

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局



(10) 国际公布号

WO 2018/081978 A1

(43) 国际公布日

2018 年 5 月 11 日 (11.05.2018)

WIPO | PCT

(51) 国际专利分类号:

C07K 19/00 (2006.01)

C12N 15/63 (2006.01)

C12N 15/52 (2006.01)

C12N 5/10 (2006.01)

RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

(21) 国际申请号:

PCT/CN2016/104450

本国际公布:

(22) 国际申请日:

2016 年 11 月 3 日 (03.11.2016)

— 包括国际检索报告 (条约第 21 条 (3))。

— 包括说明书序列表部分 (细则 5.2(a))。

(25) 申请语言:

中文

(26) 公布语言:

中文

(71) 申请人: 深圳华大基因研究院 (BGI SHENZHEN)

[CN/CN]; 中国广东省深圳市盐田区北山工业区综合楼, Guangdong 518083 (CN)。

(72) 发明人: 王健 (WANG, Jian); 中国广东省深圳市

盐田区北山工业区综合楼, Guangdong 518083

(CN)。 胡莹莹 (HU, Yingying); 中国广东省深圳市

盐田区北山工业区综合楼, Guangdong 518083

(CN)。 罗舟 (LUO, Zhou); 中国广东省深圳市盐田

区北山工业区综合楼, Guangdong 518083 (CN)。

(74) 代理人: 上海一平知识产权代理有限公司

(XU & PARTNERS, LLC); 中国上海市普陀

区真北路 958 号天地科技广场 1 号楼 106

室, Shanghai 200333 (CN)。

(81) 指定国 (除另有指明, 要求每一种可提供的国家

保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,

CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,

GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP,

KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU,

LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,

NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA,

RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST,

SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,

VC, VN, ZA, ZM, ZW。

(84) 指定国 (除另有指明, 要求每一种可提供的地区

保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,

NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,

AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,

CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,

IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,

(54) Title: METHOD AND SYSTEM FOR IMPROVING GENE EDITING EFFICIENCY

(54) 发明名称: 提高基因编辑效率的方法和系统

(57) Abstract: The present invention provides a method and system for improving a gene editing efficiency. Specifically, in the present invention, fusion protein formed after PEST short peptide and Cas9 protein are fused can remarkably improve the gene editing efficiency.

(57) 摘要: 本发明提供了一种提高基因编辑效率的方法和系统, 具体地, 本发明公开了将 PEST 短肽和 Cas9 蛋白融合后形成的融合蛋白能够显著提高基因编辑的效率。



WO 2018/081978 A1

提高基因编辑效率的方法和系统

技术领域

本发明属于生物技术领域，具体地说，本发明涉及提高基因编辑效率的方法和系统。

背景技术

CRISPR/Cas9 或 CRISPR-Cas9 系统是近年来兴起的用于基因编辑的工具。该系统由能够特异性识别 DNA 序列的 sgRNA 和能够切割 DNA 序列的 Cas9 核酸酶组成。通过 Cas9 对基因组上特定序列的切割，可以实现定点突变、片段删除或倒位等修饰。在加入外源供体 DNA 的条件下，还能够实现定点插入或替换，包括同源重组介导的精确的遗传修饰。除了对基因组进行编辑，工程化的 Cas9 还可以对 RNA、转录水平、以及表观遗传进行调控。尽管 CRISPR/Cas9 技术具有简便、快速、特异等优点，但在细胞、动物、植物的应用中，其效率尚不能充分满足基因编辑，尤其是定点插入或替换的需求。提高这一系统的效率是基因编辑技术领域关注的重点。

因此，本领域技术人员致力于开发提高基因编辑效率的方法及其应用。

发明内容

本发明的目的在于提供一种提高基因编辑效率的方法和系统。

在本发明的第一方面，提供了一种融合蛋白，所述的融合蛋白具有式 Ia 或 Ib 所述结构：



其中，

E 为核酸内切酶蛋白元件；

P 为 PEST 蛋白元件；

“-”表示连接上述各元件的肽键或肽接头。

在另一优选例中，所述的核酸内切酶蛋白元件选自下组：Cas9 蛋白（包括 SpCas9、SaCas9、NmCas9、St1Cas9）及其变体（如 SpCas9 的 VQR、EQR、VRER 等变体）、Cpf1 蛋白（包括 AsCpf1、FnCpf1、LbCpf1）、C2c2 蛋白、Argonaute 蛋白家族；TALE 蛋白、锌指蛋白、dCas9 等与 FokI 融合构成的人工核酸内切酶。

在另一优选例中，所述核酸内切酶蛋白元件为 Cas9 蛋白。

在另一优选例中，所述 SpCas9 蛋白的氨基酸序列如 SEQ ID NO.1 所示。

在另一优选例中，所述 PEST 蛋白元件选自下组：

ODC1 蛋白的 PEST 序列、GCN4 蛋白的 PEST 序列、CLN2/CLN3 蛋白的 PEST 序列、NIMA 蛋白的 PEST 序列、Cactus 蛋白的 PEST 序列、HDC 蛋白的 PEST 序列、CPEB 蛋白的 PEST 序列、NPDC1 蛋白的 PEST 序列、FOS 蛋白的 PEST 序列、NFKBIA 蛋白的 PEST 序列等。

在另一优选例中，所述 PEST 蛋白元件的氨基酸序列如 SEQ ID NO. 2 所示。

在另一优选例中，所述融合蛋白的氨基酸序列选自下组：

(A) 具有 SEQ ID NO. 2 所示氨基酸序列的多肽；

(B) 具有与 SEQ ID NO. 2 所示氨基酸序列 $\geq 80\%$ 同源性（优选地， $\geq 90\%$ 的同源性；等优选地 $\geq 95\%$ 的同源性；最优选地， $\geq 97\%$ 的同源性，如 98% 以上， 99% 以上）的多肽，且所述多肽具有基因编辑活性；

(C) 将 SEQ ID NO. 3 所示氨基酸序列经过 1-10 个氨基酸残基的取代、缺失或添加而形成的，且保留基因编辑活性的衍生多肽。

在另一优选例中，所述融合蛋白 (E-P or P-E) 中的肽键或肽接头的长度为 0-30 个氨基酸。

在另一优选例中，所述融合蛋白还包括任选地核定位信号元件 (Nuclear location signal, NLS)；优选地所述核定位信号元件添加在融合蛋白的 N 端或者 C 端，或者同时添加在 N 端和 C 端。

在本发明的第二方面，提供了一种分离的多核苷酸，所述的多核苷酸编码本发明第一方面所述的融合蛋白。

在另一优选例中，所述多核苷酸包括适于原核或真核细胞表达的密码子优化序列。

在另一优选例中，所述多核苷酸为 DNA 或 RNA。

在另一优选例中，所述多核苷酸为 mRNA。

在本发明的第三方面，提供了一种载体，所述载体含有本发明第二方面所述的多核苷酸。

在另一优选例中，所述的载体包括质粒、病毒载体等。

在另一优选例中，所述的病毒载体包括：慢病毒载体、腺病毒载体、腺相关病毒载体、逆转录病毒载体、黄热病毒载体和疱疹病毒载体等。

在另一优选例中，所述的载体包括原核和真核表达载体。

在本发明的第四方面提供了一种基因工程细胞，所述基因工程细胞含有本发明第三方面所述的载体，或含有本发明第二方面所述的多核苷酸，或含有本发明第一方面所述的融合蛋白。

在另一优选例中，所述基因工程细胞包括原核细胞和真核细胞。

在另一优选例中，所述基因工程细胞为动物细胞、植物细胞、或微生物细胞。

在另一优选例中，所述细胞为生殖细胞或受精卵。

在本发明的第五方面，提供了一种基因编辑系统，所述基因编辑系统包括本发明第一方面所述的融合蛋白、或本发明第二方面所述的多核苷酸、或本发明第三方面的载体。

在另一优选例中，所述基因编辑系统选自下组：CRISPR/Cas 基因编辑系统、CRISPR/Cpf1 基因编辑系统、CRISPR/C2c2 基因编辑系统、Argonaute/gDNA 基因编辑系统、锌指核酸酶基因编辑系统 (ZFNs)、和类转录激活因子核酸酶基因

编辑系统 (TALENs)。

本发明的第六方面，提供了一种核酸内切酶介导的基因编辑方法，所述方法包括步骤：

使用本发明第一方面所述的融合蛋白对靶基因进行编辑。

在本发明的第七方面，提供了一种提高核酸内切酶介导的基因编辑系统的基因编辑效率的方法，所述方法包括步骤：

在靶细胞中，表达 PEST 蛋白和核酸内切酶的融合蛋白。

在另一优选例中，所述方法中将所述 PEST 蛋白的编码序列融合至所述核酸内切酶基因的 5' 或 3' 端。

在另一优选例中，所述方法包括步骤：

在靶细胞中，表达本发明第一方面所述的融合蛋白。

应理解，在本发明范围内中，本发明的上述各技术特征和在下文(如实施例)中具体描述的各技术特征之间都可以互相组合，从而构成新的或优选的技术方案。限于篇幅，在此不再一一累述。

附图说明

图 1 显示了核酸内切酶蛋白元件与 PEST 蛋白元件融合的方式。PEST 元件可以位于核酸内切酶元件的 C 端或者 N 端，其中的短线代表两者之间的肽键或肽接头，为 0-30 个氨基酸残基。

图 2a 和图 2b 分别显示了示范 SpCas9 蛋白的氨基酸序列 (SEQ NO:1) 和编码多核苷酸序列 (SEQ ID NO. 4)。

图 3 显示了示范小鼠 ODC1 蛋白的 PEST 结构域的氨基酸序列 (SEQ ID NO. 2) 和编码多核苷酸序列 (SEQ ID NO. 5)。

图 4a 和图 4b 分别显示了一例将小鼠 ODC1 蛋白的 PEST 结构域融合至 SpCas9 蛋白 C 端所得融合蛋白的氨基酸序列 (SEQ ID NO. 3) 和编码多核苷酸序列 (SEQ ID NO. 6)。

图 5 显示了示范 Cas9-PEST 质粒图谱。在原有的 pGH-T7-Cas9 载体上，Cas9 编码序列与核定位信号 (NLS) 序列之间加入 PEST 编码序列的完整的质粒图谱。方框标记了 PEST 编码序列的位置。

图 6 显示了斑马鱼 *gatala* 基因的结构和 CRISPR/Cas9 靶位点信息。该基因具有 6 个外显子 (方框所示，实心部分代表编码区，空心部分代表非编码区)，本发明实施例中所选择的 CRISPR/Cas9 靶位点位于第 5 个外显子上 (背景为灰色的序列)，可用 SpeI 酶切位点来检测该靶位点的效率。

图 7 显示了通过 PCR 制备 *gatala* 靶位点的 sgRNA 模板的电泳检测结果，模板大小为 120 bp。

图 8 显示了 Cas9-PEST 提高了斑马鱼 *gatala* 基因靶位点的定点突变效率。选择斑马鱼 *gatala* 基因上的已知靶位点，注射相同剂量的 Cas9 和 Cas9-PEST，使用限制性内切酶法检测定点突变效率 (最上面的条带为突变的条带，黑色三角指示。该条带的亮度占三条带总亮度的比值即为突变效率。ImageJ 软件可以

量化条带的亮度)，计算两组各 6 个样品的平均效率分别为 22.5% 和 35.2%（对照组为未注射的样品），表明 PEST 有助于提高 Cas9 的工作效率。

图 9 显示了斑马鱼 *mstnb* 基因的结构和 CRISPR/Cas9 靶位点信息。该基因具有 3 个外显子（方框所示，实心部分代表编码区，空心部分代表非编码区），本发明实施例中所选择的 CRISPR/Cas9 靶位点位于第 1 个外显子上（背景为灰色的序列），可用 BslI 酶切位点来检测该靶位点的效率。

图 10 显示了 Cas9-PEST 可将斑马鱼 *mstnb* 基因靶位点的定点突变效率从 13.4% 提高至 56.1%。

具体实施方式

本发明人通过广泛而深入的研究后发现，将 PEST 短肽和 Cas9 蛋白融合后形成的融合蛋白能够显著提高基因编辑效率，在此基础上，完成了本发明。

在描述本发明之前，应当理解本发明不限于所述的具体方法和实验条件，因为这类方法和条件可以变动。还应当理解本文所用的术语其目的仅在于描述具体实施方案，并且不意图是限制性的，本发明的范围将仅由所附的权利要求书限制。

除非另外定义，否则本文中所用的全部技术与科学术语均具有如本发明所属领域的普通技术人员通常理解的相同含义。如本文所用，在提到具体列举的数值中使用术语“约”意指该值可以从列举的值变动不多于 1%。例如，如本文所用，表述“约 100”包括 99 和 101 之间的全部值（例如，99.1、99.2、99.3、99.4 等）。

虽然在本发明的实施或测试中可以使用与本发明中所述相似或等价的任何方法和材料，本文在此处例举优选的方法和材料。

本发明要解决的问题是提高 CRISPR 系统在受精卵分裂早期的编辑效率，从而增加在动物成体的编辑效率。本方法通过提高 Cas9 mRNA 翻译的速率，促使 Cas9 蛋白能够在受精卵发育早期中尽早地发挥功能，且更有效地编辑靶向基因。

通过在 Cas9 编码区域的 3' 端融合小鼠 ODC1 蛋白中 PEST 序列的编码序列，可有效提高 Cas9mRNA 翻译成为蛋白的速率，从而提高基因编辑的效率。

本发明中选择将 ODC1 蛋白中的 PEST 序列与 Cas9 进行融合，其他可以有效提高 Cas9 mRNA 翻译效率的序列或者结构域，比如其他蛋白的 PEST 序列，与 Cas9 进行融合后，都有可能提高在动物中 CRISPR/Cas9 介导的基因编辑效率。

另外需指出，作为本发明的一个例子使用的 Cas9 是经过斑马鱼密码子优化的 SpCas9。本发明还可能用于改造 SaCas9、NmCas9、St1Cas9、Cpf1、C2c2、Argonaute、TALEN、ZFN 等任何用于基因编辑的内切酶。

PEST 蛋白

PEST 蛋白（PEST 短肽）是一类特殊的信号肽，富含脯氨酸（P）、谷氨酸（E）、丝氨酸（S）和苏氨酸（T），可使蛋白的半衰期缩短。

已知自然界中的多种蛋白如 ODC1、HDC、NPDC、CPEB、NFKBIA 等含有 PEST 短肽（结构域），实验表明它们与蛋白的快速降解有关，降解的途径可能是通

过蛋白酶体或者钙蛋白酶。PEST 短肽的序列、长度、位置和二级结构在不同蛋白中都有所变化，有些蛋白含有多个 PEST 结构域。

在本发明的一个优选地实施方式中，所述 PEST 蛋白元件选自下组：

ODC1 蛋白的 PEST 序列、GCN4 蛋白的 PEST 序列、CLN2/CLN3 蛋白的 PEST 序列、NIMA 蛋白的 PEST 序列、Cactus 蛋白的 PEST 序列、HDC 蛋白的 PEST 序列、CPEB 蛋白的 PEST 序列、NPDC1 蛋白的 PEST 序列、FOS 蛋白的 PEST 序列、NFKBIA 蛋白的 PEST 序列等(具体可参考文献：Rogers et al. 1986, Rechsteiner and Rogers 1996, Fleming and Wang 2000, Reverte et al. 2001, Spencer et al. 2004)

在本发明的一个优选的实施方式中，所述 PEST 蛋白的氨基酸序列如 SEQ ID NO. 2 所示；其编码核苷酸序列（来自小鼠 *Odc1* 的 cDNA）如 SEQ ID NO. 5 所示。

核酸内切酶蛋白元件

本发明中所述的核酸内切酶蛋白元件是指可以用于基因编辑的核酸内切酶、其变体、及其活性片段。

典型的核酸内切酶包括：Cas9 蛋白(包括 SpCas9、SaCas9、NmCas9、St1Cas9)及其变体（如 SpCas9 的 VQR、EQR、VRER 等变体）(Cong et al. 2013, Hou et al. 2013, Mali et al. 2013, Kleinstiver et al. 2015, Ran et al. 2015)、Cpf1 蛋白（包括 AsCpf1、FnCpf1、LbCpf1）(Zetsche et al. 2015, Kim et al. 2016, Kleinstiver et al. 2016)、C2c2 蛋白(Abudayyeh et al. 2016)、Argonaute 蛋白家族(Gao et al. 2016)；TALE 蛋白、锌指蛋白、dCas9 等与 FokI 融合构成的人工核酸内切酶(Kim et al. 1996, Bibikova et al. 2002, Miller et al. 2011, Tsai et al. 2014)。

CRISPR/Cas 系统

CRISPR/Cas 系统为目前发现存在于大多数细菌与所有的古菌中的一种后天免疫系统，以消灭外来的质体或者噬菌体，并在自身基因组中留下外来基因片段作为“记忆”。全名为常间回文重复序列丛集/常间回文重复序列丛集关联蛋白系统(clustered regularly interspaced short palindromic repeats/CRISPR-associated proteins)。

目前已发现三种不同类型的 CRISPR/Cas 系统，存在于大约 40%和 90%已测序的细菌和古菌中。其中第二型的组成较为简单，以 Cas9 蛋白以及向导 RNA(gRNA)为核心的组成，由于其对 DNA 干扰(DNAi)的特性，目前被积极地应用于遗传工程中，作为基因体剪辑工具，与锌指核酸酶(ZFN)及类转录活化因子核酸酶(TALEN)同样利用非同源性末端接合(NHEJ)的机制，于基因体中产生去氧核糖核酸的双股断裂以利剪辑。二型 CRISPR/Cas 并经由遗传工程的改造应用于哺乳类细胞及斑马鱼的基因体剪辑。其设计简单以及操作容易的特性为最大的优点。未来将可应用在各种不同的模式生物当中。

称为 CRISPR 的基因组重复丛集，即原核生物拟核 DNA 链中的丛生重复序列，在 1987 关于 *E. coli* 的一份研究报告中被首次描述。2000 年，相似的重重复序列在其它真细菌和古细菌中被发现并被命名为短间隔重复序列(Short Regularly Spaced Repeats, SRSR)。2002 年 SRSR 被重命名为 CRISPR。其中一部分基因编码

的蛋白为核酸酶和解旋酶。这些关联蛋白 (CAS, CRISPR-associated proteins) 与 CRISPR 组成了 CRISPR/Cas 系统。

CRISPR/Cas 技术

本发明所称的“CRISPR/Cas 技术”、“CRISPR/Cas 基因编辑”、“CRISPR/Cas 基因编辑技术”、“CRISPR/Cas 基因编辑方法”均指利用 CRISPR/Cas 系统的原理对目的基因进行改造的基因编辑技术。

Cas9 蛋白及其变体

CRISPR/Cas 的核心就是 Cas9 蛋白以及向导 RNA (gRNA)。在不同物种中能够利用 CRISPR/Cas 系统进行基因编辑的核心技术, 首要一步即在该物种中异源表达有 DNA 剪切酶活性的 Cas9 蛋白, 第二步则是获得 gRNA 和靶点同源序列将 Cas9 引导至靶点进行 DNA 剪切。其中第二步中, 具体的操作方法是本领域技术人员所熟知的。

来源于 *Streptococcus pyogenes* 的 Cas9 蛋白是一种多结构域多功能的 Cas 蛋白, 其 N 端具有类 RuvC 核酸酶的结构域, 其中部具有 HNH 核酸酶结构域。Cas9 蛋白与 gRNA 结合能够实现在特异位点处切割 DNA, 来源于 *Streptococcus pyogenes* 的 CRISPR/Cas 系统识别序列为 23 bp, 并能靶向 20 bp, 其识别位点最末 3 位 NGG 序列被称作 PAM (protospacer adjacent motif) 序列, 其对于 DNA 切割非常重要。目前大多数真核生物 (包括家蚕、拟南芥、酵母、线虫等) 的 CRISPR/Cas 系统最初均源自 *Streptococcus pyogenes*, Cas9 蛋白则多是经过人源化改造或其它物种的密码子优化。

较佳地, 本发明所提供的 Cas9 来源于化脓链球菌 (*Streptococcus pyogenes*)。在本发明的一个优选地实施方式中, 所述 Cas9 蛋白的氨基酸序列如 SEQ ID NO. 1 所示;

其编码核酸序列 (经过斑马鱼密码子优化) 如 SEQ ID NO. 4 所示。

融合蛋白

本发明提供了 PEST 蛋白和 Cas9 蛋白的融合蛋白及其编码序列 (包括 DNA、mRNA)。

本发明的融合蛋白, 可任选地含有连接肽。连接肽大小和复杂性可能会影响蛋白的活性。通常, 连接肽应当具有足够的长度和柔韧性, 以保证连接的两个蛋白在空间上有足够的自由度以发挥其功能。同时避免连接肽中形成 α 螺旋或 β 折叠等对融合蛋白的稳定性的影响。连接肽的长度一般为 0-30 个氨基酸, 较佳地 1-5 个氨基酸。

应理解, 所述术语还包括本发明融合蛋白的衍生物, 指本发明融合蛋白在经过 1-3 个氨基酸添加或替换、1-2 个氨基酸缺失并仍具有基因编辑活性的多肽。这些保守性变异多肽最好根据表 1 进行氨基酸替换而产生。

表 1

最初的残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln

Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

一旦鉴定获得了相关的肽序列，就可以用重组法来大批量地获得相关肽序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到相关肽(融合蛋白)。

此外，还可用化学方法直接合成相关肽序列。

在本发明的一个优选地实施方式中，所述融合蛋白的氨基酸序列如SEQ ID NO. 3所示，其编码多核苷酸序列如SEQ ID NO. 6所示。

多核苷酸

本发明的多核苷酸，是编码本发明融合蛋白的DNA或RNA序列，可以全序列人工合成。

本发明的多核苷酸可以是DNA形式或RNA形式。DNA形式包括cDNA、基因组DNA或人工合成的DNA。DNA可以是单链的或是双链的。DNA可以是编码链或非编码链。RNA形式包括mRNA。

本发明还涉及上述多核苷酸的变异体，其编码与本发明有相同的氨基酸序列的蛋白质片段、类似物和衍生物。此多核苷酸的变异体可以是天然发生的等位变异体或非天然发生的变异体。这些核苷酸变异体包括取代变异体、缺失变异体和插入变异体。如本领域所知的，等位变异体是一个多核苷酸的替换形式，它可能是一个或多个核苷酸的取代、缺失或插入，但不会从实质上改变其编码多肽的功能。

如本文所用，术语“引物”指的是在与模板配对，在DNA聚合酶的作用下能以其为起点进行合成与模板互补的DNA链的寡聚核苷酸的总称。引物可以是天然的RNA、DNA，也可以是任何形式的天然核苷酸。引物甚至可以是非天然的核苷酸如LNA或ZNA等。引物“大致上”(或“基本上”)与模板上一条链上的

一个特殊的序列互补。引物必须与模板上的一条链充分互补才能开始延伸，但引物的序列不必与模板的序列完全互补。比如，在一个 3' 端与模板互补的引物的 5' 端加上一段与模板不互补的序列，这样的引物仍大致上与模板互补。只要有足够长的引物能与模板充分的结合，非完全互补的引物也可以与模板形成引物-模板复合物，从而进行扩增。

本发明融合蛋白或其元件的核苷酸全长序列或其片段通常可以用 PCR 扩增法、重组法或人工合成的方法获得。对于 PCR 扩增法，可根据已公开的有关核苷酸序列，尤其是开放阅读框序列来设计引物，并用市售的 cDNA 库或按本领域技术人员已知的常规方法所制备的 cDNA 库作为模板，扩增而得有关序列。当序列较长时，常常需要进行两次或多次 PCR 扩增，然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列，就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外，还可用人工合成的方法来合成有关序列，尤其是片段长度较短时。通常，通过先合成多个小片段，然后再进行连接可获得序列很长的片段。

应用 PCR 技术扩增 DNA/RNA 的方法被优选用于获得本发明的基因。用于 PCR 的引物可根据本文所公开的本发明的序列信息适当地选择，并可用常规方法合成。可用常规方法如通过凝胶电泳分离和纯化扩增的 DNA/RNA 片段。

如本文所用，“可操作地连于”指这样一种状况，即线性 DNA 序列的某些部分能够影响同一线性 DNA 序列其他部分的活性。例如，如果信号肽 DNA 作为前体表达并参与多肽的分泌，那么信号肽(分泌前导序列)DNA 就是可操作地连于多肽 DNA；如果启动子控制序列地转录，那么它就是可操作地连于编码序列；如果核糖体结合位点被置于能使其翻译的位置时，那么它是可操作地连于编码序列。一般“可操作地连于”意味着相邻近，而对于分泌前导序列则意味着在阅读框中相邻。

如本文所用，术语“载体”包括质粒、粘粒、表达载体、克隆载体等。

在本发明中，可选用本领域已知的各种载体如市售的载体。比如，选用市售的载体，然后将编码本发明新免疫毒素的核苷酸序列可操作地连于表达调控序列，形成蛋白表达载体。

本发明也涉及包含本发明的多核苷酸的载体，以及用本发明的载体或融合蛋白编码序列经基因工程产生的宿主细胞，以及经重组技术产生本发明所述蛋白质的方法。

本领域的技术人员熟知的方法能用于构建含本发明蛋白的编码 DNA 序列和合适的转录/翻译控制信号的表达载体。这些方法包括体外重组 DNA 技术、DNA 合成技术、体内重组技术等。所述的 DNA 序列可有效连接到表达载体中的适当启动子上，以指导 mRNA 合成。表达载体还包括翻译起始用的核糖体结合位点和转录终止子。

此外，表达载体优选地包含一个或多个选择性标记基因，以提供用于选择转化的宿主细胞的表型性状，如真核细胞培养用的二氢叶酸还原酶、新霉素抗

性以及绿色荧光蛋白(GFP)，或用于大肠杆菌的四环素或氨苄青霉素抗性。

包含上述的适当 DNA 序列以及适当启动子或者控制序列的载体，可以用于转化适当的宿主细胞，以使其能够表达蛋白质。

宿主细胞(基因工程细胞)可以是原核细胞，如细菌细胞；或是低等真核细胞，如酵母细胞；或是高等真核细胞，如哺乳动物细胞。代表性例子有：大肠杆菌，链霉菌属的细菌细胞；真菌细胞如酵母；植物细胞；果蝇 S2 或 Sf9 的昆虫细胞；CHO、NS0、COS7、或 293 细胞的动物细胞等。

用重组 DNA 转化宿主细胞可用本领域技术人员熟知的常规技术进行。当宿主为原核生物如大肠杆菌时，能吸收 DNA 的感受态细胞可在指数生长期后收获，用 CaCl_2 法处理，所用的步骤在本领域众所周知。另一种方法是使用 MgCl_2 。如果需要，转化也可用电穿孔的方法进行。当宿主是真核生物，可选用如下的 DNA 转染方法：磷酸钙共沉淀法，常规机械方法如显微注射、电穿孔、脂质体包装等。

获得的转化子可以用常规方法培养，表达本发明的基因所编码的多肽。根据所用的宿主细胞，培养中所用的培养基可选自各种常规培养基。在适于宿主细胞生长的条件下进行培养。当宿主细胞生长到适当的细胞密度后，用合适的方法(如温度转换或化学诱导)诱导选择的启动子，将细胞再培养一段时间。

在上面的方法中的蛋白质可在细胞内、或在细胞膜上表达、或分泌到细胞外。如果需要，可利用其物理的、化学的和其它特性通过各种分离方法分离和纯化蛋白。这些方法是本领域技术人员所熟知的。这些方法的例子包括但不限于：常规的复性处理、用蛋白沉淀剂处理(盐析方法)、离心、渗透破菌、超离心、分子筛层析(凝胶过滤)、吸附层析、离子交换层析、高效液相层析(HPLC)和其它各种液相层析技术及这些方法的结合。

在基因编辑过程中，可以对编码本发明融合蛋白的基因进行体外转录，获得对应的 mRNA，然后将 mRNA 注射入靶细胞中翻译产生本发明的融合蛋白，在 gRNA 配合下对靶基因进行编辑。

基因编辑方法

本发明通过无缝克隆的方法将编码 PEST 短肽的 DNA 序列融合至 Cas9 编码序列的 3' 端，然后体外转录得到 Cas9-PEST mRNA 用于受精卵注射。具体的方案如下：

1. 将 pGH-T7-Cas9 载体(该载体参考文献(Liu et al. 2014)，此处 Cas9 为经过斑马鱼密码子优化的 Cas9 编码序列)用 SphI 和 XbaI 两种限制性内切酶进行双酶切，切除 Cas9 C 端编码区及核定位信号(NLS)，电泳后回收载体骨架。

2. 从小鼠 cDNA 文库中扩增 ODC1 蛋白所含的 PEST 编码序列，在扩增的上、下游引物外侧分别加入与 Cas9、NLS 重叠的约 20 bp 序列，扩增的产物命名为片段 A。

3. 从原载体上分别扩增 Cas9 C 端编码区及 NLS，扩增时也需要在引物外侧加上与载体骨架或 PEST 编码序列重叠的约 20 bp 序列，扩增的产物命名为片段 B 和片段 C。

4. 使用诺维赞公司的无缝克隆试剂盒(CloneExpressMultiS One Step

Cloning Kit) 将 A、B、C 三个片段与载体骨架连接, 由于片段间的重叠序列的设置, 三个片段会按照 B→A→C 的顺序连入载体骨架, 即成功将 PEST 编码序列插入 Cas9 和 NLS 之间(图 5), 获得 Cas9-PEST 质粒。

5. 质粒经测序验证后, 使用 XbaI 进行线性化。回收线性化产物作为模板, 利用 T7 RNA 聚合酶进行体外转录, 获得 Cas9-PEST mRNA。可将此 mRNA 与 sgRNA 按照终浓度分别为 200~300 ng/ μ L 和 40~60 ng/ μ L 混合, 注射斑马鱼受精卵, 每枚受精卵注射 2 nL。

本发明的主要优点在于:

- (1) 本发明提供了能够显著提高基因编辑效率的CRISPR/Cas9系统。
- (2) 与使用 Cas9 蛋白相比, 本发明成本更低, 灵活性更高。
- (3) 本发明可以应用于其他的内切酶, 不受合成蛋白的限制。

下面结合具体实施例, 进一步详陈本发明。应理解, 这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明详细条件的实验方法, 通常按照常规条件如美国 Sambrook. J 等著《分子克隆实验室指南》(黄培堂等译, 北京: 科学出版社, 2002 年) 中所述的条件, 或按照制造厂商所建议的条件。除非另外说明, 否则百分比和份数按重量计算。以下实施例中所用的实验材料和试剂如无特别说明均可从市售渠道获得。

实施例 1

以斑马鱼 *gatala* 基因第 5 个外显子中的 CRISPR/Cas9 靶位点(图 6)为例, 比较 Cas9 mRNA 和 Cas9-PEST mRNA 在斑马鱼胚胎中的基因编辑效率。步骤如下:

1. 制备 sgRNA: 首先以 pMD19-gRNA scaffold 质粒(Chang et al. 2013)为模板, PCR 扩增用于体外转录的 sgRNA 模板。PCR 的正向引物 T7-*gatala*E5-sfd: 5' -taatacgactcactataGTAGTGTGTAGTACTAGTGgttttagagctagaaatagc-3'

(其中小写部分为 T7 启动子序列和 scaffold, 固定不变; 大写部分为靶位点序列); 反向引物 tracr rev: 5' -aaaaaagcaccgactcggtgccac-3'。使用 2×TaqPlatinum Mix (TIANGEN 公司) 进行 PCR, 电泳检测为单一条带后(图 7), 使用超薄 DNA 产物纯化试剂盒(TIANGEN 公司)对 PCR 产物进行纯化, 得到用于体外转录 sgRNA 的模板。该模板的全序列为:

5' -taatacgactcactataGTAGTGTGTAGTACTAGTGgttttagagctagaaatagcaagtta (SEQ ID NO. 7);

aaataaggctagtcgcgttatcaacttgaaaaagtggcaccgagtcggtgctttttttt-3' (SEQ ID NO. 8)。

然后使用 T7 RNA 聚合酶体系(Takara 公司)体外转录 sgRNA, 反应体系一般为 20 μ L。转录完成后加入 1 μ L DNase I (Takara 公司)去除模板。

最后通过乙醇沉淀法纯化 sgRNA: 用 RNase-free H₂O 将反应体系稀释至 200 μ L, 再加入 550 μ L 无水乙醇和 20 μ L 3M 醋酸钠, 混合均匀后于冰上放置 10 分钟。接着在 4℃离心机以 13000 g 的转速离心 15 分钟, 小心弃去上清后, 加入 500 μ L RNase-free 的 70%乙醇, 以 4℃, 13000 g 离心 5 分钟, 再次弃去

上清。室温晾干后,加入 30 μ L RNase-free H_2O 溶解,得到 sgRNA 储液。Nanodrop 测定浓度后于 $-20^{\circ}C$ 保存,注射时进行稀释。

2. 制备 Cas9 mRNA 和 Cas9-PEST mRNA: 将改造前的 pGH-T7-Cas9 质粒和改造的 Cas9-PEST 质粒都用 XbaI (NEB 公司) 进行线性化,电泳检测后使用普通 DNA 产物纯化试剂盒 (TIANGEN 公司) 回收线性化产物。以各 1 μ g 线性化产物为模板,用 T7 mMESSAGE mMACHINE 试剂盒 (Ambion 公司) 体外转录两种 mRNA,反应体系一般为 20 μ L。转录完成后加入 1 μ L TURBO DNase (Ambion 公司) 去除模板。

通过氯化锂沉淀法纯化两种 mRNA: 向 20 μ L 反应体系中加入 30 μ L 氯化锂和 30 μ L RNase-free H_2O , 混合均匀后于 $-20^{\circ}C$ 放置 2 小时以上。接着在 $4^{\circ}C$ 离心机以 13000 g 的转速离心 15 分钟,小心弃去上清后,加入 500 μ L RNase-free 的 70%乙醇,以 $4^{\circ}C$, 13000 g 离心 5 分钟,再次弃去上清。室温晾干后,加入 30 μ L RNase-free H_2O 溶解,得到 Cas9 mRNA 储液和 Cas9-PEST mRNA 储液。Nanodrop 测定浓度后于 $-20^{\circ}C$ 保存,注射时进行稀释。

3. 显微注射: 配置两种注射液各 5 μ L, Cas9 mRNA (或 Cas9-PEST mRNA) 的终浓度为 200 ng/ μ L, sgRNA 的终浓度为 50 ng/ μ L, 注射液中加入约 5% 体积的酚红溶液作为指示剂。将斑马鱼的受精卵分为三组, 第一组注射 Cas9 mRNA+sgRNA, 第二组注射 Cas9-PEST mRNA+sgRNA, 第三组不注射作为对照。确保注射的时期都是单细胞期, 注射量都是每枚卵 2 nL。

4. 检测基因编辑效率: 待受精卵发育至 24 小时, 从两个实验组中各取 6×3 枚, 对照组中取 2×3 枚提取基因组。用 *gatala* 基因靶位点的上、下游引物, 通过 PCR 扩增基因组上覆盖靶位点的序列。引物序列信息见表 2。

表2 实施例中所用基因靶位点的检测引物

引物名称	序列	SEQ ID NO.
<i>gatala</i> 正向引物	5' -GCATTTAGTTCACCAGAAGCG-3'	13
<i>gatala</i> 反向引物	5' -CCTGGGTTCAGAGAATACGC-3'	14
<i>mstnb</i> 正向引物	5' -ATGGAGATATAACGGCGCAC-3'	15
<i>mstnb</i> 反向引物	5' -ATGCGTAAAATTGCTGTGGC-3'	16

使用 $2 \times$ HotstartTaq PCR StarMix (GenStar 公司) 进行 PCR, 电泳检测为单一条带后, 用 SpeI (NEB 公司) (靶位点上含有限制性内切酶位点且在 PCR 产物上是唯一的) 对 PCR 产物进行酶切, 再通过电泳检测酶切产物。由于 Cas9 在靶位点的切割引发定点突变, 可能破坏 SpeI 酶切位点, 因此实验组的 PCR 产物不能完全被 SpeI 切开, 未酶切的比例可代表定点突变的效率, 即 Cas9 的基因编辑效率, 而对照组的 PCR 产物则能完全被切开。

通过计算, 注射 400 pg Cas9 mRNA+100 pg sgRNA 引起的定点突变效率为 22.5%, 而注射相同剂量的 Cas9-PEST mRNA+sgRNA 可将定点突变效率提升至 35.2% (图 8), 基因编辑效率提高了约 56%。

实施例 2

以斑马鱼 *mstnb* 基因第 1 个外显子中的 CRISPR/Cas9 靶位点 (图 9) 为例, 比较 Cas9 mRNA 和 Cas9-PEST mRNA 在斑马鱼胚胎中的基因编辑效率。步骤与实施例 1 相似。经计算, 注射 400 pg Cas9 mRNA+100 pg sgRNA 引起的定点突变效率为 13.4%, 而注射相同剂量的 Cas9-PEST mRNA+sgRNA 可将定点突变效率提升至 56.1% (图 10), 与对照组相比突变效率提高了约 4 倍。

在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。

参考文献

- Abudayyeh, O. O., J. S. Gootenberg, S. Konermann, J. Joung, I. M. Slaymaker, D. B. Cox, S. Shmakov, K. S. Makarova, E. Semenova, L. Minakhin, K. Severinov, A. Regev, E. S. Lander, E. V. Koonin and F. Zhang (2016). "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector." *Science* **353**(6299): aaf5573.
- Bibikova, M., M. Golic, K. G. Golic and D. Carroll (2002). "Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases." *Genetics* **161**(3): 1169-1175.
- Chang, N., C. Sun, L. Gao, D. Zhu, X. Xu, X. Zhu, J. W. Xiong and J. J. Xi (2013). "Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos." *Cell Res* **23**(4): 465-472.
- Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini and F. Zhang (2013). "Multiplex genome engineering using CRISPR/Cas systems." *Science* **339**(6121): 819-823.
- Fleming, J. V. and T. C. Wang (2000). "Amino- and carboxy-terminal PEST domains mediate gastrin stabilization of rat L-histidine decarboxylase isoforms." *Mol Cell Biol* **20**(13): 4932-4947.
- Gao, F., X. Z. Shen, F. Jiang, Y. Wu and C. Han (2016). "DNA-guided genome editing using the *Natronobacterium gregoryi* Argonaute." *Nat Biotechnol* **34**(7): 768-773.
- Hou, Z., Y. Zhang, N. E. Propson, S. E. Howden, L. F. Chu, E. J. Sontheimer and J. A. Thomson (2013). "Efficient genome engineering in human pluripotent stem cells using Cas9 from *Neisseria meningitidis*." *Proc Natl Acad Sci U S A* **110**(39): 15644-15649.
- Kim, D., J. Kim, J. K. Hur, K. W. Been, S. H. Yoon and J. S. Kim (2016). "Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells." *Nat Biotechnol* **34**(8): 863-868.
- Kim, Y. G., J. Cha and S. Chandrasegaran (1996). "Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain." *Proc Natl Acad Sci U S A* **93**(3): 1156-1160.
- Kleinstiver, B. P., M. S. Prew, S. Q. Tsai, V. V. Topkar, N. T. Nguyen, Z. Zheng, A. P. Gonzales, Z. Li, R. T. Peterson, J. R. Yeh, M. J. Aryee and J. K. Joung (2015). "Engineered CRISPR-Cas9 nucleases with altered PAM specificities." *Nature* **523**(7561): 481-485.
- Kleinstiver, B. P., S. Q. Tsai, M. S. Prew, N. T. Nguyen, M. M. Welch, J. M. Lopez, Z. R. McCaw, M. J. Aryee and J. K. Joung (2016). "Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells." *Nat Biotechnol* **34**(8): 869-874.
- Liu, D., Z. Wang, A. Xiao, Y. Zhang, W. Li, Y. Zu, S. Yao, S. Lin and B. Zhang (2014). "Efficient gene targeting in zebrafish mediated by a zebrafish-codon-optimized cas9 and evaluation of off-targeting effect." *J Genet Genomics* **41**(1): 43-46.
- Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville and G. M. Church (2013). "RNA-guided human genome engineering via Cas9." *Science* **339**(6121): 823-826.
- Miller, J. C., S. Tan, G. Qiao, K. A. Barlow, J. Wang, D. F. Xia, X. Meng, D. E. Paschon, E. Leung, S. J. Hinkley, et al. (2011). "A TALE nuclease architecture for efficient genome editing." *Nat Biotechnol* **29**(2): 143-148.
- Ran, F. A., L. Cong, W. X. Yan, D. A. Scott, J. S. Gootenberg, A. J. Kriz, B. Zetsche, O. Shalem, X. Wu, K. S. Makarova, E. V. Koonin, P. A. Sharp and F. Zhang (2015). "In vivo genome editing using *Staphylococcus aureus* Cas9." *Nature* **520**(7546): 186-191.
- Rechsteiner, M. and S. W. Rogers (1996). "PEST sequences and regulation by proteolysis." *Trends*

Biochem Sci **21**(7): 267-271.

Reverte, C. G., M. D. Ahearn and L. E. Hake (2001). "CPEB degradation during *Xenopus* oocyte maturation requires a PEST domain and the 26S proteasome." Dev Biol **231**(2): 447-458.

Rogers, S., R. Wells and M. Rechsteiner (1986). "Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis." Science **234**(4774): 364-368.

Spencer, M. L., M. Theodosiou and D. J. Noonan (2004). "NPDC-1, a novel regulator of neuronal proliferation, is degraded by the ubiquitin/proteasome system through a PEST degradation motif." J Biol Chem **279**(35): 37069-37078.

Tsai, S. Q., N. Wyvekens, C. Khayter, J. A. Foden, V. Thapar, D. Reyon, M. J. Goodwin, M. J. Aryee and J. K. Joung (2014). "Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing." Nat Biotechnol **32**(6): 569-576.

Zetsche, B., J. S. Gootenberg, O. O. Abudayyeh, I. M. Slaymaker, K. S. Makarova, P. Essletzbichler, S. E. Volz, J. Joung, J. van der Oost, A. Regev, E. V. Koonin and F. Zhang (2015). "Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system." Cell **163**(3): 759-771.

权 利 要 求 书

1. 一种融合蛋白，其特征在于，所述的融合蛋白具有式 Ia 或 Ib 所述结构：



其中，

E 为核酸内切酶蛋白元件；

P 为 PEST 蛋白元件；

“-”表示连接上述各元件的肽键或肽接头。

2. 如权利要求 1 所述的融合蛋白，其特征在于，所述的核酸内切酶蛋白元件选自下组：Cas9 蛋白（包括 SpCas9、SaCas9、NmCas9、St1Cas9）及其变体（如 SpCas9 的 VQR、EQR、VRER 等变体）、Cpf1 蛋白（包括 AsCpf1、FnCpf1、LbCpf1）、C2c2 蛋白、Argonaute 蛋白家族；TALE 蛋白、锌指蛋白、dCas9 等与 FokI 融合构成的人工核酸内切酶。

3. 如权利要求 1 所述的融合蛋白，其特征在于，所述 PEST 蛋白元件选自下组：

ODC1 蛋白的 PEST 序列、GCN4 蛋白的 PEST 序列、CLN2/CLN3 蛋白的 PEST 序列、NIMA 蛋白的 PEST 序列、Cactus 蛋白的 PEST 序列、HDC 蛋白的 PEST 序列、CPEB 蛋白的 PEST 序列、NPDC1 蛋白的 PEST 序列、FOS 蛋白的 PEST 序列、NFKBIA 蛋白的 PEST 序列等。

4. 一种分离的多核苷酸，所述的多核苷酸编码权利要求 1 所述的融合蛋白。

5. 一种载体，所述载体含有权利要求 4 所述的多核苷酸。

6. 一种基因工程细胞，其特征在于，所述基因工程细胞含有权利要求 5 所述的载体，或含有权利要求 4 所述的多核苷酸，或含有权利要求 1 所述的融合蛋白。

7. 如权利要求 6 所述的基因工程细胞，其特征在于，所述基因工程细胞为动物细胞、植物细胞、或微生物细胞。

8. 一种基因编辑系统，其特征在于，所述基因编辑系统包括权利要求 1 所述的融合蛋白、或权利要求 4 所述的多核苷酸、或权利要求 5 的载体。

9. 一种核酸内切酶介导的基因编辑方法，其特征在于，所述方法包括步骤：使用权利要求 1 所述的融合蛋白对靶基因进行编辑。

10. 一种提高核酸内切酶介导的基因编辑系统的基因编辑效率的方法，其特征在于，所述方法包括步骤：

在靶细胞中，表达 PEST 蛋白和核酸内切酶的融合蛋白。

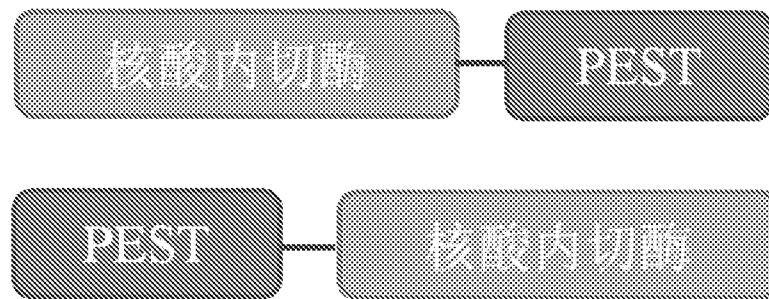


图 1

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET
 AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHER
 HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIEG
 DLNPDNSDVKLFIQLVQTYNQLFEEPNINASGVDAKAILSARLSKSRLENLIAQL
 PGEKKNGLFGNLIASLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQY
 ADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
 PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
 KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLAR
 GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL
 YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKK
 IECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFED
 REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
 DGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIIKKGILQTVK
 VVDELVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH
 PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKV
 LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
 AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
 QFYKVVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
 EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
 KVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
 SVLVVAKVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEVKKDILIKL
 PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEQK
 QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH RD KPIREQAENI IHL
 FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD
 (SEQ ID NO.1)

图 2a

GATAAAAAGTACAGCATTGGACTGGACATCGGAACAAATAGCGTGGGCTGGGCTGTGATTACTGACGAATATAAGGTGCCTAGCAAA
AAGTTTAAAGTGCTGGGAAACACCGACAGACACAGCATCAAAAAAACCTGATCGGCGCTCTGCTGTTTGATAGCGGTGAAACTGCC
GAGGCTACTAGACTGAAGAGAACTGCTAGAAGAAGATATACCAGAAGAAAGAATAGAATTTGTTACCTGCAAGAAATCTTTAGCAAT
GAGATGGCAAAGGTTGACGATAGCTTCTTTATAGACTGGAGGAGAGCTTCTGCTGGTCGAGGAGGACAAGAAGCACGAGAGACACCCC
ATCTTCGGAATATCGTGGACGAGGTGGCATACCATGAAAAGTATCCTACCATTACCACCTGAGAAAAAGCTGGTGGACAGCACA
GACAAGGCCGATCTGAGACTGATCTACCTGGCACTGGCCACATGATCAAATTTAGAGGCCATTCTCTGATTGAAGGAGACCTGAAC
CCCGATAACAGCGATGTTGATAAACTGTTTATCCAACCTGGTTTACAGCCTATAACCAACTGTTTGAGGAGAACCCCTATTAACGCCAGC
GGAGTGGATGCAAAGGCCATCCTGAGCGCTAGACTGAGCAAAAGCAGAAGACTGGAAAATCTGATCGCCACGCTGCCCGCGCAAAAA
AAGAATGGACTGTTTCGCAATCTGATTGCACTGAGCCTGGGACTGACACCTAACTTCAAGAGCAATTTGATCTGGCTGAGGACGCC
AACTGCACTGAGCAAAGACACATATGATGACGACCTGGATAACCTGCTGGCACAAATTTGGTGACCAATACGCTGACCTGTTCTCTG
GCTGCTAAGAATCTGAGCGATGCCATTCTGCTGAGCGACATCCTGAGAGTGAACACAGAGATTACCAAGGCACCCCTGAGCGCAAGC
ATGATTAAGAGATACGACGAGCACCACCAAGATCTGACCCTGCTGAAGGCCCTGGTCAGACAACAACCTGCCAGAGAAGTATAAAGAA
ATTTTCTTTGACCAAAGCAAGAACGGTTACGCTGGCTACATTGACGGCGGTGCAAGCCAAGAGGAGTTCTATAAGTTCATTAAGCCA
ATCCTGGAGAAAATGGATGGAAGTGGAGAGCTGCTGGTTAAGCTGAATAGAGAGGATCTGCTGAGAAAAACAAAGAACATTCGACAAC
GGTAGCATCCACACCAGATTCTCTGGGTGAGCTGCACGCAATTCTGAGAAGACAGGAAGACTTTTATCCATTCTCTGAAGGACAAC
AGAGAAAAGATCGAGAAGATTCTGACATTTAGAATCCCTACTACGTGGGACCTCTGGCTAGAGGCAATAGCAGATTGCGATGGATG
ACTAGAAAAGAGCGAGGAGACAATTACCCCTTGGAACCTTTGAAGAAGTGGTGGATAAGGGAGCAAGCGCCAAAGCTTCATTGAGAGA
ATGACAACTTCGATAAGAACCTGCCTAACGAGAAGGTTCTGCCAAGCATAGCCTGCTGTATGAATATTTACAGTGTACAACGAG
CTGACAAAGGTCAAGTACGTACAGAGGGCATGAGAAAGCCCGCCTTTCTGAGCGGAGAACAAAAGAAGGCTATTGTTGACCTGCTG
TTCAAGACCAACAGAAAAGTTACAGTTAAACAGCTGAAAGAGGACTACTTCAAAAAGATTGAATGTTTTGACAGCGTGGAAATCAGC
GGCGTTGAGGACAGATTTAACGCTAGCCTGGGCACCTACCACGATCTGCTGAAAATCATCAAAGATAAGGACTTTCTGGACAACGAA
GAAAACGAGGACATTCTGGAAGACATTGTGCTGACACTGACTCTGTTTCAAGATAGAGAAATGATCGAGGAAAGACTGAAAACCTTAT
GCACATCTGTTTACGACGACAAAAGTATGAAGCAACTGAAGAGAAGAAGATACACTGGATGGGGCAGACTGAGCAGAAAAGCTGATCAAC
GGAATCAGAGACAAGCAAAGCGGAAAAACTATTCTGGATTTTCTGAAAAGCGACGGTTTCGCCAATAGAAACTTCATGCAACTGATT
CACGATGACAGCCTGACTTTCAAGGAGGATATTCAAAGGCACAGGTGAGCGGCCAGGGCGATAGCCTGCACGAACACATCGCAAT
CTGGCCGTTAGCCCTGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTTGTTGATGAAGTGGTCAAGGTGATGGGTAGACACAAG
CCCAGAAATATTGTATCGAGATGGCTAGAGAGAACCAACAACACAAAAGGGACAGAAGAATAGCAGAGAAAAGAAATGAAAAGAATT
GAGGAGGGAATCAAGGAGCTGGGTAGCCAGATCCTGAAAAGAACCCCTGTGAGAATACACAACCTGCAAAACGAAAAGCTGTACCTG
TACTACCTGCAAAATGGCAGAGACATGTACGTGGACCAAGAGCTGGATATTAACAGACTGAGCGACTACGATGTCGACCACATCGTG
CCTCAAAGCTTCTGAAGGATGACAGCATCGACAATAAAGTGTGCTGACTAGAAGCGACAAGAACAGAGGAAAAAGCGACAACGTGCC
AGCGAGGAAGTGGTTAAAAAGATGAAGAACTACTGGAGACAGCTGCTGAATGCCAAGCTGATCACACAAAGAAAATTCGACAACCTG
ACCAAAGCCGAGAGAGGAGGTCTGAGCGAAGTGGACAAGGCTGGATTCTTAAGAGACAACCTGGTTGAAACCAGACAGATTACAAAG
CACGTGGCTCAAATCCTGGACAGCAGAATGAATACCAATATGACGAGAACGACAACTGATTAGAGAGGTGAAGGTTATTACTCTG
AAGAGCAAAGTGGTCAGCGACTTCAGAAAGGACTTCCAATTCTACAAGGTGAGAGAGATCAACAATTACCACCACGCACACGACGCT
TACCTGAACGCTGTGGTGGGCACAGCTCTGATCAAAAAGTATCCAAAACCTGGAAGCGAGTTTGTGTACGGTGACTATAAAGTTTAT
GATGTGAGAAAAATGATCGCTAAGAGCGAGCAGGAGATCGGAAAGGCTACAGCCAAGTATTTCTTTTACAGCAACATTATGAACTTT
TTCAAGACTGAAATCACCTGGCAAACGGTGAGATCAGAAAAAGACCCTGATCGAAACAAATGGCGAGACAGGCGAGATCGTGTGG
GATAAGGGAAGAGACTTCGCTACCGTTAGAAAGGTTCTGAGCATGCCACAGGTTAACATTGTGAAGAAAACCTGAGGTGCAGACAGGA
GGTTTTCAGCAAGGAGAGCATCCTGCCTAAGAGAAACAGCGATAAGCTGATTGCAAGAAAAAAGGATTGGGACCCTAAGAAGTACGGC
GGTTTTGACAGCCCTACTGTGGCTTACAGCGTGTGGTGGTGGCTAAAGTGGAGAAGGGCAAAAGCAAGAAGCTGAAAAGCGTGAAG
GAACTGCTGGGAATTACAATCATGGAGAGAAGCAGCTTCGAGAAGAACCCTGACTTCTGGAGGCTAAGGGATACAAGGAAGTT
AAGAAGGACCTGATCATCAAGCTGCCCCAAGTACAGCCTGTTTCGAGCTGGAAAATGGTAGAAAGAGAATGTGGCTAGCGCTGGTGAG
CTGCAGAAGGGAAATGAACTGGCACTGCCTAGCAAGTACGTTAACTTTCTGTATCTGGCAAGCCATTACGAGAACTGAAAGGAAGC
CCCAGGACAATGAGCAGAAACAAGTTCGTGGAACAGCACAAACACTATCTGGACGAGATTATCGAGCAGATCAGCGAATTTAGC
AAAAGAGTGATCCTGGCTGATGCTAACCTGGATAAAGTCTGAGCGCTTACAACAAACATAGAGATAAGCCTATCAGAGAGCAGGCC
GAAAACATCATCCACCTGTTTCACTGACAAACCTGGGCGCTCTGCCGCTTTCAAGTACTTTGATACCACTATTGATAGAAAGAGA
TATACTAGCACCAAAGAGGTGCTGGACGCCACCCTGATTCACCAGAGCATTACCGGACTGTACGAAAAGTAGAATCGACCTGAGCCAA
CTGGGAGGAGAC (SEQ ID NO. 4)

图 2b

SHGFPPEVEEQDDGTLPMSCAQESGMDRHPAACASARINV

(SEQ ID NO.2)

AGCCATGGCTTCCCGCCGGAGGTGGAGGAGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGG
AGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGGATCAATGTG (SEQ ID NO.5)

图 3

DKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGET
AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHER
HPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHAHMIKFRGHFLIEG
DLNPDNSDVKLFIQLVQTYNQLFEEPNINASGVDAKAILSARLSKSRLENLIAQL
PGEKKNGLFGNLIALLSLGLTPNFKSNFDLAEDAKLQLSKDITYDDDLNLLAQIGDQY
ADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL
PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYVGPLAR
GNSRFAWMTRKSEETITPWNFEVVVDKGASQSFIERMTNFDKNLPNEKVLPHKSL
YEFYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKK
IECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFED
REMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS
DGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDLSLHEHIANLAGSPAIIKKGILQTVK
VVDELVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH
PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKV
LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDF
QFYKVREINNYHHAHDAYLNAVVGITALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ
EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR
KVLSPQVNIIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
SVLVVAKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDEQK
QLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RD KPIREQAENIIHL
FTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGDS
HGFPPEVEEQDDGTLPMSCAQESGMDRHPAACASARINV (SEQ ID NO.3)

图 4a

GATAAAAAGTACAGCATTGGACTGGACATCGGAACAAATAGCGTGGGCTGGGCTGTGATTACTGACGAATATAAGGTGCCTAGCAAA
AAGTTTAAAGTGCTGGGAAACACCGACAGACACAGCATCAAAAAAACCTGATCGGCGCTCTGCTGTTTGATAGCGGTGAAACTGCC
GAGGCTACTAGACTGAAGAGAACTGCTAGAAGAAGATATACCAGAAGAAAGAATAGAATTTGTTACCTGCAAGAAATCTTTAGCAAT
GAGATGGCAAAGGTTGACGATAGCTTCTTTATAGACTGGAGGAGAGCTTCTGCTGGTCGAGGAGGACAAGAAGCACGAGAGACACCCC
ATCTTCGGAATATCGTGGACGAGGTGGCATACCATGAAAAGTATCCTACCATTACCACCTGAGAAAAAGCTGGTGGACAGCACA
GACAAGGCCGATCTGAGACTGATCTACCTGGCACTGGCCACATGATCAAATTTAGAGGCCATTCTCTGATTGAAGGAGACCTGAAC
CCCGATAACAGCGATGTTGATAAACTGTTTCATCCAACCTGGTTCAGACCTATAACCAACTGTTTGAGGAGAACCTATTAACGCCAGC
GGAGTGGATGCAAAGGCCATCCTGAGCGCTAGACTGAGCAAAAGCAGAAGACTGGAAAATCTGATCGCCACGCTGCCCGCGAAAAA
AAGAATGGACTGTTTCGCAATCTGATTGCACTGAGCCTGGGACTGACACCTAACTTCAAGAGCAATTTGATCTGGCTGAGGACGCC
AACTGCACTGAGCAAAGACACATATGATGACGACCTGGATAACCTGCTGGCACAAATTGGTGACCAATACGCTGACCTGTTCTCTG
GCTGCTAAGAATCTGAGCGATGCCATTCTGCTGAGCGACATCCTGAGAGTGAACACAGAGATTACCAAGGCACCCCTGAGCGCAAGC
ATGATTAAGAGATACGACGAGCACCACCAAGATCTGACCCTGCTGAAGGCCCTGGTCAGACAACAACCTGCCAGAGAAGTATAAGAA
ATTTTCTTTGACCAAAGCAAGAACGGTTACGCTGGCTACATTGACGGCGGTGCAAGCCAAGAGGAGTTCTATAAGTTCATTAAGCCA
ATCCTGGAGAAAATGGATGGAAGTGGAGAGCTGCTGGTTAAGCTGAATAGAGAGGATCTGCTGAGAAAAACAAAGAACATTCGACAAC
GGTAGCATCCACACCAGATTCTCTGGGTGAGCTGCACGCAATTCTGAGAAGACAGGAAGACTTTTATCCATTCTCTGAAGGACAAC
AGAGAAAAGATCGAGAAGATTCTGACATTTAGAATCCCTACTACCTGGGACCTCTGGCTAGAGGCAATAGCAGATTGCGATGGATG
ACTAGAAAAGAGCGAGGAGACAATTACCCCTTGAACTTTGAAGAAGTGGTGGATAAGGGAGCAAGCGCCAAAGCTTCATTGAGAGA
ATGACAACTTCGATAAGAACCTGCCTAACGAGAAGGTTCTGCCAAGCATAGCCTGCTGTATGAATATTTACAGTGTACAACGAG
CTGACAAAGGTCAAGTACGTCACAGAGGGCATGAGAAAGCCCGCTTTCTGAGCGGAGAACAAAAGAGGCTATTGTTGACCTGCTG
TTCAAGACCAACAGAAAAGTTACAGTTAAACAGCTGAAAGAGGACTACTTCAAAAAGATTGAATGTTTTGACAGCGTGGAAATCAGC
GGCGTTGAGGACAGATTTAACGCTAGCCTGGGCACCTACCACGATCTGCTGAAAATCATCAAAGATAAGGACTTTCTGGACAACGAA
GAAAACGAGGACATTCTGGAAGACATTGTGCTGACACTGACTCTGTTTCAAGATAGAGAAATGATCGAGGAAAGACTGAAAACCTAT
GCACATCTGTTTCGACGACAAAAGTATGAAGCAACTGAAGAGAAGAAGATACACTGGATGGGGCAGACTGAGCAGAAAAGCTGATCAAC
GGAATCAGAGACAAGCAAAGCGGAAAAACTATTCTGGATTTTCTGAAAAGCGACGGTTTCGCCAATAGAACTTCATGCAACTGATT
CACGATGACAGCCTGACTTTCAAGGAGGATATTCAAAAGGCACAGGTGAGCGGCCAGGGCGATAGCCTGCACGAACACATCGCAAT
CTGGCCGTTAGCCCTGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTTGTTGATGAAGTGGTCAAGGTGATGGGTAGACACAAG
CCCAGAAATATTGTATCGAGATGGCTAGAGAGAACCAACAACACAAAAGGGACAGAAGAATAGCAGAGAAAAGAAATGAAAAGAATT
GAGGAGGGAATCAAGGAGCTGGGTAGCCAGATCCTGAAAAGAACCCCTGTGAGAATACACAACCTGCAAAACGAAAAGCTGTACCTG
TACTACCTGCAAAATGGCAGAGACATGTACGTGGACCAAGAGCTGGATATTAACAGACTGAGCGACTACGATGTCGACCACATCGTG
CCTCAAAGCTTCCTGAAGGATGACAGCATCGACAATAAAGTGTGACTAGAAGCGACAAGAACAGAGGAAAAAGCGACAACGTGCC
AGCGAGGAAGTGGTTAAAAAGATGAAGAACTACTGGAGACAGCTGCTGAATGCCAAGCTGATCACACAAAGAAAATTCGACAACCTG
ACCAAAGCCGAGAGAGGAGGTCTGAGCGAAGTGGACAAGGCTGGATTCTTAAGAGACAACCTGGTTGAAACCAGACAGATTACAAAG
CACGTGGCTCAAATCCTGGACAGCAGAATGAATACCAATATGACGAGAACGACAAACTGATTAGAGAGGTGAAGGTTATTACTCTG
AAGAGCAAAGTGGTCAGCGACTTCAGAAAGGACTTCCAATTCTACAAGGTGAGAGAGATCAACAATTACCACCACGCACACGACGCT
TACCTGAACGCTGTGGTGGGCACAGCTCTGATCAAAAAGTATCCAAAAGTGGAAAAGCGAGTTTGTGTACGGTGACTATAAAGTTTAT
GATGTGAGAAAAATGATCGCTAAGAGCGAGCAGGAGATCGGAAAGGCTACAGCCAAGTATTTCTTTTACAGCAACATTATGAACTTT
TTCAAGACTGAAATCACCTGGCAAACGGTGAGATCAGAAAAAGACCCTGATCGAAACAAATGGCGAGACAGGCGAGATCGTGTGG
GATAAGGGAAGAGACTTCGCTACCGTTAGAAAGGTTCTGAGCATGCCACAGGTTAACATTGTGAAGAAAAGTGGGTGCAGACAGGA
GGTTTCAGCAAGGAGAGCATCCTGCCTAAGAGAAACAGCGATAAGCTGATTGCAAGAAAAAGGATTGGGACCCTAAGAAGTACGGC
GGTTTTGACAGCCCTACTGTGGCTTACAGCGTGTGGTGGTGGCTAAAGTGGAGAAGGGCAAAAGCAAGAAGCTGAAAAGCGTGAAG
GAACTGCTGGGAATTACAATCATGGAGAGAAGCAGCTTCGAGAAGAACCCTGACTTCTGGAGGCTAAGGGATACAAGGAAGTT
AAGAAGGACCTGATCATCAAGCTGCCCCAAGTACAGCCTGTTTCGAGCTGGAAAATGGTAGAAAGAGAATGCTGGCTAGCGCTGGTGAG
CTGCAGAAGGGAAATGAACTGGCACTGCCTAGCAAGTACGTTAACTTTCTGTATCTGGCAAGCCATTACGAGAACTGAAAGGAAGC
CCCAGGACAATGAGCAGAAACAAGTTCGTGGAACAGCACAAACACTATCTGGACGAGATTATCGAGCAGATCAGCGAATTTAGC
AAAAGAGTGATCCTGGCTGATGCTAACCTGGATAAAGTCTGAGCGCTTACAACAAACATAGAGATAAGCCTATCAGAGAGCAGGCC
GAAAACATCATCCACCTGTTTCACTGACAAACCTGGGCGCTCTGCCGCTTTCAAGTACTTTGATACCACTATTGATAGAAAGAGA
TATACTAGCACCAAAGAGGTGCTGGACGCCACCCTGATTACCAGAGCATTACCGGACTGTACGAAAAGTAAATCGACCTGAGCCAA
CTGGGAGGAGACAGCCATGGCTTCCCGCCGGAGGTGGAGGAGCAGGATGATGGCACGCTGCCCATGTCTTGTGCCCAGGAGAGCGGG
ATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGGATCAATGTG (SEQ ID NO. 6)

图 4b

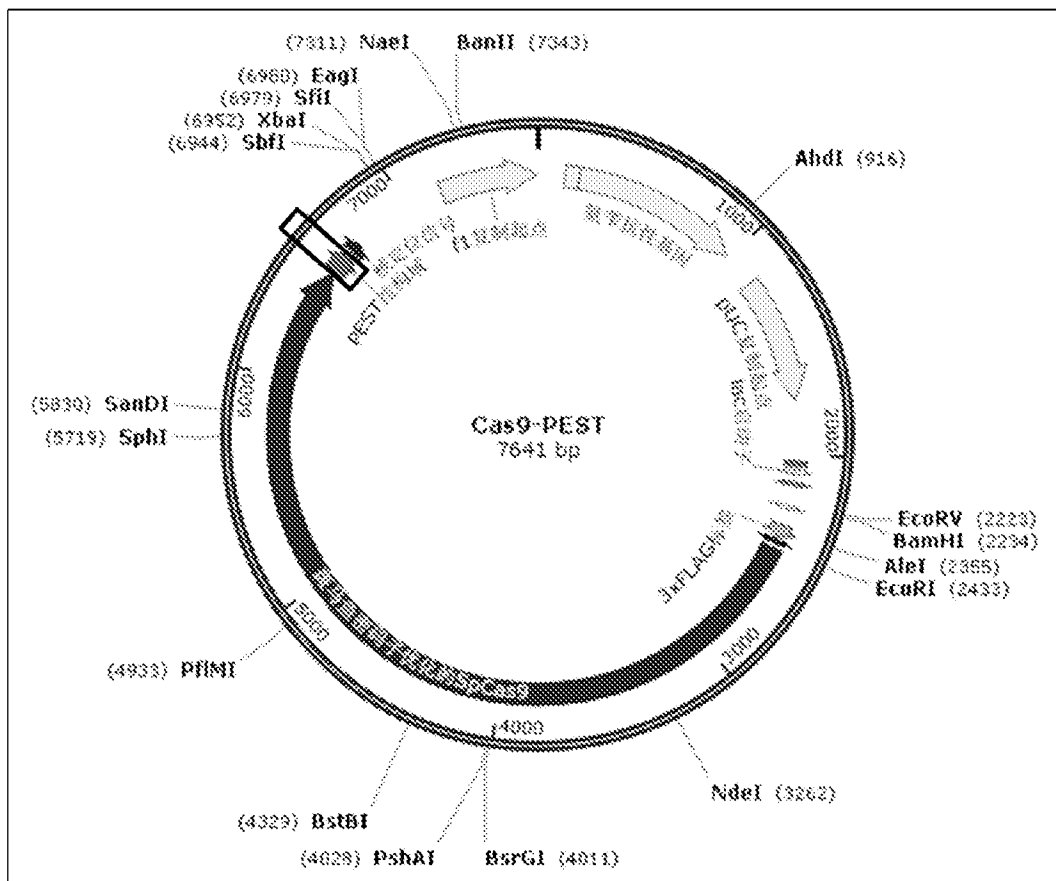


图 5

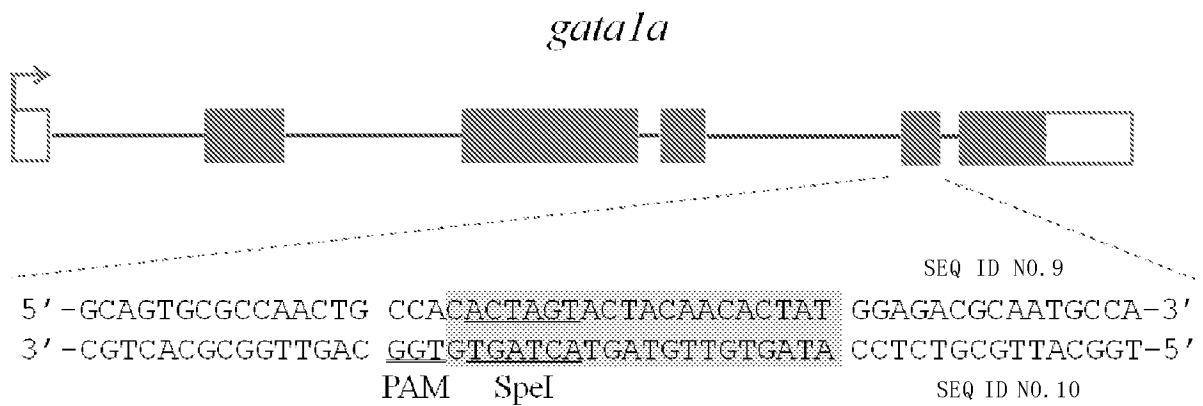


图 6

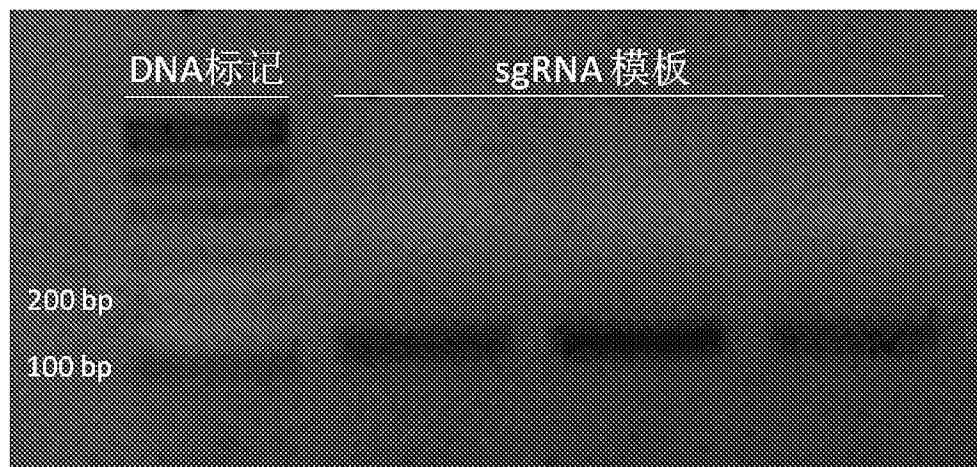


图 7

gata1a 基因靶位点效率检测

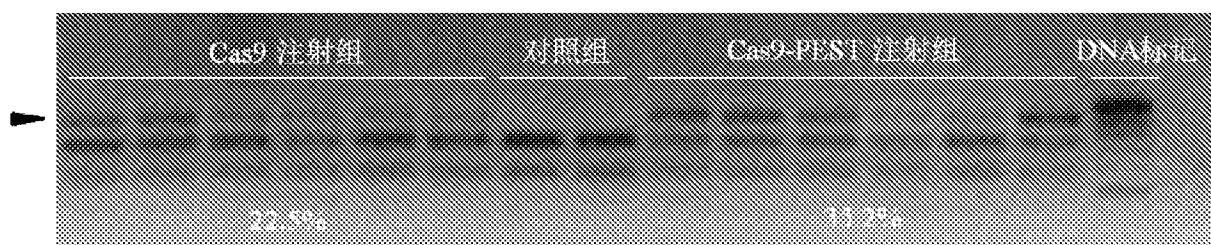


图 8

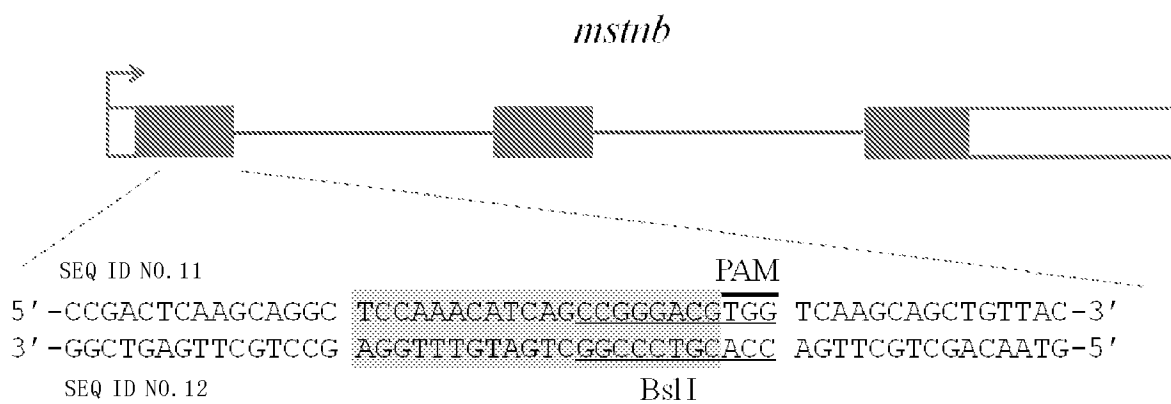


图 9

mstnb 基因靶位点效率检测

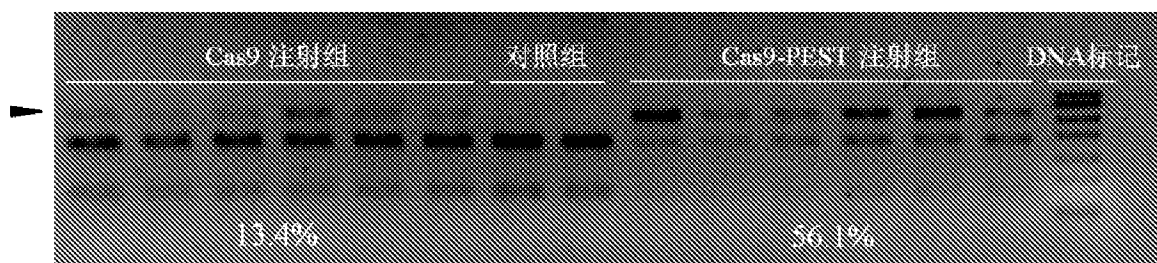


图 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/104450

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing filed or furnished:
- a. (means)
- ☐ on paper
- ☒ in electronic form
- b. (time)
- ☒ in the international application as filed
- ☐ together with the international application in electronic form
- ☐ subsequently to this Authority for the purposes of search
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/104450

A. CLASSIFICATION OF SUBJECT MATTER

C07K 19/00 (2006.01) i; C12N 15/52 (2006.01) i; C12N 15/63 (2006.01) i; C12N 5/10 (2006.01) i
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

C07K; C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CPRSABS; CNABS; CNKI; CNTXT; VEN; MEDLINE; EPTXT; USTXT; WOTXT and keywords: 成簇有规间隔短回文重复序列, 基因, 编辑, PEST, gene, edit+, CRISPR, Cas9, etc.;

NATIONAL BIO-SEQUENCE DATABASE OF CHINESE PATENT; Genbank; EMBL; STN and searched sequences: SEQ ID NO: 1-6

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104531632 A (LI, Yunying), 22 April 2015 (22.04.2015), see the abstract, embodiments 1-4, description, paragraphs [0003]-[0012], and claims 1-6	1-10
A	WO 2014197748 A2 (UNIV DUKE), 11 December 2014 (11.12.2014), see entire document	1-10
A	WO 2014039702 A2 (DOW AGROSCIENCES LLC et al.), 13 March 2014 (13.03.2014), see entire document	1-10
A	CN 104805099 A (SECOND MILITARY MEDICAL UNIVERSITY), 29 July 2015 (29.07.2015), see entire document	1-10
A	WO 2014039684 A1 (DOW AGROSCIENCES LLC et al.), 13 March 2014 (13.03.2014), see entire document	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E" earlier application or patent but published on or after the international filing date	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&" document member of the same patent family
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 24 July 2017	Date of mailing of the international search report 07 August 2017
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer WANG, Jinfeng Telephone No. (86-10) 62412282

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/CN2016/104450

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104531632 A	22 April 2015	None	
WO 2014197748 A2	11 December 2014	CA 2914519 A1	11 December 2014
		JP 2016521555 A	25 July 2016
		EP 3004370 A4	11 January 2017
		WO 2014197748 A3	19 February 2015
		AU 2014274840 A2	21 January 2016
		CN 105658805 A	08 June 2016
		AU 2014274840 A1	21 January 2016
		US 2016201089 A1	14 July 2016
		EP 3004370 A2	13 April 2016
		KR 20160023765 A	03 March 2016
WO 2014039702 A2	13 March 2014	CN 105025702 A	04 November 2015
		HK 1213736 A1	15 July 2016
		IL 237536 A	30 April 2015
		EP 2893023 A4	21 September 2016
		US 2017044559 A1	16 February 2017
		RU 2015112584 A	27 October 2016
		UY 35020 A	31 March 2014
		US 2014090116 A1	27 March 2014
		MX 2015002990 A	28 September 2015
		HK 1212153 A1	10 June 2016
		IL 237537 D0	30 April 2015
		AU 2013312465 A1	12 March 2015
		IL 237536 D0	30 April 2015
		CA 2884084 A1	13 March 2014
		AR 092480 A1	22 April 2015
		AU 2013312465 A2	12 March 2015
		JP 2015527082 A	17 September 2015
		EP 2892321 A2	15 July 2015
		CA 2883846 A1	13 March 2014
		TW 201414838 A	16 April 2014
		AR 092481 A1	22 April 2015
		IL 237537 A	30 April 2015
		WO 2014039702 A3	24 July 2014
		KR 20150043539 A	22 April 2015
		RU 2015112583 A	27 October 2016
		CN 104780756 A	15 July 2015
		US 9493779 B2	15 November 2016
		EP 2892321 A4	02 March 2016
		WO 2014039692 A2	13 March 2014
		CL 2015000563 A1	02 October 2015
		WO 2014039692 A3	16 July 2015
		AU 2013312455 A1	12 March 2015
		UY 35021 A	31 March 2014
		CL 2015000561 A1	02 October 2015
		US 2014090112 A1	27 March 2014
		JP 2015529464 A	08 October 2015
		KR 20150043538 A	22 April 2015
		EP 2893023 A2	15 July 2015
CN 104805099 A	29 July 2015	None	

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/104450

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2014039684 A1	13 March 2014	HK 1217732 A1	20 January 2017
		EP 2893006 A4	13 April 2016
		IL 237567 A	30 April 2015
		UY 35019 A	31 March 2014
		RU 2015112578 A	27 October 2016
		KR 20150043540 A	22 April 2015
		AR 092478 A1	22 April 2015
		IL 237567 D0	30 April 2015
		CL 2015000564 A1	02 October 2015
		CA 2884162 A1	13 March 2014
		CN 105264067 A	20 January 2016
		JP 2015527081 A	17 September 2015
		AU 2013312538 A1	12 March 2015
		EP 2893006 A1	15 July 2015
		US 2015067921 A1	05 March 2015

国际检索报告

国际申请号

PCT/CN2016/104450

A. 主题的分类

C07K 19/00(2006.01)i; C12N 15/52(2006.01)i; C12N 15/63(2006.01)i; C12N 5/10(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

C07K; C12N

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CPRSABS; CNABS; CNKI; CNTXT; VEN; MEDLINE; EPTXT; USTXT; WOTXT和关键词: 成簇有规间隔短回文重复序列, 基因, 编辑, PEST, gene, edit+, CRISPR, Cas9, 等 中国专利生物序列检索系统; Genbank; EMBL; STN和检索的序列: SEQ ID NO:1-6

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 104531632 A (李云英) 2015年 4月 22日 (2015 - 04 - 22) 参见摘要、实施例1-4、说明书第[0003]-[0012]段、权利要求1-6	1-10
A	WO 2014197748 A2 (UNIV DUKE) 2014年 12月 11日 (2014 - 12 - 11) 参见全文	1-10
A	WO 2014039702 A2 (DOW AGROSCIENCES LLC等) 2014年 3月 13日 (2014 - 03 - 13) 参见全文	1-10
A	CN 104805099 A (中国人民解放军第二军医大学) 2015年 7月 29日 (2015 - 07 - 29) 参见全文	1-10
A	WO 2014039684 A1 (DOW AGROSCIENCES LLC等) 2014年 3月 13日 (2014 - 03 - 13) 参见全文	1-10

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2017年 7月 24日

国际检索报告邮寄日期

2017年 8月 7日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

王金凤

电话号码 (86-10)62412282

第I栏

核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何对要求保护的发明必要的核苷酸和/或氨基酸序列，国际检索是在下列基础上进行的：
- a. (提交提供)
- ☐ 纸件形式
- ☒ 电子形式
- b. (提交时间)
- ☒ 含在申请提交时的国际申请中
- ☐ 以电子形式与国际申请一起提交
- ☐ 为检索之用随后提交本单位
2. ☐ 另外，在提交/提供了多个版本或副本的序列表的情况下，提供了关于随后提交的或附加的副本中的信息与申请时提交的申请中的信息相同或未超出申请时提交的申请中的信息范围（如适用）的所需声明。
3. 补充意见：

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/104450

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	104531632	A	2015年 4月 22日	无			
WO	2014197748	A2	2014年 12月 11日	CA	2914519	A1	2014年 12月 11日
				JP	2016521555	A	2016年 7月 25日
				EP	3004370	A4	2017年 1月 11日
				WO	2014197748	A3	2015年 2月 19日
				AU	2014274840	A2	2016年 1月 21日
				CN	105658805	A	2016年 6月 8日
				AU	2014274840	A1	2016年 1月 21日
				US	2016201089	A1	2016年 7月 14日
				EP	3004370	A2	2016年 4月 13日
				KR	20160023765	A	2016年 3月 3日
WO	2014039702	A2	2014年 3月 13日	CN	105025702	A	2015年 11月 4日
				HK	1213736	A1	2016年 7月 15日
				IL	237536	A	2015年 4月 30日
				EP	2893023	A4	2016年 9月 21日
				US	2017044559	A1	2017年 2月 16日
				RU	2015112584	A	2016年 10月 27日
				UY	35020	A	2014年 3月 31日
				US	2014090116	A1	2014年 3月 27日
				MX	2015002990	A	2015年 9月 28日
				HK	1212153	A1	2016年 6月 10日
				IL	237537	D0	2015年 4月 30日
				AU	2013312465	A1	2015年 3月 12日
				IL	237536	D0	2015年 4月 30日
				CA	2884084	A1	2014年 3月 13日
				AR	092480	A1	2015年 4月 22日
				AU	2013312465	A2	2015年 3月 12日
				JP	2015527082	A	2015年 9月 17日
				EP	2892321	A2	2015年 7月 15日
				CA	2883846	A1	2014年 3月 13日
				TW	201414838	A	2014年 4月 16日
				AR	092481	A1	2015年 4月 22日
				IL	237537	A	2015年 4月 30日
				WO	2014039702	A3	2014年 7月 24日
				KR	20150043539	A	2015年 4月 22日
				RU	2015112583	A	2016年 10月 27日
				CN	104780756	A	2015年 7月 15日
				US	9493779	B2	2016年 11月 15日
				EP	2892321	A4	2016年 3月 2日
				WO	2014039692	A2	2014年 3月 13日
				CL	2015000563	A1	2015年 10月 2日
				WO	2014039692	A3	2015年 7月 16日
				AU	2013312455	A1	2015年 3月 12日
				UY	35021	A	2014年 3月 31日
				CL	2015000561	A1	2015年 10月 2日
				US	2014090112	A1	2014年 3月 27日
				JP	2015529464	A	2015年 10月 8日
				KR	20150043538	A	2015年 4月 22日
				EP	2893023	A2	2015年 7月 15日
CN	104805099	A	2015年 7月 29日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/104450

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
WO	2014039684	A1	2014年 3月 13日	HK	1217732	A1	2017年 1月 20日
				EP	2893006	A4	2016年 4月 13日
				IL	237567	A	2015年 4月 30日
				UY	35019	A	2014年 3月 31日
				RU	2015112578	A	2016年 10月 27日
				KR	20150043540	A	2015年 4月 22日
				AR	092478	A1	2015年 4月 22日
				IL	237567	D0	2015年 4月 30日
				CL	2015000564	A1	2015年 10月 2日
				CA	2884162	A1	2014年 3月 13日
				CN	105264067	A	2016年 1月 20日
				JP	2015527081	A	2015年 9月 17日
				AU	2013312538	A1	2015年 3月 12日
				EP	2893006	A1	2015年 7月 15日
				US	2015067921	A1	2015年 3月 5日

表 PCT/ISA/210 (同族专利附件) (2009年7月)