



(12) 发明专利

(10) 授权公告号 CN 102770544 B

(45) 授权公告日 2015. 04. 15

(21) 申请号 201180006977. 9

(51) Int. Cl.

(22) 申请日 2011. 01. 24

C12N 15/82(2006. 01)

(30) 优先权数据

C12N 9/10(2006. 01)

10000712. 9 2010. 01. 25 EP

A01H 5/00(2006. 01)

61/297, 973 2010. 01. 25 US

A01H 5/10(2006. 01)

审查员 徐丹

(85) PCT国际申请进入国家阶段日

2012. 07. 24

(86) PCT国际申请的申请数据

PCT/EP2011/000365 2011. 01. 24

(87) PCT国际申请的公布数据

W02011/089021 EN 2011. 07. 28

(73) 专利权人 拜尔作物科学公司

地址 比利时迪海姆

专利权人 文森特·布洛内 杰亚·圭里罗

(72) 发明人 文森特·布洛内 杰亚·圭里罗

马克·德布洛克

弗兰克·莫伊勒韦特

(74) 专利代理机构 北京市中咨律师事务所

11247

代理人 史文静 黄革生

权利要求书1页 说明书20页

序列表39页 附图2页

(54) 发明名称

制备包含几丁质的植物细胞壁的方法

(57) 摘要

本发明提供了改变植物次生细胞壁的反应性,尤其棉花纤维中存在的棉花细胞壁的反应性的方法和工具。这可通过在棉花植物中表达编码同丝水霉几丁质合酶的嵌合基因方便地实现。

1. 编码几丁质合酶多肽的分离的核酸分子,其中所述分离的核酸分子的核苷酸序列由 SEQ ID NO:1 的核苷酸序列或它的互补序列组成。
2. 编码几丁质合酶多肽的分离的核酸分子,所述几丁质合酶多肽由 SEQ ID NO:2 的氨基酸序列组成。
3. 嵌合基因,其包含以下有效地连接的 DNA 区:
 - a) 植物可表达的启动子,
 - b) 编码根据权利要求 1 或 2 的任一项中的几丁质合酶多肽的 DNA 区,
 - c) 转录终止以及多聚腺苷酸化区。
4. 根据权利要求 3 的嵌合基因,其中所述启动子是组成型启动子。
5. 根据权利要求 3 的嵌合基因,其中所述启动子是纤维选择性启动子。
6. 制备包含带正电多糖的植物细胞壁的方法,所述方法包括
 - a) 在植物中表达根据权利要求 3 至 5 的任一项中所述的嵌合基因,
 - b) 和分离所述植物细胞壁。
7. 根据权利要求 6 的方法,其中所述植物细胞壁是棉花植物细胞壁。
8. 根据权利要求 7 的方法,其中所述棉花植物细胞壁存在于棉花纤维中。

制备包含几丁质的植物细胞壁的方法

发明领域

[0001] 本发明涉及改变植物细胞壁（例如次生植物细胞壁）的反应性。具体地，本发明提供了包含带正电多糖（例如几丁质和壳聚糖）的棉纤维。这些棉纤维具有改变的反应性，所述改变的反应性可被利用以使用纤维反应性染料将纤维染色。另外，可应用所述改变的反应性来提高纤维与反应物（如阻燃剂、水、拒油剂以及拒污剂、软化剂、抗静电剂、荧光增白剂等）的反应性。

[0002] 发明背景

[0003] 人类数千年来一直使用天然纤维，包括含有纤维素的、来自植物（例如棉花和亚麻）的天然纤维，以生产许多不同种类的纺织品。随着要求的增加以及纤维质量的改进，基于棉花的全球纺织工业发展迅速。现今，棉花提供了约 45% 的全球服装，并且许多最好的时装店都使用由高质量棉花制成的纺织品。棉纤维由纤维素组成，所述纤维素是由许多葡萄糖分子组成的天然聚合物。棉纤维独特的结构理想地适合于纺织品生产。每根纤维基本上都是几厘米长的中空管，当纺织和编织时提供非常特别的特征性棉“感”。然而，因为由 β -1-4 连接的葡萄糖单体组成的纤维素相对惰性的性质，所以含有纤维素的天然纤维不具有合成纤维的化学多功能性。这种相对惰性性质在棉纤维和织物的染色过程中是明显的。直接染料和纤维反应性染料是用于棉花着色的两类阴离子染料。棉花本身在水中产生阴离子电荷，使得不用特殊处理，纤维或织物即可非常巧妙地吸收染料。直接染料与纤维素聚合物建立了相对较弱的氢键，形成半永久性连接，这是因为它经不起彻底洗涤。纤维反应性染料是将发色团与反应性基团组合在一起的分子，所述反应性基团通过与羟基反应而与纤维形成强共价键。共价键提供了染色的纤维对洗涤的良好的抗性。在染色过程期间，需要大量电解质来保护这些阴离子染料免受阴离子纤维电荷的影响。未反应的经水解的染料（多达 40%）需要通过洗涤步骤去除，从而产生也含有上文提到的电解质的大量废水。

[0004] 对纤维素纤维提供正电荷（例如通过掺入带正电的化合物）可因此改进天然纤维素纤维的可染性，并且改进改性纤维素纤维与带负电的化合物的任何化学反应。还将可能使用酸性染料。

[0005] 一些出版物已经描述了将壳聚糖寡聚物涂布在纤维素纤维上以产生壳聚糖 / 纤维素掺合物、纱线或织物。壳聚糖是带正电的葡糖胺聚合物，其可以通过几丁质的脱乙酰作用，例如通过碱处理来获得。几丁质本身是 β -1-4 连接的 N-乙酰葡糖胺 (GlcNAc) 的聚合物。

[0006] Liu 等人 (Carbohydrate Polymers 44 (2003) 233-238) 描述了用壳聚糖对棉纤维进行表面涂布的化学方法。具有了这种壳聚糖涂层，棉纤维表面变得具有生理学和生物学活性，这是因为氨基的化学反应性导致对纤维进行随后的化学修饰的可能性更大。基于壳聚糖抑制例如皮肤真菌的生理功能，许多功能性衣服、织物以及纤维都采用纤维素-壳聚糖掺合纤维、纤维素纤维-壳聚糖缀合物、以及涂有含壳聚糖树脂的织物。

[0007] W000/09729 描述了在植物中表达几丁质合酶和几丁质脱乙酰酶基因以改变细胞壁，供工业应用并且提高抗病性。特别提到的用途是：提供纤维素、几丁质以及壳聚糖的单

个植物来源,增加拉伸强度并增加脆断。特别提出的几丁质合酶基因源自真菌生物。既没提供关于在植物中产生几丁质或壳聚糖的实验数据,也没提供关于在植物细胞壁中掺入几丁质或壳聚糖的实验数据。W02006/136351 显示,W000/09729 中所提出的策略并不能导致将几丁质功能性地掺入到植物细胞壁中。而 W02006/136351 公开了只有当真菌几丁质合酶(来自粗糙脉孢菌(*Neurospora crassa*))主动重新定位至高尔基体时,才在棉纤维的次生细胞壁中有效地产生几丁质。通过将这种真菌几丁质合酶与对高尔基体特异的信号锚序列有效地融合,并且在植物中表达所得的嵌合基因,实现几丁质合酶主动重新定位至高尔基体。虽然几丁质可在棉花植物的细胞壁中有效地产生,但还观察到转基因植物依然较小,推定是因为嵌合几丁质合酶在棉花中的棉纤维外部的表达具有毒性。

[0008] 因此,仍然需要备选方法,以产生包含带正电多糖的植物细胞壁(例如次生细胞壁)。具体地,存在提供用于产生纤维的方法的需要,所述纤维可以直接从植物中收获、并且含有带正电的化学基团和/或反应性比纤维素的羟基更强的基团,并且不需引入此类的化学基团的进一步化学处理即可直接应用,并且其中转基因植物不显示生长迟缓。如下文在不同实施方案、实例以及权利要求书中所描述的解决了这些以及其他问题。

[0009] 发明概述

[0010] 本发明显示,在植物细胞中表达源自同丝水霉(*Saprolegnia monoica*)的几丁质合酶(这种酶没有与高尔基体靶向信号有效地连接),可使 N-乙酰葡萄糖胺聚合物有效地掺入到植物细胞壁、尤其植物次生细胞壁中。在不同的实施方案、实例以及权利要求书中所描述的本发明提供了水霉属几丁质合酶以及包含于表达盒中的它的嵌合基因,所述表达盒可以用于用带正电多糖修饰植物细胞壁。因此,本发明的第一目标提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子,其中所述核苷酸序列与 SEQ ID NO:1 或它的互补序列具有至少 97% 的同一性。在另一个实施方案中,本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子,所述几丁质合酶多肽与 SEQ ID NO:2 具有至少 80% 的同一性。在又一个实施方案中,本发明提供了嵌合基因,其包含以下有效连接的 DNA 区:a) 植物可表达的启动子,b) 与 SEQ ID NO:1 或它的互补序列具有至少 97% 的同一性的核苷酸序列,或编码与 SEQ ID NO:2 具有至少 80% 的同一性的几丁质合酶多肽的核苷酸序列,以及 c) 转录终止以及多聚腺苷酸化区。在另一个实施方案中,嵌合基因包含组成型启动子。在又一个实施方案中,嵌合基因包含纤维选择性启动子。在又一个实施方案中,嵌合基因包含扩展蛋白启动子。在又一个实施方案中,本发明提供了如之前实施方案中任一个所定义的嵌合基因。在又一个实施方案中,本发明提供了基本上由植物细胞组成的植物,植物细胞包含如之前实施方案中任一个所定义的嵌合基因。在特定实施方案中,植物是棉花植物。在又一个实施方案中,本发明提供了从根据本发明的植物获得的纤维。在又一个实施方案中,本发明提供了包含如之前实施方案中任一个所定义的嵌合基因的转基因种子。在又一个实施方案中,本发明提供了制备包含带正电多糖的植物细胞壁的方法,所述方法包括在植物中表达根据之前实施方案中任一个的嵌合基因,并且分离所述植物细胞壁。在特定的实施方案中,所述方法提供了棉花植物的细胞壁。在又一个特定的实施方案中,所述方法提供了存在于棉纤维中的棉花植物的细胞壁。。

[0011] 附图

[0012] 图 1:

[0013] 图 1a 显示了经罗丹明缀合的几丁质结合探针染色的未转化的成熟毛状体（拟南芥），而图 1b 显示了经罗丹明缀合的几丁质结合探针染色的包含 SmCHS2 嵌合基因的重组成熟毛状体（拟南芥）。观察到与几丁质结合探针的明显反应，这是因为重组毛状体（图 1b）的染色比源自对照植物的毛状体（图 1a）更强。重组毛状体中红点的存在，证明有几丁质（图 1b），而未转化的毛状体中无红点（图 1a）。

[0014] 本发明的不同实施方案的详细描述

[0015] 本发明是基于以下值得注意的发现：在植物细胞中表达编码几丁质合酶的同丝水霉基因导致 N-乙酰葡萄糖胺 (GlcNAc) 聚合物被有效地掺入到植物细胞壁中。编码几丁质合酶的同丝水霉基因不需要与高尔基体靶向定位信号有效地连接（或融合），而在 WO2006/136351 中的粗糙脉孢菌几丁质合酶是需要的。在本发明中，观察到几丁质（即 N-乙酰葡萄糖胺的聚合物）是在转基因植物的细胞壁、特别是次生细胞壁中产生的。编码几丁质合酶的同丝水霉基因可以用于例如在棉花植物中表达以产生反应性更强的棉纤维。

[0016] 因此，在第一个实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，其中所述核苷酸序列与 SEQ ID NO:1 或它的互补序列具有至少 97% 的同一性。SEQ ID NO:1 描述了同丝水霉几丁质合酶 2 (SmCHS2) 的核苷酸序列。

[0017] 在另一个实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，所述几丁质合酶多肽与 SEQ ID NO:2 具有至少 80% 的同一性。SEQ ID NO:2 描述了同丝水霉几丁质合酶 2 (SmCHS2) 的氨基酸序列。在具体的实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，所述几丁质合酶多肽与 SEQ ID NO:2 具有至少 85%、至少 90%、至少 95% 或至少 98% 或甚至至少 99% 的同一性。在又一个实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，其中所述核苷酸序列与 SEQ ID NO:3 或它的互补序列具有至少 90% 的同一性。在具体的实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，其中所述核苷酸序列与 SEQ ID NO:3 或它的互补序列具有至少 95%、至少 98 或甚至至少 99% 的同一性。SEQ ID NO:3 是同丝水霉几丁质合酶 1 (SmCHS1)。在又一个实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，所述几丁质合酶多肽与 SEQ ID NO:4 具有至少 80% 的同一性。在具体的实施方案中，本发明提供了包含编码几丁质合酶多肽的核苷酸序列的分离的核酸分子，所述几丁质合酶多肽与 SEQ ID NO:4 具有至少 85%、至少 90% 同一性、至少 95%、至少 98% 或甚至至少 99% 的同一性。SEQ ID NO:4 是同丝水霉几丁质合酶 1 的氨基酸序列。

[0018] 核酸可以是 DNA 或单链或双链的 RNA。核酸可以化学合成或通过体外或甚至体内生物表达来产生。

[0019] 可使用适当保护的核糖核苷亚磷酰胺以及常规 DNA/RNA 合成仪化学合成核酸。RNA 合成试剂的供应商是 Proligo (Hamburg, 德国)、Dharmacon Research (Lafayette, CO, USA)、Pierce Chemical (Perbio Science 的分部, Rockford, IL, USA)、Glen Research (Sterling, VA, USA)、ChemGenes (Ashland, MA, USA)、以及 Cruachem (Glasgow, UK)。

[0020] 与本发明的嵌合基因有关的 DNA 包括 cDNA 和基因组 DNA。

[0021] 在酶学中，几丁质合酶 (EC2.4.1.16) 是催化以下化学反应的酶：

[0022]
$$\text{UDP-N-乙酰-D-葡萄糖胺} + [1,4-(\text{N-乙酰}-\beta\text{-D-葡萄糖胺基})]_n \rightarrow \text{UDP} + [1,4-(\text{N-乙酰}-\beta\text{-D-葡萄糖胺基})]_n$$

酰-β-D-葡萄糖胺基)]_{n+1}

[0023] 因此,这种酶的一种底物是 UDP-N-乙酰-D-葡萄糖胺和 [[[1,4-(N-乙酰-β-D-葡萄糖胺基)]_n]],而它的两种产物是 UDP 和 [[[1,4-(N-乙酰-β-D-葡萄糖胺基)]_{n+1}]].这种酶属于糖基转移酶家族,具体地是己糖基转移酶家族。这种酶类别的系统名称是 UDP-N-乙酰-D-葡萄糖胺:几丁质 4-β-N-乙酰葡萄糖胺基-转移酶。其他常用名称包括几丁质-UDP N-乙酰葡萄糖胺基转移酶、几丁质-尿苷二磷酸乙酰葡萄糖胺基转移酶、几丁质合成酶、几丁质合酶以及反式-N-乙酰葡萄糖胺苷酶。这种酶参与氨基糖代谢。因此,几丁质是 β-1,4-连接的 N-乙酰-D-葡萄糖胺 (GlcNAc) 的聚合物。

[0024] 几丁质合酶 1 和 2(分别描述于 SEQ ID NO:3 和 1 中)都源自同丝水霉。水霉属于卵菌纲(Oomycetes)家族。根据一般形态和生活方式,这一组群传统上被划分为真菌。然而,基于现代观点的分支分类法支持它与真核生物组群—长短鞭毛体中的光合生物诸如褐藻和硅藻的关系相对密切。源自同丝水霉的几丁质合酶 1 和 2 的序列包含约 70 个氨基酸的保守氨基末端结构域,也称为 MIT 结构域。MIT 结构域是包含约 70 个氨基酸的保守结构域,具有交替的保守疏水区与极性区,和预测的螺旋二级结构,所述二级结构在涉及微管结合与胞内运输的蛋白质中首先被鉴定出。将这类结构域命名为 MIT 结构域(包含在微管相互作用和运输分子内)。MIT 结构域还存在于以下蛋白质中:分选连接蛋白、核硫醇蛋白酶 PalBH、具有类似结构域结构的 AAA 蛋白痉挛蛋白(spastin)和古细菌蛋白、液泡分选蛋白以及其他。MIT 结构域的确切分子功能尚不清楚。并非意在将本发明局限于特定的作用模式,认为同丝水霉的几丁质合酶中存在的 MIT 结构域是转基因植物的细胞壁中几丁质的有效合成的原因。之前在 W02006/136351 中已显示,真菌来源的几丁质合酶(例如粗糙脉孢菌几丁质合酶)必须通过将真菌几丁质合酶与高尔基定位信号有效地偶联以适应于获得转基因植物的细胞壁中几丁质的有效合成。因为 MIT 结构域存在于几个其他几丁质合酶中,尤其存在于一些源自卵菌纲的几丁质合酶中,所以本发明提供了天然包含 MIT 结构域的几丁质合酶用于在植物次生细胞壁中制备几丁质的用途。此类天然包含 MIT 结构域的几丁质合酶包括寄生水霉(*Saprolegnia parasitica*)的几丁质合酶,其具有 SEQ ID NO:22 和 SEQ ID NO:24 中描述的氨基酸序列,或由 SEQ ID NO:21 和 SEQ ID NO:23 中描述的核酸序列编码的氨基酸序列。

[0025] 在另一个实施方案中,本发明提供了 MIT 结构域与不包含 MIT 结构域的几丁质合酶(例如真菌几丁质合酶,例如非卵菌纲几丁质合酶)的有效偶联,以生成用于在植物次生细胞壁中有效生产几丁质的嵌合几丁质合酶。换言之,公开的是嵌合基因,其包含编码不含 MIT 结构域的几丁质合酶多肽的 DNA 区(该 DNA 区有效地连接至编码 MIT 结构域的 DNA 区,使得表达时产生融合蛋白)、植物可表达的启动子、以及转录终止以及多聚腺苷酸化区。还公开了用于制备包含带正电多糖的植物细胞壁的方法,其包括 a) 表达嵌合基因,其包含编码不含 MIT 结构域的几丁质合酶多肽的 DNA 区(该 DNA 区有效地连接至一个编码 MIT 结构域的 DNA 区,使得表达时产生融合蛋白);和 b) 分离 a) 中所获得的植物细胞壁。

[0026] MIT 结构域被指定为 Pfam 家族,编号 PF04212。目前,已知 495 种蛋白质现获悉或预测包含 MIT 结构域。这些蛋白质的序列可例如从 UniProt 服务器(<http://www.uniprot.org/>)用以下 UniProt 登录号检索,登录号之后是(预测的)MIT 结构域在序列中的位置。将所有序列以及尤其是(预测的)MIT 结构域的序列以其整体并入作为参考。

[0027] C5L7B3/6-74 ; C5K7I8/6-74 ; B9PJ69/9-77 ; B6K9M2/9-77 ; B9QA65/9-77 ; A7ASR9/7-74 ; B6AJD9/4-72 ; Q5CFS7/1-69 ; Q5CSB4/3-71 ; Q4UC87/4-82 ; B3L9J0/6-74 ; A5K3I1/6-74 ; Q8IKQ5/6-74 ; Q4Z291/6-74 ; Q4X5E3/6-74 ; Q7RRP6/6-74 ; A8IAJ1/7-74 ; Q6ETH5/5-73 ; B8AI60/5-73 ; B9SG62/5-73 ; B9HVV7/5-73 ; B9HL02/5-73 ; B9NGD1/5-73 ; Q3EDG2/1-41 ; Q9SGD4/1-41 ; Q8LKV4/5-73 ; B6TLN7/5-73 ; B8A2I4/5-73 ; B6T3Y2/5-73 ; A2WKH8/5-73 ; A2ZP36/5-73 ; B8A2W9/5-73 ; A2WKI0/5-73 ; Q5ZEN9/5-73 ; Q8LAK9/5-73 ; Q9ZNT0/5-73 ; Q9SEA8/5-73 ; Q1W2L1/5-73 ; B9HQW8/5-73 ; B9H1R8/5-73 ; A5BIG1/5-73 ; A7R0D5/5-73 ; A9SGM2/5-73 ; A9TBU2/5-73 ; A9P2N1/5-73 ; B9SCR4/5-73 ; A4S3E8/8-75 ; C1ECR7/11-78 ; C1NA06/11-78 ; C1E2F1/103-169 ; B6K5C2/6-74 ; Q09803/6-74 ; A8PSV3/1-33 ; Q5KC30/6-74 ; Q4PDZ4/6-74 ; A8N0F3/6-74 ; B0DXQ0/6-74 ; A7F3H9/6-74 ; Q7S0H4/6-74 ; Q2GQ74/6-74 ; B2AFE6/6-74 ; C5FLK6/6-74 ; C5JDP2/6-74 ; C5GXE6/6-74 ; C0NGS1/6-74 ; C0SHS5/6-74 ; C1H9G7/6-74 ; C1GCX1/6-74 ; C4JW95/6-74 ; Q1E6C4/6-74 ; B6QQZ4/6-74 ; B8M727/6-74 ; Q2UQD2/6-74 ; B8MZP8/6-74 ; A1D7B7/7-75 ; B0XY62/7-75 ; Q4WXF8/7-75 ; Q5B8R9/6-74 ; Q0CXN9/6-74 ; B6GYF9/6-74 ; A2R7C1/6-74 ; A1CK47/6-74 ; B2VXZ4/6-65 ; Q0U7R6/6-74 ; Q6CEE2/6-74 ; Q5YKJ0/7-75 ; C4R134/4-72 ; Q6FQG5/6-74 ; Q9C1F4/6-74 ; B3LKD3/6-74 ; A6ZX48/6-74 ; P52917/6-74 ; B5VTV5/1-34 ; C5DUT4/6-74 ; Q6CVM8/6-74 ; C5DBA6/6-74 ; A7TH89/6-74 ; Q758U9/6-74 ; C4Y9U8/7-75 ; Q5AGH7/7-75 ; Q5AG40/7-75 ; B9WHM5/7-75 ; A5E2L0/55-123 ; C5MHK4/5-73 ; A5DQ68/6-74 ; A3LVF1/7-75 ; Q6BPY2/7-75 ; A9V5Z2/5-73 ; Q57V58/5-74 ; Q4E658/5-74 ; Q4D9C2/5-74 ; Q4FXF2/5-74 ; A4I4W4/5-74 ; A4HHP9/5-74 ; Q54PT2/6-74 ; Q4RKZ3/1-69 ; A8QBQ7/7-76 ; C3YEH0/5-74 ; C3ZYE4/31-98 ; C4Q408/2-71 ; Q5DBH6/2-71 ; A8Y1H3/5-74 ; Q9BL83/5-74 ; B3RJ28/5-74 ; A7SK75/5-74 ; B3MXW2/6-75 ; Q29H77/6-75 ; B4NPI4/6-75 ; B4L2B2/6-75 ; B3NWZ3/6-75 ; B4M6S6/6-75 ; B4R6Q7/6-75 ; B4I6L5/6-75 ; Q9Y162/6-75 ; B4Q2M1/6-75 ; Q7QFR0/6-75 ; B0XJH8/6-75 ; Q17GP3/6-75 ; B7PVD7/6-75 ; Q66IY7/5-74 ; B2GUK1/3-72 ; B2RCB7/5-74 ; Q9UN37/5-74 ; Q08BZ6/5-74 ; Q4SKA0/1-69 ; Q2HJB1/5-74 ; Q793F9/5-74 ; Q8VEJ9/5-74 ; Q6IRG3/5-74 ; Q3TDX2/5-74 ; Q5U4Y4/7-76 ; Q6DJK7/7-76 ; Q8AVB9/7-76 ; B5X7N1/5-74 ; VPS4B/7-76 ; A8K4G7/7-76 ; Q69HW4/7-76 ; A8K5D8/7-76 ; 075351/7-76 ; Q3TN07/7-76 ; Q6PJZ4/7-76 ; P46467/7-76 ; Q4KLL7/7-76 ; Q3U8P5/7-76 ; Q0VD48/7-76 ; Q4RVG5/6-75 ; Q5ZMI9/4-73 ; C0H991/95-164 ; B5X1U4/6-75 ; Q7SXY0/6-75 ; A7YYH5/6-75 ; B2GU12/6-75 ; A5WWM0/6-75 ; C0H991/9-78 ; Q8T127/7-75 ; B7P0K9/283-324 ; C3Z907/321-389 ; A7S4I8/283-324 ; Q4S8A8/285-352 ; A4IG43/277-345 ; A4IID6/279-347 ; Q4V7Q6/279-347 ; Q5ZJH6/280-348 ; B4DFS6/190-258 ; B4DRQ7/280-348 ; B2RXK3/280-348 ; B4DDG2/163-231 ; B4DQN3/163-231 ; B4DFT0/291-359 ; Q3U3Q1/280-348 ; B2RXB9/280-348 ; B3RSI2/276-344 ; B4GNS6/275-343 ; Q29BD1/275-343 ; B4NAJ4/275-343 ; B3M073/275-343 ; Q86P98/275-343 ; Q8T0L6/275-343 ; Q9VHF6/275-343 ; B3P1R4/275-343 ; B4QX75/275-343 ; B4HKF5/275-343 ; B4PUL9/275-343 ; B4K4Y3/275-343 ; B4LW06/275-343 ; B4JYT0/275-343 ; B0XAE0/178-246 ; Q7QE2/275-343 ; Q5C1I1/163-221 ; C4Q8U0/273-338 ; B0E7C2/4-72 ; C4LYN8/4-72 ; A3DP09/6-74 ; B8D2W2/9-77 ; A9A4K6/8-76 ; B3TCM2/8-75 ; A0RVT9/8-75 ; Q877H3/7-75 ; Q972B3/7-75 ; Q97ZJ7/7-75 ; C3MPU4/7-75 ; C3NE34/7-75 ; C3NHM9/7-75 ; C4KH36/7-75 ;

C3MUV6/7-75 ;C3N5H0/7-75 ;A4YHC5/7-75 ;A2BKZ1/9-75 ;Q9YDF3/7-75 ;C4LZH7/3-71 ;
A8BSU6/8-77 ;Q8T6L4/8-77 ;B8C9Z5/5-73 ;B8LDI1/5-73 ;B7GCY6/5-73 ;Q54CX7/4-72 ;
Q22143/5-70 ;A8XST3/5-70 ;Q54CX7/246-316 ;Q55GN8/5-73 ;A4QZC1/8-77 ;A2D8M7/10-80 ;
Q54RJ5/10-79 ;C4LYH7/2-70 ;Q8I8X1/16-84 ;Q54CX7/120-188 ;Q0UZT0/375-444 ;
C4JEZ5/18-85 ;Q1E904/19-86 ;C5FCB8/23-90 ;C1GDJ4/63-130 ;C0SAK1/19-86 ;
C1HDA2/19-86 ;C5JPG7/78-145 ;C5G855/19-86 ;A6RHR0/1-68 ;C0NMG2/1-68 ;
A2QKF2/239-307 ;B6H2K6/100-168 ;Q0CFY4/216-284 ;B0XW44/210-278 ;A1DG63/193-261 ;
Q4X270/211-279 ;A1CSH7/266-334 ;B8NC13/82-150 ;Q2TZ95/233-301 ;Q5B0X3/178-246 ;
B8M0Z2/44-113 ;B6Q8T3/44-113 ;A7EZ31/1-69 ;A6S2Q3/293-361 ;Q7S0P7/402- 470 ;
B2ADU1/368-424 ;Q2GTY0/352-420 ;C1E4H7/3-68 ;C1MPZ7/3-69 ;C3ZZL4/6-70 ;
B7P5B6/8-73 ;Q9R1S8/6-71 ;Q499T5/6-71 ;Q9Y6W3/6-71 ;B2RAM2/6-71 ;Q7Z479/6-71 ;
A0FKG7/6-71 ;B0FGU2/6-71 ;B4F6P1/6-71 ;C0H9M1/6-71 ;Q4S467/6-71 ;B3RQK1/8-73 ;
A7RRP7/7-70 ;Q54JG4/13-81 ;Q55GN8/176-244 ;C4LYH7/369-438 ;Q8I8X1/383-452 ;
B0EV39/371-440 ;A9V3C1/6-74 ;A7URZ9/6-72 ;B0X0P8/1-69 ;Q16FL8/1-69 ;Q0IFG1/1-69 ;
B4KNS6/15-76 ;B4MS07/21-82 ;B3M8E0/5-71 ;B4IZ41/4-72 ;B4KXT1/5-71 ;B4MFW3/3-71 ;
B4MLA7/3-71 ;Q29EY7/4-71 ;B3NG09/3-71 ;B4PH91/3-71 ;Q9VZK1/3-71 ;Q7YU93/3-71 ;
B4QQ18/3-62 ;B4HTX3/3-71 ;B3RKL7/450-515 ;Q7T332/322-390 ;A9JT54/322-390 ;
A4IIB7/268-336 ;Q28CB9/261-329 ;Q63ZH0/269-337 ;Q0P3R2/270-338 ;A7SVK8/274-341 ;
Q20821/252-320 ;A8Y3X4/266-339 ;Q148E7/275-343 ;B3KU31/268-336 ;B2RDR2/268-336 ;
Q9NRS6/268-336 ;B0CM75/268-336 ;Q4V896/269-337 ;Q91WE1/268-336 ;Q9CW22/63-131 ;
Q6NU31/257-325 ;KS6C1/239-307 ;B1APS8/227-295 ;B3KVM4/239-307 ;B4DRK0/58-126 ;
Q5R5U6/239-307 ;KS6C1/238-306 ;B8JLT3/232-300 ;Q4RR16/253-353 ;B4NG93/233-301 ;
B3MTR0/236-304 ;B3NZ90/246-314 ;Q9VEA9/242-310 ;B4QUH5/242-310 ;B4IB83/242-310 ;
B4PL88/246-314 ;B4GLM4/255-323 ;Q293U2/255-323 ;B4JI50/203-271 ;B4K7H6/227-295 ;
B4M0W2/211-279 ; Q7PYF1/252-320 ;B0X4C2/216-284 ;Q16ZH6/224-292 ;Q4RLU8/6-73 ;
Q8R2S1/49-117 ;B4DSP6/49-117 ;Q5RA67/49-117 ;Q9Y6S9/49-117 ;Q32PG2/49-117 ;
A9RII8/29-104 ;B8B1Z6/49-124 ;Q658G8/49-124 ;B4FVB5/49-124 ;Q944N4/90-165 ;
A8MRR2/55-130 ;Q0WMJ4/55-130 ;O64630/55-130 ;B9RDF4/52-127 ;B9IC21/13-88 ;
A7PY13/56-131 ;A5BB69/56-131 ;A5BIC4/76-120 ;A9V1R3/29-100 ;B3S0V7/15-93 ;
C3XZ15/1-79 ;C3ZJS8/19-88 ;B7PXE3/98-174 ;B4NBP4/237-313 ;B4M0H8/223-299 ;
B4K799/217-293 ;B4JII0/234-310 ;B4G437/239-315 ;B4QSF0/232-308 ;B3M301/235-311 ;
B4HGG6/232-308 ;B3P8A3/232-308 ;Q298L4/239-315 ;B4PL32/232-308 ;Q8I0P1/232-308 ;
Q16IA7/179-255 ;Q7QBW0/230-306 ;B0X289/2-60 ;Q4TCF6/6-91 ;Q4TCD9/16-91 ;
Q6NW58/83-158 ;Q6AZT2/110-185 ;Q05AS3/113-188 ;Q9QYY8/117-193 ;Q9UBP0/119-195 ;
A2VDN5/117-193 ;Q719N1/116-192 ;Q5ZK92/116-192 ;B2RYN7/117-192 ;C3XW83/29-96 ;
B7Q7T2/7-83 ;A7RQX2/15-88 ;Q4RRG4/9-85 ;C0H9Y4/20-96 ;Q0P3W0/20-96 ;Q4SG05/11-87 ;
A8WFZ3/15-90 ;B3DLA2/21-97 ;Q8R1X6/19-95 ;Q3TVW1/19-95 ;A0JNJ3/19-95 ;
Q4R7V2/19-95 ;Q8N0X7/19-95 ;B3KMI3/19-95 ;A8K6Q9/19-95 ;A2Q8K6/7-77 ;Q00204/6-77 ;
B8MY71/5-77 ;Q9Y6Z8/5-77 ;Q0CUT6/6-77 ;A1CRW2/6-77 ;Q1DXZ5/5-72 ;B6QNP8/12-76 ;

B0DXY2/11-78 ;Q7QGN5/7-67 ;B0WDD1/4-67 ;Q179Z1/4-67 ;Q179Z0/4-67 ;B7P5B6/89-157 ;
Q5DBL1/86-142 ; A8P1B5/37-127 ;Q9R1S8/86-154 ;Q7Z479/86-154 ;B2RAM2/86-154 ;
Q9Y6W3/86-154 ;A0FKG7/86-154 ;B0FGU2/86-154 ;Q499T5/86-154 ;C0H9M1/86-154 ;
B4F6P1/86-154 ;Q4S467/86-149 ;Q6NVT4/42-110 ;C1BMZ4/11-74 ;A8PMC5/4-73 ;
Q5D999/6-77 ;Q5D9R4/6-77 ;C4Q8N9/35-106 ;A7RL55/2-70 ;Q5I0J5/12-79 ;Q8VDV8/12-79 ;
Q8WV92/12-79 ;B8ZZL5/12-79 ;A8YXZ4/12-79 ;Q6DJ62/7-75 ;C3XWE6/1-69 ;B0S8J9/9-77 ;
B5XBB0/7-75 ;Q4S0N8/9-77 ;Q54KQ7/168-232 ;B3RXW5/5-72 ;A8PZJ1/282-350 ;
C3Z907/423-490 ;Q4V7Q6/374-442 ;Q3U3Q1/375-444 ;B2RXB9/375-444 ;B4DFT0/386-455 ;
B4DRQ7/375-444 ;B4DDG2/258-327 ;B4DFS6/285-354 ;B4DQN3/258-327 ;B2RXK3/375-444 ;
A4IG43/372-441 ;Q5ZJH6/375-444 ;Q7QE2/417-485 ;B0XAE0/287-363 ;B3M073/404-479 ;
B4HKF5/399-474 ;B3P1R4/399-474 ;B4PUL9/399-474 ;B4QX75/399-474 ;B4JYT0/404-479 ;
Q86P98/399-474 ;Q9VHF6/399-474 ;Q8T0L6/399-464 ;Q29BD1/399-474 ;B4NAJ4/403-478 ;
B4LW06/400-475 ;B4K4Y3/397-472。另外的 MIT 结构域包括以下文献中所公开的那些
MIT 结构域:Ciccarelli 等人 (2003 ;Genomics81 :437-41) 或 Patel 等人 (2002 ;Nature
Genetics31 :347-8) 或 pdb (蛋白质数据库) 文档 (或相关出版物) :1wfd (Suetake 等人 ;
Solution structure of mouse MIT domain) ;1wr0 (Takasu 等人 (2005). Biochem Biophys
Res Commun, 334, 460-465.) ;1yxr (Scott 等人 (2005) Proc Natl Acad Sci U S A. 102 :
13813-13818) ;2cpt (Suetake 等人 ;Solution structure of MIT domain from human
skd1) ;2dl1 (Suetake 等人 ;Solution structure of MIT domain from human skd1) ;2jq9
和 2jqh (Stuchell-Brereton 等人 (2007). Nature, 449, 740-744.) , 2k3w (C. Kieffer 等人
(2008). Dev Cell, 15, 62-73.) , 2v6x, 2v6y (T. Obita 等人 (2007). Nature, 449, 735-739.) ,
2w2u (Samson 等人 (2008). Science, 322, 1710-1713.) , 2zam, 2zan, 2zao (均为 Inoue 等
人 (2008). Traffic, 9, 2180-2189.) 或 3eab (Yang 等人 (2008). Nat Struct Biol, 15,
1278-1286.)。

[0028] 在此所鉴定的几丁质合酶中, MIT 结构域位于 SEQ ID NO:4 (CHS1) 中从 103 位到
171 位的氨基酸序列段 (对应于 SEQ ID NO:20), 以及 SEQ ID NO:2 (CHS2) 中从 141 位到 209
位 (对应于 SEQ ID NO:19) 的氨基酸序列段。

[0029] 其他 MIT 结构域实例显示出与 CHS1 (SEQ ID NO:4) 中存在的 MIT 结构域 (对应于
SEQ ID NO:20) 以及 CHS2 (SEQ ID NO:2) 中存在的 MIT 结构域 (对应于 SEQ ID NO:19)、或
UniProt 序列或以上所列参考文献中存在的任一个 MIT 结构域的序列同一性为至少 80%、至
少 85%、至少 90%、至少 95%、至少 98%、至少 99% 或甚至 100%。在此公开的所有 MIT 结构域都
具有 CHS1 (SEQ ID NO:4) 以及 CHS2 (SEQ ID NO:2) 中存在的 MIT 结构域的生物学功能。
所述生物学功能可以例如如实例中所概述的通过检查包含所述 MIT 结构域的融合蛋白在
植物细胞中的定位来进行评估。

[0030] 如在此所使用的, 术语“序列同一性 %”是指最佳比对 DNA 窗的两个区段之间相
同核苷酸的百分比。用于比对的比较窗的最佳序列 (Isotianil)、代森锰锌 (Mancozeb)、
代森锰 (Maneb)、叉氨苯酰胺 (Metominostrobin)、吡噻菌胺 (Penthiopyrad)、啉氧菌
酯 (Picoxystrobin)、甲基代森锌 (Propineb)、丙硫菌唑 (Prothioconazole)、百克
敏 (Pyraclostrobin)、五氯硝基苯 (Quintozene)、戊唑醇 (Tebuconazole)、氟醚唑

(Tetraconazole)、甲基托布津(Thiophanate-methyl)、布洛芬(Trifloxystrobin)。

[0031] MIT 结构域可以例如与以下几丁质合酶(几丁质 UDP-乙酰-葡萄糖胺基转移酶)中任一项有效地连接,所述几丁质合酶(包括氨基酸)可见于不同数据库,具有以下标识符(登录号):CHS1_AJECA(P30576)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(I 类几丁质合酶 1),来自荚膜组织胞浆菌(*Ajellomyces capsulata* 或 *Histoplasma capsulatum*);CHS1_AJEDE(P30579)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-氨基葡萄糖基转移酶 1)(I 类几丁质合酶 1),来自皮炎组织胞浆菌(*Ajellomyces dermatitidis*)(皮炎芽生菌(*Blastomyces dermatitidis*));CHS1_ASPNG(P30581)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(I 类几丁质合酶 1),来自黑曲霉;CHS1_BOTCI(P49603)几丁质合酶 1(EC 2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(I 类几丁质合酶 1),来自灰葡萄孢(贵腐菌)(富氏葡萄孢盘菌);CHS1_CANAL(P23316)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1),来自白假丝酵母(酵母);CHS1_CRYNV(013356)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(IV 类几丁质合酶 1)。{基因:名称=CHS1}-新型隐球酵母变种格鲁比(*Cryptococcus neoformans* var. *grubii*)(新生线黑粉菌变种格鲁比(*Filobasidiella neoformans* var. *grubii*));CHS1_EMENI(P30583)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-氨基葡萄糖基转移酶 1)(I 类几丁质合酶 1)(片段)。{基因:名称=chs1}-构巢裸胞壳(构巢曲霉);CHS1_EXODE(P30600)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(II 类几丁质合酶 1)。{基因:名称=CHS1}-皮炎外瓶霉(皮炎万氏霉(*Wangiella dermatitidis*));CHS1_EXOJE(P30585)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(I 类几丁质合酶 1)(片段)。{基因:名称=CHS1}-甄氏外瓶霉;CHS1_NEUCR(P29070)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(III 类几丁质合酶 3)。{基因:名称=chs-1;ORF 名称=B11H24.170,NCU03611.1}-粗糙脉孢菌;CHS1_PHAEX(P30590);几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(片段)。{基因:名称=CHS1}-外瓶暗色球菌(*Phaeococcomyces exophialae*);CHS1_PHYBL(P87073)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(II 类几丁质合酶 1)。{基因:名称=chs1}-布拉克须霉(*Phycomyces blakesleeenans*);CHS1_RHIAT(P30592)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(I 类几丁质合酶 1)(片段)。{基因:名称=CHS1}-暗绿喙枝孢(*Rhinocladiella atrovirens*);CHS1_RHIOL(P30594)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)。{基因:名称=CHS1}-少孢根霉;CHS1_RHIRA(Q12632)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(II 类几丁质合酶 1)。{基因:名称=CHS1}-多变根毛霉(*Rhizomucor racemosus*)(毛霉科真菌卷枝毛霉(*Mucor circinelloides* t. *lusitanicus*));CHS1_SCHCO(P30596);几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)(片段)。{基因:名称=CHS1}-裂褶菌(弧状菌);CHS1_SCHPO(P30597)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-葡萄糖胺基转移酶 1)。{基因:名称=chs1;ORF 名称=SPAC13G6.12c,SPAC24B11.01c}-粟酒裂殖酵母(裂殖酵母);CHS1_TUBUN(P55003)几丁质合酶 1(EC2.4.1.16)(几丁质-UDP 乙酰-氨基葡萄糖基转移酶 1)(片段)。{基因:名称=CHS1}-具钩块菌(*Tuber uncinatum* 或 *Burgundy truffle*);

CHS1_USTMA(P30598) 几丁质合酶 1(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 1) (片段)。{基因:名称=CHS1}-玉蜀黍黑粉菌(黑粉菌);CHS1_XYLBA(P30603) 几丁质合酶 1(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 1) (片段)。{基因:名称=CHS1}-班替木丝霉(*Xylohypha bantiana*);CHS1_酵母(P08004) 几丁质合酶 1(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 1)。;CHS2_AJECA(P30577) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (III 类几丁质合酶 2) 荚膜组织胞浆菌;CHS2_AJEDE(P30580) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (II 类几丁质合酶 2) {基因:名称=CHS2}-皮炎芽生菌 CHS2_ASPNG(P30582);几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (II 类几丁质合酶 2) (片段)。{基因:名称=chs2}-黑曲霉;CHS2_CANAL(P30572) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) {基因:名称=CHS2}-白色念珠菌(酵母);CHS2_EXODE(P30601) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (I 类几丁质合酶 2)。{基因:名称=CHS2}-皮炎外瓶霉(皮炎万氏霉);CHS2_EXOJE(P30586) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (片段)。{基因:名称=CHS2}-甄氏外瓶霉;CHS2_NEUCR(P30589);几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2)。{基因:名称=chs-2;ORF 名称=NCU05239.1}-粗糙脉孢菌;CHS2_PARBR(Q92444) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (II 类几丁质合酶 2)。{基因:名称=CHS2}-巴西副球孢子菌;CHS2_PHAEX(P30591);几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (II 类几丁质合酶 2) (片段)。{基因:名称=CHS2}-外瓶暗色球菌;CHS2_RHIAT(P30593) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (III 类几丁质合酶 2) (片段)。{基因:名称=CHS2}-暗绿喙枝孢;CHS2_RHIOL(P30595) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) {基因:名称=CHS2}-少孢根霉;CHS2_SCHPO(074756) 几丁质合酶样蛋白质 2。{基因:名称=chs2;ORF 名称=SPBC1709.01, SPBC1734.17}-粟酒裂殖酵母(裂殖酵母);CHS2_USTMA(P30599) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (片段)。{基因:名称=CHS2}-玉蜀黍黑粉菌(黑粉菌);CHS2_XYLBA(P30604) 几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2) (II 类几丁质合酶 2) (片段)。{基因:名称=CHS2}-班替木丝霉;CHS2_YEAST(P14180);几丁质合酶 2(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2)。{基因:名称=CHS2;顺序基因座名称=YBR038W;ORF 名称=YBR0407}-酿酒酵母(面包酵母);CHS3_AJECA(P30578) 几丁质合酶 3(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 3) (II 类几丁质合酶 3) (片段)。{基因:名称=CHS3}-荚膜组织胞浆菌;CHS3_CANAL(P30573) 几丁质合酶 3(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 3) (IV 类几丁质合酶 3)。{基因:名称=CHS3}-白假丝酵母(酵母);CHS3_EXODE(P30602) 几丁质合酶 3(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 3) (III 类几丁质合酶 3)。{基因:名称=CHS3}-皮炎外瓶霉(皮炎万氏霉);CHS3_EXOJE(P30587);几丁质合酶 3(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 3) (III 类几丁质合酶 3) (片段)。{基因:名称=CHS3}-甄氏外瓶霉;CHS3_NEUCR(P30588) 几丁质合酶 3(EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 3)。{基因:名称=chs-3;ORF 名称=G65A3.040}-粗糙脉孢菌;CHS3_酵母(P29465) 几丁质合酶 3(EC2. 4. 1. 16) (几

丁质 -UDP 乙酰 - 葡糖胺基转移酶 3) (IV 类几丁质合酶 3)。{ 基因 : 名称 = CHS3 ; 同义词 = CAL1、CSD2、DIT101、KIT2 ; 顺序基因座名称 = YBR023C ; ORF 名称 = YBR0305 } - 酿酒酵母 (面包酵母) ; CHS4_MAGGR(013353) ; 几丁质合酶 4 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 4) (IV 类几丁质合酶 4)。{ 基因 : 名称 = CHS4 } - 稻瘟病菌 (水稻稻瘟病菌 (Rice blast fungus 或 Pyricularia grisea)) ; CHS4_NEUCR(Q01285) 几丁质合酶 4 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 4) (IV 类几丁质合酶 4)。{ 基因 : 名称 = chs-4 ; ORF 名称 = NCU09324. 1 } - 粗糙脉孢菌 ; CHS5_USTMA(013394) 几丁质合酶 5 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 5) (IV 类几丁质合酶 5)。{ 基因 : 名称 = CHS5 } - 玉蜀黍黑粉菌 (黑粉菌) ; CHS6_USTMA(013395) 几丁质合酶 6 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 6) (V 类几丁质合酶 6)。{ 基因 : 名称 = CHS6 } - 玉蜀黍黑粉菌 (黑粉菌) ; CHSA_AMPQU(Q12564) ; 几丁质合酶 A (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 A) (I 类几丁质合酶 A)。{ 基因 : 名称 = CHSA } - 宿白粉菌 ; CHSA_EMENI(P30584) 几丁质合酶 A (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 A) (II 类几丁质合酶 A)。{ 基因 : 名称 = chsA ; 同义词 = chs2 } - 构巢裸胞壳 (构巢曲霉) ; CHSB_EMENI(Q00757) 几丁质合酶 B (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 B) (III 类几丁质合酶 B)。{ 基因 : 名称 = chsB } - 构巢裸胞壳 (构巢曲霉) ; CHSC_ASPFU(Q92197) 几丁质合酶 C (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 C) (III 类几丁质合酶 C)。{ 基因 : 名称 = chsC } - 烟曲霉 (萨托菌属 (Sartorya fumigata)) ; CHSD_ASPFU(P78746) 几丁质合酶 D (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 D) (VI 类几丁质合酶 D)。{ 基因 : 名称 = chsD } - 烟曲霉 (萨托菌属) ; CHSD_EMENI(P78611) 几丁质合酶 D (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 D) (V 类几丁质合酶 D)。{ 基因 : 名称 = chsD ; 同义词 = chsE } - 构巢裸胞壳 (构巢曲霉) ; CHSG_ASPFU(P54267) ; 几丁质合酶 G (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 G) (III 类几丁质合酶 G)₁ { 基因 : 名称 = chsG } - 烟曲霉 (萨托菌属) ; CHSX_USTMA(Q99126) 几丁质合酶 1 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 1)。{ 基因 : 名称 = CHS1 } - 玉蜀黍黑粉菌 (黑粉菌) ; 或 CHSY_USTMA(Q99127) 几丁质合酶 2 (EC2. 4. 1. 16) (几丁质 -UDP 乙酰 - 葡糖胺基转移酶 2)。{ 基因 : 名称 = CHS2 } - 玉蜀黍黑粉菌 (黑粉菌)。所有序列并入本文作为参考。

[0032] 有效地连接可导致 MIT 结构域存在于几丁质合酶序列的 N 末端、C 末端或内部, 只要这一位置基本上不损伤所得蛋白质的几丁质合酶活性。就这一点而言, “基本上不损伤”指的是当表达于植物细胞中时, 几丁质合酶的活性仍足以在所述植物的细胞壁中提供可检测测量值的带正电多糖。这指的是几丁质合酶活性是 CHS1 或 CHS2 的几丁质合酶活性的至少 20%、至少 30%、至少 40%、至少 50%、至少 60%、至少 70%、至少 80%、至少 90% 或至少 100%。测定几丁质合酶活性的方法在本领域是已知的, 并且在本申请中也已公开。

[0033] 在特定的实施方案中, 本发明还提供了 SmCHS1 和 SmCHS2 的变体序列, 其可以通过体外或体内诱导产生变异来获得。本领域中可获得体外诱导产生变体核苷酸序列的几种方法, 其包括但不限于 DNA 改组或定向进化技术, 如美国专利 5605793、美国专利 5, 811, 238 以及美国专利 5, 830, 721 中所述。体内诱导产生变异核苷酸序列的方法在本领域中也是熟知的, 并且例如, 如在本描述书的其他地方所述, 可以包括将同丝水霉暴露于诱变剂, 如电离辐射、EMS、MMS 等, 之后分离几丁质合酶 1 或 2 的编码核酸。在另一个特定的实施方案中,

可以通过使密码子用法适应于植物中特异表达来修饰 SmCHS1 和 SmCHS2、尤其是编码它们的核酸。此类用于密码子修饰（或密码子适应，它是等效用语）以便在特定植物中表达的方法在本领域中为熟知的。SEQ ID NO:18 描述了变体 SmCHS2 核苷酸序列，其是经过优化用于在棉花中表达的经密码子修饰的变体。SEQ ID NO:18 编码了在 SEQ ID NO:2 中描述的几丁质合酶多肽。

[0034] 在另一个实施方案中，本发明提供了嵌合基因，其包含以下有效连接的 DNA 区：1) 植物可表达的启动子，2) 编码了几丁质合酶多肽的 DNA 区，其中 DNA 区的核苷酸序列与 SEQ ID NO:1 具有至少 97% 的同源性，或其中 DNA 区的核苷酸序列与 SEQ ID NO:3 至少 90% 的同源性，或其中 DNA 区的核苷酸序列包含编码与 SEQ ID NO:2 具有至少 80% 同源性的几丁质合酶多肽的核苷酸序列，或其中 DNA 区的核苷酸序列包含编码与 SEQ ID NO:4 具有至少 90% 同源性的几丁质合酶多肽的核苷酸序列，以及 3) 转录终止以及多聚腺苷酸化区。

[0035] 根据本发明的嵌合基因包含植物可表达的启动子。如在此所使用的，术语“启动子”表示在转录起始期间被 DNA 依赖的 RNA 聚合酶识别并（直接或间接地）结合的任何 DNA。启动子包括转录起始位点、以及转录起始因子与 RNA 聚合酶的结合位点，并且可包含基因表达调控蛋白可结合的多种其他位点（例如增强子）。

[0036] 如在此所使用的，术语“植物可表达的启动子”指的是能够在植物细胞中控制（起始）转录的 DNA 序列。这包括植物来源的任何启动子，也包括能够在植物细胞中指导转录的任何非植物来源的启动子，即某些病毒或细菌来源的启动子（如 CaMV35S）、地三叶病毒启动子 No4 或 No7、或 T-DNA 基因启动子等。

[0037] 优先在纤维细胞中控制转录的起始和维持的植物可表达的启动子是驱动有效连接的 DNA 区在纤维细胞和底层表皮细胞中的转录水平高于在植物的其他细胞或组织中的转录水平的启动子。较高的水平可以是例如在纤维细胞以及任选地在底层表皮细胞中的转录比在植物的其他细胞或组织中的转录多至少 2 倍、至少 5 倍、至少 10 倍、至少 20 倍、至少 50 倍、至少 100 倍或至少 500 倍。这种定义也适用于术语“纤维选择性启动子”。换言之，该术语包括这样的启动子，其在纤维细胞以及任选地在底层表皮细胞中驱动的转录比在植物的其他细胞或组织中的转录多至少 2 倍、至少 5 倍或至少 10 倍。此类启动子包括来自棉花的纤维特异性 β -微管蛋白基因的启动子（如 W00210377 中所描述）、来自棉花的纤维特异性肌动蛋白基因的启动子（如 W00210413 中所描述）、来自棉花的纤维特异性脂质转移蛋白基因的启动子（如 US5792933 中所描述）、来自棉花的扩展蛋白基因的启动子（W09830698）、或来自棉花的几丁质酶基因的启动子（US2003106097）、或在 US6259003 或 US6166294 中所描述的纤维特异性基因的启动子、或如在 US6096950 中所公开的源自 E6 家族的启动子。在特定的实施方案中，植物可表达的启动子是组成型启动子。如实施例 3 最后 6 行中所讨论，与未转化的植物相比，表达本发明的几丁质合酶并不引起植物生长的差异。

[0038] 在特定的实施方案中，本发明提供了植物细胞，其包含在此之前描述的嵌合基因。

[0039] 该嵌合基因可通过本领域熟知的方法引入到植物细胞中。

[0040] “引入”与本申请结合时涉及通过人工方法将遗传信息放置于植物细胞或植物中。这可通过本领域中已知的用于将 RNA 或 DNA 引入到植物细胞、组织、原生质体或整株植物中的任何方法来实现。

[0041] 许多方法可用于将 DNA 转移到植物细胞或植物中。例如美国专利 5,004,863、美国

专利 6,483,013 或 W02000/71733 中已描述了农杆菌属介导的棉花转化。

[0042] 植物或植物细胞也可以通过粒子轰击来转化：用 DNA 涂布金或钨颗粒，然后射入年幼植物细胞或植物胚芽中。这一方法也可允许植物质体的转化。例如 W092/15675 中报导了通过粒子轰击的棉花转化。

[0043] 病毒转化（转导）可以用于基因的瞬时或稳定表达，这取决于病毒基因组的性质。所希望的遗传物质被包装到适合的植物病毒中，并且允许经修饰的病毒感染植物或植物细胞。经感染的植物的子代无病毒，并且无插入基因。在示例性病毒转化方案中，DNA（例如包含在此所述嵌合基因的 DNA）以及辅助病毒可通过单独用包含在此所描述的嵌合基因的 DNA 或用含有所述 DNA 的病毒体进行机械接种被引入整株植物中。最适合的转化方案的选择（包括转化方式和参数、转化计时等）完全属于本领域普通技术人员知识范围内。适合的方法描述或进一步详述于例如 W090/12107、W003/052108 或 W02005/098004。

[0044] 在另一个特定的实施方案中，本发明提供了基本由植物细胞组成的植物，所述植物细胞包含在此之前所描述的嵌合基因。个特定的实施方案中，该植物是转基因棉花植物。

[0045] 除了用于引入以上描述的嵌合基因的方法以外，所述嵌合基因还可渐渗（introgress）到植物中。“渐渗”指的是通过自然方法将基因整合到植物的基因组中，即通过将包含在此描述的嵌合基因的植物与不包含所述嵌合基因的植物进行杂交。可选择包含该嵌合基因的后代。

[0046] 植物可以是任何产生纤维的植物，如棉花。其他实例包括其他产生纤维的植物，如大麻、黄麻、亚麻以及木本植物，包括但不限于松属（*Pinus* spp.）、杨属（*Populus* spp.）、云杉属（*Picea* spp.）、桉树属（*Eucalyptus* spp.）等。

[0047] 在另一个特定的实施方案中，本发明提供了转基因种子，其包含如在此之前描述的嵌合基因。

[0048] 在又一个实施方案中，本发明提供了制备包含带正电多糖的植物细胞壁的方法，所述方法包括 1) 在植物或植物细胞中表达如本文之前所概述的嵌合基因，以及 2) 分离所述植物细胞壁。所述分离是从进行步骤 1) 之后获得的植物分离。

[0049] 在特定的实施方案中，所述制备（或产生，它是等效用语）方法在植物次生细胞壁中产生了带正电多糖。在另一个特定的实施方案中，所述制备方法在植物次生细胞壁中产生了几丁质。在另一个特定的实施方案中，所述制备方法在植物次生细胞壁中产生了壳聚糖，这是因为例如如下文进一步描述的包含几丁质脱乙酰酶的另外的嵌合基因的存在和表达。在特定的实施方案中，所述植物细胞壁或所述植物次生细胞壁是棉花植物细胞壁或棉花次生植物的细胞壁。在特定的实施方案中，所述棉花植物细胞壁或所述棉花次生细胞壁存在于棉纤维中。

[0050] 本发明进一步提供了包括使用根据本发明的方法从植物细胞获得的细胞壁的植物细胞壁，以及包含所述细胞壁的纤维。此类植物细胞壁包含带正电多糖，如 N-乙酰葡萄糖胺聚合物或几丁质，这些带正电多糖被嵌入到纤维素中。这些植物细胞壁可经进一步修饰，例如部分或完全脱乙酰化，这样就获得了包含葡萄糖胺残基的寡聚物。所得葡萄糖胺的氨基基团在化学上的反应性强于 N-乙酰葡萄糖胺的氨基乙酰基或纤维素的羟基。在特定的实施方案中，包含几丁质的植物细胞壁、尤其是植物次生细胞壁，可通过化学脱乙酰作用的步骤来进一步修饰。在另一个特定的实施方案中，在植物细胞壁中制备带正电多糖的方法可通过

以下方法进行：在植物或植物细胞中表达两种嵌合基因，一种嵌合基因是本发明的几丁质合酶 加上嵌合的几丁质脱乙酰酶。在酶学中，几丁质脱乙酰酶是催化以下化学反应的酶 (EC3. 5. 1. 41)：

[0051]



[0052] 因此，这种酶的两个底物是几丁质和 H_2O ，而它的两个产物是壳聚糖和乙酸盐。这种酶属于水解酶家族，它们作用于除肽键以外的碳氮键、尤其是线性氨基化物中的碳氮键。这种酶类别的系统名称是几丁质酰胺水解酶。

[0053] 在特定的实施方案中，根据本发明获得的植物细胞壁、尤其是已进行了脱乙酰作用的步骤（例如按照以上描述在体内，或在收集细胞壁以后）的那些植物细胞壁，可进一步进行化学修饰。含有此类植物细胞壁的产物（例如纤维、纱线或织物）其质量类似于本领域中所描述的纤维素-壳聚糖掺合物的质量，包括改进的可染性、改进的对例如皮肤真菌的抑制作用、受控药物释放等。

[0054] 在特定的实施方案中，本发明提供了棉纤维，其根据本发明的方法获自或可以获自棉花植物。换言之，棉纤维是由棉花植物提供，所述棉花植物的细胞基因组（例如核基因组）中包含含有被有效地连接至编码根据本发明的几丁质合酶基因 1 或 2 的 DNA 区的植物可表达的启动子的嵌合基因，或包含编码根据本发明的几丁质合酶基因 1 和 2 的两个嵌合基因的组合。将嵌合基因中所包含的 DNA 编码区或启动子转移到棉花植物中的特定实施方案在本文件的其他地方有所描述。

[0055] 根据本发明的棉纤维与天然存在的棉纤维（即从不包含根据本发明的嵌合基因的同基因系获得的棉纤维）可通过以下各项进行区别：这种纤维具有增强的用阴离子染料（包括例如刚果红）染色 的能力、这种纤维具有增强的用胺活性染料（包括例如四氟苯基酯）染色的能力。根据本发明的棉纤维还具有结合小麦胚芽凝集素的能力，而小麦胚芽凝集素结合几丁质。根据本发明的棉纤维与天然存在的棉纤维还可通过以下方式进行区别：任选地在用几丁质酶处理纤维细胞壁材料后，直接检测 N-乙酰葡萄糖胺和 GlcNAc 聚合物。根据本发明的棉纤维还可通过其增加的氮含量进行区别。

[0056] 根据本发明的棉纤维还可与现有技术中壳聚糖涂布的纤维相互区别，因为带正电聚合物均匀分布于构成纤维的次生植物细胞壁中，这与现有技术中表面涂布着壳聚糖的纤维相反。因此，在例如如以下所述用 WGA 或用刚果红或用四氟苯基染色的棉纤维的显微镜切片中，染料将均匀分布在构成棉纤维的整个细胞壁中，而在经壳聚糖涂布的纤维中，染色将集中在以片层形式位于经处理的纤维表面上的壳聚糖涂层中。

[0057] 阴离子染料（例如刚果红）对根据本发明的植物细胞壁材料的增强的染色可通过以下方式定量，例如在标准条件下将统一量的材料染色，在标准化区域（例如多孔板的孔中）内铺展该材料，针对材料着色层的灰度标度使该区域的照片数字化。灰色越浅，染色的植物细胞壁材料越多。以此方式，与来自等基因植物系但无几丁质合酶编码基因的对照细胞壁材料或纤维相比，根据本发明的棉纤维和细胞壁材料显示刚果红的染色增加了至少约 5%。在特定的实施方案中，根据本发明获得的植物细胞材料可用商业染料进行染色，商业染料包括棉花活性染料（例如活性红 120 (Reactive Red120)、丽华实蓝 (Levafix Blue CA)、酸性染料（酸性橙 7 (Acid Orange7)、酸性蓝 281 (Acid Blue281)）以及羊毛活性染料

(例如活性红 116、丽雅伦琥珀 (Realan Amber EHF)。在一个实例中,与不包含几丁质合酶编码基因的所述纤维和材料相比,根据本发明的棉纤维和细胞壁材料显示染色增加了至少约 5%、至少约 10%、至少 20%、至少 30%、至少 40%、至少 50%、至少 60%、至少 70%、至少 80%、至少 90%或甚至至少 100%,如 2 倍、5 倍或 10 倍。

[0058] 新型棉纤维特异性地结合小麦胚芽凝集素的能力(可由偶联的荧光基团检测)是所提供的新型棉纤维与天然存在的棉纤维之间的明显的区别特征。除极低的背景荧光以外,当用 WGA-alexa fluor488 或 555 处理时,天然存在的棉纤维不染色/发荧光。当几丁质聚合物存在时,棉纤维的荧光提高了至少 5 倍。因此,本发明提供了能够特异性地结合小麦胚芽凝集素或与荧光基团偶联的 WGA(例如 WGA Alexa488 或 WGA Alexa555)的棉纤维,或当用 WGA Alexa488 或 WGA Alexa555 处理时在 UV 光下提供亮荧光的棉纤维。这种荧光不局限于棉纤维的表面,而是分布在纤维细胞的整个细胞壁中。

[0059] 根据本发明的棉纤维与仅用壳聚糖涂布的纤维还可以通过应用 W02010/015423 中公开的检测方法和检查纤维中本发明的嵌合基因的存在来进行区分。

[0060] 虽然本发明的方法涉及将嵌合基因引入植物细胞中,但应当清楚的是还可将这种方法应用于将植物细胞掺入到成熟植物中的情况。例如,根据所建立的方法,可在转基因植物中再生转基因细胞。

[0061] 对植物细胞和植物进行转化的方法在本领域中是熟知的。对棉花植物进行转化的方法在本领域中也是熟知的。例如美国专利 5,004,863 中或美国专利 6,483,013 中已经描述了农杆菌属介导的棉花的转化,并且例如 W092/15675 中报导了通过粒子轰击进行棉花转化。

[0062] 可以通过在源自胚愈伤组织的棉花植物中进行转化而引入嵌合基因,所述棉花植物例如 Coker312、Coker310、Coker5AcalaSJ-5、GSC25110、FiberMax819、Siokra1-3、T25、GSA75、Acala SJ2、Acala SJ4、Acala SJ5、Acala SJ-C1、Acala B1644、Acala B1654-26、Acala B1654-43、Acala B3991、Acala GC356、Acala GC510、Acala GAM1、Acala C1、Acala Royale、Acala Maxxa、Acala Prema、Acala B638、Acala B1810、Acala B2724、Acala B4894、Acala B5002、非 Acala “采摘者”Siokra(non Acala “picker” Siokra)、“摘取机”品种 FC2017、Coker315、STONEVILLE506、STONEVILLE825、DP50、DP61、DP90、DP77、DES119、McN235、HBX87、HBX191、HBX107、FC3027、CHEMBRED A1、CHEMBRED A2、CHEMBRED A3、CHEMBRED A4、CHEMBRED B1、CHEMBRED B2、CHEMBRED B3、CHEMBRED C1、CHEMBRED C2、CHEMBRED C3、CHEMBRED C4、PAYMASTER145、HS26、HS46、SICALA、PIMA S6 以及 ORO BLANCO PIMA、**Fibermx**® FM5013、FM5015、FM5017、FM989、FM832、FM966 以及 FM958、FM989、FM958、FM832、FM991、FM819、FM800、FM960、FM966、FM981、FM5035、FM5044、FM5045、FM5013、FM5015、FM5017 或 FM5024、以及具有源自它们的基因型的植物。

[0063] 如在此所使用的,“棉花”包括陆地棉 (*Gossypium hirsutum*)、海岛棉 (*Gossypium barbadense*)、亚洲棉 (*Gossypium arboreum*) 以及草本棉 (*Gossypium herbaceum*) 或此类物种与其他物种杂交的子代或此类物种之间杂交的子代。

[0064] 本发明的方法和工具还可用于其他植物种类,如大麻、黄麻、亚麻以及木本植物,包括但不限于松属、白杨属、云杉属、桉树属等。

[0065] 所获得的经转化的植物可用于常规育种方案中,以产生具有相同特征的更多的经

转化的植物,或在相同或相关植物种类的其他品种中或在杂种植物中引入根据本发明的嵌合基因。从经转化的植物获得的种子含有作为稳定的基因组插入片段的本发明的嵌合基因,且为本发明所涵盖。

[0066] 本发明的嵌合基因与编码赋予了此类植物有用的农艺学性质的蛋白质或 RNA 的其他基因有利地在植物中组合。在编码赋予了经转化植物有用的农艺学性质的蛋白质或 RNA 的基因中,可提及这样的 DNA 序列,所述 DNA 序列编码赋予了对一种或多种除草剂的耐受性的蛋白质以及赋予对某些昆虫的耐受性的其他蛋白质或赋予了对某些疾病的耐受性的那些蛋白质。

[0067] 此类基因描述于例如公开的 PCT 专利申请 W091/02071 和 W095/06128 中。

[0068] 在编码赋予了经转化的植物细胞和植物对某些除草剂的耐受性的蛋白的 DNA 序列中,可提及 bar 或 PAT 基因或天蓝色链霉菌 (*Streptomyces coelicolor*) 基因(该基因描述于 W02009/152359,它赋予了对除草剂草铵膦的耐受性)、编码合适的 EPSPS 的基因(该基因赋予了对以 EPSPS 作为靶标的除草剂的耐受性,例如草甘膦及其盐)(US4, 535, 060、US4, 769, 061、US5, 094, 945、US4, 940, 835、US5, 188, 642、US4, 971, 908、US5, 145, 783、US5, 310, 667、US5, 312, 910、US5, 627, 061、US5, 633, 435、US6, 566, 587 以及 W097/04103)、编码草甘膦氧化还原酶的基因(US5, 463, 175)、或编码对 HPPD 抑制剂除草剂(如异噁唑类)的耐受性的 HPPD 酶的基因(W096/38567)。

[0069] 如在此使用的,“包含”应当被解释为指定正如所提及的经声明的特征、整数、步骤或组分的存在,但是并不排除存在或添加一个或多个特征、整数、步骤或组分、或它们的组群。因此,例如包含核苷酸序列或氨基酸序列的核酸或蛋白质可以包含比实际引用的序列更多的核苷酸或氨基酸,即被包含在更大的核酸或蛋白质中。包含功能或结构已定义的 DNA 区的嵌合基因可以包含额外的 DNA 区等。

[0070] 以下非限制性实例描述了用于改变植物细胞壁的方法。除非实例中另外指出,否则所有重组 DNA 技术将根据以下文献中描述的标准方案进行: Sambrook 等人(1989) 分子克隆:实验室手册第二版,冷泉港出版社,纽约,以及 Ausubel 等人(1994) 现代分子生物学方案,现代方案(Current Protocols), USA 的第 1 卷和第 2 卷(A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, NY and in Volumes 1 and 2 of Ausubel et al. (1994) Current Protocols in Molecular Biology, Current Protocols, USA)。用于植物分子工作的标准材料和方法描述于由 R. D. D. Croy 撰写的植物分子生物学(Plant Molecular Biology Labfax(1993)) 实验指导(1993) 中,该书由 BIOS Scientific Publications Ltd(UK) 以及 Blackwell Scientific Publications, UK 联合出版。

[0071] 通过在此描述的方法获得的经转化的植物细胞和植物可进一步用于本领域所熟知的育种方法中,例如杂交、自交、以及回交。育种程序可涉及杂交产生 F1 代(子一代),之后进行几代自交(产生 F2、F3 等)。育种程序还可涉及回交(BC) 步骤,其中将后代与称为回归亲本的亲本品系之一回交。

[0072] 由在此所公开的方法获得的转基因植物细胞和植物还可进一步用于随后的转化步骤中,例如用来引入另外的嵌合基因。

[0073] 包含在此公开的嵌合基因的植物可还用以下各项处理:棉花除草剂,如敌草隆(Diuron)、伏草隆(Fluometuron)、MSMA、氟硝草醚(Oxyfluorfen)、扑草净(Prometryn)、

氟乐灵 (Trifluralin)、氟酮唑草 (Carfentrazone)、烯草酮 (Clethodim)、精吡氟禾草灵 (Fluazifop-butyl)、草甘膦 (Glyphosate)、达草灭 (Norflurazon)、胺硝草 (Pendimethalin)、嘧硫苯甲酸 (Pyriithiobac-sodium)、三氟啶磺隆 (Trifloxysulfuron)、得杀草 (Tepaloxymid)、草铵膦 (Glufosinate)、氟噁唑酮 (Flumioxazin)、噻二唑素 (Thidiazuron) ; 棉花杀昆虫剂, 如高灭磷 (Acephate)、涕灭威 (Aldicarb)、毒死蜱 (Chlorpyrifos)、氯氰菊酯 (Cypermethrin)、溴氰菊酯 (Deltamethrin)、阿维菌素 (Abamectin)、吡虫清 (Acetamiprid)、因灭汀 (Emamectin Benzoate)、吡虫啉 (Imidacloprid)、噁二唑虫 (Indoxacarb)、 λ -赛洛宁 (Lambda-Cyhalothrin)、赐诺杀 (Spinosad)、硫敌克 (Thiodicarb)、 γ -氯氟氰菊酯 (Gamma-Cyhalothrin)、螺甲螨酯 (Spiromesifen)、啉虫丙醚 (Pyridalyl)、氟尼胺 (Flonicamid)、氟虫酰胺 (Flubendiamide)、杀虫隆 (Triflumuron)、氯虫酰胺 (Rynaxypyr)、 β -氟氯氰菊酯 (Beta-Cyfluthrin)、螺虫乙酯 (Spirotetramat)、可尼丁 (Clothianidin)、噻虫嗪 (Thiamethoxam)、噻虫啉 (Thiacloprid)、呋虫胺 (Dinotofuran)、氟虫酰胺 (Flubendiamide)、氰虫酰胺 (Cyazypyr)、艾克敌 (Spinosad)、赐诺特 (Spinotoram)、 γ -氯氟氰菊酯 (gamma Cyhalothrin)、4-[[(6-氯吡啶-3-基) 甲基] (2,2-二氟乙基) 氨基] 呋喃-2(5H)-酮、硫双威 (Thiodicarb)、齐墩螨素 (Avermectin)、氟尼胺 (Flonicamid)、啉虫丙醚 (Pyridalyl)、螺甲螨酯 (Spiromesifen)、虱虫啉 (Sulfoxaflor) ; 以及棉花杀真菌剂, 如腈嘧菌酯 (Azoxystrobin)、双噁芬 (Bixafen)、白克列 (Boscalid)、多菌灵 (Carbendazim)、百菌清 (Chlorothalonil)、铜 (Copper)、环唑醇 (Cyproconazole)、恶醚唑 (Difenoconazole)、醚菌胺 (Dimoxystrobin)、氧唑菌 (Epoconazole)、咪唑菌酮 (Fenamidone)、氟啉胺 (Fluazinam)、氟吡菌酰胺 (Fluopyram)、氟嘧菌酯 (Fluoxastrobin)、氟唑菌酰胺 (Fluxapyroxad)、异丙定 (Iprodione)、吡唑萘菌胺 (Isopyrazam)、异噻菌胺 (Isotianil)、代森锰锌 (Mancozeb)、代森锰 (Maneb)、叉氨苯酰胺 (Metominostrobin)、吡噻菌胺 (Penthiopyrad)、啉氧菌酯 (Picoxystrobin)、甲基代森锌 (Propineb)、丙硫菌唑 (Prothioconazole)、百克敏 (Pyraclostrobin)、五氯硝基苯 (Quintozene)、戊唑醇 (Tebuconazole)、氟醚唑 (Tetraconazole)、甲基托布津 (Thiophanate-methyl)、布洛芬 (Trifloxystrobin)。

[0074] 在整个说明书以及实施例中, 对序列表中所示出的以下序列进行引用 :

[0075] SEQ ID No :1 :同丝水霉几丁质合酶 2 基因的核苷酸序列

[0076] SEQ ID No :2 :同丝水霉几丁质合酶 2 基因的氨基酸序列

[0077] SEQ ID No :3 :同丝水霉几丁质合酶 1 基因的核苷酸序列

[0078] SEQ ID No :4 :同丝水霉几丁质合酶 1 基因的氨基酸序列

[0079] SEQ ID NO :5-17 是合成引物, 其序列描述在表 1 中。

[0080] SEQ ID NO :18 :同丝水霉几丁质合酶 2 的经修饰的核苷酸序列 (密码子被优化用于在棉花中表达)

[0081] SEQ ID NO:19 :同丝水霉几丁质合酶 2 的 MIT 结构域的氨基酸序列

[0082] SEQ ID NO:20 :同丝水霉几丁质合酶 1 的 MIT 结构域的氨基酸序列

[0083] SEQ ID NO :21 :寄生水霉几丁质合酶基因的核苷酸序列

[0084] SEQ ID NO :22 :由寄生水霉几丁质合酶基因 SEQ ID NO :21 编码的氨基酸序列

[0085] SEQ ID NO :23 :寄生水霉几丁质合酶基因的核苷酸序列

[0086] SEQ ID NO :24 :由寄生水霉几丁质合酶基因 SEQ ID NO :23 编码的氨基酸序列

实施例

[0087] 1. 来自同丝水霉的几丁质合酶基因的克隆

[0088] 卵菌纲同丝水霉 Pringsheim53-967Dick 获自“Centraal Bureau voor Schimmel Culturen”(CBS, Baarn, The Netherlands), 并且在 90 毫米培养皿内的马铃薯葡萄糖琼脂上培养。实验中所用的菌丝体在黑暗中以 25℃ 下在玛驰 (Machlis) 液体培养基 (Machlis L. (1953) Am. J. Bot. 40, 449-460) 中生长 3-5 天。

[0089] 使用 RNeasy 植物迷你试剂盒 (RNeasy Plant Mini Kit) (Qiagen), 并联合 DNaseI 的柱上消化, 从 100mg3 天龄的同丝水霉菌丝体中提取总 RNA。按照制造商的说明书, 使用 Superscript III 第一链 cDNA 合成试剂盒 (Superscript III First Strand cDNA Synthesis kit) (Invitrogen), 用 4.5 μg 总 RNA 进行逆转录。将 1 μl 合成的 cDNA 用于 PCR 反应, 使用引物 CHS2Fwd 和 CHS2Rev 对 CHS2 序列进行扩增, 所述引物是根据以登录号 U19946 储存的序列 (该序列还公开于 Mort-Bontemps M. 等人 (1997) Microbiology 143, 2009-2020 中, 其中几丁质合酶 2 (CHS2) 序列描述于所述参考文献的图 3 中) 而设计的。使用 RNA 连接酶介导的 RACE 试剂盒 (RLM-RACE, Ambion) 来分离并且测定 CHS2 的 3' 端序列。用引物 CHS2Fwd1 和 CHS2Fwd2 获得 CHS2 的 3' 端。使用引物 CHS2Fwd 和 CHS2FLRev, 从同丝水霉菌丝体 cDNA 扩增完全的全长 CHS2 编码序列 (描述于 SEQ ID NO :1 中)。根据制造商的说明书, 使用 Gateway 技术将这种所得到的序列克隆到 pENTR-D-TOPO (Invitrogen) 中。将构建体转化入 One Shot TOP10 化学感受态细胞 (Invitrogen) 中, 并且测定核苷酸序列 (MWG, Germany)。观察到, 所公开的同丝水霉 CHS2 的编码序列 (登录号 U19946) 与 SEQ ID NO :1 中描述的分离的 CHS2 序列仅具有 96% 的同一性。核苷酸序列的差异是因为以登录号 U19946 公开的几丁质合酶 2 的编码序列中存在许多小缺失。因此, 描述于 SEQ ID NO :1 中的编码序列的翻译序列与登录号 U19946 的经翻译的编码序列仅具有 79% 的同一性。后者是由 U19946 的核苷酸序列中的小缺失引起的, 这导致当 U19946 的编码序列翻译时, 相对于 SEQ ID NO :1 的翻译发生移码。显然, SEQ ID NO :2 包含保守的微管相互作用和运输分子结构域 (MIT 结构域), 而从保藏号 U19946 翻译的几丁质合酶蛋白质中并无该结构域。

[0090] 使用软件 BioEdit (Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis - <http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) 编辑这些序列。使用 ClustalW (Larkin MA 等人 (2007) Bioinformatics 23, 2947-2948) 进行比对。使用 SmCHS1 和 SmCHS2 全长序列对来自美国国家生物技术中心 (NCBI, <http://www.ncbi.nlm.nih.gov>) 的非冗余蛋白质数据库进行 BLASTp 搜索。

[0091] 表 1 描述了引物的序列:

[0092]

CHS2Fwd, SEQ ID NO :5	CACCATGAGTGACCAGCTCGACCTCGCGGC
CHS2Rev, SEQ ID NO :6	TGCTCTCTGCACGGGCAACCACAACCCGAC
CHS1Fwd, SEQ ID NO :7	AATGAGGACGAGAACGAGCTCCGGTCG
CHS1Rev, SEQ ID NO :8	AGCTTGTAAGGACGACTTGTTGGC

[0093]

CHS1Rev1, SEQ ID NO :9	TTCCTTGGTCATTTGCAGCGAGTGTTTC
CHS1Rev2, SEQ ID NO :10	CAGAACGTTGTTGCAAACCTTGCGGAGT
CHSFwd1, SEQ ID NO :11	TCCGGTCGACACTCCGCAAGGTTTGCAA
CHS1Fwd2, SEQ ID NO :12	ACTACACGGTCCTCCTCGATGTTGGGAC
CHS2Fwd1, SEQ ID NO :13	TGTCGGTGGCTTGATTGTCTTTGC
CHS2Fwd2, SEQ ID NO :14	TTTGGCTCTACGTTGTGACGGACT
CHS1FLFwd, SEQ ID NO :15	CACCATGCCGCCCAAGCGACCGACGCCGA
CHS1FLRev, SEQ ID NO :16	CTAGCGCATGCGGTTGTACGGCGCTTGG
CHS2FLRev, SEQ ID NO :17	TTAGACTTGTTGGTAGGCGCCGCCGCGG

[0094] 两种引物 (CHS1Fwd 和 CHS1Rev) 是根据同丝水霉几丁质合酶 1 (CHS1) 的部分序列 (对应于登录号 U42304) 而设计的,用于 CHS1 的保守区的扩增。使用 RNA 连接酶介导的 RACE 试剂盒 (RLM-RACE, Ambion) 扩增 CHS1 的 5' 端和 3' 端。使用以下巢式反向引物组来获得 CHS1 的 5' 端:CHS1Rev1 和 CHS1Rev2。使用巢式正向引物 CHSFwd1 和 CHS1Fwd2 分离 3' 端。使用用于扩增 CHS1 基因的编码序列的引物 CHS1FLFwd 和 CHS1FLRev,从同丝水霉菌丝体 cDNA 扩增全长 CHS1,以证实完全序列。根据制造商的说明书,使用 Gateway 技术,将全长基因克隆到 pENTR-D-TOPO (Invitrogen) 中。将构建体转化至 One Shot TOP10 化学感受态细胞 (Invitrogen) 中,并且通过测序确定核苷酸序列 (MWG, Germany)。所有 PCR 反应都是根据制造商的说明书使用 Phusion High-Fidelity DNA 聚合酶 (Finnzymes) 进行。几丁质合酶 1 核苷酸序列的编码区描述于 SEQ ID NO :3 中。几丁质合酶 1 的氨基酸序列描述于 SEQ ID NO :4 中。

[0095] 2. 编码同丝水霉几丁质合酶 2 (CHS2) 蛋白的嵌合的植物可表达的基因的构建

[0096] 使用标准的重组 DNA 技术,构建了植物可表达的同丝水霉 CHS2 嵌合基因,其含有以下有效地连接的 DNA 片段:

- [0097] • 来自 CaMV 的 35S 启动子区
- [0098] • 编码未翻译的前导序列 (5' Cab22L) 的 DNA 片段
- [0099] • 编码同丝水霉的 CHS2 (描述于 SEQ ID NO :1 中) 的 DNA 片段
- [0100] • 来自章鱼碱合酶基因的转录终止以及多聚腺苷酸化信号

[0101] 简言之,根据制造商的说明书,使用定向 TOPO 克隆试剂盒 (Invitrogen),通过将 CHS2 基因 PCR 克隆到 Gateway 进入载体 pENTR/D/TOPO 中来产生该构建体。使用 Phusion DNA 聚合酶 (Finnzymes) 进行 PCR 反应。使用 Gateway LR Clonase Enzyme Mix (Invitrogen),将得到的进入载体重组到目标载体 pEarleyGate103 中 (Earley KW 等人 (2006) Plant J. 45 :616-629)。所有构建体均通过序列分析来核实序列。用于植物转化的选择标记是提供对草胺磷的抗性的 bar 基因。使用包含 SmCHS2 基因的重组双元载体 pEarleyGate103 对农杆菌属菌株 C58C1 进行转化 (Voinnet O. 等人 (2003) Plant J. 33 : 949-956),所述菌株 C58C1 用于通过花浸染法对拟南芥进行转化 (Clough SJ 和 Bent AF (1998) Plant J. 16 :735-743)。与未转化的拟南芥植物相比,包含嵌合的同丝水霉几丁质合酶 2 的经转化的拟南芥植物在生长方面没有差异。这与这样的转基因拟南芥植物生长减慢形成对比,所述植物包含与高尔基定位信号有效地连接的粗糙脉孢菌几丁质合酶 (如 W02006/136351 中所公开的,参见实施例 9)。

[0102] 3. 重组拟南芥的毛状体的组织化学分析

[0103] 拟南芥中的毛状体 (植物毛发) 是具有特征性三维结构的大型非分泌表皮细胞,。

因为毛状体很容易地被遗传学、细胞生物学以及分子方法组合使用,所以我们使用源自实施例 2 中产生的重组拟南芥植物的成熟毛状体用于对经过 SmCHS2 转化的拟南芥植物的植物细胞壁进行分析。在最近的研究过程中,我们观察到毛根培养物的细胞壁分析(参见 W02006/136351,实例 2)导致了与成熟毛状体的细胞壁分析类似的结果(成熟毛状体更容易建立,并且比毛根培养物更容易分离)。

[0104] 将毛状体进行组织化学染色使细胞的不同化合物可见,并且用显微镜进行分析。

[0105] 例如,在与针对 N-乙酰葡萄糖胺的 IgM 单克隆抗体(BIODESIGN)进行免疫反应之后,可检测到 N-乙酰葡萄糖胺。几丁质可以通过例如使用小麦胚芽凝集素-Alexa Fluor555 或备选地通过使用罗丹明缀合的结合几丁质的探针进行检测。后者是特异性地结合几丁质的重组融合蛋白。融合蛋白包含小的(5kDa)几丁质结合结构域(CBD),其源自与来自大肠杆菌的麦芽糖结合蛋白质(43kDa)的 C 末端融合的环状芽孢杆菌 WL-12 的几丁质酶 A1 的 C 末端区。该融合蛋白按照标准方法使用四甲基罗丹明异硫氰酸酯(TRITC)标记。我们从 New England BioLabs 购买了罗丹明缀合的几丁质结合探针(目录号 P5210)。

[0106] 通过荧光显微术,使用配备有 Apotome(Zeiss)用于允许光学切片的 Axioplan2 显微镜(Zeiss, Jena, Germany)来检查经组织化学染色的毛状体。使用 Axio4.2 版(Zeiss)进行图像处理。

[0107] 植物细胞壁的 Calcofluor 染色可用以下方案进行。Calcofluor White(或荧光增白剂 28)是无色有机化合物,它在紫外光照射($\lambda_{\max} = 350\text{nm}$)下发出清楚的淡蓝色荧光。将待染色的样品(即毛状体)在包含最终浓度为 $50 \mu\text{g/mL}$ 的荧光增白剂 28 的培养基或 PBS(缓冲溶液)中浸渍 15 到 30 分钟。然后用培养基或缓冲液洗涤样品,并且使用配备用于荧光显微术且使用 Zeiss 滤镜组 18 的显微镜检查样品。细胞壁发出清楚的淡蓝色荧光。对包含 SmCHS2 嵌合基因的重组成熟毛状体进行免疫组织化学染色以检查 N-乙酰葡萄糖胺的存在,并且随后用 Calcofluor 染色以使细胞壁可见。正如由光学切片的重叠可观察到,仅在重组毛状体的细胞壁中检测到 N-乙酰葡萄糖胺的存在。

[0108] 小麦胚芽凝集素-Alexa Fluor555染色或用罗丹明缀合的几丁质结合探针染色的方法:

[0109] 在约 3 周后,从体外产生的转基因芥菜属植物(即根据实施例 2 的包含 SmCHS2 的重组植物)以及从对照芥菜属植物(即非转化植物)中分离包含毛状体的成熟叶子。首先将叶子在固定剂(含 10%福尔马林、0.25%戊二醛的 PBS)中固定 1 小时。真空浸润样品,并且重新更换固定剂并与叶子再温育 3 小时。最后,固定后可对叶子进一步处理或存储在 4°C 下若干天。下一步骤,用 PBS 洗涤固定过的叶子(60 分钟),之后用 PBT(即 PBS+0.1% Tween20)洗涤 60 分钟。随后将叶子与小麦胚芽凝集素-Alexa Fluor555(Alexa fluor555TFP 酯是可从 Molecular Probes 公司获得的试剂盒)在 PBT(3mg/ml)中温育 4 至 6 小时。或者备选地,将叶子与罗丹明-几丁质结合探针一起温育 16 小时。温育后,洗去荧光染料或罗丹明-几丁质结合探针(用 PBT 洗涤两次并且用 PBS 洗涤一次)。

[0110] 使用荧光显微术例如用 Zeiss 滤镜 38,检查叶子边缘处的经染色的重组毛状体。观察到来自重组毛状体的细胞壁材料可再现的染色比来自对照植物的细胞壁材料更强烈(参见图 1)。

[0111] 4. 几丁质合酶在棉花中的纤维特异性表达

[0112] 使用如在 US6483013 中所述的转化方法已经产生了包含如实施例 2 中所概述、在 F286 纤维选择性启动子（在 US2003/1060971 中公开）控制下的嵌合几丁质合酶 2 基因的转基因棉花植物，以及包含两个嵌合基因（即，在 F286 启动子控制下的所述嵌合几丁质合酶 2 基因以及在所述 F286 启动子控制下的谷氨酰胺：果糖 -6- 磷酸酰胺转移酶（gfa）基因，或者在 E6 启动子控制下的所述嵌合几丁质合酶 2 基因以及在 E6 启动子控制下的 gfa 基因）的转基因植物。将来自这些转基因棉花植物的纤维分离，并且用于产生具有改进的反应性（例如改进的可染性）的纱线和织物。从转基因植物的棉铃分离的纤维具有增加量的均匀分布于细胞壁中的 N- 乙酰葡萄糖胺聚合物。

[0113] 5. 反应性提高的棉纤维

[0114] 正如实施例 4 中所述，已经产生包含有效地连接到纤维特异性启动子的嵌合 SmCHS2 编码区的转基因棉花植物。从这些植物收集成熟的棉纤维，并且可将其用刚果红进行染色，或可与 WGA-Alexa fluor555 进行反应。此外，可用商业染料对得到的成熟棉纤维进行染色，所述商业染料包括棉活性染料（例如活性红 120、丽华实蓝 CA）、酸性染料（酸性橙 7、酸性蓝 281）、以及羊毛活性染料（例如活性红 116、丽雅伦琥珀 EHF）。

[0115] WGA-Alexa555染色

[0116] 棉纤维不需要脱水或经透化处理。相反，通过将纤维在氯仿：甲醇混合物（1：1）中处理 3 次 10 分钟，之后在丙酮中处理 2 次 10 分钟，并且在乙醚中处理 2 次 5 分钟来去除脂质和蜡。允许将纤维风干。

[0117] 可用 WGA-Alexa555、WGA-Alexa488 或 WGA- 四甲基罗丹明对纤维进行染色。

[0118] 将纤维放入封闭溶液（pH7.4 的 150mM NaCl、10mM 磷酸钠缓冲液、0.1% Tween20、以及 1% 牛血清白蛋白）中，并且温育 1 小时。此后，用包含 WGA- 荧光染料的相同缓冲液替换该缓冲液，并且温育 4 小时。用封闭溶液替换 WGA- 荧光染料溶液，洗涤 10 分钟，之后用无 BSA 的封闭溶液洗涤 3 次 10 分钟，并且用无 BSA 并且无 Tween 的封闭溶液洗涤 2 次 5 分钟。将染色的纤维在显微镜载玻片上封固，并且通过荧光显微术（Axioplan2（Zeiss, Jena, Germany））使用滤镜组 38（激发：BP470/40；发射：BP525/50）评估 Alexa fluor488 缀合物或使用滤镜组 20（激发：BP546/12；发射：BP575-640）评估 Alexa fluor555 或四甲基罗丹明缀合物。虽然在来自非转基因棉花植物的棉纤维中没有检测到特异性荧光，但是在包含嵌合 SmCHS2 基因的转基因棉花植物的棉纤维中可检测到亮荧光。棉纤维的实际显微切片显示，WGA-fluor555 均匀分布于棉纤维细胞的次生细胞壁中。

[0001]

序列表

<110> Bayer Bioscience N.V.
Bulone, Vincent
Guerriero, Gea

<120> 制备包含几丁质的植物细胞壁的方法

<130> bes10-2001

<150> EP10000712.9

<151> 2010-01-25

<150> US61/297,973

<151> 2010-01-25

<160> 24

<170> PatentIn version 3.5

<210> 1

<211> 2889

<212> DNA

<213> 同丝水霉(*Saprolegnia monoica*)

<400> 1

atgtctgacc agctcgacct cgcggcacgc ctccgcgcc tgcgtgaggg caacgccgcg	60
ccagccgacc ccgaggcacc gccaccacac cagcctgcgc ccgcgccgca gtaccatcca	120
cagcggctgc cgcgctgta tacacaagag tcgctcgagt teggaggagc gtatgccacg	180
ggcagccccg tgggcgccga ggccgagggg tcctactcgc aagtacctgt gtggaaggac	240
tccaaggaga cgcgaagaag ttacctggac gatgagccga cgcgcagcc ccaatcgtc	300
attaacatgg ctaacacctt ggtccagcgc caagcgtcga atcaatcctt ccggcggcag	360
catacagcga actlccgacc gctgcccaac accgtcgagg agcttctgga cggcacgcca	420
acgtacgaag gagcgtttcg cctcgtgcag ctccgggtgc agatggagca agacggcgac	480
ccaggtgctg caattaactt gtacgtggac gctggcacga cgctcgtcga agtcggttaag	540

[0002]

cgtgaggtcg accctcttct gcagaagggc atccagcaaa aggcctttga gctcctgcaa	600
cgtgccgagg agctcggcac gtggatgaac acggtggccg aggaagcgcg gaaagctgcg	660
ttgccaccgc agctcaagat cgcgcgcacc aacgtcccca cggtcgagca agcgtggaag	720
ggtcgcactc cgcctttcca tgacgccgac gagttccgac ttatgcggtg cacggccgtg	780
gcaacgaaag acccgatcca gttctcgaac gacggatacg tgctccgggt gcaccagcta	840
catcggcgca tcaaggtctt catcacaatc actatgtaca acgaagaagg ctacagagatc	900
ttgggcacgc tcaactggtc cgccaagggc ctcgggtata tgtgcaagga gtatggccag	960
gatttttggc aagaagttgc tgtggtatc gtctcagacg gccgcaccaa agctagcaag	1020
acgtgtctcg agtacctcaa cggcctcggc gcatttgacg aagagatcat gacggtcacg	1080
agcctcgggtg tcgacgtgca gatgcacctc ttcgagtcca cactccagct ggttigagaac	1140
cagacgtttg aaaactactt tccgccgctc caagtgatct acgcgctcaa agagaacaac	1200
ggtggcaagc tcaactcgca tctgtggttc ttcaacgctt tcagcgagca attgaacccc	1260
aagtacactg tactcgtgga cgtcggcacc attcccgcg aaacgtccgt tttccgcttg	1320
atccgcagca tggagcgcaa ctaccagatc ggccggcgtt cgggggagat tgcagtcgaa	1380
gcacctaaact acttcaatcc tgcattgcc gcacagcact tcgagtacaa gatctcaaat	1440
atcatggaca agtcgcttga gtcgctctt ggttttatct cggtgctccc gggggccttc	1500
tcagctacc ggtacgaagc catccgtget gtcaagggtg tggggccgct gccagagtac	1560
tttaagagcc tcacgtcgac gaccaaagag ctcgggccat ttcagggcaa tatgtacctc	1620
gccgaagatc gtattttatg ttttgaattg ctggcgcgca aacataaaca gtggacaatg	1680
cactatgtca aggacggat cggccgcact gacgttcccg agacgctcgt agacctgac	1740
aagcagcgcc ggcggtgget caacgggtct ttctttgccg gcctctttgc catcgggcac	1800
tttggcgcg tctggagcca gagtccccc accatgtccc ggaagcttgt gttcacgttt	1860

[0003]

cagttctttt accttgcctt acagaacctg ctcagttggt tctctcttgag caatttggtc	1920
cttacattct actttgtttt gacgctcgcc ttacggact cagcaccggc ccttctccaa	1980
gcgatgctga cgctgtatct ggccattgtc ggtggcttga ttgtctttgc gctcggaac	2040
aagcccgaac ctgcgacggc cagcttttac ctcttcagct gcctttacat gggcatcctc	2100
atgatgcttg tgaccggcat ttccatctac ggccttgctg gcaagggcac aagcgctgtg	2160
aaagaccac ggggtgatcac gggggccctt ggcaactgta ctgtctctga aggggagctt	2220
gtcgggtggtg tcgtcaccic gcttggettg atcttctct ctgccttctt ccatggcgag	2280
tttagcctcc tctcagcgt catccagtac ttcttcctgc tcccaacgtt tgtgaatgtg	2340
ctgggcatct atgcctacag caatttgac gacttgagct ggggtaccaa gggcctggag	2400
tctgggggctg gccacggacc aacgaaaacc ggcgggcgta acgtcaagga cgtcgtcgag	2460
cagcagaaga agctcgaagc ccaacgccag gcccgggcta aggagaaaga agatgtggac	2520
aacagcttcc gggcggttccg gtcgacactg ctgctgtcgt ggctcacgac caacggcatt	2580
tggtctctacg ttgtgacgga ctacatgtcg agcgggtgct acttgaaggc cctcagcttc	2640
gtcgtcggtt tctttaacgt catccgtttc acgggctgct tcgtatttat tatcttgcgc	2700
atctttcgcc gattcggcct caactgctgt gcaatggggg ccacccacga tacctatgag	2760
cgcaacctgc cgcccgactg gcagactcat tacaatgtac agaaccaagc cgacggctcg	2820
gttgtggtcg cccgtgcaga gagcataaac ccagcaacgc ccccgggcgg cgcctaccaa	2880
caagtctag	2889

<210> 2

<211> 962

<212> PRT

<213> 同丝水霉

<400> 2

[0004]

Met Ser Asp Gln Leu Asp Leu Ala Ala Arg Leu Arg Ala Leu Arg Glu
1 5 10 15

Gly Asn Ala Ala Pro Ala Asp Pro Glu Ala Pro Pro Pro Thr Gln Pro
20 25 30

Ala Pro Ala Pro Gln Tyr His Pro Gln Arg Leu Pro Pro Leu Tyr Thr
35 40 45

Gln Glu Ser Leu Glu Phe Gly Gly Thr Tyr Ala Thr Gly Ser Pro Val
50 55 60

Gly Ala Glu Ala Glu Gly Ser Tyr Ser Gln Val Pro Val Trp Lys Asp
65 70 75 80

Ser Lys Glu Thr Arg Arg Ser Tyr Leu Asp Asp Glu Pro Thr Pro Gln
85 90 95

Pro Gln Ser Leu Ile Asn Met Ala Asn Thr Leu Val Gln Arg Gln Ala
100 105 110

Ser Asn Gln Ser Phe Arg Arg Gln His Thr Ala Asn Phe Arg Pro Leu
115 120 125

Pro Asn Thr Val Glu Glu Leu Leu Asp Gly Thr Pro Thr Tyr Glu Gly
130 135 140

Ala Phe Arg Leu Val Gln Leu Ala Val Gln Met Glu Gln Asp Gly Asp
145 150 155 160

Pro Gly Ala Ala Ile Asn Leu Tyr Val Asp Ala Gly Thr Thr Leu Val
165 170 175

[0005]

Glu Val Gly Lys Arg Glu Val Asp Pro Leu Leu Gln Lys Gly Ile Gln
180 185 190

Gln Lys Ala Phe Glu Leu Leu Gln Arg Ala Glu Glu Leu Gly Thr Trp
195 200 205

Met Asn Thr Val Ala Glu Glu Ala Arg Lys Ala Ala Leu Pro Pro Gln
210 215 220

Leu Lys Ile Ala Arg Thr Asn Val Pro Thr Val Glu Gln Ala Trp Lys
225 230 235 240

Gly Arg Thr Pro Pro Phe His Asp Ala Asp Glu Phe Arg Leu Met Arg
245 250 255

Tyr Thr Ala Val Ala Thr Lys Asp Pro Ile Gln Phe Ser Asn Asp Gly
260 265 270

Tyr Val Leu Arg Val His Gln Leu His Arg Arg Ile Lys Val Phe Ile
275 280 285

Thr Ile Thr Met Tyr Asn Glu Glu Gly Ser Glu Ile Leu Gly Thr Leu
290 295 300

Thr Gly Leu Ala Lys Gly Leu Gly Tyr Met Cys Lys Glu Tyr Gly Gln
305 310 315 320

Asp Phe Trp Gln Glu Val Ala Val Ala Ile Val Ser Asp Gly Arg Thr
325 330 335

Lys Ala Ser Lys Thr Cys Leu Glu Tyr Leu Asn Gly Leu Gly Ala Phe
340 345 350

[0006]

Asp Glu Glu Ile Met Thr Val Thr Ser Leu Gly Val Asp Val Gln Met
355 360 365

His Leu Phe Glu Ser Thr Leu Gln Leu Val Glu Asn Gln Thr Phe Glu
370 375 380

Asn Tyr Phe Pro Pro Leu Gln Val Ile Tyr Ala Leu Lys Glu Asn Asn
385 390 395 400

Gly Gly Lys Leu Asn Ser His Leu Trp Phe Phe Asn Ala Phe Ser Glu
405 410 415

Gln Leu Asn Pro Lys Tyr Thr Val Leu Val Asp Val Gly Thr Ile Pro
420 425 430

Ala Glu Thr Ser Val Phe Arg Leu Ile Arg Ser Met Glu Arg Asn Tyr
435 440 445

Gln Ile Gly Gly Val Ala Gly Glu Ile Ala Val Glu Ala Pro Asn Tyr
450 455 460

Phe Asn Pro Val Ile Ala Ala Gln His Phe Glu Tyr Lys Ile Ser Asn
465 470 475 480

Ile Met Asp Lys Ser Leu Glu Ser Val Phe Gly Phe Ile Ser Val Leu
485 490 495

Pro Gly Ala Phe Ser Ala Tyr Arg Tyr Glu Ala Ile Arg Ala Val Lys
500 505 510

Gly Val Gly Pro Leu Pro Glu Tyr Phe Lys Ser Leu Thr Ser Thr Thr
515 520 525

[0007]

Lys Glu Leu Gly Pro Phe Gln Gly Asn Met Tyr Leu Ala Glu Asp Arg
530 535 540

Ile Leu Cys Phe Glu Leu Leu Ala Arg Lys His Lys Gln Trp Thr Met
545 550 555 560

His Tyr Val Lys Asp Ala Ile Ala Arg Thr Asp Val Pro Glu Thr Leu
565 570 575

Val Asp Leu Ile Lys Gln Arg Arg Arg Trp Leu Asn Gly Ser Phe Phe
580 585 590

Ala Gly Leu Phe Ala Ile Gly His Phe Gly Arg Val Trp Ser Gln Ser
595 600 605

Ser His Thr Met Ser Arg Lys Leu Val Phe Thr Phe Gln Phe Phe Tyr
610 615 620

Leu Ala Leu Gln Asn Leu Leu Ser Trp Phe Leu Leu Ser Asn Leu Phe
625 630 635 640

Leu Thr Phe Tyr Phe Val Leu Thr Leu Ala Phe Thr Asp Ser Ala Pro
645 650 655

Ala Leu Leu Gln Ala Met Leu Thr Leu Tyr Leu Ala Ile Val Gly Gly
660 665 670

Leu Ile Val Phe Ala Leu Gly Asn Lys Pro Glu Pro Arg Thr Ala Ser
675 680 685

Phe Tyr Leu Phe Ser Cys Leu Tyr Met Gly Ile Ile Met Met Leu Val
690 695 700

[0008]

Thr Gly Ile Ser Ile Tyr Gly Leu Val Gly Lys Gly Thr Ser Ala Val
705 710 715 720

Lys Asp Pro Arg Val Ile Thr Gly Ala Leu Gly Asn Cys Thr Val Ser
725 730 735

Glu Gly Glu Leu Val Gly Gly Val Val Thr Ser Leu Gly Leu Ile Phe
740 745 750

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

Gln Tyr Phe Phe Met Leu Pro Thr Phe Val Asn Val Leu Gly Ile Tyr
770 775 780

Ala Tyr Ser Asn Leu His Asp Leu Ser Trp Gly Thr Lys Gly Leu Glu
785 790 795 800

Ser Gly Gly Gly His Gly Pro Thr Lys Thr Gly Gly Gly Asn Val Lys
805 810 815

Asp Val Val Glu Gln Gln Lys Lys Leu Glu Ala Gln Arg Gln Ala Ala
820 825 830

Ala Lys Glu Lys Glu Asp Val Asp Asn Ser Phe Arg Ala Phe Arg Ser
835 840 845

Thr Leu Leu Leu Ser Trp Leu Thr Thr Asn Gly Ile Trp Leu Tyr Val
850 855 860

Val Thr Asp Tyr Met Ser Ser Gly Cys Tyr Leu Lys Gly Leu Ser Phe
865 870 875 880

[0009]

Val Val Gly Phe Phe Asn Val Ile Arg Phe Thr Gly Cys Val Val Phe
885 890 895

Ile Ile Leu Arg Ile Phe Arg Arg Phe Gly Leu Asn Cys Cys Ala Met
900 905 910

Gly Ala Thr His Asp Thr Tyr Glu Arg Asn Leu Pro Pro Asp Trp Gln
915 920 925

Thr His Tyr Asn Val Gln Asn Gln Ala Asp Gly Arg Val Val Val Ala
930 935 940

Arg Ala Glu Ser Ile Asn Pro Ala Thr Pro Arg Gly Gly Ala Tyr Gln
945 950 955 960

Gln Val

<210> 3

<211> 2733

<212> DNA

<213> 同丝水霉

<400> 3

atgccgccca agcgaccgac gaccgacggc cgccgcgcgt acaatgcagg caatacgact 60

gttcgcgcac ccacaaagcg cacacagcgg cgcggaacaa teggcagcgg ggcateccat 120

aacccgagcg cggcgagcat gcaagcatac gaatttgaat acgactacaa cagcgcaatg 180

atgccaatgt tgcagccacc caagagccaa cctacgttcc tcaacaatat tgcgcccatc 240

tcgtccaaag aagcgagtat gaagagctca aatgcaatgc agctcttget ccagagcact 300

tccttcacta ttgacgatgc ctccgtgcc atcgacgcg caattcaagc tgagaacgaa 360

ggccggtacc gcgaagccct caagcatttc ttggacggtg gcgagatgat tgtgactgcc 420

[0010]

gccgagaagg aagcgtcgca aaaggtgcgc aacttgcttt tgcacaaagg caaggaagtg	480
ttggagtggg ccgagcacct cgccgagtgg attgagcgct ataatacgca ctccgcacca	540
gttcgcgttg ctaagcccat ggctgtcgaa gtcacgtacg accgcacgat gaattctccg	600
gatttagacg aaacggaagc gcgcacaatg ttttacacgc cagtgtgctg caccgccaa	660
gcttttactg aaacggggta ccggctccaa tgcattccagt cgggccggcg cccgcggctt	720
atggtcgtca ttaccatgta caacgaagac gagaacgagc tccggtcgac actccgcaag	780
gtttgcaaca acgttctgta cctcaagcag cagagcctgc ctggctacga aggtgacgac	840
gcgtggaagc aggtgctcgt ggttatcgtc teggacgggc gcaccaaagc caataaaggc	900
acgcttgagt ggctttcgaa cgtggcctg tatgacgaag atgtcatgaa cattacgtcg	960
acgggtgtga aagtgcaatg ccattctattt gaacactcgc tgcaaatgac caaggagaac	1020
tccgacccgt ttccgcgcgt tcaagtgacc tttgactaa aagagcaca cgcgggaaa	1080
cttgactcgc atctttggta ttccgacgca tttgctgaac aagtgtgccc cgactacacg	1140
gtcctctctg atgttgggac gatgcccaacc aagtcgtcct tttacaagct cctcacggcc	1200
ctcgagatca accgcagat tggcgggtgc tgcggcgaga tcgtgtcga taagccgctg	1260
ccaaatatgt gcaactgggt cattgccgcg cagcacttcg agtacaagat cagcaacatt	1320
ctggacaagt cactcgagtc gtgctttggc tttatatctg tgctacctgg tgcattctcc	1380
gcctatcgtc acaaagcgat ccgtgggtgc cgtttgcagg cctactttaa gagtcttacg	1440
acgccgatgg cggagctcgg gccatttgcg ggcaatatgt acctggctga agaccgaatc	1500
ttgtgctttg agctgctcgc gcgcaaggat tgcaactgga cgatgcacta tgtcaaggac	1560
gcgatagcac gcaccgacgt gccacaatat ttgattgatt tggtcgggca gcgcggcg	1620
tggtcaatg gcctctctt tgcgacactc tttgcgattt ggaactgggg tcgctgtac	1680
acggagagca accactcctt cagcgcaag atggcactgc tcgtgcagta cgtctataac	1740

[0011]

```

gtgctccaag ttatcttttc gtggttttta cccgccaact tttacctgc tctgtatttt 1800
gtcatctttc aaggctttta agacaatcgg tggaacttca tcgatacgtc caagtacccg 1860
gcgctgcttt tagacggcct cccgacggcg tttaacgtgt tttacgccgt caccgtgttt 1920
accaagtca cgattggcct cgggaacaag cccaagcatg taaaaggcac gcactacctc 1980
atctcggtgc tctttggtat cctcatgctc attgcgtcca ccattgctat agtcatcttc 2040
gtcacggcgc acaagaccgt cgaagccatc atcctcgctg tettaatcct cggcacgttc 2100
ttcattggtt cggccatgca ctgcgaagtg catcatattg tctcaegt tgtgcagtat 2160
acggcgctga tgccgagctt cgtcaatata ctcattggtc actcgttctg caaccttcac 2220
gacctgagct ggggcaccaa aggcattcac acgggtcatg agcacaagag tgacggcgcc 2280
gtcgggcagt acaaggacat cgtcgcgcgg caaaaagcgt tggaagccaa gaaggcggaa 2340
gacgccagga accaggacga gctcaagaaa cgattcgact cgttcgggtc gaatttgctc 2400
ttgatttggg ttatgtccaa tatggccatg gtggtcattt gcgtcaatac aatcggcgcc 2460
gacagctacc tcccgtttct ctatgcgttt gttgccgcgt tcaacggcat ccggtcctc 2520
ggctgcatcg ggtacctctt gtactacgcg cgccagtitt tgctctttaa tacgtgagt 2580
gccactgggg tgctacacaa gcgccatgaa gcgcgcaagc acaagaaagc cgaagaagac 2640
ccggacccga tcgactttga aatgggcact ttccagaacg accttcctga cgtcgcggtg 2700
ccgatccaag cgcggtacaa ccgcatgcgc tag 2733

```

<210> 4

<211> 910

<212> PRT

<213> 同丝水霉

<400> 4

```

Met Pro Pro Lys Arg Pro Thr Thr Asp Gly Arg Arg Ala Tyr Asn Ala
1           5           10           15

```

[0012]

Gly Asn Thr Thr Val Arg Ala Pro Thr Lys Arg Thr Gln Pro Arg Gly
20 25 30

Lys Ile Gly Ser Arg Ala Ser Asn Asn Pro Ser Ala Ala Ser Met Gln
35 40 45

Ala Tyr Glu Phe Glu Tyr Asp Tyr Asn Ser Ala Met Met Pro Met Leu
50 55 60

Gln Pro Pro Lys Ser Gln Pro Thr Phe Leu Asn Asn Ile Ala Pro Ile
65 70 75 80

Ser Ser Lys Glu Ala Ser Met Lys Ser Ser Asn Ala Met Gln Leu Leu
85 90 95

Leu Gln Ser Thr Ser Phe Thr Ile Asp Asp Ala Phe Arg Ala Ile Glu
100 105 110

Arg Ala Ile Gln Ala Glu Asn Glu Gly Arg Tyr Arg Glu Ala Leu Lys
115 120 125

His Phe Leu Asp Gly Gly Glu Met Ile Val Thr Ala Ala Glu Lys Glu
130 135 140

Ala Ser Gln Lys Val Arg Asn Leu Leu Leu His Lys Gly Lys Glu Val
145 150 155 160

Leu Glu Trp Ala Glu His Leu Ala Glu Trp Ile Glu Arg Tyr Asn Thr
165 170 175

His Ser Ala Pro Val Arg Val Ala Lys Pro Met Ala Val Glu Val Thr
180 185 190

[0013]

Tyr Asp Arg Thr Met Asn Ser Pro Asp Leu Asp Glu Thr Glu Ala Arg
195 200 205

Thr Met Phe Tyr Thr Pro Val Cys Cys Thr Pro Gln Ala Phe Thr Glu
210 215 220

Thr Gly Tyr Arg Leu Gln Cys Ile Gln Ser Gly Arg Arg Pro Arg Leu
225 230 235 240

Met Val Val Ile Thr Met Tyr Asn Glu Asp Glu Asn Glu Leu Arg Ser
245 250 255

Thr Leu Arg Lys Val Cys Asn Asn Val Leu Tyr Leu Lys Gln Gln Ser
260 265 270

Leu Pro Gly Tyr Glu Gly Asp Asp Ala Trp Lys Gln Val Leu Val Val
275 280 285

Ile Val Ser Asp Gly Arg Thr Lys Ala Asn Lys Gly Thr Leu Glu Trp
290 295 300

Leu Ser Asn Val Gly Leu Tyr Asp Glu Asp Val Met Asn Ile Thr Ser
305 310 315 320

Thr Gly Val Lys Val Gln Cys His Leu Phe Glu His Ser Leu Gln Met
325 330 335

Thr Lys Glu Asn Ser Ile Arg Phe Pro Pro Leu Gln Val Thr Phe Ala
340 345 350

Leu Lys Glu His Asn Ala Gly Lys Leu Asp Ser His Leu Trp Tyr Phe
355 360 365

[0014]

Asp Ala Phe Ala Glu Gln Val Met Pro Asp Tyr Thr Val Leu Leu Asp
370 375 380

Val Gly Thr Met Pro Thr Lys Ser Ser Phe Tyr Lys Leu Leu Thr Ala
385 390 395 400

Leu Glu Ile Asn Ala Gln Ile Gly Gly Val Cys Gly Glu Ile Ala Val
405 410 415

Asp Lys Pro Leu Pro Asn Met Cys Asn Trp Val Ile Ala Ala Gln His
420 425 430

Phe Glu Tyr Lys Ile Ser Asn Ile Leu Asp Lys Ser Leu Glu Ser Cys
435 440 445

Phe Gly Phe Ile Ser Val Leu Pro Gly Ala Phe Ser Ala Tyr Arg Tyr
450 455 460

Lys Ala Ile Arg Gly Ala Pro Leu Gln Ala Tyr Phe Lys Ser Leu Thr
465 470 475 480

Thr Pro Met Ala Glu Leu Gly Pro Phe Ala Gly Asn Met Tyr Leu Ala
485 490 495

Glu Asp Arg Ile Leu Cys Phe Glu Leu Leu Ala Arg Lys Asp Cys Asn
500 505 510

Trp Thr Met His Tyr Val Lys Asp Ala Ile Ala Arg Thr Asp Val Pro
515 520 525

Thr Asn Leu Ile Asp Leu Val Gly Gln Arg Arg Arg Trp Leu Asn Gly
530 535 540

[0015]

Ser Phe Phe Ala Thr Leu Phe Ala Ile Trp Asn Trp Gly Arg Val Tyr
545 550 555 560

Thr Glu Ser Asn His Ser Phe Thr Arg Lys Met Ala Leu Leu Val Gln
565 570 575

Tyr Val Tyr Asn Val Leu Gln Val Ile Phe Ser Trp Phe Leu Pro Ala
580 585 590

Asn Phe Tyr Leu Ala Leu Tyr Phe Val Ile Phe Gln Gly Phe Lys Asp
595 600 605

Asn Arg Trp Asn Phe Ile Asp Thr Ser Lys Tyr Pro Ala Leu Leu Leu
610 615 620

Asp Gly Leu Pro Thr Ala Phe Asn Val Phe Tyr Ala Val Thr Val Phe
625 630 635 640

Thr Gln Val Thr Ile Gly Leu Gly Asn Lys Pro Lys His Val Lys Gly
645 650 655

Thr His Tyr Leu Ile Ser Val Leu Phe Gly Ile Leu Met Leu Ile Ala
660 665 670

Ser Thr Ile Ala Ile Val Ile Phe Val Thr Ala His Lys Thr Val Glu
675 680 685

Ala Ile Ile Leu Ala Val Leu Ile Leu Gly Thr Phe Phe Ile Gly Ser
690 695 700

Ala Met His Cys Glu Val His His Ile Val Leu Thr Phe Val Gln Tyr
705 710 715 720

[0016]

Thr Ala Leu Met Pro Ser Phe Val Asn Ile Leu Met Val Tyr Ser Phe
725 730 735

Cys Asn Leu His Asp Leu Ser Trp Gly Thr Lys Gly Ile Asp Thr Gly
740 745 750

His Glu His Lys Ser Asp Gly Ala Val Gly Gln Tyr Lys Asp Ile Val
755 760 765

Ala Arg Gln Lys Ala Leu Glu Ala Lys Lys Ala Glu Asp Ala Arg Asn
770 775 780

Gln Asp Glu Leu Lys Lys Arg Phe Asp Ser Phe Arg Ser Asn Leu Leu
785 790 795 800

Leu Ile Trp Val Met Ser Asn Met Ala Met Val Val Ile Cys Val Asn
805 810 815

Thr Ile Gly Ala Asp Ser Tyr Leu Pro Phe Leu Tyr Ala Phe Val Ala
820 825 830

Ala Phe Asn Gly Ile Arg Leu Leu Gly Cys Ile Gly Tyr Leu Leu Tyr
835 840 845

Tyr Ala Arg Gln Phe Leu Leu Phe Asn Thr Leu Ser Ala Thr Gly Val
850 855 860

Leu His Lys Arg His Glu Ala Arg Lys His Lys Lys Ala Glu Glu Asp
865 870 875 880

Pro Asp Pro Ile Asp Phe Glu Met Gly Thr Phe Gln Asn Asp Leu Pro
885 890 895

[0017]

Asp Val Ala Val Pro Ile Gln Ala Pro Tyr Asn Arg Met Arg
900 905 910

<210> 5

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 引物 CHS2Fwd

<400> 5

caccatgagt gaccagctcg acctcgcggc

30

<210> 6

<211> 30

<212> DNA

<213> 人工序列

<220>

<223> 引物 CHS2Rev

<400> 6

tgctctctgc acgggcaacc acaacccgac

30

<210> 7

<211> 27

<212> DNA

<213> 人工序列

<220>

<223> 引物 CHS1Fwd

<400> 7

aatgaggacg agaacgagct ccggtcg

27

<210> 8

<211> 27

[0018]

<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS1Rev	
<400>	8	
	agcttgtaaa aggacgactt ggttggc	27
<210>	9	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS1Rev1	
<400>	9	
	tctccttggt catttgcagc gagtgttc	28
<210>	10	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS1Rev2	
<400>	10	
	cagaacgttg ttgcaaacct tgcggagt	28
<210>	11	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHSFwd1	
<400>	11	
	tccggtcgac actccgcaag gtttgcaa	28

[0019]

<210>	12	
<211>	28	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS1Fwd2	
<400>	12	
	actacacggt cctcctcgat gttgggac	28
<210>	13	
<211>	24	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS2Fwd1	
<400>	13	
	tgtcggtgge ttgattgtct ttgc	24
<210>	14	
<211>	24	
<212>	DNA	
<213>	人工序列	
<220>		
<223>	引物 CHS2Fwd2	
<400>	14	
	tttggctcta cgttgtgacg gact	24
<210>	15	
<211>	30	
<212>	DNA	
<213>	人工序列	
<220>		

[0020]

<223> 引物 CHS1FLFwd

<400> 15

caccatgccg cccaagcgac cgacgaccga

30

<210> 16

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> 引物 CHS1FLRev

<400> 16

ctagcgcatg cggttgtagc ggccttgg

28

<210> 17

<211> 28

<212> DNA

<213> 人工序列

<220>

<223> CHS2FLRev

<400> 17

ttagacttgt tggtaggcgc cgccgcgg

28

<210> 18

<211> 2889

<212> DNA

<213> 人工序列

<220>

<223> 同丝水霉几丁质合成酶 2 的经修饰的核苷酸序列（密码子被优化用于在棉花中表达）

<400> 18

atgagtgatc aacttgacct cgcagcaagg ttacgagcat taagagaggg taatgcagct

60

cctgtgatc ctgaagcacc tccacctaca caacctgctc cagcacccca ataccatcca

120

cagagattac caccctctta tacacaagag agtctggagt ttggtgggac ttaigctact

180

[0021]

ggatctcctg teggagctga agctgaagga tcctatagtc aagtccegggt gtggaaggat	240
tctaaagaaa ctctgcgctc ttacttggat gatgagccaa ctccgcaacc acagtcactc	300
attaacatgg caaacacctt ggtccagcgt caagccctcta atcaatcctt caggcgtcag	360
catactgcta atttccgacc tcigccaaac actgtcgaag agcttttga cggaacacca	420
acctacgaag gtgcttttcg attggttcaa cttgtctgtc agatggagca agatggcgac	480
ccagggtcag caattaacct ttacgtcgat gctggtacaa ctctcgttga agtcggtaaa	540
cgtgaggtag accctctact tcagaagggt atccagcaga aggcctttga acttctgcaa	600
cgtgctgaag agtiggggac ttggatgaat actgttgccg aagaagcaag gaaagctgct	660
ttgccacctc aactcaagat tgctagaaca aacgtccga ctgtcgaaca agcatggaaa	720
ggaagaactc ctcttttcca tgacgtgac gagtttcgct tgatgaggtg tactgtgtg	780
gtaccaaag atccgateca gttctetaac gacggatacg tgcttcgagt tcaccaacta	840
catcggagga tcaaggtctt catcacaatc actatgtaca acgaagaagg gtccgagatc	900
ttgggaactc ttactggact tgctaaggga ctcggtata tgtcgaagga gtatggccaa	960
gatttttggc aagaggtcgc tgttgcatac gtgtcagatg ggagaaccaa agctagtaag	1020
acatgtctgg agtacttgaa cggcctcgggt gcttttgacg aagagataat gaccgttact	1080
togettggig tcgacgtgca gatgcacttg ttgagagca cactccaact ggttgagaac	1140
cagacattcg aaaactactt tccgcccttg caagtgatct acgctctgaa ggagaacaat	1200
ggagggaac tgaatagcca tccttggttc ttcaacgctt ttccggagca attgaatccc	1260
aagtatactg tgctcgtgga cgttggaaac attcccgtg aaacctccgt ttttcggttg	1320
atccgaagca tggaaaggaa ctaccagatt ggaggggttg ctggagagat agcagtcgaa	1380
gtccccaact acttcaatcc tgtgattgcc gctcagcact tcgagtacaa galcagcaat	1440
atcatggaca agagccttga aagcgtgttc gggtttatat ccgttcttcc tggggctttt	1500

[0022]

tcggttattc gttacgaagc catacgtgct gtcaagggtg tcggaccttt gccagagtac	1560
ttcaagagcc tcacticaac caccaaagaa ctcggtccat tccaggggaa tatgtatctc	1620
gccgaagatc ggattctctg ttctgaactg ctcgcaagga aacataaaca gtggaccatg	1680
cactatgtga aggacgcaat cgctagaact gatgttcctg agacactcgt cgatctgac	1740
aagcaacgac gtaggtggtt gaacggatct ttctttgccg gactctttgc aatcggacac	1800
tttggaaagg tatggtccca gtcttcacac acaatgtcca gaaagctcgt cttcaccttc	1860
cagttctctt accttgccct tcagaatctc ctcaagttgt tcctcttgag caatctgttc	1920
ctgaccttct acttcgtctt taactctgct tttaccgatt cagctcctgc tctactccaa	1980
gctatgctta cctgtatctt ggccattgtt ggtggttga ttgtctttgc tctcggaac	2040
aagcccgaa ctagaacagc ctcatctac ctgttcagct gcctttacat ggggatcata	2100
atgatgctcg tgactggaat ttcatctac ggtcttgtt gcaaaggtag aagcgtgtg	2160
aaagatccaa ggtgataac tggagccctt gggaattgta ctgtctctga aggtgaactt	2220
glgggagggt tcgttacttc acctggttgc atcttctgt cagcattcgt tcatggcgag	2280
ttctccatcc tcttttcagt gatccagtag ttcttcatgc ttccgacctt tgtcaatgtg	2340
ctcgggatct atgcctattc caatctgcac gatttgagct ggggtacca gggactcgaa	2400
tctgggggtg gacatggctc aacaaaaact ggtgggggaa atgtgaaaga cgtcgtcgag	2460
cagcaaaaga aactcgaagc tcaaagacaa gccgcagcta aggagaaaga ggatgtggac	2520
aattcgttcc gcgtttccg atctactctg ttgttgagtt ggtcactac caatgggatt	2580
tggtcttacg ttgtgacaga ctacatgtca tctgggtgct acttgaaggg actcagtttc	2640
gtgtcggtct tcttcaacgt catcagattc actggctgcg tcgtctttat aatccttcga	2700
atcttttaggc ggtttggctt gaactgtgtt gcaatgggtg caacacacga tacctatgag	2760
cgtaacctac ctcccgattg gcaaacccat tacaatgtcc agaaccaagc tgacggtcgg	2820

[0023]

gtggtagtcg caagagccga gagtataaac ccagcaactc caagaggagg tgcttaccaa 2880

caagtctaa 2889

<210> 19

<211> 69

<212> PRT

<213> 人工序列

<220>

<223> 来自同丝水霉几丁质合成酶 2 基因的 MIT 结构域

<400> 19

Thr Tyr Glu Gly Ala Phe Arg Leu Val Gln Leu Ala Val Gln Met Glu
1 5 10 15

Gln Asp Gly Asp Pro Gly Ala Ala Ile Asn Leu Tyr Val Asp Ala Gly
20 25 30

Thr Thr Leu Val Glu Val Gly Lys Arg Glu Val Asp Pro Leu Leu Gln
35 40 45

Lys Gly Ile Gln Gln Lys Ala Phe Glu Leu Leu Gln Arg Ala Glu Glu
50 55 60

Leu Gly Thr Trp Met
65

<210> 20

<211> 69

<212> PRT

<213> 人工序列

<220>

<223> 来自同丝水霉几丁质合成酶 1 基因的结构 MIT 域

[0024]

<400> 20

Thr Ile Asp Asp Ala Phe Arg Ala Ile Glu Arg Ala Ile Gln Ala Glu
1 5 10 15

Asn Glu Gly Arg Tyr Arg Glu Ala Leu Lys His Phe Leu Asp Gly Gly
20 25 30

Glu Met Ile Val Thr Ala Ala Glu Lys Glu Ala Ser Gln Lys Val Arg
35 40 45

Asn Leu Leu Leu His Lys Gly Lys Glu Val Leu Glu Trp Ala Glu His
50 55 60

Leu Ala Glu Trp Ile
65

<210> 21

<211> 2676

<212> DNA

<213> 寄生水霉(Saprolegnia parasitica)

<400> 21

atgccaccca agcgcccgac cgaggccagt ggccgccgct acgcgcccgc cgccggtcgt 60
ccgtccaaca acgcccgaat tgccaagccg cgcgcgccgc gaaagggcgt gagcagccgt 120
gcgtccaacg tcccatccgc cgctcgtcg tacgagtag actacgagta caacatgatg 180
ccgatgatgc aggcgcgcgc caagtcgcag ccaaccttc tctccaacat tgcgcccatc 240
tcggcgaaag aggcgagcat gaagggtcgc aacgcgatgc aactcctgct gcagggcacg 300
tcgttcacga tcgacgatgc gtccgcgcgc atcgagcgtg cgatccaagc cgagaacgag 360
ggccgggttc gcgaggcgt caagcacttt ctggacggcg gcgagatgat tgtgacggcc 420
gccgagaagg aagcgtccca gaaggtccgc aacctcctcc tccataaagg caaggaggtg 480

[0025]

ctcgagtggg ccgagcacct cgccgagtgg atcgagcget acaacacgtc gacgccgccg	540
gtccgcacatg ccaagccgat ggccgtcgag gtcacgtacg accgcacgat gaactcgcca	600
gacctcgacg aaaccgagge gcgcacgatg ttttacacgc ccgtgtgctc gggcccgaag	660
gcgttcaccc agacggggta ccggctgcag tgcacccaga gcggccgccg ccgcggctc	720
atggtgggtca tcaccatgta caacgaagac gagaacgagc tccggtcgac gctccgcaag	780
gtctgcaaca acgtcctgta cctcaagcag cacagcctcc cggggtacga aggcgacgac	840
gcgtggaagc aggtgctggt cgtcgtcgtc tcggatggcc gcacgaaagc caacaaaggc	900
acgtctgagt ggetcgcacg cgtcggcctc tacgacgaag acgttatgaa catcacgtcg	960
accggcgctca aggtgcagtg ccacctcttc gagcattcgc ttcagatgac gaaagagaac	1020
tcgatccggt tcccgccact ccaactcgac tcgcacctct gglactttga cgcgtttgcc	1080
gagcagatca tgcccacta caccgtgctc ctgcacgtcg gcacgatgcc gaccaagtcg	1140
tcgtttctaca agctgctcac ggcgctcgag atcaacgcac agatcgggcg cgtctcgggc	1200
gagattgcgg tcgacaagcc gctgcccaac atgtgcaact gggtcacgc ggccgacat	1260
ttcgaataca agatttccaa catcctcgac aagtgcgtcg agtcgtgctt cggcttcac	1320
tcggtgctgc ccggcgctt ctcgccctac cgctacaagg cgattcgcg cgcgccactg	1380
caagcgtaact ttaagagtct cacaaccgac atggccgagc tcggaccgtt tgctggcaac	1440
atgtacctcg ccgaagatcg catcctgtgc tttgagctcc tcgcgcgcaa agactgcaac	1500
tggaccatgc actacgtcaa ggacgcgac gcgcgcacag acgtgccgac caacttgatc	1560
gacctcgtgg gccagcgccg gcgttggtc aacgggtcgt ttttcgcac gttgtttgcg	1620
atctggaact ggggcgcgt gtacaccgag tcgaaccact cgtcacgcg caagctcgcg	1680
ctcctcgtgc atgctcttct cggcgtctca gggcccaact tttacctgc ctgtacttt	1740
gtcatcttcc aaggcttccg cgacaaccgg tggaaattca tcgacacgtc cgagtaccgc	1800

[0026]

```

cagtgggtcc tcgacggcct cccgacggcg tttaacgtct ttacgcggt cacggtcttt 1860
acccaagtca cgatcggcct cggcaacaag cccaaacacg tcaaaggcac ccattacctc 1920
atctcgggtgc tctttggctt gctcatgttg ctgcctccg gtgtcgccat cgteatcttc 1980
atcacgtcct cgaaagacgc catggcgatc gtgctcgcag tgcctattct cggcacgttc 2040
ttcatlgggt cagcgctgca ctgcgaagtg caccacattg tgcctacgtt cgtccagtac 2100
acggcgcctca tgccgagctt tgtcaacatc ctcatggctt actcgttctg caacctccac 2160
gacctcagct ggggcacaaa gggcatcgac acggggccacg agggccataa aaccgaagcc 2220
gtcggccagt acaaggacat tgtcgcgcgc caaaaagctc tggaggccaa gaaagcccaa 2280
gacgcgcgca accaagacga gctcaagaag cgcttcgact cgttcggtc caacttgcta 2340
ctgggtgtggg tcatgtccaa catgtcgatg gtcctcatct gcgtcaacac ggtcggcgcc 2400
gactcgttct tgccgttctt gtacgcgttc gtcgccgcct tcaacggcat ccggtcctc 2460
ggctgcctcg ggtacctcat ctactacgcg cggcagttcc tctctttcaa tacgtgagc 2520
gccacgggcg tcttcacaaa gcgccacgaa gcgcgcaagc acaagaagc cgaggaccg 2580
gaccgatcg acatggaact cggcacgttc aacgaaccgg cgacgtccga gattggcgcg 2640
cccatgatgc aagcgccgta caaccgatg cgttag 2676

```

<210> 22

<211> 891

<212> PRT

<213> 寄生水霉

<400> 22

```

Met Pro Pro Lys Arg Pro Thr Glu Ala Ser Gly Arg Arg Tyr Ala Pro
1           5           10           15

```

```

Pro Ala Gly Arg Pro Ser Asn Asn Ala Ala Asn Ala Lys Pro Arg Ala
20           25           30

```

[0027]

Pro Arg Lys Gly Val Ser Ser Arg Ala Ser Asn Val Pro Ser Ala Ala
35 40 45

Ser Ser Tyr Glu Tyr Asp Tyr Glu Tyr Asn Met Met Pro Met Met Gln
50 55 60

Ala Pro Pro Lys Ser Gln Pro Thr Phe Leu Ser Asn Ile Ala Pro Ile
65 70 75 80

Ser Ala Lys Glu Ala Ser Met Lys Gly Ser Asn Ala Met Gln Leu Leu
85 90 95

Leu Gln Gly Thr Ser Phe Thr Ile Asp Asp Ala Phe Arg Ala Ile Glu
100 105 110

Arg Ala Ile Gln Ala Glu Asn Glu Gly Arg Phe Arg Glu Ala Leu Lys
115 120 125

His Phe Leu Asp Gly Gly Glu Met Ile Val Thr Ala Ala Glu Lys Glu
130 135 140

Ala Ser Gln Lys Val Arg Asn Leu Leu Leu His Lys Gly Lys Glu Val
145 150 155 160

Leu Glu Trp Ala Glu His Leu Ala Glu Trp Ile Glu Arg Tyr Asn Thr
165 170 175

Ser Thr Pro Pro Val Arg Ile Ala Lys Pro Met Ala Val Glu Val Thr
180 185 190

Tyr Asp Arg Thr Met Asn Ser Pro Asp Leu Asp Glu Thr Glu Ala Arg
195 200 205

[0028]

Met Met Phe Tyr Thr Pro Val Cys Ser Gly Pro Lys Ala Phe Thr Glu
210 215 220

Thr Gly Tyr Arg Leu Gln Cys Ile Gln Ser Gly Arg Arg Pro Arg Leu
225 230 235 240

Met Val Val Ile Thr Met Tyr Asn Glu Asp Glu Asn Glu Leu Arg Ser
245 250 255

Thr Leu Arg Lys Val Cys Asn Asn Val Leu Tyr Leu Lys Gln His Ser
260 265 270

Leu Pro Gly Tyr Glu Gly Asp Asp Ala Trp Lys Gln Val Leu Val Val
275 280 285

Val Val Ser Asp Gly Arg Thr Lys Ala Asn Lys Gly Thr Leu Glu Trp
290 295 300

Leu Ala Asn Val Gly Leu Tyr Asp Glu Asp Val Met Asn Ile Thr Ser
305 310 315 320

Thr Gly Val Lys Val Gln Cys His Leu Phe Glu His Ser Leu Gln Met
325 330 335

Thr Lys Glu Asn Ser Ile Arg Phe Pro Pro Leu Gln Leu Asp Ser His
340 345 350

Leu Trp Tyr Phe Asp Ala Phe Ala Glu Gln Ile Met Pro Asp Tyr Thr
355 360 365

Val Leu Leu Asp Val Gly Thr Met Pro Thr Lys Ser Ser Phe Tyr Lys
370 375 380

[0029]

Leu Leu Thr Ala Leu Glu Ile Asn Ala Gln Ile Gly Gly Val Cys Gly
385 390 395 400

Glu Ile Ala Val Asp Lys Pro Leu Pro Asn Met Cys Asn Trp Val Ile
405 410 415

Ala Ala Gln His Phe Glu Tyr Lys Ile Ser Asn Ile Leu Asp Lys Ser
420 425 430

Leu Glu Ser Cys Phe Gly Phe Ile Ser Val Leu Pro Gly Ala Phe Ser
435 440 445

Ala Tyr Arg Tyr Lys Ala Ile Arg Gly Ala Pro Leu Gln Ala Tyr Phe
450 455 460

Lys Ser Leu Thr Thr Asp Met Ala Glu Leu Gly Pro Phe Ala Gly Asn
465 470 475 480

Met Tyr Leu Ala Glu Asp Arg Ile Leu Cys Phe Glu Leu Leu Ala Arg
485 490 495

Lys Asp Cys Asn Trp Thr Met His Tyr Val Lys Asp Ala Ile Ala Arg
500 505 510

Thr Asp Val Pro Thr Asn Leu Ile Asp Leu Val Gly Gln Arg Arg Arg
515 520 525

Trp Leu Asn Gly Ser Phe Phe Ala Thr Leu Phe Ala Ile Trp Asn Trp
530 535 540

Gly Arg Val Tyr Thr Glu Ser Asn His Ser Leu Thr Arg Lys Leu Ala
545 550 555 560

[0030]

Leu Leu Val His Ala Leu Leu Gly Val Ser Ala Ala Asn Phe Tyr Leu
565 570 575

Ala Leu Tyr Phe Val Ile Phe Gln Gly Phe Arg Asp Asn Arg Trp Asn
580 585 590

Phe Ile Asp Thr Ser Glu Tyr Pro Gln Trp Val Leu Asp Gly Leu Pro
595 600 605

Thr Ala Phe Asn Val Phe Tyr Ala Val Thr Val Phe Thr Gln Val Thr
610 615 620

Ile Gly Leu Gly Asn Lys Pro Lys His Val Lys Gly Thr His Tyr Leu
625 630 635 640

Ile Ser Val Leu Phe Gly Leu Leu Met Leu Leu Ala Ser Gly Val Ala
645 650 655

Ile Val Ile Phe Ile Thr Ser Ser Lys Asp Ala Met Ala Ile Val Leu
660 665 670

Ala Val Leu Ile Leu Gly Thr Phe Phe Ile Gly Ser Ala Leu His Cys
675 680 685

Glu Val His His Ile Val Leu Thr Phe Val Gln Tyr Thr Ala Leu Met
690 695 700

Pro Ser Phe Val Asn Ile Leu Met Val Tyr Ser Phe Cys Asn Leu His
705 710 715 720

Asp Leu Ser Trp Gly Thr Lys Gly Ile Asp Thr Gly His Glu Ala His
725 730 735

[0031]

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

Ala Leu Glu Ala Lys Lys Ala Gln Asp Ala Arg Asn Gln Asp Glu Leu
755 760 765

Lys Lys Arg Phe Asp Ser Phe Arg Ser Asn Leu Leu Leu Val Trp Val
770 775 780

Met Ser Asn Met Ser Met Val Ile Ile Cys Val Asn Thr Val Gly Ala
785 790 795 800

Asp Ser Phe Leu Pro Phe Leu Tyr Ala Phe Val Ala Ala Phe Asn Gly
805 810 815

Ile Arg Leu Leu Gly Cys Ile Gly Tyr Leu Ile Tyr Tyr Ala Arg Gln
820 825 830

Phe Leu Leu Phe Asn Thr Leu Ser Ala Thr Gly Val Leu His Lys Arg
835 840 845

His Glu Ala Arg Lys His Lys Lys Ala Glu Asp Pro Asp Pro Ile Asp
850 855 860

Met Glu Leu Gly Thr Phe Asn Glu Pro Ala Thr Ser Glu Ile Gly Ala
865 870 875 880

Pro Met Met Gln Ala Pro Tyr Asn Arg Met Arg
885 890

<210> 23

<211> 2880

[0032]

<212> DNA

<213> 寄生水霉

<400> 23

```

atgagcgaca gcaacctoga cctcgccgcc cgcctccggg cgtgcgcga gggcggcgcc      60
gagcccgccc ctgcgcgggc accgacgccc tacatgcatt cgccaccatc gcggacccgg      120
cccacgccgc tgtacacgca agagtcgctc gagtttggcg ggacgtatac gactggtagc      180
cccgtgggcg ccgaggcgga cggcgtgtac acgcagggtc ctgtgtggaa ggactcgaag      240
gagaagacgt acggctactt ggacgacgag ccggcgccgc aggcccagac gctcctgaac      300
aaggccaatg acctcgtgca gcgccaagcc tccaacaagg cgttctggcg tcagcacaca      360
gccgcgtttc ggccgtgcc caacactgtc gaggagctcc tcgatggctc gccgacgtac      420
gaaggcgcgt tccggctcgt gcagctgcc gtgcagatgg agcaggatgg cgatccgcaa      480
gcccgcatca acttgtaacg agacgcgggc gcgacgcttg tcgaggtcgg gcgcaaagaa      540
gtcgaccgcg ttttgcaaaa ggcatccgc caaaaggccc aagagctcct ccagcgtgcc      600
gaagacctcg aggcgtggat gaacggcgtc gccgaggagg cgcgcaaggc cgctttgccg      660
ccgagcctcc gcctcgcgcg caccaacgtg ccgaccgtcg agcagacgtg ggccggccga      720
ccgccgccct tccatgacgc caacgaattc aaattgatgc ggtacacggc ggtggcgacc      780
aaggaccgga tccagttctc ggacgacggc tacgtgctcc gcgtgcacga gctgcaacga      840
cccatcaagg tctttataac gatcaccatg tacaacgaag aaggtcggga gatcaagggc      900
acgtctacgg gccttgccaa aggtctcgcg tacatgtgca aggagtacgg cgacgacttt      960
tggcagcaag tcgcggtggc catcgtctcg gacggccgga cgaaagcgtc caagacgtgt     1020
ctcgagtacc tcaaggccgt cggcgcgttt gacgaagaga tcattgacgtt cactcgtctc     1080
ggtgtcgacg tccagatgca tctcttcgag tcgacgtccc agctgggtcg gaaccagaac     1140
tttgaggcct attaccacc gctgcaagtg atttacgcgc tcaaggagaa caatggcggg     1200

```

[0033]

aagctcaact cgcctctctg gtctttcaac gcgttctcgg agcagctcaa cccaaagtac	1260
acggtgctcg tcgacgtcgg gacgattccc gccgagacct cggctcttcg cttgatccgg	1320
agcatggagc gcaacgcgca gattggcggg gtcgtcggcg agattgcagt tgaagccccg	1380
aattttctca accctgtgat tgccgcccag catttcgagt acaagatctc caacattatg	1440
gacaagtcgc tcgagtcggg gtttggtttt atctcgggtg tgcccgggtg gttctcagcg	1500
taccggtaag aagcgattcg ggccgtcaag ggcgctcggc cgtgcccga gtactttaaa	1560
agcctcaagt cgacgaccaa ggagctcggg ccgttccaag gcaacatgta cttggccgaa	1620
gaccgcatct tgtgtttga gctcctcgg cgcaagcagc ggcgctggac catgcactac	1680
gtcaaggacg cgattgcggc caccgacgtg cccgaaacgc tcgtcgacct catcaagcag	1740
cgccggcgct ggctcaacgg gtcgttcttc gcgggcctct ttgcgatcgg gcactttggc	1800
cgcgctctga gccaaagcag ccactcggtt ggctcgcaagc tcgtgtttac gttccagttt	1860
gtctacctcg cgtccagaa cctcctctcg tggtttctgt tgagcaacct cttcttgacg	1920
ttttactttg tcttgacgt ggccctttacc gagtcggcgc cggcgctgct ccagacgatg	1980
ctgaccgtct acctcggat cattggcggc ctcatcgctt ttgcccttgg caacaagccc	2040
gagccgcgca eggccagctt ttacctcttt agctgcctct acatgggcat catcatgctg	2100
ctcgtgacgg gcatttccat ctacggcttg attggcaagg gtacgagcgc cgtcaaggac	2160
ccgcgcacga tcaccgggat ctctctcaac tgcaaggta gcgacgcga gctcgcgggc	2220
ggcgctatca cgtcgtcgg gctcatcttc ctctcggcct tcgtccacgg cgagtttggc	2280
atcctctca gctttgtgca gtatttcttc atgtccccga cgtttgtcaa cgtgctgggc	2340
atctatgcgt acagcaacct tcattgatctg agctggggca cgaaaggcct cgagtcgggc	2400
ggcggccacg ggcccgccaa ggccggcggc ggcaacgtca aggacgtcgt cgagcagcaa	2460
aagaagatag aggccgcgcg gcaagccgcc gccagagaga aggaagatgt cgacaacagc	2520

[0034]

ttccgcgcgt tccggtcgac gttgctgctc tcgtggctca cgaccaacgg catttggtg 2580
tacgtcgtca cggactacat gtcgagcggg tgctacctca agggcctcag ctacattgtc 2640
ggtttcttca acgtcgtgcg ctttacgggc tgcgtcgtgt ttgtgatcct gcgcatgttt 2700
cggcgggttcg gctgcggcgc gcgcgcgtcc cgcgacaact accaagaggc gctgccggcc 2760
gagtggcaga cgcactacaa cgtgacgaac cgaaccgatg gccgcgtcgc gccgccgcc 2820
aagcacgcgg cgtcgatgga cccgaccacg ccccatggcg gcgtctacca gcaagtgtaa 2880

<210> 24
<211> 959
<212> PRT
<213> 寄生水霉

<400> 24

Met Ser Asp Ser Asn Leu Asp Leu Ala Ala Arg Leu Arg Ala Leu Arg
1 5 10 15

Glu Gly Gly Ala Glu Pro Ala Pro Ala Pro Ala Pro Thr Pro Tyr Met
20 25 30

His Ser Pro Pro Ser Arg Thr Arg Pro Thr Pro Leu Tyr Thr Gln Glu
35 40 45

Ser Leu Glu Phe Gly Gly Thr Tyr Thr Thr Gly Ser Pro Val Gly Ala
50 55 60

Glu Ala Asp Gly Val Tyr Thr Gln Val Pro Val Trp Lys Asp Ser Lys
65 70 75 80

Glu Lys Thr Tyr Gly Tyr Leu Asp Asp Glu Pro Ala Pro Gln Ala Gln
85 90 95

[0035]

Thr Leu Leu Asn Lys Ala Asn Asp Leu Val Gln Arg Gln Ala Ser Asn
100 105 110

Lys Ala Phe Arg Arg Gln His Thr Ala Ala Phe Arg Pro Leu Pro Asn
115 120 125

Thr Val Glu Glu Leu Leu Asp Gly Ser Pro Thr Tyr Glu Gly Ala Phe
130 135 140

Arg Leu Val Gln Leu Ala Val Gln Met Glu Gln Asp Gly Asp Pro Gln
145 150 155 160

Ala Ala Ile Asn Leu Tyr Ala Asp Ala Gly Ala Thr Leu Val Glu Val
165 170 175

Gly Arg Lys Glu Val Asp Pro Leu Leu Gln Lys Gly Ile Arg Gln Lys
180 185 190

Ala Gln Glu Leu Leu Gln Arg Ala Glu Asp Leu Glu Ala Trp Met Asn
195 200 205

Gly Val Ala Glu Glu Ala Arg Lys Ala Ala Leu Pro Pro Ser Leu Arg
210 215 220

Ile Ala Arg Thr Asn Val Pro Thr Val Glu Gln Thr Trp Ala Gly Arg
225 230 235 240

Pro Pro Pro Phe His Asp Ala Asn Glu Phe Lys Leu Met Arg Tyr Thr
245 250 255

Ala Val Ala Thr Lys Asp Pro Ile Gln Phe Ser Asp Asp Gly Tyr Val
260 265 270

[0036]

Leu Arg Val His Glu Leu Gln Arg Pro Ile Lys Val Phe Ile Thr Ile
275 280 285

Thr Met Tyr Asn Glu Glu Gly Ser Glu Ile Lys Gly Thr Leu Thr Gly
290 295 300

Leu Ala Lys Gly Leu Ala Tyr Met Cys Lys Glu Tyr Gly Asp Asp Phe
305 310 315 320

Trp Gln Gln Val Ala Val Ala Ile Val Ser Asp Gly Arg Thr Lys Ala
325 330 335

Ser Lys Thr Cys Leu Glu Tyr Leu Lys Ala Val Gly Ala Phe Asp Glu
340 345 350

Glu Ile Met Thr Val Thr Ser Leu Gly Val Asp Val Gln Met His Leu
355 360 365

Phe Glu Ser Thr Leu Gln Leu Val Glu Asn Gln Asn Phe Glu Ala Tyr
370 375 380

Tyr Pro Pro Leu Gln Val Ile Tyr Ala Leu Lys Glu Asn Asn Gly Gly
385 390 395 400

Lys Leu Asn Ser His Leu Trp Phe Phe Asn Ala Phe Ser Glu Gln Leu
405 410 415

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

Thr Ser Val Phe Arg Leu Ile Arg Ser Met Glu Arg Asn Ala Gln Ile
435 440 445

[0037]

Gly Gly Val Ala Gly Glu Ile Ala Val Glu Ala Pro Asn Phe Phe Asn
450 455 460

Pro Val Ile Ala Ala Gln His Phe Glu Tyr Lys Ile Ser Asn Ile Met
465 470 475 480

Asp Lys Ser Leu Glu Ser Val Phe Gly Phe Ile Ser Val Leu Pro Gly
485 490 495

Ala Phe Ser Ala Tyr Arg Tyr Glu Ala Ile Arg Ala Val Lys Gly Val
500 505 510

Gly Pro Leu Pro Glu Tyr Phe Lys Ser Leu Thr Ser Thr Thr Lys Glu
515 520 525

Leu Gly Pro Phe Gln Gly Asn Met Tyr Leu Ala Glu Asp Arg Ile Leu
530 535 540

Cys Phe Glu Leu Leu Ala Arg Lys Gln Arg Arg Trp Thr Met His Tyr
545 550 555 560

Val Lys Asp Ala Ile Ala Arg Thr Asp Val Pro Glu Thr Leu Val Asp
565 570 575

Leu Ile Lys Gln Arg Arg Arg Trp Leu Asn Gly Ser Phe Phe Ala Gly
580 585 590

Leu Phe Ala Ile Gly His Phe Gly Arg Val Trp Ser Gln Ser Ser His
595 600 605

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

[0038]

Leu Gln Asn Leu Leu Ser Trp Phe Leu Leu Ser Asn Leu Phe Leu Thr
625 630 635 640

Phe Tyr Phe Val Leu Thr Leu Ala Phe Thr Glu Ser Ala Pro Ala Leu
645 650 655

Leu Gln Thr Met Leu Thr Val Tyr Leu Ala Ile Ile Gly Gly Leu Ile
660 665 670

Val Phe Ala Leu Gly Asn Lys Pro Glu Pro Arg Thr Ala Ser Phe Tyr
675 680 685

Leu Phe Ser Cys Leu Tyr Met Gly Ile Ile Met Leu Leu Val Thr Gly
690 695 700

Ile Ser Ile Tyr Gly Leu Ile Gly Lys Gly Thr Ser Ala Val Lys Asp
705 710 715 720

Pro Arg Thr Ile Thr Gly Ile Phe Ser Asn Cys Thr Val Ser Asp Ala
725 730 735

Glu Leu Ala Gly Gly Val Ile Thr Ser Leu Gly Leu Ile Phe Leu Ser
740 745 750

Ala Phe Val His Gly Glu Phe Gly Ile Leu Leu Ser Phe Val Gln Tyr
755 760 765

Phe Phe Met Leu Pro Thr Phe Val Asn Val Leu Gly Ile Tyr Ala Tyr
770 775 780

Ser Asn Leu His Asp Leu Ser Trp Gly Thr Lys Gly Leu Glu Ser Gly
785 790 795 800

[0039]

Gly Gly His Gly Pro Ala Lys Ala Gly Gly Gly Asn Val Lys Asp Val
805 810 815

Val Glu Gln Gln Lys Lys Ile Glu Ala Ala Arg Gln Ala Ala Ala Arg
820 825 830

Glu Lys Glu Asp Val Asp Asn Ser Phe Arg Ala Phe Arg Ser Thr Leu
835 840 845

Leu Leu Ser Trp Leu Thr Thr Asn Gly Ile Trp Leu Tyr Val Val Thr
850 855 860

Asp Tyr Met Ser Ser Gly Cys Tyr Leu Lys Gly Leu Ser Tyr Ile Val
865 870 875 880

Gly Phe Phe Asn Val Val Arg Phe Thr Gly Cys Val Val Phe Val Ile
885 890 895

Leu Arg Met Phe Arg Arg Phe Gly Cys Gly Ala Arg Ala Ser Arg Asp
900 905 910

Asn Tyr Gln Glu Ala Leu Pro Ala Glu Trp Gln Thr His Tyr Asn Val
915 920 925

Thr Asn Arg Thr Asp Gly Arg Val Ala Pro Pro Pro Lys His Ala Ala
930 935 940

Ser Met Asp Pro Thr Thr Pro His Gly Gly Val Tyr Gln Gln Val
945 950 955

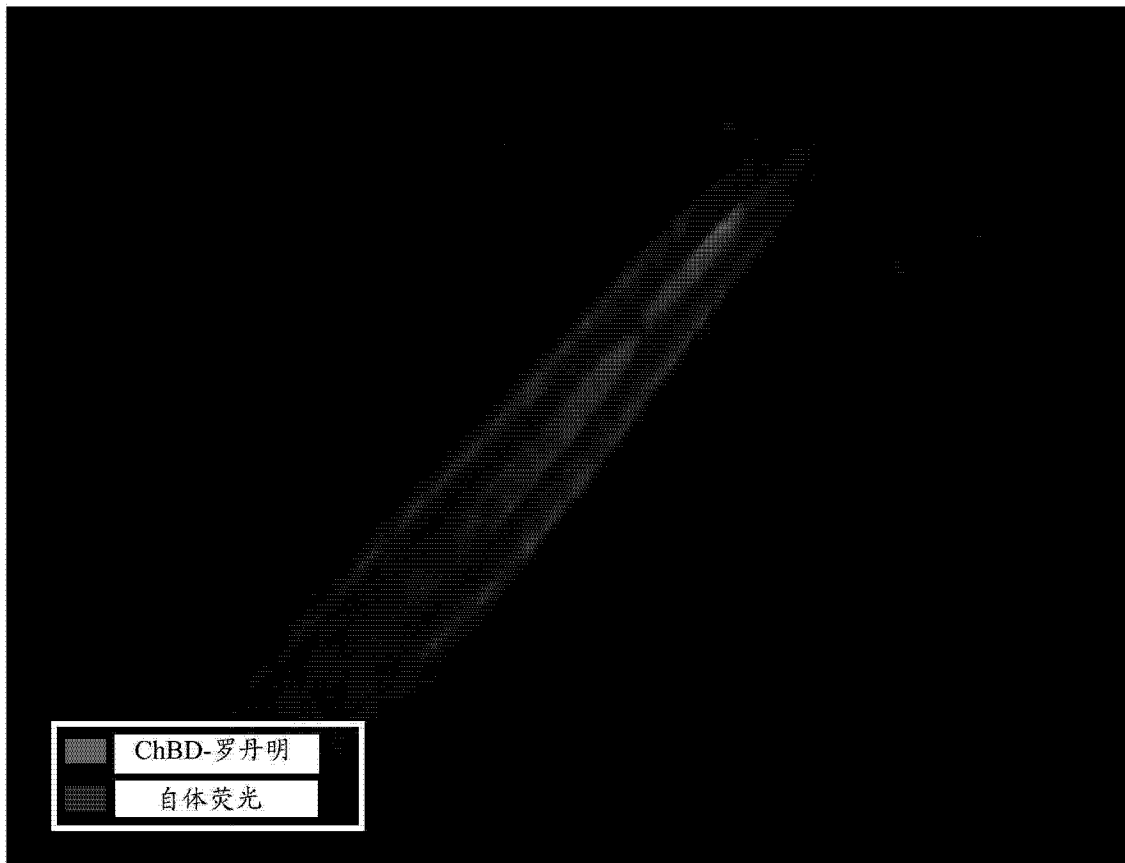


图 1a

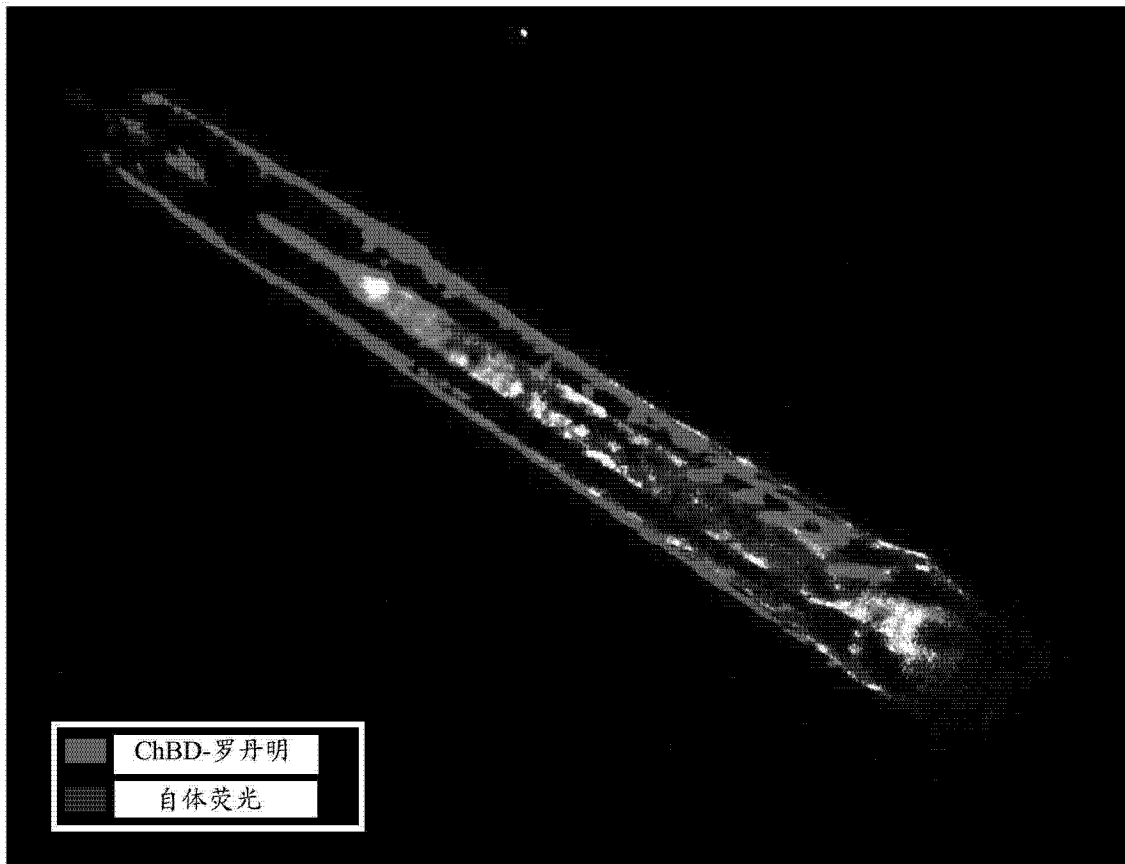


图 1b