



(12)发明专利

(10)授权公告号 CN 105132425 B

(45)授权公告日 2018.10.02

(21)申请号 201510543571.4

(22)申请日 2010.08.11

(65)同一申请的已公布的文献号

申请公布号 CN 105132425 A

(43)申请公布日 2015.12.09

(30)优先权数据

09169019.8 2009.08.31 EP

61/238,230 2009.08.31 US

(62)分案原申请数据

201080038708.6 2010.08.11

(73)专利权人 巴斯夫植物科学有限公司

地址 德国路德维希港

(72)发明人 J·M·库恩 L·P·洛亚尔

M·西伯特 E·杜维尼哥

(74)专利代理机构 北京市中咨律师事务所

11247

代理人 柴云峰 黄革生

(51)Int.Cl.

C12N 15/113(2010.01)

C12N 15/82(2006.01)

C12N 1/21(2006.01)

C12N 1/15(2006.01)

C12N 1/19(2006.01)

A01H 5/00(2018.01)

A01H 5/10(2018.01)

A23K 10/18(2016.01)

(56)对比文件

EP 1645633 A2,2006.04.12,

CN 101405394 A,2009.04.08,

US 7179904 B2,2007.02.20,

审查员 汪未申

权利要求书2页 说明书37页

序列表40页 附图5页

(54)发明名称

用于增强植物中组成型基因表达的调节性
核酸分子

(57)摘要

本发明属于植物分子生物学领域并且提供了用于产生高表达的组成型启动子的方法和产生具有增强的核酸组成型表达的植物的方法,其中将增强核酸表达的核酸(NEENAs)与所述启动子功能性连接和/或导入植物中。

1. 用于产生高表达组成型植物启动子的方法,包括将一个或多个增强核酸表达的核酸(NEENA)分子与启动子功能性连接,其中所述增强核酸表达的核酸(NEENA)分子相对于所述启动子是异源的,所述增强核酸表达的核酸的核苷酸序列为SEQ ID NO:2。

2. 用于产生植物或其部分的方法,所述植物或其部分与相应的对照植物或其部分相比具有一种或多种核酸分子的增强的组成型表达,所述方法包括步骤

a) 将一个或多个SEQ ID NO:2所示的NEENA导入所述植物或其部分;

和

b) 将所述一个或多个NEENA与组成型启动子和与处在所述组成型启动子控制下的核酸分子功能性连接,其中NEENA对所述核酸分子为异源。

3. 根据权利要求1和2任一项所述的方法,包括步骤

a) 将一个或多个SEQ ID NO:2所示的NEENA导入植物或其部分,和

b) 将所述一个或多个NEENA整合至所述植物或其部分的基因组中,从而所述一个或多个NEENA与相对所述一个或多个NEENA为异源的内源组成型表达的核酸功能性连接,和任选地

c) 从所述转化的细胞再生出包含所述一个或多个NEENA的植物或其部分。

4. 根据权利要求1至2任一项所述的方法,包括步骤

a) 提供表达构建体,所述表达构建体包含一个或多个SEQ ID NO:2所示的NEENA,所述NEENA功能性连接到组成型启动子和一个或多个核酸分子,所述一个或多个核酸分子与所述一个或多个NEENA为异源的并处在所述组成型启动子控制下,和

b) 将包含所述一个或多个NEENA的所述表达构建体整合至所述植物或其部分的基因组中,和任选地

c) 从所述转化的植物或其部分再生出包含所述一个或多个表达构建体的植物或其部分。

5. 根据权利要求1至2任一项所述的方法,其中所述植物是单子叶或双子叶植物。

6. 根据权利要求1至2任一项所述的方法,其中所述一个或多个NEENA与靠近所述异源核酸分子的转录起始位点的组成型启动子功能性连接。

7. 根据权利要求6所述的方法,其中所述一个或多个NEENA与距离所述异源核酸分子的转录起始位点2500bp或更少的组成型启动子功能性连接。

8. 根据权利要求6所述的方法,其中所述一个或多个NEENA与距离所述异源核酸分子的转录起始位点2000bp或更少的组成型启动子功能性连接。

9. 根据权利要求6所述的方法,其中所述一个或多个NEENA与距离所述异源核酸分子的转录起始位点1500bp或更少的组成型启动子功能性连接。

10. 根据权利要求6所述的方法,其中所述一个或多个NEENA与距离所述异源核酸分子的转录起始位点1000bp或更少的组成型启动子功能性连接。

11. 根据权利要求6所述的方法,其中所述一个或多个NEENA与距离所述异源核酸分子的转录起始位点500bp或更少的组成型启动子功能性连接。

12. 根据权利要求1至2任一项所述的方法,其中所述一个或多个NEENA与核酸分子的翻译起始位点上游的组成型启动子功能性连接,所述核酸分子的表达处在所述组成型启动子的控制下。

13. 根据权利要求1至2任一项所述的方法, 其中所述一个或多个NEENA与核酸分子的5' UTR内部的组成型启动子功能性连接, 所述核酸分子的表达处在所述组成型启动子的控制下。

14. 包含一个或多个NEENA的重组表达构建体, 所述的NEENA的核苷酸序列为SEQ ID NO:2。

15. 根据权利要求14所述的包含一个或多个SEQ ID NO:2所示的NEENA的重组表达构建体, 所述NEENA功能性连接到一个或多个组成型启动子和一个或多个表达的核酸分子, 所述一个或多个表达的核酸分子与所述一个或多个NEENA为异源的。

16. 重组表达载体, 其包含根据权利要求14或15所述的一个或多个重组表达构建体。

17. 转基因细菌或真菌细胞, 其包含权利要求16的重组表达载体或权利要求14或15的重组表达构建体。

18. 衍生自根据权利要求17所述的转基因细菌或真菌细胞的转基因细胞培养物, 其包含权利要求1所定义的所述异源NEENA、权利要求14或15的重组表达构建体或权利要求16的重组载体。

19. 权利要求1中所定义的NEENA或权利要求14至16中任一项所定义的重组构建体或重组载体的用途, 用于增强在植物或其部分中的表达。

20. 权利要求18中所述的转基因细胞培养物的用途, 用于食物、动物饲料、种子或精细化学品的生产。

21. 权利要求20的用途, 用于药物的生产。

用于增强植物中组成型基因表达的调节性核酸分子

[0001] 本申请是申请日为2010年8月11日、发明名称为“用于增强植物中组成型基因表达的调节性核酸分子”的中国专利申请201080038708.6 (国际申请号PCT/EP2010/061659) 的分案申请。

[0002] 发明描述

[0003] 本发明属于植物分子生物学领域并且提供了用于产生高表达的组成型启动子和产生具有增强的核酸组成型表达的植物的方法,其中将增强核酸表达的核酸 (NEENAs) 与所述启动子功能性连接和/或导入植物中。

[0004] 转基因在植物中的表达强烈地受多种外部和内部因素影响,从而导致变动和不可预测水平的转基因表达。经常不得不产生大量的转化体并进行分析,以鉴定具有合乎需要的表达强度的株系。由于转化并筛选表达强度合乎需要的株系是昂贵和耗费人力的,故需要一个或多个转基因在植物中高表达。当不得不在转基因植物中协调表达几个基因以实现特定效应时,鉴于必须鉴定其中每个基因均强烈表达的植物,这个问题尤其严重。

[0005] 例如,取决于构建体设计和各个转化事件中T-DNA插入基因座的位置效应,转基因的表达可以显著地变动。强启动子可以部分地克服这些难题。然而,显示强烈表达同时特异性合乎需要的合适启动子的可获得性经常是有限的。为了确保可获得具有合乎需要的表达特异性的充足启动子,额外启动子的鉴定和表征可能有助于弥合这种缺口。然而,具有相应特异性和强度的启动子的天然可获得性和费时的候选启动子表征阻碍了合适的新启动子的鉴定。

[0006] 为了克服这些难题,已经显示多样的遗传元件和/或基序积极地影响基因表达。在它们当中,已经认可一些内含子作为具有改善基因表达的强大潜能的遗传元件。虽然机制大多未知,但是已经显示一些内含子积极地影响成熟mRNA的稳态量,这可能通过增强转录活性、改善mRNA成熟、增强胞核mRNA输出和/或改善翻译起始来影响 (例如Huang和Gorman, 1990; Le Hir等人, 2003; Nott等人, 2004)。由于显示仅所选的内含子增加表达,故剪接本身很可能解释不了观察到的效果。

[0007] 将内含子与启动子功能性连接时观察到的基因表达增加称作内含子介导的基因表达增强 (IME) 并且已经在多种单子叶植物 (例如Callis等人, 1987; Vasil等人, 1989; Bruce等人, 1990; Lu等人, 2008) 和双子叶植物 (例如Chung等人, 2006; Kim等人, 2006; Rose等人, 2008) 得到显示。在这个方面,显示内含子相对于翻译起始位点 (ATG) 的位置对于内含子介导的基因表达增强至关重要 (Rose等人, 2004)。

[0008] 继一些内含子增强基因表达的潜能之后,显示这些内含子还在植物中于其原有核苷酸环境下影响组织特异性。发现报道基因表达依赖于含有多达2个内含子的基因组区域的存在 (Sieburth等人, 1997; Wang等人, 2004)。也已经报道5' UTR内含子对启动子元件的恰当功能性是重要的,这可能归因于内含子中存在的组织特异性基因控制元件 (Fu等人, 1995a; Fu等人, 1995b; Vitale等人, 2003; Kim等人, 2006)。然而,这些研究还显示内含子与异源启动子的组合可能对基因表达的强度和/或特异性产生强烈的不利影响 (Vitale等人, 2003; Kim等人, 2006; W02006/003186、W02007/098042)。例如,花椰菜花叶病毒CaMV35S组成

型强启动子因与芝麻SeFAD25' UTR内含子组合而不利地影响表达(Kim等人,2006)。与这些观察结果相反,一些文献显示通过IME增强核酸的表达而不影响相应启动子的组织特异性(Schünmann等人,2004)。

[0009] 在本申请中,描述了与组成型启动子功能性连接时增强所述启动子的表达同时不影响其特异性的其他核酸分子。在本申请中将这些核酸分子描述为“增强核酸表达的核酸”(NEENA)。内含子具有从分别的前mRNA剪接下来的内在特征。与其相反,本申请中即将陈述的核酸不必要一定包含于mRNA中,或如果存在于mRNA中,没有必要一定从mRNA中剪接下来,以增强源自与NEENA功能性连接的启动子的表达。

[0010] 发明详述

[0011] 本发明的第一实施方案包括用于产生高表达的组成型启动子的方法,所述高表达的组成型启动子包含与启动子功能性连接的一个或多个增强核酸表达的核酸(NEENA)分子,所述方法包括

[0012] i) 具有如SEQ ID NO:1至19、优选地SEQ ID NO:1至9的任一者中所定义序列的核酸分子,或

[0013] ii) 具有下述序列的核酸分子,所述序列与如SEQ ID NO:1至19、优选地SEQ ID NO:1至9所定义的任一序列具有80%或更大的同一性,优选地,该同一性是85%或更大,更优选地,该同一性是90%或更大,甚至更优选地,该同一性是95%或更大、96%或更大、97%或更大、98%或更大或99%或更大,在最优选的实施方案中,该同一性相对于如SEQ ID NO:1至19、优选地SEQ ID NO:1至9所定义的任一序列是100%,或

[0014] iii) i) 或 ii) 的核酸分子的100个或更多连续碱基、优选地150个或更多连续碱基、更优选地200个或更多连续碱基、或甚至更优选地250个或更多连续碱基的片段,所述片段具有如具备SEQ ID NO:1至19、优选地SEQ ID NO:1至9所定义任一序列的相应核酸分子那样的增强表达活性,例如65%或更大、优选地70%或更大、更优选地75%或更大,甚至更优选地80%或更大、85%或更大或90%或更大,在一个最优选的实施方案中,95%或更大的增强表达活性,或

[0015] iv) 作为前文在i) 至iii) 下提及的核酸分子中任一者的互补物或反向互补物的核酸分子,或

[0016] v) 使用由SEQ ID NO:20至57、优选地如表2中所示SEQ ID NO:20/21;26/27;30/31;38/39;42/43;44/45;46/47;50;51和56/57描述的寡核苷酸引物,通过PCR可获得的核酸分子,或

[0017] vi) 100个或更多核苷酸、150个或更多核苷酸、200个或更多核苷酸或250个或更多核苷酸的核酸分子,所述核酸分子在等同于在50℃于7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA杂交,以及在50℃或65℃、优选地65℃于2×SSC、0.1%SDS中洗涤的条件下与包含由SEQ ID NO:1至19、优选地SEQ ID NO:1至9描述的增强转录的核苷酸序列或其互补物的至少50个、优选地至少100、更优选地至少150、甚至更优选地至少200、最优选地至少250个连续核苷酸的核酸分子杂交。优选地,所述核酸分子在等同于在50℃于7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA杂交,以及在50℃或65℃、优选地65℃于1×SSC、0.1%SDS中洗涤的条件下与包含由SEQ ID NO:1至19、优选地SEQ ID NO:1至9描述的增强转录的核苷酸序列或其互补物的至少50个、优选地至少100、更优选地至少150、甚至更优选地至少

200、最优选地至少250个连续核苷酸的核酸分子杂交；更优选地，所述核酸分子在等同于在50℃于7%十二烷基硫酸钠(SDS)、0.5M NaPO₄、1mM EDTA杂交，以及在50℃或65℃、优选地65℃于0.1×SSC、0.1%SDS中洗涤的条件下与包含由SEQ ID NO:1至19、优选地SEQ ID NO:1至9描述的增强转录的核苷酸序列或其互补物的至少50个、优选地至少100、更优选地至少150、甚至更优选地至少200、最优选地至少250个连续核苷酸的核酸分子杂交。

[0018] 在一个实施方案中，一个或多个NEENA相对于与NEENA功能性连接的启动子是异源的。

[0019] 如以上在v)下所述，使用如表2中所示由SEQ ID NO:20至57、优选地SEQ ID NO:20/21;26/27;30/31;38/39;42/43;44/45;46/47;50/51和56/57所定义的寡核苷酸，通过PCR可获得的核酸分子是例如使用如下文实施例1中所述的条件，从来自拟南芥属(*Arabidopsis*)植物如拟南芥(*A.thaliana*)的基因组DNA中可获得的。

[0020] 技术人员知晓变动温度特征、循环数和/或缓冲液组成或浓度以获得相应的NEENA分子。表2中描述在用于获得相应NEENA分子的相应PCR反应中待使用的寡核苷酸的特定组合。

[0021] 本领域技术人员知晓用于使单向启动子变成双向启动子的方法和使用启动子序列的互补物或反向互补物以产生与原始序列具有相同启动子特异性的启动子的方法。用于组成型及诱导型启动子的此类方法例如由Xie等人(2001)“*Bidirectionalization of polar promoters in plants*(植物中极性启动子的双向化)”*nature biotechnology* 19第677—679页描述。作者描述了添加最小启动子至任意给定启动子的5'引发端足以在两个方向获得启动子控制表达，同时具有相同的启动子特异性。因而，与如上文所述NEENA功能性连接的高表达启动子在互补或反向互补情况下是有功能的，并且因此所述NEENA在互补或反向互补情况下也是有功能的。

[0022] 如本文中所述的组成型启动子意指在基本上全部植物组织中贯穿植物或其部分的基本上整个生活期限表达的启动子。在基本上全部植物组织中表达的启动子也可以包括在主要植物组织如叶、茎和/或根的至少两者中并且可以在或可以不在一些或全部次要组织或细胞如表皮、气孔、毛状体、花、种子或分生组织中表达的启动子。在一个优选的实施方案中，如本文中所意指的组成型启动子至少在绿色组织如叶和茎中表达。

[0023] 贯穿植物或其部分的基本上整个生活期限表达的启动子也可以包括在嫩的和已发育的组织中表达，但是可以在植物生活期限中的特定时间点或在特定条件下如在萌发和/或衰老期间或在生物性和/或非生物性胁迫条件如真菌性或细菌性感染、干旱、热或寒冷下缺少表达的启动子。在一个优选的实施方案中，贯穿植物的基本上整个生活期限表达的组成型启动子至少在充分扩大的组织中表达直至开始衰老。

[0024] 原则上，NEENA可以与任意的启动子如组织特异性、诱导型、发育特异性或组成型启动子功能性连接。相应的NEENA将导致在与至少一种NEENA功能性连接的相应启动子控制下的异源核酸的增强表达。增强除组成型启动子之外的启动子例如组织特异性启动子的表达将给予这些启动子的特异性。核酸在相应启动子控制下的表达将是在无NEENA的情况下检测不到所述核酸的转录物的额外组织或发育阶段中可检测的。因此，组织特异性或发育特异性或任意其他启动子可以通过将至少一种如上文所述的NEENA分子与所述启动子功能性连接而变成组成型启动子。因此，本发明的另一个实施方案是提供用于将植物中有功能

的任意给定启动子的特异性通过相应启动子与NEENA分子连接而赋予组成型启动子的方法,其中所述NEENA分子包含如以上在i)至vi)下所述的序列。

[0025] 优选地,一个或多个NEENA与任意的组成型启动子功能性连接并且将增强在所述启动子控制下的核酸分子的表达。待用于本发明任意方法中的组成型启动子可以源自植物例如单子叶或双子叶植物、源自细菌和/或病毒或可以是合成性启动子。待使用的组成型启动子例如是来自欧芹 (*P. crispum*) 的PcUbi-启动子(WO 2003102198)、来自玉米的ZmUbi-启动子、来自编码腈水解酶1的拟南芥基因At3g44310的AtNit-启动子、来自玄参 (*figwort*) 花叶病毒的34S-启动子、来自烟草花叶病毒的35S-启动子、源自农杆菌的nos启动子和ocs启动子、ScBV-启动子(US 5994123)、SUPER-启动子(Lee等人2007, *Plant. Phys.*)、来自编码铁氧还蛋白NADH还原酶的拟南芥基因At5g66190的AtFNR-启动子、来自豌豆 (*Pisum sativum*) 的ptxA启动子(WO2005085450)、来自编码磷酸三糖易位蛋白的拟南芥基因At5g46110的AtTPT-启动子、来自拟南芥基因At4g14880和At4g14890的双向AtOASTL-启动子、来自编码甘油醛-3-磷酸脱氢酶的拟南芥基因的At1g13440的PR00194启动子、来自编码二磷酸果糖醛缩酶的拟南芥基因At3g52930的PR00162启动子、AHAS-启动子(WO2008124495)或咖啡酰辅酶A-MT启动子和来自稻的OsCP12(WO2006084868)。

[0026] 与NEENA功能性连接的本发明高表达的组成型启动子可以用于任意植物中,所述植物包括例如苔藓、蕨类、裸子植物或被子植物,例如单子叶或双子叶植物。在一个优选的实施方案中,与NEENA功能性连接的本发明所述启动子可以用于单子叶或双子叶植物中,优选地是作物植物如谷物、大豆、卡诺拉油菜、棉属植物、马铃薯、甜菜、稻、小麦、高粱、大麦、芭蕉属植物、甘蔗、芒属植物等。在本发明的一个优选的实施方案中,与NEENA功能性连接的所述启动子可以用于单子叶作物植物如谷物、稻、小麦、高粱、芭蕉属植物、芒属植物、甘蔗或大麦中。在一个特别优选的实施方案中,与NEENA功能性连接的启动子可以用于双子叶作物植物如大豆、卡诺拉油菜、棉属植物、甜菜或马铃薯中。

[0027] 如本申请中所用的高表达的组成型启动子意指例如与NEENA功能性连接的启动子,所述NEENA引起该启动子在植物或其部分中增强的组成型表达,其中源自功能性连接于NEENA的相应启动子控制下的核酸分子的RNA积累或RNA合成速率比由缺少本发明NEENA的相同启动子引起的表达高,优选地显著更高。优选地,与相同条件下培育的包含不与本发明NEENA功能性连接的相同组成型启动子的同龄对照植物相比,植物中相应核酸的RNA的量和/或RNA合成速率和/或RNA稳定性增加50%或更大,例如100%或更大、优选地200%或更大、更优选地5倍或更大、甚至更优选地10倍或更大、最优选地20倍或更大,例如50倍。

[0028] 在本文中使用时,显著更高指技术人员知晓如何确定的统计显著性,例如通过对相应的数据集应用统计检验如t-检验。

[0029] 用于检测由启动子赋予的表达的方法是本领域已经的。例如,该启动子可以与标记基因如GUS、GFP或萤光素酶基因功能性连接,并且可以在植物或其部分中测定由相应标记基因编码的相应蛋白质的活性。作为代表性实例,下文详细描述用于检测萤光素酶的方法。其他方法例如是通过本领域已知的方法,例如RNA印迹分析法、qPCR、连缀(run-on)测定法或本领域所述的其他方法测量受该启动子控制的核酸分子的RNA稳态水平或合成速率。

[0030] 技术人员知晓多种用于功能性连接两个或多个核酸分子的方法。此类方法可以包括限制性切割/连接法、不依赖连接酶的克隆法、重组工程法、重组或合成法。其他方法可以

用来功能性连接两个或多个核酸分子。

[0031] 本发明的又一个实施方案是用于产生植物或其部分的方法,所述植物或其部分与相应的对照植物或其部分相比具有一种或多种核酸分子的增强的组成型表达,所述方法包括步骤:将包含如上文在i)至vi)下所定义核酸分子的一个或多个NEENA导入所述植物或其部分,和将所述一个或多个NEENA与启动子、优选地组成型启动子并且与处在所述启动子、优选地组成型启动子控制下的核酸分子功能性连接,其中NEENA对所述核酸分子为异源。

[0032] NEENA可以相对于处在与NEENA功能性连接的所述启动子的控制下的核酸分子为异源,或它可以相对于该启动子和处在该启动子控制下的核酸分子均为异源。

[0033] 就核酸分子或DNA而言,术语“异源的”指这样的核酸分子,所述核酸分子有效连接于,或受到操作以变得有效连接于自然界中不与该核酸分子有效连接或自然界中与该核酸分子在不同位置有效连接的第二种核酸分子。例如,本发明的NEENA在其天然环境中与其天然启动子功能性连接,而在本发明中,它与可以源自相同生物、不同生物或合成性启动子如SUPER-启动子的另一个启动子连接。它也可以意指本发明的NEENA与其天然启动子连接,但是处于所述启动子控制下的核酸分子相对于包含其天然NEENA的启动子为异源。此外,应当理解,启动子和/或处在与本发明NEENA功能性连接的所述启动子控制下的核酸分子相对于所述NEENA为异源,原因是它们的序列已经受到例如突变如插入、缺失等操作,从而启动子和/或处在所述启动子控制下的核酸分子的天然序列被修饰并且因此已经变得相对于本发明的NEENA为异源。也可以理解,当NEENA与其天然启动子功能性连接,其中NEENA的位置相对于所述启动子改变,从而该启动子在这种操作后显示更高表达时,该NEENA相对于功能性连接至NEENA的核酸为异源。

[0034] 如本文中所示意的显示增强的核酸分子组成型表达的植物意指与相同条件下培育的没有与相应核酸分子功能性连接的相应NEENA的对照植物相比,具有更高的、优选地统计显著更高的核酸分子组成型表达的植物。这种对照植物可以是野生型植物或是包含控制与本发明植物中相同的基因的相同启动子的转基因植物,其中所述启动子不与本发明的NEENA连接。

[0035] 产生如本文中所述的植物包括用于稳定转化的方法,如借助农杆菌介导的转化、原生质体转化、粒子轰击等将重组DNA构建体导入植物或其部分并且任选地随后再生出转基因植物。它还包括用于瞬时转化植物或其部分的方法,如病毒感染法或农杆菌浸润法。技术人员知晓用于稳定和/或瞬时转化植物或其部分的其他方法。方法如育种方法或原生质体融合法可以用于本发明植物的产生并且由本发明涵盖。

[0036] 本发明的方法可以应用于任意植物,例如裸子植物或被子植物,优选地是被子植物,例如双子叶或单子叶植物,优选地是双子叶植物。优选的单子叶植物例如是谷物、小麦、稻、大麦、高粱、芭蕉属植物、甘蔗、芒属植物和短柄草属植物(Brachypodium),特别优选的单子叶植物是谷物、小麦和稻。优选的双子叶植物是例如大豆、油菜籽、卡诺拉油菜、亚麻、棉属植物、马铃薯、甜菜、万寿菊和拟南芥属植物(Arabidopsis),特别优选的双子叶植物是大豆、油菜籽、卡诺拉油菜和马铃薯。

[0037] 在本发明的一个实施方案中,如上文定义的方法包括以下步骤:

[0038] a) 将包含如上文的i)至vi)中所定义核酸分子的一个或多个NEENA导入植物或其部分,和

[0039] b) 将所述一个或多个NEENA整合至所述植物或其部分的基因组中,从而所述一个或多个NEENA与相对所述一个或多个NEENA为异源的优选地组成型表达的内源核酸功能性连接,和任选地

[0040] c) 从所述转化的细胞再生出包含所述一个或多个NEENA的植物或其部分。

[0041] NEENA可以相对于处在与NEENA功能性连接的所述启动子的控制下的核酸分子为异源,或它可以相对于该启动子和处在该启动子控制下的核酸分子均为异源。

[0042] 可以将一个或多个NEENA分子借助粒子轰击法、原生质体电穿孔法、病毒感染法、农杆菌介导的转化法或本领域已知的任意其他方法导入植物或其部分。可以将NEENA分子导入整合例如至质粒或病毒DNA或病毒RNA中。NEENA分子也可以在导入植物或植物部分之前包含于BAC、YAC或人工染色体上。也可以将NEENA分子作为包含NEENA序列的线性核酸分子导入,其中额外的序列可以紧邻该核酸分子上的NEENA序列存在。毗邻NEENA序列的这些序列可以长约20bp,例如20bp至数百碱基对,例如100bp或更多,并且可以促进整合至基因组中,例如通过同源重组。可以使用用于基因组整合的任何其他方法,无论它是定向整合法,如同源重组,或随机整合法,如非常规(illegitimate)重组。

[0043] 可以与NEENA分子功能性连接的优选地组成型表达的内源核酸可以是任意核酸,优选地是任意的组成型表达的核酸分子。该核酸分子可以是编码蛋白质的核酸分子或非编码性分子如反义RNA、rRNA、tRNA、miRNA、ta-siRNA、siRNA、dsRNA、snRNA、snoRNA或本领域已知的任何其他非编码性RNA。

[0044] 技术人员知晓用于鉴定组成型表达的核酸分子的方法,例如通过微阵列芯片杂交、qPCR、RNA印迹分析、下一代测序法等,其中本发明的方法可以优选地应用于所述核酸分子。

[0045] 实施本发明方法的另一种方式可以是

[0046] a) 提供包含一个或多个NEENA的表达构建体,所述NEENA包含如上文i)至vi)中所定义的与启动子、优选地组成型启动子并且与相对于所述一个或多个NEENA为异源并处在所述启动子、优选地组成型启动子控制下的一种或多种核酸分子功能性连接的核酸分子,和

[0047] b) 将包含所述一个或多个NEENA的所述表达构建体整合至所述植物或其部分的基因组中,和任选地

[0048] c) 从所述转化的植物或其部分再生出包含所述一个或多个表达构建体的植物或其部分。

[0049] NEENA可以相对于处在与NEENA功能性连接的所述启动子的控制下的核酸分子为异源,或它可以相对于该启动子和处在该启动子控制下的核酸分子均为异源。

[0050] 表达构建体可以借助本领域已知的任何方法整合至相应植物的基因组中。使用如粒子轰击法或农杆菌介导的转化法的方法时,整合可以是随机的。在一个优选的实施方案中,整合借助定向整合,例如通过同源重组进行。后一种方法将允许包含与NEENA功能性连接的高表达启动子的表达构建体整合至有利的基因组区域中。有利的基因组区域例如是已知包含例如在种子中高度表达的基因的基因组区域,并且因而与显示无转录活性的基因组区域相比,可以增加源自所述表达构建体的表达。

[0051] 在另一个优选的实施方案中,所述一个或多个NEENA与靠近所述异源核酸分子的

转录起始位点的启动子、优选地组成型启动子功能性连接。

[0052] 如本文中所意指的靠近转录起始位点包含一个或多个NEENA与距离所述异源核酸分子的转录起始位点2500bp或更少、优选地2000bp或更少、更优选地1500bp或更少、甚至更优选地1000bp或更少并且最优选地500bp或更少的启动子、优选地组成型启动子功能性连接。应当理解,该NEENA可以距相应启动子的转录起始位点以相应距离在上游或下游整合。因而,一个或多个NEENA不必一定包含于处在与所述一个或多个NEENA功能性连接的优选地组成型启动子控制下的相应异源核酸的转录物中。优选地,一个或多个NEENA整合在相应启动子、优选地组成型启动子的转录起始位点的下游。转录起始位点下游的整合位点可以位于5' UTR、3' UTR、外显子或内含子中,或它可以替换内含子或部分或完全地替换处于优选地组成型启动子控制下的异源核酸的5' UTR或3' UTR。优选地,一个或多个NEENA整合在5' UTR或内含子中,或所述NEENA替换内含子或部分或全部5' UTR,最优选地,它整合在相应异源核酸的5' UTR中。

[0053] 本发明的又一个实施方案包括含有一个或多个NEENA的重组表达构建体,其中所述的NEENA包含上文i)至vi)中所定义的核酸分子。

[0054] 重组表达构建体可以还包含与一个或多个NEENA功能性连接的一个或多个启动子,优选地组成型启动子和任选地一个或多个表达的核酸分子,所述核酸分子相对于所述一个或多个NEENA为异源。

[0055] NEENA可以相对于处在与NEENA功能性连接的所述启动子的控制下的核酸分子为异源,或它可以相对于该启动子和处在该启动子控制下的核酸分子均为异源。

[0056] 该表达构建体可以包含与NEENA功能性连接的启动子、优选地组成型启动子和待表达核酸分子的1个或多个,例如2个或更多个,例如5个或更多个,如10个或更多个组合,所述的待表达核酸分子相对于相应的NEENA为异源。该表达构建体也可以包含其他启动子,所述的其他启动子不包含与相对于相应的启动子为同源或异源的待表达核酸分子功能性连接的NEENA。

[0057] 包含如上文定义的一个或多个重组表达构建体的重组表达载体是本发明的另一个实施方案。可以在本发明中使用的多种表达载体是技术人员已知的。用于将包含这种表达构建体的此种载体导入植物基因组中的方法和用于从转化的细胞恢复转基因植物的方法也是本领域熟知的,其中所述的表达构建体包含例如与NEENA功能性连接的启动子和任选地其他元件如终止子。取决于用于转化植物或其部分的方法,完整载体可以整合至所述植物或其部分的基因组中,或者载体的某些组分可以整合至所述基因组中,例如T-DNA。

[0058] 本发明中还包括转基因植物或其部分,其包含如上文i)至vi)中所定义的一种或多种异源NEENA。如果NEENA是合成的、源自另一种生物或源自相同生物,但是其天然基因组位置与对照植物例如野生型植物相比被改变,则将NEENA理解为相对于该植物是异源的。应当理解,改变的基因组位置意指,NEENA位于另一条染色体上或位于同一条染色体上,但是脱离例如在野生型植物中其天然基因组位置10kb或更远,例如10kb,优选地5kb或更远,例如5kb,更优选地1000bp或更远,例如1000bp,甚至更优选地500bp或更远,例如500bp,特别优选地100bp或更远,例如100bp,最优选地10bp或更远,例如10bp。

[0059] 包含如上文定义的重组表达载体或如上文定义的重组表达构建体的转基因细胞或转基因植物或其部分是本发明的又一个实施方案。转基因细胞、转基因植物或其部分可

以选自细菌、真菌、酵母或植物、昆虫细胞或哺乳动物细胞或植物。优选地，转基因细胞是细菌、真菌、酵母或植物细胞。优选的细菌是肠杆菌属细菌如大肠杆菌 (*E.coli*) 和农杆菌属的细菌，例如根癌农杆菌 (*Agrobacterium tumefaciens*) 和发根农杆菌 (*Agrobacterium rhizogenes*)。优选的植物是单子叶或双子叶植物，例如单子叶或双子叶作物植物，如谷物、大豆、卡诺拉油菜、棉属植物、马铃薯、甜菜、稻、小麦、高粱、大麦、芒属植物、芭蕉属植物、甘蔗等。优选的作物植物是谷物、稻、小麦、大豆、卡诺拉油菜、棉属植物或马铃薯。尤其优选的双子叶作物植物是大豆、卡诺拉油菜、棉属植物或马铃薯。尤其优选的单子叶作物植物是谷物、小麦和稻。

[0060] 源自如上文所定义的转基因细胞或植物或其部分中的转基因细胞培养物、转基因种子、部分或繁殖材料是本发明的其他实施方案，其中所述的转基因细胞或植物或其部分包含如上文在 i) 至 vi) 中定义的所述异源 NEENA 或如上文定义的所述重组表达构建体或所述重组载体。

[0061] 如本文中所意指的转基因部分或繁殖材料包含含有相应 NEENA、重组表达构建体或重组载体的全部组织和器官，例如叶、茎和果实以及用于植物繁殖和/或再生的材料如插条、接穗、压条、枝条或幼苗。

[0062] 本发明的又一个实施方案是如上文在 i) 至 vi) 中定义的 NEENA 或如上文定义的重组构建体或重组载体用于增强植物或其部分中表达的用途。

[0063] 因而，本申请即将提供种子特异的和/或种子优先的增强基因表达的核酸分子，其包含一个或多个启动子，优选地是与一个或多个 NEENA 功能性连接的种子特异的和/或种子优先的启动子。另外，提供了此类增强基因表达的核酸分子和包含此类增强基因表达的核酸分子的表达构建体、表达载体、转基因植物或其部分和转基因细胞的用途。

[0064] 本发明中还包括源自如上文定义的转基因细胞或植物或其部分中的转基因细胞培养物、转基因种子、部分或繁殖材料用于生产食物、动物饲料、种子、药物或精细化学品的用途。

[0065] 定义

[0066] 缩写: NEENA-增强核酸表达的核酸, GFP-绿色荧光蛋白, GUS- β -葡萄糖苷酸酶, BAP-6-苄氨基嘌呤, 2,4-D-2,4-二氯苯氧乙酸, MS-Murashige-Skoog 培养基, NAA-1-萘乙酸, MES, 2-(N-吗啉代)-乙磺酸, IAA: 吲哚乙酸, Kan: 硫酸卡那霉素, GA3-赤霉素, TimentinTM: 替卡西林二钠/克拉维酸钾, microl: 微升。

[0067] 应当理解，本发明不限于具体的方法或方案。还应当理解本文所用的术语目的仅在于描述具体实施方案，并且不意图限制本发明，本发明将仅受所附权利要求书限制。必须指出，如本文中和所附权利要求中所用，单数形式“一个”、“一种”和“该”包括复数指称，除非上下文另外明确地指明并非如此。因此，例如，对“一个载体”的提及是对一个或多个载体的提及并且包括本领域技术人员已知的其等同物等。术语“约”在本文中用来指大约、大致、左右和在……范围内。当术语“约”与一个数字范围联合使用时，它通过使界限值扩展高于和低于所述数值而修饰该范围。通常而言，术语“约”在本文中用来通过 20%、优选地 10% 之上或之下 (更高或更低) 变异而修饰高于和低于所述值的数值。如本文中所用，词汇“或”意指特定列出的任何一成员并且还包括该列出的成员的任意组合。在本说明书中及以下权利要求中使用，词汇“包含”、“包含着”、“包括”、“包括着”和“包括了”意图指明一个或多个

所述特征、整数、组分或步骤的存在,但是它们不排除一个或多个其他特征、整数、组分、步骤或其组的存在或添加。为清晰起见,将本说明书中使用的某些术语如下定义并使用。

[0068] 反平行:“反平行”在本文中指经互补性碱基残基之间氢键配对的两个核苷酸序列,其中磷酸二酯键在一个核苷酸序列中以5′-3′方向分布并且在另一个核苷酸序列中以3′-5′方向分布。

[0069] 反义:术语“反义”指相对于其转录或发挥作用的正常方向为反向并且从而表达下述RNA转录物的核苷酸序列,其中所述的RNA转录物与宿主细胞内部表达的靶基因mRNA分子互补(例如,它可以借助Watson-Crick碱基配对作用与靶基因mRNA分子或单链基因组DNA杂交)或与靶DNA分子(例如宿主细胞中存在的基因组DNA)互补。

[0070] 编码区:如本文中所用,术语“编码区”在谈及结构基因使用时,指编码在作为mRNA分子翻译结果的新生多肽中存在的氨基酸的核苷酸序列。在真核生物中,编码区在5′侧以编码起始物甲硫氨酸的核苷酸三联体“ATG”为界并且在3′侧以特定终止密码子的3种三联体(即TAA、TAG、TGA)为界。除含有内含子之外,基因的基因组形式也可以包含位于RNA转录物中存在的序列的5′-及3′-端的序列。这些序列称作“侧翼”序列或区(这些侧翼序列位于mRNA转录物上存在的非翻译序列的5′或3′)。5′-侧翼区可以含有控制或影响基因转录的调节序列如启动子和增强子。3′-侧翼区可以含有指导转录终止、转录后剪切和聚腺苷酸化的序列。

[0071] 互补的:“互补的”或“互补性”指包含反平行核苷酸序列的两个核苷酸序列,其中一旦在反平行核苷酸序列中的互补性碱基残基之间形成氢键,则所述反平行核苷酸序列能够相互配对(通过碱基配对原则)。例如,序列5′-AGT-3′与序列5′-ACT-3′互补。互补性可以是“部分的”或“全部的”。“部分”互补性是其中一个或多个核酸碱基根据碱基配对规则未匹配的情况。核酸分子之间的“全部”或“完全”互补性是其中每个核酸碱基按照碱基配对规则与另一个碱基匹配的情况。核酸分子链之间互补性的程度对核酸分子链之间杂交的效率和强度具有明显影响。如本文中所用的核酸序列“互补物”指这样的核苷酸序列,其核酸分子显示与该核酸序列的核酸分子的全部互补性。

[0072] 双链RNA:“双链RNA”分子或“dsRNA”分子包含核苷酸序列的有义RNA片段和该核苷酸序列的反义RNA片段,二者均包含彼此互补的核苷酸序列,因而允许有义RNA片段和反义RNA片段配对并形成双链RNA分子。

[0073] 内源的:“内源的”核苷酸序列指存在于未转化植物细胞的基因组中的核苷酸序列。

[0074] 增强的表达:“增强”或“增加”核酸分子在植物细胞中的表达在本文中同等地使用,并且意指该核酸分子在应用本发明方法之后在植物、植物部分或植物细胞中的表达水平比其在应用该方法之前在植物、植物部分或植物细胞中的表达更高,或与缺少本发明重组核酸分子的参照植物相比更高。例如,参照植物包含仅缺少相应NEENA的相同构建体。如本文中所用的术语“增强”或“增加”是同义的,并且在本文中意指待表达的核酸分子的更高、优选地显著更高的表达。如本文中所用的,“增强”或“增加”某物质(如蛋白质、mRNA或RNA)的水平意指相对于基本上相同条件下培育的、缺少本发明重组核酸分子(例如缺少本发明的NEENA分子、重组构建体或重组载体)的基本上相同植物、植物部分或植物细胞,该水平增加。如本文中所用的,“增强”或“增加”某物质(例如由靶基因表达的前RNA、mRNA、rRNA、

tRNA、snoRNA、snRNA和/或由其编码的蛋白质产物)的水平意指相对于缺少本发明重组核酸分子的细胞或生物,该水平增加50%或更多,例如100%或更多,优选地200%或更多,更优选地5倍或更多倍,甚至更优选地10倍或更多倍,最优选地20倍或更多倍,例如50倍。可以通过技术人员熟悉的方法测定所述增强或增加。因而,可以例如通过蛋白质的免疫学检测法确定核酸或蛋白质量的增强或增加。另外,可以使用技术如蛋白质测定法、荧光法、RNA杂交法、核酸酶保护测定法、逆转录法(定量RT-PCR)、ELISA(酶联免疫吸附测定法)、蛋白质印迹法、放射免疫测定法(RIA)或其他免疫测定法和荧光激活的细胞分析(FACS)来测量植物或植物细胞中的特定蛋白质或RNA。取决于所诱导蛋白质产物的类型,也可以确定其活性或对生物或细胞表型的影响。用于确定蛋白质数量的方法是技术人员已知的。可以提到的实例是:微量Biuret法(Goa J(1953) Scand J Clin Lab Invest 5:218-222)、Folin-Ciocalteu法(Lowry OH等人(1951) J Biol Chem 193:265-275)或测量CBB G-250的吸光度(Bradford MM(1976) Analyt Biochem 72:248-254)。作为用于量化蛋白质的活性的一个实例,在下文实施例中描述荧光素酶活性检测法。

[0075] 表达:“表达”指基因产物的生物合成,优选地指细胞中核苷酸序列例如内源基因或异源基因的转录和/或翻译。例如,在结构基因的情况下,表达涉及结构基因转录成mRNA并且任选地随后mRNA翻译成一种或多种多肽。在其他情况下,表达可以仅指携带RNA分子的DNA的转录。

[0076] 表达构建体:如本文中所用的“表达构建体”意指能够指导特定核苷酸序列在适宜植物部分或植物细胞中表达的DNA序列,该DNA序列包含在导入此DNA序列的所述植物部分或植物细胞中有功能的启动子,所述启动子与任选地有效连接至终止信号的目的核苷酸序列有效连接。如果需要翻译,该DNA序列一般还包含为正确翻译所述核苷酸序列所需的序列。编码区可以编码目的蛋白,但也可以以有义或反义方向编码功能性目的RNA,例如RNAa、siRNA、snoRNA、snRNA、microRNA、ta-siRNA或任何其他非编码的调节性RNA。包含目的核苷酸序列的表达构建体可以是嵌合的,这意指该表达构建体的组件中一者或多者相对于该表达构建体的其他组件中一者或多者是异源的。该表达构建体也可以是一种这样的表达构建体,它天然地存在,但已经以用于异源表达的重组形式获得。然而,一般而言,该表达构建体相对于宿主是异源,即,该表达构建体的特定DNA序列不天然存在于宿主细胞中并且必须已经通过转化事件导入宿主细胞或该宿主细胞的祖先中。该表达构建体中的核苷酸序列的表达可以处于组成型启动子或处于仅在宿主细胞暴露于一些特定外部刺激时才启动转录的诱导型启动子的控制下。在植物的情况下,该启动子也可以是针对特定组织或器官或发育阶段特异的。

[0077] 外来:术语“外来”指任何核酸分子(例如,基因序列),所述核酸分子通过实验操作导入细胞的基因组中并且可以包括该细胞中存在的序列,只要导入的序列含有一些修饰(例如,点突变、存在可选择标记基因等)和因此相对于天然存在序列是不同的。

[0078] 功能性连接:将术语“功能性连接”或“功能性连接的”理解为意指例如调节性元件(例如启动子)与待表达的核酸序列并且根据需要进行与其他调节性元件(例如,终止子或NEENA)以如此方式依次排列,从而每种调节性元件可以履行其目的功能以允许、修饰、促进或影响所述核酸序列的表达。作为同义词,可以使用“有效连接”或“有效连接的”。可以根据所述核酸序列相对于有义或反义RNA的排列,产生表达。为此目的,不是必需要求化学意义

上的直接连接。遗传控制序列例如增强子序列也能够从远离的位置或甚至从其它DNA分子对靶序列产生其作用。优选的排列是这样的排列,其中待表达的核酸序列重组地位于充当启动子的序列之后,从而这两个序列相互共价地连接。该启动子序列与待重组表达的核酸序列之间的距离优选地小于200碱基对、特别优选地小于100碱基对、非常特别优选地小于50碱基对。在一个优选的实施方案中,待转录的核酸序列以如此方式位于启动子之后,其中转录起点与本发明嵌合RNA的合乎需要的开端相同。功能性连接和表达构建体可以借助如(例如,在Maniatis T,Fritsch EF和Sambrook J(1989)Molecular Cloning:A Laboratory Manual,第二版,Cold Spring Harbor Laboratory,Cold Spring Harbor (NY);Silhavy等人(1984)Experiments with Gene Fusions,Cold Spring Harbor Laboratory,Cold Spring Harbor (NY);Ausubel等人(1987)Current Protocols in Molecular Biology, Greene Publishing Assoc.and Wiley Interscience;Gelvin等人(编著)(1990)Plant Molecular Biology Manual;Kluwer Academic Publisher,Dordrecht,The Netherlands)所述的惯用重组和克隆技术产生。然而,其他序列,例如充当带限制性酶特定切割位点的接头或充当信号肽的序列,也可以位于这两个序列之间。序列的插入也可以导致融合蛋白的表达。优选地,由调节性区域例如启动子和待表达核酸序列的连接组成的表达构建体可以以载体整合的形式存在并且被插入植物基因组,例如通过转化法。

[0079] 基因:术语“基因”指与能够以某种方式调节基因产物(例如,多肽或功能性RNA)表达的适宜调节序列有效连接的区域。基因包括DNA中位于编码区(可读框,ORF)之前(上游)和之后(下游)的不翻译的调节区(例如启动子、增强子、阻抑物等),以及在可用的情况下,在各个编码区(例如外显子)之间的间插序列(例如内含子)。如本文中所用的术语“结构基因”意图指转录成mRNA的DNA序列,其中所述的mRNA随后翻译成作为具体多肽的特征的氨基酸序列。

[0080] 基因组和基因组DNA:术语“基因组”或“基因组DNA”指宿主生物的可遗传信息。所述基因组DNA包括胞核的DNA(也称作染色体DNA),还包括质体(例如,叶绿体)和其他细胞器(例如,线粒体)的DNA。优选地,术语“基因组”或“基因组DNA”指胞核的染色体DNA。

[0081] 异源的:就核酸分子或DNA而言,术语“异源的”指这样的核酸分子,所述核酸分子有效连接于,或受到操作以变得有效连接于自然界中不与该核酸分子有效连接或自然界中与该核酸分子在不同位置有效连接的第二种核酸分子。包含核酸分子和与之连接的一个或多个调节性核酸分子(如启动子或转录终止信号)的异源表达构建体例如是源自实验操作的构建体,在所述构建体中,a)所述核酸分子或b)所述调节性核酸分子或c)二者(即(a)和(b))不位于其天然(原有)遗传环境中或已经因实验操作受到修饰,修饰的实例是一个或多个核苷酸残基的置换、添加、缺失、倒位或插入。天然遗传环境指源生物中的天然染色体基因座或指存在于基因组文库中。在基因组文库的情况下,优选地保留,至少部分地保留该核酸分子的序列的天然遗传环境。该环境分布在所述核酸序列的至少一侧并且具有至少50bp,优选地至少500bp,特别优选地至少1,000bp,非常特别优选地至少5,000bp序列长度。天然存在的表达构建体例如启动子与相应基因的天然存在组合-在通过非天然的合成性“人工”方法例如诱变法修饰时变成转基因表达构建体。已经描述了此类方法(US 5,565,350;WO 00/15815)。例如,与启动子有效连接的编码蛋白质的核酸分子相对于该启动子视为异源,其中所述的启动子不是该核酸分子的天然启动子。优选地,异源DNA相对于导入该

异源DNA的细胞不是内源的或不天然与该细胞相关,但是已经从另一种细胞获得或已经合成。异源DNA还包括含有一些修饰的内源DNA序列、不天然存在的多拷贝内源DNA序列或这样的DNA序列,该DNA序列不与同物理连接于所述DNA序列的另一个DNA序列天然接合。通常地但不是必需地,异源DNA编码在细胞中表达的RNA或蛋白质,而所述RNA或蛋白质正常情况下不被所述细胞产生。

[0082] 高表达组成型启动子:如本文中所用的“高表达组成型启动子”意指在植物或其部分中引起组成型表达的启动子,其中衍生自处于相应启动子控制下的核酸分子中的RNA积累或合成速率或RNA稳定性比缺少本发明NEENA的启动子所引起的表达更高,优选地显著更高。优选地,相对于缺少本发明NEENA的组成型启动子,RNA的量和/或RNA合成速率和/或RNA稳定性增加50%或更大,例如100%或更多,优选地200%或更多,更优选地5倍或更多倍,甚至更优选地10倍或更多倍,最优选地20倍或更多倍,例如50倍。

[0083] 杂交:如本文中所用,术语“杂交”包括“核酸分子的链通过碱基配对作用与互补链结合的任意过程”。(J.Coombs (1994) Dictionary of Biotechnology, Stockton Press, New York)。杂交和杂交的强度(即,核酸分子之间结合的强度)受此类因素影响,如核酸分子之间的互补性程度、所涉及条件的严格性、所形成杂合体的 T_m 和核酸分子内部的G:C比。如本文中所用,术语“ T_m ”用来指“解链温度”。解链温度是双链核酸分子群体一半解离成单链的温度。用于计算核酸分子的 T_m 的等式是本领域熟知的。如标准参考文献所示,当核酸分子处于1M NaCl的水溶液中时,可以通过等式: $T_m = 81.5 + 0.41 (\%G+C)$ 计算来进行 T_m 值的简单估计[见例如,Anderson和Young, Quantitative Filter Hybridization, in Nucleic Acid Hybridization (1985)]。其他参考文献包括更复杂算法,这些算法考虑了结构特征及序列特征以计算 T_m 。严格条件是本领域技术人员已知的并且可以在Current Protocols in Molecular Biology, John Wiley&Sons, N.Y. (1989), 6.3.1-6.3.6中找到。

[0084] “同一性”:在比较两个或多个核酸或氨基酸分子使用时,“同一性”意指所述分子的序列共有某种程度的序列相似性,即所述序列是部分相同的。

[0085] 为确定两个氨基酸序列或两个核酸分子的同一性(同源性在本文可互换地使用)百分数,将所述序列以一个序列在另一个序列下的方式书写以最佳比较(例如,可以将空位插入蛋白质的序列或核酸的序列,旨在与另一种蛋白质或另一种核酸产生最佳比对)。

[0086] 随后在对应氨基酸位置或核苷酸位置处比较氨基酸残基或核酸分子。当一个序列中的位置由另一个序列中对应位置处的相同氨基酸残基或相同核酸分子占据时,则所述分子在这个位置处是同源的(即,如本上下文中所用的氨基酸或核酸“同源性”对应于氨基酸或核酸“同一性”。这两个序列之间的同一性百分数是所述序列共有的相同位置的数值的函数(即同源性百分数=相同位置的数值/位置总数 $\times 100$)。术语“同源性”和“同一性”因而视为同义词。

[0087] 为确定两个或多个氨基酸序列或两个或多个核苷酸序列的同一性百分数,已经开发了几个计算机软件程序。两个或更多序列的同一性可以用例如软件fasta计算,其中软件fasta已经以fasta 3版本使用(W.R.Pearson和D.J.Lipman, PNAS 85, 2444 (1988); W.R.Pearson, Methods in Enzymology 183, 63 (1990); W.R.Pearson和D.J.Lipman, PNAS 85, 2444 (1988); W.R.Pearson, Enzymology 183, 63 (1990))。用于计算不同序列的同一性的另一种有用程序是标准blast程序,其中该程序包含于Biomax pedant软件(联邦德国慕尼

黑Biomax)中。不幸地,该程序有时产生次优结果,因为blast不总是包含主题和查询对象的完整序列。然而,由于该程序非常高效,因而它可以用于庞大数目序列的比较。以下设置通常用这样的序列比较:

[0088] -p程序名称[字符串];-d数据库[字符串];默认=nr;-i查询文件[FileIn];默认=stdin;-e期望值(E)[Real];默认=10.0;-m比对浏览选项:0=配对;1=查询-比上区域(query-anchored),显示同一性;2=查询-比上区域,不显示同一性;3=查询-比上区域的屏文形式,显示同一性;4=查询-比上区域的屏文形式,不显示同一性;5=查询-比上区域,不显示同一性和突然结束;6=查询-比上区域的屏文形式,不显示同一性和突然结束;7=XML Blast输出;8=TAB格式输出;9带注释行的TAB格式输出[整数];默认=0;-o BLAST报告输出文件[File Out]任选;默认=stdout;-F过滤软件查询序列(DUST用blastn,SEG用其他方法)[字符串];默认=T;-G空位开口成本(零激发默认行为)[整数];默认=0;-E开口延伸成本(零激发默认行为)[整数];默认=0;-X X空位比对的下降值(比特)(零激发默认行为);blastn 30,megablast 20,tblastx 0,全部其他方法15[整数];默认=0;-I在定义行中显示GI'[T/F];默认=F;-q核苷酸错配罚分(仅blastn)[整数];默认=-3;-r核苷酸匹配回报(仅blastn)[整数];默认=1;-v显示对(V)的单-线描述的数据库序列数[整数];默认=500;-b显示对(B)的比对结果的数据库序列数[整数];默认=250;-f延伸命中的阈值,若为零,则默认;blastp 11,blastn 0,blastx 12,tblastn 13;tblastx 13,megablast 0[整数];默认=0;-g执行空位比对(对tblastx不可用)[T/F];默认=T;-Q查询使用的遗传密码[整数];默认=1;-D DB遗传密码子(仅用tblast[nx])[整数];默认=1;-a使用的处理器数目[整数];默认=1;-O SeqAlign文件[File Out]任选;-J相信查询定义行[T/F];默认=F;-M矩阵[字符串];默认=BLOSUM62;-W字大小,如果为零,则默认(blastn 11,megablast 28,全部其他方法3)[整数];默认=0;-z数据库的有效长度(对于真实大小,使用零)[Real];默认=0;-K来自保持区域的最佳命中数(默认关闭;若使用,则推荐值为100)[整数];默认=0;-P 0用于多重命中;1用于单一命中[整数];默认=0;-Y搜索空间的有效长度(对真实大小,使用零)[Real];默认=0;-S针对数据库搜索的查询链(对于blast[nx]和tblastx);3用于两者,1是顶端,2是底部[整数];默认=3;-T产生HTML输出结果[T/F];默认=F;-l限制数据库搜索至GI's列表[字符串]任选;-U使用FASTA序列的较低事件过滤作用[T/F]任选;默认=F;-y无空位延伸的X下降值(比特)(0.0激发默认行为);blastn 20,megablast 10,全部其他方法7[Real];默认=0.0;-Z最终空位比对的X下降值(比特)(0.0激发默认行为);blastn/megablast50,tblastx 0,全部其他方法25[整数];默认=0;-R PSI-TBLASTN检查点文件[File In]任选;-n MegaBlast搜索[T/F];默认=F;-L查询序列上的位置[字符串]任选;-A多重命中窗口大小,如果为零,则默认(blastn/megablast 0,全部其他方法40[整数];默认=0;-w移码罚分(00F算法用于blastx)[整数];默认=0;-t允许tblastn中用于联系HSP的最大内含子的长度(0取消连接联系)[整数];默认=0。

[0089] 通过使用Needleman和Wunsch或者Smith和Waterman算法获得高质量的结果。因而,优选基于所述算法的程序。有利地,可以用程序PileUp(J.Mol.Evolution.,25,351(1987),Higgins等人,CABIOS 5,151(1989))或优选地用均基于Needleman和Wunsch算法(J.Mol.Biol.48;443(1970))的程序“Gap”与“Needle”和基于Smith和Waterman(Adv.Appl.Math.2;482(1981))算法的“BestFit”进行序列的比较。“Gap”和“BestFit”是

GCG软件包[Genetics Computer Group,575Science Drive,Madison,Wisconsin,USA 53711 (1991);Altschul等人(Nucleic Acids Res.25,3389 (1997)),"Needle"是欧洲分子生物学开放软件包(The European Molecular Biology Open Software Suite (EMBOSS))的部分(Trends in Genetics 16 (6),276 (2000))。因而,优选地,用程序"Gap"或"Needle"在所述序列的整个范围内进行确定序列同一性百分数的计算。用于比较核酸序列的以下标准调整用于"Needle":矩阵:EDNAFULL,空位罚分:10.0,延伸罚分:0.5。用于比较核酸序列的以下标准调整用于"Gap":空位权重:50,长度权重:3,平均匹配:10.000,平均错配:0.000。

[0090] 例如,将声称在核酸水平与序列SEQ ID NO:1具有80%同一性的序列理解为意指,通过上述程序"Needle"以设置的上述参数与序列SEQ ID NO:1所代表的序列比较时具有80%同一性的序列。优选地,在查询序列例如SEQ ID NO:1的完整长度上计算同一性。

[0091] 内含子:指基因内部的DNA区段(间插序列),该DNA区段不编码该基因产生的蛋白质的部分,并且从该基因转录的mRNA中在该mRNA从细胞核输出之前剪切下来。内含子序列指内含子的核酸序列。因而,内含子是DNA序列的这些区域,它们随编码序列(外显子)一起转录但是在成熟mRNA形成期间被除去。内含子可以位于实际编码区内部或位于前mRNA(未剪接的mRNA)的5'或3'非翻译前导序列中。初级转录物中的内含子被切下并且编码序列同时且精确地连接以形成成熟的mRNA。内含子和外显子的交界形成剪接位点。内含子的序列始于GU并止于AG。另外,在植物中,已经描述AU-AC内含子的两个实例:来自拟南芥(*Arabidopsis thaliana*)的RecA样蛋白基因第14内含子和G5基因第7内含子是AT-AC含子。含有内含子的前mRNA具有3种短序列,连同其他序列,这些短序列对于精确剪接内含子是必需的。这些序列是5'剪接位点、3'剪接位点和分支点。mRNA剪接是除去初级mRNA转录物中存在的间插序列(内含子)并接合或连接外显子序列。这又称作顺式剪接作用,所述顺式剪接作用将相同RNA上的两个外显子接合,同时除去间插序列(内含子)。内含子的功能性元件包含由剪接体的特定蛋白质组分(例如其剪接内含子末端处的共有序列)识别并结合的序列。功能性元件与剪接体的相互作用导致从不成熟mRNA除去内含子序列和外显子序列的再接合。内含子具有3种短序列,这些短序列对于精确剪接内含子是必需的,但是并非足够的。这些序列是5'剪接位点、3'剪接位点和分支点。分支点序列是在植物中剪接过程和剪接位点选择方面重要的。分支点序列通常位于3'剪接位点上游10-60个核苷酸处。

[0092] 同基因的:除可以因存在或不存在异源DNA序列而不同之外,在遗传上相同的生物(例如植物)。

[0093] 分离的:如本文中所用的术语"分离"意指材料已经通过人工取出并且离开其原来的天然环境存在并且因此不是自然界的产物。分离的材料或分子(如DNA分子或酶)可以以纯化的形式存在或可以存在于非天然环境中如例如存在于转基因宿主细胞中。例如,活植物中存在的天然存在多核苷酸或多肽不是分离的,然而与该天然系统中一些或全部共存物质分开的相同多核苷酸或多肽是分离的。此类多核苷酸可以是载体的一部分和/或此类多核苷酸或多肽可以是组合物的一部分,和是分离的,因为这种载体或组合物不是其最初环境的一部分。优选地,术语"分离的"相对于核酸分子使用时,如在"分离的核酸序列"中,指被鉴定并且与该核酸序列天然来源中通常与其接合的至少一种杂质性核酸分子分开的核酸序列。分离的核酸分子是这样的核酸分子,其在与自然界中找到该核酸分子的形式或环

境不同的形式或环境下存在。相反,未分离的核酸分子是以其在自然界中存在的状态所找到的核酸分子如DNA和RNA。例如,在宿主细胞染色体上相邻基因附近找到给定的DNA序列(例如,基因);作为与编码多种蛋白质的众多其他mRNA的混合物,在细胞中找到RNA序列,如编码特定蛋白质的特定mRNA序列。然而,包含例如SEQ ID NO:1的分离的核酸序列包括例如在细胞中通常含有SEQ ID NO:1的此类核酸序列,其中所述核酸序列位于不同于天然细胞的染色体或染色体外位置中,或侧翼分布有不同于自然界中存在的核酸序列。分离的核酸序列可以以单链或双链形式存在。当分离的核酸序列用来表达蛋白质时,该核酸序列将最少含有有义链或编码链的至少一部分(即,该核酸序列可以是单链的)。备选地,它可以含有有义链和反义链(即,该核酸序列可以是双链的)。

[0094] 最小启动子:在缺少上游激活情况下无活性或启动子活性大大降低的启动子元件,尤其是TATA元件。在合适的转录因存在下,最小启动子发挥作用以引起转录。

[0095] NEENA:参见“增强核酸表达的核酸”。

[0096] 非编码:术语“非编码”指核酸分子中不编码所表达蛋白质的部分或全部的序列。非编码序列包括但不限于内含子、增强子、启动子区、3'非翻译区和5'非翻译区。

[0097] 增强核酸表达的核酸(NEENA):术语“增强核酸表达的核酸”指这样的序列和/或核酸分子,其是具有在与NEENA功能性连接的启动子的控制下增强核酸表达的内在特性的特定序列。不同于启动子序列,NEENA本身不能驱动表达。为了履行增强与NEENA功能性连接的核酸分子表达的职能,NEENA本身应当与启动子功能性连接。与本领域已知的增强子序列的区别在于,NEENA以顺式方式而不以反式方式起作用,并且必需靠近待表达核酸的转录起始位点存在。

[0098] 核酸和核苷酸:术语“核酸”和“核苷酸”指天然存在的或合成的或人工核酸或核苷酸。术语“核酸”和“核苷酸”包括处于单链或双链、有义或反义形式的脱氧核糖核苷酸或核糖核苷酸或其任何核苷酸类似物和聚合物或杂合体。除非另外说明,特定核酸序列也隐含地包括其保守方式修饰的变体(例如简并密码子置换)和互补序列,以及明确指出的序列。术语“核酸”在本文中“基因”、“DNA”、“mRNA”、“寡核苷酸”和“多核苷酸”互换地使用。核苷酸类似物包括在碱基、糖和/或磷酸酯的化学结构中具有修饰的核苷酸,所述修饰包括但不限于5-位置嘧啶修饰、8-位置嘌呤修饰、胞嘧啶环外环胺处的修饰、5-溴-尿嘧啶置换等;和2'-位置糖修饰,包括但不限于糖修饰的核糖核苷酸,其中2'-OH由选自H、OR、R、卤素、SH、SR、NH₂、NHR、NR₂或CN的基团替换。短发夹RNA(shRNA)也可以包含非天然元件如非天然碱基,例如,肌苷和黄嘌呤,非天然糖,例如,2'-甲氧基核糖,或非天然磷酸二酯键,例如甲基磷酸酯、硫代磷酸酯和肽。

[0099] 核酸序列:短语“核酸序列”指从5'-端至3'-端读取的脱氧核糖核苷酸或核糖核苷酸的单链或双链聚合物。它包括染色体DNA、自我复制型质粒、DNA或RNA的感染性聚合物、和主要发挥结构性作用的DNA或RNA。“核酸序列”也指代表核苷酸的缩写、字母、字符或字的连续串。在一个实施方案中,核酸可以是“探针”,所述探针是相对短的核酸,通常长度小于100个核苷酸。经常地,核酸探针具有约50个核苷酸长度至约10个核苷酸长度。核酸的“靶区域”是核酸中被鉴定为有目标的部分。核酸的“编码区”是核酸的部分,其中置于适宜的调节性序列控制下时,所述部分以序列特异性方式转录和翻译以产生特定的多肽或蛋白质。称该编码区编码这种多肽或蛋白质。

[0100] 寡核苷酸:术语“寡核苷酸”指核糖核酸(RNA)或脱氧核糖核酸(DNA)或其模拟物的低聚物或聚合物,以及具有类似发挥作用的非天然存在部分的寡核苷酸。此类修饰或取代的寡核苷酸因合乎需要的特性,例如增强的细胞摄取、增强的核酸靶亲和力和在核酸酶存在下增加的稳定性而经常优选地胜过其天然形式。寡核苷酸优选地包括通过键(例如,磷酸二酯键)或取代键(substitute linkage)相互共价偶联的两个或多个核苷酸单体(nucleomonomer)。

[0101] 突出端:“突出端”是双链寡核苷酸分子的5'-或3'-羟基端上相对短的单链核苷酸序列(也称作“延伸”、“伸出端”或“粘末端”)。

[0102] 植物:通常理解为意指能够光合作用的任何真核单细胞或多细胞生物或其细胞、组织、器官、部分或繁殖材料(如种子或果实)。为本发明的目的,包括植物界(Plant Kingdom)的全部属和物种的高等和低等植物。优选地一年生、多年生、单子叶和双子叶植物。该术语包括成熟植物、种子、幼苗和籽苗及其衍生部分、繁殖材料(如种子或微孢子)、植物器官、组织、原生质体、愈伤组织和其他培养物(例如细胞培养物),和归并成产生功能单元或结构单元的任何其他类型的植物细胞。成熟植物指在除籽苗之外的任何目的发育阶段的植物。籽苗指在早期发育阶段的年幼不成熟植物。一年生、二年生、单子叶和双子叶植物是用于产生转基因植物的优选宿主生物。基因的表达在全部观赏植物、用材树或观赏树、花、切花、灌木或草坪草中是更进一步有利的。可以用举例方式、但非限制方式提到的植物是被子植物、苔藓植物例如獐耳细辛属(Hepaticae)(地钱类(liverwort))和藓纲(Musci)(藓类植物);蕨类植物如蕨类、木贼类和石松类;裸子植物如松柏类植物、苏铁植物、银杏(ginkgo)和买麻藤科(Gnetaceae)植物;藻类如绿藻纲(Chlorophyceae)、褐藻纲(Phaeophyceae)、红藻纲(Rhodophyceae)、粘藻纲(Myxophyceae)、黄藻纲(Xanthophyceae)、硅藻纲(Bacillariophyceae)(硅藻类)和裸藻纲(Euglenophyceae)。优选地用于食物或饲料目的的植物,如豆科(Leguminosae)如豌豆、苜蓿和大豆;禾本科(Gramineae)如稻、玉米、小麦、大麦、高粱、黑麦、黑小麦或燕麦;伞形科(Umbelliferae),具体地是胡萝卜属(Daucus)(非常特别地是物种胡萝卜(carota))和芹属(Apium)(非常特别地是物种旱芹(Graveolens dulce(celery)))及众多其它植物;茄科(Solanaceae),特别地是番茄属(Lycopersicon),非常特别地是物种番茄(tomato),和茄属(Solanum),非常特别地是物种马铃薯(tomato)和茄子(aubergine)和众多其它植物(如烟草(tobacco)),和辣椒属(Capsicum),非常特别地是物种辣椒(Capsicum annum(pepper))物种及众多其它植物;豆科(Leguminosae),特别地是大豆属(Glycine),非常特别地是物种大豆(soybean)、苜蓿、豌豆、紫花苜蓿、菜豆或花生及众多其它植物;和十字花科(Brassicaceae),特别地是芸苔属(Brassica),非常特别地是物种欧洲油菜(oilseed rape)、芸苔(beet)、甘蓝(Brassica oleracea cvTastie(卷心菜))、花椰菜(Brassica oleracea cv Snowball Y(花椰菜))和花茎甘蓝(oleracea cv Emperor(broccoli));和拟南芥属,非常特别地是物种拟南芥及众多其它植物;菊科(Compositae),具体地是莴苣属(Lactuca),非常特别地是物种莴苣(lettuce)及众多其它植物;菊科(Asteraceae)如向日葵、万寿菊、莴苣或金盏花及众多其它植物;葫芦科(Cucurbitaceae)如甜瓜、西葫芦/南瓜或夏南瓜,和亚麻。更优选地是棉属植物、甘蔗、大麻、亚麻、辣椒和多种树、坚果和藤本(wine)物种。

[0103] 多肽:术语“多肽”、“肽”、“寡肽”、“多肽”、“基因产物”、“表达产物”和“蛋白质”在

文中互换地使用,用来指连续氨基酸残基的聚合物或低聚物。

[0104] 前蛋白:正常情况下靶向细胞器如叶绿体并且仍包含其转运肽的蛋白质。

[0105] 初级转录物:如本文中所用的术语“初级转录物”指基因的不成熟RNA转录物。“初级转录物”例如仍包含内含子和/或仍不包含聚腺苷酸尾或帽结构和/或是缺少对其作为转录物正常发挥作用必需的其他修饰,例如修剪或编辑。

[0106] 启动子:术语“启动子”或“启动子序列”是等同物并且如本文中所用的,指与目的核苷酸序列连接时能够控制目的核苷酸序列转录成RNA的DNA序列。此类启动子可以例如在以下公共数据库中<http://www.grassius.org/grasspromdb.html>、<http://mendel.cs.rhul.ac.uk/mendel.php?topic=plantprom>、<http://ppdb.gene.nagoya-u.ac.jp/cgi-bin/index.cgi>找到。其中所列的启动子可以用于本发明的方法并且在此引用而包括。启动子位于由该启动子控制转录成mRNA的目的核苷酸序列的5' (即,上游),靠近其转录起始位点,并且为RNA聚合酶和用于转录起始的其他转录因子的特异性结合提供位点。所述启动子包含靠近转录起始位点例如至少10kb,例如5kb或2kb。它也可以包含靠近转录起始位点至少1500bp,优选地至少1000bp,更优选地至少500bp,甚至更优选地至少400bp,至少300bp,至少200bp或至少100bp。在又一个优选的实施方案中,启动子包含靠近转录起始位点至少50bp,例如至少25bp。启动子不包含外显子和/或内含子区或5' 非翻译区。启动子可以例如相对于相应植物为异源或同源。如果多核苷酸序列源自外来物种,或如果来自相同物种,但从其原有形式中被修饰,则它相对于某生物或第二多核苷酸序列为“异源”。例如,与异源编码序列有效连接的启动子指编码序列来自不同于衍生该启动子的物种中,或,如果来自相同物种,编码序列与该启动子不天然接合(例如基因修饰的编码序列或来自不同生态型或品种的等位基因)。合适的启动子可以源自其中应当出现表达的宿主细胞的基因或源自该宿主细胞的病原体(例如,植物或植物病原体如植物病毒)。植物特异性启动子是适于调节植物中表达的启动子。这种启动子可以源自植物,也可以源自植物病原体,或它可以是由人设计的合成性启动子。如果启动子是诱导型启动子,则转录的速率应答复于诱导剂而增加。另外,启动子可以以组织特异性或组织优先的方式受到调节,从而它仅仅或优势地在特定组织类型如叶、根或分生组织中转录所接合的编码区方面有活性。用于启动子时,术语“组织特异性”指这样的启动子,该启动子能够指导目的核苷酸序列在特定类型的组织(例如,花瓣)中选择性表达,同时相同目的核苷酸序列在不同类型组织(例如,根)中相对缺少的表达。可以例如通过这样的方式评价启动子的组织特异性:将报道基因与该启动子序列有效连接以产生报道构建体,将报道构建体导入植物的基因组,从而该报道构建体整合至所产生的转基因植物的每种组织中,并且检测报道基因在转基因植物的不同组织中的表达(例如,检测mRNA、蛋白质或由报道基因编码的蛋白质的活性)。相对于报道基因在其他组织中的表达水平,检测到该报道基因在一种或多种组织中的更大表达水平表明该启动子对其中检测到更大表达水平的组织是特异性的。用于启动子时,术语“细胞类型特异的”指这样的启动子,该启动子能够指导目的核苷酸序列在特定类型的细胞中选择性表达,同时相同目的核苷酸序列在相同组织内部的不同类型细胞中相对缺少的表达。用于启动子时,术语“细胞类型特异的”还意指能够促进目的核苷酸序列在单一组织内部的某区域中选择性表达的启动子。使用本领域熟知的方法,例如GUS活性染色法、GFP蛋白法或免疫组织化学染色法,可以评估启动子的细胞类型特异性。谈及启动子或源自启动子的表达时,术语“组

成型”意指能够指导有效连接的核酸分子在刺激物(例如,热休克、化学品、光等)不存在下在大部分植物组织和细胞中贯穿植物或植物部分的基本上整个生活期限转录的启动子。一般而言,组成型启动子能够指导转基因在基本上任何细胞和任何组织中的表达。

[0107] 启动子特异性:谈及启动子时,术语“特异性”意指由相应启动子引起的表达的样式。特异性描述了植物或其部分的组织和/或发育状态,在其中该启动子引起处于相应启动子控制下的核酸分子表达。启动子的特异性也可以包含环境条件,在所述环境条件下可以激活或下调该启动子,如由生物胁迫或环境胁迫如寒冷、干旱、受伤或感染诱导或阻遏。

[0108] 纯化:如本文中所用,术语“纯化的”指从其天然环境取出、分离或分开的分子,即核酸序列或氨基酸序列。“基本上纯化的”分子至少60%没有、优选地至少75%没有和更优选地至少90%没有与它们天然结合在一起的其他组分。纯化的核酸序列可以是分离的核酸序列。

[0109] 重组:就核酸分子而言,术语“重组”指通过重组DNA技术产生的核酸分子。重组核酸分子也可以包括本身不存在于自然界中但是被人修饰、改变、突变或操纵的分子。优选地,“重组核酸分子”是在序列方面与天然存在的核酸分子有至少一个核酸不同的非天然存在的核酸分子。“重组核酸分子”也可以包括“重组构建体”,其包含不天然地以这种顺序存在的一系列核酸分子,所述核酸分子优选地有效地连接。用于产生所述重组核酸分子的优选方法可以包括克隆技术、定向或非定向诱变法、合成或重组技术。

[0110] 有义:术语“有义”理解为意指核酸分子,其具有与靶序列互补或相同的序列,例如与蛋白质转录因子结合和参与给定基因表达的序列。根据一个优选的实施方案,该核酸分子包含目的基因和允许表达该目的基因的元件。

[0111] 显著的增加或减少:大于测量技术中固有误差幅度的增加或减少,例如在酶活性或在基因表达方面,优选地是对照酶的活性或对照细胞中的表达增加或减少约2倍或更多,更优选地增加或减少约5倍或更多,并且最优选地增加或减少约10倍或更多。

[0112] 小核酸分子:“小核酸分子”理解为由核酸或其衍生物如RNA或DNA组成的分子。它们可以是双链或单链的,并且其长度在约15和约30bp之间,例如在15和30bp之间,更优选地在约19和约26bp之间,例如在19和26bp之间,甚至更优选地在约20和约25bp之间,例如在20和25bp之间。在一个特别优选的实施方案中,寡核苷酸的长度在约21和约24bp之间,例如在21和24bp之间。在一个最优选的实施方案中,小核酸分子的长度是约21bp和约24bp,例如21bp和24bp。

[0113] 基本上互补的:在最广意义上,本文中就核苷酸序列相对于参比或靶核苷酸序列而言使用时,术语“基本上互补的”意指这样的核苷酸序列,其具有在基本上互补的核苷酸序列和所述参比或靶核苷酸序列的完全互补序列之间至少60%,更希望地至少70%,更希望地至少80%或85%,优选地至少90%,更优选地至少93%,仍更优选地至少95%或96%,仍旧更优选地至少97%或98%,依旧更优选地至少99%或最优选地100%的同一性百分数(后者等同于本上下文中的术语“同一的”)。优选地,在至少19个核苷酸长度、优选地至少50个核苷酸长度、更优选地核酸序列的全部长度范围内针对所述参比核苷酸序列评估同一性(如果不是,则另外在下文说明)。使用威斯康辛大学GCG的默认GAP分析,GAP的SEQWEB应用,基于Needleman和Wunsch算法(Needleman和Wunsch (1970) J Mol.Biol.48:443-453;如上文定义)实施序列比较。与参比核苷酸序列“基本上互补的”核苷酸序列与该参比核苷酸序列

在低严格性条件、优选地中等严格性条件、最优选地高严格性条件(如上文定义)下杂交。

[0114] 转基因:如本文中所用的术语“转基因”指通过实验操作导入细胞的基因组中的任何核酸序列。转基因可以是“内源DNA序列”或“异源DNA序列”(即,“外来DNA”)。术语“内源DNA序列”指这样的核苷酸序列,其天然存在于导入该核苷酸序列的细胞中,只要它相对于天然存在序列不含有一些修饰(例如,点突变、存在可选择标记基因等)。

[0115] 转基因的:当提及生物时,术语“转基因的”意指用重组DNA分子转化生物,优选地稳定转化生物,其中所述重组DNA分子优选地包含与目的DNA序列有效连接的合适启动子。

[0116] 载体:如本文中所用,术语“载体”指能够运输已经与之连接的另一个核酸分子的核酸分子。一种类型的载体是基因组整合的载体,或“整合的载体”,所述载体可以整合至宿主细胞的染色体DNA中。另一个类型的载体是游离型载体,即,能够进行染色体外复制的核酸分子。本文中将能够指导与它们有效连接的基因表达的载体称作“表达载体”。在本说明书中,“质粒”和“载体”互换地使用,除非从上下文显而易见并非如此。设计旨在体外或体内产生如本文所述RNA的表达载体可以含有被任何RNA聚合酶识别的序列,所述的RNA聚合酶包括线粒体RNA聚合酶、RNA pol I、RNA pol II和RNA pol III。这些载体可以用来在根据本发明的细胞中转录想要的RNA分子。将植物转化载体理解为适用于植物转化过程中的载体。

[0117] 野生型:就生物、多肽或核酸序列而言,术语“野生型”、“天然”或“天然来源”意指所述生物是天然存在的或是在至少一种天然存在生物中可获得的,其没有被改变、突变或由人类操作。

实施例

[0118] 化学品和常用方法

[0119] 除非另外说明,如(Sambrook等人,1989)所述为本发明的目的进行克隆过程,包括限制性消化、琼脂糖凝胶电泳、核酸的纯化、核酸的连接、细菌细胞的转化、选择和培养。使用Sanger技术(Sanger等人,1977),用激光荧光DNA测序仪(Applied Biosystems,Foster City,CA,USA)进行重组DNA的序列分析。除非另外描述,从Sigma Aldrich(Sigma Aldrich, St.Louis,USA)、Promega(Madison,WI,USA)、Duchefa(Haarlem,The Netherlands)或Invitrogen(Carlsbad,CA,USA)获得化学品和试剂。限制性核酸内切酶来自New England Biolabs(Ipswich,MA,USA)或Roche Diagnostics GmbH(Penzberg,Germany)。寡核苷酸由Eurofins MWG Operon(Ebersberg,Germany)合成。

[0120] 实施例1:从具有组成型表达的基因鉴定增强核酸表达的核酸(NEENA)

[0121] 1.1从拟南芥(*A.thaliana*)基因鉴定NEENA分子

[0122] 使用公开可获得的基因组DNA序列(例如<http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>)和转录物表达数据(例如<http://www.weigelworld.org/resources/microarray/AtGenExpress/>),选择源自高度表达性组成型基因的拟南芥(*Arabidopsis thaliana*)转录物中的一组18个潜在NEENA候选物用于详细分析。此外,该分析中还包括源自欧芹的推定的NEENA分子。候选物命名如下:

[0123] 表1:组成型NEENA候选物(NEENAc)。

[0124]

NEENA 名称	基因座	注释	SEQ ID NO
NEENAc24		欧芹 (<i>Petroselinum crispum</i>) 多聚遍在蛋白基因 Pcubi4-2	1
NEENAc17	At2g47170	ADP-核糖基化因子 1(ARF1)	2
NEENAc5	At1g56070	推定的延伸因子 2/推定的 EF-2	3
NEENAc18	At5g54760	推定的真核生物翻译起始因子 SUI1	4
NEENAc7	At4g02890	多聚遍在蛋白(UBQ14)	5
NEENAc13	At3g03780	AtMS2(拟南芥甲硫氨酸合酶 2)	6
NEENAc1	At5g60390	推定的延伸因子 1- α /推定的 EF-1- α	7
NEENAc21	At1g14400	遍在蛋白缀合酶 1(UBC1)	8
NEENAc16	At4g14880	半胱氨酸合酶/O-乙酰丝氨酸硫解酶/O-乙酰丝氨酸硫化氢解酶(OAS1)	9
NEENAc2	At4g27960	遍在蛋白缀合酶 E2-17kDa 9(UBC9)	10
NEENAc14	At1g64230	推定的遍在蛋白缀合酶	11
NEENAc4	At2g37270	40S 核糖体蛋白 S5(RPS5A)	12
NEENAc6	At4g05050	多聚遍在蛋白(UBQ11)	13
NEENAc8	At1g43170	60S 核糖体蛋白 L3(RPL3A)	14
NEENAc11	At1g01100	60S 酸性核糖体蛋白 P1(RPPIA)	15
NEENAc12	At5g04800	40S 核糖体蛋白 S17(RPS17D)	16
NEENAc19	At4g34110	聚腺苷酸结合蛋白 2(PABP2)	17
NEENAc22	At2g34770	脂肪酸羧化酶(FAH1)(预测的 IME 效应)	18
NEENAc23	At5g17920	钴胺素非依赖性甲硫氨酸合酶(CIMS)	19

[0125]

[0126] 1.2NEENA候选物的分离

[0127] 使用Qiagen DNeasy Plant Mini试剂盒(Qiagen,Hilden,Germany)从拟南芥绿色组织提取基因组DNA。对于具有SEQ ID NO1的推定的NEENA分子,使用载体构建体

1bXPcUbi4-2GUS (WO 2003102198) 的DNA。通过常规的聚合酶链反应 (PCR) 分离含有推定的 NEENA 分子的基因组DNA片段。该聚合酶链反应包含19组引物 (表2)。基于具有多种NEENA候选物的拟南芥基因组序列设计出引物。将载体构建体1bXPcUbi4-2GUS (WO 2003102198) 的核苷酸序列用于设计引物 (SEQ ID NO56和57), 所述引物用于扩增带有SEQ ID NO1的NEENA候选物 (表2)。该聚合酶链反应遵从由Phusion高保真DNA聚合酶 (目录编号F-540L, New England Biolabs, Ipswich, MA, USA) 概述的方案。使用分离的DNA作为利用以下引物的PCR扩增中的模板DNA:

[0128] 表2: 引物序列

引物名称	序列	SEQ ID	PCR 产生的 SEQ ID
		NO	NO
NEENAc1_for	tttatggtaccagcgcgaagactccttcagattct	20	7
NEENAc1_rev	aaattccatggttagctgtcaaaacaaaaacaaaaatcga	21	
NEENAc2_for	aaaaagggtacctcgaagaacacaaaaacaaaaacgtga	22	10
NEENAc2_rev	ttttccatgggttatttatccaaaaatccacgatccaaattcca	23	
NEENAc4_for	tttttggtaccgatccctacttctctcgacact	24	12
NEENAc4_rev	ttttaccatggtgactggaggatcaatagaagat	25	

[0130]

NEENAc5_for	tttttggtacctttctctctgttctcatctttctctct	26	3
NEENAc5_rev	taatagatatctttgtcaaaccttttgattgtcacct	27	
NEENAc6_for	tataaggtaccaaatcaatctctcaaatctctca	28	13
NEENAc6_rev	tttatecatggtctgttaatecagaaaaaccgagat	29	
NEENAc7_for	tatatggtaccaaatcggtctttcaaatctctca	30	5
NEENAc7_rev	ttataccatggtctgttaattcacaaaaaactgaga	31	
NEENAc8_for	tttttggtacctcatcggttgagcttagaagc	32	14
NEENAc8_rev	ttttccatggtctctctctctctctctctacatca	33	
NEENAc11_for	tatatggtaccaaagcattttcgatcttactcttaggt	34	15
NEENAc11_rev	ttttccatggttttttctctgaaacgattca	35	
NEENAc12_for	tttttggtaccttttgacgcgcgcgtctctctct	36	16
NEENAc12_rev	ttttccatggtctttcagttacctgtgtgacttacct	37	
NEENAc13_for	tttaaggtaacctctctcatctccactctct	38	6
NEENAc13_rev	ttttgatactttttgtttgtttttgtttttact	39	
NEENAc14_for	ttataggtaccaagtgaatcgtaaaaccgagtt	40	11
NEENAc14_rev	ttttccatggtctctcaacaaaaaaaactcct	41	
NEENAc16_for	tttttggtaccacgattcgggtcaaggtattga	42	9
NEENAc16_rev	ttttccatggtgatteaagcttactgcttaattcaca	43	
NEENAc17_for	tttttggtaccttagatctctgctgccgtctgcca	44	2
NEENAc17_rev	ttttccatggtttgatcaagcctgttcaca	45	
NEENAc18_for	aaaaaggtaacctcatcagatcttcaaaaccccaa	46	4
NEENAc18_rev	aaaaaccatggtgatttgagggtagtactaaccgggaa	47	
NEENAc19_for	ttttaggtaccatacggttaacttcaccaatcccaa	48	17
NEENAc19_rev	ttttccatggttaattaatgcagtctttgtggtcgatgga	49	
NEENAc21_for	ttttccgggatctttacctcaacaacgagat	50	8
NEENAc21_rev	ttttccatggttatcctctctttcttctaataaacaaccca	51	
NEENAc22_for	tttttggtacctctcttccgtctcgagtcgctgaga	52	18
NEENAc22_rev	ttttccatggtttgcagaccttttactgat	53	

[0131]	NEENAc23_for	tttttggtaccttctctctctctcgattcttct	54	19
	NEENAc23_rev	tttttccatgggtattgatttttttactgcat	55	
	NEENAc24_for	tttttggtaccttaagaaatctctcttctct	56	1
	NEENAc24_rev	tttttccatgggtctgcacatacataacatatca	57	

[0132] PCR期间按以下组成实施扩增(50微升)：

[0133] 3.00微升拟南芥基因组DNA(50ng/微升基因组DNA,5ng/微升载体构建体)

[0134] 10.00微升5x Phusion HF缓冲液

[0135] 4.00微升dNTP(2.5mM)

[0136] 2.50微升正向引物(10μM)

[0137] 2.50微升反向引物(10μM)

[0138] 0.50微升Phusion HF DNA聚合酶(2U/微升)

[0139] 将递降方法用于以下参数的PCR:98.0℃30秒(1个循环),98.0℃30秒,56.0℃30秒和72.0℃60秒(4个循环),额外4个循环,每个循环的复性温度是:54.0℃,51.0℃和49.0℃,随后20个循环:98.0℃30秒,46.0℃30秒和72.0℃60秒(4个循环)及72.0℃5分钟。将扩增产物载于2%(w/v)琼脂糖凝胶上并且在80V分离。将PCR产物从该凝胶切下并且用Qiagen凝胶提取试剂盒(Qiagen,Hilden,Germany)纯化。用KpnI(10U/微升)和NcoI(10U/微升)或EcoRV(10U/微升)限制性核酸内切酶进行DNA限制性消化后,消化的产物再次用Qiagen凝胶提取试剂盒(Qiagen,Hilden,Germany)纯化。

[0140] 1.3载体构建

[0141] 1.3.1产生带有潜在NEENA分子的载体构建体

[0142] 使用Multisite Gateway系统(Invitrogen,Carlsbad,CA,USA),将启动子::NEENA::报道基因盒装配入用于植物转化的二元构建体。将拟南芥p-AtNit1(At3g44310,GenBank X86454;W003008596,前缀p-指启动子)启动子用于报道基因构建体中,并且将萤火虫萤光素酶(Promega,Madison,WI,USA)用作报道蛋白以定量地测定待分析的推定的NEENA分子的表达增强作用。

[0143] 根据制造商手册(Invitrogen,Carlsbad,CA,USA),在任一端处带有位点特异性重组位点的基因组DNA(见上文)上,用引物p-AtNit1-for和p-AtNit1-rev(表3),通过pDONR/A载体和p-AtNit1扩增产物之间的位点特异性重组(BP反应)克隆携带有p-AtNit1启动子的pENTR/A载体。对阳性pENTR/A克隆进行序列分析以确保p-AtNit1启动子的正确性。

[0144] 表3:引物序列(p-AtNit1)

[0145]	引物名称	序列	SEQ ID NO.
	p-AtNit1-for	ggggacaactttgtatagaaaagttgtcgagaccagatgttttacacttga	58
	p-AtNit1-rev	ggggactgctttttgtacaaacttggacactcagagacttgagagaagc a	59

[0146] 产生ENTR/B载体,其含有萤火虫萤光素酶编码序列(Promega, Madison, WI, USA),后接t-nos胭脂碱合酶转录终止子(Genbank V00087)。使用KpnI和NcoI或EcoRV限制酶,将NEENA候选物PCR片段(见上文)分别克隆在萤火虫萤光素酶编码序列的上游。表4中汇总所产生的pENTR/B载体,其中启动子分子具有前缀p-,编码序列具有前缀c-,并且终止子分子具有前缀t-。

[0147] 表4:带有和不带有NEENA候选物的全部pENTR/B载体

[0148]

部分表达盒的组成	
pENTR/B 载体	SEQ ID NO::报道基因::终止子
LJK1	MCS::c-LUC::t-nos
LJK4	SEQ ID NO1::c-LUC::t-nos
LJK40	SEQ ID NO7::c-LUC::t-nos
LJK41	SEQ ID NO10::c-LUC::t-nos
LJK43	SEQ ID NO12::c-LUC::t-nos
LJK44	SEQ ID NO3::c-LUC::t-nos
LJK46	SEQ ID NO13::c-LUC::t-nos

[0149]

LJK47	SEQ ID NO5::c-LUC::t-nos
LJK48	SEQ ID NO14::c-LUC::t-nos
LJK51	SEQ ID NO15::c-LUC::t-nos
LJK52	SEQ ID NO16::c-LUC::t-nos
LJK53	SEQ ID NO6::c-LUC::t-nos
LJK54	SEQ ID NO11::c-LUC::t-nos
LJK56	SEQ ID NO9::c-LUC::t-nos
LJK57	SEQ ID NO2::c-LUC::t-nos
LJK58	SEQ ID NO4::c-LUC::t-nos
LJK59	SEQ ID NO17::c-LUC::t-nos
LJK61	SEQ ID NO8::c-LUC::t-nos
LJK62	SEQ ID NO18::c-LUC::t-nos
LJK63	SEQ ID NO19::c-LUC::t-nos

[0150] 通过经KpnI和HindIII限制性位点导入一个多克隆位点(SEQ ID NO60)构建

pENTR/C载体。通过进行位点特异性重组 (LR反应), 根据制造商 (Invitrogen, Carlsbad, CA, USA) 的 Multisite Gateway 手册, 将产生的 pENTR/A、pENTR/B 和 pENTR/C 与 pSUN 目的载体 (pSUN 衍生物) 组合。所述反应产生了具有 p-AtNit1 启动子、萤火虫萤光素酶编码序列 c-LUC 和 t-nos 终止子的 1 个双元载体和紧邻萤火虫萤光素酶编码序列上游携带 SEQ ID NO1、NO2、NO3、NO4、NO5、NO6、NO7、NO8、NO9、NO10、NO11、NO12、NO13、NO14、NO15、NO16、NO17、NO18 和 NO19 的 19 个载体 (表5), 对于所述载体, 给出与 SEQ ID NO1 的组合作为示例 (SEQ ID NO61)。除了 SEQ ID NO2 至 NO19 变化外, 全部载体中的核苷酸序列是相同的 (表5)。表5中汇总所产生的植物转化载体。

[0151] 表5: 用于拟南芥转化的植物表达载体

[0152]	植物表达载体	表达盒的组成	
		启动子::SEQ ID NO::报道基因::终止子	SEQ ID NO
	LJK132	p-AtNit1::-::c-LUC::t-nos	

[0153]

LJK133	p-AtNit1::SEQ ID NO1::c-LUC::t-nos	61
LJK91	p-AtNit1::SEQ ID NO7::c-LUC::t-nos	
LJK92	p-AtNit1::SEQ ID NO10::c-LUC::t-nos	
LJK94	p-AtNit1::SEQ ID NO12::c-LUC::t-nos	
LJK95	p-AtNit1::SEQ ID NO3::c-LUC::t-nos	
LJK97	p-AtNit1::SEQ ID NO13::c-LUC::t-nos	
LJK98	p-AtNit1::SEQ ID NO5::c-LUC::t-nos	
LJK99	p-AtNit1::SEQ ID NO14::c-LUC::t-nos	
LJK102	p-AtNit1::SEQ ID NO15::c-LUC::t-nos	
LJK103	p-AtNit1::SEQ ID NO16::c-LUC::t-nos	
LJK104	p-AtNit1::SEQ ID NO6::c-LUC::t-nos	
LJK105	p-AtNit1::SEQ ID NO11::c-LUC::t-nos	
LJK107	p-AtNit1::SEQ ID NO9::c-LUC::t-nos	
LJK108	p-AtNit1::SEQ ID NO2::c-LUC::t-nos	
LJK109	p-AtNit1::SEQ ID NO4::c-LUC::t-nos	
LJK110	p-AtNit1::SEQ ID NO17::c-LUC::t-nos	
LJK112	p-AtNit1::SEQ ID NO8::c-LUC::t-nos	
LJK113	p-AtNit1::SEQ ID NO18::c-LUC::t-nos	
LJK114	p-AtNit1::SEQ ID NO19::c-LUC::t-nos	

[0154] 所产生的载体随后用来瞬时转化拟南芥叶原生质体。

[0155] 1.3.2海肾 (Renilla) 萤光素酶对照构建体

[0156] 使用10ng来自Promega (Madison, WI, USA) 的质粒pRL-null作为DNA模板并且使用引物R-LUC_for和R-LUC_rev (表6), 以如上文所述的PCR参数, 扩增海肾萤光素酶cDNA。

[0157] 表6: 引物序列 (c-RLUC)

[0158]

引物名称	序列	SEQ ID NO
RLUC_for	aaaaaggtaccatgacttcgaaagtftatgatc	62
RLUC_rev	aaattgagetcttattgttcatttttgagaactc	63

[0159]

[0160] 用KpnI (10U/微升) 和SacI (10U/微升) 限制性核酸内切酶进行DNA限制性消化后, 消化的产物再次用Qiagen凝胶提取试剂盒 (Qiagen, Hilden, Germany) 纯化。

[0161] 将片段通过KpnI和SacI限制性位点克隆至含有胭脂碱合酶组成型启动子p-nos (Genbank V00087) 接续t-nos胭脂碱合酶转录终止子 (Genbank V00087) 的ENTR/B载体中, 从而产生pENTR/B克隆, 对所述pENTR/B克隆进行序列分析以确保含有海肾萤光素酶的表达盒的正确性。

[0162] 实施例2: 筛选在瞬时转化的拟南芥叶原生质体中增强基因表达的NEENA候选分子

[0163] 本实施例说明仅选择的NEENA分子能够增强基因表达。

[0164] 2.1 拟南芥叶原生质体的分离和瞬时转化

[0165] 根据建立的方案 (Damm和Willmitzer, 1988; Damm等人, 1989) 修订拟南芥叶原生质体的分离和瞬时转化。将4周龄拟南芥植物的叶用剃须刀片切成小片并且转移至具有1.5% 纤维素酶R10 (Duchefa, Haarlem, The Netherlands)、0.3% 离析酶R10 (Duchefa, Haarlem, The Netherlands)、400mM甘露醇、20mM KCl、20mM MES、10mM CaCl₂、pH5.7的溶液, 并且在室温孵育过夜。鉴于瞬时拟南芥叶原生质体转化法的变异性, 使用海肾萤光素酶 (双萤光素酶®报道分子测定系统, Promega, Madison, WI, USA) 来归一化上述构建体的萤火虫萤光素酶表达能力。使用PEG (聚乙二醇) 和 1×10^4 个原生质体, 一式三份用6微克质粒DNA进行无NEENA的载体构建体 (LJK132) 和每种含NEENA的载体构建体 (LJK66-LJK114) 的瞬时转化, 所述质粒在转化前与25微克含有海肾萤光素酶的构建体混合。

[0166] 2.2 双萤光素酶报道基因测定法

[0167] 通过以100g离心收集转染的拟南芥原生质体并且在除去上清液后冷冻于液氮中。根据制造商 (Promega, Madison, WI, USA) 的双萤光素酶报道分子分析系统手册进行转染细胞中检测萤火虫和海肾萤光素酶活性的测定法。发光测量在MicroLumat Plus LB96V (Berthold Technologies, Bad Wildbad, Germany) 中实施, 在添加萤光素酶底物后记录。通过产生萤火虫萤光素酶和海肾萤光素酶之间的比率, 使两种萤光素酶记录的仪器读数归一化。来自三个实验的数据对每种构建体平均化, 并且基于这些平均表达值, 计算倍数变化值以评估推定的NEENA的存在对缺少相应推定的NEENA的报道基因构建体的影响。与仅有p-AtNit1启动子缺少NEENA的报道基因构建体相比, 19个含有NEENA候选物的受检构建体显示阴性作用以及阳性作用, 范围从0.1倍至18.1倍萤火虫萤光素酶活性诱导作用 (图1)。总计, 基于萤光素酶报道基因活性, 与缺少NEENA仅有启动子的报道基因构建体相比 (图1), 包含SEQ ID NO1、NO2、NO3、NO4、NO5、NO6、NO7、NO8和NO9序列的9个推定的NEENA分子引起基因表达增加大于2倍, 并且因而是有功能的NEENA分子。由于众多的受检NEENA候选分子对基因表达增强作用产生边际的或甚至负性影响, 故并非全部推定的NEENA分子介导共同的刺激性作用, 而是选择的NEENA序列传递显著的基因表达增强作用 (SEQ ID NO.1至9)。

[0168] 实施例3: 在油籽油菜植物中测试用于增强基因表达的NEENA分子

[0169] 本实施例说明与缺少NEENA仅有启动子的方法相比, NEENA分子可以跨物种使用以增强全部受检组织中的基因表达。

[0170] 选择预筛选中介导最强基因表达增强作用的NEENA分子 (参考实施例2, SEQ ID NO1、NO2、NO3、NO4和NO5) 用于测定对转基因油籽油菜植物中基因表达水平的增强作用。

[0171] 3.1 构建用于欧洲油菜植物转化的载体

[0172] 为了转化油籽油菜植物,将具有或没有基因表达控制分子(SEQ IDs N01-N05)的报道基因表达盒与pENTR/C载体内部携带用于检测转基因植物品系的选择标记基因的基因表达盒组合。如前文所述(见上文1.4),通过进行位点特异性重组(LR反应),根据制造商(Invitrogen, Carlsbad, CA, USA)的Multisite Gateway手册,将携带所述选择标记盒的pENTR/A、pENTR/B和pENTR/C与pSUN目的载体组合。所述反应产生了具有p-AtNit1启动子、萤火虫萤光素酶编码序列c-LUC、t-nos终止子和所述选择标记盒的1个双元载体,以及紧邻萤火虫萤光素酶编码序列上游携带SEQ ID N01、N02、N03、N04和N05的5个载体(表7),对于所述载体,给出与SEQ ID N01的组合作为示例(SEQ ID N064)。除了SEQ ID N02至N05变化外,全部载体中的核苷酸序列是相同的(表7)。表7中汇总所产生的植物转化载体:

[0173] 表7:用于欧洲油菜(B.napus)转化的植物表达载体

植物表达载体	表达盒的组成	SEQ ID
	启动子::SEQ ID NO::报道基因::终止子	NO
LJK138	p-AtNit1::c-LUC::t-nos	
LJK139	p-AtNit1::SEQ ID N01::c-LUC::t-nos	64
LJK141	p-AtNit1::SEQ ID N03::c-LUC::t-nos	
LJK142	p-AtNit1::SEQ ID N05::c-LUC::t-nos	
LJK143	p-AtNit1::SEQ ID N02::c-LUC::t-nos	
LJK144	p-AtNit1::SEQ ID N04::c-LUC::t-nos	

[0175] 3.2转基因油菜籽植物的产生(根据Moloney等人,1992,Plant Cell Reports,8:238-242的修订方案)

[0176] 在产生转基因油菜籽植物的准备工作中,将所述双元载体转化至根癌农杆菌C58C1:pGV2260(Deblaere等人,1985,Nucl.Acids.Res.13:4777-4788)中。将携带相应双元构建体的农杆菌的过夜培养物的1:50稀释物在Murashige-Skoog培养基(Murashige和Skoog,1962,Physiol.Plant 15,473)的补充有3%蔗糖的培养基(3MS培养基)中培育。为转化油菜籽植物,将无菌植物的叶柄或下胚轴与1:50农杆菌溶液孵育5-10分钟,随后在黑暗下25℃于补充有0.8%细菌培养用琼脂的3MS培养基上共孵育3日。在3日后,将外植体转移至含有500mg/l凯福隆(头孢噻肟-钠)、100nM咪唑烟(Imazetapyr)、20μM苄氨基嘌呤(BAP)和1.6g/l葡萄糖的MS-培养基中于16小时光照/8小时黑暗的光照方案下,该过程以每周期间重复。将正在生长的幼苗转移至含有2%蔗糖、250mg/l凯福隆和0.8%细菌培养用琼脂的MS-培养基。3周后,将生长激素2-吡啶丁酸添加至该培养基以促进根形成。根发育后,将幼苗转移至土壤,在生长室中培育2周并且在温室条件下培育直至成熟。

[0177] 3.3植物分析

[0178] 组织样品从产生的转基因植物中自叶、花和长角果收集,贮藏在-80℃的冰箱中,进行萤光素酶报道基因测定(根据Ow等人,1986的修订方案)。在研磨后,将冷冻的组织样品重悬于800微升缓冲液I(0.1M磷酸盐缓冲液pH7.8,1mM DTT(Sigma Aldrich,St.Louis,MO,

USA), 0.05% Tween 20 (Sigma Aldrich, St. Louis, MO, USA)) 中, 随后以10000g离心10分钟。将75微升含水上清液转移至96孔平板。在添加25微升缓冲液II (80mM甘氨酸, 20mM Tris-HCl, 10mM MgSO₄, 10mM DTT, 10mM ATP (Sigma Aldrich, St. Louis, MO, USA), pH 7.8) 和D-荧光素至终浓度0.5mM (目录编号:L-8220, BioSynth, Staad, Switzerland) 后, 在MicroLumat Plus LB96V (Berthold Technologies, Bad Wildbad, Germany) 中记录发光, 从而产生单位, 是相对光照单位RLU/分钟 (RLU/min)。

[0179] 为了归一化样品之间的荧光素酶活性, 与荧光素酶活性平行地测定上清水液中的蛋白质浓度 (改编自Bradford, 1976, Anal. Biochem. 72, 248)。将缓冲液I中的5微升细胞水质提取物与250微升Bradford试剂 (Sigma Aldrich, St. Louis, MO, USA) 混合, 在室温孵育10分钟。在平板读数仪 (Thermo Electron Corporation, Multiskan Ascent 354) 中在595nm测定吸收。用事先产生的标准浓度曲线计算出样品中的总蛋白量。从RLU/min和mg蛋白质/ml样品的比率中产生的值对携带相同构建体的转基因植物平均化, 并且计算倍数变化以评估NEENA分子存在对缺少NEENA的报道基因构建体的影响。

[0180] 3.4NEENA序列在油籽油菜植物中介导强烈的基因表达增强作用

[0181] 为评估所选择NEENA分子 (SEQ ID NO: 1、2、3、4和5) 在油籽油菜植物中增强基因表达的潜力, 收集具有相同发育阶段并且在同等生长条件下培育的植物的叶、花和携带种子的长角果。样品取自独立的转基因油籽油菜植物品系, 所述的转基因油籽油菜植物品系携带仅有启动子的报道基因构建体或含有NEENA (SEQ ID NO: 1、2、3、4和5) 的荧光素酶报道基因构建体。从每个转基因事件收集10粒种子, 加工并如上文所述 (实施例3.3) 分析荧光素酶活性。

[0182] 与仅有组成型p-AtNit1启动子缺少NEENA的报道基因构建体相比, 5个受检NEENA分子均在叶组织中介导强烈的基因表达增强作用 (图2, a)。在油籽油菜的花和包含种子的长角果中检测到由NEENA (SEQ ID NO: 1、2、3、4和5) 介导的可比较的基因表达增强作用 (图2, b和c)。

[0183] 实施例4: 在大豆植物中分析基因表达的组成型增强作用

[0184] 本实施例说明与缺少NEENA仅有启动子的方法相比, NEENA分子可以在广泛类型的植物物种中并且跨越不同植物科的物种界限使用以增强全部组织中的基因表达。

[0185] 选择预筛选中介导最强基因表达增强作用的NEENA序列分子 (参考实施例2, SEQ ID NO: 1、2、3、4和5) 用于测定对转基因大豆植物中基因表达水平的增强作用。将植物表达载体LJK138、LJK139、LJK141、LJK142、LJK143和LJK144 (参考实施例3.1) 用于稳定的大豆转化。

[0186] 4.1转基因大豆植物的产生 (根据W02005/121345; Olhoft等人, 2007的修订方案)

[0187] 大豆种子萌发、繁殖、发根农杆菌和腋生分生组织外植体制备和接种如先前所述进行 (W02005/121345; Olhoft等人, 2007), 除了使用的构建体LJK138、LJK139、LJK141、LJK142、LJK143和LJK144 (参考实施例3.1) 各自含有受欧芹遍在蛋白启动子PcUbi4-2驱动的突变AHAS基因, 所述的突变AHAS基因介导耐受用于选择的咪唑啉酮除草剂。

[0188] 4.2NEENA序列在大豆植物中介导强烈的基因表达增强作用

[0189] 从产生的转基因植物中自叶、花和种子收集组织样品。组织样品如上文所述 (参考

实施例3.3)加工和分析。

[0190] 与仅有组成型p-AtNit1启动子缺少NEENA的报道基因构建体相比,5个受检NEENA分子均在叶中介导强烈的基因表达增强作用(图3a)。在大豆的花和长角果中检测到由NEENA(SEQ ID NO1至NO 5)介导的可比较的基因表达增强作用(图3,b和c)。

[0191] 实施例5:在单子叶植物中分析NEENA活性

[0192] 本实施例描述在单子叶植物中对具有SEQ ID NO 1、2、3、4和5的NEENA序列的分析。

[0193] 5.1载体构建

[0194] 为在单子叶植物中分析具有SEQ ID NO 1、2、3、4和5的NEENA序列,将携带表达盒的基于pUC的表达载体与接续胭脂碱合酶(NOS)转录终止子的β-葡糖苷酸酶(GUS)基因的编码序列组合,其中所述的表达盒包含缺少NEENA的来自玉米(Z.mais)的组成型单子叶启动子p-Ubi。使用Qiagen DNeasy Plant Mini试剂盒(Qiagen,Hilden,Germany)从拟南芥绿色组织提取基因组DNA。通过常规的聚合酶链反应(PCR)分离含有NEENA分子的基因组DNA片段。基于具有多种NEENA候选物的拟南芥基因组序列设计出引物。该反应包含5组引物(表8)并且使用以下引物,遵从由Phusion高保真DNA聚合酶(目录编号F-540L,New England Biolabs,Ipswich,MA,USA)概述的方案:

[0195] 表8:引物序列

	引物名称	序列	SEQ ID	PCR 产生的
			NO	SEQ ID NO
[0196]	NEENAc5_forII	tttttggegcgcctttctctegttctcatctttctctct	65	3
	NEENAc5_revII	taataggcgcgcctttgtcaaaacttttgattgtcact	66	
	NEENAc7_forII	tatatggcgcgcgcaaatcggtcttcaaatctctca	67	5
	NEENAc7_revII	ttataggcgcgcctctgttaattcacaaaaaactgaga	68	
	NEENAc17_forII	tttttggegcgccttagatctctgtgccgtctgtgga	69	2
	NEENAc17_revII	tttttggegcgcctttgatcaagcctgttcaca	70	
	NEENAc18_forII	aaaaaggcgcgcctcatcgatcttcaaaaccccaa	71	4
	NEENAc18_revII	aaaaaggcgcgcctgatttgagggtagtactaaccgg	72	
[0197]		gaa		
	NEENAc24_forII	tttttggegcgccttaagaaatctctctctctct	73	1
	NEENAc24_revII	tttttggegcgcctgcacatacataacatatca	74	

[0198] PCR期间的扩增和扩增产物的纯化如上述那样(实施例1.2)实施。用AscI (10U/微升)限制性核酸内切酶进行DNA限制性消化后,消化的产物用Qiagen凝胶提取试剂盒(Qiagen,Hilden,Germany)纯化。

[0199] 使用AscI限制性位点,将NEENA PCR片段(见上文)分别克隆在β-葡糖苷酸酶编码序列的上游。所述反应产生了具有p-Ubi启动子、β-葡糖苷酸酶编码序列c-GUC和t-nos终止

子的1个双元载体,以及紧邻 β -葡萄糖苷酸酶编码序列上游携带SEQ ID NO1、NO2、NO3、NO4和NO5的5个载体(表9),对于所述载体,给出与SEQ ID NO1的组合作为示例(SEQ ID NO75)。除了SEQ ID NO2至NO5变化外,全部载体中的核苷酸序列是相同的(表9)。表9中汇总所产生的载体,其中启动子分子具有前缀p-,编码序列具有前缀c-,并且终止子分子具有前缀t-。

[0200] 表9:植物表达载体

	植物表达载体	表达盒的组成	SEQ ID NO
		启动子::SEQ ID NO::报道基因::终止子	
[0201]	RTP2940	p-Ubi::c-GUS::t-nos	
	LJK361	p-Ubi::SEQ ID NO1::c-GUS::t-nos	75
	LJK362	p-Ubi::SEQ ID NO2::c-GUS::t-nos	
	LJK363	p-Ubi::SEQ ID NO3::c-GUS::t-nos	
	LJK364	p-Ubi::SEQ ID NO4::c-GUS::t-nos	
	LJK365	p-Ubi::SEQ ID NO5::c-GUS::t-nos	

[0202] 使用所产生的载体在下文概述的实验中分析NEENA分子(实施例5.2)

[0203] 5.2在单子叶植物组织中分析增强基因表达的NEENA分子

[0204] 这些实验通过轰击单子叶植物组织或培养细胞(实施例6.2.1)、通过PEG介导(或相似方法学)将DNA导入植物原生质体(实施例6.2.2)或通过农杆菌介导的转化(实施例6.3.3)进行。用于这些实验的靶组织可以是植物组织(例如叶组织)、培养的植物细胞(例如墨西哥黑色甜玉米(Black Mexican Sweet Corn,BMS)或用于农杆菌方案的植物胚。

[0205] 5.2.1使用微抛射体轰击的瞬时测定法

[0206] 使用Qiagen质粒试剂盒(目录号12143)分离质粒构建体。根据Sanford等人,(1993) (Optimizing the biolistic process for different biological applications (为不同生物学应用优化生物射弹方法).Methods in Enzymology,217:483-509)描述的方案,将DNA沉淀在0.60 μ M金粒子(Bio-Rad目录编号165-2262),并且使用PDS-1000/He系统装置(Bio-Rad)加速到靶组织(例如2周龄玉米叶、BMS培养的细胞等)。

[0207] 全部DNA沉淀步骤和轰击步骤在无菌条件下在室温进行。悬浮培养的墨西哥黑色甜玉米(BMS)细胞在BMS细胞培养液体培养基[Murashige和Skoog (MS) 盐(4.3g/L),3% (w/v) 蔗糖,肌醇(100mg/L),3mg/L 2,4-二氯苯氧乙酸(2,4-D),酪蛋白水解物(1g/L),硫胺素(10mg/L)和L-脯氨酸(1.15g/L),pH 5.8]中增殖。每周转移10mL稳定期细胞的培养物至40mL新鲜培养基并且在110转/分钟运转的回转摇床上的250mL烧瓶中于27℃培养。

[0208] 将硅化Eppendorf管中的60mg金粒子重悬于100%乙醇中,随后离心30秒。将沉淀在100%乙醇中淋洗1次并且在无菌水中淋洗2次,每次洗涤后离心。沉淀最终重悬于1mL无菌50%甘油中。随后将金悬液分成50微升等份并且在4℃贮藏。添加以下试剂至一等份中:5微升的1微克/微升总DNA,50微升2.5M CaCl₂,20微升0.1M亚精胺,游离碱。将DNA溶液涡旋混合1分钟并置于-80℃持续3分钟,随后离心10秒。除去上清液。通过轻击该管小心地使沉

淀重悬于1mL 100%乙醇中,随后离心10秒。除去上清液,并且将沉淀小心地重悬于50微升100%乙醇中并置于-80℃直至使用(轰击前30分钟至4小时)。如果在溶液中可见到金聚集体,则将管在临用前才在水浴超声波仪中超声处理1秒。

[0209] 为了轰击,将2周龄的玉米叶切成长度大约1cm的条并且近轴侧向上放置在渗透性诱导培养基M-N6-702[N6盐(3.96g/L),3% (w/v) 蔗糖,1.5mg/L 2,4-二氯苯氧乙酸(2,4-D),酪蛋白水解物(100mg/L)和L-脯氨酸(2.9g/L),MS维生素母液(1mL/L),0.2M甘露醇,0.2M山梨醇,pH 5.8]上。将叶条孵育1-2小时。

[0210] 在BMS培养的细胞的情况下,在Beckman/Coulter Avanti J25离心机中以1000g沉淀1周龄的悬浮细胞,并且弃去上清液。使用刮铲,将细胞置于一张作为1/16英寸厚层的圆形、无尘42号Whatman滤纸上。将承载植物材料的滤纸安置在27℃黑暗下的渗透性诱导培养基上持续1-2小时,随后轰击。就在轰击前,将滤纸从培养基取下并置于上一叠无菌滤纸上以允许愈伤组织表面部分地干燥。

[0211] 对于叶材料,每个平板用6μL金-DNA溶液以1800psi射击2次,并且对于BMS培养的细胞,以1100psi射击2次。为保持植物材料的位置,将一张消毒的金属网铺在样品的顶部。轰击后,将承载样品的滤纸转移到缺少甘露醇和山梨醇的M-N6-702培养基上并且在瞬时测定法之前于27℃在黑暗下孵育2日。

[0212] 使用例如Wang等人,1988(Transient expression of foreign genes in rice, wheat and soybean cells following particle bombardment(粒子轰击后外来基因在稻细胞、小麦细胞和大豆细胞中的瞬时表达).Plant Molecular Biology,11(4),433-439),Christou,1997(Rice transformation:bombardment(稻转化:轰击).Plant Mol Biol.35(1-2))中描述的技术,借助于微抛射体轰击其他单子叶植物实施瞬时转化。

[0213] 使用本领域方法,通过GUS染色、发光/荧光定量、RT-PCR和蛋白丰度(由特异性抗体检测)测定上述构建体(实施例5.1)中表达的基因的表达水平。GUS染色通过如下方式进行:将植物材料在GUS溶液[100mMNaHPO₄,10mM EDTA,0.05%Triton×100,0.025%X-Gluc溶液(溶解于DMSO中的5-溴-4-氯-3-吡啶基-β-D-葡萄糖醛酸),10%甲醇,pH 7.0]中在37℃孵育16-24小时。将植物组织真空浸润2次持续15分钟以辅助均匀染色。萤光素酶活性的分析如上文所述(实施例2和3.3)进行。

[0214] 与仅有组成型p-Ubi启动子缺少NEENA的报道基因构建体相比,所述NEENA分子均在这些测定法中介导强烈的基因表达增强作用。

[0215] 5.2.2使用原生质体的瞬时测定法

[0216] 通过遵循Sheen(1990)(Metabolic Repression of Transcription in Higher Plants(高等植物中转录的代谢抑制).The Plant Cell 2(10),1027-1038)开发的方案,实施原生质体的分离。将玉米籽苗在黑暗中于25℃培育10日并且在原生质体制备之前光照20小时。将叶的中间部分切成0.5mm的条(长度约6cm)并且在23℃在含有1% (w/v) 纤维素RS、0.1% (w/v) 离析酶R10(两者均来自日本西宫Yakult Honsha)、0.6M甘露醇、10mM Mes(pH 5.7)、1mM CaCl₂、1mM MgCl₂、10mMB-巯基乙醇和0.1%BSA(w/v)的酶溶液中孵育3小时,随后以80转/分钟轻轻振摇以释放原生质体。将原生质体通过100x g离心2分钟收集,在冷的0.6M甘露醇溶液中洗涤1次,离心并且重悬于冷的0.6M甘露醇(2x 10⁶/mL)中。

[0217] 将总体积100微升无菌水中的总计50微克质粒DNA添加至0.5mL玉米原生质体悬液

(1×10^6 个细胞/mL) 并且轻柔地混合。将 0.5 mL PEG 溶液 (40% PEG 4000, 100 mM CaNO_3 , 0.5 甘露醇) 添加并在 70°C 预温, 伴以轻柔振摇, 随后添加 4.5 mL MM 溶液 (0.6 M 甘露醇, 15 mM MgCl_2 和 0.1% MES)。该混合物在室温孵育 15 分钟。将原生质体通过 600 转/分钟沉淀 5 分钟并重悬于 1.0 mL 的 MMB 溶液 [0.6 M 甘露醇, 4 mM Mes (pH 5.7) 和雀麦花叶病毒 (BMV) 盐 (任选)] 中洗涤 2 次, 并且在 25°C 于黑暗下温育 48 小时。在终末洗涤步骤后, 将原生质体收集于 3 mL MMB 培养基中, 并且在 25°C 于黑暗下温育 48 小时。

[0218] 使用例如在 Hodges 等人, 1991 (Transformation and regeneration of rice protoplasts (稻原生质体的转化和再生). *Biotechnology in agriculture No.6, Rice Biotechnology*. International Rice Research Institute, ISBN: 0-85198-712-5) 或 Lee 等人, 1990 (Transient gene expression in wheat (*Triticum aestivum*) protoplasts (小麦 (*Triticum aestivum*) 原生质体中的瞬时基因表达). *Biotechnology in agriculture and forestry 13-Wheat*. Springer Verlag, ISBN-10: 3540518096) 中描述的技术, 实施其他单子叶植物原生质体的瞬时转化。

[0219] 使用本领域方法, 通过 GUS 染色、发光/荧光定量、RT-PCR 或蛋白丰度 (由特异性抗体检测) 测定上述构建体 (实施例 5.1) 中表达的基因的表达水平。GUS 染色通过如下方式进行: 将植物材料在 GUS 溶液 [100 mM NaH_2PO_4 , 10 mM EDTA, 0.05% Triton X-100, 0.025% X-Gluc 溶液 (溶解于 DMSO 中的 5-溴-4-氯-3-吲哚基- β -D-葡萄糖醛酸), 10% 甲醇, pH 7.0] 中在 37°C 孵育 16-24 小时。萤光素酶活性的分析如上文所述 (实施例 2 和 3.3) 进行。

[0220] 与仅有组成型 p-Ubi 启动子缺少 NEENA 的报道基因构建体相比, 所述 NEENA 分子在这些测定法中介导强烈的基因表达增强作用。

[0221] 5.2.3 单子叶作物植物的转化和再生

[0222] 也可以为了转化作物植物的目的而实施使用标准转化和再生技术的农杆菌介导的植物转化法 (Gelvin 和 Schilperoort, 1995, *Plant Molecular Biology Manual*, 第 2 版, Dordrecht: Kluwer Academic Publ. ISBN 0-7923-2731-4; Glick 和 Thompson (1993) *Methods in Plant Molecular Biology and Biotechnology*, Boca Raton: CRC Press, ISBN 0-8493-5164-2)。可以使用例如 US 5,591,616 中描述的技术实施玉米或其他单子叶植物的转化。使用粒子轰击、聚乙二醇介导的 DNA 摄取或借助碳酸硅纤维技术的植物转化法例如由 Freeling 和 Walbot (1993) “The maize handbook (玉米手册)” ISBN 3-540-97826-7, Springer Verlag New York) 描述。

[0223] 使用本领域方法, 通过 GUS 染色、发光或荧光定量、RT-PCR、蛋白丰度 (由特异性抗体检测) 测定上述构建体 (实施例 5.1) 中表达的基因的表达水平。GUS 染色通过如下方式进行: 将植物材料在 GUS 溶液 [100 mM NaH_2PO_4 , 10 mM EDTA, 0.05% Triton X-100, 0.025% X-Gluc 溶液 (溶解于 DMSO 中的 5-溴-4-氯-3-吲哚基- β -D-葡萄糖醛酸), 10% 甲醇, pH 7.0] 中在 37°C 孵育 16-24 小时。将植物组织真空浸润 2 次持续 15 分钟以辅助均匀染色。萤光素酶活性的分析如上文所述 (实施例 2 和 3.3) 进行。

[0224] 与仅有组成型 p-Ubi 启动子缺少 NEENA 的报道基因构建体相比, 所述 NEENA 分子在植物中介导强烈的基因表达增强作用。

[0225] 实施例 6: 在谷物植物中定量分析 NEENA 活性

[0226] 本实施例描述在谷物植物中对具有 SEQ ID NO 1 和 2 的 NEENA 序列的分析。

[0227] 6.1载体构建

[0228] 为在单子叶植物中定量地分析具有SEQ ID NO 1和2的NEENA序列,将携带表达盒的基于pUC的表达载体与接续胭脂碱合酶(NOS)转录终止子的萤火虫萤光素酶(LUC)基因(Promega, Madison, WI, USA)的编码序列组合,其中所述的表达盒包含缺少NEENA的来自玉米(Z.mais)的组成型单子叶启动子p-Ubi。使用Qiagen DNeasy Plant Mini试剂盒(Qiagen, Hilden, Germany)从拟南芥绿色组织提取基因组DNA。通过常规的聚合酶链反应(PCR)分离含有NEENA分子的基因组DNA片段。基于具有多种NEENA候选物的拟南芥基因组序列设计出引物。该反应包含2组引物(表10)并且使用以下引物,遵从由Phusion高保真DNA聚合酶(目录编号F-540L, New England Biolabs, Ipswich, MA, USA)概述的方案:

[0229] 表10: 引物序列

引物名称	序列	SEQ	PCR 产生
		ID NO	的 SEQ ID NO
[0230] NEENAc17_forIII	atatacgcggttagatctcgtgcgctcg	76	2
	NEENAc17_revIII	atattggcgcgcccttgatcaagcctgttcaca	77
NEENAc24_forIII	atatacgcggttaagaaatcctctctctctc	78	1
	NEENAc24_revIII	atatggcgcgccctgcacataacatatcaagatc	79

[0231] PCR期间的扩增和扩增产物的纯化如上述那样(实施例1.2)实施。用MluI (10U/微升)和AscI (10U/微升)限制性核酸内切酶进行DNA限制性消化后,消化的产物用Qiagen凝胶提取试剂盒(Qiagen, Hilden, Germany)纯化。

[0232] 使用AscI限制性位点,将NEENA PCR片段(见上文)分别克隆在萤火虫萤光素酶编码序列的上游。所述反应产生了具有p-Ubi启动子、萤火虫萤光素酶编码序列c-LUC和t-nos终止子的1个二元载体,以及紧邻萤火虫萤光素酶编码序列上游携带SEQ ID NO1和NO2的2个载体(表11),对于所述载体,给出与SEQ ID NO1的组合作为示例(SEQ ID NO80)。除了SEQ ID NO2变化外,所述载体中的核苷酸序列是相同的(表11)。表11中汇总所产生的载体,其中启动子分子具有前缀p-,编码序列具有前缀c-,并且终止子分子具有前缀t-。

[0233] 表11: 植物表达载体

植物表达载体	表达盒的组成	SEQ ID NO
	启动子::SEQ ID NO::报道基因::终止子	
[0234]	LJK309	p-Ubi::c-LUC::t-nos
	LJK327	p-Ubi::SEQ ID NO1::c-LUC::t-nos
	LJK326	p-Ubi::SEQ ID NO2::c-LUC::t-nos
		80

[0235] 使用所产生的载体在下文概述的实验中分析NEENA分子(实施例6.2)

[0236] 6.2转基因玉米植物的产生

[0237] 玉米萌发、繁殖、根癌农杆菌制备和接种如先前所述进行(W02006136596, US20090249514),除了使用的构建体LJK309,LJK326和LJK327(参考实施例6.1)各自含有受谷物遍在蛋白启动子p-Ubi驱动的突变AHAS基因,所述的突变AHAS基因介导耐受用于选择的咪唑啉酮除草剂。

[0238] 6.3NEENA序列在玉米植物中介导强烈的基因表达增强作用

[0239] 从产生的转基因植物中自叶和谷粒收集组织样品。组织样品如上文所述(参考实施例3.3)加工和分析。

[0240] 与仅有组成型p-Ubi启动子缺少NEENA的报道基因构建体相比,2个受检NEENA分子在叶中介导强烈的基因表达增强作用(图4a)。在玉米谷粒中检测到由NEENA(SEQ ID NO1至NO 2)介导的可比较的基因表达增强作用(图4b)。

[0241] 实施例7:在稻植物中定量分析NEENA活性

[0242] 本实施例描述在稻植物中对具有SEQ ID NO 1的NEENA序列的分析。

[0243] 7.1载体构建

[0244] 为在稻植物中定量地分析具有SEQ ID NO 1的NEENA序列,根据制造商(Invitrogen,Carlsbad,CA,USA)Gateway手册,使用位点特异性重组(LR反应),将pENTR/B载体LJK1和LJK4(比较实施例1.3)与重组位点上游携带组成型PR00239启动子的目的载体组合。所述反应产生了具有PR00239启动子、萤火虫萤光素酶编码序列c-LUC和t-nos终止子的1个二元载体,以及紧邻萤火虫萤光素酶编码序列上游携带SEQ ID NO1的1个载体(表12)。表12中汇总所产生的载体,其中启动子分子具有前缀p-,编码序列具有前缀c-,并且终止子分子具有前缀t-。

[0245] 表12:植物表达载体

植物表达载体	表达盒的组成	SEQ ID NO
	启动子::SEQ ID NO::报道基因::终止子	
[0246] CD30963	p-PRO0239 ::-::c-LUC::t-nos	
CD30964	p-PRO0239 ::SEQ ID NO1::c-LUC::t-nos	-

[0247] 使用所产生的载体在下文概述的实验中分析NEENA分子(实施例7.2)

[0248] 7.2转基因稻植物的产生

[0249] 使用含有相应表达载体的农杆菌来转化稻(*Oryza sativa*)植物。将粳稻栽培品种日本晴(Nipponbare)的成熟干燥种子脱壳。通过如下方式实施消毒:在70%乙醇中孵育1分钟,随后在0.2% HgCl_2 中孵育30分钟,随后用无菌蒸馏水洗涤6次15分钟。无菌的种子随后在含有2,4-D的培养基(愈伤组织诱导培养基)上萌发。在黑暗下孵育4周后,将生胚性盾片衍生的愈伤组织切下并在相同的培养基上增殖。2周后,将所述愈伤组织通过在同一种培养

基上传代培养另外2周进行繁殖或增殖。生胚性愈伤组织片在新鲜培养基上传代培养3日，随后共培育(以助长细胞分裂活性)。

[0250] 将含有相应表达载体的农杆菌菌株LBA4404用于共培育。将农杆菌接种在含有适宜抗生素的AB培养基上并在28℃培养3日。随后收集细菌并且在液体共培育培养基中悬浮至密度(OD₆₀₀)约1。随后将混悬液转移至培养皿内，并且将所述愈伤组织浸入该混悬液中15分钟。将所述愈伤组织随后在滤纸上蘸干并转移至固化的共培育培养基，并且在25℃于黑暗下孵育3日。共培育的愈伤组织在含2,4-D的培养基上在28℃于黑暗下在选择剂存在时培育4周。在此期间，迅速生长的抗性愈伤组织团发育。在转移这种材料至再生培养基并在光照下孵育后，生胚潜能释放并且幼苗在随后4至5周内发育。将幼苗从愈伤组织切下并且在含有植物生长素的培养基上孵育2至3周，其中从所述培养基转移幼苗至土壤。硬化的幼苗在温室中在高湿度和短日照下培育。

[0251] 对于一个构建体，产生大约35个独立的T0稻转化体。将原代转化体从组织培养箱转移至温室。在验证T-DNA插入物拷贝数的定量PCR分析后，仅保留显示所述选择剂抗性的单拷贝转基因植物用于收获T1种子。随后在移栽后3至5个月收获种子。该方法以超过50%的比例产生单基因座转化体(Aldemita和Hodges1996, Chan等1993, Hiei等1994)。

[0252] 7.3NEENA序列在稻植物中介导强烈的基因表达增强作用

[0253] 从产生的转基因植物中自叶和谷粒收集组织样品。组织样品如上文所述(参考实施例3.3)加工和分析。

[0254] 与仅有组成型p-PR0239启动子缺少NEENA的报道基因构建体相比，受检NEENA分子(SEQ ID NO 1)在叶中介导强烈的基因表达增强作用(图5a)。在稻种子中检测到由NEENA(SEQ ID NO1)介导的强烈的基因表达增强作用(图5b)。

[0255] 附图简述

[0256] 图1:在缺少NEENA(LJK132)和含有NEENA的构建体(LJK91-LJK133)瞬时转化的拟南芥叶原生质体中的萤光素酶报道基因表达分析，其中所述的构建体代表处于p-AtNit1启动子控制下的源自组成型表达的基因的推定的NEENA分子。归一化通过海肾萤光素酶的共转化和分析来进行，并且显示了相对于缺少NEENA的对照构建体(LJK132=1)的表达值。显示了相对于缺少NEENA的对照构建体(LJK134=1)的表达值。

[0257] 图2:独立转基因油籽油菜植物品系的显示为相对光照单位(RLU)并且针对每份样品的蛋白质含量(平均值,20个所分析的独立转基因植物的组织)归一化后的萤光素酶报道基因活性的柱状图，其中所述的独立转基因油籽油菜植物品系携带缺少NEENA(LJK138)或含有NEENA的报道基因构建体，所述的构建体代表在p-AtNit1启动子控制下组成型表达的基因(LJK139-LJK144)中的NEENA分子。A)叶组织,B)花,C)长角果。

[0258] 图3:独立转基因大豆植物品系的显示为相对光照单位(RLU)并且针对每份样品的蛋白质含量(平均值,10个所分析的独立转基因植物的组织)归一化后的萤光素酶报道基因活性的柱状图，其中所述的独立转基因大豆植物品系携带缺少NEENA(LJK138)或含有NEENA的报道基因构建体，所述的构建体代表在p-AtNit1启动子控制下组成型表达的基因(LJK139-LJK144)中的NEENA分子。A)叶组织,B)花,C)种子。

[0259] 图4:独立转基因玉米植物品系的显示为相对光照单位(RLU)(对数尺度)并且针对每份样品的蛋白质含量(平均值,15个所分析的独立转基因植物的组织)归一化后的萤光素

酶报道基因活性的柱状图,其中所述的独立转基因玉米植物品系携带缺少NEENA (LJK309) 或含有NEENA的报道基因构建体,所述的构建体代表在p-ZmUbi 启动子控制下组成型表达的基因 (LJK326-LJK327) 中的NEENA分子。A) 叶组织,B) 谷粒。

[0260] 图5:独立转基因稻植物品系的显示为相对光照单位 (RLU) 并且针对每份样品的蛋白质含量(平均值,15个所分析的独立转基因植物的组织) 归一化后的萤光素酶报道基因活性的柱状图,其中所述的独立转基因稻植物品系携带缺少NEENA (CD30963) 或含有NEENA的报道基因构建体,所述的构建体代表在组成型p-PRO239启动子控制下组成型表达的基因 (CD30964) 中的NEENA分子。A) 叶组织,B) 种子。

<110> 巴斯夫植物科学有限公司

<120> 用于增强植物中组成型基因表达的调节性核酸分子

<130> PF 62515

<160> 80

<170> PatentIn版本3.5

<210> 1

<211> 623

<212> DNA

<213> 欧芹 (*Petroselinum crispum*)

<400> 1

```

ttaagaaatc ctctctctc ctcttcattt tcaaggtaaa tctctctctc tctctctctc      60
tcgtttatfc attgttttaa ttaggtatgt attattgcta gtttgttaat ctgcttatct      120
tatgtatgcc ttatgtgaat atctttatct tgttcacctc atccgtttag aagctataaa      180
tttgttgatt tgactgigta tctacacgtg gttatgttta tatctaatca gatatgaatt      240
tcttcataat gttgcgttg tgtgtaccaa tccgaaatcg ttgatttttt tcatttaafc      300
gtgtagctaa ttgtacgtat acatatggat ctacgtatca attgttcacg tgtttgtgtt      360
tgtatgtata cagatctgaa aacatcactt ctctcatctg attgtgtgtt tacatacata      420
gatatagatc tgttatatca ttttttttat taattgtgta tatatatatg tgcatacatc      480
tggtatcatg gattgtgatt atttacaatg tttgtttatt taccgtatgta tatatgtaga      540
tctggacttt ttggagtgtg tgacttgatt gtatttgigt gtgtatatgt gigtctctgat      600
cttgataatg tatgtatgtg cag                                     623

```

[0001]

<210> 2

<211> 519

<212> DNA

<213> 拟南芥 (*Arabidopsis thaliana*)

<400> 2

```

ttagatctcg tgcgctcgtg cgacgttggt ttccggtacg ttatttccctg ttgattccct      60
ctctgtctct ctgattcac tgcacttctt gtttggattc ctttccgcgc afctctggat      120
ccgtgcgtta ttcatggctt cgtcgttttc agatctggtg cgtttctctt gttttctggt      180
atgagtggat gcgttttctt gtgattcgtt tgtttgtaat gctggatctg tatctgcgtc      240
gtgggaattc aaagtgatag tagttgatat tttttccaga tcaggcatgt tctcgtataa      300
tcagggtctaa tggttgatga ttctgcggaa ttatagatct aagatcttga ttgatttaga      360
tttgaggata tgaatgagat tcttaggtcc aaaaaggctt tgttatctct gctgctagat      420
agatgattat ccaattgcgt ttcttagtta tttttatgga ttcaaggaaat tgcgtgtaat      480
tgagagtttt actctgtttt gtgaacaggc ttgatcaaa                                     519

```

<210> 3

<211> 618

<212> DNA

<213> 拟南芥

<400> 3

```

ctttctctcg ttctcatctt tctctctaag gttagtcctt tgtgtagatc ctctcttcat      60
cgattcaatt tcatcttctt tggtttgatt taggaatttt gtcacgattt ggtgtttaga      120

```

	tgatggagtc ttgctgttga tctatgttta tagcaetttg tttecatggg tctgttggtt	180
	cttaccgttt gtgtgcate gatcigggtt tggtagatec ctctagata ataactgaa	240
	taigtcaaag ttttcttaca agtcgttgtt gtttcagtt ttgattgata ttgttatcca	300
	tgtagtagtc gatcatccaa aaactgggtc agtgtgattc tccgatcacc gccgctgatt	360
	ttataattat catcatgtat agatttcag gttttatttt gtctgtgtg tgtttttgtt	420
	ctttgttaaa ttttaacctct atgcatttga tgtacatgtt ttgtgttgg tctcagtttc	480
	tcttattctg tataagctaa gctaaagtct ccatggttag tgtctgttta gtattaaaat	540
	atttcaattt gctaatctct tatgtttttg tgatgatatt gcagagattt ttaggtgac	600
	aatcaaaaagt ttgacaaa	618
	<210> 4	
	<211> 772	
	<212> DNA	
	<213> 拟南芥	
	<400> 4	
	tcatcagatc ttcaaaaccc caaaatttgc aaatcaggta ataatcttct ccgtatgtat	60
	ctcttcgacc accgattgtc tgtttctttg ttgtcttttc tcatccgtt attgaatttt	120
	tgagttcgct gttaaatttt gattacatca atttttcttt tcatccgat tcagagagat	180
	tgaattgcaa attatgtgcc agatcttcgt gttctgtatc tgattccgga tcttggttat	240
	tagttttttt ttgtgtttg gaatgaattc aacaacgatg attttigatt ggtaattga	300
	tttttaatg ctatagtttt aaccttatag gaatgtaatt gaactggacc agctttactt	360
[0002]	gaattgtttg gatctagatt attcagacaa catcttttta actaaatiga aacatatttg	420
	tgtttgtgtt ctgtatgacc atgtgaattt ggtagcataa tttctttgtt gagaatattt	480
	ctgagatggg tttttgtagt ttaattttct tactcgatga accatctcgt gcaccttgca	540
	aaatcttttg ttcctatatct catgttttga taacaaacta atgagggtta cgttttttgg	600
	tgaagcgaac agttatgtaa catgtttgta actatcttca tgtttgaacc gattttgatt	660
	ggatgttttc atgttctatc atattctgag gttttgaacc ttctgaattc taacgtgagg	720
	ttatctttct tgtgcagacc tcatgttccc ggtagttact acctcaaat ca	772
	<210> 5	
	<211> 332	
	<212> DNA	
	<213> 拟南芥	
	<400> 5	
	aaatcgttct ttczaatctc tcaatcggga tcaaggtaaa ttctcgggtt ctatttcagt	60
	ttcttcgatt attttccatt ccatcgaatt ttctccctg ttctttgagt ttgattcggg	120
	taatctttga tccgagataa atatctgagt ttcatcgtca gatccatct atattctgaa	180
	ctagggttcg tagatttcac agatttgtat atttagcgtt cgattctttg aatttgatcc	240
	gatatgtgat ttatattcag atctggtggg agctttttta atcggatttg agattatcca	300
	agcgattaat ctcatgtttt tgtgaattac ag	332
	<210> 6	
	<211> 501	
	<212> DNA	
	<213> 拟南芥	

	<400> 6	
	ccccatctctc atctccactc ttctcatctc gtaaggtaag ctcccgagtt tccatggctt	60
	cgcgatgaact attcttgcctc tctcccttcc cactctctta ttaccttcc gtctcttcgt	120
	ttctcttcca atccaatttt cgtttgattc cgtttctctt tcttctgta atacitgat	180
	cctaatecgg tgagtgaat tctatgaaac acagatctgg aitttttata cgatcagatt	240
	cgtaaaaatct ctgttttgcg atctttctct gcgttagatc tctgttttcc ctagtgtga	300
	ttctgtttga tctttaaact ctctcact ctgttttcaa taaatacgac aaaaaaact	360
	gcaaaagttt ttactttttt taacgattgg tggcttttac catatgacaa gaagtacatt	420
	ttaggattgt gtttggttta ctgactgatt ttggttttgt ttgtggttca aatgtagtaa	480
	aaaaacaaaa aacaaacaaa a	501
[0003]	<210> 7	
	<211> 535	
	<212> DNA	
	<213> 拟南芥	
	<400> 7	
	agccgcaga ctcttttcag attcttactt gcaggitaga tattttctct ctttagtgc	60
	tccgatcttc atcttcttat gattattgta gctgtttagg gtttagattc ttagttttag	120
	ctctatatgt actgtgatta tgccttattc ttgtctgtt ttatactgt ttgattctc	180
	tagctttaga tccgtttact cgtcgataca taigtctct attgagctg atgtataatc	240
	ctctgattaa ttgatagcgt ttagttttga tctctcttc gcattgtttt tctcatgtcg	300
	atctgtatct gctctggita tagttgattc tgaigtattt ggttgggat gttctttaga	360
	ttgatatac ctgttgcctc gttgtttgat atgatagctc aactgggat atgtgtttt	420
	gttctagtg atctgtgttt gattatatgt ttgacgtttt ggttgttga tagttgatgg	480
	ttgatgtatt ttgttgatt ctgatgttc gattttgtt ttgttttga cagct	535
	<210> 8	
	<211> 346	
	<212> DNA	
	<213> 拟南芥	
	<400> 8	
	atctttacct caacaacgag atctgttcc acggtaccat ttctatccga tctctgttct	60
	tctatgtttg ttacttggat ccgttgattt tactttcttc ttatcatttg tgtgtggcta	120
	tttgggtgaa ttgaigcgt cttaaagctt ttctaggttc aagccttata tgttctctt	180
	aagetaatca taggtcttag acttgtagg aagattcgaa tctagtacaa agttagcaac	240
	tttagcatac gtaaactcgtt agtagtaatt tctagattga tagaggttcg agacctaaa	300
	agggtatgaa tatttgggtg gttttgttta ttgaaagaa aggagg	346
	<210> 9	
	<211> 316	
	<212> DNA	
	<213> 拟南芥	
	<400> 9	
	cgattcgggt caaggttatt gactttctca ttcaggtatc atcttcttcc tctctccata	60
	ctttttgatt cacttcttca ttctcttccc ggttctgatt atatttgttt ggatcggata	120
	attgtatggg cttaaagctta ttctctgttt aaatctgttg atttgggttt aagagtgttg	180

	tataactgga ttggtgctca tacatctgga gatttatggt tgatgatgtt gtatgatgtg	240
	tttttgggtt gaggatgatt tattctcttg gtcttacttt tctctttttg tgaatttaag	300
	cagigaagct tgaate	316
	<210> 10	
	<211> 165	
	<212> DNA	
	<213> 拟南芥	
	<400> 10	
	ttcgaagaac caaaaccaaa aacgtgaatt cggagitatg attttggatc ttcttccgt	60
	ctttatataa aattttctgg gttatgtttt tcttctcttt aattcatctg accattggat	120
	tggttttcca ttgcagagtc ttcaattgga atttggatcg tggga	165
	<210> 11	
	<211> 155	
	<212> DNA	
	<213> 拟南芥	
	<400> 11	
	atcgtcaaaa cegagtttgg aggatctgtc ttgaggtaaa aaagctcaat cctttatcgg	60
	aaattttcaa ccttttcttt tcttttggtt gttgcgttgc tgaatctatc tagggtttgg	120
	ttttttggat ctaggagttt ttttttggtt gagaa	155
	<210> 12	
	<211> 116	
	<212> DNA	
	<213> 拟南芥	
[0004]	<400> 12	
	gaccccaact tctctcgaca cttatagagt ttccaggtcca ttctctgatt cctgtttttt	60
	cattttctaa ttctctatct acgggttttt actaatcttc tattgacctt ccagtc	116
	<210> 13	
	<211> 369	
	<212> DNA	
	<213> 拟南芥	
	<400> 13	
	aaatcaatct ctcaaatctc tcaaccgtga tcaaggtaga ttcttgagtt cttattgtat	60
	ttcttcgatt tgtttcgttc gatcgcaatt taggetctgt tctttgattt tgatectgtt	120
	aatctctgat cggaggcaaa ttacatagtt tcatcggttag atctcttctt atttctgat	180
	tagggttctg atttttccga gatctgttta ttttcttggt gtttccctgt atttgatccg	240
	atttgtigaa agaatttgtg tgttctcgat tatttacgtt ttgactctgt attttatct	300
	agatttgggt ttagtttctt gtttgtgaga tcgaatttgt cgattaatct cgtttttct	360
	gattaacag	369
	<210> 14	
	<211> 235	
	<212> DNA	
	<213> 拟南芥	
	<400> 14	
	ctcatcgttg gaccttagaa gccgtcgga caaggttcgt ctttctctca tccctttct	60
	cttcttttgc gtgaattcac cagaaaaagt atatctacct tagcgagtct gtctagtta	120
	gcgaatctag gataatctcc ggaagacct ttigaatcat cactgttttt gatctctgtt	180

	tctttctcta attgtgtgtg tgtattactg ttgatgtaga agaagaagaa gaaga	235
	<210> 15	
	<211> 204	
	<212> DNA	
	<213> 拟南芥	
	<400> 15	
	ccitgcgaaa agcattttcg atettactct taggtagtct atgattctcc atttgatccg	60
	tttatacttg ttattggcat ttcatcatcg tctgggttgt tcgacttact ctaattttgg	120
	tttaaaacgg atttaaagtt tgtttttttg tgaataagca attaatctat tgttactggt	180
	tttgaatcgt ttcaggataa aaaa	204
	<210> 16	
	<211> 674	
	<212> DNA	
	<213> 拟南芥	
	<400> 16	
	ttttgacgcc gccgcttctt cttcttctct gcaggtaact tctgggtaac ttagggaattg	60
	cgtcgtttgtc ggaaactgca aaagctttgt cgttaggttt cgatttgttg ttttttttt	120
	tgctctcgtg gcttttgaat cttaggtttg gtttagattcg ttaataagat aatggaaaag	180
	ttatgaaatg ggttgatttt tattaccgga tactgccgct gcattctcat tcttttagtt	240
	ttattgtatc aatggettcg ttgtttttat aatcacatta ggctctctc aagttcagtg	300
	ttatggtgat gtatatttct tgattcattt gcaattctct agacattatt tgacattgca	360
[0005]	tttacaatgt attcctgtgt ttgttttagt ttacacctga atgttgatgt ttaagacaca	420
	aatcaactat acacttgggt gtctctcat ttgtcatatg cttatggatc tctattgttt	480
	ctatttgaca ttgctgttca tatgtttctt tgaactatga ttataacaat cgatgcttta	540
	ggtgcatctc ttgatacata aatggtttta atttcatgat gtaattgaat caccgatatt	600
	gaaccctaac taaagacatg gaatttcttt ggtttttccc ctttcagcag gtaagtcaca	660
	caggttaactg aaag	674
	<210> 17	
	<211> 603	
	<212> DNA	
	<213> 拟南芥	
	<400> 17	
	gatacgttaa cttcaccaat ccccaagacg gttagtctcc ttaattttat ttggacttct	60
	tettacaatt tccaattttt agggtttgtc cgatagtacc agtactggat ttgaatgagg	120
	tgtgttttaa atgttgatac ttatgggtgt ttgtatgtat gcaatgttgt tgggtctgtt	180
	ggaaagctta tccccgtctt aatcaatctt tttaatgcta ctctctctt ctaatggta	240
	tttatctcgt gtctttgttg cagctgcaag agcgatccaa gaactgaatt acatacctct	300
	ttatggaaaa cctattaggg ttatgtattc tcatcgtgat cctagtgttc gccgaagtgg	360
	cgttggcaac atttttatta agtaecgttg tctctatcct tctgacattt gaggttctgt	420
	aagacgagat gattgtttgc aaaaaatttg tatcttgtat tattcattga aaagtctggc	480
	aaaagtagtt gatttcttta gttatgtaag tttttttatt gggtgttctt tctctcttct	540
	tgatgctaaa cctggctata atatgcagaa ttggacgag tccatcgacc acaaagcaet	600

	gca	603
<210>	18	
<211>	919	
<212>	DNA	
<213>	拟南芥	
<400>	18	
	ctctcttccg tctcgagtcg ctgagatcca tcaggttcgt ttctattcga atctattctc	60
	ttgaatccgt atttgtticc ttctctctct agatagattt catttttctg gaaacatiga	120
	gatttccatt gatttgaatg ttagtggttt agatcggata atttttgtaa ttgactctca	180
	aaaaacgaat cataactatc ttaaaacttt gtttacgcat tgttattgaa ggctaanataa	240
	aaaatgctca attgttgctc gaaaggaagg tgtctcgaaa acatgtgaat ttatittttg	300
	agcctcattg ttctcactg caacaatagt gaatgattca attgaaaaga acttttggtt	360
	aagccatttt tctctttaat agattatgtg tctcaactca attaggaact gtgtgttttt	420
	ttgtttcctt gatgttaact tgtgagtga tttatttggg tagtttactt gtatatatat	480
	ttggltggtg gtccgtgagg gaaalgagtg agagggagct ctgttcatac attctatggt	540
	tttaaatgat tagtattttt ctctcatctc atacttgctt tttgtctag gaatttcgat	600
	ttttgggact aatttatcga gtatatataa ttggataagt gagtttggta ttaatatataa	660
	gagcaattag gttggttate ttgccaacaa gtagaaccct gtcttttctc tagtcattta	720
	agaagttaag ctttaagattc aggcctctaa tctggttatg atagtgaaga ggaatctgaa	780
	atttacctat acgacctgtt gaaaaggacc tctgggtgtt gttactatgc gttctctgac	840
[0006]	aatatagaca tagttggatc atattcttct ctgacgcgc tgtttcatit tttttcaaaa	900
	tcagtaaaag gttctgcaaa	919
<210>	19	
<211>	601	
<212>	DNA	
<213>	拟南芥	
<400>	19	
	tttcttctct cctctccgat tcttctctcc tcaggttcgt ctccgacttt ctatgtcttc	60
	cttacgtctc tcttatttag ctccctttcc ttcttctcc acgttattat tactacttcg	120
	cttttagtgt gattcgttcc attctcgttt ttttatattc ctcgatctgt ttgcctattt	180
	gttgagatct attcgcctatg tgagttcatt tgactcagat ctggatattt cgtgttggtc	240
	gatttataga tctggtttct ggatctgttt acgatctatc gtcacttttc ctttgaaaat	300
	gattgggtgt tctggttctg tattcggtta gatctaaaagt ttttgatcga tgaatgtcgc	360
	atgtgttttt atctgaaaagt ttctgattac agtatcaagt ggtggtagta gtagtagtag	420
	actcaaaaaa ctgcacaaaac tttttataca cgtgaattgt gattgcctta cggttttctt	480
	ggagtttgtt aattaaatca tttaatatta agaagttaat gaattagag aacgttattt	540
	tatactatga ttttgatttt gatttggttt gtgtgtttta atgcagtaaa agaaaatcaa	600
	a	601
<210>	20	
<211>	36	
<212>	DNA	
<213>	人工的	

	<220>		
	<223>	合成的序列	
	<400>	20	
		tttatggtag cagccgcaag acicctttca gattct	36
	<210>	21	
	<211>	39	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	21	
		aaattccatg gtagctgtca aaacaaaaac aaaaatcga	39
	<210>	22	
	<211>	37	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	22	
		aaaaaggtag ctggaagaac caaaaccaaa aacgtga	37
	<210>	23	
	<211>	44	
	<212>	DNA	
	<213>	人工的	
[0007]	<220>		
	<223>	合成的序列	
	<400>	23	
		tttttccatg gttatttata caaaatecca cgatccaaat tcca	44
	<210>	24	
	<211>	33	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	24	
		tttttggtag cgatccctac ttctctegac act	33
	<210>	25	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	25	
		ttttaccatg gtgactggag gatcaataga agat	34
	<210>	26	
	<211>	37	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	26	

	tttttggtag ctttctctcg ttctcatctt tctctct	37
	<210> 27	
	<211> 36	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 27	
	taatagatat ctttgtcaaa cttttgattg tcacct	36
	<210> 28	
	<211> 34	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 28	
	tataaggtag caaatcaatc tctcaaatct ctca	34
	<210> 29	
	<211> 35	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 29	
	tttatecatg gtctgttaat cagaaaaacc gagat	35
[0008]	<210> 30	
	<211> 34	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 30	
	tatatggtag caaatcgttc ttcaaatct ctca	34
	<210> 31	
	<211> 35	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 31	
	ttataccatg gtctgtaatt cacaaaaaac tgaga	35
	<210> 32	
	<211> 32	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 32	
	tttttggtag ctcatcggtg gagcttagaa gc	32
	<210> 33	

[0009]	<211>	35	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	33	
		tttttccatg gttttcttct tctttcttcta catca	35
	<210>	34	
	<211>	38	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	34	
		tatatgggtac caaagcattt tcgatcttac tcttaggt	38
	<210>	35	
	<211>	33	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	35	
		tttttccatg gttttttatc ctgaaacgat tca	33
	<210>	36	
	<211>	36	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	36	
		tttttgggtac cttttgacgc cgccgcttct tcttct	36
	<210>	37	
	<211>	38	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	37	
		tttttccatg gtctttcagt taccigtgtg acttacct	38
	<210>	38	
	<211>	33	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	38	
		tttaagggtac ccactctctca tctccactct tct	33
	<210>	39	
	<211>	37	
	<212>	DNA	
	<213>	人工的	

	<220>		
	<223>	合成的序列	
	<400>	39	
		tttttgatat cttttgtttg ttttttgttt ttttact	37
	<210>	40	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	40	
		ttataggtac caagtgaatc gtcaaaaccg agtt	34
	<210>	41	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	41	
		tttttccatg gttctcaacc aaaaaaaac tcct	34
	<210>	42	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
[0010]	<220>		
	<223>	合成的序列	
	<400>	42	
		tttttggtag cacgattcgg gtcaaggtaa ttga	34
	<210>	43	
	<211>	40	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	43	
		tttttccatg gtgattcaag cttcactgcc taaattcaca	40
	<210>	44	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	44	
		tttttggtag cttagatctc gtgccgtcgt gcca	34
	<210>	45	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	45	

	tttttccatg gtttgatcaa gctgtttcac a	31
	<210> 46	
	<211> 34	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 46	
	aaaaaggtag ctcacagat cttcaaaacc ccaa	34
	<210> 47	
	<211> 38	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 47	
	aaaaaccatg gtgatttgag ggtagtacta accgggaa	38
	<210> 48	
	<211> 36	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 48	
	tttttaggtac catagcttaa cttcaccaat ccccaa	36
[0011]	<210> 49	
	<211> 42	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 49	
	tttttccatg gtttaattaat gcagtgtttt gtggfcatg ga	42
	<210> 50	
	<211> 33	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 50	
	tttttccgg gatctttacc tcaacaacga gat	33
	<210> 51	
	<211> 44	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 51	
	tttttccatg gttatcctcc tttcttteta ataaacaaaa ccca	44
	<210> 52	

[0012]	<211>	36	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	52	
		ttttttggtac ctctcttccg tctcgagtcg ctgaga	36
	<210>	53	
	<211>	31	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	53	
		tttttccatg gtttcagac cttttactga t	31
	<210>	54	
	<211>	37	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	54	
		tttttggtac ctctcttccf cctctccgat tcttctt	37
	<210>	55	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	55	
		tttttccaig gttattgatt ttcttttact gcac	34
	<210>	56	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	56	
		ttttttggta ccttaagaaa tctctcttct tctt	34
	<210>	57	
	<211>	34	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	57	
		ttttttccat ggctcgcaca tacataacat atca	34
	<210>	58	
	<211>	51	
	<212>	DNA	
	<213>	人工的	

[0013]	<220>		
	<223>	合成的序列	
	<400>	58	
		ggggacaact ttgtatagaa aagttgicga gaccagatgt ttacacttg a	51
	<210>	59	
	<211>	51	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	59	
		ggggactgct tttttgtaca aacttggaca ctacagact tgagagaagc a	51
	<210>	60	
	<211>	73	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	60	
		ggtacccggg gatcctctag catatgetcg aggcggccgc agatatacaga tctggtcgac	60
		ggcatgcaag ctt	73
	<210>	61	
	<211>	12176	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	载体序列	
	<400>	61	
		gacatacaaa tggacgaacg gataaacctt ttcacgccc tttaaaatate cgattatctt	60
		aataaacgct cttttctctt aggtttaccc gccaatatat cctgtcaaac actgatagtt	120
		taaaactgaag gcggaacg acaatcagat ctataggaa acagctatga ccatgattac	180
		gccaaagctat cgattacgcc aagctatcaa ctttgtatag aaaagttgca tcaagatett	240
		ggtgatgtag caagagctaa gttgtacttc gattcggttg gacattactc gagaccagat	300
		gttttacact tgaccgtaaa tgagcaccgc aagaaaccgg tcacattcat ttcgaaggig	360
		gagaaagcgg aagatgactc aaacaagtaa tcggttgta ttcgtcagtt catgtcactc	420
		ctatgaagga gtcaagttca aaatgttatg ttgagtttca aacttttatg ctaaactttt	480
		tttttttatt ttctttaafa atggaagaga accaattctc ttgtatctaa agattateca	540
		tctatcatcc aatttgagtg ttcaattctg gatgttgtgt taccctacat tctacaacca	600
		tgtagccaat tattatgaat ctggtttga ttacagttgt gttcttttct tttttttctt	660
		tgcataattg catttagaat gttaataat taagtacttg tatttccaea tacattagtt	720
		ccaagaatat acatatatta atttattttt cttaaaaatg ttttggaatg actaatattg	780
		acaacgaaaa tagaagctat gctaaaccat tacgtatatg tgacttcaca tgttgttgtt	840
		ttacattccc tatatatatg gatggctgtc acaatcagaa acgtgatcga aaaaagacaa	900
		acagtgtttg caaaaaaaga ctatttcggtt tcattgacaa ttgtgtttta ttgtaaaaga	960
		aaagtggcaa agtggaaatt gagttccctg aagtaagaaa gatgaaataa aagacttgag	1020

[0014]

tggtgtttt tttttttat ctgaaagcig caatgaaata ttcttaccac gcccggttga	1080
ttattaatgg gggtttggtt ttcttgaigc gaactaatgg gtatataag aaactataca	1140
atccatgita attcaaaaat ttgatttct ctgttaggaa tatgatttac tatatgagac	1200
ttttttttcg ccaataatag taaatccaaa gatatttgac cggaccacaaa cacattgatc	1260
tatttttttag ttattttaat ccagtttctc tgagataatt cattaaggaa aacttagtat	1320
taaccacatc taagattaaa taggagccaa acacacatti caaatattaa ataacataaa	1380
atggatttaa aaaaatctata cgtcaaatit tatttatgac atttcttatt taaatttata	1440
tttaatgaaa tacagctaag acaaaccaaa aaaaaatac ttcttaagtg gtccaaaaca	1500
tcaattccgt tcaatattat taggtagaat cgtacgacca aaaaaaggtg ggtaaatagc	1560
aattagaaac atatctataa calagtatat attattacct attatgagga atcaaaaagc	1620
atcaaatatg gatttaagga atccataaaa gaataaatc tacgggaaaa aaaaatggaat	1680
aaattctttt aagtttttta ttgtttttt atttggtagt tctccatttt gttttatttc	1740
gitttgattt atttggtcca aaatcttctt aaaccacgtt tgtaattctt aaacgggggtt	1800
ttcacttctt ttttatatt agacataaag catcggttgg tttaataaat caatagattt	1860
tatttttctt ctcaattatt agtaggttg atgtgaactt tacaacaaaa acaaaaacaa	1920
atcaatgcag agaaaagaaa ccacgtgggc tagtccacc ttgtttcatt tccaccacag	1980
gttcgatctt cgttaccgtc tcaatagga aaataaacgt gaccacaaa aaaaaacaaa	2040
aaaaagctta tatattgctt ctctcaagtc tctgagtgtc atgaacacaa gtaaaaaaca	2100
aagactcgag tcaagtttgt acaaaaaagc aggttggtac cttaagaaat cctctcttct	2160
ctcttctatt ttcaaggtaa atctctctct ctctctctct ctctgttatt cttgttttta	2220
attaggatg tattattgct agtttgtaa tctgcttate ttatgtatgc cttatgtgaa	2280
tatctttate ttgttctct catccgttta gaagctataa atttgittgat ttgactgtgt	2340
atctacacgt ggttatgttt atatctaatc agataagaat ttcttctat ttgtgcgttt	2400
gtgtgtacca atccgaaatc gttgattttt ttcatttaat cgtgtagcta attgtacgta	2460
tacatatgga tctacgtatc aattgttcat ctgtttgtgt ttgtatgtat acagatctga	2520
aaacatcact tctctctctt gattgtgttg ttacatacat agatatagat ctgtttatct	2580
atttttttta ttaatttgtt atatataat gtgcatagat ctggattaca tgattgtgat	2640
tatttacatg attttgttat ttacgtatgt atatatgtag atctggactt ttggagtgtg	2700
ttgacttgat tttatttgtg ttgttatatg ttgtttctga tcttgataat ttatgtatgt	2760
gcagaccaig gaagacgcca aaaaacataa gaaaggcccg gcgccattct atccgttggc	2820
agatggaacc gctggagagc aactgcataa ggetatgaag agatacgccc tggttcctgg	2880
aaacattgct ttacagatg cacatatcga ggtagacatc acttacgtg agtacttcca	2940
aatgtccgtt cggttggcag aagctatgaa acgatatggg ctgaatacaa atcacagaat	3000
cgtcgtatgc agtgaanaat ctcttcaatt ctttatgccg gtgttggggc cgttatttat	3060
cggagtgtga gttagcggcg cgaacgacat ttataatgaa cgtgaattgc tcaacagtat	3120
gggcatttgc cagcctaccg ttgtgttctt ttccaaaaag gggttgcaaa aaattttgaa	3180
cgtgcacaaa aagctcccaa tcatccaaa aattattatc atggattcta aaacggatta	3240
ccagggattt cagtcatgtt acacgttctt cacatctcat ctacctcccg gtttaaatga	3300

[0015]

atacgatttt gtgccagagt ccttcgatag ggacaagaca attgcactga tcatgaactc 3360
 ctctggatct actggctcgc ctaaaggigt cgcctcgcct catagaactg cctgcgtgag 3420
 attctcgcat gccagagatc ctatttttgg caatcaaate attccggata cfcgattttt 3480
 aagtgllgtt ccattccalc acggtttggg aalgtttact acacleggat attgatalg 3540
 tggatttcga gtcgtcttaa tgfatagatt tgaagaagag ctgtttctga ggagcctica 3600
 ggattacaag attcaaatg cgcctcgtgt gccaaccceta ttctcttct tgcocaaaag 3660
 cactctgatt gacaaatag atttatctaa ttacacgaa attgcttctg gtggcgtctc 3720
 cctctctaag gaagtcgggg aagcgttgc caagagttc catctgccag gtatcaggca 3780
 aggataatgg ctactgaga ctacatcagc tattctgatt acacccgagg gggatgataa 3840
 accgggcgcg gtcggtaaag ttgttccatt tttgaagcg aaggttctgg atctggatac 3900
 cgggaaaacg ctgggcgtta atcaaagagg cgaactgtgt gtgagaggtc ctatgattat 3960
 gtccggttat gtaacaate cggaaagcgc caacgccttg attgacaagg atggatggct 4020
 acattctgga gacatagctt actgggacga agacgaacac ttcttcacgc ttgaccgcct 4080
 gaagtctcgc attaagtaca aagcctatca ggtggctccc gctgaattgg aatccatctt 4140
 gctccaacac cccaacatct tcgacgcagg tctcgcaggt cttcccgacg atgacgccgg 4200
 tgaacttccc gccgccttg ttgttttga gcacggaaag acgatgacgg aaaaagagat 4260
 cgtggattac gtcgccagtc aagtaacaac cgcgaaaaag ttgcgcggag gatttgtgtt 4320
 tgtggacgan gtaccgaag gtcttaccgg aanaactcgc gcaagaaaaa tcagagagat 4380
 cctcataaag gccaaagagg gcggaaagat cgcctgttaa ctcgagcata tgggctcgaa 4440
 tttcccgcgt cgttcaaca ttggcaata aagtttctta agattgaatc ctgttgcggg 4500
 tcttgcatg attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat 4560
 gtaatgcatg acgttattta tgagatgggt tttatgatt agagtcccgc aattatacat 4620
 ttaatacgcg atagaaaaa aaatalagcg cgcuaactag galaanltat cgcgcgcggt 4680
 gtcatctatg ttactagatc gggaattcaa gcttggcgta atcatggacc cagctttctt 4740
 gtacaaagtg gggatcccg ggatcctcta gcatatgctc gaggcggcgc cagatatacg 4800
 atctggtcga cggcatgcaa gcttggcgta atcatggcaa ctttattata catagttgat 4860
 aattcactgg ccgataatt cactggcgtt cgttttaca cgaactcagga tccgttcaaa 4920
 cactgatagt ttaaaactga ggccgggaaac gacaactga tcatgagcgg agaattaagg 4980
 gagtcacgtt atgaccccg ccgatgacgc gggacaagcc gttttacgtt tggaaactgac 5040
 agaaccgcaa cgttgaagga gccactcagc cgcgggttcc tggagittaa tgagctaagc 5100
 acatacgtca gaaaccatta ttgcgcgttc aaaagtgcgc taaggtcaat atcagctagc 5160
 aaatatttct tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt 5220
 agagtctcat attcactctc aatccaaata atctgcaccg gatctggatc gtttgcgatg 5280
 attgaacaag atggattgca cgcaggttct ccggccgctt gggtagagag gctattcggc 5340
 tatgactggg cacaacagac aatcggtcgc tctgatgccg ccgtgttccg gctgicagcg 5400
 caggggcgcc cggttctttt tgtcaagacc gacctgtccg gtgccctgaa tgaactgcag 5460
 gacgagcgag cgcggtatc gtggctggcc acgacggggc ttcccttgcgc agctgtgctc 5520
 gacgttgtca ctgaagcggg aagggaactg ctgctattgg gegaagtgc ggggcaggat 5580

	ctcctgtcat etcaccttgc tectgccgag aaagtatcca tcatggetga tgcaatgcgg	5640
	cggttcgata cgcttgatcc ggctaccitgc ccattcgacc accaagcgaa acatcgcatc	5700
	gagcgagcac gtactcggat ggaagccggt ctigtcgatc aggatgatct ggacgaagag	5760
	caicaggggc tgcgccagc cgaactgllc gccaggctca aggcgcgcac gcccgacggc	5820
	gaggatctcg tcttgaccca tggcgatgcc tgcctgccga atatcatggt ggaanaatggc	5880
	cgcctttctg gattcatcga ctgtggccgg ctgggtgtgg cggaccgcta tcaggacata	5940
	gcgttggeta cccgtgatat tgcctgaagag cttggcggcg aatgggetga ccgcttcctc	6000
	gtgctttacg gtatcgccgc tcccgattcg cagcgcacgc cttctatcg cttcttgac	6060
	gagttcttct gagcgggacc caagctctag atcttgcitg gtccggatat tttcgtggag	6120
	ttcccgcac agaccggat gatecccgat cgttcaaca ttggcaafa aagtttctta	6180
	agattgaatc ctgttgccgg tcttgcatg attatcatat aatttctggt gaattacgtt	6240
	aagcatgtaa taattaacat gtaatgcacg acgttattta tgagatgggt tttatgatt	6300
	agagtccgc aattatacat ttaatacgcg atagaaaaca aaatatagcg cgcaaacatg	6360
	gataaaitat cgcgcgcgtt gtcactatg ttaactagatc ggccctctg tcaagctctg	6420
	cttgtaata attgtcatta gattgttttt atgcatagat gcactcgaaa tcagccaatt	6480
	ttagacaagt atcaaacgga tgttaaitca gtacattaaa gacgtccgca atgtgttatt	6540
	aagttgtcta agcgtcaatt tgtttacacc acaatatate ctgccaccag ccagccaaca	6600
	gtccccgcac cggcagctcg gcacaaaate accacgcgtt accaccacgc cggccggccg	6660
[0016]	catggtgttg aecgtgticg ccggcattgc cgagttcgag cgttccttaa tcatcgaccg	6720
	caccgccgagc gggcgcgagg ccgccaagge ccgagggcgtg aagtttggcc cccgccctac	6780
	cctcaccccg gcacagatcg cgcacgcccg cgagctgacg gaccaggaag gccgcaccgt	6840
	gaaagagggc getgcaicgc ttggcgtgca tgcctcgacc ctgtaccgcg cactigagcg	6900
	cagcgaggaa gtagcgcaca ccgagggcag gggcgcggtt gccttcggtg aggacgcatt	6960
	gaccgaggcc gacgccctgg cggccgccga gaatgaacgc caagaggaa aagcatgaaa	7020
	ccgcaccagg aegggcagga cgaaccgtt ttattaccg aagagatcga ggcggagatg	7080
	atcgccggcg ggtacgtgtt cgagccgcc cgcacgtct caaccgtcg gctgcatgaa	7140
	atcctggccg gtttgtctga tgcgaagctg gcggccctgg ccgccagctt ggccgctgaa	7200
	gaaaccgagc gccgccgtct aaaaaggta tgtgtatttg agtaaaacag cttgcgtcat	7260
	gcggtcgctg cgtatatgat gcgatgagta aataaacaac tacgcaaggg gaacgcaiga	7320
	aggttatcgc tgtacttaac cagaaaggcg ggicaggcaa gacgaccate gcaaccatc	7380
	tagcccgccg cctgcaactc gccggggccg atgttctgtt agtcgatcc gatccccagg	7440
	gcagtgcggc cgatiggggc gccgtgcggg aagatcaacc gctaaccgtt gtcggcatcg	7500
	accgcgcgac gattgacgc gacgtgaagg ccattcgccg gcgcgacttc gtagtgatcg	7560
	acggagcgcc ccaggcgcg gacttgctg tctcccgat caaggcagcc gacttcgtgc	7620
	tgattccggt gcagccaagc ctttacgaca tatgggccac cgcgcacctg gtggagctgg	7680
	ttaagcagcg cattgaggtc acggatggaa ggctacaagc ggctttgtc gtgtcgcggg	7740
	cgatcaaaag cagcgcgcatc ggcggtgagg ttgccgagc gctggccggg tacgagctgc	7800
	ccattcttga gtcccgatc acgcagcgcg tgagctaccc aggcactgcc gccgcggca	7860

[0017]

caaccgttct tgaatcagaa cccgagggcg acgtgcccg cgaggteccag gcgctggccg	7920
ctgaaattaa atcaaaactc atttgagtta atgaggtaaa gaaaaaatga gcaaaagcac	7980
aaacacgcta agtgccggcc gtccgagcgc acgcagcagc aaggctgcaa cgttggccag	8040
cttggcagac acgcagacca tgaagcgggt caactttcag ttgccggcgg aggatcacac	8100
caagctgaag atgtacgctg iacgccaagg caagaccatt accgagctgc tatctgaata	8160
catcgccag ctaccagagt aaatgagcaa atgaataaat gagtagatga attttagcgg	8220
ctaaaggagg cggcatggaa aatcaagaac aaccaggcac cgacgccgig gaatgcccc	8280
tgtgtggagg aacgggcgggt tggccaggcg taagcggctg gttgtctgc cgccctgca	8340
atggcactgg aacccccaa cccgaggaat cggcgtgagc ggtcgcaaac catccggccc	8400
ggtacaaatc ggcgccggcg tgggtgatga cctggaggag aagttgaagg ccgcgcaggc	8460
cgcacagcgg caacgcctcg aggcagaagc acgccccggt gaatcgtggc aagcggccgc	8520
tgatcgaatc cgcaagaat cccggcaacc gccggcagcc ggtgcgccgt cgattaggaa	8580
gcgcaccaag ggagcagcgc aaccagattt ttctgtccg atgctctatg acgtgggcac	8640
cgcgcatagt cgcagcatca tggacgtggc cgttttcgt cgtcgaagc gtgaccgacg	8700
agctggcgag gtgatecgt acgagctcc agacgggcac gtagaggttt ccgcaggccc	8760
ggccggcatg gccagtgtgt gggattacga cctggtactg atggcggttt cccatctaac	8820
cgaatccatg aaccgatacc ggaagggaa gggagacaag cccggccgcg tgttccgtcc	8880
acacgttgcg gacgtactca agttctgccg gcgagccgat ggcggaaagc agaaagacga	8940
ccctggtagaa acctgcattc ggttaaacac cagcacgtt gccatgcagc gtacgaagaa	9000
ggccaagaac ggccgcctgg tgacggtatc cgagggtaga gccctgatta gccgtacaa	9060
gatcgtaaaag agcgaaccgc ggcggccgga gtacatcgag atcgagctag ctgattggat	9120
gtaccgcgag atcacagaag gcaagaacct ggacgtgctg acggttcacc ccgattactt	9180
tttgatcgat cccggcatcg gccgtttct ctaccgctg gcacgccgcg ccgcaggcaa	9240
ggcagaagcc agatggttgt tcaagacgat ctacgaacgc agtggcagcg ccggagagtt	9300
caangatttc tgtttacccg tgcgaagct gatcgggtca aatgacctgc cggagtacga	9360
tttgaaggag gaggcggggc aggcctggccc gatcctagtc atgcgtacc gcaacctgat	9420
caggggcgaa gcatccgccg gttcctaatt tacggagcag atgctagggc aaattgccct	9480
agcaggggaa aaaggctgaa aaggctcttt tccgtggat agcacgtaca ttgggaacct	9540
aaagccgtac attgggaacc ggaacctgta catggggaac ccaaagccgt acattgggaa	9600
ccggtcacac atgtaagtga ctgataaaa agagaaaaaa ggcgattttt ccgcctaaaa	9660
ctctttaaaa ctattaaaa ctcttaaaac ccgcctggcc tgtgcataac tgtctggcca	9720
ggcacagacc gaagagctgc aaaaagcgcc tacccttcgg tcgtgcgct cctacgccc	9780
cgcgcttcg cgtcgcccta tcggggccgc tggccgctca aaaaatgctg gcctacggcc	9840
aggcaatcta ccagggcgcg gacaagccgc gccgtcgcca ctgcaccgcc ggcgccaca	9900
tcaaggcacc ctgcctcgcg cgtttcggtg atgacgtga aaacctctga cacatgcagc	9960
tcccgagac ggtcacagct tgtctgtaag cggatgccgg gacgagacaa gcccgtagg	10020
gcgcgtcagc gggtgttggc ggtgtcggg gcgcagccat gacccagica cgtagcgata	10080
gggagtgta tactggctta actatgcggc atcagagcag attgtactga gagtgcacca	10140

tatgcgggtgt gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggcgctcttc	10200
cgcttccctcg ctcaactgact cgtcgcgctc ggtcgcttcgg ctgcggcgag cggatcagc	10260
tcactcaaaag gcggfaatac ggtaaccac agaatacagg gataacgcag gaaagaacat	10320
gtgagcaaaa gccccagaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt	10380
ccataggetc cgcceccctg acgagcatca caaaaatcga cgtcaagtc agaggtggcg	10440
aaacccgaca ggactataaa gataccagge gtttccccct ggaagctccc tcgtcgctc	10500
tccgttccg accctgcgc ttaccggata cctgcgcgc tttctccctt cgggaagcgt	10560
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggctg ttcgctccaa	10620
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaaacta	10680
tegtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccaactgtaa	10740
caggattagc agagcgaggt atgtagggcg tgcacagag ttcctgaagt ggtagcttaa	10800
ctacggetac actagaagga cagtatttgg tatctgcgt ctgctgaagc cagttacctt	10860
cggaaaaaga gttggtagct ctgtatccgg caaacaaacc accgctggta gcggtgggtt	10920
ttttgtttgc aagcagcaga ttacgcgcag aaaaaagga tctcaagaag atcccttgat	10980
cttttctacg gggctctgacg ctacagtgaa cgaaaactca cgtaaggga ttttggtcat	11040
gcattgatata tcctcccaatt tggtagggc ttattatgca cgcttaaaaa taataaaagc	11100
agacttgacc tgatagtttg gctgtgagca attatgtgct tagtgcactc aacgcttgag	11160
ttaagccgcg ccgcgaagcg gcgtcgctt gaacgaattt ctactagac attatttgcc	11220
gactaacctg gtatctcgc ctticacgta gtggacaact tcttccaact gatctgcgcg	11280
cgaggccaag cgatcttctt cttgtccaag ataagcctgt ctacttcaa gtatgacggg	11340
ctgatactgg gccggcagcg gctccattgc ccagtcggca gcgacatcct tcggcgcgat	11400
tttgccggtt actgcgctgt accaaatgcg ggacaacgta agcactacat ttcgctcact	11460
gccagcccg tcggcgcgcg agttccatag cgttaagggt tcatttagcg cctcaantag	11520
atctgttca ggaaccggat caaagagttc ctccgccgt ggacctacca aggcaacgct	11580
atgttctctt gcttttgta gcaagatagc cagatcaatg tcgactcgtg ctggctcgaa	11640
gatacctgca agaattgcat tgcgtgccca ttctccaaat tgcagttcgc gcttagctgg	11700
ataacgccac ggaatgatgt cgtcgtgcac aacaatgggt acttctacag cgcggagaat	11760
ctcgtctctt ccaggggag ccgaagtct caaaaggctg ttgatcaaag ctgcgccgt	11820
tgtttcaica agctttagg tcaccglaac cagcaaatca atactactgt gtggcttcag	11880
gcgcacatcc actgcggagc cgtacaaatg taegccagc aacgtcgggt cgagatggcg	11940
ctcgatgaag ccaactacct ctgatagttg agtcgatact tcggcgatca ccgcttccc	12000
catgatgttt aactttgttt tagggcgact gccctgcgc gtaacatcgt tgctgtctca	12060
taacatcaaa catcgaccca cggcgtaacg cgttgctgc ttggatgcc gaggcataga	12120
ctgtacccca aaaaaacagt cataacaagc catgaaaacc gccactgcgt tccatg	12176

[0018]

- <210> 62
 <211> 33
 <212> DNA
 <213> 人工的
 <220>
 <223> 合成的序列

	<400> 62	
	aaaaaggtac catgacttcg aaagtttaig atc	33
	<210> 63	
	<211> 34	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 63	
	aaattgagct cttattgttc atttttgaga actc	34
	<210> 64	
	<211> 14880	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 载体序列	
	<400> 64	
	gacatacaaa tggacgaacg gataaacctt ttcaagccct tttaaatatc cgattatici	60
	ataaaacgct cttttctctt aggtttaccc gccaatatct cctgtcaaac actgatagtt	120
	taaaactgaag gcgggaaacg acaatcagat ctagtaggaa acagctaiga ccatgattac	180
	gccaaagctat cgattacgcc aagctatcaa cttgtatag aaaagttgca tcaagatctt	240
	ggtgatgtag caagagctaa gtigtacttc gattcgggtg gacattactc gagaccagat	300
	gttttacact tgaccgtaaa tgagcaccgc aagaaccgg tcacattcat ttcaaggtg	360
	gagaaagcgg aagaigactc aaacaagtaa tcggttgtga ttcgtcagtt catgtcactc	420
	ctatgaagga gtcaagtcca aaatgtttat ttgagtttca aactttttat ctaaacittt	480
	tttcctttatt ttcggttaata atggaagaga accaattctc ttgtatctaa agattatcca	540
	tcatacatcc aaillgagtg ttcaattcig gatgllgtgt taccctacat tcataacca	600
	tgtagccaat tattatgaat ctggccttga tticagttgt gttectttct ttttttctt	660
	tgcataattg catttagaat gttaataaat taagtactg tatttccaca tacattagtt	720
	ccaagaatat acatatatta atttattttt cttaaaaatg ttttggaatg actaatattg	780
	acaacgaaaa tagaagctat gctaaacat tacgtatatg tgacttcaca tgttgttgtt	840
	ttacattccc tatatataig gatggctgtc acaatcagaa acgtgatega aaaaagacaa	900
	acagtgtttg cataaaaaga ctatttcgtt tcattgacaa tttgtgttta ttgtaaaaga	960
	aaagtggcaa agtgggaattt gagttcctgc aagtaagaaa gatgaataa aagaattgag	1020
	tgtgigtgtt tttcttttat ctgaaagctg caatgaaata ttcttaccaa gcccgtttga	1080
	ttattaattg gggtttggtt ttcttgatgc gaactaattg gttatataag aaactataca	1140
	atccatgtta attcaaaaat ttgtatttct ctgttaggaa tatgatttac tatatgagac	1200
	tttccttttc ccaataatag taaatccaaa gatatttgac cggaccaaaa cacattgac	1260
	tattttttag ttattttaat ccagtttctc tgagataatt cattaaggaa aacttagtat	1320
	taaccatccc taagattaaa taggagecaa actcacattt caaatattaa ataacataaa	1380
	atggatttaa aaaatctata cgtcaaattt tatttatgac atttcttatt taaatttata	1440
	tttaatgaaa tacagctaaag acaaaccaaa aaaaaaatat tttctaagtg gtccaaaaca	1500

[0019]

[0020]

tcaattccgt tcaatattat taggtagaat cgtacgacca aaaaaaggta ggtaaatacg	1560
aattagaaac atatctataa catagtatat attattacct attatgagga atcaaaatgc	1620
atcaaatatg gatttaagga atccataaaa gaataaaatc tacgggaaaa aaaaiggaat	1680
aaattctttt aagtttttlll tttgtttttt atttggttagt tctccatttt gttttatttc	1740
gtttggattt attggttcca aatactttgt aaaccaccgt tgtaatcttt aaacggggtt	1800
ttcacttctt ttttatattc agacataaag catcggcctg tttaatcaat caatagattt	1860
tatttttctt ctcaattatt agtaggtttg atgtgaactt tacaaaaaaa acaaaaacaa	1920
atcaatgcag agaaaagaaa ccacgtgggc tagtcccacc ttgtttcatt tccaccacag	1980
gttcgatctt cgttaccgtc tccaatagga aaataaacgt gaccacaaaa aaaaaacaaa	2040
aaaaagtcta tatattgctt ctctcaagtc tctgagtgtc atgaacacaa gtaaaaaaca	2100
aagactcgag tcaagtttgt acaaaaaagc aggcctgtac ctttaagaaat cctctcttct	2160
cctcttcatt ttcaaggtaa atctctctct ctctctctct ctctgttatt ccttggttta	2220
attaggtatg tattattgct agtttgtaa tctgcctate ttatgtatgc cttatgtgaa	2280
tatctttate ttgttctct catccgttta gaagctataa aattgttgat ttgactgtgt	2340
atctacacgt ggttatgttt atatctaate agatatgaat tcttccatat tgttgctttt	2400
gtgtgtacca atccgaaatc gttgattttt ttcatttaat cgtgtagcta attgtacgia	2460
tacatatgga tctacgtate aattgttcat ctgtttgtgt ttgtatgtat acagatciga	2520
aaacateact tctctcatct gattgtgttg ttacatacat agatatagat ctgttatate	2580
atttttttta ttaattgtgt atatataat gtgcatagat ctggattaca tgatttgtat	2640
tattttacatg attttgttat ttacgtatgt atataatgtag atctggactt tttggagtgt	2700
ttgacttgat tgtatttgtg tgggtatgtg tgggttctga tcttgatatg ttatgtatgt	2760
gcagaccaatg gaagacgcca aaaacataaa gaaaggcccg gcgccattct atccgcigga	2820
agatggaacc gctggagagc aactgcataa ggctatgaag agatagccc tgggtccctg	2880
aacaattgct ttacagatg cacatatcga ggtggacatc acttaccttg agtacttga	2940
aatgtccgtt cgggtggcag aagctatgaa acgatatggg ctgaatacaa atcacagaat	3000
cgtcgtatgc agtgaanaact ctcttcaatt ctttatgccg gtgttggcg cgttatttat	3060
cggagtigca gttgcgcccg cgaacgacat ttataatgaa cgtgaattgc tcaacagtat	3120
gggcatttgc cagcctaccg tgggtttcgt ttccaaaaag gggttgcaaa aaattttgaa	3180
cgtgcacaaa aagctcccaa tcatccaaaa aattattatc atggattcta aaacggatta	3240
ccagggatit cagtcgatgt acacgttcgt cacatctcat ctaccctccg gttttaatga	3300
atacgttttt gtgccagagt ccttcgatag ggacaagaca attgcactga tcatgaactc	3360
ctctggatct actggtctgc cttaagggtg cgtcttgcct catagaactg cctgcgtgag	3420
attctcgcct gccagagatc ctatttttgg caatcaaatc attccggata ctgcgatit	3480
aagtgttgtt ccattccatc acggttttgg aatgtttact acactcggat atttgatatg	3540
tggatttcca gtctctttaa tgtatagatt tgaagaagag ctgtttctga ggagccttca	3600
ggattacaag attcaaatg cgtctgttgt gccaacctta ttctctctct tcgccaaaag	3660
cactctgatt gacaaatagc atttatctaa ttacacgaa attgttcttg gtgcgcctcc	3720
cctctetaag gaagtcgggg aagcgttgc caagagggtc catctgccag gtatcaggca	3780

[0021]

```

aggatatggg ctactgaga ctacatcagc tattctgatt acacccgagg gggatgataa 3840
accgggcgcg gtcggtaaag ttgttccatt tttgaagcg aaggttgtgg atctggatac 3900
cgggaaaacg ctgggcgtta atcaaagagg cgaactgtgt gtgagaggic ctatgattat 3960
gtccggttat gtaacaalc cggaaagcag caacgccttg atlgacaagg aiggaigcct 4020
acattctgga gacatagctt acfgggacga agacgaacac ttcttcacg ttgaccgcct 4080
gaagtctctg attaagtaca aaggctatca ggtggctccc getgaattgg aatccatctt 4140
gtccaacac cccaacatct tgcacgcagg tctgcagggt cttcccgaag atgacgccgg 4200
tgaacttccc gccgccgttg ttgttttga gcacggaaag acgatgacgg aaaaagagat 4260
cgtggattac gtcgccagtc aagtaacaac cgcgaaaaag ttgcgcggag gagtgtgtt 4320
tgtggacgaa gtaccgaaag gtcttaccgg aaaactcgac gcaagaaaaa tcagagagat 4380
cctcataaag gccaaagaagg gcgaaagat cgcctgttaa ctcgagcata tgggctcgaa 4440
tttcccgat cgttcaaca tttggcaata aagtttctta agattgaatc ctgttgcgg 4500
tcttgcgat attatcatat aatttctgtt gaattacgtt aagcatgtaa taattaacat 4560
gtaatgcatg acgtatttta tgagatgggt tttatgatt agagtccgc aattatcat 4620
ttaatacgcg atagaaaaa aaatatagcg cgcaactag gataaattat cgcgcgcggt 4680
gtcatctatg ttactagatc gggaattcaa gcttggcgta atcatggacc cagcttctt 4740
gtacaaagtg gggtaaccaat tcgaatccaa aaattacgga tatgaatata ggcatatccg 4800
tatecgaatt atcgtttga cagctagcaa cgattgiaca attgcttctt taaaaagga 4860
agaaagaaag aaagaaaaga atcaacatca gcgttaacaa acggccccgt tacggeccaa 4920
acggtcatat agagtaacgg cgttaagcgt tgaaagactc ctatcgaaat acgtaaccgc 4980
aaacgtgta tagtcagatc cctcttctt ctaccgcctc aaacacaaaa ataacttct 5040
acagcctata tatacaaccc ccccttctat ctctcttctc tcacaattca tcatcttct 5100
ttcttacc ccaattttaa gaaatctct ctctctctc tcaatttcaa ggtaaatctc 5160
tctctctctc tctctctctg ttattcttg ttttaattag gtatgtatga ttgctagttt 5220
gttaatctgc ttattctatg tatgccttat gtgaatatct ttatcttgtt catctcatec 5280
gtttagaagc tataaatttg ttgatttgac tgtgtatcta cagctggtta tgtttatctc 5340
taatcagata tgaatttctt catattgttg cgtttgtgtg taccaatccg aaatcgtiga 5400
tttttttcat ttaatcgtgt agctaattgt acgtatacat atggatctac gtatcaattg 5460
ttcatctgtt tgtgtttgta tgiatacaga tctgaaacaa tcacttctct catctgattg 5520
tgttgttaca tacatagata tagatctgtt atatcatitt tttattaat tgggtatata 5580
tatatgtgca tagatctgga ttacatgatt gtgattatit acatgatttt gttatttacg 5640
tatgtatata tgtagatctg gacttttgg agttgttgac ttgattgtat ttgtgtgtgt 5700
atatgtgtgt tctgacttgg atatgttatg tatgtgcage tgaacctagg cggcggcaac 5760
aacaacaaca acaacatctt ctctgatctc ctctctcacc aaacctctc ctctctctc 5820
caaatcacca ttaccaatct ccagattctc cctccattc tccetaaacc ccaacaaatc 5880
atctctctcc tcccgcgcgc gcggtatcaa atccagctct cctctctcca tctccgcggt 5940
gtcaaacaca accaccaatg tccaaaccac tccctctcca accaaacctt ccaaaaccga 6000
aacattcact tccgatttgc ctccagatca accccgcaaa ggcgttgata tctctgtgga 6060

```

[0022]

```

agcttttagaa cgtaaggcg tagaaaccgt attcgcttac cctggaggta catcaatgga 6120
gattcaccaa gecttaaccc gctcttcctc aatccgiaac gtccttcctc gtcacgaaca 6180
aggaggigtat ttgcagcag aaggatacgc tgcattctca ggtaaaccag gtaictgtat 6240
agccacttca ggccccggag ctacaaatct cgttagcgga ttaggcgatg cgttltttaga 6300
tagtgttccct ctgttagcaa tcacaggaca agtccctcgt cgtatgattg gtacagatgc 6360
gtttcaagag actccgattg ttgaggtaac ggttcgatt acgaagcata actatcttgt 6420
gatggatgtt gaagatatcc ctaggattat tgaggaagct ttcttttttag ctacttctgg 6480
tagacctgga cctgttttgg ttgatgttcc taaagatatt caacaacagc ttgcgattcc 6540
taattgggaa caggctatga gattacctgg ttatatgtct aggatgccta aacctccgga 6600
agatttcicat ttggagcaga ttgttaggtt gatttcigag tctaagaagc ctgtgttgtat 6660
tgttgggtgt ggttgtttga attctagcga tgaattgggt aggtttgttg agcttacggg 6720
gataccctgtt gcgagtacgt tgatggggct gggatcttat ccttgtgatg atgagttgtc 6780
gttacatatg ctgtgaatgc atgggactgt gtatgcaaat tacctgttgg agcatagtga 6840
tttgttgttg gcgtttgggg taaggtttga tgatcgtgtc acgggtaagc ttgaggcttt 6900
tgctagttag gctaagattg ttcataatga tattgactcg gctgagattg ggaagaataa 6960
gactcctcat gtgtctgtgt gtgtgatgt taagctggct ttgcaaggga tgaataaggt 7020
tcttgagaac cgagcggagg agcttaagct tgattttgga gtttgaggga atgagttgaa 7080
cgtacagaaa cagaagtct cgttgagctt taagncgttt gggaagcta ttctccaca 7140
gtatgcgatt aaggctcttg atgagttgac tgatggaaaa gccataataa gtactgtgt 7200
cgggcaacat caaatgtggg cggcgcagtt ctacaattac aagaaaccaa ggcagtggct 7260
atcateagga gcccttggag ctatgggatt tggacttccct gctgcgattg gacgtctgt 7320
tgctaaccct gatgcgatag ttgtggatat tgacggagat ggaagcttta taatgaatgt 7380
gcaagagcga gccactatc gtgtagagaa tcttccagtg aaggtaactt tatlaaaca 7440
ccagcatctt ggcatgttta tgcaatggga agatcggttc tacaagcta accgagctca 7500
cacatttctc ggggatccgg ctcaggagga cgagatattc ccgaacatgt tgctgtttgc 7560
agcagcttgc gggattccag cggcgagggt gacaaagaaa gcagatctcc gagaagctat 7620
tcagacaatg ctggatacac caggacctta cctgttggat gtgatttgc cgcaccaaga 7680
acatgtgttg ccgatgatcc cgaatggtag cactttcaac gatgtcataa cggaaggaga 7740
tgcccggatt aaatactgag agatgaaacc ggtgattatc agaacctttt atggctttg 7800
tatgcataig gtaaaaaaac ttagtgttga atttctgtt ttgttttgga atttgagttt 7860
cttttagttg ttgatctgcc tgccttttgg tttaactcag actactactg ctgtgttgt 7920
ttggttctct tcttttcat ttataaataa ataatecgtt tcggtttact cctgttgact 7980
ggctcagttt ggttatttgc aaatgcgaat ggtaaaatga gtaattgaaa ttctgtatta 8040
gggttctaag ctgttttaac agtcactggg ttaatatctc tcgaatcttg catggaaaat 8100
gctcttacca ttggttttta attgaaatgt gctcataatg gccgttggtt ccaaattaaa 8160
taaaactacg atgtcatcga gaagtaaaat caactgtgtc cacattatca gttttgtgta 8220
tacgatgaaa tagggtaatt caaaatctag ctigatatgc cttttgggtc attttaacct 8280
ctgtgaaaca ttttttcaga ttttgaacaa gtaaatccaa aaaaaaaaaa aaaaatctc 8340

```

[0023]

aactcaacac taaattatit taatgtataa aagatgetta aaacatttgg cttaaaagaa	8400
agaagctaaa aacatagaga actcittgiaa atgaagtat gaaaatatac tgaattgggt	8460
attatatgaa tttttctgat ttaggattca catgaacca aaaggaaatc cagaagcact	8520
aatcagacat tggaagtagg aatatttcaa aaagtttttt ttttttaagt aagtgacaaa	8580
agcttttaaa aaatagaaaa gaaactagia ttaaagttgt aaatttaata aacaaaagaa	8640
attttttata ttttttcatt tctttttcca gcatgagggt aigatggcag gatgtggatt	8700
tcattttttt ctttttgata gccttttaat tgatctatta taattgaaga aaaaatatta	8760
gttaattata gatataatit aggtagtatt agcaatttac acttccaaaa gactatgiaa	8820
gttgtaataa tgatgcgttg atctcttcat cattcaatgg ttagtcaaaa aaataaaagc	8880
tlaactagta aactaaagta gtaaaaaat gtactttagt ttaaaatati acafgaataa	8940
tcacaaaaga catttatgtg aaacaaaaac aatatctaga gtcgacttaa ttaactagt	9000
ggcgcgccaa ttgactagta ggccctatga ttaattaagg ccgcccagag catatgggca	9060
tgcaagcttg gcgtaatcat ggcaacttfa ttatacatag ttgataattc actggccgga	9120
tctgcttggt aataattgtc attagatgtt tttatgcat agatgcactc gaaatcagcc	9180
aattttagac aagtatcaaa cggatgttaa ttcagtacat taaagacgtc cgcaatgtgt	9240
tattaagttg tctaaagctc aatttgttta caccacaata tatcctgcca ccagccagcc	9300
aacagctccc cgaccggcag ctccgcacaa aatcaccacg cgttaccacc acgcccggccg	9360
gcgcgatggt gttgaccgtg ttccgcccga ttgccgagtt cgagcgttcc ctaatcctcg	9420
accgcacccg gagcgggcgc gagcccgcca aggcccgagg cgtgaagttt ggcccccgcc	9480
ctaccctcac cccggcacag atcgcgacg cccgcgagct gatcgaccag gaaggccgca	9540
ccgtgaaaga ggccgctgca ctgcttgccg tgcctgcctc gacctgttac cgcgcacttg	9600
agcgcagcga ggaagtgacg cccaccgagg ccaggcggcg cgggtgcctc cgtgaggacg	9660
cattgaccga ggccgacgcc ctggcgcccg ccgagaatga acgccaagag gaacaagcat	9720
gaaaccgcac caggacggcc aggacgaacc gtttttcatt accgaagaga tccaggcgga	9780
gatgatcgcg gccgggtacg tgttcgagcc gccgcgcac gtctcaaccg tgcggctgca	9840
tgaatcctg gccggtttgt ctgatgcaa gctggcgccc tggccggcca gcttgccgc	9900
tgaagaaacc gagcgccgcc gtctaaaaag gtgatgtgta tttagttaa acagcttgcg	9960
tcattcggtc gtgcgtata tgatcgatg agtaataaaa caaatagca aggggaacgc	10020
aigaaggtta tgcctgtact taaccagaaa ggccggctag gcaagacgac catcgcaacc	10080
catctagccc gcgccctgca acfcgccggg gccgatgttc tgttagtcga ttccgatccc	10140
cagggcagtg cccgcgattg ggccggcgtg ccgggaagatc aaccgctaac cgttgtcgcc	10200
atcgaccgcc cgacgatga ccgcgacgtg aaggccatcg gccggcgca cttcgtagt	10260
atcgacggag cgcgccaggc ggccgacttg gctgtgtccg cgatcaaggc agccgacttc	10320
gtctctatc cgggtcagcc aagcccttac gacatatggg ccaccgccga cctggtggag	10380
ctgggttaagc agcgcatga ggtcacggt ggaaggctac aagcgccctt tgcgtgtc	10440
cggcgcatca aaggcacgcg catcgccggt gaggttgccg aggcgctggc cgggtacgag	10500
ctgcccttc tttagtcccg tatcacgcag ccgctgagct acccaggcac tgcgcgcgc	10560
ggcacaaccg ttcttgaatc agaaccgag ggcgacgctg ccgcgaggt ccaggcgctg	10620

[0024]

```

gcccgtgaaa ttaaatcaaa atcatttga gttaatgagg taaagagaaa atgagcaaaa 10680
gcacaaacac gctaagtgcc ggccgtccga ggcacgcag cagcaaggct gcaacgttgg 10740
ccagccitggc agaacgcga gccatgaagc gggtaactt tcagtigccg gcggaggatc 10800
acaccaagcl gaagalgac gggtagcc aaggcaagac catlaecgag ctcatactg 10860
aafacatgc gcagctacca gagtaaatga gcaaatgaat aaatgagtag atgaatttta 10920
gggctaaag gagggcgcat gaaaaatcaa gaacaaccag gcaccgacgc cgtggaatgc 10980
cccatgtgtg gaggaacggg cgttgccca ggcgtaagcg gctgggttgt ctgccggccc 11040
tgcaatggca ctggaacccc caagcccag gaatcgccgt gagcggtcgc aaaccatccg 11100
gcccgttaca aatcggcgcg gcctgggtg atgacctggt ggagaagtig aaggccgcgc 11160
aggccgccc gcggcaacgc atcaggcag aagcacccc cgtgaatcg tggcaagcgg 11220
ccgtgatcg aatccgcaa gaatcccgc aaccgcgcgc agccggtgcg ccgtcgatta 11280
ggaagccgc caaggcgac gagcaaccag atttttctg tccatgctc tatgacgtg 11340
gcaccccgca tagtcgcagc atcatggacg tggccgttt ccgtctgctg aagcgtgacc 11400
gacgagctgg cagggtagc cgtacgagc tccagacgg gcacgtagag gttccgcag 11460
ggccggccgg catggccagt gtgtgggatt acgacctggt actgatggcg gttcccatc 11520
taaccgaate catgaaccga taccgggaag ggaagggaga caagcccgc cgcgigtcc 11580
gtccacacgt tgggacgta ctcaagtct gccggcgagc cgtggcgga aagcagaaag 11640
acgacctggt agaacctgc attcggttaa acaccacgca cgttgccatg cagcgtacga 11700
agaaggccaa gaacggcgc ctggtgacgg tatecgaggg tgaagcctg attagccgt 11760
acaagatcgt aaagagcgaa accgggcgc cggagtacat cgagatcgag ctactgatt 11820
ggatgtaccg cgagatcaca gaaggcaaga acccgacgt gctgacggtt caccocgatt 11880
actttttgat cgateccgcg atcgccgtt ttctctaccg cctggcacgc cgcgccgcag 11940
gcaaggcaga agccagalg ttttcaaga cgtctacga acgcaglgc agcccgagg 12000
agttcaagaa gttctgttc accgtgcga agctgatecg gtcaaatgac ctgccggagt 12060
acgatttgaa ggaggaggcg gggcaggctg gccgatcct agtcatgcgc taccgcaacc 12120
tgatcgaggg cgaagcatcc gccgttccct aatgtacgga gcagatgcta ggcaaatig 12180
ccctagcagg gaaaaaagg cgaagggtc ttttctctg ggatagcacg tacattggga 12240
aaccaaagcc gtacattggg aaccggaacc cgtacattgg gaacccaaag ccgtacattg 12300
ggaaccggtc acacatgtaa gtgacigata taaaagagaa aaaaggcgat ttttcgcct 12360
aaaactcttt aaaacttatt aaaactctta aaacccgcct ggctgtgca taactgtctg 12420
gccagcgac agccgaagag ctgcaaaaag cgcctaccct tgggtcgctg cgtccctac 12480
gccccgcgc ttcgcgtcg cctatcgcg ccgtggccg ctcaaaaatg gctggcctac 12540
ggccaggcaa tctaccagg cgggacaaag ccgcgcgtc gccactgac cgcggcgcc 12600
cacatcaagg caccctgct cgcgcgttc ggtgatgac gtgaaaacct ctgacacatg 12660
cagctcccg agacggcac agctgtctg taagcggatg ccgggagcag acaagccgt 12720
caggcgcgct cagcggtgt tggcggtgt cggggcgag ccatgacca gtcacgtagc 12780
gatagcggg tgtatctgg cttaactat cggcatcaga gcagattgta ctgagagtgc 12840
accatgatcg ggtgaaata ccgcacgat gcgtaaggag aaaataccgc atcaggcgct 12900

```

	cttccgcttc ctgcctcact gaetcgctgc gctcggctgt tcggctgcgg cgagcggtat	12960
	cagctcacte aaaggcggtg atacggttat ccacagaatc aggggataac gcaggaaaga	13020
	acatgtgagc aaaaggccag caaaaggcca ggaaccgtaa aaaggccgcg ttgcigggct	13080
	ttttccatag gtccgcccc cctgaagagc atcacaaaaa tcgaagctca agtcagaggt	13140
	ggcgaaaccc gacaggacta taaagatacc aggegtttcc ccttggaagc tccctcgtgc	13200
	gtctctctgt tccgacctg ccgcttaacc gataacctgc cgcctttcct ccttcgggaa	13260
	gcgtggcgct ttctcatagc tcacgtgtga ggtatctcag ttccgtgtag gtcgttcgt	13320
	ccaagctggg ctgtgtgac gaacccccg ttccagccga ccgctgcgc ttatccggtg	13380
	actatcgtct tgagtccaac ccggtaac acgacttate gccactggca gcagccactg	13440
	gtaacaggat tagcagagcg aggtatgtag gcggtgtac agagttcttg aagtgggtgc	13500
	ctaactacgg ctacactaga aggacagtat ttggtactg cgtctcgtg aagccagtta	13560
	ccttcggaaa aagagttggt agctcttgat ccggcaaaac aaccaccgct ggtagcgggt	13620
	gttttttgt ttgcaagcag cagattaagc gcagaaaaaa aggatctcaa gaagatcctt	13680
	tgatcttttc tacgggggtct gaagctcagt ggaacgaaaa ctacagttaa gggattttgg	13740
	tcatgcatga tatatctccc aatttggtga ggctttatta tgcacgctta aaaataataa	13800
	aagcagactt gacctgatag ttggctgtg agcaattatg tgcctagtgc atctaacgt	13860
	tgagttaagc cgcgcgcga agcggcgctc gcttgaacga atttctagct agacattatt	13920
	tgcggaactc ctgggtgac tcgcttttca cgtagtggac aaattcttcc aactgatctg	13980
	cgcgcgagge caagcgatct tctctgtgct caagataagc ctgtctagct tcaagtatga	14040
[0025]	cgggctgata ctgggcgggc aggcgtccca ttgccagtc gccagcgaca tccctcggcg	14100
	cgattttgcc ggttaactgc ctgtaccana tgcgggacaa cgtatgcctt acatttcgt	14160
	catcgccagc ccagtcgggc ggcgagttcc atagcgttaa ggtttcattt agcgcctcaa	14220
	atagatccig ttcaggaacc ggalcaaaag gtlcctccgc cgtcggacct accaaggcaa	14280
	cgtatgttcc tcttcttttt gtcagcaaga tagccagatc aatgtcgtc gtggttggt	14340
	cgaagatacc tgcagaatg tcattgcgt gccattctcc aaattgcagt tcgcgcttag	14400
	ctggataacg ccacggaatg atgtcgtcgt gcacaacaat ggtgacttct acagcggga	14460
	gaatctcgt ctctccaggg gaagccgaag ttccaaaag gtcgttgatc aaagctcgcc	14520
	gcgttgttcc atcaagcctt acggtcaccg taaccagcaa atcaatatca ctgtgtggt	14580
	tcaggccgcc atccactgcg gagccgtaca aatgtacggc cagcaacgtc ggttcgagat	14640
	ggcgctcgat gaagccaaact acctctgata gttagctega tacttcggcg atcaccgctt	14700
	cccccatgat gttaacttt gttttagggc gaactgccctg ctgcgttaaca tcgttctgc	14760
	tccataacat caaacatcga cccacggcgt aacgcgcttg ctgcttggat gcccaggca	14820
	tagactgtac cccaaaaaaa cagtcataac aagccatgaa aaccgccact gcgttccatg	14880
	<210> 65	
	<211> 39	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 65	

	tttttggcgc gcctttctct cgttctcacc tttctctct	39
	<210> 66	
	<211> 38	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 66	
	taataggcgc gcctttgtca aacttttgat tgtcacct	38
	<210> 67	
	<211> 36	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 67	
	tatatggcgc gccaaatcgt tctttcaaat ctctca	36
	<210> 68	
	<211> 37	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 68	
	ttataggcgc gcctctgtaa ttcacaaaaa actgaga	37
[0026]	<210> 69	
	<211> 36	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 69	
	tttttggcgc gccttagatc tcgtgccgtc gtgcga	36
	<210> 70	
	<211> 33	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 70	
	tttttggcgc gcctttgatc aagcctgttc aca	33
	<210> 71	
	<211> 36	
	<212> DNA	
	<213> 人工的	
	<220>	
	<223> 合成的序列	
	<400> 71	
	aaaaaggcgc gcctcatcag atcttcaaaa ccccaa	36
	<210> 72	

[0027]	<211>	40	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	72	
		aaaaaggcgc gccctgatttg agggtagtac taaccgggaa	40
	<210>	73	
	<211>	36	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	73	
		ttttttggcg cgcccttaaga aalectctct tctcct	36
	<210>	74	
	<211>	36	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	合成的序列	
	<400>	74	
		ttttttggcg cgccctcgcga catacataac atatca	36
[0027]	<210>	75	
	<211>	14019	
	<212>	DNA	
	<213>	人工的	
	<220>		
	<223>	载体序列	
	<400>	75	
		gigattttgt gccgagctgc cggtegaggga gctgttgget ggctgggtgc aggatatatt	60
		gigggtgtaa caaattgacg cttagacaac ttaataacac attgcggacg tctttaatgt	120
		actgaattta gttactgac actgattaag tactgatata ggtaccaagc ttcccgget	180
		gcagtgcagc gtgaccgggt cgtgccctc tctagagata atgagcattg catgtctaag	240
		ttataaaaaa ttaccacata tttttttgt cacacttgtt tgaagtgcag ttatctatac	300
		ttataacata tattttaaact ttactctacg aataatataa tctatagtag tacaataata	360
		tcagtgtttt agagaatcat ataaatgaac agttagacat ggcttaagg acaattgagt	420
		attttgacaa caggactcta cagttttatc tttttaggt gcatgtgttc tctttttttt	480
		ttgcaaatag cttcacctat ataatacttc atccatttta ttagtacata catttagggt	540
		ttagggttaa tggtttttat agactaattt ttttagtaca tetattttat tetatttttag	600
		cctctaaatt aagaaaacta aaactctatt ttagttttt tatttaaatag tttagatata	660
		aaatagaata aaataaagtgc actaaaaatt aaacaaatac cctttaagaa attaaaaaaa	720
		ctaaggaaac atttttcitg ttccgagtag ataatgccag cctgttaaac gccgtcgacg	780
		agctaaacgg acaccaacca gcgaaccagc agcgtcgcgt cgggccaagc gaagcagacg	840
		gcacggcacc tcgtgcgtg cctctggacc cctctcgaga gtcccgctcc accgttggac	900
		tgtctcget gtccgcatcc agaaattgcg tggcggagcg gcagacgtga gccggcacgg	960

[0028]

cagggggcct cctctctctc tcacggcacc ggcagctacg ggggattccct tteccaccgc 1020
 tecttcgctt tcccttcctc gcccgccgia ataaatagac accccctcca caccctcttt 1080
 ccccaacctc gtgtgtttcg gagcgcacac acacacaacc agatctcccc caaatccacc 1140
 cgtcggcacc tcgcttcaa ggtacgccgc tegtctccc ccccccccc cctctctacc 1200
 ttctctagat cggcgttccg gtccatggtt agggcccggt agttctactt ctgttcatgt 1260
 ttgtgttaga tccgtgtttg ttttagatcc gtgctgetag cgttcgtaca cggatgcgac 1320
 ctgtacgtea gacacgttct gattgctaac ttgccagtgt ttctcttttg ggaatcctgg 1380
 gatggcteta gccgttccgc agacgggacg gatttcata tttttttgt ttctgttgeat 1440
 agggtttgggt ttgccctttt cctttatttc aatatatgcc gtgcacttgt ttgtcgggtc 1500
 atcttttcat gctttttttt gctttgtttg tgalgaiglg gtcgtgttgg gcgttcgttc 1560
 tagatcggag tagaattctg ttcaaaacta cctggaggat ttattaattt tggatctgta 1620
 tgtgtgtgcc atacatatcc atagttacga attgaagatg atggatggaa atatcgatct 1680
 aggataggta tacaigtiga tgcgggtttt actgatgcat atacagagat gctttttgtt 1740
 cgttgggttg tgatgatgtg gtgtgtttgg gccgttcgtc attcgttcta gatcggagta 1800
 gaatactgtt tcaaaactacc tgggtatatt ataattttg gaactgtatg tgtgtgtcat 1860
 acatcttcat agttacgagt ttaagatgga tggaaatacc gatctaggat aggtatacat 1920
 gttgatgtgg gttttactga tgcataatac tgatggcata tgcagcatct attcatatgc 1980
 tctaaccttg agtacctatc tattataata aacaagtatg ttttataatt attcgaatct 2040
 tgatataact ggatgatggc atatgcagca gctatatgtg gattttttta gccctgcctt 2100
 catacgctat ttatttgcct ggtactgttt ctlttgtcga tgetcaccct gttgtttgggt 2160
 gttactcttg cagcccgggg gatccactag ttctagaaac catggccacc gccgcgcccg 2220
 cgtctaccgc gctcactggc gccactaccg ctgcgcccaa ggcgagggcg cgggcgcacc 2280
 tctggccacc ccgcgcgcc ctcgccgcgc ccacaggtg ctcagcggcg tcacccgcca 2340
 tgccgatggc tccccggcc accccgctcc gcccggtggg ccccaccgat ccccgcaagg 2400
 gcccgacat cctcgtcgag tccctcgagc gctggcgctt cgcgcagctc ttccctacc 2460
 ccggcggcac gtccatggag atccaccagg cactcaccgc ctccccgtc atcgccaacc 2520
 acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggctacgcg cgtctctcgg 2580
 gccgcgtcgg cgtctgcate gccacctcgc gccccggcgc caccaacctt gtctccgcgc 2640
 tgcgcgacgc gctcctcgat tccgtcccca tggctcccat caccggacag gtgcgcgac 2700
 gcatgattgg caccgaagcc ttctaggaga cggccatcgt cgaggtcaac cgtcccatca 2760
 ccaagcacia ctacctggtc ctcgacgtcg acgacatccc ccgcgtcgtg caggagcctt 2820
 tcttctctcg cctctctggt cgaccgggce cggctcttgt cgacatcccc aaggacatcc 2880
 agcagcagat ggcgggtcct gtctgggaca agcccatgag tctgcctggg tacattgcgc 2940
 gcccttccaa gcccccgcgc actgagttgc ttgagcaggt gctgcgtctt gtttgtgaat 3000
 cccggcgccc tgttctttat gtggcggtg gctgcgcagc atctggtgag gatttgcgac 3060
 gctttgtgga gctgactgga atcccgtca caactactct tatgggcctc ggcaacttcc 3120
 ccagcgacga cccactgtct ctgcgcatgc taggtatgca tggcacgggt tatgcaaat 3180
 atgcagtgga taaggccgat ctgttgcctt caetttgtgt gccgtttgat gatcgtgtga 3240

[0029]

cagggaagat tgaggctttt gcaagcaggg ctaagattgt gcacgttga attgatccgg 3300
 ctgagattgg caagaacaag cagccacatg tgcacatctg tgcagatgtt aagcttgctt 3360
 tgcaggcat gaatgctctt ctgaaggaa gcacatcaaa gaagagcttt gactttggct 3420
 catggaacga tgagttggat cagcagaaga gggaattccc ctttgggtat aaaacatcta 3480
 atgaggagat ccagccacaa tatgtattc aggttcttga tgagctgaag aaaggcgagg 3540
 ccatcatcgg cacaggtgtt gggcagcacc agatgtgggc ggcacagtac tacacttaca 3600
 agcggccaag gcagtggtt tcttcagctg gtcttggggc tatgggattt ggtttgccgg 3660
 ctgctgctgg tgccttcttg gccaaaccag gtgttaactgt tgttgacatc gatggagatg 3720
 gtagctttct catgaacgtt caggagctag ctatgatccg aatigagaac ctcccgtga 3780
 aggtctttgt gctaaacaac cagcaccctg ggalggigt gcagtgagg gacaggctct 3840
 ataagcccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatatc 3900
 cagatttctg gacgatcgcc aaagggttca acattccagc ggtccgtgtg acaaagaaga 3960
 acgaagtcgg cgcagcgata aagaagaagc tcgagactcc agggccgtac ctcttgata 4020
 taatctccc acaccaggag catgtgttc ctatgatccc taatggtggg gctttcaagg 4080
 atatgatcct ggatggtgat ggcaggactg tgtactgac taaaatccag caagcaactg 4140
 atctaaaac cagcaagcac cgcctccctg ctagtacaag ggtgatatgt tttatctgt 4200
 gtgatgttct cctgtattct atctttttt gtaggccgtc agctatctgt tatggtaatc 4260
 ctatgtagct tccgacctg taattgtga gtctgtgtt ttccttctgg catgtgcat 4320
 aagagatcat ttaagtgcct ttgctacat ataaataaga taataagcac tgcfatgcag 4380
 tggttctgaa ttggttctg ttgccaaatt taagtgtcca acfggfcttt gcttttgttt 4440
 tctctatttt ttctttttt tagttattat tatattggta atttcaactc aacatatgat 4500
 gtatggaata atgetagggc tgcaatttca aactatttta caaaccagaa tggcattttc 4560
 gtggttgag gggagtga aaatgagg catttgactg aattagtac ctgatccatt 4620
 ttctgggtt ggatcatagg aattaaatc cattctaata atagiaattt tggcatatat 4680
 caattaagtt aattcgttt tatgcaaat atattgtat actattatta tcaagatgtc 4740
 ggagataatt atatgtaca tttttactat acaggagtga gatgaagagt gtcagtgaag 4800
 ttacacagta gaaacaaatt ctattaatgc ataaatcat ttccatcatc caccctatga 4860
 atttgagata gacctatate taaacttga aaagtgggtg aatatcaaat tccaaattaa 4920
 ataagttatt ttatgagt aafttcaatt tctctaaaac gaagggaact aaagccctc 4980
 taaagctaat ttggaactc aaactttct agcattggag gggattgaga aaaaatatta 5040
 atcattttc atctcaatc ttcaatctcc aaagagattt gatttctta ttagtctgtt 5100
 ccatgcatca aatcggtc atgtgtcatt atttccatg acgattgacg agttgtctg 5160
 gggcctagcg ctttccacgc cgatgtgctg gggcctggct ctggagaaga cagcttgata 5220
 tttaaagcta tcaattgttt caattgatc ccaattcatt tttctaatg tagaaaacgg 5280
 tgacgtataa gaaaaagaat gaattaggac tttattccg tacactaatc tagagcgccc 5340
 caagcttga cactagtac cgtaattga tttaaattta attaatgcag tgcagcgtga 5400
 cccggtctg cccctctcta gagataatga gcattgcatg tctaagttat aaaaattac 5460
 cacatatttt tttgtcaca ctgtttgaa gtgcagtta tctatctta tacatatatt 5520

	ttaacttttac tctacgaata atataatcfa tagtactada ataataacag tgttttagag	5580
	aatcatataa atgaacagtt agacatggic taaaggacaa ttgagtattt tgacaacagg	5640
	acctacagtt ttatctttt tagtgtcat gtgttctctt tttttttgc aaatagcttc	5700
	acctatataa tacttcatcc attttattag tacatccatt tagggtttag ggtaaatggt	5760
	ttttatagac taatittttt agtacatcia tttatitcta ttttagcctc taaattaaga	5820
	aaactaaaac tctatttttag tttttttatt taatagtitta gatataaaa atagaataaaa	5880
	aaagtacta aaaaattaac aaataccctt taagaaatta aaaaaactaa ggaacattt	5940
	ttcttggttc gtagtagataa tgcagcctg tfaaacgccg tcgacgagtc taacggacac	6000
	caaccagcga accagcagcg tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg	6060
	tcgttgcttc tggaccctc tcgagagttc cgtccaccg ttggacttgc tccgtctgcg	6120
	gcattccaga attgcgtggc ggagcggcag acgtgagccg gcacggcagg cggcctcctc	6180
	ctctctcac ggcaccgga getacggggg attccttcc caccgtctct tcgtttccc	6240
	ttcttcgccc gccgaataa atagacaccc cctccacacc ctctttccc aacctcgtgt	6300
	tgctcggagc gcacacacac acaaccagat cggcgcgcct taagaaatcc tctctctcc	6360
	tctctattt caaggtaaat ctctctctct ctctctctct ctgttatcc ttgttttaat	6420
	taggtatgta ttatgtctag ttgttaatc tcttatctt atgtatcct tatgtgaata	6480
	tctttatctt gttcatctca tcgttttaga agctataaat ttgttgattt gactgtgtat	6540
	ctacacgtgg ttatgtttat atctaatcag atatgaattt ctccatattg ttgcgtttgt	6600
[0030]	gtgtaccaat ccgaatcgt tgattttttt catttaatcg tctagctaat tctacgtata	6660
	catatggatc taccgtatcaa ttgttcatct gttgtgttt gtatgtatac agatctgaaa	6720
	acatcacttc tctcatctga ttgtgttgtt acatcacatg atatagatct gttatatcat	6780
	ttttttatt aatttgttat atatatatgt gcctagatct ggattacatg atttgtatta	6840
	ttcatatgat ttgtttattt acgtatgtat atatgtatg ctggactttt tggagtgttt	6900
	gacttgattg tatttgtgtg tctatattgt tttctgac ttgatattt atgtatgtc	6960
	aggcgcgcgc ggatcccccg gtgttcagtc cttatgtta cgtctgttag aaaccccaac	7020
	ccgtgaaatc aaaaaactcg acggcctgtg ggcattcagt ctggatcgcg aaaactgttg	7080
	aattgatcag cgttggtggg aaagcgcgtt acaagaaagc cgggcaattg ctgtgccagg	7140
	cagttttaac gatcagttcg ccgatgcaga taicgtaat tatgcgggca acgtctggta	7200
	tcagcgcgaa gcttttatc cgaagggtg ggcagccag cgtatctgtc tgcgtttcga	7260
	tgcggtcact cattacggca aagtgtgggt caataatcag gaagtgatgg agcatcaggg	7320
	cggctatacg ccatttgaag ccgatgtcac gccgtatgtt attgccgga aaagtgtacg	7380
	taagtctctg ctctctactt tgatataat ataataatta tcatiaatta gtagiaatat	7440
	aatattcaa atattttttt caaaataaaa gaatgtatga tatagcaatt gttttctgt	7500
	agtttataag tgtgtatatt ttaatttata acttttctaa tatatgacca aaatttgtt	7560
	atgtgcaggt atcaccgttt gtgigaacaa cgaactgaac tggcagacta tccgcgcggg	7620
	aatgggtatt accgacgaaa acggcaagaa aaagcagctt tacttccatg atttctttaa	7680
	ctatgcggga atccatcgca gegtaatgct ctacaccacg ccgaacacct ggggtggcga	7740
	tatcacctgt gtgacgcatg tcgcgaaga ctgtaacac gcgtctgtt actggcaggt	7800

[0031]

```

gggtggccaat ggtgatgtca gcgttgaact gcgtgatgcg gatcaacagg tggttgcaac 7860
tggacaaggc actagcggga ctltgcaagt ggtgaatccg cacctctggc aacgggtga 7920
aggttatctc tatgaactgt gcgtcacagc caaaagccag acagagtgtg atatctaccc 7980
gcttcggctc ggcaiccggt cagtggcagt gaaggcgcaa cagttccctga ttaaccacaa 8040
accgttctac ttacttggtt ttggtcgtca tgaagatgcg gacttgcgtg gcaaaggatt 8100
cgataacgtg ctgatgtgtc acgaccacgc attaatggac tggatigggg ccaacictca 8160
ccgtacctcg cattaccctt acgtgaaga gatgctcgac tgggcagatg aacatggcat 8220
cgfggtgatt gatgaaactg ctgctgtcgg cttaacctc tctttaggca ttggtttcga 8280
agcggggcaac aagccgaaag aactgtacag cgaagaggca gtcaacgggg aaactcagca 8340
agcgcacitc caggcgatta aagagctgat agcgcgtgac aaaaaccacc caagcgtggt 8400
gafgtggagt attgccaacg aacgggatac ccgtccgcaa ggtgcacggg aatatttcgc 8460
gccactggcg gaagcaacgc gtaaacctga cccgacgcgt ccgatacctt gcgtcaatgt 8520
aatgttctgc gacgctcaca ccgataccat cagcgatctc tttagtgtc tgtgcctgaa 8580
ccgttatctc ggaatggtatg tccaaaggcg cgatttggaa acggcagaga aggtactgga 8640
aaaagaactt ctggcctggc aggagaaact gcacagccg attatcatac cgaatacgg 8700
cgtggatacg ttagccgggc tgcactcaat gtacaccgac atgtggagtg aagagtatca 8760
gtgtgcatgg ctggatatgt atcaccgcgt ctctgategc gtcagcgccg tctcgggtga 8820
acaggtatgg aatttcgccg attttgcgac ctgcgaaggc atattgcgag ttggcggtan 8880
caagaaaggg atcttcactc gcgaccgcaa accgaagtcg gcggcttttc tgcgtcaaaa 8940
acgciggact ggcatgaaat tgggtgaaaa accgcagcag ggaggcaaac aatgaatcaa 9000
caactctcct ggcgcaacct cgtcggtctc agcctcggga attgctaccg agctcctgca 9060
ggcctaggat cgttcaaaaa ttigccaata aagtttctta agattgaatc ctgttgcggg 9120
ctltgcgatg attatcalat aatttctgtt gaattacgtt aagcaigtat taattaacat 9180
gtaatgcatt acgttatctc tgagatgggt ttttatgatt agagtccgcg aattatacat 9240
ttaatacgcg atagaaaaaa aaatatagcg cgcaaacatg gataaattat cgcgcgcggt 9300
gtcatctatg ttactagatc ggccggccgt ttaaacctag ttactaatca gtgatcagat 9360
tgtcgtttcc cgccttcaat ttaactatc agtgtttgac aggatattat ggccggtaaa 9420
cctaagagaa aagagcggtt attagaataa tgggatattt aaaaggcgct gaaaagggtt 9480
atccgttcgt ccatttgtat gtcaatatig gggggggggg aaagccacgt tgtgtctcaa 9540
aatctctgat gttacattgc acaagataaa aatataatcat catgaacaat aaaactgtct 9600
gcttacataa acagtaatac aagggtgtt cgcacacatg agccatatcc agcgtgaaac 9660
ctcgtgctcc cgcgccgcgc tcaattccaa tatggatgcc gacctttatg gctacaagtg 9720
ggcgccggac aacgtggccc agtcgggcgc gaccatttat cggttttatg gcaaaccgca 9780
tgccccggaa ctgttctga agcacggcaa aggcagcgtc gcaaacgatg tcaccgatga 9840
gatggtcgcg ctgaactggc ttaccgagtt catgccgtcg ccgacgatta agcatttcat 9900
ccgtacccec gacgatgcct ggctcttgac cacggccatt ccgggcaaaa cggcctttca 9960
ggctcctgaa gagtaccggg actccggcta gaatatcgtg gacgcccctg cggcttctct 10020
ccgccgtttg catagcatcc ccgtgtgcaa ctgccccttc aactcggacc gggttttccg 10080

```

[0032]

```

cctggcacag gccagtcgc gcatgaataa cggcctcgtt gacgcgagcg atttcgacga 10140
tgaacggaat ggctggccgg tggacaggt ttggaaggaa atgcacaaac tgcttcctgt 10200
ctcgccggat tcggttgta cgcattgta ttttccctg gataatctga tcttgacga 10260
gggcaagctg atcggctgca tcgacgtggg tcgctcgtt atcgcgcgac gctalcagga 10320
cctggcgatc ttgtggaatt gccctggcga gttctcgcgc tcgctccaga agcgcctgtt 10380
ccagaagtac ggcatcgaca acccgatgat gaacaagctc cagttccacc tcatgctgga 10440
cgaatttttt tgaacagaat tggtaattg gttgtaacac tggcagagca ttacgtgac 10500
ttgacgggac ggcgctttg ttgaataaat cgaactttg ctgagttgaa ggategatga 10560
gttgaaggac ccgtagaaa agatcaaagg atctcttga gatcctttt ttctgcgcgt 10620
aatctgtgc ttgcaacaa aaaaaccacc gctaccagcg gtggtttgtt tgccggaaca 10680
agagctacca actcttttc cgaaggtaac tggcttcagc agagcgaga taccaaatac 10740
tgtcttcta gtgtagcgt agttaggcca ccacttcaag aactctgtg caccgcctac 10800
atacctcgct ctgctaatac tgttaccagt ggcctgcgc agtggcgata agtctgtct 10860
taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg gctgaacggg 10920
gggttcgtgc acacagccca gcttgagcg aacgacctac accgaactga gatactaca 10980
gcgtgagcta tgagaaagcg ccacgttcc cgaagggaga aaggcggaca ggtatccgtt 11040
aagcggcagg gtcggaacag gagagcgac gaggagctt ccagggggaa acgcttgta 11100
tctttatagt cctgtcgggt ttcgccacct ctgactigag cgtcgatttt tgtgatgctc 11160
gtcagggggg cggagcctat ggaaaaacgc cagcaacgcg gcctttttac ggttcctggc 11220
cttttctggt ccttttgcic acatgttctt tcttgctta tccctgatt ctgtggataa 11280
ccgtattacc gcctttgagt gagctgata cgtcgcgcgc agccgaacga ccgagcgag 11340
cgagtcagtg agcaggaag cgaagagcg cctgatcgcg tattttctcc ttacgatct 11400
gtgcggtatt tcacacgca taggcgcga taggcgcac cgaagcggcg ggcgttaggg 11460
agcgcagcga ccgaagggtt ggcgttttt gcagctctc ggctgtgcgc tggccagaca 11520
gttatgcaca ggccagcgg gtttaagag tttaataag tttaaaagag ttttagggg 11580
aaaaatcgcc tttttctct tttatctag tcaattacat gtgtgaccgg ttcccaatgt 11640
acggctttgg gtcccaatg tacgggttcc ggttcccaat gtacggcttt gggttccaa 11700
tgtacgtgct atccacagga aagagacctt ttgacctt ttccctgct agggcaattt 11760
gccctagcat ctgcctccta cattaggaac cggcggatgc ttccctcgc atcaggttgc 11820
ggtagcgcat gactaggatc gggccagcct gccccgcctc ctcttcaaa tegtactcgc 11880
gcagtcatt tgaccgata agcttgcga cggtaaaaca gaactcttg aactctcgg 11940
cgctgccact gcgttcgtg atcgtctga acaacctct ggcttctgce ttgcctcgg 12000
cgcggcgtgc caggcggtag agaaaacggc cgtatgcggg gtcgatcaaa aagttaatcg 12060
ggtgaacctg cagcacgtcc ggttcttgc ctctgtgat ctgcggtaac atccaatcag 12120
caagctcgat ctgatgtac tccggccgce cggtttcgt ctttaccata ttgtagcggc 12180
taatcaagcg ttcaacctcg gataccgtca ccaggcggcc gttcttggcc ttcttggta 12240
gtcgtatgce aacgtgcgtg gtgtttaacc gaatgcaggt ttctaccagg tegtcttct 12300
gttttccgce atcggtcgc cggcagaact tgagtacgtc cgcaacgtgt ggacggaaca 12360

```

cgcggccggg	cttgtctccc	ttccctccc	ggatcgggtt	catggattcg	gttagatggg	12420	
aaaccgcat	cagtaccagg	tcgtaatccc	acacactgga	catgccggcg	gggcctgcgg	12480	
aaacctctac	gtgcccgtct	ggaagctcgt	agcggatcac	ctgccagct	cgicggtcac	12540	
gcttcgacag	acggaaaacg	gccacgtcca	tgatgctgcg	actatcgagg	gtgcccacgt	12600	
catagagcat	cggaaacgaa	aaatctgggt	gctcgtgcc	cttgggcggc	ttcctaatac	12660	
acggcgcaac	ggctgcggcg	ggttgcgggg	attctttgcg	gattcgatca	gcggcccctt	12720	
gccacgattc	accggggcgt	gcttctgcct	cgatgcgttg	ccgctggggc	gcctgcggcg	12780	
cttcaacttt	ctccaccagg	tcatacccca	gcgcgcgcgc	gatttgtacc	gggcgggatg	12840	
gtttgcgacc	gctcacgccg	attcctcggg	cttgggggtt	ccagtcccat	tgcagggccg	12900	
gcagacaacc	cagccgccta	cgcttgcca	accgcccgtt	ctccacaca	tggggcattc	12960	
cacggcgtcg	gtgcctgggt	gttttgatt	ttccatgccg	ctctctttag	ccgctaaaat	13020	
tcactacttc	atttattcat	ttgctcattt	actctgtag	ctgcgcgatg	tattcagata	13080	
gcagctcggt	aatggtcttg	cttggcgta	ccgcgtacat	cttcagcttg	gtgtgatcct	13140	
ccgcggcnaa	ctgaaagtig	accgccttca	tggctggcgt	gtctgccagg	ctggccaacg	13200	
ttgcagcctt	gctgctgcgt	gcgctcgga	ggccggcact	tagcgtgttt	gtgcttttgc	13260	
tcattttctc	ttacctcat	taactcaaat	gagttttgat	ttaatttcag	cggccagcgc	13320	
ctggaectcg	cgggcagcgt	cgccctcggg	ttctgattca	agaacggttg	tgccggcggc	13380	
ggcagtgctt	gggtagctca	cgcgtgcgt	gatacgggac	tcaagaatgg	gcagctcgta	13440	
cccggccagc	gcctcgga	ctcaccgc	gatgcgcgtg	ctttgatcg	cccgcgacac	13500	
[0033]	gacaaaggcc	gcttgtagcc	ttccatccgt	gacctcaatg	cgctgcttaa	ccagctccac	13560
	caggtcggcg	gtggcccaaa	tgtcgtaagg	gcttggctgc	accggaatca	gcacgaagtc	13620
	ggctgccttg	atcgccgaca	cagccaagtc	cgccgcctgg	ggcgtccgt	cgataactac	13680
	gaagtccgc	cggccgatgg	ctttcacgtc	gcggtcaatc	gtcggcggt	cgatgccgac	13740
	aacggttagc	ggttgatctt	ccgcacggc	cgcccaatcg	cgggcactgc	cttggggatc	13800
	ggaatcgact	aacagaacat	cggccccggc	gagttgcagg	gcgcgggcta	gatgggttgc	13860
	gatggtcgtc	ttgcctgacc	cgcctttctg	gttaagtaca	gcgataacct	tcatgcgttc	13920
	cccttgcgta	ttgttttatt	tactcatcgc	atcatatacg	cagcgaccgc	atgacgcaag	13980
	ctgttttact	caaatacaca	tcaccttttt	agatgatca			14019

<210> 76

<211> 28

<212> DNA

<213> 人工的

<220>

<223> 合成的序列

<400> 76

atatacgctt ttagatctcg tgcgctcg

28

<210> 77

<211> 32

<212> DNA

<213> 人工的

<220>

<223> 合成的序列

	<400> 77 atatggcgcg cctttagatca agcctgttca ca	32
	<210> 78 <211> 33 <212> DNA <213> 人工的 <220> <223> 合成的序列	
	<400> 78 atatacgcgt ttaagaaatc ctctcttctc etc	33
	<210> 79 <211> 38 <212> DNA <213> 人工的 <220> <223> 合成的序列	
	<400> 79 atatggcgcg cctgcacat acataacata tcaagatc	38
	<210> 80 <211> 13590 <212> DNA <213> 人工的 <220> <223> 载体序列	
[0034]	<400> 80 gtgattttgt gccgagctgc cggtcgggga gctgttggct ggctggtggc aggatatatt gtggtgtaaa caaattgacg cttagacaac ttaataaac acattgcggacg tctttaatgt actgaattta gttactgac actgattaa tactgatac ggtaccaagc ttccgcggct gcagtgacg gtgacccggt cgtgccctc tctagagata atgagcattg catgtctaag ttataaaaaa ttaccacata tttttttgt cacacttgtt tgaagtgcag ttatctatc ttataacata tatttaaac ttaactctac aataatataa tctatagtac tacaataata tcagtgtttt agagaatcat ataaaagaac agttagacat ggtctaaagg acaattgagt attttgacaa caggactcta cagttttatc tttttagtgt gcattgtgtc tcctttttt ttgcaaatag cttcacetat ataatactc atccatttta ttagtacatc catttagggt ttagggttaa tggtttttat agactaattt ttttagtaca tctattttat tctattttag cctctaaatt aagaaaacta aaactctati ttagttttti tatttaatag tttagatata aaatagaata aaataaagtg actaaaaatt aaacaaatac cttttaangaa attaaaaaaa ctaaggaaac atttttcttg ttctgagtag ataattgccag cctgttaaac gccgtcgacg agtetaacgg acaccaacca gcgaaccagc agcgtcgct cgggccaagc gaagcagacg gcacggcacc tctgtcgctg cctctggacc cctctcgaga gtcccgctcc accgttggac ttgctecgct gtcggcatcc agaaattgct tggcggagcg gcagacgtga gccggcacgg caggcggcct cctcctctc tcaaggcacc ggcagctacg ggggattcct ttcccaccgc tctctcgctt tccctctcct gcccgccgta ataaatagac acccctctca caccctctt ccccaacctc gtgtgttcg gagegcacac acacacaacc agatctcccc caaatccacc	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140

	cgctggcacc tccgcttcaa ggtacgcgcg tcgtcctccc cccccccc cctctctacc	1200
	ttctctagat cggcgttccg gtccatgggt agggcccggt agttctactt ctgttcatgt	1260
	ttgtgttaga tccggttttg tgttagaicc gtgtgtctag cgttcgtaca cggatgcgac	1320
	ctgtacgtca gacacgttct gattgtctaac ttgccagtgt ttctcttttg ggaatccctg	1380
	gatggctcta gccgttccgc agacgggacg gatttcatga ttttttttgg ttctgttcat	1440
	agggttttgg ttgccctttt cctttatttc aatataatgc gtgcacttgt ttgtcgggtc	1500
	atcttttcat gctttttttt gtcttggttg tgatgatgtg gtctggttgg gcggtcgttc	1560
	tagatcggag tagaattctg ttcaaaccta cctgggtggat ttattaatit tggatctgia	1620
	tgtgtgtgcc atacatatc atagttaaga attgaagatg atggatggaa atatcgatct	1680
	aggataggta tacatgttga tgcgggtttt actgatgcac atacagagat gctttttgtt	1740
	cgtttgttgg tgatgatgtg gtgtgttgg gcggtcgttc attcgttcta gatcggagta	1800
	gaatactgtt tcaaacctac tgggtatatt attaatitg gaactgtatg tgtgtgtcat	1860
	acaacttcat agttacgagt ttaagatgga tggaaatate gatctaggat aggtatacat	1920
	gttgafttgg gtttttacta tgcatafaca tgatggcata tgcagcatct attcatatgc	1980
	tctaaccttg agtaacctac tattataata aacaagtatg tttataaatt atttcgatct	2040
	tgatatactt ggtatgatgc atatgcagca gctatagtgt gattttttta gccctgcctt	2100
	catacgctat ttatttgcct ggtactgttt cttttgttga tgcctacctt gttgtttggt	2160
	gttactttct cagcccgggg gatccactag ttctagaaac catggccacc gcccccgccg	2220
	cgctctaccg gctcacttgc gccactaccg ctgcgcccaa ggcgaggcgc cgggcgcacc	2280
[0035]	tccttgccac ccgcccgcgc ctcccgccgc ccacaggtg ctccagcgcg tcaccgccta	2340
	tgcgcagtgc tcccccgccc accccgctcc ggccttgggg ccccaccgat ccccgcaagg	2400
	gcgccgacat cctcgtcgag tcccctgagc gctgcggcgt ccgcgacgtc ttcccttacc	2460
	ccggcgccac gtccatggag atccaccagg cactcaccgc ctcccccgtc atcgccaacc	2520
	acctcttccg ccacgagcaa ggggaggcct ttgcggcctc cggtacgcg cgtctctcgg	2580
	gcccgctcgg cgtctgcate gccacctcgc gccccgcgc caccacactt gtctccgcgc	2640
	tcgccgacgc gctgctcgat tccgtcccca tggctgccat caccggacag gtcccgac	2700
	gcatgatttg caccgacgcc ttccaggaga cgcctatcgt cgaggtcacc cgtctcatca	2760
	ccaagcacia ctacctgttc ctccagctgc acgacatccc ccgcgtcgtg caggaggctt	2820
	tcttctcgc ctctcttggc cagccggggc cgttcttctt cagacatccc aaggacatcc	2880
	agcagcagat ggcggtgcct gtctgggaca agcccatgag tctgcctggg tacattgcgc	2940
	gccttcccaa gccctctgc actgagtgc ttgagcaggt gctgcgtctt gttgttgaat	3000
	ccccgcgccc tgttctttat gttggcgggt gctgcgcagc atctggtgag gatttgcgac	3060
	gctttgttga gctgaactga atcccggtca caactactct tatgggcctc ggcacattcc	3120
	ccagcgacga cccactgtct ctgcgcagc taggtatgca tggcacgggt tatgcaaatt	3180
	atgcagtgga taaggccgat ctgttcttgc cacttgggtg gcggtttgat gatcgttga	3240
	cagggaagat tgaggctttt gcaagcaggg ctaagattgt gcacgttgat attgatccgg	3300
	ctgagatttg caagaacaag cagccacatg tgtccatctg tgcagatgtt aagcttgcct	3360
	tgcagggcat gaatgtctct ctigaaggaa gcacatcaaa gaagagcttt gactttggct	3420

[0036]

```

catggaacga tgagttggat cagcagaaga gggaaattccc ccttgggtat aaaacatcta 3480
atgaggagat ccagccacaa tatgctattc aggttcttga tgagctgacg aaaggcgagg 3540
ccatcatcgg cacaggigtg ggcagcacc agatgtgggc ggcacagtac tacacttaca 3600
agcggccaag gcagtgggtg tcttcagctg gtcttggggc tatgggattt ggtttgccgg 3660
ctgctgctgg tgccttcttg gccaaaccag gtgttactgt tgttgacatc gatggagatg 3720
gtagctttct catgaacgtt caggagetag ctatgatccg aattgagaac ctcccgtga 3780
aggtctttgt gctaaacaac cagcacctgg gcatgggtgt gcagtgggag gacaggttct 3840
ataaggccaa cagagcgcac acatacttgg gaaaccaga gaatgaaagt gagatatc 3900
cagatttctg gacgatgcgc aaaggttca acattccagc ggtccgtgtg acaaagaaga 3960
acgaagtcgg cgcagcgata aagaagatgc tcgagatccc agggccgtac ccttggata 4020
taatcgtccc acaccaggag catgtgttgc ctatgatccc taatggtggg gctttcaagg 4080
atatgacctt ggatggtgat ggcaggactg tgtactgate taaaatccag caagcaactg 4140
atctaaaatc cagcaagcac cgcctcccctg ctagtacaag ggtagatgtt ttttctctgt 4200
gtgagtctct cctgtattct atcttttttt gtaggcctgc agctatctgt tatgtaatc 4260
ctatgtaget tccgaccttg taattgtgta gtctgttgtt ttcttcttg catgtgtcat 4320
aagagatcat ttaagtgcct ttgtctacat ataaataaga taataagcac tgcgatgcag 4380
tggttctgaa ttggcttctg ttgccaaatt taagtgtcca actggctcctt gcttttgttt 4440
tcgtattttt ttctcttttt tagttattat tatattggta atttcaactc aacataigat 4500
gtatggaata atctaggggc tgcaatttca aactatttta caaaccagaa tggcatttct 4560
gtggtttgag gggagtgaaa aaaaatgagg calttgactg aattagttac ctgatccatt 4620
ttctgtgttt ggatcattgg aattaaatc cattctaata atagtaattt tggcatatat 4680
caattaaagt aattcggttt tatgcaaaat atatttgtat actattattt tcaagatgtc 4740
ggagatatit atatgtaca tttttactat acaggagtga gatgaagagt gtcatgtaag 4800
ttacacagta gaaacaaatt ctattaatgc ataaaatcat ttccatcact caccctatga 4860
atttgagata gacctatctc taaactttga aaagtgttg aatatcaaat tccaaattaa 4920
ataagttatt ttattgagtg aattctaatt tctctaaaac gaagggatct aaacgcctc 4980
taaagctaat ttggaaactc aaactttctt agcattggag gggattgaga aaaaatatta 5040
attcattttc atctcaatca ttcaatctcc aaagagattt gatttcttta ttagtctgtt 5100
ccatgcatac aatcggctca atggttcatt atttgcattg acgattgacg agttgttctg 5160
gggcctagcg ctctccacgc cgtgtgtctg gggcctgggc ctggagaaga cagcttgata 5220
tttaaagcta tcaattgttt caattgattc ccacttcatt ttctaaatg tagaaaacgg 5280
tgacgtataa gaaaaagaat gaattaggac tttatttccg tacactaatc tagagcggcc 5340
caagcttgta cactagtacg cgtcaattga tttaaattta attaatgcag tgcagcgtga 5400
cccgtctgtg cccctctcta gagataatga gcattgcatg tctaagttat aaaaattac 5460
cacatatit tttgttcaca ctgtttgaa gtgcagtita tctatcttta tacatatatt 5520
taaaccttac tctacgaata atataatcta tagtaactaca ataatatcag tgttttagag 5580
aatcatataa atgaacagtt agacatggtc taaaggacaa ttgagtattt tgacaacagg 5640
actctacagt tttatctttt tagtgtcat gtgttctctt tttttttgc aaatagcttc 5700

```

[0037]

```

acctatataa tacttcatcc attttattag tacatecatt tagggtttag ggtaaatggt 5760
ttttatagac taattttttt agfacateta ttttttcta ttttagcctc taaattaaga 5820
aaactaaaaa tctatttttag tttttttatt taatagttta gatataaaat agaataaaat 5880
aaagtgacta aaaallaaac aaataccctt taagaaalla aaaaaaclaa ggaacattt 5940
ticttgtttc gagtagataa tgccagcctg ttaaacgccg tcgacgagtc taacggacac 6000
caaccagcga accagcagcg tcgcgtcggg ccaagcgaag cagacggcac ggcatctctg 6060
tcgctgcctc tggacccttc tcgagagttc cgtccaccg ttggacttgc tccgctgtcg 6120
gcattcagaa attgcgtggc ggagcggcag acgtgagcgg gcacggcagg cgccctcctc 6180
ctctctcac gccaccggca gctacggggg attcctttcc caccgtctct tcgctttccc 6240
ttctctgccc gccgtaataa atagacaccc cctccacacc ctctttcccc aacctcgtgt 6300
tgttcggagc gcacacacac acaaccagat cggcgcgttt aagaaatcct ctcttctct 6360
cttcattttc aaggtaaatc tctctctctc tctctctctc tgttattcct tgttttaatt 6420
aggtatgtat tattgctagt ttgttaatct gcttatctta tgtatgcctt atgtgaatat 6480
ctttatcttg ttcattctcat ccgtttagaa gctataaatt tgttgatttg actgtgtatc 6540
tacacgtggt tatgtttata tctaateaga tatgaatttc ttcattttgt tgcgtttgtg 6600
tgtaccaatc cgaatcgtt gatttttttc attaatcgt gtagetaatt gtacgtatac 6660
atatggatct acgtatcaat tgttcatctg tttgtgttg tatgtatata gatctgaaaa 6720
cactacttct ctcatctgat tgtgttgta cactacataga tatagatctg ttatatcatt 6780
ttttttatta attgtgtata tatatatgtg catagactg gattacatga ttgtgattat 6840
ttacatgatt ttgttttta cgtatgtata tatgtagatc tggacttttt ggagtgttg 6900
acttgatigt atttgtigt gtatatgtgt gtctgatct tgatatgtta tgtatgtgca 6960
gggcgcgcca ccatggaaga cgccaaaaac ataaagaaag gcccgscgcc attctatccg 7020
clggaagaig gaaccgcig agagcaacig cataaggcta tgaagagata cggcclggtt 7080
cctggaacaa ttgcttttac agatgcacat atcgaggtgg acatcactta cgtgagtac 7140
ttcgaatgt ccgttcggtt ggcagaagct atgaaacgat atgggctgaa tacaateac 7200
agaatcgtcg tatgcagtga aaactctctt caattcttta tgcgggtgtt gggcgcgta 7260
tttatcgag ttgcagtgc gcccgcaac gacatttata atgaacgtga attgctcaac 7320
agtatgggca tttcgcagcc taccttggtg ttcttttcca aaaaggggtt gcaaaaaatt 7380
ttgaacgtgc aaaaaaagct ccaatcacc caaaaaatta ttatcatgga ttctaaaacg 7440
gattaccagg gatttcagtc gatgtacacg ttctgtacat ctcatctacc tcccggtttt 7500
aatgaatacg attttgtgcc agagtccttc gatagggaca agacaattgc actgatcatg 7560
aactctctcg gatctactgg tctgcctaaa ggtgtcgtc tcctcatag aactgcctgc 7620
gtgagattct cgcattgccg agatctctat ttggcaate aaatcattcc ggatactgcg 7680
attttaagtg ttgttcatt ccatcacggt ttggaatgt ttactacact cggatatitg 7740
atatgtggat ttcgagtcgt cttaatgtat agatttgaag aagagctgtt tctgaggagc 7800
cttcaggatt acaagattca aagtgcctg ctggtgccaa cctattctc ctctctcgcc 7860
aaaagcactc tgatigacaa atacgattta tctaatttac acgaaattgc ttctgggtgc 7920
gtctccctct ctaaggaagt cggggaagcg gttgccaa ggttccatct gccaggtatc 7980

```

[0038]	aggcaaggat atgggctcac tgagactaca tcagctattc tgattacacc cgagggggat	8040
	gataaaccgg gcgcggtcgg taaagttgit ccattttttg aagegaaggi tgggatctg	8100
	gataccggga aaacgtcggg cgtaataca agaggcgaac tgtgtgtgag aggtcctatg	8160
	attatgtccg gttatgtaaa caatecggaa gcgaccaacg ccttgattga caaggatgga	8220
	tggctacatt ctggagacat agcttactgg gacgaagacg aacacttcct catcgttgac	8280
	cgcctgaagt ctctgattaa gtacaaaggc tctcaggtgg cccccgtga attggaatec	8340
	atcttgctec aacaccccaa catcttcgac gcaggtgtcg caggtcttcc cgacgatgac	8400
	gccggtgaac ttcccgcgcg cgttgttgtt ttggagcacg gaaagacgat gacggaaaaa	8460
	gagatcgtgg attacgtcgc cagtcaagta acaaccgcga aaaagttgcg cggaggagtt	8520
	gltgttggg acgaagtcac gaaaggctct accggaaaaa tcgacgcaug aaaaatcaga	8580
	gagatcccca taaaggccaa gaaggcgga aagatcgccg tgtaacctgc aggcctagga	8640
	tcgttcaaac atttggcaat aaagtttctt aagattgaat cctgttgcg gtcttgcgat	8700
	gattatcata taatttcigt tgaattacgt taagcatgta ataattaaac tgtaatcat	8760
	gacgttatit atgagatggg tttttatgat tagagtcccg caattatata tttaatacgc	8820
	gatagaaaac aaaatatagc gcgcaacta ggataaatta tcgcgcgcgg tgcctctat	8880
	gttactagat cggccggccg tttaactta gttaacta acgtgacaga ttgtcgttcc	8940
	ccgccttcac tttaactat cagtgtttga caggtatata tggcgggtaa acctagaga	9000
	aaagagcgtt tattagaata atcggtatatt taaaaggcgg tgaanaaggt tatccgttcg	9060
	tcattttgta tgcataatatt gggggggggg gaaagccacg ttgtgtctca aaatctciga	9120
	tgttacattg cacaagataa aaatatata tcatgaacaa taaaactgtc tgcctacata	9180
	aacagtaata caaggggtgt tcgccaccat gagccatata cagcgtgaaa cctcgtgctc	9240
	ccgccgcgcg ctcaattcca atatggatgc cgacctttat ggtacaaagt gggcgcgcga	9300
	caacgtcgcg cagtcggcgc cgaccattta tcggccttat ggcaaacccg atgccccgga	9360
	actgttcttg aagcacggca aaggcagcgt cgcaaacgat gtcaccgat agatggtcgc	9420
	cctgaactgg cttaaccagt tcatgcgcgt gccgacgatt aagcatttca tcgtacccc	9480
	ggacgatgcc tggctcttga ccacggccat tcggggcaaa acggccttcc aggtccttga	9540
	agagtaccgg gactccggtg agaatatcgt ggacgccttc gcggtcttcc tcgcccgttt	9600
	gcatagcata cccgtgtgca actgcctctt caactcggac cgggttttcc gcctggcaca	9660
	ggcccagtcg cgcataata acggccctgt tgacgcgagc gatttcgacg atgaacggaa	9720
	tggctggccg gtggaacagg ttggaagga aaigcacaac ctgcttccgt tctgcgcgga	9780
	ttcgggtggt acgcatggtg atttttccct ggataatctg atctttgacg agggcaagct	9840
	gatcggtctg atcgacgtgg gtcgcgtcgg tctgcgcgac cgtatcagg acctggcgat	9900
	cttgtgaat tgcctcggcg agttctcgcg ctgctccag aagcgcctgt tcagaagta	9960
	cggcatcgac aaccgggata tgaacaagct ccagttccac ctcatgctgg acgaattttt	10020
	ttgaacagaa ttggttaatt ggttgtaaca ctggcagagc attacgctga ctgacggga	10080
	cggcgcgttt gttgaataaa tcgaactttt gctgagttga aggatcgatg agttgaagga	10140
	ccccgtagaa aagatcaaaag gatcttcttg agatcttttt ttctcgcgcg taatctcgtg	10200
	cttgcacaaa aaaaaaccac cgtacacgac ggtggtttgt ttgccggatc aagagctacc	10260

	aactcttttt ccgaaggtaa ctggtttcag cagagcgcag ataccaaata ctgtcettct	10320
	agtgtagccg tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc	10380
	tctgttaate ctgttaccag tggctgcigc cagtggcgat aagtcgtgic ttaccgggtt	10440
	ggatlcgaaga cgatagllac cggataaggc gcagcgcigc ggctgaacgg ggggllcgtg	10500
	cacacagccc agcttggagc gaacgacctt caccgaactg agatacctac agcgtgagct	10560
	atgagaaage gccacgcttc ccgaaggag aaaggcggac aggtatccgg taagcggcag	10620
	ggtcggaaaca ggagagcgca cgaggagct tccaggggga aacgcctggt atctttatag	10680
	tcctgtcggg tttegccacc tctgactiga gcgtcgattt ttgtgatgct cgtcaggggg	10740
	gcggagccta tggaaaaacg ccagcaacgc ggcttttita cggttcctgg ctttttctg	10800
	gccttttict cacatgttct ttctgcgtt atccccgat tctgttgata accgtattac	10860
	cgccttttag tgagctgata ccgtcgcgc cagccgaacg accgagcgca gcgagtcagt	10920
	gagcgaggaa gcggaagage gcctgatgcg gtattttctc cttacgcate tgtcgggtat	10980
	ttcacaccgc atagcccgcg atagcccgac gcgaagcgge ggggcgtagg gagcgcagcg	11040
	accgaagggt aggcgctttt tgcagctctt cggctgtgcg ctggccagac agttatgcac	11100
	aggccaggcg ggttttaaga gttttaataa gttttaaga gttttaggcg gaaaaatcgc	11160
	cttttttctc ttttataatc gtcacttaca tgtgtgaccg gttcccaatg tacggcttfg	11220
	ggttcccaat gtacgggttc cggttcccaa tgtacggctt tgggttccca atgtacgtgc	11280
	tatecacagg aaagagacct tttegaectt ttccccctgc tagggcaatt tgccttagca	11340
[0039]	tcgtctccgt acattaggaa ccggcggatg cttegccttc gatcaggttg cggtagcgca	11400
	tgactaggat cgggcccagc tgccccctct cctecttcaa atcgtactcc ggcaggctat	11460
	ttgaccgat cagcttgcgc accgtgaaac agaacttctt gaactctccg gcgtgccac	11520
	tgcgttcgta gatcgtcttg aacaaccate tggcttctgc cttagcctgc gcggcgctg	11580
	ccaggcggia gagaaaaacg ccgalgccg ggctgatcaa aaagtaatcg gggfgaaccg	11640
	tcagcacgtc cgggttcttg cttctctgta tctcgcggta catccaatca gcaagctega	11700
	tctcgtatga ctccggccgc ccggtttcgc tccttacgat ctgttagcgg ctaatcaagg	11760
	ctteaccctc ggataaccgt accagcgccg cgttcttggc cttcttggta cgcctgcatg	11820
	caacgtgcgt ggtgtttaac cgaatgcagg ttcttaccag gtcgtcttct tgccttccgc	11880
	catcggtctg ccgagcagaac ttgagtacgt ccgcaacgtg tggacggaac acgcggccgg	11940
	gcctgtctcc ctteccctcc cggatcgggt tcatggatc ggttagatgg gaaaccgcca	12000
	tcagtaccag gtcgtaatcc cacacactgg ccatgccggc ggggcctcgc gaaacctcta	12060
	cgtgcccgtc tggaaagctg tagcggatca cctcgccagc tcgtcggtea cgtctgaca	12120
	gacggaaaaa ggccacgtcc atgatgtgc gactatcgc ggtgcccaac tcatagagca	12180
	tcggaacgaa aaaatctggt tgcctcgcgc ctttgggcgg cttectaate gacggcgcac	12240
	cggctgccgg cggttgccgg gattctttgc ggattcgatc agcggccctt tgccacgatt	12300
	caccggggcg tgccttctgc tcgatgcgtt gccgctgggc ggcttgcgcg gccttcaact	12360
	tcctcaccag gtcatcaccc agcgcgcgc cgatttgtac cgggcgggat ggtttgcgac	12420
	cgtcaccgce gattctctcg gcttgggggt tccagtgcga ttgcagggcc ggcagacaa	12480
	ccagccgctt acgcctggcc aaccgccgtt tcctccacac atggggcatt ccacggcgtc	12540

[0040]

```

ggtgcctggt tgttcttgat ttccatgcc gcctccttta gccgctaaaa ttcacttact 12600
catttattea ttgtctcatt tactctggta gctgegcgat gtattcagat agcagctcgg 12660
taatggtctt gccctggcgt acegcgtaca tcttcagctt ggtgtgatcc tccgccggca 12720
actgaaagtt gaccegttcc atggctggcg tgtctgccag gctggccaac gttgcagcct 12780
tgctgctgcg tgcgctcgga cgcccgccac ttagecgtgt ttgtcttttg ctcatittct 12840
ctttacctea ttaactcaaa tgagttttga tttaatttea gcggccagcg cctggacctc 12900
gcgggcagcg tcgcccctgg gttctgattc aagaacggtt gtgccggcgg cggcagtgcc 12960
tggttagctc acgcgctcgc tgatacggga ctcaagaatg ggcagctcgt acccgccag 13020
cgctcggca acctcaccgc cgatgcgcgt gcctttgate gcccgcgaca cgacaaagcc 13080
cgctttagc cttccatccg tgacctcaat gcctgctta accagctcca ccaggtcgcc 13140
ggtggcccaa atgtcgtaag gcttggctg caccggaatc agcaagaagt cggctgcctt 13200
gatcgccgac acagccaagt ccgccgctg gggcgctccg tcgatcacta cgaagtcgcg 13260
ccggccgatg gccctcacgt cgggtcaat cgtcgggcgg tcgatccga caacggttag 13320
cggttgatct tcccgcacgg ccgcccaatc gcgggcactg ccctggggat cggaaatcac 13380
taacagaaca tcggcccccg cgagttgcag ggcgcggt agatgggttg cgatggtcgt 13440
cttgccctgac ccgcctttct ggttaagtac agcgataacc ttcatgcgtt ccccttgctt 13500
atttgtttat ttactcatcg catcatatac gcagcgaccg catgacgcaa gctgttttac 13560
tcaatacac atcaccitit tagatgatca 13590

```

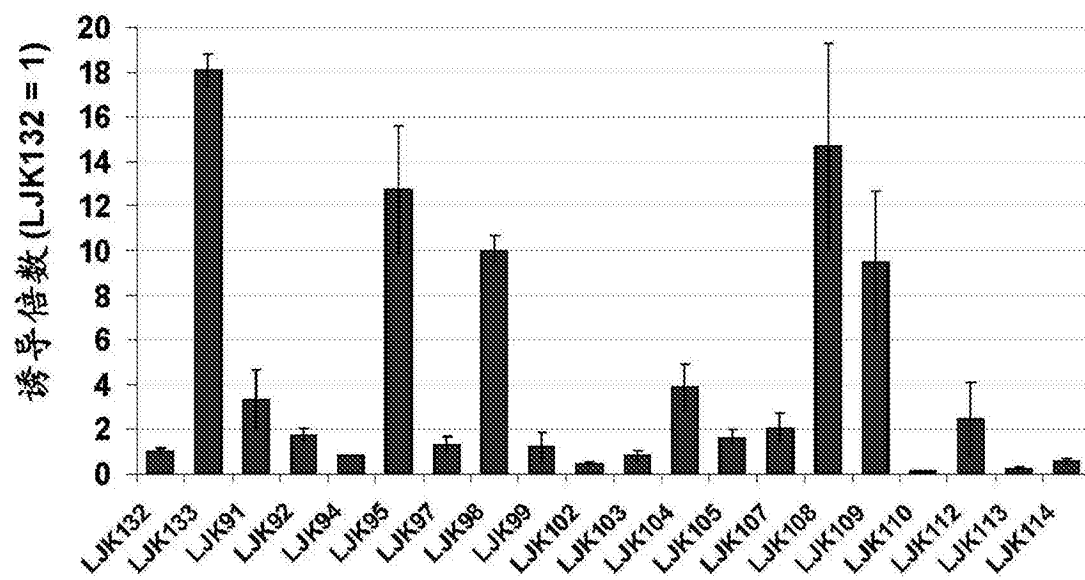


图1

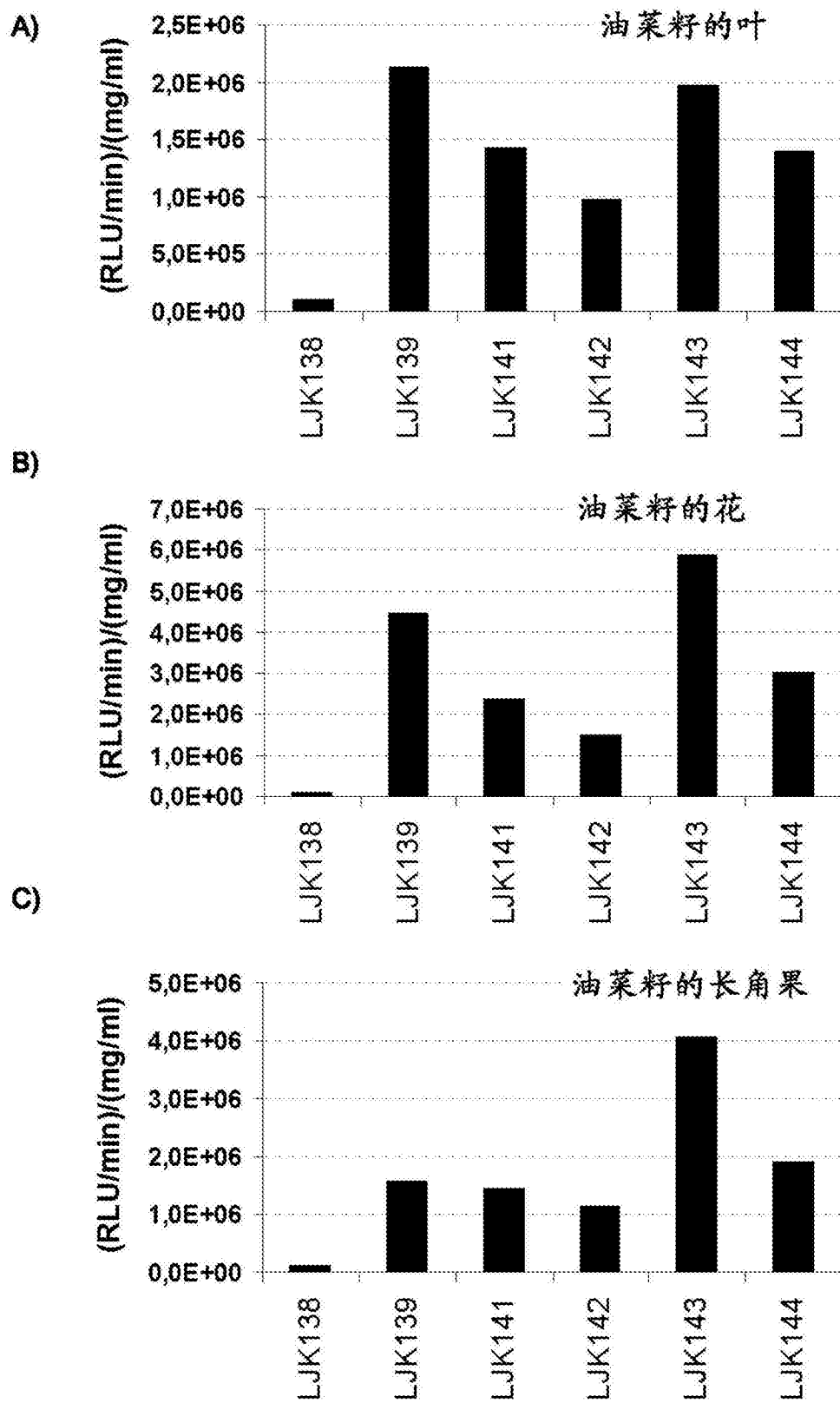


图2

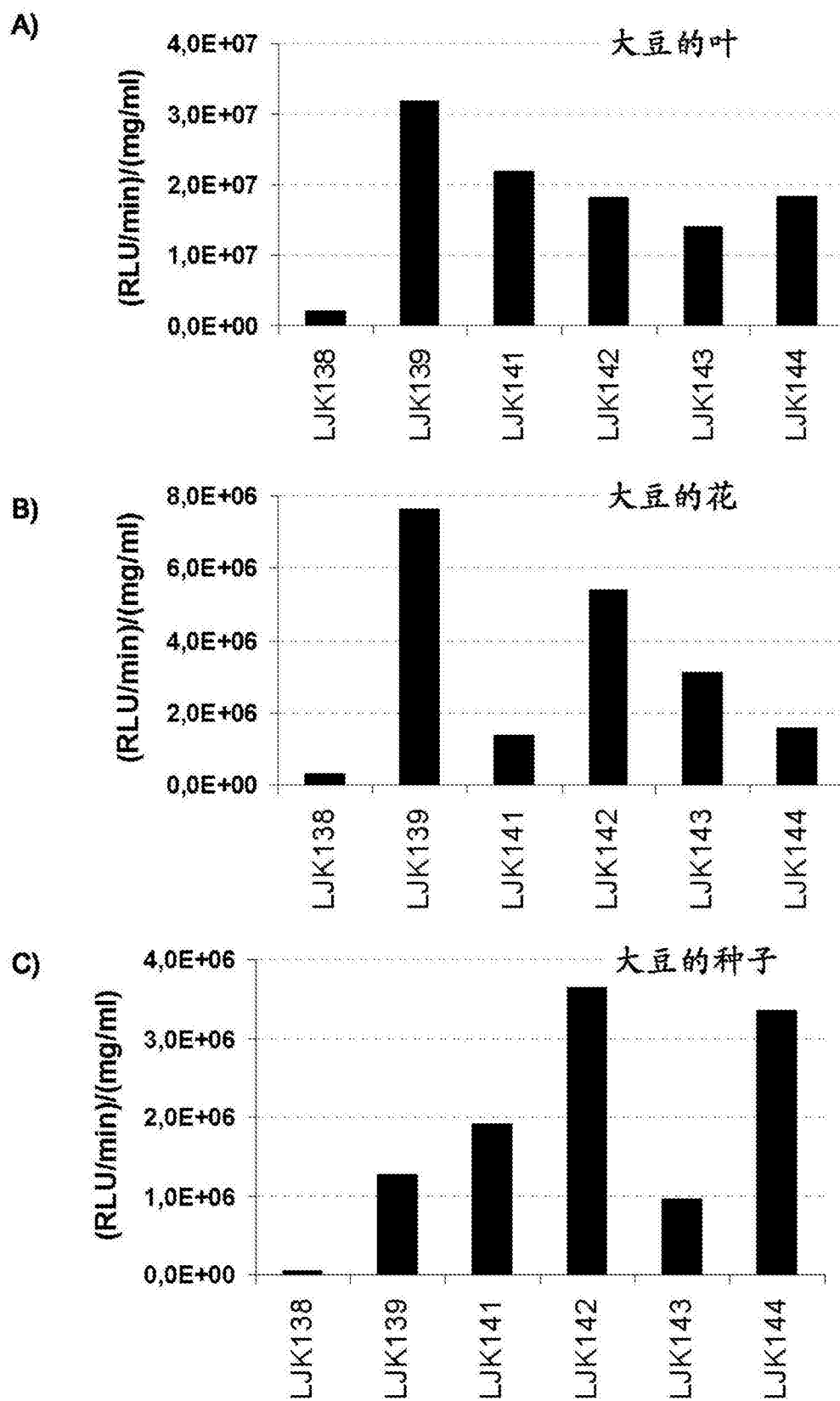
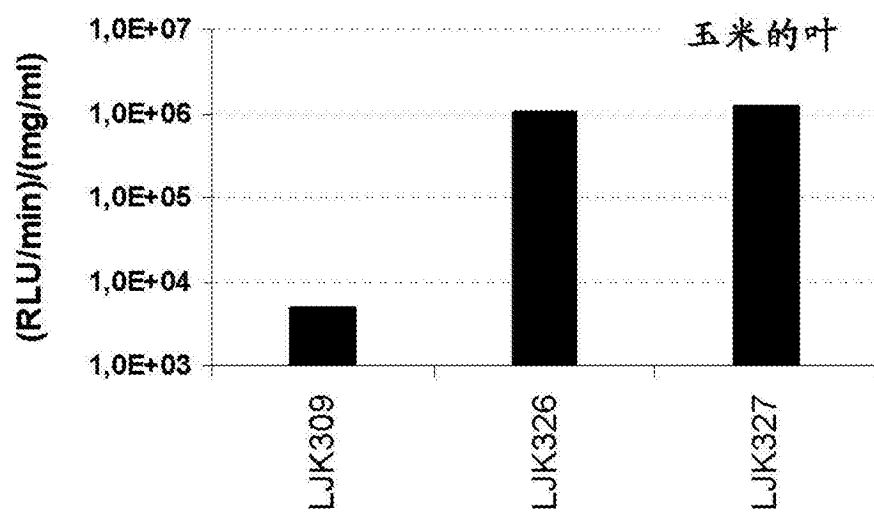


图3

A)



B)

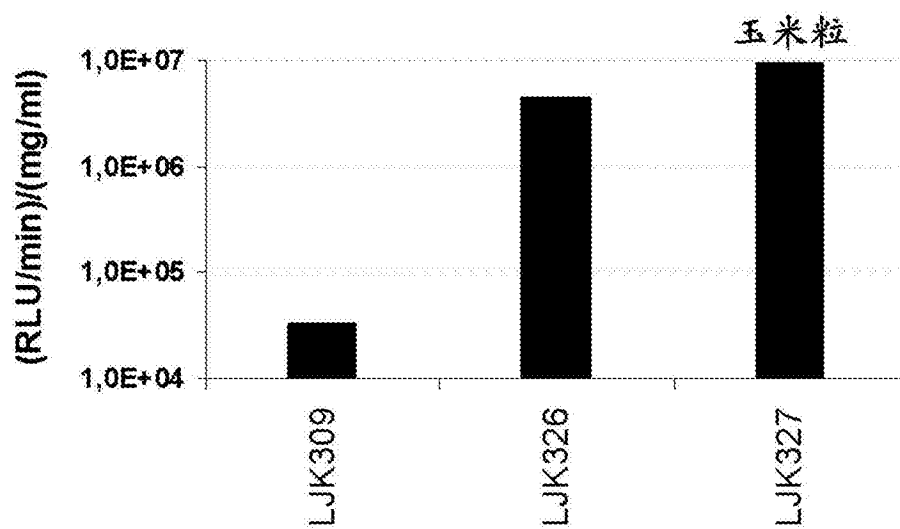


图4

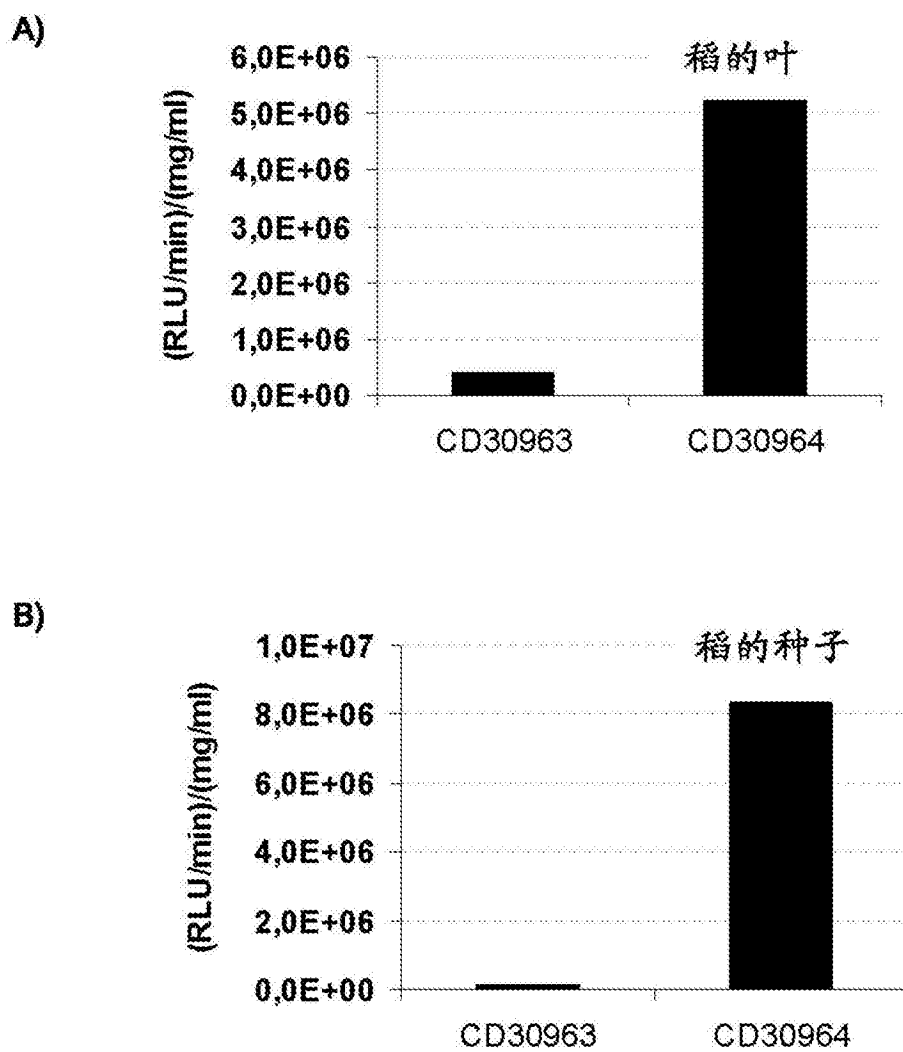


图5