



(12)发明专利申请

(10)申请公布号 CN 110938610 A

(43)申请公布日 2020.03.31

(21)申请号 201911295765.1

C12Q 1/6806(2018.01)

(22)申请日 2019.12.16

C12N 15/10(2006.01)

(71)申请人 广东菲鹏生物有限公司

C40B 50/06(2006.01)

地址 523000 广东省东莞市松山湖高新技术
产业开发区花莲路5号1号厂房7楼

G01N 33/68(2006.01)

G01N 33/53(2006.01)

(72)发明人 张志毅 章瑞程 贺婷 杜翔斐
钟淑瑶

(74)专利代理机构 北京超凡宏宇专利代理事务
所(特殊普通合伙) 11463

代理人 周文波

(51)Int.Cl.

C12N 9/12(2006.01)

C12N 15/54(2006.01)

C12N 15/62(2006.01)

C07K 19/00(2006.01)

权利要求书2页 说明书9页
序列表15页 附图1页

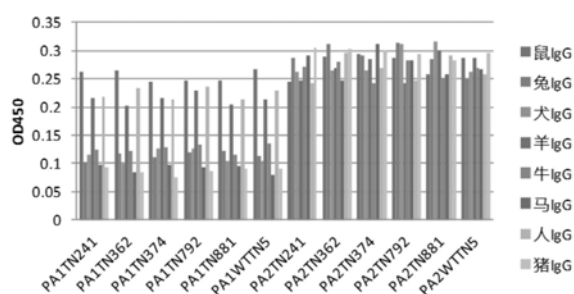
(54)发明名称

一种转座酶突变体、融合蛋白、其制备方法和应用

(57)摘要

本发明公开了一种转座酶突变体、融合蛋白、其制备方法和应用,涉及生物技术领域,所述转座酶突变体相较于野生型转座酶具有至少一种突变形式,所述野生型转座酶的氨基酸序列如SEQ ID No.1所示,该转座酶突变体的稳定性高,能长期于-20℃的条件下储存,可以显著降低储存成本,使转座酶或其构建的融合蛋白的使用更方便,简化实验流程。

与抗体结合能力



1. 一种转座酶突变体,其特征在于,所述转座酶突变体相较于野生型转座酶具有如下突变中的至少一种或多种的组合:R27T、M56G、Y41R、E190A、K120Q以及E146N;所述野生型转座酶的氨基酸序列如SEQ ID No.1所示;

优选地,所述转座酶突变体相较于所述野生型转座酶具有如下(1)~(3)所示突变组合中的任意一种;

(1) R27T和M56G;

(2) Y41R和M56G;

(3) K120Q和E146N。

2. 根据权利要求1所述的转座酶突变体,其特征在于,所述转座酶突变体相较于野生型转座酶还具有突变E54K和/或K120L;

优选地,所述转座酶突变体相较于野生型转座酶具有如下(4)和/或(5)所示突变组合:

(4) Y41R、E54K和M56G;

(5) K120L和E190A。

3. 一种融合蛋白,其特征在于,其包括权利要求1或2所述的转座酶突变体;

优选地,所述融合蛋白还包括用于结合免疫球蛋白Fc段的结合蛋白,所述结合蛋白与所述转座酶突变体相连;

优选地,所述结合蛋白包括金黄色葡萄球菌A蛋白或其变体、链球菌G蛋白或其变体和链球菌L蛋白或其变体中的至少一种;

优选地,所述结合蛋白包括金黄色葡萄球菌A蛋白或其变体;

优选地,所述结合蛋白的氨基酸序列如SEQ ID No.2或3所示;

优选地,所述结合蛋白位于所述融合蛋白的氨基酸序列的N端,所述转座酶突变体位于所述融合蛋白的氨基酸序列的C端;

优选地,所述结合蛋白与所述转座酶突变体通过连接肽连接;

优选地,所述连接肽为刚性连接肽或柔性连接肽;

优选地,所述连接肽为柔性连接肽;

优选地,所述连接肽的长度为6~59个氨基酸;

优选地,所述连接肽的长度为20~35个氨基酸;

优选地,所述连接肽的序列如SEQ ID No.4所示。

4. 一种分离的核酸,其特征在于,其编码如权利要求1或2所述的转座酶突变体或如权利要求3所述的融合蛋白。

5. 一种载体,其特征在于,其含有如权利要求4所述的分离的核酸。

6. 一种宿主细胞,其特征在于,其含有权利要求5所述的载体。

7. 制备如权利要求1所述的转座酶突变体或如权利要求3所述的融合蛋白的方法,其特征在于,其包括培养如权利要求6所述的宿主细胞。

8. 一种用于构建测序文库的试剂盒,其特征在于,其包括如权利要求1或2所述的转座酶突变体或如权利要求3所述的融合蛋白。

9. 如权利要求1或2所述的转座酶突变体或如权利要求3所述的融合蛋白在DNA建库中的应用。

10. 如权利要求1或2所述的转座酶突变体或如权利要求3所述的融合蛋白在分析蛋白

与染色质的相互作用中的应用；

优选地，所述应用包括孵育含有待分析的目标蛋白、目标染色质、以及所述转座酶突变体的混合物体系；

优选地，所述混合物体系还包括目标蛋白的抗体。

一种转座酶突变体、融合蛋白、其制备方法和应用

技术领域

[0001] 本发明涉及生物技术领域,具体而言,涉及一种转座酶突变体、融合蛋白、其制备方法和应用。

背景技术

[0002] 随着体外转座相关技术的大量应用,新型转座酶的开发和相关应用也越来越多,用转座酶介导的各种新型技术来替代旧技术的案例层出不穷。

[0003] 转座酶不仅可以应用于一般的二代测序建库过程,而且在多组学研究中的应用也越来越广泛。但现有的转座酶蛋白存在一些缺陷,其蛋白稳定性较差,于-20℃存放时,容易出现活性显著降低或丧失的情况,需要相对苛刻的保存条件,通常将其与经过专门设计序列的核酸共孵育形成复合体来提高其稳定性,这种使用条件上的限制通常提高了实验本身的成本,并且增加了实验整体流程的复杂程度。

[0004] 鉴于此,特提出本发明。

发明内容

[0005] 本发明的目的在于提供转座酶突变体、融合蛋白、分离的核酸、载体、宿主细胞、融合蛋白的制备方法、用于构建测序文库的试剂盒、转座酶突变体或融合蛋白在DNA建库中的应用以及融合蛋白在染色质共沉淀中的应用。

[0006] 本发明是这样实现的:

[0007] 第一方面,实施例提供一种转座酶突变体,所述转座酶突变体相较于野生型转座酶具有如下突变中的至少一种:R27T、M56G、Y41R、E190A、K120Q以及E146N;

[0008] 所述野生型转座酶的氨基酸序列如SEQ ID No.1所示。

[0009] 第二方面,实施例提供一种融合蛋白,其包括有上述转座酶突变体。

[0010] 第三方面,实施例提供一种分离的核酸,其编码如前述实施例所述的转座酶突变体或融合蛋白。

[0011] 第四方面,实施例提供一种载体,其含有如前述实施例所述的分离的核酸。

[0012] 第五方面,实施例提供一种宿主细胞,其含有如前述实施例所述的载体。

[0013] 第六方面,实施例提供制备转座酶突变体或融合蛋白的方法,其包括培养如前述实施例所述的宿主细胞。

[0014] 第七方面,实施例提供一种用于构建测序文库的试剂盒,其包括有如前述实施例所述的转座酶突变体或融合蛋白。

[0015] 第八方面,实施例提供一种前述实施例所述的转座酶突变体或融合蛋白在DNA建库中的应用。

[0016] 第九方面,实施例提供前述实施例所述的转座酶突变体或融合蛋白在分析蛋白与染色质的相互作用中的应用。

[0017] 本发明具有以下有益效果:

[0018] 本发明实施例提供一种转座酶突变体,该转座酶突变体的稳定性高,能够长期于-20℃的条件下储存,可以显著降低转座酶的储存成本,使转座酶或其构建的融合蛋白的使用更方便,简化实验流程。

附图说明

[0019] 为了更清楚地说明本发明实施例的技术方案,下面将对实施例中所需要使用的附图作简单地介绍,应当理解,以下附图仅示出了本发明的某些实施例,因此不应被看作是对范围的限定,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他相关的附图。

[0020] 图1为本发明验证例1中转座酶1与野生型转座酶的稳定性检测结果;

[0021] 图2为本发明验证例3中不同融合蛋白和多种酶标抗体结合的效果。

具体实施方式

[0022] 为使本发明实施例的目的、技术方案和优点更加清楚,下面将对本发明实施例中的技术方案进行清楚、完整地描述。实施例中未注明具体条件者,按照常规条件或制造商建议的条件进行。所用试剂或仪器未注明生产厂商者,均为可以通过市售购买获得的常规产品。

[0023] 本发明通过对野生型Tn5序列进行功能、蛋白结构和关键位点进行分析,并设计了大量突变进行稳定性改造,最终获得如下具有稳定性的转座酶突变体,并且该转座酶突变体与蛋白A构建成融合蛋白后也不影响其稳定性。

[0024] 本发明实施例提供了一种转座酶突变体,所述转座酶突变体相较于野生型转座酶具有如下突变中的至少一种或多种的组合:R27T、M56G、Y41R、E190A、K120Q以及E146N;所述野生型转座酶的氨基酸序列如SEQ ID No.1所示。该转座酶突变体的稳定性高,能显著降低储存成本,简化实验流程。

[0025] 需要说明的是,上述“如下突变中的至少一种或多种的组合”选自:上述6种突变中的任意1种突变、任意2种突变、任意3种突变、任意4种突变、任意5种突变以及以上6种突变的组合。

[0026] 在可选实施方式中,所述转座酶突变体相较于所述野生型转座酶具有如下(1)~(3)所示突变组合中的任意一种;

[0027] (1) R27T和M56G;

[0028] (2) Y41R和M56G;

[0029] (3) K120Q和E146N。

[0030] 在可选实施方式中,所述转座酶突变体相较于野生型转座酶还具有突变E54K和/或K120L。需要说明的是“还具有突变E54K和/或K120L”是指:在转座酶突变体相较于野生型转座酶具有上述6种突变中的任意1种突变、任意2种突变、任意3种突变、任意4种突变、任意5种突变以及以上6种突变的组合的基础上,还具有突变E54K和/或K120L。

[0031] 在可选实施方式中,所述转座酶突变体相较于野生型转座酶具有如下(4)~(5)所示突变组合中的任意一种:

[0032] (4) Y41R、E54K和M56G;

[0033] (5) K120L和E190A。

[0034] 相对于其他突变组合,具有上述(1)~(5)中任意突变的转座酶突变体的稳定性得到进一步提高。需要说明的是,“转座酶突变体相较于野生型转座酶具有(1)所述突变组合”是指,转座酶突变体的氨基酸序列为:野生型转座酶的氨基酸序列上第27位的精氨酸突变为苏氨酸的同时,野生型转座酶的氨基酸序列上第56位的蛋氨酸突变为甘氨酸,(2)~(5)以此类推。

[0035] 本发明实施例提供了一种融合蛋白,其包括前述任一实施方式所述的转座酶突变体。

[0036] 在可选实施方式中,所述融合蛋白还包括有用于结合免疫球蛋白Fc段的结合蛋白,所述结合蛋白与所述转座酶突变体相连。

[0037] 在可选实施方式中,所述结合蛋白包括金黄色葡萄球菌A蛋白(Protein A,PA)或其变体、链球菌G蛋白(Protein G)或其变体和链球菌L蛋白(Protein L)或其变体中的至少一种。

[0038] 在可选实施方式中,所述结合蛋白包括金黄色葡萄球菌A蛋白或其变体。

[0039] 在可选实施方式中,所述结合蛋白的氨基酸序列如SEQ ID No.2或3所示。

[0040] 在可选实施方式中,所述结合蛋白的氨基酸序列位于所述融合蛋白的氨基酸序列的N端,所述转座酶突变体的氨基酸序列位于所述融合蛋白的氨基酸序列的C端。

[0041] 在可选实施方式中,所述结合蛋白与所述转座酶突变体通过连接肽连接,融合蛋白的结构为:结合蛋白-连接肽-转座酶突变体。

[0042] 在可选实施方式中,所述连接肽为刚性连接肽或柔性连接肽。

[0043] 在可选实施方式中,所述连接肽为柔性连接肽。

[0044] 在可选实施方式中,所述连接肽的长度为6~59个氨基酸。

[0045] 在可选实施方式中,所述连接肽的长度为20~35个氨基酸。

[0046] 在可选实施方式中,所述连接肽的序列如SEQ ID No.4所示。

[0047] 本发明实施例提供了一种分离的核酸,其编码如前述任一实施方式所述的转座酶突变体或如前述任一实施方式所述的融合蛋白。

[0048] 本发明实施例提供了一种载体,其含有如前述实施例所述的分离的核酸。

[0049] 本发明实施例提供了一种宿主细胞,其含有前述实施例所述的载体。

[0050] 本发明实施例提供了一种制备前述任一实施例所述的转座酶突变体或融合蛋白的方法,其包括培养如前述任一实施例所述的宿主细胞。

[0051] 本发明实施例提供了一种用于构建测序文库的试剂盒,其包括如前述任一实施方式所述的转座酶突变体或如前述任一实施方式所述的融合蛋白。

[0052] 在可选实施方式中,所述试剂盒还包括Tn5转座酶缓冲液和/或PCR反应液。

[0053] 本发明实施例提供了如前述任一实施方式所述的转座酶突变体或如前述任一实施方式所述的融合蛋白在DNA建库中的应用。

[0054] 在可选实施方式中,所述应用包括将所述转座酶突变体或所述融合蛋白中的转座酶突变体进行接头(Adapter)包埋,形成Tn5转座体,将得到的转座体与目的基因混合共孵育,激活转座体以片段化目的基因,得到两端带有接头的DNA文库。

[0055] 在可选实施方式中,还包括采用PCR技术对片段后的目的基因进行PCR扩增。

[0056] 本发明实施例提供了如前述任一实施方式所述的转座酶突变体或如前述任一实施方式所述的融合蛋白在分析蛋白与染色质的相互作用中的应用。

[0057] 在可选实施方式中,所述应用包括孵育含有待分析的目标蛋白、目标染色质、以及所述转座酶突变体的混合物体系。

[0058] 在可选实施方式中,所述混合物体系还包括目标蛋白的抗体,所述转座酶为结合有用于结合免疫球蛋白Fc段的结合蛋白的所述融合蛋白,所述应用包括先将目标蛋白、目标染色质与目标蛋白的抗体结合,再加入融合蛋白共孵育。

[0059] 在混合体系中,目标蛋白与目标染色质结合,目标蛋白的抗体与目标蛋白结合,融合蛋白中的结合蛋白与目标蛋白的抗体的Fc段结合,融合蛋白中的转座酶突变体激活后,用于切割目标染色质,得到片段化的目标序列,用于测序。

[0060] 在可选实施方式中,采用金属离子激活转座酶突变体,金属离子优选为钙离子或镁离子。

[0061] 以下结合实施例对本发明的特征和性能作进一步的详细描述。

[0062] 实施例1

[0063] 本实施例提供了5种转座酶突变体,具体如下。

[0064] 转座酶1的氨基酸序列如SEQ ID No.5所示,核酸序列如SEQ ID No.6所示;

[0065] 转座酶2的氨基酸序列如SEQ ID No.7所示,核酸序列如SEQ ID No.8所示;

[0066] 转座酶3的氨基酸序列如SEQ ID No.9所示,核酸序列如SEQ ID No.10所示;

[0067] 转座酶4的氨基酸序列如SEQ ID No.11所示,核酸序列如SEQ ID No.12所示;

[0068] 转座酶5的氨基酸序列如SEQ ID No.13所示,核酸序列如SEQ ID No.14所示。

[0069] 实施例2

[0070] 本实施例提供一种融合蛋白,该融合蛋白包括有实施例1提供的转座酶突变体、用于结合免疫球蛋白Fc段的结合蛋白以及连接肽,结合蛋白与转座酶的氨基酸序列通过连接肽连接,该融合蛋白的结构为:结合蛋白-连接肽-转座酶突变体。

[0071] 其中,结合蛋白为蛋白A (PA),其包括有PA1和PA2。

[0072] 具体地,PA1的序列如SEQ ID No.2所示,PA2的序列如SEQ ID No.3所示,连接肽的序列如SEQ ID No.4所示。

[0073] 融合蛋白的制备方法如下。

[0074] 将上述转座酶突变体(实施例1中的转座酶1~5)和野生型转座酶Tn5,分别与PA1和PA2进行融合蛋白构建,所有序列通过人工基因合成,并将上述序列的片段分别连接至质粒载体上,用于构建表达载体,分别获得融合蛋白1~12。

[0075] 融合蛋白1:PA1-连接肽-转座酶1;

[0076] 融合蛋白2:PA1-连接肽-转座酶2;

[0077] 融合蛋白3:PA1-连接肽-转座酶3;

[0078] 融合蛋白4:PA1-连接肽-转座酶4;

[0079] 融合蛋白5:PA1-连接肽-转座酶5;

[0080] 融合蛋白6:PA1-连接肽-野生型转座酶Tn5;

[0081] 融合蛋白7:PA2-连接肽-转座酶1;

[0082] 融合蛋白8:PA2-连接肽-转座酶2;

[0083] 融合蛋白9:PA2-连接肽-转座酶3;

[0084] 融合蛋白10:PA2-连接肽-转座酶4;

[0085] 融合蛋白11:PA2-连接肽-转座酶5;

[0086] 融合蛋白12:PA2-连接肽-野生型转座酶Tn5。

[0087] 将所构建的表达载体转化到感受态大肠杆菌表达菌株内,将表达菌株添加至1mL LB培养基中,并在37℃培养30min,培养后取200μl菌液涂布于含卡那霉素的LB固体平板上,置于37℃恒温箱中过夜培养16h,并挑选白色菌落的单克隆菌株接种于3ml含有卡那霉素的液体LB培养基中,置于恒温摇床中以37℃过夜培养16h。

[0088] 次日取培养后的菌液,在500ml的三角培养瓶中按照1:100的比例接入到200ml含有卡那霉素的液体LB培养基中,并在37℃恒温摇床中培养,在OD600达到0.6左右时,加入终浓度为0.25mM的IPTG进行诱导,并在18℃过夜培养16h。对于发酵得到的表达目的蛋白的基因工程菌菌液,分别取500ml使用低温高速离心机进行离心(4℃,12000rpm,5min),弃去上清,对收集到的菌体分别加入50ml裂解缓冲液(50mM HEPES-KOH at PH 7.5,800mM NaCl,1mM EDTA,0.25%Tween-20)重悬菌体,对菌体进行超声破碎(30s超声破碎,10次),在相同条件下再次离心,收集离心得到的上清。对上清进行纯化以得到纯度高于90%的目的蛋白(融合蛋白)。

[0089] 验证例1

[0090] 验证实施例1提供的转座酶突变体的稳定性。

[0091] 分别取3×100μl实施例1提供的转座酶1~5和野生型转座酶至1.5mL的EP管中,对照设置8组试验例,试验例1~8的转座酶突变体相对于野生型转座酶而言依次具有以下突变:R27T、M56G、Y41R、E54K、K120L、E190A、K120Q以及E146N;

[0092] 然后将转座酶1~5、野生型转座酶以及试验例1~8提供的转座酶突变体分别置于25℃、4℃、-20℃环境中持续7天,以进行目的蛋白的稳定性检测。

[0093] 当目的蛋白在各考核温度条件下放置7天后,对其进行DNA片段化活性的检测,检测方法如下。

[0094] 制备Adapter Mix

[0095] 将100μM的Primer M分别与100μM Primer T1、100μM Primer T2等体积混匀后退火,然后将M-T1和M-T2的退火产物等体积混合均匀,所得混合物即为Adapter Mix。Primer序列如下:

[0096] Primer M:5'-pCTGTCTCTTATACACATCT-3' (p指磷酸化修饰);

[0097] Primer T1:5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3';

[0098] Primer T2:5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3';

[0099] 制备转座体

[0100] 将30μM样品(转座酶1~5以及野生型转座酶)分别与60μM退火的Adapter Mix在Coupling buffer(50mM Tris-Ac at PH8.2,400mM NaCl,0.1mM EDTA,0.25%Tween-20,25%甘油)中混匀置于室温孵育1h,得到转座体。

[0101] DNA片段化

[0102] 取1μl上述转座体,加入到含有2×Tnp buffer(50mM TAPS at PH7.9,5mM MgCl₂,2mM海藻糖、0.1%NP-40)以及200ng人基因组的体系中,并补加水至20μl,然后在热循环仪

中以55℃孵育10min(Lid 99℃),取反应产物10μl,通过琼脂糖凝胶电泳检测片段化产物的分布范围,结果请参照图1和表1。

[0103] 表1转座酶的稳定性结果

条件	25℃, 7 天	4℃, 7 天	-20℃, 7 天
转座酶 1	++	+++	+++
转座酶 2	++	++	+++
转座酶 3	++	+++	+++
转座酶 4	++	+++	+++
转座酶 5	++	++	+++
[0104] 野生型转座酶	-	+	+++
R27T	+	++	+++
M56G	-	++	+++
Y41R	+	++	+++
E190A	-	+++	+++
E146N	-	++	+++
K120Q	-	+++	+++
[0105] K120L	-	+	+++
E54K	-	+	+++

[0106] 备注:1、+++表示片段化程度高,条带弥散主要在100-500kb,无2000kb以上的片段;

[0107] 2、++表示片段化程度高,条带弥散主要在500-1000kb,少量2000kb以上片段;

[0108] 3、+表示片段化程度低,条带主要弥散在2000kb以上;

[0109] 4、-表示基本无法片段化,条带未呈现弥散状。

[0110] 图1为转座酶1与野生型转座酶的稳定性检测结果,其中,1~3为野生型转座酶分别在-20℃、-4℃、25℃的条件下储存7天的检测结果,4~6为转座酶1分别在-20℃、-4℃、25℃的条件下储存7天的结果。

[0111] 由表1可知,试验例中的单突变K120Q、E146N、E190A、Y41R、M56G以及R27T相对于野生型转座酶而言,稳定性得到了提高;

[0112] 结合表1和图1,本发明实施例1提供的转座酶突变体的稳定性显著优于野生型和试验例1~8中的转座酶突变体(单突变),其可以于-20℃的条件下进行长期储存,甚至可以在4℃条件下进行短期储存,避免了存放于-80℃条件下导致的取用不便的情况,更无需将其构建为转座体进行储存,极大地提高了使用和体系调整的便利性。

[0113] 验证例2

[0114] 对实施例2制备得到的融合蛋白进行稳定性检测。

[0115] 检测方法同验证例1。

[0116] 检测结果请参照表2。

[0117] 表2稳定性评价结果

考核条件	25℃ 7天	4℃ 7天	-20℃ 7天
融合蛋白 1	++	+++	+++
融合蛋白 2	++	++	+++
融合蛋白 3	++	+++	+++
融合蛋白 4	++	+++	+++
融合蛋白 5	++	++	+++
融合蛋白 6	-	+	+++
融合蛋白 7	++	+++	+++
融合蛋白 8	++	+++	+++

[0118]

融合蛋白 9	++	+++	+++
融合蛋白 10	++	+++	+++
融合蛋白 11	++	+++	+++
融合蛋白 12	-	++	+++

[0119]

[0120] 由表2可知,由转座酶突变体构建的融合蛋白的稳定性显著高于由野生型转座酶构建的融合蛋白的稳定性。

[0121] 验证例3

[0122] 验证实施例2制备得到的融合蛋白结合抗体的能力。

[0123] 使用酶联免疫反应(ELISA)原理检测不同融合蛋白与多种酶标抗体(IgG)结合的能力,具体检测流程如下。

[0124] 用稀释液将实施例2提供的融合蛋白1~12均稀释至统一浓度(0.8nmol/ml即0.6mg/ml)。将稀释好的融合蛋白各100μl加入酶标板的对应孔中,每个样品均设置24个平行组,同时有3个反应孔对应一种酶标抗体以保证数据准确性,同时设置阴性对照并以ddH₂O代替,封板后放置37℃孵育30min;移去孔中液体并用洗液洗涤4次并彻底去除孔中残留的洗液,然后向每个孔中加入250μl封闭液,封板后置于37℃孵育1h,用稀释液将多种酶标抗体(鼠、兔、犬、羊、牛、马、人、猪)稀释至(0.4nmol/ml),分别向对应的反应孔中加入250μl封板后置于37℃孵育30min;移去孔中液体并用洗液洗涤4次并彻底去除孔中残留的洗液,然后向每个孔中加入100μl显色液,封板后置于室温(20-25℃)避光孵育15min,然后向每孔中加入50μl终止液,终止后立即用酶标仪测定OD₄₅₀数值。

[0125] 以OD450数值为纵坐标,以多种酶标抗体类型为横坐标,可以看出不同融合蛋白和多种酶标抗体结合的效果,结果请参见图2。

[0126] 由图2可知,融合蛋白7~12对于不同种类的抗体均能有效结合,可见PA2构建的融合蛋白对于不同物种来源的抗体结合能力没有明显的偏倚,其应用场景相对而言更加广泛。

[0127] 验证例4

[0128] 采用实施例2提供的融合蛋白进行蛋白-染色质互作的分析(ChIP-Seq技术),其包括以下步骤。

[0129] 制备Adapter Mix

[0130] 将100 μ M的Primer M分别与100 μ M Primer T1、100 μ M Primer T2等体积混匀后退火,然后将M-T1和M-T2的退火产物等体积混合均匀,所得混合物即为Adapter Mix,Primer序列如下:

[0131] Primer M:5'-pCTGTCTCTTATACACATCT-3' (p指磷酸化修饰);

[0132] Primer T1:5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3';

[0133] Primer T2:5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3'。

[0134] 制备转座体

[0135] 将30 μ M目的蛋白(融合蛋白1~5)分别与60 μ M混合均匀的Adapter Mix在Coupling buffer(50mM Tris-Ac at PH8.2,400mM NaCl,0.1mM EDTA,0.25%Tween-20,25%甘油)中混匀置于室温孵育1h,得到转座体。

[0136] DNA片段化

[0137] 取1 μ l上述转座体,加入到含有2 $\times$ Tnp buffer(50mM TAPS at PH7.9,5mM MgCl₂,2mM海藻糖、0.1%NP-40)、50ng人基因组的体系中,并补加水至20 μ l,然后在热循环仪中孵育10min(Lid 99 $^{\circ}$ C),向反应后的体系中加入1 μ l消化缓冲液,然后在热循环仪中以55 $^{\circ}$ C孵育5min(Lid 99 $^{\circ}$ C)进行终止。

[0138] 随后将反应体系补水至50 μ l并用1 $\times$ FP Beads进行纯化,使用FP NM008试剂盒对纯化产物进行扩增,并对扩增产物以0.6 $\times$ /0.15 $\times$ 的FP Beads进行双轮片段分选,得到最终的文库,对文库用illumina Hiseq测序,测序策略为PE150。测序数据对照见表3。

[0139] 表3测序数据对照结果

[0140]

文库编号	测序深度	覆盖度(%)	有效数据率(%)	碱基错误(%)	Q30(%)	Dup(%)	GC含量
融合蛋白1	9.27×	98.95	97.53	0.015	93.42	17	41.75
融合蛋白2	9.56×	89.63	90.16	0.023	89.37	30	42.23
融合蛋白3	10.35×	98.26	95.42	0.017	94.06	19	41.91
融合蛋白4	12.27×	94.64	89.53	0.021	90.71	27	40.84
融合蛋白5	12.83×	92.02	91.27	0.02	87.13	24	41.21
融合蛋白6	11.64×	98.21	91.42	0.02	91.59	22	43.18

[0141] 由表3可知,本发明实施例提供的融合蛋白均能够用于建库测序,其中,从覆盖度和有效数据率分析可以发现,融合蛋白1和3的效果较优。

[0142] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

SEQUENCE LISTING

<110> 广东菲鹏生物有限公司

<120> 一种转座酶突变体、融合蛋白、其制备方法和应用

<160> 14

<170> PatentIn version 3.5

<210> 1

<211> 476

<212> PRT

<213> 人工序列

<400> 1

```

Met Ile Thr Ser Ala Leu His Arg Ala Ala Asp Trp Ala Lys Ser Val
1           5           10           15
Phe Ser Ser Ala Ala Leu Gly Asp Pro Arg Arg Thr Ala Arg Leu Val
           20           25           30
Asn Val Ala Ala Gln Leu Ala Lys Tyr Ser Gly Lys Ser Ile Thr Ile
           35           40           45
Ser Ser Glu Gly Ser Glu Ala Met Gln Glu Gly Ala Tyr Arg Phe Ile
           50           55           60
Arg Asn Pro Asn Val Ser Ala Glu Ala Ile Arg Lys Ala Gly Ala Met
65           70           75           80
Gln Thr Val Lys Leu Ala Gln Glu Phe Pro Glu Leu Leu Ala Ile Glu
           85           90           95
Asp Thr Thr Ser Leu Ser Tyr Arg His Gln Val Ala Glu Glu Leu Gly
           100          105          110
Lys Leu Gly Ser Ile Gln Asp Lys Ser Arg Gly Trp Trp Val His Ser
           115          120          125
Val Leu Leu Leu Glu Ala Thr Thr Phe Arg Thr Val Gly Leu Leu His
           130          135          140
Gln Glu Trp Trp Met Arg Pro Asp Asp Pro Ala Asp Ala Asp Glu Lys
145          150          155          160
Glu Ser Gly Lys Trp Leu Ala Ala Ala Ala Thr Ser Arg Leu Arg Met
           165          170          175
Gly Ser Met Met Ser Asn Val Ile Ala Val Cys Asp Arg Glu Ala Asp
           180          185          190
Ile His Ala Tyr Leu Gln Asp Lys Leu Ala His Asn Glu Arg Phe Val
           195          200          205
Val Arg Ser Lys His Pro Arg Lys Asp Val Glu Ser Gly Leu Tyr Leu
210          215          220

```

Tyr Asp His Leu Lys Asn Gln Pro Glu Leu Gly Gly Tyr Gln Ile Ser			
225	230	235	240
Ile Pro Gln Lys Gly Val Val Asp Lys Arg Gly Lys Arg Lys Asn Arg			
	245	250	255
Pro Ala Arg Lys Ala Ser Leu Ser Leu Arg Ser Gly Arg Ile Thr Leu			
	260	265	270
Lys Gln Gly Asn Ile Thr Leu Asn Ala Val Leu Ala Glu Glu Ile Asn			
	275	280	285
Pro Pro Lys Gly Glu Thr Pro Leu Lys Trp Leu Leu Leu Thr Ser Glu			
	290	295	300
Pro Val Glu Ser Leu Ala Gln Ala Leu Arg Val Ile Asp Ile Tyr Thr			
305	310	315	320
His Arg Trp Arg Ile Glu Glu Phe His Lys Ala Trp Lys Thr Gly Ala			
	325	330	335
Gly Ala Glu Arg Gln Arg Met Glu Glu Pro Asp Asn Leu Glu Arg Met			
	340	345	350
Val Ser Ile Leu Ser Phe Val Ala Val Arg Leu Leu Gln Leu Arg Glu			
	355	360	365
Ser Phe Thr Leu Pro Gln Ala Leu Arg Ala Gln Gly Leu Leu Lys Glu			
	370	375	380
Ala Glu His Val Glu Ser Gln Ser Ala Glu Thr Val Leu Thr Pro Asp			
385	390	395	400
Glu Cys Gln Leu Leu Gly Tyr Leu Asp Lys Gly Lys Arg Lys Arg Lys			
	405	410	415
Glu Lys Ala Gly Ser Leu Gln Trp Ala Tyr Met Ala Ile Ala Arg Leu			
	420	425	430
Gly Gly Phe Met Asp Ser Lys Arg Thr Gly Ile Ala Ser Trp Gly Ala			
	435	440	445
Leu Trp Glu Gly Trp Glu Ala Leu Gln Ser Lys Leu Asp Gly Phe Leu			
	450	455	460
Ala Ala Lys Asp Leu Met Ala Gln Gly Ile Lys Ile			
465	470	475	
<210> 2			
<211> 58			
<212> PRT			
<213> 人工序列			
<400> 2			
Val Asp Asn Lys Phe Asn Lys Glu Gln Gln Asn Ala Phe Tyr Glu Ile			
1	5	10	15

Leu His Leu Pro Asn Leu Asn Glu Glu Gln Arg Asn Ala Phe Ile Gln
 20 25 30
 Ser Leu Lys Asp Asp Pro Ser Gln Ser Ala Asn Leu Leu Ala Glu Ala
 35 40 45
 Lys Lys Leu Asn Asp Ala Gln Ala Pro Lys
 50 55

<210> 3

<211> 57

<212> PRT

<213> 人工序列

<400> 3

Ala Ser Gln Arg Phe Asn Lys Asp Gln Gln Leu Ala Phe Tyr Glu Ile
 1 5 10 15
 Ile Gln Pro Pro Asn Leu Glu Cys Thr Gln Arg Asn Gly Phe Ile Gln
 20 25 30
 Ser Leu Lys Asp Asp Phe Ser Ser Ser Gly Asn Ile Ile Val Glu Ala
 35 40 45
 Lys Lys Ile Asn Ser Thr Gln Ala Pro
 50 55

<210> 4

<211> 27

<212> PRT

<213> 人工序列

<400> 4

Asp Asp Asp Lys Glu Phe Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
 1 5 10 15
 Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser His
 20 25

<210> 5

<211> 476

<212> PRT

<213> 人工序列

<400> 5

Met Ile Thr Ser Ala Leu His Arg Ala Ala Asp Trp Ala Lys Ser Val
 1 5 10 15
 Phe Ser Ser Ala Ala Leu Gly Asp Pro Arg Thr Thr Ala Arg Leu Val
 20 25 30
 Asn Val Ala Ala Gln Leu Ala Lys Tyr Ser Gly Lys Ser Ile Thr Ile
 35 40 45

Ser Ser Glu Gly Ser Glu Ala Gly Gln Glu Gly Ala Tyr Arg Phe Ile		
50	55	60
Arg Asn Pro Asn Val Ser Ala Glu Ala Ile Arg Lys Ala Gly Ala Met		
65	70	75 80
Gln Thr Val Lys Leu Ala Gln Glu Phe Pro Glu Leu Leu Ala Ile Glu		
85	90	95
Asp Thr Thr Ser Leu Ser Tyr Arg His Gln Val Ala Glu Glu Leu Gly		
100	105	110
Lys Leu Gly Ser Ile Gln Asp Lys Ser Arg Gly Trp Trp Val His Ser		
115	120	125
Val Leu Leu Leu Glu Ala Thr Thr Phe Arg Thr Val Gly Leu Leu His		
130	135	140
Gln Glu Trp Trp Met Arg Pro Asp Asp Pro Ala Asp Ala Asp Glu Lys		
145	150	155 160
Glu Ser Gly Lys Trp Leu Ala Ala Ala Ala Thr Ser Arg Leu Arg Met		
165	170	175
Gly Ser Met Met Ser Asn Val Ile Ala Val Cys Asp Arg Glu Ala Asp		
180	185	190
Ile His Ala Tyr Leu Gln Asp Lys Leu Ala His Asn Glu Arg Phe Val		
195	200	205
Val Arg Ser Lys His Pro Arg Lys Asp Val Glu Ser Gly Leu Tyr Leu		
210	215	220
Tyr Asp His Leu Lys Asn Gln Pro Glu Leu Gly Gly Tyr Gln Ile Ser		
225	230	235 240
Ile Pro Gln Lys Gly Val Val Asp Lys Arg Gly Lys Arg Lys Asn Arg		
245	250	255
Pro Ala Arg Lys Ala Ser Leu Ser Leu Arg Ser Gly Arg Ile Thr Leu		
260	265	270
Lys Gln Gly Asn Ile Thr Leu Asn Ala Val Leu Ala Glu Glu Ile Asn		
275	280	285
Pro Pro Lys Gly Glu Thr Pro Leu Lys Trp Leu Leu Leu Thr Ser Glu		
290	295	300
Pro Val Glu Ser Leu Ala Gln Ala Leu Arg Val Ile Asp Ile Tyr Thr		
305	310	315 320
His Arg Trp Arg Ile Glu Glu Phe His Lys Ala Trp Lys Thr Gly Ala		
325	330	335
Gly Ala Glu Arg Gln Arg Met Glu Glu Pro Asp Asn Leu Glu Arg Met		
340	345	350
Val Ser Ile Leu Ser Phe Val Ala Val Arg Leu Leu Gln Leu Arg Glu		

355	360	365
Ser Phe Thr Leu Pro Gln Ala Leu Arg Ala Gln Gly Leu Leu Lys Glu		
370	375	380
Ala Glu His Val Glu Ser Gln Ser Ala Glu Thr Val Leu Thr Pro Asp		
385	390	395
Glu Cys Gln Leu Leu Gly Tyr Leu Asp Lys Gly Lys Arg Lys Arg Lys		
405	410	415
Glu Lys Ala Gly Ser Leu Gln Trp Ala Tyr Met Ala Ile Ala Arg Leu		
420	425	430
Gly Gly Phe Met Asp Ser Lys Arg Thr Gly Ile Ala Ser Trp Gly Ala		
435	440	445
Leu Trp Glu Gly Trp Glu Ala Leu Gln Ser Lys Leu Asp Gly Phe Leu		
450	455	460
Ala Ala Lys Asp Leu Met Ala Gln Gly Ile Lys Ile		
465	470	475

<210> 6

<211> 1428

<212> DNA

<213> 人工序列

<400> 6

atgatcacct ctgctctgca ccgtgctgct gactgggcta aatctgtttt ctcttctgct 60
gctctgggtg acccgctac caccgctcgt ctggtaacg ttgctgctca gctggctaaa 120
tactctggta aatctatcac catctcttct gaaggttctg aagctggta ggaaggtgct 180
taccgtttca tccgtaaccc gaacgtttct gctgaagcta tccgtaaagc tgggtgctatg 240
cagaccgtta aactggctca ggaatttccg gaactgctgg ctatcgaaga caccacctct 300
ctgtcttacc gtcaccaggt tgctgaagaa ctgggtaaac tgggttctat ccaggacaaa 360
tctcgtaggtt ggtgggttca ctctgttctg ctgctggaag ctaccacctt ccgtaccgtt 420
ggtctgctgc accaggaatg gtggatgcgt ccggacgacc cggctgacgc tgacgaaaaa 480
gaatctggta aatggctggc tgctgctgct acctctcgtc tgcgtatggg ttctatgatg 540
tctaacgtta tcgctgtttg cgaccgtgaa gctgacatcc acgcttacct gcaggacaaa 600
ctggctcaca acgaacgttt cgttgttcgt tctaaacacc cgcgtaaaga cgttgaatct 660
ggtctgtacc tgtacgacca cctgaaaaac cagccggaac tgggtggtta ccagatctct 720
atcccgcaaga aaggtgttgt tgacaaacgt ggtaaacgta aaaaccgtcc ggctcgtaaa 780
gcttctctgt ctctgcgttc tggctgtatc acctgaaac agggtaacat caccctgaac 840
gctgttctgg ctgaagaaat caaccgccg aaaggtgaaa ccccgctgaa atggctgctg 900
ctgacctctg aaccggttga atctctggct caggctctgc gtgttatcga catctacacc 960
caccgttggc gtatcgaaga atttcacaaa gcttgaaaaa ccggtgctgg tgctgaacgt 1020
cagcgatagg aagaaccgga caacctggaa cgtatggttt ctatcctgtc tttcgttgct 1080
gttcgtctgc tgcagctgcg tgaatctttc acctgccgc aggctctgcg tgctcagggt 1140

ctgctgaaag aagctgaaca cggtgaatct cagtctgctg aaaccgttct gaccccgac 1200
 gaatgccagc tgctgggtta cctggacaaa ggtaaactga aacgtaaaga aaaagctggt 1260
 tctctgcagt gggcttacat ggctatcgct cgtctgggtg gtttcatgga ctctaaacgt 1320
 accggtatcg cttcttgggg tgctctgtgg gaaggttggg aagctctgca gtcgaagctg 1380
 gacggcttcc tggttgctaa ggacctgatg gctcagggtg tcaaaatc 1428

<210> 7

<211> 476

<212> PRT

<213> 人工序列

<400> 7

Met	Ile	Thr	Ser	Ala	Leu	His	Arg	Ala	Ala	Asp	Trp	Ala	Lys	Ser	Val
1				5					10					15	
Phe	Ser	Ser	Ala	Ala	Leu	Gly	Asp	Pro	Arg	Arg	Thr	Ala	Arg	Leu	Val
			20					25					30		
Asn	Val	Ala	Ala	Gln	Leu	Ala	Lys	Arg	Ser	Gly	Lys	Ser	Ile	Thr	Ile
		35					40					45			
Ser	Ser	Glu	Gly	Ser	Glu	Ala	Gly	Gln	Glu	Gly	Ala	Tyr	Arg	Phe	Ile
	50					55					60				
Arg	Asn	Pro	Asn	Val	Ser	Ala	Glu	Ala	Ile	Arg	Lys	Ala	Gly	Ala	Met
65				70					75					80	
Gln	Thr	Val	Lys	Leu	Ala	Gln	Glu	Phe	Pro	Glu	Leu	Leu	Ala	Ile	Glu
			85					90					95		
Asp	Thr	Thr	Ser	Leu	Ser	Tyr	Arg	His	Gln	Val	Ala	Glu	Glu	Leu	Gly
			100				105						110		
Lys	Leu	Gly	Ser	Ile	Gln	Asp	Lys	Ser	Arg	Gly	Trp	Trp	Val	His	Ser
		115				120						125			
Val	Leu	Leu	Leu	Glu	Ala	Thr	Thr	Phe	Arg	Thr	Val	Gly	Leu	Leu	His
		130				135					140				
Gln	Glu	Trp	Trp	Met	Arg	Pro	Asp	Asp	Pro	Ala	Asp	Ala	Asp	Glu	Lys
145				150					155					160	
Glu	Ser	Gly	Lys	Trp	Leu	Ala	Ala	Ala	Ala	Thr	Ser	Arg	Leu	Arg	Met
			165						170				175		
Gly	Ser	Met	Met	Ser	Asn	Val	Ile	Ala	Val	Cys	Asp	Arg	Glu	Ala	Asp
		180				185						190			
Ile	His	Ala	Tyr	Leu	Gln	Asp	Lys	Leu	Ala	His	Asn	Glu	Arg	Phe	Val
		195				200						205			
Val	Arg	Ser	Lys	His	Pro	Arg	Lys	Asp	Val	Glu	Ser	Gly	Leu	Tyr	Leu
	210					215						220			
Tyr	Asp	His	Leu	Lys	Asn	Gln	Pro	Glu	Leu	Gly	Gly	Tyr	Gln	Ile	Ser

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

<210> 8

<211> 1428

<212> DNA

<213> 人工序列

<400> 8

atgatcacct ctgctctgca ccgtgctgct gactgggcta aatctgtttt ctcttctgct 60
gctctgggtg acccgctcg taccgctcgt ctggttaacg ttgctgctca gctggctaaa 120
cgttctggta aatctatcac catctcttct gaaggttctg aagctggtca ggaaggtgct 180

taccgtttca tccgtaaccc gaacgtttct gctgaagcta tccgtaaagc tggtgctatg 240
 cagaccgtta aactggctca ggaatttccg gaactgctgg ctatcgaaga caccacctct 300
 ctgtcttacc gtcaccaggt tgctgaagaa ctgggtaaac tgggttctat ccaggacaaa 360
 tctcgtgggtt ggtgggttca ctctgttctg ctgctggaag ctaccacctt ccgtaccgtt 420
 ggtctgctgc accaggaatg gtggatgcgt ccggacgacc cggctgacgc tgacgaaaaa 480
 gaatctggta aatggctggc tgctgctgct acctctcgtc tgcgtatggg ttctatgatg 540
 tctaacgtta tcgctgtttg cgaccgtgaa gctgacatcc acgcttacct gcaggacaaa 600
 ctggctcaca acgaacgttt cgttgttcgt tctaaacacc cgcgtaaaga cgttgaatct 660
 ggtctgtacc tgtacgacca cctgaaaaac cagccggaac tgggtgggta ccagatctct 720
 atccccgaga aaggtgttgt tgacaaacgt ggtaaagcta aaaaccgtcc ggctcgtaaa 780
 gcttctctgt ctctgcgttc tggctgtatc acctgaaac agggtaacat caccctgaac 840
 gctgttctgg ctgaagaaat caaccgccg aaaggtgaaa ccccgctgaa atggctgctg 900
 ctgacctctg aaccggttga atctctggct caggctctgc gtgttatcga catctacacc 960
 caccgttggc gtatcgaaga atttcacaaa gcttgaaaaa ccggtgctgg tgctgaacgt 1020
 cagcgatagg aagaaccgga caacctggaa cgtatggttt ctatcctgtc tttcgttgc 1080
 gttcgtctgc tgcagctgcg tgaatctttc acctgccgc aggtctctgc tgctcagggt 1140
 ctgctgaaag aagctgaaca cgttgaatct cagtctgctg aaaccgttct gaccccgac 1200
 gaatgccagc tgctgggta cctggacaaa ggtaaagcta aacgtaaaga aaaagctggt 1260
 tctctgcagt gggcttacat ggctatcgct cgtctgggtg gtttcatgga ctctaaacgt 1320
 accggtatcg cttcttgggg tgctctgtgg gaaggttggg aagctctgca gtcgaagctg 1380
 gacggcttcc tggctgctaa ggacctgatg gctcagggtg tcaaaatc 1428

<210> 9

<211> 476

<212> PRT

<213> 人工序列

<400> 9

Met	Ile	Thr	Ser	Ala	Leu	His	Arg	Ala	Ala	Asp	Trp	Ala	Lys	Ser	Val
1				5					10					15	
Phe	Ser	Ser	Ala	Ala	Leu	Gly	Asp	Pro	Arg	Arg	Thr	Ala	Arg	Leu	Val
			20					25					30		
Asn	Val	Ala	Ala	Gln	Leu	Ala	Lys	Arg	Ser	Gly	Lys	Ser	Ile	Thr	Ile
			35				40						45		
Ser	Ser	Glu	Gly	Ser	Lys	Ala	Gly	Gln	Glu	Gly	Ala	Tyr	Arg	Phe	Ile
		50				55					60				
Arg	Asn	Pro	Asn	Val	Ser	Ala	Glu	Ala	Ile	Arg	Lys	Ala	Gly	Ala	Met
65				70					75					80	
Gln	Thr	Val	Lys	Leu	Ala	Gln	Glu	Phe	Pro	Glu	Leu	Leu	Ala	Ile	Glu
			85						90					95	
Asp	Thr	Thr	Ser	Leu	Ser	Tyr	Arg	His	Gln	Val	Ala	Glu	Glu	Leu	Gly

100	105	110
Lys Leu Gly Ser Ile Gln Asp	Lys Ser Arg Gly Trp Trp	Val His Ser
115	120	125
Val Leu Leu Leu Glu Ala Thr	Thr Phe Arg Thr Val Gly	Leu Leu His
130	135	140
Gln Glu Trp Trp Met Arg Pro	Asp Asp Pro Ala Asp Ala	Asp Glu Lys
145	150	155
Glu Ser Gly Lys Trp Leu Ala	Ala Ala Ala Thr Ser Arg	Leu Arg Met
165	170	175
Gly Ser Met Met Ser Asn Val	Ile Ala Val Cys Asp Arg	Glu Ala Asp
180	185	190
Ile His Ala Tyr Leu Gln Asp	Lys Leu Ala His Asn Glu	Arg Phe Val
195	200	205
Val Arg Ser Lys His Pro Arg	Lys Asp Val Glu Ser Gly	Leu Tyr Leu
210	215	220
Tyr Asp His Leu Lys Asn Gln	Pro Glu Leu Gly Gly Tyr	Gln Ile Ser
225	230	235
Ile Pro Gln Lys Gly Val Val	Asp Lys Arg Gly Lys Arg	Lys Asn Arg
245	250	255
Pro Ala Arg Lys Ala Ser Leu	Ser Leu Arg Ser Gly Arg	Ile Thr Leu
260	265	270
Lys Gln Gly Asn Ile Thr Leu	Asn Ala Val Leu Ala Glu	Glu Ile Asn
275	280	285
Pro Pro Lys Gly Glu Thr Pro	Leu Lys Trp Leu Leu Leu	Thr Ser Glu
290	295	300
Pro Val Glu Ser Leu Ala Gln	Ala Leu Arg Val Ile Asp	Ile Tyr Thr
305	310	315
His Arg Trp Arg Ile Glu Glu	Phe His Lys Ala Trp Lys	Thr Gly Ala
325	330	335
Gly Ala Glu Arg Gln Arg Met	Glu Glu Pro Asp Asn Leu	Glu Arg Met
340	345	350
Val Ser Ile Leu Ser Phe Val	Ala Val Arg Leu Leu Gln	Leu Arg Glu
355	360	365
Ser Phe Thr Leu Pro Gln Ala	Leu Arg Ala Gln Gly Leu	Leu Lys Glu
370	375	380
Ala Glu His Val Glu Ser Gln	Ser Ala Glu Thr Val Leu	Thr Pro Asp
385	390	395
Glu Cys Gln Leu Leu Gly Tyr	Leu Asp Lys Gly Lys Arg	Lys Arg Lys
405	410	415

Glu Lys Ala Gly Ser Leu Gln Trp Ala Tyr Met Ala Ile Ala Arg Leu
 420 425 430
 Gly Gly Phe Met Asp Ser Lys Arg Thr Gly Ile Ala Ser Trp Gly Ala
 435 440 445
 Leu Trp Glu Gly Trp Glu Ala Leu Gln Ser Lys Leu Asp Gly Phe Leu
 450 455 460
 Ala Ala Lys Asp Leu Met Ala Gln Gly Ile Lys Ile
 465 470 475

<210> 10

<211> 1428

<212> DNA

<213> 人工序列

<400> 10

atgatcacct ctgctctgca ccgtgctgct gactgggcta aatctgtttt ctcttctgct 60
 gctctgggtg acccgctcg taccgctcgt ctggttaacg ttgctgctca gctggctaaa 120
 cgttctggta aatctatcac catctcttct gaaggttcta aagctggctca ggaaggtgct 180
 taccgtttca tccgtaaccc gaacgtttct gctgaagcta tccgtaaagc tggtgctatg 240
 cagaccgtta aactggctca ggaatttccg gaactgctgg ctatcgaaga caccacctct 300
 ctgtcttacc gtcaccaggt tgctgaagaa ctgggttaac tgggttctat ccaggacaaa 360
 tctcgtgggtt ggtgggttca ctctgttctg ctgctggaag ctaccacctt ccgtaccgtt 420
 ggtctgctgc accaggaatg gtggatgctg ccggacgacc cggctgacgc tgacgaaaaa 480
 gaatctggta aatggctggc tgctgctgct acctctcgtc tgcgtatggg ttctatgatg 540
 tctaacgtta tcgctgtttg cgaccgtgaa gctgacatcc acgcttacct gcaggacaaa 600
 ctggctcaca acgaacgttt cgttgttcgt tctaaacacc cgcgtaaaga cgttgaatct 660
 ggtctgtacc tgtacgacca cctgaaaaac cagccggaac tgggtgggta ccagatctct 720
 atccccgaga aaggtgttgt tgacaaacgt ggtaaacgta aaaaccgtcc ggctcgtaaa 780
 gcttctctgt ctctgcgttc tggctgtatc acctgaaac agggtaacat caccctgaac 840
 gctgttctgg ctgaagaaat caaccgccc aaaggtgaaa ccccgctgaa atggctgctg 900
 ctgacctctg aaccggttga atctctggct caggtctctg gtgttatcga catctacacc 960
 caccgttggc gtatcgaaga atttcacaaa gcttgaaaaa ccggtgctgg tgctgaacgt 1020
 cagcgatatg aagaaccgga caacctggaa cgtatggttt ctatcctgtc tttcgttgct 1080
 gtctgtctgc tgcagctcgc tgaatctttc acctgccgc aggctctgcg tgctcagggt 1140
 ctgctgaaaag aagctgaaca cgttgaatct cagtctgctg aaaccgttct gaccccgac 1200
 gaatgccagc tgctgggtta cctggacaaa ggtaaacgta aacgtaaaga aaaagctggt 1260
 tctctgcagt gggcttacat ggctatcgtc cgtctgggtg gtttcatgga ctctaaacgt 1320
 accggtatcg cttcttgggg tgctctgtgg gaaggttggg aagctctgca gtcgaagctg 1380
 gacggcttcc tggctgctaa ggacctgatg gctcagggtg tcaaaatc 1428

<210> 11

<211> 476

<212> PRT

<213> 人工序列

<400> 11

```

Met Ile Thr Ser Ala Leu His Arg Ala Ala Asp Trp Ala Lys Ser Val
1           5           10           15
Phe Ser Ser Ala Ala Leu Gly Asp Pro Arg Arg Thr Ala Arg Leu Val
          20           25           30
Asn Val Ala Ala Gln Leu Ala Lys Tyr Ser Gly Lys Ser Ile Thr Ile
          35           40           45
Ser Ser Glu Gly Ser Glu Ala Met Gln Glu Gly Ala Tyr Arg Phe Ile
          50           55           60
Arg Asn Pro Asn Val Ser Ala Glu Ala Ile Arg Lys Ala Gly Ala Met
65           70           75           80
Gln Thr Val Lys Leu Ala Gln Glu Phe Pro Glu Leu Leu Ala Ile Glu
          85           90           95
Asp Thr Thr Ser Leu Ser Tyr Arg His Gln Val Ala Glu Glu Leu Gly
          100          105          110
Lys Leu Gly Ser Ile Gln Asp Leu Ser Arg Gly Trp Trp Val His Ser
          115          120          125
Val Leu Leu Leu Glu Ala Thr Thr Phe Arg Thr Val Gly Leu Leu His
          130          135          140
Gln Glu Trp Trp Met Arg Pro Asp Asp Pro Ala Asp Ala Asp Glu Lys
145          150          155          160
Glu Ser Gly Lys Trp Leu Ala Ala Ala Ala Thr Ser Arg Leu Arg Met
          165          170          175
Gly Ser Met Met Ser Asn Val Ile Ala Val Cys Asp Arg Ala Ala Asp
          180          185          190
Ile His Ala Tyr Leu Gln Asp Lys Leu Ala His Asn Glu Arg Phe Val
          195          200          205
Val Arg Ser Lys His Pro Arg Lys Asp Val Glu Ser Gly Leu Tyr Leu
          210          215          220
Tyr Asp His Leu Lys Asn Gln Pro Glu Leu Gly Gly Tyr Gln Ile Ser
225          230          235          240
Ile Pro Gln Lys Gly Val Val Asp Lys Arg Gly Lys Arg Lys Asn Arg
          245          250          255
Pro Ala Arg Lys Ala Ser Leu Ser Leu Arg Ser Gly Arg Ile Thr Leu
          260          265          270
Lys Gln Gly Asn Ile Thr Leu Asn Ala Val Leu Ala Glu Glu Ile Asn
          275          280          285

```

Pro	Pro	Lys	Gly	Glu	Thr	Pro	Leu	Lys	Trp	Leu	Leu	Leu	Thr	Ser	Glu
290						295				300					
Pro	Val	Glu	Ser	Leu	Ala	Gln	Ala	Leu	Arg	Val	Ile	Asp	Ile	Tyr	Thr
305					310					315					320
His	Arg	Trp	Arg	Ile	Glu	Glu	Phe	His	Lys	Ala	Trp	Lys	Thr	Gly	Ala
				325						330					335
Gly	Ala	Glu	Arg	Gln	Arg	Met	Glu	Glu	Pro	Asp	Asn	Leu	Glu	Arg	Met
				340						345					350
Val	Ser	Ile	Leu	Ser	Phe	Val	Ala	Val	Arg	Leu	Leu	Gln	Leu	Arg	Glu
				355						360					365
Ser	Phe	Thr	Leu	Pro	Gln	Ala	Leu	Arg	Ala	Gln	Gly	Leu	Leu	Lys	Glu
				370						375					380
Ala	Glu	His	Val	Glu	Ser	Gln	Ser	Ala	Glu	Thr	Val	Leu	Thr	Pro	Asp
385						390									400
Glu	Cys	Gln	Leu	Leu	Gly	Tyr	Leu	Asp	Lys	Gly	Lys	Arg	Lys	Arg	Lys
					405										415
Glu	Lys	Ala	Gly	Ser	Leu	Gln	Trp	Ala	Tyr	Met	Ala	Ile	Ala	Arg	Leu
				420											430
Gly	Gly	Phe	Met	Asp	Ser	Lys	Arg	Thr	Gly	Ile	Ala	Ser	Trp	Gly	Ala
				435											445
Leu	Trp	Glu	Gly	Trp	Glu	Ala	Leu	Gln	Ser	Lys	Leu	Asp	Gly	Phe	Leu
				450											460
Ala	Ala	Lys	Asp	Leu	Met	Ala	Gln	Gly	Ile	Lys	Ile				
465						470									475

<210> 12

<211> 1428

<212> DNA

<213> 人工序列

<400> 12

```

atgatcacct ctgctctgca cctgtctgct gactgggcta aatctgtttt ctcttctgct 60
gctctgggtg acccgctcg taccgctcgt ctggttaacg ttgctgctca gctggctaaa 120
tactctggta aatctatcac catctcttct gaaggttctg aagctatgca ggaaggtgct 180
taccgtttca tccgtaaccc gaacgtttct gctgaagcta tccgtaaagc tggtgctatg 240
cagaccgtta aactggctca ggaatttccg gaactgctgg ctatcgaaga caccacctct 300
ctgtcttacc gtcaccaggt tgctgaagaa ctgggtaaac tgggttctat ccaggacctg 360
tctcgtaggtt ggtgggttca ctctgttctg ctgctggaag ctaccacctt ccgtaccgtt 420
ggtctgctgc accaggaatg gtggatgcgt ccggacgacc cggctgacgc tgacgaaaaa 480
gaatctggta aatggctggc tgctgctgct acctctcgtc tgcgtatggg ttctatgatg 540
tctaacgtta tcgctgtttg cgaccgtgct gctgacatcc acgcttacct gcaggacaaa 600

```


ctggctcaca acgaacgttt cgttggtcgt tctaaacacc cgcgtaaaga cgttgaatct 660
 ggtctgtacc tgtacgacca cctgaaaaac cagccggaac tgggtgggta ccagatctct 720
 atccccgaga aagggtgtgt tgacaaacgt ggtaaacgta aaaaccgtcc ggctcgtaaa 780
 gcttctctgt ctctgcgttc tggtcgtatc accctgaaac agggtaacat caccctgaac 840
 gctgttctgg ctgaagaaat caaccgcgcg aaagggtgaaa ccccgctgaa atggctgctg 900
 ctgacctctg aaccggttga atctctggct caggctctgc gtgttatcga catctacacc 960
 caccgttggc gtatcgaaga atttcacaaa gcttggaaaa ccggtgctgg tgctgaacgt 1020
 cagcgatagg aagaaccgga caacctggaa cgtatggttt ctatcctgtc tttcgttgc 1080
 gttcgtctgc tgcagctgcg tgaatctttc accctgccgc aggctctgcg tgctcagggt 1140
 ctgctgaaaag aagctgaaca cgttgaatct cagtctgctg aaaccgttct gaccccgac 1200
 gaatgccagc tgctgggta cctggacaaa ggtaaacgta aacgtaaaga aaaagctggt 1260
 tctctgcagt gggcttacat ggetatcgt cgtctgggtg gtttcatgga ctctaaacgt 1320
 accggtatcg cttcttgggg tgctctgtgg gaaggttggg aagctctgca gtcgaagctg 1380
 gacggcttcc tggctgctaa ggacctgatg gctcagggtg tcaaaatc 1428

<210> 13

<211> 476

<212> PRT

<213> 人工序列

<400> 13

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

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

465
 <210> 14
 <211> 1428
 <212> DNA
 <213> 人工序列
 <400> 14
 atgatcacct ctgctctgca ccgtgctgct gactgggcta aatctgtttt ctcttctgct 60
 gctctgggtg acccgcgctcg taccgctcgt ctggttaacg ttgctgctca gctggctaaa 120
 tactctggta aatctatcac catctcttct gaaggttctg aagctatgca ggaagggtgct 180
 taccgtttca tccgtaaccc gaacgtttct gctgaagcta tccgtaaagc tggtgctatg 240
 cagaccgtta aactggetca ggaatttccg gaactgctgg ctatcgaaga caccacctct 300
 ctgtcttacc gtcaccaggt tgetgaagaa ctgggtaaac tgggttctat ccaggaccag 360
 tctctgggtt ggtgggttca ctctgttctg ctgctggaag ctaccacctt ccgtaccgtt 420
 ggtctgctgc accagaactg gtggatgcgt ccggacgacc cggctgacgc tgacgaaaaa 480
 gaatctggta aatggctggc tgctgctgct acctctcgtc tgcgtatggg ttctatgatg 540
 tctaacgtta tcgctgtttg cgaccgtgaa gctgacatcc acgcttacct gcaggacaaa 600
 ctggctcaca acgaacgttt cgttgttcgt tctaaacacc cgcgtaaaga cgttgaatct 660
 ggtctgtacc tgtacgacca cctgaaaaac cagccggaac tgggtgggta ccagatctct 720
 atcccgcaaga aagggtgtgt tgacaaacgt ggtaaacgta aaaaccgtcc ggctcgtaaa 780
 gcttctctgt ctctgcgttc tggctgtatc accctgaaac agggtaacat caccctgaac 840
 gctgttctgg ctgaagaaat caaccgccg aaagggtgaaa ccccgctgaa atggctgctg 900
 ctgacctctg aaccggttga atctctggct caggctctgc gtgttatcga catctacacc 960
 caccgttggc gtatcgaaga atttcacaaa gcttggaaaa ccggtgctgg tgctgaacgt 1020
 cagcgtatgg aagaaccgga caacctggaa cgtatgggtt ctatcctgtc tttcgttgct 1080
 gtctgtctgc tgcagctgcg tgaatctttc acctgccgc aggctctgcg tgctcagggt 1140
 ctgctgaaag aagctgaaca cgttgaatct cagtctgctg aaaccgttct gaccccgac 1200
 gaatgccagc tgctgggtta cctggacaaa ggtaaacgta aacgtaaaga aaaagctggt 1260
 tctctgcagt gggcttacat ggctatcgtc cgtctgggtg gtttcatgga ctctaaacgt 1320
 accggtatcg cttcttgggg tgctctgtgg gaagggtggg aagctctgca gtcgaagctg 1380
 gacggcttcc tggctgctaa ggacctgatg gctcagggtg tcaaaatc 1428

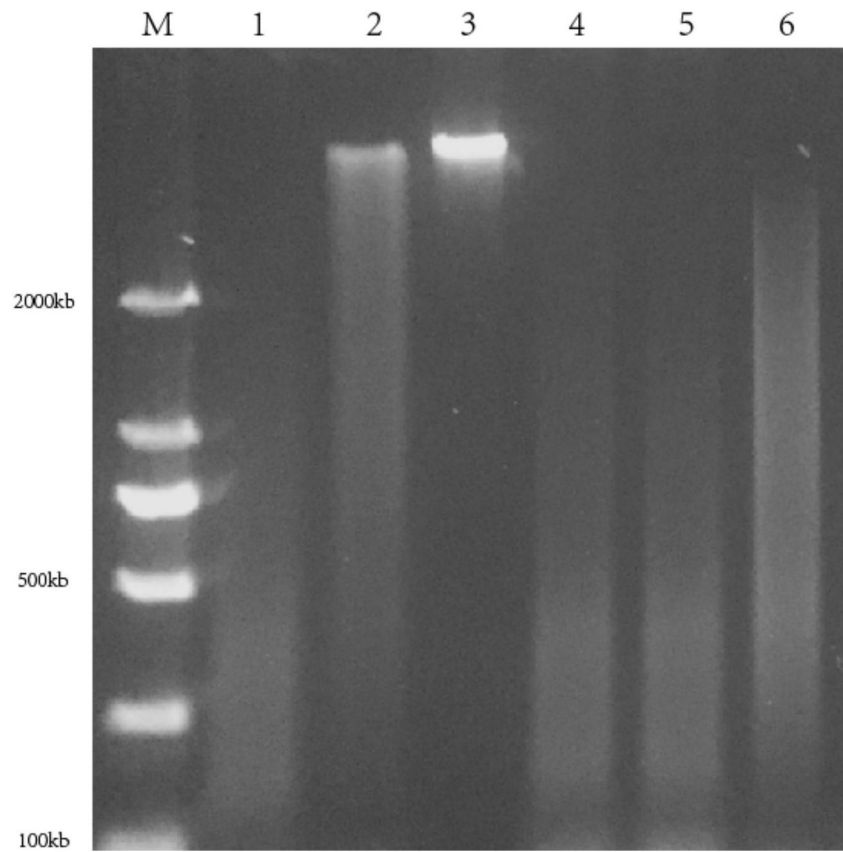


图1

与抗体结合能力

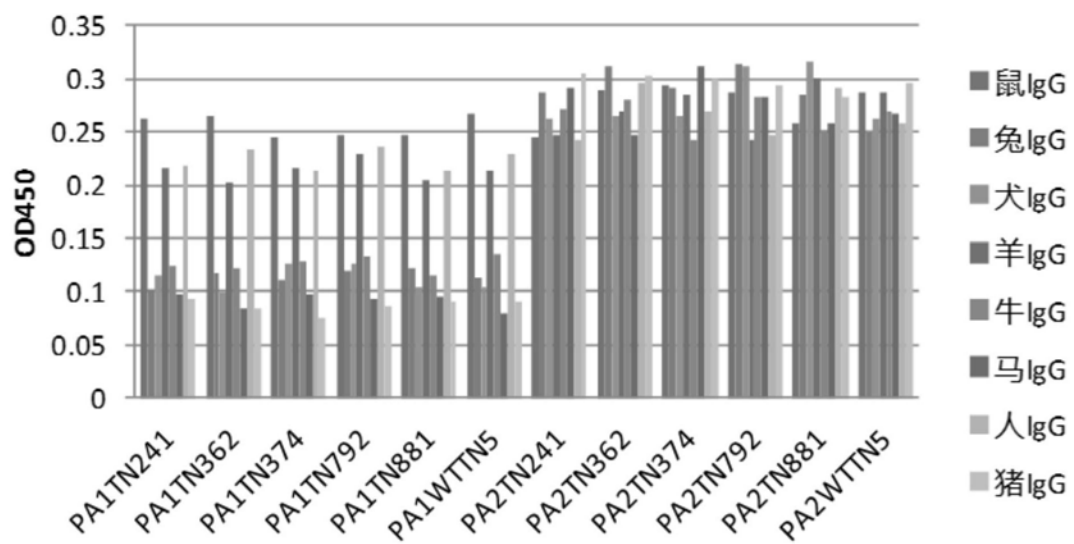


图2

专利名称(译)	一种转座酶突变体、融合蛋白、其制备方法和应用		
公开(公告)号	CN110938610A	公开(公告)日	2020-03-31
申请号	CN201911295765.1	申请日	2019-12-16
[标]申请(专利权)人(译)	广东菲鹏生物有限公司		
申请(专利权)人(译)	广东菲鹏生物有限公司		
当前申请(专利权)人(译)	广东菲鹏生物有限公司		
[标]发明人	张志毅 章瑞程 贺婷 杜翔斐 钟淑瑶		
发明人	张志毅 章瑞程 贺婷 杜翔斐 钟淑瑶		
IPC分类号	C12N9/12 C12N15/54 C12N15/62 C07K19/00 C12Q1/6806 C12N15/10 C40B50/06 G01N33/68 G01N33/53		
CPC分类号	C07K2319/705 C12N9/1241 C12N15/1093 C12Q1/6806 C40B50/06 G01N33/53 G01N33/68		
代理人(译)	周文波		
外部链接	Espacenet SIPO		

摘要(译)

本发明公开了一种转座酶突变体、融合蛋白、其制备方法和应用，涉及生物技术领域，所述转座酶突变体相较于野生型转座酶具有至少一种突变形式，所述野生型转座酶的氨基酸序列如SEQ ID No.1所示，该转座酶突变体的稳定性高，能长期于-20℃的条件下储存，可以显著降低储存成本，使转座酶或其构建的融合蛋白的使用更方便，简化实验流程。

与抗体结合能力

