

[19] 中华人民共和国国家知识产权局

[51] Int. Cl⁷



[12] 发明专利说明书

专利号 ZL 00805720.6

C12N 9/10

C12N 15/54

C12N 15/29

C12N 15/82

C12N 15/63

C12N 5/04

A01H 5/00

[45] 授权公告日 2005 年 12 月 21 日

[11] 授权公告号 CN 1232638C

[22] 申请日 2000.3.29 [21] 申请号 00805720.6

[30] 优先权

[32] 1999. 3. 31 [33] US [31] 09/282,995

[32] 2000. 2. 11 [33] US [31] 09/502,852

[86] 国际申请 PCT/EP2000/002769 2000.3.29

[87] 国际公布 WO2000/060061 英 2000.10.12

[85] 进入国家阶段日期 2001.9.28

[71] 专利权人 辛根塔参与股份公司

地址 瑞士巴塞尔

[72] 发明人 T·M·赫恩 C·彼特斯

J·M·萨米隆 J·N·瑞德

J·L·道森

审查员 徐益君

[74] 专利代理机构 中国国际贸易促进委员会专利
商标事务所

代理人 樊卫民

权利要求书 2 页 说明书 55 页

[54] 发明名称 对真菌毒素具有抗性的转基因植物
和方法

[57] 摘要

本发明涉及包含异源多核苷酸的植物细胞，所述多核苷酸编码在所述植物细胞中表达的基因产物，其中所述基因产物具有单端孢烯抗性。

I S S N 1 0 0 8 - 4 2 7 4

1.一种植物细胞,其对产生包含C-3羟基的单端孢烯的真菌具有抗性,所述的植物细胞包含编码SEQ ID NO: 2的氨基酸序列的异源多核苷酸。

2.权利要求1的植物细胞,其中植物细胞对镰孢属具有抗性。

3.权利要求2的植物细胞,其中所述的植物细胞对禾本科镰孢具有抗性。

4.权利要求1的植物细胞,其中异源多核苷酸包含SEQ ID NO: 1的核酸序列。

5.权利要求1-4任一项的植物细胞,其中植物细胞是小麦、玉米、大麦、或稻植物细胞。

6.权利要求5的植物细胞,其中植物细胞是小麦植物细胞。

7.一种用于产生包含C-3羟基的单端孢烯抗性植物的方法,包括下列步骤:

a) 用编码SEQ ID NO: 2的氨基酸序列的异源基因转化植物细胞;

b) 以生物学显著水平表达基因产物;

c) 将植物细胞再生成植物; 并

d) 选择对包含C-3羟基的单端孢烯抗性增加的植物。

8.权利要求7的方法,其中所述的单端孢烯由属于镰孢属的真菌产生。

9.权利要求7的方法,其中异源基因包含SEQ ID NO: 1的序列。

10.一种用于在真菌侵染地区预防真菌毒素污染植物,包括植物种子的方法,其中包括培养由权利要求7-9任一项的方法产生的植物。

11.一种用于减少和/或预防在植物上镰孢属的真菌的生长的方法,其中所述真菌产生包含C-3羟基的单端孢烯,包括培养由权利要求7-9任一项的方法产生的植物。

12.权利要求7-11任一项的方法,其中植物是农作物。

13.权利要求12的方法,其中农作物是小麦。

14.一种用于产生对产生包含C-3羟基的单端孢烯的真菌具有抗性的

植物的方法，包括

- a) 用编码SEQ ID NO: 2的氨基酸序列的异源基因转化植物细胞;
- b) 以生物学显著水平表达基因产物;
- c) 将植物细胞再生成植物;
- d) 选择对包含C-3羟基的单端孢烯抗性增加的植物; 并任选的,
- e) 将步骤d)中得到的植株进行自交或异型杂交。

15.权利要求14的方法，其中真菌属于镰孢属。

16.权利要求14或15的方法，其中的异源基因包含SEQ ID NO: 1的序列。

17.一种用于产生种子的方法，其中由种子生长形成的植物具有真菌抗性,包括将由权利要求14-16任一项的方法产生的植物自交或异型杂交,并收集种子,其中所述的种子包含编码SEQ ID NO: 2的氨基酸序列的异源基因。

18.权利要求17的方法，其中种子对镰孢属真菌具有抗性。

对真菌毒素具有抗性的转基因植物和方法

本发明涉及转基因宿主，特别是具有单端孢烯(trichothecene)抗性的转基因植物、植物组织、种子、和细胞，及其产生和使用方法。本发明还涉及用于在植物、植物组织、种子、或植物细胞上预防和/或减少真菌生长的方法。本发明还涉及预防和/或减少植物、植物组织、或种子真菌毒素污染。本发明还涉及使用单端孢烯作为转化方案中的选择剂。

很多真菌对重要的经济农作物具有严重危害。而且，真菌毒素对农作物的污染是世界范围的主要农业难题。真菌毒素是经常在农业产品中发现的有毒真菌代谢物，它能够引起脊椎动物的健康问题。单端孢烯是由镰孢属 (*Fusarium*)、单端孢属 (*Trichothecium*)、和漆斑菌属 (*Myrothecium*) 物种产生的倍半萜烯环氧化物真菌毒素，是真核蛋白质合成的强抑制剂。产生这些单端孢烯的镰孢属物种包括尖圭镰孢 (*F. acuminatum*)、*F. crookwellense*、大刀镰孢 (*F. culmorum*)、木贼镰孢 (*F. equiseti*)、禾本科镰孢 (*F. graminearum*) (玉蜀黍赤霉 *Gibberella zae*)、桑砖红镰孢 (*F. lateritium*)、早熟禾镰孢 (*F. poae*)、接骨木镰孢 (*F. sambucinum*) (虱状赤霉 *G. pulicaris*)、和拟分枝孢镰孢 (*F. sporotrichioides*) (W.F.O.Marasas, P.E.Nelson, 和 T.A.Toussoun, 1984)。

正如先前描述的 (A.E.Desjardins和 T.M.Hohn, 植物发病机理中的真菌毒素, 植物与微生物的分子相互作用 (Mol. Plant-Microbe Interact.)10(2): 147-152, 1997), 农场动物和人的急性和慢性真菌毒素中毒症都与食用被产生单端孢烯真菌毒素的镰孢属物种污染的小麦、黑麦、大麦、燕麦、稻、和玉米有关。使用低剂量水平化学纯的单端孢烯的实验再现了在动物发霉谷粒中毒症中观察到的许多特征, 包括贫血、免疫抑制、出血、呕吐和拒绝进食。人群的病史和传染病学资料指出了

某些传染病与食用被产生单端孢烯的镰孢属物种感染的谷物之间的联系。特别提到的是，十九世纪俄国爆发的致命的消化器官中毒性白细胞缺乏症与食用被产生单端孢烯T-2毒素的镰孢属物种污染的过冬谷物有关。在日本，称为akakabi-byo或赤霉病的类似疾病的爆发与被产生单端孢烯，脱氧瓜萎镰孢醇（以下称“DON”）的镰孢属物种感染的谷物有关。在对印度和日本最近爆发的人类疾病负有责任的有毒谷物样品中检测到单端孢烯。因此，存在对用于预防的农业方法和真菌毒素污染水平减少的农作物的需要。

此外，产生单端孢烯的镰孢属物种是破坏性的病原体，它们攻击广泛的植物物种。单端孢烯的急性植物毒性及其在植物组织中的存在还说明，这些真菌毒素在镰孢属对植物的发病中具有作用。这意味着真菌毒素在疾病中具有作用，因此，减少其对植物的毒性同样可能预防或减少植物疾病。此外，疾病水平的降低可能还有其它好处，即降低真菌毒素对植物（特别是谷物）的污染，这里的植物指谷类植物。

已经用于控制植物疾病（诸如玉米穗腐烂、原种腐烂、或小麦尖枯萎）的各种方法获得了不同程度的成功。控制植物疾病的一种方法是对农作物施用抗微生物的化学药品。这种方法有很多本领域周知的问题。或者，最近的方法包括使用生物学控制生物（“生物控制”），即有害生物的天然竞争者或抑制者。然而，大面积应用生物控制是困难的，使那些活的生物长时间停留在治理地区甚至更困难。最近，重组DNA技术提供了将克隆基因插入植物细胞从而表达抗微生物化合物的机会。然而，这种技术使得人们更加关注微生物最终对众所周知的、天然存在的抗菌剂的抗性。因此，存在鉴定天然存在的抗微生物试剂的持续需要，诸如可以通过单基因转化由植物细胞直接产生的蛋白质。

已经在拟分枝孢镰孢中鉴定了催化许多不同镰孢单端孢烯（包括DON的C3羟基）的乙酰化反应的单端孢烯3-O-乙酰基转移酶（S.P.McCormick、N.J.Alexander、S.C.Trapp、和T.M.Hohn，来自拟分枝孢镰孢、编码单端孢烯3-O-乙酰基转移酶的基因，TRI101的断裂，

应用和环境微生物学(*Applied. Environ. Microbiol.*)65(12): 5252 - 5256, 1999)。单端孢烯在C3-OH的乙酰化显著降低其在脊椎动物和植物中的毒性，并产生反应产物3-乙酰脱氧瓜萎镰孢醇（以下“3ADON”）。参阅Kimura等人，见下文。

已经发布了来自禾本科镰孢和拟分枝孢镰孢之编码单端孢烯3-O-乙酰基转移酶的结构基因的序列，以及其它正同型物的序列。参阅Kimura等人，*生物科学、生物技术、和生物化学 (Biosci. Biotechnol. Biochem.)*62(5): 1033 - 1036, 1998)；和Kimura等人，*FEBS通讯(FEBS Letters)*435: 163 - 168, 1998)。此外，推测来自拟分枝孢镰孢、编码单端孢烯3-O-乙酰基转移酶的基因可用于发展镰孢属抗性增加的植物变种。参阅T.M.Hohn等人，《植物疾病中宿主特异毒素的分子遗传学》(*Molecular Genetics of Host-Specific Toxins in Plant Disease*), 17 - 24, 1998; 和Kimura等人，*生物化学杂志(J. Biological Chemistry)*273(3): 1654 - 1661, 1998。

然而，在本发明之前，许多不确定性使得在植物中表达单端孢烯3-O-乙酰基转移酶是否会实际产生单端孢烯抗性植株变得不明确。例如，由拟分枝孢镰孢单端孢烯3-O-乙酰基转移酶催化的反应是可逆的，因此有可能不会为植物细胞提供针对单端孢烯（诸如DON）的保护。植物细胞中是否存在与3-O-乙酰基转移酶由3ADON产生有毒DON的活性竞争的酯酶同样不确定。反应产物3ADON的新陈代谢如何影响植物同样不确定，如单端孢烯3-O-乙酰基转移酶的导入是否会以否定乙酰基转移酶的任何积极贡献的方式改变植物的生长和发育，例如通过妨碍植物的天然疾病抗性机制尚不确定。3ADON能否由植物新陈代谢成具有毒性作用的新的次级代谢物同样不确定。如果由侵入真菌产生的DON有效转变成3ADON，这种转变是否会赋予植物以增强的病原体抗性甚至同样不确定。上述只是在本发明之前技术中的一些不确定性。

本发明的一个目的是提供包含异源多核苷酸的植物细胞或细胞，其

中所述多核苷酸编码在所述植物细胞中表达的基因产物，该基因产物具有单端孢烯抗性。

本发明的另一个目的是提供包含上述植物细胞的植物，其中植物对单端孢烯具有抗性。

本发明的另一个目的是提供对含C-3羟基的单端孢烯具有抗性的本发明植物。

本发明的另一个目的是提供对产生单端孢烯，优选产生含C-3羟基的单端孢烯的真菌具有抗性的本发明植物。

本发明的另一个目的是提供对镰孢属、单端孢属、或漆斑菌属具有抗性的本发明植物。

本发明的另一个目的是提供对镰孢属具有抗性的本发明植物，特别是（但不限于）禾本科镰孢、大刀镰孢、拟分枝孢镰孢、早熟禾镰孢、接骨木镰孢、木贼镰孢、尖圭镰孢、桑砖红镰孢、和拟禾本科镰孢（*Fusarium pseudograminearum*）。

本发明的另一个目的是提供对禾本科镰孢具有抗性的本发明植物。

本发明的另一个目的是提供包含异源多核苷酸的本发明植物，其中由所述异源多核苷酸编码在植物细胞中表达的基因产物，其中该基因产物具有单端孢烯抗性，其中异源多核苷酸是微生物多核苷酸，优选酵母或真菌多核苷酸。

本发明的另一个目的是提供本发明植物，其中真菌多核苷酸是镰孢属的多核苷酸，优选禾本科镰孢或拟分枝孢镰孢多核苷酸。

本发明的另一个目的是提供本发明植物，其中异源多核苷酸包含与SEQ ID NO: 1、5、或7的核酸序列基本上相似的序列。

本发明的另一个目的是提供本发明植物，其中异源多核苷酸包含SEQ ID NO: 1、5、或7的核酸序列或其同源序列。

本发明的另一个目的是提供包含异源多核苷酸的本发明植物，所述多核苷酸包含与SEQ ID NO: 1、5、或7所述的连续80个碱基对部分序列上相同的至少连续80个碱基对部分，优选与SEQ ID NO: 1、5、或7所述

的连续50个碱基对片段部分相同的至少连续50个碱基对部分，更优选与SEQ ID NO: 1、5、或7所述的连续21个碱基对部分序列上相同的至少连续21个碱基对部分，最优选与SEQ ID NO: 1、5、或7所述的连续18个碱基对部分序列相同的至少连续18个碱基对部分。

本发明的另一个目的是提供对单端孢烯具有抗性的本发明植物，其中单端孢烯选自T-2毒素、HT-2毒素、异木霉菌醇、4,15-二乙酸基

蔗草镰孢烯醇（diacetoxyscripenol）（以下“DAS”）、3-deacetylcalonectrin、3,15-dideacetylcalonectrin、蔗草镰孢烯三醇（scirpentriol）、新腐皮镰孢醇（neosolaniol）；B型：15-乙酰脱氧瓜萎镰孢醇、瓜萎镰孢醇、4-乙酰瓜萎镰孢醇（镰孢酮（fusarenone）-X）、4,15-二乙酰瓜萎镰孢醇、4,7,15-乙酰瓜萎镰孢醇、和DON。

本发明的另一个目的是提供对DAS或DON具有抗性的本发明植物。

本发明的另一个目的是提供本发明植物，其中基因产物由本发明多核苷酸编码。

本发明的另一个目的是提供本发明植物，其中基因产物是3-O-乙酰基转移酶。

本发明的另一个目的是提供本发明植物，其中基因产物是包含与SEQ ID NO: 2、6、或8基本上相似的序列的多肽。

本发明的另一个目的是提供本发明的任一植物的种子。

本发明的另一个目的是提供上述植物的任一项，其中植物是小麦、玉米、大麦、或稻。

本发明的另一个目的是提供用于产生单端孢烯抗性植物的方法，包括下列步骤：

- a)用异源基因转化植物细胞，所述异源基因编码增加单端孢烯抗性的基因产物；
- b)以生物学显著水平表达基因产物；
- c)将植物细胞再生成植株；并
- d)选择单端孢烯抗性增加的植株。

本发明的另一个目的是提供上述方法，还包括选择植物上产生单端孢烯的真菌生长减少的植株的步骤。

本发明的另一个目的是提供上述方法，其中真菌属于镰孢属。

本发明的另一个目的是提供由上述方法获得的单端孢烯抗性植物。

本发明的另一个目的是提供由上述本发明植物通过自交或异型杂交产生的种子，其中由种子生长形成的植物具有增加的单端孢烯抗性。

本发明的另一个目的是提供用于预防真菌毒素污染植物（包括植物种子）的方法，包括培养上述本发明植物，优选在真菌侵染情况为中等至严重的地区，其中植物优选农作物。

本发明的另一个目的是提供用于在植物上预防真菌生长，优选镰孢属真菌的生长的方法，包括培养产生单端孢烯（优选上述含C-3羟基的单端孢烯）的本发明植物，优选在真菌侵染情况为中等至严重的地区，其中植物优选农作物。

本发明的另一个目的是提供用于产生真菌抗性植物的方法，包括

a)用异源基因转化植物细胞，所述异源基因编码的基因产物增加对单端孢烯抗性；

b)以生物学显著水平表达基因产物；

c)将植物细胞再生成植株；

d)选择单端孢烯抗性增加的植株；并

e)任选的，将步骤(d)中得到的植株进行自交或异型杂交。

本发明的另一个目的是提供方法，其中真菌属于镰孢属。

本发明的另一个目的是提供用于产生种子的方法，其中由种子生长形成的植物具有真菌抗性，包括将本发明植物自交或异型杂交。

本发明的另一个目的是提供上述方法，其中种子对镰孢属的真菌具有抗性。

本发明的另一个目的是提供用于选择经转化的宿主细胞的方法，包括：

a)用编码单端孢烯3-O-乙酰基转移酶的核酸构建物转化宿主细胞；并

b)在存在单端孢烯选择剂的情况下培养经转化宿主细胞。

本发明的另一个目的是提供用于选择经转化宿主细胞的方法，其中宿主细胞是植物细胞或微生物细胞，特别是微生物的真菌细胞。

本发明的另一个目的是提供用于选择经转化宿主细胞的上述方法，其中宿主细胞用第二种感兴趣的多核苷酸进一步转化。

本发明的另一个目的是提供用于选择经转化宿主细胞的方法，其中单端孢烯选自 T-2 毒素、HT-2 毒素、异木霉菌醇、DAS、3-deacetylcalonectrin、3,15-dideacetylcalonectrin、蔗草镰孢三烯醇、新腐皮镰孢醇；B型：15-乙酰脱氧瓜萎镰孢醇、瓜萎镰孢醇、4-乙酰瓜萎镰孢醇(镰孢酮-X)、4,15-二乙酰瓜萎镰孢醇、4,7,15-乙酰瓜萎镰孢醇、和DON。

本发明的另一个目的是提供由本发明方法获得的单端孢烯抗性植物。

本发明的另一个目的是提供由本发明方法获得的真菌抗性植物。

本发明的另一个目的是提供由本发明方法获得的植物种子。

为了清楚起见，下文列出并定义了详细说明书中所用的一些术语：

“表达”指植物中内源基因或转基因的转录和/或翻译。在反义构建物的情况中，表达可以仅指反义DNA的转录。

“可操作连接/相连”当用于指调控DNA序列与编码RNA或蛋白质的DNA序列“可操作连接”或“可操作相连”时，指两种序列的排列位置使得调控DNA序列可影响编码DNA序列的表达。

术语“异源多核苷酸”或“异源DNA”，如此处所用，指与其所导入的宿主细胞非天然相关的核酸分子，包括基因构建物，天然存在的核酸分子的非天然存在的多拷贝，和与非天然核酸分子可操作连接的其它同源核酸分子。因此，宿主细胞中的异源基因包括对于特定宿主细胞是内源的、但是例如通过DNA改组(DNA shuffling)经修饰的基因。因此，该术语涵盖对于细胞是外来的或异源的DNA片段，或者对于细胞是同源的，但是所处宿主细胞核酸的位置非正常存在该元件。

术语“核酸”或“多核苷酸”指脱氧核糖核酸或核糖核酸及其单链或双链形式的聚合物。除非明确限制，该术语涵盖包含已知的天然核苷酸类似物、与参考核酸具有相似结合特性、且新陈代谢方式与天然存在的核苷酸相似的核酸。除非特别指出，正如明确指出的序列，特定核酸序列还不言而喻地涵盖其保守性修饰变体（如简并密码子替代）和互补序列。具体的说，可以通过产生其中一个或多个选定（或所有）密码子的第三位用混和碱基和/或脱氧次黄苷残基替代的序列来实现简并密码子替代（Batzer等人，核酸研究(Nucleic Acid Res.)19: 5081, 1991; Ohtsuka等人，生物化学杂志(J. Biol. Chem.)260: 2605 - 2608, 1985; Rossolini等人，分子和细胞探针(Mol. Cell. Probes)8: 91 - 98, 1994。术语“核酸”或“核酸序列”或“多核苷酸”还可以与基因、cDNA、和由基因编码的mRNA互换使用。

当此处用于核酸分子时，术语“基本上相似”在其最广泛的含义中指与参考核苷酸序列相应的核酸分子，其中相应的核酸分子编码的多肽与参考核苷酸序列编码的多肽具有基本上相同的结构和功能，如只有不影响多肽功能的氨基酸发生变化。希望基本上相似的核酸分子编码由参考核苷酸序列编码的多肽。术语“基本上相似”明确的意欲包括其序列经修饰以优化在特定细胞（如植物细胞）中的表达的核酸分子。基本上相似的核酸分子与参考核苷酸序列之间的同一性百分比希望是至少45%，更希望是至少65%，更希望是至少75%，优选至少85%，更优选至少90%，依然更优选至少95%，仍然更优选至少99%。同一性百分比优选跨越至少大约50个残基的序列区域，更优选跨越至少大约100个残基的区域，最优选跨越至少大约150个残基的序列基本上相似。在最优选的实施方案中，跨越完整编码区的序列基本上相似。可以使用Smith - Waterman序列排列算法进行序列比较，下文的描述更为详细（参阅M.S.Waterman, 《计算生物学导论：图谱、序列、和基因组》(Introduction to Computational Biology: Maps, sequences and genomes), Chapman & Hall, 伦敦, 1995, ISBN 0-412-99391-0, 或者 <http://www->

hto.usc.edu/software/seqaln/index.html)。以下列参数使用local S程序1.16版：配对 1，错配罚 0.33，开口罚 2，延伸缺口罚 2。

是本发明核酸的基本上相似的核酸的另一种标志是该核酸序列与本发明核酸分子在严谨条件下发生杂交。短语“特异杂交”指当序列存在于复杂的混和（如总细胞）DNA或RNA中时，分子只与特定核苷酸序列在严谨条件下结合、形成双链体、或发生杂交。“基本上结合”指探针核酸与靶核酸之间的互补杂交，并包括可以通过降低杂交介质的严谨度来包容较小的错配以实现靶核酸序列的期望检测。

核酸杂交实验（诸如Southern和Northern杂交）内容中的“严谨杂交条件”和“严谨杂交清洗条件”具有序列依赖性，而且在不同的环境参数下不同。较长序列在较高温度特异杂交。对于核酸杂交的广泛指南见Tijssen, 1993, 《生化和分子生物学的实验室技术-核酸探针的杂交》(Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes), 第一部分, 第2章, 杂交原理及核酸探针实验策略概论, Elsevier, 纽约。通常, 高度严谨的杂交和清洗条件选择比特定序列在定义的离子强度和pH下的热熔点(T_m)低大约 5°C 。具有代表性的是, 在“严谨条件”下, 探针将与其靶亚序列而非其它序列发生杂交。

T_m 是在定义的离子强度和pH下50%的靶序列与完全配对的探针发生杂交的温度。非常严谨的条件选择与特定探针的 T_m 相同。用于超过100个互补残基的互补核酸在滤膜上进行Southern或Northern杂交的严谨杂交条件的范例是50% 甲酰胺和1mg肝素, 42°C , 杂交过夜。高度严谨的清洗条件的范例是0.15M NaCl, 72°C , 大约15分钟。严谨清洗条件的范例是用0.2x SSC于 65°C 清洗15分钟(参阅Sambrook, 见下文, SSC缓冲液的描述)。通常, 高度严谨的清洗之前进行低度严谨的清洗以除去背景探针信号。用于双链体(如超过100个核苷酸)的中等严谨的清洗是用1x SSC于 45°C 清洗15分钟。用于双链体(如超过100个核苷酸)的低度严谨的清洗的范例是用4-6x SSC于 40°C 清洗15分钟。对于短探针(如大约10

- 50个核苷酸)，严谨条件通常包括盐浓度低于大约1.0M Na离子，通常是Na离子（或其它盐）浓度大约0.01 - 1.0M，pH7.0 - 8.3，温度通常为至少大约30℃。严谨条件还可以通过加入去稳定剂（诸如甲酰胺）来实现。通常，特定杂交实验中观察到的无关探针的2倍（或更高）信号-噪音比指示特异杂交的检测。如果编码的蛋白质基本上相似，那么在严谨条件下不相互杂交的核酸仍然基本上相似。如使用遗传密码允许的最大密码子简并性产生核酸拷贝。

下文是可用于鉴定与本发明参考核苷酸序列基本上相似的同源核苷酸序列的杂交/清洗条件组的范例：待测序列与参考核苷酸序列在7% 十二烷基磺酸钠（SDS）、0.5M NaPO₄、1mM EDTA中于50℃杂交，在2x SSC、0.1% SDS中于50℃清洗；更希望在7% 十二烷基磺酸钠（SDS）、0.5M NaPO₄、1mM EDTA中于50℃杂交，在1x SSC、0.1% SDS中于50℃清洗；依然更希望在7% 十二烷基磺酸钠（SDS）、0.5M NaPO₄、1mM EDTA中于50℃杂交，在0.5x SSC、0.1% SDS中于50℃清洗；优选在7% 十二烷基磺酸钠（SDS）、0.5M NaPO₄、1mM EDTA中于50℃杂交，在0.1x SSC、0.1% SDS中于50℃清洗；更优选在7% 十二烷基磺酸钠（SDS）、0.5M NaPO₄、1mM EDTA中于50℃杂交，在0.1x SSC、0.1% SDS中于65℃清洗。在上述条件下杂交的本发明多核苷酸优选包含至少80个碱基对，更优选至少50个碱基对，特别是至少21个碱基对，更特别是18个碱基对。本发明中所用的优选同源物包括编码据测量（使用下文所述参数）与SEQ ID NO: 2、6、或8至少45%同一的氨基酸序列的核酸分子，其中由同源物编码的氨基酸序列具有单端孢烯抗性，如3-O-乙酰基转移酶活性。

当此处用于蛋白质时，术语“基本上相似”指与参考蛋白质相应的蛋白质，其中蛋白质与参考蛋白质具有基本上相同的结构和功能，如只有不影响多肽功能的氨基酸序列发生变化。当用于蛋白质或氨基酸序列时，基本上相似的蛋白质与参考蛋白质或氨基酸序列之间的同一性百分比希望是至少45%，更希望是至少65%，更希望是至少75%，优选至少85

%，更优选至少90%，依然更优选至少95%，仍然更优选至少99%，使用缺省BLAST分析参数，下文的描述更为详细。

本发明中所用的优选多肽同源物包括那些所含氨基酸序列与SEQ ID NO: 2、6、或8至少45%同一的多肽，其中由同源物编码的氨基酸细胞具有单端孢烯抗性，如3-O-乙酰基转移酶活性。

可以如上所述进行用于比较核酸或蛋白质序列的最佳排列，如通过Smith和Waterman（应用数学进展(Adv. Appl. Math.)2: 482, 1981）的局部同源性算法、通过Needleman和Wunsch（分子生物学杂志(J. Mol. Biol.)48: 443, 1970）的同源性排列算法、通过Pearson和Lipman（美国国家科学院进展(Proc. Natl. Acad. Sci. USA)85: 2444, 1988）的相似性搜寻方法、通过这些算法的计算机化执行（威斯康辛遗传学软件包中的GAP、BESTFIT、FASTA、和TFASTA, Genetics Computer Group, 575 Science Dr., Madison, WI）、或者通过视觉检查（通常参阅Ausubel等人，见下文）。

适用于测定序列同一性和序列相似性百分比的算法的一个范例是由Altschul等人（分子生物学杂志(J. Mol. Biol.)215: 403 - 410, 1990）描述的BLAST算法。用于进行BLAST分析的软件可以由国家生物技术信息中心（<http://www.ncbi.nlm.nih.gov>）公开获得。这种算法包括首先通过鉴定查询序列中长度W的短词来鉴定高得分序列对（HSP），当与数据库序列中相同长度的词一起排列时，查询序列匹配或满足正阈值得分T。T指相邻词的得分阈值（Altschul等人，1990）。这些初始相邻词数据作为开始搜寻包含其之更长HSP的种子。然后将词数据在每个序列上向两个方向延伸，只要累积排列得分增加。使用参数M（一对匹配残基的奖分；总是>0）和N（错配残基的罚分；总是<0）计算核苷酸序列的累积得分。对于氨基酸序列，使用得分矩阵计算累积得分。当累积排列得分由其达到的最大值下降数量X时，当由于一个或多个负得分残基的排列，累积得分达到或低于零时，或者当达到任一序列的末端时，停止词数据每个方向的延伸。BLAST算法参数W、T、和X决定了算法的灵敏度和速度。

以缺省值使用BLASTN程序（用于核苷酸序列），词长（W） 11，期望值（E） 10，截止 100，M= 5，N= -4，比较两条链。对于氨基酸序列，以缺省值使用BLASTP程序，词长（W） 3，期望值（E） 10，BLOSUM62得分矩阵（参阅Henikoff和Henikoff，美国国家科学院进展(Proc. Natl. Acad. Sci. USA)89: 10915, 1989）。

除了计算序列同一性百分比，BLAST算法还进行两种序列之间的相似性统计分析（参阅Karlin和Altschul，美国国家科学院进展(Proc. Natl. Acad. Sci. USA)90: 5873 - 5787, 1993）。由BLAST算法提供的一种相似性量度是最小总几率（P(N)），指示两种核苷酸或氨基酸序列之间发生偶然匹配的几率。例如，如果待测核酸序列与参考核酸序列比较的最小总几率小于大约0.1，更优选小于大约0.01，最优选小于大约0.001，那么认为待测核酸序列与参考序列相似。

两种核酸序列或蛋白质基本上相似的另一种标志是由第一种核酸编码的蛋白质与由第二种核酸编码的蛋白质发生免疫交叉反应或特异结合抗体。因此，当两种蛋白质仅仅由于保守替换而不同时，这两种蛋白质典型的基本上相似。

当用于蛋白质或肽时，短语“特异（或选择性）结合抗体”或“特异（或选择性）免疫反应”指决定蛋白质或其它生物制剂的异质群中存在蛋白质的结合反应。因此，在指定的免疫实验条件下，特定抗体结合特定蛋白质，且不以显著量结合样品中存在的其它蛋白质。在这些条件下与抗体特异结合可能需要根据其特定蛋白质的特异性来选择抗体。例如，针对具有由任何本发明核酸序列编码的氨基酸序列的蛋白质制备的抗体可以进行选择，以获得与这种蛋白质而非其它蛋白质（除了多形变体）发生特异免疫反应的抗体。各种免疫实验形式可以用于选择与特定蛋白质发生特异免疫反应的抗体。例如，固相ELISA免疫实验、Western印迹、或免疫组织化学常规用于选择与蛋白质发生特异免疫反应的单克隆抗体。可用于测定特异免疫反应性的免疫实验形式和条件的描述参阅Harlow和Lane（1988，《抗体，实验室手册》(Antibodies, A Laboratory

Manual), Cold Spring Harbor Publications, 纽约, “Harlow and Lane”)。通常, 特异性或选择性反应至少是背景信号或噪音的两倍, 更典型的是高于背景10 - 100倍。

特定核酸序列的“保守性修饰的变异”指那些编码同一或基本上同一的氨基酸序列的核酸序列, 或者当核酸序列不编码氨基酸序列时, 指基本上同一的序列。由于遗传密码的简并性, 大量功能同一的核酸编码任何给定的多肽。例如, 密码子CGT、CGC、CGA、CGG、AGA、和AGG都编码精氨酸。因此, 在由密码子说明是精氨酸的每一个位置, 密码子可以变成此处所述任何相应密码子而不改变所编码的蛋白质。这些核酸变体是“保守性修饰的变异”的一种, “静默变异”。此处所述编码蛋白质的每一种核酸序列同样描述了任何可能的静默变异, 除非特别注明。技术人员将认识到核酸中的每一个密码子(除了ATG, 通常是甲硫氨酸的唯一密码子)可以通过标准技术经修饰产生功能同一的分子。因此, 编码蛋白质的核酸的每一种“静默变异”隐含于每一种所述序列中。

此外, 本领域技术人员将认识到在所编码的序列中改变、删除、或添加单个氨基酸或小比例氨基酸(通常少于5%, 更典型的少于1%)的个别替代、删除、或添加是“保守性修饰的变异”, 改变导致用化学上相似的氨基酸替代氨基酸。提供功能相似氨基酸的保守替代代表在本领域是众所周知的。下列五组每组包含相互之间可以保守替代的氨基酸: 脂族: 甘氨酸(G)、丙氨酸(A)、缬氨酸(V)、亮氨酸(L)、异亮氨酸(I); 芳香族: 苯丙氨酸(F)、酪氨酸(Y)、色氨酸(W); 含硫: 甲硫氨酸(M)、半胱氨酸(C); 碱性: 精氨酸(R)、赖氨酸(K)、组氨酸(H); 酸性: 天冬氨酸(D)、谷氨酸(E)、天冬酰胺(N)、谷氨酰胺(Q)。还可以参阅Creighton(1984, 《蛋白质》(Proteins), W.H.Freeman and Compny)。另外, 在所编码的序列中改变、删除、或添加单个氨基酸或小比率氨基酸的个别替代、删除、或添加也是“保守性修饰的变异”。

“亚序列”指包含较长核酸或氨基酸序列（如蛋白质）的一部分的核酸或氨基酸序列。

当向核酸掺入额外核苷酸（或其它类似分子）时，核酸被“延长”。最常见的是，使用聚合酶（如DNA聚合酶）进行这种延长，如在核酸3'末端添加序列的聚合酶。

当来自两种核酸的序列在后代核酸中联合时，两种核酸“重组”。当两种核酸作为重组底物时，两种序列“直接”重组。当序列使用中间体（诸如交换寡核苷酸）重组时，两种序列“间接”重组。对于直接重组，不超过一种序列作为重组底物，在某些情况中，没有序列作为重组底物。

两种分子（例如配体与受体）之间的“特异结合亲和力”指优先结合分子混合物中的一种分子。如果结合亲和力为大约 $1 \times 10^4 \text{ M}^{-1}$ 至大约 $1 \times 10^6 \text{ M}^{-1}$ 或更大，那么认为分子结合是特异的。

“底物”指由酶天然识别并在酶天然发挥其功能的生化途径中转变成产物的分子，或分子的修饰形式，同样由酶识别并在与天然反应相似的酶反应中由酶转变成产物。

“转化”指将异源DNA导入细胞、组织、或昆虫的过程。经转化细胞、组织、或昆虫理解为不仅包括转化过程的终产物，还包括其转基因后代。

“经转化的”、“转基因的”、和“重组（体）的”宿主指导入了异源核酸分子的宿主生物，诸如细胞或植物。核酸分子可以稳定整合到宿主基因组中，或者核酸分子也可以作为染色体外分子存在。这种染色体外分子可以自主复制。经转化细胞、组织、或植物理解为不仅包括转化过程的终产物，还包括其转基因后代。“未转化的”、“非转基因的”、或“未重组的”宿主指不包含异源核酸分子的野生型生物，如细菌或植物。

本发明涉及转基因宿主，特别是包含异源多核苷酸的转基因植物、植物组织、植物种子、和植物细胞，其中由所述异源多核苷酸编码的基

因产物具有单端孢烯抗性，及其产生和使用方法。如此处所用，单端孢烯抗性指降低或抑制单端孢烯的植物毒性（特别是对真菌和/或植物）的活性，在本发明的特定实施方案中，单端孢烯抗性指向单端孢烯的C-3位置转移乙酸的活性。

本发明还涉及转基因宿主，特别是表达异源多核苷酸的转基因植物、植物组织、植物种子、和植物细胞，其中由所述异源多核苷酸编码的基因产物具有单端孢烯抗性，基因产物特别是乙酰基转移酶基因产物，更特别的是3-O-乙酰基转移酶基因产物，更特别的是单端孢烯3-O-乙酰基转移酶基因产物，及其产生和使用方法。本发明异源多核苷酸的表达包括RNA的合成，可以通过Northern印迹分析来检测。具体而言，将由本发明异源多核苷酸（在特定实施方案中，指SEQ ID NO: 1、5、或7）衍生的标记探针与由本发明转基因植物分离得到的RNA在7% 十二烷基磺酸钠（SDS）、0.5M 磷酸钠（pH7.0）、1mM EDTA、10mg/ml BSA中于65℃杂交，并在0.5% BSA(成分V)、5% SDS、40mM 磷酸钠(pH7.0)、1mM EDTA、0.25M 氯化钠中于65℃清洗，优选在1% SDS、40mM 磷酸钠（pH7.0）、1mM EDTA、0.125M 氯化钠中于65℃清洗，优选在1% SDS、40mM 磷酸钠（pH7.0）、1mM EDTA中于65℃清洗，由此可以检测本发明异源多核苷酸的表达。

本发明还涉及表达本发明异源多核苷酸的转基因植物、植物组织、植物种子、和植物细胞，其中植物、植物细胞、植物组织、或植物种子具有单端孢烯抗性。如此处所用，当存在单端孢烯时（可以如下文实施例7所述测定），单端孢烯抗性植物、植物细胞、植物组织、和植物种子能够进行新陈代谢。在特定实施方案中，在饱和底物水平温育15分钟后，单端孢烯抗性植物、植物组织、植物细胞、和植物种子具有比野生型对照中天然存在的背景活性水平高至少10nmol 三乙酰氧基蔗草镰孢烯醇（triacetoxyscirpenol）（以下“TAS”）/μg蛋白质的比酶活性，更特别的是至少5nmol TAS/mg蛋白质，更特别的是至少1nmol TAS/μg蛋白质，更特别的是至少0.8nmol TAS/μg蛋白质，更特别的是至少0.5nmol

TAS/ μ g蛋白质, 更特别的是至少0.25nmol TAS/ μ g蛋白质, 更特别的是至少0.1nmol TAS/ μ g蛋白质, 更特别的是至少0.05nmol TAS/ μ g蛋白质, 甚至更特别的是至少0.01nmol TAS/ μ g蛋白质的比活性, 测定实验如下文实施例6所述。

本发明的单端孢烯抗性植物包括那些当存在单端孢烯时, 种子的发芽和生根百分率比野生型对照高的那些植物, 其中存在的单端孢烯浓度为至少5 μ g/ml, 更优选至少10 μ g/ml, 更优选至少15 μ g/ml, 更优选至少20 μ g/ml, 更优选至少25 μ g/ml。在特定的优选实施方案中, 本发明的单端孢烯抗性植物包括当存在单端孢烯时, 种子的发芽和生根百分率比野生型对照的高至少10%, 更优选至少20%, 更优选至少30%, 更优选至少40%, 更优选至少50%, 更优选至少60%, 更优选至少70%, 更优选至少80%, 更优选至少90%的那些植物。

单端孢烯通常分成几个不同的结构组。本发明的特定实施方案致力于针对A组和B组单端孢烯的抗性。A组和B组包括镰孢属单端孢烯, 并主要根据C-8处羰基官能团的存在与否(A组无, B组有)来区别。因此, B组单端孢烯DON在C-8处包含羰基。

本发明更具体的致力于针对含C-3氢氧根的单端孢烯的抗性。这些单端孢烯包括T-2毒素、HT-2毒素、异木霉菌醇、DAS、3-deacetylcalonectrin、3,15-dideacetylcalonectrin、蔗草镰孢烯三醇、新腐皮镰孢醇; 15-乙酰脱氧瓜萎镰孢醇、瓜萎镰孢醇、4-乙酰瓜萎镰孢醇(镰孢酮-X)、4,15-二乙酰瓜萎镰孢醇、4,7,15-乙酰瓜萎镰孢醇、和DON, 及其各种乙酰化衍生物。

在特定实施方案中, 单端孢烯抗性植物及其细胞、组织、或种子对由产生单端孢烯的真菌(特别是镰孢属的真菌)具有抗性。如此处所用, 真菌抗性指接种真菌后不发生感染, 或者与野生型对照相比, 接种真菌后感染的蔓延减少。

在优选实施方案中, 具有真菌抗性的本发明转基因植物是谷类植物, 而且在真菌攻击下, 与野生型对照相比, 包含较少的受感染谷粒或种子。

与相同数目的评价的野生型对照的谷粒或种子相比，受感染谷粒或种子优选减少至少10%，更优选减少至少20%，更优选减少至少40%，更优选减少至少50%。具有真菌抗性的本发明转基因谷类植物包括（但不限于）玉米、小麦、大麦、稻、和燕麦。

在小麦中，可以如下文实施例9所述，通过对每个接种麦尖的有症状和无症状的麦穗进行计数，并计算每个麦尖有症状麦穗所占的百分比，从而评价真菌在麦尖的蔓延。在优选实施方案中，在真菌攻击下，本发明真菌抗性小麦受感染的麦穗比野生型对照少。与相同数目的评价的野生型对照的麦穗相比，受感染麦穗优选减少至少10%，更优选减少至少20%，更优选减少至少40%，更优选减少至少50%。

在玉米中，可以如下文实施例9所述，通过视觉评价受感染谷粒所占百分比来评价真菌在穗中的蔓延。在优选实施方案中，在真菌攻击下，本发明真菌抗性玉米受感染的谷粒比野生型对照少。与相同数目的视觉评价的野生型对照的玉米穗相比，受感染谷粒优选减少至少10%，更优选减少至少20%，更优选减少至少30%，更优选减少至少40%，更优选减少至少50%。在玉米中，可以通过劈开茎并评估变色的量来评价真菌在茎内部的蔓延。在本发明的优选实施方案中，本发明转基因玉米的茎内部和/或外部变色比野生型对照少。

在另一个优选实施方案中，本发明真菌抗性植物包括当存在真菌攻击时，种子发芽百分率比野生型对照高的那些植物。在特定的优选实施方案中，本发明真菌抗性植物包括当存在镰孢属时种子发芽百分率比野生型对照的种子高至少10%，更优选高至少20%，更优选高至少30%，更优选高至少40%，更优选高至少50%，更优选高至少60%，更优选高至少70%，更优选高至少80%，更优选高至少90%，更优选高至少100%，更优选高至少150%的那些植物。

在另一个优选实施方案中，提供了具有真菌抗性的转基因植物，所述转基因植物所产生的种子或谷粒所含有的真菌毒素（如单端孢烯污染）比野生型对照的种子少。在特定的优选实施方案中，提供了农作物植物，

更特别的是谷类植物，所述植物所产生的种子所含有的单端孢烯污染比野生型对照少至少10%，更优选少至少20%，更优选少至少30%，更优选少至少40%，更优选少至少50%，更优选少至少60%，更优选少至少70%，更优选少至少80%。可以如下文实施例10所述测定单端孢烯污染。

本发明中所用多核苷酸包括编码乙酰基转移酶的异源多核苷酸，特别是编码能够赋予单端孢烯抗性的乙酰基转移酶，更特别的是编码单端孢烯3-O-乙酰基转移酶。在特定实施方案中，本发明的异源多核苷酸可以由（但不限于）真菌起源衍生，更具体具体的说是镰孢属、单端孢属、和漆斑菌属起源，更具体的说是镰孢属物种起源，诸如尖圭镰孢、*F. crookwellense*、大刀镰孢、木贼镰孢、禾本科镰孢（玉蜀黍赤霉）、桑砖红镰孢、早熟禾镰孢、接骨木镰孢（虱状赤霉）、和拟分枝孢镰孢。本发明中所用异源多核苷酸包括SEQ ID NO: 1、5、和/或7及与SEQ ID NO: 1、5、和/或7基本上相似的序列。

可以使用传统的重组DNA技术将本发明中所用多核苷酸掺入宿主细胞，诸如植物、真菌、或细菌细胞。通常，这包括使用本领域知道的标准克隆步骤将多核苷酸插入对多核苷酸而言是异源的表达系统。载体包含在含该载体宿主细胞中转录和翻译本发明中所用多核苷酸所需元件。可以使用本领域知道的大量的载体系统，诸如质粒、噬菌体、病毒、和其它经修饰病毒。还可以变更表达系统的成分以增加表达。例如，可以采用截短序列、核苷酸替代、核苷酸优化、或其它修饰。事实上，在适当条件下可以将本领域知道的表达系统用于转化任何农作物植物细胞。本发明的异源多核苷酸优选稳定转化并整合到宿主细胞基因组中。在另一个优选实施方案中，本发明的异源多核苷酸位于自主复制载体上。自主复制载体的范例是病毒，特别是双生病毒。经转化细胞可以再生成完整植株，使得本发明多肽的选择形式赋予转基因植物以单端孢烯抗性。

首先将意欲在转基因植物中表达的本发明多核苷酸装配到位于可在植物中表达的适当启动子之后的表达盒中。表达盒还可以包含需要或选择用于表达本发明异源多核苷酸的任何其它序列。这些序列包括（但不

限于) 转录终止子、用于增强表达的外来序列(诸如内含子)、生存序列、和意欲使基因产物靶向特异细胞器和细胞区室的序列。可以容易的将这些表达盒转移至下文所述植物转化载体。下文是关于典型表达盒的各种成分的描述。

表达盒中所用启动子的选择将决定本发明异源多核苷酸在经转化植物中的空间和时间表达模式。选择的启动子将在特定细胞类型(诸如叶表皮细胞、叶肉细胞、根外皮细胞)中或在特定组织或器官(诸如根、叶、或花)中表达本发明的异源多核苷酸,而且这一选择将反映期望基因产物累积的位置。或者,选择的启动子可以在各种诱导条件下驱动基因的表达。不同的启动子的强度(即启动转录的能力)是不同的。根据利用的宿主细胞系统,可以使用本领域知道的大量的适当启动子中的任一种。例如,对于组成型表达,可以使用CaMV 35S启动子、稻肌动蛋白启动子、或遍在蛋白质启动子。对于调控型表达,可以使用来自烟草或拟南芥的化学诱导的PR-1启动子(参阅美国专利号5,689,044)。

多种转录终止子可用于表达盒。它们负责本发明异源多核苷酸转录的终止及其正确的多聚腺苷酸化。适当的转录终止子是那些已知在植物中发挥功能的转录终止子,包括CaMV 35S终止子、*tml*终止子、胭脂碱合酶终止子、和豌豆*rbcS E9*终止子。它们可以用于单子叶和双子叶植物。

已经发现大量的序列能够增强转录单位内基因的表达,而且这些序列可以与本发明多核苷酸相连以增加它们在转基因植物中的表达。例如,多种内含子序列(诸如玉米*AdhI*基因的内含子)显示增强表达,特别是在单子叶细胞中。另外,还知道大量由病毒衍生的非翻译前导序列能够增强表达,而且它们在双子叶细胞中特别有效。

所选基因的编码序列任选的通过改变编码序列进行了基因工程改造,以优化在感兴趣的农作物物种中的表达。众所周知用于修饰编码序列以实现特定农作物物种中的最佳表达的方法(参阅Perlak等人,美国国家科学院进展(Proc. Natl. Acad. Sci. USA)88: 3324, 1991; Koziel等人,生物技术(Bio/technol.)11: 194, 1993; Fennoy和Bailey-Serres, 核酸研

究(Nucl. Acids Res.)21: 5294 - 5300, 1993)。众所周知通过考虑植物基因、高等植物、绿藻、和蓝细菌中的密码子使用率来修饰编码序列的方法(参阅Murray等人, 核酸研究(Nucl. Acids Res.)17: 477 - 498, 1989中的表4; Campbell和Gowri, 植物生理学(Plant Physiol.)92: 1 - 11, 1990)。

已知植物中存在使基因产物具有靶向的各种机制, 而且已经详细鉴定了控制这些机制发挥功能的序列。例如, 使基因产物靶向叶绿体受到在各种蛋白质氨基末端发现的信号序列的控制, 这种信号序列在进入叶绿体的过程中被切除而产生成熟蛋白质(如Comai等人, 生物化学杂志(J. Biol. Chem.)263: 15104 - 15109, 1988)。其它基因产物定位于其它细胞器, 诸如线粒体和过氧化物酶体(如Unger等人, 植物分子生物学(Plant Mol. Biol.)13: 411 - 418, 1989)。也可以对编码这些产物的cDNA进行操作, 从而将由DNA序列编码的异源产物靶向这些细胞器。另外, 已经鉴定了引起由DNA序列编码的产物靶向其它细胞分区的序列。氨基末端序列负责靶向ER、质外体、糊粉细胞的细胞外分泌(Koehler和Ho, 植物细胞(Plant Cell)2: 769 - 783, 1990)。另外, 氨基末端序列与羧基末端序列一起负责将基因产物靶向液泡(Shinshi等人, 植物分子生物学(Plant Mol. Biol.)14: 357 - 368, 1990)。通过融合上述适当的靶向序列与本发明的异源多核苷酸, 将产生的产物定向任何细胞器或细胞分区是有可能的。

植物转化领域的普通技术人员知道大量的可用于植物转化的转化载体, 与本发明有关的多核苷酸可以与这些载体的任一种一起使用。载体的选择将取决于优选的转化技术和转化的靶物种。对于某些靶物种, 可能优选不同的选择标记。在转化中常规使用的选择标记包括赋予卡那霉素和相关抗生素抗性的*nptII*基因(Messing和Vierra, 基因(Gene)19: 259 - 268, 1982; Bevan等人, 自然(Nature)304: 184 - 187, 1983)、赋予除草剂膦丝菌素抗性的*bar*基因(White等人, 核酸研究(Nucl. Acids Res.)18: 1062, 1990; Spencer等人, 理论和应用遗传学(Theor. Appl.

Genet.)79: 625 - 631, 1990)、赋予抗生素潮霉素抗性的 *hph* 基因 (Blochinger和Diggelmann, 分子和细胞生物学(Mol. Cell. Biol.)4: 2929 - 2931)、赋予氨甲蝶呤抗性的 *dhfr* 基因 (Bourouis等人, 欧洲分子生物学杂志(EMBO J.)2(7): 1099 - 1104, 1983)、赋予草甘磷抗性的EPSPS 基因 (美国专利号4,940,935和5,188,642)、当存在甘露糖时赋予选择性新陈代谢优势的磷酸甘露糖异构酶基因 *manA* (美国专利号5,767,378; 此处完整引用作为参考; 和Miles和Guest, 基因(Gene)32: 41 - 48, 1984)、赋予BASTA抗性的PAT选择标记 (Sung H. Park等人, 体外细胞发育生物学-植物(In Vitro Cell. Dev. Biol.-Plant)34: 117 - 121, 1998)。

许多载体可用于使用根瘤土壤杆菌 (*Agrobacterium tumefaciens*) 的转化。这些载体通常至少携带一个T-DNA边界序列, 包括诸如pBIN19 (Bevan, 核酸研究(Nucl. Acids Res.), 1984) 等载体。适用于土壤杆菌转化的典型载体包括二元载体pCIB200和pCIB2001, 以及二元载体pCIB10及其潮霉素选择衍生物 (参阅美国专利号5,639,949)。

不使用根瘤土壤杆菌的转化在选择的转化载体中不再需要T-DNA序列, 因此, 除了上述包含T-DNA序列的载体, 还可以利用缺乏这些序列的载体。不依赖土壤杆菌的转化技术包括微粒轰击、原生质体摄取 (如PEG和电穿孔)、和微注射。载体的选择主要取决于待转化物种的优选选择。适用于非土壤杆菌转化的典型载体包括pCIB3064、pSOG19、和pSOG35 (参阅美国专利号5,639,949)。

一旦将感兴趣的多核苷酸克隆到表达系统中, 然后将其转化到植物细胞中。本领域众所周知用于植物转化和再生的方法。例如, Ti质粒载体用于投递外来DNA, 以及直接DNA摄取、脂质体、电穿孔、微注射、和微弹另外, 可以利用土壤杆菌属的细菌来转化植物细胞。

本领域众所周知用于双子叶植物的转化技术, 包括基于土壤杆菌的技术和不需要土壤杆菌的技术。非土壤杆菌技术包括由原生质体或细胞直接摄取外源遗传物质。这可以通过PEG或电穿孔介导的摄取、微粒轰击介导的投递、或微注射来实现。在每种情况中, 使用本领域知道的标

准技术，将经转化细胞再生成完整植株。

大多数单子叶物种的转化现在也已经成为常规工作。优选技术包括使用PEG或电穿孔技术将基因直接转移到原生质体中，使用微粒轰击将基因直接转移到愈伤组织中，以及土壤杆菌介导的转化。靶组织可以由这些来源衍生，如小麦栽培变种UC703或玉米基因型CG000526。例如，可以如美国专利申请09/089,111（相应的以WO 98/54961发表）所述进行土壤杆菌介导的玉米转化，此处完整引用作为参考；可以如下所述进行大麦的转化：M.Cho、J.Wong、C.Marx、W.Jiang、P.Lemaux、和B.Buchanan，硫氧还蛋白h在大麦谷粒中的过度表达导致淀粉脱支酶（支链淀粉酶）活性增强，美国国家科学院进展(PNAS)96: 14641 - 14646, 1999; S.Zhang、M.Cho、T.Koprek、R.Yun、P.Bregitzer、和P.Lemaux，燕麦（*Avena sativa L.*）和大麦（*Hordeum vulgare L.*）的商业性栽培变种的基因转化使用由发芽幼苗衍生的体外芽分生组织培养物，植物细胞报告(Plant Cell Rep.)18: 959 - 966, 1999; P.Bregitzer、S.Harlbert、和P.Lemaux，转基因大麦后代中的体细胞克隆变异，理论和应用遗传学(TAG)96: 421 - 425, 1998; M.Cho、W.Jiang、和P.Lemaux，通过改进再生性和减少白化病转化难对付的大麦栽培变种，植物科学(Plant Sci.)138: 229 - 244, 1998; P.Lemaux、M.Cho、S.Zhang、和P.Bregitzer，转基因谷类：Hordeum vulgare L. - 当前的状况及对未来的展望，在《谷类农作物的分子改良》(Molecular Improvement of Cereal Crops)中，I.Vasil和R.Phillips编，Kluwer Academic Publ, Dordrecht, 荷兰，pp 255-316, 1998; S.Zhang、R.Williams-Carrier、D.Jackson、和P.Lemaux，玉米（*Zea mays L.*）和大麦中CDC2Zm和KNOTTED1在体外叶腋芽分生组织增殖过程和不定芽分生组织形成过程中的表达，植物(Planta)204: 542 - 549, 1998; D.McElroy、J.Louwerse、S.McElroy、和P.Lemaux，Ac/Ds在完整大麦组织细胞中的简单瞬时实验的发展，植物杂志(Plant J.)11: 157 - 165, 1997; S.Tingay、D.McElroy、R.Kalla、S.Fieg、M.Wang、S.Thornton、和R.Brettell，根瘤土壤杆菌介导的大麦转化，植物杂志(The

Plant J.)11: 1369 - 1376, 1997; J.Qureshi, Z.Basri, R.Singh, R.Burton, M.Dalton, J.Kollmorgen, 和 G.Fincher, 1988, 土壤杆菌介导的两种大麦变种的转化, Proc.澳大利亚生化和分子生物学学会第42次会议, 1998年9月28日至10月1日, Adelaide, 澳大利亚; J.Qureshi, R.Singh, Z.Basri, R.Stewart, R.Burton, J.Kollmorgen, 和 G.Fincher, 1997, 用于 elite 澳大利亚大麦变种的遗传转化的策略, Proc.澳大利亚大麦技术 8th. Aust.Barley Technical symp., Gold Coast, Queensland, 1997年9月7-12日, 2: 8.9 - 11; P.Lemaux, M.Cho, J.Louwerse, R.Williams, 和 Y.Wan, 轰击介导的大麦转化方法, Bio-Rad US/EG Bull 2007: 1 - 6, 1996; T.Koprek, R.Hansch, A.Nerlich, R.Mendel, 和 J.Schulze, 通过调整轰击条件与组织应答获得能育的转基因大麦的不同栽培变种, 植物科学 (Plant Sci.)119: 79 - 91, 1996; T.Hagio, T.Hirabayashi, H.Machii, 和 H.Tomutsune, 使用潮霉素抗性标记产生能育的转基因大麦植株, 植物细胞报告 (Plant Cell Rep.)14: 329 - 334, 1995; H.Funatsuki, H.Kuroda, M.Kihara, P.Lazzeri, E.Muller, H.Lorz, 和 I.Kishinami, 通过将DNA直接转移至原生质体而再生的能育的转基因大麦, 理论和应用遗传学 (TAG)91: 707 - 712, 1995; A.Jahne, D.Becker, R.Brettschneider, 和 H.Lorz, 由小孢子衍生的转基因不育大麦的再生, 理论和应用遗传学 (TAG)89: 525 - 533, 1994; Y.Wan 和 P.Lemaux, 大量独立转化的大麦能育植株的产生, 植物生理学 (Plant Physiol.)104: 37 - 48, 1994.

本发明多核苷酸可用于赋予广泛的植物细胞以单端孢烯抗性, 包括裸子植物、单子叶植物、和双子叶植物。虽然本发明的异源多核苷酸可以插入 (如转化) 这些宽范围纲中的任何植物细胞, 但是它特别适用于农作物植物细胞, 诸如稻、小麦、大麦、黑麦、玉米、燕麦、马铃薯、甘薯、芜菁、南瓜、西葫芦、小胡瓜、甜瓜、大豆、和高粱。赋予植物以单端孢烯抗性的本发明中所用多核苷酸可以与对产量和质量重要的其它特性一起使用。可以通过本领域知道的育种方法和技术, 将本发明多

核苷酸掺入植物品系。

当通过转化到农作物植物或转化到可以再生成农作物植物的植物细胞培养物获得单端孢烯抗性基因的等位基因时，使用传统育种技术将其转移到商业性变种中以发展单端孢烯抗性农作物，而不需要对等位基因进行基因工程改造和将其转化到植物中。

多种育种步骤的特征为详细描述的人类介入，诸如选择杂交系，指导亲本系的授粉，或选择适当的后代植株。根据期望的特性，选择不同的育种措施。本领域众所周知相关技术，包括（但不限于）杂交、近交、回交、多系育种、品种混合、种间杂交、非整倍体技术、等等。杂交技术还包括植物不育化以通过机械的、化学的、或生化的方法产生雄性或雌性不育植株。用不同品系的花粉对雄性不育植株进行交叉授粉，确保了雄性不育的基因组；但是，雌性能育植株将一律获得双亲系的特性。因此，本发明的转基因种子和植物可用于改良植物品系的育种，在植物、植物组织、或种子上不表现真菌生长或真菌生长减少，由此预防和/或减少所述植物、植物组织、或种子的真菌毒素污染。

上述转基因种子和植物经改造获得的单端孢烯抗性通过有性繁殖或营养生长获得传递，由此可以在后代植株中获得维持和传播。通常所说的维持和传播利用了发展用于特定目的（诸如耕作、播种、或收获）的已知农业方法。由于生长的农作物易受昆虫或感染引起的攻击和伤害，需要采取措施来控制植物疾病、昆虫、线虫、和其它不利条件以提高产量。这些包括机械措施，诸如耕地或除去受感染植物，以及施用农药，诸如杀真菌剂、杀配子剂、杀线虫剂、生长调节剂、成熟剂、和杀昆虫剂。

在种子生产中，种子的发芽质量和均一性是重要的产品特性，而由农夫收获并出售的种子的发芽质量和均一性并不重要。由于难以将已知农作物与其它农作物和杂草种子隔离，难以控制种子传播疾病，难以产生发芽较好的种子，在培育、调节、和销售纯种子领域有经验的生产者已经发展了相当广泛且详细描述的种子生产实践。因此，购买达到特定

质量标准的合格种子，而非使用由其自己的农作物收获的种子，对于农夫来说是普通实践。通常将所用繁殖物（如种子）用包含除草剂、杀昆虫剂、杀真菌剂、杀细菌剂、杀线虫剂、灭螺药、及其混和物的保护剂涂层进行处理。通常所用的保护剂涂层包含诸如克菌丹、萎锈灵、福美双（TMTD®）、methalaxyl（Apron®）、和pirimiphos-methyl（Actellic®）等化合物。如果希望，可以将这些化合物与配制领域通常采用的其它载体、保护剂或施用促进佐剂一起配制，从而针对由细菌、真菌、或动物害虫一起的伤害提供保护。可以通过将繁殖物在液体制剂中浸泡，或者通过涂上湿的或干的组合制剂，来施用保护剂涂层。其它施用方法也是有可能的，诸如针对芽或果的处理。

本发明的另一个方面提供了新的农业方法，诸如上文例示的方法，其特征为使用本发明的转基因植物、转基因植物材料、或转基因种子。

在另一个实施方案中，本发明涉及由转基因植物获得的转基因植物细胞、组织、种子或植物部分。本发明还包括植物的转基因后代，以及由后代获得的转基因植物细胞、组织、器官、种子、和植物部分。

在另一个实施方案中，本发明中所用异源多核苷酸还可以作为转化步骤中的选择标记使用。在这方面，使用上文例示的、本领域知道的表达盒和转化技术，用本发明的异源多核苷酸（编码的基因产物具有单端孢烯抗性）以及感兴趣的第二种异源多核苷酸转化宿主细胞。转化之后，根据暴露于单端孢烯（特别是DAS或DON或T-2毒素）的存活能力选择经转化细胞。使用本领域知道的转化和表达系统，宿主细胞可以是真核的或原核的宿主细胞。宿主细胞可以是植物细胞、真菌细胞、细菌细胞、酵母细胞、动物细胞、或昆虫细胞。

在本发明特定的优选实施方案中，将所编码的基因产物具有单端孢烯活性的多核苷酸在植物细胞转化方法中作为选择标记使用。例如，还可以转化至少表达感兴趣的第二种异源DNA序列的植物、植物组织、植物种子、或植物细胞，以表达所编码的多肽包含与SEQ ID NO: 2、6、或8基本上相似的序列的序列。将经转化细胞转移至含足量植物毒素单端

孢烯（特别是DAS和/或DON和/或T-2毒素）的培养基，以抑制不表达氨基酸序列与SEQ ID NO: 2、6、或8基本上相似的多肽的植物细胞的生长或存活，其中只有经转化细胞将生长或不受阻碍。用于选择表达氨基酸序列与SEQ ID NO: 2、6、或8基本上相似的多肽的植物的单端孢烯浓度范围为1 - 90 μ g/ml。这一方法可用于能够表达所含核苷酸序列与SEQ ID NO: 1、5、或7基本上相似的多核苷酸的任何植物细胞，而且可以与任何感兴趣的异源DNA序列一起使用。第二种异源DNA序列和本发明异源多核苷酸的表达可以由在植物细胞中发挥功能的同一启动子或不同启动子驱动。

序列的描述

- SEQ ID NO: 1 是来自拟分枝孢镰孢、编码具有单端孢烯抗性的本发明多肽的cDNA序列。
- SEQ ID NO: 2 是由SEQ ID NO: 1编码、具有单端孢烯抗性的本发明多肽。
- SEQ ID NO: 3 是DNA引物。
- SEQ ID NO: 4 是DNA引物。
- SEQ ID NO: 5 是来自禾本科镰孢、编码具有单端孢烯抗性的本发明多肽的DNA序列。
- SEQ ID NO: 6 是由SEQ ID NO: 5编码、具有单端孢烯抗性的多肽。
- SEQ ID NO: 7 是来自酿酒酵母、编码具有单端孢烯抗性的本发明多肽的DNA序列。
- SEQ ID NO: 8 是由SEQ ID NO: 7编码、具有单端孢烯抗性的多肽。
- SEQ ID NO: 9 是pCIB9818的DNA序列。
- SEQ ID NO: 10 是pAgroTRlr的DNA序列。
- SEQ ID NO: 11 是pNOV1704的DNA序列。

实施例

下列实施例进一步描述了进行本发明所用的材料和方法，及随后的结果。我们以例示的形式提供这些实施例，它们的叙述不应当理解为对本发明的限制。

实施例1: pNOV1700、pCIB9818、pAgroTRIr、和pNOV1704的组成

1. pNOV1700:

pNOV1700根据布达佩斯条约于1999年5月19日收藏于北方研究机构保藏中心(NRRL)，1815 Northern University Street, Peoria, Illinois, 61604, 美国, 编号NRRL-30117。

pNOV1700包含与玉米遍在蛋白质启动子(包括一部分外显子和内含子)和胭脂碱合酶多聚腺苷酸化信号可操作连接的SEQ ID NO: 1。

2. pCIB9818:

质粒pCIB9818是6111个碱基对的环状质粒, 包含SEQ ID NO: 9的DNA序列。包括一部分外显子(第896-1011位碱基)和内含子(第1047-1993位碱基)的玉米遍在蛋白质启动子(SEQ ID NO: 9的第12-1993位碱基)与磷酸甘露糖异构酶选择标记(第2090-3193位碱基对)、反向PEPC内含子#9(第3248-3355位碱基)、和CaMV 35S基因的终止序列(第3357-3434位碱基)可操作连接。

3. pAgroTRIr:

质粒pAgroTRIr是13737个碱基对的环状二元载体, 包含SEQ ID NO: 10的DNA序列。pAgroTRIr包含与启动子和终止序列可操作连接的选择标记, 以及SEQ ID NO: 1的多核苷酸区, 其位于拟南芥(*Arabidopsis thaliana*) UBI3启动子(S.Norris, S.Meyer, 和J.Callus, 植物分子生物学(Plant Molecular Biology)22: 895-906, 1993)之后且读码框一致, 位于nos多聚腺苷酸化信号之前且读码框一致。

4. pNOV1704:

质粒pNOV1704是12949个碱基对的环状二元载体，包含SEQ ID NO: 11的DNA序列。包括外显子1（第895 - 1010位）和内含子1（第1046 - 1992位）的玉米遍在蛋白质启动子（SEQ ID NO: 11的第11 - 1992位碱基）与磷酸甘露糖异构酶选择标记（第2089 - 3192位碱基）和胭脂碱合酶终止序列（第3557 - 3688位碱基）可操作连接。pNOV1704还包含与SEQ ID NO: 1的单端孢烯3-O-乙酰基转移酶序列（第11234 - 12662位碱基）和nos终止序列（第12667 - 12935位）可操作连接的玉米遍在蛋白质启动子（第9218 - 11218位碱基，包含外显子1: 第10110 - 10224位和内含子1: 10225 - 11218位）。

实施例2: 小麦的转化、选择、和再生

转化:

由表面消毒的小麦颖果（10% Chlorox, 10分钟）分离未成熟合子胚（0.75 - 1.25mm），然后将盾片朝上置于基于MS的培养基（Murashige和Skoog, 植物生理学(Physiol. Plant)15: 437 - 439, 1962）上，培养基添加3% 蔗糖、3mg/ml 2,4-D（二氯苯氧乙酸）、150mg/ml 谷氨酰胺、75mg/ml 天冬酰胺，并用0.7% 植物用琼脂（3MS3S培养基）固化。在进行轰击之前，将胚在黑暗中于28℃温育5 - 10天。轰击的最佳时间是铺板后6 - 7天。轰击前4小时，将胚置于质壁分离培养基（与上述培养基相同，只是加入了15% 麦芽糖而非蔗糖）上，并将盾片朝上排列成直径2.5cm的圈。

用PvuII和XmnI消化上文实施例1中所述pNOV1700，并分离包含具有SEQ ID NO: 1的序列的多核苷酸区以及遍在蛋白质启动子和nos多聚腺苷酸化信号的4117bp片段。用AscI消化上文实施例1中所述pCIB9818，并分离包含UBI玉米启动子、选择标记、和CaMV 35S终止序列的4246bp片段。

使用标准Sanford法，将分离得到的DNA片段沉淀在0.3 μm金微粒上。在持续振荡中，向装有50 μl 50% 甘油和3mg 金微粒的Eppendorf

管中加入5 μ g 每种DNA构建物的分离片段、50 μ l 2.5M CaCl₂、和20 μ l 0.1M 亚精胺。将DNA/金微粒混合物用乙醇清洗两次。最后一次清洗弃去上清液后，加入乙醇至终体积70 μ l。每管金微粒/DNA可以提供6次发射。对靶板发射两次，使得向每个靶板投递了大约3 μ g DNA（如果两种构建物都使用了5 μ g，这是通常的情况）和1.0mg 金微粒。所用破裂压是1100psi。轰击后，将靶板回到黑暗中放置过夜。质壁分离大约24小时后，将胚转移至3MS3S，并回到黑暗中放置3周，进行愈伤组织起始。在此过程中不进行传代培养。

选择/再生:

与非胚发生组织相分离在3周起始期间形成的胚发生组织，并置于再生/选择培养基上。基本的再生培养基是3MS3S，不含2,4-D，但是加入1mg/l GA3（赤霉素A₃）和NAA（1-萘乙酸）（称为NG培养基）。加入10g/l 甘露糖和5g/l蔗糖（NG1M.5S）。将组织进行这一再生和选择初始阶段2周。对于这两周的大部分时间，将组织置于光照室中。芽和根的发育在这一阶段开始，2周后，将所有组织带入下一阶段。

对于使用甘露糖选择的再生和选择的第二阶段，将甘露糖降至5g/l，而将蔗糖升至20g/l（MS2S.5M）。通常将组织在这些培养基上放置大约4周时间，在此过程中，发生芽和根的发育。

将板中旺盛生长、颜色鲜艳、根与芽发育的幼苗从板上转移至较大容器GA7。这是选择和再生的最后阶段。培养基只含1/2MS盐和15g/l 甘露糖。植物已被转化的最好指示是观察到根积极生长到培养基中。收集活泼生长幼苗的叶组织，并在转移至温室之前，对感兴趣的基因和选择标记都进行PCR。

实施例3: 芥属的转化

通过电穿孔（W.J.Dower，分子生物学报告(Mol. Biol. Rep.)1:5,1987），将上文实施例1中所述二元载体pAgroTRIr构建物转化到根瘤土壤杆菌菌株GV3101（N.Bechtold等人，CR Acad. Sci. Paris，

Sciences de la vie, 316: 1194 - 1199, 1993) 中。含pAgroTRIR质粒的GV3101土壤杆菌单菌落在25ml添加雷帕霉素 (Rifampin) 100和卡那霉素 (Kanomycin) 100的YEB培养基中于30℃温育过夜。用5ml小量培养液接种500ml相同培养基, 并于30℃温育过夜, 由此进行大量培养。测定培养液于600nm的OD值, 然后将培养液在GSA转头中以5K离心15分钟。将细胞重悬于“IM修饰的渗透培养基”, 使600nm的OD值达到0.08。向每升悬浮细胞中加入200 μl Silwet。将3盆开花的拟南芥哥伦比亚变种 (每盆大约4株植物) 倒置于大约500ml细胞悬浮液中。将花在细胞悬浮液中振摇, 以除去气泡, 并将植物在细胞悬浮液中温育15分钟。给盘子加盖, 保持植物的湿度, 过夜。

让植物生长大约3 - 4周, 然后停止浇水1周。收集荚果, 并在干燥室中干燥大约1周半。播种并生长大约2周。向植物喷洒选择剂, 2天后和4天后再次喷洒。大约3天后, 将存活植物转移至新盆。

实施例4: 玉米的biolistic转化

由来自温室生长物的1.5 - 2.0mm未成熟玉米胚起始I型胚发生愈伤组织培养 (Green等人, 1983, 玉米中的体细胞遗传系统, A.Fazelahmad, K.Downey, J.Schultz, RW Voellmy编, 《基因技术进展: 植物和动物的分子遗传学》(Advances in Gene Technology: Molecular Genetics of Plants and Animals), 迈阿密冬季系列座谈会, 第20卷, Academic Press, NY)。授粉后大约14天, 由表面消毒的玉米穗无菌分离胚。将胚置于添加了2% 蔗糖和5mg/L 氯氮苯的D愈伤组织起始培养基 (Duncan et等人, 植物(Planta)165: 322 - 332, 1985)。随后在黑暗中培养胚和胚发生培养物。大约14天后由外植体切下胚发生应答。将应答置于添加了2% 蔗糖和0.5mg/L 2,4-D的D愈伤组织维持培养基上。每周用新鲜的维持培养基进行选择性传代培养, 大约6周后, 建立了高质量的紧密胚发生培养物。选择积极生长的胚发生愈伤组织块作为基因投递的靶组织。基因投递前大约4小时, 将愈伤组织块置于装有添加了12% 蔗糖的维持培养基的靶

板中。

将愈伤组织块排列成圈，据靶板中心的直径为8和10mm。

用PvuII和XmnI消化上文实施例1中所述pNOV1700，并分离包含具有序列SEQ ID NO: 1的多核苷酸以及启动子和多聚腺苷酸化信号的4117bp片段。用AscI消化上文实施例1中所述pCIB9818，并分离包含标记基因、启动子、和终止信号的4246bp片段。如DuPont Biolistics手册所述，将分离得到的DNA片段沉淀在金微载体上。6次微载体试剂发射中每次使用2-3 μg每种质粒。使用PDS-1000He Biolistics设备，将本发明多核苷酸投递给靶组织细胞。Biolistics设备上的设置如下：破裂盘和微载体之间相距8mm，微载体和终止屏之间相距10mm，终止屏和靶之间相距7cm。使用650psi破裂盘对每个靶板发射2次。将200x200不锈钢网（McMaster Carr, New Brunswick, NJ）置于终止屏和靶组织之间。基因投递后7天，将靶组织块由高渗透压培养基转移至选择培养基。

将靶组织置于不含蔗糖、含1%甘露糖的维持培养基上。3-5周后，用不含蔗糖、含1.5%甘露糖的维持培养基传代培养生长中的愈伤组织块。每2周对在选择培养基上生长的胚发生愈伤组织进行传代培养，6-10周，直到产生的愈伤组织足以产生10-20株植物。将原始靶组织块经历选择后残存的组织以单集落传代培养，并命名为独立的转化事件。将集落转移至经修改的MS培养基（Murashige和Skoog，用于烟草组织培养物的快速生长和生物实验的修正培养基，植物生理学(Physiol. Plant)15: 473-497, 1962），含2%蔗糖和1%甘露糖（MS2S+1M），添加0.25mg/L环丙嘧啶醇和0.5mg/L激动素。2周后，将再生集落转移至不含激素的MS2S+1M。将所有集落的有根或无根的再生苗转移至装有MS3S培养基的Magenta盒，回收有根的小植株并转移至温室土壤。

实施例5: 转基因植物表达的分析

a. PCR分析

对来自经转化植物的组织分析包含序列SEQ ID NO: 1的多核苷酸的

存在。按照标准方案，由经转化植物提取并进行PCR分析。用于扩增基因构建物的引物是(5'-acgaatcattcaccgaggag-3') (SEQ ID NO: 3) 和 (5'-ctcacactctcaggcttacc-3') (SEQ ID NO: 4)。检测到上文实施例2中获得的小麦序列SEQ ID NO: 1中的650nt片段。

b. Northern分析

通过Northern印迹杂交，对经转化植物分析RNA的存在。对于Northern印迹分析，将由植物组织提取的RNA进行大小分离，并转移到尼龙膜上。随后将该膜与由SEQ ID NO: 1的多核苷酸的429nt StyI片段衍生的放射性探针杂交。检测按照上文实施例2和3转化的小麦和拟南芥中的RNA。

实施例6: 用于鉴定单端孢烯3-O-乙酰基转移酶活性的酶实验

a) 用于酶实验的植物组织的提取: 选择来自本发明转基因植物(包括按照上文实施例2-4转化和再生的植物)的3片1x1/8英寸叶片(大约50mg)。

b) 玻璃珠研磨: 将组织置于2ml圆底管中，关上盖子。将管子浸入液氮，并于-80℃保温过夜。将管子在saws-all上振摇24秒钟，加入0.4ml磷酸钠缓冲液。将管子振荡10秒钟，并置于冰上。将管子再振荡5分钟，然后以14,000rpm在Eppendorf离心机中离心5分钟。经上清液转移至干净的管中。

混和下列成分:

单端孢烯底物，2 μl DAS (20% 丙酮，50mM 磷酸钠缓冲液，pH7.0)。

也可以使用DON。

乙酰辅酶A底物，2 μl [¹⁴C]-乙酰辅酶A NEN产品目录号NEC313 (60mCi/mmol和0.02mCi/ml)。

缓冲液，磷酸钠缓冲液 (pH7.0) 至终体积50 μl。

加入下面的酶制剂来起始实验，并于30℃温育15分钟。

酶制剂，10 μl溶于磷酸钠缓冲液 (pH7.0) 的植物提取物

15分钟后，加入100 μl乙酸乙酯，并经管子振荡2次，每次数秒钟。

将管子以14,000rpm在Eppendorf离心机中离心2分钟。将50 μ l乙酸乙酯相转移至装有闪烁混合液的小瓶中。使用闪烁计数器对管子计数2分钟。

按照上文实施例2获得20株分离的小麦植株，鉴定的比活性为0.60 - 13.4nmol乙酰化产物/ μ g蛋白质/15分钟。经转化小麦植株的比活性显著高于阴性对照。阴性对照是未转化的小麦栽培变种，其比活性为0.1 - 0.2nmol乙酰化产物/ μ g蛋白质/15分钟。

按照上文实施例3获得5株分离的拟南芥植株，鉴定的比活性范围为3.8 - 28nmol乙酰化产物/ μ g蛋白质/15分钟。经转化植株的比活性显著高于阴性对照。阴性对照是用用于表达选择标记的核酸构建物转化的拟南芥哥伦比亚变种，其比活性低于0.1nmol乙酰化产物/ μ g蛋白质/15分钟。

鉴定了来自按照上文实施例4获得的至少2株不同转化体的玉米植株，其比活性范围为11.1 - 17.9nmol乙酰化产物/ μ g蛋白质/15分钟。经转化植株的比活性显著高于阴性对照。阴性对照是未转化的玉米基因型，其比活性低于0.2nmol乙酰化产物/ μ g蛋白质/15分钟。

鉴定了来自使用土壤杆菌介导的pNOV1704转化获得的至少16株不同转化体的玉米植株，其比活性范围为17 - 183nmol乙酰化产物/ μ g蛋白质/15分钟。

实施例7: 用于鉴定转基因植物中的单端孢烯抗性的生物实验

配制250ml包含下列成分的CPR培养基，并用KOH将pH调至6.5:

1/2 MS盐	0.54g
1/2 MS维生素	1.25ml
1%蔗糖(任选)	2.50g.

向上述培养基中加入琼脂糖至1% (2.50g)，并将培养基高压灭菌。向高压灭菌后的培养基中加入25ml 50mg/ml 氯酚红。

当培养基维持在55 $^{\circ}$ C时，以各种浓度加入溶于丙酮的DAS或DON(即每1.7ml加入10mg/ml DON 4、8、或16 μ l或者50mg/ml DAS 2、4、或6 μ l)。向48孔微量滴定板的每个孔中加入大约0.5ml培养基。

向微量滴定板的孔中加入1/3x1/8英寸的经转化植物组织以及来自未转化野生型对照的对照组织。将叶片放在培养皿中，并用镊子推入微量滴定板孔的培养基中。将微量滴定板于20℃光照下温育2-4天。叶片新陈代谢导致颜色由红色变成黄色（pH下降）。转化体具有对单端孢烯抗性或对单端孢烯的敏感性降低，导致存在DAS或DON的孔中的培养基变成黄色。

与保持红色的对照相比，在本发明的小麦和玉米植物中观察到由红色到黄色的颜色变化。此外，个别叶片的萎黄病显著少于相应对照。

实施例8: 发芽实验

A. 单端孢烯抗性发芽实验

在来自选择剂的选择压力下培养来自本发明转基因植物的种子，并将获得的植物自交。将产生的种子以1000-2000粒种子/培养皿（直径100mm）的密度置于添加了DAS或DON（20mg/ml）的MS3S培养基（MS盐4.3g/L，MS维生素100x，蔗糖30g/L，植物用琼脂8g/L）上。光照下温育4天后，检验板上的幼苗生长情况。

将由按照上文实施例3获得的植物的拟南芥种子在含DAS的培养基中培养，得到大量根和芽都发育的植株。当在添加了DAS的培养基中生长时，对照种子（拟南芥亲本系，哥伦比亚变种）发芽情况很差，没有形成根。在不含DAS的同一培养基中，没有观察到经转化种子与对照种子之间的差异。

B. 用于检测幼苗枯萎病抗性的真菌抗性发芽实验

1. 小麦的真菌抗性发芽实验

基本上如R.H.Proctor、T.M.Hohn、和S.P.McCormick（单端孢烯毒素生物合成基因的断裂导致玉蜀黍赤霉的毒性降低，植物与微生物之间的分子相互作用(Mol. Plant-Microbe Interact.)8(4): 593-601, 1995)所述进行小麦的真菌抗性发芽实验，此处完整引用作为参考。

接种物包含禾本科镰孢的大型分生孢子，用水稀释至 1×10^6 个分生孢

子/ml。通过清洗来自在白光和近紫外荧光灯下生长7-10天的V-8汁琼脂培养物的大型分生孢子来制备接种物。在幼苗实验中,通过在10%漂白剂和0.05%吐温溶液中清洗大约15分钟,并用无菌蒸馏水漂洗5次,对自上文实施例2的2株不同的转基因小麦事件的种子和野生型对照进行表面消毒。将种子在大型分生孢子悬浮液中浸泡大约10分钟,然后播种于装在10cm塑料盆的蛭石中,每盆20粒种子。播种前,将盆用蛭石装大约3/4满,浇水2-4cm,直至蛭石表面也湿了。播种后,用另外1-2cm蛭石覆盖种子,将盆单个置于塑料袋中,并在生长间中于22℃温育1周,光照16小时,黑暗8小时。大约1周后,将盆由袋中取出,2周后,通过计数每个盆中出现的幼苗数目来评价疾病情况。如上所述处理对照,除了将种子在无菌水中浸泡40秒。

与用水处理的相同转基因种子相比,来自2株不同的转基因植物事件的种子有50%和43%发芽;而与用水处理的相同种子相比,野生型对照有17%发芽了。

2. 玉米的真菌抗性发芽实验

由在绿豆琼脂(用煮沸绿豆汁制成)上、光照与黑暗12小时更替、于25℃生长的禾本科镰孢培养物制备接种物。首先向板中注入无菌水,然后用玻璃棒刮板,由此收获孢子。收集溶液,并于接种当天用无菌双蒸水将孢子浓度调至 1×10^6 个/ml。

5升盆中的每升土壤接种1ml孢子溶液,土壤为土壤、泥煤、和蛭石的无菌混和物,将接种后的土壤在水泥搅拌器中进行搅拌,每个装载量2分钟。对照处理包括未感染的土壤。在盆中种植30粒本发明的转基因种子或野生型对照,并于保持在55°F的生长间中在半饱和土壤条件下温育4周。光照水平为种植后14天保持24小时黑暗,种植后第15-24天为12小时光照。向盆中插入标记桩,然后将盆移至生长间并随机化。

植物一出现就开始计数,每天或隔天进行,直至不再有变化。

测定视觉变色的症状和出土幼苗的腐烂,用于描述相对于野生型对照的植物抗性程度。选择视觉变色和/或出土幼苗腐烂比野生型少的植株。

实施例9: 真菌抗性实验

A. 转基因小麦植物的测试

小麦麦尖枯萎病或麦尖斑点病由禾本科镰孢（有性型：玉蜀黍赤霉）引起。在V-8琼脂培养基（用V-8汁制成）或绿豆琼脂（用煮沸绿豆汁制成）上，在光照与黑暗12小时更替下，于25℃培养禾本科镰孢培养物。首先向板中注入无菌水，然后用玻璃棒刮板，由此收获孢子。收集溶液，并于接种当天用无菌双蒸水将孢子浓度调至 5×10^4 个/ml。

由上文实施例2所述获得转基因植物，可以在温室中培养对照植株直至开花或抽穗。通过在每个穗头近中部的一朵花的外稃和内稃之间接种20 μ l（大约1000个孢子）接种物来进行穗头接种。留一些穗头不接种或用无菌水接种作为对照。然后将植物移至生长间，并在高湿度下温育21天，或者首先将植物在雾室中于65-70°F温育72小时，然后在温室中再温育18天。

也可以在田间培养本发明的转基因植物和对照植物，通过喷洒进行接种。

通过在有代表性数目的转基因穗头和野生型对照穗头上计数每个接种穗头中有症状和无症状的小穗，并计算每个穗头上有症状小穗所占百分比，由此评价疾病。症状包括相对于对照植物早熟和某些情况中的小穗坏死。选择有症状小穗比野生型少的植株。

在根据Southern数据具有不同的SEQ ID NO: 1拷贝数和不同的插入位点数目的6株不同的转基因植物中，每个穗头的有症状小穗的平均百分比范围为10.40% - 31.20%，而野生型对照为44.75%，如上所述，转基因植物和对照在温室中温育。这些相同的转基因植物具有范围为0.874 - 29.1nmol/ μ g/15min的酶活性，如上文实施例6所述测量。

B. 转基因玉米植物的测试（玉米穗腐烂实验）

玉米穗腐烂由禾本科镰孢（有性型：玉蜀黍赤霉）引起。在V-8琼脂培养基（用V-8汁制成）上，在光照与黑暗12小时更替下，于25℃培养禾

本科镰孢培养物。首先向板中注入无菌水，然后用玻璃棒刮板，由此收获孢子。收集溶液，并于接种当天用无菌双蒸水将孢子浓度调至 5×10^5 个/ml。

在温室或田间培养转基因植株和对照植株。当在温室中培养时，穗丝出现后，向穗丝沟（外壳腔内，穗轴以上）中注入2ml孢子悬浮液，并将转基因植株和对照植株在温室中再维持4-7天。使用装配好的18号不锈钢皮下注射器针头和大注射器进行这一步。除了穗丝沟接种，谷粒接种法也可用于测定疾病抗性。谷粒接种包括使用装配好的18号针头和注射器对一组4粒谷粒多次注射孢子悬浮液（大约0.4ml）。

通过对接种后5-7周收获的玉米穗视觉检查可见的受感染谷粒来评价疾病。去壳玉米穗的发病率标准是基于下面的玉米穗上可见的受感染谷粒所占百分比的视觉评价：1是0%；2是1-3%；3是4-10%；4是11-25%；5是26-50%；6是51-75%；7是76-100%。选择可见的受感染谷粒所占百分比比野生型对照少的玉米植株。

实施例10: 真菌毒素污染实验

如下制备用于真菌毒素浓度分析的样品。由本发明的转基因植物收集种子，一起称重量、量体积。当测定小麦种子时，由相同的本发明转基因植物的穗头收集小麦种子，一起称重量、量体积。当测定转基因玉米时，将玉米穗干燥至低含水量水平，将玉米穗手工脱粒，并将来自相同的转基因植物的玉米穗的谷粒一起称重量、量体积。将每份种子或谷粒样品充分搅拌，以促进种子的随机分布。将50g种子或谷粒样品用磨（如ZM1型Retsch超速离心磨，Brinkman设备公司，Rexdale, Ontario, Romer Series II Mill, Union, Missouri, USA）磨成细粉。然后使用商品化测试（诸如DONtest TAG™真菌毒素测试系统，VICAM, LP, 313 Pleasant Street, Watertown, MA 02472）测定感兴趣的真菌毒素（诸如DON）的浓度，或者通过商业分析公司（如Romer Labs公司，Union, Missouri, 美国；或Trilogy Analytical Laboratory公司，Beaufort, Missouri, 美国）

进行分析。分析的所有方面都遵循制造商的指示。对于DONtest TAG™真菌毒素测试系统，对DON进行最终的荧光测量。选择真菌毒素（诸如DON）比野生型对照少的产生种子或谷粒的植株。

实施例11：使用SEQ ID NO: 1的多核苷酸作为选择标记

A. 真菌细胞中的选择标记

使用标准真菌转化技术，用DNA构建物进行棉阿舒囊霉（*Ashbya gossypi*）转化，该DNA构建物所含多核苷酸包含与半乳糖苷酶启动子可操作连接的序列SEQ ID NO: 1。在含DAS浓度范围1.56ng/ml - 196pg/ml的培养基中培养经转化细胞，而未转化的野生型真菌细胞的培养基不含DAS。

B. 植物细胞中的选择标记

将按照上文实施例3转化但未进行选择的选择的拟南芥植株的种子置于含0、5、或10 μg/ml DAS的0.1% 琼脂糖培养基中。在生长间于22℃温育2周后（16小时光照，8小时黑暗），由DAS板转种较大的、未受阻碍的植株，并由对照板转种相应数目。

生长2周后，对由5mg/ml板转种的经转化拟南芥植株的叶按照实施例6测定酶活性，11株未受阻碍的植物中有11株显示酶活性，而在同一实验中，10株未受DAS选择的植株中有9株为阴性。那株显示酶活性的未选择植株比测定的任一株DAS选择植株的活性低得多。

上文公开的实施方案是例示性的。本发明的此公开书将使本领域技术人员了解本发明的许多变异。所附权利要求意欲涵盖所有这些明显的且可预知的变异。

SEQUENCE LISTING

<110> Novartis AG

<120> Transgenic Plant and Methods

<130> S-30884A/RTP 2107

<140>

<141>

<150> US 09/282,995

<151> 1999-03-31

<150> US 09/502,852

<151> 2000-02-11

<160> 11

<170> PatentIn Ver. 2.0

<210> 1

<211> 1403

<212> DNA

<213> *Fusarium sporotrichioides*

<400> 1

```

atcaaaatgg ccgcaacaag cagcacaagc agccagtctt ttgacataga gctcgacatc 60
atcggccagc aaccgcctct tctttcaatc tacaccaga tcagtctcgt ttaccccgtc 120
tctgatccct cccagtatcc caccatcgtc agcacccttg aggaaggcct aaaacgcctc 180
tctcaaacct tcccatgggt cgcgggccag gtcaagaccg agggcatcag cgaaggaaac 240
acaggaactt ccaagatcat tccatatgag gagacacccc gtcttgtggt gaaagacctc 300
cgtgatgatt cctcagcgcc aacgatcgag gggttgagaa aggcgggttt ccccttagag 360
atgtttgacg agaacgtcgt cgctccgagg aagacattag ctatcggacc tggcaatggc 420
cccaacgacc cgaagcctgt gttgctattg cagctcaact tcattaaggg cggactcatt 480
ctcaccgtca acggacaaca tggtgctatg gacatgacag gacaagatgc aattattcgt 540
cttctctcca aggcgtgccc caacgaatca ttcaccgagg aggaaatctc ggccatgaac 600
ctcgatcgca agacggtagt ccctctcctt gaaaactaca aagttggtcc tgagctagac 660
caccagatcg ccaaacctgc gcctgctggc gacgctccac ccgaccggc caaggcaagc 720
tgggcttctt tttcattcac tccaaggcc ctctcggagc tgaaagacgc agccacaaag 780
actcttgacg cgctcgccaa gtttgtgtca actgatgatg ctctttcggc gtttatctgg 840
caatcaacct cgcgctacg tctcgcaaga ttggatgctt ccacacctac tgaattctgc 900
cgcgctgtcg acatgctggg cccaatgggc gtatcaagca cataccagc ccttcttcaa 960
aacatgacct accatgactc gaccgtcgcc gaaatcgcca acgaaccact tggcgcaaca 1020
gcatcacgcc tgcgctcgga actcaacagt gatcgtttgc gcagacgaac acaagctttg 1080
gcgacgtaca tgcattggcct gcctgacaag tcgagcgtct ccctgaccgc cgatgccaat 1140
ccgtcaagca gcatcatgct gagttcctgg gccaaaggtg gatgctggga gtatgacttt 1200
gggtttggac tgggtaagcc tgagagtgtg agaagacctc gctttgaacc ttttgagagt 1260
ttgatgtact ttatgccc aaagcctgat ggggagtta cggcgtccat ttctctgagg 1320
gatgaggata tggagagact aaaggcggat gaggagtgga caaagtacgc aaagtatatt 1380
gggtagatag tttactagac tac 1403

```

<210> 2

<211> 459

<212> PRT

<213> *Fusarium sporotrichioides*

<400> 2

```

Met Ala Ala Thr Ser Ser Thr Ser Ser Gln Ser Phe Asp Ile Glu Leu
  1                               5           10           15

```

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

	340		345		350										
Arg	Arg	Thr	Gln	Ala	Leu	Ala	Thr	Tyr	Met	His	Gly	Leu	Pro	Asp	Lys
		355					360					365			
Ser	Ser	Val	Ser	Leu	Thr	Ala	Asp	Ala	Asn	Pro	Ser	Ser	Ser	Ile	Met
	370					375					380				
Leu	Ser	Ser	Trp	Ala	Lys	Val	Gly	Cys	Trp	Glu	Tyr	Asp	Phe	Gly	Phe
385					390					395					400
Gly	Leu	Gly	Lys	Pro	Glu	Ser	Val	Arg	Arg	Pro	Arg	Phe	Glu	Pro	Phe
				405					410					415	
Glu	Ser	Leu	Met	Tyr	Phe	Met	Pro	Lys	Lys	Pro	Asp	Gly	Glu	Phe	Thr
			420					425					430		
Ala	Ser	Ile	Ser	Leu	Arg	Asp	Glu	Asp	Met	Glu	Arg	Leu	Lys	Ala	Asp
		435					440					445			
Glu	Glu	Trp	Thr	Lys	Tyr	Ala	Lys	Tyr	Ile	Gly					
	450					455									

<210> 3

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: DNA Primer

<400> 3

acgaatcatt caccgaggag

20

<210> 4

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: DNA Primer

<400> 4

ctcacactct caggcttacc

20

<210> 5

<211> 1356

<212> DNA

<213> Fusarium graminearum

<400> 5

```

atggctttca agatacagct cgacaccctc ggccagctac caggcctcct ttcgatctac 60
acccaaatca gtctcctcta ccccgctctc gattcctctc aatateccac tattgtcagc 120
accttcgagc aaggtcttaa gcgcttctcc gaagccgtcc catgggtcgc aggccaggtc 180
aaagccgagg gcattagcga gggaaacaca ggaacttctt ttatcgcccc ttttgaggac 240
gttctcgtg ttgtagtga agacctccgc gatgatoctt cagcgcccc gatcgagggt 300
atgagaaagg cgggataccc tatggcgatg tttgacgaga acatcatcgc gccaaaggaag 360
acgttaccta ttggacctgg tactggtocc gacgacccaa agcctgtaat tctattgcag 420
ctcaacttca tcaagggcgg actcactctc actgtcaacg gacagcacgg tgctatggat 480
atggtaggcc aagatgcggt gatccgtcta ctctccaagg cgtgcccgtaa cgaccattc 540

```

```

accgaagagg aaatgacggc catgaacctc gatcgcaaga cgatagttcc ttaccttgaa 600
aactatacga ttggccccga ggtagatcat cagattgtca aagctgatgt agctgggtgg 660
gacgctgttc tcaagccggt cagtgcgaagc tgggcgttct tcacattcag cccaaggcc 720
atgicagagc tcaaggatgc tgctaccaag actccttgacg catcaacaaa gttcgtgtcg 780
actgacgatg ctctttcggc gttcatctgg aaatcggcct ctgcggtgcg tctcgaaga 840
atcgatggct ctgcacctac cgagttctgc cgtgctgttg atgctcgacc ggcaatgggt 900
gtctcgaaca actacccagg ccttcttcaa aacatgacct accacaactc gaccatcggc 960
gaaatcgcca acgagtcact cggcgcaaca gcatcacgcc ttcgttcaga actcgacccc 1020
gcgagcatgc gccagcgaac aagaggtctc gcgacgtacc tgcacaacaa ccccgacaag 1080
tccaacgtat ccctgacggc tgatgcggac ccatctacca gcgtcatgct gagttcttgg 1140
gccaaggtgg gactctggga ttacgacttt gggctcggac tgggtaagcc cgagactgtg 1200
agacggccaa tctttgagcc tgttgagagc ttgatgtact ttatgcccaa gaagcctgat 1260
ggcgagtctt gtgcggcgct ttctctgagg gatgaggata tggaccgatt gaaggcggat 1320
aaggagtgga ccaagtatgc gcagtacgtt ggttag 1356

```

<210> 6

<211> 451

<212> PRT

<213> *Fusarium graminearum*

<400> 6

```

Met Ala Phe Lys Ile Gln Leu Asp Thr Leu Gly Gln Leu Pro Gly Leu
  1           5           10           15
Leu Ser Ile Tyr Thr Gln Ile Ser Leu Leu Tyr Pro Val Ser Asp Ser
           20           25           30
Ser Gln Tyr Pro Thr Ile Val Ser Thr Phe Glu Gln Gly Leu Lys Arg
           35           40           45
Phe Ser Glu Ala Val Pro Trp Val Ala Gly Gln Val Lys Ala Glu Gly
           50           55           60
Ile Ser Glu Gly Asn Thr Gly Thr Ser Phe Ile Val Pro Phe Glu Asp
           65           70           75           80
Val Pro Arg Val Val Val Lys Asp Leu Arg Asp Asp Pro Ser Ala Pro
           85           90           95
Thr Ile Glu Gly Met Arg Lys Ala Gly Tyr Pro Met Ala Met Phe Asp
           100          105          110
Glu Asn Ile Ile Ala Pro Arg Lys Thr Leu Pro Ile Gly Pro Gly Thr
           115          120          125
Gly Pro Asp Asp Pro Lys Pro Val Ile Leu Leu Gln Leu Asn Phe Ile
           130          135          140
Lys Gly Gly Leu Ile Leu Thr Val Asn Gly Gln His Gly Ala Met Asp
           145          150          155          160
Met Val Gly Gln Asp Ala Val Ile Arg Leu Leu Ser Lys Ala Cys Arg
           165          170          175
Asn Asp Pro Phe Thr Glu Glu Glu Met Thr Ala Met Asn Leu Asp Arg
           180          185          190
Lys Thr Ile Val Pro Tyr Leu Glu Asn Tyr Thr Ile Gly Pro Glu Val
           195          200          205
Asp His Gln Ile Val Lys Ala Asp Val Ala Gly Gly Asp Ala Val Leu

```

210	215	220
Thr Pro Val Ser Ala 225	Ser Trp Ala Phe Phe 230	Thr Phe Ser Pro Lys Ala 235 240
Met Ser Glu Leu Lys 245	Asp Ala Ala Thr Lys 250	Thr Leu Asp Ala Ser Thr 255
Lys Phe Val Ser Thr 260	Asp Asp Ala Leu Ser 265	Ala Phe Ile Trp Lys Ser 270
Ala Ser Arg Val Arg 275	Leu Glu Arg Ile Asp 280	Gly Ser Ala Pro Thr Glu 285
Phe Cys Arg Ala Val 290	Asp Ala Arg Pro Ala 295	Met Gly Val Ser Asn Asn 300
Tyr Pro Gly Leu Leu 305	Gln Asn Met Thr Tyr 310	His Asn Ser Thr Ile Gly 315 320
Glu Ile Ala Asn Glu 325	Ser Leu Gly Ala Thr 330	Ala Ser Arg Leu Arg Ser 335
Glu Leu Asp Pro Ala 340	Ser Met Arg Gln Arg 345	Thr Arg Gly Leu Ala Thr 350
Tyr Leu His Asn Asn 355	Pro Asp Lys Ser Asn 360	Val Ser Leu Thr Ala Asp 365
Ala Asp Pro Ser Thr 370	Ser Val Met Leu Ser 375	Ser Trp Ala Lys Val Gly 380
Leu Trp Asp Tyr Asp 385	Phe Gly Leu Gly Leu 390	Gly Lys Pro Glu Thr Val 395 400
Arg Arg Pro Ile Phe 405	Glu Pro Val Glu Ser 410	Leu Met Tyr Phe Met Pro 415
Lys Lys Pro Asp Gly 420	Glu Phe Cys Ala Ala 425	Leu Ser Leu Arg Asp Glu 430
Asp Met Asp Arg Leu 435	Lys Ala Asp Lys Glu 440	Trp Thr Lys Tyr Ala Gln 445
Tyr Val Gly 450		

<210> 7

<211> 1425

<212> DNA

<213> *Saccharomyces cerevisiae*

<400> 7

```

atgtttagag tcaagatcat ctctcagaaa cgtacaaaaa gtgtacagat gctagaaaac 60
gatcaacttg atattttggg acaacaacct tcgctataca aactatacac tcaaataatgc 120
tctatctacc gtgtaccaga tccttctgct catgaccata tcgtaaatac cttacaaga 180
ggacttgaaa cattggctaa aaatttccag tggctagcag gaaatgctgt aatgaaggt 240
gctgacgaag gtaacactgg tacctacaga attgtcccgt cagacaaaat tccacttatc 300
gtccaagatc ttcgagaaga tctgtctgcc ccaacaatgg attcgcttga aaaagctgac 360
tttctatct acatgttaga cgaaaagact tttgcgctt gcatgactat caatccacct 420

```

```

ggaacacta taggtatggc cgccaagagt gggcctgtat ttgcagttca agcaaacttt 480
atctccggcg gcctcgtctt aactattgtc gggcagcaca atattatgga tataacagga 540
caggaaagta tcatcaactt gctcaataaa tcttgccacc aaaaaccttt ctctgatgaa 600
gaactgctca ttggaaatat agataaaaagc aaatctattc ctttgtttga tgaacttgg 660
gaacccgaca ccacgctagt tcatgaaata gtggaaacct ctagaatac aagtggagag 720
gaaaaggaac agtcttgctt ttcgaactct acttgggctt atgttgaatt ttctgctatc 780
tcattgcaga atctgaggat tttggcaatg cagacatgta cttctggcac aaaatttgtc 840
tccactgatg atatcgtcac tgctttcatc tggaaatcag tttctcgagc ccgtttatct 900
cgacttaaac cagaaacgaa atcaaattta gggcgtgctg tggatgtag aaaacggcta 960
ggactccccg aaacgtatcc agggttatta gtcaacatga cctttaatac aggttccctg 1020
aaaagcttgg atcataaaaag tttgggcggt cttgcatcac agattcgagc gaagctagac 1080
cctaaagtct tcgatttggc ctataataca tgcgcacttg ctacgctcct tagccgatgc 1140
ccggacaaga ctaagggttc tatacctcaa ccaattgata ctttatctgg aattatggtc 1200
agttcgtggg caaaagtcag cctgtatgac gttgatttca atctagggct tgggaagccc 1260
aagagtgtac gacggccgcg cttcatttcc cttgagagcc taatatattt tatgcctaga 1320
tcctccagag gtgaaatggt gtttgctcct tgccttagag ataaagattg ggagtgacct 1380
aatgcggata aagaatggac aaattatgct acacatatag gatga 1425

```

<210> 8

<211> 474

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

Met Phe Arg Val Lys Ile Ile Ser Gln Lys Arg Thr Lys Ser Val Gln
1 5 10 15

Met Leu Glu Asn Asp Gln Leu Asp Ile Leu Gly Gln Gln Pro Ser Leu
20 25 30

Tyr Lys Leu Tyr Thr Gln Ile Cys Ser Ile Tyr Arg Val Pro Asp Pro
35 40 45

Ser Ala His Asp His Ile Val Asn Thr Leu Thr Arg Gly Leu Glu Thr
50 55 60

Leu Ala Lys Asn Phe Gln Trp Leu Ala Gly Asn Val Val Asn Glu Gly
65 70 75 80

Ala Asp Glu Gly Asn Thr Gly Thr Tyr Arg Ile Val Pro Ser Asp Lys
85 90 95

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

Met Asp Ser Leu Glu Lys Ala Asp Phe Pro Ile Tyr Met Leu Asp Glu
115 120 125

Lys Thr Phe Ala Pro Cys Met Thr Ile Asn Pro Pro Gly Asn Thr Ile
130 135 140

Gly Met Ala Ala Lys Ser Gly Pro Val Phe Ala Val Gln Ala Asn Phe
145 150 155 160

Ile Ser Gly Gly Leu Val Leu Thr Ile Val Gly Gln His Asn Ile Met
165 170 175

Asp Ile Thr Gly Gln Glu Ser Ile Ile Asn Leu Leu Asn Lys Ser Cys
180 185 190

His Gln Lys Pro Phe Ser Asp Glu Glu Leu Leu Ile Gly Asn Ile Asp

	195		200		205														
Lys	Ser	Lys	Ser	Ile	Pro	Leu	Phe	Asp	Glu	Thr	Trp	Glu	Pro	Asp	Thr				
	210					215					220								
Thr	Leu	Val	His	Glu	Ile	Val	Glu	Thr	Ser	Arg	Asn	Thr	Ser	Gly	Glu				
	225				230					235					240				
Glu	Lys	Glu	Gln	Ser	Cys	Ser	Ser	Asn	Ser	Thr	Trp	Ala	Tyr	Val	Glu				
				245					250					255					
Phe	Ser	Ala	Ile	Ser	Leu	Gln	Asn	Leu	Arg	Ile	Leu	Ala	Met	Gln	Thr				
			260					265					270						
Cys	Thr	Ser	Gly	Thr	Lys	Phe	Val	Ser	Thr	Asp	Asp	Ile	Val	Thr	Ala				
		275					280					285							
Phe	Ile	Trp	Lys	Ser	Val	Ser	Arg	Ala	Arg	Leu	Ser	Arg	Leu	Lys	Pro				
	290					295					300								
Glu	Thr	Lys	Ser	Asn	Leu	Gly	Arg	Ala	Val	Asp	Val	Arg	Lys	Arg	Leu				
	305				310					315									
Gly	Leu	Pro	Glu	Thr	Tyr	Pro	Gly	Leu	Leu	Val	Asn	Met	Thr	Phe	Asn				
				325					330					335					
Thr	Gly	Ser	Leu	Lys	Ser	Leu	Asp	His	Lys	Ser	Leu	Gly	Val	Leu	Ala				
			340				345						350						
Ser	Gln	Ile	Arg	Arg	Lys	Leu	Asp	Pro	Lys	Val	Phe	Asp	Leu	Ala	Tyr				
		355					360					365							
Asn	Thr	Cys	Ala	Leu	Ala	Thr	Leu	Leu	Ser	Arg	Cys	Pro	Asp	Lys	Thr				
	370					375					380								
Lys	Val	Ser	Ile	Pro	Gln	Pro	Ile	Asp	Thr	Leu	Ser	Gly	Ile	Met	Val				
	385				390					395					400				
Ser	Ser	Trp	Ala	Lys	Val	Ser	Leu	Tyr	Asp	Val	Asp	Phe	Asn	Leu	Gly				
				405					410					415					
Leu	Gly	Lys	Pro	Lys	Ser	Val	Arg	Arg	Pro	Arg	Phe	Ile	Ser	Leu	Glu				
			420					425					430						
Ser	Leu	Ile	Tyr	Phe	Met	Pro	Arg	Ser	Ser	Arg	Gly	Glu	Met	Val	Val				
		435					440					445							
Ala	Leu	Cys	Leu	Arg	Asp	Lys	Asp	Trp	Glu	Cys	Leu	Asn	Ala	Asp	Lys				
	450					455					460								
Glu	Trp	Thr	Asn	Tyr	Ala	Thr	His	Ile	Gly										
	465				470														

<210> 9

<211> 6111

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Plasmid pCIB9818, a 6111 base

pair circular plasmid comprising the phosphate mannose isomerase
selectable marker (base pair 2090 to 3193)

<400> 9

```

aagcttgcac gacctgcagt cagcgtgacc cggtcgtgcc cctctctaga gataatgagc 60
attgcatgtc taagttataa aaaattacca catatttttt ttgtcacact tgtttgaagt 120
gcagtttata tatctttata catatattta aactttactc tacgaataat ataacttata 180
gtactacaat aatatcagt ttttagagaa tcatataaat gaacagttag acatgggtcta 240
aaggacaatt gagtattttg acaacaggac tctacagttt tatcttttta gtgtgcatgt 300
gttctccttt ttttttgcaa atagcttcac ctatataata ctatcatccat tttattagta 360
catccattta ggggttaggg ttaatggttt ttatagacta attttttttag tacatctatt 420
ttattctatt ttagcctcta aattaagaaa actaaaactc tatttttagtt tttttattta 480
ataatttaga tataaaatag aataaaataa agtgactaaa aattaaaca atacccttta 540
agaaattaaa aaaactaagg aacatttttt ctgttttcga gtagataatg ccagcctggt 600
aaacgccgtc gacgagtcta acggacacca accagcgaac cagcagcgtc gcgtcggggc 660
aagcgaagca gacggcacgg catctctgtc gctgcctctg gacctctc gagagttccg 720
ctccaccgtt ggaacttgct cgctgtcggc atccagaaat tgctggtggc agcggcagac 780
gtgagccggc acggcagggc gcctcctcct cctctcacgg caccggcagc tacgggggat 840
tcctttccca ccgctccttc gctttccctt cctcggccgc cgtaataaat agacaccccc 900
tccacaccct ctttccccaa cctcgtggtg ttcggagcgc acacacacac aaccagatct 960
cccccaaatc caccgctcgg cacctccgct tcaaggtagc ccgctcgtcc tcccccccc 1020
ccctctctca ccttctctag atcggcggtc cgggtccatgg ttagggcccg gtagttctac 1080
ttctgttcat gttgtgta gatccgtgtt tgtgttagat ccggtgctgt agcgttcgta 1140
cacggatgcg acctgtacgt cagacacggt ctgattgcta acttgccagt gtttctcttt 1200
ggggaatcct gggatggctc tagccgttcc gcagacggga tcgatttcat gatttttttt 1260
gtttcgttgc atagggtttg gtttgccctt tccttttatt tcaatatatg ccgtgcactt 1320
gtttgtcggg tcatcttttc atgctttttt ttgtcttggg tgtgatgatg tggctcgggt 1380
gggcggtcgt tctagatcgg agtagaattc tgtttcaaac tacctggtgg atttattaat 1440
tttggatctg tatgtgtgtg ccatacatat tcatagttac gaattgaaga tgatggatgg 1500
aaatatcgat ctaggatagg tatacatggt gatgcccgggt ttactgatgc atatacagag 1560
atgctttttg ttcgcttggg tgtgatgatg tgggtgtggt gggcggtcgt tcatctgttc 1620
tagatcggag tagaatactg ttcaaacta cctgggtgat ttattaattt tggaaactgta 1680
tgtgtgtgtc atacatcttc atagttacga gtttaagatg gatggaataa tcgatctagg 1740
ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcagcat 1800
ctattcatat gctctaacct tgagtaccta tctattataa taaacaagta tgttttataa 1860
ttattttgat cttagatatac ttggatgatg gcataatgcag cagctatatg tggatttttt 1920
tagccctgcc ttcatacgtt atttatttgc ttggtactgt ttcttttgtc gatgctcacc 1980
ctgttgtttg gtgttacttc tgcagggatc cccgatcatg caaaaactca ttaactcagt 2040
gcaaaactat gcctggggca gcaaaacggc gttgactgaa ctttatggta tggaaaatcc 2100
gtccagccag ccgatggccg agctgtggat gggcgcacat ccgaaaagca gttcacagat 2160
gcagaatgcc gccggagata tcgtttccact gcgtgatgtg attgagagtg ataaatcgac 2220
tctgctcggg gaggccgttg ccaaacgctt tggcgaactg cctttcctgt tcaaagtatt 2280
atgcccagca cagccactct ccattcaggt tcatccaaac aaacacaatt ctgaaatcgg 2340
ttttgccaaa gaaaatgccg caggtatccc gatggatgcc gccgagcgtg actataaaga 2400
tctaaccac aagccggagc tggtttttgc gctgacgcct ttctttgca tgaacgcgtt 2460
tcgtgaattt tccgagattg tctccctact ccagccggtc gcaggtgcac atccggcgat 2520
tgctcacttt ttacaacagc ctgatgccga acgtttaagc gaactgttcg ccagcctggt 2580
gaatatgcag ggtgaagaaa aatcccgcgc gctggcgatt ttaaaatcgg ccctcgatag 2640
ccagcagggg gaaccgtggc aaacgattcg ttttaattct gaattttacc cggaaagacg 2700
cggctctgtt tccccgtat tgctgaatgt tttgaaattg aacctggcg aagcagattt 2760
cctgttcgct gaaacaccgc acgcttaact gcaaggcgtg gcgctggaag tgatggcaa 2820
ctccgataac gtgctgcgtg cgggtctgac gcctaaatac attgatattc cggaaactgg 2880
tgccaatgtg aaattcgaag ccaaaccggc taaccagttg ttgaccagc cggtgaaaca 2940
aggtgcagaa ctggacttcc cgattccagt ggatgatttt gccttctcgc tgcatgacct 3000
tagtgataaa gaaaccacca ttagccagca gactgcccgc attttgttct gcgtcgaagg 3060
cgatgcaacc ttgtggaag gttctcagca gttacagctt aaaccgggtg aatcagcgtt 3120
tattgcccgc aacgaatcac cgggtgactg caaaggccac ggccgttttag cgcgtgttta 3180
caacaagctg taagagctta ctgaaaaaat taacatctct tgctaagctg ggagctctag 3240
atctgttctg cacaaagtgg agtagtcagt catcgatcag gaaccagaca ccagactttt 3300
attcatacag tgaagtgaag tgaagtgcag tgcagtgagt tgctggtttt tgtacaactt 3360
agtatgtatt tgtatttcta aaatacttct atcaataaaa tttctaattc ctaaaacca 3420

```

```

aatccagtgg gtaccgaatt cactggccgt cgttttaciaa cgtcgtgact gggaaaaccc 3480
tggcgttacc caacttaatc gccttgacgc acatccccct ttcgccagct ggcgtaatag 3540
cgaagaggcc cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggcg 3600
cctgatgcgg tattttctcc ttacgcatct gtgcggtatt tcacaccgca tatggtgcac 3660
tctcagtaca atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc 3720
cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gtttacagac aagctgtgac 3780
cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcgagacg 3840
aaagggcctc gtgatacgcc tatttttata ggtaaatgtc atgataataa tggtttctta 3900
gacgtcaggt ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta 3960
aatacattca aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaatggcg 4020
cgccgcggcc gcttaagaat attgaaaaag gaagagtatg agtattcaac atttccgtgt 4080
cgcccttatt cctttttttg cggcattttg ccttccgtgt tttgctcacc cagaaaacgt 4140
ggtgaaagta aaagatgctg aagatcagtt ggggtgcacga gtgggttaca tcgaaactgga 4200
tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag 4260
cacttttaaa gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca 4320
actcggtcgc cgcatacact attctcagaa tgacttggtt gactactcac cagtcacaga 4380
aaagcatcct acggatggca tgacagtaag agaattatgc agtgctgcca taacctatgag 4440
tgataacact gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc 4500
ttttttgcac aacatggggg atcatgtaac tcgcttgatc cgttggggaa cggagctgaa 4560
tgaagccata ccaaacgcgc agcgtgacac cacgatgctt gtagcaatgg caacaacgtt 4620
gcgcaacta ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg 4680
gatggaggcg gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctgggt 4740
tattgctgat aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg 4800
gccagatggt aagccctccc gtatcgtagt tatctacacg acggggagtc aggcaactat 4860
ggatgaacga aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact 4920
gtcagaccaa gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa 4980
aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt 5040
ttcgttccac tgagcgtcag acccctgaga aaagatcaaa ggatcttctt gagatccttt 5100
ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg 5160
tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca 5220
gataccaaat actgtccttc tagttagacc gtagttagc caccacttca agaactctgt 5280
agcaccgcct acatacctcg ctctgctaata cctgttacca gtggctgctg ccagtgccga 5340
taagtcgtgt cttaccgggt tggactcaag acgatagtta ccggataagg cgcagcggtc 5400
gggtgaacg ggggttcgt gcacacagcc cagcttgag cgaacgacct acaccgaaact 5460
gagataccta cagcgtgagc tatgagaaaag gcaccagctt cccgaaggga gaaaggcgga 5520
caggtatccg gtaagcggca gggtcggaac aggagagcgc acgagggagc ttccaggggg 5580
aaacgcctgg tatctttata gtctgtcgg gtttcgccac ctctgacttg agcgtcatt 5640
tttgtgatgc tcgtcagggg ggccggagcct atggaaaaac gccagcaacg cggccttttt 5700
acggttctcg gccttttgct ggctttttgc tcacatgttc tttcctgctg tatcccctga 5760
ttctgtggat aaccgtatta ccgcctttga gtgagctgat accgctcgcc gcagccgaac 5820
gaccgagcgc agcgagttag tgagcgagga agcggaaagag cttaaagcggc cgcggcgcg 5880
cgccaatac gcaaacggcc tctccccgcg cgttggccga ttcatatag cagctggcac 5940
gacaggtttc ccgactggaa agcgggcagc gagcgcaacg caattaatgt gagttagctc 6000
actcattagg caccaccggc tttacacttt atgcttccgg ctctgatgtt gtgtggaatt 6060
gtgagcggat aacaatttca cacaggaaac agctatgacc atgattaccg c 6111

```

<210> 10

<211> 13737

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Plasmid pAgroTRIr, a 13,737 base pair circular binary vector comprising a selectable marker and the polynucleotide region of SEQ ID NO: 1

<400> 10

```

gatccagaat tcgtgatcaa atggccgcaa caagcagcac aagcagccag tcttttgaca 60
tagagctcga catcatcggc cagcaaccgc ctcttctttc aatctacacc cagatcagtc 120
tcgtttacct cgtctctgat cctcccagc atccccacat cgtcagcacc cttgaggaag 180
gcctaaaacg cctctctcaa accttcccat gggctcgggg ccaggtcaag accgagggca 240

```

tcagcgaagg	aaacacagga	acttccaaga	tcattccata	tgaggagaca	ccccgtcttg	300
tggtgaaaga	cctccgtgat	gattcctcag	cgccaacgat	cgaggggttg	agaaaggcgg	360
gtttcccctt	agagatgttt	gacgagaacg	tcgtcgtccc	gaggaagaca	ttagctatcg	420
gacctggcaa	tggccccaac	gacccgaagc	ctgtgttgct	attgcagctc	aacttcatta	480
agggcggact	cattctcacc	gtcaacggac	aacatgggtgc	tatggacatg	acaggacaag	540
atgcaattat	tcgtcttctc	tccaaggcgt	gccgcaacga	atcattcacc	gaggaggaaa	600
tctcggccat	gaacctcgat	cgcaagacgg	tagtccctct	ccttgaaaac	tacaaagttg	660
gtcctgagct	agaccaccag	atcgccaaac	ctgcgcctgc	tggcgacgct	ccaccgcac	720
cggccaaggc	aagctgggcg	ttcttttcat	tcactcccaa	ggccctctcg	gagctgaaag	780
acgcagccac	aaagactctt	gacgcgtcgt	ccaagtttgt	gtcaactgat	gatgctcttt	840
cggcgtttat	ctggcaatca	acctcgcgcg	tacgtctcgc	aagattggat	gcttccacac	900
ctactgaatt	ctgccgcgct	gtcgacatgc	ggggcccaat	gggcgtatca	agcacatacc	960
caggccttct	tcaaaacatg	acctaccatg	actcgaccgt	cgccgaaatc	gccaacgaac	1020
cactttggcgc	aacagcatca	cgctcgcgct	cggaactcaa	cagtgatcgt	ttgcgcgac	1080
gaacacaagc	tttggcgcgc	tacatgcatg	gcctgcctga	caagtcgagc	gtctccctga	1140
ccgccgatgc	gaatccgtca	agcagcatca	tgctgagttc	ctgggccaag	gtgggatgct	1200
gggagtatga	ctttgggttt	ggactgggta	agcctgagag	tgtgagaaga	cctcgctttg	1260
aaccttttga	gagtttgatg	tactttatgc	ccaagaagcc	tgatggggag	ttfacggcgt	1320
ccatttctct	gagggatgag	gatatggaga	gactaaaggc	ggatgaggag	tggacaaagt	1380
acgcaaagta	tattgggtag	atagtttact	agactactgc	agggatatcg	tggatcccc	1440
gaatttcccc	gatcgttcaa	acatttggca	ataaagtttc	ttaagattga	atcctgttgc	1500
cggtcttgcg	atgattatca	tctaatttct	gttgaattac	gttaagcatg	taataattaa	1560
catgtaatgc	atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	cgcaattata	1620
catttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	tatcgcgccg	1680
ggttactctc	atgttactag	atccgggaat	tcggcgccgc	caattgattt	aaatggcgcg	1740
tgcggccaat	tctcgcagcg	ttgcggttct	gtcagttcca	aacgtaaaac	ggcttgctcc	1800
gcgtcatcgg	cgggggtcac	aacgtgactc	ccttaattct	ccgctcatga	tcagattgtc	1860
gtttcccgcc	ttcagtttaa	actatcagtg	tttgacagga	tatattggcg	ggtaaaccta	1920
agagaaaaga	gcgtttatta	gaataatcgg	atatttataa	gggcgtgaaa	aggtttatcc	1980
gttcgtccat	ttgtatgtgc	atgccaacca	cagggttccc	cagatctggc	gccggccagc	2040
gagacgagca	agattggccg	ccgccgaaa	cgatccgaca	gcgcgccccg	cacaggtgcg	2100
caggcaaatt	gcaccaacgc	atacagcgcc	agcagaatgc	catagtgggc	ggtgacgtcg	2160
ttcgagtga	ccagatcgcg	caggaggccc	ggcagcaccg	gcataatcag	gccgatccc	2220
acagcgtcga	gcgcgacagt	gctcagaatt	acgatcaggg	gtatggtggg	ttcacgctct	2280
ggcctccgga	ccagcctccg	ctggtccgat	tgaacgcgcg	gattctttat	cactgataag	2340
ttggtggaca	tattatgttt	atcagtgata	aagtgtcaag	catgacaaag	ttgcagccga	2400
atacagtgat	ccgtgccgcc	ctggacctgt	tgaacgaggt	cggcgtagac	ggtctgacga	2460
cacgcaaact	ggcggaacgg	ttgggggttc	agcagccggc	gctttactgg	cacttcagga	2520
acaagcgggc	gctgctcgac	gcactggccg	aagccatgct	ggcggagaat	catacgcatt	2580
cggtgccgag	agccgacgac	gactggcgct	catttctgat	cgggaatgcc	cgcagcttca	2640
ggcagcgct	gctcgcctac	cgcgatggcg	cgcgatcca	tgccggcacg	cgaccggcgg	2700
caccgcagat	ggaaacggcc	gacgcgcagc	ttcgcttctc	ctgcgaggcg	ggttttctcg	2760
ccggggacgc	cgtcaatgcg	ctgatgacaa	tcagctactt	cactgttggg	ccgctgcttg	2820
aggagcaggc	cggcgacagc	gatgccggcg	agcgcggcgg	caccgttgaa	caggctccgc	2880
tctcgcgcgt	gttgccggcc	gcgatagacg	ccttcgacga	agccggctcg	gacgcagcgt	2940
tcgagcaggg	actcgcggtg	attgtcgatg	gattggcgaa	aaggaggctc	gttgtcagga	3000
acgttgaagg	accgagaaag	ggtgacgatt	gatcaggacc	gctgccggag	cgcaaccac	3060
tcactacagc	agagccatgt	agacaacatc	ccctccccct	ttccaccgcg	tcagacgccc	3120
gtagcagccc	gctacgggct	ttttcatgce	ctgccctage	gtccaagcct	cacggccgcg	3180
ctcgccctct	ctggcgccct	tctggcgctc	ttccgcttcc	tcgctcactg	actcgttgcg	3240
ctcggctcgtt	cggctgcggc	gagcggtatc	agctcactca	aaggcggtaa	tacggttatc	3300
cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	3360
gaaccgtaaa	aaggccgcgt	tgctggcgct	tttccatagg	ctccgcccc	ctgacgagca	3420
tcacaaaaat	cgacgctcaa	gtcagaggtg	gcgaaaccgg	acaggactat	aaagatacca	3480
ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctggt	ccgacctgct	cgcttaccgg	3540
atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	ttccgctgca	taacctgct	3600
tcggggctcat	tatagcgatt	ttttcggtat	atccatcctt	tttcgcacga	tatacaggat	3660
tttgccaag	ggttcgtgta	gactttcctt	ggtgtatcca	acggcgtcag	ccgggcagga	3720
taggtgaagt	aggcccaccc	cgcagcgggt	gttccttctt	cactgtccct	tatcgcacc	3780
tggcggtgct	caacgggaat	ctcgtctgct	gaggtggcc	ggctaccgcc	ggcgtaacag	3840
atgagggcaa	gcggatggct	gatgaaacca	agccaaccag	gaagggcagc	ccacctatca	3900

agggtgactg	ccttccagac	gaacgaagag	cgattgagga	aaaggcggcg	gcgccggca	3960
tgagcctgtc	ggcctacctg	ctggccgtcg	gccagggtca	caaaatcacg	ggcgtcgtgg	4020
actatgagca	cgtccgcgag	ctggccccga	tcaatggcga	cctggggccgc	ctggggcgcc	4080
tgctgaaact	ctggctcacc	gacgaccgcg	gcacggcgcg	gttcggtgat	gccacgatcc	4140
tcgccctgct	ggcgaagatc	gaagagaagc	aggacgagct	tggcaaggct	atgatggcg	4200
tggtccgccc	gagggcagag	ccatgacttt	tttagccgct	aaaacggccg	gggggtgcg	4260
gtgattgcca	agcacgtccc	catgcgctcc	atcaagaaga	gcgacttcgc	ggagctggtg	4320
aagtacatca	ccgacgagca	aggcaagacc	gagcgccttt	gcgacgctca	ccgggctggt	4380
tgccctcgcc	gctgggctgg	cggccgtcta	tggccctgca	aacgcgccag	aaacgccgtc	4440
gaagccgtgt	gcgagacacc	gcgcccgcc	gccggcgttg	tggatacctc	gcggaaaact	4500
tggccctcac	tgacagatga	ggggcggacg	ttgacacttg	aggggcccag	tcaccggcg	4560
cggcgttgac	agatgagggg	caggctcgat	ttcggccggc	gacgtggagc	tggccagcct	4620
cgcaataag	cgaaaacgcc	tgattttacg	cgagtttccc	acagatgatg	tggacaagcc	4680
tggggataag	tgccctgcgg	tattgacact	tgagggcgcg	gactactgac	agatgagggg	4740
cgcgatcctt	gacacttgag	gggcagagtg	ctgacagatg	aggggcgcac	ctattgacat	4800
ttgaggggct	gtccacaggc	agaaaatcca	gcatttgcaa	gggtttccgc	ccgtttttcg	4860
gccaccgcta	acctgtcttt	taacctgctt	ttaaaccaat	atztataaac	ctgtttttta	4920
accagggctg	cgccctgtgc	gcgtgaccgc	gcacgccgaa	ggggggtgcc	cccccttctc	4980
gaaccctccc	ggcccgtata	cgcgggctc	ccatccccc	aggggctgcg	ccctcggcc	5040
gcgaacggcc	tcaccccaaa	aatggcagcg	ctggcagctc	ttgccattgc	cgggatcggg	5100
gcagtaacgg	gatgggcgat	cagcccagc	gcgacgcccg	gaagcattga	cgtgccgcag	5160
gtgctggcat	cgacattcag	cgaccaggtg	cgggcagtg	agggcgccg	cctgggtgcc	5220
ggcctgccct	tcacttcggc	cgtcggggca	ttcacggact	tcatggcggg	gcccggcaat	5280
ttaccttgg	gcattcttgg	catagtggtc	gcggtgccc	tgctcgtgtt	cgggggtgcg	5340
ataaaccag	cgaaccattt	gaggtgatag	gtaagattat	accgaggtat	gaaaacgaga	5400
attggacctt	tacagaatta	ctctatgaag	cgccatattt	aaaaagctac	caagacgaag	5460
aggatgaaga	ggatgaggag	gcagattgcc	ttgaatatat	tgacaatact	gataagataa	5520
tatatctttt	atatagaaga	tatcgccgta	tgtaaggatt	tcagggggca	aggcataggc	5580
agcgcgctta	tcaatatatc	tatagaatgg	gcaaagcata	aaaacttgca	tggactaatg	5640
cttgaaccca	aggacaataa	ccttatagct	tgtaaattct	atcataattg	ggtaatgact	5700
ccaacttatt	gatagtgttt	tatgttcaga	taatgcccga	tgactttgtc	atgcagctcc	5760
accgattttg	agaacgcag	cgacttccgt	cccagccgtg	ccaggtgctg	cctcagattc	5820
aggttatgcc	gctcaattcg	ctgcgtatat	cgcttgctga	ttacgtgcag	ctttcccttc	5880
aggcgggatt	catacagcgg	ccagccatcc	gtcatccata	tcaccacgtc	aaagggtgac	5940
agcaggctca	taagacgccc	cagcgtcgcc	atagtgcgct	caccgaatac	gtgcgcaaca	6000
accgtcttcc	ggagactgtc	atacgcgtaa	aacagccagc	gctggcgcga	tttagccccg	6060
acatagcccc	actgttcgtc	catttccgcg	cagacgatga	cgctactgcc	cggctgtatg	6120
cgcgaggtta	ccgactgcgg	cctgagtttt	ttaagtgcag	taaaatcgtg	ttgaggccaa	6180
cgcccataat	gcgggctggt	gcccggcatc	caacgccatt	catggccata	tcaatgattt	6240
tctggtgcgt	accgggttga	gaagcgggtg	aagtgaactg	cagttgccat	gttttacggc	6300
agtgagagca	gagatagcgc	tgatgtccgg	cggtgctttt	gcccgttacg	accaccccgt	6360
cagtagctga	acaggagggg	cagctgatag	acacagaagc	cactggagca	cctcaaaaac	6420
accatcatac	actaaatcag	taagttggca	gcatcaccca	taattgtggt	ttcaaaatcg	6480
gctccgtcga	tactatgtta	tacgccaact	ttgaaaacaa	ctttgaaaaa	gctgttttct	6540
ggtatttaag	gtttttagaat	gcaaggaaca	gtgaattgga	gttcgtcttg	ttataattag	6600
cttcttgggg	tatctttaa	tactgtagaa	aagaggaagg	aaataataaa	tggctaaaat	6660
gagaatatca	ccggaattga	aaaaactgat	cgaaaaatac	cgctgcgtaa	aagatacggg	6720
aggaatgtct	cctgctaagg	tatataagct	ggtgggagaa	aatgaaaacc	tatatttaa	6780
aatgacggac	agccggtata	aagggaccac	ctatgatgtg	gaacgggaaa	aggacatgat	6840
gctatggctg	gaaggaaagc	tgctgttcc	aaaggtcctg	cactttgaac	ggcatgatgg	6900
ctggagcaat	ctgctcatga	gtgaggccga	tggcgtcctt	tgctcggaag	agtatgaaga	6960
tgaacaaagc	cctgaaaaga	ttatcgagct	gtatgcccgg	tgcatcaggc	tctttcactc	7020
catcgacata	tcggattgtc	cctatacgaa	tagcttagac	agccgcttag	ccgaattgga	7080
ttacttactg	aataacgatc	tggccgatgt	ggattgcgaa	aactgggaag	aagacactcc	7140
atttaaatg	ccgcgcgagc	tgatgatatt	tttaaagacg	gaaaagcccg	aagaggaact	7200
tgtcttttcc	cacgcgcgac	tgggagacag	caacatcttt	gtgaaagatg	gcaaagtaag	7260
tggctttatt	gatcttggga	gaagcggcag	ggcggacaag	tggtatgaca	tggccttctg	7320
cgtccggtcg	atcagggagg	atatcgggga	agaacagtat	gtcgagctat	ttttgactt	7380
actggggatc	aagcctgatt	gggagaaaa	aaaatattat	attttactgg	atgaattggt	7440
ttagtacct	gatgtggcgc	aacgatgccg	gcgacaagca	ggagcgcacc	gacttcttcc	7500
gcatcaagtg	ttttggctct	caggccgagg	cccacggcaa	gtatttgggc	aaggggtcgc	7560

tggtattcgt	gcagggcaag	attcgaata	ccaagtacga	gaaggacggc	cagacggctct	7620
acgggaccga	cttcattgcc	gataaggtgg	attatctgga	caccaaggca	ccaggcgggt	7680
caaatcagga	ataagggcac	attgccccgg	cgtgagtcgg	ggcaatcccg	caaggagggt	7740
gaatgaatcg	gacgtttgac	cggaaggcat	acaggcaaga	actgatcgac	gcggggtttt	7800
ccgccgagga	tgccgaaacc	atcgcaagcc	gcaccgcat	gcgtgcgccc	cgcgaaacct	7860
tccagtccgt	cggtctgatg	gtccagcaag	ctacggccaa	gatcgagcgc	gacagcgtgc	7920
aactggctcc	cctgcctctg	cccgcgccat	cggccgccgt	ggagcgttcg	cgctgtctcg	7980
aacaggaggc	ggcaggtttg	gcgaagtcga	tgaccatcga	cacgcgagga	actatgacga	8040
ccaagaagcg	aaaaaccgcc	ggcgaggacc	tggcaaaaca	ggtcagcgag	gccaaagcag	8100
ccgcgttgct	gaaacacacg	aagcagcaga	tcaaggaaat	gcagctttcc	ttgttcgata	8160
ttgcgccgtg	gccggacacg	atgcgagcga	tgccaaacga	cacggccccg	tctgccctgt	8220
tcaccacgcy	caacaagaaa	atcccgcgcy	aggcgtgca	aaacaaggtc	attttccacg	8280
tcaacaagga	cgtgaagatc	acctacaccg	gcgtcgagct	gcgggcccag	gatgacgaac	8340
tggtgtggca	gcagggtgtg	gagtacgcga	agcgcacccc	tatcggcgag	ccgatcacct	8400
tcacgttcta	cgagctttgc	caggacctgg	gctggtcgat	caatggcccg	tattacacga	8460
aggccgagga	atgcctgtcg	cgctacagg	cgacggcgat	gggcttcacg	tccgaccgcy	8520
ttgggcacct	ggaatcggty	tcgctgctgc	accgcttccg	cgctctggac	cgtaggcaaga	8580
aaacgtcccg	ttgccaggtc	ctgatcgacg	aggaaatcgt	cgtagctgtt	gctggcgacc	8640
actacacgaa	attcatatgg	gagaagtacc	gcaagctgtc	gccgacggcc	cgacggatgt	8700
tcgactatit	cagctcgcac	cgggagccgt	acccgctcaa	gctggaaacc	ttccgcctca	8760
tgtgcggatc	ggattccacc	cgcgtgaaga	agtggcgcyga	gcaggctcgg	gaagcctgcy	8820
aagagtctcg	aggcagcggc	ctggtggaac	acgcctgggt	caatgatgac	ctggtcatt	8880
gcaaacgcta	gggcttctgt	gggtcagttc	cggctggggg	ttcagcagcc	agcctttac	8940
tgccatttca	ggaacaagcy	ggcactgctc	gacgcacttg	cttcgctcag	tatcgtcgg	9000
gacgcacggc	gcgctctacg	aactgccgat	aaacagagga	ttaaaattga	caattgtgat	9060
taaggctcag	attcgacggc	ttggagcggc	cgacgtgcag	gatttccgcy	agatccgatt	9120
gtcggccctg	aagaaagctc	cagagatggt	cgggtcccgt	tacgagcacy	aggagaaaa	9180
gccccatggag	gcgttcgctg	aacggttcg	agatgccgtg	gcattcggcg	cctacatcga	9240
cggcgagatc	attgggctgt	cggtcttcaa	acaggaggac	ggccccagg	acgctcaaca	9300
ggcgcatctg	tccggcgttt	tcgtggagcc	cgaacagcga	ggccgagggg	tcgcccgtat	9360
gctgctgcyg	gcgttgccgg	cgggtttatt	gctcgtgatg	atcgtccgac	agattccaac	9420
gggaatctgg	tggtatgcgca	tcttcatcct	cggcgcaact	aatatttcgc	tattctggag	9480
cttgttgttt	atttcggctc	accgcctgcc	ggcgggggtc	gcggcgacgg	taggcgctgt	9540
gcagccgctg	atggctcgtg	tcctctctgc	cgctctgcta	ggtagcccga	tacgattgat	9600
ggcggctctg	ggggctattt	gcccgaactgc	ggcgctggcg	ctgttgggtg	tgacaccaaa	9660
cgagcgccta	gatcctgtcg	gcgtcgcagc	gggctggcg	ggggcggttt	ccatggcggt	9720
cggaaaccgtg	ctgaccccga	agtggcaacc	tcccgtgcct	ctgctcacct	ttaccgcctg	9780
gcaactggcg	gcccggaggac	ttctgtcctg	tccagtagct	ttagtgtttg	atccgccaat	9840
cccgatgcct	acaggaacca	atgttctcgg	ctcggcctga	ctcggcctga	tcggagcggg	9900
tttaacctac	ttcctttggg	tccgggggat	ctcgcgactc	gaacctacag	ttgtttcctt	9960
actgggcttt	ctcagcccca	gatctggggg	cgatcagccg	gggatgcac	aggccgacag	10020
tcggaacttc	gggtccccga	cctgtaccat	tcggtagcga	atggataggg	gagttgatat	10080
cgtaacgctt	cacttctaaa	gaaatagcgc	cactcagctt	cctcagcggc	tttatccagc	10140
gatttccat	tatgtcggca	tagttctcaa	gatcgacagc	ctgtcacggg	taagcgagaa	10200
atgaataaga	aggctgataa	ttcggatctc	tgcgagggag	atgatatttg	atcacaggca	10260
gcaacgctct	gtcatcgtta	caatcaacat	gctaccctcc	gcgagatcat	ccgtgtttca	10320
aacccggcag	cttagttgcc	gttcttccga	atagcatcgg	taacatgagc	aaagtctgcc	10380
gccttacaac	ggctctcccc	ctgacgccgt	cccggactga	tgggctgcct	gtatcgagtg	10440
gtgattttgt	gccgagctgc	cggtcgggga	gctgttggct	ggctgggtggc	aggatatatt	10500
gtggtgtaaa	caaattgacg	cttagacaac	ttaataacac	attgcggacg	tttttaattg	10560
actgcggtac	ggccatgctg	gcccggggg	caccggtaaa	tttctgcag	ggctagcgaa	10620
ttcgagctcg	gtaccctctg	attttgggtt	taggaattag	attattgata	gaagtatttt	10680
acaaatacaa	atacatacta	agggtttctt	atatgctcaa	cacatgagcg	aaaccctata	10740
agaaccctaa	ttcccttatc	tgggaactac	tcacacatta	ttatagagag	agatagattt	10800
gtagagagag	actggtgatt	tcagcgggca	tgctcagcag	tcgactcaga	ctggggtaac	10860
tggcctaact	ggccttggag	gagctggcaa	ctcaaaatcc	ctttgccaaa	aaccaacatc	10920
atgccatcca	ccatgcttgt	atccagctgc	gcgcaatgta	ccccgggctg	tgtatcccaa	10980
agcctcatgc	aacctaacag	atggatcgtt	tggaaaggct	ataacagcaa	ccacagactt	11040
aaaaccttgc	gcctccatag	acttaagcaa	atgtgtgtac	aatgtggatc	ctaggcccaa	11100
cctttgatgc	ctatgtgaca	cgtaaacagt	actctcaact	gtccaatcgt	aagcgttctt	11160
agccttccag	ggcccagcgt	aagcaatacc	agccacaaca	ccctcaacct	cagcaacca	11220

```

ccaagggtat ctatcttgca acctctctag atcatcaatc cactcttggtg gtgtttgtgg 11280
ctctgtccta aagttcactg tagacgtctc aatgtaatgg ttaacgatat cacaaaccgc 11340
ggccatatca gctgctgtag ctggccta atcaactggt ctctctccg gagacatgctc 11400
gactctagag gatccccggg taccctgtcc tctccaaatg aaatgaactt ccttatatag 11460
aggaagggtc ttgcaagga tagtgggatt gtgcgtcatc ccttacgtca gtggagatat 11520
cacatcaatc cacttgcttt gaagacgtgg ttggaacgtc ttctttttcc acgatgctcc 11580
tcgtgggtgg gggccatctt ttgggaccac tgtcggcaga ggcatcttca acgatggcct 11640
ttcttttata gcaatgatgg catttgtagg agccaccttc cttttccact atcttcacaa 11700
taaagtgaca gatagctggg caatggaatc cgaggagggt tccggatatt accctttggt 11760
gaaaagtctc aattgccctt tggctctctg agactgtatc tttgatattt ttggagtaga 11820
caagcgtgct gtgctccacc atgttgacga agatattctt ctgtcattg agtcgtaaga 11880
gactctgtat gaactgttcg ccagtcttta cggcgagttc tgttggctct ctatattgat 11940
ctttgactcc atgggaattg agatctctcg aggtttaa ac gggccacgcc tgcggccgcc 12000
tcgaggtagc ggatctggag ccaagtctca taaacgccat tgtggaagaa agtcttgagt 12060
tgctgggcaa caaaaatcct gaacatctta ttttagcaaa gagaaagagt tccgagctcg 12120
tagcagaaga gtgaggagaa atttaagctc ttggacttgt gaattgttcc gcctcttgaa 12240
tacttcttca atcctcatat attcttcttc tatgttacct gaaaaccggc atttaatctc 12300
gcgggtttat tccggttcaa catttttttt gttttgagtt attatctggg cttaataacg 12360
caggcctgaa ataaattcaa ggccaactg ttttttttt taagaagtgg ctgttaaaaa 12420
aaaaaaaaag gaattaacaa caacaacaaa aaaagataaa gaaaataata acaattactt 12480
taattgtaga ctaaaaaaac atagatttta tcatgaaaaa aagagaaaag aaataaaaac 12540
ttgatcaaa aaaaaaaaca tacagatctc ctaattatta acttttctta aaaattaggt 12600
cctttttccc aacaattagg tttagagttt tggaattaaa ccaaaaagat tgttctaaaa 12660
aatactcaaa tttggtagat aagtttcctt attttaatta gtcaatggta gattcttttt 12720
ttctttttct ttattagagt agattagaat cttttatgcc aagttttgat aaattaaatc 12780
agaagataa actatcataa tcaacatgaa attaaaagaa aaatctcata tatagtatta 12840
gtattctcta tatatattat gattgcttat tcttaatggg ttgggttaac caagacatag 12900
tcttaatgga aagaatcttt tttgaacttt ttccttattg attaaattct tctatagaaa 12960
agaaagaaat tatttgagga aaagtatata caaaaagaaa aatagaaaaa tgtcagtgaa 13020
gcagatgtaa tggatgacct aatccaacca ccaccatagg atgtttctac ttgagtcggt 13080
cttttaaaa cgcacggtgg aaaatatgac acgtatcata tgattccttc ctttagtttc 13140
gtgataataa tcctcaactg atatcttctt tttttgttt tggctaaaga tattttatta 13200
tcattaatag aaaagacggt tttgggcttt tggtttgca tataaagaag accttcgtgt 13260
ggaagataa aattcatcct ttcgtctttt tctgactctt caatctctcc caaagcctaa 13320
agcgatctct gcaaactctc cgcgactctc tctttcaagg tatattttct gattcttttt 13380
gtttttgatt cgtatctgat ctccaatttt tgttatgtgg attattgat cttttgtata 13440
aattgctttt gacaatattg ttcgtttctg caatccagct tctaaatttt gtccgtgatta 13500
ctaagatctc gattcgtagt gtttacatct gtgtaatttc ttgcttgatt gtgaaattag 13560
gattttcaag gacgatctat tcaatttttg tgttttcttt gttcgattct ctctgtttta 13620
ggtttcttat gtttagatcc gtttctcttt ggtgttggtt tgatttctct tacggctttt 13680
gatttggtat atgttcgctg attgggttctt acttgttcta ttgttttatt tcaggtg 13737

```

<210> 11

<211> 12949

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Plasmid pNOV1704, a 12949 base pair circular binary vector comprising the phosphate mannose isomerase selectable marker sequence (base 2089 to 3192), and the trichothecene 3-O-acetyl transferase sequence of SEQ ID NO.1 (base 11234 to 12662)

<400> 11

```

agcttgcagt cctgcagctg agcgtgacct ggtcgtgccc ctctctagag ataatgagca 60
ttgcattgct aagttataaa aaattaccac atattttttt tgtcacactt gtttgaagtg 120
cagttttctt atctttatca atatatgtaa actttactct acgaataata taatctatag 180
tactacaata atatcagtggt tttagagaat catataaatg aacagttaga catggctcaa 240
aggacaattg agtattttga caacaggact ctacagtttt atcttttttag tgtgcatgtg 300
ttctcctttt tttttgcaaa tagcttcacc tatataatac ttcattccatt ttattagtag 360

```

atccatttag	ggtttagggt	taatggtttt	tatagactaa	tttttttagt	acatctattt	420
tattctattt	tagcctctaa	attaagaaaa	ctaaaactct	attttagttt	ttttatftaa	480
taatttagat	ataaaataga	ataaaaataaa	gtgactaaaa	atftaaacaaa	taccctftaa	540
gaaattaaaa	aaactaagga	aacattttttc	ttgtttcgag	tagataatgc	cagcctgtta	600
aacgccgtcg	acgagtctaa	cggacaccaa	ccagcgaacc	agcagcgtcg	cgtcgggcca	660
agcgaagcag	acggcacggc	atctctgtcg	ctgcctctgg	accctctctg	agagttccgc	720
tccaccgttg	gacttgctcc	gctgtcggca	tccagaaatt	gcgtggcgga	gcggcagacg	780
tgagccggca	cggcaggcgg	cctcctctct	ctctcacggc	accggcagct	acgggggatt	840
cctttcccac	cgctccttcg	ctttcccttc	ctcgcccgcc	gtaataaata	gacaccctc	900
ccacaccctc	tttccccaac	ctcgtgttgt	tccggagcga	cacacacaca	accagatctc	960
ccccaaatcc	accctgcggc	acctcctctt	caaggtagcg	cgctcgtcct	cccccccccc	1020
ccctctctac	cttctctaga	tccggcgttc	ggccatgggt	tagggcccgg	tagttctact	1080
tctgttcatg	tttgtgttag	atccgtgttt	gtgttagatc	cgctgtgcta	gcgttcgtac	1140
acggatgcga	cctgtacgtc	agacacgttc	tgattgctaa	cttgccagtg	tttctctttg	1200
gggaatcctg	ggatggctct	agccgttccg	cagacgggat	cgatttcatg	atftttttttg	1260
tttcgttgca	tagggttttg	tttgcccttt	tctttatftt	caataatgc	cggtcacttg	1320
ttgtcgggt	catcttttca	tgcttttttt	tgtcttggtt	gtgatgatgt	ggtctggttg	1380
ggcggctgtt	ctagatcggga	gtagaattct	gtttcaaact	acctggtgga	ttftattaatt	1440
ttggatctgt	atgtgtgtgc	catacatatt	catagttacg	aattgaagat	gatggatgga	1500
aatatcgatc	taggataggt	atacatgttg	atgcgggttt	tactgatgca	tatacagaga	1560
tgctttttgt	tcgcttggtt	gtgatgatgt	ggtgtggttg	ggcggctcgtt	cattcgttct	1620
agatcggagt	agaataactgt	ttcaaactac	ctggtgtatt	tattaatftt	ggaactgtat	1680
gtgtgtgtca	tacatcttca	tagttacgag	tttaagatgg	atggaaatat	cgatctagga	1740
taggtataca	tgttgatgtg	ggttttactg	atgcataatac	atgatggcat	atgcagcatc	1800
tattcatatg	ctctaacttt	gagtacctat	ctattataat	aaacaagtat	gttttataat	1860
tattttgatc	ttgatataact	tgatgatagg	catatgcagc	agctatatgt	ggattttttt	1920
agccctgcct	tcatacgccta	tttatttgct	tggtactggt	tcttttgctg	atgctcacc	1980
tgttgtttgg	tgttacttct	gcagggatcc	ccgatcatgc	aaaaactcat	taactcagtg	2040
caaaactatg	cctggggcag	caaaacggcg	ttgactgaac	tttatgggat	ggaaaatccg	2100
tccagccaagc	cgatggccga	gctgtggatg	ggcgcacatc	cgaaaagcag	ttcacgagtg	2160
cagaatgccg	ccggagatat	cgtttcaactg	cgatgatgta	ttgagagtga	taaactcgact	2220
ctgctcggag	aggccgttgc	caaacgcttt	ggcgaactgc	ctttcctggt	caaagtatta	2280
tgcgcagcag	agccactctc	cattcaggtt	catccaaaaca	aacacaattc	tgaaaatcggg	2340
tttgccaaaag	aaaatgccgc	aggtatcccg	atggatgccg	ccgagcgtaa	ctataaagat	2400
cctaaccaca	agccggagct	ggtttttgct	ctgacgcctt	tcttgccgat	gaacgcgttt	2460
cgatgaatftt	ccgagattgt	ctccctactc	cagccggctg	caggtgcaca	tccggcgatt	2520
gctcactftt	tacaacagcc	tgatgccgaa	cgtttaagcg	aactgttcgc	cagcctgttg	2580
aatatgcagg	gtgaagaaaa	atcccgcgcg	ctggcgattt	taaaatccgc	cctcgatagc	2640
cagcaggggtg	aaccgtggca	aacgattcgt	ttaatftctg	aattttacc	ggaagacagc	2700
ggtctgttct	ccccgctatt	gctgaatgtg	gtgaaattga	accctggcga	agcgatgttc	2760
ctgttcgctg	aaacaccgca	cgcttacctg	caaggcgtgg	cgctggaagt	gatggcaaac	2820
tccgataacg	tgctgcgtgc	gggtctgagc	cctaaataca	ttgatattcc	ggaactgggt	2880
gccaatgtga	aattcgaagc	caaaccgctt	aaccagttgt	tgaccagcc	ggtgaaacaa	2940
ggtgcagaac	tggacttccc	gattccagtg	gatgatfttg	ccttctcgct	gcatgacctt	3000
agtataaaag	aaaccacat	tagccagcag	agtgccgcca	ttttgttctg	cgctgaaggc	3060
gatgcaacgt	tgtggaagag	ttctcagcag	ttacagctta	aaccgggtga	atcagcgttt	3120
attgccgcca	acgaatcacc	ggtgactgtc	aaaggccacg	gccgttttagc	gcgtgtttac	3180
acaagctgt	aagagcttac	tgaaaaaatt	aacatctctt	gctaagctgg	gagctcgatc	3240
cgctcagctg	cagatcgttc	aaacatttgg	caataaagtt	tcttaagatt	gaatcctggt	3300
gccggtcttg	cgatgattat	catataatft	ctgttgaatt	acgttaagca	tgtaataatt	3360
aacatgtaat	gcatgacgtt	atftatgaga	tgggtfttta	tgattagagt	cccgaatta	3420
tacattftaat	acgcgataga	aaacaaaata	tagcgcgcaa	actaggataa	attatcgcg	3480
gcggtgtcat	ctatgttact	agatctgcta	gccctgcagg	aaatttaccg	gtgcccgggc	3540
ggccagcatg	gccgatccg	caatgtgtta	ttaaagttgtc	taagcgtcaa	ttgttttaca	3600
ccacaatata	tcttgccacc	agccagccaa	cagctccccg	accggcagct	cggcacaaaa	3660
tcaccactcg	atacaggcag	cccatcagaa	ttaatftctca	tgfttgacag	cttatcatcg	3720
actgcacggg	gcaccaatgc	ttctggcgct	aggcagccat	cggaagctgt	ggtatggctg	3780
tgcaagtcgt	aaatcactgc	ataattcgtg	tcgctcaagg	cgactccccg	ttctgggataa	3840
tgftttttgc	gccgacaatca	taacggttct	ggcaaatatt	ctgaaatgag	ctgttgacaa	3900
ftaatcatcc	ggctcgtata	atgtgtggaa	ttgtgagcgg	ataacaattt	cacacaggaa	3960
acagaccatg	agggaaagcgt	tgatcgcgca	agtatcgact	caactatcag	aggtagttgg	4020

cgtcacgag cyccatctcg aaccgacgtt gctggccgta catttgtag gctccgcagt 4080
 ggatggcggc ctgaagccac acagtgatat tgatttgcct gttacgggta ccgtaaggct 4140
 tgatgaaaca acgcggcgag ctttgatcaa cgaccttttg gaaacttcg cttcccctgg 4200
 agagagcgag attctccgcg ctgtagaagt caccattgtt gtgcacgacg acatcattcc 4260
 gtggcgttat ccagctaagc gcgaactgca atttggagaa tggcagcgca atgacattct 4320
 tgcaggtatc ttcgagccag ccacgatcga cattgatctg gctatcttgc tgacaaaagc 4380
 aagagaacat agcgttgccct tggtaggtcc agcggcggag gaactctttg atccggttcc 4440
 tgaacaggat ctatttgagg cgctaaatga aaccttaacg ctatggaact cgccgcccga 4500
 ctgggctggc gatgagcgaa atgtagtgtc tacgttgtcc cgcatttggg acagcgcagt 4560
 aaccggcaaa atcgcgccga aggatgtcgc tgccgactgg gcaatggagc gctgcccggc 4620
 ccagtatcag cccgtcatac ttgaagctag gcaggcttat ctggacaag aagatcgctt 4680
 ggctcgcgcg gcagatcagt tggagaatc tgttcactac gtgaaaggcg agatcaccaa 4740
 agtagctggc aaataaagct ctagtggatc tccgtacccc cgggggatct ggctcgcggc 4800
 ggacgcacga cgccggggcg agaccatagg cgatctccta aatcaatagt agctgtaacc 4860
 tcgaagcgtt tcaactgtaa caacgattga gaatttttgt cataaaattg aaatacttgg 4920
 ttcgcatttt tgcacccgcg ggtcagccgc aattctgacg aactgcccat ttagctggag 4980
 atgattgtac atccttcacg tgaaaatttc tcaagcgtg tgaacaaggg ttcgattttt 5040
 agattgaaag gtgagccgtt gaaacacgtt cttcttgcgc atgacgacgt cgctatgcgg 5100
 catcttatta ttgaatacct tacgatccac gccttcaaag tgaccgcggg agccgacagc 5160
 acccagttca caagagtact ctcttccgcg acggtcgatg tcgtggttgt tgatctaaat 5220
 ttaggctcgt aagatgggct cgagatcgtt cgtaatctgg cggcaaagtc tgatattcca 5280
 atcataatta tcaagtggca ccgccttgag gcagcggata aagttgttgc actcgagcta 5340
 ggagcaagtg attttatcgc taagccgttc agtatcagag agtttctagc acgcattcgg 5400
 gttgccttgc gcgtgcgccc caacgttgtc cgctccaaag accgacggtc tttttgtttt 5460
 actgactgga cacttaatct caggcaacgt cgcttgatgt ccgaagctgg cggtgagggtg 5520
 aaacttacgg caggtgagtt caatcttctc ctccgctttt tagagaaacc ccgcgacgtt 5580
 ctatcgcgcg agcaacttct cattgccagt cgagtacgcg acgaggaggt ttatgacagg 5640
 agtatagatg ttctcatttt gaggtgcgcg cgaaacttg aggcagatcc gtcaagcctt 5700
 caactgataa aaacagcaag aggtgcccgt tattaacttg acgcggacgt gcaggttctg 5760
 caggggggga cgatggcagc ctgagccaat tcccagatcc ccgaggaatc ggctgagcg 5820
 gtcgaaaacc atccggcccc gtacaaatcg gcgcgccgct ggggtgatgac ctggtggaga 5880
 agttgaaggc cgcgcaggcc gccagcggc aacgcacga ggcagaagca cgccccggtg 5940
 aatcgtggca agcggccgct gatcgaatcc gcaaagaatc ccggcaaccg ccggcagccg 6000
 gtgcgcccgc gattaggaag ccgcccgaag gcgacgagca accagatttt ttcgttccga 6060
 tgctctatga cgtgggcacc cgcgatagtc gcagcatcat ggacgtggcc gttttccgct 6120
 tgcgaagcg tgaccgacga gctggcgagg tgatccgcta cgagcttcca gacgggacg 6180
 tagaggtttc cgcagggccc gccggcatgg ccagtggtg ggattacgac gagggtgag 6240
 tggcggtttc ccatctaacc gaatccatga aactgaccg ggaagggag ggagacaagc 6300
 ccggccgctg ttcccgcca caggttgcgg acgtactcaa gttctgcccg cgagccgatg 6360
 gcggaaagca gaaagacgac ctggtagaaa cctgcattcg gttaaacacc acgcacgtt 6420
 ccatgcagcg tacgaagaag gccagaacg gccgcctggt gacggtatcc gagggtgag 6480
 ccttgattag ccgctacaag atcgtaaaga gcgaaaccgg gcggccggag tacatcgaga 6540
 tcgagctagc tgattggatg taccgcgaga tcacagaagg caagaaccgg gacgtgctga 6600
 cggttcacc cgcattacttt ttgatcgatc ccggcatcgg ccgttttctc taccgcctgg 6660
 cacgccgctg cgcaggcaag gcagaagcca gatggttgtt caagacgatc tacgaacgca 6720
 gtggcagcgc cggagagttc aagaagtctt gtttcaccgt gcgcaagctg atccgggtcaa 6780
 atgacctgcc ggagtacgat ttgaaggagg aggcggggca ggctggcccg atccagtca 6840
 tgcgctaccg caacctgatc gagggcgaag catccgcccg ttccctaatgt acggagcaga 6900
 tgctagggca aattgcccta gcaggggaaa aaggtcgaaa aggtctcttt cctgtggata 6960
 gcacgtacat tgggaaccca aagccgtaca ttgggaaccg gaacccttac attgggaacc 7020
 caaagccgta cattgggaac cgttcacaca tgtaagtac tgatataaaa gagaaaaaag 7080
 gcgatttttc cgcctaaaac tctttaaacc ttattaaaac tctttaaacc cgcctggcct 7140
 gtgcataact gtctggccag cgcacagccc aagagctgca aaaagcgcct acccttcggt 7200
 cgctgcgctc cctacgcccc gccgcttccg gtcggcctat cgcggccgct ggcctgcaa 7260
 aaatggctgg cctacggcca ggcaatctac cagggcctg acaagccgcg ccgtcgccac 7320
 tcgaccgccc cgcctgaggt ctgcctcgtg aagaaggtgt tgctgactca taccaggcct 7380
 gaatcgcgcc atcatccagc cagaaagtga gggagccacg gttgatgaga gctttgttgt 7440
 aggtggacca gttggtgatt ttgaactttt gctttgccac ggaacggtct gcgttgcgg 7500
 gaagatgctg gatctgatcc ttcaactcag caaaagttcg atttattcaa caaagccgcc 7560
 gtcccgtcaa gtcagcgtaa tgctctgcca gtgttacaac caattaacca attctgatta 7620
 gaaaaactca tcgagcatca aatgaaactg caattttattc atatcaggat tatcaatacc 7680

atatttttga	aaaagccgtt	tctgtaatga	aggagaaaac	tcaccgaggc	agttccatag	7740
gatggcaaga	tcctgggtatc	ggtctgcgat	tccgactcgt	ccaacatcaa	tacaacctat	7800
taatttcccc	tcgtcaaaaa	taaggttatc	aagtgagaaa	tcaccatgag	tgacgactga	7860
atccggtgag	aatggcaaaa	gctctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	7920
gtttgctgat	tgggcgctct	tccgcttccct	cgctcactga	ctcgtctgctc	tcggctgcttc	7980
ggctgcgccg	agcgggtatca	gctcactcaa	agggcggtaat	acggttatcc	acagaatcag	8040
gggataacgc	aggaaagaac	atgtgagcaa	aaggccagca	aaaggccagg	aaccgtaaaa	8100
aggccgcgct	gctggcgctt	ttccataggc	tccgcccccc	tgacgagcat	cacaaaaatc	8160
gacgctcaag	tcagagggtg	cgaaaccgga	caggactata	aagataccag	gcgtttcccc	8220
ctggaagctc	cctcgtgctc	tctcctgttc	cgaccctgcc	gcttaccgga	tacctgtccg	8280
cctttctccc	ttcgggaagc	gtggcgcttt	ctcatagctc	acgctgtagg	tatctcagtt	8340
cggtgtaggt	cgttcgcctc	aagctgggct	gtgtgcacga	acccccgctt	cagcccagcc	8400
gctgcgcctt	atccggtaac	tatcgtcttg	atccaacc	ggtaagacac	gacttatcgc	8460
cactggcagc	agccaactgt	aacaggatta	gcagagcgag	gtatgtaggc	ggtgctacag	8520
agttcttgaa	gtggtggcct	aactacggct	acactagaag	aacagtattt	ggtatctgctg	8580
ctctgctgaa	gccagttacc	ttcggaaaaa	gagttggtag	ctcttgatcc	ggcaaaaaa	8640
ccaccgctgg	tagegggtgt	ttttttgttt	gcaagcagca	gattacgcgc	agaaaaaaag	8700
gatctcaaga	agatcctttg	atcttttcta	cggggtctga	cgctcagtgg	aacgaaaact	8760
cacgttaagg	gattttggct	atgagattat	caaaaaggat	cttcacctag	atccttttga	8820
tccggaatta	attcctgtgg	ttggcatgca	catacaaatg	gacgaacgga	taaacctttt	8880
caagcccttt	taaatatccg	attattctaa	taaagctctc	tttctcttag	gtttaccgca	8940
caatataatc	tgtcaaacac	tgatagttta	aactgaaggc	gggaaacgac	aatctgatca	9000
tgagcggaga	attaagggag	tcacgttatg	accccgcgcg	atgacgcggg	acaagccgct	9060
ttacgtttgg	aactgacaga	accgcaacgc	tgcaggaatt	ggccgcagcg	gccatttaaa	9120
tggtacctta	attaacgtac	gaagcttgca	tgcacgcggg	ctagagcggc	cgctcagagg	9180
taccgggccc	cccctcgagg	tcgacgggat	cgataagctt	gcatgcctgc	agtgcagcgt	9240
gacccggtcg	tgcccctctc	tagagataat	gagcattgca	tgtctaagtt	ataaaaaatt	9300
accacatatt	ttttttgtca	cacttgtttt	aagtgcagtt	tatctatctt	tatacatata	9360
tttaaacttt	actctacgaa	taatataatc	tatagtacta	caataatata	agtgttttag	9420
agaatcattt	aaatgaacag	ttagacatgg	tctaaggagc	aattgagtat	tttgacaaca	9480
ggactctaca	gtttttatctt	tttagtgtgc	atgtgttctc	cctttttttt	gcaaatagct	9540
tcacctatat	aataacttcat	ccatttttatt	agtacatcca	tttagggttt	agggttaatg	9600
gtttttatag	actaattttt	ttagtagatc	tatttttatc	tatttttagcc	tctaaattaa	9660
gaaaactaaa	actctatttt	agttttttta	tttaataatt	tagatataaa	atagaataaa	9720
ataaagtgac	taaaaattaa	acaaatacc	tttaagaaat	taaaaaaact	aaggaaacat	9780
ttttcttggt	tcgagtagat	aatgccagcc	tgttaaacgc	cgtcgacgag	tctaaccggac	9840
accaaccagc	gaaccagcag	cgctcgcctc	ggccaagcga	agcagacggc	acggcatctc	9900
tgctcgtgcc	tctggacccc	tctcgagagt	tccgctccac	cgttggactt	gctcgcctgt	9960
cggcatccag	aaattgcgctg	gcccagcggc	agacgtgagc	cggcacggca	gctccgctcc	10020
tcctcctctc	acggcacggc	agctacgggg	gattcctttc	ccaccgctcc	ttcgttttcc	10080
cttccctgcc	cgccgtaata	aatagacacc	ccctccacac	cctctttccc	caacctcgtg	10140
ttgttcggag	cgcacacaca	cacaaccaga	tctcccccaa	atccaccctg	cggcacctcc	10200
gcttcaaggt	acgcccctcg	tcctcccccc	ccccccctct	ctaccttctc	tagatcggcg	10260
ttccgggtcca	tgggttagggc	ccggtagttc	tacttctggt	catgtttgtg	ttagatccgt	10320
gtttgtggtta	gatccgtgct	gctagcgttc	gtacacggat	gcgacctgta	cgtcagacac	10380
gttctgattg	ctaacttgcc	agtgtttctc	tttggggaat	cctgggatgg	ctctagccgt	10440
tccgcagacg	ggatcgattt	catgattttt	tttgtttcgt	tgcatagggt	ttggtttggc	10500
cctttctctt	atttcaatat	atgccgtgca	cctgtttgtc	gggtcatctt	ttcatgcttt	10560
tttttgtctt	ggttgtgatg	atgtggtctg	gttggggcgt	cgttctagat	cggagtagaa	10620
ttctgtttca	aactacctgg	tggatttatt	aattttggat	ctgtatgtgt	gtgccataca	10680
tattcatagt	tacgaattga	agatgatgga	tggaaatata	gatctaggat	aggtatacat	10740
gttgatgcgg	gttttactga	tgcataataca	gagatgcttt	ttgttcgctt	ggttgtgatg	10800
atgtgggtg	gttggggcgt	cgttcattcg	ttctagatcg	gagtagaata	ctgtttcaaa	10860
ctacctggtg	tatttattaa	ttttggaact	gtatgtgtgt	gtcatacatc	ttcatagtta	10920
cgagtttaag	atggatggaa	atatacatct	aggataggta	tacatgttga	tgtgggtttt	10980
actatgata	atacatgatg	gcatatgcag	catctattca	tatgctctaa	ccttgagtac	11040
ctatctatta	taataaacia	gtatgtttta	taattatttt	gatcttgata	tacttggatg	11100
atggcatatg	cagcagctat	atgtggattt	tttttagcct	gccttcatac	gctattttatt	11160
tgcttggtag	tgtttctttt	gtogatgctc	accctgttgt	ttgggtgttac	ttctgcaggt	11220
cgactctaga	ggatccagaa	ttcgtgatca	aatggccgca	acaagcagca	caagcagcca	11280
gtcttttgac	atagagctcg	acatcatcgg	ccagcaaccg	cctcttcttt	caatctacac	11340

```

ccagatcagt ctcgtttacc ccgtctctga tccctcccag tatcccacca tcgtcagcac 11400
ccttgaggaa ggcctaaaac gcctctctca aaccttccca tgggtcgcgg gccagggtcaa 11460
gaccgagggc atcagcgaag gaaacacagg aacttccaag atcattccat atgaggagac 11520
accccgcttt gtggtgaaaag acctccgtga tgattcctca gcgccaacga tcgaggggtt 11580
gagaaaggcg ggtttcccct tagagatgtt tgacgagaac gtcgtcgctc cgaggaagac 11640
attagctatc ggacctggca atggcccaa cgaccgaag cctgtgttgc tattgcagct 11700
caacttcatt aagggcggac tcattctcac cgtcaacgga caacatggtg ctatggacat 11760
gacaggacaa gatgcaatta ttcgtcttct ctccaaggcg tgccgcaacg aatcattcac 11820
cgaggaggaa atctcggcca tgaacctcga tcgcaagacg gtagtccctc tccttgaaaa 11880
ctacaaagtt ggtcctgagc tagaccacca gatcgccaaa cctgcgctg ctggcgacgc 11940
tccacccgca ccggccaagg caagctggc gttcttttca ttactccca aggccctctc 12000
ggagctgaaa gacgcagcca caaagactct tgacgcgtcg tccaagttg tgtcaactga 12060
tgatgctctt tcggcgttta tctggcaatc aacctcgcgc gtacgtctcg caagattgga 12120
tgcttcaca cctactgaat tctgccgcgc tgtcgacatg cggggcccaa tgggcgtatc 12180
aagcacatac ccaggccttc ttcaaaacat gacctaccat gactcgaccg tcgccgaaat 12240
cgccaacgaa ccacttggcg caacagcatc acgectgcgc tcggaactca acagtgatcg 12300
tttgcgcaga cgaacacaag ctttggcgac gtacatgcat ggectgcctg acaagtcgag 12360
cgtctccctg accgccgatg cgaatccgtc aagcagcatc atgctgagtt ccfgggccaa 12420
ggtaggatgc tgggagtatg actttgggtt tggactgggt aagcctgaga gtgtgagaag 12480
acctcgcttt gaaccttttg agagtttgat gtactttatg cccaagaagc ctgatgggga 12540
gtttacggcg tccatttctc tgagggatga ggatatggag agactaaagg cggatgagga 12600
gtggacaaaag tacgcaaagt atattgggta gatagtttac tagactactg caggatatcg 12660
tggatccccg aatttccccg atcgttcaaa catttggcaa taaagtttct taagattgaa 12720
tcctgttgcc ggtcttgcgga tgattatcat ataatttctg ttgaattacg ttaagcatgt 12780
aataattaac atgtaatgca tgacgttatt tatgagatgg gtttttatga ttagagtccc 12840
gcaattatac atttaatacg cgatagaaaa caaaatatag cgcgcaaact aggataaatt 12900
atcgcgcgcg gtgtcatcta tgttactaga tcgggaattc ggcgcgcca 12949

```