

# (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

(43) 国际公布日  
2018年1月18日 (18.01.2018)



(10) 国际公布号  
**WO 2018/010516 A1**

(51) 国际专利分类号:  
*C12N 15/00* (2006.01) *A61P 43/00* (2006.01)  
*A61P 35/00* (2006.01)

(21) 国际申请号: PCT/CN2017/088281

(22) 国际申请日: 2017年6月14日 (14.06.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:  
201610550293.X 2016年7月13日 (13.07.2016) CN

(72) 发明人: 及

(71) 申请人: 陈奇涵 (CHEN, Qihan) [CN/CN]; 中国江苏省南京玄武区苜蓿园东街28号7-202, Jiangsu 211100 (CN)。

(74) 代理人: 北京华睿卓成知识产权代理事务所 (普通合伙) (CHENG & PENG INTELLECTUAL PROPERTY LAW OFFICE); 中国北京市东城区东直门外大街42号宇飞大厦821, Beijing 100027 (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, JO, JP, KE, KG, KH, 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, 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)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(54) Title: METHOD FOR SPECIFICALLY EDITING GENOMIC DNA AND APPLICATION THEREOF

(54) 发明名称: 一种基因组DNA特异性编辑方法和应用

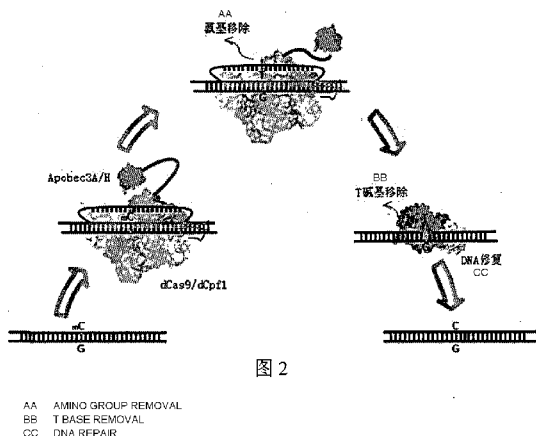


图2

(57) Abstract: A method for modulating a methylation/demethylation state of a nucleic acid, more specifically, a method for site-removing one or more methylated bases from a genome guided by a sgRNA sequence in a cell.

(57) 摘要: 一种调节核酸的甲基化/去甲基化状态的方法, 具体为: 一种可以在细胞内由sgRNA序列引导的在基因组上定点去除一个或多个甲基化碱基的方法。



WO 2018/010516 A1

## 一种基因组 DNA 特异性编辑方法和应用

### 技术领域

本发明属于生物工程技术领域，尤其涉及一种特异性调节基因组 DNA 的甲基化/去甲基化状态的方法及其应用。

### 背景技术

DNA 甲基化是表观遗传调控中的重要修饰之一，被称为哺乳动物 DNA 中除 ATCG 四种碱基以外的“第五种碱基”。作为一种共价修饰，DNA 甲基化在正常的分化发育和疾病发生中发挥重要作用，同时在更高等真核生物器官的细胞分化中可以稳定遗传，并且在斑马鱼中发现可以通过精子传给下一代。在细胞分化、疾病和环境影响下，DNA 甲基化状态会发生巨大变化。

研究表明，DNA 甲基化与肿瘤的发生发展有着密切的联系。DNA 甲基化状态的改变包括高甲基化（hypermethylation）和低甲基化（hypomethylation）两种。一般来说，基因启动子区的 DNA 高甲基化具有沉默基因表达的作用，而低甲基化则激活基因的表达。对不同肿瘤细胞的 DNA 分析表明，癌变细胞中出现基因突变的概率要远远低于预期。而在转录组范围内检测结肠直肠癌中由启动子高甲基化引起的基因表达抑制，发现高达 5%的已知基因在肿瘤细胞中发生了异常的启动子高甲基化。因此可以推测，与基因突变相比，DNA 甲基化改变在细胞恶变过程中可能发挥了更大的作用。

目标特异性的核酸编辑技术，尤其是基因组 DNA 的特异性编辑，一直是基因治疗的重要技术基础。而随着表观遗传学研究的深入，越来越多的研究表明了基因组的甲基化直接参与了转录调控和基因组的其他调控中，而处于表达活跃状态的基因其启动子和增强子区域的甲基化水平往往处于比较低的状态。因此，一种能够特异性去甲基化的核苷酸编辑技术对于沉默基因的转录激活有非常重要的意义。

目前，特定位点以及区域化的去甲基化都已有报道。比如生殖细胞的基因组重塑往往伴随着大规模的去甲基化。此外，某些酶(比如 Tet)能够将 5mC 氧化为 5hmC 并进一步通过 BER 或 NER 途径来达到去甲基化的目的。徐国良等人已在 2015 年报道并申请了通过 Tet 双加氧酶和胸腺嘧啶 DNA 糖基化酶等试剂去甲基化的方法的专利，但该方法没办法准确编辑某个位点，成为了应用于基因疗法或实验技术工具的重要瓶颈。

Apobec 蛋白家族的某些成员具有在单链的 DNA 中将 5mC 脱氨基从而转变为 T 的能力。借助这样的特性以及 CRISPR 蛋白家族精确定位的能力，开发一种能够精确编辑基因组中特定位点甲基化的系统成为了可能。

#### 发明内容

为了解决上述问题，本发明提供了一种精确地对核酸分子进行编辑的方法，所述方法包括如下步骤：

(1) 获得编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体，所述融合蛋白 (A) 包括位于氮端的 Apobec 家族蛋白结构域和位于碳端的核酸酶活性失活的 Cas9 家族或 Cpf1 蛋白结构域，单链导向 RNA 与目标核酸分子的目标编辑区域具有互补区，其中目标核酸分子的目标编辑区域中包含至少一个甲基化的胞嘧啶核苷酸；

(2) 使步骤 (1) 获得的编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体与目标核酸分子接触。

上述步骤中的重组载体，可以是两个载体分别编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体，也可以是一个重组载体同时编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体。

在一种优选情况下，融合蛋白中位于氮端的 Apobec 家族蛋白选自人 Apobec3A 或 Apobec3H，或与人 Apobec3A 或 Apobec3H 具有 95%以上同源性的具有脱氨基活性的蛋白。更优选的 Apobec 蛋白为 Apobec 3H 或 Apobec 3A。

在另一种优选情况下，融合蛋白中位于碳端的核酸酶活性失活的 Cas9 家族蛋白为在野生型 Cas9 蛋白中将第 10 和 840 位点的天冬氨酸和组氨酸突变为丙氨酸和丙氨酸，或者位于碳端的核酸酶活性失活的 Cpf1 家族但被为在野生型 Cpf1 蛋白中将第 908 位点的天冬氨酸突变为丙氨酸。

为使融合蛋白的两个蛋白结构域具有更好的空间结构上的灵活性，融合蛋白的两个结构域之间可增加由 3-14 个基序组成的连接体。所述基序选自 (GGG)。连接体越长，蛋白的空间灵活性越高，可编辑的目标区域越大。

为方便融合蛋白的表达纯化，还可包括纯化标签序列。通常使用的纯化标签为 6xHis。

在更优选的情况下，所述融合蛋白选自 SEQ ID No 201-207 的任一序列。

本发明还提供了编码上述融合蛋白序列的基因序列，优选情况下所属基因序列选自 SEQ ID No 301-307 所示。

本发明还提供了包含上述任一基因序列的重组载体，所述载体视后续试验目可以为原核表达载体或真核表达载体，包括但不限于质粒载体、病毒载体等。

本发明另一方面提供了单链导向 RNA 分子。在一种优选情况下，所述单链导向 RNA 长度为 60-80bp。在另一种优选情况下，所述单链导向 RNA 与目标核酸分子的互补区长度为 18-25bp，优选为 20bp。

一种体外编辑目标核酸分子的方法，包括如下步骤：

(1) 获得融合蛋白(A) 和单链导向 RNA (sgRNA) (B)，所述融合蛋白(A) 包括位于氮端的 Apobec 家族蛋白结构域和位于碳端的核酸酶活性失活的 Cas9 家族蛋白结构域，单链导向 RNA 与目标核酸分子的目标编辑区域具有互补区，其中目标核酸分子的目标编辑区域中包含至少一个甲基化的胞嘧啶核苷酸；

(2) 将目标核酸分子与融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 接触；

(3) 高温终止反应之后加入有效量的 TDG，42°C 反应 6-8 小时；

(4) 加入有效量的 EDTA、甲酰胺和 NaOH 共同作用，90-95°C 反应 5-10 分钟。

本发明还提供了所述编辑目标核酸分子的方法在特异性调节基因组 DNA 甲基化/去甲基化状态中的应用。

在本发明中所述的编辑目标核酸分子的方法，目标核酸分子含有至少一个甲基化的胞嘧啶核苷酸，所述甲基化的胞嘧啶核苷酸与癌症、基因紊乱性遗传疾病、发育错误等疾病相关，该编辑目标核酸分子的方法可用于与胞嘧啶核苷酸甲基化有关的疾病的治疗，所述疾病包括但不限于与细胞分化异常相关的疾病。

本发明的有益效果：

本发明通过 sgRNA 的引导作用和突变体 Cas9 或 Cpf1 的特异结合功能，将具有脱氨基活性的 Apobec 蛋白引导至目标核酸分子的甲基化胞嘧啶位置，对甲基化胞嘧啶进行修饰。再通过体内 DNA 修复机制，去除甲基化胞嘧啶，从而实现目标核酸分子的特异性编辑。本发明的基因编辑方法，特异性高，对目标位点上下游序列无依赖性，因而具有普遍适用性。而且本发明的基因编辑方法仅对目标靶点进行编辑，不会产生脱靶效应，在编辑过程中也不会引入插入或缺失突变，具有较低的毒副作用。

附图说明

图 1，融合蛋白细胞外编辑示意图

图 2，融合蛋白细胞内编辑示意图

图 3，几种融合蛋白在体外实验中的活性强度及范围测试

图 4，编辑靶点前一位碱基对编辑效率的影响



(GGG-GGS-GGS-GGS-GGS-GGS-GGS-GGS-GGS-GGS-GGS-GGS-GGS) - dCas9(Asp10Ala/His840Ala)、6His-NLS-Apobec3H-linker (GGG-GGS-GGS-GGS-GGS-GGS-GGS) -dCpf1(Asp908Ala)基因序列, 分别为 Seq ID No 301、 No 302、 No 303、 No 304、 No 305、 No 306 和 No 307, 并在基因片段 5'端引入 Nco I内切酶位点, 3'端引入 Hind III内切酶位点。将合成的基因片段和 pET28a (+) 载体分别经 Nco I与 Hind III双酶切, T4 DNA 连接酶连接基因片段和载体片段, 常规转化 DH5 $\alpha$  感受态细胞(天根生化科技(北京)有限公司), 根据卡那霉素抗性筛选阳性克隆, 提取质粒。重组质粒经 Nco I与 Hind III双酶切和琼脂糖凝胶电泳鉴定, 同时委托 Invitrogen 公司对重组质粒进行序列测定, 使用 BioEdit 软件对测序结果进行分析, 结果与设计序列相同, 说明重组菌构建成功。

将获得的阳性克隆质粒转化到 E.Coli.BL21(DE3)感受态细胞(天根生化科技(北京)有限公司)内, 在含有 100ug/ml 卡那霉素的 LB 培养基中 37 度培养过夜, 之后转移到 1L 的同样的 LB 培养基中 37 度培养到 OD=0.6 左右。之后培养基被降温到 4 度, 加入 0.5mM 的 IPTG 诱导表达约 16 小时。菌体通过 4000g 离心收集, 用裂解缓冲液(50mM Tris pH=7.0, 1M NaCl, 20% glycerol, 10mM TCEP)重新悬浮。菌体由超声方法裂解(6W 输出 8 分钟, 20 秒开 20 秒关), 上清液通过 25000g 离心分离。将上清液和 Nickel resin(ThermoFisher)在 4 度下共置 1 小时, 然后通过重力柱并用 40ml 裂解缓冲溶液洗。重组蛋白用含有 285mM 的裂解缓冲溶液洗脱, 稀释至 0.1M 的 NaCl 并用离心管浓缩至合适浓度。重组蛋白的质量和浓度由 SDS-Page 决定。

重组蛋白序列为 Seq ID No 201-207。

### 实施例二 sgRNA 体外转录

根据拟测试的 34 条 dsDNA 底物序列 (Seq ID No 39-54 及其互补链 55-70, 71-85 及其互补链 86-100, 101-104 及其互补链 105-108) 和提供 sgRNA 通用序列的 pFYF320 载体序列, 分别设计 sgRNA 正向引物 (Seq ID No 2-17,18-34,35-38) 和反向引物 (Seq ID No 1)。sgRNA 由含有 T7 启动子的线性 DNA 片段通 TranscriptAid T7 High Yield Transcription Kit (ThermoFisher Scientific) 获得, 用 DpnI 去除模版 DNA, 然后使用 MEGAclean Kit (ThermoFisher Scientific) 纯化, 由 UV 吸收来检测质量。

### 实施例三 底物制备

委托 Invitrogen 公司分别合成底物序列的正反寡核酸链序列，其中正链序列的 5'端标记 FAM 荧光标记。分别将 2OD 的单链寡核酸链溶于 500ul 水中，将等量的正反链溶液混合，静置 5 分钟，获得双链底物 (dsDNA)。

用于 dCas9 融合蛋白去甲基化范围测试 15 个，序列号为 SEQ ID No.39-54。

用于 dCpf1 融合蛋白去甲基化范围测试 15 个，序列号为 SEQ ID No.71-85。

用于测试靶点前一位碱基对活性影响的 4 个，序列号为 SEQ ID No.101-104。

#### 实施例四 体外活性测试

将实施例一获得的重组融合蛋白，按照摩尔比 1:1 的比例分别与实施例二获得的 sgRNA 混合，在室温下放置 5 分钟。加入对应的 dsDNA 底物至终浓度 125nM，37°C 下反应 2 小时。使用 EconoSpin micro spin column (Epoch Life Science) 纯化获得反应过的 dsDNA 之后加入 1 单位的 TDG(NEB)37 度反应 1 小时。反应后加入 10ul Formamide、1ul 0.5M EDTA 以及 0.5ul 5M NaOH，在 95°C 下反应 5 分钟。产物在 10%的 TBE-尿素胶上分离。

目标 DNA 链中包含目标 Met-C，并且 3'端被荧光基团 FAM 标记。在重组蛋白作用下，Met-C 转变为 T 从而无法与互补链的 G 配对。在 TDG 的作用下，无法配对的 T 会被切除，留下碱基缺失位点。在 Formamide 和 NaOH 的作用下双链会变成单链并进一步在碱基缺失位点被切断，从而形成被荧光基团 FAM 标记的短链。长短链标记 DNA 在尿素胶中被分离。如果跑胶结果出现长短条带，则说明重组蛋白有活性。

#### 实施例五 焦磷酸测序 dsDNA 底物制备

委托 Invitrogen 公司分别合成底物序列的正反寡核酸链序列，其中正链序列的 5'端标记 FAM 荧光标记。分别将 2OD 的单链寡核酸链溶于 500ul 水中，将等量的正反链溶液混合，静置 5 分钟，获得双链底物 (dsDNA)。将实施例一获得的重组融合蛋白，按照摩尔比 1:1 的比例分别与实施例二获得的 sgRNA 混合，在室温下放置 5 分钟。加入对应的 dsDNA 底物至终浓度 125nM，37°C 下反应 2 小时。使用 EconoSpin micro spin column (Epoch Life Science) 纯化获得反应过的 dsDNA 之后交由华大基因通过亚硫酸盐处理及其设计的引物扩增后进行焦磷酸测序。

#### 实施例六 体内活性测定

##### (1) 细胞培养

HEK293 细胞系或 PC3 细胞系在 Dulbecco's Modified Eagle's Medium plus 中在 37°C 5%二氧化碳环境下维持。

(2) PX330 重组蛋白表达载体构建

委托 Invitrogen 公司合成 6His-NLS-Apobec3H-linker (GGG-GGG-GGG) - dCas9(Asp10Ala/His840Ala)、6His-NLS-Apobec3H-linker (GGG-GGG-GGG-GGG-GGG-GGG-GGG) 、6His-NLS-Apobec3H-linker (GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG) - dCas9(Asp10Ala/His840Ala)、6His-NLS-Apobec3A-linker (GGG-GGG-GGG) - dCas9(Asp10Ala/His840Ala) -dCas9(Asp10Ala/His840Ala)、6His-NLS-Apobec3A-linker (GGG-GGG-GGG-GGG-GGG-GGG-GGG) - dCas9(Asp10Ala/His840Ala) 、6His-NLS-Apobec3A-linker (GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG-GGG) - dCas9(Asp10Ala/His840Ala)、6His-NLS-Apobec3H-linker (GGG-GGG-GGG-GGG-GGG-GGG-GGG) -dCpf1(Asp908Ala)基因序列，分别为 Seq ID No 301、No 302、No 303、No 304、No 305、No 306 和 No 307，并在基因片段 5'端引入 BamHI 内切酶位点，3'端引入 AgeI 内切酶位点。将合成的基因片段和 pX330 载体 (Addgene) 分别经 BamHI 与 AgeI 双酶切，T4 DNA 连接酶连接基因片段和载体片段。经测序确认重组载体构建正确。

对应五组细胞内实验的 sgRNA 载体将对应的 PCR 产物 (由正向引物 121、123、125、127、129 及反向引物 1、122、124、126、128、130 通过 PCR 获得) 经 MluI 和 SpeI 双酶切插入。

(3) 转染

A. 转染前一天，HEK293 细胞或 PC3 细胞在不包含抗生素的培养基中接种，至转染时细胞的汇合度为 30~50%。

B. 制备转染样品：

分别将 1 $\mu$ l 浓度为 20 $\mu$ M 的 pX330 重组载体和 1.5 $\mu$ l 细胞转染试剂 Lipofectamine™ 2000 (Invitrogen) 稀释于 0.05ml Opti-MEM (Invitrogen)，轻轻混匀，孵育 5 分钟。对照组采用未克隆任何外源基因的空白 pX330 载体。

取稀释的 pX330 重组载体和 Lipofectamine™ 2000 (Invitrogen)，在室温下孵育 20 分钟，形成重组载体-Lipofectamine 2000 (Invitrogen) 复合物和空白载体-Lipofectamine 2000

(Invitrogen) 复合物。孵育时间不超过 30 分钟，更长的孵育时间可能会降低活性。

将载体- Lipofectamine™2000 复合物加入到每一个包含细胞和培养基的孔中，通过轻轻地前后摇动培养板混合，在 37°C，CO<sub>2</sub> 培养箱孵育 72 小时。

转染后的细胞 3 天后收获，基因组 DNA 由 Agencourt DNAdvance Genomic DNA Isolation Kit (Beckman Coulter) 纯化获得。通过实施例五的方法进行样品制备，所得样品交由深圳华大公司进行焦磷酸测序。

#### 实施例七 去甲基化作用位点范围测定

根据实施例二，发明人合成了 30 个（15 个针对 dCas9 的融合蛋白，15 个针对 dCpf1 的融合蛋白）长度为 59 个碱基的 ssDNA 做为反应底物、它们的互补 ssDNA、以及相应的 sgRNA 引物。其中反应底物 ssDNA 的 5'端被荧光基团 FAM 修饰，中间有一个甲基化修饰的 C (Met-C)，也就是编辑的目标。ssDNA 与其互补链形成 dsDNA 底物之后，重组蛋白的 Cas9 区域在相应的 sgRNA 引导下会结合到 dsDNA 中间的对应该区域，并且将该区域约 20 个碱基解链，也就是在 dsDNA 中间形成一个单链区。目标 Met-C 就在这个区域中，并且根据它到 5'端双链区的距离（4-20 个碱基）命名为底物 4-20。当重组蛋白与不同 sgRNA 结合后再与相应的 dsDNA 底物作用一定时间后，一部分的目标 Met-C 在脱氨基作用下会变成 T，与互补链上的 G 不配对形成一个突起。通过高温终止反应之后加入 1Unit 的 TDG，会移除没有办法正常配对的 T 碱基，从而在底物的编辑目标位置形成一个缺失。之后在 EDTA（0.5ul 浓度为 0.5M）、Formamide（10ul）和 NaOH（1ul 5M）的共同作用下 dsDNA 会重新变回 ssDNA 并且在碱基缺失位点断裂。由于断裂的 5'端短链 ssDNA 和未被作用的 ssDNA 底物在 5'端均有特定的 FAM 荧光基团标记，可以准确地测算两者的相对比例，并以此推断重组蛋白在该位点将 Met-C 变为 T 的作用效率。

如图 3，通过对 15 个不同底物的实验结果，可以看到：对于 linker 为 (GGS)<sub>3</sub> 的 dCas9 融合蛋白，在与目标区域双链解链后单链区域的 5'端第一个碱基距离为 7-10 个碱基内的 Met-C 是可以被变成 T 的，而超出这个范围的则不会；而对于 linker 为 (GGS)<sub>7</sub> 和 (GGS)<sub>14</sub> 的融合蛋白，这一编辑区间的距离为 6-11 个碱基和 5-13 个碱基。这个范围会因为 linker 的长度变长而稍微变宽。这个范围将是我们之后实验设计以及未来基因疗法设计 sgRNA 的重要依据。

从结果中也可看到 A3H 的活性比 A3A 要略强一些。

从结果可以看出，linker 长度为 (GGS)<sub>7</sub> 的 dCpf1 融合蛋白也有类似的活性，作用区间的距离为 7-12 个碱基。

对照组实验，阳性用合成的 T，阴性的用错误 sgRNA 和没有 sgRNA 的 Cas-9 或者 Cpf1。

对照实验主要为了证明两个问题：第一，我们的方法是可行的。选择其中某一个明显看到短链 DNA 形成的组，我们合成同样的 ssDNA 底物，但是把中间的 Met-C 变成 T，也就是人为完成了重组蛋白的功能。之后采用同样的操作。结果同样看到了短链 DNA 的出现。证明实验结果中的短链 DNA 确实是由重组蛋白在目标 DNA 上作用产生的；第二，通过让重组蛋白不结合或结合不配对的 sgRNA 再继续接下来的实验流程，并没有短链 DNA 产生，证明这样的编辑是定向的。

#### 实施例八 作用位点上下游碱基的影响

采用重组蛋白（linker 为 GGS\*7，Apobec 蛋白为 A3H）作用编辑靶点前一位碱基对去甲基化的活性影响的研究对象。

基于前人对 Apobec 蛋白家族的研究，编辑靶点前一位的碱基对其活性有直接的影响。我们选取了 Met-C 位于 7 号位的底物，并分别改变前一位碱基为 A、T、C 和 G。如图 4，试验结果显示，前一位碱基的序列对编辑效率基本没有影响，证明本技术的通用性。

#### 实施例九 细胞内去甲基化的效率

当我们已经证明了该重组蛋白在细胞外具有理想的将 Met-C 变为 T 的能力后，我们希望进一步验证在细胞内这样的活性是否依然存在、活性强度、以及反应后的 T 是否会被细胞自身的 DNA 修复机制修复成普通的 C，从而达到定点去甲基化的作用。申请人设计了三组细胞内实验，我们分别选取了三个不同基因的启动子区进行去甲基化测试。

第一个细胞内编辑目标为 HEK293 细胞系内，11 号染色体的 17741472 和 17741474 位点的两个甲基化 C，位于基因 MYOD1 的启动子区。如图 5，本实验证明了本系统可在两个位置很近的甲基化修饰中精确地编辑我们选择的那一个。

第二个编辑目标为 HEK293 细胞系内，6 号染色体 31138558 位点的一个甲基化 C，位于基因 POUF1 的启动子区。如图 5，本实验也获得了理想的编辑效果。

第三个编辑目标为 PC3 细胞系内，2 号染色体的 113875226 位点的一个甲基化 C，位于基因 IL1RN 的启动子区。如图 6，本系统通过合理的 sgRNA 设计可以相邻的两个甲基化位点中的一个或者两个位点。

采用实施例六所述的方法分别构建重组载体并转染至细胞，然后通过焦磷酸测序对编辑结果进行评估。

实施例十 编辑后细胞内出现 indel (insertion and deletion) 的比例

基于上述实验的测序结果，我们同时统计了在整个过程中在目标位点附近出现的碱基插入 (insertion) 和缺失 (deletion) 的情况。从测序结果来看周围没有插入和缺失碱基的现象出现。

实施例中所使用的核酸序列具体如下表所示。

Seq ID no.	名称	序列 (5'-3')
1	Rev_sgRNA_T7	AAAAAAAGCACCGACTCGGTG
2	Fwd_sgRNA_T7_ds DNA_4	TAATACGACTCACTATAGGTATCGGATTTATTTATTTAAGTTT TAGAGCTAGAAATAGC
3	Fwd_sgRNA_T7_ds DNA_5	TAATACGACTCACTATAGGTTATCGGATTTATTTATTTAGTTT TAGAGCTAGAAATAGC
4	Fwd_sgRNA_T7_ds DNA_6	TAATACGACTCACTATAGGTTTATCGGATTTATTTATTAGTTT TAGAGCTAGAAATAGC
5	Fwd_sgRNA_T7_ds DNA_7	TAATACGACTCACTATAGGATTTATCGGATTTATTTATTGTTT TAGAGCTAGAAATAGC
6	Fwd_sgRNA_T7_ds DNA_8	TAATACGACTCACTATAGGTATTTATCGGATTTATTTATGTTT TAGAGCTAGAAATAGC
7	Fwd_sgRNA_T7_ds DNA_9	TAATACGACTCACTATAGGTTATTTATCGGATTTATTTAGTTT TAGAGCTAGAAATAGC
8	Fwd_sgRNA_T7_ds DNA_10	TAATACGACTCACTATAGGATTATTTATCGGATTTATTTGTTT TAGAGCTAGAAATAGC
9	Fwd_sgRNA_T7_ds DNA_11	TAATACGACTCACTATAGGTATTATTTATCGGATTTATTGTTT TAGAGCTAGAAATAGC
10	Fwd_sgRNA_T7_ds DNA_12	TAATACGACTCACTATAGGATTATTATTATCGGATTTATGTTT TAGAGCTAGAAATAGC
11	Fwd_sgRNA_T7_ds DNA_13	TAATACGACTCACTATAGGTATTATATTTATCGGATTTAGTTT TAGAGCTAGAAATAGC
12	Fwd_sgRNA_T7_ds DNA_14	TAATACGACTCACTATAGGTTATTATATTTATCGGATTTGTTT TAGAGCTAGAAATAGC

13	Fwd_sgRNA_T7_dsDNA_15	TAATACGACTCACTATAGGATTATTATATTTATCGGATTGTTT TAGAGCTAGAAATAGC
14	Fwd_sgRNA_T7_dsDNA_16	TAATACGACTCACTATAGGTATTATTATATTTATCGGATGTTT TAGAGCTAGAAATAGC
15	Fwd_sgRNA_T7_dsDNA_17	TAATACGACTCACTATAGGATTATTATTATTATATCGGAGTTT TAGAGCTAGAAATAGC
16	Fwd_sgRNA_T7_dsDNA_20	TAATACGACTCACTATAGGATTATTATTATTATTATATCGTTT TAGAGCTAGAAATAGC
17	Fwd_sgRNA_T7_dsDNA_noC	TAATACGACTCACTATAGGTATAGGATTTATTTATTTAAGTTT TAGAGCTAGAAATAGC
18	Fwd_crRNA_T7	TAATACGACTCACTATAGGAATTTCTACTGTTGTAGATG
19	Rev_crRNA_T7_dsDNA_4	TTAAATAAATAAATCCGATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
20	Rev_crRNA_T7_dsDNA_5	TAAATAAATAAATCCGATAACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
21	Rev_crRNA_T7_dsDNA_6	TAATAAATAAATCCGATAAACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
22	Rev_crRNA_T7_dsDNA_7	AATAAATAAATCCGATAAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
23	Rev_crRNA_T7_dsDNA_8	ATAAATAAATCCGATAAATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
24	Rev_crRNA_T7_dsDNA_9	TAAATAAATCCGATAAATAACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
25	Rev_crRNA_T7_dsDNA_10	AAATAAATCCGATAAATAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
26	Rev_crRNA_T7_dsDNA_11	AATAAATCCGATAAATAATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
27	Rev_crRNA_T7_dsDNA_12	ATAAATCCGATAAATAAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
28	Rev_crRNA_T7_dsDNA_13	TAAATCCGATAAATATAATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
29	Rev_crRNA_T7_dsDNA_14	AAATCCGATAAATATAAATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA

30	Rev_crRNA_T7_dsDNA_15	AATCCGATAAATATAATAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
31	Rev_crRNA_T7_dsDNA_16	ATCCGATAAATATAATAATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
32	Rev_crRNA_T7_dsDNA_17	TCCGATATAATAATAATAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
33	Rev_crRNA_T7_dsDNA_20	GATATAATAATAATAATAATCATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
34	Rev_crRNA_T7_dsDNA_noC	TTAAATAAATAAATCCTATACATCTACAACAGTAGAAATTCC TATAGTGAGTCGTATTA
35	Fwd_sgRNA_6T	TAATACGACTCACTATAGGTTATTTTCGTGGATTTATTTAGTTT TAGAGCTAGAAATAGC
36	Fwd_sgRNA_6A	TAATACGACTCACTATAGGTTATTTTCGTGGATTTATTTAGTTT TAGAGCTAGAAATAGC
37	Fwd_sgRNA_6C	TAATACGACTCACTATAGGTTATTTTCGTGGATTTATTTAGTTT TAGAGCTAGAAATAGC
38	Fwd_sgRNA_6G	TAATACGACTCACTATAGGTTATTTTCGTGGATTTATTTAGTTT TAGAGCTAGAAATAGC
39	dCas9_ds_4	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATmet- CGGATTTATTTATTTAAT GGATGACCTCTGGATCCATG
40	dCas9_ds_5	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTTATmet- CGGATTTATTTATTTAT GGATGACCTCTGGATCCATG
41	dCas9_ds_6	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTTTATmet- CGGATTTATTTATTAT GGATGACCTCTGGATCCATG
42	dCas9_ds_7	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTTATmet- CGGATTTATTTATTT GGATGACCTCTGGATCCATG
43	dCas9_ds_8	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATTTATmet -CGGATTTATTTATT GGATGACCTCTGGATCCATG
44	dCas9_ds_9	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTTATTTATm et-CGGATTTATTTAT GGATGACCTCTGGATCCATG

45	dCas9_ds_10	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTATTTAT met-CGGATTTATTTT GGATGACCTCTGGATCCATG
46	dCas9_ds_11	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATTATTTA Tmet-CGGATTTATTT GGATGACCTCTGGATCCATG
47	dCas9_ds_12	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTATTATT ATmet-CGGATTTATT GGATGACCTCTGGATCCATG
48	dCas9_ds_13	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATTATATT TATmet-CGGATTTAT GGATGACCTCTGGATCCATG
49	dCas9_ds_14	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTTATTATAT TTATmet-CGGATTTT GGATGACCTCTGGATCCATG
50	dCas9_ds_15	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTATTATA TTTATmet-CGGATTT GGATGACCTCTGGATCCATG
51	dCas9_ds_16	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATTATTAT ATTTATmet-CGGATT GGATGACCTCTGGATCCATG
52	dCas9_ds_17	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTATTATT ATTATATmet-CGGAT GGATGACCTCTGGATCCATG
53	dCas9_ds_20	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCATTATTATT ATTATTATATmet-CT GGATGACCTCTGGATCCATG
54	dCas9_ds_noC	FAM- GGTAGTTAGGATGAATGGAAGGTTGGTATAGCCTATAGGATT TATTTATTTAAT GGATGACCTCTGGATCCATG
55	dCas9_ds_com_4	CATGGATCCAGAGGTCATCCATTAATAAATAAATCCGATAG GCTATACCAACCTTCC ATTCATCCTAACTACC
56	dCas9_ds_com_5	CATGGATCCAGAGGTCATCCATAAATAAATAAATCCGATAA GGCTATACCAACCTTCC ATTCATCCTAACTACC
57	dCas9_ds_com_6	CATGGATCCAGAGGTCATCCATAAATAAATAAATCCGATAAA GGCTATACCAACCTTCC ATTCATCCTAACTACC
58	dCas9_ds_com_7	CATGGATCCAGAGGTCATCCAAATAAATAAATCCGATAAAT GGCTATACCAACCTTCC ATTCATCCTAACTACC

59	dCas9_ds_com_8	CATGGATCCAGAGGTCATCCAATAAATAAATCCGATAAATA GGCTATACCAACCTTCC ATTCATCCTAACTACC
60	dCas9_ds_com_9	CATGGATCCAGAGGTCATCCATAAATAAATCCGATAAATAA GGCTATACCAACCTTCC ATTCATCCTAACTACC
61	dCas9_ds_com_10	CATGGATCCAGAGGTCATCCAAAATAAATCCGATAAATAAT GGCTATACCAACCTTCC ATTCATCCTAACTACC
62	dCas9_ds_com_11	CATGGATCCAGAGGTCATCCAAATAAATCCGATAAATAATA GGCTATACCAACCTTCC ATTCATCCTAACTACC
63	dCas9_ds_com_12	CATGGATCCAGAGGTCATCCAATAAATCCGATAATAATAATG GCTATACCAACCTTCC ATTCATCCTAACTACC
64	dCas9_ds_com_13	CATGGATCCAGAGGTCATCCATAAATCCGATAAATAATAAG GCTATACCAACCTTCC ATTCATCCTAACTACC
65	dCas9_ds_com_14	CATGGATCCAGAGGTCATCCAAAATCCGATAAATAATAATA GGCTATACCAACCTTCC ATTCATCCTAACTACC
66	dCas9_ds_com_15	CATGGATCCAGAGGTCATCCAAATCCGATAAATAATAATAATG GCTATACCAACCTTCC ATTCATCCTAACTACC
67	dCas9_ds_com_16	CATGGATCCAGAGGTCATCCAATCCGATAAATAATAATAATAG GCTATACCAACCTTCC ATTCATCCTAACTACC
68	dCas9_ds_com_17	CATGGATCCAGAGGTCATCCATCCGATATAATAATAATAATG GCTATACCAACCTTCC ATTCATCCTAACTACC
69	dCas9_ds_com_20	CATGGATCCAGAGGTCATCCAGATATAATAATAATAATAATG GCTATACCAACCTTCC ATTCATCCTAACTACC
70	dCas9_ds_com_noC	CATGGATCCAGAGGTCATCCATTAAATAAATAAATCCTATAG GCTATACCAACCTTCC ATTCATCCTAACTACC
71	dCpf1_ds_4	FAM-GGTACCCGGGGATCCTTTATATmet- CGGATTTATTTATTTAAGTTAAAAAGCTTGGCGTAAT
72	dCpf1_ds_5	FAM-GGTACCCGGGGATCCTTTATTATmet- CGGATTTATTTATTTAGTTAAAAAGCTTGGCGTAAT
73	dCpf1_ds_6	FAM-GGTACCCGGGGATCCTTTATTTATmet- CGGATTTATTTATTAGTTAAAAAGCTTGGCGTAAT
74	dCpf1_ds_7	FAM-GGTACCCGGGGATCCTTTAATTTATmet- CGGATTTATTTATTGTTAAAAAGCTTGGCGTAAT
75	dCpf1_ds_8	FAM-GGTACCCGGGGATCCTTTATATTTATmet- CGGATTTATTTATGTTAAAAAGCTTGGCGTAAT

76	dCpf1_ds_9	FAM-GGTACCCGGGGATCCTTTATTATTTATmet- CGGATTTATTTAGTTAAAAAGCTTGGCGTAAT
77	dCpf1_ds_10	FAM-GGTACCCGGGGATCCTTTAATTATTTATmet- CGGATTTATTTGTTAAAAAGCTTGGCGTAAT
78	dCpf1_ds_11	FAM-GGTACCCGGGGATCCTTTATATTATTTATmet- CGGATTTATTGTTAAAAAGCTTGGCGTAAT
79	dCpf1_ds_12	FAM-GGTACCCGGGGATCCTTTAATTATTATTATmet- CGGATTTATGTTAAAAAGCTTGGCGTAAT
80	dCpf1_ds_13	FAM-GGTACCCGGGGATCCTTTATATTATATTTATmet- CGGATTTAGTTAAAAAGCTTGGCGTAAT
81	dCpf1_ds_14	FAM-GGTACCCGGGGATCCTTTATTATTATATTTATmet- CGGATTTGTTAAAAAGCTTGGCGTAAT
82	dCpf1_ds_15	FAM-GGTACCCGGGGATCCTTTAATTATTATATTTATmet- CGGATTGTTAAAAAGCTTGGCGTAAT
83	dCpf1_ds_16	FAM-GGTACCCGGGGATCCTTTATATTATTATATTTATmet- CGGATGTTAAAAAGCTTGGCGTAAT
84	dCpf1_ds_17	FAM-GGTACCCGGGGATCCTTTAATTATTATTATTATATmet- CGGAGTTAAAAAGCTTGGCGTAAT
85	dCpf1_ds_20	FAM- GGTACCCGGGGATCCTTTAATTATTATTATTATATmet- CGTTAAAAAGCTTGGCGTAAT
86	dCpf1_ds_com_4	ATTACGCCAAGCTTTTTAACTTAAATAAATAAATCCGATATA AAGGATCCCCGGGTACC
87	dCpf1_ds_com_5	ATTACGCCAAGCTTTTTAACTAATAAATAAATCCGATAATA AAGGATCCCCGGGTACC
88	dCpf1_ds_com_6	ATTACGCCAAGCTTTTTAACTAATAAATAAATCCGATAAATA AAGGATCCCCGGGTACC
89	dCpf1_ds_com_7	ATTACGCCAAGCTTTTTAACATAAATAAATCCGATAAATTA AAGGATCCCCGGGTACC
90	dCpf1_ds_com_8	ATTACGCCAAGCTTTTTAACATAAATAAATCCGATAAATATA AAGGATCCCCGGGTACC
91	dCpf1_ds_com_9	ATTACGCCAAGCTTTTTAACTAATAAATAAATCCGATAAATAATA AAGGATCCCCGGGTACC
92	dCpf1_ds_com_10	ATTACGCCAAGCTTTTTAACAAATAAATAAATCCGATAAATAATTA AAGGATCCCCGGGTACC

93	dCpf1_ds_com_11	ATTACGCCAAGCTTTTTAACAATAAATCCGATAAATAATATA AAGGATCCCCGGGTACC
94	dCpf1_ds_com_12	ATTACGCCAAGCTTTTTAACATAAATCCGATAATAATAATTA AAGGATCCCCGGGTACC
95	dCpf1_ds_com_13	ATTACGCCAAGCTTTTTAACTAAATCCGATAAATAATATA AAGGATCCCCGGGTACC
96	dCpf1_ds_com_14	ATTACGCCAAGCTTTTTAACAAATCCGATAAATAATAATA AAGGATCCCCGGGTACC
97	dCpf1_ds_com_15	ATTACGCCAAGCTTTTTAACAATCCGATAAATAATAATTA AAGGATCCCCGGGTACC
98	dCpf1_ds_com_16	ATTACGCCAAGCTTTTTAACATCCGATAAATAATAATATA AAGGATCCCCGGGTACC
99	dCpf1_ds_com_17	ATTACGCCAAGCTTTTTAACTCCGATATAATAATAATAATTA AAGGATCCCCGGGTACC
100	dCpf1_ds_com_20	ATTACGCCAAGCTTTTTAACGATATAATAATAATAATAATTA AAGGATCCCCGGGTACC
101	dCas9_ds_6T	ACGTAAACGGCCACAAGTTCTTATTTmet- CGTGGATTTATTTATGGCATCTTCTTCAAGGAC
102	dCas9_ds_6A	ACGTAAACGGCCACAAGTTCTTATTmet- CGTGGATTTATTTATGGCATCTTCTTCAAGGAC
103	dCas9_ds_6C	ACGTAAACGGCCACAAGTTCTTATTCmet- CGTGGATTTATTTATGGCATCTTCTTCAAGGAC
104	dCas9_ds_6G	ACGTAAACGGCCACAAGTTCTTATTGmet- CGTGGATTTATTTATGGCATCTTCTTCAAGGAC
105	dCas9_ds_com_6T	GTCCTTGAAGAAGATGCCATAAATAAATCCACGAAATAAGA ACTTGTGGCCGTTTACGT
106	dCas9_ds_com_6A	GTCCTTGAAGAAGATGCCATAAATAAATCCACGTAATAAGA ACTTGTGGCCGTTTACGT
107	dCas9_ds_com_6C	GTCCTTGAAGAAGATGCCATAAATAAATCCACGGAATAAGA ACTTGTGGCCGTTTACGT
108	dCas9_ds_com_6G	GTCCTTGAAGAAGATGCCATAAATAAATCCACGCAATAAGA ACTTGTGGCCGTTTACGT
109	ds_6_F	CGTAAACGGCCACAAGTTCTTAT

110	ds_6_R	GTCCTTGAAGAAGATGCCATAAA
111	ds_6_S	CGGCCACAAGTTCTTAT
112	HEK293T-T1-F	GGATTTGYGTTTTTTYGAAGATTTGG
113	HEK293T-T1-R	AAATACRAATACTCTTCRAATTTCAAAAAC
114	HEK293T-T1-S	GTTTTTTAGAAAGATTTGGAT
115	HEK293T-T2-F	GTTTTGAATGAATGTGTGTATATATGTATG
116	HEK293T-T2-R	CTAACAAAAACCAAATAATTCTTATCTAC
117	HEK293T-T2-S	ATGAATGTGTGTATATATGTATGAG
118	PC3-F	TAAGGGTTTTTYGGAAYGGGGT
119	PC3-R	CCAAACAAAACATCCCTCAAC
120	PC3-S	GGGTTGTGTGAGTGGG
121	HEK293T-gRNA1-F	CACCG GGACCCGCGCCTGATGCACG
122	HEK293T-gRNA1-R	AAAC CGTGCATCAGGCGCGGGTCC C
123	HEK293T-gRNA2-F	CACCG GAGCTGGCGGCAGTCGGGGT
124	HEK293T-gRNA2-R	AAAC ACCCCGACTGCCGCCAGCTC C
125	Gfap-gRNA-F	CACCG TTCCGAGAAGTCTATTGAGC
126	Gfap-gRNA-R	AAAC GCTCAATAGACTTCTCGGAA C

127	PMP24-gRNA-F	CACCG TGGGGCCGTCGGGGCCGGGCT
128	PMP24-gRNA-R	AAAC AGCCCGGCCCGACGGCCCCA C
129	C/EBP $\delta$ -gRNA-F	CACCG TCAGCCGGGGCTAGAAAAGG

蛋白结构域序列如下所示:

>APOBEC3A

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVERLDNGT SVKMDQHRGFLHNQAK  
 NLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTFWISWSPCFSWGCAGEVRAFLQENTHVR  
 LRIFAARIYDYDPLYKEALQMLRDAGAQVSIMTYDEFKHCWDTFVDHQGCPFQPWDGLD  
 EHSQALSGRLRAILQNQGN

>APOBEC3H Hyplotype II

MALLTAETFRLQFNNKRRLRRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFIN  
 EIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLRIFASRLYYHWCKPQQD  
 GLRLLCGSQVPVEVMGFPEFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLDRIKS

>Cas9

MDKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEAT  
 RLKRTARRRYTRRKNRICYLQEFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVD  
 EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFI  
 QLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGLT  
 PNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT  
 EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF  
 YKFIKPILEKMDGTEELLVKLNRDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFKLD  
 NREKIEKILTRIPYYVGPLARGNSRFAMTRKSEETITPWNFEEVVDK GASAQSFIERMTN  
 FDKNLPNEK VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKAIVDLLFKTNRK  
 VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVL  
 TLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF  
 LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV  
 DELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGKELGSQILKEHPVENTQ  
 LQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRG  
 KSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR  
 QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYK VREINNYHHAHD  
 AYLN AVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK  
 TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESI  
 LPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEK GKSKKLSVKELLGITIMER  
 SSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV  
 NFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN  
 KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETR  
 IDLSQLGGDPPKKRKY

>Cpf1

MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYAD  
 QCLQLVQLDWENLSAIDSYRKEKTEETRNLALIEEQATYRNAIHDFYFIGRTDNLTDAINKRH  
 AEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFS AEDIST  
 AIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQ  
 TQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTL  
 SFILEEFKSDEEVIQSFCKYKTLRNENVLETAEALFNELNSIDLTHIFISHKKLETISSALCDH  
 WDTLRNALYERRISELTGKITKSAKEK VQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILS  
 HAHAALDQPLPTTLKKQEEKEILKSQLDSLGLYHLLDWFVAVDESNEVDPEFSARLTGIKLE  
 MEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYLYG  
 IMPKQKGRYKALSFEPTEKTSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTHHTPIL  
 LSNNFIEPLEITKEIYDLNNPEKEPKKFQ TAYAKKTGDQKGYREALCKWIDFTRDFLSKYTK  
 TTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFAK  
 GHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGKMLNKKL  
 KDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKFFFHVP  
 ITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQFD  
 YQKKLDNREKERRAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVLENLNF GF  
 KSKRTGIAEKAVYQFQEKMLIDKLNCLVLKDYPAEKVGGLVLPYQLTDQFTSFAKMGTS  
 GFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLGDFDLHYDVKTGDFILHFKMNR  
 NLSFQRGLPGFMPAWDIVFEKNETQF DAKGTPFIAGKRIVVIENHRFTGRYRDLYPANELIA  
 LLEEKGIVFRDGSNILPKLLENDSDHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDLNG  
 VCFDSRFQNPWPMDADANGAYHIALKGQLLNHLKESKDLKLQNGISNQDWLAYIQELR  
 N

Seq ID NO201:

>6his-NLS-A3A-GGS3-dCas9

HHHHHH-SSGLVPRGSHM-PKKKRKV-

MEASPASGPRHLMDPHIFTSNFNNGIGRHKTYLCYEVEERLDNGT SVKMDQHRGFLHNQAK  
 NLLCGFYGRHAELRFLDLVPSLQLDPAQIYRVTWFIWSPCF SWGCAGEVRAFLQENTHVR  
 LRIFAARIYDYDPLYKEALQMLRDAGA QVSIMTYDEFKHCWDTFVDHQGCPFQPDGDL  
 EHSQALSGRLRAILQNQGN-GGSGGSGGS-  
 MDKKYSIGLAIGTNSVGWAVITDEYK VPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA  
 TRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIV  
 DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKL  
 FIQLVQTYNQLFEEPNINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLGL  
 LTPNFKSNFDLAEDAQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV  
 NTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ  
 EEFYKFIKPILEKMDGTEELLVKLNREDDL RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF  
 LKDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPWNFEEVVDKGASAQSFIER  
 MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK  
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL  
 EDIVLTLTLFEDREMIEERLKYAHLFDDK VMKQLKRRRYTGWGRLSRKLINGIRDKQSGK  
 TILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQ  
 TVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHP  
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS  
 DKNRGKSDNPSEE VVKMKNYWRQLLNAKLITQRKFDNLTKAERGG LSELDKAGFIKR  
 QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYK VREINN  
 YHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYK VYDVRKMIKSEQEIGKATAKYFFYS



FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLG  
 LTPNFKSNFDLAEDAQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV  
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ  
 EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF  
 LKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER  
 MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK  
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL  
 EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK  
 TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQ  
 TVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHP  
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRS  
 DKNRGKSDNPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR  
 QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKREINN  
 YHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVVRKMIKSEQEIGKATAKYFFYS  
 NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQT  
 GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE  
 LLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGN  
 ELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADAN  
 LDKVLSAYNKHHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLI  
 HQSITGLYETRIDLSQLGGDPPKKKRKV

Seq ID NO204:

>6his-NLS-A3H-GGS3-dCas9

HHHHHH-SSGLVPRGSHM-PKKKRKV-

MALLTAETFRLQFNKRRLRRPYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFIN  
 EIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDLNLRIFASRLYYHWCKPQQD  
 GLRLLCGSQVPVEVMGFPEFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLDRIKS  
 -GGSGGSGGS-

MDKKYSIGLAIGTNSVGWAVITDEYKVPKSKFKVLGNTDRHSIKKNLIGALLFDSGETAEA  
 TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVVDSFFHRLEESFLVEEDKKHERHPIFGNIV  
 DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKL  
 FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLG  
 LTPNFKSNFDLAEDAQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRV  
 NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ  
 EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF  
 LKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER  
 MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK  
 TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL  
 EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK  
 TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQ  
 TVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEIEGKELGSQILKEHP  
 VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVAIVPQSFLKDDSIDNKVLTRS  
 DKNRGKSDNPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR  
 QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKREINN  
 YHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVDVVRKMIKSEQEIGKATAKYFFYS  
 NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQT  
 GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE  
 LLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGN  
 ELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR VILADAN  
 LDKVLSAYNKHHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLI  
 HQSITGLYETRIDLSQLGGDPPKKKRKV

Seq ID NO205:

>6his-NLS-A3H-GGS7-dCas9

HHHHHH-S SGLVPRGSHM-PKKKRKV-

MALLTAETFRLQFNNKRRLLRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFIN
EIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHNLNRIFASRLYYHWCKPQQD
GLRLLCGSQVPVEVMGFPEFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLDRIKS
-GGSGGSGGSGGSGGSGGSGGSGGS-

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDV DKL
FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLG
LTPNFKSNFDLAEDAQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYEIFFDQSKNGYAGYIDGGASQ
EEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF
LKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER
MTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIKDKDFLDNEENEDIL
EDIVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK
TILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGIQ
TVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEIGIKELGSQILKEHP
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS
DKNRGKSDNPSEE VVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINN
YHHAHDAYLNAVVG TALIKKYPKLESEFVYG DYKVYDVRKMIAKSEQEIGKATAKYFFYS
NIMNFFKTEITLANGEIRKRIETNGETGEIVWDKGRDFATVRKVL SMPQVNIKKTEVQT
GGFSKESILPKRNSDKLIARKKDWD PKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE
LLGITIMERSSEK NPIDFLEAKGYKEVKKD LIKLPKYSLFELENGRKRMLASAGELQKGN
ELALPSKYVNFLYLASHYEKLKGPEDNEQKQLFVEQH KHYLDEIIEQISEFSKR VILADAN
LDKVL SAYNKH RDKPIREQAENIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLI
HQSITGLYETRIDL SQLGGDPPKPKKRKV

Seq ID NO206:

>6his-NLS-A3H-GGS14-dCas9

HHHHHH-S SGLVPRGSHM-PKKKRKV-

MALLTAETFRLQFNNKRRLLRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFIN
EIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHNLNRIFASRLYYHWCKPQQD
GLRLLCGSQVPVEVMGFPEFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLDRIKS
-GGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGSGGS-

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA
TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIV
DEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLA LAHMIKFRGHFLIEGDLNPDNSDV DKL
FIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIASLG
LTPNFKSNFDLAEDAQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLS DAILLSDILRV
NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYEIFFDQSKNGYAGYIDGGASQ
EEFYKFIKPILEKMDGTEELLVKNLREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF
LKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER
MTNFDKNLPNEKVLPKHSLLEYEFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK

TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL  
EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGK  
TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIIKKGILQ  
TVKVVDDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEHP  
VENTQLQNEKLYLYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRS  
DKNRGKSDNPSEEVVKMKNYWRQLLNAKLITQRKFDNLTKAERGGELSEDKAGFIKR  
QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINN  
YHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVVYDVRKMIKSEQEIGKATAKYFFYS  
NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLVSMQVNVIVKKTEVQT  
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGGKSKKLKSVKE  
LLGITIMERSSEKPNIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQKGN  
ELALPSKYVNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADAN  
LDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLI  
HQSITGLYETRIDLSQLGGDPPKKKRKV

Seq ID NO207:

>6his-NLS-A3H-GGS7-dCpf1 基因序列

HHHHHH-SSGLVPRGSHM-PKKKRKV-

MALLTAETFRLQFNNKRRRLRPYYPRKALLCYQLTPQNGSTPTRGYFENKKKCHAEICFIN  
EIKSMGLDETQCYQVTCYLTWSPCSSCAWELVDFIKAHDHLNLRIFASRLYYHWCKPQQD  
GLRLLCGSQVPVEVMGFPEFADCWENFVDHEKPLSFNPYKMLEELDKNSRAIKRRLDRIKS  
-GGSGGSGGSGGSGGSGGSGGS-

KLTFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYA  
DQCLQLVQLDWENLSAAIDSYRKEKTEETRNLIEEQATYRNAIHDFIGRTDNLTDANK  
RHAEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFS  
AEDISTAIPHRIQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFY  
NQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQIL  
SDRNLTSFILEEFKSDEEVIQSFCKYKTLLRNENVLETAELFNELNSIDLTHIFISHKKLET  
ISSALCDHWDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIISAAGKELSEAFK  
QKTSEILSHAHAALDQPLPTTLKKQEEKEILKSQDLSLLGLYHLLDWFAVDESNEVDPEFS  
ARLTGIKLEMPSLSFYNKARNYATKPKYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVK  
NGLYYLGIMPQKQGRYKALSFEPTKTSEGFDKMYDYFPDAAKMIPKCSTQLKAVTAHFQTH  
TTPILLSNNFIEPLEITKEIYDLNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLS  
KYTKTTSIDLSSLRPSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNK  
DFAKGGHGGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKML  
NKKLKDQKTPIDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKF  
FFHVPITLNYQAANSKFNQRVNAYLKEHPETPIIGIARGERNLIYITVIDSTGKILEQRSLN  
TIQQFDYQKKLDNREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVLE  
NLNFGFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGLNPNYQLTDQFTSFA  
KMGTQSGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLGDFDLHYDVKTGDFIL  
HFKMNRNLSFQRGLPGFMPAWDIVFEKNETQFDAQGTPFIAGKRIVPVIENTHRFTGRYRDL  
YPANELIALLEEKGIVFRDGSNILPKLLENDSDHAIDTMVALIRSVLQMRNSNAATGEDYIN  
SPVRDLNGVCFDSRFQNPWPMDADANGAYHIALKGQLLNHLKESKDLKLQNGISNQD  
WLAYIQELRN

Seq ID NO301:

>6his-NLS-A3A-GGS3-dCas9 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtgccgcggcagccatgccaagaagaagcggaggtcGAA  
 GCCAGCCCAGCATCCGGGCCAGACACTTGATGGATCCACACATATTCACCTTCCAACCTT  
 TAACAATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGAC  
 AATGGCACCTCGGTCAAGATGGACCAGCACAGGGGCTTTCTACACAACCAGGCTAAGA  
 ATCTTCTCTGTGGCTTTTACGGCCGCCATGCGCAGCTGCGCTTCTTGGACCTGGTTCCCTT  
 CTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTTCATCTCCTGGAGCCCC  
 TGCTTCTCCTGGGGCTGTGCCGGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACG  
 TGAGACTGCGTATCTTCGCTGCCCGCATCTATGATTACGACCCCCTATATAAGGAGGCA  
 CTGCAAATGCTGCGGGATGCTGGGGCCCAAGTCTCCATCATGACCTACGATGAATTTA  
 AGCACTGCTGGGACACCTTTGTGGACCACCAGGGATGTCCCTTCCAGCCCTGGGATGG  
 ACTAGATGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAG  
 GGAAACGGAGGAAGTGGAGGAAGTGGAGGAAGTaaagcttgacaagaagtacagcatcgccctggccatcgg  
 caccaactctgtgggctgggcccgtgatccagcaggtacaaggtgccagcaagaattcaaggtgctgggcaacaccgaccggcacagc  
 atcaagaagaacctgatcggagccctgctgttcgacagcggcgaacagccgaggccaccggctgaagagaaccgccagaagaagatac  
 accagacggaagaaccggatctgctatctgcaagagatctcagcaacagatggccaaggtggacgacagcttctccacagactggaaga  
 gtccttctggtggaagaggataagaagcacgagcggcaccatcttcggcaacatcgtggacgaggtggcctaccacgagaagtaccca  
 ccatctaccactgagaaagaaactggtggacagcaccgacaaggccgacctgcccgtgatctatctggccctggcccacatgatcaagtcc  
 ggggccaacttctgatcagggcgacctgaaccccgacaacagcagcgtggacaagctgttcatccagctggtgcagacctacaaccagctg  
 ttcgaggaaaacccatcaacgccagcggcgtggacgccaagccatctgtctgcccagactgagcaagagcagacggctggaanaatctga  
 tcgcccagctgcccgcgagaagaagaatggcctgttcggaaacctgattgccctgagcctgggctgaccccaactcaagagcaacttcg  
 acctggccgaggatgccaactgcagctgagcaaggacacctacgacgacacctggacaacctgctggcccagatcggcgaccagtacg  
 ccgacctgttctgcccgaagaacctgtccgacgccatctctgtagcgacatctgagagtgaaaccagagatcaccaaggccccctga  
 ggcctctatgatcaagagatacagcagcaccaccaggacctgacctgctgaaagctctcgtgcccagcagctgacctgagaagtacaaa  
 gagatcttctgaccagagcaagaacggctacgcccgtacattgacggcggagccagccaggaagagttctacaagttcatcaagccatc  
 ctggaanaagatggacggcaccgaggaactgctcgtgaagctgaacagagaggacctgctcgggaagcagcggaccttcgacaacggcagc  
 atccccaccagatccacctgggagagctgcacgccattctgcccggcagggaagattttaccattctgaaggacaaccgggaaaagatc  
 gagaagatcctgaccttcgcatcccactactgctggccctctggccaggggaaacagcagattcgcctggatgaccagaaagagcagga  
 aaccatccccctggaacttcgaggaagtgtggacaagggcgttccgccagagcttcatcagcggatgaccaacttcgataagaacct  
 gccaacgagaaggtgctgccaagcacagcctgctgtacgagtaactaccgtgtataacgagctgaccaaaagtgaatacgtgaccgagg  
 gaatgagaagccccttctgagcggcagcagaaaaaggccatctggacctgctgtcaagaccaaccggaaagtaccgtgaagca  
 gctgaaagaggactactcaagaaaatcagtgcttcgactccgtggaaatctccggcgtggaagatcggttcaacgctccttgggcacatac  
 cacgatctgctgaaaattatcaaggacaaggacttctggacaatgaggaaaacgaggacattctggaagatctgctgacctgacactgtt  
 tgaggacagagagatgatcaggaacggctgaaaacctatgccacctgttcgacgacaaaagtgatgaagcagctgaagcggcggagatac  
 accggctggggcaggctgagccggaagctgatcaacggcatccgggacaagcagctccggcaagacaatcctggattcctgaagtccgacg  
 gcttcgccaacagaaacttcatgcagctgatccacgacgacgacctgaccttaaaaggacatccagaaagcccaggtgtccggccagggc  
 gatagcctgcacgagcattgccaatctgcccggcagccccgccattaagaagggcatcctgcagacagtgaaagtggtggacgagctcgt  
 gaaagtgatggccggcacaagcccgagaacatcgtgatcgaatggccagagagaaccagaccaccagaaggacagagaagaacagcc  
 gcgagagaatgaagcggatcgaagagggcatcaaagagctgggcagccagatcctgaaagaacaccgctggaanaacaccagctgcag  
 aacgagaagctgtacctgactacctgcagaatggcggggatatgtacgtggaccaggaactggacatcaaccggctgtccgactacgatgtg  
 gacgctatcgtgcctcagagcttctgaaggacgactccatcgacaacaaggtgctgaccagaagcacaagaaccggggcaagagcgaca  
 acgtgccctcgaagaggtcgtgaagaagatgaagaactactggcggcagctgctgaacgccaagctgattaccagagaaagttcgacaat  
 ctgaccaaggccgagagagggcggcctgagcgaactggataaggccggcttcatcaagagacagctggtggaanaaccggcagatcacaag  
 cacgtggcacagatcctggactcccggatgaactaagtacgacgagaatgacaagctgatccgggaagtgaagtgatcacctgaagtc  
 caagctggtgtccgatttccggaaggatttccagttttacaagtgccgagatcaacaactaccaccagcccagcgcctacctgaacgcc  
 gtcgtgggaaccgacctgatcaaaaagtaccctaagctggaagcaggtcgtgtacggcgactacaaggtgtacgacgtgcccgaagatgatc  
 gccaaagagcagcaggaatcggcaaggctaccgccaagtacttctctacagcaacatcatgaacttttcaagaccgagattaccctggcca  
 acggcgagatccggaagcggcctctgatcgagacaacggcgaacccggggagatcgtgtgggataagggccgggattttgccaccctgctc  
 ggaaagtctgagcatgcccgaagtgaatatcgtgaaaaagaccgaggtgcagacagggcgttcagcaaaagagctatctgcccgaaggg  
 aacagcgataagctgatccagaaagaaggactgggaccctaagaagtacggcggcttcgacagccccaccgtggcctattctgtgctggt  
 ggtggccaaaagggaaggcaagtccaagaaactgaagagtgtaaaagagctgctgggatcaccatcatggaagaagcagcttcgag  
 aagaatccatcgacttctggaagccaagggtcacaagaagtgaanaaggacctgatcatcaagctgcctaagctcctgttcgagctgg  
 aaaaacggccggaagagaatgctggcctctgcccggcaactgcagaagggaaacgaactggccctgcctccaatatgtgaacttctgtac

ctggccagccactatgagaagctgaaggctccccgaggataatgagcagaaacagctgtttgtggaacagcacaagcactacctggacga  
gatcatcgagcagatcagcgagttctccaagagagtgatcctggccgacgctaactctggacaaagtgtgtccgctacaacaagcaccggga  
taagccatcagagagcagggcagaatatcatccacctgtttaccctgaccaatctgggagcccctgccgcttcaagtactttgacaccacca  
tcgaccgggaagaggtacaccagcaccaaaagaggtgctggacgccaccctgatccaccagagcatcaccggcctgtacgagacacggatcg  
acctgtctcagctgggagggcactaactcgag

Seq ID NO302:

>6his-NLS-A3A-GGS7-dCas9 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtgccgcgccgagccatattgcaaagaagaagcggaaaggtcGAA  
GCCAGCCCAGCATCCGGGCCAGACACTTGATGGATCCACACATATTCACCTTCCAACCT  
TAACAATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGAC  
AATGGCACCTCGGTCAAGATGGACCAGCACAGGGGCTTTCTACACAACCAGGCTAAGA  
ATCTTCTCTGTGGCTTTTACGGCCGCCATGCGCAGCTGCGCTTCTTGGACCTGGTTCCCT  
CTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTTCATCTCTGGAGCCCC  
TGCTTCTCTGGGGCTGTGCCGGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACACG  
TGAGACTGCGTATCTTCGCTGCCCGCATCTATGATTACGACCCCCTATATAAGGAGGCA  
CTGCAAATGCTGCGGGATGCTGGGGCCCAAGTCTCCATCATGACCTACGATGAATTTA  
AGCACTGCTGGGACACCTTTGTGGACCACCAGGGATGTCCCTTCCAGCCCTGGGATGG  
ACTAGATGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAG  
GGAAACGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGG  
AAGTGGAGGAAGTaaagcttgacaagaagtagcagcatcggcctggccatcggcaccactctgtgggctgggcccgtgatcaccga  
cgagtacaaggtgccagcaagaattcaaggtgctgggcaacaccgaccggcacagcatcaagaagaacctgatcggagcccctgctgtc  
gacagcggcgaaacagccgagccaccggctgaagagaaccgccagaagaagatacaccagacggaagaaccggatctgctatctgca  
agagatcttcagcaacagatggccaaggtggacgacagcttctccacagactggaagagtccttctggtggaagaggataagaagcacg  
agcggcaccccatcttcggcaacatcgtggacgaggtggcctaccacgagaagtacccaccatctaccacctgagaaagaaactggtggac  
agcaccgacaaggccgacctgcccgtgatctatctggccctggcccacatgatcaagttcggggccacttctgatcagggcgacctgaac  
cccgacaacagcagctgggacaagctgttcatccagctggtgcagacctacaaccagctgttcgaggaaaaccccatcaacgccagcggcgt  
ggacgccaaggccatctgtctgccagactgagcaagagcagacggctggaaaatctgatcggcagctgccggcgagaagaagaatgg  
cctgttcggaacctgattgcctgagcctgggctgacccccaaactcaagagcaacttcgacctggccgaggatgccaactgcagctgag  
caaggacacctacgacgacacctggacaacctgctggcccagatcggcgaccagtagccgacctgtttctggccccaagaacctgtccg  
acgcatctgctgagcagatcctgagagtgaaaccagagatcaccagggccccctgagcgcctctatgatcaagagatacagcagcagc  
caccaggacctgacctgctgaaagctctctgctggcagcagctgcctgagaagtacaagagatttctcgaccagagcaagaacggctac  
gcccgtacattgacggcgagccagccaggaaggttctacaagttcatcaagccatcctggaaaagatggacggcaccgaggaactgct  
cgtgaagctgaacagagaggacctgctgcccgaagcagcggaccttcgacaacggcagcatccccaccagatccacctgggagagctgca  
cgccattctgcccggcaggaagatcttaccattctgaaggacaaccgggaaaagatcgagaagatcctgaccttccgcatccctactac  
gtgggcccctctggccaggggaaacagcagattcgcctggatgaccagaaagagcaggaacatcaccctggaacttcaggaagtg  
tggacaagggcgcttccgcccagactcagcagcggatgaccaactcgataagaacctgccaacgagaaggtgctgccaagcagcagc  
ctgctgtacgagtacttaccgtgtataacgagctgaccaaagtgaatacgtgaccgagggatgagaagccccgcttctgagcggcgag  
cagaaaaggccatcgtggacctgctgttcaagaccaaccggaaaagtgacctgaaagcagctgaaagaggacttcaagaaaatcgagt  
cttcgactccgtggaaatcctggcgtggaagatcggttcaacgcctcctgggcacataccagatctgctgaaaattatcaaggacaaggac  
ttcctggacaatgaggaaaacgaggacattctggaagatctgctgacctgacctgtttgaggacagagatgatcaggaacggcgtg  
aaaacctatgccacctgttcgacgacaaagtgatgaagcagctgaagcggcggagatacaccggctggggcaggctgagccggaaagctga  
tcaacggcatcgggacaagcagtcggcaagacaatcctggatttctgaagtcgacggcttcgccaacagaaacttcatgagctgatcc  
acgacgacagcctgacctttaaaggacatccagaaaagcccaggtgtccggccagggcgatagcctgcagcagcacttccaatctggcc  
ggcagccccgccattaagaaggcctcctgcagacagtgaaggtggtggacgagctcgtgaaagtgatggcggcggcacaagcccgagaac  
atcgtgatcgaatggccagagagaaccagaccaccagaaggacagaagaacagccgcgagagaatgaagcggatcgaagagggcat  
caagagctgggagccagatcctgaaagaacccccgtggaacacccagctgcagaacgagaagctgtacctgtactacctgcagaat  
ggcggggatgtacgtggaccaggaactggacatcaaccggctgtccgactacgatgtggacgctatcgtgcctcagagcttctgaaggac  
gactccatcgacaacaaggtgctgaccagaagcgacaagaaccggggcaagagcgacaacgtgcctcctgaaagagctcgtgaagaagatg  
aagaactactggcggcagctgctgaacgccaagctgattaccagagaaagttcacaatctgaccaaggccgagagagggcctgagcg

aactggataaggccggcttcatcaagagacagctggtggaaccggcgagatcaaaagcacgtggcacagatcctggactcccggatgaa  
 cactaagtagcagcagaatgacaagctgatccgggaagtgaagtgtaccctgaagtccaagctggtgtccgattccggaaggattccag  
 tttacaaagtgcgcgagatcaacaactaccaccacgcccacgacgcctacctgaacgccgtcgtgggaaccgacctgatcaaaaagtacct  
 aagctggaaagcgagttcgtgtacggcgactacaaggtgtacgacgtgcggaagatgatcgccaagagcagcaggaaatcggcaaggcta  
 ccgccaagtacttctctacagcaacatcatgaacttttcaagaccgagattaccctggccaacggcgagatccggaagcggcctctgatcgag  
 acaaacggcgaaaccggggagatcgtgtgggataagggccgggatttggcaccgtgcggaaagtgtgagcatgccccagtgaatatcgt  
 gaaaagaccgaggtgcagacaggcggctcagcaagagcttatctgccaaagaggaaacagcgataagctgatcgccagaaaaggga  
 ctgggaccctaagaagtacggcgcttcgacagccccaccgtggcctattctgtgctggtgggccaagtggaaaagggaagtccaaga  
 aactgaagagtgtgaaagagctgctggggatcaccatcatgaaagaagcagcttcgagaagaatccatcgacttctggaagccaagggt  
 acaaagaagtgaaaaaggacctgatcatcaagctgcctaagtactccctgttcgagctggaaaacggccggaagagaatgctggcctctgccg  
 gcgaactgcagaagggaacgaactggcctgcctccaaatgtgaacttctgtacctggccagccactatgagaagctgaagggtccc  
 ccgaggataatgagcagaaacagctgtttgtggaacagcacaagcactacctggacgagatcatcgagcagatcagcgagttctcaagaga  
 gtgatctggccgacgctaacttgacaaaagtgtctccgcctacaacaagcaccgggataagcccatcagagagcaggccgagaatatcat  
 ccacctgtttacctgaccaatctgggagccccctgcccttcaagtaactttgacaccacctgacccggaagaggtacaccagcaccaaaga  
 ggtgctggacgccaccctgatccaccagatcaccggcctgtacgagacacggatcgacctgtctcagctgggagggcactaactcgag

Seq ID NO303:

>6his-NLS-A3A-GGS14-dCas9 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtccgcgcggcagccatatgcaaagaagaagcggaaaggtcGAA  
 GCCAGCCCAGCATCCGGGCCAGACACTTGATGGATCCACACATATTCACCTTCCAACCT  
 TAACAATGGCATTGGAAGGCATAAGACCTACCTGTGCTACGAAGTGGAGCGCCTGGAC  
 AATGGCACCTCGGTCAAGATGGACCAGCACAGGGGCTTTCTACACAACCAGGCTAAGA  
 ATCTTCTCTGTGGCTTTTACGGCCGCCATGCGCAGCTGCGCTTCTTGGACCTGGTTCCTT  
 CTTTGCAGTTGGACCCGGCCAGATCTACAGGGTCACTTGGTTCATCTCCTGGAGCCCC  
 TGCTTCTCCTGGGGCTGTGCCGGGGAAGTGCCTGCGTTCCTTCAGGAGAACACACAG  
 TGAGACTGCGTATCTTCGCTGCCGCATCTATGATTACGACCCCTATATAAGGAGGCA  
 CTGCAAATGCTGCGGGATGCTGGGGCCC AAGTCTCCATCATGACCTACGATGAATTTA  
 AGCACTGCTGGGACACCTTTGTGGACCACCAGGGATGTCCCTTCCAGCCCTGGGATGG  
 ACTAGATGAGCACAGCCAAGCCCTGAGTGGGAGGCTGCGGGCCATTCTCCAGAATCAG  
 GGAAACGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGG  
 AAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGA  
 GTGGAGGAAGTGGAGGAAGTaaagcttgacaagaagtacagcatcggcctggccatcggcaccaactctgtgggctgggccc  
 gtgatcaccgacgagtagcaaggtgcccagcaagaattcaaggtgctgggcaacaccgaccggcacagcatcaagaagaacctgatcggag  
 ccctgctgttcgacagcggcgaaacagccgagggccaccggctgaagagaaccgccagaagaagatacaccagacggaagaaccggatct  
 gctatctgcaagagatctcagcaacagatggccaaggtggacgacagcttctccacagactggaagagtcttctcgtggaagaggataa  
 gaagcagcagcggcaccatcttcggcaacatcgtggacgaggtggcctaccacgagaagtacccaccatctaccactgagaagaaa  
 ctggtggacagcaccgacaaggccgacctgcggctgatctatctggccctggccacatgatcaagttccggggccacttctgatcagggg  
 gacctgaaccccgacaacagcgacgtggacaagctgttcatccagctggtgcagacctacaaccagctgttcgaggaaaaccccatcaacgc  
 cagcggcgtggacgccaagccatctgtctgcccagactgagcaagagcagcggctggaaaatctgatcggccagctgcccggcgagaa  
 gaagaatggcctgttcggaaacctgattgcctgagcctggcctgacccccaaactcaagagcaacttcgacctggccgaggtgccaact  
 gcagctgagcaaggacacctacgacgacacctggacaacctgctggcccagatcggcgaccagtacgccgacctgttctggccgccaag  
 aacctgtccgacgccatcctgctgagcgacatcctgagagtgaacaccgagatcaccaaggccccctgagcgcctctatgatcaagagatac  
 gacgagcaccaccaggacctgacctgctgaaagctctcgtgcggcagcagctgctgagaagtacaaagagatttcttcgaccagagcaa  
 gaacggctacgccggctacattgacggcggagccagccaggaagagttctacaagttcatcaagcccactctggaaaagatggacggcacc  
 gaggaactgctcgtgaagctgaacagagaggacctgctgcggaagcagcggaccttcgacaacggcagcatccccaccagatccacctgg  
 gagagctgcacgccattctgcggcggcaggaagattttaccattcctgaaggacaaccgggaaaagatcgagaagatcctgacctccgca  
 tcccctactacgtgggcccctggccaggggaaacagcagattcgctggatgaccagaaagagcaggaaacatccccctggaactc  
 gaggaagtgtggacaagggcgcttccgcccagagcttcatcgagcggatgaccaacttcgataagaacctgcccacagagaaggtgctgc  
 ccaagcacagcctgctgtacgagtctaccgtgtataacgagctgaccaaagtgaatacgtgaccgaggggaatgagaagccccgcttcc  
 tgagcggcgagcagaaaaggccatcgtggacctgctgttcaagaccaaccggaagtgacctgaaagcagctgaaagaggactactca



accgtacgccgacctgtttctggccgccaagaacctgtccgacgccatcctgctgagcgacatcctgagagtgaacaccgagatcaccaagg  
ccccctgagcgctctatgatcaagagatacgcagcaccaccaggacctgacctgctgaaagctctctgctggcgagcagctgcctgag  
aagtacaaagagattttctgaccagagcaagaacggctacgccggctacattgacggcggagccagccaggaagagtctacaagttcatc  
aagcccacctggaagatggacggcaccgaggaactgctcgtgaaagctgaacagagaggacctgctgcggaagcagcggaccttcgac  
aacggcagcatccccaccagatccacctgggagagctgcacgccattctgcggcggcaggaagattttaccattcctgaaggacaaccg  
ggaaaagatcgagaagatcctgacctccgcacccctactactctggggcccttggccaggggaaacagcagattcgcttgatgaccagaa  
agagcggagaaacctacccccctggaacttcgaggaagtggggacaaggcgtccgcccagagcttcatcgagcggatgaccaacttc  
gataagaacctgccaacgagaaggtgctgccaagcacagcctgctgtacgagtacttaccctgtataacgagctgaccaaagtgaatac  
gtgaccgagggaatgagaagccccgcttctgagcggcgagcagaaaaaggccatcgtggacctgctgttcaagaccaaccggaaaagtga  
cctggaagcagctgaaagaggactacttcaagaaaatcagtgcttcgactccgtgaaatctccggcgtggaagatcggttcaacgcctccct  
gggcacataccacgatctgctgaaaattatcaaggacaaggacttctggacaatgaggaacgaggacattctggaagatcgtgctgacc  
ctgacactgtttgaggacagagagatgatcgaggaacggctgaaaacctatgccacctgttcgacgacaaaagtgtgaagcagctgaagcg  
gaggagatacaccggctggggcaggctgagccggaagctgatcaacggcatccgggacaagcagctccggcaagacaatcctggatttctg  
aagtcggacggcttcgccaacagaaacttcatgcagctgatccacgacgacagcctgacctttaaaggagcatccagaaagcccaggtgtcc  
ggccagggcgatagcctgcacgagcaccattgccaatctggcggcagccccgcattaagaaggccatcctgcagacagtgaaagtggtgg  
acgagctcgtgaaagtgtgggcccgcacaagcccgagaacatcgtgatcgaatggccagagagaaccagaccaccagaaggagacag  
aagaacagccgcgagagaatgaagcggatcgaagaggcgcataaagagctgggcagccagatcctgaaagaacccccctggaaaacac  
ccagctgcagaacgagaagctgtacctgtactacctgagaatggcggggatgtactgctggaccaggaactggacatcaaccggctgtccg  
actacgatgtggacgctatcgtgcctcagagctttctgaaggacgactccatcgacaacaagtgctgaccagaagcgacaagaaccggggc  
aagagcgacaacgtgccctccgaagaggtcgtgaagaagatgaagaactactggcggcagctgctgaacgccaagctgattaccagagaa  
agttcgacaatctgaccaaggccgagagagcggcctgagcgaactggataaggccggcttcatcaagagacagctggtgaaaccgggca  
gatcacaagcacgtggcacagatcctggactcccggatgaactaagtacgacgagaatgacaagctgatccgggaaagtgaagtgatca  
ccctgaagtccaagctggtgtccgatttccggaaggatttccagttttacaaagtgcgagatcaacaactaccaccacgcccacgacgccta  
cctgaacgccgtcgtgggaaccgcccctgatcaaaaagtacctaagctggaaagcagttcgtgtacggcgactacaaggtgtacgacgtgc  
ggaagatgatcgccaagagcagcaggaatcggcaaggctaccgccaagtacttctctacagcaacatcatgaacttttcaagaccgagat  
taccctggccaacggcgagatccggaagcggcctctgatcgagacaacggcgaaaccggggagatcgtgtgggataaggccgggattt  
gccaccgtgcggaagtgctgagcatgccccagtgaatctgtgaaaaagaccgaggtgcagacaggcggcttcagcaaaagctctatcct  
gccccagaggaacagcgataagctgatcgccagaaagaaggactgggaccctaagaagtacggcggcttcgacagccccaccgtggcctat  
tctgtgctggtggtggccaaagtggaaaagggaagtcacaagaactgaagagtgtaagagctgctggggatcaccatcatgaaagaag  
cagttcgagaagaatccatcactttctggaagccaagggtacaaagaagtgaaaaaggacctgatcatcaagctgcctaagtactccctg  
ttcagctggaaaacggccggaagagaatgctggcctctgcccggcaactgcagaagggaaacgaactggcctgcctccaaatgatgaa  
cttctgtacctggccagccactatgagaagctgaagggtcctcccgaggataatgagcagaacagctgtttgtggaacagcacaagcacta  
cctggacgagatcatcgacagatcagcaggttctcaagagagtgatcctggccagcctaacttggaacaaagtgtgtccgctacaacaa  
gcaccgggataagccatcagagagcaggccgagaatatcatccactgtttaccctgaccaatctgggagcccctgcccctcaagtacttt  
gacaccaccatcgaccggaagaggtacaccagcaccaaagaggtgctggacgccacctgatccaccagagcatcaccggcctgtacgag  
acacggatcgacctgtctcagctgggagggcgactaactcgag

Seq ID NO305:

>6his-NLS-A3H-GGS7-dCas9 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtgccgcgccgagccatgatgcaaagaagaagcggaaaggtcGCT  
CTTCTTACTGCTGAAACTTTTCGTCTCCAATTTAATAATAAACGCCGTCTGCGTCGCCCG  
TATTACCCGCGCAAGGCGCTGCTGTGTTACCAACTGACCCACAAAACGGTTCCACCCC  
GACTCGCGGTTACTTTGAGAATAAGAAAAAATGTCACGCTGAGATCTGTTTCATTAAC  
GAAATCAAATCTATGGGCCTGGATGAACTCAGTGCTACCAGGTCACCTGCTACCTGA  
CCTGGAGCCCGTGTAGCTCTTGCGCGTGGGAACTGGTTGACTTCATCAAAGCGCACGA  
CCATCTGAACCTGCGTATCTTCGCTTCCCGCCTGTACTATCACTGGTGCAAGCCGCAAC  
AGGATGGCCTGCGCCTGCTGTGTGGTTCTCAGGTTCCGGTTGAAGTTATGGGTTTCCCG  
GAGTTTGCGGACTGCTGGGAAAACTTTGTGACCATGAGAAGCCACTGTCCTTTAACCC  
GTATAAAATGCTGGAAGAGCTGGACAAAACTCTCGTGCTATCAAGCGCCGTCTGGAT



>6his-NLS-A3H-GGS14-dCas9 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtgccgcgaggcagccatgccaagaagaagcggaaagtcGCT  
CTTCTTACTGCTGAAACTTTTCGTCTCCAATTAATAATAAACGCCGTCTGCGTCGCCCG  
TATTACCCGCGCAAGGCGCTGCTGTGTTACCAACTGACCCACAAAACGGTTCCACCCC  
GACTCGCGGTTACTTTGAGAATAAGAAAAAATGTCACGCTGAGATCTGTTTCATTAAC  
GAAATCAAATCTATGGGCCTGGATGAAACTCAGTGCTACCAGGTCACCTGCTACCTGA  
CCTGGAGCCCCTGTAGCTCTTGC GCGTGGGAACTGGTTGACTTCATCAAAGCGCACGA  
CCATCTGAACCTGCGTATCTTCGCTTCCCGCCTGTACTACTGTTGCAAGCCGCAAC  
AGGATGGCCTGCGCCTGCTGTGTGGTTCTCAGGTTCCGGTTGAAGTTATGGGTTTCCCG  
GAGTTTGC GGACTGCTGGGAAAAC TTTGTTGACCATGAGAAGCCACTGTCCTTTAACCC  
GTATAAAATGCTGGAAGAGCTGGACAAAAC TCTCGTGCTATCAAGCGCCGTCTGGAT  
CGTATCAAGTCTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGG  
GGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGG  
AGGAAGTGGAGGAAGTGGAGGAAGTaaagcttgacaagaagtacagcatcggcctggccatcggcaccactctgtgg  
gctgggcccgtgatcaccgacgagtacaaggtgccagcaagaatcaaggtgctgggcaacaccgaccggcacagcatcaagaagaacct  
gatcggagcccctgctgttcgacagcggcgaacagccgagggcaccggctgaagagaaccgccaagaagaatataccagacgggaaga  
accggatctgctatctgcaagagatcttcagcaacgagatggccaaggtggacgacagcttcttcacagactggaagagctctcctggtgga  
agaggataagaagcagcagcggcaccatcttcggcaacatcgtggacgaggtggcctaccacgagaagtaccacccatctaccactg  
agaaagaactggtggacagcaccgacaaggccgacctgaggctgatctatctggccctggccacatgatcaagttccggggccacttctg  
atcgagggcgacctgaaccccgacaacagcgcagctggacaagctgttcacagctggtgcagacctacaaccagctgttcgaggaaaacc  
catcaacgccagcggcgtggacgccaaggccatctgtctgacagactgagcaagagcagacggctggaaaatctgatcggcagctgcc  
ggcgagaagaagaatggcctgttcgaaacctgattgccctgagcctggcctgaccccaactcaagagcaacttcgacctggccgagga  
tgccaaactgcagctgagcaaggacacctacgacgacacctggacaacctgctggcccagatcggcgaccagtagccgacctgtttctg  
ccgccaagaacctgtccgacgccatctgtgagcagatcctgagagtgaacaccgagatcaccaggccccctgagcgcctctatgatc  
aagagatacagcagcaccaccaggacctgacctgctgaaagctctcgtgcggcagcagctgctgagaagtaacaagagattttcttcac  
cagagcaagaacggctacgccggctacattgacggcggagccagccaggaagagttctacaagttcatcaagccatctggaaaagatgg  
acggcaccgaggaactgctcgtgaagctgaacagagaggacctgctgaggaaagcagcggaccttcgacaacggcagcatccccaccaga  
tccacctgggagagctgcacgccattctgcggcggcaggaagattttaccattcctgaaggacaaccgggaaaagatcgagaagatcctga  
ccttccgcatcccactactcgtggccctctggccaggggaaacagcagattcgcctggatgaccagaaagagcagggaaacctaccccc  
tggaacttcgaggaagtgggtggacaagggccttcgcccagagcttcacgagcggatgaccaacttcgataagaacctgcccacagaa  
ggtgctgcccgaacgacagcctgctgtacgagtagcttaccctgtataacgagctgaccaaagtgaatacgtgaccgaggggaatgagaaagcc  
cgcttctgagcggcgagcagaaaaaggccatcgtggacctgctgttcaagaccaaccggaaagtgacctgaaagcagctgaaagaggac  
tacttcaagaaaatcgagtcttcgactcctggaaatctccggcgtggaagatcggttcaacgcctcctgggacataccacgatctgctgaa  
aattatcaaggacaaggacttctggacaatgaggaaaaagggacattctggaagatctgctgacctgacctgtttgaggacagagag  
atgatcgaggaacggctgaaaacctatgccacctgttcgacgacaagtgatgaagcagctgaagcggcggagatacaccggctggggca  
ggctgagccggaagctgatcaacggcatccgggacaagcagctccggcaagacaatcctggatttctgaagtccgacggcttcgccaacaga  
aacttcatgcagctgatccacgacgacacctgacctttaaagaggacatccagaaagcccaggtgtccggccagggcgatagcctgcacga  
gcacattgccaatctggccggcagccccgccattaagaagggcaccctgcagacagtgaaaggtggtggacgagctcgtgaaagtgatgggc  
ggcacaagcccagaaacatcgtgatcgaatggccagagagaaccagaccaccagaagggacagaagaacagccgagagagaatgaa  
gcggatcgaagagggcatcaagagctgggcagccagatcctgaaagaacaccctggaaaacaccagctgcagaacgagaagctgta  
cctgtactacctgcagaatggcgggatgtacgtggaccaggaactggacatcaaccggctgtccgactacgatgtggacgctatcgtgct  
cagagctttctgaaggacgactccatcgaacaaggtgctgaccagaagcgacaagaaccggggcaagagcagacaactgacctccgaa  
gaggtcgtgaagaagatgaagaactactggcggcagctgctgaacgcaagctgattaccagagaaagttcgacaatctgaccaaggccga  
gagagggcggcctgagcgaactggataagggcggcttcatcaagagacagctggtggaaccggcagatcacaagcagctggcacagat  
cctggactcccggatgaacactaagtagcagagaatgacaagctgatccgggaagtgaagtgatcacctgaagtccaagctggtgtccga  
ttccggaaggatttcagtttacaaagtgcgcgagatcaacaactaccaccacgcccacgacgcctacctgaacggctcgtggaaaccgccc  
ctgatcaaaaagtaccctaagctggaaagcagctcgtgtacggcgactacaaggtgtacgagctgcggaaagatgatcccaagagcagca  
ggaaatcggcaaggctaccgcaagcttcttacagcaacatcatgaacttttcaagaccgagattacctggccaacggcgagatccgg  
aagcggcctctgatcagacaaacggcgaacccggggagatcgtgtgggataagggcgggatttggccacctgctggaaagtctgagca  
tgcccaagtgaatatcgtgaaaagaccgaggtgcagacagcggcttcagcaaagagctatcctgcccagaggaacagcagataagctg  
atcgccagaaagaagactgggaccctaagaagtagcggcggcttcgacagccccaccctggcctattctgtctggtggtggccaaagtgga  
aaagggcaagtccaagaactgaagagtgtaagagagctgctggggatcacatcatgaaagaagcagcttcgagaagaatccatcgact

ttctggaagccaagggctacaaagaagtgaaaaaggacctgatcatcaagctgcctaagtactcctgttcgagctggaaaacggccggaaga  
gaatgctggcctctgccggcgaactgcagaagggaaacgaactggccctgccctccaaatgtgaacttctgtacctggccagccactatga  
gaagctgaagggctccccgaggataatgagcagaaacagctgtttgtggaacagcacaagcactacctggacgagatcatcgagcagatca  
gcgagttctccaagagagtgatcctggccgacgctaactctggacaaaagtgtctgccctacaacaagcaccgggataagcccatcagagag  
caggccgagaatatcatccacctgtttacctgaccaatctgggagccctgccgcctcaagtactttgacaccaccatcgaccggaagagg  
acaccagcaccaaagaggtgctggacgccaccctgatccaccagagcatcaccggcctgtacgagacacggatcgacctgtctcagctggg  
aggcgactaactcgag

Seq ID NO307:

>6his-NLS-A3H-GGS7-dCpf1 基因序列

ATGggcagcagccatcatcatcatcacagcagcggcctggtgccgcggcagccatgccaagaagaagcgggaaggtcGCT  
CTTCTTACTGCTGAAACTTTTCGTCTCCAATTTAATAATAAACGCCGTCTGCGTCGCCCC  
TATTACCCGCGCAAGGCGCTGCTGTGTTACCAACTGACCCACAAAACGGTTCCACCCC  
GACTCGCGGTTACTTTGAGAATAAGAAAAAATGTCACGCTGAGATCTGTTTCATTAAC  
GAAATCAAATCTATGGGCCTGGATGAAACTCAGTGCTACCAGGTCACCTGCTACCTGA  
CCTGGAGCCCGTGTAGCTCTTGCGCGTGGGAACTGGTTGACTTCATCAAAGCGCACGA  
CCATCTGAACCTGCGTATCTTCGCTTCCCGCCTGTACTACTGTTGCAAGCCGCAAC  
AGGATGGCCTGCGCCTGCTGTGTGGTTCTCAGGTTCCGGTTGAAGTTATGGGTTTCCCG  
GAGTTTGCGGACTGCTGGGAAAACCTTTGTTGACCATGAGAAGCCACTGTCCTTTAACCC  
GTATAAAATGCTGGAAGAGCTGGACAAAACCTCTCGTGCTATCAAGCGCCGTCTGGAT  
CGTATCAAGTCTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGTGGAGGAAGT  
GGAGGAAGTGGAGGAAGTATGACACAGTTCGAGGGCTTTACCAACCTGTATCAGGTGA  
GCAAGACACTGCGGTTTGAGCTGATCCACAGGGCAAGACCCTGAAGCACATCCAGGA  
GCAGGGCTTCATCGAGGAGGACAAGGCCCGCAATGATCACTACAAGGAGCTGAAGCC  
CATCATCGATCGGATCTACAAGACCTATGCCGACCAGTGCCTGCAGCTGGTGCAGCTG  
GATTGGGAGAACCTGAGCGCCGCCATCGACTCCTATAGAAAGGAGAAAACCGAGGAG  
ACAAGGAACGCCCTGATCGAGGAGCAGGCCACATATCGCAATGCCATCCACGACTACT  
TCATCGGCCGGACAGACAACCTGACCGATGCCATCAATAAGAGACACGCCGAGATCTA  
CAAGGGCCTGTTCAAGGCCGAGCTGTTTAATGGCAAGGTGCTGAAGCAGCTGGGCACC  
GTGACCACAACCGAGCACGAGAACGCCCTGCTGCGGAGCTTCGACAAGTTTACAACCT  
ACTTCTCCGGCTTTTATGAGAACAGGAAGAACGTGTTTACGCGCCGAGGATATCAGCAC  
AGCCATCCCACACCGCATCGTGCAGGACAACCTCCCCAAGTTTAAGGAGAATTGTCAC  
ATCTTACACGCCTGATCACCGCCGTGCCAGCCTGCGGGAGCACTTTGAGAACGTGA  
AGAAGGCCATCGGCATCTTCGTGAGCACCTCCATCGAGGAGGTGTTTTCTTCCCTTTT  
TATAACCAGCTGCTGACACAGACCCAGATCGACCTGTATAACCAGCTGCTGGGAGGAA  
TCTCTCGGGAGGCAGGCACCGAGAAGATCAAGGGCCTGAACGAGGTGCTGAATCTGGC  
CATCCAGAAGAATGATGAGACAGCCCACATCATCGCCTCCCTGCCACACAGATTCATC  
CCCCTGTTTAAGCAGATCCTGTCCGATAGGAACACCCTGTCTTTCATCCTGGAGGAGTT  
TAAGAGCGACGAGGAAGTGATCCAGTCCTTCTGCAAGTACAAGACACTGCTGAGAAAC  
GAGAACGTGCTGGAGACAGCCGAGGCCCTGTTTAACGAGCTGAACAGCATCGACCTGA  
CACACATCTTCATCAGCCACAAGAAGCTGGAGACAATCAGCAGCGCCCTGTGCGACCA  
CTGGGATACACTGAGGAATGCCCTGTATGAGCGGAGAATCTCCGAGCTGACAGGCAAG  
ATCACCAAGTCTGCCAAGGAGAAGGTGCAGCGCAGCCTGAAGCACGAGGATATCAAC  
CTGCAGGAGATCATCTCTGCCGACAGGCAAGGAGCTGAGCGAGGCCTTCAAGCAGAAA  
ACCAGCGAGATCCTGTCCCACGCACACGCCGCCCTGGATCAGCCACTGCCTACAACCC  
TGAAGAAGCAGGAGGAGAAGGAGATCCTGAAGTCTCAGCTGGACAGCCTGCTGGGCC  
TGTACCACCTGCTGGACTGGTTTGCCGTGGATGAGTCCAACGAGGTGGACCCCGAGTT  
CTCTGCCCGGCTGACCGGCATCAAGCTGGAGATGGAGCCTTCTCTGAGCTTCTACAACA

AGGCCAGAAATTATGCCACCAAGAAGCCCTACTCCGTGGAGAAGTTCAAGCTGAACTT  
TCAGATGCCTACACTGGCCTCTGGCTGGGACGTGAATAAAGGAGAAGAACAATGGCGCC  
ATCCTGTTTTGTGAAGAACGGCCTGTACTATCTGGGCATCATGCCAAAGCAGAAGGGCA  
GGTATAAGGCCCTGAGCTTCGAGCCCACAGAGAAAACCAGCGAGGGCTTTGATAAGAT  
GTAATATGACTACTTCCCTGATGCCGCCAAGATGATCCCAAAGTGCAGCACCCAGCTG  
AAGGCCGTGACAGCCCCTTTCAGACCCACACAACCCCCATCCTGCTGTCCAACAATTT  
CATCGAGCCTCTGGAGATCACAAAGGAGATCTACGACCTGAACAATCCTGAGAAGGAG  
CCAAAGAAGTTTCAGACAGCCTACGCCAAGAAAACCGGCGACCAGAAGGGCTACAGA  
GAGGCCCTGTGCAAGTGGATCGACTTCACAAGGGATTTTCTGTCCAAGTATACCAAGA  
CAACCTCTATCGATCTGTCTAGCCTGCGGCCATCCTCTCAGTATAAGGACCTGGGCGAG  
TACTATGCCGAGCTGAATCCCCTGCTGTACCACATCAGCTTCCAGAGAATCGCCGAGA  
AGGAGATCATGGATGCCGTGGAGACAGGCAAGCTGTACCTGTTCCAGATCTATAACAA  
GGACTTTGCCAAGGGCCACCACGGCAAGCCTAATCTGCACACACTGTATTGGACCGGC  
CTGTTTTCTCCAGAGAACCTGGCCAAGACAAGCATCAAGCTGAATGGCCAGGCCGAGC  
TGTTCTACCGCCCTAAGTCCAGGATGAAGAGGATGGCACACCGGCTGGGAGAGAAGAT  
GCTGAACAAGAAGCTGAAGGATCAGAAAACCCAATCCCCGACACCCTGTACCAGGA  
GCTGTACGACTATGTGAATCACAGACTGTCCCACGACCTGTCTGATGAGGCCAGGGCC  
CTGCTGCCAACGTGATCACCAAGGAGGTGTCTCACGAGATCATCAAGGATAGGCGCT  
TTACCAGCGACAAGTTCTTTTTCCACGTGCCTATCACACTGAACTATCAGGCCGCCAAT  
TCCCCATCTAAGTTCAACCAGAGGGTGAATGCCTACCTGAAGGAGCACCCCGAGACAC  
CTATCATCGGCATCGATCGGGGCGAGAGAAACCTGATCTATATCACAGTGATCGACTC  
CACCGCAAGATCCTGGAGCAGCGGAGCCTGAACACCATCCAGCAGTTTGATTACCAG  
AAGAAGCTGGACAACAGGGAGAAGGAGAGGGTGGCAGCAAGGCAGGCCTGGTCTGTG  
GTGGGCACAATCAAGGATCTGAAGCAGGGCTATCTGAGCCAGGTCATCCACGAGATCG  
TGGACCTGATGATCCACTACCAGGCCGTGGTGGTGGTGGGAGAACCTGAATTTCCGGCTTT  
AAGAGCAAGAGGACCGGCATCGCCGAGAAGGCCGTGTACCAGCAGTTCGAGAAGATG  
CTGATCGATAAGCTGAATTGCCTGGTGTGCTGAAGGACTATCCAGCAGAGAAAGTGGGAG  
GCGTGCTGAACCCATAACCAGCTGACAGACCAGTTCACCTCCTTTGCCAAGATGGGCAC  
CCAGTCTGGCTTCCCTGTTTTACGTGCCTGCCCATATACATCTAAGATCGATCCCCTGA  
CCGGCTTCGTGGACCCCTTCGTGTGGAAAACCATCAAGAATCACGAGAGCCGCAAGCA  
CTTCCTGGAGGGCTTCGACTTTCTGCACTACGACGTGAAAACCGGCGACTTCATCCTGC  
ACTTTAAGATGAACAGAAATCTGTCCCTCCAGAGGGGCCCTGCCCGGCTTTATGCCTGCA  
TGGGATATCGTGTTCGAGAAGAACGAGACACAGTTTGACGCCAAGGGCACCCCTTTCA  
TCGCCGGCAAGAGAATCGTGCCAGTGATCGAGAATCACAGATTCACCGGCAGATACCG  
GGACCTGTATCCTGCCAACGAGCTGATCGCCCTGCTGGAGGAGAAGGGCATCGTGTTT  
AGGGATGGCTCCAACATCCTGCCAAAGCTGCTGGAGAATGACGATTCTCACGCCATCG  
ACACCATGGTGGCCCTGATCCGCAGCGTGCTGCAGATGCGGAACTCCAATGCCGCCAC  
AGGCGAGGACTATATCAACAGCCCCGTGCGCGATCTGAATGGCGTGTGCTTCGACTCC  
CGTTTTCAGAACCAGAGTGGCCATGGACGCCGATGCCAATGGCGCCTACCACATCG  
CCCTGAAGGGCCAGCTGCTGCTGAATCACCTGAAGGAGAGCAAGGATCTGAAGCTGCA  
GAACGGCATCTCCAATCAGGACTGGCTGGCCTACATCCAGGAGCTGCGCAACAAAAGG  
CCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAGGGATCCTACCATAC  
GATGTTCCAGATTACGCTTATCCCTACGACGTGCCTGATTATGCATACCCATATGATGT  
CCCCGACTATGCCTAAG

## 权 利 要 求 书

- 1、一种编辑目标核酸分子的方法，所述方法包括如下步骤：
  - (1) 获得编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体，所述融合蛋白 (A) 包括位于氮端的 Apobec 家族蛋白结构域和位于碳端的核酸酶活性失活的 Cas9 家族或 Cpf1 家族蛋白结构域，单链导向 RNA 与目标核酸分子的目标编辑区域具有互补区，其中目标核酸分子的目标编辑区域中包含至少一个甲基化的胞嘧啶核苷酸；
  - (2) 使步骤 (1) 获得的编码融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 的重组载体与目标核酸分子接触。
- 2、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于融合蛋白中位于氮端的 Apobec 家族蛋白选自人 Apobec3A 或 Apobec3H，或与人 Apobec3A 或 Apobec3H 具有 95%以上同源性的具有脱氨基活性的蛋白。
- 3、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于融合蛋白中位于碳端的核酸酶活性失活的 Cas9 蛋白的蛋白序列在第 10 和 840 位点的天冬氨酸和组氨酸突变为丙氨酸和丙氨酸的突变序列，或者位于碳端的核酸酶活性失去的 Cpf1 蛋白的蛋白序列在第 908 位点的天冬氨酸突变为丙氨酸的突变序列。
- 4、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于融合蛋白的两个结构域之间为 3-14 个基序组成的连接体。
- 5、根据权利要求 4 所述的编辑目标核酸分子的方法，其特征在于所述基序选自 (GGG)。
- 6、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于所述融合蛋白还包括纯化标签序列。
- 7、根据权利要求 1-6 任一项所述的编辑目标核酸分子的方法，其特征在于所述融合蛋白选自 SEQ ID No 201-207 的任一序列。
- 8、编码权利要求 7 所述蛋白序列的基因序列。
- 9、包含权利要求 8 任一基因序列的重组载体，所述载体选自质粒载体或病毒载体。
- 10、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于所述单链导向 RNA 长度为 60-80bp。
- 11、根据权利要求 1 所述的编辑目标核酸分子的方法，其特征在于所述单链导向 RNA 与目标核酸分子的互补区长度为 18-25bp。
- 12、一种体外编辑目标核酸分子的方法，所述方法包括如下步骤：
  - (1) 获得融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 或编码融合蛋白(A) 和单链

导向 RNA (sgRNA) (B) 的重组载体, 所述融合蛋白 (A) 包括位于氮端的 Apobec 家族蛋白结构域和位于碳端的核酸酶活性失活的 Cas9 家族蛋白结构域, 单链导向 RNA 与目标核酸分子的目标编辑区域具有互补区, 其中目标核酸分子的目标编辑区域中包含至少一个甲基化的胞嘧啶核苷酸;

(2) 将目标核酸分子与融合蛋白(A) 和单链导向 RNA (sgRNA) (B) 接触;

(3) 高温终止反应之后加入有效量的 TDG, 42°C反应 6-8 小时;

(4) 加入有效量的 EDTA、甲酰胺 和 NaOH 共同作用, 90-95°C反应 5-10 分钟。

- 13、根据权利要求 1 或 12 所述的编辑目标核酸分子的方法, 其特征在于所述甲基化的胞嘧啶核苷酸与癌症、基因紊乱性遗传疾病、发育错误等疾病相关。
- 14、权利要求 1 或 12 所述编辑目标核酸分子的方法在特异性调节基因组 DNA 甲基化/去甲基化状态中的应用。
- 15、权利要求 1 所述的编辑目标核酸分子的方法在治疗癌症、基因紊乱性遗传疾病、发育错误等疾病中的应用。

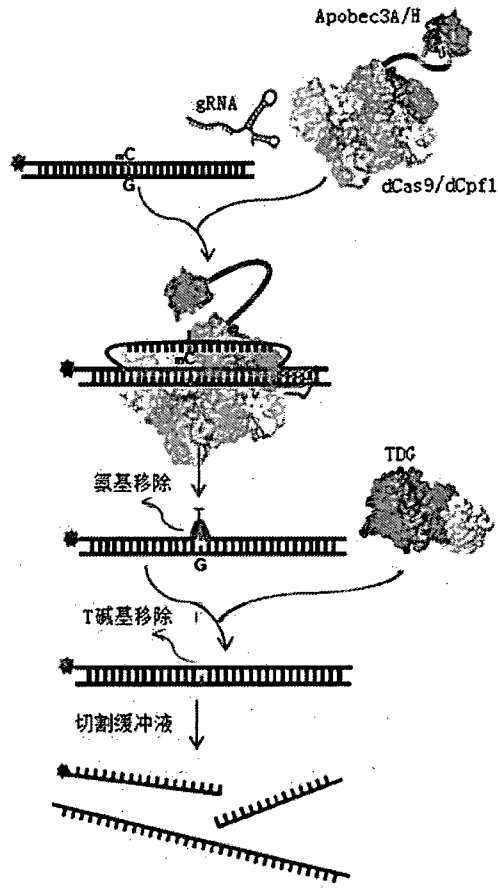


图 1

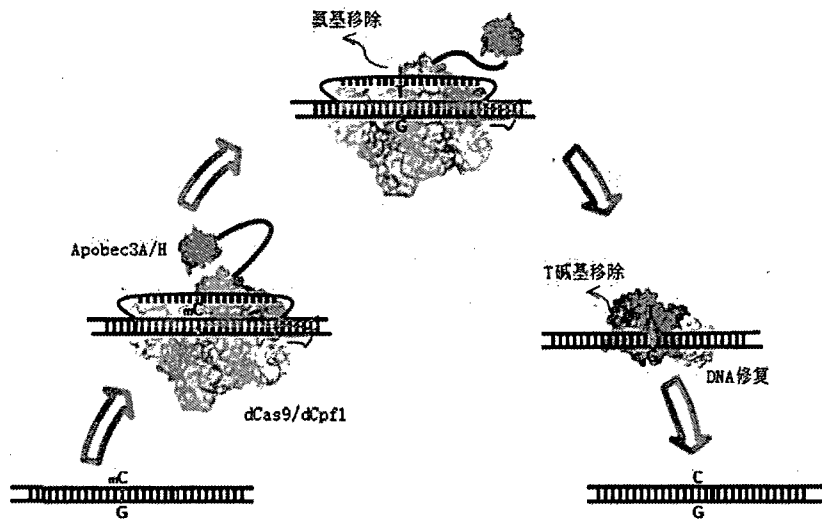


图 2

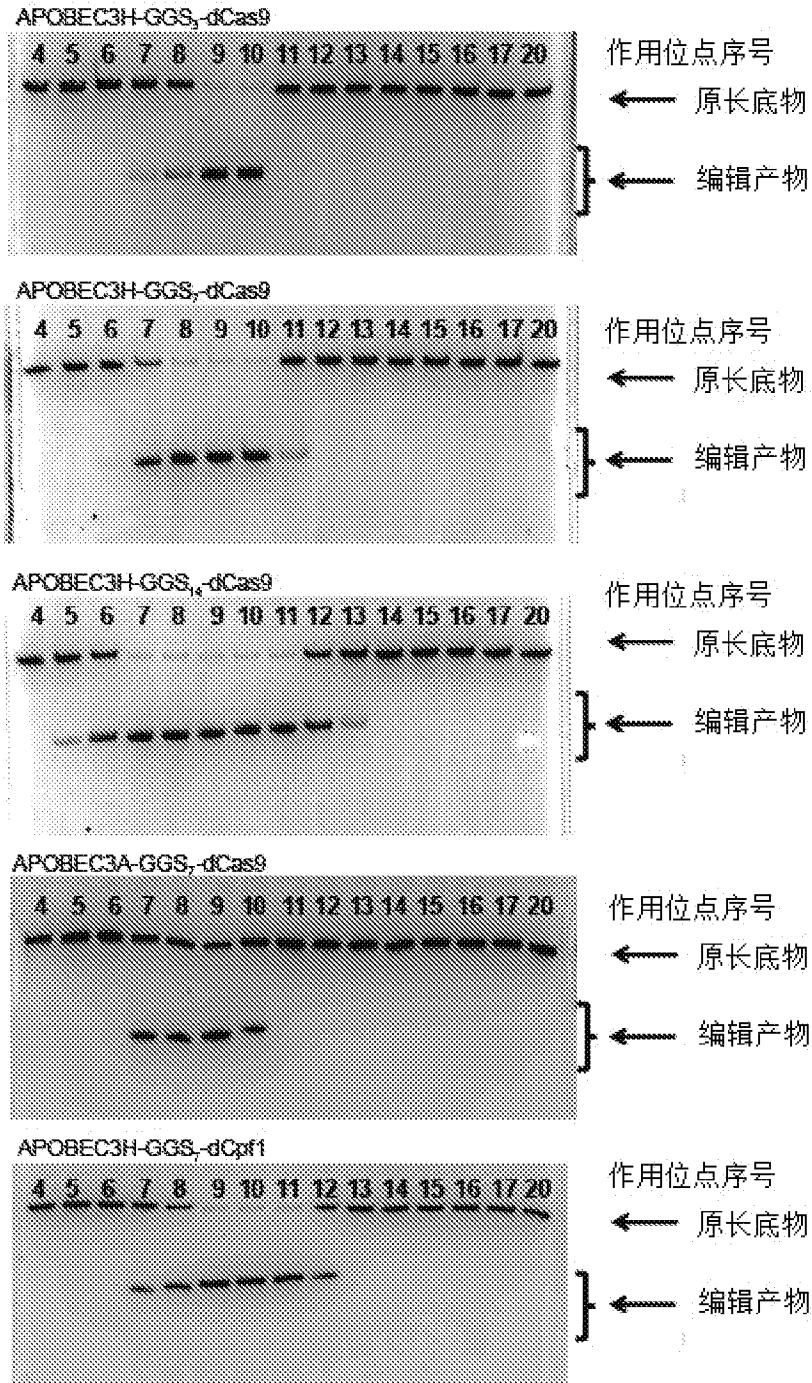


图 3

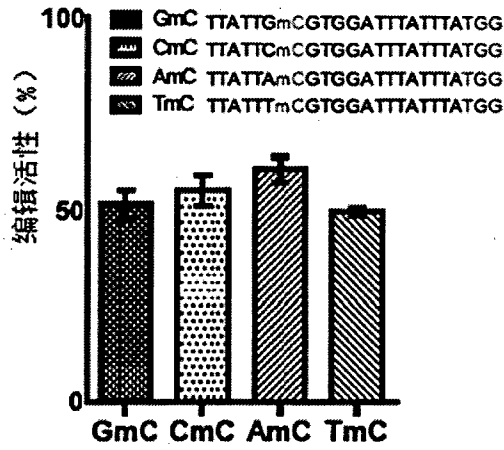


图 4

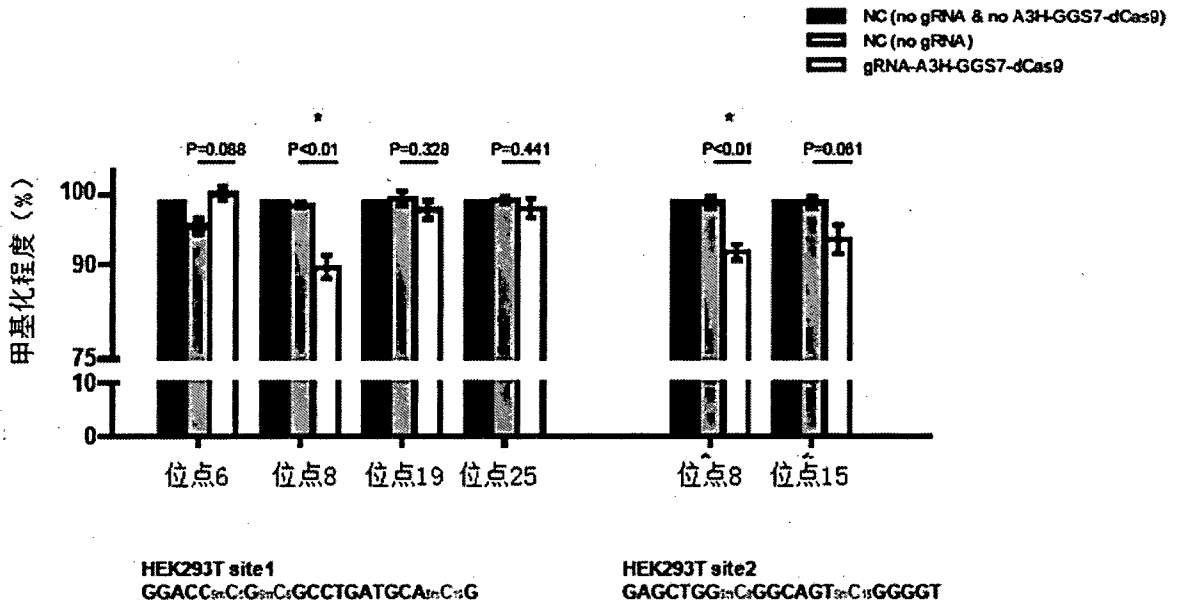
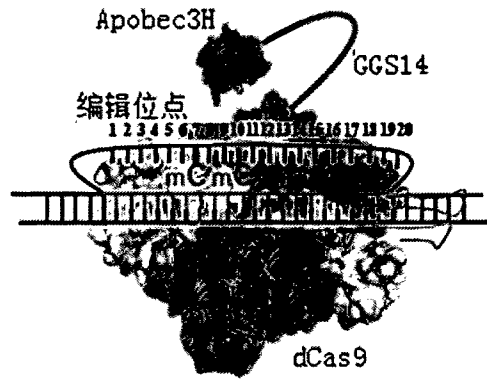
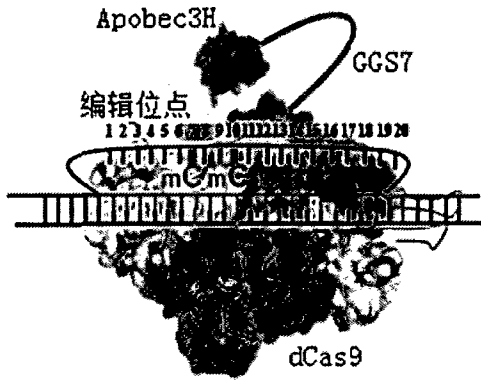
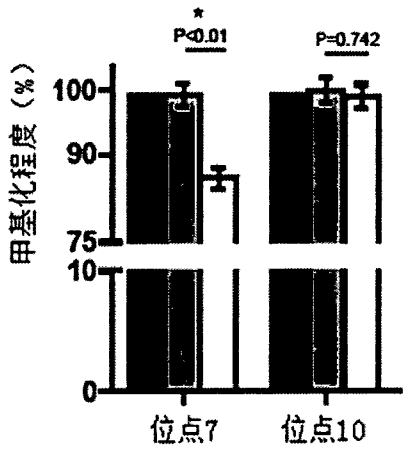


图 5



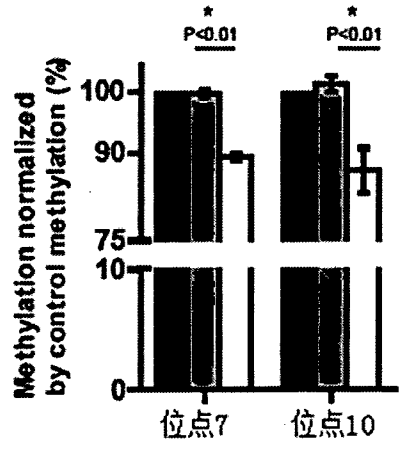
D

■ NC (no gRNA & no A3H-GGS7-dCas9)  
 ▨ NC (no gRNA)  
 ▩ gRNA-A3H-GGS7-dCas9



PC-3 site  
 TGGGGC<sub>5mC</sub>:GT<sub>5mC</sub>:GGGCCGGGCT

■ NC (no gRNA & no A3H-GGS14-dCas9)  
 ▨ NC (no gRNA)  
 ▩ gRNA-A3H-GGS14-dCas9



PC-3 site  
 TGGGGC<sub>5mC</sub>:GT<sub>5mC</sub>:GGGCCGGGCT

图 6

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2017/088281**

## A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/00 (2006.01) i; A61P 35/00 (2006.01) i; A61P 43/00 (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)

C12N; A61P

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; DWPI; SIPOABS; VEN; CNKI; CNTXT; WOTXT; EPTXT; USTXT; Pubmed; Genbank; Uniprot; Google;  
 SEQ ID NOs: 201-207, Cas enzyme, fusion protein, gene, nucleic acid, DNA, edit, deaminase, nuclease, nuclear localization signal,  
 NLS, linker, mutation, sgRNA, fusion, APOBEC, Cpfl, methylated cytosine, mC, demethylation, thymine DNA glycosylation, TDG,  
 cancer, tumor, genetic disorders, developmental error

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2015089406 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE), 18 June 2015 (18.06.2015), description, paragraphs 9-10, 24, 32, 35, 43-44, 46-48, 68-69, 127 and 140	1-15
Y	WO 2015133554 A1 (NATIONAL UNIVERSITY CORPORATION KOBE UNIVERSITY), 11 September 2015 (11.09.2015), abstract, claims 1-21, description, and figures 1-2 and 4-7	1-15
Y	WO 2016025131 A1 (E. I. DU PONT DE NEMOURS AND COMPANY), 18 February 2016 (18.02.2016), description, page 82, lines 29-33, and figure 3B	1-15
Y	CN 105112446 A (INSTITUTE OF BASIC MEDICAL SCIENCES OF CHINESE ACADEMY OF MEDICAL SCIENCES), 02 December 2015 (02.12.2015), description, paragraph 415	1-15
Y	VAZQUEZ-VILAR, M. et al., "A Modular Toolbox for Grna-Cas9 Genome Engineering in Plants Based on the Goldenbraid Standard", PLANT METHODS., vol. 12, no. 10, 01 February 2016 (01.02.2016), pages 1-12	1-15

Further documents are listed in the continuation of Box C.

See patent family annex.

<p>* Special categories of cited documents:</p> <p>"A" document defining the general state of the art which is not considered to be of particular relevance</p> <p>"E" earlier application or patent but published on or after the international filing date</p> <p>"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)</p> <p>"O" document referring to an oral disclosure, use, exhibition or other means</p> <p>"P" document published prior to the international filing date but later than the priority date claimed</p>	<p>"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</p> <p>"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</p> <p>"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</p> <p>"&amp;" document member of the same patent family</p>
---	---

<p>Date of the actual completion of the international search</p> <p style="text-align: center;">09 August 2017 (09.08.2017)</p>	<p>Date of mailing of the international search report</p> <p style="text-align: center;"><b>06 September 2017 (06.09.2017)</b></p>
<p>Name and mailing address of the ISA/CN:</p> <p>State Intellectual Property Office of the P. R. China                  No. 6, Xitucheng Road, Jimenqiao                  Haidian District, Beijing 100088, China                  Facsimile No.: (86-10) 62019451</p>	<p>Authorized officer</p> <p style="text-align: center;"><b>GUO, Tingting</b></p> <p>Telephone No.: (86-10) <b>62413879</b></p>

# INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2017/088281**

<b>C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	YIN, Xiaotong et al., "Dynamic Regulation of Mammalian DNA Methylation", vol. 34, no. 5, 31 October 2014 (31.10.2014), pages 604-613	1-15
A	CN 105408497 A (THE GENERAL HOSPITAL CORPORATION), 16 March 2016 (16.03.2016), the whole document	1-15

## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2017/088281**

### Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1.  Claims Nos.: 15  
because they relate to subject matter not required to be searched by this Authority, namely:  
[1] claims 15 relates to a method for treating a disease, and therefore does not comply with PCT Rule 39.1(iv). The following amendment is made in this report based on claim 15: application of the recombinant vector of claim 1 in the preparation of a medicament for treating cancer, genetic disorders, developmental error.
  
2.  Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
  
3.  Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/CN2017/088281**

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2015089406 A1	18 June 2015	US 2015166980 A1 US 9068179 B1 EP 3080265 A1 CN 105934516 A US 2016304846 A1 CA 2933625 A1 US 2015166985 A1 US 2015166984 A1 AU 2014362208 A1 US 2015165054 A1 US 2015166983 A1 JP 2017500035 A US 2015166982 A1 US 2015166981 A1	18 June 2015 30 June 2015 19 October 2016 07 September 2016 20 October 2016 18 June 2015 18 June 2015 18 June 2015 30 June 2016 18 June 2015 18 June 2015 05 January 2017 18 June 2015 18 June 2015
WO 2015133554 A1	11 September 2015	EP 3115457 A1 SG 11201609211V A CA 2947941 A1 US 2017073670 A1 JP WO2015133554 A1 CN 106459957 A	11 January 2017 29 December 2016 11 September 2015 16 March 2017 06 April 2017 22 February 2017
WO 2016025131 A1	18 February 2016	WO 2016025131 A8 CA 2957684 A1 KR 20170036792 A AU 2015302213 A1 EP 3180425 A1	08 June 2017 18 February 2016 03 April 2017 02 March 2017 21 June 2017
CN 105112446 A	02 December 2015	None	
CN 105408497 A	16 March 2016	WO 2014144592 A3 EP 2970208 A2 KR 20150131250 A CA 2906553 A1 CA 2907198 A1 AU 2014227653 A2 WO 2014144592 A2 EP 2971041 A1 EP 2970986 A2 EP 2970208 A4 AU 2014227653 A1 JP 2016512264 A US 2016024524 A1 EP 2970986 A4 CN 105247066 A US 2016024523 A1 EP 2971041 A4 JP 2016512691 A EP 2971125 A2 CN 105408483 A US 9567604 B2 WO 2014144288 A1 AU 2014227653 B2	31 December 2014 20 January 2016 24 November 2015 25 September 2014 18 September 2014 12 November 2015 18 September 2014 20 January 2016 20 January 2016 30 November 2016 01 October 2015 25 April 2016 28 January 2016 07 September 2016 13 January 2016 28 January 2016 07 September 2016 09 May 2016 20 January 2016 16 March 2016 14 February 2017 18 September 2014 20 April 2017

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/CN2017/088281**

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
		CA 2906724 A1	18 September 2014
		US 9567603 B2	14 February 2017
		JP 2016517276 A	16 June 2016
		EP 2971125 A4	24 August 2016
		KR 20150131251 A	24 November 2015
		US 2014295556 A1	02 October 2014
		WO 2014152432 A3	29 October 2015
		WO 2014144761 A3	29 October 2015
		US 2014295557 A1	02 October 2014
		WO 2014152432 A2	25 September 2014
		WO 2014144761 A2	18 September 2014
		AU 2014228981 A1	01 October 2015
		AU 2014239665 A1	01 October 2015
		US 2016010076 A1	14 January 2016
		KR 20150132395 A	25 November 2015
		US 2017152508 A1	01 June 2017

国际检索报告

国际申请号

PCT/CN2017/088281

<p><b>A. 主题的分类</b></p> <p>C12N 15/00(2006.01)i; A61P 35/00(2006.01)i; A61P 43/00(2006.01)i</p> <p>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</p>																				
<p><b>B. 检索领域</b></p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>C12N; A61P</p> <p>包含在检索领域中的除最低限度文献以外的检索文献</p> <p>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))</p> <p>CPRSABS; CNABS; DWPI; SIPOABS; VEN;CNKI; CNTXT; WOTXT; EPTXT; USTXT; Pubmed; Genbank; Uniprot; Google; SEQ ID NOs; 201-207, 基因, 核酸, 编辑, Cas酶, 脱氨酶, 核酸酶, 接头, 突变, 融合蛋白, 甲基化胞嘧啶, 去甲基化, 胸腺嘧啶DNA糖基化酶, 癌, 基因紊乱, 发育错误, gene, nucleic acid, DNA, edit, deaminase, nulease, nuclear localization signal, NLS, linker, mutation, sgRNA, fusion, APOBEC, Cpfl, methylated cytosine, mC, demethylation, thymine DNA glycosylation, TDG, cancer, tumor, genetic disorders, developmental error</p>																				
<p><b>C. 相关文件</b></p> <table border="1"> <thead> <tr> <th>类型*</th> <th>引用文件, 必要时, 指明相关段落</th> <th>相关的权利要求</th> </tr> </thead> <tbody> <tr> <td>Y</td> <td>WO 2015089406 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2015年 6月 18日 (2015 - 06 - 18) 说明书第9-10, 24, 32, 35, 43-44, 46-48, 68-69, 127, 140段</td> <td>1-15</td> </tr> <tr> <td>Y</td> <td>WO 2015133554 A1 (国立大学法人神户大学) 2015年 9月 11日 (2015 - 09 - 11) 摘要, 权利要求1-21, 说明书附图图1-2, 4-7</td> <td>1-15</td> </tr> <tr> <td>Y</td> <td>WO 2016025131 A1 (E. I. DU PONT DE NEMOURS AND COMPANY) 2016年 2月 18日 (2016 - 02 - 18) 说明书第82页第29-33行, 附图3B</td> <td>1-15</td> </tr> <tr> <td>Y</td> <td>CN 105112446 A (中国医学科学院基础医学研究所) 2015年 12月 2日 (2015 - 12 - 02) 说明书第415段</td> <td>1-15</td> </tr> <tr> <td>Y</td> <td>VAZQUEZ-VILAR, M. 等. "A modular toolbox for gRNA - Cas9 genome engineering in plants based on the GoldenBraid standard." PLANT METHODS., 第12卷, 第10期, 2016年 2月 1日 (2016 - 02 - 01), 第1-12页</td> <td>1-15</td> </tr> </tbody> </table>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	WO 2015089406 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2015年 6月 18日 (2015 - 06 - 18) 说明书第9-10, 24, 32, 35, 43-44, 46-48, 68-69, 127, 140段	1-15	Y	WO 2015133554 A1 (国立大学法人神户大学) 2015年 9月 11日 (2015 - 09 - 11) 摘要, 权利要求1-21, 说明书附图图1-2, 4-7	1-15	Y	WO 2016025131 A1 (E. I. DU PONT DE NEMOURS AND COMPANY) 2016年 2月 18日 (2016 - 02 - 18) 说明书第82页第29-33行, 附图3B	1-15	Y	CN 105112446 A (中国医学科学院基础医学研究所) 2015年 12月 2日 (2015 - 12 - 02) 说明书第415段	1-15	Y	VAZQUEZ-VILAR, M. 等. "A modular toolbox for gRNA - Cas9 genome engineering in plants based on the GoldenBraid standard." PLANT METHODS., 第12卷, 第10期, 2016年 2月 1日 (2016 - 02 - 01), 第1-12页	1-15
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																		
Y	WO 2015089406 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2015年 6月 18日 (2015 - 06 - 18) 说明书第9-10, 24, 32, 35, 43-44, 46-48, 68-69, 127, 140段	1-15																		
Y	WO 2015133554 A1 (国立大学法人神户大学) 2015年 9月 11日 (2015 - 09 - 11) 摘要, 权利要求1-21, 说明书附图图1-2, 4-7	1-15																		
Y	WO 2016025131 A1 (E. I. DU PONT DE NEMOURS AND COMPANY) 2016年 2月 18日 (2016 - 02 - 18) 说明书第82页第29-33行, 附图3B	1-15																		
Y	CN 105112446 A (中国医学科学院基础医学研究所) 2015年 12月 2日 (2015 - 12 - 02) 说明书第415段	1-15																		
Y	VAZQUEZ-VILAR, M. 等. "A modular toolbox for gRNA - Cas9 genome engineering in plants based on the GoldenBraid standard." PLANT METHODS., 第12卷, 第10期, 2016年 2月 1日 (2016 - 02 - 01), 第1-12页	1-15																		
<p><input checked="" type="checkbox"/> 其余文件在C栏的续页中列出。 <input checked="" type="checkbox"/> 见同族专利附件。</p>																				
<p>* 引用文件的具体类型:</p> <p>“A” 认为不特别相关的表示了现有技术一般状态的文件</p> <p>“E” 在国际申请日的当天或之后公布的在先申请或专利</p> <p>“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)</p> <p>“O” 涉及口头公开、使用、展览或其他方式公开的文件</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“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</p> <p>“&amp;” 同族专利的文件</p>																				
<p>国际检索实际完成的日期</p> <p>2017年 8月 9日</p>		<p>国际检索报告邮寄日期</p> <p>2017年 9月 6日</p>																		
<p>ISA/CN的名称和邮寄地址</p> <p>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</p> <p>传真号 (86-10)62019451</p>		<p>受权官员</p> <p>郭婷婷</p> <p>电话号码 (86-10)62413879</p>																		

## 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
Y	阴晓彤 等. "哺乳动物DNA甲基化修饰的动态调控机制." 生命的化学., 第34卷, 第5期, 2014年 10月 31日 (2014 - 10 - 31), 第604-613页	1-15
A	CN 105408497 A (通用医疗公司) 2016年 3月 16日 (2016 - 03 - 16) 全文	1-15

## 第II栏 某些权利要求被认为是不能检索的意见(续第1页第2项)

根据条约第17条(2)(a)，对某些权利要求未做国际检索报告的理由如下：

1.  权利要求： 15  
因为它们涉及不要求本单位进行检索的主题，即：  
[1] 权利要求15涉及疾病治疗的方法，因此不符合PCT细则第39.1(iv)的规定。本报告基于权利要求15修改为如下做出：权利要求1中所述的重组载体在制备治疗癌症、基因紊乱性遗传疾病、发育错误的药物中的应用。
2.  权利要求：  
因为它们涉及国际申请中不符合规定的要求的部分，以致不能进行任何有意义的国际检索，具体地说：
3.  权利要求：  
因为它们是从属权利要求，并且没有按照细则6.4(a)第2句和第3句的要求撰写。

国际检索报告  
关于同族专利的信息

国际申请号

PCT/CN2017/088281

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	2015089406	A1	2015年 6月 18日	US	2015166980	A1	2015年 6月 18日
				US	9068179	B1	2015年 6月 30日
				EP	3080265	A1	2016年 10月 19日
				CN	105934516	A	2016年 9月 7日
				US	2016304846	A1	2016年 10月 20日
				CA	2933625	A1	2015年 6月 18日
				US	2015166985	A1	2015年 6月 18日
				US	2015166984	A1	2015年 6月 18日
				AU	2014362208	A1	2016年 6月 30日
				US	2015165054	A1	2015年 6月 18日
				US	2015166983	A1	2015年 6月 18日
				JP	2017500035	A	2017年 1月 5日
				US	2015166982	A1	2015年 6月 18日
				US	2015166981	A1	2015年 6月 18日
WO	2015133554	A1	2015年 9月 11日	EP	3115457	A1	2017年 1月 11日
				SG	11201609211V	A	2016年 12月 29日
				CA	2947941	A1	2015年 9月 11日
				US	2017073670	A1	2017年 3月 16日
				JP	W02015133554	A1	2017年 4月 6日
				CN	106459957	A	2017年 2月 22日
WO	2016025131	A1	2016年 2月 18日	WO	2016025131	A8	2017年 6月 8日
				CA	2957684	A1	2016年 2月 18日
				KR	20170036792	A	2017年 4月 3日
				AU	2015302213	A1	2017年 3月 2日
				EP	3180425	A1	2017年 6月 21日
CN	105112446	A	2015年 12月 2日	无			
CN	105408497	A	2016年 3月 16日	WO	2014144592	A3	2014年 12月 31日
				EP	2970208	A2	2016年 1月 20日
				KR	20150131250	A	2015年 11月 24日
				CA	2906553	A1	2014年 9月 25日
				CA	2907198	A1	2014年 9月 18日
				AU	2014227653	A2	2015年 11月 12日
				WO	2014144592	A2	2014年 9月 18日
				EP	2971041	A1	2016年 1月 20日
				EP	2970986	A2	2016年 1月 20日
				EP	2970208	A4	2016年 11月 30日
				AU	2014227653	A1	2015年 10月 1日
				JP	2016512264	A	2016年 4月 25日
				US	2016024524	A1	2016年 1月 28日
				EP	2970986	A4	2016年 9月 7日
				CN	105247066	A	2016年 1月 13日
				US	2016024523	A1	2016年 1月 28日
				EP	2971041	A4	2016年 9月 7日
				JP	2016512691	A	2016年 5月 9日
				EP	2971125	A2	2016年 1月 20日
				CN	105408483	A	2016年 3月 16日
				US	9567604	B2	2017年 2月 14日
WO	2014144288	A1	2014年 9月 18日				
AU	2014227653	B2	2017年 4月 20日				

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/CN2017/088281

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)
		CA 2906724 A1	2014年 9月 18日
		US 9567603 B2	2017年 2月 14日
		JP 2016517276 A	2016年 6月 16日
		EP 2971125 A4	2016年 8月 24日
		KR 20150131251 A	2015年 11月 24日
		US 2014295556 A1	2014年 10月 2日
		WO 2014152432 A3	2015年 10月 29日
		WO 2014144761 A3	2015年 10月 29日
		US 2014295557 A1	2014年 10月 2日
		WO 2014152432 A2	2014年 9月 25日
		WO 2014144761 A2	2014年 9月 18日
		AU 2014228981 A1	2015年 10月 1日
		AU 2014239665 A1	2015年 10月 1日
		US 2016010076 A1	2016年 1月 14日
		KR 20150132395 A	2015年 11月 25日
		US 2017152508 A1	2017年 6月 1日

表 PCT/ISA/210 (同族专利附件) (2009年7月)