



## (12)发明专利申请

(10)申请公布号 CN 107699563 A

(43)申请公布日 2018.02.16

(21)申请号 201710949095.5

*C12N 15/56*(2006.01)

(22)申请日 2014.03.12

*A01H 5/00*(2018.01)

(30)优先权数据

*A01H 5/10*(2018.01)

61/785,245 2013.03.14 US

(62)分案原申请数据

201480026708.2 2014.03.12

(71)申请人 孟山都技术有限公司

地址 美国密苏里州

(72)发明人 S·弗拉辛斯基 章军 赵淑玲

(74)专利代理机构 北京市金杜律师事务所

11256

代理人 孟凡宏 王月

(51)Int.Cl.

*C12N 15/113*(2010.01)

权利要求书2页 说明书27页

*C12N 5/10*(2006.01)

序列表25页 附图3页

(54)发明名称

植物调控元件和其用途

(57)摘要

本发明提供用于调解植物中的基因表达的重组DNA分子和构建体,和其核苷酸序列。本发明还提供包括含有可操作地连接至异源可转录DNA分子的重组DNA分子的转基因植物、植物细胞、植物部分和种子,以及其使用方法。

1. 一种包含选自由以下组成的组的DNA序列的重组DNA分子:

a) 与SEQ ID NO:19具有至少85%序列同一性的DNA序列;

b) 包含SEQ ID NO:19的DNA序列;和

c) 包含至少约50个具有SEQ ID NO:19的邻接核苷酸的片段,其中所述片段具有基因调控活性;

其中所述序列可操作地连接至异源可转录多核苷酸分子。

2. 如权利要求1所述的重组DNA分子,其中所述DNA序列与SEQ ID NO:19的所述DNA序列具有至少90%序列同一性。

3. 如权利要求1所述的重组DNA分子,其中所述DNA序列与SEQ ID NO:19的所述DNA序列具有至少95%序列同一性。

4. 如权利要求1所述的重组DNA分子,其中所述异源可转录DNA分子包含具有农艺学重要性的基因。

5. 如权利要求4所述的重组DNA分子,其中所述具有农艺学重要性的基因在植物中赋予除草剂耐性。

6. 如权利要求4所述的重组DNA分子,其中所述具有农艺学重要性的基因在植物中赋予害虫抗性。

7. 一种转基因植物细胞,其包括包含选自由以下组成的组的DNA序列的重组DNA分子:

a) 与SEQ ID NO:19具有至少85%序列同一性的DNA序列;

b) 包含SEQ ID NO:19的DNA序列;和

c) 包含至少约50个具有SEQ ID NO:19的邻接核苷酸的片段,其中所述片段具有基因调控活性;

其中所述序列可操作地连接至异源可转录多核苷酸分子。

8. 如权利要求7所述的转基因植物细胞,其中所述转基因植物细胞是单子叶植物细胞。

9. 如权利要求7所述的转基因植物细胞,其中所述转基因植物细胞是双子叶植物细胞。

10. 一种转基因植物或其部分,其包括包含选自由以下组成的组的DNA序列的重组DNA分子:

a) 与SEQ ID NO:19具有至少85%序列同一性的DNA序列;

b) 包含SEQ ID NO:19的DNA序列;和

c) 包含至少约50个具有SEQ ID NO:19的邻接核苷酸的片段,其中所述片段具有基因调控活性;

其中所述序列可操作地连接至异源可转录多核苷酸分子。

11. 一种如权利要求10所述的转基因植物的子代植物,或其一部分,其中所述子代植物或其部分包含所述重组DNA分子。

12. 一种如权利要求10所述的转基因植物的转基因种子,其中所述种子包含所述重组DNA分子。

13. 一种产生商品产品的方法,其包括获得根据权利要求10所述的转基因植物或其部分并且从其中产生所述商品产品。

14. 如权利要求13所述的方法,其中所述商品产品是加工种子、谷粒、植物部分和粗粉。

15. 一种产生转基因植物的方法,其包括:

- a) 用如权利要求1所述的重组DNA分子转化植物细胞以产生转化植物细胞;并且  
b) 使转基因植物从所述转化植物细胞再生。

16. 一种密码子重新设计大肠杆菌 $\beta$ -葡糖醛酸酶(GUS)编码序列,其中与所述原生大肠杆菌GUS编码序列相比,所述密码子重新设计GUS编码序列在转基因植物中展示较高表达。

17. 如权利要求16所述的密码子重新设计GUS编码序列,其选自由以下的组:SEQ ID NO:29和30。

18. 如权利要求16所述的转基因植物,其中所述转基因植物是单子叶植物。

19. 如权利要求16所述的转基因植物,其中所述转基因植物是双子叶植物。

20. 如权利要求18所述的转基因植物,其中所述单子叶植物选自由以下组成的组:玉米(*Zea mays*)、水稻(*Oryza sativa*)、小麦(*Triticum*)、大麦(*Hordeum vulgare*)、高粱(*Sorghum spp.*)、小米、珍珠粟(*Pennisetum glaucum*)、粟(*Eleusine coracana*)、黍(*Panicum miliaceum*)、谷子(*Setaria italica*)、燕麦(*Avena sativa*)、小黑麦、裸麦(*Secale cereale*)、福尼奥米(*Digitaria*)、洋葱(*Allium spp.*)、菠萝(*Ananas spp.*)、草皮草、甘蔗(*Saccharum spp.*)、棕榈(*Arecaceae*)、竹子(*Bambuseae*)、香蕉(*Musaceae*)、生姜家族(*Zingiberaceae*)、百合(*Lilium*)、水仙(*Narcissus*)、鸢尾花(*Iris*)、朱顶红、兰花(*Orchidaceae*)、美人蕉、风信子(*Hyacinthoides*)和郁金香(*Tulipa*)。

21. 如权利要求19所述的转基因植物,其中所述单子叶植物选自由以下组成的组:大豆(*Glycine max*)、野生大豆(*Glycine soja*)、棉花(*Gossypium*)、番茄(*Solanum lycopersicum*)、胡椒(*Piper*)、南瓜(*Cucurbita*)、豌豆(*Pisum sativum*)、苜蓿(*Medicago sativa*)、蒺藜状苜蓿、豆类(*Phaseolus*)、雏豆(*Cicer arietinum*)、向日葵(*Helianthus annuus*)、马铃薯(*Solanum tuberosum*)、花生(*Arachis hypogaea*)、昆诺阿藜、荞麦(*Fagopyrum esculentum*)、角豆荚(*onia siliqua*)、甜菜(*Beta vulgaris*)、菠菜(*Spinacia oleracea*)和黄瓜(*Cucumis sativus*)。

## 植物调控元件和其用途

[0001] 本申请是申请日为2014年3月12日,名称为“植物调控元件和其用途”,申请号为201480026708.2的发明专利的分案申请。

[0002] 相关申请的引用

[0003] 本申请要求2013年3月14日提交的美国临时申请序列号61/785,245的权益,所述申请以引用的方式整体并入本文。

[0004] 序列表的并入

[0005] 名称为“MONS331W0.txt”的文件中含有的序列列表通过电子方式提交来与本文一起提交并且以引用的方式并入本文,所述序列列表是54.4千字节(在Microsoft Windows®中测量的大小)并且创建于2014年3月12日。

### 发明领域

[0006] 本发明涉及植物分子生物学、植物遗传工程和适用于调节植物中的基因表达的DNA分子的领域。

[0007] 背景

[0008] 调控元件是通过调节可操作连接可转录DNA分子的转录来调控基因活性的遗传元件。这类元件包括启动子、前导区、内含子和3'未翻译区域,并且适用于植物分子生物学和植物遗传工程领域。

[0009] 发明概述

[0010] 本发明提供用于包含调控元件的植物和构建体中的新颖基因调控元件。本发明还提供包含调控元件的转基因植物、植物细胞、植物部分和种子。在一个实施方案中,本发明提供可操作地连接至可转录DNA分子的本文公开的调控元件。在某些实施方案中,可转录DNA分子相对于本文提供的调控序列是异源的。本文还提供制造和使用本文公开的调控元件的方法,所述调控元件包括包含调控元件的构建体,和包含可操作地连接至相对于调控元件异源的可转录DNA分子的调控元件的转基因植物、植物细胞、植物部分和种子。

[0011] 因此,在一方面,本发明提供包含选自以下组成的组的DNA序列的重组DNA分子:(a)具有与SEQ ID NO:1-20中任何一个至少约85%序列同一性的DNA序列;(b)包含SEQ ID NO:1-20中任何一个的DNA序列;和(c)SEQ ID NO:1-20中任何一个的片段,其中所述片段具有基因调控活性;其中所述DNA序列可操作地连接至异源可转录DNA分子。“异源可转录DNA分子”是指所述可转录DNA分子相对于所述DNA序列为异源的。在具体实施方案中,重组DNA分子包含具有与SEQ ID NO:1-20中任何一个的DNA序列至少约85%、至少约86%至少约87%、至少约88%、至少约89%、至少约90%、至少约91%、至少约92%、至少约93%、至少约94%、至少约95%、至少约96%、至少约97%、至少约98%或至少约99%序列同一性的DNA序列。在具体实施方案中,异源可转录DNA分子包含具有农艺学重要性的基因,如能够在植物中提供除草剂抗性或害虫抗性的基因。在其他实施方案中,本发明提供包含如本文提供的重组DNA分子的构建体。

[0012] 在另一方面,本文提供转基因植物细胞,其包括包含选自以下组成的组的DNA序

列的重组DNA分子：(a) 具有与SEQ ID NO:1-20中任何一个至少约85%序列同一性的DNA序列；(b) 包含SEQ ID NO:1-20中任何一个的DNA序列；和(c) SEQ ID NO:1-20中任何一个的片段，其中所述片段具有基因调控活性；其中所述DNA序列可操作地连接至异源可转录DNA分子。在某些实施方案中，转基因植物细胞是单子叶植物细胞。在其他实施方案中，转基因植物细胞是双子叶植物细胞。

[0013] 在仍然另一个方面，本文进一步提供转基因植物或其部分，其包括包含选自由以下组成的组的DNA序列的重组DNA分子：(a) 具有与SEQ ID NO:1-20中任何一个至少约85%序列同一性的DNA序列；(b) 包含SEQ ID NO:1-20中任何一个的DNA序列；和(c) SEQ ID NO:1-20中任何一个的片段，其中所述片段具有基因调控活性；其中所述DNA序列可操作地连接至异源可转录DNA分子。在具体实施方案中，转基因植物是相对于起始转基因植物的任何世代的子代植物并且包含重组DNA分子。本发明还提供包含在生长时产生这类转基因植物的重组DNA分子的转基因种子。

[0014] 在另一方面，本发明提供从含有本发明的重组DNA分子的转基因植物产生商品产品的方法。本发明的商品产品含有可检测量的SEQ ID NO:1-20。如本文使用，“商品产品”是指包含来源于含有本发明的重组DNA分子的转基因植物、植物部分、植物细胞或种子的材料的任何组合物或产物。商品产品包括但不限于加工种子、谷粒、植物部分和粗粉。含有本发明的重组DNA分子的转基因植物可用于制造通常从植物获得的任何商品产品。本发明的商品产品含有可检测量的对应于本发明的重组DNA分子的DNA。检测样品中的此重组DNA分子中的一种或多种可用于确定商品产品的内含物或来源。可使用检测DNA分子的任何标准方法，包括本文公开的检测方法。

[0015] 在仍然另一个方面，本发明提供在转基因植物中表达可转录DNA分子，如具有农艺学重要性的基因的方法，所述表达是通过获得含有本发明的重组DNA分子的转基因植物并且培养所述植物来实现的。

[0016] 本文还提供一种提供转基因植物的方法，所述提供是通过用本发明的重组DNA分子转化植物细胞以产生转化植物细胞，并且使转化植物细胞再生以产生转基因植物来实现的。

[0017] 本发明还提供密码子重新设计大肠杆菌 (*Escherichia coli*) (大肠杆菌 (*E. coli*))  $\beta$ -葡糖醛酸酶 (GUS) 编码序列；其中与原生大肠杆菌GUS编码序列相比，所述密码子重新设计GUS编码序列在转基因植物中展示较高表达。在一个实施方案中，密码子重新设计GUS编码序列可选自由以下组成的组：SEQ ID NO:29和30。转基因植物可为单子叶植物。在一个实施方案中，单子叶植物选自由以下组成的组：玉米 (*Zea mays*)、水稻 (*Oryza sativa*)、小麦 (*Triticum*)、大麦 (*Hordeum vulgare*)、高粱 (*Sorghum* spp.)、小米、珍珠粟 (*Pennisetum glaucum*)、粟 (*Eleusine coracana*)、黍 (*Panicum miliaceum*)、谷子 (*Setaria italica*)、燕麦 (*Avena sativa*)、小黑麦、裸麦 (*Secale cereale*)、福尼奥米 (*Digitaria*)、洋葱 (*Allium* spp.)、菠萝 (*Ananas* spp.)、草皮草、甘蔗 (*Saccharum* spp.)、棕榈 (*Arecaceae*)、竹子 (*Bambuseae*)、香蕉 (*Musaceae*)、生姜家族 (*Zingiberaceae*)、百合 (*Lilium*)、水仙 (*Narcissus*)、鸢尾花 (*Iris*)、朱顶红、兰花 (*Orchidaceae*)、美人蕉、风信子 (*Hyacinthoides*) 和郁金香 (*Tulipa*)。转基因植物还可为双子叶植物。在一个实施方案中，双子叶植物选自由以下组成的组：大豆 (*Glycine max*)、野生大豆 (*Glycine soja*)、棉花

(*Gossypium*)、番茄(*Solanum lycopersicum*)、胡椒(*Piper*)、南瓜(*Cucurbita*)、豌豆(*Pisum sativum*)、苜蓿(*Medicago sativa*)、蒺藜状苜蓿、豆类(*Phaseolus*)、雏豆(*Cicer arietinum*)、向日葵(*Helianthus annuus*)、马铃薯(*Solanum tuberosum*)、花生(*Arachis hypogaea*)、昆诺阿藜、荞麦(*Fagopyrum esculentum*)、角豆荚(*onia siliqua*)、甜菜(*Beta vulgaris*)、菠菜(*Spinacia oleracea*)和黄瓜(*Cucumis sativus*)。

[0018] 附图简述

[0019] 图1a-1c展示原生大肠杆菌 $\beta$ -葡糖醛酸酶(GUS)编码序列(CR-Ec.uidA-1:1:4,SEQ ID NO:31)与密码子重新设计大肠杆菌GUS编码序列(CR-Ec.uidA\_nno-1:1:1,SEQ ID NO:30)之间的比对。比对中之相同核苷酸由星号指示。

[0020] 序列简述

[0021] SEQ ID NO:1、3、5、7、9、11、13、15、17、19、21和23是启动子序列。

[0022] SEQ ID NO:2、4、6、8、10、12、14、16、18、20、22和24是前导序列。

[0023] SEQ ID NO:25-28是扩增引物序列。

[0024] SEQ ID NO:29和30是密码子重新设计GUS编码序列。SEQ ID NO:29包含可加工内含子,而SEQ ID NO:30是邻接编码序列。

[0025] SEQ ID NO:31是原生大肠杆菌 $\beta$ -葡糖醛酸酶编码序列。

[0026] SEQ ID NO:32是GUS编码序列,其具有基于SEQ ID NO:31的原生大肠杆菌 $\beta$ -葡糖醛酸酶的可加工内含子。

[0027] SEQ ID NO:33、39和40是3'UTR序列。

[0028] SEQ ID NO:34-37、41和44是转录调控表达元件组(EXP)序列,其包含:5'可操作地连接至前导序列的启动子序列,所述前导序列5'可操作地连接至内含子序列;或在SEQ ID 44的情况下,5'可操作地连接至前导序列的启动子序列。

[0029] SEQ ID NO:38是内含子序列。

[0030] SEQ ID NO:42和44分别是来源于北美萤火虫和海洋腔肠的荧光素酶蛋白质的编码序列。

[0031] 发明详述

[0032] 本发明提供在植物中具有基因调控活性的DNA分子。这些DNA分子的核苷酸序列提供为SEQ ID NO:1-20。这些DNA分子,例如,能够影响可操作地连接可转录DNA分子在植物组织中的表达,因此调控可操作地连接转基因在转基因植物中的基因表达。本发明还提供修饰、产生和使用这些分子的方法。本发明还提供包含含有本发明重组DNA分子的转基因植物细胞、植物、植物部分和种子的组合物,以及其制备和使用相同方法。

[0033] 提供以下定义和方法是为了更好地界定本发明以及指导本发明的一般技术人员实践本发明。除非另外指出,否则术语应按照相关技术领域的普通技术人员的常规用法来理解。

[0034] DNA分子

[0035] 如本文使用,术语“DNA”或“DNA分子”是指基因组或合成来源的双链DNA分子,即,脱氧核苷酸碱基的聚合物。如本文使用,术语“DNA序列”是指DNA分子的核苷酸序列。本文使用的命名法对应于美国联邦法规§1.822的标题37,并且在WIPO标准ST.25(1998),附录2,表1和3的表中阐明。

[0036] 如本文使用,“重组DNA分子”是包含组合DNA分子的DNA分子,这种组合DNA分子在没有人干预的情况下完全不会自然出现。例如,重组DNA分子可为包含相对于彼此异源的至少两个DNA分子的DNA分子,包含与在自然界中存在的DNA序列有偏差的DNA序列的DNA分子,或通过遗传转化并入宿主细胞的DNA中的DNA分子。

[0037] 如本文使用,术语“序列同一性”是指两个最佳比对多核苷酸序列或两个最佳比对多肽序列相同的程度。最优序列比对通过手动地比对两个序列,例如,参考序列和另一个DNA序列,以在具有适当内部核苷酸插入、缺失或间隙的序列比对中最大限度地提高核苷酸匹配的数目来建立。如本文使用,术语“参考序列”是指提供为SEQ ID NO:1-20的DNA序列。

[0038] 如本文使用,术语“%序列同一性”或“%同一性”是同一性分率乘以100。与参考序列最佳比对的序列的“同一性分率”是最佳比对中的核苷酸匹配的数目,除以参考序列中的核苷酸总数,例如,整个参考序列全长中的核苷酸总数。因此,本发明的一个实施方案提供包含序列的DNA分子,所述序列在与本文提供为SEQ ID NO:1-20的参考序列最佳比对时,与参考序列具有至少约85%同一性、至少约86%同一性、至少约87%同一性、至少约88%同一性、至少约89%同一性、至少约90%同一性、至少约91%同一性、至少约92%同一性、至少约93%同一性、至少约94%同一性、至少约95%同一性、至少约96%同一性、至少约97%同一性、至少约98%同一性、至少约99%同一性或至少约100%同一性。

[0039] 调控元件

[0040] 调控元件如启动子、前导区、增强子、内含子和转录终止区域(或3'UTR)在活细胞中的总体基因表达中起整体作用。如本文使用,术语“调控元件”是指具有基因调控活性的DNA分子。如本文使用,术语“基因调控活性”是指例如通过影响可操作地连接可转录DNA分子的转录和/或翻译来影响可操作地连接可转录DNA分子的表达的能力。因此,在植物中起作用的调控元件,如启动子、前导区、增强子、内含子和3'UTR适用于经由遗传工程来修饰植物表现型。

[0041] 如本文使用,“调控表达元件组”或“EXP”序列可指可操作地连接调控元件,如增强子、启动子、前导区和内含子的群组。因此,调控表达元件组可包含例如5'可操作地连接至前导序列的启动子,所述前导序列进而5'可操作地连接至内含子序列。

[0042] 调控元件可通过其基因表达模式来表征,例如,正性和/或负性效应如组成性表达或时间、空间、发育、组织、环境、生理、病理、细胞周期和/或化学响应表达,和其任何组合,以及通过定量或定性指示来表征。如本文使用,“基因表达模式”是将可操作地连接DNA分子转录至转录RNA分子中的任何模式。转录RNA分子可翻译以产生蛋白质分子或可提供反义或其他调控RNA分子,如双链RNA(dsRNA)、转移RNA(tRNA)、核糖体RNA(rRNA)、小分子RNA(miRNA)等。

[0043] 如本文使用,术语“蛋白质表达”是转录RNA分子翻译至蛋白质分子中的任何模式。蛋白质表达可通过其时间、空间、发育或形态性质,以及通过定量或定性指示来表征。

[0044] 启动子适用作调节可操作地连接可转录DNA分子的表达的调控元件。如本文使用,术语“启动子”总体上是指涉及识别和结合RNA聚合酶II和其他蛋白质,如反式作用转录因子,以起始转录的DNA分子。启动子可来源于基因的5'未翻译区域(5'UTR)。或者,启动子可为合成产生或操纵的DNA分子。启动子还可为嵌合的。嵌合启动子经由两个或更多个异源DNA分子的融合来产生。适用于实施本发明的启动子包括SEQ ID NO:1、3、5、7、9、11、13、15、

17和19,包括其片段或变体。在本发明的具体实施方案中,如本文描述的这类DNA分子和其任何变体或衍生物,可进一步定义为包含启动子活性,即,能够在宿主细胞中,如在转基因植物中充当启动子。在更进一步特异性实施方案中,片段可定义为展现其来源的起始启动子分子所具有的启动子活性,或片段可包含“最小启动子”,其提供基础水平转录并且由用于识别和结合RNA聚合酶II复合物以便起始转录的TATA盒或等效DNA序列组成。

[0045] 在一个实施方案中,提供本文公开的启动子序列的片段。启动子片段可包括启动子活性,如上所述,并且可单独与其他启动子和启动子片段组合使用,如在构建嵌合启动子中。在具体实施方案中,提供启动子的片段,其包含本文公开的具有启动子活性的多核苷酸分子的至少约50、至少约75、至少约95、至少约100、至少约125、至少约150、至少约175、至少约200、至少约225、至少约250、至少约275、至少约300、至少约500、至少约600、至少约700、至少约750、至少约800、至少约900或至少约1000个邻接核苷酸,或更长。从起始启动子分子产生这类片段的方法在本领域中是熟知的。

[0046] 从表示为SEQ ID NO:1、3、5、7、9、11、13、15、17和19的任何启动子产生的组合物,如内部或5'缺失,例如,可使用在本领域中已知的改进或改变表达的方法来产生,包括移除对于表达具有正性或负性效应的元件,复制对于表达具有正性或负性效应的元件,和/或复制或移除对于表达具有组织或细胞特异性效应的元件。从表示为SEQ ID NO:1、3、5、7、9、11、13、15、17和19的任何启动子产生的包含其中TATA盒元件或其等效序列和下游序列得以移除的3'缺失的组合物可用于例如制得增强子元件。可产生进一步缺失以移除对于表达具有正性或负性;组织特异性;细胞特异性;或定时特异性(诸如但不限于,昼夜节律)效应的任何元件。表示为SEQ ID NO:1、3、5、7、9、11、13、15、17和19的任何启动子和从其中衍生的片段或增强子可用于制得嵌合转录调控元件组合物,其包含表示为SEQ ID NO:1、3、5、7、9、11、13、15、17和19的任何启动子和从其中衍生的可操作地连接至其他增强子和启动子的片段或增强子。

[0047] 根据本发明,启动子或启动子片段可针对已知启动子元件,即,DNA序列特性,如TATA盒和其他已知转录因子结合位点基元的存在来分析。这类已知启动子元件的识别可由本领域技术人员用于设计具有与原始启动子类似表达模式的启动子的变体。

[0048] 如本文使用,术语“前导区”是指来自基因的未翻译5'区域(5'UTR)的DNA分子并且总体上定义为转录起始位点(TSS)与蛋白质编码序列起始位点之间的核苷酸节段。或者,前导区可为合成产生或操纵DNA元件。前导区可用作调节可操作地连接可转录DNA分子的表达的5'调控元件。前导区分子可与异源启动子或与其原生启动子一起使用。因此,本发明的启动子分子可以可操作地连接至其原生前导区或可操作地连接至异源前导区。适用于实施本发明的前导区包括SEQ ID NO:2、4、6、8、10、12、14、16、18和20或其片段或变体。在具体实施方案中,这类DNA序列可定义为能够在宿主细胞,包括例如转基因植物细胞中充当前导区。在一个实施方案中,这类DNA序列解码为包含前导区活性。

[0049] 表示为SEQ ID NO:2、4、6、8、10、12、14、16、18和20的前导序列(5'UTR)可包含调控元件或可采用可对于可操作地连接DNA分子之转录或翻译具有效应的次级结构。表示为SEQ ID NO:2、4、6、8、10、12、14、16、18和20的前导序列可根据本发明用于制得影响可操作地连接DNA分子之转录或翻译的嵌合调控元件。另外,表示为SEQ ID NO:2、4、6、8、10、12、14、16、18和20的前导序列可用于制得影响可操作地连接DNA分子之转录或翻译的嵌合前导序列。



[0050] 如本文使用,术语“内含子”是指DNA分子可从基因识别并且可总体上定义为在翻译之前的信使RNA (mRNA) 处理期间剪接的区域。或者,内含子可为合成产生或操纵DNA元件。内含子可含有实现可操作地连接基因的转录的增强子元件。内含子可用作调节可操作地连接可转录DNA分子的表达的调控元件。构建体可包括内含子,并且内含子可或可不相对于可转录DNA分子异源。在本领域中的内含子的实例包括水稻肌动蛋白内含子和玉米HSP70内含子。

[0051] 在植物中,在基因构建体中包含一些内含子导致相对于缺乏内含子的构建体,mRNA和蛋白质积聚增加。此效应被称为基因表达的“内含子介导增强”(IME) (Mascarenhas等人,Plant Mol.Biol.15:913-920,1990)。已知刺激植物中的表达的内含子已经在玉米基因(例如,tubA1、Adh1、Sh1和Ubi1)、水稻基因(例如,tpi)和双子叶植物基因如来自矮牵牛花的那些基因(例如,rbcS)、马铃薯(例如,st-1s1)和拟南芥(例如,ubq3和pat1)中识别。已经证明内含子的剪接位点内的缺失或突变减少基因表达,指示剪接可能为IME所需要。然而,剪接本身可能不需要,因为双子叶植物中的IME已经通过来自拟南芥的pat1基因的剪接位点的点突变来展示。在一种植物中多次使用同一内含子已被证明展现缺点。在那些情况下,需要具有基本控制元件之集合来构建适当重组DNA元件。

[0052] 如本文使用,术语“3'转录终止分子”、“3'未翻译区域”或“3'UTR”在本文中是指在mRNA分子的3'部分的未翻译区域的转录期间所使用的DNA分子。mRNA分子的3'未翻译区域可通过特异性裂解和也称为聚腺苷酸尾部的3'多聚腺苷酸化来产生。3'UTR可以可操作地连接至和位于可转录DNA分子的下游并且可包括多聚腺苷酸化信号和能够影响转录、mRNA处理或基因表达的其他调控信号。聚腺苷酸尾部被认为在mRNA稳定性和起始翻译中起作用。在本领域中的3'转录终止分子的实例是胭脂碱合成酶3'区域;小麦hsp173'区域、豌豆rubisco酶小亚单位3'区域、棉花E6 3'区域和薏苡辛3'UTR。

[0053] 3'UTR通常有利地用于特异性DNA分子的重组表达。微弱3'UTR具有产生通读的潜力,其可影响位于相邻表达盒中的DNA分子的表达。适当控制转录终止可防止通读至位于下游的DNA序列(例如,其他表达盒)并且可进一步允许RNA聚合酶的有效再循环以改进基因表达。有效终止转录(RNA聚合酶II从DNA释放)是重新启动转录的先决条件,从而直接影响总体转录水平。转录终止之后,成熟mRNA从合成位点释放并且模板运输至细胞质。真核mRNA以聚(A)形式在体内积聚,使得难以通过常规方法来检测转录终止位点。此外,通过生物信息学方法来预测功能性和有效的3'UTR很难,因为没有允许容易地预测有效3'UTR的保守DNA序列。

[0054] 从实际观点来看,用于表达盒中的3'UTR通常有利地具有以下特性。3'UTR应能够高效地和有效地终止可转录DNA分子的转录并且防止转录物通读至任何相邻DNA序列中,所述相邻DNA序列可包含另一个表达盒如在多个表达盒驻留于一个传递DNA(T-DNA)中的情况下,或T-DNA插入其中的相邻染色体DNA。3'UTR不应导致由用于驱动DNA分子的表达的启动子、前导区、增强子和内含子所赋予的转录活性减少。在植物生物技术中,3'UTR经常用于引导从转化植物提取的反向转录RNA的扩增反应并且用于:(1)评估一旦整合至植物染色体中的表达盒的转录活性或表达;(2)评估插入植物DNA内的拷贝数;和(3)评估育种之后的所得种子的接合性。3'UTR还用于从转化植物提取的DNA的扩增反应中来表征插入盒的完整性。

[0055] 如本文使用,术语“增强子”或“增强子元件”是指顺式作用调控元件,也被称为顺

式元件,其赋予总体表达模式的方面,但是通常不足以单独驱动可操作地连接DNA序列的转录。不同于启动子,增强子元件通常不包括转录起始位点(TSS)或TATA盒或等效DNA序列。启动子或启动子片段可天然地包括一个或多个影响可操作地连接DNA序列的转录的增强子元件。增强子元件还可融合至启动子以产生嵌合启动子顺式元件,其赋予基因表达的总体调节的方面。

[0056] 许多启动子增强子元件被认为结合DNA结合蛋白质且/或影响DNA布局,产生选择性允许或限制RNA聚合酶接近DNA模板或促进选择性开启转录起始位点处的双螺旋的局部构型。增强子元件可起作用来结合调控转录的转录因子。一些增强子元件结合一个以上转录因子,并且转录因子可以不同亲和力与一个以上增强子域相互作用。增强子元件可通过许多技术来识别,包括缺失分析,即,缺失5'末端或启动子内部的一个或多个核苷酸;使用DNA酶I足迹的DNA结合蛋白分析,甲基化干扰,电泳迁移率变化测定,通过连接介导聚合酶链反应(PCR)的体内基因组足迹,和其他常规测定;或使用已知顺式元件基元或增强子元件作为目标序列或目标基元以常规DNA序列比较方法如BLAST的DNA序列相似性分析。增强子域的微细结构可通过一个或多个核苷酸的诱变(或取代)或在本领域中已知的其他常规方法来进一步研究。增强子元件可通过化学合成或通过从包含这类元件的调控元件分离来获得,并且其可与额外侧接核苷酸合成,所述侧接核苷酸含有适用限制酶位点以促进子序列操纵。因此,本发明涵盖根据调节可操作地连接可转录DNA分子的表达的本文公开方法的增强子元件的设计、结构和使用。

[0057] 如本文使用,术语“嵌合”是指通过将第一DNA分子融合至第二DNA分子来产生的单一DNA分子,其中第一和第二DNA分子两者通常都不存在于所述配置中,即,彼此融合。因此,嵌合DNA分子是通常并不另外包含在自然界中的新DNA分子。如本文使用,术语“嵌合启动子”是指经由DNA分子的这类操纵来产生的启动子。嵌合启动子可组合两个或更多个DNA片段,例如,将启动子融合至增强子元件。因此,本发明涵盖根据调节可操作地连接可转录多核苷酸分子的表达的本文公开方法的嵌合启动子的设计、结构和使用。

[0058] 如本文使用,术语“变体”是指与第一DNA分子在组成上类似,但是不相同的第二DNA分子,如调控元件,并且其中第二DNA分子仍然保持第一DNA分子的一般功能性,即,相同或类似表达模式。变体可为缩短或截短型式的第一DNA分子和/或改变型式的第一DNA分子的DNA序列,如具有不同限制酶位点和/或内部缺失、取代和/或插入的DNA分子。调控元件“变体”还包括由在细菌和植物细胞转化中自然发生的突变所产生的变体。在本发明中,提供为SEQ ID NO:1-20的DNA序列可用于产生与原始调控元件的DNA序列在组成上类似,但是不相同的变体,但是仍然保持原始调控元件的一般功能性,即,相同或类似表达模式。鉴于本公开,本发明的这类变体的产生完全本领域的普通技能范围内并且涵盖于本发明范围中。

[0059] 嵌合调控元件可被设计成包含各种组成元件,其可以通过在本领域中已知的各种方法来可操作地连接,如限制酶消化和连接、连接无关的克隆、扩增期间的PCR产物的模块化装配,或调控元件的直接化学合成,以及在本领域中已知的其他方法。所得各种嵌合调控元件可包含相同组成元件,或相同组成元件的变体,但是在构成允许组成部分可操作地连接的一个或多个连接DNA序列的一个或多个DNA序列方面有所不同。在本发明中,提供为SEQ ID NO:1-20的DNA序列可提供调控元件参考序列,其中构成参考序列的组成元件可通过在

本领域中已知的方法来连接并且可包括一个或多个核苷酸的取代、缺失和/或插入或在细菌和植物细胞转化中自然发生的突变。

[0060] 本文描述的修饰、复制或缺失对于具体转基因的所需表达方面的功效可以经验为主地在稳定和瞬时植物测定中测试,如本文中的工作实例中所描述的那些,以便验证结果,所述结果可取决于产生的变化和起始DNA分子中的变化的目标而不同。

[0061] 构建体

[0062] 如本文使用,术语“构建体”意谓来源于任何来源的能够基因组整合或自主复制的任何重组DNA分子如质粒、粘粒、病毒、噬菌体或线性或圆形DNA或RNA分子,包括其中至少一个DNA分子以功能操作方式连接至另一个DNA分子,即可操作地连接的DNA分子。如本文使用,术语“载体”意谓可用于转化目的,即将异源DNA或RNA引入宿主细胞中的任何构建体。构建体通常包括一个或多个表达盒。如本文使用,“表达盒”是指包含可操作地连接至一个或多个调控元件,通常至少启动子和3'UTR的至少一个可转录DNA分子的DNA分子。

[0063] 如本文使用,术语“可操作地连接”是指第一DNA分子连接至第二DNA分子,其中第一和第二DNA分子如此布置以使得第一DNA分子影响第二DNA分子的功能。两个DNA分子可或不作为单一邻接DNA分子的一部分并且可或不相邻。举例来说,如果启动子调节细胞中的感兴趣的转录DNA分子的转录,那么启动子可操作地连接至可转录DNA分子。例如,当前导区能够影响DNA序列的转录或翻译时,它可操作地连接至DNA序列。

[0064] 在一个实施方案中,本发明的构建体可以双重肿瘤诱导(Ti)质粒边界构建体形式来提供,其具有从包含T-DNA的根癌农杆菌分离的Ti质粒的右边界(RB或AGRtu.RB)和左边界(LB或AGRtu.LB)区域,其与由根癌农杆菌细胞提供的传递分子一起允许T-DNA整合至植物细胞的基因组中(参见,例如,美国专利6,603,061)。构建体还可含有在细菌细胞中提供复制功能和抗生素选择的质粒骨架DNA节段,例如,大肠杆菌复制起点如ori322,广泛宿主范围复制起点如oriV或oriRi,和可选择标记如Spec/Strp的编码区域,其编码赋予大观霉素或链霉素抗性的Tn7氨基糖苷腺苷酰转移酶(aadA),或庆大霉素(Gm,Gent)可选择标记基因。对于植物转化,宿主菌株经常为根癌农杆菌ABI、C58或LBA4404;然而,植物转化领域技术人员已知的其他菌株可在本发明中起作用。

[0065] 本领域中已知以使得可转录DNA分子转录成翻译并表达为蛋白质的功能性mRNA分子的方式将构建体组装并引入细胞中的方法。为了实践本发明,制备和使用构建体和宿主细胞的常规组合物和方法为本领域技术人员熟知的。适用于在较高植物中表达核酸的典型载体在本领域中是熟知的并且包括来自根癌农杆菌的Ti质粒的载体和pCaMVCN传递控制载体。

[0066] 各种调控元件可包含于构建体中,包括本文提供的那些元件中的任何一个。任何这类调控元件可与其它调控元件组合来提供。这类组合可设计或改良以产生所需调控特征。在一个实施方案中,本发明的构建体包括可操作地连接至可操作地连接至3'UTR的可转录DNA分子的至少一个调控元件。

[0067] 本发明的构建体可包括本文提供或在本领域中已知的任何启动子或前导区。举例来说,本发明的启动子可以可操作地连接至异源非翻译5'前导区如来自热休克蛋白质基因的前导区。或者,本发明的前导区可以可操作地连接至异源启动子如花椰菜花叶病毒35S转录物启动子。

[0068] 表达盒还可包括编码适用于亚细胞靶向输送可操作地连接蛋白质的肽的转运肽编码序列,尤其叶绿体、白色体或其它质体细胞器;线粒体;过氧化物酶体;空泡;或细胞外位置。许多叶绿体定域蛋白质从核基因作为前体来表达并且通过叶绿体转运肽(CTP)靶向输送至叶绿体。这类分离叶绿体蛋白质的实例包括但不限于与小亚单位(SSU)核酮糖-1,5,-二磷酸羧化酶、铁氧还蛋白、铁氧还蛋白氧化还原酶、集光复合蛋白质I和蛋白质II、硫氧还蛋白F和烯醇丙酮酸莽草酸磷酸合成酶(EPSPS)相关联的那些蛋白质。叶绿体转运肽描述于例如美国专利号7,193,133中。已经证明可通过可操作地连接至转基因编码非叶绿体蛋白质的异源CTP的表达将非叶绿体蛋白质靶向输送至叶绿体。

[0069] 可转录DNA分子

[0070] 如本文使用,术语“可转录DNA分子”是指能够转录至RNA分子中的任何DNA分子,包括但不限于,具有蛋白质编码序列的那些分子和产生具有序列适用于基因抑制的RNA分子的那些分子。类型DNA分子可包括,但是不限于,来自相同植物的DNA分子,来自另一种植物的DNA分子,来自不同生物体的DNA分子,或合成DNA分子,如含有基因的反义信息的DNA分子,或编码人工、合成或另外修饰型式的转基因的DNA分子。并入本发明的构建体中的示例性可转录DNA分子包括,例如,来自除了DNA分子并入其中的物种以外的物种的DNA分子或基因或来源于或存在于相同物种中但是通过遗传工程方法而非经典育种技术来并入受体细胞中的基因。

[0071] “转基因”是指至少相对于其在宿主细胞基因组中的位置对于宿主细胞异源的可转录DNA分子和/或在当前或任何以前世代的细胞中人工并入宿主细胞基因组中的可转录DNA分子。

[0072] 调控元件,如本发明的启动子,可以可操作地连接至相对于调控元件异源的可转录DNA分子。如本文使用,术语“异源”是指两个或更多个DNA分子的组合,在这类组合通常不存在于自然界中时。举例来说,两个DNA分子可来自不同物种和/或两个DNA分子可来自不同基因,例如,来自相同物种的不同基因或来自不同物种的相同基因。因此,调控元件相对于可操作地连接可转录DNA分子为异源的,如果这类组合通常不存在于自然界中,即可转录DNA分子并非自然地可操作地连接至调控元件而出现。

[0073] 可转录DNA分子总体上可为需要转录物的表达的任何DNA分子。转录物的这类表达可导致所得mRNA分子的翻译,并且由此导致蛋白质表达。或者,例如,可转录DNA分子可被设计成最终导致特异性基因或蛋白质的表达减少。在一个实施方案中,此可使用在反义方向上定向的可转录DNA分子来完成。本领域普通技术人员熟悉使用这类反义技术。任何基因可以此方式负性调控,并且在一个实施方案中,可转录DNA分子可被设计成经由dsRNA、siRNA或miRNA分子的表达来抑制特定基因。

[0074] 因此,本发明的一个实施方案是重组DNA分子,其包含本发明的调控元件,如提供为SEQ ID NO:1-20的那些调控元件,所述调控元件可操作地连接至异源可转录DNA分子以便在构建体整合于转基因植物细胞的基因组中时以所需水平或所需模式调节可转录DNA分子的转录的。在一个实施方案中,可转录DNA分子包含基因的蛋白质编码区域并且在另一个实施方案中可转录DNA分子包含反义区域基因。

[0075] 具有农艺学重要性的基因

[0076] 可转录DNA分子可为具有农艺学重要性的基因。如本文使用,术语“具有农艺学重

要性的基因”是指在表达于具体植物组织、细胞或细胞类型中,赋予所需特性的可转录DNA分子。具有农艺学重要性的基因的产物可在植物中起作用以便导致对于植物形态学、生理学、生长、发育、产量、颗粒组成、营养概况、疾病或害虫抗性,和/或环境或化学耐受性的效应或可在以植物为食的害虫的饮食中充当杀虫剂。在本发明的一个实施例中,本发明的调控元件并入构建体以使得调控元件可操作地连接至作为具有农艺学重要性的基因的可转录DNA分子中。在含有这类构建体的转基因植物中,具有农艺学重要性的基因的表达可赋予有利农艺学性状。有利农艺学性状可包括,但是不限于,除草剂耐性、昆虫防治、改良产量、抗病性、病原体抗性、改良植物生长和发育、变性淀粉含量、改良含油量、改良脂肪酸含量、改良蛋白质含量、改良果实成熟、增强动物和人营养、生物聚合物产生、环境压力抗性、药物肽、改进加工质量、改进风味、杂种种子繁育效用、改进纤维产生,和所需生物燃料产生。

[0077] 在本领域中已知的具有农艺学重要性的基因的实例包括针对如下性状的那些基因:除草剂抗性(美国专利号6,803,501;6,448,476;6,248,876;6,225,114;6,107,549;5,866,775;5,804,425;5,633,435;和5,463,175),增加产量(美国专利号USRE38,446;6,716,474;6,663,906;6,476,295;6,441,277;6,423,828;6,399,330;6,372,211;6,235,971;6,222,098;和5,716,837),昆虫防治(美国专利号6,809,078;6,713,063;6,686,452;6,657,046;6,645,497;6,642,030;6,639,054;6,620,988;6,593,293;6,555,655;6,538,109;6,537,756;6,521,442;6,501,009;6,468,523;6,326,351;6,313,378;6,284,949;6,281,016;6,248,536;6,242,241;6,221,649;6,177,615;6,156,573;6,153,814;6,110,464;6,093,695;6,063,756;6,063,597;6,023,013;5,959,091;5,942,664;5,942,658;5,880,275;5,763,245;和5,763,241),真菌病抗性(美国专利号6,653,280;6,573,361;6,506,962;6,316,407;6,215,048;5,516,671;5,773,696;6,121,436;6,316,407;和6,506,962),病毒抗性(美国专利号6,617,496;6,608,241;6,015,940;6,013,864;5,850,023;和5,304,730),线虫抗性(美国专利号6,228,992),细菌病抗性(美国专利号5,516,671),植物生长和发育(美国专利号6,723,897和6,518,488),淀粉产生(美国专利号6,538,181;6,538,179;6,538,178;5,750,876;6,476,295),改良油产生(美国专利号6,444,876;6,426,447;和6,380,462),较高油产生(美国专利号6,495,739;5,608,149;6,483,008;和6,476,295),改良脂肪酸含量(美国专利号6,828,475;6,822,141;6,770,465;6,706,950;6,660,849;6,596,538;6,589,767;6,537,750;6,489,461;和6,459,018),较高蛋白质产生(美国专利号6,380,466),果实成熟(美国专利号5,512,466),增强动物和人营养(美国专利号6,723,837;6,653,530;6,5412,59;5,985,605;和6,171,640),生物聚合物(美国专利号USRE37,543;6,228,623;和5,958,745,和6,946,588),环境压力抗性(美国专利号6,072,103),药物肽和分泌肽(美国专利号6,812,379;6,774,283;6,140,075;和6,080,560),改进加工性状(美国专利号6,476,295),改进消化性(美国专利号6,531,648)低棉子糖(美国专利号6,166,292),工业酶产生(美国专利号5,543,576),改进风味(美国专利号6,011,199),固氮(美国专利号5,229,114),杂种种子繁育(美国专利号5,689,041),纤维产生(美国专利号6,576,818;6,271,443;5,981,834;和5,869,720)和生物燃料产生(美国专利号5,998,700)。

[0078] 或者,具有农艺学重要性的基因可通过编码导致内源性基因的基因表达的靶向调节的RNA分子来影响上述植物特性或表现型,例如通过反义(参见,例如,美国专利5,107,065);抑制性RNA(“RNAi”,包括通过miRNA、siRNA、反式作用siRNA和分阶段sRNA介导机制对

于基因表达的调节,例如在公开申请U.S.2006/0200878和U.S.2008/0066206,和美国专利申请11/974,469中所描述);或共抑制介导机制。RNA也可为工程化以使所需内源性mRNA产物裂解的催化RNA分子(例如,核酶或核糖开关;参见,例如,U.S.2006/0200878)。本领域中已知以使得可转录DNA分子转录至能够导致基因抑制的分子中的方式将构建体构建并引入细胞中的方法。

[0079] 植物细胞中的可转录DNA分子的表达也可用于抑制以植物细胞为食的植物害虫,例如,从鞘翅目害虫分离的组合物和从线虫害虫分离的组合物。植物害虫包括但不限于,节肢动物害虫、线虫害虫和真菌或微生物害虫。

[0080] 选择性标记

[0081] 选择性标记转基因还可与本发明的调控元件一起使用。如本文使用,术语“可选择标记转基因”是指其在转基因植物、组织或细胞中的表达,或其缺乏,可加以筛选或以某种方式计录的任何可转录DNA分子。用于本发明实践的可选择标记基因,和其相关联选择和筛选技术在本领域中为已知的并且包括但不限于,编码 $\beta$ -葡糖醛酸酶(GUS)、绿色荧光蛋白(GFP)、赋予抗生素抗性的蛋白质,和赋予除草剂耐性的蛋白质的可转录DNA分子。

[0082]  $\beta$ -葡糖醛酸酶

[0083] 从分离大肠杆菌K-12基因的 $\beta$ -葡糖醛酸酶(GUS)是植物生物技术中的最广泛使用报道基因之一。大肠杆菌GUS基因,uidA,是细菌染色体上的GUS操纵子的一部分。它由多种 $\beta$ -D-葡糖苷酸来诱导。GUS酶是催化 $\beta$ -D-葡糖苷酸水解成D-葡萄糖醛酸和配质的外水解酶。大肠杆菌生活在脊椎动物,包括人的消化道中。脊椎动物利用葡萄糖醛酸化途径来使异生物质和内源性废物化合物如类固醇、脂族醇、苯酚、羧酸、糖和各种其他代谢物解毒。葡萄糖醛酸化涉及与D-葡萄糖醛酸偶联。这主要发生于肝中,但是也发生于其他组织和器官如肾、肾上腺和消化道中。葡萄糖醛酸可由大肠杆菌用作碳和能量的主要来源。因此,大肠杆菌GUS蛋白质提供细菌可藉以使葡萄糖醛酸化途径的产物在脊椎动物消化道中降解以产生作为碳和能量来源的葡萄糖醛酸的手段。也由GUS酶释放的配质总体上不通过细菌降解,但是用作D-葡萄糖醛酸的运送载具(Gilissen等人,Transgenic Research,7:157-163,1998)。

[0084] 大肠杆菌 $\beta$ -葡糖醛酸酶基因作为报道物的使用由Jefferson等人(Proc.Natl.Acad.Sci.,83:8447-8451,1986)首次描述并且自从其被采用以来一直以与首次描述几乎相同的方式来使用。GUS基因用于监测植物基因表达并且常常用于将启动子或其他表达元件表征。然而,一些植物启动子在极低水平下表达并且可能使用基于GUS的测定不可检测到。这些较低表达启动子可能对于开发具有需要表现型如改进产量的转基因作物有价值。

[0085] 在开发转基因作物植物的早期,提供较高组成性表达的启动子是最需要的。来源于植物病毒基因组如花椰菜花叶病毒和玄参花叶病毒的这些较高组成性启动子用于驱动赋予除草剂耐性或抗虫性的转基因。随着植物生物技术领域复杂性增加,开发需要更特异性表达模式或较低表达水平的较新转基因性状。过表达或在错误植物组织中表达可导致转化植物中的不可需要的效应。举例来说,酶基因在植物中的异位表达(基因在生物体中的异常位置处表达)可导致所需最终产物减少,这归因于代谢途径中的分支点处缺少前体(Iwase等人,Plant Biotech.26:29-38,2009)。

[0086] 因为转录因子(TF)天然地充当细胞过程的主要调控剂,其预期为改变作物植物的

复杂性状的极好候选物,并且基于TF的技术可能为下一代成功生物技术作物的主要部分。TF技术经常需要优化,以减少不必要的副作用如生长阻滞或将所需性状增强至其具有商业价值的水平。优化常常通过改变TF转基因的表达来处理。组织特异性、发育或可诱导启动子,而非常用组成性启动子,可用于将转基因的表达限于适当组织或环境条件(Century等人, *Plant Physiology*, 147:20-29, 2008)。

[0087] 部分地归因于这些开发,需要用于表达元件表征的更灵敏测定来识别提供所需表达水平和模式的表达元件。本发明提供改进、密码子重新设计的GUS编码序列,与在本领域中普遍使用的原生大肠杆菌GUS编码序列相比,其在可操作地连接至启动子时表达更好。此改进、密码子重新设计的GUS编码序列可用于在数量上和质量上提供更大测定灵敏度,并且允许表征启动子和其他表达元件,所述表征可能在其他情形下使用原生大肠杆菌GUS编码序列是不可能的。改进、密码子重新设计GUS编码序列可用于表征单子叶和双子叶植物中的表达元件。适用于实施本发明的单子叶植物包括但不限于玉米 (*Zea mays*)、水稻 (*Oryza sativa*)、小麦 (*Triticum*)、大麦 (*Hordeum vulgare*)、高粱 (*Sorghum spp.*)、小米、珍珠粟 (*Pennisetum glaucum*)、粟 (*Eleusine coracana*)、黍 (*Panicum miliaceum*)、谷子 (*Setaria italica*)、燕麦 (*Avena sativa*)、小黑麦、裸麦 (*Secale cereale*)、福尼奥米 (*Digitaria*)、洋葱 (*Allium spp.*)、菠萝 (*Ananas spp.*)、草皮草、甘蔗 (*Saccharum spp.*)、棕榈 (*Arecaceae*)、竹子 (*Bambuseae*)、香蕉 (*Musaceae*)、生姜家族 (*Zingiberaceae*)、百合 (*Lilium*)、水仙 (*Narcissus*)、鸢尾花 (*Iris*)、朱顶红、兰花 (*Orchidaceae*)、美人蕉、风信子 (*Hyacinthoides*) 和郁金香 (*Tulipa*)。适用于实施本发明的双子叶植物包括但不限于大豆 (*Glycine max*)、野生大豆 (*Glycine soja*)、棉花 (*Gossypium*)、番茄 (*Solanum lycopersicum*)、胡椒 (*Piper*)、南瓜 (*Cucurbita*)、豌豆 (*Pisum sativum*)、苜蓿 (*Medicago sativa*)、蒺藜状苜蓿、豆类 (*Phaseolus*)、雏豆 (*Cicer arietinum*)、向日葵 (*Helianthus annuus*)、马铃薯 (*Solanum tuberosum*)、花生 (*Arachis hypogaea*)、昆诺阿藜、荞麦 (*Fagopyrum esculentum*)、角豆荚 (*onia siliqua*)、甜菜 (*Beta vulgaris*)、菠菜 (*Spinacia oleracea*) 和黄瓜 (*Cucumis sativus*)。

#### [0088] 细胞转化

[0089] 本发明还针对产生包含可操作地连接至可转录DNA分子的一个或多个调控元件的转化细胞和植物的方法。

[0090] 术语“转化”是指将DNA分子引入受体宿主中。如本文使用,术语“宿主”是指细菌、真菌或植物,包括细菌、真菌或植物的任何细胞、组织、器官或子代。特别受到关注的植物组织和细胞包括原生质体、愈伤组织、根、块茎、种子、茎、叶、秧苗、胚胎和花粉。

[0091] 本文中使用的术语“转化”是指已经向其中引入外部DNA分子(如构建体)的细胞、组织、器官或有机体。优选地,引入的DNA分子可整合到受体细胞、组织、器官或有机体的基因组DNA中,以使引入的DNA分子被随后的子代继承。“转基因”或“转化”细胞或有机体也可包括细胞或有机体的子代以及由杂交中使用这种转基因植物作为亲代并且表现由外部DNA分子的存在引起的改变表现型的育种计划中产生的子代。所引入DNA分子还可瞬时引入受体细胞中以使得所引入DNA分子不由后续子代承袭。术语“转基因”是指含有一个或多个异源DNA分子的细菌、真菌或植物。

[0092] 存在本领域技术人员熟知的将DNA分子引入植物细胞中的许多方法。此过程总体

上包括选择合适宿主细胞、将宿主细胞用载体转化,并且获得转化宿主细胞的步骤。在本发明的实践中通过将构建体引入植物基因组中来转化植物细胞的方法和材料可包括任何熟知和证明方法。合适方法尤其包括但不限于细菌感染(例如农杆菌)、二元BAC载体、直接递送DNA(例如通过PEG介导的转化、干燥/抑制介导的DNA吸收、电穿孔、使用碳化硅纤维来搅拌,和DNA涂布粒子的加速)。

[0093] 宿主细胞可为任何细胞或生物体,如植物细胞、藻类细胞、藻类、真菌细胞、真菌、细菌细胞或昆虫细胞。在具体实施方案中,宿主细胞和转化细胞可包括来自作物植物的细胞。

[0094] 转基因植物随后可从细胞本发明的转基因植物再生。使用常规育种技术或自花受粉,可从这个转基因植物产生种子。这类种子,和从这类种子生长的所得子代植物含有本发明的重组DNA分子,因此为转基因的。

[0095] 本发明的转基因植物可自花受粉以提供本发明的纯合转基因植物(对于重组DNA分子为纯合的)的种子或与非转基因植物或不同转基因植物杂交以提供本发明的杂合转基因植物(对于重组DNA分子为杂合的)的种子。这类纯合和杂合转基因植物在本文中称为“子代植物”。子代植物是起源于含有本发明的重组DNA分子的原始转基因植物的转基因植物。使用本发明的转基因植物产生的种子可收获并且用于生长几个世代的本发明转基因植物,即,子代植物,其包含本发明的构建体并且表达具有农艺学重要性的基因。通常用于不同作物的育种方法的描述可发现于多个参考书籍中的一个,参见,例如,Allard,Principles of Plant Breeding,John Wiley&Sons,NY,U.,CA,Davis,CA,50-98(1960);Simmonds,Principles of Crop Improvement,Longman,Inc.,NY,369-399(1979);Sneep和Hendriksen,Plant breeding Perspectives,Wageningen(编),Center for Agricultural Publishing and Documentation(1979);Fehr,Soybeans:Improvement,Production and Uses,第2版,Monograph,16:249(1987);Fehr,Principles of Variety Development,Theory and Technique,(第1卷)和Crop Species Soybean(第2卷),Iowa State Univ.,Macmillan Pub.Co.,NY,360-376(1987)。

[0096] 转化植物可针对所关注的一个或多个基因的存在和由本发明的调控元件赋予的表达水平和/或概况来分析。本领域技术人员知道可用于分析转化植物的许多方法。举例来说,植物分析方法包括但不限于DNA印迹或RNA印迹、基于PCR的方法、生物化学分析、表型筛选方法、现场评估和免疫诊断测定。可转录DNA分子的表达可使用如制造商所描述的TaqMan®(Applied Biosystems,Foster City,CA)试剂和方法来测量并且PCR循环时间使用TaqMan® Testing Matrix来确定。或者,如制造商所描述的Invader®(Third Wave Technologies,Madison,WI)试剂和方法可用于评估转基因表达。

[0097] 本发明还提供本发明的植物的部分。植物部分包括但不限于叶、茎、根、块茎、种子、胚乳、胚珠和花粉。本发明的植物部分可为能存活的、不能存活的、可再生的和/或不可再生的。本发明还包括并且提供包含本发明的DNA分子的转化植物细胞。本发明的转化或转基因植物细胞包含可再生和/或不可再生植物细胞。

[0098] 本发明可经由参考以下实施例来更容易理解,所述实施例提供用于说明,并且不意欲限制本发明,除非指定。本领域技术人员应当了解在下面的实施例中公开的技术代表本发明者发现的在本发明实践中起较好作用的技术。但是,鉴于本公开,本领域技术人员了



解到可以在公开的具体实施方案中进行许多变化,并且仍然获得同样或类似的结果,而不背离本发明的精神和范围,因此,附图中提出或示出的所有事项应理解为例示性的,而不具有限制意义。

## 具体实施方式

[0099] 实施例1

[0100] 识别和克隆调控元件

[0101] 从单子叶物种薏苡属(*Coix lacryma-jobi*)、毛马唐(*Digitaria sanguinalis* (L.) Scop.)、细叶芒(*Miscanthus sinensis f. gracillimus*)、伽马草(*Tripsacum dactyloides*)和甘蔗(*Saccharum officinarum*)的基因组DNA来识别并克隆新颖RCc3启动子和前导区。RCc3蛋白质属于醇溶谷蛋白超家族,其名称来自谷物的醇溶性富含脯氨酸和谷氨酰胺的储存蛋白质。醇溶谷蛋白超家族(也称为蛋白酶抑制剂/脂质-转移蛋白/种子储存2S白蛋白家族;Pfam ID:PF00234)代表植物基因组中的最普遍蛋白质超家族之一。醇溶谷蛋白超家族的成员富含于各种植物的水果、坚果、种子和蔬菜中。其已知展现不同功能,包括种子储存和保护、脂质结合或传递和酶抑制。脂质传递蛋白质(LTP)属于醇溶谷蛋白超家族并且表达于各种植物组织中。水稻RCc3蛋白质是表达于水稻的根中的LTP,但是并非所有LTP蛋白质是根特异性的。

[0102] DNA扩增引物(表示为SEQ ID NO:25-28)使用来自玉蜀黍(*Zea mays*),水稻(*Oryza sativa*),高粱(*Sorghum bicolor*)和二穗短柄草(*Brachypodium distachyon*)的二十四个(24)LTP蛋白质的编码序列来设计。扩增引物与遵循制造商协议构建的GenomeWalker™(Clontech Laboratories, Inc, Mountain View, CA)文库一起使用以克隆相应基因组DNA序列的5'区域。

[0103] 进行生物信息学分析以识别扩增DNA内的调控元件。使用此分析的结果,调控元件在DNA序列内得以界定并且设计引物来扩增调控元件。每个调控元件的相应DNA分子使用标准聚合酶链反应(PCR)条件以含有从薏苡属(*C. lacryma-jobi*)、毛马唐(*D. sanguinalis* (L.) Scop.)、细叶芒(*M. sinensis f. gracillimus*)、伽马草(*T. dactyloides*)和甘蔗(*S. officinarum*)分离的独特限制酶位点和基因组DNA的引物来扩增。所得DNA片段连接至载体中并且测序。

[0104] 所识别RCc3启动子和前导区的DNA序列在表1中列出。启动子序列本文提供为SEQ ID NO:1、3、5、7、9、11、13、15、17和19。前导序列本文提供为SEQ ID NO:2、4、6、8、10、12、14、16、18和20。

[0105] 表1. RCc3启动子和前导区从各种草物种中分离。

序列描述	SEQ ID NO:	属类/物种
P-Cl.RCc3:3	1	薏苡属( <i>Coix lacryma-jobi</i> )
L-Cl.RCc3:2	2	薏苡属( <i>Coix lacryma-jobi</i> )
P-Ds.RCc3_1:1	3	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
L-Ds.RCc3_1:1	4	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
[0106] P-Ds.RCc3_2:1	5	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
L-Ds.RCc3_2:1	6	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
P-Ds.RCc3_3:1	7	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
L-Ds.RCc3_3:1	8	毛马唐( <i>Digitaria sanguinalis</i> (L.) Scop.)
P-MISgr.RCc3_1:1	9	细叶芒( <i>M. sinensis</i> f. <i>gracillimus</i> )
L-MISgr.RCc3_1:1	10	细叶芒( <i>M. sinensis</i> f. <i>gracillimus</i> )
P-MISgr.RCc3-2:2	11	细叶芒( <i>M. sinensis</i> f. <i>gracillimus</i> )

序列描述	SEQ ID NO:	属类/物种
L-MISgr.RCc3-2:1	12	细叶芒( <i>M. sinensis</i> f. <i>gracillimus</i> )
P-Td.RCc3_1:1	13	伽马草( <i>Tripsacum dactyloides</i> )
L-Td.RCc3_1:1	14	伽马草( <i>Tripsacum dactyloides</i> )
[0107] P-Td.RCc3_2:1	15	伽马草( <i>Tripsacum dactyloides</i> )
L-Td.RCc3_2:1	16	伽马草( <i>Tripsacum dactyloides</i> )
P-Td.RCc3_3:1	17	伽马草( <i>Tripsacum dactyloides</i> )
L-Td.RCc3_3:1	18	伽马草( <i>Tripsacum dactyloides</i> )
P-So.RCc3:2	19	甘蔗( <i>Saccharum officinarum</i> )
L-So.RCc3:2	20	甘蔗( <i>Saccharum officinarum</i> )

[0108] 实施例2

[0109] 分析驱动转基因玉米中的GUS的调控元件

[0110] 玉米植物用驱动 $\beta$ -葡糖醛酸酶 (GUS) 转基因的表达的包含可操作地连接至其原生RCc3前导区的RCc3启动子的载体, 尤其二元质粒构建体来转化。所得转化植物针对GUS蛋白质表达来分析。

[0111] 用于这些实验中的载体使用在本领域中已知的克隆方法来建立。所得载体包含来自根癌农杆菌的右边界区域; 用于测定可操作地连接至GUS的密码子重新设计编码序列的RCc3启动子/前导序列的第一转基因表达盒, 所述编码序列具有5'可操作地连接至来自谷子S-腺苷蛋氨酸合成酶1基因 (T-SETit.Ams1-1:1:1, SEQ ID NO:159) 的3'UTR的可加工内含子GOI-Ec.uidA+St.LS1.nno:3 (SEQ ID NO:29); 用于选择赋予除草剂草甘磷抗性的转化植物细胞的第二转基因表达盒 (由水稻肌纤维蛋白1启动子驱动); 和来自根癌农杆菌的左边界区域。所得质粒用于使用在本领域中已知的方法来转化玉米植物。由新颖RCc3启动子和前导区赋予的GUS的表达与由谷子和水稻RCc3同源物启动子和前导区驱动的表达进行比较。表2提供质粒构建体、RCc3启动子和前导序列, 和SEQ ID NO。

[0112] 表2. 二元植物转化质粒和相关联RCc3启动子/前导序列。

质粒构建体	启动子序列描述	SEQ ID NO:	前导序列描述	SEQ ID NO:
pMON264146	P-ClRCc3:3	1	L-ClRCc3:2	2
pMON264148	P-Ds.RCc3_1:1	3	L-Ds.RCc3_1:1	4
pMON264088	P-Ds.RCc3_2:1	5	L-Ds.RCc3_2:1	6
pMON264107	P-Ds.RCc3_3:1	7	L-Ds.RCc3_3:1	8
pMON264186	P-MISgr.RCc3_1:1	9	L-MISgr.RCc3_1:1	10
pMON264187	P-MISgr.RCc3-2:2	11	L-MISgr.RCc3-2:1	12
pMON264049	P-Td.RCc3_1:1	13	L-Td.RCc3_1:1	14
pMON264050	P-Td.RCc3_2:1	15	L-Td.RCc3_2:1	16
pMON264147	P-Td.RCc3_3:1	17	L-Td.RCc3_3:1	18
pMON264166	P-So.RCc3:2	19	L-So.RCc3:2	20
pMON264108	P-SEFit.Rcc3-1:1:10	21	L-SEFit.Rcc3-1:1:2	22
pMON264206	P-Os.Rcc3-1:1:24	23	L-Os.Rcc3-1:1:2	24

[0114] 在某些情况下,植物使用在本领域中已知并且如美国专利申请公布2009/0138985所描述的农杆菌介导转化方法来转化。

[0115] 组织化学GUS分析用于转化植物的定性表达分析。将全组织切片与GUS染色溶液X-Gluc (5-溴基-4-氯基-3-吡啶基-b-葡萄糖苷酸) (1mg/ml) 一起孵育适当长度时间、冲洗,并且视觉检查蓝色着色。GUS活性使用选定植物器官和组织、通过直接肉眼检查或在显微镜下的检查来定性确定。授粉之后21天 (21DAP),将R<sub>0</sub>植物针对根和叶,以及花粉囊、穗丝和发育种子和胚芽中的表达来进行检查。

[0116] 对于定量分析,从转化玉米植物的选定组织中提取总蛋白质。在50微升的总反应体积中,一微克总蛋白质与荧光底物4-甲基伞形酮-β-D-葡萄糖苷酸 (MUG) 一起使用。反应产物,4-甲基伞形酮 (4-MU) 在羟基得以电离的较高pH下最大限度地发荧光。添加碳酸钠的碱性溶液同时停止测定并调整pH用于量化荧光产物。使用具有Micromax读取器的Fluoromax-3 (Horiba;Kyoto, Japan) 以365nm下的激发,445nm下的发射来测量荧光,所述读取器具备设定为激发2nm,发射3nm的缝隙宽度。平均表达值提供为pmol 4MU/μg蛋白质/小时。

[0117] 将对于每个转化所观察到的平均R<sub>0</sub>GUS表达加以记录并且平均表达水平和标准误差基于来自多个转化事件的样品获得的测量来确定。

[0118] 实施例3

[0119] 来自调控元件的增强子

[0120] 增强子可来源于本文提供的启动子元件,如表示为SEQ ID NO:1、3、5、7、9、11、13、15、17和19的那些。这些增强子元件可包含一个或多个顺式调控元件,所述顺式调控元件在5'或3'可操作地连接至启动子元件或5'或3'可操作地连接至可操作地连接至启动子的额外增强子元件时,可增强或调节可转录DNA分子的表达,或提供可转录DNA分子特异性在细胞类型或植物器官中,或在发育或昼夜节律中的具体时间点处的表达。增强子的制造方法是将TATA盒或功能类似元件和任何下游DNA序列从启动子中移除,所述启动子允许从如上所述的本文提供启动子,包括其中TATA盒或功能类似元件和TATA盒下游的DNA序列得以移除的其片段来开始转录。

[0121] 增强子元件可来源于本文提供的启动子元件并且使用在本领域中已知的方法来克隆以5'或3'可操作地连接至启动子元件或5'或3'可操作地连接至可操作地连接至启动子的额外增强子元件。可克隆增强子元件以5'或3'可操作地连接至来自不同属类生物体的启动子元件或5'或3'可操作地连接至来自其他属类生物体或相同属类生物体的额外增强子元件,所述额外增强子元件可操作地连接至来自相同或不同属类生物体的启动子,从而产生嵌合调控元件。GUS表达载体可使用与以前实施例所描述的构建体类似的在本领域中已知的方法来构建,其中所得植物表达载体含有来自根癌农杆菌的右边界区域;用于测试调控或嵌合调控元件的第一转基因盒,其包含可操作地连接至来自玉米的HSP70热休克蛋白质的内含子(I-Zm.DnaK-1:1:1,SEQ ID NO:38)或本文提供的任何内含子或任何其他内含子的调控或嵌合调控元件,所述其他内含子可操作地连接至GUS的编码序列,所述GUS的编码序列具有可操作地连接至来自根癌农杆菌的胭脂碱合成酶3'UTR(T-AGRtu.nos-1:1:13,SEQ ID NO:39)或来自水稻脂质转移蛋白基因的3'UTR(T-0s.LTP-1:1:1,SEQ ID NO:40)的可加工内含子(GUS-2,SEQ ID NO:32)或不具有内含子(CR-Ec.uidA-1:1:4(GUS.nat),SEQ ID NO:31);用于选择赋予除草剂草甘磷(由水稻肌纤蛋白1启动子驱动)或者抗生素卡那霉素(由水稻肌纤蛋白1启动子驱动)抗性的转化植物细胞的第二转基因表达盒;和来自根癌农杆菌的左边界区域。所得质粒可用于通过如上所述方法或在本领域中已知的其他方法来转化玉米植物或其他属类植物。或者,来自玉米或其他属类植物的原生质体细胞可使用在本领域中已知的方法来转化以执行瞬时测定。

[0122] 由包含一个或多个增强子的调控元件驱动GUS表达可在稳定或瞬时植物测定中评估以确定增强子元件对于转基因的表达的效应。一个或多个增强子元件的修饰或一个或多个增强子元件的复制可基于经验实验和使用每个调控元件组合物所观察到的所得基因表达调控来执行。改变一个或多个增强子在所得调控或嵌合调控元件中的相对位置可影响转录活性或调控或嵌合调控元件的特异性并且凭经验地来确定以识别用于玉米植物或其他属类植物内的所需转基因表达概况的最佳增强子。

[0123] 实施例4

[0124] 对于密码子重新设计 $\beta$ -葡糖醛酸酶(GUS)的更大测定灵敏度

[0125] 植物启动子经常在低于许多定量测定的正常检测阈值的水平下表达,但是其表达特性对于某些转基因的表达可能非常有价值。在较早植物生物技术中,驱动较高组成性表达的启动子是合乎需要的并且用于驱动产生需要较高组成性表达的特异性表现型,如除草剂耐性或抗虫性的可转录DNA分子。这些较高组成性启动子经常来源于植物病毒基因组而非植物基因组,例如35S启动子来源于花椰菜花叶病毒和玄参花叶病毒。值得注意的是,在某些情况下,某些可转录DNA分子的较高组成性表达可导致消极后果如转基因沉默、偏差表现型或产量阻碍。举例来说,使用两个不同甘蔗衍生泛素启动子以及玉米泛素启动子来实现的GUS基因在转基因甘蔗植物中的较高表达导致GUS基因的转录后基因沉默(Wei等人, J.Plant Physiol.160:1241-1251,2003)。

[0126] 此外,最近存在对于展现特异性表达模式或在植物的特异性组织中更高度表达的启动子的需要。举例来说,酶基因在植物中的异位表达可导致所需最终产物减少,这归因于代谢途径中的分支点处缺少前体(Iwase等人,Plant Biotech.26:29-38,2009)。在这些情况下,需要使用在正确的组织或细胞类型中,或在具体发育窗口处表达可操作地连接可转

录DNA分子的启动子。植物基因组衍生启动子可经常展现合乎需要的组织、细胞或发育表达特性。由于这些植物启动子的较低表达水平,表达测定经常需要使用增强子来增强表达水平以允许在定量测定中检测。然而,使用这类增强子经常改变植物启动子的总体表达模式。

[0127] 改善在测定中所使用的报道基因的表达可消除对于增强植物衍生启动子的需要,因此提供由启动子赋予的表达模式的更精确评估。此实施例展示在表征多个不同EXP中使用密码子重新设计GUS编码序列来改进定量测定灵敏度,所述EXP包含5'可操作地连接至前导序列的启动子序列,所述前导序列5'可操作地连接至内含子序列。

[0128] 玉米植物用含有驱动原生大肠杆菌 $\beta$ -葡糖醛酸酶(GUS)转基因或密码子重新设计 $\beta$ -葡糖醛酸酶(GUS.nno)转基因的表达的EXP序列的植物表达载体来转化,并且所得植物针对GUS蛋白质表达来分析。使用在本领域中已知的方法将EXP和GUS编码序列克隆至二元质粒构建体中。

[0129] 所得植物表达构建体含有来自根癌农杆菌的右边界区域;展示两个GUS编码序列的测定灵敏度的第一转基因盒,所述GUS编码序列包括可操作地连接至原生大肠杆菌GUS编码序列(CR-Ec.uidA-1:1:4(GUS.nat),SEQ ID NO:31)或5'可操作地连接至来自水稻脂质转移蛋白基因的3'UTR(T-Os.LTP-1:1:1,SEQ ID NO:40)的密码子重新设计GUS编码序列(CR-Ec.uidA\_nno-1:1:1(GUS.nno),SEQ ID NO:30)的EXP;用于选择赋予除草剂草甘磷抗性的转化植物细胞的第二转基因表达盒(由水稻肌纤蛋白1启动子驱动);和来自根癌农杆菌的左边界区域。图1a至1c展示原生GUS编码序列(CR-Ec.uidA-1:1:4)和密码子重新设计GUS编码序列(CR-Ec.uidA\_nno-1:1:1)之间的比对。比对中之相同核苷酸由星号指示。密码子重新设计GUS序列与原生GUS编码序列77.9%相同并且被设计成在植物中更好地表达。

[0130] 使用三个(3)不同EXP类别,每一个赋予特异性表达模式。EXP EXP-SETit.Cab3+Zm.DnaK:1:1(SEQ ID NO:34)和EXP-SETit.Cab3+Zm.DnaK:1:2(SEQ ID NO:35)在玉米中赋予叶表达概况并且基本上相同,除了EXP-SETit.Cab3+Zm.DnaK:1:2的核苷酸位置1408至1412中的5'-CCGGA-3'的五个核苷酸插入以外。EXP序列EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5(SEQ ID NO:36)在玉米中提供增强根表达概况。EXP序列EXP-Zm.UbqM1:1:2(SEQ ID NO:37)在玉米中提供较高组成性表达概况。所得质粒用于使用在本领域中已知的方法来转化玉米植物。表3列出质粒构建体名称,和相应EXP和GUS序列。

[0131] 表3.用于将原生GUS相对于密码子重新设计GUS编码序列进行比较的质粒构建体、EXP序列和表达模式。

质粒构建体	EXP 描述	表达模式	SEQ ID NO:	GUS	SEQ ID NO:
pMON122599	EXP-SETil.Cab3+Zm.DnaK:1:2	叶	35	CR-Ec.uidA-1:1:4	31
pMON122595	EXP-SETil.Cab3+Zm.DnaK:1:1	叶	34	CR-Ec.uidA_mno-1:1:1	30
[0132] pMON144050	EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5	增强的根	36	CR-Ec.uidA-1:1:4	31
pMON122597	EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5	增强的根	36	CR-Ec.uidA_mno-1:1:1	30
pMON144051	EXP-Zm.UbqM1:1:2	组成性	37	CR-Ec.uidA-1:1:4	31
pMON122598	EXP-Zm.UbqM1:1:2	组成性	37	CR-Ec.uidA_mno-1:1:1	30

[0133] 在某些情况下,植物使用在本领域中已知并且如美国专利申请公布2009/0138985所描述的农杆菌介导转化方法来转化。

[0134] 组织化学GUS分析用于转化植物的定性表达分析。将全组织切片与GUS染色溶液X-Gluc (5-溴基-4-氨基-3-吡啶基- $\beta$ -葡萄糖苷酸) (1mg/ml) 一起孵育适当长度时间、冲洗,并且视觉检查蓝色着色。GUS活性使用选定植物器官和组织、通过直接肉眼检查或在显微镜下的检查来定性确定。授粉之后21天 (21DAP),将R0植物针对根和叶,以及花粉囊、穗丝和发育种子和胚芽中的表达来进行检查。

[0135] 对于定量分析,从转化玉米植物的选定组织中提取总蛋白质。在50 $\mu$ l的总反应体积中,一微克总蛋白质与荧光底物4-甲基伞形酮- $\beta$ -D-葡萄糖苷酸 (MUG) 一起使用。反应产物,4-甲基伞形酮 (4-MU) 在羟基得以电离的较高pH下最大限度地发荧光。添加碳酸钠的碱性溶液同时停止测定并调整pH用于量化荧光产物。使用具有Micromax读取器的Fluoromax-3 (Horiba;Kyoto,Japan) 以365nm下的激发,445nm下的发射来测量荧光,所述读取器具备设定为激发2nm,发射3nm的缝隙宽度。

[0136] R<sub>0</sub>世代转化株的平均GUS表达值提供于表4、5和6中。

[0137] 表4.使用具有叶表达概况的EXP来进行的原生和密码子重新设计GUS编码序列的平均R<sub>0</sub>世代GUS表达。

[0138]

组织	pMON122599 EXP- SETit.Cab3+Zm.DnaK:1:2/ CR-Ec.uidA-1:1:4	pMON122595 EXP- SETit.Cab3+Zm.DnaK:1:1/ CR-Ec.uidA_mno-1:1:1
V4叶	798	1807
V7叶	230	1863
V1叶	508	2097
V4根	0	0
V7根	0	0
V1根	14	0
花粉囊	95	1056
穗丝	154	1590
21DAP胚芽	24	31
21DAP胚乳	18	61

[0139] 表5. 使用具有增强根表达概况的EXP来进行的原生和密码子重新设计GUS编码序列的平均R<sub>0</sub>世代GUS表达。

[0140]

组织	pMON144050 EXP-CaMV.35S- enh+Os.Rcc3+Zm.DnaK:1:5/ CR-Ec.uidA-1:1:4	pMON122597 EXP-CaMV.35S- enh+Os.Rcc3+Zm.DnaK:1:5/ CR-Ec.uidA_mno-1:1:1
V4叶	0	50
V7叶	0	51
V1叶	0	82
V4根	26	486
V7根	16	257
V1根	18	343
花粉囊	19	67
穗丝	0	12
21DAP胚芽	14	125
21DAP胚乳	17	45

[0141] 表6. 使用具有组成性表达概况的EXP来进行的原生和密码子重新设计GUS编码序列的平均R<sub>0</sub>世代GUS表达。

组织	pMON144051 EXP- Zm.UbqM1:1:2/CR- Ec.uidA-1:1:4	pMON122598 EXP- Zm.UbqM1:1:2/CR- Ec.uidA_nno-1:1:1
V4叶	988	3327
V7叶	963	2771
VT叶	1777	3787
V4根	693	2149
V7根	402	1443
VT根	776	3170
花粉囊	2247	3190
穗丝	975	3324
21DAP胚芽	511	894
21 DAP胚乳	791	2298

[0142] 如可以在表4至6中看出,在与原生GUS编码序列相比,使用密码子重新设计GUS编码序列,在定量测定中存在更大灵敏度。预期GUS.nno与GUS.Nat群体之间的一些变化性,因为表达可受T-DNA的插入位点影响;然而使用GUS.nno,灵敏度的总体趋势展示大得多的灵敏度。在与GUS.Nat相比时,使用GUS.nno,由EXP-SETit.Cab3+Zm.DnaK:1:1 (SEQ ID NO:34)和EXP-SETit.Cab3+Zm.DnaK:1:2 (SEQ ID NO:35)驱动的GUS证明2.26至8.1倍更大灵敏度。同样地,与GUS.Nat相比,使用GUS.nno,由EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5 (SEQ ID NO:36)提供的增强根概况大16.06至19.06倍,使得此密码子重新设计GUS编码序列对于筛选根启动子,尤其以低水平表达的那些启动子是理想的,并且可展现等于或低于使用原生GUS编码序列的背景水平的GUS水平。与GUS.Nat相比,在使用GUS.nno时,由EXP-Zm.UbqM1:1:2 (SEQ ID NO:37)赋予的较高组成性表达概况展示大1.42至4.09倍的定量灵敏度。

[0144] 在性质上,使用密码子重新设计GUS编码序列,GUS染色更灵敏并且在组织样品中一致地观察到。总体上,与定量测定相比,定性染色观察结果倾向于不太灵敏。使用密码子重新设计GUS编码序列提供染色组织的更好并且更一致的微观检查。举例来说,在其中GUS由EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5 (SEQ ID NO:36)驱动的组织中,用密码子重新设计GUS编码序列转化的组织的组织化学染色更显著并且在皮质、表皮、内皮、根毛和侧根尖端的所有V7根样品中是可见的。相比之下,当原生GUS编码序列由EXP-CaMV.35S-enh+Os.Rcc3+Zm.DnaK:1:5驱动时,在相应V7根组织中未定性地观察到GUS染色。改进密码子重新设计GUS编码序列,(CR-Ec.uidA\_nno-1:1:1,SEQ ID NO:30),提供更大测定灵敏度并且在测量以低水平表达的启动子的表达中是尤其有价值的。

[0145] 实施例5

[0146] 分析驱动玉米叶和根原生质体中的GUS的调控元件

[0147] 玉米叶和根原生质体用驱动 $\beta$ -葡糖醛酸酶(GUS)转基因的表达的包含可操作地连接至其原生RCC3前导区的RCC3启动子的载体来转化,并且所得植物针对GUS蛋白质表达来分析。使用在本领域中已知并且如以前在实施例2中所述的方法将RCC3启动子和前导序列



克隆至二元质粒构建体中。

[0148] 用于共转化和数据标准化的两种质粒构建体也使用在本领域中已知的方法来构建。这些质粒构建体中的每一个包含由组成性EXP驱动的特异性荧光素酶编码序列。载体pMON19437包含具有组成性启动子的表达盒,所述启动子5'可操作地连接至内含子(EXP-CaMV.35S-enh+Zm.DnaK:1,SEQ ID NO:41),所述内含子5'可操作地连接至萤火虫(*Photinus pyralis*)荧光素酶编码序列(LUCIFERASE:1:3,SEQ ID NO:42),所述编码序列5'可操作地连接至来自根癌农杆菌胭脂碱合成酶基因(T-AGRtu.nos-1:1:13,SEQ ID NO:39)的3'UTR。载体pMON63934包含具有组成性EXP序列(EXP-CaMV.35S-enh-Lhcb1,SEQ ID NO:44)的表达盒,所述EXP序列5'可操作地连接至海肾(*Renilla reniformis*)荧光素酶编码序列(CR-Ren.hRenilla Lucife-0:0:1,SEQ ID NO:43),所述编码序列5'可操作地连接至来自根癌农杆菌胭脂碱合成酶基因(T-AGRtu.nos-1:1:13,SEQ ID NO:39)的3'UTR。

[0149] 玉米根和叶原生质体使用本领域中熟知的基于聚乙二醇(PEG)的转化方法来转化。原生质体细胞用pMON19437、pMON63934和表7提供的质粒之一来转化。在转化之后,转化原生质体在全黑暗中孵育过夜。随后,GUS和荧光素酶的测量通过将如以上提及的转化细胞的裂解制备物的等分试样放置于两个不同小孔托盘中来进行。一个托盘用于GUS测量并且第二个托盘用于使用双重荧光素酶报道基因测定系统(Promega Corp.,Madison,WI;参见例如Promega Notes Magazine,第57期,1996,第02页)来执行双重荧光素酶测定。

[0150] 执行每个EXP或启动子+前导区+内含子序列的四次转化。每个EXP或启动子+前导区+内含子序列的平均表达值来自每次转化的多个样品确定。样品测量使用每个EXP或启动子+前导区+内含子序列质粒构建体转化的四次重复来进行。背景GUS表达使用缺乏GUS转基因的阴性对照质粒构建体来确定。平均GUS和荧光素酶表达水平提供于表7(叶)和8(根)中。在这些表中,萤火虫荧光素酶值(例如,来自pMON19437的表达)提供于标记为“FLUC”的列中并且海肾荧光素酶值(例如,来自pMON63934的表达)提供于标记为“RLUC”的列中。在表7和8中还提供平均GUS/FLUC和GUS/RLUC比率,其提供原生质体测定中的表达强度的相对度量。

[0151]

表7. 来自转化玉米叶原生质体的平均GUS、FLUC、RLUC和RLUC值。

质粒 构建体	启动子 前导区	SEQ ID NO:	平均 GUS	平均 FLUC	平均 RLUC	平均 GUS/FLUC	平均 GUS/RLUC
pMON264146	P-Cf.RCε3-3 L-Cf.RCε3-2	1 2	5328064.75	105434	253107.5	50.73	21.15
pMON264148	P-Ds.RCε3_1:1 L-Ds.RCε3_1:1	3 4	773613	147918	338149.5	5.23	2.28
pMON264088	P-Ds.RCε3_2:1 L-Ds.RCε3_2:1	5 6	2883555.75	129947.5	309268.5	22.33	9.45
pMON264107	P-Ds.RCε3_3:1 L-Ds.RCε3_3:1	7 8	1093785	124864.75	306178.75	8.70	3.55
pMON264186	P-MISgt.RCε3_1:1 L-MISgt.RCε3_1:1	9 10	2613839.75	128887.25	303412.75	20.45	8.83
pMON264187	P-MISgt.RCε3-2:2 L-MISgt.RCε3-2:1	11 12	2370706.25	149383.5	370443.75	15.95	6.53
pMON264049	P-Td.RCε3_1:1 L-Td.RCε3_1:1	13 14	7506585.75	150939.25	368035.5	50.15	20.88
pMON264050	P-Td.RCε3_2:1 L-Td.RCε3_2:1	15 16	4447254.25	155356.25	364604.5	28.78	12.40
pMON264147	P-Td.RCε3_3:1 L-Td.RCε3_3:1	17 18	1100118.75	153451	316691.5	7.13	3.48
pMON264166	P-So.RCε3:2 L-So.RCε3:2	19 20	3062045	143684.5	332394.5	21.55	9.45
pMON264108	P-SETU.Rce3-1:1:10 L-SETU.Rce3-1:1:2	21 22	147483	129834.25	309317.25	1.15	0.50
pMON264206	P-Os.Rce3-1:1:24 L-Os.Rce3-1:1:2	23 24	184935.25	171440.75	386387.25	1.08	0.50

[0152]

表8. 来自转化玉米叶原生质体的平均GUS、FLUC和RLUC值。

质粒 构建体	启动子 前导区	SEQ ID NO:	平均 GUS	平均 FLUC	平均 RLUC	平均 GUS/FLUC	平均 GUS/RLUC
pMON264146	P-ClRCc3 L-ClRCc3	1 2	185142.3	18310	34502.5	10.18	5.43
pMON264148	P-DsRCc3_1:1 L-DsRCc3_1:1	3 4	16306.5	17008	31233	0.98	0.53
pMON264088	P-DsRCc3_2:1 L-DsRCc3_2:1	5 6	101603.8	19201.25	43298	5.23	2.33
pMON264107	P-DsRCc3_3:1 L-DsRCc3_3:1	7 8	29196	14483.5	34700.75	2.03	0.88
pMON264186	P-MISgRCc3_1:1 L-MISgRCc3_1:1	9 10	87232	18411.75	44755.75	4.89	1.95
pMON264187	P-MISgRCc3-2:2 L-MISgRCc3-2:1	11 12	510761.5	19093.75	41948.5	26.98	12.30
pMON264049	P-TdRCc3_1:1 L-TdRCc3_1:1	13 14	884517.8	23881.75	55790	37.23	16.18
pMON264050	P-TbRCc3_2:1 L-TbRCc3_2:1	15 16	91694.5	18385	43509.5	5.03	2.18
pMON264147	P-TdRCc3_3:1 L-TdRCc3_3:1	17 18	50257.25	18716.75	34489	2.65	1.45
pMON264166	P-SoRCc3-2 L-SoRCc3-2	19 20	508345.3	22335.25	51655.75	22.98	10.13
pMON264108	P-SETbRcc3-1:1:10 L-SETbRcc3-1:1:2	21 22	8123	17750.75	37872.25	0.45	0.23
pMON264206	P-OsRcc3-1:1:24 L-OsRcc3-1:1:2	23 24	336095.3	17709.5	40179.5	19.65	8.63

[0153] 如表7展示,所有RCc3同源物启动子展示驱动玉米叶原生质体中的转基因表达的能力。基于GUS/FLUC和GUS/RLUC比率,在此测定中,与其他启动子相比,一些RCc3同源物启动子更高地驱动表达。此外,如上述表8展示,所有RCc3同源物启动子不同程度地展示驱动玉米叶原生质体中的转基因表达的能力。

[0154] 实施例6

[0155] 分析驱动转基因玉米中的GUS的调控元件。

[0156] 玉米植物用驱动β-葡糖醛酸酶 (GUS) 转基因的表达的包含可操作地连接至其原生

RCc3前导区的RCc3启动子的载体来转化。所得植物针对GUS蛋白质表达来分析。

[0157] 使用在本领域中已知的方法,如在实施例2中所述的方法将RCc3启动子和前导序列克隆至二元质粒构建体中。所得二元质粒构建体是pMON264146、pMON264148、pMON264088、pMON264107、pMON264186、pMON264187、pMON264049、pMON264050、pMON264147和pMON264166。玉米植物还用pMON264108和pMON264206稳定转化。定性和定量GUS表达分析如实施例2所描述来执行。植物在V4、V7和VT发育阶段加以测定。示出R1和R3取样。表9示出稳定转化玉米植物的平均定量GUS表达。

[0158]

表9. 稳定转化玉米植物中的平均定量GUS表达。

质粒构建体	启动子前导区	SRQ ID No:	V4 叶	V4 根	V7 叶	V7 根	VT 叶	VT 根	VT 花朵/花粉囊	R1 穗轴/穗丝	R3 2HDAP 胚芽	R3 2HDAP 胚乳
pMON264146	P-CLRCc3 L-CLRCc3	1 2	25.15	61.31	30.71	42.64	35.96	98.19	298	125.12	21.97	186.52
pMON264148	P-DsRCc3_1:1 L-DsRCc3_1:1	3 4	48.34	36.81	42.49	125.25	69.76	55.44	277.93	58	67.08	115.71
pMON264088	P-DsRCc3_2:1 L-DsRCc3_2:1	5 6	38.31	51.18	59.2	149.2	70.93	158.32	214.47	130.72	141.85	164.68
pMON264107	P-DsRCc3_3:1 L-DsRCc3_3:1	7 8	67.1	327.44	85.02	365.91	161.65	202.17	787.25	103.63		
pMON264186	P-MISgRCc3_1:1 L-MISgRCc3_1:1	9 10	38.66	40.25	39.7	139.98	105.24	308.24	406.98	239.35	118.54	196.48
pMON264187	P-MISgRCc3-2:2 L-MISgRCc3-2:1	11 12	25.9	193.25	42.13	291.5	48.02	549.37	87.89	41.83		
pMON264049	P-TdRCc3_1:1 L-TdRCc3_1:1	13 14	283.86	238.31								
pMON264050	P-TdRCc3_2:1 L-TdRCc3_2:1	15 16	51.82	653.38								
pMON264147	P-TlRCc3_3:1 L-TlRCc3_3:1	17 18	42.49	55.87	41.49	197.51	117.77	262.63	1182.96	938.3	815.36	1240.92
pMON264166	P-SoRCc3:2 L-SoRCc3:2	19 20	34.11	213.86	125.91	855.23	79.33	237.25	347.99	177.13		

[0159] 如表9展示,所有RCc3启动子同源物能够在稳定转化玉米植物中驱动GUS转基因表达。此外,每个启动子具有特异性启动子独有的表达模式。举例来说,在RCc3启动子同源物中间,VT花朵/花粉囊的表达是不同的。由P-Td.RCc3\_3:1 (SEQ ID NO:17) 驱动的表达是对于所有启动子所观察到的最高表达,而由P-MISgr.RCc3-2:2 (SEQ ID NO:11) 驱动的表达是最低的。相对于R1穗轴/穗丝表达,P-Td.RCc3\_3:1 (SEQ ID NO:17) 展示这些组织中的最高表达并且P-MISgr.RCc3-2:2 (SEQ ID NO:11) 表达最少。由P-Td.RCc3\_3:1 (SEQ ID NO:17) 驱动的表达在较晚发育组织中增加。在根中,从V4至VT阶段表达增加并且在VT花朵/花粉囊、R1穗轴/穗丝和R321DAP胚胎和胚乳中甚至更高。在RCc3启动子同源物中间,由P-Td.RCc3\_3:1驱动的表达在VT花卉/花粉囊、R1穗轴/穗丝和R321DAP胚胎和胚乳中是最高的。

[0160] 相对于叶和根表达,一些RCc3启动子同源物相对于叶在根中展示较高表达。表10示出所有测定RCc3启动子的根与叶表达比率。

[0161] 表10. 稳定转化玉米植物的根/叶表达比率。

[0162]

质粒构建体	启动子前导区	SEQ ID NO:	平均根/叶		
			V4	V7	VT
pMON264146	P-Cl.RCc3:3	1			
	L-Cl.RCc3:2	2	2.44	2.06	2.65
pMON264148	P-Ds.RCc3_1:1	3			
	L-Ds.RCc3_1:1	4	0.76	2.95	0.79
pMON264088	P-Ds.RCc3_2:1	5			
	L-Ds.RCc3_2:1	6	1.81	2.52	2.23
pMON264107	P-Ds.RCc3_3:1	7			
	L-Ds.RCc3_3:1	8	4.88	4.30	1.25
pMON264186	P-MISgr.RCc3_1:1	9			
	L-MISgr.RCc3_1:1	10	1.04	3.53	2.93
pMON264187	P-MISgr.RCc3_2:2	11			
	L-MISgr.RCc3_2:1	12	7.46	6.92	11.44
pMON264049	P-Td.RCc3_1:1	13			
	L-Td.RCc3_1:1	14	0.84		
pMON264050	P-Td.RCc3_2:1	15			
	L-Td.RCc3_2:1	16	12.61		
pMON264147	P-Td.RCc3_3:1	17			
	L-Td.RCc3_3:1	18	1.31	4.76	2.40

[0163]

质粒构建体	启动子前导区	SEQ ID NO:	平均根/叶		
			V4	V7	VT
pMON264166	P-So.RCc3:2	19			
	L-So.RCc3:2	20	6.33	6.79	2.99

[0164] 如表10展示,从V4到VT阶段,每个RCc3启动子同源物展示不同根与叶表达比率和不同模式。举例来说,从V4至VT,P-Cl.RCc3:3 (SEQ ID NO:1) 保持类似表达比率,并且在V7阶段出现轻微下降。如表9所见的根中的表达从V4到V7稍微下降,然后到VT阶段时增加。启

动子P-Ds.RCc3\_3:1 (SEQ ID NO:7) 展示从V4至VT阶段表达比率的变化,并且在V4和V7阶段,相对于叶,在根中的表达较高,然后在VT阶段,转变为叶相对于根中的表达近似相等(1.25)。通过此启动子,表9示出的平均表达展示从V4到VT阶段叶中的表达增加,同时从V7到VT阶段,根中的表达下降。在V4阶段,启动子P-So.RCc3:2 (SEQ ID NO:19) 保持根与叶表达比率为6.33并且在V7阶段为6.79,然后在VT阶段下降至2.99。然而,从V4到V7阶段,使用此启动子的表达在叶和根中分别增加3.69和3.96倍,然后在VT阶段,相对于V4,减少至2.33和1.10。

[0165] 值得注意的是,并非所有启动子具有较高根与叶比率。举例来说,在V4阶段,启动子P-Ds.RCc3\_1:1 (SEQ ID NO:3) 和P-Td.RCc3\_1:1 (SEQ ID NO:13) 具有小于一的根/叶比率。然而,在V4根中,由P-Td.RCc3\_1:1驱动的表达比P-Ds.RCc3\_1:1大6.6倍。V4阶段的最高根/叶比率使用P-Td.RCc3\_2:1 (SEQ ID NO:15) 来获得。由P-Ds.RCc3\_1:1驱动的表达比率从V4(0.76)增加至V7(2.95),然后返回到与V4的比率类似的比率(0.79)。

[0166] 从V4到VT阶段,启动子P-MISgr.RCc3-2:2 (SEQ ID NO:11) 展示叶和根中的表达的增加。遍及所有三个阶段,此启动子具有大于6.9的根与叶比率,但是此比率从V4阶段的7.46到达V7阶段的6.92,然后攀升至VT阶段的11.44。从V4到VT阶段,由P-MISgr.RCc3-2:2驱动的表达在叶和根中增加。

[0167] 每个RCc3同源物启动子在稳定转化玉米中展示可能不一定依据来源于同源基因而可预测的表达模式,尤其在用于转化异源物种如玉米时更是如此。在V4、V7或VT阶段或在所有测定阶段的某一时点,相对于叶,大多数启动子在根中展示较高表达。值得注意的是,表达程度在启动子之间广泛不同。每个RCc3启动子同源物的独特表达性质使得一些启动子与其他启动子相比更适合于某些类型的可转录DNA分子表达。举例来说,可在土壤中的营养物质吸收中起关键作用并且在植物将要开始繁殖并且产生种子的后期发育阶段最佳表达的可转录DNA分子的表达可最佳获益于大约在VT阶段增加根中的表达的启动子例如P-MISgr.RCc3-2:2 (SEQ ID NO:11)。

[0168] \*\*\*\*\*

[0169] 已经示出和描述本发明的原则,本领域技术人员显而易见本发明可在安排和细节上改进而不背离这些原则。我们主张在权利要求书的精神和范围内的所有改进。本文中引用的所有公布和公开专利文件都以引用的方式并入本文,该引用的程度就如同已明确且个别地指示将各个别公布或专利申请以引用的方式并入本文一般。

## 序列表

<110> Monsanto Technology LLC

<120> 植物调控元件和其用途

<130> MONS:331W0

<150> 61/785,245

<151> 2013-03-14

<160> 44

<170> SIPOSequenceListing 1.0

<210> 1

<211> 1500

<212> DNA

<213> 薏苡属()

<400> 1

```

gttgacgtcg gaagtgatcc gaatcaaaca ccaacagggc tagaacgccc ggagagcgag 60
gttgcagttc aacctgctga gacagcaggg acctggcta aaccgacgaa aaccgatcta 120
gagaccggtt tagatcgatc tagggtttcg ggctcctgtg ttgataactc tagaactcct 180
cccgggaaga cccgtagggg agaagcacgt agaggttgtc ctaggaccag caaggccgcc 240
tagaacgccg tcaagtctcg tcgggagacg cgccgaacag caaggtagaa ggaaaaaggg 300
gtaaaagggg agtagattga ttttgatcga ttagggtcgg atgcctcaat cggccatgat 360
cctctcatat atatagaggg ggctggtctt atcccaatag gaaacatctc cggatacgat 420
ctccaagctt cctatccgga ctctatcaac atatagaatt caatccggac gtgacaaggt 480
aaccctgatt tcgccgatcc ttggatcgac cagatcggtc taatgggctt taccagccca 540
tactgatcaa caaatcgtcc ttacaagaga aggatccaca tacttagcgt cggactcagc 600
aatcatcaca accataaggc ctccaaccag ggccacctgg tcggacataa catgagtgat 660
cggaaaccca aatgatcag gccattcaa cttagaacat ctgatctctc aaaagacaaa 720
ttcgacctta atattacagg ccgaataactt cttctaaatt catttttate tgtgacactt 780
ttgagtgtca acagtatgct tttctttgca aatattcect tttttatttc taccattag 840
ttactttggc ccttccattt cattgtatgt aaaagtggat actaaagcta acgcaacaag 900
aacaanaata aatagatccg gggtatgacg tccccagga tattttacta aaatatcttc 960
tcatcagatc tagaaaatcc tcgggcctta tccatatagg gtggtatcac atccatatac 1020
ttatagtagg acagatgagg agattttttt acctcaatt ctaaaattea tcaacttaat 1080
taaaggaatt taactcagga ccagagcggg gtcgcgagag ccatatatgt aacaaaata 1140
cctattagtt tagggcagca gcaaataaac cattattcct ttgtcccctt attccgtctc 1200
cagctagcta aaagctgtac atatgattaa tcatcgacat cgtgcatgca atttgcagga 1260
aagtgcaagg agcagccagg tgtatacacg tetgtagggt tcgcgtcaca tcgatccat 1320
ggatatgcat gcatccaaaa cacacgtacg tacacttget ggtgcatgcc ttatgaaatg 1380
gaatactaca cgcattgcca caagttaagc acgcacaggc acacacacag agagacacag 1440
ttgctcgtc aatgcaattg gctggtctat aaataccacc gaagagaacc atcctaaaga 1500

```

<210> 2

<211> 86

<212> DNA

<213> 薏苡属()

<400> 2

acacaccggt agcgattcga tccttcagaa gagctactgc tagctagcta gagctatcat 60  
ctgatcggta gcagcaatat aattca 86

<210> 3

<211> 594

<212> DNA

<213> 毛马唐()

<400> 3

accaacaagc atcatgacaa tggcagcaaa gcattcgtca gagacgacca acaagcatca 60  
cgacactggc ggcaaagcat atcaaaacaa tgtaatgaga tacaatattg tttcataaag 120  
aagcctacct gcatgatcct ttctaacaaa ctcaaaatga taagggccat gctctgttcg 180  
gtgacaacct tcaaggcatc taactttgcca gaatttagct ttgttattac cagcttgta 240  
ttagtcttta ctatccagtc tcgaacaact tgggcgcctt ttgtctcacc atacctctac 300  
atactgcctt ccctgatcaa cacaacattc ttcaacccaa tcccttggea tttgcgcatg 360  
ttacaagggtg caaaacagcc agcccatatt gcaagttact aaactaaact atgggtccaaa 420  
tgcagcaatc ttgatcgtta gtactgtccg gcattatata tgaacaagt ccagatcacc 480  
catctcatca cagtcacatg cattcgcgtt cacggaaacc gttaacacac caccaactag 540  
tcagcattgc accaatcttc ctccctataa atgcagcaac aatcgagcga cacc 594

<210> 4

<211> 88

<212> DNA

<213> 毛马唐()

<400> 4

aacaccacga accatcacag gcacttatag caacaatcaa gttatttctg ccttgtgcac 60  
tcgtgggtcga gtagtaatac atagcaaa 88

<210> 5

<211> 1500

<212> DNA

<213> 毛马唐()

<400> 5

ttgccggcct gtgtgatggt gcgccccac caaccttga ttttgectgc tgccttgta 60  
gccaacgggt ggtagtcac cctettcaaa cgtegcaccg aagttctgtc gactcccaga 120  
gaaaatattg tatgtaacat ataactattac tactacagct gaacacgtaa caatatgatc 180  
ttattttgtg atgccgaaag caccgtgcta aagaccacct atcgccttgg ttgggatcga 240  
ggccttgctg ctacgccggt tgcacagac gtgtgcgtgc atcgcacgac tggcatgtga 300



gttgtgggtc ataaataatg tctaacaata ttaaggtaat tcttagtatg cgttgggatc 360  
 attttttgat gttggatcct gcggcaggtc tcacccatgc atgacggatt gagaacaagg 420  
 gagactggac cagcatgtaa aagataatga tgaaggcgag ccatgttgac ggtgtaccgg 480  
 gacaggcaga cgcgggcca catcgaagag ggagatagcg tgcgtgctaa tgtttttgtg 540  
 gctcgcgatg tcaatgactc atacagattt cggtagcttg ctaaaatcat tcagcttgctc 600  
 ggcaagcatg ggccaaacaa tctagcaaca atccatgttt gccatcgatg caataggaag 660  
 taatagaatc cacttagctt ctagatctca cctggatcct ctttttattt atatgcatat 720  
 attttgtggg agtggaggca cacatcttta tgtttcatgg ataataat ttttatatgta 780  
 tatctgtcca aataaataac gtacgtgtcc ttcaatctta gactccagtt agattgatca 840  
 atgtagtggg ctccatatt ctttttgtgt tttgtgtgcc atgtctcaag catgcatgtg 900  
 gaatgaaaag ctggaagctt ggcatcaatt gcctaaggag ccataccata aattaaacca 960  
 tttgctgata tggccacaat ttttttaaca agctatgcca tagtcattca tgtgccacgg 1020  
 ttgttgaatc gcctcaatta cgtgtggaac atgatgggtt tttaaacaac acttgggtgat 1080  
 ttctattctg gcctactctt gcacttagga tcgtgttgtt agccatgtgc acatatttag 1140  
 ctagctagta gcaaaggcat gagtgagttg ctgatgggtt tacaataacc agcataagta 1200  
 ttcatatatt tgaagtgaag cacgttcatg aacatatgag aaaacttaca ttataatttg 1260  
 cttggtcaga tgaaaaccta gctgattcct tggcgacgac aacatcctc cttgttccat 1320  
 actttcttaa taattatttt tctcccaaaa agtccatcgc gctagcatca atgaagaaat 1380  
 ataagaaaaa tctcatatc cacatagtaa atagagcatt atgcggtcca tgtaggaatc 1440  
 accttgggag cacctgcttg ctcaatcaca ctataaatac cacccataca ctgcttcaag 1500

<210> 6

<211> 104

<212> DNA

<213> 毛马唐()

<400> 6

agaaccatca cagacatacg ctacacgcac cctgtacgaa caaccaacct agctagctac 60  
 ctactgaaaa cacacataag cttgctagge agcatatcat agca 104

<210> 7

<211> 1407

<212> DNA

<213> 毛马唐()

<400> 7

atcttcaaga tgtcgtaate gatttcttga gtcaaatatt ttgttttcat attatttagg 60  
 gagtttttca tatggacaat acataaaaa atatatgcag tgcaagttat ttttgtttac 120  
 ttattcaatt tatccgttca atcccctaaa caaaatttta ccctaaacta ttttatttgg 180  
 tccaatctac accctaatta tatttctett tttatttctc tgtgtctgag ttaaattttg 240  
 acttcaaatt ttatgaatag atgcaaaaca tgatgcttta tgctaaaaat ttttactaag 300  
 aaattttctt tgttatgttc ataggtcaag aatattta atgaaattga tcttacatat 360  
 ataaaactgt acaaaaagta atcatgaaaa aaattaatat atttgttcta acatagagca 420

ttatgtataa gtcaactgac aaaatttgaa attaaaacte agcttgcattg tgaagaaaca 480  
 aaaaaagaga aatctaatta ggggtagatt ggaccaaata aatagttta agggggggta 540  
 aatgagcca aattttgttc aaggggttaa gtggatcaag taaatagatg aggggggtaa 600  
 aatggacttt tttcaacaat taacctcatc tatagaaggg taggtgcattc tcagttcgaa 660  
 aaataagatg catttggatc ttgaaaaatc aatttcccct cacaccccca aatggaaatt 720  
 gtcgtcacta ctcagataac taatgaaagt agactcttat tgtgatgatc caaaaggctt 780  
 ctgtgattgg aatttcgtcc aactatttt cacaacatcg taatgagata taatattggt 840  
 ccaccagaaa gcctacctgc atggtcattt ctaataacca actcaaaaaa tggtaaata 900  
 catgctcttt tagtgaaaaa cttaaggaa ctagcttgc caaaatttat ctttgtcatt 960  
 aacaactcag tttcaacat ccagtctca acaacttca tttttttttt gcgggtaaaa 1020  
 caacttcaact ttacatgtgg ccaaattgaa cacaacagtc agtcccatt gaggtaccta 1080  
 gcaacttggg catccttgt ctcgtacctc tacatatttt tgteectgat caacacaaca 1140  
 ttcttaaacc catttcttg gcatttccc atgttacaag gcgcaaaa accaacccat 1200  
 atggcaattt actaaactaa actatgetga tccaatgcag caatcttgat cgctagtact 1260  
 gtcacgatt acatctgaaa caagtccaga tcccctctc catcacagtc acatgcatcc 1320  
 atggtcacgg gaaccgttaa cacacaccac caactaatcg gcattgctcc aatcctccta 1380  
 taaatacagc aacgatcagg cgagaca 1407

<210> 8

<211> 90

<212> DNA

<213> 毛马唐()

<400> 8

ataccaate acacacagtg ttagcactta gcaaccgagg tatttctgag agctttgtgc 60  
 ccttcttgtg gtcgagtaat tagtagcaaa 90

<210> 9

<211> 583

<212> DNA

<213> 细叶芒()

<400> 9

tctgctagat gcaacgagct ggaagccagg cctgggacaa ctcgttgatg ggcgtgggg 60  
 tcatcaaaca gagattagct tateccattt tatgtagatg tgccatgtat gcatatgttt 120  
 tggaataatg gtgactcgtg aggaagcccc tgcacatggt ttggaatgat ttctgccacc 180  
 gatcgcagaa ttcaaaaataa ggcgagcaca tcagtgttcc gttaaaataa taaacggtag 240  
 attgaaataa tgattccagc tagcactata tattgtattt ctacgttaat aattttatat 300  
 attgcataag ggtcctccta attcatgcat cactagtage tagcacgttg agtgtatttg 360  
 gcatggtttc tcaacttget tcccctgtat cctgategat ctctccttca taactggcac 420  
 atcccaggac aaccaggcaa cgetcaaatc acacatgcgt tgctgattat ccacgctgca 480  
 tgatgcacga agccaatcac taaagcatca cttccacac catcgttcag ctataaaaac 540  
 cctgccaac tgcgtagag gaccatccca aagcacaaga gca 583

<210> 10

<211> 61

<212> DNA

<213> 细叶芒()

<400> 10

cgaagtgcag cactaaagca ccacttctg ctcacaccgt tgtagtaata atccatcacc 60  
a 61

<210> 11

<211> 1500

<212> DNA

<213> 细叶芒()

<400> 11

ttgtgagaca aaaacactgt tgcggetggg aataagccgg ctgataagtt caagcgaaca 60  
ggctgtactt tagtgettct atttgaacca ttggttctca ctctggctat gcattggtgc 120  
attctcactc tggettataa ttttttcttg cctagtaata gtttttgaaa taatgggagc 180  
aaccacttca gttataatat aatgtttgta tgttataatg ggagcaatca ctctggttaa 240  
ttgttgccctg tttataccat gcaggetgca ttatattaat tgcttagttc cttcttgttt 300  
gtatgtgact atgagatfff tgtgacctct acaggaagtg aagggaagag gattcctgct 360  
ggtttcaaca gacatagcaa gcaaagggtt tgcttctctg attgtgtatc tagatcacag 420  
tagtgtgaat gtgcaccaag caaaacttaa attaggtgct ctatccatgt acattaagcc 480  
tagctgtatc aaaaaaaatt cgggtttatc tgcgtaatat tggatgtata tataagattc 540  
tggcattgga atttatattt tttttttggt ttttaattcat taccacagac gcgtgttggg 600  
cgtgcgcgctc tgcataaagg tgcgtcaca gacgcgtttt gactatacgc gtctgtgata 660  
gggttactac agacgcgtct ttactccacc agtctgtctt ataaccta at cgagacgcg 720  
tgtttgcata acgcagacgc ttattggtaa cacgcgtctg tagaagaaac ttattacaga 780  
cacgttatta agcgcgtctg tgatgtgtct taacacgcgt ctgtaatgtg gatttattac 840  
atagtgatcg gcccatatga acgaatgatc gatctagccc caaagttagt gtatgaaatg 900  
ttcatttacg tttcaatgcg gtttctcaaa aaaaaacggt ccaatgctaa acaaaaacat 960  
aggaattata tagttttgct gtggetttagt aacttcgtcc aatcgtgcta gtttaatttg 1020  
tatataacctg taccatgcta ttctctggc cttggttctt gcgatccat tattaatgac 1080  
cacgccacgc cacgcattca ttcttaateca ccagttgett gacatccaat gtctctctcca 1140  
ctacttgccg acaccgtctg atactccaag atcccaagct aagataacac ccagtgatca 1200  
tatatataaaa aacaaaacgcc agtgcgagcc tggccatttg cggagccaac cgaagccgtg 1260  
cacaaaatat tcgataccgt atcagggaaa aactagttta tacgaggtag gcaataatcc 1320  
atgtttcaaaa aaaaacaggt aggcaataat ccagatcgga ctcttctctaa ccgggttcac 1380  
atgcataatg gaatatgatg gccgggttca catgaacgct agatctctg cctagtagtg 1440  
caccgatttc ttaatcccga ggettggaacta taagtacccc tggtaacacc gtgatcaaag 1500

<210> 12

<211> 98

<212> DNA

<213> 细叶芒()

<400> 12

catcgcaaac aagctgctaa tcacttctca agagctctct aactacatta attagctaga 60  
gtgatccgcg aggtagctgg cttgtgatcg agcaatac 98

<210> 13

<211> 740

<212> DNA

<213> 伽马草()

<400> 13

aaatggatga aacaacttgg accaatcaga gatggccacg teagctcccg atcgtcgtaa 60  
ccgaccaaac ccgatcgata acggttttagg ctccaataca ccgtcggtag cacccggtcg 120  
ctatcatctg cccccgtccc aacgetattg gtatcgtccg cccctatatac ggteggtagc 180  
ccagtccacc gtcggggcca atcgtcccet getggtccg ctcgtgtcgg taccgatcgc 240  
caaaaacgcc acgtcaacgg cactgcggta ccgaccgccg ctggcaccgg ctttagcggg 300  
ccacacgacc gatcgtggt gtacggacgt agaggatgaat catgcgattg aattttcgtc 360  
agaggaaagt tatcatctta ttatctccaa cctctcttc tacggctgga tccgacgaaa 420  
atttaccctg gacggtgcca gtaacaattg caggtctcac tcacgtgcta aatccagcaa 480  
tcaaacacga aggaatatac gtgatctggc cagaacatgc aagagaataa tacagtagtg 540  
ttagagtacg aaacctacac gattcaacga attaataat gggttcacgt tcacgggtat 600  
gctcgcgcac gtccaaaatc caacgacatt ttataagcg gcgatgcca gacgggccag 660  
ctcgagcacc acatggcgtg gctccatctc gcatecccca tcaccgctat aaataccatt 720  
ggccatgcac acccgactc 740

<210> 14

<211> 91

<212> DNA

<213> 伽马草()

<400> 14

ccacacagca caagcagcag cagcagcagc agctcgatcg aactagctta gctactacgt 60  
gcgcgtgcaa caagcagctc gatcgatcgc c 91

<210> 15

<211> 903

<212> DNA

<213> 伽马草()

<400> 15

actataattt tcatagagac catttctccc tccataccat ttaaaaaatt caaacaaaaa 60  
atatctataa gtggccctag gtttaagaaaa gtgccagtgg aatgcattt ttattgacgg 120  
ttttcttgta tctattgcca gtaaaaaatca tatatttaca ttggcgatc tttgaagcat 180  
accgtcatga caaatctgat ttctattgac agaatgetta agagaatcgc tattggtcat 240

ttagttgcag catttaattg gtatgtagt ctatgaaata gggaagtgga attttgtatt 300  
 gagagatatt gtttcattct tagttttacg acaccttga tgatgtagca gcgctggaaa 360  
 gagtgttatg aaattcctat tgagaatagc cttacaggtc aacagcatac ttttaaaatt 420  
 aaaccagaga tataccaaac ttcagtttct acgatataat atttttctac ctcataagtt 480  
 ataatcagtg gtacctcctt tttttatata ccctttatta aggacggtag agaacgtcct 540  
 ctactcgtat gttgtacacc accggtaacg caattaagta aaacatgcat gtcaagtagt 600  
 aagtatataa gttggcaact caccagatgg caggagtctt gacctatcac cggctagcta 660  
 tagctcttgt tacgtgaatt cagcatgatt gctcaacttg cttctgtatt tccacatggg 720  
 ggagaattgt catatctccc cctaacaact ggcgctcca gagacaacca tgcaatgacc 780  
 aaaccgatcg aatcacacat gtttgcagc attcatgtac caatgctgcc aagcacgaaa 840  
 gcactcctct ctccgctcca tccatcgtca gctataaaaa ccatgccaag caccctgtga 900  
 aaa 903

<210> 16

<211> 69

<212> DNA

<213> 伽马草()

<400> 16

gcacccaaa gcacacaaga gcacgcagtg cagcactcaa gcacttctg ctactgcac 60  
 accgtacca 69

<210> 17

<211> 1500

<212> DNA

<213> 伽马草()

<400> 17

aaaaggcact tttctcaaga aacagagtta gagatcaaac caaagtgcta gaaacattgt 60  
 aataggtatg tgcaacatat ttaaaggttc aaacctcaca gttttaccct ttttcacaaa 120  
 gttgcacttc ctggcctcgt ttaagcttat ttggacttag aatcttcaaa ttgtgctcga 180  
 ttgaaccttt actcatatac ttagcaaaaca agttagtcca agatgtttgt ttggacattc 240  
 aatcaccaaaa acttatggaa atggcttaat tgcccatttc ctttcatgg tacaatttgt 300  
 ttcacagatt taagcagaag ctatatcaaa cacggegett ctacaacagc ttatcagttt 360  
 agttgcttct cagaatgaaa ggtaataaac ataagctget ttttactaga tccttagata 420  
 agttgctttt ttaacaagca gctctgcca gcagggccat agtgccatgt tgtaggctgg 480  
 tcgctttccc gagccatatg aaggtcgtcg gtccatgtgg ctgaaattaa agttgatagg 540  
 cccatcgga caaatggtca attagaggec caaaatttgt gactcaatt ttcatttaca 600  
 cttccacgct agtaaacgaa aacatatata gtacctatat atcgattttt tttttctgg 660  
 gtcttagaaa cttcgtccaa taatcatget agtttaattt gtatactgc acaaagctat 720  
 tcctctggcc ttggttcttt gcgcgtccat gcttgtttat ggatgattgc agccacgcca 780  
 cgcattcatt ctcaatcacc atttgettga catctaattg cctctgcacc acttgcgcac 840  
 gctacacacc gtctgatacg ccaggatccc aagctaaaaa aacaccaat gatcatgtga 900

aaacaagtga cagtgcgagc cagcccatgc agcgatcttg gccatttgcg gagccaaccg 960  
 aagccgtgca caaaatattc gataccgtat aaggggaaac actagttata cgaggttaggc 1020  
 aaatataatc ctcttcctaa ccggcggccg ggttcacatg catatatgat ggccgccagc 1080  
 cgggttcaca tgatgaacgc tatggtgcct agtgcacgat ttattaatct cccgaggctg 1140  
 tactataaat acatcggtaa tactgtgata aaagcatcgc aaagaagatc tctaagactg 1200  
 tctccagcaa cgtcctctat atctatctc tataatctgc ctttacagtc tctctaaaa 1260  
 gattccatcc tctatatctc ctctctctc aacaacggcc tctaaatcac gtctctata 1320  
 cgcaaatacc tatattagag acatctttaa tttttaatt tttgtacata cgtatttgct 1380  
 atactctcaa atgtattgta catatcttag tttgtctaa ccggttgtt aaagtattca 1440  
 aatggataga ggagaggaga gagaaactct atatatagag gatccagcag cgtcctctaa 1500

<210> 18

<211> 172

<212> DNA

<213> 伽马草()

<400> 18

atttagagga ccgttagag gacgctgctg gagggagtag aggacgacag cgttctctaa 60  
 aatttagagt acaggatact ttagaggacc tgctggaggc agtctaaca ctgcatttg 120  
 ctagagagag tgatcgcgag gtagctagct ggcttgtgat cgatcgagca aa 172

<210> 19

<211> 645

<212> DNA

<213> 甘蔗()

<400> 19

caaattaaag ttgataggcc catcgaaga aatgatcaat gtaaggcca aaatttgtgt 60  
 ctgaaatgtt cgtttacatt ttcaagctaa ataaaaacat aggaattata tagttttgct 120  
 ggtggcttag aaacttcgct caaaatatta aaaaaaaga aaagaaactt cgtccaaaat 180  
 ggtgcttagt ttaatttgta tacctgcacc atgctattcc tctggccttg ttcttttgct 240  
 catctatcca tgcctatgga tgatcgcagc cacgccagc aattcattct taatccat 300  
 ttgcttgaca tccaatgtcc tctacaccac cacttgcgca cctacacac cggccatttg 360  
 atacgccaag atcccaagct gaaataacac ccaatgatca tatgaaaaca agcgtgagtg 420  
 cgagcctgcc catgcagcga tcttgccat ttgcggagcc aaccgaagcc cgtgcacaaa 480  
 atattcgaga ccgtataagg gaaaacta gttatacgea atgtccaca taatccagat 540  
 cgactcttcc taaccgggtt cacatgaacg ctgttgtgce tagtgcaccg atttcttaa 600  
 tccaagget cgactataaa taccctggt atttgcaccg tgatc 645

<210> 20

<211> 86

<212> DNA

<213> 甘蔗()

<400> 20

aaagcatcgc aaacaagcta aagagctctc taactacatt gattagagtg atctgcgcta 60  
 gaaggaggct agcttgtgat cgagca 86  
 <210> 21  
 <211> 1563  
 <212> DNA  
 <213> 谷子()  
 <400> 21  
 agaccagca acctaccgag ggggtaccgg aggtagtgtt ttgtggtggg gctcgtcgaa 60  
 gatcaggaac ttgaaggatga actcgaacac acgatttaga caagttcggg ctgcttatgc 120  
 cgcataatac cccatgtcat gtgtgttggg tggattgtat tgattgatca gatatttggg 180  
 gggggccctg cctcgcctta tattgccccat tgccggaggc agggctacag gtcggttgtt 240  
 gtacaagagt actagtcggg ttgaccagc gactcctact ctaattgcta caagtagttt 300  
 cctaatacctt gactagtcct tgtccgccac gtagaccagc acgtcttgca ctagtctctt 360  
 gtgtttgata catcttggg tacagtccga tattgtagga ctatccaagc ttcccagtag 420  
 gccatagat gtatggccga caactggata atgtaactct gggtcagtac tatecttate 480  
 tatatagaca caaacaacgt actatagcag aagtttaagc tcgtaacca ccaatatttg 540  
 gtggcataga ccacgtattg ctgatatagt gctcgttaacc caccaatatt tcgtggcata 600  
 gagatctctt aggcaataaa ttagcagtac gaaacaatct atgtccacgt gttgctaata 660  
 caatgttcta aaccttacag cctactggac agttctctag ccatgataca tgtgcatgct 720  
 cgaacaaata tttatgggta cccgaaagg taattttttg tagtatttat gagggggagg 780  
 ggggcgttga cgaaaaaat aacttagcta agcgttaattg gcttaaaaac atacaatgtt 840  
 gttccagcat caagcctacg tgatcatttc acaaaaccaa ctcaaaagat aggtgtcatg 900  
 ttcttttag tgcaaaactt aaggacacct accttgcaaa acttagcttt gttaccaga 960  
 atgaaccgct aagctcgagg agctctgaac ttacatgacc aatatatta aacacaaaag 1020  
 tcatgcatga ttttctttaa taagtatcga gcaataggt tcgggtgtct ttcgtctcat 1080  
 acctctattg tcctccgtga tcaacaaggg tggatccggg tggtgcaagg gggctcaagc 1140  
 ccccctacct ctccaaagg agaaagaagg gagaagaaag tgaaggaaga agaaaccccc 1200  
 tatattctaa tgctacctcc gccactgctg atcaacacaa cattcttaa accatttctt 1260  
 tggcatttgc gcatgttaca aggtacaaaa gagccagccc atatgccaag ttactaaact 1320  
 aaactatgat ccaccatgga gcgagaacaa acgtcaacag gcatcaacca atgcagcaat 1380  
 cttgatcgct agtaactgtcc ggcattatat ctgaaacaaa tccagatcac ccatctcatc 1440  
 acagtcacat gcattcatgg tcacgggaac cgtttagcaaa ccaccaacta atcagcattg 1500  
 caacactctt cctcctataa atgcagcgag cgggggacac cataaacat cacaggcact 1560  
 tag 1563  
 <210> 22  
 <211> 67  
 <212> DNA  
 <213> 谷子()  
 <400> 22

cgatcaagtt aatTTtgttt ctgctttgtg cgctgtgtt ccagtaatta ctttccgtgt 60  
 agcaaaa 67  
 <210> 23  
 <211> 1528  
 <212> DNA  
 <213> 水稻()  
 <400> 23  
 ggaagctaac tagtcacggc gaatacatga cgacatcggc ctacaacgca caacttcttg 60  
 gcataaaaagc ttcaatttca atgcccctat ctggaagccc taggcgccgc gcaaatgtaa 120  
 aacattcgct tcgcttggct tgttatecaa aatagagtat ggacctcca cagattggca 180  
 acccgtaggt aatcgaaaat ggetccatct gccctttgt cgaaggaatc aggaaacggc 240  
 cctcacctcc tggcggagt tagatatgtg aaagaatcta ggcgacctt gcagactgga 300  
 caacatgtga acaaaataaga ccaacgttat ggcaacaagc ctcgacgcta ctcaagtgg 360  
 gggaggccac cgcatttcc aacgaagcgc caaagaaagc cttgcagact ctaatgctat 420  
 tagtcgcta ggatatttgg aatgaaagga accgcagagt ttttcagcac caagagcttc 480  
 cggtagctag tctgatagcc aaaattaagg aggatgcaa aacatgggtc ttggcgggcg 540  
 cgaaacacct tgataggtgg cttacctttt aacatgttcg ggccaaaggc cttgagacgg 600  
 taaagtttcc tatttgcgct tgcgcatgta caattttatt cctctattca atgaaattgg 660  
 tggctcactg gttcattaaa aaaaaaagaa tctagcctgt tcgggaagaa gaggatttta 720  
 ttcgtgagag agagagagag agagagagag agaggagag agaaggagga ggaggatttt 780  
 caggcttcgc attgcccac ctctgcttct gttggcccaa gaagaatccc aggcgcccac 840  
 gggctggcag ttaccacgg acctacctag cctaccttag ctatctaagc gggccgacct 900  
 agtagctacg tgcctagtgt agattaaagt tggcgggcca gcaggaagcc acgctgcaat 960  
 ggcatcttcc cctgtccttc gcgtacgtga aaacaaacct aggtaagctt agaatcttct 1020  
 tgcccgttgg actgggacac ccaccaatcc caccatgccc cgatattcct ccggtctcgg 1080  
 ttcattgtgat gtctctctt gtgtgacac ggagcaagca ttcttaaagc gcaaaagaaa 1140  
 atcaccaact tgctcacgca gtcacgtgc accgcgcaa gcgacgccc ataggccaag 1200  
 atcgcgagat aaaataacaa ccaatgatca taaggaaaca agcccgcgat gtgtcgtgtg 1260  
 cagcaatctt ggtcatttgc gggatcgagt gttcacggc taaccaata ttcggccgat 1320  
 gatttaacac attatcagcg tagatgtacg tacgatttgt taattaatct acgagccttg 1380  
 ctagggcagg tgttctgcca gccaatccag atgcacctg tatgcacgt cacatgatgg 1440  
 cagggcaggg ttcacatgag ctctaacggc cgattaatta atcccggggc tcgactataa 1500  
 atacctccct aatcccctga tcaaaacc 1528  
 <210> 24  
 <211> 99  
 <212> DNA  
 <213> 水稻()  
 <400> 24  
 atctcaagca gcctaatacat ctccagctga tcaagagctc ttaattagct agctagtgt 60



tagctgcgct tgtgatcgat cgatctcggg tacgtagca 99  
<210> 25  
<211> 27  
<212> DNA  
<213> 人工序列()  
<220>  
<221> misc\_feature  
<222> (1) .. (27)  
<223> RCc3\_Inter\_1扩增引物。  
<400> 25  
cagcacgttg ggcacacacgc ccagctt 27  
<210> 26  
<211> 27  
<212> DNA  
<213> 人工序列()  
<220>  
<221> misc\_feature  
<222> (1) .. (27)  
<223> RCc3\_Outer\_1扩增引物。  
<400> 26  
gcacaccgcs gcctcsaggt cgacgag 27  
<210> 27  
<211> 27  
<212> DNA  
<213> 人工序列()  
<220>  
<221> misc\_feature  
<222> (1) .. (27)  
<223> RCc3\_Inter\_2扩增引物。  
<400> 27  
aggttgatgc cgagsatggt gsccttg 27  
<210> 28  
<211> 28  
<212> DNA  
<213> 人工序列()  
<220>  
<221> misc\_feature  
<222> (1) .. (28)  
<223> RCc3\_Outer\_2扩增引物。

<400> 28

cttgccgcag trgttgagga kgaggctg 28

<210> 29

<211> 2001

<212> DNA

<213> 人工序列()

<220>

<221> misc\_feature

<222> (1) .. (2001)

<223> 具有可加工内含子的密码子重新设计的GUS编码序列。

<400> 29

atggtgagge ccgttgagac cccgactagg gagatcaaga agctggacgg cctctgggce 60  
 ttctccctcg accgtgagaa ctgcggcate gaccagcget ggtgggagtc cgccctccag 120  
 gagtctaggg ccatcgccgt gcccggttcc ttcaacgacc agttcgccga cgccgacatc 180  
 cgcaactacg cgggcaacgt ctggtatcag cgcgaggtgt tcateccgaa gggctgggce 240  
 ggccagcgcgca tcgtgctccg cttcgacgcc gtgaccact acggcaaggt ctgggtgaac 300  
 aatcaggagg taagtctctg cttctacctt tgatataat ataataatta tcattaatta 360  
 gtagtaatat aatatttcaa atattttttt caaaataaaa gaatgtagta tatagcaatt 420  
 gcttttctgt agtttataag tgtgtatatt ttaatttata acttttctaa tatatgacca 480  
 aaatttggtg atgtgcaggt gatggagcac cagggcgggt acaccccggt cgaggccgac 540  
 gtgacgccgt acgtgatcgc cgggaagtcc gtccgcatca ccgtctcgt gaacaatgag 600  
 ctgaactggc agaccatccc gcctggcatg gtcatcaccg acgagaacgg caagaagaag 660  
 cagtcctact tccacgactt cttcaactac gctggcatec accgctcgt gatgctctac 720  
 accactccca acacctgggt ggacgacatc accgtggta cccacgtggc ccaggactgc 780  
 aaccacgcct ccgtggactg gcaagtcgtt gccaacggcg acgtcagcgt cgagctgcgc 840  
 gacgccgacc agcaagtcgt tgccaccggc cagggcacca gcggcaccct ccaagtcgtc 900  
 aaccctcacc tctggcagcc tggcgagggc tacctctacg agctgtcgt caccgccaag 960  
 agccagactg agtgcgacat ctaccctctc cgcgtcgca tcaggagcgt cgctgtcaag 1020  
 ggcgagcagt tcctcatcaa ccacaagcct ttctacttea ctggtttcgg ccgccacgag 1080  
 gacgctgacc tgaggggcaa gggtttcgac aacgtctga tgggtccacga ccacgctctg 1140  
 atggactgga tcggtgcaa cagctacagg accagteact accgtacgc tgaggagatg 1200  
 ctggactggg ctgacgagca cggtatcgtc gtgatcgacg agactgctgc ggtcggttcc 1260  
 aacctgtctc tgggcattgg tttcgagget gggaacaagc cgaaggagct gtactctgag 1320  
 gaagctgtca acggcgagac tcagcaaget catctccagg cgattaagga gctgattgcc 1380  
 agggacaaga accatccgtc tctcgtgatg tggctctatt cgaatgagcc ggacaccaga 1440  
 ccgcaagggg cgcgtgaata cttcgcgccg ctggcggagg cgactcgca actggacca 1500  
 acccgtccaa tcacgtcgt caatgtcatg ttctgcgacg ccctacgga tacgatctcg 1560  
 gacctgttcg atgttctttg tctcaatcgg tactatgggt ggtatgttca gagcggggat 1620  
 cttgagacgg cggagaaggt tcttgagaag gaactcctgg cgtggcaaga gaagctccat 1680

cagccgatca ttatcacgga gtacggggtt gacacacttg cgggccttca cagtatgtac 1740  
 acagatatgt ggtcggagga ataccagtgt gcatgggttg atatgtacca tcgtgtcttc 1800  
 gaccggggtt cagcggttgt cggcgaacaa gtctggaact tcgcagactt cgccacgagc 1860  
 caagggatac tgcgggtagg agggaacaag aagggaatct tcacacggga tcggaagccc 1920  
 aagtcagcag ccttcctggt gcagaagcga tggacaggaa tgaacttcgg agaaaagcca 1980  
 cagcaaggcg gaaagcagtg a 2001

<210> 30

<211> 1812

<212> DNA

<213> 人工序列()

<220>

<221> misc\_feature

<222> (1) .. (1812)

<223> 密码子重新设计的GUS编码序列。

<400> 30

atggtgagge ccgttgagac cccgactagg gagatcaaga agctggacgg cctctgggcc 60  
 ttctccctcg accgtgagaa ctgcggcacc gaccagcget ggtgggagtc cgccctccag 120  
 gagtctaggg ccatcgccgt gcccggttcc ttcaacgacc agttcgccga cgccgacatc 180  
 cgcaactacg cgggcaacgt ctggtatcag cgcgaggtgt tcatcccga gggctgggcg 240  
 ggccagcgca tcgtgctccg cttcgacgcc gtgaccact acggcaaggt ctgggtgaac 300  
 aatcaggagg tgatggagca ccagggcggc tacaccccg tcgaggccga cgtgacgccg 360  
 tacgtgatcg ccgggaagtc cgtccgcacc accgtctgcg tgaacaatga gctgaactgg 420  
 cagaccatcc cgctggcat ggtcatcacc gacgagaacg gcaagaagaa gcagtcctac 480  
 ttccacgact tcttcaacta cgctggcacc caccgctccg tgatgctcta caccactccc 540  
 aacacctggg tggacgacat caccgtggc acccactggg ccaggactg caaccacgcc 600  
 tccgtggact ggcaagtcgt tgccaacggc gacgtcagcg tcgagctgcg cgacgccgac 660  
 cagcaagtcg ttgccaccgg ccagggcacc agcggcacc tccaagtcgt caaccctcac 720  
 ctctggcagc ctggcgaggg ctaccttac gagctgtgcg tcaccgcaa gagccagact 780  
 gagtgcgaca tctaccctct ccgctcggc atcaggagcg tcgctgtcaa gggcgagcag 840  
 ttctcatca accacaagcc tttctacttc actggtttcg gccgccacga ggacgctgac 900  
 ctgaggggca agggtttcca caacgtctg atggctccag accacgtct gatggactgg 960  
 atcggtgcca acagctacag gaccagtcac taccgtagc ctgaggagat gctggactgg 1020  
 gctgacgagc acggtatcgt cgtgatcgac gagactgctg cggtcggttt caacctgtct 1080  
 ctgggcattg gtttcgagge tgggaacaag ccgaaggagc tgtactctga ggaagctgtc 1140  
 aacggcgaga ctcagcaagc tcattccag gcgattaagg agctgattgc cagggacaag 1200  
 aaccatccgt ctgtcgtgat gtggtctatt gcgaatgagc cggacaccag accgcaaggg 1260  
 gcgcgtgaat acttcgcgcc gctggcggag gcgactcgca aactggacc aaccgtcca 1320  
 atcacgtgcg tcaatgcat gttctgcgac gccatacgg atacgatctc ggacctgttc 1380  
 gatgttcttt gtctcaatcg gtactatggg tggatgttc agagcgggga tcttgagacg 1440

gcggaagaagg ttcttgagaa ggaactcctg gcgtggcaag agaagctcca tcagccgatc 1500  
 attatcacgg agtacgggggt tgacacactt gcgggccttc acagtatgta cacagatag 1560  
 tggtcggagg aataccagtg tgcattggtg gatatgtacc atcgtgtctt cgaccgggtt 1620  
 tcagcgggtg tcggcgaaca agtctggaac ttcgcagact tcgccacgag ccaagggata 1680  
 ctgcgggtag gagggaaaca gaagggaatc ttcacacggg atcggaagcc caagtcagca 1740  
 gccttcctgt tgcagaagcg atggacagga atgaacttcg gagaaaagcc acagcaaggc 1800  
 ggaaagcagt ga 1812

<210> 31

<211> 1812

<212> DNA

<213> 大肠杆菌()

<400> 31

atgggtccgtc ctgtagaaac cccaaccctg gaaatcaaaa aactcgacgg cctgtgggca 60  
 ttcagtctgg atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag cgcgttaca 120  
 gaaagccggg caattgctgt gccaggecgt tttaacgac agttcgcga tgcagatatt 180  
 cgtaattatg cgggcaacgt ctggtatcag cgcgaagtct ttataccgaa aggttgggca 240  
 ggccagcgtg tcgtgctgct tttcgtgctg gtcactcatt acggcaaagt gtgggtcaat 300  
 aatcaggaag tgatggagca tcagggcggc tatacgcct ttgaagccga tgtcacgccg 360  
 tatgttattg ccgggaaaag tgtacgtatc accgtttgtg tgaacaacga actgaactgg 420  
 cagactatcc cgccgggaat ggtgattacc gacgaaaac gcaagaaaa gcagtcttac 480  
 ttccatgatt tctttaacta tgccggaatc catcgcagcg taatgcteta caccacgccg 540  
 aacacctggg tggacgatat caccgtggtg acgcatgtcg cgcaagactg taaccacgccg 600  
 tctgttgact ggcaggtggt ggccaatggt gatgtcagcg ttgaactgcg tgatgcggat 660  
 caacaggtgg ttgcaactgg acaaggcact agcgggactt tgcaagtggg gaatccgcac 720  
 ctctggcaac cgggtgaagg ttatctctat gaactgtgct tcacagccaa aagccagaca 780  
 gagtgtgata tctaccgct tcgcgtcggc atccggtcag tggcagtga ggcgaacag 840  
 ttctgatta accacaaacc gttctacttt actggctttg gtcgtcatga agatgcggac 900  
 ttgcgtggca aaggattcga taactgtctg atgggtgcag accacgcatt aatggactgg 960  
 attggggcca actcctaccg tacctcgcct tacccttac ctgaagagat gctcgactgg 1020  
 gcagatgaac atggcatcgt ggtgattgat gaaactgctg ctgtcggctt taacctctct 1080  
 ttaggcattg gtttcgaagc gggcaacaag ccgaaagaac tgtacagcga agaggcagtc 1140  
 aacggggaaa ctacgaagc gcaactacag gcgattaaag agctgatagc gcgtgacaaa 1200  
 aaccacccaa gcgtggtgat gtggagtatt gccaacgaac cggatacccg tccgcaaggt 1260  
 gcacgggaat atttcgcgcc actggcggaa gcaacgcgta aactcgacc gacgcgtccg 1320  
 atcacctgcg tcaatgtaat gttctgcgac gctcacaccg ataccatcag cgatctcttt 1380  
 gatgtgctgt gcctgaaccg ttattacgga tggatgtcc aaagcggcga tttggaaacg 1440  
 gcagagaagg tactggaaaa agaacttctg gcctggcagg agaaactgca tcagccgatt 1500  
 atcatcaccg aatacggcgt ggatacgtta gccgggctgc actcaatgta caccgacatg 1560  
 tggagtgaag agtatcagtg tgcattgctg gatatgtatc accgcgtctt tgatcgcgctc 1620

agcgcctcgc tcggtgaaca ggtatggaat ttcccgatt ttgcgacctc gcaaggcata 1680  
 ttccgcgttg gcggtaacaa gaaagggatc ttactcgcg accgcaaacc gaagtcggcg 1740  
 gcttttctgc tgcaaaaacg ctggactggc atgaacttcg gtgaaaaacc gcagcaggga 1800  
 ggcaaacaat ga 1812

<210> 32

<211> 2001

<212> DNA

<213> 人工序列()

<220>

<221> misc\_feature

<222> (1) .. (2001)

<223> 包含具有可加工内含子的原生大肠杆菌GUS编码序列的嵌合编码序列。

<400> 32

atggtccgtc ctgtagaaac cccaaccgt gaaatcaaaa aactcgacgg cctgtgggca 60  
 ttcagtctgg atcgcgaaaa ctgtggaatt gatcagcgtt ggtgggaaag cgcgttacia 120  
 gaaagccggg caattgctgt gccaggcagt tttaacgatc agttcgccga tgcagatatt 180  
 cgtaattatg cgggcaacgt ctggtatcag cgcgaaagtct ttataccgaa aggttgggca 240  
 ggccagcgtc tcgtgctcgc tttcgtatgc gtcactcatt acggcaaagt gtgggtcaat 300  
 aatcaggaag tgatggagca tcagggcggc tatacgcctt ttgaagccga tgtcacgccg 360  
 tatgttattg ccgggaaaag tgtacgtaag tttctgcttc tacctttgat atatatataa 420  
 taattatcat taattagtag taatataata tttcaaatat ttttttcaaa ataaaagaat 480  
 gtagtatata gcaattgctt ttctgtagtt tataagtgtg tatattttaa tttataactt 540  
 ttctaataata tgaccaaaat ttgttgatgt gcaggtatca ccgtttgtgt gaacaacgaa 600  
 ctgaactggc agactatccc gccgggaatg gtgattaccg acgaaaacgg caagaaaaag 660  
 cagtcttact tccatgattt cttaactat gccggaatcc atcgcagcgt aatgctctac 720  
 accacgccga acacctgggt ggacgatatc accgtgggta cgcatgtcgc gcaagactgt 780  
 aaccacgcgt ctgttgactg gcaggtgggt gccaatgggt atgtcagcgt tgaactgcgt 840  
 gatgcggatc aacaggtgggt tgcaactgga caaggcacta gcgggacttt gcaagtgggtg 900  
 aatccgcacc tctggcaacc gggatgaagg tctctctatg aactgtgcgt cacagccaaa 960  
 agccagacag agtgtgatat ctaccgctt cgcgtcgca tccggtcagt ggcagtgaag 1020  
 ggcgaaacagt tctgattaa ccacaaaccg ttctacttta ctggctttgg tcgtcatgaa 1080  
 gatgcggact tgcgtggcaa aggattcgat aacgtgctga tgggtgcacga ccacgcatta 1140  
 atggactgga ttggggccaa ctctaccgt acctgcatt accttacgc tgaagagatg 1200  
 ctgcactggg cagatgaaca tggcatcgtg gtgattgatg aaactgctgc tgcggcttt 1260  
 aacctctctt taggcattgg tttcgaagcg ggcaacaagc cgaaagaact gtacagcgaa 1320  
 gaggcagtca acggggaaac tcagcaagcg cacttacagg cgattaaaga gctgatagcg 1380  
 cgtgacaaaa accaccaag cgtgggtgat tggagtattg ccaacgaacc ggatacccg 1440  
 ccgcaagggt cacgggaata tttcgcgcca ctggcggaag caacgcgtaa actcgaccg 1500  
 acgcgtccga tcacctgcgt caatgtaatg ttctgcgacg ctcacaccga taccatcagc 1560

gatctctttg atgtgctgtg cctgaaccgt tattacggat ggtatgtcca aagcggcgat 1620  
 ttggaaacgg cagagaaggt actggaaaaa gaacttctgg cctggcagga gaaactgcat 1680  
 cagccgatta tcatcaccga atacggcgtg gatacgttag ccgggctgca ctcaatgtac 1740  
 accgacatgt ggagtgaaga gtatcagtggt gcatggctgg atatgtatca ccgctctttt 1800  
 gatcgcgtca gcgccgtcgt cggatgaacag gtatggaatt tcgccgattt tgcgacctcg 1860  
 caaggcatat tgcgcgttgg cggtaacaag aaagggatct tcaactcgca ccgcaaaccg 1920  
 aagtcggcgg cttttctgct gcaaaaacgc tggactggca tgaacttcgg tgaaaaaccg 1980  
 cagcagggag gcaaaacaatg a 2001

<210> 33

<211> 435

<212> DNA

<213> 谷子()

<400> 33

ggcgcccttt gagaagaagc ttttggctctg ctgcgcttat catgttttgt ggcttctgtg 60  
 ttgtgattct tgatctgecc cttgetataca tttgtattgt actgtcctaa taagtggtag 120  
 ttgtgagggt attactgtgt ctggttattt acctagagga ggaattattg tctgctattt 180  
 ctggttttgc tgtttatgta atggtgaacc aaagaatgaa gctgcaggct actttgagaa 240  
 ggaaggggac ctgctgcttt ctatcttgtc atgcgtgatt acttgaacag tctgatgat 300  
 ctattaatgt tctttggca gtgcaagtgt ttggtgtagc tccaacaggt agtgtttatg 360  
 tttggtgaag cagcaatggc cgactgtatg tgtttggtaga agctgcaacc tgcttgtgct 420  
 aactgaacat gcaga 435

<210> 34

<211> 2219

<212> DNA

<213> 人工序列()

<220>

<221> misc\_feature

<222> (1) .. (2219)

<223> 嵌合转录调控表达元件组 (EXP)。

<400> 34

attatggcca attatttate ggectggtcc tggcaggatg agtgtagtag tgtacctaat 60  
 ttggtaggca atgggtttga gttaggccaa taaatgtgca ggtttacatt tggttgcat 120  
 gttagtcat aactcatac acttccattt ttgtacgtcg gcgatgtttc aaccagaaca 180  
 gattatgttt ccttgggact gtaaatttct tcatcgetec atcccaatga acttgcaaga 240  
 gattcgtgtt agcaacaaaa attgetctgt cettcccaaa aagcattgcc tctggatctt 300  
 ttccgatcaa gatcagaaac cctatatgta ttgttcttat gttctgctgt gggctggtgg 360  
 ttcttgtcac ccttaaactt ctctgtttgg taccagata aatacaaatt cggctggtgg 420  
 gtagaaaaac cgcctacgaa cactggtttc acctataag actactgcac tgtttctagc 480  
 aggettattt ttctcttctg tttttatatt ccaatgtata catattaatg tttatcgagc 540

tatgatctta taaatttgta catgettcaa tattttctaa aaacttgaat aacagtatgt 600  
 aaaacctagg ttgatgtttc aaatgaactt atttaacatt ttacgttgaa acagtacatc 660  
 gcgaatggca tattattttg ttgcatttat tctaaacacc ttaaaatgga atttgaaaac 720  
 gggctctaag tttgagagaa gttaagggt aatagtattc taaacacttc aagtttgaga 780  
 tccaaaataa ttaatctctc acctatcacc tccaatcaag ttgtttatca gtttatgcc 840  
 tgtacatgta tcgctggttt gttatttcac ctcatcttcg ctatgtatct actactattg 900  
 cgctaattctc aaatattaga tgacatgtaa actaaaatct ttggaaagat taataggata 960  
 cccgcgggtg ctgatatctg tttctaaaaa tgttgaatct aactatttgt aaatattgat 1020  
 atatttttca gaaatgttgg atttagtctt gtgaaatgtt gaaccgatat ctctgatatg 1080  
 gatttaaatgg gcttaaagct tccactagcc gtgtgcatgg gcgaaaaaaa atcttggcca 1140  
 agtgttactc cgtcgcggcc acacgccaca agtctctccc gccctcctc gcccttate 1200  
 ccatcgtacc cgccacacgg cgcgccacca gtggcgccgc ggatgcgect catctccccg 1260  
 gcggccacct cgcgcggttt agatttccct gggeccccct cgcggtaccg tcacatattt 1320  
 ttggcgccctc ttcttctgcg cccctctcct cccgaaccgc agataccacc gagtcggcag 1380  
 ctgaacacaa gcaacaagca agtgatcccc gaccgaccgt cttcggtacg cgtcactcc 1440  
 gccctctgcc tttgttactg ccacgtttct ctgaatgctc tcttgtgtgg tgattgctga 1500  
 gagtggttta gctggatcta gaattacact ctgaaatcgt gttctgcctg tgctgattac 1560  
 ttgccgtcct ttgtagcagc aaaatatagg gacatggtag tacgaaacga agatagaacc 1620  
 tacacagcaa tacgagaaat gtgtaatttg gtgcttagcg gtatttattt aagcacatgt 1680  
 tgggtttata gggcacttgg attcagaagt ttgctgttaa tttaggcaca ggcttcatac 1740  
 tacatgggtc aatagtatag ggattcatat tataggcgat actataataa tttgttcgctc 1800  
 tgcagagctt attatttgcc aaaattagat attcctatc tgtttttggt tgtgtgctgt 1860  
 taaattgtta acgcctgaag gaataaatat aaatgacgaa attttgatgt ttatctctgc 1920  
 tcctttattg tgaccataag tcaagatcag atgcacttgt tttaaatatt gttgtctgaa 1980  
 gaaataagta ctgacagtat tttgatgcat tgatctgctt gtttgttgta acaaaattta 2040  
 aaaataaaga gtttcccttt tgttctctc cttacctct gatggtatct agtatctacc 2100  
 aactgacact atattgcttc tctttacata cgtatcttgc tcgatgcctt ctccctagtg 2160  
 ttgaccagtg ttactcacat agtctttgct catttcattg taatgcagat accaagcgg 2219  
 <210> 35  
 <211> 2224  
 <212> DNA  
 <213> 人工序列()  
 <220>  
 <221> misc\_feature  
 <222> (1) .. (2224)  
 <223> 嵌合转录调控表达元件组 (EXP)。  
 <400> 35  
 attatggcca attatttate ggccctgtcc tggcaggatg agtgtagtag tgtacctaat 60  
 ttggtaggca atgggtttga gttaggccaa taaatgtgca ggtttacatt tggttgccat 120

gttagtcat aactcatac acttccattt ttgtacgteg gcgatgttcc aaccagaaca 180  
 gattatgttt ccttgggact gtaaatttct tcategetcc atcccaatga acttgcaaga 240  
 gattcgtgtt agcaacccaaa attgctctgt ccttcccaaa aagcattgcc tctggatcct 300  
 ttccgatcaa gatcagaaac cctatatgta ttgttcttat gttctgctgt gggctcgttg 360  
 ttcttgtcac ccttaaactt ctctgtttgg taccagata aatacaaat cggtcgttg 420  
 gtagaaaaac cgcctacgaa cactggttcc acctataag actactgcac tgtttctagc 480  
 aggcttattt ttctcttctg tttttatatt ccaatgtata catattaatg tttatcgagc 540  
 tatgatctta taaatttcta catgcttcaa tttttctaa aaacttgaat aacagtatgt 600  
 aaaacctagg ttgatgttcc aaatgaactt atttaacatt ttacgttgaa acagtacatc 660  
 gcgaatggca tattattttg ttgcatttat tctaaacacc ttaaaatgga atttgaaaac 720  
 gggctctaag tttgagagaa gttaagggt aatagtatcc taaacacttc aagtttgaga 780  
 tccaaaataa ttaatctctc acctatcacc tccaatcaag ttgtttatca gtttatgcca 840  
 tgtacatgta tcgctgggtt gttatttccac ctcttttctg ctatgtatct actactattg 900  
 cgctaatctc aaatattaga tgacatgtaa actaaaatct ttggaaagat taataggata 960  
 cccgcgggtg ctgatatctg tttctaaaaa tgttgaatct aactatttgt aaatattgat 1020  
 atatttttca gaaatgttgg atttagtctt gtgaaatgtt gaaccgatat ctctgatatg 1080  
 gatttaatgg gcttaaagct tccactagcc gtgtgcatgg gcgaaaaaaa atcttggcca 1140  
 agtgttactc cgtcgcggcc acacgccaca agtctctccg gccctcctc gcccttctc 1200  
 ccatcgtacc cgccacacgg cgcgccacca gtggcgccgc ggatgcgcct catctctccg 1260  
 gcggccacct cgcgcgggtt agatttccct gggccccct cgcggtaccg tcacatattt 1320  
 ttggcgcctc ttcttctgcg cccctctct cccgaaccgc agataccacc gagtcggcag 1380  
 ctgaacacaa gcaacaagca agtgatcccg gaccggaccg accgtctctg gtacgcgctc 1440  
 actccgccct ctgcctttgt tactgccacg tttctctgaa tgctctcttg tgtggtgatt 1500  
 gctgagagtg gtttagctgg atctagaatt aactctgaa atcgtgttct gcctgtgctg 1560  
 attacttgcc gtcctttgta gcagcaaat ataggacat ggtagtacga aacgaagata 1620  
 gaacctacac agcaatacga gaaatgtgta atttggtgct tagcggattt tatttaagca 1680  
 catgttgggtg ttatagggca cttggattca gaagtttct gtttaatttag gcacaggctt 1740  
 catactacat gggcaatag tatagggatt catattatag gcgatactat aataatttgt 1800  
 tcgtctgcag agcttattat ttgccaaaat tagatattcc tattctgttt ttgtttgtgt 1860  
 gctgttaaat tgtaaacgc tgaaggaata aatataaat acgaaatctt gatgtttatc 1920  
 tctgctcctt tattgtgacc ataagtcaag atcagatgca cttgttttaa atattgttgt 1980  
 ctgaagaaat aagtactgac agtattttga tgcattgate tgcttgtttg ttgtaacaaa 2040  
 atttaaaaat aaagagtctc cttttgttg ctctcttac ctctgatgg tatctagtat 2100  
 ctaccaactg aactatatt gcttctctt acatacgtat ctgctcgat gcctctctcc 2160  
 tagtgttgac cagtgttact cacatagctt ttgctcattt cattgtaatg cagataccaa 2220  
 gcgg 2224  
 <210> 36  
 <211> 2966  
 <212> DNA



<213> 人工序列()

<220>

<221> misc\_feature

<222> (1)..(2966)

<223> 嵌合转录调控表达元件组 (EXP)。

<400> 36

```

gggccgattg agacttttca acaaagggtg atatccggaa acctcctcgg attccattgc 60
ccagctatct gtcactttat tgtgaagata gtggaaaagg aaggtggctc ctacaaatgc 120
catcattgcg ataaaggaaa ggccatcggt gaagatgect ctgccgacag tgggtcccaa 180
gatggacccc cacccacgag gagcatcggt gaaaaagaag acgttccaac cacgtcttca 240
aagcaagtgg attgatgtga tgggccgatt gagacttttc acaaagggt aatatccgga 300
aacctcctcg gattccattg cccagctatc tgtaacttta ttgtgaagat agtggaaaag 360
gaaggtgggt cctacaaatg ccatcattgc gataaaggaa aggccatcgt tgaagatgcc 420
tctgccgaca gtggtcccaa agatggaccc ccaccacga ggagcatcgt ggaaaaagaa 480
gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg atcggaagct aactagtcac 540
ggcgaataca tgacgacatc ggccataaac gcacaacttc ttggcataaa agcttcaatt 600
tcaatgcccc tatctggaag ccctaggcgc cgcgcaaagt taaaacattc gcttcgcttg 660
gcttgttatc caaaatagag tatggacctc cgacagattg gcaaccctg ggtaatcgaa 720
aatggctcca tctgcccctt tgtcgaagga atcaggaaac ggccctcacc tcttgccgga 780
gtgtagatat gtgaaagaat ctaggcgaca cttgcagact ggacaacatg tgaacaaata 840
agaccaacgt tatggcaaca agcctcgacg ctactcaagt ggtgggagc caccgcatgt 900
tccaacgaag cgccaaagaa agccttgacg actctaattc tattagtcgc ctaggatatt 960
tggaatgaaa ggaaccgcag agtttttcag caccaagagc ttccggtggc tagtctgata 1020
gccaaaatta aggaggatgc caaaacatgg gtcttggcgg ggcgaaaca cttgatagg 1080
tggttacct ttaacatgt tcgggcaaaa ggccctgaga cggtaaagtt ttctatttgc 1140
gcttgcgcat gtacaatttt attcctctat tcaatgaaat tgggtggctca ctggttcatt 1200
aaaaaaaaaa gaatctagcc tgttcgggaa gaagaggatt ttattcgtga gagagagaga 1260
gagagagaga gagagaggga gagagaagga ggaggaggat tttcaggctt cgcattgccc 1320
aacctctgct tctgttgccc caagaagaat cccaggcgc catgggctgg cagtttacca 1380
cggacctacc tagcctacct tagctatcta agcgggcccga cctagtagct acgtgcctag 1440
tgtagattaa agttggcggg ccagcaggaa gccacgtgc aatggcatct tcccctgtcc 1500
ttcgcgtacg tgaaaacaaa cccaggtaag cttagaatct tcttgcccgt tggactggga 1560
cacccaccaa tcccaccatg ccccgatatt cctccggtct cggttcatgt gatgtcctct 1620
cttgtgtgat cacggagcaa gcattcttaa acggcaaaag aaaatcacca acttgctcac 1680
gcagtcacgc tgaccgcgc gaagcgacgc ccgatagccc aagatcgca gataaaataa 1740
caaccaatga tcataaggaa acaagcccgc gatgtgtcgt gtgcagcaat cttggtcatt 1800
tgcgggatcg agtgcttca ggetaaccaa atattcggcc gatgatttaa cacattatca 1860
gcgtagatgt acgtacgatt tgtaattaa tctacgagcc ttgctagggc aggtgttctg 1920
ccagccaatc cagatcgcgc tcgtatgcac gctcacatga tggcagggca ggttccat 1980

```

gagctctaac ggtcgattaa ttaatccccg ggctcgacta taaatacctc cctaatecca 2040  
 tgatcaaaac catctcaagc agcctaatac tctccagctg atcaagagct cttattagc 2100  
 tagctagtga ttagctgcgc ttgtgatcga tcgatctcgg gtacgtagca cggaccggac 2160  
 cgaccgtctt cggtagcgc tcaactccgc ctctgccttt gttactgcca cgtttctctg 2220  
 aatgctctct tgtgtggtga ttgctgagag tggtttagct ggatctagaa ttacactctg 2280  
 aaatcgtggt ctgcctgtgc tgattacttg ccgtcctttg tagcagcaaa atataggac 2340  
 atggtagtac gaaacgaaga tagaacctac acagcaatac gagaaatgtg taatttgggtg 2400  
 cttagcggta tttatttaag cacatgttgg tgttataggg cacttggatt cagaagtttg 2460  
 ctgttaattt aggcacagge ttcatactac atgggtcaat agtataggga ttcataattat 2520  
 aggcgatact ataataattt gttcgtctgc agagcttatt atttgccaaa attagatatt 2580  
 cctattctgt ttttgtttgt gtgctgttaa attgttaacg cctgaaggaa taaatataaa 2640  
 tgacgaaatt ttgatgttta tctctgetec tttattgtga ccataagtca agatcagatg 2700  
 cacttgtttt aaatattggt gtctgaagaa ataagtactg acagtatttt gatgcattga 2760  
 tctgcttgtt tgttgtaaca aaatttaaaa ataaagagtt tctttttgtg tgcctctctt 2820  
 acctcctgat ggtatctagt atctaccaac tgacactata ttgcttctct ttacatacgt 2880  
 atcttgcctg atgccttctc cctagtgttg accagtgtta ctcacatagt ctttgcctcat 2940  
 ttcatgttaa tgcagatacc aagcgg 2966

<210> 37

<211> 2005

<212> DNA

<213> 玉米墨西哥亚种()

<400> 37

gtcgtgcccc tctctagaga taaagagcat tgcattgteta aagtataaaa aattaccaca 60  
 tatttttttg tcacacttat ttgaagtga gtttatctat ctctatacat atattttaa 120  
 ttcactctac aaataatata gtctataata ctaaaataat attagtgttt tagaggatca 180  
 tataaataaa ctgctagaca tggctctaaag gataattgaa tattttgaca atctacagtt 240  
 ttatcttttt agtgtgcatg tgatctctct gtttttttg caaatagctt gacctatata 300  
 atacttcatc cattttatta gtacatccat ttaggattta gggttgatgg tttctataga 360  
 ctaattttta gtacatccat tttattcttt ttagtcteta aattttttaa aactaaaact 420  
 ctatttttagt tttttattta ataattttaga tataaaatga aataaaataa attgactaca 480  
 aataaaacaa atacccttta agaaataaaa aaactaagea aacatttttc ttgtttcgag 540  
 tagataatga caggctgttc aacgccgtcg acgagtctaa cggacaccaa ccagegaacc 600  
 agcagcgtcg cgtcgggcca agcgaagcag acggcacggc atctctgtag ctgcctctgg 660  
 acccctctcg agagtccgc tccaccgttg gaettgetec gctgtcggca tccagaaatt 720  
 gcgtggcgga gcggcagacg tgaggcgga cggcaggcgg cctcttctc ctctcacggc 780  
 accggcagct acgggggatt ctttcccac cgetctctg ctttccctc ctgcccgc 840  
 gtaataaata gacacccctt ccacaccctc tttcccacac ctctgttctg ttcggagcgc 900  
 acacacacgc aaccagatct ccccaaatc cagcgtcgg cacctccgt tcaaggtacg 960  
 ccgtctatcc tcccccccc cctctcteta cttctctag atcggcgatc cggctcatgg 1020

ttagggcccc gtagttctac ttctgttcat gtttgtgta gagcaaacat gttcatgttc 1080  
 atgtttgtga tgatgtggtc tggttgggcg gtcgttctag atcggagtag gatactgttt 1140  
 caagctacct ggtggattta ttaattttgt atctgtatgt gtgtgccata catcttcata 1200  
 gttacgagtt taagatgatg gatggaaata tcgatctagg ataggtatac atgttgatgc 1260  
 gggttttact gatgcatata cagagatgct ttttttctcg cttggtttgt atgatatggt 1320  
 ctggttgggc ggtcgttcta gatcggagta gaatactggt tcaaactacc tgggtgattt 1380  
 attaaaggat aaagggcgt tctagatcgg agtagaatac tgtttcaaac tacctggtgg 1440  
 atttattaaa ggatctgtat gtatgtgect acatcttcat agttacgagt ttaagatgat 1500  
 ggatggaaat atcgatctag gataggtata catgttgatg cgggttttac tgatgcatat 1560  
 acagagatgc ttttttctgc ttggttgtga tgatgtggtc tggttgggcg gtcgttctag 1620  
 atcggagtag aatactgttt caaactacct ggtggattta ttaattttgt atctttatgt 1680  
 gtgtgccata catcttcata gttacgagtt taagatgatg gatggaaata ttgatctagg 1740  
 ataggtatac atgttgatgt gggttttact gatgcatata catgatggca tatgcggcat 1800  
 ctattcatat gctctaacct tgagtaceta tctattataa taaacaagta tgttttataa 1860  
 ttattttgat cttgatatac ttggatgatg gcatatgcag cagctatatg tggatTTTT 1920  
 agccctgect tcatacgeta tttatttgct tggactggt tcttttgtcc gatgctcacc 1980  
 ctgttgttgg gtgatacttc tgcag 2005

<210> 38

<211> 804

<212> DNA

<213> 水稻()

<400> 38

accgtcttcg gtacgcgctc actccgcct ctgccttgt tactgccacg tttctctgaa 60  
 tgctctcttg tgtggtgatt gctgagagtg gtttagctgg atctagaatt acactctgaa 120  
 atcgtgttct gcctgtgctg attacttgcc gtcctttgta gcagcaaat ataggacat 180  
 ggtagtacga aacgaagata gaacctacac agcaatacga gaaatgtgta atttggtgct 240  
 tagcggattt tatttaagca catgttgggt ttatagggca cttggattca gaagtttgcct 300  
 gttaatttag gcacaggctt catactacat gggtaacatg tataggatt catattatag 360  
 gcgatactat aataatttgt tcgtctgcag agcttattat ttgccaaat tagatattcc 420  
 tattctgttt ttgtttgtgt gctgttaaat tgtaacgcc tgaaggaata aatataaatg 480  
 acgaaatTTT gatgtttate tctgetcett tattgtgacc ataagtcaag atcagatgca 540  
 cttgttttaa atattgttgt ctgaagaaat aagtaactgac agtattttga tgcattgatc 600  
 tgcttgtttg ttgtaacaaa atttaaaaat aaagagtttc ctttttgttg ctctccttac 660  
 ctctctgatgg tatctagat ctaccaactg acactatatt gcttctcttt acatacgtat 720  
 cttgctcgat gccttctccc tagtgttgac cagtgttact cacatagtct ttgctcattt 780  
 cattgtaatg cagataccaa gcgg 804

<210> 39

<211> 253

<212> DNA

<213> 根癌农杆菌()

<400> 39

gatcgttcaa acatttggca ataaagtffc ttaagattga atcctgttgc cggctcttgcg 60  
 atgattatca tataatttct gttgaattac gttgaagcatg taataattaa catgtaatgc 120  
 atgacgttat ttatgagatg ggtttttatg attagagtcc cgcaattata catttaatac 180  
 gcgatagaaa acaaaaatata gcgcgcaaac taggataaat tatcgcgcgc ggtgtcatct 240  
 atgttactag atc 253

<210> 40

<211> 300

<212> DNA

<213> 水稻()

<400> 40

attaatcgat cctccgatcc ctttaattacc ataccattac accatgcate aatatccata 60  
 tatatataaa ccctttcgca cgtacttata ctatgttttg tcatacatat atatgtgtcg 120  
 aacgatcgat ctatcactga tatgatatga ttgatccate agcctgatct ctgtatcttg 180  
 ttatttgtat accgtcaaat aaaagtcttct tccacttgtg ttaataatta gctactctca 240  
 tctcatgaac cctatatata actagttaa tttgctgtca attgaacatg atgatcgatg 300

<210> 41

<211> 1446

<212> DNA

<213> 人工序列()

<220>

<221> misc\_feature

<222> (1)..(1446)

<223> 嵌合转录调控表达元件组(EXP)。

<400> 41

ggtccgattg agacttttca acaaagggta atatccggaa acctcctcgg attccattgc 60  
 ccagctatct gtcactttat tgtgaagata gtggaaaagg aaggtggctc ctacaaatgc 120  
 catcattgcg ataaaggaaa ggccatcggt gaagatgcct ctgccgacag tggccccaaa 180  
 gatggacccc caccacgag gagcatcgtg gaaaaagaag acgttccaac cacgtcttca 240  
 aagcaagtgg attgatgtga tgggccgatt gagacttttc acaaagggt aatatccgga 300  
 aacctcctcg gattccattg cccagctatc tgtaacttta ttgtgaagat agtggaaaag 360  
 gaaggtggct cctacaaatg ccatcattgc gataaaggaa aggccatcgt tgaagatgcc 420  
 tctgccgaca gtggccccaa agatggacce ccaccacga ggagcatcgt ggaaaaagaa 480  
 gacgttccaa ccacgtcttc aaagcaagtg gattgatgtg atatctccac tgacgtaagg 540  
 gatgacgcac aatccccacta tccttcgcaa gacccttctt ctatataagg aagttcattt 600  
 catttggaga ggacacgctg acaagetgac tctagcagat ctaccgtctt cggtagcgcg 660  
 tcaactccgc ctctgccttt gttactgcca cgtttctctg aatgctctct tgtgtggtga 720  
 ttgctgagag tggtttagct ggatctagaa ttacactctg aatcgtggtt ctgcctgtgc 780

tgattacttg ccgtcctttg tagcagcaaa atataggac atggtagtac gaaacgaaga 840  
 tagaacctac acagcaatac gagaaatgtg taatttgggtg cttagcggta tttatttaag 900  
 cacatgttgg tgttataggg cacttggatt cagaagtttg ctgttaattt aggcacaggc 960  
 ttcatactac atgggtcaat agtataggga ttcataattat aggcgatact ataataattt 1020  
 gttcgtctgc agagcttatt atttgccaaa attagatatt cctattctgt ttttgtttgt 1080  
 gtgctgttaa attgttaacg cctgaaggaa taaatataaa tgacgaaatt ttgatgttta 1140  
 tctctgctcc tttatttga ccataagtca agatcagatg cacttgtttt aaatattggt 1200  
 gtctgaagaa ataagtactg acagtatfff gatgcattga tctgcttgtt tgttgtaaca 1260  
 aaatttaaaa ataaagagtt tctttttgtg tctctctctt acctctgat ggtatctagt 1320  
 atctaccaac tgacactata ttgettctct ttacatacgt atcttgctcg atgccttctc 1380  
 cctagtgttg accagtgta ctcacatagt ctttgctcat ttcattgtaa tgcagatacc 1440  
 aagcgg 1446

<210> 42

<211> 1653

<212> DNA

<213> 萤火虫()

<400> 42

atggaagacg ccaaaaacat aaagaaaggc ccggcgccat tctatcctct agaggatgga 60  
 accgctggag agcaactgca taaggctatg aagagatacg ccctggttcc tggaacaatt 120  
 gcttttacag atgcacatat cgaggtgaac atcacgtacg cggataactt cgaatgtcc 180  
 gttcggttgg cagaagctat gaaacgatat gggctgaata caaatcacag aatcgctgta 240  
 tgcagtgaaa actctcttca attctttatg ccggtgttgg gcgcgttatt tatcggagtt 300  
 gcagttgcgc ccgcaacga catttataat gaacgtgaat tgctcaacag tatgaacatt 360  
 tgcgagccta ccgtagtggt tgtttccaaa aaggggttgc aaaaaatttt gaacgtgcaa 420  
 aaaaaattac caataatcca gaaaattatt atcatggatt ctaaacgga ttaccagga 480  
 tttcagtcga tgtacacgtt cgtcacatct catctacctc ccggttttaa tgaatacgat 540  
 tttgtaccag agtcctttga tctgacaaa acaattgcac tgataatgaa ttctctgga 600  
 tctactgggt tacctaaggg tgtggccctt ccgcatagaa ctgcctgcgt cagattctcg 660  
 catgccagag atcctatfff tggcaatcaa atcattccgg atactgcgat ttttaagtgtt 720  
 gttccattcc atcacggttt tggaatgttt actaacctcg gatatttgat atgtggattt 780  
 cgagtcgtct taatgtatag atttgaagaa gagctgtttt tacgatecct tcaggattac 840  
 aaaattcaaa gtgcgttget agtaccaccc ctatfffcat tcttcgcaa aagcactctg 900  
 attgacaaat acgatttate taatttacac gaaattgctt ctgggggcgc acctctttcg 960  
 aaagaagtcg gggaagcggg tgcaaacgc ttccatcttc cagggatacg acaaggatat 1020  
 gggctcactg agactacatc agctattctg attacaccg aggggatga taaaccgggc 1080  
 gcggtcggta aagttgttcc attttttgaa gcgaaggttg tggatctgga taccgggaaa 1140  
 acgctgggag ttaatcagag aggcgaatta tgtgtcagag gacctatgat tatgtccggt 1200  
 tatgtaaaca atccggaagc gaccaacgcc ttgattgaca aggatggatg gctacattct 1260  
 ggagacatag ctactggga cgaagacgaa cacttcttca tagttgaccg cttgaagtct 1320

ttaattaaat acaaaggata tcaggtggcc cccgctgaat tggaatcgat attgttaca 1380  
 caccccaaca tcttcgacgc gggcgtggca ggtcttccc acgatgacgc cggatgaactt 1440  
 cccgccgccc ttgttgtttt ggagcacgga aagacgatga cggaaaaaga gatcgtggat 1500  
 tacgtcgcca gtcaagtaac aaccgcgaaa aagttgctgcg gaggagtgtt gtttgtggac 1560  
 gaagtaccga aaggtcttac cggaaaactc gacgcaagaa aatcagaga gatcctcata 1620  
 aaggccaaga agggcggaaa gtccaaattg taa 1653

<210> 43

<211> 936

<212> DNA

<213> 海肾 ()

<400> 43

atggcttcca aggtgtacga ccccgagcaa cgcaaacgca tgatcactgg gcctcagtgg 60  
 tgggctcgtc gcaagcaaat gaacgtgctg gactccttca tcaactacta tgattccgag 120  
 aagcacgccg agaacgccgt gatTTTTctg catggtaacg ctgcctccag ctacctgtgg 180  
 aggcacgtcg tgctcacaat cgagcccgct gctagatgca tcatecctga tctgatcgga 240  
 atgggtaagt ccggcaagag cgggaatggc tcatatcgcc tcttgatca ctacaagtac 300  
 ctcaccgctt ggctcgagct gctgaacctt ccaaagaaaa tcattcttgt gggccacgac 360  
 tggggggcct gtctggcctt tcaactactcc tacgagcacc aagacaagat caaggccatc 420  
 gtccatgctg agagtgtcgt ggacgtgatc gactcctggg acgagtggcc tgacatcgag 480  
 gaggatatcg ccctgatcaa gagcgaagag ggcgagaaaa tgggtcctga gaataacttc 540  
 ttctctgaga ccatgctccc aagcaagatc atgcggaaac tggagcctga ggagtctcgt 600  
 gcctacctgg agccattcaa ggagaagggc gaggttagac ggctaccct ctctggcct 660  
 cgcgagatcc ctctcgtaa gggaggcaag cccgacgtcg tccagattgt ccgcaactac 720  
 aacgcctacc ttcgggccag cgacgatctg cctaagatgt tcacgagtc cgaccctggg 780  
 ttcttttcca acgctattgt cgaggagct aagaagtcc ctaacaccga gttcgtgaag 840  
 gtgaagggcc tccactcag ccaggaggac gctccagatg aatgggtaa gtacatcaag 900  
 agcttcgtgg agcgcgtgct gaagaacgag cagtaa 936

<210> 44

<211> 675

<212> DNA

<213> 人工序列 ()

<220>

<221> misc\_feature

<222> (1) .. (675)

<223> 嵌合转录调控表达元件组 (EXP)。

<400> 44

gggccgatgt gagacttttc aacaaagggt aatatccgga aacctctcgt gattccattg 60  
 ccagctatc tgctacttta ttgtgaagat agtggaaaag gaagtggtcct cctacaaatg 120  
 ccatcattgc gataaaggaa aggccatcgt tgaagatgcc tctgccgaca gtggtcccaa 180

---

agatggacce ccaccacga ggagcatcgt ggaaaaagaa gacgttccaa ccacgtcttc 240  
aaagcaagtg gattgatgtg atggtccgat gtgagacttt tcaacaaagg gtaatccg 300  
gaaacctcct cggattccat tgcccagcta tctgtcactt tattgtgaag atagtggaaa 360  
aggaaggtgg ctctacaaa tgccatcatt gcgataaagg aaaggccatc gttgaagatg 420  
cctctgccga cagtgggtccc aaagatggac cccaccac gagagcatc gtggaaaaag 480  
aagacgttcc aaccacgtct tcaaagcaag tggattgatg tgatatctcc actgacgtaa 540  
gggatgacgc acaatcccac tatccttcgc aagacccttc ctctatataa ggaagttcat 600  
ttcatttgga gaggaacctt cttccacaca ctcaagccac actattggag aacacacagg 660  
gacaacacac cataa 675







CR-Ec.uida\_nno-1:1:1  
 ATCAGGTGGTCAATGTCATGTTCTGGACGCCCAATACGGATACGATCTCGGAACCIGTTC  
 CR-Ec.uida-1:1:4  
 ATCACCTGGTCAATGTAATGTTCTGGACGCTCACACCGGATACCAATCAACCGATCTCTTT  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 GATGTTCTTTGTCATAATCGGTACTATGGGTGGTATGTTTCAGAGCGGGGATCTTGAGACG  
 CR-Ec.uida-1:1:4  
 GATGTCGTGCCCTGAACCGTTAATTACGGATGGTATGTCCAAAGCGCGGATTTGGAAACG  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 GCGGAGAAGGTTCTTGAAAGGAACCTCTGGCGGTGSCAAGAGAAGCTCCCATCAGCCGATC  
 CR-Ec.uida-1:1:4  
 GCAGAGAAGGTACTGGAAAAGAACTTCTGGCCCTGGCAGGAGAAACTGTCATCAGCCGATT  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 ATTATCACGGAGTACGGGGTTGACACACTTCGGGGCTTCACAGTATGTTACACAGATAAG  
 CR-Ec.uida-1:1:4  
 ATCATCACCGAATACGSCGTGGATACGTTAGCCGGCTGCACCTCAATGHIACACCCGACATG  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 TGGTCGGAGGAATACCAGTGTCCATGGTTGGATAATGTACCATCGTGTCTTCGACCCGGTT  
 CR-Ec.uida-1:1:4  
 TGGAGTGAAGAGTATCAGTGTGCAATGGCTGGATAATGTAATCACCCGGTCTTTGATCGCGTC  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 TCAGCGGTTGTCGGGAACAAGTCTGGAACCTTCGCAGACTTCGCCAUGGAGCCAAAGGATA  
 CR-Ec.uida-1:1:4  
 AGCCCCGTCGTCGGTGAACAGGTATGGAATTTCCCCGATTTCCGACCTCGCAAGGCATA  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 CTGCGGTTAGGAGGGAACAAGAAGGAATCTTCACACGCGGATCGGAAGCCCAAGTCAGCA  
 CR-Ec.uida-1:1:4  
 TTGCGGTTGGCGGTAACAAGAAGGGAATCTTCACTCGGACCCGCAACCCGAAGTCGGCG  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 GCCTTCCTGTTGCAGAAGCGGATGGACAGGAATGAACTTCGGAGAAAGCCACAGCAAGGC  
 CR-Ec.uida-1:1:4  
 GCTTTCCTGCAAAAACCGCTGGACTGGCATGAACTTCGGTGAATAAACCCGACAGCAGGGA  
 \*\*\*\*\*  
 CR-Ec.uida\_nno-1:1:1  
 GGAAAGCAGTGA  
 CR-Ec.uida-1:1:4  
 GGCAAAACAATGA  
 \*\*\*\*\*

图1c