcc act tcc aag ttc ctt cca aag	576
Tyr Arg Met Glu Ser Asn Ala Asp Ala Thr Ser Lys Phe Leu Pro Lys	
180 185 190	
gtt gtg ctt gag aga ata gac cac tgc gtt gga aac cag gac tgg gac	624
Val Val Leu Glu Arg Ile Asp His Cys Val Gly Asn Gln Asp Trp Asp	
195 200 205	
gag atg gag cga gtc tgc gac tac tac gag aag atc etc gga ttc cac	672
Glu Met Glu Arg Val Cys Asp Tyr Tyr Glu Lys Ile Leu Gly Phe His	
210 215 220	
cgt ttc tgg tcc gtt gat gac aag gac atc tgc act gaa ttc tet gca	720
Arg Phe Trp Ser Val Asp Asp Lys Asp Ile Cys Thr Glu Phe Ser Ala	
225 230 235 240	
ctg aag agt atc gtc atg gca tct cca aat gat atc gtc aag atg ccc	768
Leu Lys Ser Ile Val Met Ala Ser Pro Asn Asp Ile Val Lys Met Pro	
245 250 255	
atc aac gag ccc gcc aag gga aag aaa caa tcc cag att gaa gaa tat	816
Ile Asn Glu Pro Ala Lys Gly Lys Lys Gln Ser Gln Ile Glu Glu Tyr	
260 265 270	
gtt gac ttc tac aat ggt get ggc gtt cag cac att get etc cga acc	864
Val Asp Phe Tyr Asn Gly Ala Gly Val Gln His Ile Ala Leu Arg Thr	
275 280 285	
aac aac atc atc gat gcc atc acc aac etc aag gcg cgc gcc acc gaa	912
Asn Asn Ile Ile Asp Ala Ile Thr Asn Leu Lys Ala Arg Gly Thr Glu	
290 295 300	
ttc atc aag gtt cca gag acc tac tat gaa gac atg aag att cgc etc	960
Phe Ile Lys Val Pro Glu Thr Tyr Tyr Glu Asp Met Lys Ile Arg Leu	
305 310 315 320	
aag aga caa ggc ctg gtc etc gat gag gac ttt gag acc ctg aag agc	1008
Lys Arg Gln Gly Leu Val Leu Asp Glu Asp Phe Glu Thr Leu Lys Ser	
325 330 335	
ctg gac atc ctt atc gac ttt gac gag aat ggg tat etc ctg cag ctt	1056
Leu Asp Ile Leu Ile Asp Phe Asp Glu Asn Gly Tyr Leu Leu Gln Leu	
340 345 350	
ttc acc aag cat etc atg gat cgc cca acc gtt ttc att gaa atc atc	1104
Phe Thr Lys His Leu Met Asp Arg Pro Thr Val Phe Ile Glu Ile Ile	
355 360 365	
caa cgc aac aac ttt tcc ggt ttc ggt gcg gcc aac ttc agg gcc etc	1152
Gln Arg Asn Asn Phe Ser Gly Phe Gly Ala Gly Asn Phe Arg Ala Leu	
370 375 380	
ttc gag get att gag cgt gag cag get etc cgt gcc acc ctt atc tag	1200
Phe Glu Ala Ile Glu Arg Glu Gln Ala Leu Arg Gly Thr Leu Ile	
385 390 395	
<210> 8	
<211> 399	
<212> PRT	
<213> 粗球孢子菌 (Coccidioides immitis)	
<400> 8	
Met Ala Pro Ala Ala Asp Ser Pro Thr Leu Gln Pro Ala Gln Pro Ser	
1 5 10 15	

[0012]

Asp Leu Asn Gln Tyr Arg Gly Tyr Asp His Val His Trp Tyr Val Gly
 20 25 30
 Asn Ala Lys Gln Ala Ala Thr Tyr Tyr Val Thr Arg Met Gly Phe Glu
 35 40 45
 Arg Val Ala Tyr Arg Gly Leu Glu Thr Gly Ser Lys Ala Val Ala Ser
 50 55 60
 His Val Val Arg Asn Gly Asn Ile Thr Phe Ile Leu Thr Ser Pro Leu
 65 70 75 80
 Arg Ser Val Glu Gln Ala Ser Arg Phe Pro Glu Asp Glu Ala Leu Leu
 85 90 95
 Lys Glu Ile His Ala His Leu Glu Arg His Gly Asp Gly Val Lys Asp
 100 105 110
 Val Ala Phe Glu Val Asp Cys Val Glu Ser Val Phe Ser Ala Ala Val
 115 120 125
 Arg Asn Gly Ala Glu Val Val Ser Asp Val Arg Thr Val Glu Asp Glu
 130 135 140
 Asp Gly Gln Ile Lys Met Ala Thr Ile Arg Thr Tyr Gly Glu Thr Thr
 145 150 155 160
 His Thr Leu Ile Glu Arg Ser Gly Tyr Arg Gly Gly Phe Met Pro Gly
 165 170 175
 Tyr Arg Met Glu Ser Asn Ala Asp Ala Thr Ser Lys Phe Leu Pro Lys
 180 185 190
 Val Val Leu Glu Arg Ile Asp His Cys Val Gly Asn Gln Asp Trp Asp
 195 200 205
 Glu Met Glu Arg Val Cys Asp Tyr Tyr Glu Lys Ile Leu Gly Phe His
 210 215 220
 Arg Phe Trp Ser Val Asp Asp Lys Asp Ile Cys Thr Glu Phe Ser Ala
 225 230 235 240
 Leu Lys Ser Ile Val Met Ala Ser Pro Asn Asp Ile Val Lys Met Pro
 245 250 255
 Ile Asn Glu Pro Ala Lys Gly Lys Lys Gln Ser Gln Ile Glu Glu Tyr
 260 265 270
 Val Asp Phe Tyr Asn Gly Ala Gly Val Gln His Ile Ala Leu Arg Thr
 275 280 285
 Asn Asn Ile Ile Asp Ala Ile Thr Asn Leu Lys Ala Arg Gly Thr Glu
 290 295 300
 Phe Ile Lys Val Pro Glu Thr Tyr Tyr Glu Asp Met Lys Ile Arg Leu
 305 310 315 320
 Lys Arg Gln Gly Leu Val Leu Asp Glu Asp Phe Glu Thr Leu Lys Ser
 325 330 335
 Leu Asp Ile Leu Ile Asp Phe Asp Glu Asn Gly Tyr Leu Leu Gln Leu
 340 345 350
 Phe Thr Lys His Leu Met Asp Arg Pro Thr Val Phe Ile Glu Ile Ile

[0013]

355	360	365	
Gln Arg Asn Asn Phe Ser Gly Phe Gly Ala Gly Asn Phe Arg Ala Leu			
370	375	380	
Phe Glu Ala Ile Glu Arg Glu Gln Ala Leu Arg Gly Thr Leu Ile			
385	390	395	
<210> 9			
<211> 1260			
<212> DNA			
<213> 小麦叶枯病菌 (<i>Mycosphaerella graminicola</i>)			
<220>			
<221> CDS			
<222> (1)..(1260)			
<400> 9			
atg gca ccc gga gca ctc etc gtc aca tca cag aat gga aga acg agc			48
Met Ala Pro Gly Ala Leu Leu Val Thr Ser Gln Asn Gly Arg Thr Ser			
1	5	10	15
ccc etc tac gac tcc gat ggc tat gta cca gcg cct gcg gct cta gta			96
Pro Leu Tyr Asp Ser Asp Gly Tyr Val Pro Ala Pro Ala Ala Leu Val			
20	25	30	
gta ggt ggt gag gtc aat tac aga ggc tac cat cat gca gaa tgg tgg			144
Val Gly Gly Glu Val Asn Tyr Arg Gly Tyr His His Ala Glu Trp Trp			
35	40	45	
gtg ggc aat gca aag cag gtg gcg caa ttc tac atc aca cgc atg ggc			192
Val Gly Asn Ala Lys Gln Val Ala Gln Phe Tyr Ile Thr Arg Met Gly			
50	55	60	
ttc gag cct gtt gca cac aaa ggt ctg gag acc gga tct cgc ttc ttt			240
Phe Glu Pro Val Ala His Lys Gly Leu Glu Thr Gly Ser Arg Phe Phe			
65	70	75	80
gcc agc cac gtt gtc cag aac aac ggc gtt cgc ttc gtc ttc aca tca			288
Ala Ser His Val Val Gln Asn Asn Gly Val Arg Phe Val Phe Thr Ser			
85	90	95	
cct gtt cgg tca teg gca cgg caa aca etc aaa gca geg cct etc gcg			336
Pro Val Arg Ser Ser Ala Arg Gln Thr Leu Lys Ala Ala Pro Leu Ala			
100	105	110	
gac caa gca cgc etc gac gaa atg tac gat cac ctc gac aag cac gga			384
Asp Gln Ala Arg Leu Asp Glu Met Tyr Asp His Leu Asp Lys His Gly			
115	120	125	
gat gga gtg aag gat gtt gcc ttc gaa gtt gac gat gtc ttg gct gtg			432
Asp Gly Val Lys Asp Val Ala Phe Glu Val Asp Asp Val Leu Ala Val			
130	135	140	
tac gag aac gca gtt gcg aat ggt geg gag tcc gtc agt tca cca cat			480
Tyr Glu Asn Ala Val Ala Asn Gly Ala Glu Ser Val Ser Ser Pro His			
145	150	155	160
acc gat tca tgc gac gaa ggc gat gtg atc tcc gcg gcg atc aag aca			528
Thr Asp Ser Cys Asp Glu Gly Asp Val Ile Ser Ala Ala Ile Lys Thr			
165	170	175	
tac gga gac acc acg cac act ttc atc caa cgc aca aca tat aca gga			576
Tyr Gly Asp Thr Thr His Thr Phe Ile Gln Arg Thr Thr Tyr Thr Gly			
180	185	190	

[0014]

cca ttt ctt cct ggc tat cga tca tgt acc aca gtg gat tgg gcc aac Pro Phe Leu Pro Gly Tyr Arg Ser Cys Thr Thr Val Asp Ser Ala Asn 195 200 205	624
aag ttc ttg cca cca gtc aat ctc gaa gcg atc gat cac tgt gtc ggc Lys Phe Leu Pro Pro Val Asn Leu Glu Ala Ile Asp His Cys Val Gly 210 215 220	672
aat caa gac tgg gac gag atg agc gat gcc tgc gac ttc tac gag cgc Asn Gln Asp Trp Asp Glu Met Ser Asp Ala Cys Asp Phe Tyr Glu Arg 225 230 235 240	720
tgt ctt gga ttc cat cgc ttc tgg agt gtc gat gac aag gac atc tgt Cys Leu Gly Phe His Arg Phe Trp Ser Val Asp Asp Lys Asp Ile Cys 245 250 255	768
acg gag ttc tcc gcg ctg aag tct atc gtt atg agt tct ccc aac cag Thr Glu Phe Ser Ala Leu Lys Ser Ile Val Met Ser Ser Pro Asn Gln 260 265 270	816
gta gtc aag atg cca atc aac gag ccc gcc cat ggc aag aag aag agc Val Val Lys Met Pro Ile Asn Glu Pro Ala His Gly Lys Lys Lys Ser 275 280 285	864
cag atc gag gag tac gtc gat ttc tac aat gga cct ggc gta caa cac Gln Ile Glu Glu Tyr Val Asp Phe Tyr Asn Gly Pro Gly Val Gln His 290 295 300	912
atc gct ctc cgt acg cca aac atc atc gag gca gta tca aac ttg cgg Ile Ala Leu Arg Thr Pro Asn Ile Ile Glu Ala Val Ser Asn Leu Arg 305 310 315 320	960
tca aga ggc gtg gag ttc atc agc gtg cca gat acg tac tac gag aac Ser Arg Gly Val Glu Phe Ile Ser Val Pro Asp Thr Tyr Tyr Glu Asn 325 330 335	1008
atg cgt ctt cgt ctc aaa gcg gca gga atg aag ctg gag gag tca ttc Met Arg Leu Arg Leu Lys Ala Ala Gly Met Lys Leu Glu Glu Ser Phe 340 345 350	1056
gac atc att caa aag ctg aac atc ctc atc gat ttc gac gaa ggt ggc Asp Ile Ile Gln Lys Leu Asn Ile Leu Ile Asp Phe Asp Glu Gly Gly 355 360 365	1104
tat ttg ctg cag ctg ttc acg aag ccg ctg atg gat cgg ccg acg gtc Tyr Leu Leu Gln Leu Phe Thr Lys Pro Leu Met Asp Arg Pro Thr Val 370 375 380	1152
ttc att gaa atc att caa cgg aac aac ttt gat ggc ttc gga get gga Phe Ile Glu Ile Ile Gln Arg Asn Asn Phe Asp Gly Phe Gly Ala Gly 385 390 395 400	1200
aac ttc aag agt ctg ttc gag gcg att gag cga gag cag gac ttg cgt Asn Phe Lys Ser Leu Phe Glu Ala Ile Glu Arg Glu Gln Asp Leu Arg 405 410 415	1248
ggc aat ctc tag Gly Asn Leu	1260

<210> 10

<211> 419

<212> PRT

<213> 小麦叶枯病菌 (Mycosphaerella graminicola)

[0015]

<400> 10

Met Ala Pro Gly Ala Leu Leu Val Thr Ser Gln Asn Gly Arg Thr Ser
1 5 10 15

Pro Leu Tyr Asp Ser Asp Gly Tyr Val Pro Ala Pro Ala Ala Leu Val
20 25 30

Val Gly Gly Glu Val Asn Tyr Arg Gly Tyr His His Ala Glu Trp Trp
35 40 45

Val Gly Asn Ala Lys Gln Val Ala Gln Phe Tyr Ile Thr Arg Met Gly
50 55 60

Phe Glu Pro Val Ala His Lys Gly Leu Glu Thr Gly Ser Arg Phe Phe
65 70 75 80

Ala Ser His Val Val Gln Asn Asn Gly Val Arg Phe Val Phe Thr Ser
85 90 95

Pro Val Arg Ser Ser Ala Arg Gln Thr Leu Lys Ala Ala Pro Leu Ala
100 105 110

Asp Gln Ala Arg Leu Asp Glu Met Tyr Asp His Leu Asp Lys His Gly
115 120 125

Asp Gly Val Lys Asp Val Ala Phe Glu Val Asp Asp Val Leu Ala Val
130 135 140

Tyr Glu Asn Ala Val Ala Asn Gly Ala Glu Ser Val Ser Ser Pro His
145 150 155 160

Thr Asp Ser Cys Asp Glu Gly Asp Val Ile Ser Ala Ala Ile Lys Thr
165 170 175

Tyr Gly Asp Thr Thr His Thr Phe Ile Gln Arg Thr Thr Tyr Thr Gly
180 185 190

Pro Phe Leu Pro Gly Tyr Arg Ser Cys Thr Thr Val Asp Ser Ala Asn
195 200 205

Lys Phe Leu Pro Pro Val Asn Leu Glu Ala Ile Asp His Cys Val Gly
210 215 220

Asn Gln Asp Trp Asp Glu Met Ser Asp Ala Cys Asp Phe Tyr Glu Arg
225 230 235 240

Cys Leu Gly Phe His Arg Phe Trp Ser Val Asp Asp Lys Asp Ile Cys
245 250 255

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

Val Val Lys Met Pro Ile Asn Glu Pro Ala His Gly Lys Lys Lys Ser
275 280 285

Gln Ile Glu Glu Tyr Val Asp Phe Tyr Asn Gly Pro Gly Val Gln His
290 295 300

Ile Ala Leu Arg Thr Pro Asn Ile Ile Glu Ala Val Ser Asn Leu Arg
305 310 315 320

Ser Arg Gly Val Glu Phe Ile Ser Val Pro Asp Thr Tyr Tyr Glu Asn
325 330 335

[0016]

Met Arg Leu Arg Leu Lys Ala Ala Gly Met Lys Leu Glu Glu Ser Phe
 340 345 350
 Asp Ile Ile Gln Lys Leu Asn Ile Leu Ile Asp Phe Asp Glu Gly Gly
 355 360 365
 Tyr Leu Leu Gln Leu Phe Thr Lys Pro Leu Met Asp Arg Pro Thr Val
 370 375 380
 Phe Ile Glu Ile Ile Gln Arg Asn Asn Phe Asp Gly Phe Gly Ala Gly
 385 390 395 400
 Asn Phe Lys Ser Leu Phe Glu Ala Ile Glu Arg Glu Gln Asp Leu Arg
 405 410 415
 Gly Asn Leu

<210> 11
 <211> 1305
 <212> DNA
 <213> 大麦 (Hordeum vulgare)

<220>
 <221> CDS
 <222> (1)..(1305)

<400> 11
 atg ccg ccc acc ccc acc acc ccc gcg gct acc ggc gcc gcc gcc gcg 48
 Met Pro Pro Thr Pro Thr Thr Pro Ala Ala Thr Gly Ala Ala Ala Ala
 1 5 10 15
 gtg acg ccg gag cac gcg cga ccg cac cga atg gtc cgc ttc aac ccg 96
 Val Thr Pro Glu His Ala Arg Pro His Arg Met Val Arg Phe Asn Pro
 20 25 30
 cgc agc gac cgc ttc cac acg ctc tcc ttc cac cac gtc gag ttc tgg 144
 Arg Ser Asp Arg Phe His Thr Leu Ser Phe His His Val Glu Phe Trp
 35 40 45
 tgc gcg gac gcc gcc tcc gcc gcc ggc cgc ttc gcg ttc gcg ctc gcc 192
 Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg Phe Ala Phe Ala Leu Gly
 50 55 60
 gcg ccg ctc gcc gcc agg tcc gac ctc tcc acg ggg aac tcc gcg cac 240
 Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser Thr Gly Asn Ser Ala His
 65 70 75 80
 gcc tcc cag ctg ctc cgc tcg gcc tcc ctc gcc ttc ctc ttc acc gcg 288
 Ala Ser Gln Leu Leu Arg Ser Gly Ser Leu Ala Phe Leu Phe Thr Ala
 85 90 95
 ccc tac gcc aac ggc tgc gac gcc gcc acc gcc tcc ctg ccc tcc ttc 336
 Pro Tyr Ala Asn Gly Cys Asp Ala Ala Thr Ala Ser Leu Pro Ser Phe
 100 105 110
 tcc gcc gac gcc gcg cgc cgg ttc tcc gcc gac cac ggg atc gcg gtg 384
 Ser Ala Asp Ala Ala Arg Arg Phe Ser Ala Asp His Gly Ile Ala Val
 115 120 125
 cgc tcc gta gcg ctg cgc gtc gca gac gcc gcc gag gcc ttc cgc gcc 432
 Arg Ser Val Ala Leu Arg Val Ala Asp Ala Ala Glu Ala Phe Arg Ala
 130 135 140

[0017]

agt cgt cga cgg ggc gcg cgc ccg gcc ttc gcc ecc gtg gac ctc ggc Ser Arg Arg Arg Gly Ala Arg Pro Ala Phe Ala Pro Val Asp Leu Gly 145 150 155 160	480
cgc gcc ttc cgg ttc gcg gag gtc gag ctc tac gcc gac gtc gtg etc Arg Gly Phe Ala Phe Ala Glu Val Glu Leu Tyr Gly Asp Val Val Leu 165 170 175	528
cgc ttc gtc agc cac ccg gac ggc acg gac gtg ecc ttc ttg ccg ggc Arg Phe Val Ser His Pro Asp Gly Thr Asp Val Pro Phe Leu Pro Gly 180 185 190	576
ttc gag gcc gta acc aac ccg gac gcc gtg gac tac gcc ctg acg ccg Phe Glu Gly Val Thr Asn Pro Asp Ala Val Asp Tyr Gly Leu Thr Arg 195 200 205	624
ttc gac cac gtc gtc ggc aac gtc ccg gag ctt gcc ccc gcc gca gcc Phe Asp His Val Val Gly Asn Val Pro Glu Leu Ala Pro Ala Ala Ala 210 215 220	672
tac atc gcc ggg ttc acg ggg ttc cac gag ttc gcc gag ttc acg gcg Tyr Ile Ala Gly Phe Thr Gly Phe His Glu Phe Ala Glu Phe Thr Ala 225 230 235 240	720
gag gac gtg gcc acg acc gag agc ggg ctc aac tcg gtg gtg ctc gcc Glu Asp Val Gly Thr Thr Glu Ser Gly Leu Asn Ser Val Val Leu Ala 245 250 255	768
aac aac tcg gag gcc gtg ctg ctg ccg ctc aac gag ccg gtg cac gcc Asn Asn Ser Glu Gly Val Leu Leu Pro Leu Asn Glu Pro Val His Gly 260 265 270	816
acc aag cgc cgg agc cag ata cag acg ttc ctg gaa cac cac gcc gcc Thr Lys Arg Arg Ser Gln Ile Gln Thr Phe Leu Glu His His Gly Gly 275 280 285	864
ccg gcc gtg cag cac atc gcg gtg gcc age agt gac gtg ctc agg acg Pro Gly Val Gln His Ile Ala Val Ala Ser Ser Asp Val Leu Arg Thr 290 295 300	912
etc agg aag atg cgt gcg cgc tcc gcc atg gcc gcc ttc gac ttc ctg Leu Arg Lys Met Arg Ala Arg Ser Ala Met Gly Gly Phe Asp Phe Leu 305 310 315 320	960
cca ccc ccg ctg ccg aag tac tac gaa gcc gtg cga cgc ctt gcc ggc Pro Pro Pro Leu Pro Lys Tyr Tyr Glu Gly Val Arg Arg Leu Ala Gly 325 330 335	1008
gat gtc ctc tcg gag gcg cag atc aag gaa tgc cag gag ctg ggt gtg Asp Val Leu Ser Gln Ala Gln Ile Lys Glu Cys Gln Glu Leu Gly Val 340 345 350	1056
etc gtc gat agg gac gac caa ggg gtg ttg etc caa atc ttc acc aag Leu Val Asp Arg Asp Asp Gln Gly Val Leu Leu Gln Ile Phe Thr Lys 355 360 365	1104
cca gta ggg gac agg ccg acc ttg ttc ctg gag atg atc cag agg atc Pro Val Gly Asp Arg Pro Thr Leu Phe Leu Glu Met Ile Gln Arg Ile 370 375 380	1152
ggg tgc atg gag aag gac gag aga ggg gaa gag tac cag aag ggt gcc Gly Cys Met Glu Lys Asp Glu Arg Gly Glu Tyr Gln Lys Gly Gly 385 390 395 400	1200
tgc gcc ggg ttc gcc aaa gcc aac ttc tcc gag ctg ttc aag tcc att	1248

[0018]

Cys Gly Gly Phe Gly Lys Gly Asn Phe Ser Glu Leu Phe Lys Ser Ile
 405 410 415

gaa gat tac gag aag tcc ctt gaa gcc aag caa tct gct gca gtt cag 1296
 Glu Asp Tyr Glu Lys Ser Leu Glu Ala Lys Gln Ser Ala Ala Val Gln
 420 425 430

gga tca tag 1305
 Gly Ser

<210> 12
 <211> 434
 <212> PRT
 <213> 大麦 (Hordeum vulgare)

<400> 12

Met Pro Pro Thr Pro Thr Thr Pro Ala Ala Thr Gly Ala Ala Ala Ala
 1 5 10 15

Val Thr Pro Glu His Ala Arg Pro His Arg Met Val Arg Phe Asn Pro
 20 25 30

Arg Ser Asp Arg Phe His Thr Leu Ser Phe His His Val Glu Phe Trp
 35 40 45

Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg Phe Ala Phe Ala Leu Gly
 50 55 60

Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser Thr Gly Asn Ser Ala His
 65 70 75 80

Ala Ser Gln Leu Leu Arg Ser Gly Ser Leu Ala Phe Leu Phe Thr Ala
 85 90 95

Pro Tyr Ala Asn Gly Cys Asp Ala Ala Thr Ala Ser Leu Pro Ser Phe
 100 105 110

Ser Ala Asp Ala Ala Arg Arg Phe Ser Ala Asp His Gly Ile Ala Val
 115 120 125

Arg Ser Val Ala Leu Arg Val Ala Asp Ala Ala Glu Ala Phe Arg Ala
 130 135 140

Ser Arg Arg Arg Gly Ala Arg Pro Ala Phe Ala Pro Val Asp Leu Gly
 145 150 155 160

Arg Gly Phe Ala Phe Ala Glu Val Glu Leu Tyr Gly Asp Val Val Leu
 165 170 175

Arg Phe Val Ser His Pro Asp Gly Thr Asp Val Pro Phe Leu Pro Gly
 180 185 190

Phe Glu Gly Val Thr Asn Pro Asp Ala Val Asp Tyr Gly Leu Thr Arg
 195 200 205

Phe Asp His Val Val Gly Asn Val Pro Glu Leu Ala Pro Ala Ala Ala
 210 215 220

Tyr Ile Ala Gly Phe Thr Gly Phe His Glu Phe Ala Glu Phe Thr Ala
 225 230 235 240

Glu Asp Val Gly Thr Thr Glu Ser Gly Leu Asn Ser Val Val Leu Ala
 245 250 255

[0019]

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

Gly Ser

<210> 13
 <211> 1332
 <212> DNA
 <213> 玉米 (Zea mays)

<220>
 <221> CDS
 <222> (1).. (1332)

<400> 13
 atg ccc ceg acc ccc aca gcc gcc gca gcc ggc gcc gcc gtc gcg gcg 48
 Met Pro Pro Thr Pro Thr Ala Ala Ala Ala Gly Ala Ala Val Ala Ala
 1 5 10 15
 gca tca gca gcg gag caa gcg gcg ttc cgc ctc gtg ggc cac cgc aac 96
 Ala Ser Ala Ala Glu Gln Ala Ala Phe Arg Leu Val Gly His Arg Asn
 20 25 30
 ttc gtc cgc ttc aac ceg cgc tcc gac cgc ttc cac acg ctc gcg ttc 144
 Phe Val Arg Phe Asn Pro Arg Ser Asp Arg Phe His Thr Leu Ala Phe
 35 40 45
 cac cac gtg gag ctc tgg tgc gcc gac geg gcc tcc gcc geg ggc cgc 192
 His His Val Glu Leu Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg
 50 55 60
 ttc tcc ttc ggc ctg gcc gcg ceg ctc gcc gca cgc tcc gac ctc tcc 240
 Phe Ser Phe Gly Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser

[0020]

65	70	75	80	
acg ggc aac tcc gcg cac geg tcc ctg ctg etc cgc tcc ggc tcc etc				288
Thr Gly Asn Ser Ala His Ala Ser Leu Leu Leu Arg Ser Gly Ser Leu	85	90	95	
tcc ttc etc ttc aeg gcg ccc tac gcg cac ggc gcc gac gct gcc acc				336
Ser Phe Leu Phe Thr Ala Pro Tyr Ala His Gly Ala Asp Ala Ala Thr	100	105	110	
gcc gcg ctg ccc tcc ttc tcc gcc gcc gcc gcg cgg cgc ttc gca gcc				384
Ala Ala Leu Pro Ser Phe Ser Ala Ala Ala Arg Arg Phe Ala Ala	115	120	125	
gac cac ggc etc geg gtg cgc gcc gtc gcg etc cgc gtc gcc gac gcc				432
Asp His Gly Leu Ala Val Arg Ala Val Ala Leu Arg Val Ala Asp Ala	130	135	140	
gag gac gcc ttc cgc gcc agc gtc gcg gcc ggg gcg cgc cgg gcg ttc				480
Glu Asp Ala Phe Arg Ala Ser Val Ala Ala Gly Ala Arg Pro Ala Phe	145	150	155	160
ggc ecc gtc gac etc ggc cgc gcc ttc cgc etc gcc gag gtc gag etc				528
Gly Pro Val Asp Leu Gly Arg Gly Phe Arg Leu Ala Glu Val Glu Leu	165	170	175	
tac ggc gac gtc gtg etc cgg tac gtg agc tac cgg gac ggc gcc gcg				576
Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr Pro Asp Gly Ala Ala	180	185	190	
ggc gag ccc ttc ctg cgg ggg ttc gag gcc gtg gcc agc ecc ggg gcg				624
Gly Glu Pro Phe Leu Pro Gly Phe Glu Gly Val Ala Ser Pro Gly Ala	195	200	205	
gcc gac tac ggg ctg agc agg ttc gac cac atc gtc gcc aac gtg cgg				672
Ala Asp Tyr Gly Leu Ser Arg Phe Asp His Ile Val Gly Asn Val Pro	210	215	220	
gag ctg gcg ccc gcc gcc gcc tac ttc gcc ggc ttc acg ggg ttc cac				720
Glu Leu Ala Pro Ala Ala Ala Tyr Phe Ala Gly Phe Thr Gly Phe His	225	230	235	240
gag ttc gcc gag ttc acg acg gag gac gtg gcc acc geg gag agc gcc				768
Glu Phe Ala Glu Phe Thr Thr Glu Asp Val Gly Thr Ala Glu Ser Gly	245	250	255	
etc aac tcc atg gtg etc gcc aac aac tcg gag aac gtg ctg etc cgg				816
Leu Asn Ser Met Val Leu Ala Asn Asn Ser Glu Asn Val Leu Leu Pro	260	265	270	
etc aac gag ccg gtg cac gcc acc aag cgc cgc agc cag ata caa acg				864
Leu Asn Glu Pro Val His Gly Thr Lys Arg Arg Ser Gln Ile Gln Thr	275	280	285	
ttc ctg gac cac cac ggc gcc ccc gcc gtg cag cac atg gcg ctg gcc				912
Phe Leu Asp His His Gly Gly Pro Gly Val Gln His Met Ala Leu Ala	290	295	300	
agc gac gac gtg etc agg acg ctg agg gag atg cag gcg cgc tcg gcc				960
Ser Asp Asp Val Leu Arg Thr Leu Arg Glu Met Gln Ala Arg Ser Ala	305	310	315	320
atg gcc gcc ttc gag ttc atg gcg cct ccc aca tcc gac tac tat gac				1008
Met Gly Gly Phe Glu Phe Met Ala Pro Pro Thr Ser Asp Tyr Tyr Asp	325	330	335	

[0021]

ggc gtg agg egg cgc gcc ggg gac gtg ctc acg gaa gca cag att aag Gly Val Arg Arg Arg Ala Gly Asp Val Leu Thr Glu Ala Gln Ile Lys 340 345 350	1056
gag tgc cag gag cta ggg gtg ctg gtg gac agg gat gac cag ggc gtg Glu Cys Gln Glu Leu Gly Val Leu Val Asp Arg Asp Asp Gln Gly Val 355 360 365	1104
ctg ctc caa atc ttc acc aag cca gtg ggg gac agg cca acg ctg ttc Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu Phe 370 375 380	1152
ttg gaa atc atc caa agg atc ggg tgc atg gag aag gat gag aag ggg Leu Glu Ile Ile Gln Arg Ile Gly Cys Met Glu Lys Asp Glu Lys Gly 385 390 395 400	1200
caa gaa tac caa aag ggt ggc tgc ggc ggg ttc ggc aag gga aac ttc Gln Glu Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe 405 410 415	1248
teg cag ctg ttc aag tcc atc gag gat tat gag aag tcc ctt gaa gcc Ser Gln Leu Phe Lys Ser Ile Glu Asp Tyr Glu Lys Ser Leu Glu Ala 420 425 430	1296
aag caa gct gct gca gca gct gca gct cag gga tcc Lys Gln Ala Ala Ala Ala Ala Ala Ala Gln Gly Ser 435 440	1332
<210> 14 <211> 444 <212> PRT <213> 玉米 (Zea mays)	
<400> 14	
Met Pro Pro Thr Pro Thr Ala Ala Ala Ala Gly Ala Ala Val Ala Ala 1 5 10 15	
Ala Ser Ala Ala Glu Gln Ala Ala Phe Arg Leu Val Gly His Arg Asn 20 25 30	
Phe Val Arg Phe Asn Pro Arg Ser Asp Arg Phe His Thr Leu Ala Phe 35 40 45	
His His Val Glu Leu Trp Cys Ala Asp Ala Ala Ser Ala Ala Gly Arg 50 55 60	
Phe Ser Phe Gly Leu Gly Ala Pro Leu Ala Ala Arg Ser Asp Leu Ser 65 70 75 80	
Thr Gly Asn Ser Ala His Ala Ser Leu Leu Leu Arg Ser Gly Ser Leu 85 90 95	
Ser Phe Leu Phe Thr Ala Pro Tyr Ala His Gly Ala Asp Ala Ala Thr 100 105 110	
Ala Ala Leu Pro Ser Phe Ser Ala Ala Ala Ala Arg Arg Phe Ala Ala 115 120 125	
Asp His Gly Leu Ala Val Arg Ala Val Ala Leu Arg Val Ala Asp Ala 130 135 140	
Glu Asp Ala Phe Arg Ala Ser Val Ala Ala Gly Ala Arg Pro Ala Phe 145 150 155 160	

[0022]

Gly Pro Val Asp Leu Gly Arg Gly Phe Arg Leu Ala Glu Val Glu Leu
 165 170 175
 Tyr Gly Asp Val Val Leu Arg Tyr Val Ser Tyr Pro Asp Gly Ala Ala
 180 185 190
 Gly Glu Pro Phe Leu Pro Gly Phe Glu Gly Val Ala Ser Pro Gly Ala
 195 200 205
 Ala Asp Tyr Gly Leu Ser Arg Phe Asp His Ile Val Gly Asn Val Pro
 210 215 220
 Glu Leu Ala Pro Ala Ala Ala Tyr Phe Ala Gly Phe Thr Gly Phe His
 225 230 235 240
 Glu Phe Ala Glu Phe Thr Thr Glu Asp Val Gly Thr Ala Glu Ser Gly
 245 250 255
 Leu Asn Ser Met Val Leu Ala Asn Asn Ser Glu Asn Val Leu Leu Pro
 260 265 270
 Leu Asn Glu Pro Val His Gly Thr Lys Arg Arg Ser Gln Ile Gln Thr
 275 280 285
 Phe Leu Asp His His Gly Gly Pro Gly Val Gln His Met Ala Leu Ala
 290 295 300
 Ser Asp Asp Val Leu Arg Thr Leu Arg Glu Met Gln Ala Arg Ser Ala
 305 310 315 320
 Met Gly Gly Phe Glu Phe Met Ala Pro Pro Thr Ser Asp Tyr Tyr Asp
 325 330 335
 Gly Val Arg Arg Arg Ala Gly Asp Val Leu Thr Glu Ala Gln Ile Lys
 340 345 350
 Glu Cys Gln Glu Leu Gly Val Leu Val Asp Arg Asp Asp Gln Gly Val
 355 360 365
 Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu Phe
 370 375 380
 Leu Glu Ile Ile Gln Arg Ile Gly Cys Met Glu Lys Asp Glu Lys Gly
 385 390 395 400
 Gln Glu Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn Phe
 405 410 415
 Ser Gln Leu Phe Lys Ser Ile Glu Asp Tyr Glu Lys Ser Leu Glu Ala
 420 425 430
 Lys Gln Ala Ala Ala Ala Ala Ala Ala Gln Gly Ser
 435 440

<210> 15
 <211> 1329
 <212> DNA
 <213> 胡萝卜 (Daucus carota)

<220>
 <221> CDS
 <222> (1).. (1329)

<400> 15

[0023]

atg ggg aaa aaa caa tcg gaa gct gaa att ctc tca agc aat tca tca Met Gly Lys Lys Gln Ser Glu Ala Glu Ile Leu Ser Ser Asn Ser Ser 1 5 10 15	48
aac acc tct cct gca aca ttc aag ctg gtc ggt ttc aac aac ttc gtc Asn Thr Ser Pro Ala Thr Phe Lys Leu Val Gly Phe Asn Asn Phe Val 20 25 30	96
cgc gcc aac ecc aag tcc gat cac ttc gcc gtg aag cgg ttc cac cac Arg Ala Asn Pro Lys Ser Asp His Phe Ala Val Lys Arg Phe His His 35 40 45	144
att gag ttc tgg tgc gcc gac gcc acc aac acg tcg cgg cgg ttc tcg Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Thr Ser Arg Arg Phe Ser 50 55 60	192
tgg gcc ctc gcc atg cct ttg gtg gcg aaa tcg gat ctc tct act gga Trp Gly Leu Gly Met Pro Leu Val Ala Lys Ser Asp Leu Ser Thr Gly 65 70 75 80	240
aac tct gtt cac gct tct tat ctt gtt cgc tcg gcg aat ctc agt ttc Asn Ser Val His Ala Ser Tyr Leu Val Arg Ser Ala Asn Leu Ser Phe 85 90 95	288
gtc ttc acc get cct tac tct ccg tcc acg acc act tcc tct ggt tca Val Phe Thr Ala Pro Tyr Ser Pro Ser Thr Thr Thr Ser Ser Gly Ser 100 105 110	336
gct gcc atc ccg tct ttt tcg gca tcg ggt ttt cac tct ttt gcg gcc Ala Ala Ile Pro Ser Phe Ser Ala Ser Gly Phe His Ser Phe Ala Ala 115 120 125	384
aaa cac gcc ctt get gtt egg get att get ctt gaa gtt get gac gtg Lys His Gly Leu Ala Val Arg Ala Ile Ala Leu Glu Val Ala Asp Val 130 135 140	432
get get gcg itt gag gcc agt gtt gcg cgt ggg gcc agg ccg get tcg Ala Ala Ala Phe Glu Ala Ser Val Ala Arg Gly Ala Arg Pro Ala Ser 145 150 155 160	480
gct cct gtt gaa ttg gac gac cag gcg tgg ttg get gag gtg gag ttg Ala Pro Val Glu Leu Asp Asp Gln Ala Trp Leu Ala Glu Val Glu Leu 165 170 175	528
tac gga gat gtg gtc ttg agg ttt gtt agt ttt ggg agg gag gag ggt Tyr Gly Asp Val Val Leu Arg Phe Val Ser Phe Gly Arg Glu Glu Gly 180 185 190	576
ttg ttt ttg cct gga ttc gag gcg gtg gag ggg acg gcg tcg ttt ccg Leu Phe Leu Pro Gly Phe Glu Ala Val Glu Gly Thr Ala Ser Phe Pro 195 200 205	624
gat ttg gat tat gga att aga aga ctt gat cat gcg gtg ggg aat gtt Asp Leu Asp Tyr Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val 210 215 220	672
acc gag ttg ggg cct gtg gtg gag tat att aaa ggg ttt acg ggg ttt Thr Glu Leu Gly Pro Val Val Glu Tyr Ile Lys Gly Phe Thr Gly Phe 225 230 235 240	720
cat gaa ttt gcg gag ttt aca gcg gag gat gtg ggg act ttg gag agt His Glu Phe Ala Glu Phe Thr Ala Glu Asp Val Gly Thr Leu Glu Ser 245 250 255	768
ggg ttg aat tcg gtg gtg ttg gcg aat aat gag gag atg gtt ctg ttg	816

[0024]

Gly Leu Asn Ser Val Val Leu Ala Asn Asn Glu Glu Met Val Leu Leu 260 265 270	
ccc ttg aat gag cct gtg tat ggg acc aag agg aag agt cag ata cag Pro Leu Asn Glu Pro Val Tyr Gly Thr Lys Arg Lys Ser Gln Ile Gln 275 280 285	864
act tac ttg gag cac aat gaa ggg gct gga gtg cag cat ttg gct tta Thr Tyr Leu Glu His Asn Glu Gly Ala Gly Val Gln His Leu Ala Leu 290 295 300	912
gtg agt gag gat att ttt agg act tta agg gag atg agg aag agg agt Val Ser Glu Asp Ile Phe Arg Thr Leu Arg Glu Met Arg Lys Arg Ser 305 310 315 320	960
tgc ctt ggt ggt ttt gag ttt atg cct tgc cca ccg cct acg tat tac Cys Leu Gly Gly Phe Glu Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr 325 330 335	1008
aag aat ttg aag aat agg gtc ggg gat gtg ttg agt gat gaa cag atc Lys Asn Leu Lys Asn Arg Val Gly Asp Val Leu Ser Asp Glu Gln Ile 340 345 350	1056
aag gag tgt gaa gat ttg ggg att ttg gtg gat agg gat gat cag ggt Lys Glu Cys Glu Asp Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly 355 360 365	1104
aca ttg ctt caa atc ttt acc aag cct gta ggt gac agg cct acc tta Thr Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu 370 375 380	1152
ttc ata gag atc att cag agg gta ggg tgc atg ctc aag gac gat gca Phe Ile Glu Ile Ile Gln Arg Val Gly Cys Met Leu Lys Asp Asp Ala 385 390 395 400	1200
ggg cag atg tac cag aag ggc ggg tgc gga gga ttt ggg aag ggg aac Gly Gln Met Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn 405 410 415	1248
ttc tca gag ctg ttc aag tcc atc gaa gaa tat gaa aaa aca ctt gaa Phe Ser Glu Leu Phe Lys Ser Ile Glu Glu Tyr Glu Lys Thr Leu Glu 420 425 430	1296
gct aaa caa atc act gga tct gct gct gca tga Ala Lys Gln Ile Thr Gly Ser Ala Ala Ala 435 440	1329
<210> 16 <211> 442 <212> PRT <213> 胡萝卜 (Daucus carota)	
<400> 16	
Met Gly Lys Lys Gln Ser Glu Ala Glu Ile Leu Ser Ser Asn Ser Ser 1 5 10 15	
Asn Thr Ser Pro Ala Thr Phe Lys Leu Val Gly Phe Asn Asn Phe Val 20 25 30	
Arg Ala Asn Pro Lys Ser Asp His Phe Ala Val Lys Arg Phe His His 35 40 45	
Ile Glu Phe Trp Cys Gly Asp Ala Thr Asn Thr Ser Arg Arg Phe Ser 50 55 60	

[0025]

Trp Gly Leu Gly Met Pro Leu Val Ala Lys Ser Asp Leu Ser Thr Gly
 65 70 75 80
 Asn Ser Val His Ala Ser Tyr Leu Val Arg Ser Ala Asn Leu Ser Phe
 85 90 95
 Val Phe Thr Ala Pro Tyr Ser Pro Ser Thr Thr Thr Ser Ser Gly Ser
 100 105 110
 Ala Ala Ile Pro Ser Phe Ser Ala Ser Gly Phe His Ser Phe Ala Ala
 115 120 125
 Lys His Gly Leu Ala Val Arg Ala Ile Ala Leu Glu Val Ala Asp Val
 130 135 140
 Ala Ala Ala Phe Glu Ala Ser Val Ala Arg Gly Ala Arg Pro Ala Ser
 145 150 155 160
 Ala Pro Val Glu Leu Asp Asp Gln Ala Trp Leu Ala Glu Val Glu Leu
 165 170 175
 Tyr Gly Asp Val Val Leu Arg Phe Val Ser Phe Gly Arg Glu Glu Gly
 180 185 190
 Leu Phe Leu Pro Gly Phe Glu Ala Val Glu Gly Thr Ala Ser Phe Pro
 195 200 205
 Asp Leu Asp Tyr Gly Ile Arg Arg Leu Asp His Ala Val Gly Asn Val
 210 215 220
 Thr Glu Leu Gly Pro Val Val Glu Tyr Ile Lys Gly Phe Thr Gly Phe
 225 230 235 240
 His Glu Phe Ala Glu Phe Thr Ala Glu Asp Val Gly Thr Leu Glu Ser
 245 250 255
 Gly Leu Asn Ser Val Val Leu Ala Asn Asn Glu Glu Met Val Leu Leu
 260 265 270
 Pro Leu Asn Glu Pro Val Tyr Gly Thr Lys Arg Lys Ser Gln Ile Gln
 275 280 285
 Thr Tyr Leu Glu His Asn Glu Gly Ala Gly Val Gln His Leu Ala Leu
 290 295 300
 Val Ser Glu Asp Ile Phe Arg Thr Leu Arg Glu Met Arg Lys Arg Ser
 305 310 315 320
 Cys Leu Gly Gly Phe Glu Phe Met Pro Ser Pro Pro Pro Thr Tyr Tyr
 325 330 335
 Lys Asn Leu Lys Asn Arg Val Gly Asp Val Leu Ser Asp Glu Gln Ile
 340 345 350
 Lys Glu Cys Glu Asp Leu Gly Ile Leu Val Asp Arg Asp Asp Gln Gly
 355 360 365
 Thr Leu Leu Gln Ile Phe Thr Lys Pro Val Gly Asp Arg Pro Thr Leu
 370 375 380
 Phe Ile Glu Ile Ile Gln Arg Val Gly Cys Met Leu Lys Asp Asp Ala
 385 390 395 400
 Gly Gln Met Tyr Gln Lys Gly Gly Cys Gly Gly Phe Gly Lys Gly Asn

[0026]

	405	410	415	
Phe Ser Glu	Leu Phe Lys Ser Ile	Glu Glu Tyr Glu Lys	Thr Leu Glu	
	420	425	430	
Ala Lys Gln	Ile Thr Gly Ser	Ala Ala Ala		
	435	440		
<210>	17			
<211>	1146			
<212>	DNA			
<213>	阿维链霉菌 (Streptomyces avermitilis)			
<220>				
<221>	CDS			
<222>	(1)..(1146)			
<400>	17			
atg acg cag acc aca cac cac act ecc gac acc gcc cgg cag gcc gac				48
Met Thr Gln Thr Thr His His Thr Pro Asp Thr Ala Arg Gln Ala Asp	5	10	15	
1				
ccc ttc cgg gtg aag gga atg gac gcg gtc gtc ttc gcc gta ggc aac				96
Pro Phe Pro Val Lys Gly Met Asp Ala Val Val Phe Ala Val Gly Asn	20	25	30	
20				
gcc aag cag gcc gcg cac tac tac tcc acc gcc ttc ggc atg cag ott				144
Ala Lys Gln Ala Ala His Tyr Tyr Ser Thr Ala Phe Gly Met Gln Leu	35	40	45	
35				
gtg geg tac tcc gga cgg gag aac ggc agc cgc gag acc gct tgg tac				192
Val Ala Tyr Ser Gly Pro Glu Asn Gly Ser Arg Glu Thr Ala Ser Tyr	50	55	60	
50				
gtc etc acc aac ggc teg gca cgc ttc gtc etc acc tcc gtc atc aag				240
Val Leu Thr Asn Gly Ser Ala Arg Phe Val Leu Thr Ser Val Ile Lys	65	70	75	80
65				
ccc gcc acc ccc tgg ggc cac ttc etc gcc gac cat gtg gcc gag cac				288
Pro Ala Thr Pro Trp Gly His Phe Leu Ala Asp His Val Ala Glu His	85	90	95	
85				
ggc gac ggc gtc gtc gac etc gcc atc gag gtc cgg gac gcc cgc gcc				336
Gly Asp Gly Val Val Asp Leu Ala Ile Glu Val Pro Asp Ala Arg Ala	100	105	110	
100				
gcc cac gcg tac gcg atc gag cac ggc gcc cgc tgg gtc gcc gag cgg				384
Ala His Ala Tyr Ala Ile Glu His Gly Ala Arg Ser Val Ala Glu Pro	115	120	125	
115				
tac gag ctg aag gac gag cac ggc acg gtc gtc etc gcc gcg atc gcc				432
Tyr Glu Leu Lys Asp Glu His Gly Thr Val Val Leu Ala Ala Ile Ala	130	135	140	
130				
acc tac ggc aag acc cgc cac acc etc gtc gac cgg acc gcc tac gac				480
Thr Tyr Gly Lys Thr Arg His Thr Leu Val Asp Arg Thr Gly Tyr Asp	145	150	155	160
145				
ggc ccc tac etc ccc ggc tac gtg gcc gcc gcc cgg atc gtc gaa cgg				528
Gly Pro Tyr Leu Pro Gly Tyr Val Ala Ala Ala Pro Ile Val Glu Pro	165	170	175	
165				
ccc gcc cac cgc acc ttc cag gcc atc gac cac tgc gtc gcc aac gtc				576
Pro Ala His Arg Thr Phe Gln Ala Ile Asp His Cys Val Gly Asn Val	180	185	190	
180				

[0027]

gag ctc ggc cgg atg aac gaa tgg gtc ggc ttc tac aac aag gtc atg 624
 Glu Leu Gly Arg Met Asn Glu Trp Val Gly Phe Tyr Asn Lys Val Met
 195 200 205

ggc ttc acg aac atg aag gag ttc gtg ggc gac gac atc gcg acc gag 672
 Gly Phe Thr Asn Met Lys Glu Phe Val Gly Asp Asp Ile Ala Thr Glu
 210 215 220

tac tcg gcg ctg atg tcg aag gtc gtg gcc gac ggc acg ctc aag gtc 720
 Tyr Ser Ala Leu Met Ser Lys Val Val Ala Asp Gly Thr Leu Lys Val
 225 230 235 240

aag ttc ccg atc aac gag ccc gcc ctc gcc aag aag aag tcc cag atc 768
 Lys Phe Pro Ile Asn Glu Pro Ala Leu Ala Lys Lys Lys Ser Gln Ile
 245 250 255

gac gag tac ctg gag ttc tac ggc ggc gcg ggc gtc cag cac atc gcg 816
 Asp Glu Tyr Leu Glu Phe Tyr Gly Gly Ala Gly Val Gln His Ile Ala
 260 265 270

ctg aac acg ggt gac atc gtc gag acg gta cgc acg atg cgc gcc gcc 864
 Leu Asn Thr Gly Asp Ile Val Glu Thr Val Arg Thr Met Arg Ala Ala
 275 280 285

ggc gtc cag ttc ctg gac acg ccc gac tcg tac tac gac acc ctc ggg 912
 Gly Val Gln Phe Leu Asp Thr Pro Asp Ser Tyr Tyr Asp Thr Leu Gly
 290 295 300

gag tgg gtg ggc gac acc cgc gtc ccc gtc gac acc ctg cgc gag ctg 960
 Glu Trp Val Gly Asp Thr Arg Val Pro Val Asp Thr Leu Arg Glu Leu
 305 310 315 320

aag atc ctc gcg gac cgc gac gag gac ggc tat ctg ctc cag atc ttc 1008
 Lys Ile Leu Ala Asp Arg Asp Glu Asp Gly Tyr Leu Leu Gln Ile Phe
 325 330 335

acc aag ccg gtc cag gac cgc ccg acg gtc ttc ttc gag atc atc gaa 1056
 Thr Lys Pro Val Gln Asp Arg Pro Thr Val Phe Phe Glu Ile Ile Glu
 340 345 350

cgc cac ggc tcg atg gga ttc ggc aag ggc aac ttc aag gcc ctg ttc 1104
 Arg His Gly Ser Met Gly Phe Gly Lys Gly Asn Phe Lys Ala Leu Phe
 355 360 365

gag gcg atc gag cgg gag cag gag aag cgg ggc aac ctg tag 1146
 Glu Ala Ile Glu Arg Glu Gln Glu Lys Arg Gly Asn Leu
 370 375 380

<210> 18

<211> 381

<212> PRT

<213> 阿维链霉菌 (Streptomyces avermitilis)

<400> 18

Met Thr Gln Thr Thr His His Thr Pro Asp Thr Ala Arg Gln Ala Asp
 1 5 10 15

Pro Phe Pro Val Lys Gly Met Asp Ala Val Val Phe Ala Val Gly Asn
 20 25 30

Ala Lys Gln Ala Ala His Tyr Tyr Ser Thr Ala Phe Gly Met Gln Leu
 35 40 45

[0028]

Val Ala Tyr Ser Gly Pro Glu Asn Gly Ser Arg Glu Thr Ala Ser Tyr
 50 55 60
 Val Leu Thr Asn Gly Ser Ala Arg Phe Val Leu Thr Ser Val Ile Lys
 65 70 75 80
 Pro Ala Thr Pro Trp Gly His Phe Leu Ala Asp His Val Ala Glu His
 85 90 95
 Gly Asp Gly Val Val Asp Leu Ala Ile Glu Val Pro Asp Ala Arg Ala
 100 105 110
 Ala His Ala Tyr Ala Ile Glu His Gly Ala Arg Ser Val Ala Glu Pro
 115 120 125
 Tyr Glu Leu Lys Asp Glu His Gly Thr Val Val Leu Ala Ala Ile Ala
 130 135 140
 Thr Tyr Gly Lys Thr Arg His Thr Leu Val Asp Arg Thr Gly Tyr Asp
 145 150 155 160
 Gly Pro Tyr Leu Pro Gly Tyr Val Ala Ala Ala Pro Ile Val Glu Pro
 165 170 175
 Pro Ala His Arg Thr Phe Gln Ala Ile Asp His Cys Val Gly Asn Val
 180 185 190
 Glu Leu Gly Arg Met Asn Glu Trp Val Gly Phe Tyr Asn Lys Val Met
 195 200 205
 Gly Phe Thr Asn Met Lys Glu Phe Val Gly Asp Asp Ile Ala Thr Glu
 210 215 220
 Tyr Ser Ala Leu Met Ser Lys Val Val Ala Asp Gly Thr Leu Lys Val
 225 230 235 240
 Lys Phe Pro Ile Asn Glu Pro Ala Leu Ala Lys Lys Lys Ser Gln Ile
 245 250 255
 Asp Glu Tyr Leu Glu Phe Tyr Gly Gly Ala Gly Val Gln His Ile Ala
 260 265 270
 Leu Asn Thr Gly Asp Ile Val Glu Thr Val Arg Thr Met Arg Ala Ala
 275 280 285
 Gly Val Gln Phe Leu Asp Thr Pro Asp Ser Tyr Tyr Asp Thr Leu Gly
 290 295 300
 Glu Trp Val Gly Asp Thr Arg Val Pro Val Asp Thr Leu Arg Glu Leu
 305 310 315 320
 Lys Ile Leu Ala Asp Arg Asp Glu Asp Gly Tyr Leu Leu Gln Ile Phe
 325 330 335
 Thr Lys Pro Val Gln Asp Arg Pro Thr Val Phe Phe Glu Ile Ile Glu
 340 345 350
 Arg His Gly Ser Met Gly Phe Gly Lys Gly Asn Phe Lys Ala Leu Phe
 355 360 365
 Glu Ala Ile Glu Arg Glu Gln Glu Lys Arg Gly Asn Leu
 370 375 380
 <210> 19
 <211> 45

[0029]

<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	
<400> 19	
ccatggctca tcaccatcac catcaccaaa acgccgcgt ttcag	45
<210> 20	
<211> 27	
<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	
<400> 20	
tctagatcat cccactaact gtttggc	27
<210> 21	
<211> 51	
<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	
<400> 21	
ccatggetca tcaccatcac catcacgcag atctatagca aaacccaatg g	51
<210> 22	
<211> 29	
<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	
<400> 22	
tctagattaa tcggcgggtca atacaccac	29
<210> 23	
<211> 33	
<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	
<220>	
<221> misc 特征	
<222> (16)..(18)	
<223> n为a、c、g或t	
<400> 23	
ggtggttttg gcaaannnaa tttctctgag ctc	33
<210> 24	
<211> 33	
<212> DNA	
<213> 人工	
<220>	
<223> 引物序列	

[0030]

<220>		
<221>	misc 特征	
<222>	(16).. (18)	
<223>	n为a、c、g或t	
<400>	24	
	gagctcagag aaattnnntt tgccaaaacc acc	33
<210>	25	
<211>	33	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<220>		
<221>	misc 特征	
<222>	(16).. (18)	
<223>	n为a、c、g或t	
<400>	25	
	cagcgccttg aagttnnnct cgcaaaacc atc	33
<210>	26	
<211>	33	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<220>		
<221>	misc 特征	
<222>	(16).. (18)	
<223>	n为a、c、g或t	
<400>	26	
	gatgggtttg gcgagnnaa ctccaaggeg ctg	33
<210>	27	
<211>	22	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	27	
	gatcttctcg gaaacctga tg	22
<210>	28	
<211>	22	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	28	
	gggattcttg tagacagaga tg	22
<210>	29	
<211>	20	

[0031]

<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	29	
	cccactaact gtttgettc	20
<210>	30	
<211>	20	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	30	
	ggcggteaac acaccaegac	20
<210>	31	
<211>	23	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	31	
	gactcgaaca gcgccttgaa gtt	23
<210>	32	
<211>	18	
<212>	DNA	
<213>	人工	
<220>		
<223>	引物序列	
<400>	32	
	ggatgtggtg gttttggc	18

```

-----MTYNN--KGFPERG-----RELFHHS
MAPAADSPFLQ-----PAQPSD-----LN-----QYRGYDH
MAPGALLVTSQNGRTSELYSDGYVP---APAALVVGGE---VNYRGYHH
MPPTFTTAA--TGAAAATVPEHARP---HR--MVFNERSDRFHTLSFHH
MPPTTAAAAGAAVAAAQAAAFVJGHRNFVRFNPRSDRFHTLAFHH
MGHQNAAVSE--NQNHDDGAAASPGFKLVGFSKVFVKNEKSKDKFKVKREHH
MGKK--QSEAE--ILSSNSNTSEATFKLVGNFVFRANPKSDHFAVKREHH
-----NTQJTHHT-----PDTAQAADP-----FPVKGMDA
-----MADLYEN-----PMG-----LMGFEE

```

小鼠
粗球孢子菌
小麦叶枯病菌
大麦
玉米
拟南芥
胡萝卜
阿维链霉菌
荧光假单胞菌

! 1
! 10

```

VTFWVGNAKQAASEYCNKMGFEFLAYRGLLETGSRVSVSHVVKRGIIVFL
VHWYVGNAKQAATYYVTRMGFEFVAYRGLLETGSKAVASHVVRNGNITFIL
AEWVYGNAKQVAQFYITRMGFEFVAVRKHGLETGSRFFASHVQVQNGVRFVF
VEFWCADAAASAAREAFALGAPLAARELSTLSTGNSAHAQLLRSGSLAFLF
VELWCADAAASAAGRFSFGLGAPLAARSDLSTGNSAHAQLLRSGSLAFLF
IEFWCGDATVTRRFRFVGLMPLVAKSDLSLSTGNSVHASLYLRSANLSEFE
IEFWCGDATVTRRFRFVGLMPLVAKSDLSLSTGNSVHASLYLRSANLSEFE
WFFAVGNAKQAA--HYSTAFGQVAYVSGPENGSRRETASYVLTNGSARFVL
IELASETENTLEIFEIFIMGETKVATHR-----SKD---VHLYRQGAINLIL

```

小鼠
粗球孢子菌
小麦叶枯病菌
大麦
玉米
拟南芥
胡萝卜
阿维链霉菌
荧光假单胞菌

! 20
! 30
! 40
! 50

荧光假单胞菌编号

```

CSALN-----PWN-----KEMGDHLVKHGDGKVDIAFEVEDC
TSPLR--SVEQA---SRFP--EDEALLKEIHAHLERHGDGKVDVAFEVDCV
TSPVRSARQT---LKAAPLADQARLDYDHLKKGDKVDVAFEVDDV
TAPIYANGCDA-----TASLPSFSAARFSAADHGIKVDVAFEVDDV
TAPIYAHGADAA-----TAALPSFSAARFSAADHGLAVRAVALRVADA
TAPYSPSLSAGEIKPTTTASIPFSDHCSRSFFSSHGLGVRVAIAEVEDA
TAPYSPSTTTSSGS-----AAIPFSAAGFHSFAKHGLAVRAIAIEVADV
TSVIK--P--A---TPWG-----HFLADHVAEHGDGVDLAIIEVDDA
NNEFH-----S-----VASYFAAEHGFPSVCGMAFRVKDS

```

小鼠
粗球孢子菌
小麦叶枯病菌
大麦
玉米
拟南芥
胡萝卜
阿维链霉菌
荧光假单胞菌

! 60
! 70
! 80

荧光假单胞菌编号

图 1

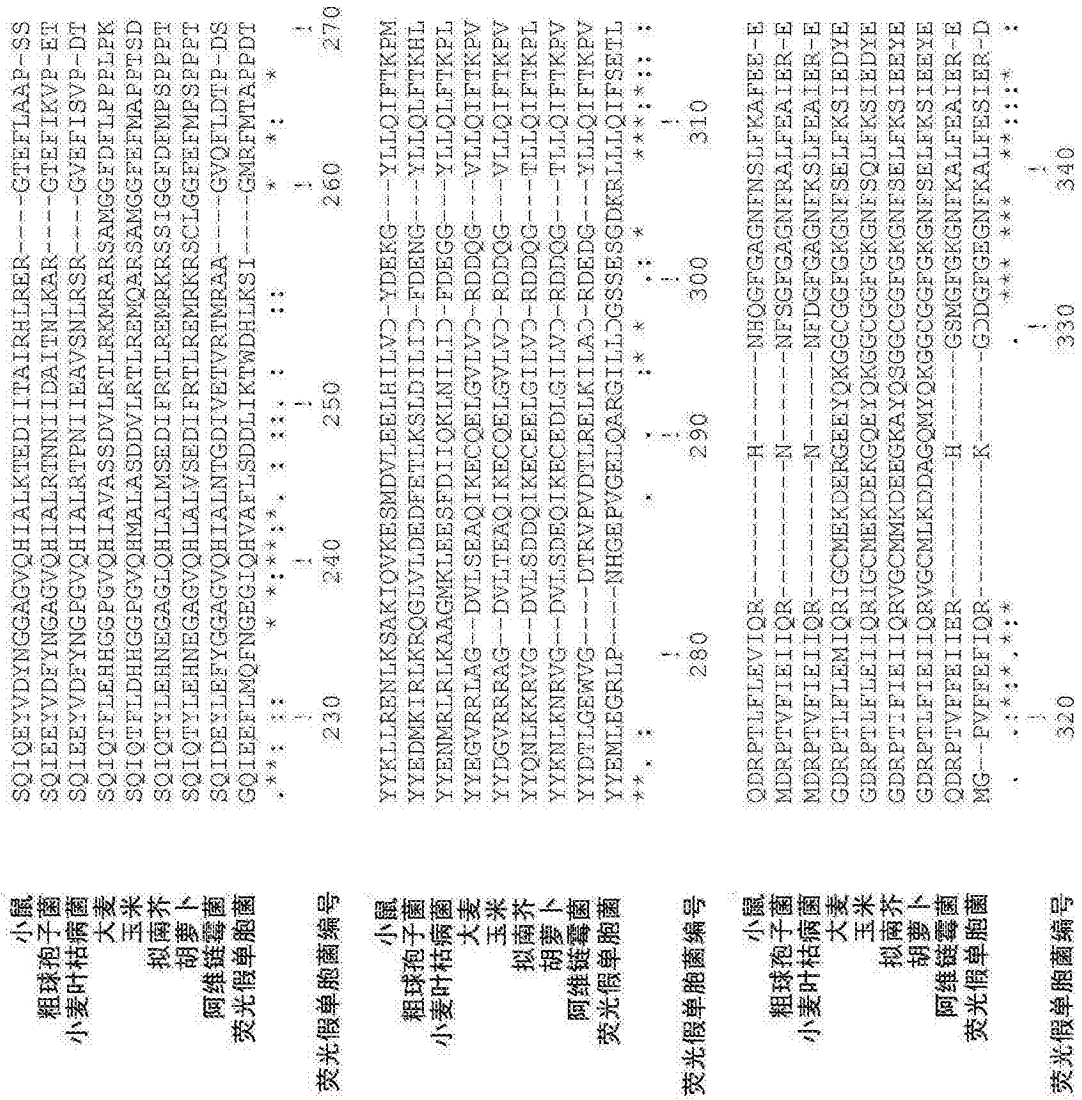


图1(续)

小鼠	QALRGNLTDLEPENGVRKSGM	
粗球孢子菌	QALRGTLI-----	
小麦叶枯病菌	QDLRGNL-----	
大麦	KSLEAKQS---AAVQGS---	
玉米	KSLEAKQAAAAAAQAQGS---	
拟南芥	KTLEAKQIVG-----	
胡萝卜	KTLEAKQITGSAAA-----	
阿维链霉菌	QEKRGNL-----	
荧光假单胞菌	QVRRGVLTD-----	
	: **	
	↓	
		350

荧光假单胞菌编号

图 1(终)