



(12) 发明专利

(10) 授权公告号 CN 111019968 B

(45) 授权公告日 2023. 06. 23

(21) 申请号 201911405272.9

(22) 申请日 2019.12.31

(65) 同一申请的已公布的文献号  
申请公布号 CN 111019968 A

(43) 申请公布日 2020.04.17

(73) 专利权人 北京市农林科学院  
地址 100097 北京市海淀区曙光花园中路9  
号北京市农林科学院玉米研究中心

(72) 发明人 徐雯 杨进孝 武莹 张成伟  
康桂婷 刘亚

(74) 专利代理机构 北京纪凯知识产权代理有限公司 11245  
专利代理师 关畅

(51) Int. Cl.  
C12N 15/82 (2006.01)  
C12N 15/54 (2006.01)  
A01H 5/00 (2018.01)  
A01H 6/46 (2018.01)

(56) 对比文件

CN 110551752 A, 2019.12.10  
Yongwei Sun等. "Engineering Herbicide-Resistant Rice Plants through CRISPR/Cas9-Mediated Homologous Recombination of Acetolactate Synthase". 《Molecular Plant》. 2016, 第9卷(第4期), 第628-631页.  
Katarzyna Duda等. "High-efficiency genome editing via 2A-coupled co-expression of fluorescent proteins and zinc finger nucleases or CRISPR/Cas9 nickase pairs". 《Nucleic acids research》. 2014, 第42卷(第10期), 第e84篇.  
郭明璋等. 细菌多元基因编辑技术研究进展及其在合成生物学中的应用. 《生物技术通报》. 2018, 第34卷(第5期),

审查员 邢维玲

权利要求书2页 说明书12页  
序列表15页 附图2页

(54) 发明名称

NTS/dNTS组合在制备植物突变体中的应用

(57) 摘要

本发明公开了NTS/dNTS组合在制备植物突变体中的应用。该方法包括如下步骤：将esgRNA、Cas9切刻酶、筛选剂抗性蛋白、供体DNA导入植物中；esgRNA靶向DNA片段甲靶点序列；DNA片段甲靶点序列位于植物基因组中的转录链上；供体DNA的转录链依次由DNA片段甲靶点序列、DNA片段乙和DNA片段甲靶点序列组成；DNA片段乙为将DNA片段甲突变后得到的DNA分子；在esgRNA引导下，Cas9切刻酶在植物基因组的非转录链上产生一个单链DNA切刻，在供体DNA的非转录链上产生两个单链DNA切刻，通过植物体内的修复机制将植物基因组中DNA片段甲替换为DNA片段乙，实现植物基因替换。

1. 一种植物基因替换的方法,包括如下步骤:将esgRNA、Cas9切刻酶、筛选剂抗性蛋白、供体DNA导入目的植物中;

所述esgRNA靶向DNA片段甲靶点序列;所述DNA片段甲靶点序列位于目的植物基因组中的转录链上;

所述esgRNA结构如下:tRNA-所述DNA片段甲靶点序列转录的RNA-esgRNA骨架;

所述tRNA为将序列1第474-550位中的T替换为U得到的RNA分子;

所述esgRNA骨架为将序列1第571-656位中的T替换为U得到的RNA分子;

所述Cas9切刻酶为Cas9 D10A切刻酶;

所述Cas9 D10A切刻酶为SpCas9n蛋白质;

所述SpCas9n蛋白质的氨基酸序列如序列3所示;

所述Cas9切刻酶带有核定位信号;

所述核定位信号为SV40 NLS;

所述SV40 NLS的氨基酸序列如序列2所示;

所述Cas9切刻酶与所述筛选剂抗性蛋白通过依次由启动子、所述Cas9切刻酶的编码基因、自切割寡肽的编码基因、所述筛选剂抗性蛋白的编码基因和终止子组成的表达盒导入目的植物中;

所述自切割寡肽为来源于猪捷申病毒-1的2A自切割寡肽;

所述来源于猪捷申病毒-1的2A自切割寡肽的氨基酸序列如序列4所示;

所述筛选剂抗性蛋白为潮霉素磷酸转移酶;

所述潮霉素磷酸转移酶的氨基酸序列如序列5所示;

所述供体DNA的转录链依次包括所述DNA片段甲靶点序列、DNA片段乙和所述DNA片段甲靶点序列;

在所述esgRNA引导下,所述Cas9切刻酶或其变体在目的植物基因组中的DNA片段甲靶点序列的非转录链上产生一个单链DNA切刻,在所述供体DNA中的DNA片段甲靶点序列的非转录链上产生两个单链DNA切刻,并通过目的植物体内的修复机制将目的植物基因组中的所述DNA片段甲替换为所述DNA片段乙,实现植物基因替换;

所述DNA片段甲为序列5第1300-1993位所示的DNA分子;

所述DNA片段乙为序列1第8513-9206位所示的DNA分子;

所述植物为水稻。

2. 根据权利要求1所述的方法,其特征在于:所述SpCas9n蛋白质的编码基因如序列1第2887-6987位所示。

3. 根据权利要求1所述的方法,其特征在于:所述潮霉素磷酸转移酶的编码基因如序列1第7195-8220位所示。

4. 根据权利要求1所述的方法,其特征在于:所述将esgRNA、Cas9切刻酶或其变体、筛选剂抗性蛋白、供体DNA导入植物中的方法包括如下步骤:将转录esgRNA的DNA分子、Cas9切刻酶或其变体的编码基因、筛选剂抗性蛋白的编码基因和供体DNA导入目的植物中。

5. 根据权利要求4所述的方法,其特征在于:所述转录esgRNA的DNA分子、所述Cas9切刻酶或其变体的编码基因、所述筛选剂抗性蛋白的编码基因和所述供体DNA通过重组表达载体导入目的植物中。

6. 根据权利要求5所述的方法,其特征在于:所述重组表达载体包括依次由启动子、所述转录esgRNA的DNA分子和终止子组成的表达盒和依次由启动子、所述Cas9切刻酶或其变体的编码基因、自切割寡肽的编码基因、筛选剂抗性蛋白的编码基因和终止子组成的表达盒。

7. 权利要求1-6任一所述的方法在植物基因编辑中的应用;所述植物为水稻。

8. 权利要求1-6任一所述的方法在制备植物突变体中的应用;所述植物为水稻。

9. 权利要求1-6任一所述的方法在提高植物基因替换效率中的应用;所述植物为水稻。

10. 权利要求1-6任一所述的方法在减少植物基因替换产生的副产物中的应用;所述植物为水稻。

11. 一种植物基因编辑的方法,包括如下步骤:按照权利要求1-6任一所述的方法对植物基因组中的目的基因片段进行替换,从而实现植物基因编辑;所述植物为水稻。

12. 一种制备植物突变体的方法,包括如下步骤:按照权利要求1-6任一所述的方法对植物基因组中的目的基因片段进行替换,获得植物突变体;所述植物为水稻。

13. 一种提高植物基因替换效率的方法,包括如下步骤:按照权利要求1-6任一所述的方法对植物基因组中的目的基因片段进行替换;所述植物为水稻。

14. 一种减少植物基因替换产生的副产物的方法,包括如下步骤:按照权利要求1-6任一所述的方法对植物基因组中的目的基因片段进行替换;所述植物为水稻。

## NTS/dNTS组合在制备植物突变体中的应用

### 技术领域

[0001] 本发明属于生物技术领域,具体涉及NTS/dNTS组合在制备植物突变体中的应用。

### 背景技术

[0002] 长链DNA模板介导的基因精确替换在细胞中的概率非常低,但是在待替换位点附近引入DNA双链断裂(dsDNA break,DSB)能够明显的增加替换的概率。CRISPR-Cas9技术已经成为强有力的基因组编辑手段,被广泛应用到很多组织和细胞中。CRISPR/Cas9protein-RNA复合物通过向导RNA(guide RNA)定位于靶点上,通过切割使DNA产生DSB,从而增加长链DNA模板介导的基因精确替换效率。在产生DSB后,生物体会本能的启动DNA修复机制。修复机制一般有两种,一种是非同源末端连接(non-homologous end joining,NHEJ),此种修复所占的比例占大多数,修复DNA后一般会产生随机的indels(insertions or deletions)。另一种是同源重组(homology-directed repair,HDR),使用姐妹染色单体或外源DNA供体(donor)作为修复模板,实现基因的精确修复。在动物细胞中,修复的具体原理是:CtIP酶在DSB处起始末端切割,产生突出的3'端单链DNA(single-stranded DNA,ssDNA)尾巴,ssDNA会被重组酶Rad51识别,结合成复合物,侵入供体DNA模板,并与其同源的片段退火结合,以供体DNA作为模板,合成新的DNA链,完成修复。当供体DNA同源臂之间的序列带有外源突变时,此突变会在修复的过程中引入到DNA链中,从而实现精确定点替换。此种HDR修复起始于DSB的产生,由于NHEJ的修复概率远大于HDR,在发生了精确替换的样品中,会有较多的副产物,比如引入indels,造成DNA大片段缺失、移位等。

[0003] 为了提高产物中精确HDR:非精确HDR的比例,人们尝试使用Cas9的一种失活突变体D10A,对DNA造成单链切刻。在动物中,相比于DSB而言,单链切刻起始的HDR的副产物较少,但同时一定程度上降低了HDR的效率。在植物中,DSB诱发的HDR在不同的基因上能够实现精确替换,但是Cas9D10A引发的切刻能否实现HDR精确替换,其效率是否比DSB诱导的HDR低,副产物是否减少,均无报道。

[0004] 来源于病毒基因组的2A自切割寡肽(self-cleaving peptide 2A,P2A),是一类长18-22个氨基酸的多肽,能够诱发重组蛋白的切割。自切割寡肽有多种类型,如:口蹄疫病毒(FMDV)(F2A)肽、马A型鼻炎病毒(ERAV)(E2A)肽、明脉扁刺蛾β四体病毒(Thosea asigna virus)(TaV)(T2A)肽、猪捷申病毒-1(PTV-1)(P2A)肽、泰勒病毒2A肽以及脑心肌炎病毒2A肽。第一个被发现的P2A序列来源于口蹄疫病毒。当被转录的一条mRNA链上含有P2A时,核糖体在此处发生跳跃事件,将P2A序列前后的两个蛋白分别翻译成两个独立的蛋白。在动物中,使用P2A将目的蛋白与荧光蛋白连接在一起共同表达,能够保证两个蛋白同时空表达,并且能够很好的对目的蛋白进行细胞定位。但是在植物中,使用P2A将目的蛋白与其它标记蛋白偶连在一起报道非常少。

### 发明内容

[0005] 本发明的目的是提供一种植物基因替换的方法。

[0006] 本发明提供的植物基因替换的方法包括如下步骤:将esgRNA、Cas9切刻酶(Cas9n)或其变体、筛选剂抗性蛋白、供体DNA导入目的植物中;

[0007] 所述esgRNA靶向DNA片段甲靶点序列;所述DNA片段甲靶点序列位于目的植物基因组中的转录链上;

[0008] 所述Cas9切刻酶或其变体与所述筛选剂抗性蛋白通过依次由启动子、所述Cas9切刻酶或其变体的编码基因、自切割寡肽的编码基因、所述筛选剂抗性蛋白的编码基因和终止子组成的表达盒或通过依次由启动子、所述筛选剂抗性蛋白的编码基因、自切割寡肽的编码基因、所述Cas9切刻酶或其变体的编码基因和终止子组成的表达盒导入目的植物中;

[0009] 所述供体DNA的转录链依次包括所述DNA片段甲靶点序列、DNA片段乙和所述DNA片段甲靶点序列;

[0010] 所述DNA片段乙为将所述DNA片段甲经过一个或几个碱基突变后得到的DNA分子;

[0011] 在所述esgRNA引导下,所述Cas9切刻酶或其变体在目的植物基因组中的DNA片段甲靶点序列的非转录链上产生一个单链DNA切刻,在所述供体DNA中的DNA片段甲靶点序列的非转录链上产生两个单链DNA切刻,并通过目的植物体内的修复机制将目的植物基因组中的所述DNA片段甲替换为所述DNA片段乙,实现植物基因替换。

[0012] 上述方法中,所述DNA片段甲可为目的植物基因组上的任意片段,将DNA片段甲上的某一个或某几个碱基突变后获得DNA片段乙,所述DNA片段乙为供体DNA上的片段。所述碱基突变可为碱基替换和/或碱基插入和/或碱基缺失。

[0013] 在实际应用中,将esgRNA/Cas9n系统和两端添加有靶点对应的靶点序列的DNA片段乙即供体DNA导入目的植物后,可实现将目的植物基因组中的DNA片段甲替换为DNA片段乙,进而实现基因替换。此处基因替换可实现将基因突变位点引入目的植物基因组中,进而实现目的植物基因组上的基因突变(如碱基替换、碱基插入或碱基缺失),从而使目的植物中表达的相应蛋白质的氨基酸功能位点和/或种类和/或活性和/或含量等发生改变,获得具有某一功能或性状的植物突变体。在本发明的具体实施例中,所述碱基突变具体可为碱基替换。

[0014] 进一步的,所述DNA片段甲和所述DNA片段乙的大小均可为200-2000bp或200-1500bp或200-1000bp。

[0015] 更进一步的,所述DNA片段甲和所述DNA片段乙的大小均为694bp。

[0016] 所述DNA片段甲为序列5第1300-1993位所示的DNA分子。

[0017] 所述DNA片段乙为序列1第8513-9206位所示的DNA分子。

[0018] 在本发明的具体实施例中,所述DNA片段甲依次由大小为636bp的ALS基因片段和其下游大小为58bp的片段组成。所述DNA片段乙为将DNA片段甲第344位由碱基G突变为碱基T,且将第581位由碱基G突变为碱基T,且将第336位由碱基G突变为碱基C,且将第339位由碱基G突变为碱基C,且将第342位由碱基A突变为碱基G,且将第396位由碱基G突变为碱基C后得到的DNA分子。将水稻基因组中的DNA片段甲替换为DNA片段乙后,由于水稻基因组中的DNA片段甲第344位由碱基G突变为碱基T,且第581位由碱基G突变为碱基T,水稻中表达的ALS蛋白氨基酸序列第548位的氨基酸由酪氨酸(Try)突变为亮氨酸(Leu),第627位的氨基酸由丝氨酸(Ser)突变为异亮氨酸(Ile),从而产生具有除草剂抗性的精确编辑植株。

[0019] 上述方法中,所述esgRNA结构如下:tRNA-所述DNA片段甲靶点序列转录的RNA-

esgRNA骨架；

[0020] 所述tRNA为a1)或a2)或a3)：

[0021] a1)将序列1第474-550位中的T替换为U得到的RNA分子；

[0022] a2)将a1)所示的RNA分子经过一个或几个核苷酸的取代和/或缺失和/或添加且具有相同功能的RNA分子；

[0023] a3)与a1)或a2)限定的核苷酸序列具有75%或75%以上同一性且具有相同功能的RNA分子；

[0024] 所述esgRNA骨架为b1)或b2)或b3)：

[0025] b1)将序列1第571-656位中的T替换为U得到的RNA分子；

[0026] b2)将b1)所示的RNA分子经过一个或几个核苷酸的取代和/或缺失和/或添加且具有相同功能的RNA分子；

[0027] b3)与b1)或b2)限定的核苷酸序列具有75%或75%以上同一性且具有相同功能的RNA分子。

[0028] 上述方法中,所述供体DNA的转录链依次由所述DNA片段甲靶点序列、PAM序列、DNA片段乙和所述DNA片段甲靶点序列和PAM序列组成。所述DNA片段甲靶点序列为序列7或序列12;所述供体DNA的转录链的序列为W548-NTS/dNTS重组表达载体序列第8490-9229位的反向互补序列或P565-NTS/dNTS重组表达载体序列第8490-9229位的反向互补序列。

[0029] 上述方法中,所述Cas9切刻酶可为Cas9D10A切刻酶或Cas9H840A切刻酶；

[0030] 所述Cas9切刻酶变体包括来源于细菌的Cas9(如SaCas9、SaCas9-KKH等),识别不同PAM的Cas9变体(如xCas9、Cas9-NG、Cas9-VQR、Cas9-VRER等),Cas9高保真酶变体(如HypaCas9、eSpCas9(1.1)、Cas9-HF1等)等。

[0031] 进一步的,所述Cas9D10A切刻酶为SpCas9n蛋白质；

[0032] 所述SpCas9n蛋白质为C1)或C2)：

[0033] C1)氨基酸序列是序列3所示的蛋白质；

[0034] C2)将序列表中序列3所示的氨基酸序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加且具有相同功能的蛋白质；

[0035] 更进一步的,所述SpCas9n蛋白质的编码基因为c1)或c2)或c3)：

[0036] c1)序列表中序列1第2887-6987位所示的cDNA分子或DNA分子；

[0037] c2)与c1)限定的核苷酸序列具有75%或75%以上同一性,且编码所述SpCas9n的cDNA分子或DNA分子；

[0038] c3)在严格条件下与c1)或c2)限定的核苷酸序列杂交,且编码所述SpCas9n的cDNA分子或DNA分子。

[0039] 上述方法中,所述Cas9切刻酶或其变体带有核定位信号。所述核定位信号可为BP NLS、VirD2NLS或SV40NLS。所述核定位信号的个数可为1个或2个或多个。

[0040] 进一步的,所述核定位信号为SV40NLS。所述SV40NLS的氨基酸序列为序列2。所述核定位信号的个数为8个。

[0041] 更进一步的,所述SV40NLS的编码序列为序列1第2752-2772位。所述Cas9切刻酶或其变体两端分别带有4个SV40NLS。

[0042] 上述方法中,所述自切割寡肽可为来源于病毒基因组的2A自切割寡肽,如口蹄疫

病毒 (FMDV) (F2A) 肽、马A型鼻炎病毒 (ERAV) (E2A) 肽、明脉扁刺蛾β四体病毒 (Thosea asigna virus) (T2A) 肽、猪捷申病毒-1 (PTV-1) (P2A) 肽、泰勒病毒2A肽以及脑心肌炎病毒2A肽。

[0043] 进一步的,所述自切割寡肽为来源于猪捷申病毒-1 (PTV-1) 的2A自切割寡肽。所述来源于猪捷申病毒-1 (PTV-1) 的2A自切割寡肽的氨基酸序列为序列4。

[0044] 更进一步的,所述来源于猪捷申病毒-1 (PTV-1) 的2A自切割寡肽的编码序列为序列1第7138-7194位。

[0045] 上述方法中,所述筛选剂抗性蛋白为潮霉素磷酸转移酶。

[0046] 进一步的,所述潮霉素磷酸转移酶为D1) 或D2) :

[0047] D1) 氨基酸序列是序列5所示的蛋白质;

[0048] D2) 将序列表中序列5所示的氨基酸序列经过一个或几个氨基酸残基的取代和/或缺失和/或添加且具有相同功能的蛋白质;

[0049] 更进一步的,所述潮霉素磷酸转移酶的编码基因为d1) 或d2) 或d3) :

[0050] d1) 序列表中序列1第7195-8220位所示的cDNA分子或DNA分子;

[0051] d2) 与d1) 限定的核苷酸序列具有75%或75%以上同一性,且编码所述潮霉素磷酸转移酶的cDNA分子或DNA分子;

[0052] d3) 在严格条件下与d1) 或d2) 限定的核苷酸序列杂交,且编码所述潮霉素磷酸转移酶的cDNA分子或DNA分子。

[0053] 上述方法中,所述将esgRNA、Cas9切刻酶或其变体、筛选剂抗性蛋白、供体DNA导入植物中的方法包括如下步骤:将转录esgRNA的DNA分子、Cas9切刻酶或其变体的编码基因、筛选剂抗性蛋白的编码基因和供体DNA导入目的植物中。

[0054] 上述方法中,所述esgRNA为tRNA-esgRNA,转录所述tRNA-esgRNA的DNA分子转录后得到的所述tRNA-esgRNA为不成熟的RNA前体,该RNA前体中的tRNA会被两种酶 (RNase P和RNase Z) 切割掉后得到成熟的RNA。一个重组表达载体中有多少个靶点,就会得到多少个独立的成熟的RNA,每个成熟的RNA依次由所述靶点序列转录的RNA和所述esgRNA骨架组成,或依次由所述靶点序列转录的RNA、所述esgRNA骨架和所述tRNA残留的个别碱基组成。

[0055] 进一步的,所述转录esgRNA的DNA分子、所述Cas9切刻酶或其变体的编码基因、所述筛选剂抗性蛋白的编码基因和所述供体DNA通过重组表达载体导入目的植物中。所述转录esgRNA的DNA分子、所述Cas9切刻酶或其变体的编码基因、所述筛选剂抗性蛋白的编码基因和所述供体DNA可通过同一个重组表达载体导入目的植物中,也可通过两个或者多个重组表达载体共同导入目的植物中。

[0056] 在本发明的具体实施例中,所述转录esgRNA的DNA分子、所述Cas9切刻酶或其变体的编码基因、所述筛选剂抗性蛋白的编码基因和所述供体DNA通过同一个重组表达载体导入目的植物中。所述重组表达载体包括依次由启动子、所述转录esgRNA的DNA分子和终止子组成的表达盒和依次由启动子、所述Cas9切刻酶或其变体的编码基因、自切割寡肽的编码基因、筛选剂抗性蛋白的编码基因和终止子组成的表达盒。

[0057] 更进一步的,所述重组表达载体为W548-NTS/dNTS重组表达载体或P565-NTS/dNTS重组表达载体。

[0058] 所述W548-NTS/dNTS重组表达载体的核苷酸序列为将序列1中的第551-570位替换

为序列7所示的AST339靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列8所示的AST339靶点序列,且保持其他序列不变后得到的序列。

[0059] 所述P565-NTS/dNTS重组表达载体的核苷酸序列为将序列1中的第551-570位替换为序列12所示的AST340靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列13所示的AST340靶点序列,且保持其他序列不变后得到的序列。

[0060] 上述方法在植物基因编辑或制备植物突变体或提高植物基因替换效率或减少植物基因替换产生的副产物中的应用也属于本发明的保护范围。

[0061] 本发明最后还提供了方法一或方法二或方法三或方法四:

[0062] 所述方法一为一种植物基因编辑的方法;所述植物基因编辑的方法包括如下步骤:按照上述方法对植物基因组中的目的基因片段进行替换,从而实现植物基因编辑。所述编辑具体可为碱基替换。

[0063] 所述方法二为一种制备植物突变体的方法;所述制备植物突变体的方法包括如下步骤:按照上述方法对植物基因组中的目的基因片段进行替换,获得植物突变体。所述植物突变体具体可为抗除草剂突变体。

[0064] 所述方法三为一种提高植物基因替换效率的方法;所述提高植物基因替换效率的方法包括如下步骤:按照上述方法对植物基因组中的目的基因片段进行替换。所述替换效率具体可为HDR替换效率。

[0065] 所述方法四为一种减少植物基因替换产生的副产物的方法;所述减少植物基因替换产生的副产物的方法包括如下步骤:按照上述方法对植物基因组中的目的基因片段进行替换。所述减少植物基因替换产生的副产物具体体现在按照上述方法对植物基因组中的目的基因片段进行替换得到的产物没有额外的Indels产生。

[0066] 上述方法或应用中,所述植物为如下m1)-m3)中的任一种:

[0067] m1) 单子叶植物或双子叶植物;

[0068] m2) 禾本科植物;

[0069] m3) 水稻(如日本晴)。

[0070] 本发明提供的植物基因替换方法的原理如下:Cas9D10A切刻酶(Cas9n)的编码基因通过P2A与筛选标记抗性基因潮霉素磷酸转移酶基因(Hpt)偶连在一起,在同一个启动子的驱动下转录,然后翻译成独立的蛋白。在靶点的引导下,Cas9n/esgRNA复合物在基因组ALS靶位点引发单链DNA切刻。当靶点对应的靶序列位于基因组中的非转录链(non-transcribed strand,NTS)时,Cas9n/esgRNA复合物会在基因组的转录链(transcribed strand,TS)上产生切刻;当靶点对应的靶序列位于基因组中的TS时,Cas9n/esgRNA复合物会在基因组的NTS上产生切刻。载体上的供体DNA含有具有除草剂抗性的突变位点,同时供体DNA 5'端和3'端各含有1个靶点对应的靶序列。当此靶序列同时位于DNA双链中的NTS,此时会在TS上产生2个切刻;当此靶序列同时位于DNA双链中的TS,此时会在NTS上产生2个切刻。在水稻内源ALS基因组的不同位点、不同DNA链上形成切刻,在供体DNA上不同DNA链形成切刻,发现组合二(基因组-TS,donor-NTS)比组合一(基因组-TS,donor-TS)的精确替换效率高,组合三(基因组-NTS,donor-NTS)与组合二的精确替换效率高,从而能够提高获得除草剂抗性的精确编辑植株的效率。

[0071] 本发明具有以下优点:



[0072] 1、效率高: T<sub>0</sub>苗中, 组合二(基因组-TS, donor-NTS)的精确替换效率是组合一(基因组-TS, donor-TS)的1.4-1.7倍, 组合三(基因组-NTS, donor-NTS)的精确替换效率是组合一的1.9-2.7倍。

[0073] 2、副产物少: 所有组合产生的精确替换植株, 均不含有随机的Indels。

[0074] 3、成本低: 此方法仅需要载体中含有Cas9n相关元件, 以及对应的供体DNA, 通过农杆菌侵染的方法即可实现精确替换。

[0075] 本发明提供了一种植物基因替换的方法, 该方法不产生DNA双链断裂。通过实验证明: 本发明提供的植物基因替换的方法在水稻中实现了对内源的乙酰乳酸合酶ALS基因的精确替换, 获得具有除草剂抗性的精确编辑植株。

## 附图说明

[0076] 图1为精确替换载体构建示意图。

[0077] 图2为针对W548位点的不同组合切刻方式在基因组和donor上的切刻分布示意图。

[0078] 图3为针对P565位点的不同组合切刻方式在基因组和donor上的切刻分布示意图。

[0079] 图4为针对W548位点的不同组合且刻方式获得的转基因植株的特异引物扩增检测结果。

[0080] 图5为针对P565位点的不同组合且刻方式获得的转基因植株的特异引物扩增检测结果。

## 具体实施方式

[0081] 以下的实施例便于更好地理解本发明, 但并不限定本发明。

[0082] 下述实施例中的实验方法, 如无特殊说明, 均为常规方法。下述实施例中所用的试验材料, 如无特殊说明, 均为自常规生化试剂商店购买得到的。下述实施例中, 如无特殊说明, 序列表中各核苷酸序列的第1位均为相应DNA/RNA的5'末端核苷酸, 末位均为相应DNA/RNA的3'末端核苷酸。

[0083] 引物对P1由引物HDR-F: 5'-gcgcccgattctctatgtc-3'和引物HDR-R: 5'-acctatcctccaactggacg-3'组成, 用于检测植株是否发生精确替换。

[0084] 引物对P2由引物gALS-F: 5'-atcccagttacaaccactctg-3'和引物gALS-R: 5'-cacttaactcagagctattgcatag-3'组成, 用于扩增基因组ALS序列并测序。

[0085] 转录链是指在基因组DNA中, 被RNA聚合酶II用来做模板, 起始转录过程的DNA链, 转录出来的mRNA与模板链互补, 与DNA的另外一条链(非转录链)序列一致。下述实施例中提供的载体序列均为非转录链序列。

[0086] HDR苗是指在获得的T<sub>0</sub>苗中, 含有donor带入的对应突变位点的植株。

[0087] T<sub>0</sub>苗HDR替换效率 = T<sub>0</sub>苗中发生了精确替换的苗数/获得的T<sub>0</sub>苗总数

[0088] 基于愈伤数量的HDR替换效率 = T<sub>0</sub>苗中发生了精确替换的苗数/起始侵染的愈伤总数

[0089] HDR中的indels效率 = 精确替换的T<sub>0</sub>苗中发生了indels的苗数/精确替换的T<sub>0</sub>苗总数

[0090] 日本晴水稻: 参考文献: 梁卫红, 王高华, 杜京尧, 等. 硝普钠及其光解产物对日本

晴水稻幼苗生长和5种激素标记基因表达的影响[J].河南师范大学学报(自然版),2017(2):48-52.;公众可以从北京市农林科学院获得。

[0091] 恢复培养基:含有200mg/L特美汀的N6固体培养基。

[0092] 筛选培养基1:含有50mg/L潮霉素的N6固体培养基。

[0093] 筛选培养基2:含有0.4uM/L双草醚的N6固体培养基。

[0094] 分化培养基:含有2mg/L KT、0.2mg/L NAA、0.5g/L谷氨酸、0.5g/L脯氨酸的N6固体培养基。

[0095] 生根培养基:含有0.2mg/L NAA、0.5g/L谷氨酸、0.5g/L脯氨酸、0.28uM/L双草醚的N6固体培养基。

[0096] 下述实施例中ALS蛋白的氨基酸序列如序列表中序列15所示,其编码基因序列如序列表中序列14第1-1935位所示。

[0097] 实施例1、TS/dTS、TS/dNTS和NTS/dNTS对应载体的构建及其在水稻基因替换中的应用

[0098] 一、重组表达载体的构建及替换原理说明

[0099] 1、重组表达载体的构建

[0100] 人工合成如下重组表达载体,各载体均为环状质粒:

[0101] 对于W548位点,含有三个载体,分别为W548-TS/dTS,W548-TS/dNTS,W548-NTS/dNTS。

[0102] 对于P565位点,含有三个载体,分别为P565-TS/dTS,P565-TS/dNTS,P565-NTS/dNTS。

[0103] 重组表达载体骨架结构示意图如图1所示。具体结构描述如下:

[0104] W548-TS/dTS重组表达载体的序列为序列表中的序列1。序列1的第131-467位为OsU3启动子序列,第474-550位为tRNA序列,第551-570位为ST215靶点序列,第571-656位为esgRNA骨架序列,第657-947位为OsU3终止子序列。序列1的第954-2667位为OsUbq3启动子序列,第2752-2772位、第2782-2802位、第2812-2832位、第2842-2862位、第7006-7026位、第7036-7056位、第7066-7086位、第7096-7116位均为SV40核定位序列(编码序列2所示的核定位信号SV40),第2887-6987位为SpCas9n蛋白质的编码序列(编码序列3所示的SpCas9n蛋白质),第7138-7194位为自切割寡肽P2A的编码序列(编码序列4所示的自切割寡肽P2A),第7195-8220位为潮霉素磷酸转移酶的核苷酸序列(编码序列5所示的潮霉素磷酸转移酶),第8227-8481位为Nos终止子的核苷酸序列。第8490-8512位、第9207-9229位均为ST215靶点靶序列(由ST215靶点序列和PAM序列组成);第8513-9206位为ALS供体DNA序列。

[0105] W548-TS/dNTS重组表达载体的序列为将序列1中的第8490-8512位、第9207-9229位均替换为序列6所示的AST215靶点靶序列,且保持其他序列不变后得到的序列。

[0106] W548-NTS/dNTS重组表达载体的序列为将序列1中的第551-570位替换为序列7所示的AST339靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列8所示的AST339靶点靶序列,且保持其他序列不变后得到的序列。

[0107] P565-TS/dTS重组表达载体的序列为将序列1中的第551-570位替换为序列9所示的ST319靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列10所示的ST319靶点靶序列,且保持其他序列不变后得到的序列。

[0108] P565-TS/dNTS重组表达载体的序列为将序列1中的第551-570位替换为序列9所示的ST319靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列11所示的AST319靶点靶序列,且保持其他序列不变后得到的序列。

[0109] P565-NTS/dNTS重组表达载体的序列为将序列1中的第551-570位替换为序列12所示的AST340靶点序列,且将序列1中的第8490-8512位、第9207-9229位均替换为序列13所示的AST340靶点靶序列,且保持其他序列不变后得到的序列。

[0110] 2、重组表达载体的精确替换原理

[0111] 1) W548位点重组表达载体的精确替换原理

[0112] W548位点重组表达载体是基于单链切刻(Nick)引导的精确替换,原理示意图如图2所示。

[0113] W548位点重组表达载体包括如下元件:esgRNA、Cas9n、供体DNA。

[0114] esgRNA靶向靶点靶序列。靶点为ST215或AST339。

[0115] 供体DNA(donor DNA):供体DNA依次由靶点靶序列、ALS供体DNA序列和靶点靶序列组成。

[0116] ALS供体DNA序列(序列1第8513-9206位)是将DNA片段甲(DNA片段甲为序列14第1300-1993位所示的大小为694bp的片段,其依次由大小为636bp的ALS基因片段和其下游大小为58bp的片段组成。DNA片段甲为水稻基因组上的片段。)进行突变后得到的DNA分子。突变包括功能位点突变和同义位点突变。

[0117] 功能位点突变:将DNA片段甲第344位由碱基G突变为碱基T(此处碱基突变可使水稻中的ALS蛋白氨基酸序列第548位的酪氨酸(Try)突变为亮氨酸(Leu)),且将第581位由碱基G突变为碱基T(此处碱基突变可使水稻中的ALS蛋白氨基酸序列第627位的丝氨酸(Ser)突变为异亮氨酸(Ile))。水稻中的ALS蛋白氨基酸序列第548位的酪氨酸(Try)突变为亮氨酸(Leu),且第627位的丝氨酸(Ser)突变为异亮氨酸(Ile)后能够抗双草醚除草剂,将水稻ALS蛋白氨基酸序列第548位记作W548L功能突变位点,将水稻ALS蛋白氨基酸序列第627位记作S627I功能突变位点。

[0118] 同义位点突变:为了方便后期设计特异检测引物检测精确替换突变体,将DNA片段甲第336位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第545位),且将第339位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第546位),且将第342位由碱基A突变为碱基G(此处碱基突变对应水稻ALS蛋白氨基酸序列第547位),且将第396位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第565位)。此处碱基突变不改变水稻ALS蛋白氨基酸序列上的相应氨基酸位点所对应的氨基酸,将水稻ALS蛋白氨基酸序列第545位记作545同义突变位点,将水稻ALS蛋白氨基酸序列第546位记作546同义突变位点,将水稻ALS蛋白氨基酸序列第547位记作547同义突变位点,将水稻ALS蛋白氨基酸序列第565位记作P565同义突变位点。

[0119] 基因组靶点选择说明:ST215靶点位于水稻ALS基因上的W548位点对应的密码子附近的非转录链上,AST339靶点位于水稻ALS基因上的W548位点对应的密码子附近的转录链上。

[0120] Cas9n/esgRNA复合物在相应靶点的引导下,在载体中供体donor上的靶序列产生两个单链切刻位点,同时在水稻基因组ALS基因序列上的靶序列产生一个单链切刻位点,在

水稻体内的修复机制下,供体DNA在水稻基因组上的切刻位点处发生精确替换(水稻基因组上的DNA片段甲替换为ALS供体DNA),使ALS供体DNA上的突变位点引入到水稻基因组中,获得基因替换后的植株,该基因替换后的植株具有除草剂抗性。

[0121] W548-TS/dTS重组表达载体中,ST215靶点对应的靶序列位于水稻基因组ALS基因的非转录链上和供体DNA的非转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的转录链上和供体DNA的转录链上产生切刻。

[0122] W548-TS/dNTS重组表达载体中,ST215靶点对应的靶序列位于水稻基因组ALS基因的非转录链上和供体DNA的转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的转录链上和供体DNA的非转录链上产生切刻。

[0123] W548-NTS/dNTS重组表达载体中,AST339靶点对应的靶序列位于水稻基因组ALS基因的转录链上和供体DNA的转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的非转录链上和供体DNA的非转录链上产生切刻。

[0124] 2)P565位点重组表达载体的精确替换原理

[0125] P565位点重组表达载体是基于单链切刻Nick引导的精确替换,原理示意图如图3所示。

[0126] P565位点重组表达载体包括如下元件:esgRNA、Cas9n、供体DNA。

[0127] esgRNA靶向靶点靶序列。靶点为ST319或AST340。

[0128] 供体DNA(donor DNA):供体DNA依次由靶点靶序列、ALS供体DNA序列和靶点靶序列组成。

[0129] ALS供体DNA序列(序列1第8513-9206位)是将DNA片段甲(DNA片段甲为序列5第1300-1993位所示的大小为694bp的片段,其依次由大小为636bp的ALS基因片段和其下游大小为58bp的片段组成。DNA片段甲为水稻基因组上的片段)进行突变后得到的DNA分子。突变包括功能位点突变和同义位点突变。

[0130] 功能位点突变:将DNA片段甲第344位由碱基G突变为碱基T(此处碱基突变可使水稻中的ALS蛋白氨基酸序列第548位的酪氨酸(Try)突变为亮氨酸(Leu)),且将第581位由碱基G突变为碱基T(此处碱基突变可使水稻中的ALS蛋白氨基酸序列第627位的丝氨酸(Ser)突变为异亮氨酸(Ile))。水稻中的ALS蛋白氨基酸序列第548位的酪氨酸(Try)突变为亮氨酸(Leu),且第627位的丝氨酸(Ser)突变为异亮氨酸(Ile)后能够抗双草醚除草剂,将水稻ALS蛋白氨基酸序列第548位记作W548L功能突变位点,将水稻ALS蛋白氨基酸序列第627位记作S627I功能突变位点。

[0131] 同义位点突变:为了方便后期设计特异检测引物检测精确替换突变体,将DNA片段甲第336位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第545位),且将第339位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第546位),且将第342位由碱基A突变为碱基G(此处碱基突变对应水稻ALS蛋白氨基酸序列第547位),且将第396位由碱基G突变为碱基C(此处碱基突变对应水稻ALS蛋白氨基酸序列第565位)。此处碱基突变不改变水稻ALS蛋白氨基酸序列上的相应氨基酸位点所对应的氨基酸,将水稻ALS蛋白氨基酸序列第545位记作545同义突变位点,将水稻ALS蛋白氨基酸序列第546位记作546同义突变位点,将水稻ALS蛋白氨基酸序列第547位记作547同义突变位点,将水稻ALS蛋白氨基酸序列第565位记作P565同义突变位点。

[0132] 基因组靶点选择说明:ST319靶点位于水稻ALS基因上的P565位点对应的密码子附近的非转录链上,AST340靶点位于水稻ALS基因上的P565位点对应的密码子附近的转录链上。

[0133] Cas9n/esgRNA复合物在相应靶点的引导下,在载体中供体donor上的靶序列产生两个单链切刻位点,同时在水稻基因组ALS基因序列上的靶序列产生一个单链切刻位点,在水稻体内的修复机制下,供体DNA在水稻基因组上的切刻位点处发生精确替换(水稻基因组上的DNA片段甲替换为ALS供体DNA),使ALS供体DNA上的突变位点引入到水稻基因组中,获得基因替换后的植株,该基因替换后的植株具有除草剂抗性。

[0134] P565-TS/dTS重组表达载体中,ST319靶点对应的靶序列位于水稻基因组ALS基因的非转录链上和供体DNA的非转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的转录链上和供体DNA的转录链上产生切刻。

[0135] P565-TS/dNTS重组表达载体中,ST319靶点对应的靶序列位于水稻基因组ALS基因的非转录链上和供体DNA的转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的转录链上和供体DNA的非转录链上产生切刻。

[0136] P565-NTS/dNTS重组表达载体中,AST340靶点对应的靶序列位于水稻基因组ALS基因的转录链上和供体DNA的转录链上,Cas9n/esgRNA复合物会在水稻基因组ALS基因的非转录链上和供体DNA的非转录链上产生切刻。

[0137] 二、水稻阳性抗性愈伤的获得

[0138] 将步骤一获得的W548-TS/dTS、W548-TS/dNTS、W548-NTS/dNTS、P565-TS/dTS、P565-TS/dNTS、P565-NTS/dNTS重组表达载体分别按照如下步骤1-7进行操作:

[0139] 1、将载体导入农杆菌EHA105(上海唯地生物技术有限公司的产品,CAT#:AC1010),得到重组农杆菌。

[0140] 2、采用培养基(含50 $\mu$ g/ml卡那霉素和25 $\mu$ g/ml利福平的YEP培养基)培养重组农杆菌,28 $^{\circ}$ C,150rpm震荡培养至OD<sub>600</sub>为1.0-2.0,室温条件下,10000rpm离心1min,用侵染液(将N6液体培养基中的糖替换为葡萄糖和蔗糖,葡萄糖和蔗糖在侵染液中的浓度分别为10g/L和20g/L)重悬菌体并稀释至OD<sub>600</sub>为0.2,得到农杆菌侵染液。

[0141] 3、水稻品种日本晴成熟种子去壳脱粒,置于100mL三角瓶中,加入70% (v/v) 乙醇水溶液浸泡30sec,再置于25% (v/v) 次氯酸钠水溶液中,120rpm震荡灭菌30min,无菌水冲洗3次,用滤纸吸干水分,然后将种子胚朝下置于N6固体培养基上,28 $^{\circ}$ C暗培养4-6周,得到水稻愈伤。

[0142] 4、完成步骤3后,将水稻愈伤浸泡置于农杆菌侵染液甲(农杆菌侵染液甲为向农杆菌侵染液中加入乙酰丁香酮得到的液体,乙酰丁香酮的添加量满足乙酰丁香酮与农杆菌侵染液的体积比为25 $\mu$ l:50ml)中浸泡10min,然后,放在铺有两层灭菌滤纸的培养皿(内含约200ml不含农杆菌的侵染液)上,21 $^{\circ}$ C暗培养1天。

[0143] 5、取步骤4得到的水稻愈伤放入恢复培养基上,25-28 $^{\circ}$ C暗培养3天。

[0144] 6、取步骤5得到的水稻愈伤,置于筛选培养基1上,28 $^{\circ}$ C暗培养2周。

[0145] 7、取步骤6得到的水稻愈伤,转移至筛选培养基2上,28 $^{\circ}$ C暗培养4周,得到水稻抗性愈伤。

[0146] 三、水稻T0苗的获得

[0147] 1、取步骤1得到的水稻抗性愈伤放入分化培养基上,25℃光照培养1个月左右。

[0148] 2、将分化出来的小苗移至生根培养基上,25℃光照培养2周,获取抗除草剂的水稻T0苗。

#### [0149] 四、精确替换植株鉴定

[0150] 1、经过含有除草剂的生根培养基筛选后,存活下来的水稻T0苗,分别提取基因组DNA并以其作为模板,采用引物HDR-F(5'-gcgcccgattctctatgtc-3')和引物HDR-R(5'-acctatcctccaactggacg-3')组成的引物对进行PCR扩增,得到PCR扩增产物;将该PCR扩增产物进行琼脂糖凝胶电泳,然后进行如下判断:如果PCR扩增产物中含有约833bp的DNA片段,则相应的水稻T0苗为发生了精确替换的阳性T0苗;如果PCR扩增产物中不含有约833bp的DNA片段,则相应的水稻T0苗为未发生精确替换的T0苗。

[0151] 2、使用引物gALS-F(5'-atcccgattacaaccactctg-3')和引物gALS-R(5'-cacttaactcagagctattgcatag-3')组成的引物对,对上述步骤1筛选出的精确替换的阳性T0苗的基因组ALS基因序列进行PCR扩增,得到PCR扩增产物;将该PCR扩增产物进行一代测序,分析对应位点是否确实发生了精确替换。

#### [0152] 五、结果分析

[0153] 1、水稻T0苗中引物初筛获得的精确替换植株

[0154] W548-TS/dTS载体,共获得55棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,24棵苗存活。经过HDR-F和HDR-R引物筛选后,有12棵转基因苗为精确替换植株,PCR检测结果如图4所示。

[0155] W548-TS/dNTS载体,共获得40棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,24棵苗存活。经过HDR-F和HDR-R引物筛选后,有15棵转基因苗为精确替换植株,PCR检测结果如图4所示。

[0156] W548-NTS/dNTS载体,共获得38棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,24棵苗存活。经过HDR-F和HDR-R引物筛选后,有23棵转基因苗为精确替换植株,PCR检测结果如图4所示。

[0157] P565-TS/dTS载体,共获得26棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,16棵苗存活。经过HDR-F和HDR-R引物筛选后,有9棵转基因苗为精确替换植株,PCR检测结果如图5所示。

[0158] P565-TS/dNTS载体,共获得32棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,16棵苗存活。经过HDR-F和HDR-R引物筛选后,有16棵转基因苗为精确替换植株,PCR检测结果如图5所示。

[0159] P565-NTS/dNTS载体,共获得40棵转基因阳性苗(独立转化事件),经过含有除草剂的生根培养基筛选后,32棵苗存活。经过HDR-F和HDR-R引物筛选后,有27棵转基因苗为精确替换植株,PCR检测结果如图5所示。

[0160] 2、水稻T0苗中测序确认的精确替换植株及对应效率

[0161] 一代测序结果显示,PCR检测结果为阳性的植株均在对应位点上发生了精确替换,且无额外的Indels碱基发生。替换效率统计结果具体如表1所示,其中W548-TS/dTS、W548-TS/dNTS、W548-NTS/dNTS载体在T0苗中发生精确替换的概率分别是21.8%(12/55)、37.5%(15/40)、60.5%(23/38);P565-TS/dTS、P565-TS/dNTS、P565-NTS/dNTS载体在T0苗中发生

精确替换的概率分别是34.6% (9/26)、50% (16/32)、67.5% (27/40)。若从整体转化效率来看,以起始侵染的840块抗性愈伤为分母计算,W548-TS/dTS、W548-TS/dNTS、W548-NTS/dNTS载体在发生精确替换的概率分别是1.4% (12/840)、1.8% (15/840)、2.7% (23/840)。P565-TS/dTS、P565-TS/dNTS、P565-NTS/dNTS载体发生精确替换的概率分别是1.1% (9/840)、1.9% (16/840)、3.2% (27/840)。

[0162] 综上所述,不论是W548位点还是P565位点,TS/dNTS方案在T0苗中的精确替换效率是TS/dTS方案的1.4-1.7倍,NTS/dNTS在T0苗中的精确替换效率是TS/dTS方案的1.9-2.7倍。说明TS/dNTS方案、NTS/dNTS方案对于水稻内源的精确替换效率均优于TS/dTS方案。

[0163] 表1

位点	编辑方案		侵染愈伤数	T0苗总数	BS阳性苗数	精确替换苗数	T0苗精确替换效率(%)	以愈伤愈基础的替换效率(%)
	基丕组链	Donor链						
W548	TS	dTS	840	55	24	12	21.8	1.4
	TS	dNTS	840	40	24	15	37.5	1.8
	NTS	dNTS	840	38	24	23	60.5	2.7
P565	TS	dTS	840	26	16	9	34.6	1.1
	TS	dNTS	840	32	16	16	50.0	1.9
	NTS	dNTS	840	40	32	27	67.5	3.2

[0165] 以上所述仅是本发明的优选实施方式,应当指出,对于本技术领域的普通技术人员来说,在不脱离本发明技术原理的前提下,还可以做出若干改进和润饰,这些改进和润饰也应视为本发明的保护范围。

[0001]	序列表	
[0002]	<110>北京市农林科学院	
[0003]	<120>NTS/dNTS组合在制备植物突变体中的应用	
[0004]	<160>15	
[0005]	<170>PatentIn version 3.5	
[0006]	<210>1	
[0007]	<211>9404	
[0008]	<212>DNA	
[0009]	<213>Artificial Sequence	
[0010]	<400>1	
[0011]	ggtggcagga tatattgtgg tgtaaactg gcactagcct caccgtcttc gcagacgagg	60
[0012]	ccgctaagtc gcagctacgc tctcaacggc actgactagg tagtttaaac gtgcacttaa	120
[0013]	ttaaggtacc gaagcaactt aaagttatca ggcattgatg gatcttggag gaatcagatg	180
[0014]	tgcagtcagg gaccatagca caagacaggc gtcttctact ggtgctacca gcaaatgctg	240
[0015]	gaagccggga aactcgggta cgttggaac cacgtgatgt gaagaagtaa gataaactgt	300
[0016]	aggagaaaag catttcgtag tgggcatga agccttcag gacatgtatt gcagtatggg	360
[0017]	ccggcccatt acgcaattgg acgacaaca agactagtat tagtaccacc tcggctatcc	420
[0018]	acatagatca aagctgattt aaaagagttg tgcagatgat ccgtggcgga tccaacaag	480
[0019]	caccagtggc ctagtggtag aatagtacc tgccacggta cagaccggg ttcgattccc	540
[0020]	ggctggtgca atttgggtat ggtggtgcaa gtttcagagc tatgctggaa acagcatagc	600
[0021]	aagttgaaat aaggetagtc cgttatcaac ttgaaaaagt ggcaccgagt cgggtgctttt	660
[0022]	tttttcggt ttgcattgag ttttctccgt cgcattgttg cagttttatt ttccgttttg	720
[0023]	cattgaaatt tctccgtctc atgtttgcag cgtgttcaaa aagtacgcag ctgtatttca	780
[0024]	cttatttacg gcgccacatt ttcatgccgt ttgtgccaac tatcccgagc tagtgaatac	840
[0025]	agcttggtct cacacaacac tggtagcccg ctgacctgct cgtacctcgt accgtcgtac	900
[0026]	ggcacagcat ttggaattaa aggggtgatg cgatactgct tgctgctaag cttacaaatt	960
[0027]	cgggtcaagg cggaagccag cgcgccacc cacgtcagca aatacggagg cgcggggttg	1020
[0028]	acggcgtcac ccggtcctaa cggcgacca caaaccagcc agaagaaatt acagtaaaaa	1080
[0029]	aaaagtaaat tgcacttga tccacctttt attacctaag tctcaatttg gatcaccctt	1140
[0030]	aaacctatct tttcaatttg ggccgggttg tggtttgac taccatgaac aacttttcgt	1200
[0031]	catgtctaac ttcccttca gcaacatat gaaccatata tagaggagat cggccgtata	1260
[0032]	ctagagctga tgtgtttaag gtcgttgatt gcacgagaaa aaaaaatcca aatcgcaaca	1320
[0033]	atagcaaat tatctggttc aaagtgtgaa gatatgttta aaggtagtcc aaagtaaac	1380
[0034]	ttatagataa taaatgtgg tccaaagcgt aattcactca aaaaaatca acgagacgtg	1440
[0035]	taccaaaccg agacaaaccg catcttctcg aaatttccca accgctcgtc cggccgctc	1500
[0036]	gtcttcccgg aaaccgcggt ggtttcagcg tggcggattc tccaagcaga cggagacgtc	1560
[0037]	acggcacggg actcctccca ccaccaacc gccataaata ccagccccct catctcctct	1620
[0038]	cctcgcatac gctccacccc cgaaaaattt ctccccaatc tcgagaggtc ctcgctcgtc	1680
[0039]	aatcgaatcc tctcgcgtcc tcaaggtacg ctgcttctcc tctcctcgtc tcgtttcgat	1740
[0040]	tcgatttcgg acgggtgagg ttgtttgtt gctagatccg attggtggtt agggttgtcg	1800
[0041]	atgtgattat cgtgagatgt ttaggggttg tagatctgat ggttgtgatt tgggcacggt	1860



[0042]	tggttcgata ggtggaatcg tggtaggtt ttgggattgg atgttggttc tgatgattgg	1920
[0043]	ggggaatttt tacggttaga tgaattgttg gatgattcga ttggggaaat cgggttagat	1980
[0044]	ctgttgggga attgtggaac tagtcatgcc tgagtattg gtgcgattg tagcgtgttc	2040
[0045]	catctttag gccttgttc gagcatgtc agatctactg ttccgctctt gattgagtta	2100
[0046]	ttggtccat gggttgtgc aaacacaggc tttaatatgt tataatctgt ttgtgtttga	2160
[0047]	tgtagatctg taggtagtt cttcttagac atggttcaat tatgtagctt gtgcgtttcg	2220
[0048]	atgtgatttc atatgtcac agattagata atgatgaact cttttaatta attgtcaatg	2280
[0049]	gtaaatagga agtcttgcg ctatatctgt cataatgatc tcatgttact atctgccagt	2340
[0050]	aatttatgct aagaactata ttagaataac atgttacaat ctgtagtaat atcatgttac	2400
[0051]	aatctgtagt tcatctatat aatctattgt ggtaatttct ttttactatc tgtgtgaaga	2460
[0052]	ttattgccac tagttcatc tacttattc tgaagttcag gatacgtgtg ctgttactac	2520
[0053]	ctatctgaat acatgtgtga tgtgcctgtt actatctttt tgaatacatg tatgttctgt	2580
[0054]	tggaatatgt ttgctgtttg atccgtgtt gtgtccttaa tcttgtgcta gttcttacc	2640
[0055]	tatctgtttg gtgattattt cttgcagtac gtaagcatgg actacaagga ccacgacggg	2700
[0056]	gattacaaag accacgacat agactacaag gatgacgatg acaaatggc accgaagaa	2760
[0057]	aaaaggaagg tcggcggctc cccgaagaaa aaaaggaagg tcggcggctc cccgaagaaa	2820
[0058]	aaaaggaagg tcggcggctc cccgaagaaa aaaaggaagg tcggaatcca tggcgttcca	2880
[0059]	gctgccgaca agaagtactc catcggctc gccatcggca ccaacagcgt cggctggcg	2940
[0060]	gtgatcaccg acgagtacaa ggtcccgtcc aagaagtca aggtcctggg caacaccgac	3000
[0061]	cgccactcca tcaagaagaa cctcatcggc gccctcctct tcgactccgg cgagacggcg	3060
[0062]	gaggcagccc gcctcaaggc caccgcccgc cgccgtaca cccgccgcaa gaaccgcatc	3120
[0063]	tgctacctcc aggagatctt ctccaacgag atggcgaagg tcgacgactc cttcttccac	3180
[0064]	cgctcgagg agtcttctc cgtggaggag gacaagaagc acgagcgcca ccccatcttc	3240
[0065]	ggcaacatcg tcgacgaggt cgcctaccac gagaagtacc ccaactatca ccaccttctg	3300
[0066]	aagaagcttg ttgactctac tgataaggct gatcttctc tcacttacct tgctctcctg	3360
[0067]	cacatgatca agttccgtgg tcaattcctt atcgagggtg accttaacce tgataactcc	3420
[0068]	gacgtggaca agctcttcat ccagctcgtc cagacctaca accagctctt cgaggagaac	3480
[0069]	cctatcaacg cttccggtgt cgacgtaag gcgatcctt ccgctaggct ctccaagtcc	3540
[0070]	aggcgtctcg agaacctcat cgcccagctc cctggtgaga agaagaacgg tcttttcggt	3600
[0071]	aacctcatcg ctctctccct cggctctgacc cctaactca agtccaactt cgacctcctg	3660
[0072]	gaggacgcta agcttcatg ctccaaggat acctacgac atgatctcga caacctctc	3720
[0073]	gctcagattg gagatcagta cgctgatctc ttccttctg ctaagaacct ctccgatctg	3780
[0074]	atctccttt cggatctct taggttaac actgagatca ctaaggctcc tctttctgct	3840
[0075]	tccatgatca agcgtacga cgagcaccac caggacctca cctctctcaa ggctcttgtt	3900
[0076]	cgtcagcagc tccccgagaa gtacaaggag atcttcttcg accagtcca gaacggctac	3960
[0077]	gccggttaca ttgacggtg agctagccag gaggagtctt acaagtctat caagccaatc	4020
[0078]	cttgagaaga tggatgttac tgaggactt ctcgttaagc ttaaccgtga ggacctctt	4080
[0079]	aggaagcaga ggacttctga taacgctct atcctcacc agatccacct tggtagctt	4140
[0080]	cacgccatcc ttcgtaggca ggaggactt tacccttcc tcaaggacaa ccgtgagaag	4200
[0081]	atcgagaaga tccttacttt ccgtattcct tactacgttg gtcctcttgc tcgtggtaac	4260
[0082]	tcccgttctg cttggatgac taggaagtcc gaggagacta tcacccttg gaacttcgag	4320
[0083]	gaggtgttg acaagggtgc ttccgccag tcttcatcg agcgcgatgac caacttcgac	4380

[0084] aagaacctcc ccaacgagaa ggtcctcccc aagcaactccc tcctctacga gtacttcacg 4440  
[0085] gtctacaacg agctcaccaa ggtcaagtac gtcaccgagg gtatgcgcaa gcctgccttc 4500  
[0086] ctctccggcg agcagaagaa ggctatcggt gacctctct tcaagaccaa ccgcaaggtc 4560  
[0087] accgtcaagc agctcaagga ggactacttc aagaagatcg agtgcttcca ctccgtcgag 4620  
[0088] atcagcggcg ttgaggaccg tttcaacgct tctctcgta cctaccacga tctcctcaag 4680  
[0089] atcatcaagg acaaggactt cctcgacaac gaggagaacg aggacatcct cgaggacatc 4740  
[0090] gtctcactc ttactctctt cgaggatagg gagatgatcg aggagaggt caagacttac 4800  
[0091] gctcatctct tcgatgacaa ggttatgaag cagctcaagc gtcgccgta caccggttgg 4860  
[0092] ggtaggctct cccgcaagct catcaacggt atcagggata agcagagcgg caagactatc 4920  
[0093] ctgacttcc tcaagtctga tggtttcgct aacaggaact tcatgcagct catccacgat 4980  
[0094] gactctctta cttcaagga ggatattcag aaggtcagg tgtccggtca gggcgactct 5040  
[0095] ctccacgagc acattgctaa ctttgcgtgt tcccctgcta tcaagaaggg catccttcag 5100  
[0096] actgttaagg ttgtcgatga gcttgtaag gttatgggtc gtcacaagcc tgagaacatc 5160  
[0097] gtcacgaga tggctcgtga gaaccagact acccagaagg gtcagaagaa ctcgaggag 5220  
[0098] cgcatgaaga ggattgagga gggatcaag gagcttggt ctcagatcct taaggagcac 5280  
[0099] cctgtcgaga acaccagct ccagaacgag aagctctacc tctactacct ccagaacggt 5340  
[0100] aggatattgt acgttgacca ggagctcgac atcaacaggc tttctgacta cgacgtcgac 5400  
[0101] cacattgttc ctgactctt ccttaaggat gactccatcg acaacaaggt cctcacgagg 5460  
[0102] tccgacaaga acaggggtaa gtcggacaac gtcccttccg aggaggttgt caagaagatg 5520  
[0103] aagaactact ggaggcagct tctcaacgct aagctcatta cccagaggaa gttcgacaac 5580  
[0104] ctcacgaagg ctgagagggg tggccttcc gagcttgaca aggttggtt catcaagagg 5640  
[0105] cagcttggtg agacgaggca gattaccaag cacgttgctc agatcctcga ttctaggatg 5700  
[0106] aacaccaagt acgacgagaa cgacaagctc atccgcgagg tcaaggtgat caccctcaag 5760  
[0107] tccaagctcg tctccgactt ccgcaaggac ttccagttct acaaggtccg cgagatcaac 5820  
[0108] aactaccacc acgtcacga tgcttacctt aacgtgtcg ttggtaccgc tcttatcaag 5880  
[0109] aagtacctta agcttgagtc cgagttcgtc tacggtgact acaaggtcta cgacgttcgt 5940  
[0110] aagatgatcg ccaagtccga gcaggagatc ggcaaggcca ccgccaagta cttcttctac 6000  
[0111] tccaacatca tgaacttctt caagaccgag atcacctcg ccaacggcga gatccgcaag 6060  
[0112] cgccctctta tcgagacgaa cggtagact ggtgagatcg tttgggacaa ggtcgcgac 6120  
[0113] ttcgtactg ttcgcaaggt ctttctatg cctcaggtta acatcgtcaa gaagaccgag 6180  
[0114] gtccagaccg gtggctctc caaggagtct atccttcaa agagaaactc ggacaagctc 6240  
[0115] atcgctagga agaaggattg ggaccctaag aagtacggtg gtttcgactc ccctactgtc 6300  
[0116] gcctactccg tcctcgtggt cgccaagggt gagaagggtta agtcgaagaa gctcaagtcc 6360  
[0117] gtcaaggagc tcctcggcat caccatcatg gacgctcct ctttcgagaa gaaccctgatc 6420  
[0118] gacttctcgc agccaagggt ctacaaggag gtcaagaagg acctcatcat caagctcccc 6480  
[0119] aagtactctc ttttcgagct cgagaacggt cgtaagagga tgctggcttc cgctgggtgag 6540  
[0120] ctccagaagg gtaacgagct tgctcttct tccaagtac tgaacttct ctacctgcc 6600  
[0121] tcccactacg agaagctcaa gggttcccct gaggataacg agcagaagca gctcttcgtg 6660  
[0122] gagcagcaca agcactacct cgacgagatc atcgagcaga tctccgagtt ctccaagcgc 6720  
[0123] gtcacctcgc ctgacgctaa cctcgacaag gtcctctccg cctacaacaa gcaccgcgac 6780  
[0124] aagcccatcc gcgagcaggc cgagaacatc atccacctct tcacgctcac gaacctcggc 6840  
[0125] gccctgctg ctttcaagta ctctgacacc accatcgaca ggaagcgtta cacgtccacc 6900

[0126]	aaggaggttc	tcgacgctac	tctcatccac	cagtccatca	ccggtcttta	cgagactcgt	6960
[0127]	atcgaccttt	cccagcttgg	tggtgatgac	gatgacaaaa	tggcaccgaa	gaaaaaaagg	7020
[0128]	aaggtcggcg	gctccccgaa	gaaaaaaagg	aaggtcggcg	gctccccgaa	gaaaaaaagg	7080
[0129]	aaggtcggcg	gctccccgaa	gaaaaaaagg	aaggtcggaa	tccatggcgg	atcaggagcc	7140
[0130]	accaacttct	ccctctcaa	gcaggccggc	gacgtggagg	agaaccggg	cccaatgaaa	7200
[0131]	aagcctgaac	tcaccgcgac	gtctgtcgag	aagtttctga	tcgaaaagtt	cgacagcgtc	7260
[0132]	tccgacctga	tgcagctctc	ggagggcgaa	gaatctctgt	ctttcagctt	cgatgtagga	7320
[0133]	ggcggtgat	atgtctcg	ggtaaatagc	tgcgccgatg	gtttctacaa	agatcgttat	7380
[0134]	gtttatcggc	actttgcatc	ggccgcgctc	ccgattccgg	aagtgcctga	cattggggag	7440
[0135]	tttagcgaga	gcctgacctc	ttgcatctcc	cgccgttcac	agggtgtcac	gttgcaagac	7500
[0136]	ctgctgaaa	ccgaactgcc	cgctgttcta	caaccggctc	cggaggctat	ggatgcgac	7560
[0137]	gctgcggccg	atcttagcca	gacgagcggg	ttcggcccat	tcggaccgca	aggaatcgg	7620
[0138]	caatacacta	catggcgtga	tttcatatgc	gcgattgtct	atccccatgt	gtatcactgg	7680
[0139]	caactgtga	tggacgacac	cgtcagtgcg	tccgtcgcgc	aggctctcga	tgagctgatg	7740
[0140]	ctttggcccg	aggactgccc	cgaagtccgg	cacctcgtgc	acgcggattt	cggtccaac	7800
[0141]	aatgtcctga	cggacaatgg	ccgcataaca	gcggctcattg	actggagcga	ggcgatgttc	7860
[0142]	ggggattccc	aatacgaggt	cgccaacatc	ttcttctgga	ggccgtggtt	ggcttgtatg	7920
[0143]	gagcagcaga	cgcgctactt	cgagcggagg	catccggagc	ttgcaggatc	gccacgactc	7980
[0144]	cgggcgtata	tgctccgcat	tggtcttgac	caactctatc	agagcttgg	tgacggcaat	8040
[0145]	ttcgatgatg	cagcttgggc	gcagggtcga	tgcgacgcaa	tcgtccgatc	cggagccggg	8100
[0146]	actgtcgggc	gtacacaaat	cgcccgcaga	agcgcggccg	tctggaccga	tggtctgtga	8160
[0147]	gaagtactcg	ccgatagtgg	aaaccgacgc	cccagcactc	gtccgagggc	aaagaaatag	8220
[0148]	actagtccc	gatcgttcaa	acatttggca	ataaagtttc	ttaagattga	atcctgttgc	8280
[0149]	cggtcttgcg	atgattatca	tataatttct	gttgaattac	gttaagcatg	taataattaa	8340
[0150]	catgtaatgc	atgacgttat	ttatgaggtg	ggtttttatg	attagagtcc	cgcaattata	8400
[0151]	catttaatac	gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	tatcgcgcgc	8460
[0152]	ggtgtcatct	atgttactag	aggcgcgcca	tttgggtatg	gtggtgcaac	gggaagagat	8520
[0153]	cccaccgcaa	tatgccattc	aggtgtgga	tgagctgacg	aaaggtgagg	caatcatcgc	8580
[0154]	tactggtgtt	gggcagcacc	agatgtgggc	ggcacaatat	tacacctaca	agcggccacg	8640
[0155]	gcagtggctg	tcttcggctg	gtctgggcgc	aatgggattt	gggctgcctg	ctgcagctgg	8700
[0156]	tgtctctgtg	gctaaccag	gtgtcacagt	tgttgatatt	gatggggatg	gtagcttctc	8760
[0157]	catgaacatt	caggagctgg	cattgatccg	cattgagaac	ctccctgtga	aggtgatggt	8820
[0158]	gttgaacaac	caacatttgg	gtatggtcgt	ccagttggag	gataggtttt	acaaggcgaa	8880
[0159]	tagggcgc	acatacttgg	gcaaccccga	atgtgagagc	gagatatatc	cagattttgt	8940
[0160]	gactattgct	aaggggttca	atattcctgc	agtccgtgta	acaagaaga	gtgaagtccg	9000
[0161]	tgccgccatc	aagaagatgc	tcgagactcc	agggccatac	ttgttgata	tcatcgtccc	9060
[0162]	gcaccaggag	catgtgctgc	ctatgatccc	aattgggggc	gcattcaagg	acatgatcct	9120
[0163]	ggatggtgat	ggcaggactg	tgtattaatc	tataatctgt	atgttgcaaa	agcaccagcc	9180
[0164]	cggcctatgt	ttgacctgaa	tgaccattt	gggtatggtg	gtgcaacggc	ctgcaggacg	9240
[0165]	cgtttaatta	agtgcacgcg	gccgcctact	tagtcaagag	cctcgcacgc	gactgtcacg	9300
[0166]	cggccaggat	cgctcgtga	gcctcgcaat	ctgtacctag	tgtttaaact	atcagtgttt	9360
[0167]	gacaggatat	attggcgggt	aaacctaaga	gaaaagagcg	ttta	9404	

[0168] <210>2  
 [0169] <211>7  
 [0170] <212>PRT  
 [0171] <213>Artificial Sequence  
 [0172] <400>2  
 [0173] Pro Lys Lys Lys Arg Lys Val  
 [0174] 1 5  
 [0175] <210>3  
 [0176] <211>1367  
 [0177] <212>PRT  
 [0178] <213>Artificial Sequence  
 [0179] <400>3  
 [0180] Asp Lys Lys Tyr Ser Ile Gly Leu Ala Ile Gly Thr Asn Ser Val Gly  
 [0181] 1 5 10 15  
 [0182] Trp Ala Val Ile Thr Asp Glu Tyr Lys Val Pro Ser Lys Lys Phe Lys  
 [0183] 20 25 30  
 [0184] Val Leu Gly Asn Thr Asp Arg His Ser Ile Lys Lys Asn Leu Ile Gly  
 [0185] 35 40 45  
 [0186] Ala Leu Leu Phe Asp Ser Gly Glu Thr Ala Glu Ala Thr Arg Leu Lys  
 [0187] 50 55 60  
 [0188] Arg Thr Ala Arg Arg Arg Tyr Thr Arg Arg Lys Asn Arg Ile Cys Tyr  
 [0189] 65 70 75 80  
 [0190] Leu Gln Glu Ile Phe Ser Asn Glu Met Ala Lys Val Asp Asp Ser Phe  
 [0191] 85 90 95  
 [0192] Phe His Arg Leu Glu Glu Ser Phe Leu Val Glu Glu Asp Lys Lys His  
 [0193] 100 105 110  
 [0194] Glu Arg His Pro Ile Phe Gly Asn Ile Val Asp Glu Val Ala Tyr His  
 [0195] 115 120 125  
 [0196] Glu Lys Tyr Pro Thr Ile Tyr His Leu Arg Lys Lys Leu Val Asp Ser  
 [0197] 130 135 140  
 [0198] Thr Asp Lys Ala Asp Leu Arg Leu Ile Tyr Leu Ala Leu Ala His Met  
 [0199] 145 150 155 160  
 [0200] Ile Lys Phe Arg Gly His Phe Leu Ile Glu Gly Asp Leu Asn Pro Asp  
 [0201] 165 170 175  
 [0202] Asn Ser Asp Val Asp Lys Leu Phe Ile Gln Leu Val Gln Thr Tyr Asn  
 [0203] 180 185 190  
 [0204] Gln Leu Phe Glu Glu Asn Pro Ile Asn Ala Ser Gly Val Asp Ala Lys  
 [0205] 195 200 205  
 [0206] Ala Ile Leu Ser Ala Arg Leu Ser Lys Ser Arg Arg Leu Glu Asn Leu  
 [0207] 210 215 220  
 [0208] Ile Ala Gln Leu Pro Gly Glu Lys Lys Asn Gly Leu Phe Gly Asn Leu  
 [0209] 225 230 235 240

[0210]	Ile Ala Leu Ser Leu Gly Leu Thr Pro Asn Phe Lys Ser Asn Phe Asp
[0211]	245 250 255
[0212]	Leu Ala Glu Asp Ala Lys Leu Gln Leu Ser Lys Asp Thr Tyr Asp Asp
[0213]	260 265 270
[0214]	Asp Leu Asp Asn Leu Leu Ala Gln Ile Gly Asp Gln Tyr Ala Asp Leu
[0215]	275 280 285
[0216]	Phe Leu Ala Ala Lys Asn Leu Ser Asp Ala Ile Leu Leu Ser Asp Ile
[0217]	290 295 300
[0218]	Leu Arg Val Asn Thr Glu Ile Thr Lys Ala Pro Leu Ser Ala Ser Met
[0219]	305 310 315 320
[0220]	Ile Lys Arg Tyr Asp Glu His His Gln Asp Leu Thr Leu Leu Lys Ala
[0221]	325 330 335
[0222]	Leu Val Arg Gln Gln Leu Pro Glu Lys Tyr Lys Glu Ile Phe Phe Asp
[0223]	340 345 350
[0224]	Gln Ser Lys Asn Gly Tyr Ala Gly Tyr Ile Asp Gly Gly Ala Ser Gln
[0225]	355 360 365
[0226]	Glu Glu Phe Tyr Lys Phe Ile Lys Pro Ile Leu Glu Lys Met Asp Gly
[0227]	370 375 380
[0228]	Thr Glu Glu Leu Leu Val Lys Leu Asn Arg Glu Asp Leu Leu Arg Lys
[0229]	385 390 395 400
[0230]	Gln Arg Thr Phe Asp Asn Gly Ser Ile Pro His Gln Ile His Leu Gly
[0231]	405 410 415
[0232]	Glu Leu His Ala Ile Leu Arg Arg Gln Glu Asp Phe Tyr Pro Phe Leu
[0233]	420 425 430
[0234]	Lys Asp Asn Arg Glu Lys Ile Glu Lys Ile Leu Thr Phe Arg Ile Pro
[0235]	435 440 445
[0236]	Tyr Tyr Val Gly Pro Leu Ala Arg Gly Asn Ser Arg Phe Ala Trp Met
[0237]	450 455 460
[0238]	Thr Arg Lys Ser Glu Glu Thr Ile Thr Pro Trp Asn Phe Glu Glu Val
[0239]	465 470 475 480
[0240]	Val Asp Lys Gly Ala Ser Ala Gln Ser Phe Ile Glu Arg Met Thr Asn
[0241]	485 490 495
[0242]	Phe Asp Lys Asn Leu Pro Asn Glu Lys Val Leu Pro Lys His Ser Leu
[0243]	500 505 510
[0244]	Leu Tyr Glu Tyr Phe Thr Val Tyr Asn Glu Leu Thr Lys Val Lys Tyr
[0245]	515 520 525
[0246]	Val Thr Glu Gly Met Arg Lys Pro Ala Phe Leu Ser Gly Glu Gln Lys
[0247]	530 535 540
[0248]	Lys Ala Ile Val Asp Leu Leu Phe Lys Thr Asn Arg Lys Val Thr Val
[0249]	545 550 555 560
[0250]	Lys Gln Leu Lys Glu Asp Tyr Phe Lys Lys Ile Glu Cys Phe Asp Ser
[0251]	565 570 575

[0252]	Val Glu Ile Ser Gly Val Glu Asp Arg Phe Asn Ala Ser Leu Gly Thr
[0253]	580 585 590
[0254]	Tyr His Asp Leu Leu Lys Ile Ile Lys Asp Lys Asp Phe Leu Asp Asn
[0255]	595 600 605
[0256]	Glu Glu Asn Glu Asp Ile Leu Glu Asp Ile Val Leu Thr Leu Thr Leu
[0257]	610 615 620
[0258]	Phe Glu Asp Arg Glu Met Ile Glu Glu Arg Leu Lys Thr Tyr Ala His
[0259]	625 630 635 640
[0260]	Leu Phe Asp Asp Lys Val Met Lys Gln Leu Lys Arg Arg Arg Tyr Thr
[0261]	645 650 655
[0262]	Gly Trp Gly Arg Leu Ser Arg Lys Leu Ile Asn Gly Ile Arg Asp Lys
[0263]	660 665 670
[0264]	Gln Ser Gly Lys Thr Ile Leu Asp Phe Leu Lys Ser Asp Gly Phe Ala
[0265]	675 680 685
[0266]	Asn Arg Asn Phe Met Gln Leu Ile His Asp Asp Ser Leu Thr Phe Lys
[0267]	690 695 700
[0268]	Glu Asp Ile Gln Lys Ala Gln Val Ser Gly Gln Gly Asp Ser Leu His
[0269]	705 710 715 720
[0270]	Glu His Ile Ala Asn Leu Ala Gly Ser Pro Ala Ile Lys Lys Gly Ile
[0271]	725 730 735
[0272]	Leu Gln Thr Val Lys Val Val Asp Glu Leu Val Lys Val Met Gly Arg
[0273]	740 745 750
[0274]	His Lys Pro Glu Asn Ile Val Ile Glu Met Ala Arg Glu Asn Gln Thr
[0275]	755 760 765
[0276]	Thr Gln Lys Gly Gln Lys Asn Ser Arg Glu Arg Met Lys Arg Ile Glu
[0277]	770 775 780
[0278]	Glu Gly Ile Lys Glu Leu Gly Ser Gln Ile Leu Lys Glu His Pro Val
[0279]	785 790 795 800
[0280]	Glu Asn Thr Gln Leu Gln Asn Glu Lys Leu Tyr Leu Tyr Tyr Leu Gln
[0281]	805 810 815
[0282]	Asn Gly Arg Asp Met Tyr Val Asp Gln Glu Leu Asp Ile Asn Arg Leu
[0283]	820 825 830
[0284]	Ser Asp Tyr Asp Val Asp His Ile Val Pro Gln Ser Phe Leu Lys Asp
[0285]	835 840 845
[0286]	Asp Ser Ile Asp Asn Lys Val Leu Thr Arg Ser Asp Lys Asn Arg Gly
[0287]	850 855 860
[0288]	Lys Ser Asp Asn Val Pro Ser Glu Glu Val Val Lys Lys Met Lys Asn
[0289]	865 870 875 880
[0290]	Tyr Trp Arg Gln Leu Leu Asn Ala Lys Leu Ile Thr Gln Arg Lys Phe
[0291]	885 890 895
[0292]	Asp Asn Leu Thr Lys Ala Glu Arg Gly Gly Leu Ser Glu Leu Asp Lys
[0293]	900 905 910

[0294]	Ala Gly Phe Ile Lys Arg Gln Leu Val Glu Thr Arg Gln Ile Thr Lys
[0295]	915 920 925
[0296]	His Val Ala Gln Ile Leu Asp Ser Arg Met Asn Thr Lys Tyr Asp Glu
[0297]	930 935 940
[0298]	Asn Asp Lys Leu Ile Arg Glu Val Lys Val Ile Thr Leu Lys Ser Lys
[0299]	945 950 955 960
[0300]	Leu Val Ser Asp Phe Arg Lys Asp Phe Gln Phe Tyr Lys Val Arg Glu
[0301]	965 970 975
[0302]	Ile Asn Asn Tyr His His Ala His Asp Ala Tyr Leu Asn Ala Val Val
[0303]	980 985 990
[0304]	Gly Thr Ala Leu Ile Lys Lys Tyr Pro Lys Leu Glu Ser Glu Phe Val
[0305]	995 1000 1005
[0306]	Tyr Gly Asp Tyr Lys Val Tyr Asp Val Arg Lys Met Ile Ala Lys
[0307]	1010 1015 1020
[0308]	Ser Glu Gln Glu Ile Gly Lys Ala Thr Ala Lys Tyr Phe Phe Tyr
[0309]	1025 1030 1035
[0310]	Ser Asn Ile Met Asn Phe Phe Lys Thr Glu Ile Thr Leu Ala Asn
[0311]	1040 1045 1050
[0312]	Gly Glu Ile Arg Lys Arg Pro Leu Ile Glu Thr Asn Gly Glu Thr
[0313]	1055 1060 1065
[0314]	Gly Glu Ile Val Trp Asp Lys Gly Arg Asp Phe Ala Thr Val Arg
[0315]	1070 1075 1080
[0316]	Lys Val Leu Ser Met Pro Gln Val Asn Ile Val Lys Lys Thr Glu
[0317]	1085 1090 1095
[0318]	Val Gln Thr Gly Gly Phe Ser Lys Glu Ser Ile Leu Pro Lys Arg
[0319]	1100 1105 1110
[0320]	Asn Ser Asp Lys Leu Ile Ala Arg Lys Lys Asp Trp Asp Pro Lys
[0321]	1115 1120 1125
[0322]	Lys Tyr Gly Gly Phe Asp Ser Pro Thr Val Ala Tyr Ser Val Leu
[0323]	1130 1135 1140
[0324]	Val Val Ala Lys Val Glu Lys Gly Lys Ser Lys Lys Leu Lys Ser
[0325]	1145 1150 1155
[0326]	Val Lys Glu Leu Leu Gly Ile Thr Ile Met Glu Arg Ser Ser Phe
[0327]	1160 1165 1170
[0328]	Glu Lys Asn Pro Ile Asp Phe Leu Glu Ala Lys Gly Tyr Lys Glu
[0329]	1175 1180 1185
[0330]	Val Lys Lys Asp Leu Ile Ile Lys Leu Pro Lys Tyr Ser Leu Phe
[0331]	1190 1195 1200
[0332]	Glu Leu Glu Asn Gly Arg Lys Arg Met Leu Ala Ser Ala Gly Glu
[0333]	1205 1210 1215
[0334]	Leu Gln Lys Gly Asn Glu Leu Ala Leu Pro Ser Lys Tyr Val Asn
[0335]	1220 1225 1230

[0336]	Phe Leu Tyr Leu Ala Ser His Tyr Glu Lys Leu Lys Gly Ser Pro
[0337]	1235 1240 1245
[0338]	Glu Asp Asn Glu Gln Lys Gln Leu Phe Val Glu Gln His Lys His
[0339]	1250 1255 1260
[0340]	Tyr Leu Asp Glu Ile Ile Glu Gln Ile Ser Glu Phe Ser Lys Arg
[0341]	1265 1270 1275
[0342]	Val Ile Leu Ala Asp Ala Asn Leu Asp Lys Val Leu Ser Ala Tyr
[0343]	1280 1285 1290
[0344]	Asn Lys His Arg Asp Lys Pro Ile Arg Glu Gln Ala Glu Asn Ile
[0345]	1295 1300 1305
[0346]	Ile His Leu Phe Thr Leu Thr Asn Leu Gly Ala Pro Ala Ala Phe
[0347]	1310 1315 1320
[0348]	Lys Tyr Phe Asp Thr Thr Ile Asp Arg Lys Arg Tyr Thr Ser Thr
[0349]	1325 1330 1335
[0350]	Lys Glu Val Leu Asp Ala Thr Leu Ile His Gln Ser Ile Thr Gly
[0351]	1340 1345 1350
[0352]	Leu Tyr Glu Thr Arg Ile Asp Leu Ser Gln Leu Gly Gly Asp
[0353]	1355 1360 1365
[0354]	<210>4
[0355]	<211>19
[0356]	<212>PRT
[0357]	<213>Artificial Sequence
[0358]	<400>4
[0359]	Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
[0360]	1 5 10 15
[0361]	Pro Gly Pro
[0362]	<210>5
[0363]	<211>341
[0364]	<212>PRT
[0365]	<213>Artificial Sequence
[0366]	<400>5
[0367]	Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile
[0368]	1 5 10 15
[0369]	Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu
[0370]	20 25 30
[0371]	Glu Ser Arg Ala Phe Ser Phe Asp Val Gly Gly Arg Gly Tyr Val Leu
[0372]	35 40 45
[0373]	Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr
[0374]	50 55 60
[0375]	Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile
[0376]	65 70 75 80
[0377]	Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ser Gln



[0378]		85		90		95
[0379]	Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu					
[0380]		100		105		110
[0381]	Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser					
[0382]		115		120		125
[0383]	Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr					
[0384]		130		135		140
[0385]	Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr					
[0386]		145		150		155
[0387]	His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln					
[0388]		165		170		175
[0389]	Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg					
[0390]		180		185		190
[0391]	His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn					
[0392]		195		200		205
[0393]	Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp					
[0394]		210		215		220
[0395]	Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala					
[0396]		225		230		235
[0397]	Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu					
[0398]		245		250		255
[0399]	Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp					
[0400]		260		265		270
[0401]	Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp					
[0402]		275		280		285
[0403]	Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val					
[0404]		290		295		300
[0405]	Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly					
[0406]		305		310		315
[0407]	Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg					
[0408]		325		330		335
[0409]	Pro Arg Ala Lys Lys					
[0410]		340				
[0411]	<210>6					
[0412]	<211>23					
[0413]	<212>DNA					
[0414]	<213>Artificial Sequence					
[0415]	<400>6					
[0416]	ccgttgacc accatacca aat 23					
[0417]	<210>7					
[0418]	<211>20					
[0419]	<212>DNA					

[0420] <213>Artificial Sequence  
[0421] <400>7  
[0422] gcaccacat acccaaatgt 20  
[0423] <210>8  
[0424] <211>23  
[0425] <212>DNA  
[0426] <213>Artificial Sequence  
[0427] <400>8  
[0428] ccaacatttg ggtatggtgg tgc 23  
[0429] <210>9  
[0430] <211>20  
[0431] <212>DNA  
[0432] <213>Artificial Sequence  
[0433] <400>9  
[0434] gcatacatatc ttgggcaacc 20  
[0435] <210>10  
[0436] <211>23  
[0437] <212>DNA  
[0438] <213>Artificial Sequence  
[0439] <400>10  
[0440] gcatacatatc ttgggcaacc cgg 23  
[0441] <210>11  
[0442] <211>23  
[0443] <212>DNA  
[0444] <213>Artificial Sequence  
[0445] <400>11  
[0446] ccgggtgcc caagtatgta tgc 23  
[0447] <210>12  
[0448] <211>20  
[0449] <212>DNA  
[0450] <213>Artificial Sequence  
[0451] <400>12  
[0452] atatctcgct ctcacattcc 20  
[0453] <210>13  
[0454] <211>23  
[0455] <212>DNA  
[0456] <213>Artificial Sequence  
[0457] <400>13  
[0458] cccggaatgt gagagcgaga tat 23  
[0459] <210>14  
[0460] <211>2936  
[0461] <212>DNA

[0462]	<213>Artificial Sequence	
[0463]	<400>14	
[0464]	atggetacga cgcgcgccc cgcggccgcc gccctgtccg ccgcccgcac ggccaagacc	60
[0465]	ggccgtaaga accaccagcg acaccacgtc cttcccgcgc gaggccgggt gggggcggcg	120
[0466]	gcggtcaggt gctcggcggt gtccccggtc accccgccgt ccccggcgcc gccggccacg	180
[0467]	ccgctccggc cgtgggggcc ggccgagccc cgcaaggcg cgacatcct cgtggaggcg	240
[0468]	ctggagcggg gcggcgtcag cgacgtgttc gcctaccgg gcggcgcgc catggagatc	300
[0469]	caccaggcgc tgacgcgctc cccggtcacc accaaccacc tcttcgccca cgagcagggc	360
[0470]	gaggcgttcg cggcgtccgg gtacgcgcgc gcgtccggcc gcgtcgggg ctgcgtcgc	420
[0471]	acctccggcc ccggggcaac caacctcgtg tccgcgctcg ccgacgcgct gctcgactcc	480
[0472]	gtcccgatgg tcgccatcac gggccaggtc ccccgccgca tgatcggcac cgacgccttc	540
[0473]	caggagacgc ccatagtcga ggtcaccgc tccatcacca agcacaatta ccttgtcctt	600
[0474]	gatgtggagg acatccccg cgtcatacag gaagccttct tctcgcgctc ctcgggccgt	660
[0475]	cctggcccgg tgctggtcga catcccgaag gacatccagc agcagatggc cgtgccggtc	720
[0476]	tgggacacct cgatgaatct accagggtac atcgcacgcc tgcccgaagc acccgcgaca	780
[0477]	gaattgcttg agcaggtctt gcgtctggtt ggcgagtcac ggcgcccgat tctctatgct	840
[0478]	ggtggtggct gctctgcacc tggtagcaga ttgcgctggt ttgttagact gactggtatc	900
[0479]	ccagttacaa cactctgat gggcctcggc aatttccca gtgacgacc gttgtccctg	960
[0480]	cgcatgcttg ggatgcatgg cacggtgtac gcaattatg ccgtggataa ggctgacctg	1020
[0481]	ttgcttgctg ttggtgtgcg gttttagatg cgtgtgacag ggaaaattga ggcttttgca	1080
[0482]	agcagggcca agattgtgca cattgacatt gatccagcag agattggaaa gaacaagcaa	1140
[0483]	ccacatgtgt caatttgccg agatgttaag cttgctttac agggcttgaa tgctctgcta	1200
[0484]	caacagagca caacaagac aagtctgat tttagtgcat ggcacaatga gttggaccag	1260
[0485]	cagaagaggg agtttctct ggggtacaaa acttttggtg aagagatccc accgcaatat	1320
[0486]	gccattcagg tgctggatga gctgacgaaa ggtgaggcaa tcacgcctac tgggtgttggg	1380
[0487]	cagcaccaga tgtggcgccg acaatattac acctacaagc ggccacggca gtggctgtct	1440
[0488]	tcggctggtc tggcgcaat gggattggg ctgcctgctg cagctggtgc ttctgtggct	1500
[0489]	aaccagggtg tcacagttgt tgatattgat ggggatggta gcttcctcat gaacattcag	1560
[0490]	gagctggcat tgatccgcat tgagaacctc cctgtgaagg tgatggtgtt gaacaacca	1620
[0491]	catttgggta tgggtgtgca atgggaggat aggttttaca aggcgaatag ggcgcataca	1680
[0492]	tacttgggca acccggaatg tgagagcgag atatatccag attttgtgac tattgctaag	1740
[0493]	gggttcaata ttctgcagt ccgtgtaaca aagaagagtg aagtccgtgc cgccatcaag	1800
[0494]	aagatgctcg agactccagg gccatacttg ttggatatca tcgtcccgca ccaggagcat	1860
[0495]	gtgctgccta tgatcccag tggggcgcca ttcaaggaca tgatcctgga tggatgatggc	1920
[0496]	aggactgtgt attaacttat aatctgtatg ttggcaaagc accagcccgg cctatgtttg	1980
[0497]	acctgaatga cccataaaga gtggtatgcc tatgatgttt gtatgtgctc tatcaataac	2040
[0498]	taagggtgca actatgaacc atatgctctt ctgttttact tgtttgatgt gcttggcatg	2100
[0499]	gtaatcctaa ttagcttct gctgtctagg tttgtagtgt gttgttttct gtaggcatac	2160
[0500]	gcatcacaaag atatcatgta agtttcttgt cctacatc aataataaga gaataaagta	2220
[0501]	cttctatgca atagctctga gtttaagtgt tcaacaattt ctgaacttct gaacttatgt	2280
[0502]	ttgtcaact gtcatcacac gaagtactct ccttgtaact acattttccc caagacttta	2340
[0503]	aatcccctca gttacagcaa aaaataaact ttgcatctac tgttttccct ctcttcggtc	2400

[0504] gatcttattg ggtactacta tagagagagg ctgcatgaag tatttccttt ttctgttttag 2460  
 [0505] ttatgccgtg taaattagca tccatgcaaa atagatgaaa aatcaagcta ttcctgactg 2520  
 [0506] ctaaggatta tttttggcat aatgtattct tatatactcc ctccgtccca tattataagg 2580  
 [0507] gattttgagt ttttgittat actgtttgac cactcgtctt attcaaaaaa ttttagaatt 2640  
 [0508] attatttatt tttttgtga cttactttat tatctaaagt actttaagca caattttcgt 2700  
 [0509] attttatatt tgcacaaatt ttttgaataa gacgaatgg caaacaatac aaataaaaat 2760  
 [0510] tcaaaaatccc ttataatatg ggacggaggt atgatagttg gtgaactget acgtattgcc 2820  
 [0511] atttgacatt ttttgatta tgcaattttg ctgtctatag tgctctaate aattcgcaat 2880  
 [0512] cccgaccttg gagtattggc ctcatggaac ccctcatctg agtaatctcc atattt 2936  
 [0513] <210>15  
 [0514] <211>644  
 [0515] <212>PRT  
 [0516] <213>Artificial Sequence  
 [0517] <400>15  
 [0518] Met Ala Thr Thr Ala Ala Ala Ala Ala Ala Ala Leu Ser Ala Ala Ala  
 [0519] 1 5 10 15  
 [0520] Thr Ala Lys Thr Gly Arg Lys Asn His Gln Arg His His Val Leu Pro  
 [0521] 20 25 30  
 [0522] Ala Arg Gly Arg Val Gly Ala Ala Ala Val Arg Cys Ser Ala Val Ser  
 [0523] 35 40 45  
 [0524] Pro Val Thr Pro Pro Ser Pro Ala Pro Pro Ala Thr Pro Leu Arg Pro  
 [0525] 50 55 60  
 [0526] Trp Gly Pro Ala Glu Pro Arg Lys Gly Ala Asp Ile Leu Val Glu Ala  
 [0527] 65 70 75 80  
 [0528] Leu Glu Arg Cys Gly Val Ser Asp Val Phe Ala Tyr Pro Gly Gly Ala  
 [0529] 85 90 95  
 [0530] Ser Met Glu Ile His Gln Ala Leu Thr Arg Ser Pro Val Ile Thr Asn  
 [0531] 100 105 110  
 [0532] His Leu Phe Arg His Glu Gln Gly Glu Ala Phe Ala Ala Ser Gly Tyr  
 [0533] 115 120 125  
 [0534] Ala Arg Ala Ser Gly Arg Val Gly Val Cys Val Ala Thr Ser Gly Pro  
 [0535] 130 135 140  
 [0536] Gly Ala Thr Asn Leu Val Ser Ala Leu Ala Asp Ala Leu Leu Asp Ser  
 [0537] 145 150 155 160  
 [0538] Val Pro Met Val Ala Ile Thr Gly Gln Val Pro Arg Arg Met Ile Gly  
 [0539] 165 170 175  
 [0540] Thr Asp Ala Phe Gln Glu Thr Pro Ile Val Glu Val Thr Arg Ser Ile  
 [0541] 180 185 190  
 [0542] Thr Lys His Asn Tyr Leu Val Leu Asp Val Glu Asp Ile Pro Arg Val  
 [0543] 195 200 205  
 [0544] Ile Gln Glu Ala Phe Phe Leu Ala Ser Ser Gly Arg Pro Gly Pro Val  
 [0545] 210 215 220

[0546]	Leu Val Asp Ile Pro Lys Asp Ile Gln Gln Gln Met Ala Val Pro Val
[0547]	225 230 235 240
[0548]	Trp Asp Thr Ser Met Asn Leu Pro Gly Tyr Ile Ala Arg Leu Pro Lys
[0549]	245 250 255
[0550]	Pro Pro Ala Thr Glu Leu Leu Glu Gln Val Leu Arg Leu Val Gly Glu
[0551]	260 265 270
[0552]	Ser Arg Arg Pro Ile Leu Tyr Val Gly Gly Gly Cys Ser Ala Ser Gly
[0553]	275 280 285
[0554]	Asp Glu Leu Arg Trp Phe Val Glu Leu Thr Gly Ile Pro Val Thr Thr
[0555]	290 295 300
[0556]	Thr Leu Met Gly Leu Gly Asn Phe Pro Ser Asp Asp Pro Leu Ser Leu
[0557]	305 310 315 320
[0558]	Arg Met Leu Gly Met His Gly Thr Val Tyr Ala Asn Tyr Ala Val Asp
[0559]	325 330 335
[0560]	Lys Ala Asp Leu Leu Leu Ala Phe Gly Val Arg Phe Asp Asp Arg Val
[0561]	340 345 350
[0562]	Thr Gly Lys Ile Glu Ala Phe Ala Ser Arg Ala Lys Ile Val His Ile
[0563]	355 360 365
[0564]	Asp Ile Asp Pro Ala Glu Ile Gly Lys Asn Lys Gln Pro His Val Ser
[0565]	370 375 380
[0566]	Ile Cys Ala Asp Val Lys Leu Ala Leu Gln Gly Leu Asn Ala Leu Leu
[0567]	385 390 395 400
[0568]	Gln Gln Ser Thr Thr Lys Thr Ser Ser Asp Phe Ser Ala Trp His Asn
[0569]	405 410 415
[0570]	Glu Leu Asp Gln Gln Lys Arg Glu Phe Pro Leu Gly Tyr Lys Thr Phe
[0571]	420 425 430
[0572]	Gly Glu Glu Ile Pro Pro Gln Tyr Ala Ile Gln Val Leu Asp Glu Leu
[0573]	435 440 445
[0574]	Thr Lys Gly Glu Ala Ile Ile Ala Thr Gly Val Gly Gln His Gln Met
[0575]	450 455 460
[0576]	Trp Ala Ala Gln Tyr Tyr Thr Tyr Lys Arg Pro Arg Gln Trp Leu Ser
[0577]	465 470 475 480
[0578]	Ser Ala Gly Leu Gly Ala Met Gly Phe Gly Leu Pro Ala Ala Ala Gly
[0579]	485 490 495
[0580]	Ala Ser Val Ala Asn Pro Gly Val Thr Val Val Asp Ile Asp Gly Asp
[0581]	500 505 510
[0582]	Gly Ser Phe Leu Met Asn Ile Gln Glu Leu Ala Leu Ile Arg Ile Glu
[0583]	515 520 525
[0584]	Asn Leu Pro Val Lys Val Met Val Leu Asn Asn Gln His Leu Gly Met
[0585]	530 535 540
[0586]	Val Val Gln Trp Glu Asp Arg Phe Tyr Lys Ala Asn Arg Ala His Thr
[0587]	545 550 555 560

[0588]	Tyr Leu Gly Asn Pro Glu Cys Glu Ser Glu Ile Tyr Pro Asp Phe Val
[0589]	565 570 575
[0590]	Thr Ile Ala Lys Gly Phe Asn Ile Pro Ala Val Arg Val Thr Lys Lys
[0591]	580 585 590
[0592]	Ser Glu Val Arg Ala Ala Ile Lys Lys Met Leu Glu Thr Pro Gly Pro
[0593]	595 600 605
[0594]	Tyr Leu Leu Asp Ile Ile Val Pro His Gln Glu His Val Leu Pro Met
[0595]	610 615 620
[0596]	Ile Pro Ser Gly Gly Ala Phe Lys Asp Met Ile Leu Asp Gly Asp Gly
[0597]	625 630 635 640
[0598]	Arg Thr Val Tyr

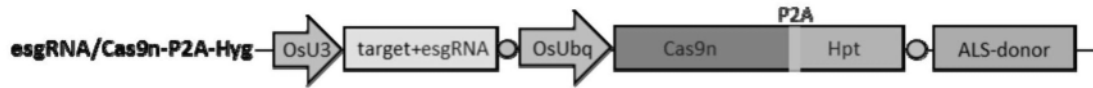


图1

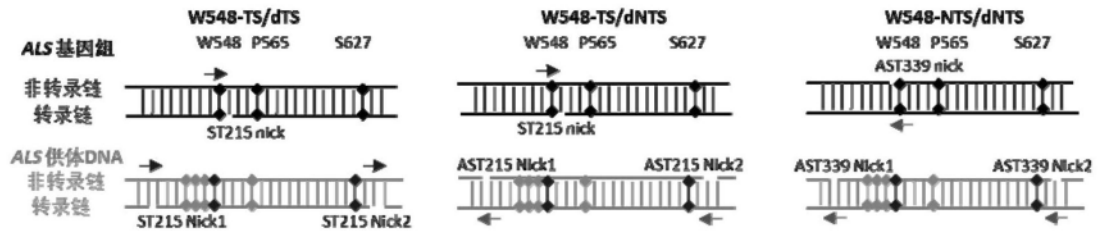


图2

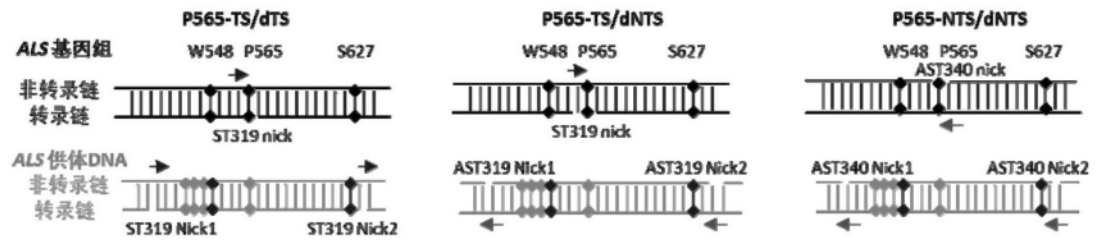


图3

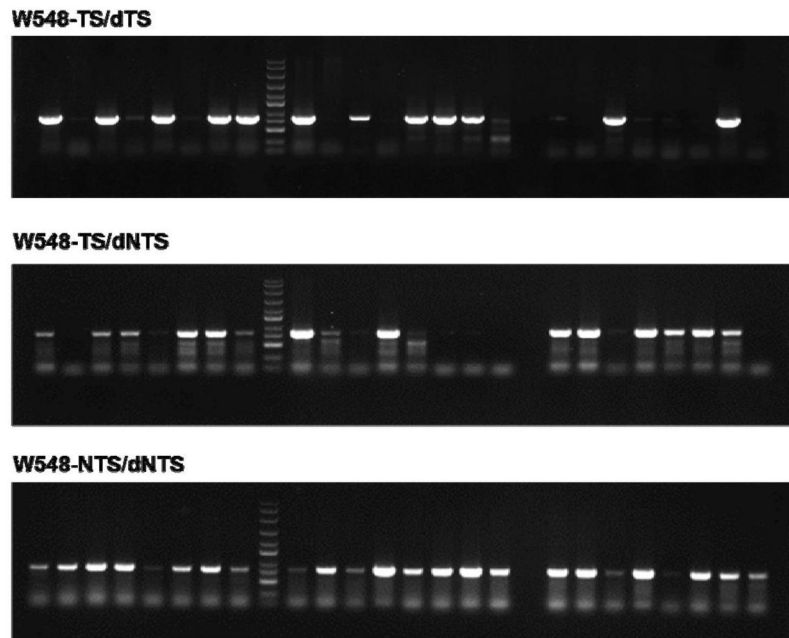


图4

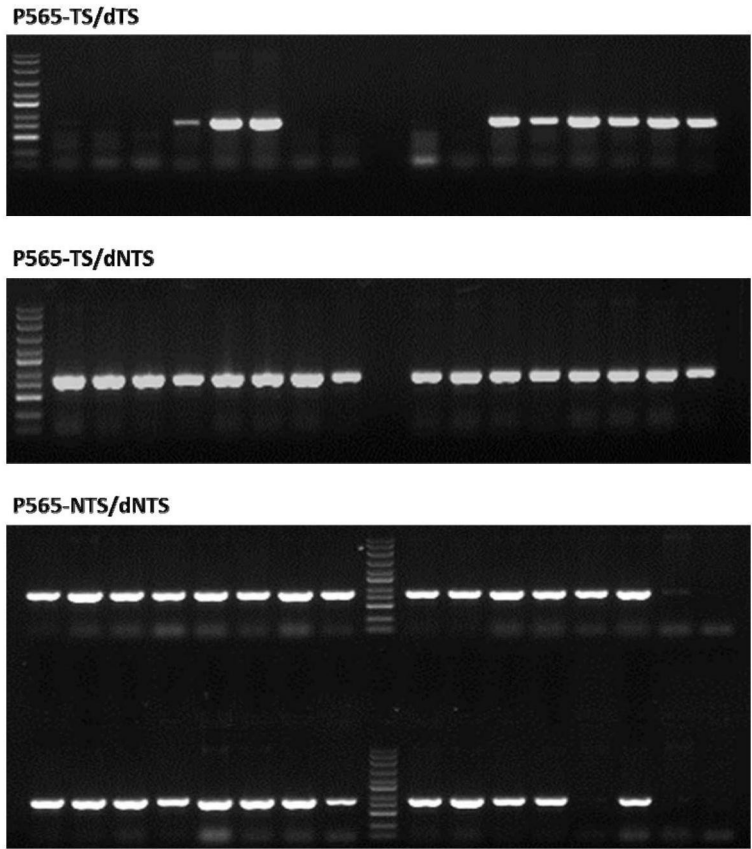


图5