



(12) 发明专利

(10) 授权公告号 CN 101289694 B

(45) 授权公告日 2013.01.02

(21) 申请号 200810089212.6

(22) 申请日 2005.12.15

(30) 优先权数据

11/043,520 2005.01.25 US

(62) 分案原申请数据

200510131877.5 2005.12.15

(73) 专利权人 国家科学和技术发展局

地址 泰国曼谷

专利权人 泰国农业大学

(72) 发明人 阿披乍特·瓦纳维奇特

宋旺·特拉古欧荣

提叻于特·图金达 萨马特·瓦乍那

温台·卡莫利苏克云勇

(74) 专利代理机构 北京市路盛律师事务所

11326

代理人 李宓

(51) Int. Cl.

C12Q 1/68 (2006.01)

G01N 33/68 (2006.01)

(56) 对比文件

CN 101061219 A, 2007.10.24, 说明书第7页

第1-2段, 第9页第7-8段, 实施例2, 附图6, 附图7.

S.Garland et al.PCR-based molecular markers for the fragrance gene in rice(oryza sativa.L.).《Theor Appl Genet》.2000, 第101卷364-371.

Qingsheng Jin et al.A single nucleotide polymorphism(SNP) marker linked to the fragrance gene in rice (Oryza sativa L.).《Plant Science》.2003, 第165卷359-364.

Giovanni M. Cordeiro et al.Identification of microsatellite markers for fragrance in rice by analysis of the rice genome sequence.《Molecular Breeding》.2002, 第9卷245-250.

审查员 曹扣

权利要求书 1 页 说明书 59 页 附图 15 页

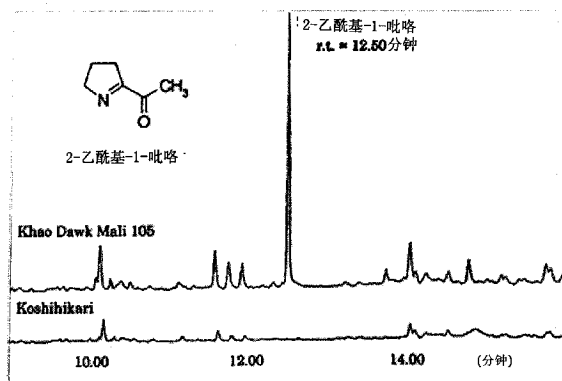
(54) 发明名称

增强植物和真菌中的2-乙酰基-1-吡咯的合成的核酸

(57) 摘要

芳香化合物2-乙酰基-1-吡咯(2AP)是所有芳香稻米的主要潜在香味成分。本发明提供了一种非天然出现的植物,在所述植物中,所合成的2-乙酰基-1-吡咯的水平大于天然出现的非芳香变种中的水平。本发明还提供了能够用作分子标记物的核酸,所述标记物用于选择合成高水平的2-乙酰基-1-吡咯的植物和真菌。本发明还提供了一种检测植物或植物部分中2-乙酰基-1-吡咯基因的芳香等位基因的存在的方法。

CN 101289694 B



1. 一种检测植物或植物部分中2-乙酰基-1-吡咯基因的芳香等位基因的存在的的方法,包括检测2-乙酰基-1-吡咯基因是否具有减小2-乙酰基-1-吡咯蛋白质的量、功能或活性的突变,其中所述2-乙酰基-1-吡咯基因如SEQ ID NO:1或SEQ ID NO:4所示,所述突变是在Os2AP的芳香等位基因中发现的8碱基对缺失。

2. 如权利要求1所述的方法,其特征在于,所述检测突变的方法包括PCR和/或微阵列的使用。

3. 如权利要求1所述的方法,其特征在于,所述检测突变的方法包括杂交或测序。

4. 如权利要求3所述的方法,其特征在于,所述检测突变的方法包括检测2-乙酰基-1-吡咯蛋白质的水平、结构或活性。

增强植物和真菌中的 2-乙酰基 -1-吡咯的合成的核酸

[0001] 本发明是中国专利申请号为 200510131877.5 申请日为 2005 年 12 月 15 日发明名称为“增强植物和真菌中的 2-乙酰基 -1-吡咯的合成的核酸”的分案申请。

[0002] 发明领域

[0003] 本发明通常涉及植物分子遗传学。尤其是涉及具有高水平的 2-乙酰基 -1-吡咯的非天然出现的植物和真菌、这些植物和真菌的制造方法,以及 2-乙酰基 -1-吡咯的合成中所涉及的核酸。

[0004] 发明背景

[0005] 米粒的芳香是高质量稻米最吸引人的特征,而高质量的稻米不仅在亚洲市场的需求日益增长,而且也在欧洲和全世界得到广泛认可。熟米的香气是由一百多种挥发性化合物组成的,这些化合物诸如烃类、醇类、醛类、酮类、酸类、酯类、酚类、吡啶类、吡唑类和其它化合物 (Yajima 等人,1978 ;Maga,1984 ;Takashi 等人,1980 ;Paule 和 Power,1989)。“爆米花样”的芳香化合物即 2-乙酰基 -1-吡咯 (2AP) 据发现是所有芳香稻米、面包麦皮和黑面包的主要有效香味成分 (Buttery 等人,1982,1983)。2-乙酰基 -1-吡咯主要负责为许多芳香稻米变种提供特征香气 (Tanchotikul 和 Hsieh,1991)。令人惊奇的是,这种稻米的香气也已经从露兜树叶 (Buttery 等人,1983)、面包花 (*Vallaris Glabra* Ktae.) (Wongpornchai 等人,2003)、湿粟 (Seitz 等人,1993)、爆米花 (Schieberle,1991)、蜡样芽孢杆菌 (Romanczyk 等人,1995) 和真菌 (Nagsuk 等人,2004) 中分离出来并进行了鉴定。2-乙酰基 -1-吡咯存在于芳香稻米植物的除根之外的所有部分 (干、叶、粒) 中 (Lorieux 等人,1996)。虽然这种香气存在于芳香米粒中,但是并不存在于所有的米粒中。

[0006] 芳香化合物 2-乙酰基 -1-吡咯具有与氨基酸脯氨酸类似的吡咯环 (图 1)。将氨基酸脯氨酸作为合成 2-乙酰基 -1-吡咯的前体的第一证据,在细胞和愈伤组织培养的实验中发现 (Suprasanna 等人,1998 ;Suprasanna 等人,2002)。此结论被采用同位素标记的实验所支持,该实验表明,谷粒 2-乙酰基 -1-吡咯的前体最可能是 Thai Hom Mali (THM) 稻米中的氨基酸脯氨酸 (Yoshihashi 等人,2002),也可能是其它芳香稻米。然而,2-乙酰基 -1-吡咯的精确生物合成路线还没有阐明。这样,就需要鉴定 2-乙酰基 -1-吡咯合成中涉及的基因,并提供一种增大植物和真菌中的 2-乙酰基 -1-吡咯水平的方法,以便增加香气。

[0007] 发明概述

[0008] 本发明通过提供非天然出现的植物和真菌,而满足这些需要,与对照植物和真菌相比,化合物 2-乙酰基 -1-吡咯在这些非天然出现的植物和真菌内是以高水平产生的。对照植物和真菌是指具有较低水平化合物 2-乙酰基 -1-吡咯的类似或相关基因型的植物和真菌。这样的对照植物可以是非天然出现的。本发明还提供用于筛选和制造这些非天然出现的植物和真菌的方法,以及由所述植物产生的种子。

[0009] 化合物 2-乙酰基 -1-吡咯的精确生物合成路线还不可知,但它是脯氨酸的衍生物。假定,脯氨酸能够转换成 2-乙酰基 -1-吡咯或谷氨酸,因此对谷氨酸合成路线的抑制,将增大脯氨酸 (或中间体) 对 2-乙酰基 -1-吡咯合成的有效性 (图 4A)。编码控制稻米香气的蛋白质的基因,即 Os2AP (*Oryza sativa* 2-乙酰基 -1-吡咯),被鉴定为在脯氨酸向谷

氨酸的转换过程中可以起关键作用的醛脱氢酶族的一原。所测试的所有芳香稻米变种在此基因中都具有 8 核苷酸缺失。该缺失产生提前终止密码子,该密码子引起其自身的 mRNA 发生无义介导的降解,从而导致功能损失的表型的产生。RNA 干扰 (RNAi) 的研究表明,Os2AP 基因转录的破坏导致植物内 2-乙酰基-1-吡咯的水平升高,以及香气的增加。

[0010] 本发明提供了非天然出现的植物和真菌,所述植物和真菌通过抑制 2-乙酰基-1-吡咯 (2AP) 基因的表达、减少 2AP 基因的 mRNA 水平、和 / 或减小 2AP 蛋白质的活性,而具有高水平的 2-乙酰基-1-吡咯。与对照植物相比,2AP 蛋白质的水平可减小 25%、50% 或 100%。通过以下手段,可实现对 2AP 基因表达的抑制或 2AP 基因的 mRNA 水平的减小: a) 反义取向的 2AP 基因或其片段的表达; b) 将部分基因克隆成 RNA 干扰构造并将此构造表达在转基因植物内;或者 c) 利用多种方法(包括靶向诱导的基因组的局部损害 (TILLING) 和 tDNA 插入诱变) 进行诱变,然后利用 PCR 或其它方法筛选芳香变种。

[0011] 本发明还提供了一种与对照的非转基因稻米植物相比,具有高水平的化合物 2-乙酰基-1-吡咯的转基因稻米植物,其中与对照的非转基因植物中 2AP 基因编码的 mRNA 或蛋白质水平相比,通过减小转基因植物中 2AP 编码的 mRNA 或蛋白质水平,而增大了该转基因植物中此化合物的水平。在一个版本中,mRNA 和蛋白质的水平由于 RNA 干扰或由于反义而得以减小。本发明还涉及由本发明的稻米产生的转基因稻米种子。

[0012] 虽然后面的实例描述的是在稻米中实施的实验,但是本发明涉及其它植物和真菌,包括(但不限于)小麦、大麦、黑麦、椰子、高粱和燕麦。

[0013] 本发明还提供了一种编码 2AP 基因的分离核酸,其中所述核酸包括:在以下杂交条件下杂交到 SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:4 或 SEQ ID NO:5 中描述的核酸序列或其补体上的核酸;所述杂交条件包括在 60-65°C 于 0.1×SSC 和 0.1% SDS 中至少洗涤一次 30 分钟;与 SEQ ID NO:1、SEQ ID NO:2、SEQ ID NO:4 或 SEQ ID NO:5 中描述的序列至少 70%相同、80%相同、90%相同、95%相同或大于 95%相同的核酸;以及对与 SEQ ID NO:3 或 SEQ ID NO:6 中描述的氨基酸序列至少 70%相同、至少 80%相同、至少 90%相同、至少 95%相同或大于 95%相同的多肽进行编码的核酸。

[0014] 稻米内的 2-乙酰基-1-吡咯的量可根据收获条件和土壤类型来改变。非芳香变种 (Nipponbare) 具有 0-0.1ppm(百万分之.....) 的 2-乙酰基-1-吡咯水平。相反,芳香稻米变种 (Thai Hom Mali) 具有 1-2.5ppm 的 2-乙酰基-1-吡咯的量。例 2 详细描述了稻米 Os2AP 基因的 RNA 干扰实验,其将 Nipponbare 稻米的 2-乙酰基-1-吡咯水平增大到高达 2.5ppm。本发明由此提供了将非芳香植物的香气水平增大到芳香水平的方法。

[0015] 本发明还提供了含有本发明核酸的重组构造和表达载体,其中所述核酸可操作地链接到启动子上。一种有用的启动子是花椰菜花叶病毒 (CaMV) 启动子,该启动子使大多数植物组织都具有高水平的表达。

[0016] 本发明还提供了含有本发明的核酸、构造和表达载体的宿主细胞。

[0017] 本发明还提供了用于筛选植物及核酸的方法,用于使 2AP 基因发生突变,从而导致 2AP 的蛋白质表达或活性减小,并因此使由于 2AP 化合物产量的增大而产生的香气增大。一种具体的突变是与稻米 2AP 基因中的芳香表型有关的八核苷酸缺失。对这种或其它突变的筛选可利用各种方法来实施,这些方法诸如有 PCR、测序、杂交或微阵列实验。另一种选择是查寻 2AP 蛋白质水平的减小或 2AP 蛋白质结构或活性的变化。诸如通过利用与活性 2AP

蛋白质结合的抗体或通过 2AP 蛋白质活性的测定,可实现这一点。

[0018] 本文描述的序列可用作核酸杂交实验中的探针或引物。可采用包括至少 14 核苷酸长邻接序列的核酸节段,所述邻接序列具有与 SEQ ID NO :1、SEQ ID NO :2、SEQ ID NO :4 或 SEQ ID NO :5 中描述的核苷酸序列的 14 核苷酸长邻接 DNA 节段相同或互补的序列。

[0019] 在某些实施例中也将采用更长邻接的相同或互补序列,例如大约 20、30、40、50、100、200、500、1000、2000、5000、10000 等长度的序列(包括所有的中间长度并高达且包括整个长度的序列)。

[0020] 这些核酸探针与 2AP 基因序列特异性杂交的能力,使其能够用于检测给定样品中互补序列的存在。

[0021] 然而,预想其它用途,包括利用序列信息制备突变物种引物,或者用于制备其它遗传构造的引物。

[0022] 附图简述

[0023] 图 1 表示出来自 GC-MS、从 Khoa Dawk Mali (KDML105)、Thai Jasmine 米(芳香的)和 Kosihikari(非芳香米)提取的 2-乙酰基-1-吡咯的化学结构。

[0024] 图 2 表示出利用所分离的谷粒香气的单个 F16 植物衍生的 186 个 F9 植物,对稻米的芳香基因进行的精细标度作图,和包括谷粒芳香基因的物理图构造。

[0025] 图 2 的第一部分表示出来自单个 F6 植物的八个 F11 植物的图形基因型。第二部分表示出利用单个 F6 植物衍生的 1116 个 F12 植物进行的超标度作图,从而使临界区向下变窄到单个 BAC 中的 27kb。第三部分表示出来自 KDML105 的基因组序列的注释,揭示出三个开放读框。第四部分表示出利用来自 KDML105 与 JHN 之间的交配的 177 个 F6 植物,鉴定未知蛋白质基因的外显子 7 内的三个双重组体,所述的未知蛋白质基因显著影响谷粒的香气和 2-乙酰基-1-吡咯的含量。未知蛋白质基因被命名为 Os2AP,即 *Oryza saliva* 2-乙酰基-1-吡咯。“Aromarker”是基于 PCR 的标记物,其限定出 8 碱基对缺失和特异于谷粒香气的 3 SNPs(单核苷酸多态性)。

[0026] 图 3A 表示出利用 RT-PCR、在 10、15、20 天之后,从芳香与非芳香同基因系之间的传粉中分离的总 RNA 的 7 个候选基因的表达。在此图中,Os2AP 被称作甜菜碱醛脱氢酶。

[0027] 图 3B 表示出在 15 天之后从芳香与非芳香同基因系之间的传粉中分离的总 RNA 的叶、干和根的 Os2AP 转录的不同表达。

[0028] 图 3C 表示出 THM(KDML 105)、芳香同基因系 117、非芳香同基因系 10 和它们的 F2(ISL117×ISL10)的谷粒中的 2-乙酰基-1-吡咯水平分析。在 F1 中,2-乙酰基-1-吡咯水平分析是在叶中进行的。

[0029] 图 3D 在上部表示出携带 Os2AP RNAi 构造的转基因 Nipponbare 中的染色体 4 和肌动蛋白的 Os2AP 同系物的表达。图 3D 在下部表示出染色体 8 上的而不是染色体 4 上的 Os2AP 基因(此处称作 BADH),表示出稻米的芳香和非芳香同基因系中的不同表达。

[0030] 图 4A 示出了脯氨酸和谷氨酸之间的代谢途径。L-脯氨酸可以利用酶 P5C 合成酶和 P5C 还原酶、用谷氨酸来合成,而谷氨酸可以利用脯氨酸脱氢酶(ProDH)和 P5C 脱氢酶(P5C DH)、用脯氨酸来合成。所建议的代谢漂移由 Os2AP 的无义突变来介导。

[0031] 图 4B 示出了利用 RASMOL 创建的 OS2AP 酶的预测蛋白质结构,RASMOL 是通过美国 Amherst 的 Massachusetts 大学的分子生物学俱乐部网站获得的程序(新的版本是蛋白质

探索者)。

[0032] 图 5A 表示出来自 KDML 105(芳香的)和 Nipponbare(非芳香的)的 Os2AP 基因的基因组序列比较。

[0033] 图 5B 表示出 DNA 的排列,此排列表明与 Nipponbare(非芳香株系)相比,芳香株系 Thai Hom Mali (THM) 中的 8 碱基对缺失,以及这两个株系之间的氨基酸序列比较。排列中的核苷酸序列包括来自 SEQ ID NO :5 的核苷酸 701-765,和来自 SEQ ID NO :2 的核苷酸 701-757。

[0034] 图 6 表示出来自单个 F6 植物的 8 个 F11 植物的图形基因型,和来自米粒的 2-乙酰基-1-吡咯水平分析。

[0035] 图 7A 表示出用于转化的 RNA 干扰构造和载体。该图的下部表示出构造表达的证实。BADH 在此处称作 Os2AP。

[0036] 图 7B 表示出 RNAi 构造、GFP 和内源性 Os2AP 的表达。该图的下部表示出 2-乙酰基-1-吡咯的水平和香气。BADH-RNAi2 是针对 Os2AP 基因的 RNAi 构造。

[0037] 图 8A 表示出利用“Aromarker”引物组,筛选与增大的香气相关的 Os2AP 变种的 F6 子代的结果。

[0038] 图 8B 表示出筛选芳香 Os2AP 等位基因的稻米的多种变种的结果。

[0039] 图 9A 表示出稻米的多种株系的 Os2AP 直向同源的 DNA 序列。排列中的核苷酸序列包括来自 SEQ ID NO :5 的核苷酸 701-765,和来自 SEQ ID NO :2 的核苷酸 701-757。

[0040] 图 9B 表示出利用 22 个 BADH(Os2AP) 直向同源的邻近-接合点 (Neighbor-Joint) 树构建的系统发育树。

[0041] 序列表中的一些序列的简单描述

[0042] SEQ ID NO :1 是芳香米株系 Thai Hom Mali 的 Os2AP 基因的基因组核苷酸序列。

[0043] SEQ ID NO :2 是 Thai Hom Mali Os2AP 的蛋白质编码的核苷酸序列。

[0044] SEQ ID NO :3 是 Thai Hom Mali Os2AP 蛋白质的氨基酸序列。

[0045] SEQ ID NO :4 是非芳香米株系 Nipponbare 的 Os2AP 基因的基因组核苷酸序列。

[0046] SEQ ID NO :5 是 Nipponbare Os2AP 的蛋白质编码的核苷酸序列。

[0047] SEQ ID NO :6 是 Nipponbare Os2AP 的氨基酸序列。

[0048] SEQ ID NO :7 至 88 是能够用来扩增 Os2AP 核苷酸序列、GFP 或肌动蛋白部分的引物(表 1)。

[0049] SEQ ID NO :89 至 95 是 2AP 基因直向同源片段的序列。

[0050] SEQ ID NO :96 是高度保存在通常的醛脱氢酶之中的脱肽。

[0051] 发明详述

[0052] 本发明涉及 2-乙酰基-1-吡咯(2AP)基因。2AP 基因的核酸序列用来增大 2-乙酰基-1-吡咯的水平,此化合物是在稻米、小麦、玉米、露兜树叶、芳香椰子和一些细菌及真菌中发现的一种主要芳香化合物。

[0053] 在详细解释本发明的至少一个实施例之前,应该理解,本发明并不局限于应用于一些描述中提到的成分的构建和布置细节。本发明能够具有其它实施例或者以多种方式进行实践或完成。而且,应该理解,此处采用的措词和术语是为了进行描述,而不应该被认为是对本发明的限定。

[0054] 在此公开文本中,参考多个不同的出版物、专利和出版的专利说明书。这些出版物、专利和出版的专利说明书在此作为参考并入本发明的公开文本中,以便更充分地描述本发明所属的技术领域状态。除非另外指出,否则本发明的实践将采用本领域技能范围内常规的植物繁殖、免疫学、分子生物学、细胞生物学和重组 DNA 技术。参见诸如, Sambrook、Fritsch 和 Maniatis 的分子克隆:实验室手册 (MOLECULAR CLONING: A LABORATORY MANUAL), 第二版 (1989); 分子生物学中的当前方案 (CURRENT PROTOCOLS IN MOLECULAR BIOLOGY) (F. M. Ausubel 等人, eds., (1987)); 植物的繁殖:原理和前景 (Plant Breeding: Principles and Prospects) (植物繁殖 (Plant Breeding), 第 1 卷) M. D. Hayward, N. O. Romagosa; Chapman & Hall, (1993); Coligan、Dunn、Ploegh、Speicher 和 Wingfeld, eds, (1995) 蛋白质科学中的当前方案 (CURRENT PROTOCOLS IN PROTEIN SCIENCE) (John Wiley & Sons, Inc.); 酶学中的系列方法 (the series METHODS IN ENZYMOLOGY) (Academic Press, Inc.); PCR2: 实践方法 (A PRACTICAL APPROACH) (M. J. MacPherson, B. D. Hames 和 G. R. Taylor eds, (1995), Harlow 和 Lane, eds, (1988)); 抗体, 实验室手册 (ANTIBODIES, A LABORATORY MANUAL) 和动物细胞培养 (ANIMAL CELL CULTURE) R. I. Freshney, ed, (1987)。

[0055] 除非另外指出,否则技术术语是按照常规用途来使用的。分子生物学中的通用术语的定义可以在以下书籍中找到:1994 年 Oxford University Press 出版的 Lewin, Genes V (SBN0-19-854287-9); 1994 年 Blackwell Science Ltd. 出版的 Kendrew 等人 (eds.) 的分子生物学百科全书 (The Encyclopedia of Molecular Biology) (SBN0-632-02182-9); 以及 1995 年 VCH Publishers, Inc. 出版的 Robert A. Meyers (ed.) 的分子生物学和生物技术 (Molecular Biology and Biotechnology), a Comprehensive Desk Reference, (ISBN 1-56081-569-8); Ausubel 等人的 (1987) 分子生物学中的当前方案 (Current Protocols in Molecular Biology), Green Publishing; Sambrook 等人的 (1989) 分子克隆:实验室手册 (Molecular Cloning: A Laboratory Manual), Cold Spring Harbor, New York。植物生物学中的通用术语定义可以在以下书籍中找到:Esau 的植物解剖学 (Plant Anatomy), 由 John Wiley & Sons 出版 (1977) (ISBN 0-471-24520-8); 以及 Solomon 等人的生物学 (Biology), Saunders College Publishing 出版 (1993)。

[0056] 定义

[0057] 为了易于阅读本发明的多种实施例,提供以下定义:

[0058] 2-乙酰基-1-吡咯 (2AP) 多核苷酸序列:按照本发明的这些基因不仅包括本文公开的整个长度的序列,而且还包括这些序列的片段,所述片段保留了本文具体解释的序列的特征活性。

[0059] 本文公开了稻米的 2AP 基因序列。本领域技术人员显而易见的是,利用稻米的 2AP 基因序列,通过若干方式能够容易地鉴定和获得另一种植物的 2AP 多核苷酸序列。具体的基因或其一部分可以从培养基存放处获得,或者诸如利用基因机器来合成构建。

[0060] 利用制造点突变的标准技术,可容易地构建这些基因的变异。而且,利用商业上购得的内切核酸酶或外切核酸酶、按照标准方案,能够制造这些基因的片段。例如,诸如 Bal31 之类的酶或针对位点的诱变可用来从这些基因的端部系统切断核苷酸。而且,利用各种其它限制酶可获得为活性片段编码的基因。多种酶可用来直接获得这些 2AP 多核苷酸序列的活性片段。

[0061] 利用本文提供的技术,还可以从株系和 / 或 DNA 文库中分离出等同的 2AP 多核苷酸序列和 / 或编码这些等同 2AP 多核苷酸序列的基因。例如,本文公开的 2AP 蛋白质的抗体可用来从蛋白质混合物中鉴定和分离出其它 2AP 蛋白质。具体地说,抗体可以提升 2AP 蛋白质最稳定和距离其它蛋白质最远的部分。然后,通过免疫沉淀法、酶联免疫测定法 (ELISA) 或 Western 印迹法,利用特征活性,用这些抗体来具体鉴定等同的 2AP 蛋白质。

[0062] 用于鉴定本发明基因的另一方法是,采用寡核苷酸探针。这些探针是具有可检测标记物的核苷酸序列。如本领域内众所周知的,如果探针分子与核酸样品通过在这两个分子之间形成强键而进行杂交,那么可以合理地假设,探针和样品实质上是相同的。探针的可检测标记物提供了用公知方式进行测定的方法,而无论杂交是否已经发生。这样的探针分析提供了用于鉴定本发明基因的快速方法。示范性探针在表 1 中有所描述。

[0063] 用作按照本发明的探针的核苷酸节段,可以利用标准方案、采用 DNA 合成器来合成。在利用核苷酸节段作为探针的过程中,用本领域技术人员公知的任何适宜标记物来标记特定的探针,所述标记物包括放射性和非放射性标记物。一般的放射性标记物包括 ^{32}P 、 ^{125}I 、 ^{35}S 等。用放射性同位素标记的探针可以用与 DNA 样品互补的核苷酸序列、通过常规的切口转录反应、利用 Dnase 和 DNA 聚合酶来构建。探针和样品然后合并并在杂交缓冲液中,并保持在合适的温度下,直到退火发生为止。其后,洗掉膜上的外源性材料,而剩下一般用自动放射显影法和 / 或液体闪烁计数检测和鉴定的样品及结合探针分子。

[0064] 非放射性标记物包括诸如配体例如生物素或甲状腺素,以及酶类例如水解酶或过氧化物酶,或者多种化学发光剂例如荧光素 (luciferin),或者诸如荧光素 (fluorescein) 及其衍生物之类的荧光化合物。探针也可以在两端用不同种类的标记物进行标记,以便容易分离,或者诸如通过在以上所述的端部使用同位素标记物、在另一端使用生物素标记物,而进行标记。

[0065] 双螺旋的形成和稳定性取决于两个杂交链之间的实质互补性,并且如上所述,可容许一定程度的错配。因此,本发明的探针包括所述序列的突变(一个或多个)、缺失、插入、及其组合,其中所述突变、插入和缺失允许和感兴趣的靶多核苷酸形成稳定的杂交。利用本领域普通技术人员目前公知的方法,以及或许是将来公知的其它方法,能够以许多方式在给定的多核苷酸序列中产生突变、插入和缺失。

[0066] 所列出的探针的潜在变异是由于(或者部分是由于)遗传密码的丰余。由于遗传密码的丰余,即多个编码核苷酸三联体(密码子)的遗传密码能够用于用来制造蛋白质的大多数氨基酸。因此,不同的核苷酸序列能够为特定的氨基酸编码。这样,用编码蛋白质或肽的相同氨基酸序列的等同核苷酸序列,能够制备 2AP 蛋白质和肽的氨基酸序列。据此,本发明包括这样等同的核苷酸序列。而且,反向或互补序列是本发明的一个方面,并且能够容易被本领域的技术人员所使用。此外,已经表明,通过改变氨基酸序列,可以构建结构和功能已鉴定的蛋白质,如果这样的变化不改变蛋白质的二级结果的话(Kaiser 和 Kezdy, 1984)。因此,本发明包括此处所述的不改变蛋白质的二级结构的氨基酸序列突变体,或者如果结构改变,则基本上保留生物活性。另外,本发明还包括宿主所有或部分 2AP 多核苷酸序列的有机体的突变体,所述 2AP 多核苷酸序列编码本发明的基因。这样的突变体可利用本领域技术人员公知的技术来制造。例如,紫外辐射可用来制备宿主有机体的突变体,可采用 tDNA 插入诱变,或者采用 TILLING(定向诱导的基因组局部损伤)。同样,这些突变体可

包括也能够用本领域公知的方案来制备的无孢子宿主细胞。

[0067] 2AP 多核苷酸序列（包括稻米中的那些）可以利用本发明的 PCR 引物来获得。这些引物在表 1 中示出，并对应于 SEQ ID NO :7 至 SEQ ID NO :79 中所描述的序列。这些引物的组合可用来扩增 Os2AP 基因的不同区域。为此扩增工作的 PCR 条件如下所示：10 μ L 的反应混合物中具有 10ng 的模板 DNA、0.1mM 的 dNTP、0.5M 的每种引物、0.5 个单位的 Taq 聚合酶、2.0mM 的 MgCl₂ 和 1x 嗜热性聚合酶缓冲液（Promega）。此混合物应该在以下时间和温度下进行 30 个循环的 PCR：94 $^{\circ}$ C, 30 秒钟（变性）；60 $^{\circ}$ C, 30 秒钟（退火）；以及 72 $^{\circ}$ C, 2 分钟（延伸反应）。

[0068] 表 1 :引物列表

[0069]

SEQ ID NO :	引物名	序列 (5' \rightarrow 3')
7	OS2AP-8L	GCCATGCCAACTGAGTAAAG
8	OS2AP-8R	CAATTTTATTCGCTCTGTGC
9	OS2AP-9L	TGCAACATCGCGTCTATTC
10	OS2AP-9R	GCAACTAGCAAGAGCATAACACC
11	OS2AP-12L	ACCTGACATCATGCCTTTGG
12	OS2AP-12R	CCGGTCATCAGCTAACTTCC
13	OS2AP-13R	CCCTTCGTCATAAAATATACTAGCAA
14	OS2AP-14L	TCCTCCAACATGCTCTTTCG
15	OS2AP-14R	CAGAGAAGTTTACGCCGTTG
16	OS2AP-15L	TTTTTAAATAAGATGAACGGTCAAA
17	OS2AP-16L	CTCTCCACCCTCTGCTTCTG
18	OS2AP-16R	CTCTCCGCTTGAACCCATC
19	OS2AP-17L	GCATGGCTGATTGTGTATCTG
20	OS2AP-17R	TTCCAAACCTACGGACAAAAG
21	OS2AP-18L	TTCCTCTTCTTTGTGCAAAC
22	OS2AP-18R	CACGGAAGCCAATTCAGATG

23	OS2AP-19L	CTATCCTCTCCTGATGGCAAC
24	OS2AP-19R	TGGCTACTAGAATGATGCTCAAAG
25	OS2AP20L	CCTTTTGTGTCGCTTTTGAG
26	OS2AP20R	AAAATAGCCTTCACTCGTTGC
27	OS2AP21L	CCATCGATTTGAGGGTAAC
28	OS2AP21R	CGCATCCGATAATATGTTG
29	OS2AP22L	GTAATTAGGAGTACGACTCTCGTC
30	OS2AP22R	GCTTATAGCCTACTGTATCCTCCTC
31	OS2AP23L	AATTGGTTAACCAGCAAGC
32	OS2AP23R	ACATTGTGAAACGGAGGAAG
33	OS2AP24L	GCTATAAGCCAGCTGCAAAC
34	OS2AP24R	GCAGTTGGTACGGACTTCG
35	OS2AP25L	CCTAAATATTTGACGCCGTTG
36	OS2AP25R	TGAAGAGGAGGGTACCGATG
37	OS2AP26L	CACCACTCCACACCTGACAC
38	OS2AP26R	GTACGGAACACACGCACAAG
39	OS2AP27L	TGTTGTTGTTGTTGCTGCTG
40	OS2AP27R	GCCGTGAGCCATATACTTG
41	023088C02 1L	AGCTCCAGCTCCTCCTCGAT
42	023088C02 2L	TATCTCTCACCGACCCAAA
43	023088C02 2R	TGTTGCCATCAGGAGAGGA
44	023088C02 3R	CTCTTGATGAAGCAGCATGG
45	023088C02 3R	CCCAGTAAATGCAACCTTGTC
46	023088C02 4R	GGCAACATGGAAGGTAGCTC

[0070]

47	023088C02 4R	CCATGCAACCATCCTTTCTT
48	023088C02 5R	TTATGGCTTCAGCTGCTCCT
49	023088C02 5R	CAATGGCTTCTTCTTCAGTGC
50	023088C02 6R	GCCCGTTGTTAGTGAAGGAC
51	023088C02 6R	GTACCATCCCCACGGCTCAT
52	023088C02 7L	CGAGCGATGCCAGAGATTA
53	023088C02 7R	AGCACATGGCAAATCAAACA
54	OS2AP-exon7. 1- del_F	TGCTCCTTTGTCATCACACC
55	OS2AP-exon7. 1- del_R	TTCCACCAAGTCCAGTGA
56	OS2AP_in1L	TTCGCTGCAGAACAGATGAC
57	OS2AP_in1R	CTGATGGTTACGCGACAATTT
58	OS2AP_inTA2L	ATTTGAACCGGGACAGAACA
59	OS2AP_inTA2R	TTTTGATGTGCCCTCTCCTT
60	OS2AP_inAAT3 L	TGGGTAATCTTGTCTGGAG
61	OS2AP_inAAT3 R	AGTGCCAAATGCATGCTAGA
62	OS2AP_inG4L	TGGGGCTCAAAAACCTACTG
63	OS2AP_inG4R	GTCCGGGCCAAGTACCTC
64	OS2AP-5-UTR- EX1-5F	ATCTCTACCGACCCCAAAT
65	OS2AP-5-UTR- EX1-5R	CCATTGGAAGAGAGACAGGTG

66	OS2AP-ATG1-600F	TGTTGTTGTTGTTGCTGCTG
67	OS2AP-ATG1-600R	TGGGGCTCAAAAACCTACTG
68	OS2AP-EX5-12F	GGTTGGTCTTCCTTCAGGTG
69	OS2AP-EX5-12R	GGTCCAAAAGCAACCAAAGA
70	Aromarker BigL	ACTGGTAAAAAGATTATGGC
71	Aromarker BigR	CAAGCCGATCAACCAGTACA
72	Aromarker SmallL	CCATGCTGCAAGCAATGTA
73	Aromarker SmallR	AACCATAGGAGCAGCTGAAATA
74	OS2AP8_OUT_F	ACCCTGGTGTAGACAAGGTA
75	OS2AP8_IN_F	GGGAGTTATGAAACTGGTATAT
76	OS2AP8_IN_R	ATAGGAGCAGCTGAAGCCAT
77	OS2AP8_OUT_R	GTCCCGCACTTCAGAATTAG
78	OS2AP8_ex2_F	CTCTGCTTCTGCCTCTGATT
79	OS2AP-exon9.1-del_RN	CTGGCTACTAGAATGATGCTC
80	Exon6to9NcoIF	AATTCCATGGGGTTGGTCTTCCTC AGGTG
81	Exon6to9SpeIR	AATTACTAGTTTCCACCAAGTTCCA GTGAA
82	Exon6to8NheIR	AATTCCATGGGGTTGGTCTTCCTC AGGTG

83	GFPU	CTTGTTGAATTAGATGGTGATGTT
84	GFPL	GTTGTGGGAGTTGTAGTTGTATTC
85	Os2APCH4U	TAGCTTCACATCCCCATGTG
86	Os2APCH4L	GCACCTTCACATCTTGCTGT
87	肌动蛋白 U	ACATCGCCCTGGACTATGAC
88	肌动蛋白 L	TGCTGAGAGATGCCAAGATG

[0071] *Oryza sativa* 2-乙酰基-1-吡咯 (Os2AP) 同系物:通过基于计算机的方法,利用公共域序列排列算式和检索序列数据库(公共域数据库包括 Genbank、EMBL、Swiss-Prot、PIR 等)的序列相似性检索工具,能够鉴定显示出与本申请所述的那些序列相似的序列。

[0072] 相似性检索找回并排列序列,以便与待分析的靶序列(即,查询序列)进行比较。比较序列局部区域之间的最佳排列被称作局部排列。序列比较算式利用记分矩阵将全部分数分配给每个排列。

[0073] 多核苷酸及多肽序列可以进行排列,并利用通过公共途径获得的计算机算式,针对多核苷酸及多肽序列,来确定特定区域中的相同残基的百分率。百分率相同性分数取决于比较序列的重叠区长度。

[0074] 两个核酸序列或两个氨基酸序列之间的相似性可以用序列的相同性来表达(或者,对于蛋白质,也可以用序列的相似性来表达)。序列的相同性经常以百分率相同性来测定;百分率越高,两个序列就越相似。如本文所述,2AP 蛋白质编码的核酸分子的同系物和变种可用于本发明。这些核酸分子的同系物和变种在利用标准方法进行排列时,将拥有相当高程度的序列相同性。这样的同系物和变种在高度严格的条件下将彼此杂交。

[0075] 为了进行比较而排列序列的方法在本领域内是公知的。多种程序和排列算式在以下有所描述:Smith and Waterman(1981);Needleman and Wunsch(1970);Pearson and Lipman(1988);Higgins and Sharp(1988);Higgins and Sharp(1989);Corpet 等人(1988);Huang 等人(1992);以及 Pearson 等人(1994)。Altschul 等人(1994)详细描述了序列的排列方法和同系物计算。

[0076] NCBI 基本局部检索工具 (BLAST) (Altschul 等人,1990) 可从几个来源获得,包括国家生物技术信息中心 (NCBI, Bethesda, MD) 和互连网,所述检索工具与序列分析程序 blastp、blastn、blastx、tblastn 和 tblastx 结合使用。它能够在 NCBI 网站上存取。有关如何利用此程序确定序列的相同性的描述可在 NCBI 网站上得到。

[0077] 所公开的蛋白质序列同系物的一般特征是,拥有利用 NCBI Blast 2.0(设定到缺席参数的 gapped blastp) 在具有所公开序列的氨基酸序列的整个长度排列上计数的至少 40% 序列相同性。可调整的参数优选以下值来设定:重叠跨度 1,重叠分数 = 0.125,字阈值 (word threshold) (T) = 11。HSP S 和 HSP S2 是动态值,并且用程序来建立,而程序本身取决于特定序列的组成和检索感兴趣序列的特定数据库的组成;然而,可调整这些值,以增大灵敏度。与参比序列具有更大相似性的蛋白质在用此方法评估时,将显示百分率相同性

增大,诸如至少约 50%、至少约 60%、至少约 70%、至少约 75%、至少约 80%、至少约 85%、至少约 90%或至少约 95%的序列相同性。

[0078] 所公开的核酸序列同系物的一般特征是,拥有利用 NCBI Blast 2.0(设定到缺席参数的 gapped blastp)在具有所公开序列的氨基酸序列的整个长度排列上计数的至少 40%序列相同性。此外,这样的序列在高度严格的条件下与同源序列杂交。一种优选方法采用 WU-BLAST-2 的 BLASTN 模块(Altschul 等人,1996);设定到缺席参数,具有分别设定到 1 和 0.125 的重叠跨度和重叠分数。与参比序列具有更大相似性的核酸序列在用此方法评估时,将显示百分率相同性增大,诸如至少约 50%、至少约 60%、至少约 70%、至少约 75%、至少约 80%、至少约 85%、至少约 90%或至少约 95%的序列相同性。

[0079] 排列包括将间隙引入到待排列的序列中。此外,对于含有比 SEQ ID NO :3 或 SEQ ID NO :6 中描述的蛋白质多或少的氨基酸的序列,应该理解,在一个实施例中,将基于与氨基酸总数相关的相同氨基酸的数目,来确定序列相同性的百分率。这样,诸如,在一个实施例中,将利用更长序列中的氨基酸数目来确定比附图中所示(如下所述)更短的序列的序列相同性。在百分率相同性的计算中,并没有将相对重量分配给序列变异的多种表现形式,例如插入、缺失、取代等。

[0080] 在一个实施例中,仅将相同性记录为正的(+1),而将包括间隙的所有序列变异形式指定为“0”值,从而无需用于序列相似性计算的加权标度或参数(如此处所述)。诸如,可如下计算序列相同性百分率:用配对的相同残基数目除以排列区域中“更短”序列的残基总数,并乘以 100。“更长”的序列是在排列区域中具有大多数实际残基的序列。

[0081] 蛋白质按照它们与相同基因组(平生同源)或不同基因组(直向同源)中的其它蛋白质的序列相关性,来进行分类。直向同源基因是由共同祖先基因的物种形成所进化的基因。这些基因通常保留了与它们进化的相同的功能。平生同源基因是在基因组内复制的基因。这些基因可获得与原始相关的新的物种形成或修正功能。系统发育分析方法对于生物学领域的普通技术人员来说是众所周知的。

[0082] 正如本领域技术人员所理解的,本发明的序列可含有排序误差。也就是说,在任何序列中可以有不正确的氨基酸序列、核苷酸移码、未知核苷酸或其它种类的排序误差;然而,正确的序列将落在此处对核酸的同源和严格定义以及针对蛋白质或多肽所述的蛋白质同源性之内。

[0083] 2-乙酰基-1-吡咯(2AP)多肽:如此处所使用的,术语“2AP 多肽”是指基本上具有 2AP 直向同源的氨基酸序列的基因产物。2AP 多肽的特征是(或部分),其表达的减少、其 mRNA 水平的减少、或者蛋白质的量或活性的减少,导致植物中的化合物 2-乙酰基-1-吡咯的水平增大。2AP 多肽的特征还在于(或部分),具有与 SEQ ID NO :3 或 SEQ ID NO :6 中所描述的氨基酸序列至少约 30%、40%、50%、60%、70%、80%、90%或 95%氨基酸相同的氨基酸序列。

[0084] 基本上相同:“基本上相同”的意思是,多肽或核酸表现出,与参比氨基酸序列(例如,SEQ ID NO :3 或 SEQ ID NO :6 中描述的氨基酸序列)或核酸序列(例如,SEQ ID NO :1、SEQ ID NO :2、SEQ ID NO :4 或 SEQ ID NO :5 中描述的核酸序列)具有至少 30%、优选 50%、更优选 80%、最优选 90%、甚至 95%的同源性。对于多肽,比较序列的长度通常至少为 16 个氨基酸,优选至少 20 个氨基酸,更优选至少 25 个氨基酸,最优选 35 个氨基酸或更

多。对于核酸,比较序列的长度通常至少为 50 个核苷酸,优选至少 60 个核苷酸,更优选至少 75 个核苷酸,最优选 110 个核苷酸或更多。

[0085] 序列的相同性一般利用序列分析软件(例如,基因计算机集团的序列分析软件包(Sequence Analysis Software Package of the Genetics Computer Group),University of Wisconsin Biotechnology Center,1710 University Avenue, Madison, Wis. 53705, BLAST,或者 PILEUP/PRETTYBOX 程序)来测定。诸如,这样的软件在设定到标准参数时,通过将同源性程度分配给多种取代、缺失和/或其它修正,来匹配相同或相似的序列。保守的取代一般包括下列组内的取代:甘氨酸、丙氨酸;缬氨酸、异亮氨酸、亮氨酸;天冬氨酸、谷氨酸、天冬酰胺、谷氨酰胺;丝氨酸、苏氨酸;赖氨酸、精氨酸;和苯丙氨酸、酪氨酸。

[0086] 高水平:高水平,如此处所使用的,是指非天然出现的植物中的化合物 2-乙酰基-1-吡咯的平均水平与相应的天然出现的植物中的化合物 2-乙酰基-1-吡咯的平均水平相比,是增大的。给出,植物中的化合物 2-乙酰基-1-吡咯的水平将依据许多变量,而在植物与植物之间是不同的,本领域的技术人员应该理解,在比较非天然出现的植物与相应的天然出现的植物中的化合物 2-乙酰基-1-吡咯的平均水平时,应该比较在相似受控条件下生长的每种植物的合理尺寸的样品群。化合物 2-乙酰基-1-吡咯的水平优选地在来自每个群体的几种植物中进行测定,并进行平均,以便确定非天然出现的植物是否含有高水平的这种化合物。本发明的非天然出现的植物中的高水平是,按平均水平计,比相应的天然出现的植物大至少约 20%、40%、60%、80%、100%、150%、200%、250%、300%、400%或 500%。

[0087] 启动子:DNA 序列或 DNA 序列组上的识别位点,它为结构基因提供表达控制要素,并且 RNA 聚合酶特异性结合到该位点上,并启动此基因的 RNA 合成(转录)。

[0088] 在 Fraley 等人的 U. S. 5, 352, 605 中发现了利用这些启动子的植物表达构造实例。在大多数转基因植物组织中, CaMV 35S 启动子是强启动子(参见例如, Ode11 等人,自然(Nature)313:810,1985)。CaMV 启动子在单子叶植物中也具有高度活性(参见例如, Dekeyser 等人,植物细胞(Plant Cell)2:591,1990;Terada 和 Shimamoto, Mol. Gen. Genet. 220:389,1990)。而且,通过 CaMV 35S 启动子的复制,此启动子的活性能够进一步增大(即,在 2-10 倍之间)(参见例如, Kay 等人,科学(Science)236:1299,1987;Ow 等人, Proc. Natl. Acad. Sci., U. S. A. 84:4870,1987;和 Fang 等人,植物细胞(Plant Cell)1:141,1989,以及 McPherson 和 Kay, U. S. Pat. No. 5, 378, 142)。

[0089] 其它有用的植物启动子包括(但不限于)胭脂氨酸合成酶(NOS)启动子(An 等人, Plant Physiol. 88:547,1988 以及 Rodgers 和 Fraley, U. S. Pat. No. 5, 034, 322)、章鱼氨酸合成酶启动子(Fromm 等人,植物细胞(Plant Cell)1:977,1989)、玄参镶嵌病毒(FMV)启动子(Rogers, U. S. Pat. No. 5, 378, 619)、以及稻米肌动蛋白启动子(Wu and McElroy, W091/09948)。

[0090] 示范性单子叶植物启动子包括(但不限于)鸭草跖黄斑驳病毒启动子、甘蔗 badna 病毒启动子、水稻 tungro bacilliform 病毒启动子、玉米划线病毒元素和小麦矮化病毒启动子。

[0091] 构造:除非另外指明,否则术语“构造”是指包括本发明的一个或多个分离多核苷酸序列的重组遗传分子。

[0092] 用于宿主有机体内的转基因表达的基因构造包括可操作地链接到开放读框上并任选地链接到开放读框下游的基因端序列 3' 上的基因启动子序列。开放读框可以在有义或反义方向上取向,这取决于基因序列的预计使用。构造还可包括可选择的标记物基因和基因表达的其他调节元素。

[0093] 载体:能够在宿主细胞中复制和 / 或另一 DNA 片段能够可操作地链接到其上从而导致附着片段复制的 DNA 分子。质粒、噬菌粒、粘粒、噬菌体、病毒、YACs 和 BACs 是所有的示范性载体。

[0094] 术语“载体”是指用来将多核苷酸序列引入宿主细胞内、借此产生转化宿主细胞的核酸分子。“载体”除了上述的基因构造之外,还可包括基因材料,例如允许在一个或多个宿主细胞内复制的一个或多个核酸序列,例如复制的起源、可选择的标记物基因以及本领域内公知的其它基因元素(例如,将基因材料并入宿主细胞的基因组内的序列,等等)。

[0095] 转化:转化细胞是已经用分子生物学技术引入核酸分子的细胞。如此处所使用的,术语转化包含将核酸分子引入这样的细胞、植物或动物细胞内的所有技术,包括利用病毒载体的转染、利用质粒载体由土壤杆菌属实施的转化、以及通过电穿孔、质脂转染和粒子枪加速引入裸 DNA,并且包括瞬时稳定的转化体。

[0096] 植物细胞的 DNA 转化方法包括土壤杆菌属介导的植物转化、原生质体的转化、基因转移到花粉内、注入再生器官、注入未成熟的胚胎内和粒子轰击法。这些每种方法都具有不同的优缺点。因此,将基因引入特定植物株系的一种特定方法可以不必是对另一种植物株系最有效的方法,但是哪一种方法对特定植物株系有用,这是公知的。

[0097] 有许多将转化 DNA 片段引入细胞内的方法,但不是所有的方法都适合将 DNA 运送到植物细胞内。合适的方法据信包括诸如通过土壤杆菌属感染、DNA 的直接运送诸如通过 REG 介导的原生质体转化 (Omirulleh 等人,1993)、通过干燥 / 抑制介导的 DNA 摄取、通过电穿孔、通过用碳化硅纤维实施的搅拌、通过 DNA 涂布的粒子的加速等,而将 DNA 引入细胞的任何实际方法。在某些实施例中,加速方法是优选的并包括诸如微粒轰击等。

[0098] 将 DNA 引入细胞的技术对于本领域的技术人员来说是公知的。将基因运送到细胞内的四种通用方法已经有所描述:(1) 化学方法 (Graham and van der Eb,1973 ;Zatloukal 等人,1992);(2) 物理方法例如微注射 (Capecci,1980)、电穿孔 (Wong 和 Neumann,1982 ;Fromm 等人,1985) 和基因枪 (Johnston 和 Tang,1994 ;Fynan 等人,1993);(3) 病毒载体 (Clapp,1993 ;Lu 等人,1993 ;Eglitis 和 Anderson,1988a ;1988b);以及 (4) 受体介导机制 (Curiel 等人,1991 ;1992 ;Wagner 等人,1992)。

[0099] 电穿孔:将短暂的高压电脉冲施加到各种动物和植物细胞上,导致在质膜上形成纳米大小的孔。DNA 通过这些孔或作为伴随着孔闭合的膜组分的重新分布结果,而直接被摄取到细胞胞质内。电穿孔是极其有效的,并可用于克隆基因的瞬时表达和建立运载感兴趣基因的一体拷贝的细胞系。电穿孔,与磷酸钙介导的转染和原生质体的融合相反,其频繁地作用于运载一个或者至多几个一体拷贝的细胞系上。

[0100] 借助于电穿孔引入 DNA 的方法对于本领域的技术人员来说是公知的。在这种方法中,利用某些细胞壁降解酶诸如果胶降解酶使靶受体细胞比未处理的细胞更易受电穿孔的转化。或者是,使受体细胞更易受机械创伤的转化。为了使电穿孔的转化有效,一种方法是采用易碎组织例如细胞或胚胎发育膀胱的敏感培养基,或者是直接转化未成熟的胚胎或其

它有机化组织。一种方法是通过将所选细胞暴露于果胶降解酶（果胶酶）中或者以受控的方式进行机械创伤，而部分降解所述细胞的细胞壁。这些细胞然后通过可以在此阶段完成的电穿孔接受 DNA 转染，并且然后用合适的选择或筛选方案（这取决于新并入的 DNA 的性质）来鉴定所转化的细胞。

[0101] 微粒轰击：将转化 DNA 片段运送到植物细胞中的另一种有益的方法是微粒轰击。在这种方法中，粒子用核酸涂布并利用推进力运送到细胞内。示范性粒子包括用钨、金、铂等组成的粒子。

[0102] 微粒轰击除了是一种有效的重复获得稳定转化的单子叶植物的方式之外，它的另一个优点是，不需要分离原生质体 (Cristou 等人, 1988)，也无需承受土壤杆菌属的转染。通过加速将 DNA 运送到玉米细胞内的一种方法的图解实施例是生物列表 (Biolistic) 粒子运送系统，该系统用来将用 DNA 或细胞涂布的粒子通过筛子（例如不锈钢或 Nytex 筛）推到用培养在悬浮液中的谷物细胞涂布的过滤器表面上。筛将这些粒子分散，以便粒子不会以大凝聚体的形式运送到受体细胞中。据信，介于抛射装置与待轰击细胞之间的筛子减小了抛射凝聚体的尺寸，并且通过减小由太大的抛射造成的施加在受体细胞上的损害，可以使转化的频率更高。

[0103] 对于轰击，悬浮液中的细胞优选地集中在过滤器或固态培养介质上。或者是，未成熟的胚胎或其它靶细胞可以安置在固态培养介质上。使待轰击的细胞位于微粒终止板之下的合适距离处。如果需要的话，一个或多个筛子也可位于加速装置与待轰击的细胞之间。通过使用本文提到的技术，可以获得高达 1000 或更多的瞬时表达标记物基因的细胞焦点。48 小时轰击后表达外源性基因产物的焦点处细胞数经常在 1-10 的范围内，平均为 1-3。

[0104] 在轰击转化中，可以优化预先轰击培养条件和轰击参数，以产生最大数目的稳定转化体。轰击的物理和生物参数在此技术中都是重要的。物理因数是包括操纵 DNA/微粒沉淀物或影响大颗粒或微粒的飞行和速度的那些因数。生物因数包括在轰击之前或轰击之后立即操纵细胞的所有步骤、靶细胞的渗透调节以有助于减轻有关轰击的损伤，还有转化 DNA 的性质，例如线性化 DNA 或完整的超螺旋质粒。据信，预先轰击的操纵对于未成熟胚胎的成功转化是重要的。

[0105] 据此，可设想在小规模的研究中调整一些轰击参数，以充分优化这些条件。特别希望调整物理参数例如间隙的距离、飞行距离、组织的距离和氢压。通过调整影响受体细胞的生理状态并因此影响转化和整合效率的条件，还可以使损失减小因数 (TRFs) 最小。例如，为了优化转化，可调整渗透状态、组织水合以及受体细胞的传代培养阶段和细胞周期。

[0106] 土壤杆菌介导的转移：土壤杆菌介导的转移是将基因引入植物细胞的广泛应用的系统，这是因为 DNA 可以引入到全部植物组织内，借此无需由原生质体再生完整的植物。利用土壤杆菌介导的植物整合载体将 DNA 引入到植物细胞内，这在本领域内是公知的。参见例如，(Fraley 等人, 1985 ;Rogers 等人, 1987) 中所述的方法。另外，Ti-DNA 的整合是导致很少重排的相当精确的过程。待转移的 DNA 区域由边界序列限定出，并且间插 DNA 经常插入植物基因组内，如 (Spielmann 等人, 1986 ;Jorgensen 等人, 1987) 所述。

[0107] 现代的土壤杆菌转化载体能够在大肠杆菌 (*E. coli.*) 以及土壤杆菌中复制从而允许便利地进行操纵，如 (Klee 等人, 1985) 所述。而且，在土壤杆菌介导的基因转移载体上的最近技术进展，已经改善了载体中的基因排列和限制位点，从而易于构建能够表达编码

基因的多种多肽的载体。(Rogers 等人,1987) 中所述的载体,具有便利的由启动子和聚腺苷酸化位点侧翼包围的多个接头区域,以便直接表达所插入的编码基因的多肽,并适合本发明的目的。此外,含有接臂和无臂 Ti 基因的土壤杆菌可用于转化。在土壤杆菌介导的转化有效的那些植物株系中,由于基因转移的易得和确定性质,这是所选择的方法。

[0108] 叶盘和其它组织例如子叶和胚轴的土壤杆菌介导的转化似乎局限于土壤杆菌天然感染的植物。土壤杆菌介导的转化在双子叶植物中最有效。很少的单子叶植物表现出是土壤杆菌的天然宿主,尽管利用天然杆菌载体已经在芦笋中产生转基因植物,如 (Bytebier 等人,1987) 所述。因此,商业上重要的谷粒诸如稻米、玉米和小麦必需经常用等同的方法来转化。

[0109] 利用土壤杆菌转化方法形成的转基因植物一般在一个染色体上含有一个基因。这样的转基因植物被称作外加基因的杂种。然而,由于词汇“杂种”的使用经常暗示在一对染色体的第二个染色体的相同基因座存在互补基因,并且在含有一个外加基因的植物中没有这样的基因存在,因此据信,对这种植物的更准确命名是一个单独分离的部分,因为外加的外源性基因在有丝分裂和减数分裂的过程中是单独分离的。

[0110] 更优选的是,转基因植物是外加结果基因的杂种;即,含有两个外加基因的转基因植物,一个基因在染色体对的每个染色体上的相同基因座处。杂合转基因植物可如下获得:使含有一个外加基因的单分离的转基因植物有性交配(自交),然后使所生成的一些种子萌发,并分析所产生的植物相对于对照物(天然、非转基因的)或单独分离的转基因植物升高的 2-乙酰基-1-吡咯水平。

[0111] 这些系统在不同植物株系中的应用取决于由原生质体再生特定植物株系的能力。图示的由原生质体再生谷类的方法在 (Fujimura 等人,1985 ;Toriyama 等人,1986 ;Yamada 等人,1986 ;Abdullah 等人,1986) 中有所描述。

[0112] 为了转化不能由原生质体成功产生的植物株系,可采用将 DNA 引入完整细胞或组织的其它方式。例如,如 (Vasil,1988) 所述,用未成熟胚胎或外植体再生谷类。此外,可采用“粒子枪”或高速微粒技术 (Vasil,1992)。

[0113] 利用后一种技术,小金属粒子表面上的 DNA 通过细胞壁并进入胞质内,如 (Klein 等人,1987 ;Klein 等人,1988 ;McCabe 等人,1988) 所述。金属粒子穿过细胞的几个层,由此使细胞在组织外植体内转化。

[0114] 分离的:“分离的”生物组分(例如核酸或蛋白质或细胞器)已经从组分在其中天然出现的细胞或有机体内的其它生物组分中基本上分离出来或提纯出来,所述的其它生物组分为其它染色体或外染色体 DNA 和 RNA、蛋白质和细胞器。已经“分离的”的核酸及蛋白质包括用标准的纯化方法提纯的核酸及蛋白质。此术语包含包括化学合成的核酸的核酸,也包含通过在体外或宿主细胞内重组表达而制备的蛋白质,和重组核酸(如下定义)。作为一个实例,大基因组 DNA 片段例如重叠群中的基因,由于在平均重叠群中发现的相当大量的外 DNA,而不能充分地将从将要考虑分离的其它生物组分中提纯出来。概括地说,以下的“重组核酸”及“重组蛋白质”也是如上所述进行“分离的”。

[0115] 重组体:“重组核酸”在此处是指这样的核酸:这种核酸具有不是天然出现的序列,或者具有通过将两个另外分离的序列片段人工组合而生成的序列。此人工组合经常通过以下方式来实现:化学合成,或者更通用的是核酸的人工操纵,例如通过基因工程技术,

诸如通过用限制酶、连接酶、重组酶和 / 或聚合酶操纵至少一个核酸。一旦重组核酸被引入宿主细胞,就由宿主细胞来复制;然而,曾经在宿主细胞内复制的重组核酸保留了用于本发明目的的重组核酸。“重组蛋白质”在此处是指,由采用重组核酸的方法产生的蛋白质。概括地说,以上的“重组核酸”及“重组蛋白质”也是如上所述进行“分离的”。大片段例如重叠群中的基因不是“重组核酸”,如果这样的人工组合与此基因无关的话。然而,如果重叠群中的基因周围或内部的序列为了与该基因相关已经被操纵(及不仅仅是因为该基因在重叠群附近),那么重叠群中的这种基因,由于核酸的重组部分相对接近可疑基因,而将构成“重组核酸”。

[0116] 非天然出现的植物:如此处所用的,术语“非天然出现”在用于指植物时,是指植物已经通过人的干涉在基因上进行了修饰,从而改变了植物的化合物 2-乙酰基-1-吡咯的水平。已经进行基因修饰的天然出现的植物被称作“控制”植物。在本发明的上下文中,控制植物可以是转基因植物。例如,控制植物可以是抗除莠的转基因植物。在此例中,非天然出现的植物是抗除莠的,并且与抗除莠的控制植物相比具有高水平的化合物 2-乙酰基-1-吡咯。本发明的转基因植物诸如是,含有编码 2AP 基因或其片段的外源性核酸分子并因此已经通过人类干涉进行基因修饰的非天然出现的植物。此外,在诸如 2AP 基因调节元件或编码序列中含有突变的植物,也被认为是非天然出现的植物,这是因为它已经通过人类干涉进行了基因修饰,其中所述突变是所计算的暴露于诱变剂例如化学诱变剂或“插入诱变剂”(诸如转座子或 T-DNA)的结果。反之,含有仅仅自发或天然出现的突变的植物不是此处所定义的“非天然出现的植物”。本领域的技术人员应该理解,虽然非天然出现的植物与天然出现的植物相比一般具有改变的核苷酸序列,但是非天然出现的植物例如通过修饰其甲基化图,也能够通过人类干涉进行基因修饰,而不会改变其核苷酸序列。

[0117] 转基因植物:即,由转化的植物细胞或原始质体衍生的植物或其子代,其中植物 DNA 含有不是原始存在于相同株系的天然、非转基因植物中的引入的外源性 DNA。术语“转基因植物”和“转化植物”有时在本领域内是作为同义词使用的,用来定义其 DNA 含有外源性 DNA 分子的植物。然而,应该更科学地将转化的植物细胞或原生质体获得的再生植物或愈伤组织成为转基因植物,并且其用途在此处如下所述。

[0118] I. 2-乙酰基-1-吡咯(2AP)基因

[0119] A. 直向同源/同系物的分离:描述稻米中 Os2AP 基因的分离和特征的以下实例部分,是分离 2AP 基因的通用方法示范。这样的 2AP 基因编码 2AP 蛋白质。所分离的基因然后用来构建减小植物中 2AP 基因表达的重组载体。

[0120] 通常,以下所述的重组 DNA 技术中的命名和方案在本领域内是公知的并普遍采用的。标准的技术用于克隆、DNA 和 RNA 的分离、扩增以及纯化。涉及 DNA 连接酶、DNA 聚合酶、限制内切核酸酶等的酶反应通常按照制造者的具体说明来实施。这些技术和多种其它技术通常按照 Sambrook 等人的分子克隆-实验室手册(Molecular Cloning-A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y., (1989)) 来实施。

[0121] 2AP 基因的分离可利用许多技术来完成。例如,基于本文公开的序列的寡核苷酸探针可用来鉴定 cDN 或基因组 ADNA 文库中的所需 DNA。为了构建基因组文库,基因组 DNA 的大节段通过诸如利用限制性内切核酸酶的随机断裂而产生,并与载体 DNA 连接,从而形成能够包装成合适载体的多联体。为了制备 cDNA 文库,将 mRNA 从所需的器官例如叶中分

离出来,并且含有 2AP 转录物的 cDNA 文库用 mRNA 来制备。或者是,cDNA 可以用从表达 2AP 基因或同系物的其它组织中提取的 mRNA 来制备。

[0122] 然后,可利用基于克隆 2AP 基因(例如本文公开的稻米 Os2AP 基因)序列的探针筛选 cDNA 或基因组文库。探针可用来与 DNA 或 cDNA 序列杂交,从而将相同或不同植物物种中的同源基因分离出来。

[0123] 具有序列区的核酸分子尤其预计作为用于诸如南北印迹法的杂交探针,其中所述序列区包括与 SEQ ID NO :1、SEQ ID NO :2、SEQ ID NO :4 或 SEQ ID NO :5 中描述的 DNA 序列相同或互补的 10-14、15-20、30、50、甚至 100-200 个核苷酸等的邻接核苷酸延伸序列。较小的片段通常发现用于杂交实施例,其中邻接互补区的长度可以诸如在约 10-14 以及 100 或 200 个核苷酸之间变化,但是较长的邻接互补延伸序列可按照希望检测的互补序列长度来使用。

[0124] 约 14 个核苷酸长度的杂交分子的使用,允许形成稳定的、可选择的双螺旋分子。在超过 14 个碱基长度的延伸序列上具有邻接互补序列的分子通常是优选的,以便增加杂交的稳定性和选择性,借此提高所得到的特定杂交分子的质量和杂交度。通常优选的是,设计具有 15-20 个邻接核苷酸、甚至更长(如果需要的话)的基因互补延伸序列的核酸分子。

[0125] 当然,片段也可利用其它技术诸如通过机械剪切或通过限制酶消化来获得。通过诸如利用化学方式直接合成片段,可容易地制备小核酸节段或片段,诸如利用自动寡核苷酸合成器所普遍实践的。而且,通过应用核酸再生技术例如 U. S. 4, 683, 195 和 4, 683, 202(每一篇文献作为参考引入本文)的 PCR. TM. 技术、通过将所选序列引入重组载体用于重组制备、以及通过分子生物学领域内的技术人员通常公知的其它重组 DNA 技术,可获得这些片段。

[0126] 据此,本发明的核苷酸序列可用于同 DNA 片段的互补延伸序列选择性形成双螺旋分子。根据设想的应用,期望采用不同的杂交条件,以获得不同选择程度的针对靶序列的探针。对于需要高度选择的应用,一般期望采用相当严格的条件来形成杂交,例如选择相当低的盐和/或高温条件,诸如约 0.02M-约 0.15M NaCl 和约 50°C -约 70°C 的温度。这样的选择性条件容许探针与模板或靶链之间具有很少的错配序列(即便有的话),并且特别适合于分离编码 DNA 节段的 2AP 多肽或蛋白质。通过杂交对 DNA 节段的检测,对于本领域的技术人员来说是公知的,并且 U. S. 4, 965, 188 和 5, 176, 995(每一篇文献作为参考引入本文)的教导是杂交分析方法的示范。诸如在 Maloy 等人,1994 ;Segal 1976 ;Prokop,1991 ;和 Kuby,1994 的正文中发现的那些教导是特别相关的。

[0127] 当然,对于一些应用,诸如当需要制备采用杂交到下面模板上的突变型引物链的突变体时,或者在寻求从相关物种、功能性等同物或类似物质中分离 2AP 蛋白质编码序列时,一般需要不太严格的杂交条件,以便形成异源双链体。在这些情况下,期望采用诸如约 0.15M-约 0.9M 的盐和约 20°C -约 55°C 的温度之类的条件。交叉杂交物种借此能够容易被鉴定为相对于对照杂交的正杂交信号。在任何情况下,通常认为,通过加入大量的甲酰胺而使条件更加严格。这样,杂交条件容易操纵,并由此通常作为根据所需结果的选择方法。

[0128] 在某些实施例中,有益的是,将本发明的核酸序列与合适的方式例如标记物结合使用,来测定杂交。各种合适的指示剂在本领域内是公知的,包括能够给出可检测信号的荧光物质、放射性物质、酶或其它配体例如亲和素/生物素。在优选的实施例中,可能期望采

用荧光标记物或酶标记例如脲酶、碱性磷酸酶或过氧化物酶,而不是放射性或其它对环境有害的试剂。在酶标记的情形中,量热指示剂底物是公知的,其能够用来提供人眼可视的或分光光度测量的方式,从而能够鉴定与含有互补核酸的样品的特定杂交。

[0129] 通常,预想,本文所述的杂交探针可用作溶液杂交中的试剂以及用于采用固相的实施例。在涉及固相的实施例中,被测 DNA(或 RNA)被吸附或固定到所选择的基质或表面上。此固定的单链核酸然后在所需的条件下接受与所选探针的杂交。所选择的条件将取决于基于所需的特定临界(诸如取决于 G+C 内容、靶核酸的类型、核酸的来源、杂交探针的尺寸等)的特定情况。在洗涤杂交表面以除去非特异结合的探针分子之后,借助于标记物对特异性杂交进行检测、甚至定量。或者是,利用扩增技术从核酸样品扩增感兴趣的核酸。例如,聚合酶链反应(PCR)技术可用来直接从 mRNA、cDNA、基因组文库或 cDNA 文库扩增 2AP 基因序列。PCR 以及其它体外扩增方法也可用于诸如克隆为待表达的蛋白质编码的核酸序列、制备用作检测样品中所需 mRNA 的存在、核酸测序或其它目的探针的核酸。

[0130] 用于从植物组织中鉴定 2AP 基因序列的核酸引物及探针,是由本文提供的序列比较产生的。对于 PCR 的综述,参见 PCR 方案:方法和应用指南(A Guide to Methods and Applications. (Innis, M, Gelfand, D., Sninsky, J. 和 White, T., eds.), Academic Press, San Diego(1990),此文献作为参考引入本文。

[0131] 多核苷酸也可以用诸如技术文献中描述的公知技术来合成。参见例如, Carruthers 等人, Cold Spring Harbor Symp. Quant. Biol. 47 :411-418(1982), 和 Adams 等人, J. Am. Chem. Soc. 105 :661(1983)。然后,通过合成互补链并在合适条件下使这些链一起退火,或者通过利用 DNA 聚合酶将互补链与合适的因为序列加成,可获得双链 DNA 片段。

[0132] B. 从中分离出 2AP 基因直向同源的有机体:可使用许多种植物和真菌,包括来自普通 Zea, Avena, Hordeum, Secale, Triticum 和 Sorghum 的物种。已经表明,2-乙酰基-1-吡咯是所有芳香稻米、小麦和黑面包(Buttery 等人,1982,1983)、湿粟(Seitz 等人,1993)、爆米花(Schieberle 等人,1991)、蜡样芽孢杆菌(Romanczyk 等人,1995)和一些真菌物种例如 Aspergillus oryzae、Aspergillus awamori 和 Sporobolus virginicus(Nagsuk 等人,2004)的主要潜在香味化合物。有几个报告表明谷类中的基因组宽的同线性以及植物和微生物中良好保存的脯氨酸代谢途径(在京都 Kyoto 大学的化学研究所生物信息中心的基因和基因组的京都百科全书的网页上联机发现的)(found online at the web page for the Kyoto Encyclopedia of Genes and Genomes, Kyoto University, Bioinformatics Center, Institute for Chemical Research)。因此,2AP 基因的直向同源在谷类植物中可以比较的方式起作用,并且本发明提供的教导在其它植物中能够以与在稻米中类似的方式增强化合物 2-乙酰基-1-吡咯的积累。

[0133] II. 减小 2AP 基因表达或 2AP 蛋白质水平或活性,以增加植物中的 2-乙酰基-1-吡咯的水平

[0134] A. 诱变:化学诱变剂例如 EMS(甲磺酸、乙酸乙酯)和 γ 射线以及快速中子辐射能够在 DNA 中产生突变。突变的一些实例是缺失、插入和错义突变。在突变之后,进行筛选,以鉴定产生终止提前密码子或非功能性 2AP 基因的缺失。通过测序或通过利用由本发明提供的特异于 2AP 基因或蛋白质的探针或引物,可实施突变体的筛选。2AP 基因的特异性突变是也可通过 TILLING(靶向诱导的基因组的局部损害)和 tDNA 插入来产生。这样的突变可

导致 2AP 基因表达的减小、2AP mRNA 稳定性的减小或者 2AP 蛋白质的活性和 / 或稳定性的减小。这样的植物如本文所定义的,是非天然出现的。

[0135] B. 反义技术:反义技术阻止靶 mRNA(例如 2AP mRNA)的翻译。如此,反义技术通常减小靶蛋白质(此处是 2AP 蛋白质)的水平。2AP 基因或其片段能够引入到反义取向的植物内。这些片段可以象 18 个核苷酸那么小,象 3000 个核苷酸那么大或更大。2AP 基因的 cDNA 片段可以被克隆到与天然基因相反取向的载体(例如, pCAMBIA1302)内。反向转录物将与天然 2AP 基因转录物形成异源双链结构,该结构然后在转录之前发生降解。有义和反义结构不必相同,甚至部分同源就足以抑制靶基因的表达。于是,来自稻米 Os2AP (SEQ ID NO :1、SEQ IDNO :2、SEQ ID NO :4 和 SEQ ID NO :5)的序列能够用来抑制其它植物中的 2AP 基因的表达,而无需知道该植物的 2AP 基因直向同源序列。

[0136] C. RNA 干扰:RNA 干扰是另一种适合消除靶 mRNA 的技术。在 RNAi 载体的构建中有若干变异,但是基本的结果是在 RNA 中形成发夹环结构 (Horiguchi2004)。在例 2 中,在与其 cDNA 相反方向上包含外显子 6、7 和 8 的 Os2AP 核苷酸序列片段被克隆到载体内,从而导致在 mRNA 中形成反向发夹环结构。利用名为 Dicer 的内切核酸酶,将这些发夹结构裂解成若干小的片段,所述核酸酶的作用是避免错误的转录被翻译 (Hamilton 和 Baulcombe, 1999, Matzke 等人, 2001)。MRNA 的小片段通常约 20-21 个核苷酸的大小,并且已经被命名为 siRNAs 或小的干涉 RNAs。也可采用更大或更小的片段。这些 siRNAs 能够下调同源基因的表达。再者,同源性不必是完全的,因此本发明详述的稻米序列能够用来产生其它植物的 RNAi 构造。例如 2 表示出靶向非芳香米株系中的 Os2AP 基因的 RNAi 实验是如何将 2-乙酰基 1-吡咯的水平 and 香气升高到芳香米中发现的水平。

[0137] III. 转基因植物的产生

[0138] 为了利用以前技术中所分离的 2AP 基因序列,可制备适合植物细胞转化的重组 DNA 载体。转化各种更高级植物物种的技术是公知的并且在科学技术文献中有所描述。参见例如, Weising 等人的 An. Rev. Genet. 22 :421-477 (1988)。

[0139] 利用各种常规技术,将这样的 DNA 构造引入所需植物宿主的基因组内。例如,利用诸如电穿孔、PEG 扩散 (poration)、粒子轰击和植物细胞原生质体或胚胎发育愈伤组织的微注射之类的技术,可以将 DNA 构造直接引入植物细胞的基因组 DNA 中,或者利用弹道方法例如 DNA 粒子轰击,将 DNA 构造直接引入到植物组织内。或者是, DNA 构造可以与合适的 T-DNA 侧翼区组合,并引入到常规的 *Agrobacterium tumefaciens* 宿主载体内。当细胞被细菌感染时, *Agrobacterium tumefaciens* 宿主的毒力作用将引导构造和相邻标记物插入到植物细胞的 DNA 内。微注射技术在本领域内是公知的,并且在科学技术文献中有较好的描述。利用聚乙二醇沉淀物引入 DNA 构造,这在 Paszkowski 等人的 Embo J. 3 :2717-2722 (1984) 中有所描述。电穿孔技术在 Fromm 等人的 Proc. Natl. Acad. Sci. USA 82 :5824 (1985) 中有所描述。弹道转化技术在 Klein 等人的 Nature 327 :70-73 (1987) 中有所描述。利用许多方法,可以转化谷类物种例如黑麦 (de la Pena 等人, Nature 325 :274-276 (1987))、玉米 (Rhodes 等人, Science 240 :204-207 (1988)) 和稻米 (Shimamoto 等人, Nature 338 :274-276 (1989), 通过电穿孔; Li 等人的 Plant Cell Rep. 12 :250-255 (1993), 通过弹道技术)。

[0140] *Agrobacterium tumefaciens* 介导的转化技术在科学文献中有较好的描述。参见

例如, Horsch 等人的 Science 233 :496-498(1984) 和 Fraley 等人的 Proc. Natl. Acad. Sci. USA 80 :4803(1983)。虽然 Agrobacterium 主要用于双子叶植物, 但是某些单子叶植物也能用 Agrobacterium 来转化。例如, 稻米的 Agrobacterium 转化在 Hiei 等人, Plant J. 6 : 271-282(1994) 中有所描述。

[0141] 可以培养用上述任何转化技术衍生的转化植物细胞, 以便再生拥有所转化的基因型并由此拥有高水平的 2-乙酰基-1-吡咯和较大香气的整个植物。这种再生技术依赖于组织培养生长介质中某些植物激素的操纵, 一般依赖于已经与 2AP 核苷酸序列一起引入的抗微生物剂和 / 或除莠剂。由所培养的原生质体再生植物, 这在 Evans 等人, 原生质体的分离和培养 (Protoplasts Isolation and Culture), Handbook of Plant Cell Culture, pp. 124-176, MacMillan Publishing Company, New York, 1983 ; 和植物、植物原生质体的结合、再生 (Binding, Regeneration of Plants, Plant Protoplasts), pp. 21-73, CRC Press, Boca Raton, 1985 中有所描述。再生也可从植物愈伤组织、外植体、器官、或其一部分得到。这样的再生技术通常在 Klee 等人的 An. Rev. of Plant Phys. 38 :467-486(1987) 中有所描述。

[0142] 本领域的技术人员将认识到, 表达盒在稳定并入转基因植物并确定是可操作的之后, 它能够通过有性交配引入到其它植物内。根据待交配的物种, 可采用许多标准繁殖技术中的任一种。

[0143] IV. 测试 2-乙酰基-1-吡咯水平

[0144] A. 感觉评估 : 感觉干种子、熟米或地面叶子的挥发性香气 (Dhulappanavar, 1976, Ghose 等人, 1952, Kadam 等人, 1938)。这样的感觉包括品尝测试。在稻米种植者中另一种受欢迎的实践是, 在水中加热叶组织, 然后施加稀 KOH 溶液 (Sood and Siddiq, 1978)。这些测试并不总是一致和可靠的, 也倾向于人的误差和偏爱。

[0145] B. 色谱学 : 建立了一种更可靠的利用气相色谱的方法, 用于定量 100 克熟米中的挥发性化合物 (Petrov 等人, 1995)。最近, 开发了气相色谱 / 质谱 (GCMS) 技术, 用于分析仅 1 克米中的十亿分之一那么小的 2-乙酰基-1-吡咯水平 (Mahatheeranont 等人, 1995)。如此, 利用 GCMS 来测试天然出现和非天然出现的植物中的 2-乙酰基-1-吡咯水平。参见图 1。

[0146] 优选实施例的描述

[0147] 例 1 : 控制稻米香气的基因的定位克隆 : 以前已经基于定性和定量香气测定方法, 在染色体 8 上对稻米的芳香特征进行了作图 (Lorieux 等人, 1996)。已经在距离 RG28 4.5 cM 处 (Ahn 等人, 1992) 并且在 RG28 与 RG1 之间的 12 cM 之内 (Lorieux 等人, 1996) 对芳香基因进行作图。最近, 已经在距离芳香基因 2cM 处, 对由公共稻米基因组序列开发的单核苷酸多态性 (SNP) 标记物 RSP04 进行作图 (Jin 等人, 2003)。

[0148] 单个 QTL 位于由染色体 8 上的 RFLP 标记物 RG1 和 RG28 侧翼包围的 4.5 cM 区域上 (Lorieux 等人, 1996 ; Lanceras 等人, 2000)。来自 KDML105 的 EcoRI BAC 文库包括 75, 264 个代表单倍体基因组的克隆。包括 25 个 BAC 克隆的初始 BAC 重叠群用该区域内的六个图标记物来选择 (Wanchana 等人, 2005), 并进一步用 HindIII 指纹法来精制。为了鉴定 2-乙酰基-1-吡咯累积中涉及的基因, 利用以下策略使临界区域向下变窄。

[0149] 最初, 将为 F7 中的粒香分离的单个 F6 植物选为在临界区内产生新的重组体的来

源。从 F7 至 F12, 对于每个生成, 基于粒香和标记物的基因型来选择杂种植物。在 F10 的生成中, 通过利用从 KDML105 BAC 末端序列开发的八个多态性标记物筛选 374 个 F10 植物, 而使临界区域向下变窄到 380kb。通过筛选 274 个 F11 植物, 该区域进一步精制到 120 kb (图 2)。基于图形基因型和 2-乙酰基-1-吡咯的量, 选择 8 个 F11 植物, 以产生芳香的、非芳香的和杂种同基因系 (图 2)。在此阶段, 为鸟枪测序选择 3 个 KDMLBAC 克隆、155L11、68L13 和 167M23。来自 KDML 的序列和相应的 Nipponbare 基因组重叠群的排列, 导致 21 个插入/缺失 (“indel”) 标记物的鉴定, 以便在 F12 的产生中实施进一步的筛选。通过利用 1, 116 个 F12 植物, 靶区在 58 kb BAC, 167M23 之内向下变窄到 27kb (图 2)。在此交配中没有鉴定出进一步的重组。开发出几对等基因系, 从而清楚地表明, 米粒中的 2-乙酰基-1-吡咯的累积差异 (图 6)。

[0150] 序列分析揭示出, BAC 167M23 含有 19 个基因。然而, 仅有 10 个基因与公知蛋白或编码序列具有相似性。在 27kb 区域之内, 三个候补基因被鉴定为: 3-甲基巴豆酰-CoA 羧基酶 (MCCase)、假拟基因和未知蛋白质。利用来自 KDML105 与 Jao Hom Nin (JHN) 之间的交配的 177 个 F6 植物, 在影响米粒香气和 2-乙酰基-1-吡咯水平的未知蛋白质的外显子 7 之内鉴定出非芳香的黑米 (一个重组位点)。为了研究来自 BAC 167M23 和 BAC 68L13 的这 7 个候补基因的表达, 当 2-乙酰基-1-吡咯在米粒中累积时, 在米粒填充过程中实施 RT-PCR。在芳香和非芳香同基因系传粉之后的 10、15 和 20 天从水稻圆锥花中收集总 RNA。对于 NBS/LRR 基因, 没有检测出表达。大多数的互补基因在芳香与非芳香同基因系之间并没有显示出不同的表达。仅仅在 0s2AP 的情形中, 在传粉之后的 15 天芳香同基因系中的基因表达急剧下降 (图 3A)。0s2AP 基因表达的急剧下降在叶、干和根中也出现了 (图 3B)。因此, 表达研究和定位克隆都支持 0s2AP 作为负责米粒香气和 2-乙酰基-1-吡咯合成的体内累积的调节剂。

[0151] 研究者以前已经报道控制米粒香气的单个隐性核基因 (Berner 和 Hoff, 1986; Yanjuan 等人, 1992; Ali 等人, 1993)。这符合本文报道的发现。比较芳香和非芳香同基因系以及 F1 植物中的 2-乙酰基-1-吡咯的累积。结果表明, 在 F1 叶中检测不到 2-乙酰基-1-吡咯。由于 F2 产生中的分离, 因此 2-乙酰基-1-吡咯的水平是在供体亲代中发现的水平四分之一 (图 3C)。因此, 经典的和分子遗传学都支持, 外显子 7 中的突变是调节植物中 2-乙酰基-1-吡咯的累积的分子机制。

[0152] 比较 KDML105、芳香同基因系、非芳香同基因系和 Nipponbare 中的 0s2AP 基因的结构。5.8kb 0s2AP 基因包括 15 个具有在外显子 2 发现的几个同义突变的外显子 (图 5A)。对于 KDML105 和芳香同基因系, 在外显子 7 鉴定出两个重要的突变事件。首先, 在位置 730 (A 至 T) 和 732 (T 至 A) 发现两个传递突变, 随后 8 碱基对缺失 ‘GATTAGGC’ 始于位置 734。在 8 个同基因系中 2-乙酰基-1-吡咯水平的分析表明, 8 碱基对缺失与米粒中的 2-乙酰基-1-吡咯的累积相关 (图 6A)。此突变导致移码翻译始于位置 729、终止提前密码子始于位置 753 (图 5B)。在 Nipponbare 中, 0s2AP 的整个长度 cDNA 被翻译成 503 个氨基酸。缺失产生 252 个氨基酸的截短肽 (图 5B)。在由 KDML105 和 JHN 衍生的 177 个 F6 植物中, 侧翼包围 8 碱基对缺失的双重重组解释在米粒中生成 2-乙酰基-1-吡咯的失败。

[0153] 此终止提前密码子对 0s2AP 的表达可具有显著作用。含有翻译终止提前密码子的大多数 mRNAs 经常不能被翻译。据此, 它们经常触发无义介导的 mRNA 的衰变 (NMD), 即其作

用是减少基因表达中的错误的监督系统 (Pulak 和 Anderson, 1993)。此现象可解释芳香米中的 Os2AP 基因表达的低水平。

[0154] 可能 Os2AP 可能对脯氨酸的代谢路径起作用。由脯氨酸合成谷氨酸需要脯氨酸脱氢酶 (ProDH) 和 δ -1-脯氨酸-5-羧基酶脱氢酶 (P5CDH)。在米粒填充期间,不能检测到脯氨酸脱氢酶的表达,同时非芳香同基因系中的 Os2AP 基因的表达上调。因此,在非芳香稻米中 Os2AP 可代替 ProDH,而在芳香稻米中,Os2AP 的降解可移动脯氨酸集合体,以便用于 2-乙酰基-1-吡咯的合成。同位素标记实验也支持这个假说 (Yoshihashi 等人, 2002)。

[0155] 例 2:利用 Os2AP 基因的植物转化:Os2AP 基因并入用来下调稻米中的 Os2AP 的基因表达并增强 2-乙酰基-1-吡咯水平的 RNAi 构造。

[0156] 含有载体的 Os2AP-ihp2AP 的构建:pCambia1302 载体用于表达 Os2AP 基因的有义-反义片段。有义和反义片段是分别利用来自 KDML105 的 DNA 和总 RNA 作为模板、通过 PCR 制备的。含有对应于外显子 6-9 的基因组 DNA 序列的有义片段,用含有 NcoI 和 SpeI 限制位点(下划线)的引物来扩增(正向: AATTCCATGGGGTTGGTCTTCCTTCAGGTG (SEQ ID NO: 80);反向: AATTACTAGTTTCCACCAAGTCCAGTGAA (SEQ ID NO: 81))。含有对应于外显子 6-8 的 cDNA 序列的反义片段,用含有 NcoI 和 NheI 限制位点的引物来扩增(正向: AATTCCATGGGGTTGGTCTTCCTTCAGGTG (SEQ ID NO: 80);反向: AATTGCTAGCGGTCCAAAAGCAACCAAAGA (SEQ ID NO: 82))。PCR 产物首先用 SpeI 和 NheI 来消化,然后用 T4 DNA 连接酶来连接。所连接的片段然后用 NcoI 来消化。将纯化的连接片段在 NcoI 克隆位点克隆成 pCambia1302 载体(图 7A)。

[0157] 稻米的转化:非芳香稻米变种 (*Oryza sativa* L. japonica 变种 Nipponbare) 的胚胎发育愈伤组织用作粒子轰击转化的靶组织 (Nimlek, 1999)。利用引物 Os2AP-外显子 7.1-del_F 和 R(U: 5'-TGCTCCTTTGTCATCACACC-3' (SEQ ID NO: 54) 和 L: 5'-TTTCCACCAAGTCCAGTGA-3' (SEQ ID NO: 55)) 以及 GFP 引物 (U: 5'-CTTGTTGAATTAGATG GTGATGTT-3' (SEQ ID NO: 83) 和 L: 5'-GTTGTGGGAGTTGTAGTTGTATTC-3' (SEQ ID NO: 84)、通过 PCR,来筛选耐潮霉素的愈伤组织。

[0158] Southern 印迹分析:经总 DNA 从转基因植物 (R0) 和对照植物 (Nipponbare) 中分离出来。基因组 DNA (10mg) 用 NcoI 消化,以便检测 Os2AP-ihpRNA 片段。作为阳性对照物,从质粒 Os2AP-RNAi 中分离的 DNA 用 NcoI 消化。在通过 0.8% 的琼脂糖凝胶进行电泳之后, DNA 被转移到 Hybond-N+Nylon 膜 (Southern 1975)。按照制造者 (Amersham) 的指令来实施与探针的杂交。利用 (α -³²P)-dCTP、通过随机引物方法来制备放射性探针。探针包括 Os2AP 基因的编码区 (pOs2AP-RNAi 的 SpeI-NcoI 片段, 210 个核苷酸)。

[0159] RT-PCR:为了调查不同生长阶段和不同组织中的转录水平,从 Nipponbare 和转基因植物的幼苗 (10 天)、成年叶 (30 天) 和根 (14 天) 以及开花圆锥中提取总 RNA,并用于 RT-PCR 分析。染色体 8 上的 Os2AP 基因的芳香基因座用 Os2AP-外显子 7.1 引物 (U: 5'-TGCTCCTTTGTCATCACACC-3' (SEQ ID NO: 54) 和 L: 5'-TTTCCACCAAGTCCAGTGA-3' (SEQ ID NO: 55))、通过 PCR 来扩增。也被称作 BADH (Os2AP-chr4; Genbank Accession No. AB001348) 的染色体 4 上的 Os2AP 基因同系物,用 Os2APch4 引物 (U: 5'-TAGCTTCACATCCCCATGTG-3' (SEQ ID NO: 85) 和 L: 5'-GCACCTTCACATCTTGCTGT-3' (SEQ ID NO: 86)、通过 PCR 来扩增。作为对照物,稻米的肌动蛋白基因 (Genbank Accession No. :

X16280) 用肌动蛋白 U :5' -ACATCGCCCTGGACTATGAC-3' (SEQ ID NO :87) 和肌动蛋白 L :5' -TGCTGAGAGATGCCAAGATG-3' (SEQ ID NO :88) 来扩增。

[0160] 实时定量 PCR : 利用 Taqman 化学和 ABI PRISM 7700 序列检测系统 (Applied Biosystems, Foster City, CA, USA)、按照制造者的指令来实施定量 PCR。将相对量作为 Os2AP 的拷贝数与稻米肌动蛋白的拷贝数之比来计算。Taqman 探针被设计成在 Os2AP 基因的芳香与非芳香等位基因之间是有差别的。

[0161] 结果 : 再生四个潮霉素 -GFP 阳性小植物。GFP 的表达在转基因稻米植物中是改变的。T₀RNAi2(10) 即具有最好的 GFP 表达的细胞系, 具有最少的内源性 Os2AP 表达 (图 7B)。2-乙酰基 -1-吡咯内酰胺的分析以及利用 KOH 方法的感觉评估良好相关。T₀RNAi2(10) 含有大量的 2-乙酰基 -1-吡咯并具有最香的叶子。表达少量干扰序列的 T₀RNAi5(20) 和 T₀RNAi3(30), 累积较少的 2-乙酰基 -1-吡咯并具有较低的香叶分数 (图 7B)。T₁RNAi2(10) 的米粒香气的分离分析证实了这个结果。因此, RNA 干涉对 Os2AP 基因表达的抑制, 增强了稻米植物中 2-乙酰基 -1-吡咯的累积。这暗示, 由 8 碱基对缺失产生的芳香稻米中的 NMD 对 Os2AP 基因表达的天然抑制, 是芳香稻米香气的直接原因。

[0162] 例 3 : 稻米的 Os2AP/BADH 基因座的特征 : Os2AP 是以前在染色体 4 上发现的 BADH (甜菜碱醛脱氢酶) 的同系物 (Nagamura 等人, 1997)。我们分别用 Genbank 序列存取号码 AB001348 和 AP004463 作为染色体 4 的 BADH 基因的基因组序列和染色体 8 的 Os2AP 基因的特征。按照基因注释, 来自 RiceGAAS (稻米基因组自动注释系统 (Rice Genome Automated Annotation System), 来自 Sakata 等人, 2002) 的结果, 即 BADH 和 Os2AP 的结构基因包括被 14 个内显子中断的 15 个外显子。ORFs 分别编码 BADH 和 Os2AP 的 505 个和 503 个氨基酸残基的蛋白质。这两种蛋白质在氨基酸水平上具有 77% 的相同性, 并与其它植物 BADH 蛋白质具有 88% -99% 的相同性。而且, 这两种蛋白质含有编码去肽 Val-Thr-Leu-Glu-Leu-Gly-Gly-Lys-Ser-Pro (SEQ ID NO :96) 的核苷酸序列, 该序列高度保存在普通的醛脱氢酶中 (Weretilnyk 等人, 1990)。按照蛋白质序列的 Pfam 分析, BADH 和 Os2AP 属于醛脱氢酶族。该族中的蛋白质在某些植物物种的糖胶甜菜碱的合成中具有推定作用。然而, 在诸如不累积糖胶甜菜碱的稻米、玉米和小麦之类的谷类中, 这些蛋白质的功能是未知的。我们惊讶, 在脯氨酸分解代谢中起作用的脯氨酸 -5-羧化脱氢酶 (P5CDH) (Genbank Accession No. P30038) 也包含在此族内, 这是由于已经报道, 脯氨酸是 2-乙酰基 -1-吡咯的前体。Os2AP 的功能可能是, 催化其它醛而不是甜菜碱醛的脱氢。

[0163] Os2AP 基因的表达和稻米的香气 : 为了鉴定 Os2AP 基因是否与香气的生物合成有关, 分析芳香和非芳香同基因系 (除了芳香基因区域之外, 它们的基因组背景相同) 的表达水平。结果, 我们发现, BADH 在两种同基因系之间没有显示出不同的表达, 同时 Os2AP 相应地在芳香同基因系的所有组织中显示出基因抑制 (图 3D)。我们还研究了同基因系和 Nipponbare 及 K105 (芳香株系) 中的蛋白质表达。Os2AP 基因表达仅在芳香同基因系和 KMLL105 中下调。反之, 在非芳香的同基因系和 Nipponbare 稻米 (非芳香株系) 中没有检测到 Os2AP 基因表达的下调。

[0164] 例 4 : 标记物有助选择

[0165] 为了利用传统的交配繁殖开发新的芳香稻米, 可以使芳香稻米供体与更高产量的受体稻米变种交配。将针对其它特征中的米粒香气分离出交配的子代。此情形使种植者困

感并且使香气更大的稻米的传统繁殖很少成功。芳香基因的发现将为传统种植者制造一个新的范例。可开发用于检测芳香基因的特异性分子标记物,以便将成千种植物以最高的准确度筛选出来。DNA 标记物技术使得种植者能够在早期阶段以高灵敏度检测出 Os2AP 的芳香等位基因。这使得种植者有更多的机会在以后阶段将优选的特征加到芳香植物中。在稻米中鉴定的香气的分子基础可能在其它谷类中发现,这也为其它谷类的 DNA 标记物的开发提供了途径。

[0166] 例 5 :Os2AP 基因的系统发育分析

[0167] 脯氨酸的生物合成在植物王国中是高度保守的途径。本文提到的 2-乙酰基-1-吡咯的生物合成可以是其它产生 2-乙酰基-1-吡咯的植物所利用的共同主题。为了阐明这一点,利用多个序列排列来比较 2AP 的氨基酸序列。所得到的系统发育树表明,来自稻米、小麦和大麦的 2AP 基因可共享共同的祖先(图 9B)。因此,谷类中的 2-乙酰基-1-吡咯的生物合成可共享共同的主题。用其它谷类中的 2AP 基因直向同源所做的实验,在阐明这一点的将来必然得以实施。

[0168] 有趣的是,往回追踪 Os2AP 基因出自于哪里。通过利用“Aromarker”PCR 引物筛选 95 个兰德瑞斯猪变种,我们发现,具有 8 碱基对缺失的等位基因与增大的香气密切相关。我们还测序包括 *Oryza nivara* 和 *O. Rufipogon* 的野生物种中的 8 碱基对缺失的外显子 7(图 9A)。我们发现,包括 Thai Jasmine、Basmat 和 Azucena 的大多数芳香变种都具有缺失,并因此共享可往回追踪到远古时代的共同祖先。我们还鉴定了芳香野生稻米中的芳香等位基因。因此,在人类耕作很久之前发生的单个突变,导致我们今天知道的芳香稻米的产生。

[0169] 参考文献

[0170] 以下的参考文献在此全文并入本文作为参考。

[0171] Abdullah 等人的生物技术 (Biotechnology), 4 :1087, 1986。

[0172] Ahn, S. N., C. N. Bollich 和 S. D. Tanksley. 1992. 水稻芳香基因的 RFLP 标记 (RFLP tagging of a gene for aroma in rice). *Theor. Appl. Genet.* 84 :825-828。

[0173] Ali, S. S., S. J. H. Jafri, M. G. Khan 和 M. A. Butt. 1993. Pakistan 的两种芳香变种的遗传性研究 (Inheritance studies for aroma in two aromatic varieties of Pakistan). *IRRN* 18 :6。

[0174] Altschul, S. F. 等人的 (1990) *J. Mol. Bio.* 215 :403-410。

[0175] Altschul, S. F. 等人的 (1994) *Nat. Genet.* 6(2) :119-29。

[0176] Altschul SF, Gish W. 局部排列统计 (Local alignment statistics). *Methods Enzymol.* 1996 ;266 :460-80。

[0177] Berner, D. R. 和 B. J. Hoff. 1986. 美国长粒稻米香气的遗传性 (Inheritance of scent in America long grain rice). *Crop sci.* 26 :876-878。

[0178] Buttery, R. G., L. C. Ling, B. O. Juliano, J. G.

[0179] Turnbaurg. 1983. 熟米的香气和 2-乙酰基-1-吡咯 (Cooked rice aroma and 2-acetyl-1-pyrroline). *J. Agric. Food Chem.* :823-826。

[0180] Buttery, R. G., L. C. Ling 和 O. B.

[0181] Juliano. 1982. 2-乙酰基-1-吡咯:熟米的一种重要的芳香化合物 (2-acetyl-1-pyrroline:an important aroma component of cooked rice). *Chem*

Ind(London). p. 958。

[0182] Callis 和 Walbot, 基因和发展 (Genes and Develop.), 1 :1183-1200, 1987。

[0183] Capecchi,

[0184] “通过将 DNA 直接微注射到所培养的哺乳动物细胞内而形成的高效转化”(" Highefficiency transformation by direct microinjection of DNA into cultured mammalian cells, ")Cell, 22(2) :479-488, 1980。

[0185] Clapp, “造血细胞内的体细胞基因治疗方法。当前状态和未来展望”(" Somatic genetherapy into hematopoietic cells. Current status and future implications, ")Clin. Perinatol., 20(1) :155-168, 1993。

[0186] Corpet, F. (1988) Nucleic Acids Res. 16 :10881-10890。

[0187] Cristou 等人的 Plant Physiol, 87 :671-674, 1988。

[0188] Curiel, Agarwal, Wagner, Cotten. “运铁蛋白 - 聚赖氨酸介导的基因运送的腺病毒增强”(" Adenovirus enhancement of transferrin-polylysine-mediated gene delivery, ")Proc Natl. Acad Sci. USA, 88(19) :8850-8854, 1991。

[0189] Curiel, Wagner, Cotten, Bimstiel, Agarwal, Li, Loechel, Hu, “由偶联到 DNA- 聚赖氨酸复合物上的腺病毒介导的高效基因转移”(" High-efficiency gene transfer mediated by adenovirus coupled to DNA-polylysine complexes, ")Hum. Gen. Ther., 3(2) :147-154, 1992。

[0190] Dhulappanavar, C. V. 1976. 稻米香气的遗传性 (Inheritance of scent in rice). Euphytica 25 :659-622。

[0191] Eglitis 和 Anderson, “将基因引入哺乳动物细胞内的逆转录病毒载体”(" Retroviral vectors for introduction of genes into mammalian cells, ") Biotechniques 6(7) :608-614, 1988。

[0192] Eglitis, Kantoff, Kohn, Karson, Moen, Lothrop, Blaese, Anderson. “逆转录病毒介导的向造血细胞内的基因转移”(" Retroviral-mediated gene transfer into hemopoietic cells, ")Adv. Exp. Med Biol., 241 :19-27, 1988a。

[0193] Fraley 等人的生物 / 技术 (Bio/Technology), 3 :629-635, 1985。

[0194] Fromm, Taylor, Walbot, “通过电穿孔转移到单子叶植物和双子叶植物细胞内的基因的表达”(" Expression of genes transferred into monocot and dicot plant cells by electroporation, ")Proc. Natl. Acad Sci. USA, 82(17) :5824-5828, 1985。

[0195] Fujimura 等人, 植物组织培养书 (Plant Tissue Culture Letters), 2 :74, 1985。

[0196] Fynan, Webster, Fuller, Haynes, Santoro,

[0197] Robinson, “DNA 疫苗: 由亲代、粘膜和基因枪接种实施的防护性免疫”(" DNA vaccines : protective immunizations by parenteral, mucosal, and gene gun inoculations, ")Proc.

[0198] Natl. Acad Sci. USA 90(24) :11478-11482, 1993。

[0199] Ghose, R. L. M. 和 W. T. Butany. 1952, 对稻米的一些特征的遗传性研究 (Study on the inheritance of some characters in rice) (Oryza sativa L.). Indian J. Genet. Plant Breed. 12 :26-30。

- [0200] Graham 和 van der Eb,“由人逆转录病毒 5 实施的对鼠细胞的转化”(“ Transformation of rat cells by DNA of human adenovirus 5, ”)Virology 54(2) : 536-539, 1973。
- [0201] Hamilton, A. J. 和 Baulcombe, D. C. 1999.
- [0202] 植物中后转录基因沉默的小反义 RNA 的物种 (Aspecies of small antisense RNA in post-transcriptional gene silencing in plants). Science 286 :950-952.
- [0203] Higgins, D. G. 和 Sharp, P. M. (1988) Gene 73(1) :237-44.
- [0204] Higgins, D. G. 和 Sharp, P. M. (1989) Comput Appl Biosci. 5(2) :151-3.
- [0205] Horiguchi, G. 2004. 植物中的 RNA 沉默:功能性分析的短切 (RNA silencing in plants :a shortcut to functional analysis). Differentiation 72 :65-73.
- [0206] Huang X. 等人的 (1992) Comput. Appl. Biosci. 8(2) :155-65.
- [0207] Jin, Q. , D. Walters, G. M. Corderio, R. J. Henry and R. F. Reinke.
- [0208] 2003. 利用稻米基因组序列分析的稻米香味的单个核苷酸多态性 (SNP) 标记物 (A single nucleotide polymorphism (SNP) markers for fragrance in rice by analysis of the rice genome sequence). Mol. Breed. 9 ;245-250.
- [0209] Johnston 和 Tang,“动物细胞的基因枪转染和遗传免疫”(“ Gene gun transfection of animal cells and genetic immunization, ”)Methods Cell. Biol, 43(A) :353-365, 1994.
- [0210] Jorgensen 等人的 Mol. Gen. Genet, 207 :471, 1987.
- [0211] Kadam, B. S. 和 V. K. Patankar. 1938. 稻米芳香的遗传性 (Inheritance of aroma in rice). India J. Genet. Breed. 40 ;327-329.
- [0212] Kaiser 等人的“两亲二级结构:肽激素的设计”(“ Amphiphilic secondary structure :design of peptide hormones. ”)Science, 223(4633) :249-255, 1984.
- [0213] Klee 等人的 Bio/Technology, 3 :637-642, 1985.
- [0214] Klein 等人的 Nature, 327 :70, 1987.
- [0215] Klein 等人的, Proc. Natl. Acad. Sci. USA, 85 :8502-8505, 1988.
- [0216] Kuby, In :免疫学 (Immunology) 第 2 版, W. H. Freeman & Company, NY, 1994.
- [0217] Lanceras, J. C. , Z. L. Huang, O. Naivikul, A. Vanavichit, V. Ruanjaichon and S. Tragoonrung. 2000. Thai 茉莉米 (KDML 105) 的煮吃质量基因图谱 (Mapping of genes for cooking and eating qualities in Thai jasmine rice (KDML 105)). DNA Res 7 :93-101.
- [0218] Lorieux, M. , M. Petrov, N. Huang, E. Guiderdoni, A.
- [0219] Ghesquiere. 1996. 稻米的芳香:定量特征的基因分析 (Aroma in rice : Genetic analysis of a quantitative trait). Theo. Appl. Genet. 93 :1145-1151.
- [0220] Lorz 等人, Mol. Gen. Genet. , 199 :178, 1985.
- [0221] Lu, Xiao, Clapp, Li, Broxmeyer,
- [0222] “高效逆转录介导的从人脐带血向单个分离的未成熟 CD34(3+) 造血干/祖细胞的基因转导”(“ High efficiency retroviral mediated gene transduction into single isolated immature and replatable CD34(3+) hematopoietic stem/progenitor cells from human umbilical cord blood, ”)J. Exp. Med, 178(6) :2089-2096, 1993.

- [0223] Maga, J. A. 1984. 稻米产物的挥发性:综述 (Rice product volatile: A review). J. Agric. Food. Chem. 32 :924-970。
- [0224] Mahatheeranont, S., S. Promdang 和 A. Chaimpiriyakul. 1995. Khao Dawk Mali 105 的挥发性芳香化合物 (Volatile aroma compound of Khao Dawk Mali 105). Kasetsart J. (Nat. sci.) 29 :508-514。
- [0225] Maloy 等人, In: 微生物遗传学 (Microbial Genetics), 第 2 版, Jones and Bartlett Publishers, Boston, Mass., 1994。
- [0226] Marcotte 等人, Nature, 335 :454, 1988。
- [0227] Matzke, M. Matzke, A. J. M. 和 Kooter, J. M. 2001. RNA: 导向基因沉默 (RNA: Guiding gene silencing). Science 293 :1080-1083。
- [0228] McCabe 等人, Biotechnology, 6 :923, 1988。
- [0229] Nagamura Y, Antonio BA, Sasaki T. 采用 RFLPs 的稻米分子基因图谱及其应用 (Rice molecular genetic map using RFLPs and its applications). Plant Mol Biol. 1997 Sep ;35(1-2) :79-87. Review。
- [0230] Nagsuk 等人,
- [0231] 2004, 真菌培养基中的主要芳香米香味化合物 2-乙酰基-1-吡咯的鉴定 (Identification of 2-acetyl-1-pyrroline, the principal aromatic rice flavor compound in fungus cultures). Proceedings: The 2nd International Conference on Medicinal Mushroom and The International Conference on Biodiversity and Bioactive Compounds, 17-19 July, 2003, PEACH, Pattaya, Thailand, p 395-400。
- [0232] Needleman, S. B. 和 Wunsch, C. D. (1970) J. Mol. Biol. 48 :443-453。
- [0233] Nimlek, E. 1999. 四种芳香米的愈伤组织培养的发展和 Khao Dawk Mali 105 的基因转化 (Development of callus culture in four aromatic rices and gene transformation in Khao Dawk Mali 105). M. S. Thesis, Kasetsart University。
- [0234] Omirulleh 等人, Plant Molecular Biology, 21 :415-428, 1993。
- [0235] Paule, C. M. 和 Powers. 1989. 芳香和非芳香米的感觉及化学检查 (Sensory and chemical examination of aromatic and nonaromatic rices). J. Food Sci. 54 :343-346。
- [0236] Pearson WR. 利用 FASTA 程序检索蛋白质和 DNA 序列数据库 (Using the FASTA program to search protein and DNA sequence databases). Methods Mol Biol. 1994 ;24 :307-31。
- [0237] Pearson, W. R., Lipman, D. J., (1988) Proc Natl Acad Sci U. S. A. 85 :2444-8。
- [0238] Petrov, M., M. Danzart, P. Giampaoli, J. Faure, 和 H. Richard. 1995. 稻米的香味分析。芳香与非芳香米之间的区别 (Rice aroma analysis. Discrimination between scented and a non-scented rice). Sci. Aliments 16 :339-352。
- [0239] Potrykus 等人, Mol. Gen. Genet., 199 :183, 1985。
- [0240] Prokop 和 Bajpai, “重组 DNA 技术 I” (“Recombinant DNA Technology I, ”) Ann. N. Y. Acad. Sci., Vol. 646, 1991。
- [0241] Pulak R, Anderson P. “mRNA surveillance by the *Caenorhabditis elegans*

smggenes”。

[0242] Genes Dev. 1993 Oct ;7(10) :1885-97。

[0243] Rogers 等人, Methods Enzymol. ,153 :253-277,1987。

[0244] Romanczyk, J. R. L. J, C. A. McClelland, L. S. Post 和 W. Martin Aitken. 1995.

[0245] 由从可可发酵盒分离的几个仙人山芽孢杆菌株系形成 2-乙酰基-1-吡咯 (Formation of 2-acetyl-1-pyrroline by several Bacillus cereus strains isolated from cacao fermentation boxes)。J. Agric. Food chem. 43(2) :469-475。

[0246] Sakata, K. , Nagamura, Y. , Numa, H. , Antonio, B. A. , Nagasaki, H. , Idonuma, A. , Watanabe, W. , Shimizu, Y. , Horiuchi, I. , Matsumoto, T. , Sasaki, T. & Higo, K. :“稻米 GAAS :稻米基因组序列的自动注释系统和数据库”(“ RiceGAAS :an automated annotation system and database for rice genome sequence ”), Nucleic Acids Res. ,30 :98-102(Jan. 2002)。

[0247] Schieberle, P. 1991. 爆米花中的主要香料 (Primary odorants in popcorn)。J. Agric. Food Chem. 39(6) :1141-1144。

[0248] Segal, In :生物化学计算 (Biochemical Calculations), 第 2 版, John Wiley & Sons, New York, 1976。

[0249] Seitz, L. M. , R. L. Wright, R. D. Waniska 和 L. W. Rooney. 1993.

[0250] 2-乙酰基-1-吡咯从湿地御谷分布到气味中 (Contribution of 2-acetyl-1-pyrroline to odors from wetted ground pearl millet)。J. Agric. Food Chem. 41(6) :955-958。

[0251] Smith, T. F. 和 Waterman, M. S. (1981) J. Mol. Biol. 147 :195-197。

[0252] Sood, B. C. , E. A. Siddiq. 1978. 稻米味道的快速区分技术 (A rapid technique for scent determination in rice)。Indian J. Genet. Plant Breed. 38 :268-271。

[0253] Spielmann 等人, Mol. Gen. Genet. , 205 :34, 1986。

[0254] P. Suprasanna, T. R. Ganapathi, N. K. Ramaswamy, K. K. Surendranathan, P. S. Rao, 稻米的细胞和愈伤组织培养的芳香合成 (Aroma synthesis in cell and callus cultures of rice), Rice Genet. Newsl. 15(1998) 123-125。

[0255] P. Suprarma, G. Bharati, T. R. Ganapathi, V. A. Bapat,

[0256] 稻米的芳香 :脯氨酸补加作用和愈伤组织培养的免疫作用 (Aroma in rice : effects of proline supplementation and immobilization of callus cultures), Rice Genet. Newsl. 19(2002) 9-11。

[0257] Takashi, T. , T. Kuata 和 H.

[0258] Kaio. 1980. 研磨到不同程度的米煮熟之后的挥发性组分 (Volatile components after cooking rice milled to different degrees)。Agric. Biol. Chem. 44(4) :835-840。

[0259] Tanchotikul, U. , 和 T. C. Y.

[0260] Hsieh. 1991. 通过高解析气相色谱 / 质谱 / 选择性离子监测对芳香米中的 2-乙酰基-1-吡咯 (“爆米花香的芳香物质”) 定量的改进方法, (An improved method for quantification of 2-acetyl-1-pyrroline, a “popcorn’-like aroma, in aromatic rice by high-resolution gas chromatography/mass spectrophotometry/selective ion

monitoring)。J. Agric. Food Chem. 39 :944-947。

[0261] Till, Bradley J. Reynolds, SH. Greene, EA. Codomo, CA. Fnnns, LC. Johson, JE. Burtner, C. Odden, AR. Yound, K. Taylor, NE. Henikoff, JG. Comai, L. 和 Henikoff, S. 2003.

[0262] 具有高通量 TILLING 的诱发点突变的大规模发现 (Large-scale discovery of induced point mutations with high-throughput TILLING)。Genome Research 13 : 524-530。

[0263] Toriyama 等人, Theor. Appl. Genet. , 73 :16, 1986。

[0264] Uchimiya 等人, Mol. Gen. Genet. , 204 :204, 1986。

[0265] Vasil 等人,

[0266] “通过可再生的胚胎发育愈伤组织的微粒轰击而获得的抗除莠的能育转基因小麦植物”(“ Herbicide-resistant fertile transgenic wheat plants obtained by microprojectile bombardment of regenerable embryogenic callus, ”)Biotechnology, 10 :667-674, 1992。

[0267] Vasil, Biotechnology, 6 :397, 1988。

[0268] Wagner, Zatloukal, Cotten, Kirlappos, Mechtler, Curiel, Bimstiel, “逆转录病毒与运铁蛋白-聚赖氨酸/DNA 复合物的偶联大大增强了转染基因的受体介导的基因运送和表达”(“ Coupling of adenovirus to transferrin-polylysine/DNA complexes greatly enhances receptor-mediated gene delivery and expression of transfected genes, ”)Proc. Natl. Acad. Sci. USA, 89(13) :6099-6103, 1992。

[0269] S. Wanchana, W. Kamolsukyong, S. Ruengphayak, T. Toojinda, S. Tragoonrung, A.

[0270] Vanavichit, 横跨米粒香气的 4.5

[0271] cM 区域的物理重叠群的快速构建易于富集定位克隆的标记物 (A rapid construction of a physical contig across a 4.5 cM region for rice grain aroma facilitates marker enrichment for positional cloning), Plant Sc. (in press)。

[0272] Weretilnyk, E. 和 Hanson A. D.

[0273] 1990. 植物甜菜碱-醛脱氢酶 (一种适合盐度和干旱的酶) 的分子克隆 (Molecular cloning of a plant betaine-aldehyde dehydrogenase, an enzyme implicated in adaptation to salinity and drought)。Proc. Natl. Acad. Sci. 87 : 2745-2749。

[0274] Wong 和 Neumann, “电场介导的基因转移”(“ Electric field mediated gene transfer, ”)Biochim. Biophys. Res. Commun. 107(2) :584-587, 1982。

[0275] S. Wongpornchai, T. Sriseadka, S. Choonvisase, 面包花 (Vallaris glabra Ktze) 中的稻米芳香化合物 2-乙酰基-1-吡咯的鉴定和定量 (Identification and quantitation of the rice aroma compound, 2-acetyl-1-pyrroline, in bread flowers (Vallaris glabra Ktze)), J. Agric. Food Chem. 51(2003) 457-462。

[0276] Yajima, I., T. Yanai, and M. Nakamura. 1978. 熟米的挥发性味道化合物 (Volatile flavor components of cooked rice)。Agric. Biol. Chem. 42 :1229。

- [0277] Yamada 等人, Plant Cell Rep, 4 :85, 1986。
- [0278] Yanjuan, D., H. Zhang 和 S.
- [0279] Shi. 1992. 良种胞质雄性无菌 (CMS) 芳香 Japonica 系上海 A 的芳香遗传研究 (Genetic studies of aroma in the elite cytoplasmic male sterile (CMS) aromatic Japonica lineshanghai A)。Int. Rice Res News 1. 17 :2。
- [0280] Yoshihashi, T. 2002.
- [0281] 利用稳定的同位素稀释方法对芳香米的 2-乙酰基-1-吡咯的定量分析以及对其在煮饭过程中的形成的模型研究 (Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method and model studies on its formation during cooking)。Food Sci. 67(2) :619-622。
- [0282] Zatloukal, Wagner, Cotten, Phillips, Plank, Steinlein, Curiel. Birnstiel, “运铁蛋白转染:表达真核细胞中的基因构造的高效途径”(Transferrin infection: a highly efficient way to express gene constructs in eukaryotic cells,) Ann. N. Y. Acad. Sci., 660 :136-153, 1992。
- [0283] 基因序列表
- [0284] 序列表
- [0285] <110> 阿披乍特. 瓦纳维奇特
- [0286] 宋旺. 特拉古欧荣
- [0287] 提叻于特. 图金达
- [0288] 萨马特. 瓦乍那
- [0289] 温台. 卡莫利苏克云勇
- [0290] <120> 增强植物和真菌中的 2-乙酰基-1-吡咯的合成的核酸
- [0291] <130> THA1F050153458
- [0292] <160> 96
- [0293] <170> FastSEQ for Windows Version 4.0
- [0294] <210> 1
- [0295] <211> 5857
- [0296] <212> DNA
- [0297] <213> oryza sativa
- [0298] <400> 1
- [0299] atggccacgg cgatcccgcg gggcagctc ttcgtcgccg gcgagtggcg cgcccccgcg 60
- [0300] ctcgggccgc gcctccccgt cgteaaccgc gccaccgagt ccccatcgg taccctcctc 120
- [0301] ttcacctct ccacctctg cttctgctc tgattagcct tttgttggt gttgttggtg 180
- [0302] ctgctgtttt ttgcgtgtcg gtgcgcaagc gagatcccgg cgggcacggc ggaggacgtg 240
- [0303] gacgcggcgg tggcggcggc gcgggagcgc cttaaaaaga acccgggccc cgactgggcg 300
- [0304] cccgcgccgg gcgccgtccg ggccaagtac attcgcgcaa tcgctgacaa agtagggtgg 360
- [0305] tgactacct taccagcctg cccgttttaa cgggagcctt gtgcgtgtgt tccgtacagg 420
- [0306] gggaggagct ccgcgtggct ctccagtagg tttttgagcc ccaaatcgat cgatatgctc 480
- [0307] tagttttaag tttgctgctt aaattctca agggtttagt ttgcaaccaa atccttattt 540

[0308]	tagcttcggt ataagcccc catatgatgt gcgtgcgtcg gcatcggag tgcgtatcct	600
[0309]	ctgttctgga ctaggaattg gccatagggt gatcgacagt tcgagtattc tgcttctggt	660
[0310]	tggaataagt tggaagcatg gctgattgtg tatctggatg ctgtttttgt ggtgattcgt	720
[0311]	ttcaagctct tgttaattga tgggttcaag cggagagggt gcgcaacaac aagtgtatat	780
[0312]	ggctcacggc catgggtgtg cacattgat tgggtgcgca caacaagtgt atattgtgtg	840
[0313]	tgcttcgta gttggcaggt cctagtcact aatcactat tggattggta ctagctactt	900
[0314]	ttgtgccttg acgatgggac tggattacta gccttttggg tgccctttgtg gtattccgtt	960
[0315]	gttatgggcc tgttgatgga tggatccctt taatttctag tgccaaatgc atgctagatt	1020
[0316]	tctcacagtt tttctcttca ggttatattt ctcgtatttc ctttcctaa aggattgctt	1080
[0317]	tttcatgtat tttctggcat atataggta ttattattat tattctccag aacaagatta	1140
[0318]	cccatattat ggatcactag tgtacacttt tttggatgaa aaacctactt actgaaagta	1200
[0319]	aaacagtgac cagtgcacac tttacttgaa ctgtcaaacc atcaattttc tagcaaagca	1260
[0320]	ggggatgcta gccttcagct ctaaagaca gtaaactact atacttttgt ccgtaggttt	1320
[0321]	ggaaatagc taatttctat cataaaaatt ttcattggcat atgcgagcat tttatgatca	1380
[0322]	ccttttccct ttttctcag ataatcgaga ggaaatctga gctggctaga ctagagacgc	1440
[0323]	ttgattgtgg gaagcctctt gatgaagcag catgggacat ggtatgtggc cagttatcca	1500
[0324]	ctgtatgaat atgtagttgc ctacacagca atctttcctg aacatgaatc ctgatgtatg	1560
[0325]	atattccatt tgcaggacg atgttgcctg atgctttgag tactttgcag atcttgcaga	1620
[0326]	atccttggac aaaaggcaaa atgcacctgt ctctcttcca atggaaaact ttaaatgcta	1680
[0327]	tcttcggaaa gagcctatcg gtgtagttgg gttgatcaca cttggattt tcacattttt	1740
[0328]	ctctcatcct gcgcttatat ttatttatga cccaagcatg gtactaaata gtactagtaa	1800
[0329]	catgcatata ctgaatgagt ttacaacttt acatgatttt tttgactat gaaagttaa	1860
[0330]	gacatttgag attttattcc tcttctcttg tgcaaacata ttattgtctc acaaatgta	1920
[0331]	cctagcagct actctctccg tttcatatta taagtcgttt gacttttttc ctagtcaaaa	1980
[0332]	tgtgttaagt ttgaccaagt ttatagaaaa atttagcaac atctaaaata tcaaatgcat	2040
[0333]	gttttagtgt ttttccaggc tctcatgtaa gcaattttga tgtgccctct ctttctctt	2100
[0334]	taatataatg atacacagct cttgtgtatt caaaggaata tatatatata tatatatata	2160
[0335]	taatgataca cacctctcct ccgtgttaat gcagctcatt tgttctgtcc cggttcaaat	2220
[0336]	atctattttt ctcatatggt gtcagcatga ttcacttaat ttagtatata gaagatgcca	2280
[0337]	ttatttatgt ctggaatctt actgcagaag ggaaaacaat tgataacgga attgattgca	2340
[0338]	ttctaatttg ttgtttcttt gttatgttct tatcgacaat tacaatttg attctgagaa	2400
[0339]	tcatgttcgg gatgtgtatt tctactgcag gaactatcct ctctgatgg caacatggaa	2460
[0340]	ggtagctcct gccctggctg ctggctgtac agctgtacta aaacctctg aattggcttc	2520
[0341]	cgtgtaagtt taacatgtta acttgtaaat gtcatacca tgctagtgtc aatgacattt	2580
[0342]	gattttaaaa tgttgtggca tgtccatgct gcaagcaatg taatttgaaa tctctctcta	2640
[0343]	tcattaatta ccaggacttg tttggagctt gctgatgtgt gtaaagaggt tggctctcct	2700
[0344]	tcagggtgtc taaacatagt gactggatta ggttctgaag ccggtgctcc tttgtcatea	2760
[0345]	caccctgggt tagacaaggt acagctattc ctctgtaat catgtatacc ccatcaatgg	2820
[0346]	aatgatatt cctctcaata catggtttat gttttctgtt aggttgcatt tactgggagt	2880

[0347]	tatgaaactg gtatatat	ttt cagctgctcc tatgg	ttaag gtttg	tttcc aaat	tttctgt	2940
[0348]	ggatattttt tgttctct	ttt ctactaactc tet	attatca attct	caatg ttgt	cctttt	3000
[0349]	cttttaactc ctttact	ttt tagaattgtg at	caagacac tttg	agcacc	attctagtag	3060
[0350]	ccagttctat cctgttt	cctt acccttttat	ggttcgtctt tt	ccttgacag	cctgtttcac	3120
[0351]	tggaacttgg tggaaaa	agt cctatagtg	gtttgatga t	gttgatgtt	gaaaaaggta	3180
[0352]	catgccactt gctatg	atta actaattctg	aagtgcggga	ctttgtaa	ag cacttaactg	3240
[0353]	agctggatgc tagac	ccccca aaagcc	cttt ttgg	tgtctt	ggccttgttg	3300
[0354]	ggcccagac gagc	aggtatg caagaaa	att aactact	tttt gccact	gatt agtatt	3360
[0355]	agaagttaca cct	caaggat tagca	actact ttct	ttaaaat	gtgctattga	3420
[0356]	gtcctgtatt at	tttgagca gatct	gttac tgg	ttgatcg	gcttgcatga	3480
[0357]	gaggattata at	gccatgcc aact	gagtaa agaaa	agagt tgtaa	aatat gttat	3540
[0358]	atgaatata at	gtgatttc at	ttttccct ttt	cttttcg	tggcaaggaa	3600
[0359]	aaggactgat gt	gaaaagca ca	agtactat t	cttagttct	ggaaaactgt	3660
[0360]	ttcctaacta	caattcacct	tgattag	tca gtaact	tgatt attgg	3720
[0361]	atgaattctg	tttatatttc	actaatt	ttg aatct	tttaac tacat	3780
[0362]	taacgttttg	tctggttatg	gactctgtt	gtattc	actc aatt	3840
[0363]	tttcattgtt	ggctccttct	ctgtacagc	tgttg	agtgg actc	3900
[0364]	gaccaatggc	cagatttgca	gtgcaac	atc gcgt	ccttatt	3960
[0365]	tatccgtcaa	tcataateta	ttgttg	tact tgatt	ttttt	4020
[0366]	agattattat	atgatgccat	tactatt	gca cagag	cgaat aaa	4080
[0367]	catgtat	ttt gagtaata	ta tgca	ttgcct att	tttaac	4140
[0368]	ttttgtgtcc	atgtctctat	gatgtt	tatt act	tttattat	4200
[0369]	gaactctatg	ttgatcttga	actaaa	ttg aaatta	aattg gctt	4260
[0370]	tagctttcag	cttcttgctc	ctgac	catga aag	ttttgca	4320
[0371]	ttcaagaaag	gatggttgca	tgggcc	aaaa atatta	aggt gtcag	4380
[0372]	gttgcaggct	tgggcccg	tt gttag	taag gacag	gtacc acat	4440
[0373]	ttcaaaaaag	aaatgccact	gatcaat	ggt aggt	ccttcc	4500
[0374]	tgcactgttt	tgtcaat	ttt gtg	taata gtt	ctgaa	4560
[0375]	tgctagt	ttgc tagtat	gtgg tac	aggtct	tcctact	4620
[0376]	cattgat	gaa aagcc	acct ttcc	gtaggt	ttatct	4680
[0377]	at	tttaattt	tttgtttg	ga ccc	atggca	4740
[0378]	tgtatat	tatt ttcaa	agcac caga	atggtt	agatg	4800
[0379]	attgc	cttgt gtcag	tatga	aaag ca	attt	4860
[0380]	gctacc	atc tgact	ggtg	gg tag	acc aag	4920
[0381]	ataggt	acc acat	atcatt	atga	agtaga	4980
[0382]	gagaa	aggtt tct	atattg	ga acc	caatc	5040
[0383]	tggag	ggaag aag	ttttt	gg tcc	agtgtc	5100
[0384]	gccatt	gaat tgg	ccaacga	tactc	agtga	5160
[0385]	ttcaat	cttg cag	catatgt	atata	ctctg	5220

[0386]	ttttgatagt tatggtctgg ctggtgctgt gctttccggt gaccgcgagc gatgccagag	5280
[0387]	attaactgag gtatatccaa gtgaaggggg ttggcattgt ttgattcata tgacatggtt	5340
[0388]	gcatcaagct gatattcaag aatctcattt attacttgca ttctatgcat ctccagttct	5400
[0389]	tccctggact ccggtcaatg ttaatatagt ttgtttgcta gtagtatgct actccaatta	5460
[0390]	agttgctctt cacttcaca tcatctgac catgacttta tatttgacce cttttttttg	5520
[0391]	caaaagagag gaaataactt aacgaaaatt tcctactgca ggagatcgat gccggaatta	5580
[0392]	tctgggtgaa ctgctcgcaa ccctgcttct gccaaagctcc atggggcggg aacaagcgca	5640
[0393]	gcggctttgg acgcgagctc ggagaagggt gggtagcaca caacaatctc actttaaaac	5700
[0394]	accatttcca tcgtctgatg atctcgacct gacatcatgc ctttggatt ttcatctact	5760
[0395]	tttcaggggc attgacaact acctaagcgt caagcaagtg acggagtacg cctccgatga	5820
[0396]	gccgtgggga tggtaacaaat ccccttccaa gctgtaa	5857
[0397]	<210>2	
[0398]	<211>1504	
[0399]	<212>DNA	
[0400]	<213>oryza sativa	
[0401]	<400>2	
[0402]	atggccacgg cgatcccgca gggcagctc ttcgtcgccg gcgagtggcg cgccccgcg	60
[0403]	ctcggccgcc gcctcccctgt cgtaacccc gccaccgagt ccccatcgg cgagatccc	120
[0404]	gcgggcacgg cggaggacgt ggacgcggcg gtggcggcgg cgcgggagc gcttaaaaag	180
[0405]	aaccgggcc gcgactgggc gccgcgccc ggccgctcc gggccaagta cattcgcgca	240
[0406]	atcgctgaca aaataatcga gaggaaatct gagctggcta gactagagac gcttgattgt	300
[0407]	gggaagcctc ttgatgaagc agcatgggac atggacgatg ttgctggatg ctttgagtac	360
[0408]	tttgagatc ttgcagaatc cttggacaaa aggcaaaatg cacctgtctc tcttccaatg	420
[0409]	gaaaacttta aatgctatct tcggaaagag cctatcggtg tagttgggtt gatcacacct	480
[0410]	tggaactatc ctctcctgat ggcaacatgg aaggtagctc ctgccctggc tgctggctgt	540
[0411]	acagctgtac taaaaccatc tgaattggct tccgtgactt gtttggagct tgctgatgtg	600
[0412]	tgtaaagagg ttggtcttcc tcaggtgtg ctaaacaatag tgactggatt aggttctgaa	660
[0413]	gccggtgctc ctttgtcatc acacctggg ttagacaagg ttgcatttac tgggagttat	720
[0414]	gaaactggta tatatttcag ctgctcctat ggttaagcct gtttactgg aacttggtag	780
[0415]	aaaaagtctc atagtgggtg ttgatgatg tgatgttgaa aaagctgtg agtggactct	840
[0416]	ctttggttgc ttttggacca atggccagat ttgcagtgca acatcgcgctc ttattcttca	900
[0417]	taaaaaaatc gctaaagaat tcaagaaag gatggttgca tgggcaaaa atattaaggt	960
[0418]	gtcagatcca cttgaagagg gttgcaggtc tgggccggtt gttagtgaa gacagtatga	1020
[0419]	gaagattaag caatttgtat ctaccgcaa aagccaaggt gctaccatc tgactggtag	1080
[0420]	ggttagacce aagcatctgg agaaaggtt ctatattgaa ccacaatca ttactgatgt	1140
[0421]	cgatacatca atgcaaattt ggagggaaga agtttttggc ccagtgtctc gtgtgaaaga	1200
[0422]	athtagcact gaagaagaag ccattgaatt ggccaacgat actcattatg gtctggctgg	1260
[0423]	tgctgtgctt tccggtgacc gcgagcgtg ccagagatta actgaggaga tcgatccgg	1320
[0424]	aattatctgg gtgaactgct cgcaacctg cttctgccaa gctccatggg gcgggaacaa	1380

[0425]	gcgcagcggc ttggacgcg agctcggaga agggggcatt gacaactacc taagcgtcaa	1440
[0426]	gcaagtgacg gactacgcct ccgatgagcc gtgggatgg tacaatccc cttccaagct	1500
[0427]	gtaa	1504
[0428]	<210>3	
[0429]	<211>251	
[0430]	<212>PRT	
[0431]	<213>oryza sativa	
[0432]	<400>3	
[0433]	Met Ala Thr Ala Ile Pro Gln Arg Gln Leu Phe Val Ala Gly Glu Trp	
[0434]	1 5 10 15	
[0435]	Arg Ala Pro Ala Leu Gly Arg Arg Leu Pro Val Val Asn Pro Ala Thr	
[0436]	20 25 30	
[0437]	Glu Ser Pro Ile Gly Glu Ile Pro Ala Gly Thr Ala Glu Asp Val Asp	
[0438]	35 40 45	
[0439]	Ala Ala Val Ala Ala Ala Arg Glu Ala Leu Lys Lys Asn Pro Gly Arg	
[0440]	50 55 60	
[0441]	Asp Trp Ala Pro Ala Pro Gly Ala Val Arg Ala Lys Tyr Ile Arg Ala	
[0442]	65 70 75 80	
[0443]	Ile Ala Asp Lys Ile Ile Glu Arg Lys Ser Glu Leu Ala Arg Leu Glu	
[0444]	85 90 95	
[0445]	Thr Leu Asp Cys Gly Lys Pro Leu Asp Glu Ala Ala Trp Asp Met Asp	
[0446]	100 105 110	
[0447]	Asp Val Ala Gly Cys Phe Glu Tyr Phe Ala Asp Leu Ala Glu Ser Leu	
[0448]	115 120 125	
[0449]	Asp Lys Arg Gln Asn Ala Pro Val Ser Leu Pro Met Glu Asn Phe Lys	
[0450]	130 135 140	
[0451]	Cys Tyr Leu Arg Lys Glu Pro Ile Gly Val Val Gly Leu Ile Thr Pro	
[0452]	145 150 155 160	
[0453]	Trp Asn Tyr Pro Leu Leu Met Ala Thr Trp Lys Val Ala Pro Ala Leu	
[0454]	165 170 175	
[0455]	Ala Ala Gly Cys Thr Ala Val Leu Lys Pro Ser Glu Leu Ala Ser Val	
[0456]	180 185 190	
[0457]	Thr Cys Leu Glu Leu Ala Asp Val Cys Lys Glu Val Gly Leu Pro Ser	
[0458]	195 200 205	
[0459]	Gly Val Leu Asn Ile Val Thr Gly Leu Gly Ser Glu Ala Gly Ala Pro	
[0460]	210 215 220	
[0461]	Leu Ser Ser His Pro Gly Val Asp Lys Val Ala Phe Thr Gly Ser Tyr	
[0462]	225 230 235 240	
[0463]	Glu Thr Gly Ile Tyr Phe Ser Cys Ser Tyr Gly	

[0464]	245	250	
[0465]	<210>4		
[0466]	<211>5859		
[0467]	<212>DNA		
[0468]	<213>oryza sativa		
[0469]	<400>4		
[0470]	atggccacgg cgatccccgca gcggcagctc ttcgtcgccg gcgagtggcg cgcccccgcg		60
[0471]	ctcggcgcgc gcctccccgt cgtcaacccc gccaccgagt cccccatcgg taccctcctc		120
[0472]	ttcacctctt ccacctctg cttctgcctc tgattagcct ttttgttgtt gttgttgttg		180
[0473]	ctgctgtttt ttgcgtgtcg gtgcgcaggc gagatcccgg cgggcacggc ggaggacgtg		240
[0474]	gacgcggcgg tggcggcggc gcgggaggcg ctgaagagga accggggccg cgactgggcg		300
[0475]	cgcgcgccgg gcgccgtccg ggccaagtac ctccgcgcaa tcgcggccaa ggtagggtgg		360
[0476]	tgactacccc cacccccccc cccccccaa cgcgaccgcg gtgcgtgtgt tccgtacagg		420
[0477]	gggaggagct ccgcgtggct ctccagtagg tttttgagcc ccaaatcgat cgatatgctc		480
[0478]	tagttttaag tttgctgctt aaattcctca agggtttagt ttgcaaccaa atccttattt		540
[0479]	ta9cttcggt ataagcccc catatgatgt gcgtgcgtcg gcacggaag tgcgtatcct		600
[0480]	ctgttctgga ctaggaattg gccatagggt gatcgacagt tcgagtattc tgcttctgtt		660
[0481]	tggaataagt tggaagcatg gctgatttg tatctggatg ctgtttttgt ggtgattcgt		720
[0482]	ttcaagctct tgtaattga tgggttcaag cggagagggt gcgcaacaac aagtgtatat		780
[0483]	ggctcacggc catgggtgtg cacattgat tgggtgcgcaa caacaagtgt atattgtttg		840
[0484]	tgtgcttctg tagttggcag gtccatgca ctaaactact attggattgg tactagttac		900
[0485]	ttttgtgctt tgacgatggg actggattac tagccttttg gttgcctttg tggattccg		960
[0486]	ttgttatggg cctgttgatg gatggatccc ttttaatttct agtgccaaat gcattgctaga		1020
[0487]	tttctcacag tttttctctt caggttatat ttctcgtatt tcttttccct aaaggattgc		1080
[0488]	tttttcatgt attttctggc atatataggt tattattatt attattctcc agaacaagat		1140
[0489]	tacctatatt atggatcact agtgtacact tttttggatg aaaaacctac ttactgaaag		1200
[0490]	taaaacagtg accagtgcac actttacttg aactgtcaaa ccatcaattt tctagcaaag		1260
[0491]	caggggatgc tagccttcca gtctaaatga cagtaaacta ctatactttt gtccgtaggt		1320
[0492]	ttggaatat gctaatttct atcataaaaa ttttcatggc atatgcgagc attttatgat		1380
[0493]	caccttttcc ctttttcttc agataatcga gaggaaatct gagctggcta gactagagac		1440
[0494]	gcttgattgt gggaagcctc ttgatgaagc agcatgggac atgggatgtg gccagttatc		1500
[0495]	cactgtatga atatgtagtt gcctacacag caatctttcc tgaacatgaa tcttgatgta		1560
[0496]	tgatatcca tttgtcagga cgatgttctt ggatgctttg agtactttgc agatcttgea		1620
[0497]	gaatccttgg acaaaaggca aatgcacct gtctctcttc caatggaaaa ctttaaatgc		1680
[0498]	tatcttcgga aagagcctat cgggttagtt gggttgatca caccttggta tttcacattt		1740
[0499]	ttctctcacc ctgcgcttat atttatttat gaccaagca tggtaactaaa tagtactagt		1800
[0500]	aacatgcata tactgaatga gtttacaact ttacatgatt tttttgaact atgaaagtgtg		1860
[0501]	aagacatttg agattttatt cctctctctt tgtgcaaaaca tattattgtc tcacaaattg		1920
[0502]	tacctagcag ctactctctc cgtttcatal tataagtcgt ttgacttttt tcttagtcaa		1980

[0503]	aatgtgtaa gtttgaccaa gtttatagaa aaatttagca acatctaaaa tatcaaagtc	2040
[0504]	atgttttagt gttttttcag gctctcatgt aagcaatfff gatgtgccct ctcctttctt	2100
[0505]	cttaataata tgatacacag ctcttggtga ttcaaaggaa aatataata tatataatga	2160
[0506]	tacacacctc tcctccgtgt taatgcagct catttgttct gtcccgggtc aaatatctat	2220
[0507]	ttttctcata tgttgtcagc atgattcact taatttagta tatagaagat gccattatft	2280
[0508]	atgtctggaa tcttactgca gaagggaaaa caattgataa cgggaattgat tgcattctaa	2340
[0509]	tttgttgfff ctttgttatg ttcttatcga caattacaaa ttgattctg agaatcatgt	2400
[0510]	tcgggatgtg tatttctact gcaggaacta tcctctcctg atggcaacat ggaaggtagc	2460
[0511]	tcctgccctg gctgctggct gtacagctgt actaaaacca tctgaattgg cttccgtgta	2520
[0512]	agtttaacat gttaacttgt taatgcata cccatgctag ttgcaatgac atfttgatftt	2580
[0513]	aaaatgftt ggcattgtcca tgctgcaagc aatgtaatft gaaatctctc tctatcatta	2640
[0514]	attaccagga cttgfttggg gcttgctgat gtgtgtaaag aggttggctc tccttcaggt	2700
[0515]	gtgctaaaca tagtgactgg attaggttct gaagccggtg ctcctttgtc atcacaccct	2760
[0516]	ggtgtagaca aggtacagct attcctcctg taatcatgta taccatca atggaaatga	2820
[0517]	tattcctctc aatacatggt ttatgfttfc tgttaggttg catttactgg gagttatgaa	2880
[0518]	actggtaaaa agattatggc ttcagctgct cctatggfta aggttggftt ccaaatttct	2940
[0519]	gtggatattt ttgfttctct ttctactaac tctctattat caattctcaa tgttgctcft	3000
[0520]	ttctfttaac tcctfttactt tftagaattg tgatcaagac actftgagca tcattctagt	3060
[0521]	agccagfttct atcctgfttfc ttactfttft atggftcgtc tfttcttgac agcctgfttfc	3120
[0522]	actggaactt ggtggaaaaa gtccatagat ggtgfttggat gatgfttggat ttgaaaaagg	3180
[0523]	tacatgccac ttgctatgat taactaatc tgaagtgcgg gactfttggaa agcacttaac	3240
[0524]	tgagctggat gctagacccc caaaagccct tfttggftgtc ttgggcttgt tgcagaaata	3300
[0525]	ctggftccag acgagcagga tgcaagaaaa ttaactactt ttgccactga ttagtatttc	3360
[0526]	ttagaagtta cacctcaagg attagcaata cfttctftaa atgtgctatt gattaaaaag	3420
[0527]	atgtcctgta ftattfttgag cagatcttgt actggttggat cggcttggat gaaaatattg	3480
[0528]	ttgaggatta taatgccatg ccaactgagt aaagaaaaa gfttgaaaat atgttatgca	3540
[0529]	acatgaatat atatgtgatt tcattfttcc tfttctfttft cgtggcaagg aaggcagfta	3600
[0530]	ggaaggactg atgtgaaaag cacaagtact attcttagtt ctggaaaact gtgttctfta	3660
[0531]	tfttctaac tacaattcac cfttgattagt cagtaacttg atattggcaa tfttagctga	3720
[0532]	ttatgaattc tgtttatatt tcactaatft tgaatctfta attacattft atggttgaaa	3780
[0533]	tftaacgftt tgtctggfta tggactctgt ttgtattcac tcaatttggg tcttccatta	3840
[0534]	gatttcaattg ttggftccttc tcttgtaca gctgfttggat ggactctctt tggfttgcftt	3900
[0535]	tggaccaatg gccagatttg cagtgcaca tcgctctta tcttcatgt aagcattgaa	3960
[0536]	tatatccgtc aatcataatc tattgttggta cfttgatfttft tftctgatca actcctgagt	4020
[0537]	tcagattatt atatgatgcc attactattg cacagagcga ataaaattgt atfttatgcac	4080
[0538]	agcatgtatt ttgagtaata tatgcattgc ctattatfta atatatagat tgtagcactt	4140
[0539]	aattfttgtg ccatgtctct atgatgftta ftactfttatt atftccggca tgaagcaact	4200
[0540]	ttgaacteta tgttgatctt gaactaaaat tgaattaat tggcttattg ctattaatga	4260
[0541]	tatagcttfc agcttcttgc tcctgacat gaaagfttftg cagaaaaaaa tcgctaaaga	4320

[0542]	atccaagaa aggatggttg catgggcca aatattaag gtgtcagatc cacttgaaga	4380
[0543]	gggttgccagg cttgggcccg ttgttagtga aggacaggta ccacatgtaa actttttcta	4440
[0544]	aattcaaaaa agaaatgcca ctgatcaatg gtaggtcctt ccaagcctta ttgctggatt	4500
[0545]	gttgactgt tttgtcaatt ttgtgtaata tagttctgaa tgaattagtc ggtgtatgct	4560
[0546]	cttgctagt gctagtatgt ggtacagggt cttcctactt tgagcaaatt cgtgttaaaa	4620
[0547]	tgcattgatg aaaaggccac cttccgtag gtttatcttg tcataattta aacccaata	4680
[0548]	aaattttaat tttttgtttt gacccatgg cactttaatg aatcactta gccatgagct	4740
[0549]	tttgtatata ttttcaaagc accagaatgt ttagatggtt tttggaaat cttacacatc	4800
[0550]	ctattgcctt gtgtcagtat gagaagatta agcaatttgt atctaccgcc aaaagccaag	4860
[0551]	gtgctacat tctgactggt ggggttagac ccaaggtaat aatctactac acggttgtat	4920
[0552]	atataggtac ccacatatca ttatgaagta gaaataatct tgtatgtttt tgctcagcatc	4980
[0553]	tggagaaagg tttctatatt gaaccacaa tcattactga tgcgataca tcaatgcaa	5040
[0554]	tttgaggga agaagttttt ggtccagtgc tctgtgtgaa agaatttagc actgaagaag	5100
[0555]	aagccattga attggccaac gatactcagt gagttttttt ttaatacag ttcattgtcc	5160
[0556]	tgttcaatct tgcagcatat gtatatactc tgtggcatat gaacttattc tgctactact	5220
[0557]	acttttgata gttatggtct ggctggtgct gtgctttccg gtgaccgca gcgatgccag	5280
[0558]	agattaactg aggtatatcc aagtgaagg ggttggcatt gtttgattca tatgacatgg	5340
[0559]	ttgcatcaag ctgatattca agaatctcat ttattacttg cattctatgc atctccagtt	5400
[0560]	ctccctgga ctccggctca tgtaataata gtttgtttgc tagtagtatg ctactccaat	5460
[0561]	taagttgctc ttcactteca catcatctga tccatgactt tataattgac ccctttttt	5520
[0562]	tgcaaaagaa agggaaatac ttaacgaaaa tttcctactg caggagatcg atgccggaat	5580
[0563]	tatctgggtg aactgctcgc aacctgctt ctgccaagct ccatggggcg ggaacaagcg	5640
[0564]	cagcggcttt ggacgcgagc tcggagaagg gtgggtagca cacaacaatc tcaacttaaa	5700
[0565]	acaccattc gatcgtctga tgatctgac ctgacatcat gcctttgta ttttcattca	5760
[0566]	cttttcaggg gcattgacaa ctacctaagc gtcaagcaag tgacggagta cgcctccgat	5820
[0567]	gagccgtggg gatggtacaa atcccttcc aagctgtaa	5859
[0568]	<210>5	
[0569]	<211>1512	
[0570]	<212>DNA	
[0571]	<213>oryza sativa	
[0572]	<400>5	
[0573]	atggccacgg cgatcccgca gcggcagctc ttcgtcgccg gcgagtggcg cgccccgcg	60
[0574]	ctcggccgcc gcctccccgt cgtaacccc gccaccgagt ccccatcgg cgagatcccg	120
[0575]	gcgggcacgg cggaggacgt ggacgcggcg gtggcggcgg cgcgggagc gctgaagagg	180
[0576]	aaccggggcc gcgactgggc gcgcgcgccg ggcccgctcc gggccaagta cctccgcga	240
[0577]	atcgcggcca agataatcga gaggaaatct gagctggcta gactagagac gcttgattgt	300
[0578]	gggaagcctc ttgatgaagc agcatgggac atggacgatg ttgctggatg ctttgagtac	360
[0579]	tttgagatc ttgcagaatc cttggacaaa aggcaaaatg cacctgtctc tttccaatg	420
[0580]	gaaaacttta aatgctatct tcggaaagag cctatcggtg tagttgggtt gatcacacct	480

[0581]	tggaactatc ctctcctgat ggcaacatgg aaggtagctc ctgccctggc tgctggctgt	540
[0582]	acagctgtac taaaaccatc tgaattggct tccgtgactt gtttgagct tgctgatgtg	600
[0583]	tgtaaagagg ttggtcttcc ttcagggtg ctaaacaatag tgactggatt aggttctgaa	660
[0584]	gccggtgctc ctttgtcacc acaccctggg gtagacaagg ttgcatttac tgggagttat	720
[0585]	gaaactggta aaaagattat ggcttcagct gctcctatgg ttaagcctgt ttcactggaa	780
[0586]	cttgggtggaa aaagtcctat agtgggtgtt gatgatgttg atgttgaaaa agctgttgag	840
[0587]	tggactctct ttggttgctt ttggaccaat ggccagattt gcagtgaac atcgcgtctt	900
[0588]	attcttcata aaaaaatcgc taaagaattt caagaaagga tggttgcatg ggccaaaaat	960
[0589]	attaaggtgt cagatccact tgaagagggt tgcaggcttg ggcccgttgt tagtgaagga	1020
[0590]	cagtatgaga agattaagca atttgtatct accgcaaaa gccaaagtgc taccattctg	1080
[0591]	actggtgggg ttagacccaa gcactctggag aaaggtttct atattgaacc cacaatcatt	1140
[0592]	actgatgtcg atacatcaat gcaaatttgg agggaagaag ttttggctcc agtgctctgt	1200
[0593]	gtgaaagaat ttagcactga agaagaagcc attgaattgg ccaacgatac tcattatggt	1260
[0594]	ctggctgggt ctgtgctttc cggtgaccgc gagcgatgcc agagattaac tgaggagatc	1320
[0595]	gatgccggaa ttatctgggt gaactgctcg caaccctgct tctgccaage tccatggggc	1380
[0596]	gggaacaagc gcagcggctt tggacgcgag ctcggagaag ggggcattga caactaccta	1440
[0597]	agcgtcaagc aagtgacgga gtacgcctcc gatgagccgt ggggatggta caaatccct	1500
[0598]	tccaagctgt aa	1512
[0599]	<210>6	
[0600]	<211>503	
[0601]	<212>PRT	
[0602]	<213>Oryza sativa	
[0603]	<400>6	
[0604]	Met Ala Thr Ala Ile Pro Gln Arg Gln Leu Phe Val Ala Gly Glu Trp	
[0605]	1 5 10 15	
[0606]	Arg Ala Pro Ala Leu Gly Arg Arg Leu Pro Val Val Asn Pro Ala Thr	
[0607]	20 25 30	
[0608]	Glu Ser Pro Ile Gly Glu Ile Pro Ala Gly Thr Ala Glu Asp Val Asp	
[0609]	35 40 45	
[0610]	Ala Ala Val Ala Ala Ala Arg Glu Ala Leu Lys Arg Asn Arg Gly Arg	
[0611]	50 55 60	
[0612]	Asp Trp Ala Arg Ala Pro Gly Ala Val Arg Ala Lys Tyr Leu Arg Ala	
[0613]	65 70 75 80	
[0614]	Ile Ala Ala Lys Ile Ile Glu Arg Lys Ser Glu Leu Ala Arg Leu Glu	
[0615]	85 90 95	
[0616]	Thr Leu Asp Cys Gly Lys Pro Leu Asp Glu Ala Ala Trp Asp Met Asp	
[0617]	100 105 110	
[0618]	Asp Val Ala Gly Cys Phe Glu Tyr Phe Ala Asp Leu Ala Glu Ser Leu	
[0619]	115 120 125	

[0620]	Asp Lys Arg Gln Asn Ala Pro Val Ser Leu Pro Met Glu Asn Phe Lys
[0621]	130 135 140
[0622]	Cys Tyr Leu Arg Lys Glu Pro Ile Gly Val Val Gly Leu Ile Thr Pro
[0623]	145 150 155 160
[0624]	Trp Asn Tyr Pro Leu Leu Met Ala Thr Trp Lys Val Ala Pro Ala Leu
[0625]	165 170 175
[0626]	Ala Ala Gly Cys Thr Ala Val Leu Lys Pro Ser Glu Leu Ala Ser Val
[0627]	180 185 190
[0628]	Thr Cys Leu Glu Leu Ala Asp Val Cys Lys Glu Val Gly Leu Pro Ser
[0629]	195 200 205
[0630]	Gly Val Leu Asn Ile Val Thr Gly Leu Gly Ser Glu Ala Gly Ala Pro
[0631]	210 215 220
[0632]	Leu Ser Ser His Pro Gly Val Asp Lys Val Ala Phe Thr Gly Ser Tyr
[0633]	225 230 235 240
[0634]	Glu Thr Gly Lys Lys Ile Met Ala Ser Ala Ala Pro Met Val Lys Pro
[0635]	245 250 255
[0636]	Val Ser Leu Glu Leu Gly Gly Lys Ser Pro Ile Val Val Phe Asp Asp
[0637]	260 265 270
[0638]	Val Asp Val Glu Lys Ala Val Glu Trp Thr Leu Phe Gly Cys Phe Trp
[0639]	275 280 285
[0640]	Thr Asn Gly Gln Ile Cys Ser Ala Thr Ser Arg Leu Ile Leu His Lys
[0641]	290 295 300
[0642]	Lys Ile Ala Lys Glu Phe Gln Glu Arg Met Val Ala Trp Ala Lys Asn
[0643]	305 310 315 320
[0644]	Ile Lys Val Ser Asp Pro Leu Glu Glu Gly Cys Arg Leu Gly Pro Val
[0645]	325 330 335
[0646]	Val Ser Glu Gly Gln Tyr Glu Lys Ile Lys Gln Phe Val Ser Thr Ala
[0647]	340 345 350
[0648]	Lys Ser Gln Gly Ala Thr Ile Leu Thr Gly Gly Val Arg Pro Lys His
[0649]	355 360 365
[0650]	Leu Glu Lys Gly Phe Tyr Ile Glu Pro Thr Ile Ile Thr Asp Val Asp
[0651]	370 375 380
[0652]	Thr Ser Met Gln Ile Trp Arg Glu Glu Val Phe Gly Pro Val Leu Cys
[0653]	385 390 395 400
[0654]	Val Lys Glu Phe Ser Thr Glu Glu Glu Ala Ile Glu Leu Ala Asn Asp
[0655]	405 410 415
[0656]	Thr His Tyr Gly Leu Ala Gly Ala Val Leu Ser Gly Asp Arg Glu Arg
[0657]	420 425 430
[0658]	Cys Gln Arg Leu Thr Glu Glu Ile Asp Ala Gly Ile Ile Trp Val Asn

[0659]	435	440	445
[0660]	Cys Ser Gln Pro Cys Phe Cys Gln Ala Pro Trp Gly Gly Asn Lys Arg		
[0661]	450	455	460
[0662]	Ser Gly Phe Gly Arg Glu Leu Gly Glu Gly Gly Ile Asp Asn Tyr Leu		
[0663]	465	470	475 480
[0664]	Ser Val Lys Gln Val Thr Glu Tyr Ala Ser Asp Glu Pro Trp Gly Trp		
[0665]	485	490	495
[0666]	Tyr Lys Ser Pro Ser LyS Leu		
[0667]	500		
[0668]	<210>7		
[0669]	<211>20		
[0670]	<212>DNA		
[0671]	<213>Artificial Sequence		
[0672]	<220>		
[0673]	<223>Primer		
[0674]	<400>7		
[0675]	gccatgcaa ctgagtaaag		20
[0676]	<210>8		
[0677]	<211>20		
[0678]	<212>DNA		
[0679]	<213>Artificial Sequence		
[0680]	<220>		
[0681]	<223>Primer		
[0682]	<400>8		
[0683]	caattttatt cgctctgtgc		20
[0684]	<210>9		
[0685]	<211>20		
[0686]	<212>DNA		
[0687]	<213>Artificial Sequence		
[0688]	<220>		
[0689]	<223>Primer		
[0690]	<400>9		
[0691]	tgcaacatcg cgtcttattc		20
[0692]	<210>10		
[0693]	<211>22		
[0694]	<212>DNA		
[0695]	<213>Artificial Sequence		
[0696]	<220>		
[0697]	<223>Primer		

[0698]	<400>10	
[0699]	gcaactagca agagcataca cc	22
[0700]	<210>11	
[0701]	<211>20	
[0702]	<212>DNA	
[0703]	<213>Artificial Sequence	
[0704]	<220>	
[0705]	<223>Primer	
[0706]	<400>11	
[0707]	acctgacatc atgcctttgg	20
[0708]	<210>12	
[0709]	<211>20	
[0710]	<212>DNA	
[0711]	<213>Artificial Sequence	
[0712]	<220>	
[0713]	<223>Primer	
[0714]	<400>12	
[0715]	ccggtcatca gctaacttcc	20
[0716]	<210>13	
[0717]	<211>26	
[0718]	<212>DNA	
[0719]	<213>Artificial Sequence	
[0720]	<220>	
[0721]	<223>Primer	
[0722]	<400>13	
[0723]	cccttcgtca taaaatatac tagcaa	26
[0724]	<210>14	
[0725]	<211>20	
[0726]	<212>DNA	
[0727]	<213>Artificial Sequence	
[0728]	<220>	
[0729]	<223>Primer	
[0730]	<400>14	
[0731]	tectccaaca tgctctttcg	20
[0732]	<210>15	
[0733]	<211>20	
[0734]	<212>DNA	
[0735]	<213>Artificial Sequence	
[0736]	<220>	

[0737]	<223>Primer	
[0738]	<400>15	
[0739]	cagagaagtt tacgccgttg	20
[0740]	<210>16	
[0741]	<211>25	
[0742]	<212>DNA	
[0743]	<213>Artificial Sequence	
[0744]	<220>	
[0745]	<223>Primer	
[0746]	<400>16	
[0747]	tttttaaata agatgaacgg tcaaa	25
[0748]	<210>17	
[0749]	<211>20	
[0750]	<212>DNA	
[0751]	<213>Artificial Sequence	
[0752]	<220>	
[0753]	<223>Primer	
[0754]	<400>17	
[0755]	ctctccacce tctgcttctg	20
[0756]	<210>18	
[0757]	<211>19	
[0758]	<212>DNA	
[0759]	<213>Artificial sequence	
[0760]	<220>	
[0761]	<223>Primer	
[0762]	<400>18	
[0763]	ctctccgctt gaaccatc	19
[0764]	<210>19	
[0765]	<211>21	
[0766]	<212>DNA	
[0767]	<213>Artificial Sequence	
[0768]	<220>	
[0769]	<223>Primer	
[0770]	<400>19	
[0771]	gcatggctga ttgtgtatct g	21
[0772]	<210>20	
[0773]	<211>21	
[0774]	<212>DNA	
[0775]	<213>Artificial Sequence	

[0776]	<220>	
[0777]	<223>Primer	
[0778]	<400>20	
[0779]	ttccaaacct acggacaaaa g	21
[0780]	<210>21	
[0781]	<211>21	
[0782]	<212>DNA	
[0783]	<213>Artificial Sequence	
[0784]	<220>	
[0785]	<223>Primer	
[0786]	<400>21	
[0787]	ttcctcttct cttgtgcaaa c	21
[0788]	<210>22	
[0789]	<211>20	
[0790]	<212>DNA	
[0791]	<213>Artificial Sequence	
[0792]	<220>	
[0793]	<223>Primer	
[0794]	<400>22	
[0795]	cacggaagcc aattcagatg	20
[0796]	<210>23	
[0797]	<211>21	
[0798]	<212>DNA	
[0799]	<213>Artificial Sequence	
[0800]	<220>	
[0801]	<223>Primer	
[0802]	<400>23	
[0803]	ctatcctctc ctgatggcaa c	21
[0804]	<210>24	
[0805]	<211>24	
[0806]	<212>DNA	
[0807]	<213>Artificial Sequence	
[0808]	<220>	
[0809]	<223>Primer	
[0810]	<400>24	
[0811]	tggtactag aatgatgctc aaag	24
[0812]	<210>25	
[0813]	<211>20	
[0814]	<212>DNA	

[0815]	<213>Artificial Sequence	
[0816]	<220>	
[0817]	<223>Primer	
[0818]	<400>25	
[0819]	ccttttgtgt cgcttttgag	20
[0820]	<210>26	
[0821]	<211>21	
[0822]	<212>DNA	
[0823]	<213>Artificial Sequence	
[0824]	<220>	
[0825]	<223>Primer	
[0826]	<400>26	
[0827]	aaaatagcct tcactcgttg c	21
[0828]	<210>27	
[0829]	<211>20	
[0830]	<212>DNA	
[0831]	<213>Artificial Sequence	
[0832]	<220>	
[0833]	<223>Primer	
[0834]	<400>27	
[0835]	ccatcgattt cgagggtaac	20
[0836]	<210>28	
[0837]	<211>19	
[0838]	<212>DNA	
[0839]	<213>Artificial Sequence	
[0840]	<220>	
[0841]	<223>Primer	
[0842]	<400>28	
[0843]	cgcatccgat aatatgttg	19
[0844]	<210>29	
[0845]	<211>24	
[0846]	<212>DNA	
[0847]	<213>Artificial Sequence	
[0848]	<220>	
[0849]	<223>Primer	
[0850]	<400>29	
[0851]	gtaattagga gtacgactct cgtc	24
[0852]	<210>30	
[0853]	<211>25	

[0854]	<212>DNA	
[0855]	<213>Artificial Sequence	
[0856]	<220>	
[0857]	<223>Primer	
[0858]	<400>30	
[0859]	gcttatagcc tactgtatcc tcctc	25
[0860]	<210>31	
[0861]	<211>20	
[0862]	<212>DNA	
[0863]	<213>Artificial Sequence	
[0864]	<220>	
[0865]	<223>Primer	
[0866]	<400>31	
[0867]	aattggttaa cccagcaagc	20
[0868]	<210>32	
[0869]	<211>20	
[0870]	<212>DNA	
[0871]	<213>Artificial Sequence	
[0872]	<220>	
[0873]	<223>Primer	
[0874]	<400>32	
[0875]	acattgtgaa acggaggaag	20
[0876]	<210>33	
[0877]	<211>20	
[0878]	<212>DNA	
[0879]	<213>Artificial Sequence	
[0880]	<220>	
[0881]	<223>Primer	
[0882]	<400>33	
[0883]	gctataagcc agctgcaaac	20
[0884]	<210>34	
[0885]	<211>19	
[0886]	<212>DNA	
[0887]	<213>Artificial Sequence	
[0888]	<220>	
[0889]	<223>Primer	
[0890]	<400>34	
[0891]	gcagttggta cggacttcg	19
[0892]	<210>35	

[0893]	<211>21	
[0894]	<212>DNA	
[0895]	<213>Artificial Sequence	
[0896]	<220>	
[0897]	<223>Primer	
[0898]	<400>35	
[0899]	cctaaatatt tgacgccgtt g	21
[0900]	<210>36	
[0901]	<211>20	
[0902]	<212>DNA	
[0903]	<213>Artificial Sequence	
[0904]	<220>	
[0905]	<223>Primer	
[0906]	<400>36	
[0907]	tgaagaggag ggtaccgatg	20
[0908]	<210>37	
[0909]	<211>20	
[0910]	<212>DNA	
[0911]	<213>Artificial Sequence	
[0912]	<220>	
[0913]	<223>Primer	
[0914]	<400>37	
[0915]	caccactcca cacctgacac	20
[0916]	<210>38	
[0917]	<211>20	
[0918]	<212>DNA	
[0919]	<213>Artificial Sequence	
[0920]	<220>	
[0921]	<223>Primer	
[0922]	<400>38	
[0923]	gtacggaaca cacgcacaag	20
[0924]	<210>39	
[0925]	<211>20	
[0926]	<212>DNA	
[0927]	<213>Artificial Sequence	
[0928]	<220>	
[0929]	<223>Primer	
[0930]	<400>39	
[0931]	tgttggttgtt gttgctgctg	20

[0932]	<210>40	
[0933]	<211>21	
[0934]	<212>DNA	
[0935]	<213>Artificial Sequence	
[0936]	<220>	
[0937]	<223>Primer	
[0938]	<400>40	
[0939]	gccgtgagcc atatacactt g	21
[0940]	<210>41	
[0941]	<211>20	
[0942]	<212>DNA	
[0943]	<213>Artificial Sequence	
[0944]	<220>	
[0945]	<223>Primer	
[0946]	<400>41	
[0947]	agctccagct cctcctcgat	20
[0948]	<210>42	
[0949]	<211>20	
[0950]	<212>DNA	
[0951]	<213>Artificial Sequence	
[0952]	<220>	
[0953]	<223>Primer	
[0954]	<400>42	
[0955]	tatctctcac cgaccccaaa	20
[0956]	<210>43	
[0957]	<211>19	
[0958]	<212>DNA	
[0959]	<213>Artificial Sequence	
[0960]	<220>	
[0961]	<223>Primer	
[0962]	<400>43	
[0963]	tgttgccatc aggagagga	19
[0964]	<210>44	
[0965]	<211>20	
[0966]	<212>DNA	
[0967]	<213>Artificial Sequence	
[0968]	<220>	
[0969]	<223>Primer	
[0970]	<400>44	

[0971]	ctcttgatga agcagcatgg	20
[0972]	<210>45	
[0973]	<211>21	
[0974]	<212>DNA	
[0975]	<213>Artificial Sequence	
[0976]	<220>	
[0977]	<223>Primer	
[0978]	<400>45	
[0979]	cccagtaaat gcaaccttgt c	21
[0980]	<210>46	
[0981]	<211>20	
[0982]	<212>DNA	
[0983]	<213>Artificial Sequence	
[0984]	<220>	
[0985]	<223>Primer	
[0986]	<400>46	
[0987]	ggcaacatgg aaggtagctc	20
[0988]	<210>47	
[0989]	<211>20	
[0990]	<212>DNA	
[0991]	<213>Artificial Sequence	
[0992]	<220>	
[0993]	<223>Primer	
[0994]	<400>47	
[0995]	ccatgcaacc atcctttctt	20
[0996]	<210>48	
[0997]	<211>20	
[0998]	<212>DNA	
[0999]	<213>Artificial Sequence	
[1000]	<220>	
[1001]	<223>Primer	
[1002]	<400>48	
[1003]	ttatggcttc agctgctcct	20
[1004]	<210>49	
[1005]	<211>21	
[1006]	<212>DNA	
[1007]	<213>Artificial Sequence	
[1008]	<220>	
[1009]	<223>Primer	

[1010]	<400>49	
[1011]	caatggcttc ttcttcagtg c	21
[1012]	<210>50	
[1013]	<211>20	
[1014]	<212>DNA	
[1015]	<213>Artificial Sequence	
[1016]	<220>	
[1017]	<223>Primer	
[1018]	<400>50	
[1019]	gcccgttggtt agtgaaggac	20
[1020]	<210>51	
[1021]	<211>20	
[1022]	<212>DNA	
[1023]	<213>Artificial Sequence	
[1024]	<220>	
[1025]	<223>Primer	
[1026]	<400>51	
[1027]	gtaccatccc cacggctcat	20
[1028]	<210>52	
[1029]	<211>19	
[1030]	<212>DNA	
[1031]	<213>Artificial Sequence	
[1032]	<220>	
[1033]	<223>Primer	
[1034]	<400>52	
[1035]	cgagcgatgc cagagatta	19
[1036]	<210>53	
[1037]	<211>20	
[1038]	<212>DNA	
[1039]	<213>Artificial Sequence	
[1040]	<220>	
[1041]	<223>Primer	
[1042]	<400>53	
[1043]	agcacatggc aaatcaaaca	20
[1044]	<210>54	
[1045]	<211>20	
[1046]	<212>DNA	
[1047]	<213>Artificial Sequence	
[1048]	<220>	

[1049]	<223>Primer	
[1050]	<400>54	
[1051]	tgctcctttg tcatacacacc	20
[1052]	<210>55	
[1053]	<211>20	
[1054]	<212>DNA	
[1055]	<213>Artificial Sequence	
[1056]	<220>	
[1057]	<223>Primer	
[1058]	<400>55	
[1059]	tttccaccaa gttccagtga	20
[1060]	<210>56	
[1061]	<211>20	
[1062]	<212>DNA	
[1063]	<213>Artificial Sequence	
[1064]	<220>	
[1065]	<223>Primer	
[1066]	<400>56	
[1067]	ttcgctgcag aacagatgac	20
[1068]	<210>57	
[1069]	<211>21	
[1070]	<212>DNA	
[1071]	<213>Artificial Sequence	
[1072]	<220>	
[1073]	<223>Primer	
[1074]	<400>57	
[1075]	ctgatggtta cgcgacaatt t	21
[1076]	<210>58	
[1077]	<211>20	
[1078]	<212>DNA	
[1079]	<213>Artificial Sequence	
[1080]	<220>	
[1081]	<223>Primer	
[1082]	<400>58	
[1083]	atttgaaccg ggacagaaca	20
[1084]	<210>59	
[1085]	<211>20	
[1086]	<212>DNA	
[1087]	<213>Artificial Sequence	

[1088]	<220>	
[1089]	<223>Primer	
[1090]	<400>59	
[1091]	ttttgatgtg ccctctcctt	20
[1092]	<210>60	
[1093]	<211>20	
[1094]	<212>DNA	
[1095]	<213>Artificial Sequence	
[1096]	<220>	
[1097]	<223>Primer	
[1098]	<400>60	
[1099]	tgggtaatct tgttctggag	20
[1100]	<210>61	
[1101]	<211>20	
[1102]	<212>DNA	
[1103]	<213>Artificial Sequence	
[1104]	<220>	
[1105]	<223>Primer	
[1106]	<400>61	
[1107]	agtgccaaat gcatgctaga	20
[1108]	<210>62	
[1109]	<211>20	
[1110]	<212>DNA	
[1111]	<213>Artificial Sequence	
[1112]	<220>	
[1113]	<223>Primer	
[1114]	<400>62	
[1115]	tggggctcaa aaacctactg	20
[1116]	<210>63	
[1117]	<211>18	
[1118]	<212>DNA	
[1119]	<213>Artificial Sequence	
[1120]	<220>	
[1121]	<223>Primer	
[1122]	<400>63	
[1123]	gtccgggcca agtacctc	18
[1124]	<210>64	
[1125]	<211>20	
[1126]	<212>DNA	

[1127]	<213>Artificial Sequence	
[1128]	<220>	
[1129]	<223>Primer	
[1130]	<400>64	
[1131]	atctctcacc gaccccaaat	20
[1132]	<210>65	
[1133]	<211>21	
[1134]	<212>DNA	
[1135]	<213>Artificial Sequence	
[1136]	<220>	
[1137]	<223>Primer	
[1138]	<400>65	
[1139]	ccattggaag agagacaggt g	21
[1140]	<210>66	
[1141]	<211>20	
[1142]	<212>DNA	
[1143]	<213>Artificial Sequence	
[1144]	<220>	
[1145]	<223>Primer	
[1146]	<400>66	
[1147]	tgttggtggt gttgctgctg	20
[1148]	<210>67	
[1149]	<211>20	
[1150]	<212>DNA	
[1151]	<213>Artificial Sequence	
[1152]	<220>	
[1153]	<223>Primer	
[1154]	<400>67	
[1155]	tggggctcaa aaacctactg	20
[1156]	<210>68	
[1157]	<211>20	
[1158]	<212>DNA	
[1159]	<213>Artificial Sequence	
[1160]	<220>	
[1161]	<223>Primer	
[1162]	<400>68	
[1163]	ggttggtctt ccttcaggtg	20
[1164]	<210>69	
[1165]	<211>20	

[1166]	<212>DNA	
[1167]	<213>Artificial Sequence	
[1168]	<220>	
[1169]	<223>Primer	
[1170]	<400>69	
[1171]	ggtccaaaag caaccaaaga	20
[1172]	<210>70	
[1173]	<211>20	
[1174]	<212>DNA	
[1175]	<213>Artificial Sequence	
[1176]	<220>	
[1177]	<223>Primer	
[1178]	<400>70	
[1179]	actggtaaaa agattatggc	20
[1180]	<210>71	
[1181]	<211>20	
[1182]	<212>DNA	
[1183]	<213>Artificial Sequence	
[1184]	<220>	
[1185]	<223>Primer	
[1186]	<400>71	
[1187]	caagccgate aaccagtaca	20
[1188]	<210>72	
[1189]	<211>19	
[1190]	<212>DNA	
[1191]	<213>Artificial Sequence	
[1192]	<220>	
[1193]	<223>Primer	
[1194]	<400>72	
[1195]	ccatgctgca agcaatgta	19
[1196]	<210>73	
[1197]	<211>22	
[1198]	<212>DNA	
[1199]	<213>Artificial Sequence	
[1200]	<220>	
[1201]	<223>Primer	
[1202]	<400>73	
[1203]	aaccatagga gcagctgaaa ta	22
[1204]	<210>74	

[1205]	<211>20	
[1206]	<212>DNA	
[1207]	<213>Artificial Sequence	
[1208]	<220>	
[1209]	<223>Primer	
[1210]	<400>74	
[1211]	accctggtgt agacaaggta	20
[1212]	<210>75	
[1213]	<211>22	
[1214]	<212>DNA	
[1215]	<213>Artificial Sequence	
[1216]	<220>	
[1217]	<223>Primer	
[1218]	<400>75	
[1219]	gggagttatg aaactggtat at	22
[1220]	<210>76	
[1221]	<211>20	
[1222]	<212>DNA	
[1223]	<213>Artificial Sequence	
[1224]	<220>	
[1225]	<223>Primer	
[1226]	<400>76	
[1227]	ataggagcag ctgaagccat	20
[1228]	<210>77	
[1229]	<211>20	
[1230]	<212>DNA	
[1231]	<213>Artificial Sequence	
[1232]	<220>	
[1233]	<223>Primer	
[1234]	<400>77	
[1235]	gtcccgcact tcagaattag	20
[1236]	<210>78	
[1237]	<211>20	
[1238]	<212>DNA	
[1239]	<213>Artificial Sequence	
[1240]	<220>	
[1241]	<223>Primer	
[1242]	<400>78	
[1243]	ctctgcttct gcctctgatt	20

[1244]	<210>79	
[1245]	<211>21	
[1246]	<212>DNA	
[1247]	<213>Artificial Sequence	
[1248]	<220>	
[1249]	<223>Primer	
[1250]	<400>79	
[1251]	ctggctacta gaatgatgct c	21
[1252]	<210>80	
[1253]	<211>30	
[1254]	<212>DNA	
[1255]	<213>Artificial Sequence	
[1256]	<220>	
[1257]	<223>Primer	
[1258]	<400>80	
[1259]	aattccatgg ggttggcttt ccttcaggtg	30
[1260]	<210>81	
[1261]	<211>30	
[1262]	<212>DNA	
[1263]	<213>Artificial Sequence	
[1264]	<220>	
[1265]	<223>Primer	
[1266]	<400>81	
[1267]	aattactagt ttccaccaag ttccagtgaa	30
[1268]	<210>82	
[1269]	<211>30	
[1270]	<212>DNA	
[1271]	<213>Artificial Sequence	
[1272]	<220>	
[1273]	<223>Primer	
[1274]	<400>82	
[1275]	aattgctagc ggtccaaaag caaccaaaaga	30
[1276]	<210>83	
[1277]	<211>24	
[1278]	<212>DNA	
[1279]	<213>Artificial Sequence	
[1280]	<220>	
[1281]	<223>Primer	
[1282]	<400>83	

[1283]	cttgttgaat tagatggtga tggt	24
[1284]	<210>84	
[1285]	<211>24	
[1286]	<212>DNA	
[1287]	<213>Artificial Sequence	
[1288]	<220>	
[1289]	<223>Primer	
[1290]	<400>84	
[1291]	gttgtgggag ttgtagttgt attc	24
[1292]	<210>85	
[1293]	<211>20	
[1294]	<212>DNA	
[1295]	<213>Artificial Sequence	
[1296]	<220>	
[1297]	<223>Primer	
[1298]	<400>85	
[1299]	tagcttcaca tcccatgtg	20
[1300]	<210>86	
[1301]	<211>20	
[1302]	<212>DNA	
[1303]	<213>Artificial Sequence	
[1304]	<220>	
[1305]	<223>Primer	
[1306]	<400>86	
[1307]	gcaccttcac atcttgctgt	20
[1308]	<210>87	
[1309]	<211>20	
[1310]	<212>DNA	
[1311]	<213>Artificial Sequence	
[1312]	<220>	
[1313]	<223>Primer	
[1314]	<400>87	
[1315]	acatgcacct ggactatgac	20
[1316]	<210>88	
[1317]	<211>20	
[1318]	<212>DNA	
[1319]	<213>Artificial Sequence	
[1320]	<220>	
[1321]	<223>Primer	

[1322]	<400>88	
[1323]	tgctgagaga tgccaagatg	20
[1324]	<210>89	
[1325]	<211>56	
[1326]	<212>DNA	
[1327]	<213>oryza rufi pogon	
[1328]	<400>89	
[1329]	tgcatctact gggagttatg aaactggtat atatttcagc tgctcctatg gttaag	56
[1330]	<210>90	
[1331]	<211>56	
[1332]	<212>DNA	
[1333]	<213>oryza rufi pogon	
[1334]	<400>90	
[1335]	tgcatctact gggagttatg aaactggtat atatttcagc tgctcctatg gttaag	56
[1336]	<210>91	
[1337]	<211>56	
[1338]	<212>DNA	
[1339]	<213>oryza rufi pogon	
[1340]	<400>91	
[1341]	tgcatctact gggagttatg aaactggtat atatttcagc tgctcctatg gttaag	56
[1342]	<210>92	
[1343]	<211>64	
[1344]	<212>DNA	
[1345]	<213>oryza rufi pogon	
[1346]	<400>92	
[1347]	tgcatctact gggagttatg aaactggtaa aaagattatg gettcagctg ctcctatggt	60
[1348]	taag	64
[1349]	<210>93	
[1350]	<211>56	
[1351]	<212>DNA	
[1352]	<213>oryza nivara	
[1353]	<400>93	
[1354]	tgcatctact gggagttatg aaactggtat atatttcagc tgctcctatg gttaag	56
[1355]	<210>94	
[1356]	<211>64	
[1357]	<212>DNA	
[1358]	<213>oryza nivara	
[1359]	<400>94	
[1360]	tgcatctact gggagttatg aaactggtaa aaagattatg gettcagctg ctcctatggt	60

[1361]	taag	64
[1362]	<210>95	
[1363]	<211>64	
[1364]	<212>DNA	
[1365]	<213>oryza sativa	
[1366]	<400>95	
[1367]	tgcatttact gggagttatg aaactggtaa aaagattatg gcttcagctg ctcctatggt	60
[1368]	taag	64
[1369]	<210>96	
[1370]	<211>10	
[1371]	<212>PRT	
[1372]	<213>oryza sativa	
[1373]	<400>96	
[1374]	Val Thr Leu Glu Leu Gly Gly Lys Ser Pro	
[1375]	1 5 10	

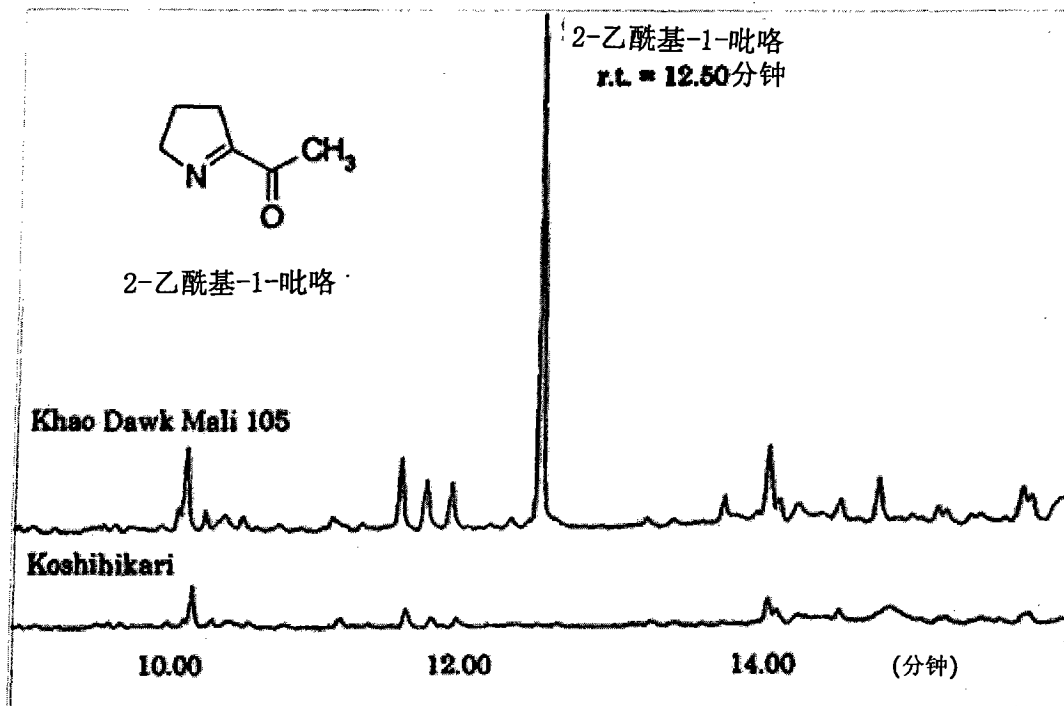


图 1

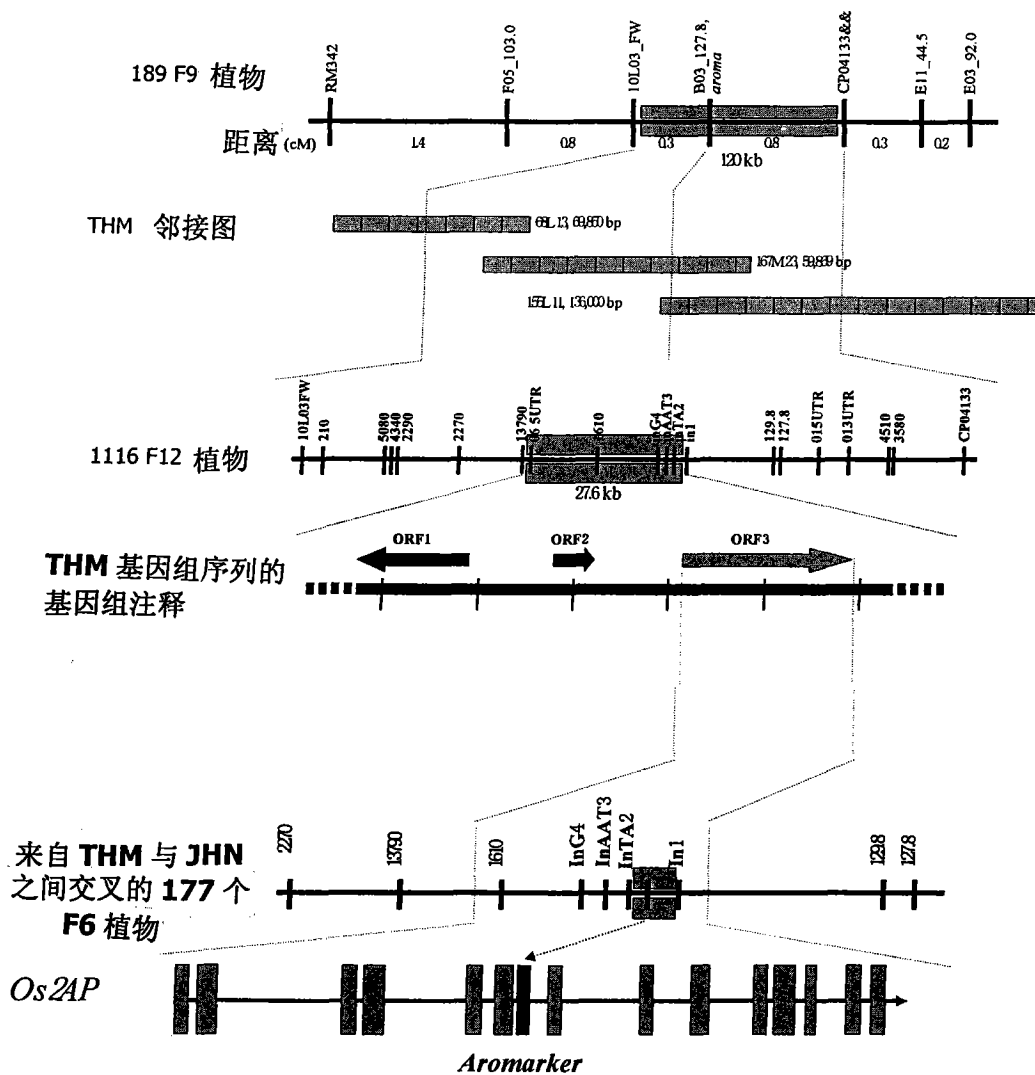


图 2

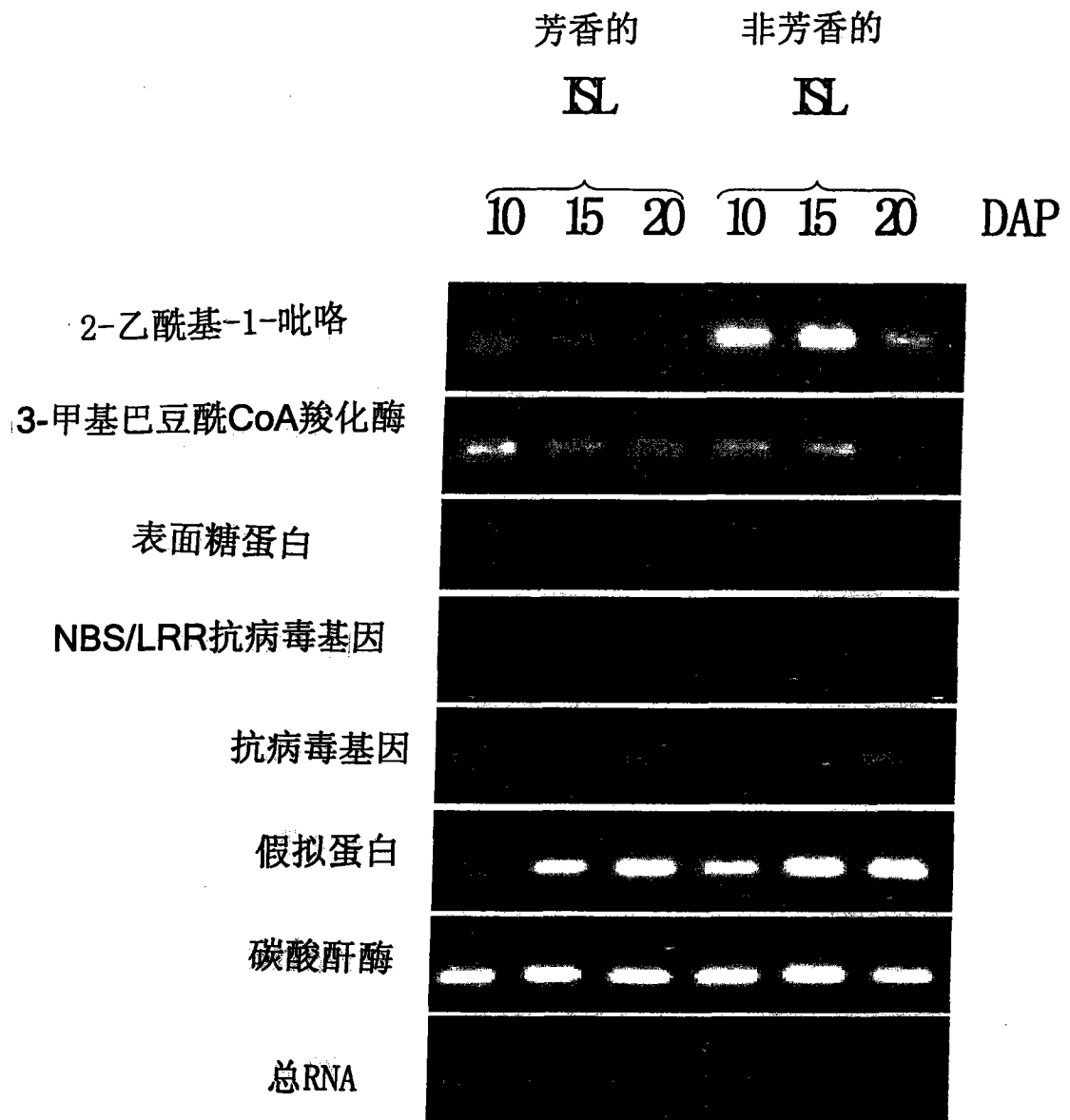


图 3A

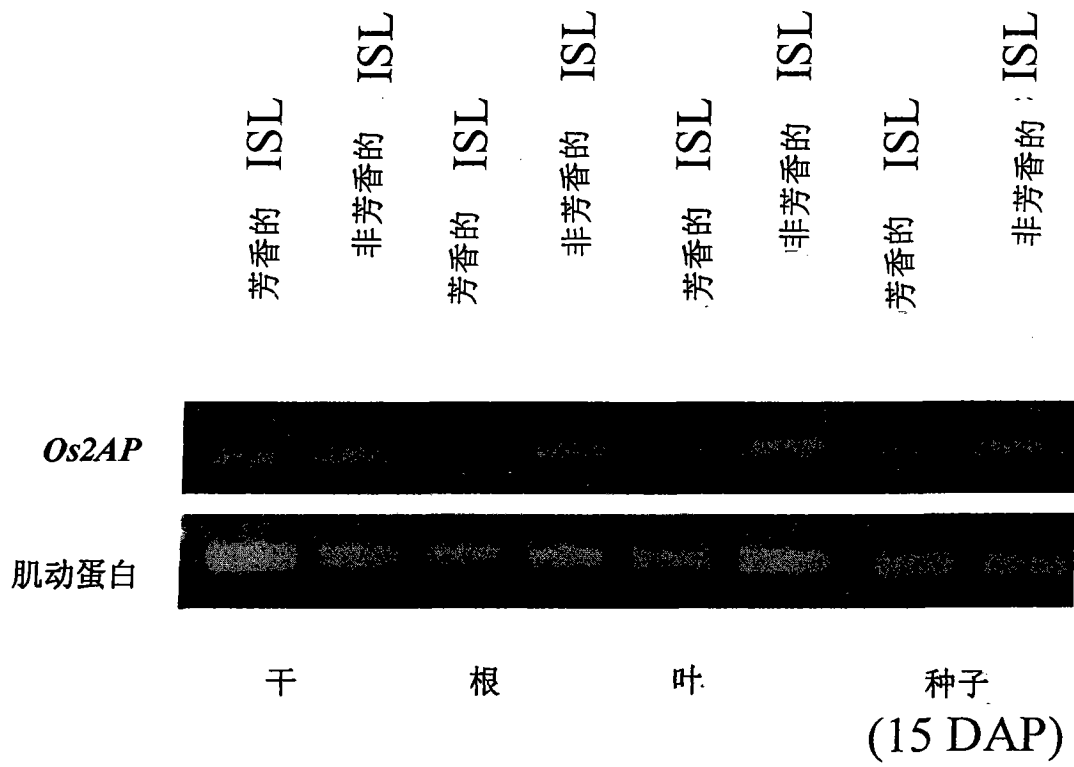


图 3B

2AP 累积分析

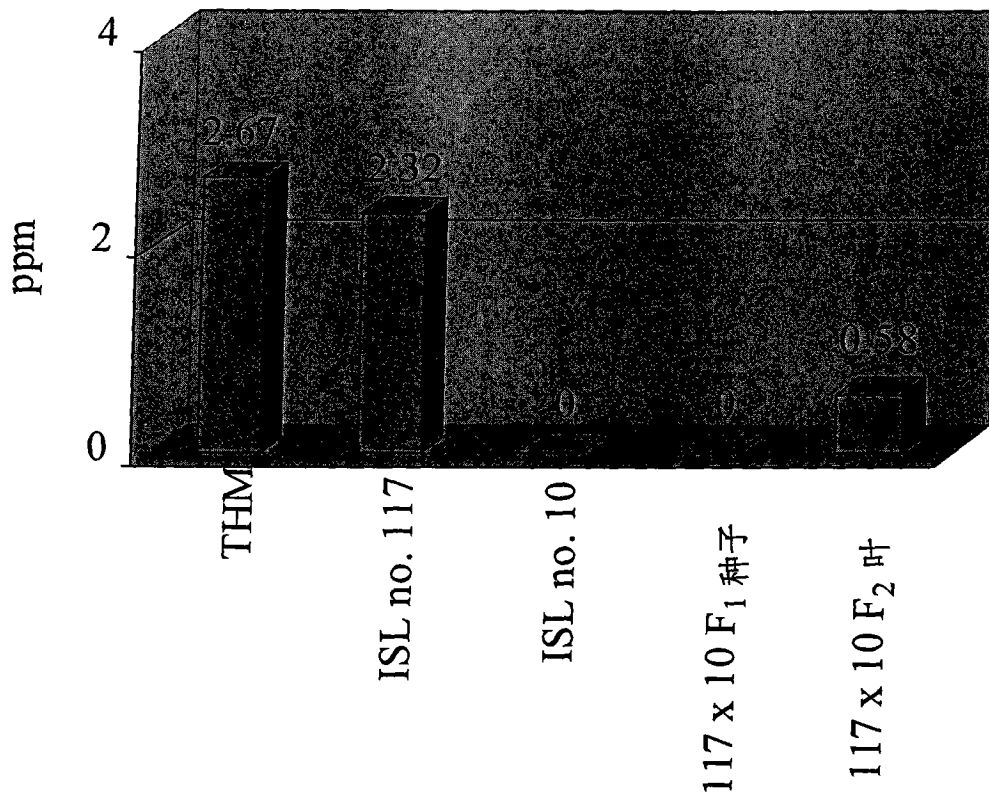


图 3C

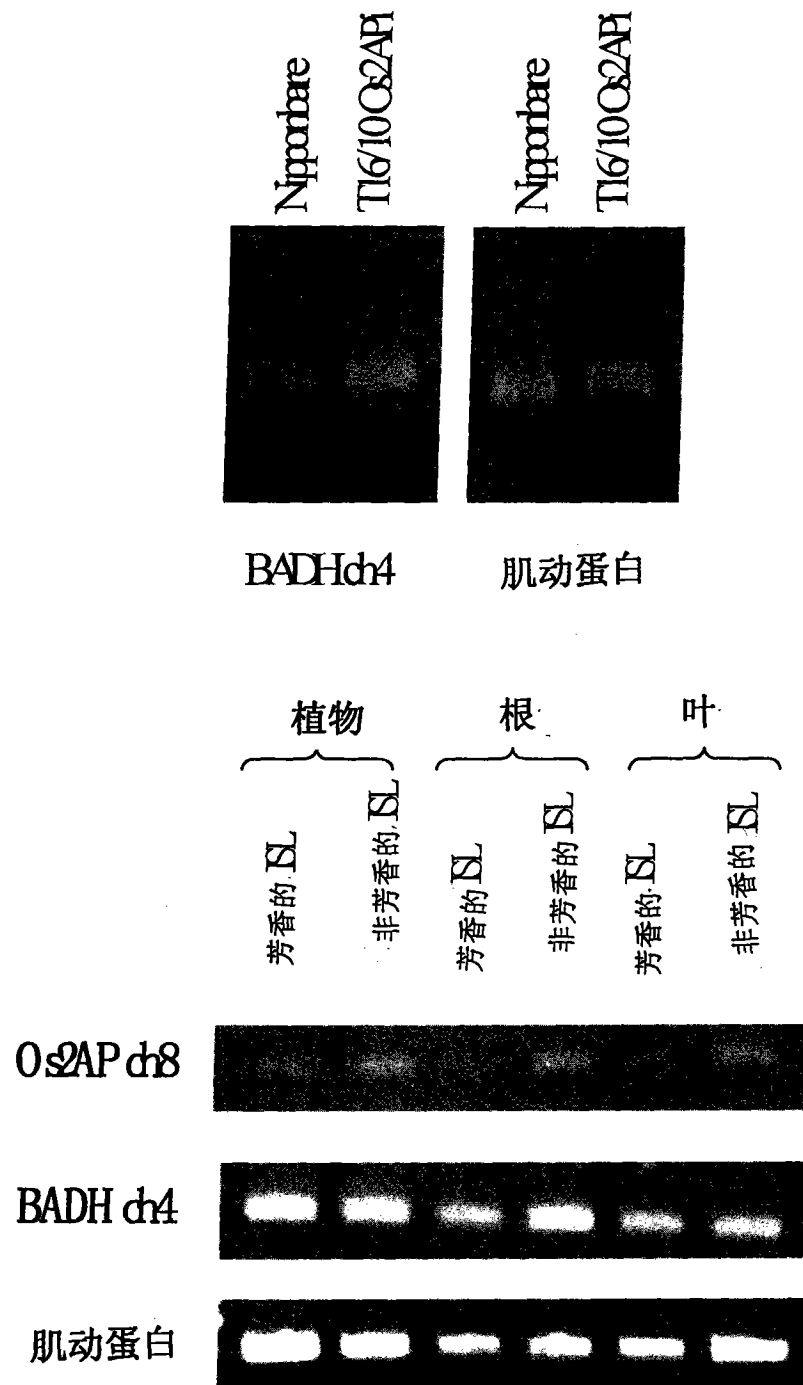


图 3D

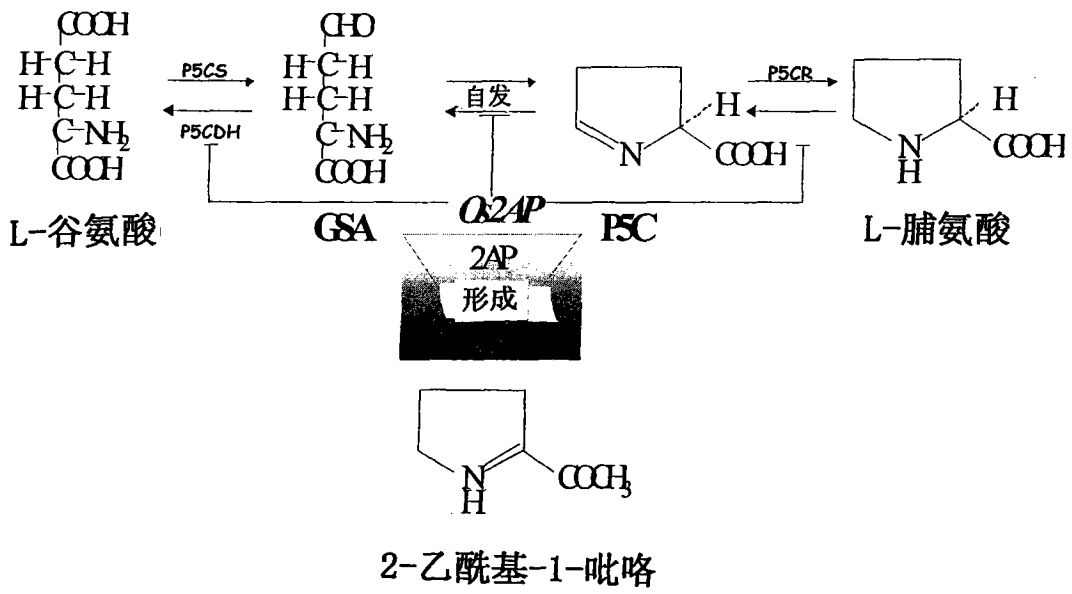


图 4A

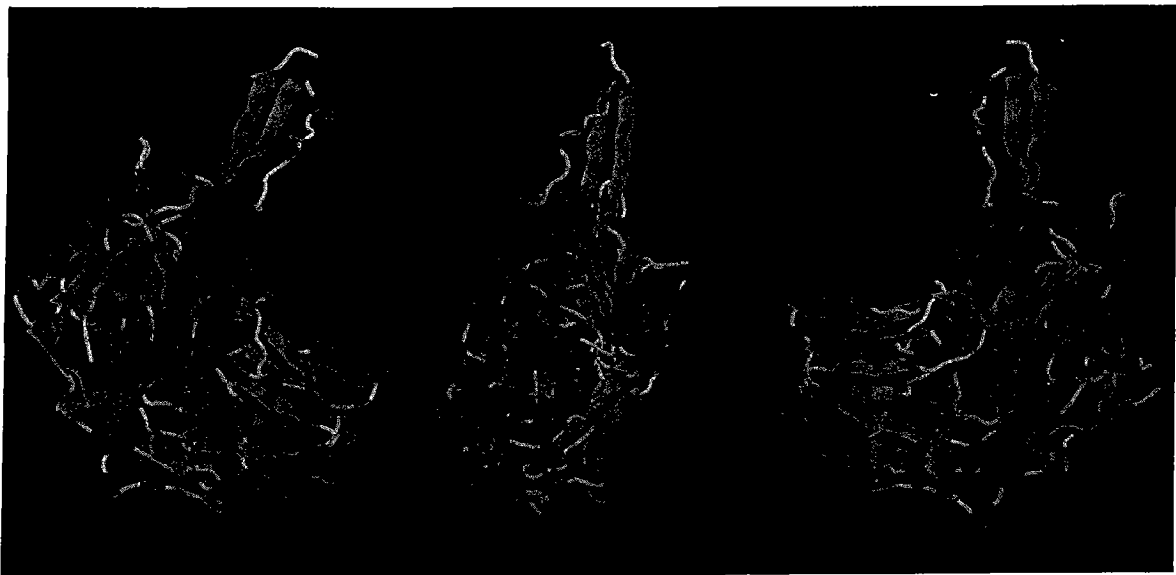


图 4B

Nipponbare THM	ATGGCCACGGCGATCCCGCAGCGGCAGCTCTTCGTGCGCCGGCGAGTGGCGCGCCCCCGCG ATGGCCACGGCGATCCCGCAGCGGCAGCTCTTCGTGCGCCGGCGAGTGGCGCGCCCCCGCG
Nipponbare THM	CTCGGCCCGCCCTCCCCGTCGTCAACCCCGCCACCGAGTCCCCCATCGGGGAGATCCCC CTCGGCCCGCCCTCCCCGTCGTCAACCCCGCCACCGAGTCCCCCATCGGGGAGATCCCC
Nipponbare THM	GCGGGCACGGCGGAGGACGTGGACGCGGGCGGTGGCGGGCGGGGAGGCGCTAAGAGG GCGGGCACGGCGGAGGACGTGGACGCGGGCGGTGGCGGGCGGGGAGGCGCTAAGAGG
Nipponbare THM	AACCGGGCCGCGACTGGGCGCGCGCCGGGCGCCGTCCGGGCCAAGTAC TCCGCGCA AACCGGGCCGCGACTGGGCGCGCGCCGGGCGCCGTCCGGGCCAAGTAC TCCGCGCA
Nipponbare THM	ATCGCGCCAAGATAATCGAGAGGAAATCTGAGCTGGCTAGACTAGAGACGCTTGATTGT ATCGCTGACAAAATAATCGAGAGGAAATCTGAGCTGGCTAGACTAGAGACGCTTGATTGT
Nipponbare THM	GGGAAGCCTCTTGATGAAGCAGCATGGGACATGGACGATGTTGCTGGATGCTTTGAGTAC GGGAAGCCTCTTGATGAAGCAGCATGGGACATGGACGATGTTGCTGGATGCTTTGAGTAC
Nipponbare THM	TTTGAGATCTTGACAGAACTCTTGACAAAAGGCAAAATGCACCTGTCTCTCTTCCAATG TTTGAGATCTTGACAGAACTCTTGACAAAAGGCAAAATGCACCTGTCTCTCTTCCAATG
Nipponbare THM	GAAAACTTTAAATGCTATCTTCGGAAAGAGCCTATCGGTGTAGTTGGTTGATCACACCT GAAAACTTTAAATGCTATCTTCGGAAAGAGCCTATCGGTGTAGTTGGTTGATCACACCT
Nipponbare THM	TGGAATATCCTCTCCTGATGGCAACATGGAAGGTAGCTCCTGCCCTGGCTGCTGGCTGT TGGAATATCCTCTCCTGATGGCAACATGGAAGGTAGCTCCTGCCCTGGCTGCTGGCTGT
Nipponbare THM	ACAGCTGTACTAAAACCATCTGAATTGGCTTCCGTGACTTGTTTGGAGCTTGCTGATGTG ACAGCTGTACTAAAACCATCTGAATTGGCTTCCGTGACTTGTTTGGAGCTTGCTGATGTG
Nipponbare THM	TGTAAAGAGGTTGGTCTTCCTTCAGGTGTGCTAAACATAGTACTGGATTAGGTTCTGAA TGTAAAGAGGTTGGTCTTCCTTCAGGTGTGCTAAACATAGTACTGGATTAGGTTCTGAA
Nipponbare THM	GCCGGTGCTCCTTTGTCATCACACCCTGGTGTAGACAAGGTTGCATTTACTGGGAGTTAT GCCGGTGCTCCTTTGTCATCACACCCTGGTGTAGACAAGGTTGCATTTACTGGGAGTTAT
Nipponbare THM	GAAACTGGTAAAAAGATTATGGCTTCAGCTGCTCCTATGGTTAAGCCTGTTTCACTGGAA GAAACTGGTAFATA-----TTTCAGCTGCTCCTATGGTTAAGCCTGTTTCACTGGAA
Nipponbare THM	CTTGGTGAAAAAGTCCATAGTGGTGTGTTGATGATGTTGATGTTGAAAAAGCTGTTGAG CTTGGTGAAAAAGTCCATAGTGGTGTGTTGATGATGTTGATGTTGAAAAAGCTGTTGAG

图 5A

Nipponbare THM	TGGACTCTCTTTGGTTGCTTTTGGACCAATGGCCAGATTTGCAGTGCAACATCGCGTCTT TGGACTCTCTTTGGTTGCTTTTGGACCAATGGCCAGATTTGCAGTGCAACATCGCGTCTT
Nipponbare THM	ATTCTTCATAAAAAAATCGCTAAAGAATTTCAAGAAAGGATGGTTGCATGGGCCAAAAAT ATTCTTCATAAAAAAATCGCTAAAGAATTTCAAGAAAGGATGGTTGCATGGGCCAAAAAT
Nipponbare THM	ATTAAGGTGTCAGATCCACTTGAAGAGGTTGCAGGCTTGGGCCCGTTGTTAGTGAAGGA ATTAAGGTGTCAGATCCACTTGAAGAGGTTGCAGGCTTGGGCCCGTTGTTAGTGAAGGA
Nipponbare THM	CAGTATGAGAAGATTAAGCAATTTGTATCTACCGCCAAAAGCCAAGGTGCTACCATTCCTG CAGTATGAGAAGATTAAGCAATTTGTATCTACCGCCAAAAGCCAAGGTGCTACCATTCCTG
Nipponbare THM	ACTGTTGGGGTTAGACCCAAGCATCTGGAGAAAGGTTCTATATTTGAACCCAAATCATT ACTGTTGGGGTTAGACCCAAGCATCTGGAGAAAGGTTCTATATTTGAACCCAAATCATT
Nipponbare THM	ACTGATGTCGATACATCAATGCAAAATTTGGAGGGAAGAAGTTTGTGTCAGTCTCTGT ACTGATGTCGATACATCAATGCAAAATTTGGAGGGAAGAAGTTTGTGTCAGTCTCTGT
Nipponbare THM	GTGAAGAATTTAGCACTGAAGAAGAAGCCATTGAATTGGCCAACGATATCATTATGGT GTGAAGAATTTAGCACTGAAGAAGAAGCCATTGAATTGGCCAACGATATCATTATGGT
Nipponbare THM	CTGGCTGGTGCTGTGCTTTCCGGTGACCGCAGC GATGCCAGAGATTAACCTGAGGAGATC CTGGCTGGTGCTGTGCTTTCCGGTGACCGCAGC GATGCCAGAGATTAACCTGAGGAGATC
Nipponbare THM	GATGCCGGAAATTA TCTGGGTGAACCTGCTCGCAAC CCTGCTTCTGCCAAGCTCCATGGGGC GATGCCGGAAATTA TCTGGGTGAACCTGCTCGCAAC CCTGCTTCTGCCAAGCTCCATGGGGC
Nipponbare THM	GGGAACAAGCGCAGCGCTTTGGACGCGAGCTCGGAGAAGGGGGCA TTGACAACCTACCTA GGGAACAAGCGCAGCGCTTTGGACGCGAGCTCGGAGAAGGGGGCA TTGACAACCTACCTA
Nipponbare THM	AGCGTCAAGCAAGTGACGGAGTACGCCTCCGATGAGCCGTGGGGATGGTACAAATCCCCT AGCGTCAAGCAAGTGACGGAGTACGCCTCCGATGAGCCGTGGGGATGGTACAAATCCCCT
Nipponbare THM	TCCAAGCTGTAA TCCAAGCTGTAA

图 5A(继续)

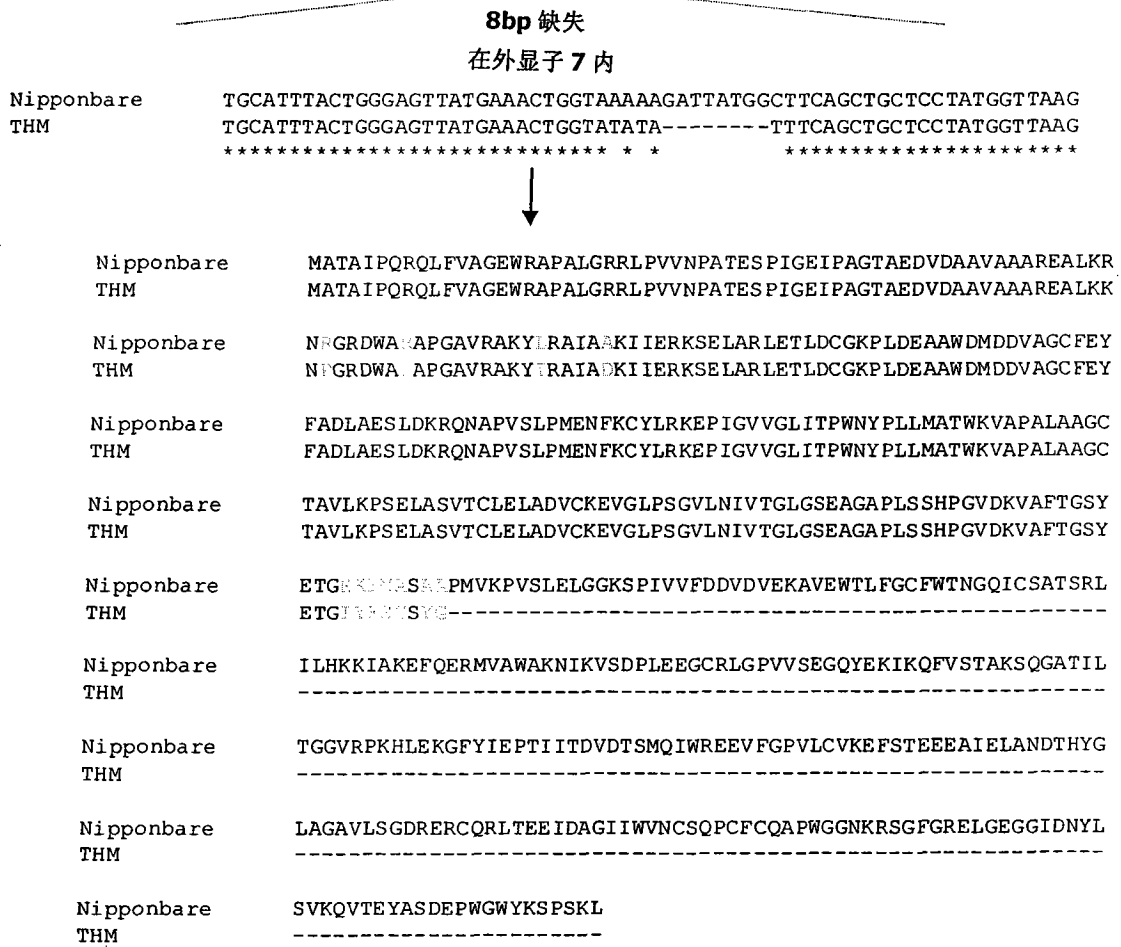


图 5B

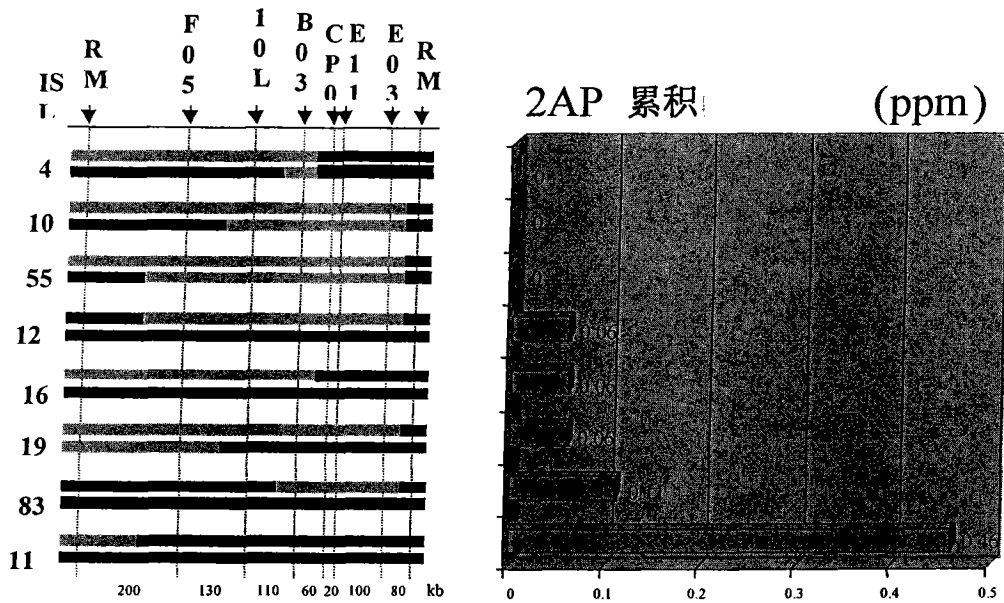


图 6

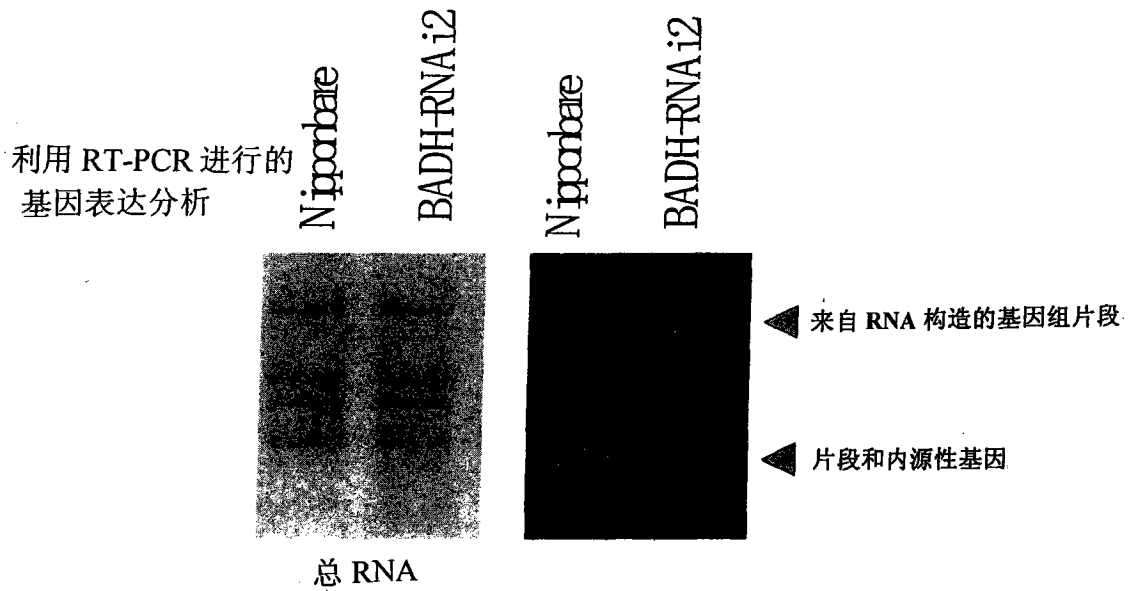
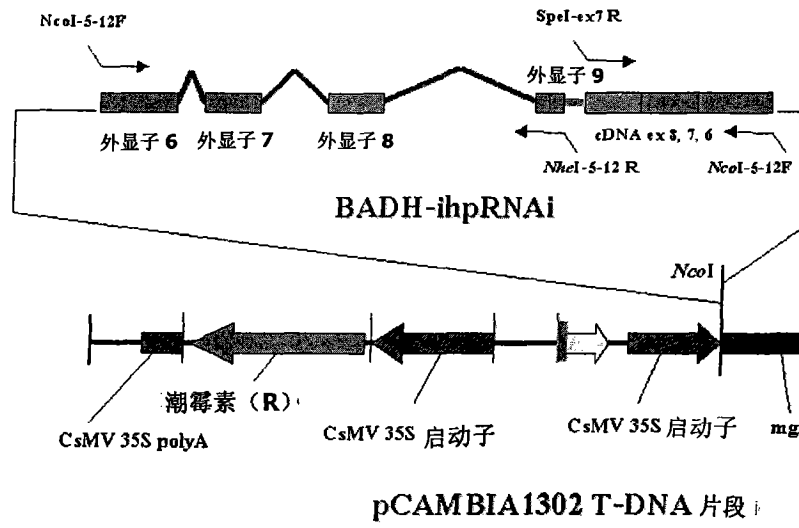


图 7A

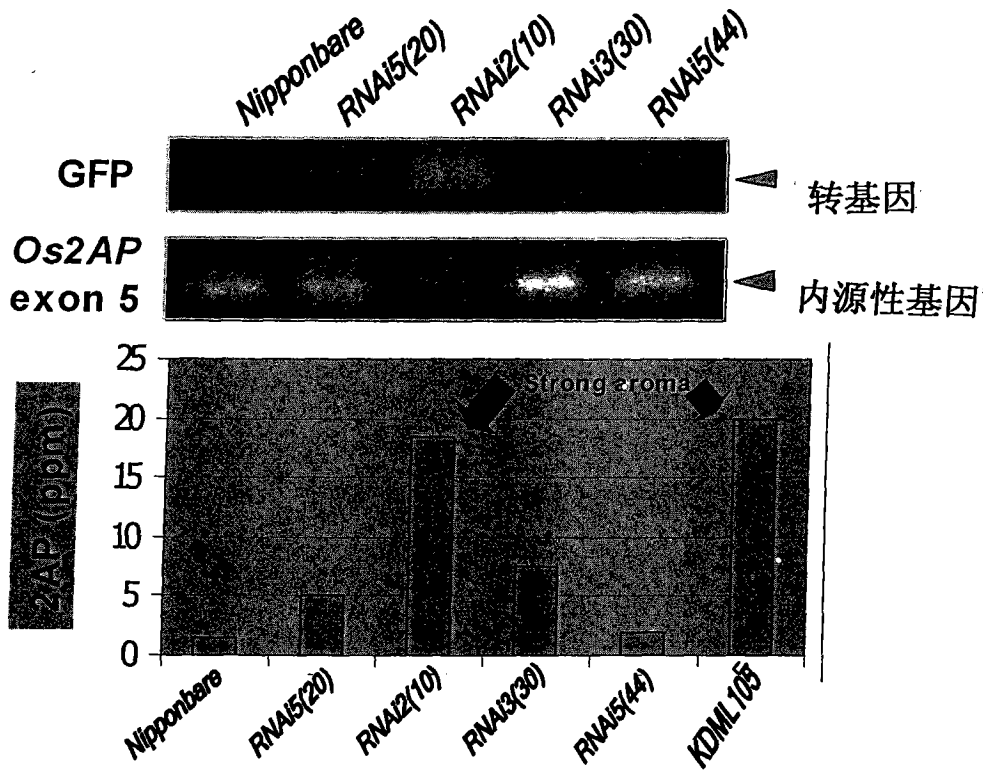


图 7B

野生稻的 8bp indel 基于 PCR 的标记物 Aromarker

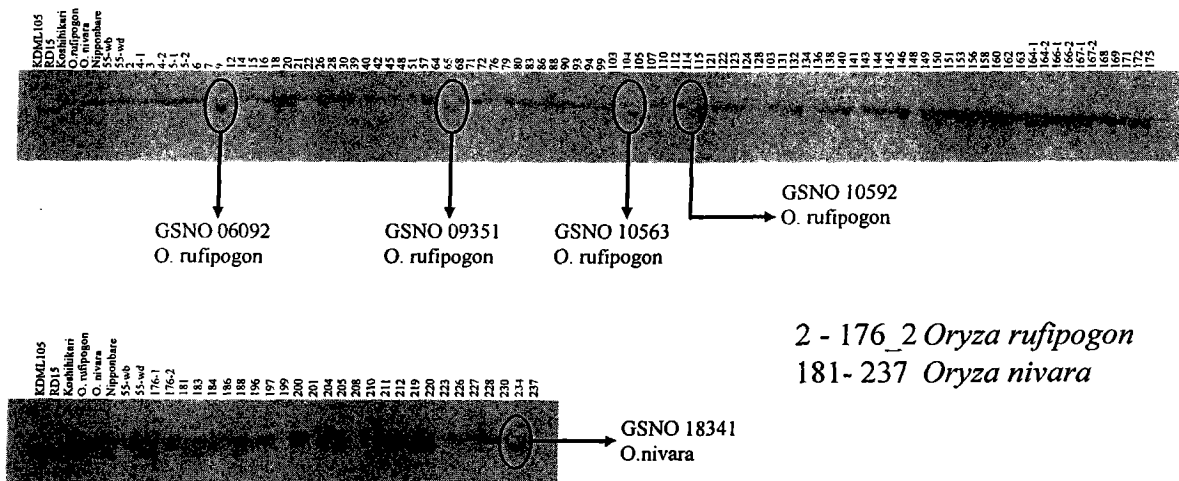


图 8A

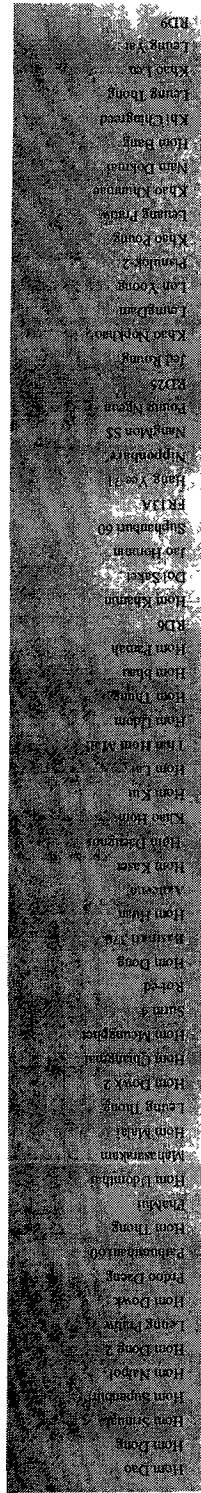


图 8B

Thai Hom Mali					
O. Rufipogon 06092	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 2)	芳香	(A)	
O. Rufipogon 09351	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 89)	芳香	(A)	
O. Rufipogon 10563	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 90)	芳香	(A)	
O. rufipogon 07919	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 91)	芳香	(A)	
O. nivara 18341	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 92)	非芳香	(A)	(NA)
O. nivara 18261	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 93)	非芳香	(A)	(NA)
Nipponbare	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 94)	非芳香	(A)	(NA)
JHN	TGCATTTACTGGGAGTTATGAAACTGGTATATA-----T---TTCAGCTGCTCCTATGGTTAAG	(SEQ ID NO: 95)	非芳香	(A)	(NA)

图 9A

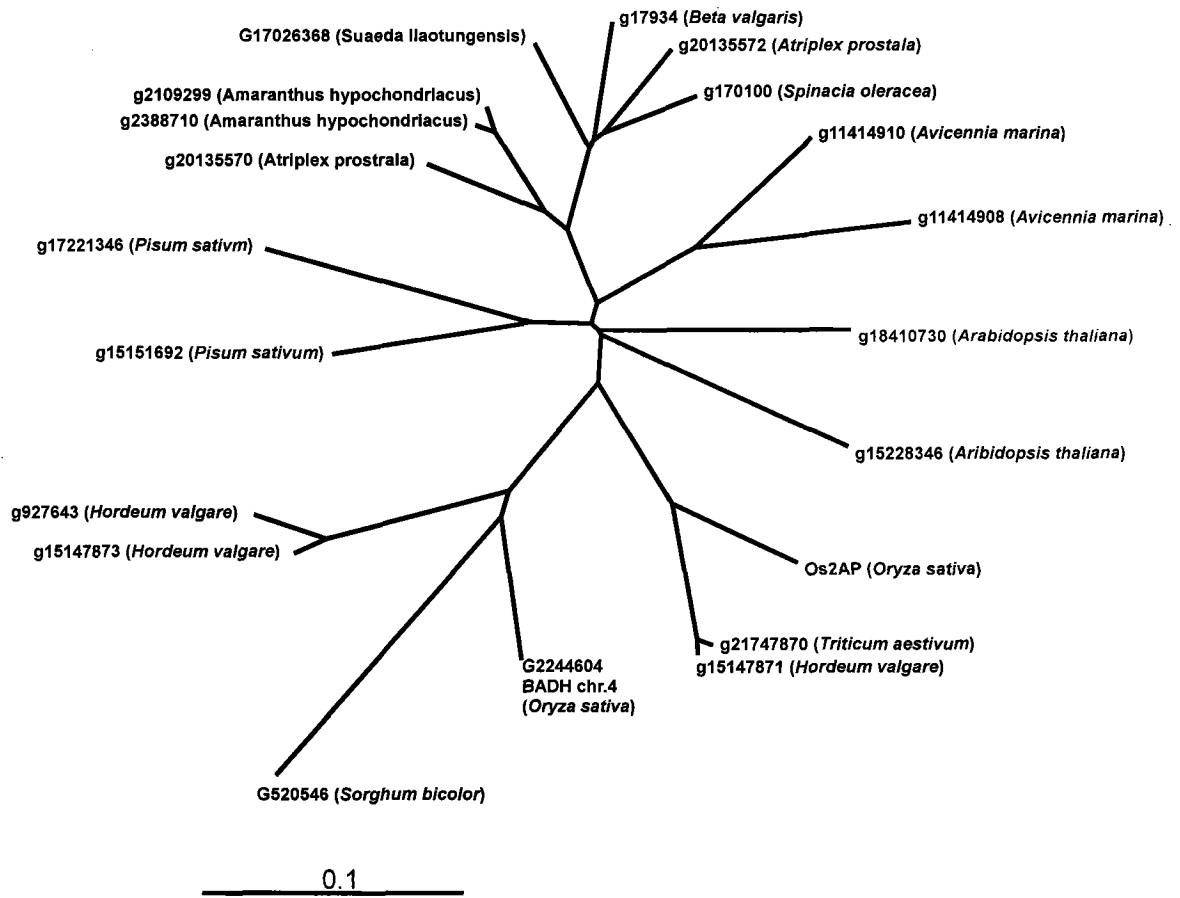


图 9B