



(12) 发明专利申请

(10) 申请公布号 CN 114438100 A

(43) 申请公布日 2022. 05. 06

(21) 申请号 202210192367.2

C12N 15/10 (2006.01)

(22) 申请日 2022.03.01

C12N 15/11 (2006.01)

C12Q 1/6869 (2018.01)

(71) 申请人 云南省农业科学院生物技术与种质资源研究所

地址 650205 云南省昆明市北京路2238号  
云南省农业科学院

(72) 发明人 陈玲 张敦宇 程在全 殷富有  
钟巧芳 王波 卢源达 肖素勤  
张云 王玲仙 柯学 余腾琼  
蒋聪 刘丽

(74) 专利代理机构 北京方圆嘉禾知识产权代理有限公司 11385

专利代理师 崔玥

(51) Int. Cl.

C12N 15/29 (2006.01)

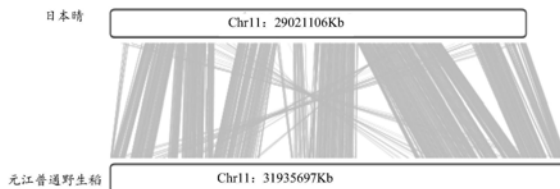
权利要求书2页 说明书15页  
序列表32页 附图5页

(54) 发明名称

一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法

(57) 摘要

本发明提供了一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法,属于分子生物学技术领域。本发明提供了一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法,利用来自元江普通野生稻与栽培稻杂交获得的多样化抗谱后代材料,高效快速地克隆到Xa47(t)基因及其家族成员,不仅为分离新的基因提供了较好的借鉴方法,同时避免了常规图位克隆的繁琐和耗时费力,利用元江普通野生稻杂交后代(渗入系)获得的家族成员,其遗传背景清晰,为后续研究家族成员的功能以及育种利用奠定了坚实的基础。



1. 一种高效分离带有野生稻血缘的抗白叶枯病基因的方法,其特征在于,所述方法包括:

通过对元江普通野生稻渗入系双亲进行建库和高通量测序获得连续基因组,将所述连续基因组与参考基因组进行共线性分析,根据目的基因定位情况,筛选出所述抗白叶枯病基因的候选基因,根据所述候选基因设计引物进行PCR扩增、PCR产物直接测序和抗病功能分析,从而分离出带有野生稻血缘的抗白叶枯病基因。

2. 根据权利要求1所述的方法,其特征在于,所述抗白叶枯病基因为Xa47(t),所述基因的核苷酸序列如SEQ ID NO.13所示;所述基因Xa47(t)的CDS序列如SEQ ID NO.14所示。

3. 根据权利要求1所述的方法,其特征在于,所述元江普通野生稻渗入系双亲为元江普通野生稻和栽培稻合系35;所述参考基因组为日本晴基因组;所述目的基因定位和所述PCR扩增的模板为元江普通野生稻渗入系G252。

4. 根据权利要求1所述的方法,其特征在于,所述引物为214QC-9F/R;所述214QC-9F的核苷酸序列如SEQ ID NO.11所示,所述214QC-9R的核苷酸序列如SEQ ID NO.12所示。

5. 根据权利要求1所述的方法,其特征在于,所述抗病功能分析包括功能互补试验和基因编辑试验分析候选基因的抗病功能。

6. 一种高效分离带有野生稻血缘的抗白叶枯病基因家族成员的方法,其特征在于,所述方法包括:

筛选获得含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料;根据权利要求2所述Xa47(t)基因序列,获得来自元江普通野生稻和合系35中的Xa47(t)基因,分别命名为Xa47(t)<sup>YP</sup>和Xa47(t)<sup>HX</sup>;

利用引物对所述元江普通野生稻渗入系不同材料进行扩增、测序,得到4种Xa47(t)基因的基因型:Xa47(t)<sup>G252</sup>、Xa47(t)<sup>YP</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>;对由4种基因型编码得到的蛋白质进行相似性分析,从而确定其是否为带有野生稻血缘的抗白叶枯病基因家族成员。

7. 根据权利要求6所述的方法,其特征在于,所述筛选的步骤包括:

利用Xa47(t)基因的共分离标记Hxjy-1对元江普通野生稻渗入系进行PCR扩增,得到扩增片段大小为167bp的渗入系,结合渗入系的抗病谱筛选出含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料。

8. 根据权利要求6所述的方法,其特征在于,根据权利要求2所述的Xa47(t)基因序列,利用blast程序从元江普通野生稻和合系35基因组R13I14和Hxjy-14区段中进行检索,即获得来自元江普通野生稻和合系35中的Xa47(t)基因。

9. 根据权利要求6所述的方法,其特征在于,所述引物包括214QC-2F/R、48QC-9F/R和48QC-11F/R,所述214QC-2F的核苷酸序列如SEQ ID NO.17所示,所述214QC-2R的核苷酸序列如SEQ ID NO.18所示;所述48QC-9F的核苷酸序列如SEQ ID NO.19所示,所述48QC-9R的核苷酸序列如SEQ ID NO.20所示;所述48QC-11F的核苷酸序列如SEQ ID NO.21所示,所述48QC-11R的核苷酸序列如SEQ ID NO.22所示。

10. 根据权利要求6所述的方法,其特征在于,所述Xa47(t)<sup>G252</sup>、Xa47(t)<sup>YP</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>基因即以其供体材料进行命名的Xa47(t)基因;所述Xa47(t)<sup>G252</sup>为权利要求2所述的Xa47(t)基因;所述Xa47(t)<sup>YP</sup>基因的核苷酸序列如SEQ ID NO.15所示;所述Xa47(t)<sup>YP</sup>基因的CDS序列如SEQ ID NO.30所示;所述Xa47(t)<sup>HX</sup>基因的核苷酸序列如SEQ ID NO.16

所示;所述Xa47(t)<sup>HX</sup>基因的CDS序列如SEQ ID NO.32所示;所述Xa47(t)<sup>L234</sup>基因的核苷酸序列如SEQ ID NO.27所示;所述Xa47(t)<sup>L234</sup>基因的CDS序列如SEQ ID NO.28所示。

## 一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法

### 技术领域

[0001] 本发明涉及分子生物学技术领域,具体涉及一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法。

### 背景技术

[0002] 水稻(*Oryza sativa* L.)是全世界近一半人口的主粮,然而水稻生长发育过程中时常遭遇各种病原物的危害,严重制约着水稻的高产稳产以及稻米品质。由稻黄单胞杆菌(*Xanthomonas oryzae* pv.*oryzae*,简称Xoo)引起的水稻白叶枯病(bacterial blight)是世界范围内最严重的细菌性病害,也是我国水稻三大病害之一,我国除新疆、西藏和东北的北部以外均有不同程度的发生,长江以南和江淮平原籼稻产区为常发区,遭受白叶枯病危害后易造成水稻减产20%~30%,严重时可达50%,甚至绝收。随着全球气候逐渐变暖,水稻白叶枯病发生面积在逐年增加,水稻产量损失严重,已严重威胁到国家粮食安全(陈功友,徐正银,杨阳阳,等.我国水稻白叶枯病菌致病型划分和水稻抗病育种中应注意的问题.上海交通大学学报(农业科学版),2019,37(1):71-77.)。从长远途径来说,要想在水稻白叶枯病害常发区获得理想的水稻产量,就必须不断培育抗病新品种,以抵御白叶枯病菌的不断变异。尤其是多渠道、多来源发掘出抗白叶枯病基因,并服务于育种实践,已成为目前水稻分子育种的重要目标。

[0003] 回顾水稻新品种的培育历程,育种家们大都将来自栽培品种中的优良基因不断重组、聚合,导致栽培品种原有的遗传资源利用日趋饱和,许多改良品种具有相同或相似的遗传来源,造成了遗传上的单一性,尤其由于早期育种家们对栽培稻抗病特性的忽视,我们现有的栽培稻不具有抗病特性或携有的抗病基因单一。从远缘种群中引入优良基因创作新的水稻品种是扩大亲本间遗传差异的重要途径之一。栽培稻有20多个野生近缘种,包含有极其丰富的遗传变异。云南省是亚洲栽培稻的遗传多样性中心和起源中心之一,拥有中国的全部三种野生稻,即普通野生稻(*Oryza rufipogon* Griff.)、药用野生稻(*Oryza officinalis* Wall.)和疣粒野生稻(*Oryza granulata* Baill.),而且每种野生稻具有许多亚种或生态类型,这些种质资源在表型和分子水平上都具有较高的遗传多样性,拥有显著的抗逆、抗病虫害、耐寒、耐旱和耐贫瘠等许多优良性状(陈玲,张敦宇,陈越,等.云南药用野生稻种质资源的白叶枯病抗性评价.南方农业学报,2019,50(7):1417-1425;程在全.云南野生稻遗传特性及其优良基因克隆研究.四川:四川大学,2006.)。因此,利用拥有的广谱高抗白叶枯病的野生资源培育水稻新品种,这不仅将大大降低生产成本,而且具有重要的生态意义。

[0004] 前期我们利用广谱高抗白叶枯病的元江普通野生稻渗入系(元江普通野生稻与栽培稻合系35杂交后代BC<sub>2</sub>F<sub>16</sub>)精细定位到了一个位于水稻11号染色体的显性抗白叶枯病基因Xa47(t)(陈玲,钟巧芳,王玲仙,等.与水稻广谱高抗白叶枯病基因Xa45(t)紧密连锁的分子标记Hxjy-14,2020a,中国,ZL201910825566.0.;陈玲,王波,程在全,等.水稻广谱高抗白

叶枯病基因Xa45(t)的共分离分子标记Hxjy-1,2020b,中国,ZL201910825603.0.;陈玲,程在全,王波,等.与水稻广谱高抗叶枯病基因Xa45(t)紧密连锁分子标记R13I14,2020c,中国,ZL201910825560.3.),之前暂命名为Xa45(t),因与位于4号染色体的水稻抗白叶枯病基因Xa1的等位基因Xa45(t)重名(Ji CH,Ji ZY,Liu B,et al.Xa1 allelic R genes activate rice blight resistance suppressed by interfering TAL effectors.Plant Comm,2020,1(4),100087.),所以后将其改为Xa47(t)。利用Xa47(t)的共分离标记检测发现该基因带有元江普通野生稻血缘,目前从野生稻中克隆到的基因只有3个,分别为来自长雄野生稻(*Oryza longistaminata*A.chev et Roehr)的Xa21(Song WY,Wang GL,Chen LL,et al.A receptor kinase-like protein encoded by the rice disease resistance gene,Xa21.Science,1995,270(5243):1804-1806.)、来自普通野生稻的Xa23(ZhouYL,Uzokwe VNE,Zhang CH.Improvement of bacterial blight resistance of hybrid rice in China using the Xa23 gene derived from wild rice(*Oryza rufipogon*).Crop Prot,2011,30(6):637-644.)和来自小粒野生稻(*Oryza minuta* Presl.)的Xa27基因(Gu K,Tian D,Yang F,et al.High-resolution genetic mapping of Xa27(t),a new bacterial blight resistance gene in rice,*Oryza sativa* L.Theor Appl Genet,2004,108(5):800-807.),这3个基因的克隆方式基本都沿用了常规的图位克隆法,但该方法费时费力,随着测序技术的到来,利用测序技术高效且快速地分离Xa47(t)基因已成为可能。经接种9个白叶枯病菌发现,359份元江普通野生稻渗入系抗病谱多样化,存在67种抗病谱,抗病谱不同携带的抗白叶枯病基因可能有所不同,或者这些抗病基因很可能为同一个家族成员,由于这些渗入系遗传背景相似,参照来自渗入系G252中的Xa47(t)基因序列,利用元江普通野生稻的其他渗入系可大量地分离到Xa47(t)家族成员。那么,如何充分利用现代测序技术以及分子生物学技术从元江普通野生稻及其渗入系中高效地分离Xa47(t)基因及其家族成员成为了急需要解决的科学问题。

## 发明内容

[0005] 有鉴于此,本发明的目的在于提供一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法,能够高效快速地克隆到Xa47(t)基因及其家族成员。

[0006] 为解决上述技术问题,本发明提供了以下技术方案:

[0007] 本发明提供了一种高效分离带有野生稻血缘的抗白叶枯病基因的方法,所述方法包括:

[0008] 通过对元江普通野生稻渗入系双亲进行建库和高通量测序获得连续基因组,将所述连续基因组与参考基因组进行共线性分析,根据目的基因定位情况,筛选出所述抗白叶枯病基因的候选基因,根据所述候选基因设计引物进行PCR扩增、PCR产物直接测序和抗病功能分析,从而分离出带有野生稻血缘的抗白叶枯病基因。

[0009] 优选的,所述抗白叶枯病基因为Xa47(t),所述基因的核苷酸序列如SEQ ID NO.13所示;所述基因Xa47(t)的CDS序列如SEQ ID NO.14所示。

[0010] 优选的,所述元江普通野生稻渗入系双亲为元江普通野生稻和栽培稻合系35;所述参考基因组为日本晴基因组;所述目的基因定位和所述PCR扩增的模板为元江普通野生稻渗入系G252。

[0011] 优选的,所述引物为214QC-9F/R;所述214QC-9F的核苷酸序列如SEQ ID NO.11所示,所述214QC-9R的核苷酸序列如SEQ ID NO.12所示。

[0012] 优选的,所述抗病功能分析包括功能互补试验和基因编辑试验分析候选基因的抗病功能。

[0013] 本发明还提供了一种高效分离带有野生稻血缘的抗白叶枯病基因家族成员的方法,所述方法包括:

[0014] 筛选获得含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料;根据上述Xa47(t)基因序列,获得来自元江普通野生稻和合系35中的Xa47(t)基因,分别命名为Xa47(t)<sup>YP</sup>和Xa47(t)<sup>HX</sup>;

[0015] 利用引物对所述元江普通野生稻渗入系不同材料进行扩增、测序,得到4种Xa47(t)基因的基因型:Xa47(t)<sup>G252</sup>、Xa47(t)<sup>YP</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>;对由4种基因型编码得到的蛋白质进行相似性分析,从而确定其是否为带有野生稻血缘的抗白叶枯病基因家族成员。

[0016] 优选的,所述筛选的步骤包括:

[0017] 利用Xa47(t)基因的共分离标记Hxjy-1对元江普通野生稻渗入系进行PCR扩增,得到扩增片段大小为167bp的渗入系,结合渗入系的抗病谱筛选出含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料。

[0018] 优选的,根据上述Xa47(t)基因序列,利用blast程序从元江普通野生稻和合系35基因组R13I14和Hxjy-14区段中进行检索,即获得来自元江普通野生稻和合系35中的Xa47(t)基因。

[0019] 优选的,所述引物包括214QC-2F/R、48QC-9F/R和48QC-11F/R,所述214QC-2F的核苷酸序列如SEQ ID NO.17所示,所述214QC-2R的核苷酸序列如SEQ ID NO.18所示;所述48QC-9F的核苷酸序列如SEQ ID NO.19所示,所述48QC-9R的核苷酸序列如SEQ ID NO.20所示;所述48QC-11F的核苷酸序列如SEQ ID NO.21所示,所述48QC-11R的核苷酸序列如SEQ ID NO.22所示。

[0020] 优选的,所述Xa47(t)<sup>G252</sup>、Xa47(t)<sup>YP</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>基因即以其供体材料进行命名的Xa47(t)基因;所述Xa47(t)<sup>G252</sup>为上述的Xa47(t)基因;所述Xa47(t)<sup>YP</sup>基因的核苷酸序列如SEQ ID NO.15所示;所述Xa47(t)<sup>YP</sup>基因的CDS序列如SEQ ID NO.30所示;所述Xa47(t)<sup>HX</sup>基因的核苷酸序列如SEQ ID NO.16所示;所述Xa47(t)<sup>HX</sup>基因的CDS序列如SEQ ID NO.32所示;所述Xa47(t)<sup>L234</sup>基因的核苷酸序列如SEQ ID NO.27所示;所述Xa47(t)<sup>L234</sup>基因的CDS序列如SEQ ID NO.28所示。

[0021] 本发明提供了一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法,利用PacBio SMRT技术和Dovetail Hi-C技术成功地将元江普通野生稻和栽培稻合系35的基因组对应到水稻染色体上,并与日本晴基因组进行共线性分析,根据测序组装序列以及Xa47(t)基因的定位情况,快速获得了元江普通野生稻和合系35中Xa47(t)基因型。利用根据该基因型设计的引物扩增后的目的产物可直接进行测序获得目的基因的完整编码区,无需通过繁琐的载体连接、蓝白斑筛选以及酶切验证等步骤,该技术高效、快捷、省时省力,为分离新的基因提供了较好的借鉴方法。同时,本发明利用来自同一个亲本的多样化抗病谱材料,高效快速地克隆到Xa47(t)基因及其家族成员,避免了常规图位克隆的繁琐和耗时

费力,利用元江普通野生稻渗入系获得的家族成员,其遗传背景清晰,为后续研究家族成员的功能以及育种利用奠定了坚实的基础。

### 附图说明

[0022] 图1为合系35与日本晴11号染色体共线性图。Chr11表示11号染色体;29021106Kb表示日本晴11号染色体总长度;29113644Kb表示合系35中11号染色体总长度。

[0023] 图2为元江普通野生稻5与日本晴11号染色体共线性图。Chr11表示11号染色体;29021106Kb表示日本晴11号染色体总长度;31935697Kb表示元江普通野生稻11号染色体总长度。

[0024] 图3为214QC-9F/R在元江普通野生稻渗入系G252中的扩增电泳图。M为DL5000的Marker,由5000bp、3000bp、2000bp、1000bp、750bp、500bp、250bp和100bp共8条链组成。

[0025] 图4为Xa47(t)基因部分区域测序峰图。

[0026] 图5为Xa47(t)<sup>YP</sup>、Xa47(t)<sup>G252</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>4个成员的蛋白结构域。

[0027] 图6为Xa47(t)<sup>L234</sup>成员与Xa47(t)<sup>YP</sup>和Xa47(t)<sup>HX</sup>成员的比对图;其中,Xa47\_HX表示Xa47(t)<sup>HX</sup>成员,Xa47\_YP表示Xa47(t)<sup>YP</sup>成员,Xa47\_L234表示Xa47(t)<sup>L234</sup>成员。

[0028] 图7为Xa47(t)<sup>G252</sup>成员与Xa47(t)<sup>YP</sup>和Xa47(t)<sup>HX</sup>成员的比对图;其中,Xa47\_HX表示Xa47(t)<sup>HX</sup>成员,Xa47\_YP表示Xa47(t)<sup>YP</sup>成员,Xa47\_G252表示Xa47(t)<sup>G252</sup>成员。

### 具体实施方式

[0029] 本发明提供了一种高效分离带有野生稻血缘的抗白叶枯病基因的方法,所述方法包括:

[0030] 通过对元江普通野生稻渗入系双亲进行建库和高通量测序获得连续基因组,将所述连续基因组与参考基因组进行共线性分析,根据目的基因定位情况,筛选出所述抗白叶枯病基因的候选基因,根据所述候选基因设计引物进行PCR扩增、PCR产物直接测序和抗病功能分析,从而分离出带有野生稻血缘的抗白叶枯病基因。

[0031] 本发明通过对元江普通野生稻渗入系双亲进行建库和高通量测序获得连续基因组。本发明中,所述建库和高通量测序之前优选的包括对元江普通野生稻渗入系双亲的DNA提取;所述提取的方法优选为CTAB法;所述元江普通野生稻渗入系双亲优选为元江普通野生稻和栽培稻合系35。本发明中,所述建库和高通量测序优选的包括PacBio Sequel\_SMRT平台、Dovetail Hi-C建库技术和MECAT组装系统。

[0032] 本发明将所述连续基因组与参考基因组进行共线性分析,筛选出所述抗白叶枯病基因的候选基因。本发明中,所述参考基因组为日本晴基因组;所述共线性分析优选的采用MCSanX软件进行。本发明中,所述筛选的方法优选包括:利用Soft Berry (<http://linux1.softberry.com/>)中的FGENESH程序对R13I14和Hxjy-14两个标记之间的序列进行基因注释分析,在Smart数据库 (<http://smart.embl-heidelberg.de/>) 预测注释基因的功能,根据注释的功能筛选出Xa47(t) 候选基因。本发明中,所述抗白叶枯病基因为Xa47(t);所述基因的核苷酸序列如SEQ ID NO.13所示;所述基因Xa47(t)的CDS序列如SEQ ID NO.14所示。

[0033] 本发明根据所述候选基因设计引物进行PCR扩增、PCR产物直接测序和抗病功能分

析,从而分离出带有野生稻血缘的抗白叶枯病基因。本发明中,所述目的基因定位和所述PCR扩增的模板优选为元江普通野生稻渗入系G252。所述抗病功能分析优选的包括功能互补试验和基因编辑试验分析候选基因的抗病功能。本发明中,所述引物优选为214QC-9F/R;所述214QC-9F的核苷酸序列如SEQ ID NO.11所示,所述214QC-9R的核苷酸序列如SEQ ID NO.12所示。本发明中,所述测序的平台选择本领域常规测序平台即可。在本发明的实施例中,所述测序优选由华大生物科技(武汉)有限公司完成,PCR扩增产物测序优选由北京擎科生物技术有限公司昆明分公司完成。

[0034] 本发明还提供了一种高效分离带有野生稻血缘的抗白叶枯病基因家族成员的方法,所述方法包括:

[0035] 筛选获得含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料;根据上述Xa47(t)基因序列,获得来自元江普通野生稻和合系35中的Xa47(t)基因,分别命名为Xa47(t)<sup>YP</sup>和Xa47(t)<sup>HX</sup>;

[0036] 利用引物对所述元江普通野生稻渗入系材料进行扩增、测序,得到4种Xa47(t)基因的基因型:Xa47(t)<sup>YP</sup>、Xa47(t)<sup>G252</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>;对由4种基因型编码得到的蛋白质进行相似性分析,从而确定其是否为带有野生稻血缘的抗白叶枯病基因家族成员。

[0037] 本发明中,所述筛选的步骤包括:利用Xa47(t)基因的共分离标记Hxjy-1对元江普通野生稻渗入系进行PCR扩增,得到扩增片段大小为167bp的渗入系,结合渗入系的抗病谱筛选出含有Xa47(t)基因或其同源基因的元江普通野生稻渗入系材料。本发明中,根据所述Xa47(t)基因序列,利用blast程序从元江普通野生稻和合系35基因组R13I14和Hxjy-14区段中进行检索,即获得来自元江普通野生稻和合系35中的Xa47(t)基因。本发明中,所述引物包括214QC-2F/R、48QC-9F/R和48QC-11F/R,所述214QC-2F的核苷酸序列如SEQ ID NO.17所示,所述214QC-2R的核苷酸序列如SEQ ID NO.18所示;所述48QC-9F的核苷酸序列如SEQ ID NO.19所示,所述48QC-9R的核苷酸序列如SEQ ID NO.20所示;所述48QC-11F的核苷酸序列如SEQ ID NO.21所示,所述48QC-11R的核苷酸序列如SEQ ID NO.22所示。

[0038] 本发明中,所述Xa47(t)<sup>G252</sup>、Xa47(t)<sup>YP</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>基因即以其供体材料进行命名的Xa47(t)基因;所述Xa47(t)<sup>G252</sup>为上述的Xa47(t)基因;所述Xa47(t)<sup>YP</sup>基因的核苷酸序列如SEQ ID NO.15所示;所述Xa47(t)<sup>YP</sup>基因的CDS序列如SEQ ID NO.30所示;所述Xa47(t)<sup>HX</sup>基因的核苷酸序列如SEQ ID NO.16所示;所述Xa47(t)<sup>HX</sup>基因的CDS序列如SEQ ID NO.32所示;所述Xa47(t)<sup>L234</sup>基因的核苷酸序列如SEQ ID NO.27所示;所述Xa47(t)<sup>L234</sup>基因的CDS序列如SEQ ID NO.28所示。

[0039] 本发明中,所述的栽培稻合系35为我国选育的高产优质粳稻品种,公布于文献(殷富有,张敦宇,叶玉,等.普通野生稻栽培稻杂交后代白叶枯病抗性评价.江西农业学报,2010,22(8):81-84.)。

[0040] 所述的HZhj19、PX099、PB、ScYc-b、T7147、Y8、YM1、YM187、YN24、YJDP-2和YJWS-2白叶枯病菌在非专利文献公开,申请人有保存,自本专利申请日起20年内可提供。

[0041] 所述的HZhj19、YM1、YM187、YN24、YJDP-2和YJWS-2菌株为本实验室在2013~2017年从自然发病的病叶上分离而来,公布于文献(陈玲,张敦宇,陈越,等.云南药用野生稻种质资源的白叶枯病抗性评价.南方农业学报,2019,50(7):1417-1425.)。

[0042] 所述的PX099菌株为菲律宾标准菌株6,采集自菲律宾,公布于文献(Ji CH, Ji ZY,



Liu B, et al. Xa1 allelic R genes activate rice blight resistance suppressed by interfering TAL effectors. *Plant Comm*, 2020, 1(4), 100087.)。

[0043] 所述的PB菌株,为PX099菌株的突变株系,公布于文献(陈玲,张敦宇,陈越,等.云南药用野生稻种质资源的白叶枯病抗性评价.南方农业学报,2019,50(7):1417-1425.)。

[0044] 所述的Y8菌株为云南强致病型生理小种,采集自中国云南,公布于文献(殷富有,张敦宇,叶玉,等.普通野生稻栽培稻杂交后代白叶枯病抗性评价.江西农业学报,2010,22(8):81-84.)。

[0045] 所述的ScYc-b(中国标准菌株5号)和YN24(中国标准菌株9号)菌株,采集自中国东北稻区,公布于文献(吴宪,许晶,温嘉伟,等.东北水稻白叶枯病菌株遗传多样性分析及品种对白叶枯病抗性评价.吉林农业大学学报,2015,37(3):290-295.)。

[0046] 所述的T7147(日本标准菌株2号),采集自日本,公布于文献(周永力,翟文学,章琦,等.Xa21转基因水稻对白叶枯病的抗性及其遗传.植物病理学报,2001,31(2):123-129.)

[0047] 所述的栽培稻JG30,是我国选育的矮秆中晚熟籼稻(栽培稻),公布于文献(金旭伟,王春连,杨清,等.水稻抗白叶枯病近等基因系CBB30的培育及Xa30(t)的初步定位.中国农业科学,2007,40(6):1094-1100.)。

[0048] 为使本发明的目的、技术方案和优点更加清楚明白,下面结合实施例对本发明进行详细的说明,但是不能把它们理解为对本发明保护范围的限定。

[0049] 下述实施例中,为了避免因不必要的细节而模糊了本发明专利的技术方案,在实施例中仅出示了与本发明专利方案密切相关的技术方案和/或处理步骤,如无特殊说明,省略的其他细节方法均为常规方法。

[0050] 下述实施例中所用的材料、试剂等,如无特殊说明,均可从商业途径得到。

[0051] 实施例1元江普通野生稻和栽培稻合系35基因组高通量测序

[0052] 1.元江普通野生稻和栽培稻合系35的基因组DNA提取

[0053] 基于CTAB方法提取基因组DNA:采取样品的幼嫩叶片约200mg,用液氮研磨成粉状,将组织转移到预热65℃的2.0mL管中,加入900μL 2%CTAB裂解缓冲液,涡旋混合。将离心管在65℃下孵育60分钟,然后在室温(RT)下以10,000rpm离心5分钟。用900μL体积的苯酚/氯仿/异戊醇(25:24:1)提取上清液,重复2次,然后在新管中以10,000rpm的速度在室温下离心10分钟。加入2/3体积的预冷(-20℃)异丙醇,-20℃放置2小时以上沉淀DNA,室温12000rpm离心15分钟。加入75%无水乙醇洗涤沉淀并离心去除,然后将DNA沉淀风干3-5分钟。用200μL ddH<sub>2</sub>O溶解沉淀后进行后续实验。

[0054] 所述2%CTAB裂解缓冲液的配制方法为:取4g十六烷基三甲基溴化铵(CTAB),16.364gNaCl,1M Tris-HCl 20mL(pH8.0),0.5M EDTA8mL,先用70mL ddH<sub>2</sub>O溶解,再定容至200mL灭菌备用。

[0055] 2.元江普通野生稻和栽培稻合系35染色体级别的连续基因组获取

[0056] 将步骤1提取得到的DNA送华大生物科技(武汉)有限公司,通过PacBio Sequel<sub>2</sub> SMRT平台进行建库和全基因组测序,结合Dovetail Hi-C建库技术以及MECAT组装系统,获得元江普通野生稻和栽培稻合系35染色体级别的连续基因组。

[0057] 实施例2Xa47(t)的克隆

[0058] 1.Xa47(t) 候选基因的预测

[0059] 选用MCSanX软件将实施例1中元江普通野生稻和栽培稻合系35的11号染色体的连续基因组与高质量参考基因组即日本晴基因组进行共线性分析。随后利用Soft Berry (<http://linux1.softberry.com/>) 中的FGENESH程序对R13I14和Hxjy-14两个标记之间的序列进行基因注释分析,在Smart数据库 (<http://smart.embl-heidelberg.de/>) 预测注释基因的功能,根据注释的功能筛选出Xa47(t) 候选基因。

[0060] 根据Xa47(t) 基因位于分子标记R13I14和Hxjy-14之间的事实,通过共线性分析得到图1和图2。从图1和图2可看出,Xa47(t) 基因的此定位区段在元江普通野生稻、合系35和日本晴之间排列顺序比较一致,表明可参照目前研究较为清楚的日本晴基因组从该区段进行Xa47(t) 基因的预测。进而根据上述步骤从该区域注释到3个候选基因,其中一个是含有NBS-LRR结构域的LOC\_0s11g46200基因,其余两个基因与转座相关,与抗病不相关,因此将LOC\_0s11g46200作为Xa47(t) 基因的候选基因。

[0061] 2.Xa47(t) 候选基因的克隆

[0062] 根据步骤1得到的候选基因的编码区上下游区域,设计10对特异引物,采用CTAB法提取Xa47(t) 的供体亲本G252的基因组DNA,利用PCR进行扩增,取7 $\mu$ LPCR扩增产物在1%琼脂糖凝胶上5V/cm恒压电泳,凝胶成像系统成像后,根据PCR产物,筛选出一个特异引物214QC-9F/R,其中,214QC-9F的核苷酸序列如SEQ ID NO.11所示,214QC-9R的核苷酸序列如SEQ ID NO.12所示,该引物的扩增片段大小正确且带型清晰,结果如图3所示。将该PCR产物直接送北京擎科生物技术有限公司昆明分公司进行测序,测序结果如图4所示。由图4可见,测序峰图信号强,峰型完整且为单峰,表明测序结果较好。

[0063] 所述的PCR扩增采用45 $\mu$ LPCR反应体系进行;45 $\mu$ LPCR反应体系为:2 $\times$ Phanta<sup>®</sup>Max Master Mix 24 $\mu$ L,10 $\mu$ mol/L上、下游引物各2.4 $\mu$ L,20ng/ $\mu$ L模板DNA 6 $\mu$ L,ddH<sub>2</sub>O 10.2 $\mu$ L;PCR反应条件为:94 $^{\circ}$ C预变性3min;94 $^{\circ}$ C变性15sec,55 $^{\circ}$ C复性15sec,68 $^{\circ}$ C延伸5min,共35个循环;68 $^{\circ}$ C延伸10min。

[0064] 测序结果显示,Xa47(t) 候选基因的核苷酸(DNA) 长度为4240bp,编码区(CDS) 序列为2409bp,编码802个氨基酸,具体核苷酸序列见序列表SEQ ID NO.13,具体CDS序列见序列表SEQ ID NO.13编码获得的氨基酸序列如SEQ ID NO.34所示。

[0065] 实施例3候选基因功能分析

[0066] 1. 功能互补试验验证候选基因功能

[0067] 提取Xa47(t) 的供体亲本G252的总RNA,反转录为cDNA后,参照Xa47(t) 候选基因编码区序列,利用软件BioXM2.6从Xa47(t) 的DNA序列中预测出其编码区(CDS区),在预测出的编码区两端设计特异引物G252-CDS-F/R,利用G252的cDNA进行PCR扩增,将预期大小正确的片段与载体pCE2连接,将连接产物送北京擎科生物技术有限公司昆明分公司进行测序,获得Xa47(t) 编码区序列,同时获得pCE2-Xa47CDS质粒。

[0068] 选择pCamBIA1305质粒作为植物表达载体骨架,用XbaI和BstEII对其进行双酶切,利用引物Ubi-F/R扩增质粒pJET-Ubi中的Ubiquitin启动子,利用Xa47-OE-F/R引物扩增质粒pCE2-Xa47CDS中的Xa47(t) CDS序列,切胶回收目的片段后,通过同源重组的方法构建Ubi-Xa47(t) 过表达载体,通过农杆菌介导法转化感白叶枯病的栽培稻JG30材料。本实施例所采用的引物序列如下表1所示。

[0069] 表1引物序列

引物对	引物名称	编号
G252-CDS-F/R	G252-CDS-F	SEQ ID NO.1
	G252-CDS-R	SEQ ID NO.2
Ubi-F/R	Ubi-F	SEQ ID NO.3
	Ubi-R	SEQ ID NO.4
Xa47-OE-F/R	Xa47-OE-F	SEQ ID NO.5
	Xa47-OE-R	SEQ ID NO.6

[0071] 对上述转化栽培稻JG30材料后的 $T_1$ 代阳性转基因苗进行孕穗期接种HZhj19、PX099、PB、ScYc-b、T7147、Y8、YM1、YM187、YN24、YJDP-2和YJWS-2共11个白叶枯病强致病菌，每个菌株接种3株转化苗，同时以感病材料金刚30作为对照，接种21天后调查病斑长度，结果如表2所示。其中，病斑长度的单位为cm。

[0072] 表2  $T_1$ 代遗传转化苗Ubi-Xa47(t)表型鉴定

供体材料	白叶枯病菌										
	HZhj1	PX09	PB	ScYc-b	T7147	Y8	YM1	YM18	YN24	YJDP	YJWS
	9	9		b				7		-2	-2

[0074]

JG30 (感病 对照, 转基因 受体)	16.00 ±1.87	15.13 ±2.46	17.60 ±3.44	13.70 ±2.38	18.75 ±0.25	14.38 ±4.13	16.00 ±4.37	16.17 ±0.94	14.58 ±2.68	12.75 ±2.75	13.50 ±2.12
G252 (抗病 对照, 转基因 供体)	3.07± 1.46	5.00± 0.84	2.50± 1.08	0.87± 0.19	3.83± 1.52	0.93± 0.09	4.83± 1.18	2.10± 0.14	2.40± 0.58	3.88± 1.84	4.00± 1.58
Ubi-Xa 47(t)-1 (阳性 转化 苗)	5.95± 0.85	5.83± 0.11	3.35± 0.65	5.10± 0.40	3.27± 0.66	3.15± 0.35	5.50± 2.00	4.06± 1.67	3.06± 1.53	6.50± 2.00	6.85± 1.25
Ubi-Xa 47(t)-2 (阳性 转化 苗)	2.65± 0.25	5.40± 0.10	2.70± 0.50	1.70± 0.82	3.33± 0.24	4.60± 1.60	4.40± 1.60	2.50± 1.22	3.73± 1.23	6.25± 3.25	6.00± 0.50
Ubi-Xa 47(t)-3 (阳性 转化 苗)	2.25± 0.85	5.87± 0.74	5.20± 2.00	3.13± 1.07	4.35± 2.85	1.85± 0.65	1.90± 0.30	5.67± 1.89	3.95± 1.05	5.18± 0.97	3.65± 2.85
Ubi-Xa 47(t)-4 (阳性 转化 苗)	1.60± 1.28	5.10± 0.90	2.17± 1.31	4.48± 2.77	5.25± 3.95	3.75± 0.75	3.25± 0.45	3.40± 0.60	3.03± 2.52	2.70± 3.37	4.75± 4.25
Ubi-Xa 47(t)-5 (阳性 转化 苗)	5.53± 1.45	5.75± 0.25	4.45± 0.25	4.50± 2.00	3.15± 1.15	3.05± 0.65	5.10± 1.36	4.10± 0.10	1.45± 0.05	2.37± 2.01	5.55± 3.95
Ubi-Xa 47(t)-6 (阳性 转化 苗)	6.15± 4.55	5.85± 0.65	6.85± 5.15	2.60± 0.50	5.65± 1.65	1.65± 1.35	5.00± 1.00	2.00± 0.30	3.00± 1.00	3.30± 2.81	3.75± 0.75
Ubi-Xa 47(t)-7 (阳性 转化)	3.30± 1.50	5.12± 2.75	4.35± 2.85	2.50± 0.50	3.25± 1.25	2.85± 0.75	1.35± 0.65	1.25± 0.75	2.00± 0.90	5.85± 2.55	3.70± 1.80

[0075]	苗)											
	Ubi-Xa 47(t) (阴性 转化 苗)	11.08 ±3.45	12.18 ±1.67	10.88 ±0.54	10.75 ±0.25	10.78 ±2.76	10.80 ±1.06	14.33 ±2.25	16.46 ±2.31	12.00 ±0.70	12.2± 2.23	11.55 ±1.61

[0076] 由表2可以看出,感病材料金刚30病斑在12cm以上,表现较强的感病特性,转化苗的病斑在5cm以下,对11个菌株达到抗到高抗的水平,说明遗传转化株系的白叶枯病抗性是由Xa47(t)基因提供,该基因具有很强的白叶枯病抗性,对于抗白叶枯病水稻育种具有重要应用价值。

[0077] 2. 基因编辑实验评估候选基因功能

[0078] 根据CRISPR-P2.0在线软件(<http://crispr.hzau.edu.cn/CRISPR2>)在Xa47(t)整个DNA序列(正反链)里找出含有GN20GG或N20GG的PAM序列,选择中靶效率评分大于0.5的特异性较高的作为sgRNA候选序列,通过查找共找到3个sgRNA序列,将3个sgRNA序列进行体外验证。将体外验证较好的sgRNA-2的PAM前面20bp的序列作为Spacer序列,分别设计上、下游sgRNA的Oligo序列sgF/R,并在sgF的5'端加入同源臂GGCA,sgR的5'端加入同源臂AAAC,送北京擎科生物技术有限公司昆明分公司合成。

[0079] 利用合成的Oligo序列制备两端具有粘性末端的双链DNA片段(Oligo-Xa47(t))。利用CIP酶切10ug p0s-sgRNA载体,制作出线性化p0s-sgRNA载体。将线性化p0s-sgRNA载体与Oligo-Xa47(t)相连,转化大肠杆菌,获得阳性克隆,构建出p0s-sgRNA-Oligo-Xa47(t)载体。利用LR Clonase™ II enzyme mix将p0s-sgRNA-Oligo-Xa47(t)质粒与pH-Ubi-cas9-7质粒进行LR反应,最终获得CRISPR/Cas9-Xa47(t)载体,并用引物Seq-U3和Seq-Cas对CRISPR/Cas9-Xa47(t)载体进行测序来验证其正确性,测序公司为北京擎科生物技术有限公司昆明分公司。将验证正确的载体通过农杆菌介导法转化Xa47(t)的供体亲本G252。本实施例所用引物序列如表3所示。

[0080] 表3引物序列

[0081]	引物对	引物名称	编号
[0082]	sgF/R	sg-F	SEQ ID NO.7
		sg-R	SEQ ID NO.8
	Seq-U3	Seq-U3	SEQ ID NO.9
	Seq-Cas	Seq-Cas	SEQ ID NO.10

[0083] 对转化后的T<sub>1</sub>代阳性基因敲除苗进行孕穗期接种上述11个白叶枯病菌株,供体亲本G252作为对照,结果显示G252对11个菌株均抗病,敲除苗对11个菌株都感病,该结果与过表达植株的表型相对应,表明本发明所述基因Xa47(t)抗不同的白叶枯病菌,具有广谱高抗的特性。

[0084] 实施例4 Xa47(t)家族成员的分离

[0085] 本实施例所述的PCR扩增程序和体系与实施例1相同。

[0086] 1.Xa47(t) 基因家族成员的供体亲本筛选

[0087] 利用Xa47(t) 的共分离标记Hxjy-1在359份元江普通野生稻渗入系中进行PCR扩增,筛选出扩增片段大小为167bp的渗入系有262份,且这262份渗入系抗病谱依然为67种类型,为了获得对白叶枯病菌具有广谱高抗的抗性基因,本实施例从262份渗入系中筛选出抗6-9个白叶枯病菌的渗入系,这些渗入系呈现出16个抗病谱,从16个抗病谱中选出1~3份代表材料用于Xa47(t) 基因家族成员的分离,具体的材料抗病谱如表4。

[0088] 表4部分渗入系抗病谱

渗入系编号	抗谱类型	白叶枯病菌								
		Y8	YM1	ScYc-b	YN24	PXO99 <sup>A</sup>	PB	T7147	YM187	HZhj19
[0089] L96	抗谱 1	R	R	R	S	S	S	R	R	R
L27	抗谱 2	R	S	R	R	S	S	R	R	R
L50		R	S	R	R	S	S	R	R	R
L174		R	S	R	R	S	S	R	R	R
L43	抗谱 3	R	S	R	R	R	S	R	R	R

[0090]

L39	抗 谱 4	R	R	R	R	S	S	R	S	R
L252		R	R	R	R	S	S	R	S	R
P104		R	R	R	R	S	S	R	S	R
L171	抗 谱 5	R	R	R	R	S	S	R	R	R
L179		R	R	R	R	S	S	R	R	R
L173		R	R	R	R	S	S	R	R	R
L309	抗 谱 6	R	S	S	S	S	R	R	R	R
L334		R	S	S	S	S	R	R	R	R
L40	抗 谱 7	R	R	R	S	S	R	R	R	R
L7		R	R	R	S	S	R	R	R	R
L128		R	R	R	S	S	R	R	R	R
L23	抗 谱 8	R	R	R	S	S	R	R	R	S
L150	抗 谱 9	R	R	S	S	S	R	R	R	R
L214	抗 谱 10	R	R	R	R	R	R	R	R	R
P165		R	R	R	R	R	R	R	R	R
L41		R	R	R	R	R	R	R	R	R
L186	抗 谱 11	R	R	R	R	R	S	R	R	R
L279		R	R	R	R	R	S	R	R	R
L126		R	R	R	R	R	S	R	R	R
L33	抗 谱 12	R	R	R	R	S	R	R	R	S
L113	抗 谱 13	R	R	R	R	S	R	R	R	R
L119		R	R	R	R	S	R	R	R	R
L275		R	R	R	R	S	R	R	R	R
L61	抗 谱 14	R	S	R	R	S	R	R	R	R
L156		R	S	R	R	S	R	R	R	R
L215		R	S	R	R	S	R	R	R	R
L114	抗 谱 15	R	S	R	R	S	R	R	S	R
L17	抗	S	S	S	R	R	R	R	R	R

[0091]		谱								
		16								

[0092] 2. 元江普通野生稻和合系35基因的Xa47(t)基因获取

[0093] 根据Xa47(t)基因序列,利用blast程序从元江普通野生稻和合系35基因组R13I14和Hxjy-14区段中进行检索,获得来自元江普通野生稻和合系35基因的Xa47(t)基因。得到的Xa47(t)<sup>YP</sup>基因序列如SEQ ID NO.15所示,Xa47(t)<sup>HX</sup>基因如SEQ ID NO.16所示。利用DNAMAN软件比对发现,Xa47(t)<sup>HX</sup>基因与LOC\_0s11g46200的编码区完全一致,相似性为100%。

[0094] 3. Xa47(t)基因家族成员的基因分离

[0095] 根据Xa47(t)<sup>YP</sup>和Xa47(t)<sup>G252</sup>基因的编码区上下游序列,设计特异引物,以及利用之前参照LOC\_0s11g46200编码区上下游设计的10对特异引物,通过PCR技术在元江普通野生稻渗入系DNA中进行扩增,取7 $\mu$ LPCR扩增产物在1%琼脂糖凝胶上5V/cm恒压电泳,凝胶成像系统成像后,根据跑胶结果,共筛选出4对可从元江普通野生稻渗入系DNA中扩增的引物,引物序列如下表5,选取片段大小正确的PCR产物直接送北京擎科生物技术有限公司昆明分公司进行测序。

[0096] 表5引物序列

引物对	引物名称	编号
214QC-2-F/R	214QC-2F	SEQ ID NO.17
	214QC-2R	SEQ ID NO.18
214QC-9-F/R	214QC-9F	SEQ ID NO.11
	214QC-9R	SEQ ID NO.12
48QC-9-F/R	48QC-9F	SEQ ID NO.19
	48QC-9R	SEQ ID NO.20
48QC-11-F/R	48QC-11F	SEQ ID NO.21
	48QC-11R	SEQ ID NO.22

[0097] 根据测序结果发现渗入系中出现了4种基因型,包括Xa47(t)<sup>YP</sup>、Xa47(t)<sup>G252</sup>、Xa47(t)<sup>HX</sup>以及Xa47(t)<sup>L234</sup>基因型,通过比对,Xa47(t)<sup>HX</sup>与日本晴中的LOC\_0s11g46200基因完全一致,而LOC\_0s11g46200基因的CDS区域已经明确。Xa47(t)<sup>L234</sup>基因的核苷酸序列如SEQ ID NO.27所示。利用软件BioXM2.6从Xa47(t)<sup>YP</sup>和Xa47(t)<sup>L234</sup>基因型的DNA序列中预测出其CDS区,在预测出的编码区两端设计特异引物对YP-CDS-F/R(其中YP-CDS-F的核苷酸序列如SEQ ID NO.23所示,YP-CDS-R的核苷酸序列如SEQ ID NO.24所示)和L234-CDS-F/R(其中L234-CDS-F的核苷酸序列如SEQ ID NO.25所示,L234-CDS-R的核苷酸序列如SEQ ID NO.26所示),这两对引物分别在元江普通野生稻和渗入系L234的cDNA中进行扩增,将预期大小正确的片段与载体pCE2连接,将连接产物进行测序,获得Xa47(t)<sup>YP</sup>和Xa47(t)<sup>L234</sup>编码区序列。根据4个成员的CDS序列通过BioXM2.6软件将其分别翻译成蛋白质序列,具体的序列如表6所



示。继续利用DNAMAN软件分析4个蛋白之间的相似性,结果如表7所示。进一步利用在线软件SMART预测4个蛋白结构,结果如图5所示。

[0099] 表6不同成员的序列

	基因型	CDS 序列	编码蛋白质的氨基酸序列
[0100]	Xa47(t) <sup>L234</sup>	SEQ ID NO. 28	SEQ ID NO.29
	Xa47(t) <sup>YP</sup>	SEQ ID NO. 30	SEQ ID NO.31
	Xa47(t) <sup>HX</sup>	SEQ ID NO. 32	SEQ ID NO.33
	Xa47(t) <sup>G252</sup>	SEQ ID NO. 14	SEQ ID NO.34

[0101] 表7不同基因型蛋白相似性分析

	Xa47 (t) <sup>HX</sup>	Xa47 (t) <sup>YP</sup>	Xa47 (t) <sup>G252</sup>	Xa47 (t) <sup>L234</sup>
Xa47 (t) <sup>HX</sup>	100%	100%	/	/
Xa47 (t) <sup>YP</sup>	95.2%	100%	/	/
Xa47 (t) <sup>G252</sup>	89.0%	91.9%	100%	/
Xa47 (t) <sup>L234</sup>	94.8%	99.0%	92.1%	100%

[0103] 从表7可见,4个蛋白之间相似性较高,说明它们极有可能互为同一个家族成员。从图5可见,4个蛋白结构域基本一致,综合说明4个基因型同为一个基因家族。那4个家族成员中,因Xa47 (t)<sup>YP</sup>和Xa47 (t)<sup>HX</sup>分别来自渗入系的双亲,说明Xa47 (t)<sup>YP</sup>是完全带有野生稻血缘的基因型,Xa47 (t)<sup>HX</sup>是来自栽培稻合系35。

[0104] 随后利用DNAMAN软件将Xa47 (t)<sup>G252</sup>和Xa47 (t)<sup>L234</sup>蛋白分别与Xa47 (t)<sup>YP</sup>和Xa47 (t)<sup>HX</sup>进行多重比对,确定Xa47 (t)<sup>G252</sup>和Xa47 (t)<sup>L234</sup>是否带有野生稻血缘,结果如图6和图7所示。

[0105] 可以看出,相对于Xa47 (t)<sup>YP</sup>和Xa47 (t)<sup>HX</sup>蛋白,Xa47 (t)<sup>G252</sup>和Xa47 (t)<sup>L234</sup>蛋白已发生变异,其中Xa47 (t)<sup>L234</sup>蛋白变异较Xa47 (t)<sup>G252</sup>小,但2个成员的大部分蛋白区域还是与Xa47 (t)<sup>YP</sup>一致,说明Xa47 (t)<sup>G252</sup>和Xa47 (t)<sup>L234</sup>基因型携带了野生稻血缘。

[0106] 本实施例从元江普通野生稻渗入系中获得了4个Xa47 (t) 家族成员,其中的3个成员都带有野生稻血缘,是具有利用价值的基因资源。所述的基因家族指来源于同一个祖先,由一个基因通过基因重复而产生两个或更多的拷贝而构成的一组基因,它们在结构和功能上具有明显的相似性,编码相似的蛋白质产物。

[0107] 4. 获得的家族成员在元江普通野生稻渗入系中的分布情况

[0108] 根据上述结果分析4个成员在元江普通野生稻渗入系中分布情况,结果如表8所示。

[0109] 表8 Xa47 (t) 基因不同家族成员在元江普通野生稻中的分布情况

渗入系编号	家族成员	抗谱类型
L96	Xa47(t) <sup>G252</sup>	抗谱 1
L27	Xa47(t) <sup>G252</sup> 和 Xa47(t) <sup>HX</sup>	抗谱 2
L50	Xa47(t) <sup>G252</sup>	
L174	Xa47(t) <sup>YP</sup> 和 Xa47(t) <sup>L234</sup>	
L43	Xa47(t) <sup>L234</sup>	抗谱 3
L39	Xa47(t) <sup>L234</sup>	抗谱 4
L252	Xa47(t) <sup>G252</sup>	
P104	Xa47(t) <sup>G252</sup>	
L171	Xa47(t) <sup>G252</sup>	抗谱 5
L179	Xa47(t) <sup>G252</sup>	
L173	Xa47(t) <sup>G252</sup>	
L309	Xa47(t) <sup>YP</sup>	抗谱 6
L334	Xa47(t) <sup>G252</sup>	
L40	Xa47(t) <sup>L234</sup>	抗谱 7
L7	Xa47(t) <sup>G252</sup> 和 Xa47(t) <sup>HX</sup>	
L128	Xa47(t) <sup>G252</sup>	
L23	Xa47(t) <sup>HX</sup>	抗谱 8
L150	Xa47(t) <sup>L234</sup>	抗谱 9
L214	Xa47(t) <sup>G252</sup>	抗谱 10
P165	Xa47(t) <sup>YP</sup>	
L41	Xa47(t) <sup>HX</sup>	
L186	Xa47(t) <sup>G252</sup>	抗谱 11
L279	Xa47(t) <sup>G252</sup>	
L126	Xa47(t) <sup>G252</sup>	
L33	Xa47(t) <sup>HX</sup>	抗谱 12
L113	Xa47(t) <sup>G252</sup>	抗谱 13
L119	Xa47(t) <sup>G252</sup>	
L275	Xa47(t) <sup>G252</sup>	
L61	Xa47(t) <sup>G252</sup>	抗谱 14
L156	Xa47(t) <sup>G252</sup>	
L215	Xa47(t) <sup>YP</sup>	
L114	Xa47(t) <sup>G252</sup>	抗谱 15
L17	Xa47(t) <sup>G252</sup> 和 Xa47(t) <sup>HX</sup>	抗谱 16

[0112] 可以看出,这些家族成员在渗入系中以纯合或杂合状态存在,表明元江普通野生稻渗入系中的Xa47(t)具有多样化,即存在一定变异也存在一定的保守型,它们互为基因家族,表明利用本发明所述方法还可从其余的渗入系材料中分离出新的家族成员。

[0113] 以上所述仅为本发明的实施例,并非因此限制本发明的专利范围,凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换,或直接或间接运用在其他相关的技术领域,均同理包括在本发明的专利保护范围内。

## 序列表

<110>	云南省农业科学院生物技术与种质资源研究所	
<120>	一种高效分离带有野生稻血缘的抗白叶枯病基因及其家族成员的方法	
<160>	34	
<170>	SIPOSequenceListing 1.0	
<210>	1	
<211>	18	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	1	
	atgacggggg aggaagtc	18
<210>	2	
<211>	19	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	2	
	atgtatgatg catgtcacg	19
<210>	3	
<211>	37	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	3	
	ggtacccggg gatcctctag actgcagtgc agcgtct	37
<210>	4	
<211>	31	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	4	
	ccccggtcat ctgcagaagt aacaccaaac a	31
<210>	5	
<211>	26	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	5	
	acttctgcag atgacggggg aggaag	26
<210>	6	
<211>	50	
<212>	DNA	

<213> 人工序列 (Artificial Sequence)	
<400> 6	
gggaaattcg agctggtcac ctaatacca tatacatgta tgatgcatgt	50
<210> 7	
<211> 24	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 7	
ggcacaaggt gccggaaaaa agaa	24
<210> 8	
<211> 24	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 8	
aaacttcttt tttccggcac cttg	24
<210> 9	
<211> 20	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 9	
taccacctcg gctatccaca	20
<210> 10	
<211> 19	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 10	
gacaagggca gggatttcg	19
<210> 11	
<211> 20	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 11	
tttagctgct ctaagttggt	20
<210> 12	
<211> 20	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 12	
agtcattagt ttcgaacggg	20

<210>	13	
<211>	4240	
<212>	DNA	
<213>	人工序列(Artificial Sequence)	
<400>	13	
atgacggggg	aggaagtcga	tgctttgtgc aaggatgagt tgatggcggg ggtgcgtgag 60
ctgtcctacg	acatggacga	cgccatcgac gaattcttct tagaggagcc catggcgggc 120
ggcgacgggtg	gccctttcga	tgagctcaag acaagagttg aggatgtctc caagcggttc 180
tccgacagcc	ggcgggtggag	gccacaggtg gagcaacatc aaccatccct aaccgccgca 240
accgtagact	gtccacctcc	tcacgctcgc ttcgtccaca acatgatgga tgtgtcagag 300
ctcgtggaga	tggacaaaact	acatgagaca gagctcatca aattgctgga acaaggtgcg 360
gacacaagca	tatatgettc	ccggtggcgc atcgcaacac catggcatga taaggaggta 420
aagacgacat	ccttttattt	cttttttate tctacttctc tttttatata ttatattata 480
aaaatttaaa	atgttttcgc	tgtgtgtatt ttggtacgtc gtggctatct tcgtatcgta 540
ttcgatctcc	cgttcagtat	gttttgtgtt gtacgtctct agcttcaga tatatcatat 600
atctttccat	tcctatgta	ttctttcttt ccaaattcca attattaatt atacctcatt 660
taatgagggg	catttatacc	attttatttc agtagtaatt tcctcctttc tcctagtgtc 720
agaagtgcac	ttgcatgagg	atggaggcaa agctgaaata cccattcgtt gtgatgaaat 780
tttaaaagcg	aactctaccg	taaactctaa attaagcccc aaaaattata atgcaagtat 840
tcctgtttta	atcaatattt	ggcatcattt tttttacaat gtataattca gctacacaat 900
tttctcattt	ttttttacaa	ttgcttataa ttaagctaca caagctacca attcagctac 960
acaattggta	catactacta	tactagccaa ataccctgta tttgctacgt attaaaacaa 1020
attaatagtg	agatttttgg	ggcaattaat ttggttttgt agaagttata tcatagagaa 1080
attattttat	aacggtaaag	ttggttgaat atatgatgga taatgtggta aataaaaaaa 1140
gtactatagt	tgggtggcaga	ttcactgcca ttgccctctt ttgaaaggaa tatagaattt 1200
tatctttag	aagtatatt	gtacaagtga aatatgatgg aatcatatat gtagaataaa 1260
acattaaagt	atgtgggggt	atttggttga aatatgatg gagtatgtgg taaatagaaa 1320
aaaaatacta	tacttgatga	tggggcgatg atagattcac tgccaccacc attgcatttt 1380
ttttaaaagg	agtatataaa	catatatagg aacctatcta ctctactt ggtaagaggt 1440
cttacttgg	aattgtgctg	gacggtagca tgccagttta ccattcatca tcattattgg 1500
attccttttt	tttcttaaaa	aaaggtataa tatgatatgc aacttcttaa ttgctttatt 1560
tcttttctag	attaatttta	gataaaaatt ttattggata tggatcagct agcgtagtaa 1620
aaagtgaacg	atacatgaga	aaaaagattg atttgacaaa acaaaaacac aaccattaa 1680
attggagcgt	cttattcccc	tagactgtaa cagaatggtt cggatgatatg atcaactaat 1740
gttttttgtt	gcagcaaagt	attgtggtca aggtgccgga aaaaagaagg gacgacatga 1800
acgatgatgc	attgcaactg	gcggtgagtt cgttgcatgg agtgccctcg ggtggtacgt 1860
ctggagattg	tagtaggttg	cagttggatg gtgaaggcgc gaacatccgc aagctcttgt 1920
ccaccctccg	gaataaggtg	ggccacgcc agttggtgca ggtcgaggat aagagaaaaa 1980
gggtagagga	ggcgacgaag	ccttgtgaat ttcacaggt caaacaata tgcctcttg 2040

gattgcctgg	cgcaggcaaa	acaactcttg	caaaactggt	gtactcccat	cactcaacga	2100
cagagcagca	attccaacac	cgggctttcg	tgctactctc	tccgggtgcc	aatctcaccg	2160
acactcttac	tgatatttta	ttgcaagtag	gagcatataa	tgatgatgca	acaccatatt	2220
gtgggaccgg	aacaccgcac	caacagtatc	tcattgacaa	catatcagct	tatctcattg	2280
gcaaaaagta	agcagagttc	tttagaatga	tgttatttta	aataatatta	ttttttttta	2340
aaaaaaatta	acaatgtgta	tttgatggaa	ttaataaaaa	tatgttttaa	gagaaattaa	2400
taaaaaatat	ttgatatcaa	attctgcagg	atggctctta	ttagaatttc	tataaaaaga	2460
gagtagatga	gaaacaccca	gggctcttc	tgctagctc	cacaagccaa	cctattatgt	2520
ttgaagcctc	accctacct	atttaatat	aggtctttct	ctaataatcg	ctattttatt	2580
gatattaaat	ccttcctaa	tattcgtgtt	tttaaagag	agtagatgac	aaacagacat	2640
caaattaagc	tgattgtttt	tcgatcatct	caaaggggaa	gcttctcatg	tggggtggact	2700
catatcttcg	aaattattat	atagttgcat	gtattagtgc	taatataatg	aggcttattt	2760
actttttttc	aacttctaaa	gatatcttat	tataattgat	gacgtttggc	gctgggaaga	2820
gtgggaagtc	atcagaaaagt	ccattccaaa	gaatgatctg	gtagcagaa	taatcatgac	2880
tactcgtctt	aattcaatag	ctgagaagtg	tcgcaatgat	gacatggatg	cgtttgttta	2940
tgaaactgag	gctctggatt	atgtggatgc	ttggctgttg	tgtgacaagg	tagcaagaaa	3000
gtctgtcaca	tgatgaaca	ttaatccatg	ctatgatatc	gtggacgtgt	gctatggtat	3060
gccgttagca	ctaattcgtg	tgctgcagc	attggcagaa	gagatacaag	ctttagacag	3120
tgatgaatgg	caaatatgga	gggctctgag	acgggtagag	gatggtatth	tggacatccc	3180
atccttgaag	ccattggcag	agagtttatg	ccttggttac	gaccatcttc	ctctctatct	3240
gaggactttg	ttgttatgth	gtagtgtgta	ccattggctt	gatggtggga	ttgttcaaag	3300
gggccgtttg	gtcacaaggt	ggattgctga	aggattgtth	tcagaagaga	aagcagcaga	3360
aggttactth	gatgagcttg	tcgacagagg	atggattaag	catagagggt	ggaacgagta	3420
tgagatctac	cctatgatgc	tggccatcct	tagatacaag	tcgaaggagt	acaattttgt	3480
aacttgthttg	ggtacgggat	ttgatacttg	tactagtgca	tctctatcct	actcctctcc	3540
aacaatggcg	attcgcggcg	ttgtcttca	aaggggttac	ccaatgaaat	gcttctcaag	3600
tatggatgtg	tcacacactc	gcagccttgt	tatccttggc	gacgtgatag	gagtcccctt	3660
ggatatgthtt	aaaagattgc	gagtgttggga	ccttgaagat	aatatcggta	tagaggactc	3720
ccacctgaag	aagatatgtg	agcagctaga	gagcctcagg	ctgctcaagt	acctaggtct	3780
caagggtacg	cgaatcacta	agctcccaca	ggagatacag	aagctgaagc	atctggagat	3840
tttgtagctg	aggagcacag	gcatcaaaga	gctcccacgg	gagatcgggg	aagtgaaca	3900
actgcggact	ctggacgtga	ggaacacgcg	gatcagcgag	ctcccgtcgc	agatcgggga	3960
gctcaaacat	ctgcggactc	tggacgtgag	gaacacgcgg	atcagcgagc	tctgtcgca	4020
gatcggggag	ctcaaacatc	tgcggactct	ggacgtgagg	aacacgcgga	tcagcgagct	4080
cccgtcgag	atcggggagc	tcaaacatct	gcggactctg	gacgtgagga	acacgcggac	4140
ttctatattt	ttttattcta	gaagaagaat	aaaaaaatat	agaagtactg	atatctggct	4200
ctctgcacgt	gacatgcac	atacatgtat	atggtattaa			4240

<210> 14

<211> 2409

<212> DNA

<213> 人工序列(Artificial Sequence)

<400> 14

atgacggggg	aggaagtcga	tgctttgtgc	aaggatgagt	tgatggcggg	ggtgcgtgag	60
ctgtcctacg	acatggacga	cgccatcgac	gaattcttct	tagaggagcc	catggcgggc	120
ggcgacgggtg	gccctttcga	tgagctcaag	acaagagttg	aggatgtctc	caagcggttc	180
tccgacagcc	ggcgggtggag	gccacaggtg	gagcaacatc	aaccatccct	aaccgccgca	240
accgtagact	gtccacctcc	tcacgctcgc	ttcgtccaca	acatgatgga	tgtgtcagag	300
ctcgtggaga	tggacaaaact	acatgagaca	gagctcatca	aattgctgga	acaaggtgcg	360
gacacaagca	tatatgcttc	ccggtggcgc	atcgcaacac	catggcatga	taaggagcaa	420
agtattgtgg	tcaaggtgcc	ggaaaaaaga	agggacgaca	tgaacgatga	tgcattgcac	480
tgggcgggtga	gttcgttgca	tggagtgcc	tcgggtggta	cgtctggaga	ttgtagtagg	540
ttgcagttgg	atggtgaagg	cgcaacatc	cgcaagctct	tgtccacct	ccggaataag	600
gtgggccacg	cccagttggt	gcaggtcgag	gataagagaa	aaagggtaga	ggaggcgacg	660
aagccttggtg	aatttcacga	ggtcaaaaaca	atatgcatcc	ttggattgcc	tggcgcaggc	720
aaaacaactc	ttgcaaaaact	gttgtactcc	catcactcaa	cgacagagca	gcaattccaa	780
caccgggctt	tcgtgtcact	ctctccgggt	gccaatctca	ccgacactct	tactgatatt	840
ttattgcaag	taggagcata	taatgatgat	gcaacacat	attgtgggac	cggaacaccg	900
caccaacagt	atctcattga	caacatatca	gcttatctca	ttggcaaaaa	gtatcttatt	960
ataattgatg	acgtttggcg	ctgggaagag	tgggaagtca	tcagaaagtc	cattccaaag	1020
aatgatctgg	gtagcagaat	aatcatgact	actcgtctta	attcaatagc	tgagaagtgt	1080
cgcaatgatg	acatggatgc	gtttgtttat	gaaactgagg	ctctggatta	tgtggatgct	1140
tggctgttgt	gtgacaaggt	agcaagaaag	tctgtccat	gtatgaacat	taatccatgc	1200
tatgatatcg	tggacgtgtg	ctatggtatg	ccgtagcac	taattcgtgt	gtcgtcagca	1260
ttggcagaag	agatacaagc	tttagacagt	gatgaatggc	aaatatggag	ggctctgaga	1320
cgggtagagg	atggtatfff	ggacatccca	tccttgaagc	cattggcaga	gagtttatgc	1380
cttggttacg	accatcttcc	tctctatctg	aggactttgt	tgttatgttg	tagtgtgtac	1440
cattggccttg	atggtgggat	tgttcaaagg	ggccgttttg	tcacaaggtg	gattgctgaa	1500
ggattttgttt	cagaagagaa	agcagcagaa	ggttactttg	atgagcttgt	cgacagagga	1560
tggattaagc	atagagggtg	gaacgagtat	gagatctacc	ctatgatgct	ggccatcctt	1620
agatacaagt	cgaaggagta	caatfffgtg	acttgtttgg	gtacgggatt	tgatacttgt	1680
actagtgc	ctctataccta	ctctctcca	acaatggcga	ttcgccggct	ttgtcttcaa	1740
agggggtacc	caatgaaatg	cttctcaagt	atggatgtgt	cacacactcg	cagccttgtt	1800
atccttggcg	acgtgatagg	agtccccttg	gatatgttta	aaagattgcg	agtgttggac	1860
cttgaagata	atatcggtat	agaggactcc	cacctgaaga	agatatgtga	gcagctagag	1920
agcctcaggc	tgctcaagta	cctaggtctc	aagggtacgc	gaatcactaa	gctcccacag	1980
gagatacaga	agctgaagca	tctggagatt	ttgtacgtga	ggagcacagg	catcaaagag	2040
ctcccacggg	agatcgggga	agtgaaacaa	ctgcggactc	tggacgtgag	gaacacgcgg	2100
atcagcgagc	tcccgtcgca	gatcggggag	ctcaaacatc	tgcggactct	ggacgtgagg	2160

aacacgcgga tcagcgagct cctgtcgcag atcggggagc tcaaacatct gcggactctg	2220
gacgtgagga acacgcggat cagcgagctc ccgtcgcaga tcggggagct caaacatctg	2280
cggactctgg acgtgaggaa cacgcggact tctatatttt tttattctag aagaagaata	2340
aaaaaatata gaagtactga tatctggctc tctgcacgtg acatgcatca tacatgtata	2400
tggtattaa	2409
<210> 15	
<211> 4786	
<212> DNA	
<213> 人工序列(Artificial Sequence)	
<400> 15	
atgatggatg tgtcagagct cgtggagatg gacaaacaac atgagaaaga gctcatcaaa	60
ttgctggaac aaggtgcgga cacaagcata tatgcttccc ggtggcgcac cgcaacacca	120
tgcatgata aggaggtaaa gacgacatcc ttttatttct tttttatctc tacttctcta	180
tttatatatt atattatatt atattataaa aatttaaaat gttttcgcctg tgtgtatttt	240
ggtacgtcgt ggctatcttc gtatcgtatt cgatctcccg ttcagtatgt tttgtgttgt	300
acgtctctag cttccagata tatcatatat ctttccattc ccatgttatt ctttctttcc	360
aaattccaat tattaattat acctcattta atgagggaca tttataccat tttatttcag	420
tagtaatttc atcctttctc ctagtgctag aagtgcactt gcatgaggat ggggatggag	480
gcaaagctga aatacccatt cgttgtgatg aaattttaaa agcgaaccct accgtaaact	540
ctaaattaag ccccaaaaat tataatgcaa gtattcctgt tttaatcaat atttggcatc	600
atTTTTTTtac aatgtataat tcagctacac aatttgcctc ttttttttta caattgctta	660
taattaagct acacaagcta ccaattcagc tacacaattg gtacatacta ctatactagc	720
caaatacccg tgatttgcta cgtattaaaa caaattaata gtgagatttt tggggcaatt	780
aatttggttt tgtagaagtt atatcataga gaaattattt tataacggta aagtgggttg	840
aaaatatgat ggagaatgtg gtaaataaaa aaaatacta tagttgggtg cagattcact	900
gccaccgcca ttgccctctt ttgaaaggaa tatagaattt tatctttagt aagtattatt	960
gtacaagtga aatatgatgg aatcatatat gtagaataaa acattaaagt atgtgggggt	1020
atTTGGTTga aatatgatg gagtatgtgg taaatagaaa aaaatacta tacttgatga	1080
tggggcgatg atagattcac tgccaccacc attgcatttt tttaaaagga gtatataaac	1140
atatatagga acccaccaac tcattatttg gtaagaggtc ttacttggtg attgtgctgg	1200
acggtagcat gccagtttac cattcatcat cattattgga ttctttttt ttcttaaaaa	1260
aaggtataat atgatatgca acttcttaat tgctttattt cttttctaga ttaattttag	1320
ataaaaaattt tattggatat ggatcagcta gcgtagtaa aagtgaacga tacatgagaa	1380
aaaagattga tttgacaaaa caaaaacaca accattaaa ttggagcgtc ttattcccgt	1440
agactgtaac agaatggttc ggtgatatga tcaactaatg ttttttgttg cagcaaagta	1500
ttgtggtcaa ggtgccggaa aaaagaaggg acgacatgta cgatgatgca ttgcaactgg	1560
cggtgagttc gttgcatgga gtgccctcgg gtggtgcgtc tggagattgt agtaggttc	1620
agttggatgg tgaaggcgcg aacatccgca agctcttgct caccctccgg aataaggtgg	1680
gccgcgcca gttggtgcag gtcgaggata agagaaaaag ggtagaggag gcgacgaagc	1740



cttgtgaatt	tcacgaggtc	aaaacaatat	gcatccttgg	attgcctggc	gcaggcaaaa	1800
caactcttgc	aaaactgttg	tactcccatc	actcaacgac	agagcagcaa	ttccaacacc	1860
gggctttcgt	gtcactctct	ccgggtgcc	atctcaccga	cactcttact	gatattttat	1920
tgcaagtagg	agcatataat	gatgatgcaa	caccatattg	tgggaccgga	acaccgcacc	1980
aacagtatct	cattgacaac	atatcagctt	atctcattgg	caaaaagtaa	gcagagttct	2040
ttagaatgat	gttattttaa	ataatattat	tttttaaaa	aaaattaaca	aacatgttga	2100
tgtgtatttg	atggaattaa	taaaaatatg	ttttaagaga	aattaataaa	aaaatatttg	2160
atatcaaatt	ctgcaggatg	gtcttaatta	gaatttctat	aaaaagagag	tagatgagaa	2220
acaccaggga	gttcttctga	ctagctccac	aagccaacct	atgtttgaag	cctcaccctt	2280
acctatttat	ttaatattag	gtctttctct	aatattcgct	atatttttga	tattaaatcc	2340
ttccctaata	ttcgtgtttt	taaaagagag	tagatgacaa	acataatca	aattaagctg	2400
attgtttttc	gatcatctca	aaggggaagc	ttctcatgtg	ggtggactca	tatcttcgaa	2460
attattatata	agttgcatgt	attagtgcta	atatattgag	gcttattttac	tttttttcaa	2520
cttctaaagg	tatcttatta	taattgatga	cgtttggcgc	tgggaagagt	gggaagtcat	2580
cagaaagtcc	attccaaaga	atgatctggg	tagcagaata	atcatgacta	ctcgtcttaa	2640
ttcaatagct	gagaagtgtc	gcaatgatga	catggatgcg	tttgtttatg	aaactgaggc	2700
tctggattat	gtggatgctt	ggttgttgtg	tgacaaggta	gcaagaaagt	ctgtcacatg	2760
tatgaacatt	aatccatgct	atgatatcgt	ggacatgtgc	tatggtatgc	cgttagcact	2820
aattcgtgtg	tcgtcagcat	tggcagaaga	gatacaagct	ttagacagtg	atgaacggca	2880
aatatggagg	gctctgagac	gggtagagga	tggtattttg	gacatcccat	ccttgaagcc	2940
attggcagag	agtttatgcc	ttggttacga	ccatcttctt	ctctatctga	ggactttggt	3000
gttatgttgt	agtgtgtacc	attggcttga	tggtgggatt	gttcaaaggg	gccgtttggt	3060
cacaaggtgg	attgctgaag	gatttgtttc	agaagagaaa	gcagcagaag	gttactttga	3120
tgagcttgtc	ggcagaggat	ggatgaagca	tagagagttg	aacgagtatg	agatccaccc	3180
tatgatgctg	gccatcctta	gatacaagtc	gaaggagtac	aattttgtaa	cttgtttggg	3240
tacgggatct	gatacttght	ctagtgcac	tctatcttac	tcctctccaa	caatggcgat	3300
tcgccggctt	tgtcttcaaa	gggggtacc	aatgaaatgc	ttctcaagta	tggatgtgtc	3360
acacactcgc	agccttgtca	tccttggcga	cgatgatagga	gtccccttgg	atatgtttaa	3420
aagattgcga	gtgttggacc	ttgaagacaa	tctcgatata	gatgactctc	acctgaagaa	3480
gatatgtgag	cagctagaga	gcctcaggct	gctcaagtac	ctgggtatca	agggtacacg	3540
gatcactaag	ctcccacagg	agatacagaa	gctgaagcat	ctggagattt	tgtacgtgag	3600
gagcacaggc	atcaaagagc	tcccacggga	gatcggggaa	ttgaaacaac	tgcggactct	3660
ggacatgagg	aacacgcgga	tcagcgagct	cccgtcgcag	atcggggagc	tcaaacatct	3720
gcggactctg	gacgtgagta	acaacatgtg	gaatatcagc	gagctgccgt	cgcaaactcg	3780
ggagctgaag	catctacaaa	ctctggatgt	gaggaatacg	tcggtgagag	agctgccatc	3840
gcaaatcggg	gagctgaagc	atctgcggac	tctggatgtg	aggaacacgg	gggtgagaga	3900
gctgccatgg	caagctggcc	agatctcggg	atcgtgcac	gtgcatacag	atgacagtga	3960
cgagggcatg	cggctgccag	aaggcgtatg	cgaagatctg	atcaagggta	ttccaaggc	4020
tgagctcgca	aagtgcagtg	aggtcctatc	catcaatatt	gtcgtatcgtt	taggatctcc	4080

ccctattggc atattcaagg ttattggctt gcacaagagt atcccgaagc tgatcaaaga	4140
tcatttcaat gttctttctt ccctagacat caggcgtac aacaagctag aggaggatga	4200
ccatgagttt ctagccaaca atatgcctaa cctccagatg cttgtactga ggttcgaggc	4260
cccacaaaga gagcccatca acattaaccg cacaggcttc cagatgatgg agagattcct	4320
tgtggagagc cgggtgccac ggataacctt ccaggaagga gccatgcccga agctcaagca	4380
tctcgagttc aagtctctac ctggcccacc aagcaaagat cccataggaa tcaccacct	4440
caagagcctc caaaaggtgg tctttcgtg ctccaaatgg tacaagagcg acaaccctgg	4500
catcaaggct gccattgacg tcgtgaagaa agaagcaagg cagcatcca accggccgat	4560
cagccttctc atcactgagg gcgataagga ggtaccaaatt attgaggcac acgggagcag	4620
tgaaaacatt gtcgttgctc acgctgctc tgacgacgc atcagttgct ctagctgcgg	4680
ccgaaccagc actagtatcc aagagggaac agtccgagat cgaataccag ctatggattt	4740
gttctggccg gagttaaca gctatgaaa agcaaaaaga aactag	4786
<210> 16	
<211> 5230	
<212> DNA	
<213> 人工序列(Artificial Sequence)	
<400> 16	
atggccgtat acagcgtcgc cacgggggcc ttggctcccg tcctatcgaa gctctccgct	60
ttgctgggcg acgagcactt ggatcttgcg gagaggacc ggagcgacgc catgttcatc	120
aggtcccagc tggaggccgt gcactctctc ctctcccga ggattagttg gggaatgacg	180
ggggaggaag tcgatgcttt gtgcaaggat gattgatgg cggaggtgcg tgagctgtcc	240
tacgacatgg acgacgcat cgacgaattc ttcttagagg agcccatggc gggcggcgac	300
ggtggccctt tcgatgagct caagacaaga gttgaggatg tctccaagcg gttctccgac	360
agccggcggg ggaggccacc ggtggagcaa catcaacct ccctaaccgc cgcaaccgta	420
gactgtccac ctctcacgc tcgcttcgtc cacaacatga tggatgtgtc agagctcgtg	480
gagatggaca aacaacatga gaaagagctc atcaaattgc tggacaaggg tgcggacaca	540
agcatatatg cttcccgtg gcgcatcgca acaccatggc atgataagga ggtaaagacg	600
acatcctttt atttctttt tatcttact tctctattha tatattatat tatattatat	660
tataaaaatt taaaatgttt tcgctgtgtg ttttttgta cgtcgtggct atcttcgtat	720
cgtattcgat ctcccgttca gtatattttg tgtgtacgt ctctagcttc cagatatac	780
atatactttt ccattcccat gttattctt ctttccaaat tccaattatt aattatacct	840
catttaatga gggacattta taccattta tttcagtagt aatttcatcc tttcttagtg	900
ctagaagtgc acttgcata ggatggaggc aaagctgaaa taccattcg ttgtgatgaa	960
atthtaaaag cgaaccctac cgtaaaactet aaattaagcc ccaaaaatta taatgcaagt	1020
attcctgttt taatcaatat ttggcatcat ttttttaca tgtataattc agctacacaa	1080
tttgetcatt ttttttaca attgcttata attaagctac acaagctacc aattcagcta	1140
cacaattggc acatactact ataactagcca aataccggtt attttctacg tattaaaca	1200
aattaatagt gagatthttg gggcaattaa tttggttttg tagaagttat atcatagaga	1260
aattatthta taacggtaaa gttggttgaa aatatgatgg agaattgtgg aaataaaaa	1320

aatactatag	ttggtggcag	attcactgcc	accgccattg	ccctcttttg	aaaggaatat	1380
agaattttat	cttgtagaag	ttatattgta	caagtgaaat	atgatggaat	catatatgta	1440
gaataaaaca	ttaaagtatg	tgggggtatt	tggttgaaaa	tatgatggag	tatgtggtaa	1500
atagaaaaaa	aatactatac	ttgatgatgg	ggcgatgata	gattcactgc	caccaccatt	1560
gcattttttt	aaaaggagta	tataaacata	tataggaacc	cacctactca	ttacttggta	1620
agaggcttta	cttggttaatt	gtgctggacg	gtagtatgcc	agtttaccat	tcatcatcat	1680
tattggattc	cttttttttc	ttaaaaaaag	gtataatatg	atatgcaact	tcttaattgc	1740
tttatttctt	ttctagatta	atthtagata	aaaattttat	tggatatgga	tcagctagcg	1800
tagtaaaaaag	tgaacgatac	atgagaaaaa	agattgattt	gacaaaacaa	aaacacaacc	1860
cattaaattg	gagtgtctta	ttcccgtaga	ctgcagtaac	agaatggttc	gatgatcaac	1920
taatgttttt	tgctgcagca	aagtactgtg	gtcaagggtc	cggaaagaga	gtggggcttc	1980
ccggacaatc	ggaacagtcc	atthtatatg	gcgagtgatt	cgthtgaacg	attgcgttcg	2040
ggaagtthgt	gtggagatac	gthtgcgtht	gatggtgaag	gcgcgaacat	ccgcaagctc	2100
thgtccacc	tccggaataa	gthtggccc	gcccagtht	tgcaggtcga	ggataagaga	2160
aaaagggtag	aggaggcgac	gaagccttht	gaatttcacg	aggtcaaac	aatatgcatc	2220
ctthgattgc	ctggcgcagg	caaaacaact	ctthcaaac	tthtgtactc	ccatcactca	2280
acgacagagc	agcaattcca	acaccgggct	thtctgtcac	tctctccggg	tccaatctc	2340
accgacactc	ttactgatat	thtattgcaa	gtaggagcat	ataatgatga	tgcaacacca	2400
tattgtggga	ccggaacacc	gcaccaacag	tatctcattg	acaatatatc	agcttatctc	2460
attggcaaaa	agtaagcaga	gthtctthaga	atgatthtat	thtaataat	aatatttttt	2520
thtaaaaaaa	aattaacaaa	catgtht	tgtatttht	ggaattaata	aaaatatht	2580
thtaagagaaa	thtaataaaaa	atattthgata	tcaattctg	caggatggtc	thtaattagaa	2640
thtctataaaa	aagagagtag	atgagaaata	cccagggtht	ctthtggcta	gctccacaag	2700
ccaatctatg	thtgaagcct	caccctacc	tattthatta	atattaggtc	thtcccta	2760
atthcgtatt	tattthgatat	thaatcctc	cctaatht	gtthttthta	aagagagtag	2820
atgacaaaaca	gacatcaaat	thagctgatt	gthttctgat	catctcaag	gggaagcttc	2880
tcatgtgggt	ggactcatat	ctthcgaatt	attatatht	tgcathtatt	agthtcaata	2940
tattgaggct	tattthactth	thtthcaact	ctgaaggtat	ctthattata	thtathgact	3000
thtggcactgg	gaagagtht	aagthtctag	aaagthtct	ccaagaatht	atctgggtag	3060
cagaataatc	atgactactc	gtctthattc	aathgctgag	aagthtctgca	atgathgacat	3120
ggathgctth	gthtthgaaa	ctgaggctct	ggaththt	gathgcttht	tgtthtthgaa	3180
caaggtagca	agaaagtht	thcatgtht	gaacatht	ccathgctat	atathgctgga	3240
catgthgctat	gthtathgctg	thgactaat	thtthttht	thgacttht	cagaagagat	3300
acaagctth	gacathgctg	aacggcaaat	atggagggct	ctgagacggg	thgaggatgg	3360
tattthtggac	atcccatcct	tgaagcct	ggcagagagtht	thtathgctth	gthtathgctca	3420
tctthcctctc	tathctgagga	ctthtthtth	atgthttht	gtthtaccat	ggctthgctg	3480
tgggathtth	caaagggg	gthtthtct	aagthttht	gctgaaggt	thtthtctgaa	3540
agagaaaagca	gcagaaggt	actthgctg	gctthtctg	agaggtgga	tgaagctag	3600
agggtthgaa	gathtathgaa	tccaccctat	gathgctgg	atctthgctg	acaathgctg	3660

ggagtacaat tttgtaactt gtttgggtac gggatctgat acttgtacta gtgcatctct	3720
atcctactcc tctccaacaa tggcgattcg ccggtttgt cttcaaaggg ggtaccaat	3780
gaaatgcttc tcaagtatgg atgtgtcaca cactcgcagc cttgtcatcc ttggcgacgt	3840
gataggagtc cccttgata tgtttaaaag attgcgagtg ttggacctg aagataatat	3900
cggtatagag gactcccacc tgaagaagat atgtgagcag ctagagagcc tcaggctgct	3960
caagtaccta ggtctcaagg gtacgcgaat cactaagctc ccacaggaga tacagaagct	4020
gaagcaactg gagattttgt acgtgaggag cacaggcatc gaagagctcc catgggagat	4080
cggggaattg aaacaactgc ggactctgga cgtgaggaac acgcggatca gcgagctccc	4140
gtcgcagatc ggggagctca aacatctgcg gactctggac gtgagtaaca tgtggaatat	4200
cagcgagctg ccgtcgcaaa tcggggagct gaagcatcta caaactctgg atgtgaggaa	4260
cacgtcagtg agagagctgc catcgcaaat cggggagctg aagcatctgc ggactctgga	4320
tgtgaggaac acgggggtga gagagctgcc atggcaagct ggccagatct cgggatcgct	4380
gcacgtgcat acagatgaca gtgacgaggg catgcggctg ccagaaggcg tatgcgaaga	4440
tctgatcaag ggtattccca aggetgagct cgcaaagtc agtgaggtcc tatccatcaa	4500
tattgtcgat cgtttaggat ctccccctat tggcatatc aaggttattg gcttgcaaa	4560
gagtatcccg aagctgatca aagatcattt caatgttctt tcttcctag acatcaggcg	4620
gtacaacaag ctagaggagg atgaccatga gtttctagcc aacaatatgc ctaacctcca	4680
gatgcttgta ctgaggttcg aggccccaca aagagagccc atcatcatta accgcacagg	4740
cttcagatg ctggagagat tccttgtaga gagccgggtg ccacggataa ctttcagga	4800
aggagccatg cccaagctca agcatctcga gtttaagttc tacgctggcc caccaagcaa	4860
agatcccata ggaatcacc acctcaagag cctccaaaag gtggtctttc gctgctccaa	4920
atggtacaag agcgacaacc ctggcatcaa ggctgccatt gacgtcgtga agaaagaagc	4980
aaggcagcat cccaaccggc cgatcagcct tctcatcact gagggcgata aggaggtacc	5040
gaatattgag gcacacggga gcagtgaaa cattgtcgtt gtccacgctg ctctgacga	5100
cgccatcagt tgctctagct gcggccgaac cagcactagt atccaagagg gaacagtccg	5160
agatcgaata ccagctatgg atttgttctg gccggagttt aacagctatg aaaaagcaaa	5220
aagaaactag	5230
<210> 17	
<211> 20	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 17	
aagcagcacc tatagctaac	20
<210> 18	
<211> 20	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 18	
agcgtggcac tataaatgaa	20

<210>	19	
<211>	20	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	19	
	cctgtgcata atagcttctc	20
<210>	20	
<211>	21	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	20	
	atcctagga tgaatgtggt c	21
<210>	21	
<211>	20	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	21	
	atcggatgct cacttaactc	20
<210>	22	
<211>	20	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	22	
	cgccccaatt ttgtttatcg	20
<210>	23	
<211>	19	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	23	
	atgatgatg tgtcagagc	19
<210>	24	
<211>	22	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	24	
	ctagtttctt tttgcttttt ca	22
<210>	25	
<211>	18	
<212>	DNA	

<213> 人工序列 (Artificial Sequence)	
<400> 25	
atggccgtat acagcgtc	18
<210> 26	
<211> 23	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 26	
ctagtttctt tttgcttttt cat	23
<210> 27	
<211> 5232	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 27	
atggccgtat acagcgtcgc cacgggggccc ttggtcccc tcctatcgaa gctctccgct	60
ttgctgggccc acgagcactt ggatcttgcg gagaggacct ggagcgacgc catgttcatc	120
aggctcccagc tggaggccgt gcactctctc ctctctccga ggattagttg gggaatgacg	180
ggggaggaag tcgatgcttt gtgcaaggat gagttgatgg cggaggtgcg tgagctgtcc	240
tacgacatgg acgacgccat cgacgaattc ttcttagagg agcccatggc gggcggcgac	300
ggtggccctt tcgatgagct caagacaaga gttgaggatg tctccaagcg gttctccgac	360
agccggcggg ggaggccacc ggtggagcaa catcaacct ccctaaccgc cgcaaccgta	420
gactgtccac ctctcacgc tcgcttcgtc cacaacatga tggatgtgtc agagctcgtg	480
gagatggaca aactacatga gaaagagctc atcaaattgc tggacaagg tcgagacaca	540
agcatatatg cttcccgggt gcgcatcgca acaccatggc atgataagga ggtaaagacg	600
acatcctttt atttctttt tatctctact tctctattta tatattatat tatattataa	660
aaatttaaaa tgttttcgct gtgtgtattt tggtagctcg tggctatctt cgtatcgtat	720
tcgatctccc gttcagtatg tttgtgttg tacgtctcta gttccagat atatcatata	780
tctttccatt cccatgttat tctttcttc caaattcca ttattaatta tacctcattt	840
aatgagggac atttatacca ttttattca gtagtaattt catcctttct tctagtgcta	900
gaagtgcact tgcattgagga tggaggcaaa gctgaaatac ccattcgttg tgatgaaatt	960
ttaaaagcga accctaccgt aaactctaaa ttaagcccca aaaattataa tgcaagtatt	1020
cctgttttaa tcaatatttg gcatcatttt ttacaatgt ataattcagc tacacaattt	1080
gctcattttt ttttacaatt gcttataatt aagctacaca agctaccaat tcagctacac	1140
aattggtaca tactactata ctageccaat acccgtgatt tgctacgtat taaaacaaat	1200
taatagtgag atttttgggg caattaattt ggtttttag aagttatata atagagaaat	1260
tattttataa cggtaaagtt ggttgaaaat atgatggaga atgtggtaaa taaaaaaaa	1320
tactatagtt ggtggcagat tcaactgccac cgccattgcc ctcttttgaa aggaatatag	1380
aattttatct tgtagaagtt atattgtaca agtgaaatat gatggaatca tatatgtaga	1440
ataaaacatt aaagtatgtg ggggtatttg gttgaaaata tgatggagta tgtggtaaat	1500

agaaaaaaaa	tactatactt	gatgatgggg	cgatgataga	ttcactgcca	ccaccattgc	1560
atTTTTTTaa	aaggagtata	taaacaata	taggaacca	ccaactcatt	atttggttaag	1620
aggTcttact	tggttaattgt	gctggacggt	agcatgccag	tttaccattc	atcatcatta	1680
ttggattcct	TTTTTTtctt	aaaaaaaggt	ataatatgat	atgcaacttc	ttaattgctt	1740
tatttctttt	ctagattaat	tttagataaa	aattttattg	gatatggatc	agctagcgtta	1800
gtaaaaagtg	aacgatacat	gagaaaaaag	attgatttga	caaaacaaaa	acacaacca	1860
ttaaattgga	gCGTcttatt	cccgtagact	gtaacagaat	ggttcggtga	tatgatcaac	1920
taatgttttt	tgTtgcagca	aagtattgtg	gtcaaggtgc	cggaaaaaag	aagtgacgac	1980
atgaacgatg	atgcattgca	ctgggcggtg	agtttgtcgc	atggagtgcc	ctcgggtggt	2040
acgtctggag	attgtagtag	gTtgcagttg	gatggtgaag	gcgcgaacat	ccgcaagctc	2100
ttgtccacc	tccggaataa	ggtgggccac	gccagttgg	tgcaggtcga	ggataagaga	2160
aaaagggtag	aggaggcgac	gaagccttgt	gaatttcacg	aggtcaaac	aatatgcatc	2220
cttgattgc	ctggcgagg	caaaacaact	cttcaaac	tgttgactc	ccatcactca	2280
acgacagagc	agcaattcca	acaccgggct	ttcgtgtcac	tctctccggg	tgccaatctc	2340
accgacactc	ttactgatat	tttattgcaa	gtaggagcat	ataatgatga	tgcaacacca	2400
tattgtggga	ccggaacacc	gcaccaacag	tatctcattg	acaacatatac	agcttatctc	2460
attggcaaaa	agtaagcaga	gttctttaga	atgatgttat	tttaataat	attatTTTT	2520
ttaaaaaaaaa	ttaacaaaca	tgTtgatgtg	tatttgatgg	aattaataaa	aatatgtttt	2580
aagagaaatt	aataaaaaaa	tatttgatat	caaattctgc	aggatggtct	taattagaat	2640
ttctataaaa	agagagtaga	tgagaaacac	ccaggggttc	ttctgactag	ctccacaagc	2700
caacctatgt	ttgaagcctc	accctacct	atTTatTTaa	tattaggtct	ttctctaata	2760
ttcgtatTT	atttgatatt	aaatccttc	ctaattctgc	tgTTTTTaa	agagagtaga	2820
tgacaaacat	atatcaaatt	aagctgattg	TTTTtcgatc	atctcaaagg	ggaagcttct	2880
catgtgggtg	gactcatatac	ttcgaaatta	ttatatagtt	gcatgtatta	gtgctaatat	2940
attgaggctt	atttactttt	tttcaacttc	taaaggtatc	ttattataat	tgatgacgtt	3000
tggcgtggg	aagagtggga	agtcatcaga	aagtccattc	caaagaatga	tctgggtagc	3060
agaataatca	tgactactcg	tcttaattca	atagctgaga	agtgtcgcaa	tgatgacatg	3120
gatgcgtttg	tttatgaaac	tgaggctctg	gattatgtgg	atgcttggtt	gttgtgtgac	3180
aaggtagcaa	gaaagtctgt	cacatgtatg	aacattaatc	catgctatga	tatcgtggac	3240
atgtgctatg	gtatgccgtt	agcactaatt	cgtgtgtcgt	cagcattggc	agaagagata	3300
caagcttttag	acagtgatga	acggcaaata	tggagggtc	tgagacgggt	agaggatggt	3360
atTTTtgaca	tcccatcctt	gaagccattg	gcagagagtt	tatgccttgg	ttacgaccat	3420
cttctctct	atctgaggac	ttgtttgta	tgTttagtg	tgtaccattg	gcttgatggt	3480
gggattgttc	aaaggggccc	tttggtcaca	aggtggattg	ctgaaggatt	tgTttcgaa	3540
gagaaagcag	cagaaggtta	ctttgatgag	cttgtcggca	gaggatgat	gaagcataga	3600
gagttgaacg	agtatgagat	ccaccctatg	atgctggcca	tccttagata	caagtcgaag	3660
gagtacaatt	tcgtaacttg	tttgggtacg	ggatctgata	cttgtactag	tgcattctta	3720
tcctactcct	ctccaacaat	ggcgattcgc	cggtttgtc	ttcaaagggg	gtaccaaatg	3780
aaatgcttct	caagtatgga	tgtgtcacac	actcgcagcc	ttgtcatcct	tggcgacgtg	3840

ataggagtcc ccttggatat gtttaaaaga ttgcgagtgt tggaccttga agacaatctc	3900
gatatagatg actctcacct gaagaagata tgtgagcagc tagagagcct caggctgctc	3960
aagtacctgg gtatcaaggg tacacggatc actaagctcc cacaggagat acagaagctg	4020
aagcatctgg agatthttgta cgtgaggagc acaggcatca aagagctccc acgggagatc	4080
ggggaattga aacaactgcg gactctggac atgaggaaca cgcggatcag cgagctcccg	4140
tcgcagatcg gggagctcaa acatctgctg actctggacg tgagtaacaa catgtggaat	4200
atcagcgcgc tgccgtcgca aatcggggag ctgaagcatc tacaactct ggatgtgagg	4260
aatacgtcgg tgagagagct gccatcgcaa atcggggagc tgaagcatct gcggactctg	4320
gatgtgagga acacgggggt gagagagctg ccatggcaag ctggccagat ctgggatcgc	4380
ctgcacgtgc atacagatga cagtgcagc ggcctgcggc tgccagaagg cgtatgcaaa	4440
gatctgatca aggtatttcc caagctgag ctgcgaaagt gcagtgaggt cctatccatc	4500
aatattgtcg atcgtttagg atctccccct attggcatat tcaaggttat tggcttgca	4560
aagagtatcc cgaagctgat caaagatcat ttcaatgttc tttcttccct agacatcagg	4620
cggtacaaca agctagagga ggatgacct gagtttctag ccaacaatat gcctaacctc	4680
cagatgcttg tactgagggt cgaggcccca caaagagagc ccatcatcat taaccgcaca	4740
ggcttccaga tgatggagag attccttgtg gagagccggg tgccacggct aaccttccag	4800
gaaggagcca tgcccaagct caagcatctc gatttcaagt tctacgctgg cccaccaagc	4860
aaagatccca taggaatcac ccacctcaag agcctccaaa aggtggtctt tcgctgctcc	4920
aaatggtaca agagcgacaa ccctggcatc aaggctgcca ttgacgtcgt gaagaaagaa	4980
gcaaggcagc atcccaaccg gccgatcagc cttctcatca ctgagggcga taaggaggt	5040
ccaaatattg aggcacacgg gagcagtga aacattgtcg ttgtccacgc tgctcctgac	5100
gacgccatca gttgctctag ctgcggccga accagcacta gtatccaaga gggaacagtc	5160
cgagatcgaa taccagctat ggattgttc tggccggagt ttaacagcta tgaaaaagca	5220
aaaagaaact ag	5232
<210> 28	
<211> 3381	
<212> DNA	
<213> 人工序列 (Artificial Sequence)	
<400> 28	
atggccgtat acagcgtcgc cacggggggc ttggtcccc tcctatcgaa gctctccgct	60
ttgctgggcg acgagcactt ggatcttgcg gagaggacc ggagcgacgc catgttcatc	120
aggctcccagc tggaggccgt gcaactctct ctctctccga ggattagttg gggaatgacg	180
ggggaggaag tcgatgcttt gtgcaaggat gatttgatgg cggagggtgc tgagctgtcc	240
tacgacatgg acgacgcat cgacgaatc ttcttagagg agcccatggc gggcgggcagc	300
ggtggccctt tcgatgagct caagacaaga gttgaggatg tctccaagcg gttctccgac	360
agccggcggg ggaggccacc ggtggagcaa catcaacct ccctaaccgc cgcaaccgta	420
gactgtccac ctctcacgc tcgcttcgtc cacaacatga tggatgtgtc agagctcgtg	480
gagatggaca aactacatga gaaagagctc atcaaattgc tggacaaggg tgcggacaca	540
agcatatatg cttcccgggt gcgcatcgca acacctggc atgataagga gcaagatt	600



gtggtcaagg	tgccgaaaa	aagaagtac	gacatgaacg	atgatgcatt	gacttgggcg	660
gtgagtttgt	cgcatggagt	gccctcgggt	ggtacgtctg	gagattgtag	taggttgcag	720
ttggatgggtg	aaggcgcgaa	catccgcaag	ctcttgtcca	ccctccgga	taaggtgggc	780
cacgcccagt	tgggtcaggt	cgaggataag	agaaaaagg	tagaggaggc	gacgaagcct	840
tgtgaatttc	acgaggtcaa	aacaatatgc	atccttggat	tgcctggcgc	aggcaaaaca	900
actcttgcaa	aactgttcta	ctcccatcac	tcaacgacag	agcagcaatt	ccaacaccgg	960
gctttcgtgt	cactctctcc	gggtgccaat	ctcaccgaca	ctcttactga	tattttattg	1020
caagtaggag	catataatga	tgatgcaaca	ccatattgtg	ggaccggaac	accgcaccaa	1080
cagtatctca	ttgacaacat	atcagcttat	ctcattggca	aaaagtatct	tattataatt	1140
gatgacgttt	ggcgttggga	agagtgggaa	gtcatcagaa	agtccattcc	aaagaatgat	1200
ctgggtagca	gaataatcat	gactactcgt	cttaattcaa	tagctgagaa	gtgtcgcaat	1260
gatgacatgg	atgcgtttgt	ttatgaaact	gaggctctgg	attatgtgga	tgcttggttg	1320
ttgtgtgaca	aggtagcaag	aaagtctgtc	acatgtatga	acattaatcc	atgctatgat	1380
atcgtggaca	tgtgctatgg	tatgccgtta	gactaattc	gtgtgtcgtc	agcattggca	1440
gaagagatac	aagctttaga	cagtgatgaa	cggcaaatat	ggagggctct	gagacgggta	1500
gaggatggta	ttttggacat	cccatccttg	aagccattgg	cagagagttt	atgccttggg	1560
tacgaccatc	ttcctctcta	tctgaggact	ttgttgttat	gttgtagtgt	gtaccattgg	1620
cttgatgggtg	ggattgttca	aaggggccgt	ttggtcacia	ggtggattgc	tgaaggattt	1680
gtttcagaag	agaaagcagc	agaaggttac	tttgatgagc	ttgtcggcag	aggatggatg	1740
aagcatagag	agttgaacga	gtatgagatc	caccctatga	tgctggccat	ccttagatac	1800
aagtogaagg	agtacaatth	cgtaacttgt	ttgggtacgg	gatctgatac	ttgtactagt	1860
gcatctctat	cctactcctc	tccaacaatg	gcgattcgcc	ggctttgtct	tcaaaggggg	1920
tacccaatga	aatgcttctc	aagtatggat	gtgtcacaca	ctcgcagcct	tgtcatcctt	1980
ggcgacgtga	taggagtccc	cttggatatg	tttaaagat	tgcgagtgtt	ggaccttgaa	2040
gacaatctcg	atatagatga	ctctcacctg	aagaagatat	gtgagcagct	agagagcctc	2100
aggctgctca	agtacctggg	tatcaagggt	acacggatca	ctaagctccc	acaggagata	2160
cagaagctga	agcatctgga	gattttgtac	gtgaggagca	caggcatcaa	agagctccca	2220
cgggagatcg	gggaattgaa	acaactgcgg	actctggaca	tgaggaacac	gcggatcagc	2280
gagctcccgt	cgcatcgg	ggagctcaaa	catctgcgga	ctctggacgt	gagtaacaac	2340
atgtggaata	tcagcgagct	gccgtcgcaa	atcggggagc	tgaagcatct	acaactctg	2400
gatgtgagga	atacgtcggg	gagagagctg	ccatcgcaaa	tcggggagct	gaagcatctg	2460
cggactctgg	atgtgaggaa	cacgggggtg	agagagctgc	catggcaagc	tggccagatc	2520
tcgggatcgc	tgcacgtgca	tacagatgac	agtgacgagg	gcatgcggct	gccagaaggc	2580
gtatgcgaag	atctgatcaa	gggtattccc	aaggctgagc	tcgcaaagtg	cagtgaggtc	2640
ctatccatca	atattgtcga	tcgttttaga	tctccccta	ttggcatatt	caaggttatt	2700
ggcttgcaaca	agagtatccc	gaagctgatc	aaagatcatt	tcaatgttct	ttcttccta	2760
gacatcaggc	ggtacaacaa	gctagaggag	gatgacctg	agtttctagc	caacaatatg	2820
cctaacctcc	agatgcttgt	actgaggttc	gaggccccac	aaagagagcc	catcatcatt	2880
aaccgcacag	gcttccagat	gatggagaga	ttccttgtgg	agagccgggt	gccacggcta	2940

accttccagg aaggagccat gcccaagctc aagcatctcg agttcaagtt ctacgctggc 3000  
 ccaccaagca aagatcccat aggaatcacc cacctcaaga gcctccaaaa ggtggtcttt 3060  
 cgctgctcca aatggtacaa gagcgacaac cctggcatca aggctgccat tgacgtcgtg 3120  
 aagaaagaag caaggcagca tcccaaccgg ccatcagcc ttctcatcac tgagggcgat 3180  
 aaggaggtac caaatattga ggcacacggg agcagtgaaa acattgtcgt tgtccacgct 3240  
 gctcctgacg acgccatcag ttgctctagc tgcggccgaa ccagcactag tatccaagag 3300  
 ggaacagtcc gagatcgaat accagctatg gatttgttct ggccggagtt taacagctat 3360  
 gaaaaagcaa aaagaaacta g 3381

<210> 29

<211> 1126

<212> PRT

<213> 人工序列 (Artificial Sequence)

<400> 29

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

Ser Asp Asp Met Asn Asp Asp Ala Leu His Trp Ala Val Ser Leu Ser  
 210 215 220  
 His Gly Val Pro Ser Gly Gly Thr Ser Gly Asp Cys Ser Arg Leu Gln  
 225 230 235 240  
 Leu Asp Gly Glu Gly Ala Asn Ile Arg Lys Leu Leu Ser Thr Leu Arg  
 245 250 255  
 Asn Lys Val Gly His Ala Gln Leu Val Gln Val Glu Asp Lys Arg Lys  
 260 265 270  
 Arg Val Glu Glu Ala Thr Lys Pro Cys Glu Phe His Glu Val Lys Thr  
 275 280 285  
 Ile Cys Ile Leu Gly Leu Pro Gly Ala Gly Lys Thr Thr Leu Ala Lys  
 290 295 300  
 Leu Leu Tyr Ser His His Ser Thr Thr Glu Gln Gln Phe Gln His Arg  
 305 310 315 320  
 Ala Phe Val Ser Leu Ser Pro Gly Ala Asn Leu Thr Asp Thr Leu Thr  
 325 330 335  
 Asp Ile Leu Leu Gln Val Gly Ala Tyr Asn Asp Asp Ala Thr Pro Tyr  
 340 345 350  
 Cys Gly Thr Gly Thr Pro His Gln Gln Tyr Leu Ile Asp Asn Ile Ser  
 355 360 365  
 Ala Tyr Leu Ile Gly Lys Lys Tyr Leu Ile Ile Ile Asp Asp Val Trp  
 370 375 380  
 Arg Trp Glu Glu Trp Glu Val Ile Arg Lys Ser Ile Pro Lys Asn Asp  
 385 390 395 400  
 Leu Gly Ser Arg Ile Ile Met Thr Thr Arg Leu Asn Ser Ile Ala Glu  
 405 410 415  
 Lys Cys Arg Asn Asp Asp Met Asp Ala Phe Val Tyr Glu Thr Glu Ala  
 420 425 430  
 Leu Asp Tyr Val Asp Ala Trp Leu Leu Cys Asp Lys Val Ala Arg Lys  
 435 440 445  
 Ser Val Thr Cys Met Asn Ile Asn Pro Cys Tyr Asp Ile Val Asp Met  
 450 455 460  
 Cys Tyr Gly Met Pro Leu Ala Leu Ile Arg Val Ser Ser Ala Leu Ala  
 465 470 475 480  
 Glu Glu Ile Gln Ala Leu Asp Ser Asp Glu Arg Gln Ile Trp Arg Ala  
 485 490 495  
 Leu Arg Arg Val Glu Asp Gly Ile Leu Asp Ile Pro Ser Leu Lys Pro  
 500 505 510  
 Leu Ala Glu Ser Leu Cys Leu Gly Tyr Asp His Leu Pro Leu Tyr Leu

515	520	525
Arg Thr Leu Leu Leu Cys Cys Ser Val Tyr His Trp Leu Asp Gly Gly		
530	535	540
Ile Val Gln Arg Gly Arg Leu Val Thr Arg Trp Ile Ala Glu Gly Phe		
545	550	555
Val Ser Glu Glu Lys Ala Ala Glu Gly Tyr Phe Asp Glu Leu Val Gly		
565	570	575
Arg Gly Trp Met Lys His Arg Glu Leu Asn Glu Tyr Glu Ile His Pro		
580	585	590
Met Met Leu Ala Ile Leu Arg Tyr Lys Ser Lys Glu Tyr Asn Phe Val		
595	600	605
Thr Cys Leu Gly Thr Gly Ser Asp Thr Cys Thr Ser Ala Ser Leu Ser		
610	615	620
Tyr Ser Ser Pro Thr Met Ala Ile Arg Arg Leu Cys Leu Gln Arg Gly		
625	630	635
Tyr Pro Met Lys Cys Phe Ser Ser Met Asp Val Ser His Thr Arg Ser		
645	650	655
Leu Val Ile Leu Gly Asp Val Ile Gly Val Pro Leu Asp Met Phe Lys		
660	665	670
Arg Leu Arg Val Leu Asp Leu Glu Asp Asn Leu Asp Ile Asp Asp Ser		
675	680	685
His Leu Lys Lys Ile Cys Glu Gln Leu Glu Ser Leu Arg Leu Leu Lys		
690	695	700
Tyr Leu Gly Ile Lys Gly Thr Arg Ile Thr Lys Leu Pro Gln Glu Ile		
705	710	715
Gln Lys Leu Lys His Leu Glu Ile Leu Tyr Val Arg Ser Thr Gly Ile		
725	730	735
Lys Glu Leu Pro Arg Glu Ile Gly Glu Leu Lys Gln Leu Arg Thr Leu		
740	745	750
Asp Met Arg Asn Thr Arg Ile Ser Glu Leu Pro Ser Gln Ile Gly Glu		
755	760	765
Leu Lys His Leu Arg Thr Leu Asp Val Ser Asn Asn Met Trp Asn Ile		
770	775	780
Ser Glu Leu Pro Ser Gln Ile Gly Glu Leu Lys His Leu Gln Thr Leu		
785	790	795
Asp Val Arg Asn Thr Ser Val Arg Glu Leu Pro Ser Gln Ile Gly Glu		
805	810	815
Leu Lys His Leu Arg Thr Leu Asp Val Arg Asn Thr Gly Val Arg Glu		
820	825	830

Leu Pro Trp Gln Ala Gly Gln Ile Ser Gly Ser Leu His Val His Thr  
 835 840 845  
 Asp Asp Ser Asp Glu Gly Met Arg Leu Pro Glu Gly Val Cys Glu Asp  
 850 855 860  
 Leu Ile Lys Gly Ile Pro Lys Ala Glu Leu Ala Lys Cys Ser Glu Val  
 865 870 875 880  
 Leu Ser Ile Asn Ile Val Asp Arg Leu Gly Ser Pro Pro Ile Gly Ile  
 885 890 895  
 Phe Lys Val Ile Gly Leu His Lys Ser Ile Pro Lys Leu Ile Lys Asp  
 900 905 910  
 His Phe Asn Val Leu Ser Ser Leu Asp Ile Arg Arg Tyr Asn Lys Leu  
 915 920 925  
 Glu Glu Asp Asp His Glu Phe Leu Ala Asn Asn Met Pro Asn Leu Gln  
 930 935 940  
 Met Leu Val Leu Arg Phe Glu Ala Pro Gln Arg Glu Pro Ile Ile Ile  
 945 950 955 960  
 Asn Arg Thr Gly Phe Gln Met Met Glu Arg Phe Leu Val Glu Ser Arg  
 965 970 975  
 Val Pro Arg Leu Thr Phe Gln Glu Gly Ala Met Pro Lys Leu Lys His  
 980 985 990  
 Leu Glu Phe Lys Phe Tyr Ala Gly Pro Pro Ser Lys Asp Pro Ile Gly  
 995 1000 1005  
 Ile Thr His Leu Lys Ser Leu Gln Lys Val Val Phe Arg Cys Ser Lys  
 1010 1015 1020  
 Trp Tyr Lys Ser Asp Asn Pro Gly Ile Lys Ala Ala Ile Asp Val Val  
 1025 1030 1035 1040  
 Lys Lys Glu Ala Arg Gln His Pro Asn Arg Pro Ile Ser Leu Leu Ile  
 1045 1050 1055  
 Thr Glu Gly Asp Lys Glu Val Pro Asn Ile Glu Ala His Gly Ser Ser  
 1060 1065 1070  
 Glu Asn Ile Val Val Val His Ala Ala Pro Asp Asp Ala Ile Ser Cys  
 1075 1080 1085  
 Ser Ser Cys Gly Arg Thr Ser Thr Ser Ile Gln Glu Gly Thr Val Arg  
 1090 1095 1100  
 Asp Arg Ile Pro Ala Met Asp Leu Phe Trp Pro Glu Phe Asn Ser Tyr  
 1105 1110 1115 1120  
 Glu Lys Ala Lys Arg Asn  
 1125

<210> 30

<211>	2920	
<212>	DNA	
<213>	人工序列 (Artificial Sequence)	
<400>	30	
atgatggatg	tgtcagagct	cgtggagatg gacaaacaac atgagaaaga gctcatcaaa 60
ttgctggaac	aaggtgcgga	cacaagcata tatgcttccc ggtggcgcac cgcaacacca 120
tggcatgata	aggagcaaag	tattgtggtc aaggtgccgg aaaaaagaag ggacgacatg 180
tacgatgatg	cattgcactg	ggcggtagt tcgttgcatg gactgccctc gggtagtgcg 240
tctggagatt	gtagtaggtt	gcagttggat ggtgaaggcg cgaacatccg caagctcttg 300
tccaccctcc	ggaataaggt	gggccgcgcc cagttggtag aggtcgagga taagagaaaa 360
aggtagagg	aggcgacgaa	gccttgtgaa ttacagagg tcaaaacaat atgcatcctt 420
ggattgcctg	gcgcaggcaa	aacaactctt gcaaaactgt tgtactcca tcaactcaacg 480
acagagcagc	aattccaaca	ccgggcttcc gtgtcactct ctccgggtgc caatctcacc 540
gacactctta	ctgatatttt	attgcaagta ggagcatata atgatgatgc aacaccatat 600
tgtgggaccg	gaacaccgca	ccaacagtat ctattgaca acatctcagc ttatctcatt 660
ggcaaaaagt	atcttattat	aattgatgac gtttggcgct gggaagagtg ggaagtcac 720
agaaagtcca	ttccaaagaa	tgatctgggt agcagaataa tcatgactac tcgtcttaat 780
tcaatagctg	agaagtgtcg	caatgatgac atggatgcgt ttgtttatga aactgaggct 840
ctggattatg	tggatgcttg	gttgttgtgt gacaagtag caagaaagtc tgtcacaatg 900
atgaacatta	atccatgcta	tgatatcgtg gacatgtgct atggtatgcc gttagcacta 960
attcgtgtgt	cgtcagcatt	ggcagaagag atacaagctt tagacagtga tgaacggcaa 1020
atatggaggg	ctctgagacg	gtagaggat ggtattttgg acatcccatc cttgaagcca 1080
ttggcagaga	gtttatgcct	tggttacgac catcttctc tctatctgag gactttgttg 1140
ttatgttgta	gtgtgtacca	ttggcttgat ggtgggattg ttcaaagggg ccgtttggtc 1200
acaaggtgga	ttgctgaagg	atttgtttca gaagagaaag cagcagaagg ttactttgat 1260
gagcttgtcg	gcagaggatg	gatgaagcat agagagtga acgagtatga gatccaccct 1320
atgatgctgg	ccatccttag	atacaagtcg aaggagtaca attttgtaac ttgtttgggt 1380
acgggatctg	atacttgtac	tagtgcactc ctatctact cctctccaac aatggcgatt 1440
cgccggcttt	gtcttcaaag	gggtaccaca atgaaatgct tctcaagtat ggatgtgtca 1500
cacactcgca	gccttgtcat	ccttggcgac gtgataggag tccccttga tatgtttaa 1560
agattgcgag	tgttgacact	tgaagacaat ctcgatatag atgactctca cctgaagaag 1620
atatgtgagc	agctagagag	cctcaggtcg ctcaagtacc tgggtatcaa ggttacacgg 1680
atcactaagc	tcccacagga	gatacagaag ctgaagcacc tggagatttt gtacgtgagg 1740
agcacaggca	tcaaagagct	cccacgggag atcggggaat tgaacaact gcggactctg 1800
gacatgagga	acacgcggat	cagcgagctc ccgtcgcaga tcggggagct caaacatctg 1860
cggactctgg	acgtgagtaa	caacatgtgg aatatcagcg agctgccgtc gcaaatcggg 1920
gagctgaagc	atctacaaac	tctggatgtg aggaatacgt cggtagaga gctgccatcg 1980
caaatcgggg	agctgaagca	tctgcggact ctggatgtga ggaacacggg ggtgagagag 2040
ctgccatggc	aagctggcca	gatctcggga tcgctgcacg tgcatacaga tgacagtgc 2100

gagggcatgc ggctgccaga aggcgtatgc gaagatctga tcaagggtat tcccaaggct 2160  
 gagctcgcaa agtgcagtga ggtcctatcc atcaatattg tcgatcgttt aggatctccc 2220  
 cctattggca tattcaaggt tattggcttg cacaagagta tcccgaagct gatcaaagat 2280  
 catttcaatg ttctttcttc cctagacatc aggcggtaca acaagctaga ggaggatgac 2340  
 catgagtttc tagccaacaa tatgcctaac ctccagatgc ttgtactgag gttcgaggcc 2400  
 ccacaaagag agcccatcaa cattaaccgc acaggcttcc agatgatgga gagattcctt 2460  
 gtggagagcc ggggtgccacg gataacctc caggaaggag ccatgcccaa gctcaagcat 2520  
 ctcgagttca agttctacgc tggcccacca agcaaagatc ccataggaat caccacctc 2580  
 aagagcctcc aaaagtggt ctttcgctgc tccaatggt acaagagcga caaccctggc 2640  
 atcaaggctg ccattgacgt cgtgaagaaa gaagcaaggc agcatcccaa cgggccgatc 2700  
 agccttctca tcaactgaggc cgataaggag gtaccaata ttgaggcaca cgggagcagt 2760  
 gaaaacattg tcgttgcca cgetgctct gacgacgcca tcagttgctc tagctgcggc 2820  
 cgaaccagca ctagtatcca agaggaaca gtccgagatc gaataccagc tatggatttg 2880  
 ttctggccgg agtttaacag ctatgaaaaa gcaaaaagaa 2920

<210> 31

<211> 974

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 31

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

Thr Glu Gln Gln Phe Gln His Arg Ala Phe Val Ser Leu Ser Pro Gly  
 165 170 175  
 Ala Asn Leu Thr Asp Thr Leu Thr Asp Ile Leu Leu Gln Val Gly Ala  
 180 185 190  
 Tyr Asn Asp Asp Ala Thr Pro Tyr Cys Gly Thr Gly Thr Pro His Gln  
 195 200 205  
 Gln Tyr Leu Ile Asp Asn Ile Ser Ala Tyr Leu Ile Gly Lys Lys Tyr  
 210 215 220  
 Leu Ile Ile Ile Asp Asp Val Trp Arg Trp Glu Glu Trp Glu Val Ile  
 225 230 235 240  
 Arg Lys Ser Ile Pro Lys Asn Asp Leu Gly Ser Arg Ile Ile Met Thr  
 245 250 255  
 Thr Arg Leu Asn Ser Ile Ala Glu Lys Cys Arg Asn Asp Asp Met Asp  
 260 265 270  
 Ala Phe Val Tyr Glu Thr Glu Ala Leu Asp Tyr Val Asp Ala Trp Leu  
 275 280 285  
 Leu Cys Asp Lys Val Ala Arg Lys Ser Val Thr Cys Met Asn Ile Asn  
 290 295 300  
 Pro Cys Tyr Asp Ile Val Asp Met Cys Tyr Gly Met Pro Leu Ala Leu  
 305 310 315 320  
 Ile Arg Val Ser Ser Ala Leu Ala Glu Glu Ile Gln Ala Leu Asp Ser  
 325 330 335  
 Asp Glu Arg Gln Ile Trp Arg Ala Leu Arg Arg Val Glu Asp Gly Ile  
 340 345 350  
 Leu Asp Ile Pro Ser Leu Lys Pro Leu Ala Glu Ser Leu Cys Leu Gly  
 355 360 365  
 Tyr Asp His Leu Pro Leu Tyr Leu Arg Thr Leu Leu Leu Cys Cys Ser  
 370 375 380  
 Val Tyr His Trp Leu Asp Gly Gly Ile Val Gln Arg Gly Arg Leu Val  
 385 390 395 400  
 Thr Arg Trp Ile Ala Glu Gly Phe Val Ser Glu Glu Lys Ala Ala Glu  
 405 410 415  
 Gly Tyr Phe Asp Glu Leu Val Gly Arg Gly Trp Met Lys His Arg Glu  
 420 425 430  
 Leu Asn Glu Tyr Glu Ile His Pro Met Met Leu Ala Ile Leu Arg Tyr  
 435 440 445  
 Lys Ser Lys Glu Tyr Asn Phe Val Thr Cys Leu Gly Thr Gly Ser Asp  
 450 455 460  
 Thr Cys Thr Ser Ala Ser Leu Ser Tyr Ser Ser Pro Thr Met Ala Ile



465	470	475	480
Arg Arg Leu Cys	Leu Gln Arg Gly Tyr	Pro Met Lys Cys Phe	Ser Ser
	485	490	495
Met Asp Val Ser	His Thr Arg Ser	Leu Val Ile Leu Gly	Asp Val Ile
	500	505	510
Gly Val Pro Leu	Asp Met Phe Lys Arg	Leu Arg Val Leu Asp	Leu Glu
	515	520	525
Asp Asn Leu Asp	Ile Asp Asp Ser His	Leu Lys Lys Ile Cys	Glu Gln
	530	535	540
Leu Glu Ser Leu	Arg Leu Leu Lys Tyr	Leu Gly Ile Lys Gly	Thr Arg
545	550	555	560
Ile Thr Lys Leu	Pro Gln Glu Ile Gln	Lys Leu Lys His Leu	Glu Ile
	565	570	575
Leu Tyr Val Arg	Ser Thr Gly Ile Lys	Glu Leu Pro Arg	Glu Ile Gly
	580	585	590
Glu Leu Lys Gln	Leu Arg Thr Leu Asp	Met Arg Asn Thr Arg	Ile Ser
	595	600	605
Glu Leu Pro Ser	Gln Ile Gly Glu Leu	Lys His Leu Arg Thr	Leu Asp
	610	615	620
Val Ser Asn Asn	Met Trp Asn Ile Ser	Glu Leu Pro Ser Gln	Ile Gly
625	630	635	640
Glu Leu Lys His	Leu Gln Thr Leu Asp	Val Arg Asn Thr Ser	Val Arg
	645	650	655
Glu Leu Pro Ser	Gln Ile Gly Glu Leu	Lys His Leu Arg Thr	Leu Asp
	660	665	670
Val Arg Asn Thr	Gly Val Arg Glu Leu	Pro Trp Gln Ala Gly	Gln Ile
	675	680	685
Ser Gly Ser Leu	His Val His Thr Asp	Asp Ser Asp Glu Gly	Met Arg
	690	695	700
Leu Pro Glu Gly	Val Cys Glu Asp Leu	Ile Lys Gly Ile Pro	Lys Ala
705	710	715	720
Glu Leu Ala Lys	Cys Ser Glu Val Leu	Ser Ile Asn Ile Val	Asp Arg
	725	730	735
Leu Gly Ser Pro	Pro Ile Gly Ile Phe	Lys Val Ile Gly Leu	His Lys
	740	745	750
Ser Ile Pro Lys	Leu Ile Lys Asp His	Phe Asn Val Leu Ser	Ser Leu
	755	760	765
Asp Ile Arg Arg	Tyr Asn Lys Leu Glu	Glu Asp Asp His Glu	Phe Leu
	770	775	780

Ala Asn Asn Met Pro Asn Leu Gln Met Leu Val Leu Arg Phe Glu Ala	
785	790 795 800
Pro Gln Arg Glu Pro Ile Asn Ile Asn Arg Thr Gly Phe Gln Met Met	
	805 810 815
Glu Arg Phe Leu Val Glu Ser Arg Val Pro Arg Ile Thr Phe Gln Glu	
	820 825 830
Gly Ala Met Pro Lys Leu Lys His Leu Glu Phe Lys Phe Tyr Ala Gly	
	835 840 845
Pro Pro Ser Lys Asp Pro Ile Gly Ile Thr His Leu Lys Ser Leu Gln	
	850 855 860
Lys Val Val Phe Arg Cys Ser Lys Trp Tyr Lys Ser Asp Asn Pro Gly	
865	870 875 880
Ile Lys Ala Ala Ile Asp Val Val Lys Lys Glu Ala Arg Gln His Pro	
	885 890 895
Asn Arg Pro Ile Ser Leu Leu Ile Thr Glu Gly Asp Lys Glu Val Pro	
	900 905 910
Asn Ile Glu Ala His Gly Ser Ser Glu Asn Ile Val Val Val His Ala	
	915 920 925
Ala Pro Asp Asp Ala Ile Ser Cys Ser Ser Cys Gly Arg Thr Ser Thr	
	930 935 940
Ser Ile Gln Glu Gly Thr Val Arg Asp Arg Ile Pro Ala Met Asp Leu	
945	950 955 960
Phe Trp Pro Glu Phe Asn Ser Tyr Glu Lys Ala Lys Arg Asn	
	965 970

- <210> 32
- <211> 3378
- <212> DNA
- <213> 人工序列 (Artificial Sequence)
- <400> 32

atggccgtat acagcgtcgc cacgggggccc ttggtcccc tcctatcgaa gctctccgct	60
ttgctgggcg acgagcactt ggatcttgcg gagaggacct ggagcgacgc catgttcac	120
aggccccagc tggaggccgt gcaactctctc ctctcccga ggattagttg gggaatgacg	180
ggggaggaag tcgatgcttt gtgcaaggat gagttgatgg cggaggtgcg tgagctgtcc	240
tacgacatgg acgacgcat cgacgaattc ttcttagagg agcccatggc gggcggcgac	300
ggtggccctt tcgatgagct caagacaaga gttgaggatg tctccaagcg gttctccgac	360
agccggcggt ggagccacc ggtggagcaa cataacat ccctaaccgc cgcaaccgta	420
gactgtccac ctctcacgc tcgcttcgtc cacaacatga tggatgtgtc agagctcgtg	480
gagatggaca aacaacatga gaaagagctc atcaaattgc tggacaagg tgcggacaca	540
agcatatatg cttcccgggtg gcgcatcgca acaccatggc atgataagga gcaaagtact	600

gtggtcaagg	tgccgaaaag	agagtggggc	ttcccggaca	atcggaacag	tccatttata	660
tgggcgagtg	attcgtttga	acgattgcgt	tcgggaagtt	tgtgtggaga	tacgttgcgg	720
ttggatgggtg	aaggcgcgaa	catccgcaag	ctcttgtcca	ccctccggaa	taaggtgggc	780
cgcgcccagt	tgggtcaggt	cgaggataag	agaaaaaggg	tagaggaggc	gacgaagcct	840
tgtgaatttc	acgaggtcaa	aacaatatgc	atccttggat	tgcttggcgc	aggcaaaaca	900
actcttgcaa	aactgttga	ctcccatcac	tcaacgacag	agcagcaatt	ccaacaccgg	960
gctttcgtgt	cactctctcc	gggtgccaat	ctcaccgaca	ctcttactga	tattttattg	1020
caagtaggag	catataatga	tgatgcaaca	ccatattgtg	ggaccggaac	accgcaccaa	1080
cagtatctca	ttgacaatat	atcagcttat	ctcattggca	aaaagtatct	tattataatt	1140
gatgacgttt	ggcactggga	agagtgggaa	gtcatcagaa	agtccattcc	aaagaatgat	1200
ctgggtagca	gaataatcat	gactactcgt	cttaattcaa	tagctgagaa	gtgtcgcaat	1260
gatgacatgg	atgcgtttgt	ttatgaaact	gaggctctgg	attatgtgga	tgcttggttg	1320
ttgtgtgaca	aggtagcaag	aaagtctgtc	acatgtatga	acattaatcc	atgctatgat	1380
atcgtggaca	tgtgctatgg	tatgccgtta	gcactaatc	gtgtgtcgtc	agcattggca	1440
gaagagatac	aagctttaga	cagtgatgaa	cggcaaatat	ggagggctct	gagacgggta	1500
gaggatggta	ttttggacat	cccatccttg	aagccattgg	cagagagttt	atgccttggg	1560
tacgaccatc	ttcctctcta	tctgaggact	ttgttgttat	gttgtagtgt	gtaccattgg	1620
cttgatgggtg	ggattgttca	aaggggccgt	ttggtcacaa	ggtggattgc	tgaaggattt	1680
gtttcagaag	agaaagcagc	agaaggttac	tttgatgagc	ttgtcggcag	aggatggatg	1740
aagcatagag	ggttgaacga	gtatgagatc	caccctatga	tgctggccat	ccttagatac	1800
aaatcgaagg	agtacaatth	tgtaacttgt	ttgggtacgg	gatctgatac	ttgtactagt	1860
gcatctctat	cctactcctc	tccaacaatg	gcgattcgcc	ggctttgtct	tcaaaggggg	1920
tacccaatga	aatgcttctc	aagtatggat	gtgtcacaca	ctcgcagcct	tgtcatcctt	1980
ggcgacgtga	taggagtccc	cttggatatg	tttaaagat	tgcgagtgtt	ggaccttgaa	2040
gataatatcg	gtatagagga	ctcccacctg	aagaagatat	gtgagcagct	agagagcctc	2100
aggctgctca	agtacctagg	tctcaagggt	acgcgaatca	ctaagctccc	acaggagata	2160
cagaagctga	agcaactgga	gattttgtac	gtgaggagca	caggcatcga	agagctccca	2220
tgggagatcg	gggaattgaa	acaactgcgg	actctggacg	tgaggaacac	gcggatcagc	2280
gagctcccgt	cgcagatcgg	ggagctcaaa	catctgcgga	ctctggacgt	gagtaacatg	2340
tggaatatca	gcgagctgcc	gtcgcaaatc	ggggagctga	agcatctaca	aactctggat	2400
gtgaggaaca	cgctcagtga	agagctgcca	tcgcaaatcg	gggagctgaa	gcatctgcgg	2460
actctggatg	tgaggaacac	gggggtgaga	gagctgceat	ggcaagctgg	ccagatctcg	2520
ggatcgctgc	acgtgcatac	agatgacagt	gacgagggca	tgcggtgcc	agaaggcgta	2580
tgcgaagatc	tgatcaaggg	tattcccaag	gctgagctcg	caaagtgcag	tgaggtccta	2640
tccatcaata	ttgtcgatcg	tttaggatct	ccccctattg	gcatattcaa	ggttattggc	2700
ttgcacaaga	gtatcccgaa	gctgatcaaa	gatcatttca	atgttctttc	ttccctagac	2760
atcaggcggg	acaacaagct	agaggaggat	gacatgagt	ttctagccaa	caatatgcct	2820
aacctccaga	tgcttgtact	gaggttcgag	gccccacaaa	gagagcccat	catcattaac	2880
cgcacaggct	tccagatgct	ggagagattc	cttgtggaga	gccgggtgcc	acggataacc	2940

ttccaggaag gagccatgcc caagctcaag catctcgagt ttaagttcta cgctggccca 3000  
 ccaagcaaaag atcccatagg aatcacccac ctcaagagcc tccaaaaggt ggtcttttcgc 3060  
 tgctccaaat ggtacaagag cgacaaccct ggcatcaagg ctgccattga cgtcgtgaag 3120  
 aaagaagcaa ggcagcatcc caaccggccg atcagccttc tcactactga gggcgataag 3180  
 gaggtaccga atattgaggc acacgggagc agtgaaaaca ttgtcgttgt ccacgctgct 3240  
 cctgacgacg ccatcagttg ctctagctgc ggccgaacca gcactagtat ccaagaggga 3300  
 acagtccgag atcgaatacc agctatggat ttgttctggc cggagtttaa cagctatgaa 3360  
 aaagcaaaaa gaaactag 3378

<210> 33

<211> 1125

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 33

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

Trp Gly Phe Pro Asp Asn Arg Asn Ser Pro Phe Ile Trp Ala Ser Asp  
 210 215 220  
 Ser Phe Glu Arg Leu Arg Ser Gly Ser Leu Cys Gly Asp Thr Leu Arg  
 225 230 235 240  
 Leu Asp Gly Glu Gly Ala Asn Ile Arg Lys Leu Leu Ser Thr Leu Arg  
 245 250 255  
 Asn Lys Val Gly Arg Ala Gln Leu Val Gln Val Glu Asp Lys Arg Lys  
 260 265 270  
 Arg Val Glu Glu Ala Thr Lys Pro Cys Glu Phe His Glu Val Lys Thr  
 275 280 285  
 Ile Cys Ile Leu Gly Leu Pro Gly Ala Gly Lys Thr Thr Leu Ala Lys  
 290 295 300  
 Leu Leu Tyr Ser His His Ser Thr Thr Glu Gln Gln Phe Gln His Arg  
 305 310 315 320  
 Ala Phe Val Ser Leu Ser Pro Gly Ala Asn Leu Thr Asp Thr Leu Thr  
 325 330 335  
 Asp Ile Leu Leu Gln Val Gly Ala Tyr Asn Asp Asp Ala Thr Pro Tyr  
 340 345 350  
 Cys Gly Thr Gly Thr Pro His Gln Gln Tyr Leu Ile Asp Asn Ile Ser  
 355 360 365  
 Ala Tyr Leu Ile Gly Lys Lys Tyr Leu Ile Ile Ile Asp Asp Val Trp  
 370 375 380  
 His Trp Glu Glu Trp Glu Val Ile Arg Lys Ser Ile Pro Lys Asn Asp  
 385 390 395 400  
 Leu Gly Ser Arg Ile Ile Met Thr Thr Arg Leu Asn Ser Ile Ala Glu  
 405 410 415  
 Lys Cys Arg Asn Asp Asp Met Asp Ala Phe Val Tyr Glu Thr Glu Ala  
 420 425 430  
 Leu Asp Tyr Val Asp Ala Trp Leu Leu Cys Asp Lys Val Ala Arg Lys  
 435 440 445  
 Ser Val Thr Cys Met Asn Ile Asn Pro Cys Tyr Asp Ile Val Asp Met  
 450 455 460  
 Cys Tyr Gly Met Pro Leu Ala Leu Ile Arg Val Ser Ser Ala Leu Ala  
 465 470 475 480  
 Glu Glu Ile Gln Ala Leu Asp Ser Asp Glu Arg Gln Ile Trp Arg Ala  
 485 490 495  
 Leu Arg Arg Val Glu Asp Gly Ile Leu Asp Ile Pro Ser Leu Lys Pro  
 500 505 510  
 Leu Ala Glu Ser Leu Cys Leu Gly Tyr Asp His Leu Pro Leu Tyr Leu

515	520	525
Arg Thr Leu Leu Leu Cys Cys Ser Val Tyr His Trp Leu Asp Gly Gly		
530	535	540
Ile Val Gln Arg Gly Arg Leu Val Thr Arg Trp Ile Ala Glu Gly Phe		
545	550	555
Val Ser Glu Glu Lys Ala Ala Glu Gly Tyr Phe Asp Glu Leu Val Gly		
565	570	575
Arg Gly Trp Met Lys His Arg Gly Leu Asn Glu Tyr Glu Ile His Pro		
580	585	590
Met Met Leu Ala Ile Leu Arg Tyr Lys Ser Lys Glu Tyr Asn Phe Val		
595	600	605
Thr Cys Leu Gly Thr Gly Ser Asp Thr Cys Thr Ser Ala Ser Leu Ser		
610	615	620
Tyr Ser Ser Pro Thr Met Ala Ile Arg Arg Leu Cys Leu Gln Arg Gly		
625	630	635
Tyr Pro Met Lys Cys Phe Ser Ser Met Asp Val Ser His Thr Arg Ser		
645	650	655
Leu Val Ile Leu Gly Asp Val Ile Gly Val Pro Leu Asp Met Phe Lys		
660	665	670
Arg Leu Arg Val Leu Asp Leu Glu Asp Asn Ile Gly Ile Glu Asp Ser		
675	680	685
His Leu Lys Lys Ile Cys Glu Gln Leu Glu Ser Leu Arg Leu Leu Lys		
690	695	700
Tyr Leu Gly Leu Lys Gly Thr Arg Ile Thr Lys Leu Pro Gln Glu Ile		
705	710	715
Gln Lys Leu Lys Gln Leu Glu Ile Leu Tyr Val Arg Ser Thr Gly Ile		
725	730	735
Glu Glu Leu Pro Trp Glu Ile Gly Glu Leu Lys Gln Leu Arg Thr Leu		
740	745	750
Asp Val Arg Asn Thr Arg Ile Ser Glu Leu Pro Ser Gln Ile Gly Glu		
755	760	765
Leu Lys His Leu Arg Thr Leu Asp Val Ser Asn Met Trp Asn Ile Ser		
770	775	780
Glu Leu Pro Ser Gln Ile Gly Glu Leu Lys His Leu Gln Thr Leu Asp		
785	790	795
Val Arg Asn Thr Ser Val Arg Glu Leu Pro Ser Gln Ile Gly Glu Leu		
805	810	815
Lys His Leu Arg Thr Leu Asp Val Arg Asn Thr Gly Val Arg Glu Leu		
820	825	830

Pro Trp Gln Ala Gly Gln Ile Ser Gly Ser Leu His Val His Thr Asp  
 835 840 845  
 Asp Ser Asp Glu Gly Met Arg Leu Pro Glu Gly Val Cys Glu Asp Leu  
 850 855 860  
 Ile Lys Gly Ile Pro Lys Ala Glu Leu Ala Lys Cys Ser Glu Val Leu  
 865 870 875 880  
 Ser Ile Asn Ile Val Asp Arg Leu Gly Ser Pro Pro Ile Gly Ile Phe  
 885 890 895  
 Lys Val Ile Gly Leu His Lys Ser Ile Pro Lys Leu Ile Lys Asp His  
 900 905 910  
 Phe Asn Val Leu Ser Ser Leu Asp Ile Arg Arg Tyr Asn Lys Leu Glu  
 915 920 925  
 Glu Asp Asp His Glu Phe Leu Ala Asn Asn Met Pro Asn Leu Gln Met  
 930 935 940  
 Leu Val Leu Arg Phe Glu Ala Pro Gln Arg Glu Pro Ile Ile Ile Asn  
 945 950 955 960  
 Arg Thr Gly Phe Gln Met Leu Glu Arg Phe Leu Val Glu Ser Arg Val  
 965 970 975  
 Pro Arg Ile Thr Phe Gln Glu Gly Ala Met Pro Lys Leu Lys His Leu  
 980 985 990  
 Glu Phe Lys Phe Tyr Ala Gly Pro Pro Ser Lys Asp Pro Ile Gly Ile  
 995 1000 1005  
 Thr His Leu Lys Ser Leu Gln Lys Val Val Phe Arg Cys Ser Lys Trp  
 1010 1015 1020  
 Tyr Lys Ser Asp Asn Pro Gly Ile Lys Ala Ala Ile Asp Val Val Lys  
 1025 1030 1035 1040  
 Lys Glu Ala Arg Gln His Pro Asn Arg Pro Ile Ser Leu Leu Ile Thr  
 1045 1050 1055  
 Glu Gly Asp Lys Glu Val Pro Asn Ile Glu Ala His Gly Ser Ser Glu  
 1060 1065 1070  
 Asn Ile Val Val Val His Ala Ala Pro Asp Asp Ala Ile Ser Cys Ser  
 1075 1080 1085  
 Ser Cys Gly Arg Thr Ser Thr Ser Ile Gln Glu Gly Thr Val Arg Asp  
 1090 1095 1100  
 Arg Ile Pro Ala Met Asp Leu Phe Trp Pro Glu Phe Asn Ser Tyr Glu  
 1105 1110 1115 1120  
 Lys Ala Lys Arg Asn  
 1125

<210> 34





275	280	285
Asp Asp Ala Thr Pro Tyr Cys Gly Thr Gly Thr Pro His Gln Gln Tyr		
290	295	300
Leu Ile Asp Asn Ile Ser Ala Tyr Leu Ile Gly Lys Lys Tyr Leu Ile		
305	310	315
Ile Ile Asp Asp Val Trp Arg Trp Glu Glu Trp Glu Val Ile Arg Lys		
325	330	335
Ser Ile Pro Lys Asn Asp Leu Gly Ser Arg Ile Ile Met Thr Thr Arg		
340	345	350
Leu Asn Ser Ile Ala Glu Lys Cys Arg Asn Asp Asp Met Asp Ala Phe		
355	360	365
Val Tyr Glu Thr Glu Ala Leu Asp Tyr Val Asp Ala Trp Leu Leu Cys		
370	375	380
Asp Lys Val Ala Arg Lys Ser Val Thr Cys Met Asn Ile Asn Pro Cys		
385	390	395
Tyr Asp Ile Val Asp Val Cys Tyr Gly Met Pro Leu Ala Leu Ile Arg		
405	410	415
Val Ser Ser Ala Leu Ala Glu Glu Ile Gln Ala Leu Asp Ser Asp Glu		
420	425	430
Trp Gln Ile Trp Arg Ala Leu Arg Arg Val Glu Asp Gly Ile Leu Asp		
435	440	445
Ile Pro Ser Leu Lys Pro Leu Ala Glu Ser Leu Cys Leu Gly Tyr Asp		
450	455	460
His Leu Pro Leu Tyr Leu Arg Thr Leu Leu Leu Cys Cys Ser Val Tyr		
465	470	475
His Trp Leu Asp Gly Gly Ile Val Gln Arg Gly Arg Leu Val Thr Arg		
485	490	495
Trp Ile Ala Glu Gly Phe Val Ser Glu Glu Lys Ala Ala Glu Gly Tyr		
500	505	510
Phe Asp Glu Leu Val Asp Arg Gly Trp Ile Lys His Arg Gly Trp Asn		
515	520	525
Glu Tyr Glu Ile Tyr Pro Met Met Leu Ala Ile Leu Arg Tyr Lys Ser		
530	535	540
Lys Glu Tyr Asn Phe Val Thr Cys Leu Gly Thr Gly Phe Asp Thr Cys		
545	550	555
Thr Ser Ala Ser Leu Ser Tyr Ser Ser Pro Thr Met Ala Ile Arg Arg		
565	570	575
Leu Cys Leu Gln Arg Gly Tyr Pro Met Lys Cys Phe Ser Ser Met Asp		
580	585	590

Val Ser His Thr Arg Ser Leu Val Ile Leu Gly Asp Val Ile Gly Val		
595	600	605
Pro Leu Asp Met Phe Lys Arg Leu Arg Val Leu Asp Leu Glu Asp Asn		
610	615	620
Ile Gly Ile Glu Asp Ser His Leu Lys Lys Ile Cys Glu Gln Leu Glu		
625	630	635
Ser Leu Arg Leu Leu Lys Tyr Leu Gly Leu Lys Gly Thr Arg Ile Thr		
	645	650
		655
Lys Leu Pro Gln Glu Ile Gln Lys Leu Lys His Leu Glu Ile Leu Tyr		
	660	665
		670
Val Arg Ser Thr Gly Ile Lys Glu Leu Pro Arg Glu Ile Gly Glu Val		
	675	680
		685
Lys Gln Leu Arg Thr Leu Asp Val Arg Asn Thr Arg Ile Ser Glu Leu		
	690	695
		700
Pro Ser Gln Ile Gly Glu Leu Lys His Leu Arg Thr Leu Asp Val Arg		
705	710	715
		720
Asn Thr Arg Ile Ser Glu Leu Leu Ser Gln Ile Gly Glu Leu Lys His		
	725	730
		735
Leu Arg Thr Leu Asp Val Arg Asn Thr Arg Ile Ser Glu Leu Pro Ser		
	740	745
		750
Gln Ile Gly Glu Leu Lys His Leu Arg Thr Leu Asp Val Arg Asn Thr		
	755	760
		765
Arg Thr Ser Ile Phe Phe Tyr Ser Arg Arg Arg Ile Lys Lys Tyr Arg		
	770	775
		780
Ser Thr Asp Ile Trp Leu Ser Ala Arg Asp Met His His Thr Cys Ile		
785	790	795
		800
Trp Tyr		

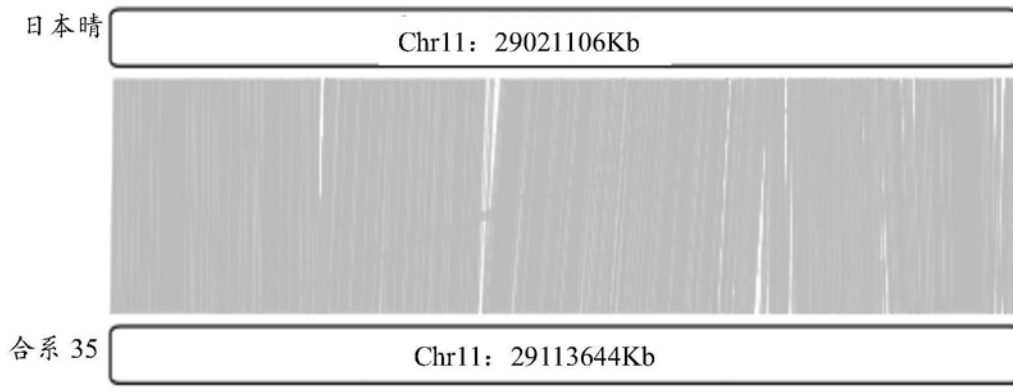


图1

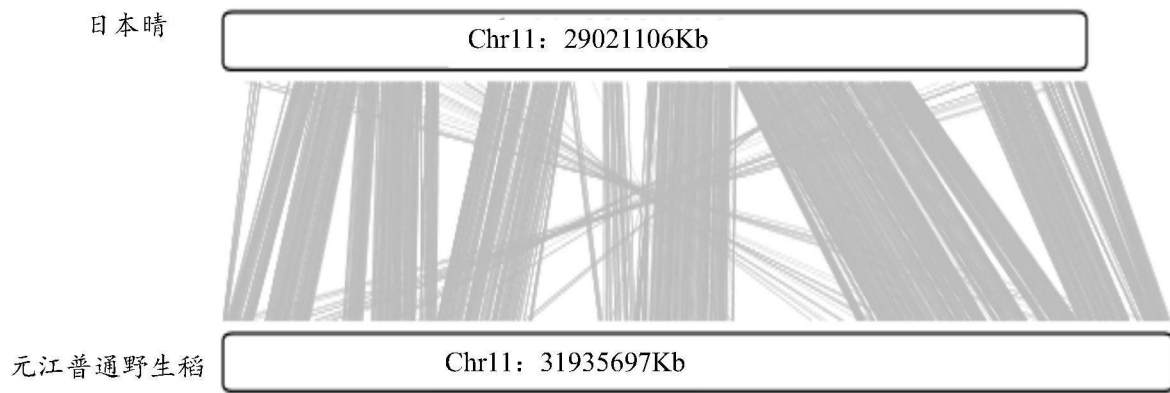


图2

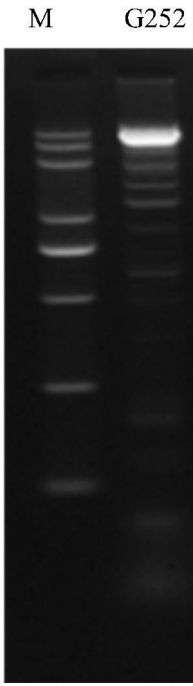


图3

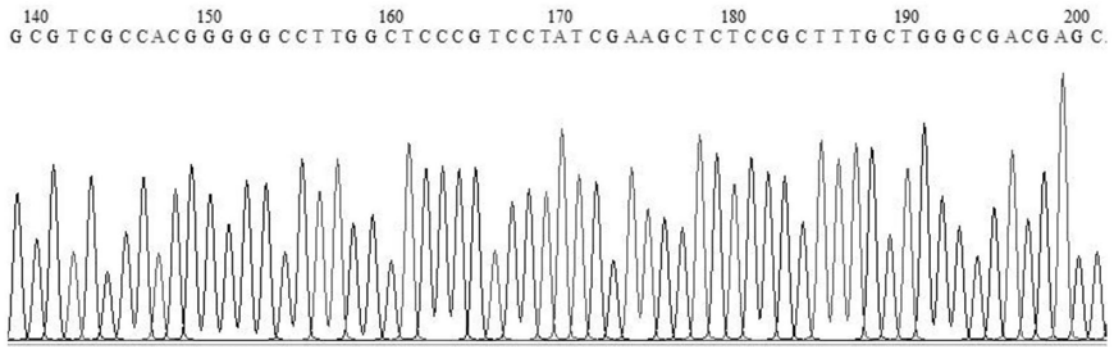


图4

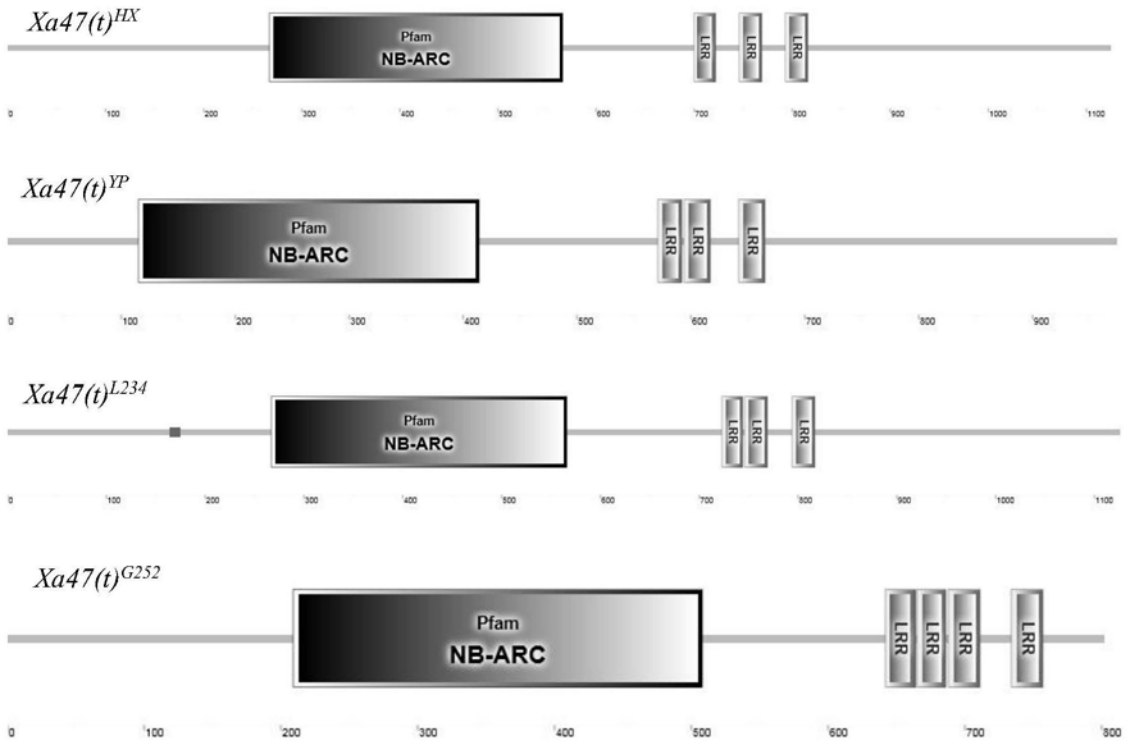


图5

Xa47_HX	MNAVYSVATGALAPVLSKLSALLGDEHLDAERTSRDAMFIRSQLEAVHSLLLPRIISWGMTGEEVDALCKDELMAEVRELSYDMDDAIDEFFLEEPMAGGDDGGPFDELKTRVEDVSKRFS	120
Xa47_YP	.....	0
Xa47_L234	MNAVYSVATGALAPVLSKLSALLGDEHLDAERTSRDAMFIRSQLEAVHSLLLPRIISWGMTGEEVDALCKDELMAEVRELSYDMDDAIDEFFLEEPMAGGDDGGPFDELKTRVEDVSKRFS	120
Consensus		
Xa47_HX	SRRWRRPPEVQHQPSTLAATVDCPPPHARFVHMMDVSELVEMDKHEKELIKLLEQGADTSIYASRWRIATPWHDKBQSTVVKVPEREWGFPDNRNSPFIWASDSFERLRGSLCGIT..	238
Xa47_YP	.....MMDVSELVEMDKHEKELIKLLEQGADTSIYASRWRIATPWHDKBQSTVVKVPERE...DMYDDALHNAVSVLSHGVPSSGASGDCSR	86
Xa47_L234	SRRWRRPPEVQHQPSTLAATVDCPPPHARFVHMMDVSELVEMDKHEKELIKLLEQGADTSIYASRWRIATPWHDKBQSTVVKVPERE...DMYDDALHNAVSVLSHGVPSSGASGDCSR	238
Consensus	mmdvselvemdkehkeliklleqgadtisyasrwriatpwhdkeqstvkvperewgfpdnrnsppiwasdsferlrgslcgit...dmyddalhnnavsvlshgvpssgsgdcscr	
Xa47_HX	LRLDGEGANIRKLLSTLRNKVGRQQLVQVEDKRRKVEEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHSSTTEQQFQHRAFVLSPGANLTDLTDLILLQVGAYNDATPYCGTGTP	358
Xa47_YP	LRLDGEGANIRKLLSTLRNKVGRQQLVQVEDKRRKVEEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHSSTTEQQFQHRAFVLSPGANLTDLTDLILLQVGAYNDATPYCGTGTP	206
Xa47_L234	LRLDGEGANIRKLLSTLRNKVGRQQLVQVEDKRRKVEEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHSSTTEQQFQHRAFVLSPGANLTDLTDLILLQVGAYNDATPYCGTGTP	358
Consensus	lrdgeganirkllstlrnkvgraqqlvqvedkrrkveeatkpcfhevkticilglpgagkttlakllyshsstteqqfghrafvlspganltdltdillqvgayndatpycgtgtp	
Xa47_HX	HQQYLIDNISAYLIGKYYLIIIDVVRWEWEVIRKSIKPNDLGSRIMTTRLNSTAEKCRNDMDAFVYETEALDYDAWLLCDKVARVSVTCMNPICYDIVMCMPLALIRVSSA	478
Xa47_YP	HQQYLIDNISAYLIGKYYLIIIDVVRWEWEVIRKSIKPNDLGSRIMTTRLNSTAEKCRNDMDAFVYETEALDYDAWLLCDKVARVSVTCMNPICYDIVMCMPLALIRVSSA	326
Xa47_L234	HQQYLIDNISAYLIGKYYLIIIDVVRWEWEVIRKSIKPNDLGSRIMTTRLNSTAEKCRNDMDAFVYETEALDYDAWLLCDKVARVSVTCMNPICYDIVMCMPLALIRVSSA	478
Consensus	hqqylidnisayligkyyliiidvvrweewirksipkndlgsriimttrlnstaeckrndmdafvetealdydwllcdkvarvsvtcmninpicydivmcmplalirvssa	
Xa47_HX	LAEEIQALDSDERQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHLDGGIVQGRGLVTRWIAEGFVSEKAAEGYFDELVGRGWMKHLNEYEIHPMMLAIL	598
Xa47_YP	LAEEIQALDSDERQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHLDGGIVQGRGLVTRWIAEGFVSEKAAEGYFDELVGRGWMKHLNEYEIHPMMLAIL	446
Xa47_L234	LAEEIQALDSDERQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHLDGGIVQGRGLVTRWIAEGFVSEKAAEGYFDELVGRGWMKHLNEYEIHPMMLAIL	598
Consensus	laeeiqaldsderqiwralrrvedgildipskplaeslclgydhplylrtlccsvyhlldggivqgrglvtrwiaegfvseekaegyfdelvgrgwmkhlneyeihpmmail	
Xa47_HX	RYKSKYFNVTCLGTGSDTCSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDVSHRSLVILGDIVGVLDMFKRLRVLDLEDNLDIDSHLKKICEQLESRLKYLIGKTRITKLPQ	718
Xa47_YP	RYKSKYFNVTCLGTGSDTCSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDVSHRSLVILGDIVGVLDMFKRLRVLDLEDNLDIDSHLKKICEQLESRLKYLIGKTRITKLPQ	566
Xa47_L234	RYKSKYFNVTCLGTGSDTCSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDVSHRSLVILGDIVGVLDMFKRLRVLDLEDNLDIDSHLKKICEQLESRLKYLIGKTRITKLPQ	718
Consensus	rykskeyfnvtclgtgsdtsaslsyssptmairrclqrgypmkcfssmdvshrsilvilgdivgvlmfmkrlrvldlednldidshlkkiceqlesrlkyligktritklpq	
Xa47_HX	EIQKLELILYVRSTGIEELPWEIGELKQLRRLDVRNTRISELPSQIGELKHLRRLDVSNNMNISELPSQIGELKHLQTLDRVNTSVRELPSQIGELKHLRRLDVRNTGVRELPOAG	837
Xa47_YP	EIQKLELILYVRSTGIEELPWEIGELKQLRRLDVRNTRISELPSQIGELKHLRRLDVSNNMNISELPSQIGELKHLQTLDRVNTSVRELPSQIGELKHLRRLDVRNTGVRELPOAG	686
Xa47_L234	EIQKLELILYVRSTGIEELPWEIGELKQLRRLDVRNTRISELPSQIGELKHLRRLDVSNNMNISELPSQIGELKHLQTLDRVNTSVRELPSQIGELKHLRRLDVRNTGVRELPOAG	838
Consensus	eigkllleilyvrstgieelpweigelkqlrtrldvrntriselpsqigelkhlrtlrvsnmnielpsqigelkhlqtldrvntsvrelpsqigelkhlrtldrvntgvrelpwag	
Xa47_HX	QISGSLHVHTDDSDGMRLEPGVCEDLIKGIPKAELAKCSEVLSINIVDRLGSPPIGPKVIGLHKSIPKLIKDHFNVLSSLDIRRYNKEEDDHEFLANMNPQLMVLRFEPAPQREPI	957
Xa47_YP	QISGSLHVHTDDSDGMRLEPGVCEDLIKGIPKAELAKCSEVLSINIVDRLGSPPIGPKVIGLHKSIPKLIKDHFNVLSSLDIRRYNKEEDDHEFLANMNPQLMVLRFEPAPQREPI	806
Xa47_L234	QISGSLHVHTDDSDGMRLEPGVCEDLIKGIPKAELAKCSEVLSINIVDRLGSPPIGPKVIGLHKSIPKLIKDHFNVLSSLDIRRYNKEEDDHEFLANMNPQLMVLRFEPAPQREPI	958
Consensus	qisgslhvhtddsdegmrlepgvcedlikgipkaelakcsevlsvinivdrlgspbigpkvighksipklikdhfnvlsldirrynkeeddheflanmnpqlmvlrfeapqrepi	
Xa47_HX	IINRTGFQMLERFLVESRVPRITPQEGAMPKPKHLEFKFYAGPPSKDPGITHLSLQKVVFRCSKWKYSDNPGIKAAIDVVKKEARQHPNRPSILLITEGDKVNPVIEAHGSSSENIVVV	1077
Xa47_YP	IINRTGFQMLERFLVESRVPRITPQEGAMPKPKHLEFKFYAGPPSKDPGITHLSLQKVVFRCSKWKYSDNPGIKAAIDVVKKEARQHPNRPSILLITEGDKVNPVIEAHGSSSENIVVV	926
Xa47_L234	IINRTGFQMLERFLVESRVPRITPQEGAMPKPKHLEFKFYAGPPSKDPGITHLSLQKVVFRCSKWKYSDNPGIKAAIDVVKKEARQHPNRPSILLITEGDKVNPVIEAHGSSSENIVVV	1078
Consensus	inrtgfmqlerflvesrvpritpqegampkpkhlefkfyagppskdpigithlskqvvrscskwyksdnpigikaidvkkearqhpnrpsillitegdkevpnieahgssenivvv	
Xa47_HX	HAAPDDAISCSGCGRTSTSIQEGTVRDRIPAMDLPWFENFSEYAKKR	1124
Xa47_YP	HAAPDDAISCSGCGRTSTSIQEGTVRDRIPAMDLPWFENFSEYAKKR	973
Xa47_L234	HAAPDDAISCSGCGRTSTSIQEGTVRDRIPAMDLPWFENFSEYAKKR	1125
Consensus	haapddaiscscgtrstsiqegtvrdripamdlpwfefnsyekakr	

图6

Xa47 HX	MAVYSVATGALAPVLSKLSALLGDEHLDAERTRSDAMPFIRSQLEAVHSLLLPRISSWMTGEEVDALCKDELMAEVRELSYDMDDAIDEFPLEEPMAGDGGPFDELKTRVEVSKRFS	120
Xa47 YP	.....	0
Xa47 G252	.....MTGEEVDALCKDELMAEVRELSYDMDDAIDEFPLEEPMAGDGGPFDELKTRVEVSKRFS	62
Consensus	.....	
Xa47 HX	SRWRPVEQHQPSLTAATVDCPPHARFVHNMDVSELVEMDKHEKELIKLLEQGAOTSIYASRWRIATFWHDKEQSTVVKVPEREWGFPINRNSPFIWASDFERLRSGSLCGT...	238
Xa47 YP	.....MMDVSELVEMDKHEKELIKLLEQGAOTSIYASRWRIATFWHDKEQSTVVKVPEKRR...DMYDDALHMAVSSLHGVPGGGSGDCSR	86
Xa47 G252	SRWRPVEQHQPSLTAATVDCPPHARFVHNMDVSELVEMDKHEKELIKLLEQGAOTSIYASRWRIATFWHDKEQSTVVKVPEKRR...DMYDDALHMAVSSLHGVPGGGSGDCSR	180
Consensus	mmdvselvemdkeheliklleqgadtstiyasrwriatfwhdkeqs vvkvp d w a s s g g d	
Xa47 HX	LDLDEGANIRKLLSTLRNKVGRAGLQVQVEDKRRKRVBEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHHSTTEQQQHRAFVLSLPGANLDTLTDILLQVGAYNDDATPYCGTGT	358
Xa47 YP	LDLDEGANIRKLLSTLRNKVGRAGLQVQVEDKRRKRVBEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHHSTTEQQQHRAFVLSLPGANLDTLTDILLQVGAYNDDATPYCGTGT	206
Xa47 G252	LDLDEGANIRKLLSTLRNKVGRAGLQVQVEDKRRKRVBEATKPCFHEVKTICILGLPGAGKTTLAKLLYSHHSTTEQQQHRAFVLSLPGANLDTLTDILLQVGAYNDDATPYCGTGT	300
Consensus	lddeganirkllstlrnkvgraqlvqvvedkrkrvbeatkpcfhevkticilgpgagkttlakllyshhstteqqqhrafvslsplanltdtldillqvgaynddatpycgtgt	
Xa47 HX	HQQYLIDNISAYLIGKYYLIIIDDVWHEWEVIRKSIKPNDLGSRIMTTRINSIAEKCRNDMDAFVYETEALDYVDAMLLCDKVARVSVTCMNPICYDIVMCGMPLALIRVSSA	478
Xa47 YP	HQQYLIDNISAYLIGKYYLIIIDDVWHEWEVIRKSIKPNDLGSRIMTTRINSIAEKCRNDMDAFVYETEALDYVDAMLLCDKVARVSVTCMNPICYDIVMCGMPLALIRVSSA	326
Xa47 G252	HQQYLIDNISAYLIGKYYLIIIDDVWHEWEVIRKSIKPNDLGSRIMTTRINSIAEKCRNDMDAFVYETEALDYVDAMLLCDKVARVSVTCMNPICYDIVMCGMPLALIRVSSA	420
Consensus	hqqylidnisayligkyyliiddvwheewevirkspkndlgsriimttrlnsiaekcrndmdafvetealdyvdamllcdkvarvsvtcmnncydivd cymplalirvssa	
Xa47 HX	LAEEIQALDSDEQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHWLDGGIVQGRGLVTRWIAEGFVSEKAAGYFDELVRGWMKRGNEYEIHPMLLAIL	598
Xa47 YP	LAEEIQALDSDEQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHWLDGGIVQGRGLVTRWIAEGFVSEKAAGYFDELVRGWMKRGNEYEIHPMLLAIL	446
Xa47 G252	LAEEIQALDSDEQIWRALRRVEDGILDIPSLKPLAESLCLGYDHLPLYLRTLLCCSVYHWLDGGIVQGRGLVTRWIAEGFVSEKAAGYFDELVRGWMKRGNEYEIHPMLLAIL	540
Consensus	laeeiqaldsdeqiwralrrvedgildipskplaeslclgydhplplylrtlccsvyhwldggivqgrglvtrwiaegfvsekaaggyfdelvrwgmkrgneyei pmlail	
Xa47 HX	RYKSKEYNFVTCLGTFDTCTSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDSVSHRSLVILGDVIGVPLMFKRRLVLDLEDNIGLDSHLKKICEQLESRLLLKYLGLKTRITKLPQ	718
Xa47 YP	RYKSKEYNFVTCLGTFDTCTSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDSVSHRSLVILGDVIGVPLMFKRRLVLDLEDNIGLDSHLKKICEQLESRLLLKYLGLKTRITKLPQ	566
Xa47 G252	RYKSKEYNFVTCLGTFDTCTSASLSYSSPTMAIRRLCLQRGYPMKCFSSMDSVSHRSLVILGDVIGVPLMFKRRLVLDLEDNIGLDSHLKKICEQLESRLLLKYLGLKTRITKLPQ	660
Consensus	rykskeynfvtclgtf dtctsaslsyssptmairrclqrgypmkcfssm dsvshrs lvi lgdvigvpldmfkrirvldledn i dshlkkiceqlesrlllkylg kqtritklpq	
Xa47 HX	EIQKLELEILYVRSTGIELEPWEIGELKQLRTLDMRNTRISELPSQIGELKHLRTLQVSNMNISELPSQIGELKHLTLQVSRVRELSQIGELKHLRTLQVSRVRELPWQAG	837
Xa47 YP	EIQKLELEILYVRSTGIELEPWEIGELKQLRTLDMRNTRISELPSQIGELKHLRTLQVSNMNISELPSQIGELKHLTLQVSRVRELSQIGELKHLRTLQVSRVRELPWQAG	686
Xa47 G252	EIQKLELEILYVRSTGIELEPWEIGELKQLRTLDMRNTRISELPSQIGELKHLRTLQVSNMNISELPSQIGELKHLTLQVSRVRELSQIGELKHLRTLQVSRVRELPWQAG	778
Consensus	eieklkleilyvrstgielepweigelkqlrtldmrntriselpsqigelkhlrtlqvs n n isel psqigelkhl tldvrnt svrelpsqigelkhlrtlqvs r vrelpwqag	
Xa47 HX	QISGSLHVHTDDDEGRLEPFGVCEDLIKGPKAELAKCSEVLSINIVDRLGSPPIGIFKVIKGLHKSIPKLIKDHFNVLSSLDIRRYNKLLEDDEHFLANMNPQLMVLRFEPAPQREPI	957
Xa47 YP	QISGSLHVHTDDDEGRLEPFGVCEDLIKGPKAELAKCSEVLSINIVDRLGSPPIGIFKVIKGLHKSIPKLIKDHFNVLSSLDIRRYNKLLEDDEHFLANMNPQLMVLRFEPAPQREPI	806
Xa47 G252	RKKYRSTDIWLRARMEHTCIWY.....	802
Consensus	i s m	
Xa47 HX	IINRTGFQMLERFLVESRVPRIITFQEGAMPKLEHLEFKFYAGPPSKDPIGITHLKSQKVVFRCSKWYKSDNPGIAAIDVVKKEARQHPNRPISLLITEGDKVEPNIEAHGSSSENIWVV	1077
Xa47 YP	IINRTGFQMLERFLVESRVPRIITFQEGAMPKLEHLEFKFYAGPPSKDPIGITHLKSQKVVFRCSKWYKSDNPGIAAIDVVKKEARQHPNRPISLLITEGDKVEPNIEAHGSSSENIWVV	926
Xa47 G252	.....	802
Consensus	.....	
Xa47 HX	HAAPDDAISCSSCGRTSTSIQEGTVRDRIIPAMDLPWPEFNSYKAKR	1124
Xa47 YP	HAAPDDAISCSSCGRTSTSIQEGTVRDRIIPAMDLPWPEFNSYKAKR	973
Xa47 G252	.....	802
Consensus	.....	

图7