

(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
2023年9月7日 (07.09.2023)



(10) 国际公布号
WO 2023/165598 A1

(51) 国际专利分类号:

C12N 9/22 (2006.01) *C12N 15/86* (2006.01)
C12N 15/11 (2006.01) *C07K 14/00* (2006.01)

SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(21) 国际申请号: PCT/CN2023/079545

(22) 国际申请日: 2023年3月3日 (03.03.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:

PCT/CN2022/079302

2022年3月4日 (04.03.2022) CN

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SC, 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, ME, 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)。

(71) 申请人: 益杰立科(上海)生物科技有限
公司(EPIGENIC THERAPEUTICS, INC.) [CN/CN];
中国上海市浦东新区中国(上海)自由
贸易试验区巴圣路160号4号楼一单元
5楼, Shanghai 200131 (CN)。

根据细则4.17的声明:

— 关于申请人有权要求在先申请的优先权(细则
4.17(iii))

(72) 发明人: 周昌阳(ZHOU, Changyang); 中国上海市
浦东新区中国(上海)自由贸易试验区巴
圣路160号4号楼一单元5楼, Shanghai 200131
(CN)。 孙怡迪(SUN, Yidi); 中国上海市浦东新
区中国(上海)自由贸易试验区巴圣路160
号4号楼一单元5楼, Shanghai 200131 (CN)。 陈
珊珊(CHEN, Shanshan); 中国上海市浦东新
区中国(上海)自由贸易试验区巴圣路160号4
号楼一单元5楼, Shanghai 200131 (CN)。 沈逸
涵(SHEN, Yihan); 中国上海市浦东新区中国
(上海)自由贸易试验区巴圣路160号4号楼
一单元5楼, Shanghai 200131 (CN)。

本国际公布:

— 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))。

(74) 代理人: 上海巛石知识产权代理事务所(普通合
伙)(SHANGHAI DIANSHI PARTNERS, P.C.); 中
国上海市浦东新区盛夏路608号2号楼103
室张琤, Shanghai 201210 (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,
CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI,
GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ,
IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ,
LA, LC, LK, LR, LS, LU, LY, MA, MD, 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,

(54) Title: CAS PROTEIN, USE THEREOF AND METHOD THEREFOR

(54) 发明名称: Cas蛋白及其用途和方法

(57) Abstract: A reduced Cas9 protein, use thereof and a method therefor.

(57) 摘要: 一种缩减型Cas9蛋白及其用途和方法。



WO 2023/165598 A1

Cas 蛋白及其用途和方法

技术领域

本申请属于蛋白质进化领域，涉及一种缩减型 Cas9 蛋白及其用途和方法。

背景技术

CRISPR/Cas 系统是来源于细菌和古细菌的天然免疫系统，其由 Cas 核酸酶和可编程的向导 RNA (gRNA) 组成。Cas 蛋白与 gRNA 形成的复合物，可对靶基因进行基因编辑。快速发展的基因编辑技术为基因治疗提供了一个多功能工具箱，可以直接修复人类疾病的相关突变。基因编辑工具箱包括碱基编辑器，先导编辑器，表观遗传编辑器等。

单碱基编辑器是基于 CRISPR/Cas 系统开发的新型靶基因编辑技术，可以在不切断 DNA 双链的情况下实现单点突变。单碱基编辑器由活性受损的 Cas 蛋白和单链脱氨酶组成。碱基编辑器主要可分为 DNA 编辑器和 RNA 碱基编辑器。DNA 碱基编辑器可以进一步分类为胞嘧啶碱基编辑器 (CBE) 或腺嘌呤碱基编辑器 (ABE)；其中 CBE 能够将靶位点的 C•G 向 T•A 转变，而 ABE 则可将靶位点的 A•T 转变成 G•C。

基于 CRISPR 系统开发的体内治疗方法，有望治疗多种人类疾病，具有潜在的巨大治疗价值。尽管目前的 CRISPR 工具箱能够对 DNA 和 RNA 进行编辑以及在基因表达水平和基因表观遗传水平进行调节，但其药物的运输仍是限制基因编辑疗法发展的瓶颈。目前，腺相关病毒(AAV)是最常用的基因治疗载体。AAV 不能独立复制，只有在辅助病毒的参与下才能进行复制。AAV 无致病性。AAV 免疫原性低。迄今为止，共发现了 12 种天然 AAV 衣壳血清型，不同血清型 AAV 提供了多样的组织趋向性。AAV 宿主范围广，对分裂细胞和非分裂细胞均具有感染能力。AAV 的转染和表达效率高。AAV 能介导基因的长期稳定表达。AAV 可以在不整合的情况下，在被感染细胞内形成伴随体，随染色体复制；或者可以介导外源基因发生低频率定点整合到 19 号染色体短臂的 AAVS1 位点。AAV 可以在宿主细胞内长期表达。这一系列优点使得 AAV 成为基因治疗中一种最有潜力的递送载体。

然而，AAV 在临床应用中最突出的问题是容量有限，只有 4.7Kb。因此，除去启动子，sgRNA，polyA，ITR 等基本元件外，留给基因编辑工具的空间不多。而基因编辑工具的体积普遍较大，大部分工具超出了 AAV 的容量，这极大地限制了基因治疗在临床上的应用。因此，体积更小的基因编辑工具具有迫切需求。

目前，化脓链球菌来源的最 Cas 蛋白为广泛使用的基因编辑效应器。对其进行理性设计，使其删减。然而，理性设计得到的删减版 Cas 蛋白和全长 SpCas9 相比存在的编辑效率不高的

问题。除了对已开发的 Cas 蛋白进行理性设计，人们也在不断寻找自然界存在的体积更小的 Cas 蛋白。然而，尽管有这些研究进展，新开发的 Cas 蛋白存在编辑效率总体低下，PAM 位点复杂，脱靶率高等问题依旧限制了基因编辑工具箱的治疗发展。

因此，在本领域中，仍然存在对目前活性最高，最好用的 SpCas 蛋白开发体积更小和活性更高的碱基编辑器的显著需要。

发明内容

本申请提供了一种用于开发蛋白质缩小进化的碱基编辑器的基于转座子插入的进化系统，该进化的碱基编辑器具有高效的编辑活性和显著缩小的体积，因此，解决了现有技术中的问题。

一方面，本申请提供了一种构建删减蛋白文库的方法，所述方法包括：使用多转座系统修饰靶核酸，所述靶核酸包含编码待删减的所述蛋白的核酸序列。

在某些实施方式中，所述多转座系统包括双转座系统。

在某些实施方式中，其中所述转座系统包含转座酶和与转座子，所述转座子包含被转座酶识别用于转座的末端序列。

在某些实施方式中，其中所述转座酶包括转座酶选自：Tn5 转座酶，MuA 转座酶，Tn552 转座酶，Ty1 转座酶，Tn7 转座酶，Tn10 转座酶，IS10 转座酶，Tc1 转座酶，Tn3 转座酶，P-element 转座酶，Mariner 转座酶和睡美人转座酶以及它们的变体。

在某些实施方式中，其中所述转座子能够在任意两个碱基之间随机插入。

在某些实施方式中，其中所述转座酶为 MuA，所述转座子末端为 Mu 转座子末端。

在某些实施方式中，其中所述转座子包含标记基因，所述的标记基因包括抗性基因、复制原点(origin of replication)、启动子元件(Promoter element)和毒素-抗毒素系统(Toxin-Antitoxin system)等其他基因。

在某些实施方式中，所述标记基因使包含所述转座子的细胞与不含所述转座子的细胞区分。

在某些实施方式中，所述细胞包括原核细胞和/或真核细胞。

在某些实施方式中，所述标记基因包括抗生素抗性基因、复制原点(origin of replication)、启动子元件(Promoter element)或毒素-抗毒素系统(Toxin-Antitoxin system)等其他基因。

在某些实施方式中，所述标记基因选自由青霉素、氨苄青霉素、羧苄青霉素、卡那霉素、新霉素、四环素、氯霉素、红霉素、庆大霉素、利福平、链霉素、奇霉素、两性霉素 B、博来霉素、嘌呤霉素、杀稻瘟菌素、Geneticin (G-418)潮霉素 B 和甲氧苄啶的抗性基因组成的组。

在某些实施方式中，其中所述不同的转座子包含不同的标记基因，所述的标记基因包括抗性基因、复制原点(origin of replication)、启动子元件(Promoter element)和毒素-抗毒素系统(Toxin-Antitoxin system)等其他基因。

在某些实施方式中，其中所述转座子的 5'端和 3'端均含有第一切割位点。

在某些实施方式中，其中所述靶核酸包括 DNA 和 RNA。

在某些实施方式中，其中所述靶核酸包括环状 DNA 质粒、线性化 DNA、环状 RNA 和线性化 RNA。

在某些实施方式中，其中所述编码待删减的所述蛋白的核酸序列的 5'端和 3'端均含有第二切割位点，所述第二切割位点与第一切割位点切割方式不同。

在某些实施方式中，其中所述切割位点包括限制酶酶切位点。

在某些实施方式中，其中所述限制酶包括I类限制性核酸内切酶、II类限制性核酸内切酶、III类限制性核酸内切酶，非限制性的，示例的限制酶包括 RNase H、NotI、EcoRI、BamHI、HindIII、TaqI、EcoRV 或 XbaI。

在某些实施方式中，所述的方法包括：a) 将含有第一标记基因的第一转座子插入靶核酸；b) 用识别所述编码待删减的所述蛋白的核酸序列的第二限制酶消化由(a)产生的核酸；c) 将由(b)酶切后产生的包含第一转座子的所述编码待删减的所述蛋白的核酸序列连接空白载体形成包含第一转座子的靶核酸；d) 将含有第二标记基因的第二转座子插入包含第一转座子靶核酸；e) 用识别所述编码待删减的所述蛋白的核酸序列的第二限制酶消化由(d)产生的核酸；f) 将由(e)酶切后产生的包含第一转座子和第二转座子的所述编码待删减的所述蛋白的核酸序列连接空白载体形成包含第一转座子和第二转座子的靶核酸；g) 用第一限制酶对包含第一转座子和第二转座子的靶核酸进行酶切；h) 回收包含编码删减的所述蛋白的核酸序列，得到删减蛋白文库。

在某些实施方式中，所述(b)还包括通过电泳分离酶切后的片段，回收包含第一转座子的所述编码待删减的所述蛋白的核酸序列。

在某些实施方式中，所述(e)还包括通过电泳分离酶切后的片段，回收包含第一转座子和第二转座子的所述编码待删减的所述蛋白的核酸序列。

在某些实施方式中，所述(e)中包含编码删减的所述蛋白的核酸序列自连接形成环状。

在某些实施方式中，所述(a)，(c)，(d)，(f)和/或(h)还包括对产生的靶核酸进行扩增。

在某些实施方式中，所述扩增包括将所述靶核酸转化到宿主细胞中进行扩增。

在某些实施方式中，其中所述蛋白包括抗体、激素、酶、受体、报告分子或选择标记。

在某些实施方式中，其中所述蛋白包括 Cas 蛋白。Cas 蛋白的非限制性实例包括：Cas1、Cas1B、Cas2、Cas3、Cas4、Cas5、Cas6、Cas7、Cas8、Cas9（也称为 Csn1 和 Csx12）、Cas10、Csy1、Csy2、Csy3、Cse1、Cse2、Csc1、Csc2、Csa5、Csn2、Csm2、Csm3、Csm4、Csm5、Csm6、Cmr1、Cmr3、Cmr4、Cmr5、Cmr6、Csb1、Csb2、Csb3、Cxl7、Cxl4、Cxl10、Cxl6、CsaX、Cxl3、Cxl1、Cxl5、Csf1、Csf2、Csf3、Csf4，和/或他们的同系物、或其修饰形式。

在某些实施方式中，其中所述蛋白包括 Cas9 蛋白。所述 Cas9 蛋白是选自棒状杆菌属、萨特氏菌属、军团菌属、密螺旋体属、产线菌属、真杆菌属、链球菌属、乳杆菌属、支原体属、拟杆菌属、Flaviivola、黄杆菌属、Sphaerochaeta、固氮螺菌属、葡糖醋杆菌属、奈瑟菌属、罗氏菌属、细小棒菌属、葡萄球菌属、Nitratifactor、支原体属或弯曲杆菌属的属的 Cas9 直向同源物。

在某些实施方式中，Cas9 蛋白选自：化脓链球菌 Cas9、沃兹沃思萨特氏菌 Cas9、靛沟产线菌 Cas9、约氏乳杆菌 Cas9、红嘴鸥弯曲杆菌 Cas9、白喉棒状杆菌 Cas9、食清洁剂细小棒菌 Cas9、鸡毒支原体 Cas9、金黄色葡萄球菌 Cas9、巴黎嗜肺军团菌 Cas9、齿垢密螺旋体 Cas9、伪中间型葡萄球菌 Cas9、或灰色奈瑟菌 Cas9。

在某些实施方式中，其中所述蛋白包括 SpCas9 蛋白。

在某些实施方式中，其中所述靶核酸包括质粒。

另一方面，本申请提供了一种具有功能活性 Cas 蛋白的筛选方法，所述方法包括将待筛选的所述 Cas 蛋白进行原核筛选和/或真核筛选；

其中，所述原核筛选包括：a) 将编码待筛选的所述 Cas 蛋白的核酸分子连接到原核载体上，所述原核载体包含无义突变的第一报告基因，第二报告基因；b) 将由(a)得到的原核载体转入原核宿主细胞；c) 通过第二报告基因的表达筛选包含原核载体的原核宿主细胞；d) 通过检测第一报告基因的修复，筛选得到具有功能活性的 Cas 蛋白；

所述真核筛选包括：a) 将编码待筛选的所述 Cas 蛋白的核酸分子连接到第二真核载体上，所述原核载体包含第三报告基因；b) 将由(a)得到的真核载体转入包含含有无义突变的第一报告基因的真核宿主细胞；c) 通过第三报告基因的表达筛选包含真核载体的真核宿主细胞；d) 检测第一报告基因的表达，筛选得到具有功能活性的 Cas 蛋白。

在某些实施方式中，其中所述原核载体还包含编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子。

在某些实施方式中，其中所述原核载体还包含编码脱氨酶的核酸分子。

在某些实施方式中，其中所述原核载体编码碱基编辑器。

在某些实施方式中，其中所述原核载体还包含至少一个原核启动子。

在某些实施方式中，其中所述原核载体包括质粒。

在某些实施方式中，其中所述原核宿主细胞包括细菌，如大肠杆菌。

在某些实施方式中，在原核筛选中，所述(c)通过将原核宿主细胞在固体培养基上培养进行筛选，所述固体培养基包含针对第二报告基因的抑制要素，当所述原核宿主细胞不含第二报告基因时，其在所述抑制要素存在的培养中无法存活。

在某些实施方式中，在原核筛选中，所述(d)通过将原核宿主细胞在固体培养基上培养进行筛选，所述固体培养基包含针对第一报告基因的抑制要素，仅在所述原核宿主细胞第一报告基因修复时，其在所述抑制要素存在的培养中存活。

在某些实施方式中，所述固体培养基还包含针对第二报告基因的抑制要素。

在某些实施方式中，在原核筛选中，还包括在所述(d)之前诱导包含原核载体的原核宿主细胞的外源基因的表达。

在某些实施方式中，所述诱导包括向所述原核宿主细胞施用异丙基硫代半乳糖苷 (Isopropyl β -D-Thiogalactoside, IPTG)。

在某些实施方式中，其中在原核筛选中，所述第一报告基因和第二报告基因包括抗生素抗性基因，所述第一报告基因与第二报告基因不同。

在某些实施方式中，其中所述原核筛选包括：a) 将编码待筛选的所述 Cas 蛋白的核酸分子连接到质粒上，所述质粒编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子，脱氨酶，无义突变的第一抗生素抗性基因和第二抗生素抗性基因，得到原核进化质粒；b) 将由(a)得到的原核进化质粒转入蛋白表达型大肠杆菌感受态，并在含有第二抗生素的固体培养基中培养，获得包含原核进化质粒的大肠杆菌；c) 用 IPTG 诱导(b)中得到的包含原核进化质粒的大肠杆菌；d) 将(c)中诱导后的包含原核进化质粒的大肠杆菌置于包含第一抗生素和第二抗生素的固体培养基中培养，筛选得到存活的大肠杆菌，其包含有原核编辑活性的 Cas 蛋白。

在某些实施方式中，在真核筛选中，(b)中所述包含含有无义突变的第一报告基因的真核宿主细胞的制备包括：将包含含有无义突变的第一报告基因和第二报告基因的第一真核载体转入真核宿主细胞，通过第二报告基因的表达筛选包含含有无义突变的第一报告基因的真核宿主细胞。

在某些实施方式中，其中所述第一真核载体和第二真核载体共同编码碱基编辑器。

在某些实施方式中，其中所述真核宿主细胞还包含编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子。

在某些实施方式中，其中所述编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子通过第一真核载体或第二真核载体转入真核细胞。

在某些实施方式中，其中所述第一真核载体和/或第二真核载体还包含编码脱氨酶的核酸分子。

在某些实施方式中，其中所述第一真核载体还包含编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子。

在某些实施方式中，其中所述第一真核载体包含编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子，含有无义突变的第一报告基因和第二报告基因。

在某些实施方式中，其中所述第二真核载体包含编码待筛选的所述 Cas 蛋白的核酸分子，编码脱氨酶的核酸分子和第三报告基因。

在某些实施方式中，其中所述真核载体还包括至少一个真核启动子。

在某些实施方式中，其中所述真核载体还包含 poly A。

在某些实施方式中，其中所述脱氨酶包括腺苷脱氨酶(TadA)。

在某些实施方式中，在所述真核筛选中，不同的所述报告基因编码不同性质的产物，如荧光蛋白所述第一报告基因、第二报告基因与第三报告基因编码不同的荧光蛋白。

在某些实施方式中，其中所述真核宿主细胞包括动物细胞、植物细胞、真菌和昆虫细胞。

在某些实施方式中，其中所述真核宿主细胞包括哺乳动物细胞。

在某些实施方式中，所述真核载体包括质粒、mRNA 和蛋白。

在某些实施方式中，所述真核筛选包括：a) 将包含编码靶向第一报告基因无义突变位置的 sgRNA 的核酸分子，含有无义突变的第一报告基因和第二报告基因的第一真核载体转入真核宿主细胞，通过检测第二报告基因的表达分选得到包含第一真核载体的真核宿主细胞；b) 将包含编码待筛选的所述 Cas 蛋白的核酸分子，编码脱氨酶的核酸分子和第三报告基因的第二真核载体转入(a)中获得的包含第一真核载体的真核宿主细胞，得到包含第一真核载体和第二真核载体的真核宿主细胞；c) 通过检测第一报告基因、第二报告基因和第三报告基因的表达，筛选得到包含具有真核编辑活性的 Cas 蛋白的真核宿主细胞。

在某些实施方式中，其中所述第一报告基因、第二报告基因和第三报告基因编码不同的荧光蛋白。

在某些实施方式中，其中所述细胞通过流式细胞仪分析。

在某些实施方式中，其中还可以通过第一报告基因的表达水平评价待筛选 Cas 蛋白的真核编辑活性。

在某些实施方式中，包括先进行本申请所述的原核筛选，然后再进行本申请所述的真核筛选。

另一方面，本申请提供了一种筛选具有功能活性的删减 Cas 蛋白的方法，包括通过本申请所述的方法构建删减 Cas 蛋白文库，然后通过本申请所述的方法对构建的删减 Cas 蛋白进行原核和/或真核筛选，获得具有原核和/或真核编辑活性的方法。

在某些实施方式中，包括：

i) 构建删减 Cas 蛋白文库：

a) 将含有第一标记基因的第一转座子插入靶核酸；b) 用识别所述编码待删减的所述 Cas 蛋白的核酸序列的第二限制酶消化由(a)产生的核酸；c) 将由(b)酶切后产生的包含第一转座子的所述编码待删减的所述 Cas 蛋白的核酸序列连接空白载体形成包含第一转座子的靶核酸；d) 将含有第二标记基因的第二转座子插入包含第一转座子靶核酸；e) 用识别所述编码待删减的所述 Cas 蛋白的核酸序列的第二限制酶消化由(d)产生的核酸；f) 将由(e)酶切后产生的包含第一转座子和第二转座子的所述编码待删减的所述蛋白的核酸序列连接空白载体形成包含第一转座子和第二转座子的靶核酸；g) 用第一限制酶对包含第一转座子和第二转座子的靶核酸进行酶切；h) 回收包含编码删减的所述蛋白的核酸序列，得到删减 Cas 蛋白文库；

ii) 高通量筛选具有功能活性的删减 Cas 蛋白，所述方法包括将待筛选的所述 Cas 蛋白进行原核筛选和/或真核筛选：

其中，所述原核筛选包括：a) 将编码待筛选的所述 Cas 蛋白文库的核酸分子连接到原核载体上，所述原核载体包含无义突变的第一报告基因，第二报告基因；b) 将由(a)得到的原核载体转入原核宿主细胞；c) 通过第二报告基因的表达筛选包含原核载体的原核宿主细胞；d) 通过检测第一报告基因的修复，筛选得到具有功能活性的 Cas 蛋白。

所述真核筛选包括：a) 将编码待筛选的所述 Cas 蛋白文库的核酸分子连接到第二真核载体上，所述原核载体包含第三报告基因；b) 将由(a)得到的真核载体转入包含含有无义突变的第一报告基因的真核宿主细胞；c) 通过第三报告基因的表达筛选包含真核载体的真核宿主细胞；d) 检测第一报告基因的表达，筛选得到具有功能活性的 Cas 蛋白。

另一方面，本申请提供了一种核酸分子，其编码一种缩减型 Cas9 (mini-Cas9)，所述 mini-Cas9 为缺失部分保守功能结构域但保留 DNA 结合活性的尺寸缩小的化脓链球菌 Cas9，所述缺失部分是位于 RuvCI 区和 REC1 区之间的任意片段，并且除去缺失部分的剩余部分通过接头连接。

在某些实施方式中，所述核酸分子是从编码所述化脓链球菌 Cas9 的核酸分子自 N 端起

第 56 位-第 1041 位之间移除任意长度的核苷酸片段获得。

在某些实施方式中，所述核酸分子是从编码所述化脓链球菌 Cas9 的核酸分子自 N 端起第 56 位-第 1043 位之间移除长度为 217-630 个核苷酸的片段获得。

在某些实施方式中，所述移除片段具有下列任一项所述的长度：217、248、293、333、359、362、378、400、403、437、461、500、518、547 或 630 个核苷酸。

在某些实施方式中，所述移除片段位于编码所述化脓链球菌 Cas9 的核酸分子自 N 端起的下列任一所述区域：第 554 位-第 956 位、第 518 位-第 765 位、第 381 位-第 927 位、第 56 位-第 433 位、第 412 位-第 1041 位、第 589 位-第 1025 位、第 226 位-第 686 位、第 194 位-第 711 位、第 271 位-第 770 位、第 437 位-第 798 位、第 693 位-第 909 位、第 604 位-第 936 位、第 530 位-第 822 位、第 542 位-第 900 位或第 517 位-第 916 位。本申请所述的移除片段均包含其所位于的核酸区域的两个端点核苷酸。

在某些实施方式中，所述核酸分子包含 SEQ ID NOs: 1-15 中任一项所述的序列。

在某些实施方式中，所述核酸分子包含与 SEQ ID NO: 1-15 中任一项所述的序列具有至少约 80%，81%，82%，83%，84%，85%，86%，87%，88%，89%，90%，91%，92%，93%，94%，95%，96%，97%，98%，99%，99.1%，99.1%，99.2%，99.4%，99.5%，99.6%，99.7%，99.8%或 99.9%同源性的核苷酸序列。

另一方面，本申请提供了一种载体，其包含本申请所述的核酸分子。

另一方面，本申请提供了一种缩减型 Cas9 (mini-Cas9)，所述 mini-Cas9 为缺失部分保守功能结构域但保留 DNA 结合活性的尺寸缩小的化脓链球菌 Cas9，所述缺失部分是位于 RuvCI 区和 REC1 区之间的任意片段，并且除去缺失部分的剩余部分通过接头连接。

在某些实施方式中，所述 mini-Cas9 是从所述化脓链球菌 Cas9 的氨基酸序列自 N 端起第 1 位-第 347 位氨基酸之间缺失任意长度的片段获得。

在某些实施方式中，所述 mini-Cas9 是从所述化脓链球菌 Cas9 的氨基酸序列自 N 端起第 1-第 347 位之间缺失长度为 73-320 个氨基酸的片段获得。

在某些实施方式中，所述缺失片段具有下列任一项所述的长度：73、84、100、111、120、130、134、135、145、210、240 或 320 个氨基酸。

在某些实施方式中，所述缺失片段位于所述化脓链球菌 Cas9 的氨基酸序列自 N 端起的下列任一所述区域：第 185 位-第 319 位、第 173 位-第 256 位、第 1 位-第 320 位、第 1 位-第 145 位、第 138 位-第 347 位、第 197 位-第 326 位、第 1 位-第 240 位、第 231 位-第 303 位、第 202 位-第 312 位、第 177 位-第 276 位、第 181 位-第 300 位或第 173 位-第 306 位。本申请

所述的缺失片段均包含其所位于区域的两个端点氨基酸。

在某些实施方式中，所述 mini-Cas9 包含 SEQ ID NOs: 16-27 中任一项所述的序列。

在某些实施方式中，所述 mini-Cas9 包含与 SEQ ID NO: 16-27 中任一项所述的氨基酸序列具有至少约 80%，81%，82%，83%，84%，85%，86%，87%，88%，89%，90%，91%，92%，93%，94%，95%，96%，97%，98%，99%，99.1%，99.1%，99.2%，99.4%，99.5%，99.6%，99.7%，99.8%或 99.9%同源性的氨基酸序列。

在某些实施方式中，所述 mini-Cas9 包含 RuvCII，HNH，RuvCIII 和 PI 结构域。

在某些实施方式中，所述 mini-Cas9 还包含其变体。

在某些实施方式中，所述 mini-Cas9 至少一个催化结构域被突变，使得所述突变后获得的蛋白变体基本上缺乏 DNA 双链切割活性，但保留 DNA 单链切割活性或者仅保留结合活性。

在某些实施方式中，所述 mini-Cas9 蛋白变体的至少两个催化结构域被突变，使得所述突变后获得的蛋白变体基本上缺乏 DNA 切割活性。

本申请提供一种包含经任选的介入接头与异源功能结构域融合的本申请所述的分离的 mini-Cas9 蛋白或本申请所述的 mini-Cas9 蛋白变体的融合蛋白，其中该接头不干扰该融合蛋白的活性。

在某些实施方式中，其中所述异源功能结构域选自解旋酶、核酸酶、解旋酶-核酸酶、DNA 甲基化酶、组蛋白甲基化酶、乙酰基转移酶、磷酸酶、激酶、转录(共)活化物、转录阻遏物、DNA 结合蛋白、DNA 结构蛋白、标志物蛋白、报告物蛋白、荧光蛋白、配体结合蛋白、信号肽、亚细胞定位序列、抗体表位或亲和纯化标签，例如绿色荧光蛋白(GFP)。

在某些实施方式中，其中所述异源功能结构域是转录激活结构域。

在某些实施方式中，其中所述异源蛋白质结构域具有以下活性中的一种或多种：甲基酶活性、脱甲基酶活性、脱氨酶活性、转录激活活性、转录抑制活性、转录释放因子活性、逆转录酶活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。

在某些实施方式中，所述的融合蛋白还包含核定位序列、细胞穿透肽序列和/或亲和标签中的一种或多种。

另一方面，本申请提供了一种分离的核酸分子，其包含编码本申请所述缩减型 Cas9 蛋白或其变体，或本申请所述融合蛋白的核苷酸序列。

在某些实施方式中，所述核酸分子包含 SEQ ID NO: 16-30 中任一项所示的核苷酸序列或与 SEQ ID NO: 16-30 中任一项所示的氨基酸序列具有至少约 80%，81%，82%，83%，84%，85%，86%，87%，88%，89%，90%，91%，92%，93%，94%，95%，96%，97%，98%，99%，

99.1%，99.1%，99.2%，99.4%，99.5%，99.6%，99.7%，99.8%或99.9%同源性的核苷酸序列。

另一方面，本申请提供一种载体，其包含本申请所述的核酸分子。

另一方面，本申请提供一种工程化的、非天然存在的 CRISPR-Cas 系统，其包括：

(a)本申请所述的 mini-Cas9 蛋白或编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)CRISPR-Cas 系统嵌合 RNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列。

另一方面，本申请提供一种工程化的、非天然存在的 CRISPR-Cas 系统，其包括：

(a) 本申请所述的 mini-Cas9 蛋白或编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)CRISPR-Cas 系统 crRNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和

(c)包含 tracr 序列的 tracrRNA。

另一方面，本申请提供一种工程化的、非天然存在的 CRISPR-Cas 载体系统，其包括一个或多个载体，所述载体包含：

(a)编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)编码 CRISPR-Cas 系统嵌合 RNA 的多核苷酸，其中所述嵌合 RNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列。

另一方面，本申请提供一种工程化的、非天然存在的 CRISPR-Cas 载体系统，其包括一个或多个载体，所述载体包含：

(a)编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)编码 CRISPR-Cas 系统 crRNA 的多核苷酸，其中所述 crRNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列；和

(c)编码 tracrRNA 的多核苷酸，其中所述 tracrRNA 包含 tracr 序列。

在某些实施方式中，其中所述 mini-Cas9 蛋白的长度为约 900-1300 个氨基酸。例如，约 900-1250，900-1200，900-1150，900-1100，900-1050，900-1000，950-1250，950-1200，950-1150，1000-1250，1000-1200，1000-1150，1000-1100，1050-1250，1050-1200，1050-1150 或 1050-1100 个氨基酸。

在某些实施方式中，其中所述 tracr 序列的长度为至少 30、40 或 50 个核苷酸。

在某些实施方式中，其中所述指导序列的长度为 15-30 个核苷酸。

在某些实施方式中，其中所述 mini-Cas9 蛋白是能够切割两条 DNA 链的核酸酶。

在某些实施方式中，其中所述 mini-Cas9 蛋白的至少一个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 双链切割活性，但保留 DNA 单链切割活性或者仅保留结合活性。

在某些实施方式中，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A，H840A，N854A 或 N863A 的突变。

在某些实施方式中，其中所述 mini-Cas9 蛋白的至少两个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 切割活性。

在某些实施方式中，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A 的突变和对应于化脓链球菌 Cas9 的 H840A，N854A 或 N863A 的突变。

在某些实施方式中，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。

在某些实施方式中，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。

在某些实施方式中，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录阻抑活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。

在某些实施方式中，其中所述异源功能结构域选自解旋酶、核酸酶、解旋酶-核酸酶、DNA 甲基化酶、组蛋白甲基化酶、乙酰基转移酶、磷酸酶、激酶、转录(共)活化物、转录阻遏物、DNA 结合蛋白、DNA 结构蛋白、标志物蛋白、报告物蛋白、荧光蛋白、配体结合蛋白、信号肽、亚细胞定位序列、抗体表位或亲和纯化标签，例如绿色荧光蛋白(GFP)。

在某些实施方式中，其中编码所述 mini-Cas9 蛋白的所述多核苷酸被密码子优化以在哺乳动物细胞或人细胞中表达。

在某些实施方式中，其中所述系统包括编码所述 mini-Cas9 蛋白的 mRNA。

在某些实施方式中，其中所述一个或多个载体是病毒载体，任选地其中组分(a)和(b)位于同一载体上。

在某些实施方式中，其中所述一个或多个载体是腺相关病毒(AAV)载体。

另一方面，本申请提供了本申请所述的缩减型 Cas 蛋白，本申请所述的核酸分子，本申请所述的载体和/或本申请所述的系统在制备用于基因疗法的药物中的用途。

本申请所述的缩减型 Cas 蛋白及其变体，本申请所述融合蛋白，本申请所述的核酸分子，

本申请所述的载体和/或本申请所述的系统用于离体或体外修饰细胞的用途。

另一方面，本申请提供了一种细胞，其包含本申请所述的缩减型 Cas 蛋白，本申请所述的核酸分子，本申请所述的载体和/或本申请所述的系统。

在某些实施方式中，所述细胞包括真核细胞。

另一方面，本申请提供了一种非人类生物，其包含本申请所述的细胞，其中所述非人类生物不是动物或植物品种。

另一方面，本申请提供了一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 系统引入真核细胞，所述系统包括：

(a)本申请所述的 mini-Cas9 蛋白或编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)CRISPR-Cas 系统嵌合 RNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列，

从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。

另一方面，本申请提供了一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 系统引入真核细胞，所述系统包括：

(a)本申请所述的 mini-Cas9 蛋白或编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)CRISPR-Cas 系统 crRNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和

(c)包含 tracr 序列的 tracrRNA，

从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。

另一方面，本申请提供了一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 载体系统引入真核细胞，所述 CRISPR-Cas 载体系统包括一个或多个载体，所述载体包含：

(a) 编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b)编码 CRISPR-Cas 系统嵌合 RNA 的多核苷酸，其中所述嵌合 RNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列，

从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。

另一方面，本申请提供了一种用于产生修饰的真核细胞的离体或体外方法，其包括将工

程化的、非天然存在的 CRISPR-Cas 载体系统引入真核细胞，所述 CRISPR-Cas 载体系统包括一个或多个载体，所述载体包含：

(a) 编码本申请所述 mini-Cas9 蛋白的多核苷酸；

(b) 编码 CRISPR-Cas 系统 crRNA 的多核苷酸，其中所述 crRNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和

(c) 编码 tracrRNA 的多核苷酸，其中所述 tracrRNA 包含 tracr 序列，

而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。

在某些实施方式中，其中所述 mini-Cas9 蛋白的长度为约 900-1300 个氨基酸。例如，约 900-1250，900-1200，900-1150，900-1100，900-1050，900-1000，950-1250，950-1200，950-1150，1000-1250，1000-1200，1000-1150，1000-1100，1050-1250，1050-1200，1050-1150 或 1050-1100 个氨基酸。

在某些实施方式中，其中所述 tracr 序列的长度为至少 30、40 或 50 个核苷酸。

在某些实施方式中，其中所述指导序列的长度为 15-30 个核苷酸。

在某些实施方式中，其中所述 mini-Cas9 蛋白是能够切割两条 DNA 链的核酸酶。

在某些实施方式中，其中所述 mini-Cas9 蛋白的至少一个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 双链切割活性，但保留单链切割活性或仅保留 DNA 结合活性。

在某些实施方式中，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A，H840A，N854A 或 N863A 的突变。

在某些实施方式中，其中所述 mini-Cas9 蛋白的至少两个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 切割活性。

在某些实施方式中，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A 的突变和对应于化脓链球菌 Cas9 的 H840A，N854A 或 N863A 的突变。

在某些实施方式中，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。

在某些实施方式中，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。

在某些实施方式中，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录抑制活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。

在某些实施方式中，其中编码所述 mini-Cas9 蛋白的所述多核苷酸被密码子优化以在哺乳动物细胞或人细胞中表达。

在某些实施方式中，其中所述系统包括编码所述 mini-Cas9 蛋白的 mRNA。

在某些实施方式中，其中所述一个或多个载体是病毒载体，任选地其中组分(a)和(b)位于同一载体上。

在某些实施方式中，其中所述一个或多个载体是腺相关病毒(AAV)载体。

在某些实施方式中，其中所述真核细胞是哺乳动物细胞。

在某些实施方式中，其进一步包括从所述修饰的真核细胞或其后代产生基因修饰的非人类动物或植物。

通过根据本申请所述的方法产生的修饰的真核细胞或其后代，其中所述修饰的真核细胞或其后代包含靶基因组基因座中与相应的野生型细胞相比的一个或多个核苷酸的插入，缺失或置换。

通过根据本申请所述的方法生产的基因修饰的非人类动物或植物，其中所述非人类动物或植物不是动物或植物品种，并且其中所述修饰的真核细胞或其后代包含靶基因组基因座中与相应的野生型细胞相比的一个或多个核苷酸的插入，缺失或置换。

另一方面，本申请提供了一种改变细胞的基因组的方法，该方法包括在该细胞中表达本申请所述的缩减型 Cas 蛋白以及具有与该细胞的基因组的所选部分互补的区域的指导 RNA，或使该细胞与所述的缩减型 Cas 蛋白以及具有与该细胞的基因组的所选部分互补的区域的指导 RNA 接触。

在某些实施方式中，其中该细胞是干细胞。

在某些实施方式中，其中该细胞是胚胎干细胞、间充质干细胞或诱导多能干细胞；处于活着的动物中；或处于胚胎中。

另一方面，本申请提供了一种改变双链 DNA(dsDNA)分子的方法，该方法包括使该 dsDNA 分子与本申请所述的分离的蛋白或融合蛋白以及具有与该 dsDNA 分子的所选部分互补的区域的指导 RNA 接触。

在某些实施方式中，其中该 dsDNA 分子是在体外。

在某些实施方式中，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。

在某些实施方式中，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。

在某些实施方式中，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录阻抑活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。

本领域技术人员能够从下文的详细描述中容易地洞察到本申请的其它方面和优势。下文

的详细描述中仅显示和描述了本申请的示例性实施方式。如本领域技术人员将认识到的，本申请的内容使得本领域技术人员能够对所公开的具体实施方式进行修改而不脱离本申请所涉及发明的精神和范围。相应地，本申请的附图和说明书中的描述仅仅是示例性的，而非为限制性的。

附图说明

本申请所涉及的发明的具体特征如所附权利要求书所显示。通过参考下文中详细描述示例性实施方式和附图能够更好地理解本申请所涉及发明的特点和优势。对附图简要说明如下：

图 1 显示的是本申请所述基于双转座子插入构建功能性随机删减文库的方法流程。

图 2A 显示的是单转座子的随机插入文库的单克隆的一代测序结果图。

图 2B 显示的是单转座子的随机插入文库 1 的二代测序分析结果图。

图 2C 显示的是单转座子的随机插入文库 1 的双酶切鉴定琼脂糖凝胶电泳图。

图 3A 显示的是单转座子的随机插入文库连空白载体的单克隆的一代测序结果图。

图 3B 显示的是单转座子的随机插入文库连空白载体的单克隆的二代测序分析结果图

图 4A 显示的是双转座子的随机插入文库的单克隆的一代测序结果图。

图 4B 显示的是双转座子的随机插入文库的二代测序分析结果图。

图 4C 显示的是双转座子的随机插入文库的双酶切鉴定琼脂糖凝胶电泳图。

图 5A 显示的是双转座子的随机插入文库连空白载体的单克隆的一代测序结果图。

图 5B 显示的是双转座子的随机插入文库连空白载体的单克隆的二代测序分析结果图。

图 5C 显示的是双转座子的随机插入文库连空白载体的双酶切鉴定琼脂糖凝胶电泳图。

图 6A 显示的是 Cas9 随机删减文库的构建流程。

图 6B 显示的是 Cas9 随机删减文库的一代测序分析结果图。

图 6C 显示的是 Cas9 随机删减文库的双酶切鉴定琼脂糖凝胶电泳图。

图 7A 显示的是 Cas9 的原核进化流程。

图 7B 显示的是原核进化质粒示意图。

图 7C 显示的是 Cas9 蛋白诱导表达后 12 个功能性删减版变体的原核编辑结果。

图 8 显示的是全长 SpCas9 蛋白结构域示意图和删减版 SpCas9 的 DNA 序列比对和蛋白质序列比对示意图。

图 9A 显示的是 Cas9 真核进化流程。

图 9B 显示的是真核双荧光报告系统的质粒示意图。

图 9C 显示的是真核双荧光报告系统的质粒图谱。

图 9D 显示的是真核进化载体的质粒示意图。

图 9E 显示的是真核进化载体的质粒图谱。

图 9F 显示的是真核慢病毒进化载体的质粒图谱。

图 9G 显示的是包含全长 Cas9 的碱基编辑器在双荧光报告细胞系中编辑 72 小时后流式分选时的细胞群划分图。

图 9H 显示的是包含全长 Cas9 的碱基编辑器在双荧光报告细胞系中编辑 72 小时的结果分析。

图 9I 显示的是包含全长 Cas9 的碱基编辑器和变体 MINI 在双荧光报告细胞系中编辑 72 小时分选 5% 转染阳性细胞的结果分析。

图 10 显示的是变体 MINI1 在 293 细胞内源基因上验证的结果。

具体实施方式

以下由特定的具体实施例说明本申请发明的实施方式，熟悉此技术的人士可由本说明书所公开的内容容易地了解本申请发明的其他优点及效果。

不欲被任何理论所限，下文中的实施例仅仅是为了阐释本申请的文库构建及筛选、用途及缩减型 Cas9 蛋白等，而不用用于限制本申请发明的范围。

实施例

实施例 1 基于双转座子插入构建功能性随机删减文库

实施例 1.1 构建基于双转座子插入构建随机删减文库

第一步，单转座子的随机插入文库构建：

以包含要进行缩小进化的片段的质粒为模版，用体外转座的方式往质粒随机插入带有卡纳霉素 (Kanamycin, Kan) 抗性基因的转座子，制备带有 Kan 抗性基因的单转座子的随机插入文库；将单转座子随机插入文库用片段两端相应的限制性核酸内切酶（非 NotI 限制性核酸内切酶）对文库进行酶切，然后进行琼脂糖凝胶电泳，获得含有单个转座子的片段，然后将含有单个转座子的片段插入空白质粒载体获得单转座子的随机插入文库；

第二步，双转座子的随机插入文库构建：

然后，在单转座子的随机插入文库的基础上，进行第二次转座，用体外转座的方式往质粒随机插入带有氯霉素 (chloramphenicol, Cam) 抗性基因的转座子，构建同时具有 Kan 抗性基因转座子和 Cam 抗性基因转座子的双转座子随机插入文库；其中 Kan 抗性基因转座子和

Cam 抗性基因转座子的 5'端和 3'端均含有 NotI 限制性核酸内切酶的酶切位点；将双转座子随机插入文库用片段两端相应的限制性核酸内切酶（非 NotI 限制性核酸内切酶）对文库进行酶切，然后进行琼脂糖凝胶电泳，获得含有双转座子的片段，然后将含有双转座子的片段插入空白质粒载体获得双转座子的随机插入文库；

第三步，随机删减文库构建

用 NotI 限制性核酸内切酶对双转座子的随机插入连空白载体的文库进行酶切，然后回收带有被截短的 Cas9 片段的载体部分，将该部分片段自连，得到 Cas9 随机删减文库（图 5C 和图 6A）。将文库转化到 DH5a 克隆型感受态中，于 37 度过夜培养。次日，挑选单克隆，摇菌，送测。Cas9 随机删减文库的一代测序分析结果指示 Cas9 随机删减版变体构建成功（图 6B）。对 Cas9 随机删减文库的用片段两端相应的限制性核酸内切酶对文库进行酶切，然后进行琼脂糖凝胶电泳。文库会被切割成两种条带，分别是载体和被随机删减的片段（图 6C）。

实施例 1.2 单转座子的随机插入文库的鉴定

以包含要进行缩小进化的片段的质粒为模版，用体外转座的方式往质粒随机插入带有 Kan 抗性基因的转座子，制备带有 Kan 抗性基因的单转座子的随机插入文库。将文库转化到 DH5a 克隆型感受态中，于 37°C 过夜培养。次日，挑选单克隆，摇菌，送测。单转座子随机插入文库的一代测序分析结果指示转座子随机插入到质粒中（图 2A）。随后，将单转座子随机插入文库进行二代测序，将单转座子随机插入文库的二代测序结果与 DNA 图谱进行比对，筛选同时包含转座子和片段的测序片段，并统计插入位置和插入次数。结果显示转座子随机插入到质粒中每对碱基对之间（图 2B）。将单转座子的随机插入文库用片段两端相应的限制性核酸内切酶对文库进行酶切，然后进行琼脂糖凝胶电泳。文库会被切割成四个条带，分别是片段和载体，以及包含有单个转座子的片段和载体（图 2C）。回收包含有单个转座子的片段，并将其连接到新的空白载体上。单转座子的随机插入文库连空白载体的单克隆的一代测序分析结果。结果指示转座子仅随机存在于目的片段中（图 3A）。将单转座子的随机插入文库连空白载体的二代测序结果与目的片段的 DNA 图谱进行比对，筛选同时包含转座子和片段的测序片段，并统计插入位置和插入次数。结果显示了转座子插入的随机性，目的片段每对碱基对间均有转座子插入（图 3B）。

实施例 1.3 双转座子的随机插入文库的鉴定

在具有 Kan 抗性基因单转座子随机插入库的基础上，用具有 Cam 抗性基因转座子进行第二次体外转座，构建同时具有 Kan 抗性基因转座子和 Cam 抗性基因转座子的双转座子随机插入文库质粒。将文库转化到 DH5a 克隆型感受态中，于 37°C 过夜培养。次日，挑选单克隆，

摇菌，送测。双转座子随机插入文库的一代测序分析结果指示转座子随机插入到质粒中，其中，第一种转座子和第二种转座子的比对结果如图 4A 所示。结果显示，两种转座子均随机插在目的片段的碱基对间。双转座子随机插入文库的二代测序分析结果。将单转座子随机插入文库的二代测序结果与 DNA 图谱进行比对，筛选同时包含转座子和片段的测序片段，并统计插入位置和插入次数。结果显示了转座子插入的随机性，目的片段每对碱基对间均有转座子插入（图 4B）。将双转座子随机插入文库用片段两端相应的限制性核酸内切酶对文库进行酶切，然后进行琼脂糖凝胶电泳。文库会被切割成四个条带，分别是空白载体、包含有单个转座子的载体、包含单个转座子的片段和双转座子的片段（图 4C）。回收包含有双转座子的片段，并将其连接到新的空白载体上。双转座子的随机插入文库连空白载体的单克隆的一代测序分析结果指示双转座子仅随机存在于目的片段中（图 5A）。将双转座子的随机插入文库连空白载体的二代测序结果与目的片段的 DNA 图谱进行比对，筛选同时包含转座子和片段的测序片段，并统计插入位置和插入次数。结果显示了双转座子插入的随机性，目的片段每对碱基对间均有转座子插入（图 5B）。将双转座子随机插入文库用限制性核酸内切酶 NotI 对文库进行酶切，然后进行琼脂糖凝胶电泳。文库会被切割成三种条带，分别是带有被截短的片段的载体，单个转座子和被删除的片段（图 5C）。

实施例 2 随机删减文库在原核生物中的功能性筛选

图 7A 为原核进化的流程示意图（图 7A）。将 Cas9 随机删减版文库连接到碱基编辑器原核进化载体（图 7B）上，得到具有随机删减版-进化载体-文库，并将其转化到蛋白表达型大肠杆菌感受态，在诱导剂的存在下，表达随机删减版蛋白，并将其涂布在 Amp 抗性的平板上，进行功能性筛选系统，如果删减版 Amp 基因有原核编辑活性，将能在 Amp 抗性平板上生长，从而得到具有功能活性的随机删减版文库（图 6C 和图 7A）。将连接到原核进化载体上 Cas9 随机删减版文库转化到表达型大肠杆菌感受态中，用 IPTG 诱导随机删减版蛋白表达。用 IPTG 诱导 Cas9 蛋白表达 5 小时，于 37 度过夜培养，具有原核编辑活性的删减版变体可以高活性修复 AmpR 抗性基因上的无义突变，在 AmpR 抗性平板上存活（图 7C）。

如图 8 和表 1 所示，对筛选得到的删减版 SpCas9 蛋白进行测序并与全长的 SpCas9 蛋白进行比对，图 8 中透明位置为 Cas9 DNA 或蛋白质删减的长度和位置以及相应 linker 的插入位置，表 1 中的“-”表示氨基酸被删除。

表 1 删减版 SpCas9 蛋白与全长删减版 SpCas9 蛋白序列比对

LABEL	长度/缺失片段信息	SEQ ID NO:	SEQ
-------	-----------	------------	-----

<p>全长 SpCas9n</p>	<p>1368 aa/ -</p>	<p>28</p>	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILS RLSKSRLENLIAQLPGEKKNGLFGNLIASLGLTPNFKSNFDL AEDAKLQLSKDYYDDDLNLLAQIGDQYADLFLAAKNLSDA ILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED FYPFLKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEE TITPWNFEVVDKGASQSFIERMTNFDKNLPNEKVLPHKSHL YEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHL FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK SDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN LAGSPAIKKILQTVKVVDELVKVMGRHKPENIVIEMARENQ TTQKGQKNSRERMKRIEKGELGSQILKEHPVENTQLQNEKL YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID NKVLTRSDKNRGKSDNVPSEEVKMKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLELKDAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVIN NYHHAHDAYLNAVVGTAIKKYPKLESEFVYGDYKVYDVR KMIKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSPQVNVKKTVEVQTGG FSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSEKPNIDFLEAKGYKE VKKDLIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFYLAHYEKLKGPEDNEQKQLFVEQHKHYLDEIEIQIS</p>
-----------------------	-------------------	-----------	---

			<p>EFSKRVLADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQISITGLYETRID LSQLGGD</p>
Mini1	1237 aa/第 185-319 位 aa	16	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPFIGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVDKL----- -----LRPH----- SMIKRYDEHHQDLTLLKALVRQQLPKEYKE IFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFKLD NREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITPWNF EEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDK DFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDDKV MKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFA NRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLT RSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKACFIKRQLVETRQITKHVAQILDSRMNT KYDENDKLIREVKVITLKSCLVSDFRKDFQFYKVVREINNYHH AHDAYLNAVVGITALIKKYPKLESEFVYGDYKVVYDVRKMIAK SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSPQVNVIVKKTEVQTGGFSKESI LPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLLSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDL</p>

			<p>IIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLY LASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVL SAYNKH RDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD</p>
Mini2	1288 aa/第 173-256 位 aa	17	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCY LQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE G-----AAA----- ----FDLAEDA KLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTBELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFL KDNREKIEKILTRIPYYVGPLARGNSRF AWMTRKSEETITPW NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYF TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKV TVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKYAHLFDD KVMKQLKRRRYTGWRLSRKLINGIRDKQSGKTILDFLKSDG FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ KGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK VLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINN YHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYK VYDVRK MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGF</p>

			<p>SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQSI TGLYETRIDL SQLGGD</p>
<p>Mini3、 Mini8、 Mini9 和 Mini11（由 不同核酸分 子转录获 得）</p>	<p>1048 aa/第 1-320 位 aa</p>	<p>18</p>	<p>----- ----- ----- ----- MIKRYDEHHQDLTLLKALVRQQLPEK YKEIFFDQSKNGYAGYIDGGASQBEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPF LKDNREKIEKILTRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLKII KDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKK GILQTVKVVDELVKVMGRHKPENIVIAMARENQTT QKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYK VREINN YHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRK MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGF</p>

			<p>SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLKSVKELLGITIMERSSSFENPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDL SQLGGD</p>
Mini4	1228 aa/第 1-145 位 aa	19	<p>-----MGAA----- ----- ATDKADRLRIYLALAHM IKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASG VDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSGL TPNFKSNFDLAEDAQLQSKD TYDDDLNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLT LLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI KPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFY PFLKDNREKIEKILTRIPYYVGPLARGNSRFA WMTRKSEETITPWNFEVVDKGASQSFIERMTNFDKNLPNE KVLPHKSHLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNA SLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGK TILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKK GILQTVKVVDDELVKVMGRHKPENI VIEMARENQTTQKGQKNSRERMKRIE EGKELGSQILKEHPVE NTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNR GKS DNV PSEEVVKKMKNYWR QLLNAKLITQRKFDNLTKAERGGSEL DKA GFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKD FQFYKVINNYHHAH DAYLNAVVG TALIKKYPKLESEFVYG</p>

			<p>DYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEITLA NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS TVAYSVLVVAKVEKGSKKLKSVKELLGITIMERSSEKPNID FLEAKGYKEVKKDLIILPKYSLFEENGRKRMLASAGELQK GNELALPSKYVNFLYLASHYEKLGSPEDNEQQLFVEQHK HYLDEIIEQISEFSKRVILADANLKVLSAYNKHRRDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLIHQ ITGLYETRIDLSQLGGD</p>
Mini5	1161 aa/第 138-347 位 aa	20	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPFGNIVDEVAY HEKYPTIYHCGR----- ----- -----KEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL RKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKI LTRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGG ASAQSFIERMTNFDK NLPNEKVLPHSLLYEYFTVYNELTKV KYVTEGMRKPAFLSGEQKKAIVDLLFKTRKVTVKQLKEDY FKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEEN EDILEDIVLTTLTFEDREMIEERLKYAHLFDDKVMKQLKRRR YTGWRLSRKLINGIRDKQSGKTILDFLKSDFANRNFMLIH DDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV KVVDELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMK RIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKS DNVPSSEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSCLVSDFRKDFQFYKVVREINNYHHAHDAYLNAV</p>

			<p> VGTALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIA RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLSVK ELLGITIMERSSEKPNIDFLEAKGYKEVKKDLIKLPKYSLFEL ENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKV LSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDTLHQSTGLYETRIDLSQLGGD </p>
Mini6	1259aa/第 197-326 位 aa	21	<p> MDKKYSIGLAIGTNSVGWAVITDEYKVPSPKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVKLFIQLVQTYNQLFGDRLLRAPGVHDTLIRS DRTHEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYPLKDNREKIEKILFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEVVVDKGASAQSFIER MTNFDKLNPNKVLPHSLLEYEFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVL TLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRLL SRKLINGIRDKQSGKTILDFLKSDFANRNFMLIHDDSLTFK EDIQKAQVSGQGDSLHEHIANLAGSPAIKKILQTVKVVDEL VKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIIEGIK ELGSQILKEHPVENTQLQNEKLYLYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNPSE EVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGSELDKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKREINNYHHAHDAYLNAVVGTAI </p>

			<p>KKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKEVGKSKKLSVKELLGI TIMERSSEFKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPED NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAY NKH RDKPIREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTS TKEVLDATLIHQSI TGLYETRIDLSQLGGD*</p>
Mini10	1134aa/第 1-240 位 aa	22	<p>MRC AALIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDDLD NLLAQIGDQYADLFLAAKNLSDAILSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEBFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIP YYVGPLARGNSRFAWMTRKSEETITPWNFEVV DKGASAQS FIERMTNFDKNLPNEKVLPHSLLEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIEC FDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILED IVLTLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGW GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSL TFKEDIQKAQVSGQDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL DINRLSDYDVHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNV PSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSCLVSDFRKDFQFYK VREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKY FFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARK</p>

			<p>KDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELL GITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGKSP EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVLS AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSIITGLYETRIDLSQLGGD*</p>
Mini12	1299aa/第 231-303 位 aa	23	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPHFNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILS RLSKSRLENLIAQLPGAAADILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPLKDNREKIEKILTRIPYVYGPLA RGNRSFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPHSLLEYFTVYNELTKVKYVTEGMRKPAF LSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISG VEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLF EDREMIERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQK AQVSGQGDSLHEHIANLAGSPAIKKILQTVKVVDLKVVMG RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDY DVDHIVPQSFLKDDSIDNKVLTRSDKNRKGSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCL VSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAIIKKYPK LESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL</p>

			<p>MPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKK YGGFDSPTVAYSVLVVAKVEKGKSKLKSVKELLGITIMERS SFEKNPIDFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLA SAGELQKGNELALPSKYVNFLYLASHYEKLLKGSPEDEQKQL FVEQHKHYLDEIIEQISEFVKRVLADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSTGLYETRIDLSQLGGD</p>
Mini13	1260aa/第 202-312 位 aa	24	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCICYLQEIF SNEMAKVDDSSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRILIYLAHAMIKFRGHFLIE GDLNPDNSDVKLFIQLVQTYNQLFEENPICGRTKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKBIFFDQSKNGYAGYID GGASQEEFYKFIKPILEKMDGTEELLVKNREDLLRKQRTFDN GSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEVVDKGASAQSFIER MTNFDKNLPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVL TLTLFEDREMIEERLKYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFK EDIQKAQVSGGDSLHEHIANLAGSPAIAKKGILQTVKVVDEL VKVMGRHKPENIVIAMARENQTTQKGQKNSRERMKRIEEGIK ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSE EVVKMKMNYWRQLLNAKLITQRKFDNLTKAERGGSELDKA GFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVINNYHHAHDAYLNAVVGTAI KKYPKLESEFVYGDYKVDVRKMIKSEQEIGKATAKYFFYS NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV</p>

			<p>RKVLSPQVNVKKTQVGGFSKESILPKRNSDKLIARKKD WDPKKGFFDSPTVAYSVLVVAKEVGKSKKLSVKELLGI TIMERSSEKPNIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLGSPED NEQKQLFVEQHKHYLDEIEQISEFSKRVLADANLDKVL SAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFYFDTTIDRKRYTS TKEVL DATLIHQ SITGLYETRIDLSQLGGD</p>
Mini14	1274aa/第 177-276 位 aa	25	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCYQLQEIF SNEMAKVDDSFHRLEESFLVEEDKKHERHPFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPA AAAALDNLLAQIGDQYADFLAAKNLSDAILLSDILR VNTEITKAPLSASMIKRYDEHHQDLTLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDN REKIEKILTFRIPYVVGPLARGNSRFAWMTRKSEETITPWNFEE VVDKGASASQFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTRNKVTVKQ LKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKHKDKDFL DNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMKQ LKRRTYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIAKK GILQTVKVVDELVKVMGRHKPENIVEMARENQTTQKGQKN SRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLYLYYLQNG RDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLITQRKFDNLTK AERGLSELKAGFIKQRLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSCLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQ EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI</p>

			<p>VWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEVEKGSK KLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIILP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLGSPEDNEQKQLFVEQHKHYLDEIEIQSEFSKRVLAD ANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYF DTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD</p>
Mini15	1251aa/第 181-300 位 aa	26	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVPSKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCICYLQEIF SNEMAKVDDSSFFHRLEESFLVEEDKKHERHPFIGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GDLNPDNSDAAALLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLEPKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKNREDLLRKQRTFDNGSIPHQIH LGELHAILRRQEDFYFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETIPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGE QKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR FNASLGTYHDLLKIKDKDFLDNEENEDILEDIVLTLTLFEDRE MIEERLKYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKILQTVKVVDELVKVMGRHKP ENIVIAMARENQTTQKGQKNSRERMKRIIEGKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGELSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFR KDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFV YGDYKVYDVRKMIKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNI</p>

			<p>VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVVAKVEKGKSKLKSVKELLGITIMERSSEFKNPI DFLEAKGYKEVKKDLIILPKYSLFELENGRKRMLASAGELQ KGNELALPSKYVNFYLYLASHYEKLGSPEDNEQQLFVEQH KHYLDEIEEQISEFSKRVLADANLDKVL SAYNKH RDKPIREQ AENIIHLFTLNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGD</p>
Mini16	1238aa/第 173-306 位 aa	27	<p>MDKKYSIGLAIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRNRCYQLQEIF SNEMAKVDDSSFFHRLEESFLVEEDKKHERHPFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE GCGRMRVNTTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGT EELLVKNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDF YPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI TPWNFEVVVDKGASAQSFIERMTNFDKNLPNEKVLPHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLK IKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFD DKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKQILQTVKVVDELVKVMGRHKPENIVIAMARENQTT QKGQKNSRERMKRIEIEGKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDN KVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKREINN YHHAHDAYLNAVVGTAIHKYKLESEFVYGDYKVVYDVRK MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGF</p>

			SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVA KVEKGKSKKLSVKELLGITIMERSSEFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLGKSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVL SAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSI TGLYETRIDL SQLGGD
--	--	--	---

表 1 中编码各删减版 SpCas9 蛋白序列的 DNA 序列及其从全长 SpCas9 DNA 中删减的片段信息汇总如下：Mini1 (SEQ ID NO: 1, DNA 序列删减信息：第 554-956 位)；Mini2 (SEQ ID NO: 2, DNA 序列删减信息：第 518-765 位)；Mini3 (SEQ ID NO: 3, DNA 序列删减信息：第 381-927 位)；Mini4 (SEQ ID NO: 4, DNA 序列删减信息：第 56-433 位)；Mini5 (SEQ ID NO: 5, DNA 序列删减信息：第 412-1041 位)；Mini6 (SEQ ID NO: 6, DNA 序列删减信息：第 589-1025 位)；Mini8 (SEQ ID NO: 7, DNA 序列删减信息：第 226-686 位)；Mini9 (SEQ ID NO: 8, DNA 序列删减信息：第 194-71 位)；Mini10 (SEQ ID NO: 9, DNA 序列删减信息：第 271-770 位)；Mini11 (SEQ ID NO: 10, DNA 序列删减信息：第 437-798 位)；Mini12 (SEQ ID NO: 11, DNA 序列删减信息：第 693-909 位)；Mini13 (SEQ ID NO: 12, DNA 序列删减信息：第 604-936 位)；Mini14 (SEQ ID NO: 13, DNA 序列删减信息：第 530-822 位)；Mini15 (SEQ ID NO: 14, DNA 序列删减信息：第 542-900 位)；Mini16 (SEQ ID NO: 15, DNA 序列删减信息：第 517-916 位)。

结果如表 1 和图 7C、图 8 和图 9I 所示，在编码全长 SpCas9 的基因 DNA 上可进行不同位置上任意长度的删减，所获得的删减版 DNA 序列编码的缩减型 Cas9 蛋白变体仍保持良好的编辑活性。

实施例 3 随机删减文库在真核生物中的功能性筛选

如图 9B-9C 所示，构建双荧光报告系统质粒，其包括以下元件：U6 启动子、J2319 启动子、sgRNA、CMV 启动子、mCherry 红色荧光蛋白基因、带有无义突变的 GFP 绿色荧光蛋白基因和 poly A。如图 9D-9E 所示构建真核进化载体的质粒。真核进化载质粒包括以下元件：CMV 启动子、TadA 脱氧腺苷脱氨酶、Cas9 随机删减版文库、BFP 蓝色荧光蛋白基因和 poly A。如图 9F 所示构建真核慢病毒进化载体。真核慢病毒进化载体包括以下元件：EFS 启动子、TadA 脱氧腺苷脱氨酶、Cas9 随机删减版文库、BFP 蓝色荧光蛋白基因、WPRE 和 poly A。

用慢病毒构建带有双荧光报告系统的稳定细胞株。将双荧光报告系统克隆到具有慢病毒质粒载体上（图 9C），转染至 HEK 293FT 细胞中，包装慢病毒。收集转染 48 小时后包含病毒颗粒的上清液。用病毒上清液感染 HEK 293 细胞，使双荧光报告系统整合到宿主染色体上，用流式分选红光阳性绿光阴性的阳性细胞，即可得到双荧光报告系统的稳转细胞株。

将带有删减版 Cas9 的 ABE 文库克隆到慢病毒质粒载体上（图 9F），转染至 HEK 293FT 细胞中，包装慢病毒。收集转染 48 小时后包含病毒颗粒的上清液。用病毒上清液感染 HEK 293 细胞，使删减版 Cas9 的 ABE 文库整合到稳转细胞株的染色体上，具有真核碱基编辑活性的变体将把修复 GFP 上的无义突变，使得稳转细胞株正常表达 GFP 基因，通过流式分选红绿蓝三阳的细胞，得到具有功能活性的删减版 Cas9 的文库（图 9A）。

将带有包含全长 Cas9 的碱基编辑器克隆到真核进化载体上（图 9E）。用包含全长 Cas9 的碱基编辑器测试双荧光报告细胞系的荧光强度和修复效率的联系。用包含全长 Cas9 的碱基编辑器在双荧光报告细胞系中编辑 72 小时后流式分选时的细胞群划分图。分选出红蓝绿三阳的细胞，并根据绿光强度将阳性细胞分为 4 群：P9、P10、P11 和 P12（绿色荧光强度依次递增）（图 9G）。结果显示 GFP 荧光强度越高，则 A 至 G 的编辑效率越高（图 9H）。用包含全长 Cas9, Mini13, Mini14, Mini15, Mini16 Cas9 的碱基编辑器在双荧光报告细胞系中编辑 72 小时后检测编辑效率，结果显示包含 Mini13, Mini14, Mini15, Mini16 Cas9 的碱基编辑器在双荧光报告细胞系中具有活性（图 9I）。

实施例 4 编辑真核哺乳动物细胞的内源性基因

将带有 Cas9 随机删减版变体 Mini1 的真核进化质粒和靶向 293 细胞内源基因基因 site1 的 sgRNA (SEQ ID NO: 35) 质粒共转染到 HEK293 细胞中，48 小时后用 PCR 扩增相应片段，用一代测序结果分析编辑效率。结果（图 10）显示 Mini1 与野生型 Cas9 蛋白相比，具有野生型 81% 的活性。

1. 一种核酸分子，其编码一种缩减型 Cas9 (mini-Cas9)，所述 mini-Cas9 为缺失部分保守功能结构域但保留 DNA 结合活性的尺寸缩小的化脓链球菌 Cas9，所述缺失部分是位于 RuvCI 区和 REC1 区之间的任意片段，并且除去缺失部分的剩余部分通过接头连接。
2. 根据权利要求 1 所述的核酸分子，其是从编码所述化脓链球菌 Cas9 的核酸分子自 N 端起第 56 位-第 1041 位之间移除任意长度的核苷酸片段获得。
3. 根据权利要求 1-2 中任一项所述的核酸分子，其是从编码所述化脓链球菌 Cas9 的核酸分子自 N 端起第 56 位-第 1043 位之间移除长度为 217-630 个核苷酸的片段获得。
4. 根据权利要求 2-3 中任一项所述的移除片段，其具有下列任一项所述的长度：217、248、293、333、359、362、378、400、403、437、461、500、518、547 或 630 个核苷酸。
5. 根据权利要求 2-4 中任一项所述的移除片段，其位于编码所述化脓链球菌 Cas9 的核酸分子自 N 端起的下列任一所述区域：第 554 位-第 956 位、第 518 位-第 765 位、第 381 位-第 927 位、第 56 位-第 433 位、第 412 位-第 1041 位、第 589 位-第 1025 位、第 226 位-第 686 位、第 194 位-第 711 位、第 271 位-第 770 位、第 437 位-第 798 位、第 693 位-第 909 位、第 604 位-第 936 位、第 530 位-第 822 位、第 542 位-第 900 位或第 517 位-第 916 位。
6. 根据权利要求 1-5 中任一所述的核酸分子，其包含 SEQ ID NOs: 1-15 中任一项所述的序列。
7. 一种载体，其包含权利要求 1-6 中任一项所述的核酸分子。
8. 一种缩减型 Cas9 (mini-Cas9)，所述 mini-Cas9 为缺失部分保守功能结构域但保留 DNA 结合活性的尺寸缩小的化脓链球菌 Cas9，所述缺失部分是位于 RuvCI 区和 REC1 区之间的任意片段，并且除去缺失部分的剩余部分通过接头连接。
9. 根据权利要求 8 所述的 mini-Cas9，其是从所述化脓链球菌 Cas9 的氨基酸序列自 N 端起第 1 位-第 347 位氨基酸之间缺失任意长度的片段获得。
10. 根据权利要求 8-9 中任一项所述的 mini-Cas9，其是从所述化脓链球菌 Cas9 的氨基酸序列自 N 端起第 1-第 347 位之间缺失长度为 73-320 个氨基酸的片段获得。
11. 根据权利要求 9-10 中任一项所述的缺失片段，其具有下列任一项所述的长度：73、84、100、111、120、130、134、135、145、210、240 或 320 个氨基酸。
12. 根据权利要求 9-11 中任一项所述的缺失片段，其位于所述化脓链球菌 Cas9 的氨基酸序列自 N 端起的下列任一所述区域：第 185 位-第 319 位、第 173 位-第 256 位、第 1 位-第

- 320 位、第 1 位-第 145 位、第 138 位-第 347 位、第 197 位-第 326 位、第 1 位-第 240 位、第 231 位-第 303 位、第 202 位-第 312 位、第 177 位-第 276 位、第 181 位-第 300 位或第 173 位-第 306 位。
13. 根据权利要求 8-12 中任一项所述的 mini-Cas9，其包含 SEQ ID NOs: 16-27 中任一项所述的序列。
14. 一种工程化的、非天然存在的 CRISPR-Cas 系统，其包括：
- (a) 权利要求 8-13 中任一项所述的 mini-Cas9 蛋白或编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) CRISPR-Cas 系统嵌合 RNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列。
15. 一种工程化的、非天然存在的 CRISPR-Cas 系统，其包括：
- (a) 权利要求 8-13 中任一项所述的 mini-Cas9 蛋白或编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) CRISPR-Cas 系统 crRNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和
 - (c) 包含 tracr 序列的 tracrRNA。
16. 一种工程化的、非天然存在的 CRISPR-Cas 载体系统，其包括一个或多个载体，所述载体包含：
- (a) 编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) 编码 CRISPR-Cas 系统嵌合 RNA 的多核苷酸，其中所述嵌合 RNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列。
17. 一种工程化的、非天然存在的 CRISPR-Cas 载体系统，其包括一个或多个载体，所述载体包含：
- (a) 编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) 编码 CRISPR-Cas 系统 crRNA 的多核苷酸，其中所述 crRNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶

- 序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列；和
- (c)编码 tracrRNA 的多核苷酸，其中所述 tracrRNA 包含 tracr 序列。
18. 根据权利要求 14-17 中任一项所述的系统，其中所述 mini-Cas9 蛋白的长度为约 900-1300 个氨基酸。
 19. 根据权利要求 14-18 中任一项所述的系统，其中所述 tracr 序列的长度为至少 30、40 或 50 个核苷酸。
 20. 根据权利要求 14-19 中任一项所述的系统，其中所述指导序列的长度为 15-30 个核苷酸。
 21. 根据权利要求 14-20 中任一项所述的系统，其中所述 mini-Cas9 蛋白是能够切割两条 DNA 链的核酸酶。
 22. 根据权利要求 14-20 中任一项所述的系统，其中所述 mini-Cas9 蛋白的至少一个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 双链切割活性，但保留 DNA 单链切割活性或者仅保留结合活性。
 23. 根据权利要求 22 所述的系统，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A, H840A, N854A 或 N863A 的突变。
 24. 根据权利要求 14-20 中任一项所述的系统，其中所述 mini-Cas9 蛋白的至少两个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 切割活性。
 25. 根据权利要求 24 所述的系统，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A 的突变和对应于化脓链球菌 Cas9 的 H840A, N854A 或 N863A 的突变。
 26. 根据权利要求 14-25 中任一项所述的系统，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。
 27. 根据权利要求 26 所述的系统，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。
 28. 根据权利要求 26-27 中任一项所述的系统，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录阻抑活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。
 29. 根据权利要求 14-28 中任一项所述的系统，其中编码所述 mini-Cas9 蛋白的所述多核苷酸被密码子优化以在哺乳动物细胞或人细胞中表达。
 30. 根据权利要求 14-29 中任一项所述的系统，其中所述系统包括编码所述 mini-Cas9 蛋白的 mRNA。

31. 根据权利要求 16-30 中任一项所述的系统，其中所述一个或多个载体是病毒载体，任选地其中组分(a)和(b)位于同一载体上。
32. 根据权利要求 16-31 中任一项所述的系统，其中所述一个或多个载体是腺相关病毒(AAV)载体。
33. 权利要求 8-13 中任一项所述的缩减型 Cas 蛋白，权利要求 1-6 所述的核酸分子，权利要求 7 所述的载体和/或权利要求 14-32 中任一项所述的系统在制备用于基因疗法的药物中的用途。
34. 权利要求 8-13 中任一项所述的缩减型 Cas 蛋白，权利要求 1-6 所述的核酸分子，权利要求 7 所述的载体和/或权利要求 14-32 中任一项所述的系统用于离体或体外修饰细胞的用途。
35. 细胞，其包含权利要求 8-13 中任一项所述的缩减型 Cas 蛋白，权利要求 1-6 所述的核酸分子，权利要求 7 所述的载体和/或权利要求 14-32 中任一项所述的系统。
36. 根据权利要求 35 所述的细胞，所述细胞包括真核细胞。
37. 一种非人类生物，其包含权利要求 35-36 中任一项所述的细胞，其中所述非人类生物不是动物或植物品种。
38. 一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 系统引入真核细胞，所述系统包括：
 - (a) 权利要求 8-13 中任一项所述的 mini-Cas9 蛋白或编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) CRISPR-Cas 系统嵌合 RNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列，从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。
39. 一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 系统引入真核细胞，所述系统包括：
 - (a) 权利要求 8-13 中任一项所述的 mini-Cas9 蛋白或编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
 - (b) CRISPR-Cas 系统 crRNA，其包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和

- (c)包含 tracr 序列的 tracrRNA，
从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。
40. 一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 载体系统引入真核细胞，所述 CRISPR-Cas 载体系统包括一个或多个载体，所述载体包含：
- (a) 编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
(b) 编码 CRISPR-Cas 系统嵌合 RNA 的多核苷酸，其中所述嵌合 RNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列，能够与 tracr 序列杂交的 tracr 配对序列，和 tracr 序列，
从而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。
41. 一种用于产生修饰的真核细胞的离体或体外方法，其包括将工程化的、非天然存在的 CRISPR-Cas 载体系统引入真核细胞，所述 CRISPR-Cas 载体系统包括一个或多个载体，所述载体包含：
- (a) 编码权利要求 8-13 中任一项所述 mini-Cas9 蛋白的多核苷酸；
(b) 编码 CRISPR-Cas 系统 crRNA 的多核苷酸，其中所述 crRNA 包含能够与真核细胞的核中与原型间隔子邻近基序(PAM)相邻的靶序列杂交并引导 CRISPR-Cas 复合物与所述靶序列来序列特异性结合的指导序列和能够与 tracr 序列杂交的 tracr 配对序列；和
(c) 编码 tracrRNA 的多核苷酸，其中所述 tracrRNA 包含 tracr 序列，
而获得修饰的真核细胞，其中所述真核细胞不是人生殖细胞。
42. 根据权利要求 38-41 中任一项所述的方法，其中所述 mini-Cas9 蛋白的长度为约 900-1300 个氨基酸。
43. 根据权利要求 38-42 中任一项所述的方法，其中所述 tracr 序列的长度为至少 30、40 或 50 个核苷酸。
44. 根据权利要求 38-43 中任一项所述的方法，其中所述指导序列的长度为 15-30 个核苷酸。
45. 根据权利要求 38-44 中任一项所述的方法，其中所述 mini-Cas9 蛋白是能够切割两条 DNA 链的核酸酶。
46. 根据权利要求 38-44 中任一项所述的方法，其中所述 mini-Cas9 蛋白的至少一个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 双链切割活性，但保留单链切割活性或

- 仅保留 DNA 结合活性。
47. 根据权利要求 46 所述的方法，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A，H840A，N854A 或 N863A 的突变。
 48. 根据权利要求 38-44 中任一项所述的方法，其中所述 mini-Cas9 蛋白的至少两个催化结构域被突变，使得所述 Cas9 蛋白基本上缺乏 DNA 切割活性。
 49. 根据权利要求 48 所述的方法，其中所述 mini-Cas9 蛋白包含对应于化脓链球菌 Cas9 的 D10A 的突变和对应于化脓链球菌 Cas9 的 H840A，N854A 或 N863A 的突变。
 50. 根据权利要求 38-49 中任一项所述的方法，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。
 51. 根据权利要求 50 所述的方法，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。
 52. 根据权利要求 50-51 中任一项所述的方法，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录阻抑活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。
 53. 根据权利要求 38-52 中任一项所述的方法，其中编码所述 mini-Cas9 蛋白的所述多核苷酸被密码子优化以在哺乳动物细胞或人细胞中表达。
 54. 根据权利要求 53 所述的方法，其中所述系统包括编码所述 mini-Cas9 蛋白的 mRNA。
 55. 根据权利要求 40-54 中任一项所述的方法，其中所述一个或多个载体是病毒载体，任选地其中组分(a)和(b)位于同一载体上。
 56. 根据权利要求 40-55 中任一项所述的方法，其中所述一个或多个载体是腺相关病毒(AAV)载体。
 57. 根据权利要求 38-56 中任一项所述的方法，其中所述真核细胞是哺乳动物细胞。
 58. 根据权利要求 38-57 中任一项所述的方法，其进一步包括从所述修饰的真核细胞或其后代产生基因修饰的非人类动物或植物。
 59. 通过根据权利要求 38-58 中任一项所述的方法产生的修饰的真核细胞或其后代，其中所述修饰的真核细胞或其后代包含靶基因组基因座中与相应的野生型细胞相比的一个或多个核苷酸的插入，缺失或置换。
 60. 通过根据权利要求 38-58 中任一项所述的方法生产的基因修饰的非人类动物或植物，其中所述非人类动物或植物不是动物或植物品种，并且其中所述修饰的真核细胞或其后代包含靶基因组基因座中与相应的野生型细胞相比的一个或多个核苷酸的插入，缺失或置

换。

61. 一种改变细胞的基因组的方法，该方法包括在该细胞中表达权利要求 8-13 中任一项所述的缩减型 Cas 蛋白以及具有与该细胞的基因组的所选部分互补的区域的指导 RNA，或使该细胞与所述的缩减型 Cas 蛋白以及具有与该细胞的基因组的所选部分互补的区域的指导 RNA 接触。
62. 根据权利要求 61 所述的方法，其中该细胞是干细胞。
63. 根据权利要求 62 所述的方法，其中该细胞是胚胎干细胞、间充质干细胞或诱导多能干细胞；处于活着的动物中；或处于胚胎中。
64. 一种改变双链 DNA(dsDNA)分子的方法，该方法包括使该 dsDNA 分子与权利要求 8-13 中任一项所述的缩减型 Cas 蛋白以及具有与该 dsDNA 分子的所选部分互补的区域的指导 RNA 接触。
65. 根据权利要求 64 所述的方法，其中该 dsDNA 分子是在体外。
66. 根据权利要求 61-65 中任一项所述的方法，其中所述 mini-Cas9 蛋白与至少一个异源蛋白质结构域连接。
67. 根据权利要求 66 所述的方法，其中所述异源蛋白质结构域具有表观遗传修饰和/或转录调控的功能。
68. 根据权利要求 66-67 中任一项所述的方法，其中所述异源蛋白质结构域具有以下活性中的一种或多种：脱氨酶活性、甲基酶活性、脱甲基酶活性、转录激活活性、转录阻抑活性、转录释放因子活性、组蛋白修饰活性、RNA 切割活性和核酸结合活性。

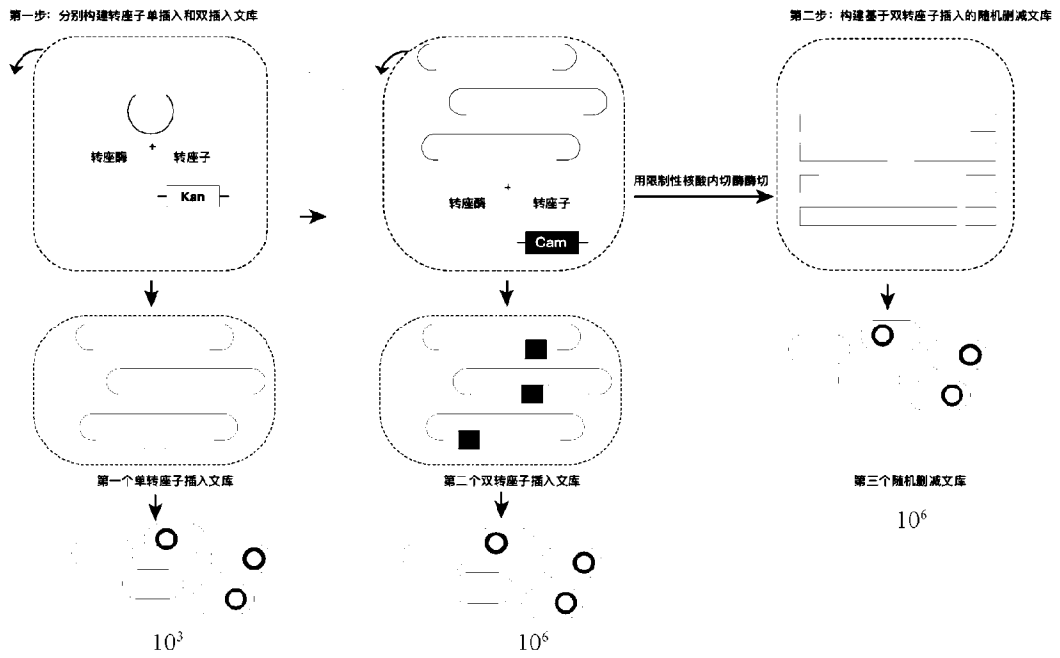


图 1



图 2A

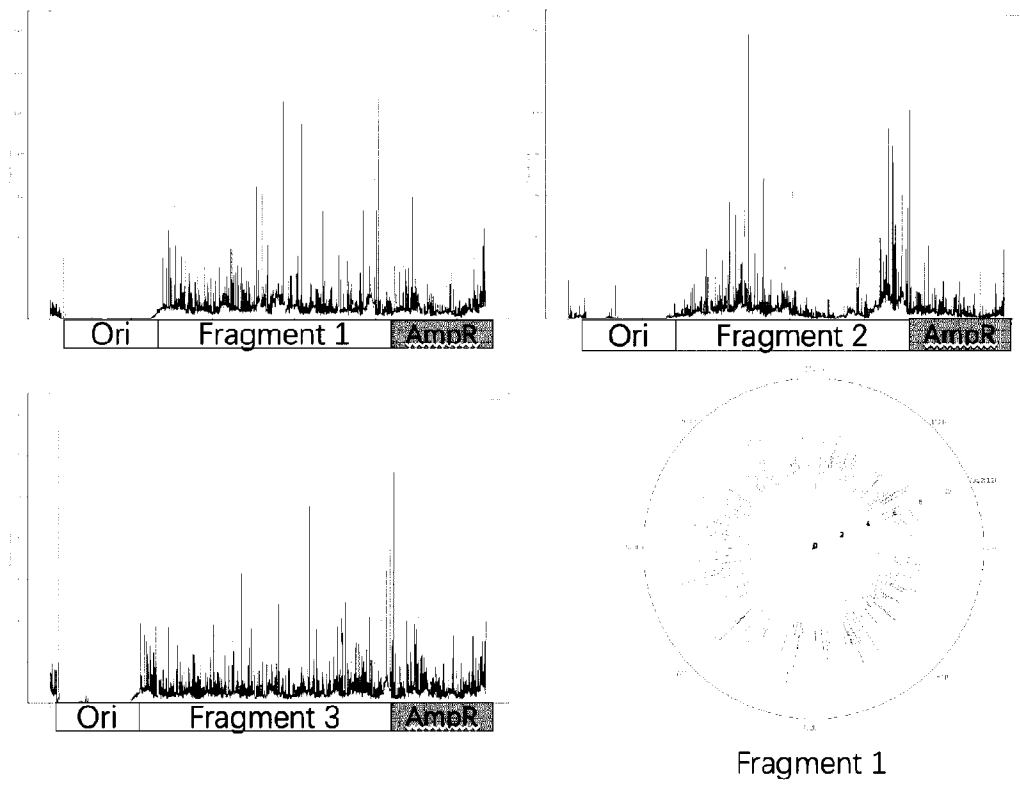


图 2B

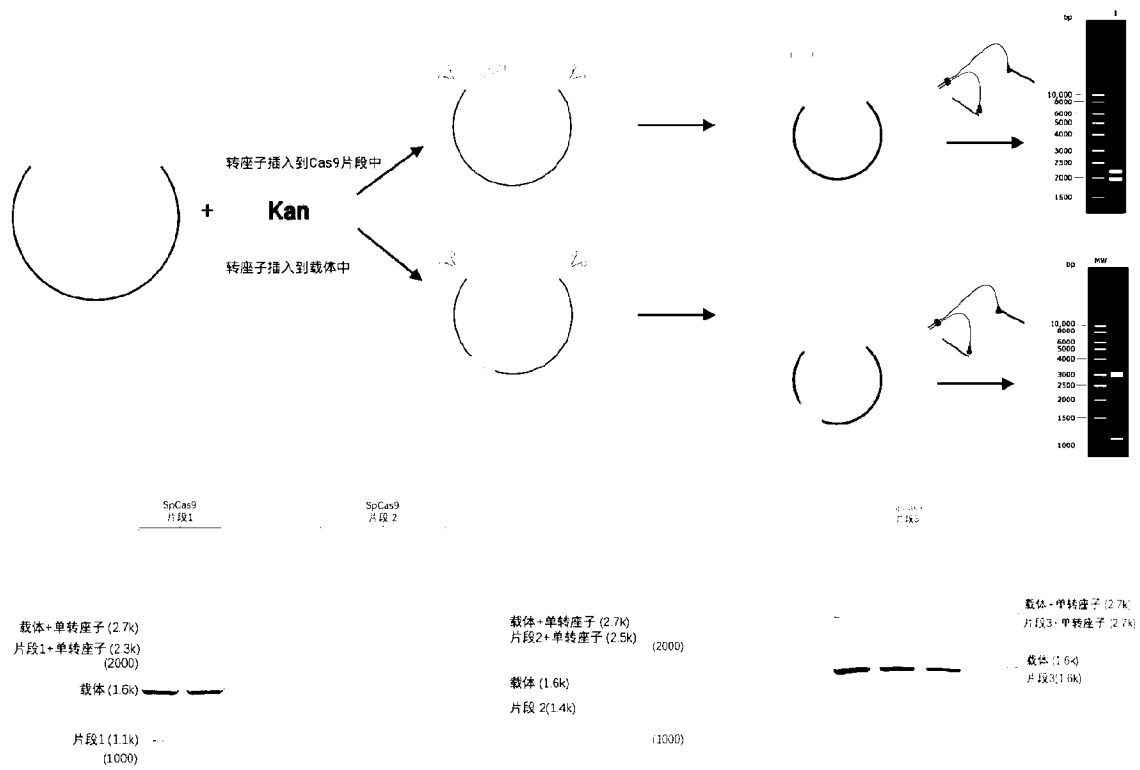


图 2C

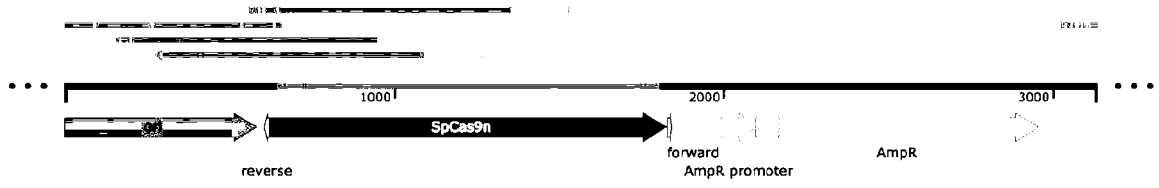


图 3A

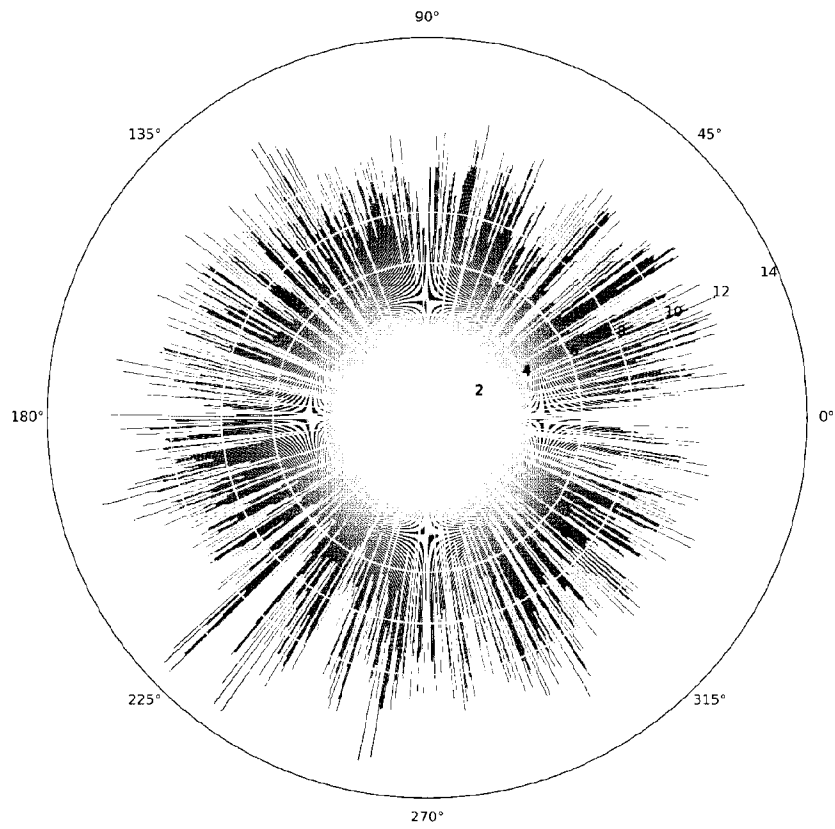


图 3B

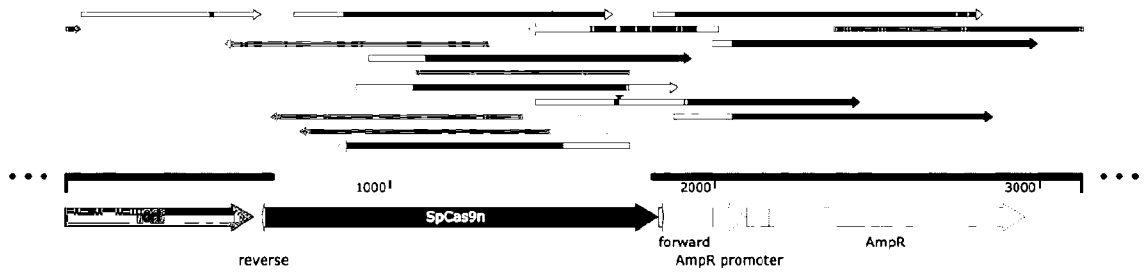
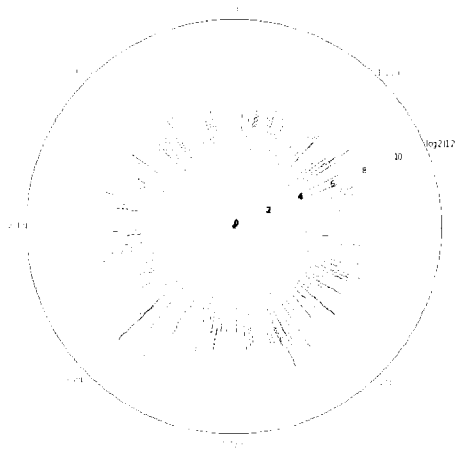
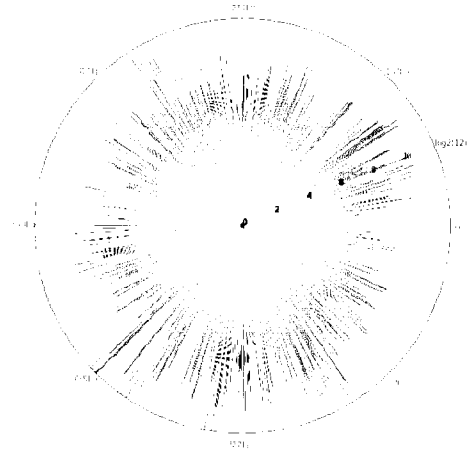


图 4A



Kan抗性转座子插入分布图



Cam抗性转座子插入分布图

图 4B

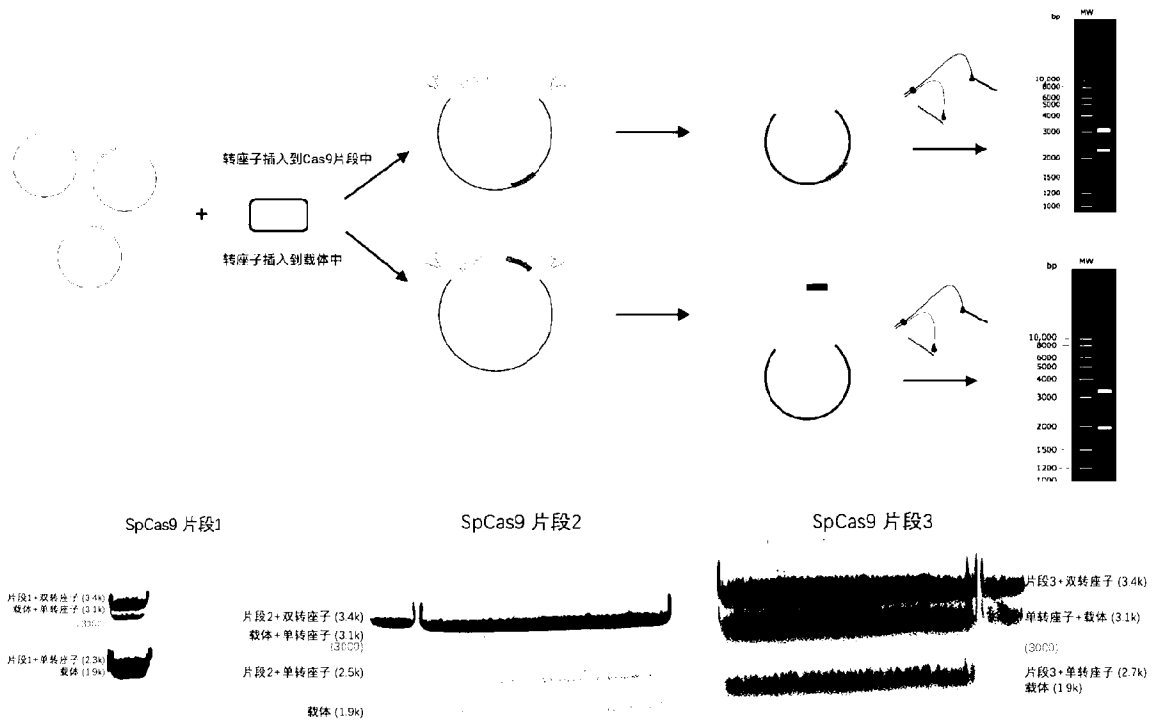


图 4C

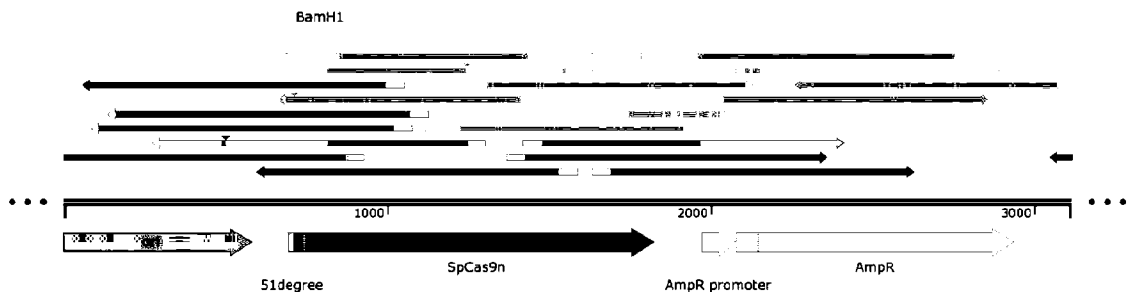
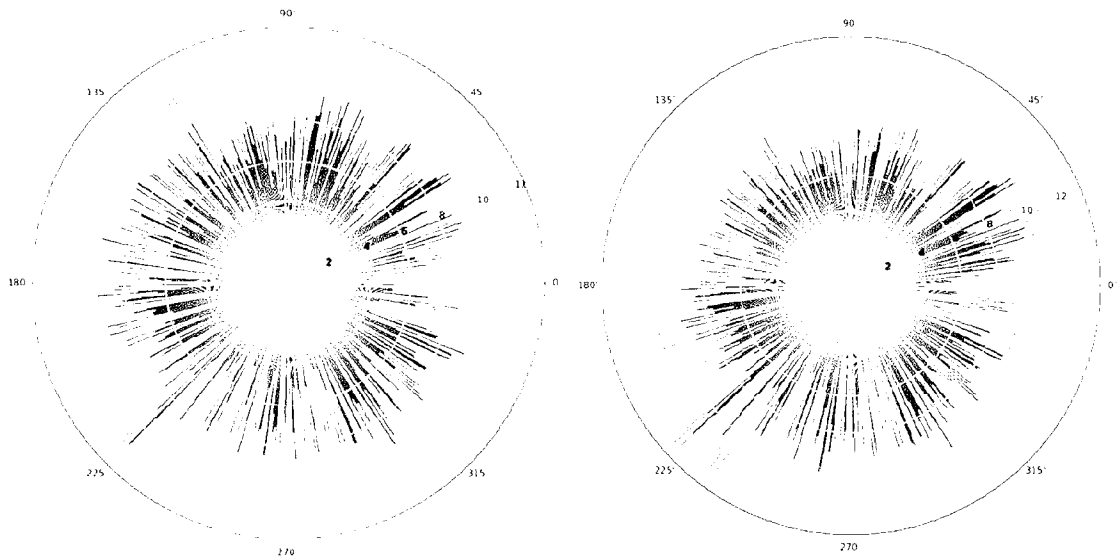


图 5A



Cam抗性转座子插入分布图

Kan抗性转座子插入分布图

图 5B

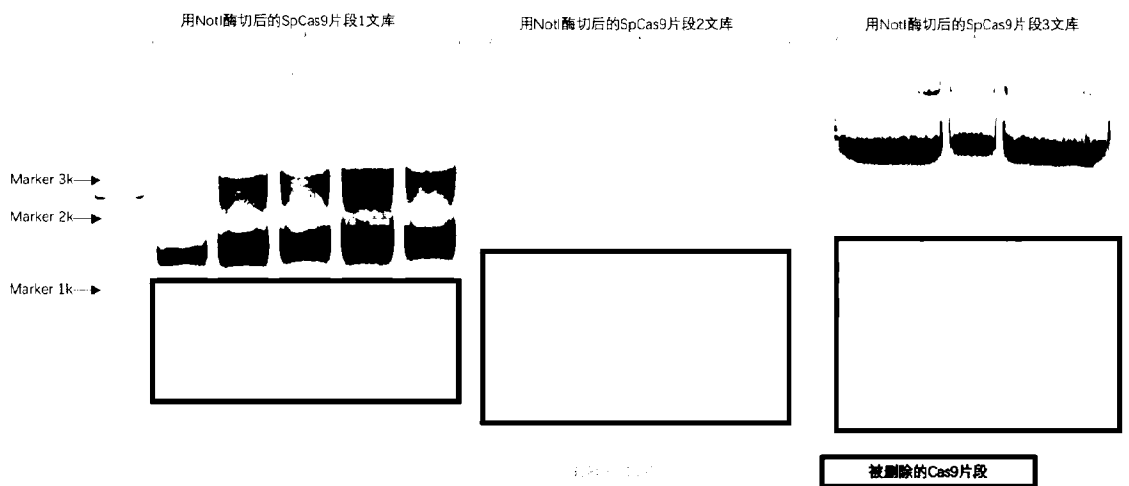


图 5C

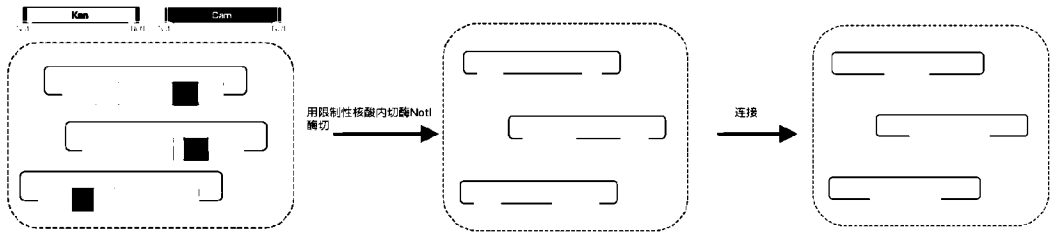


图 6A

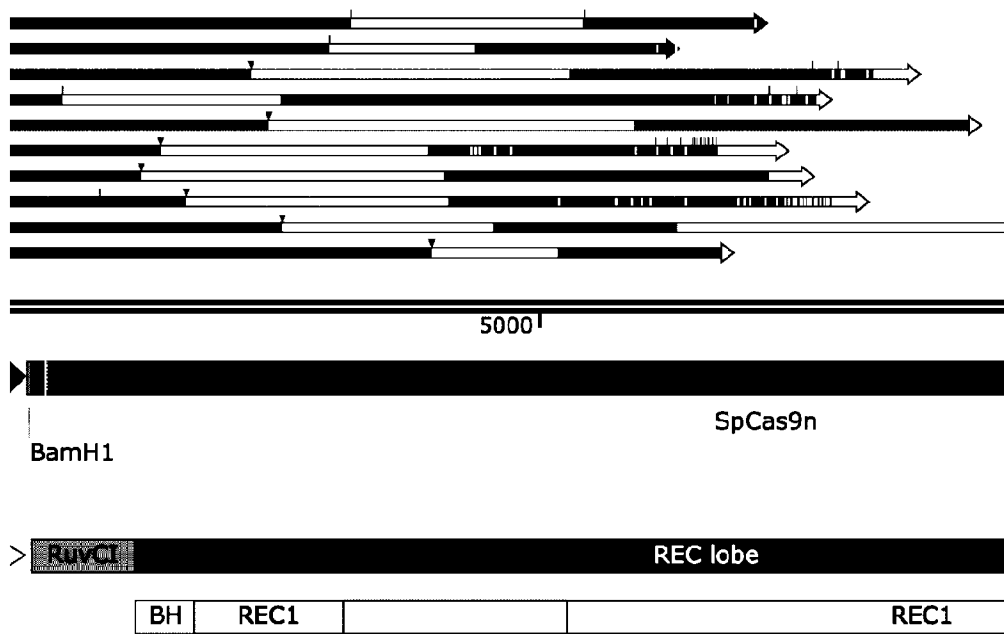


图 6B

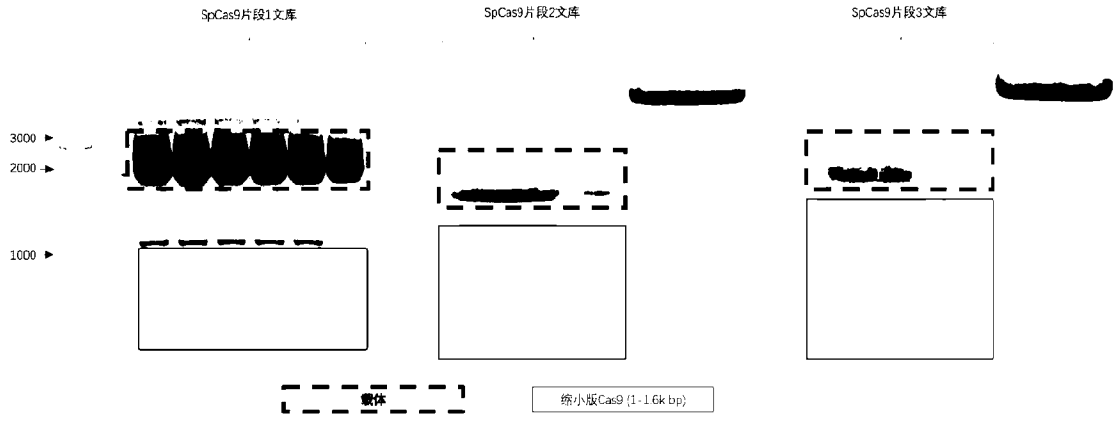


图 6C

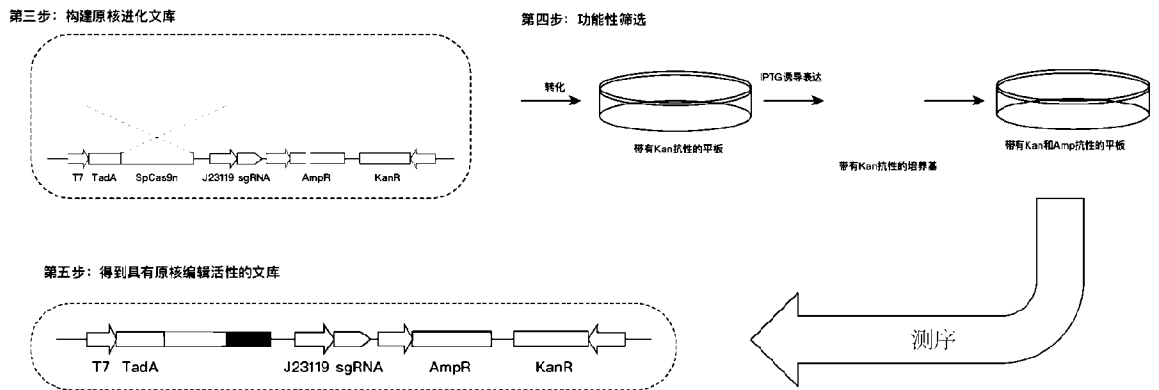


图 7A

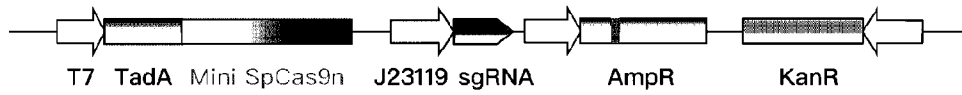


图 7B

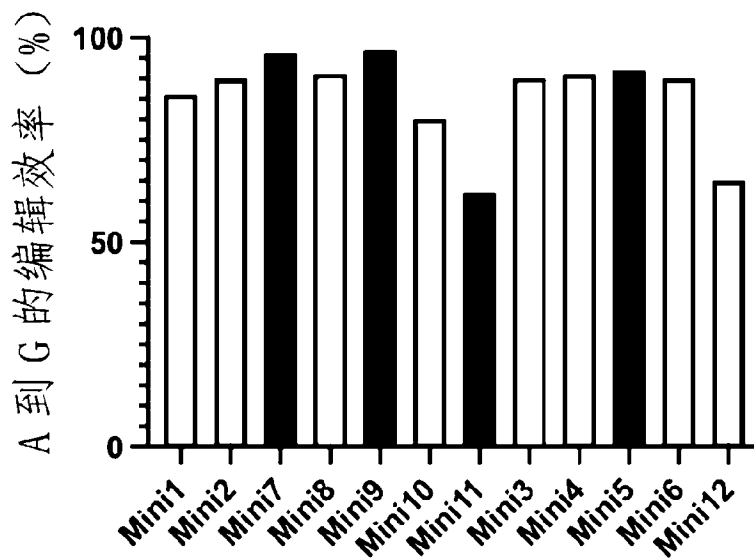


图 7C

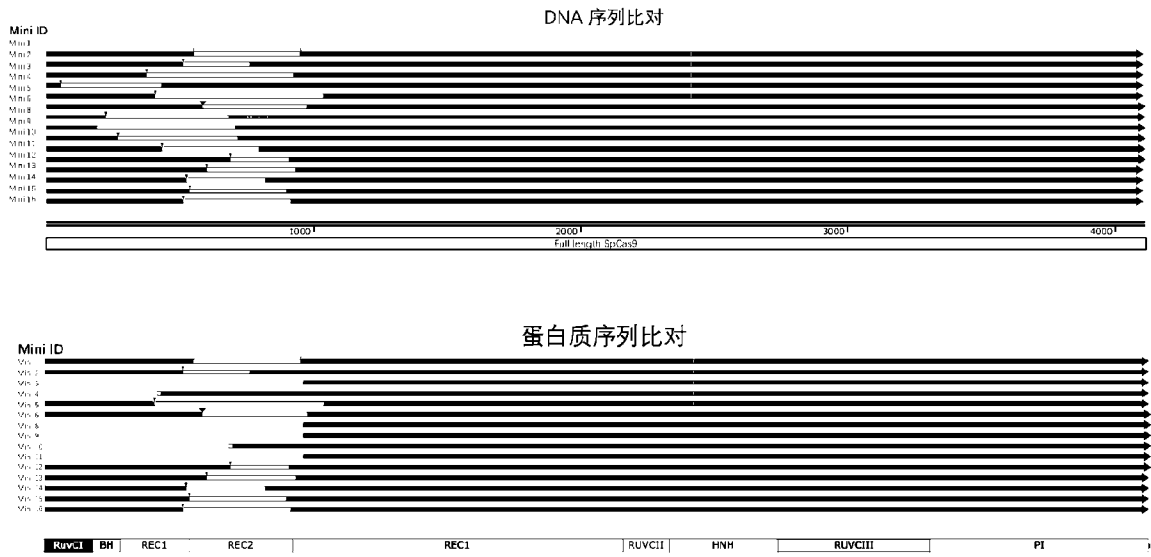


图 8

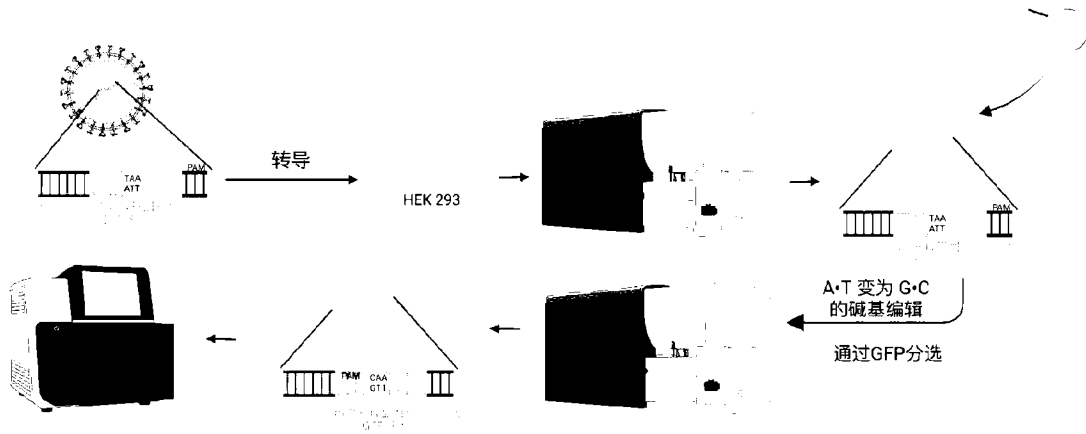


图 9A



图 9B

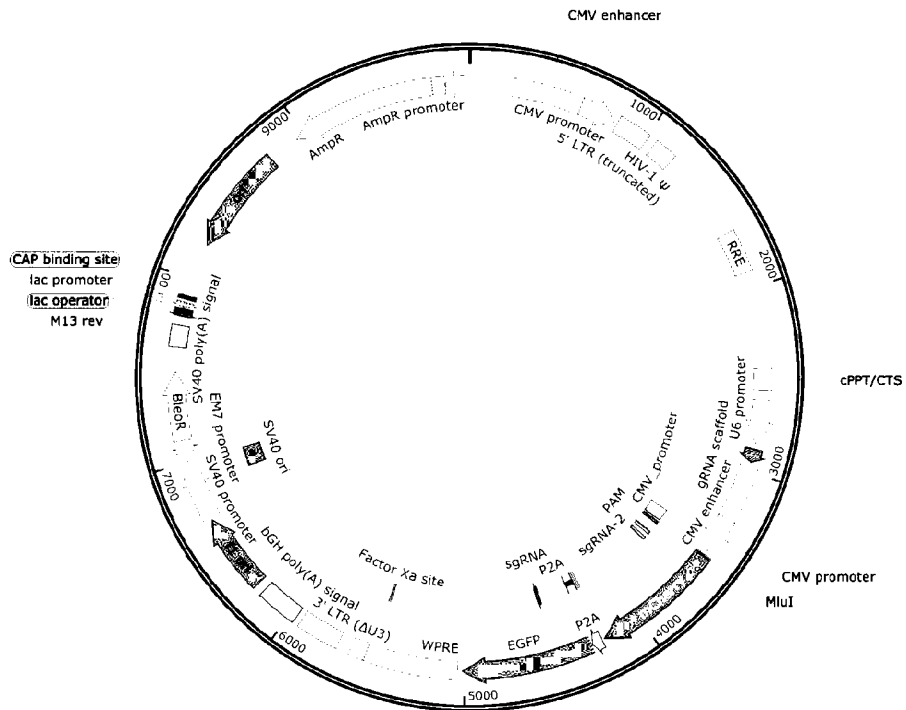


图 9C

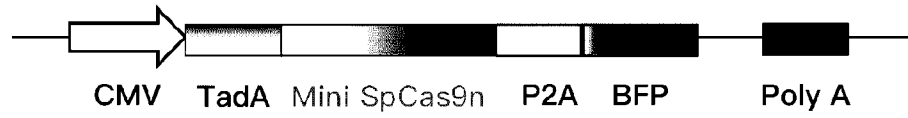


图 9D

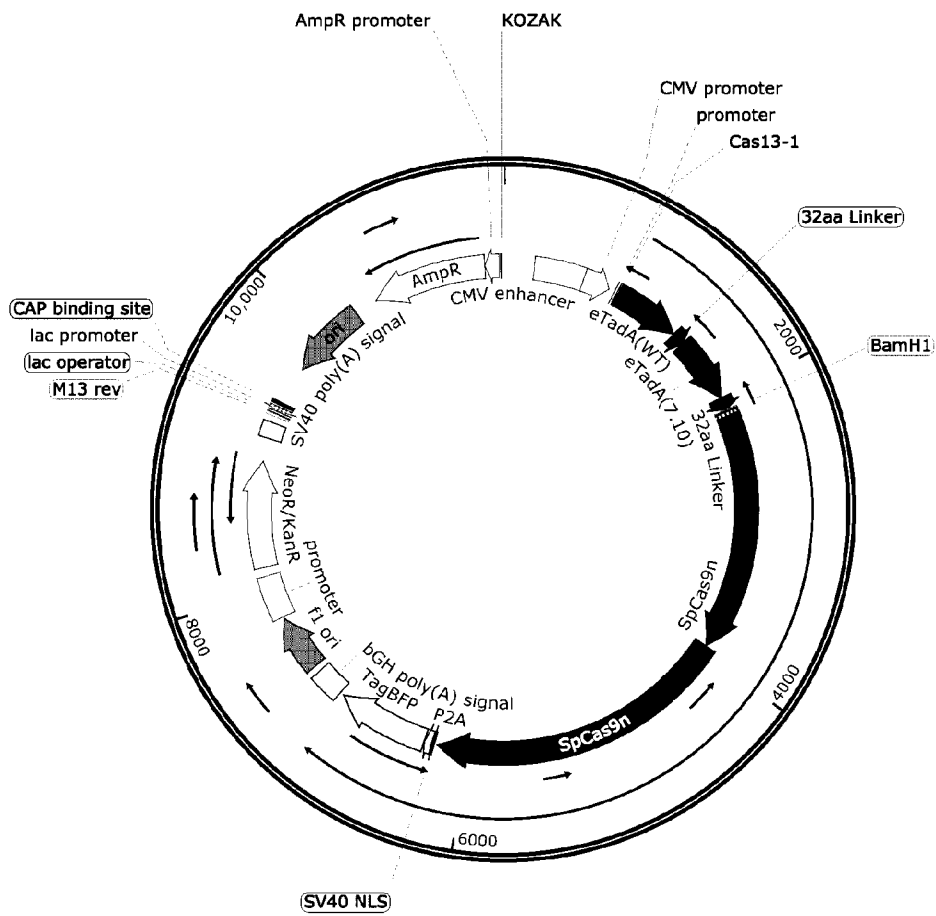


图 9E

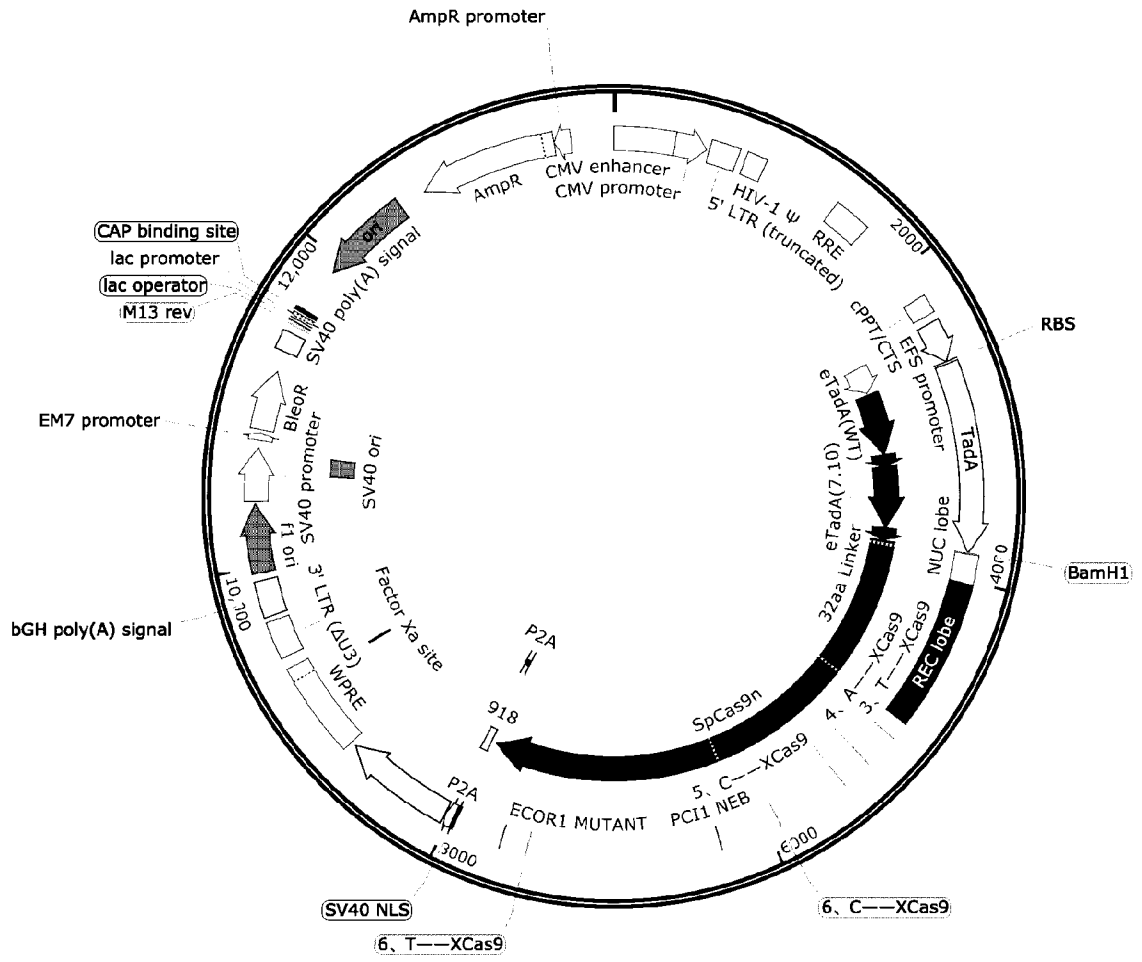


图 9F

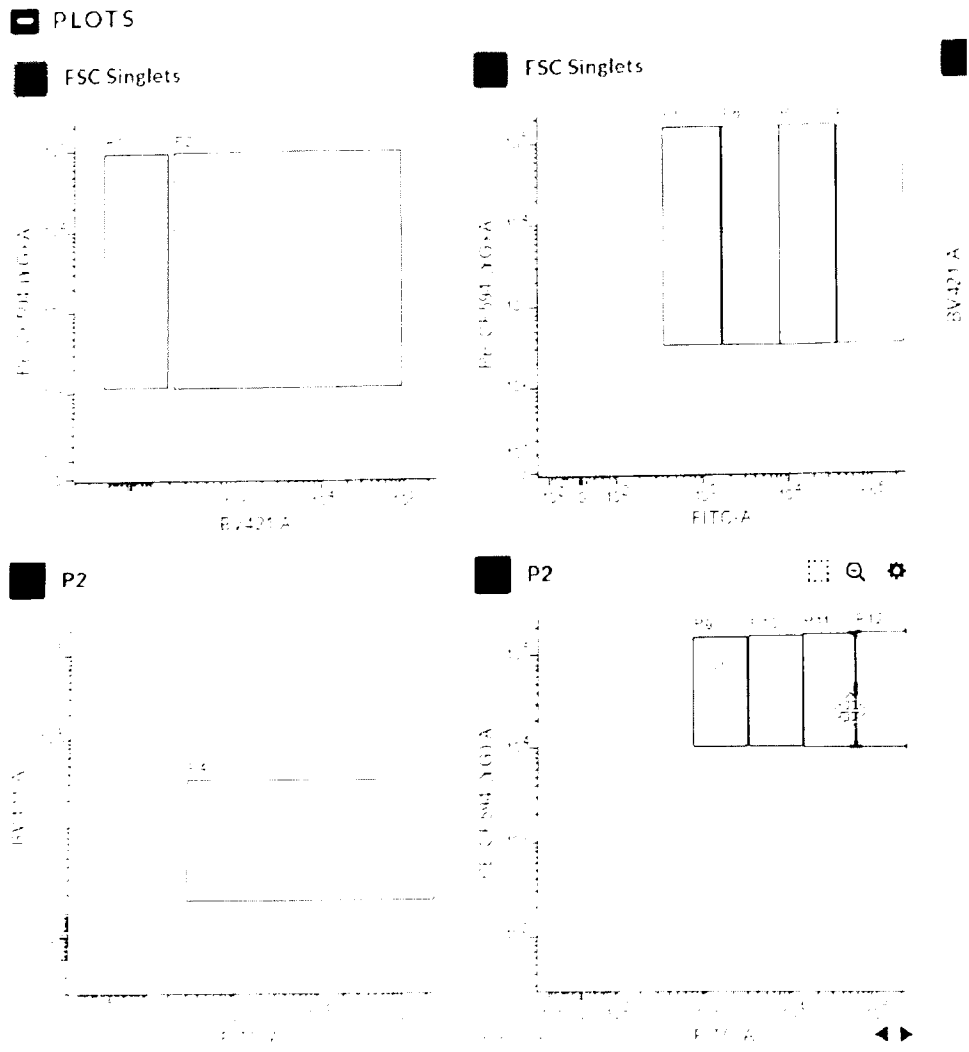


图 9G

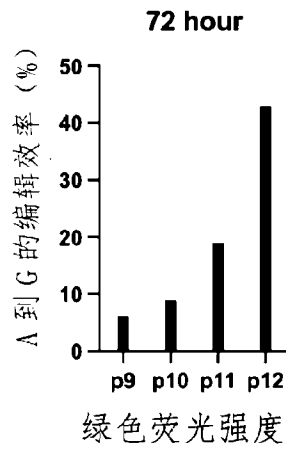


图 9H

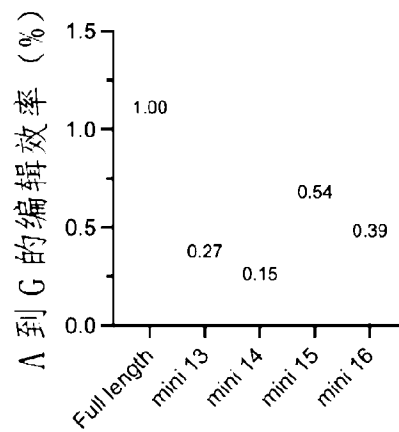
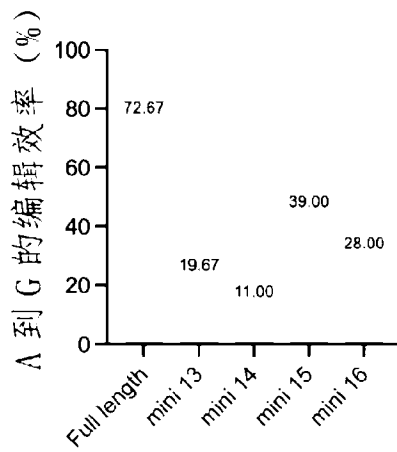


图 9I

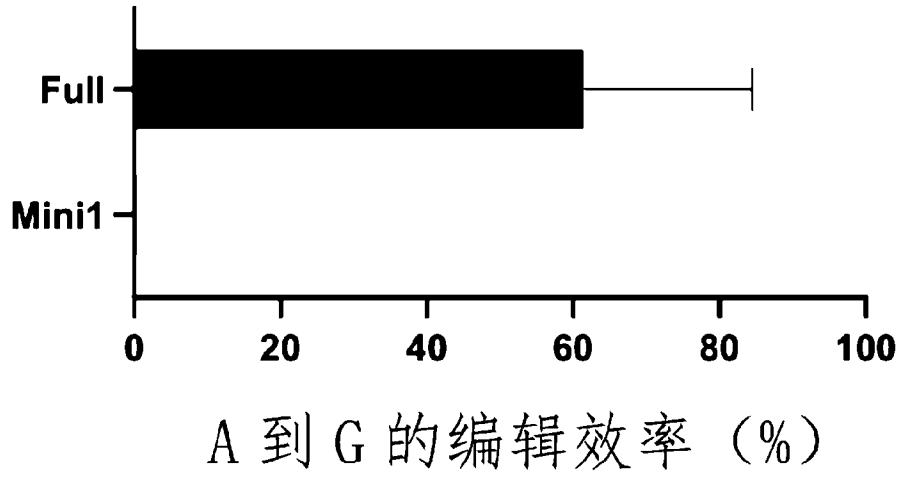


图 10

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2023/079545

A. CLASSIFICATION OF SUBJECT MATTER		
C12N9/22(2006.01)i;C12N15/11(2006.01)i;C12N15/86(2006.01)i;C07K14/00(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,C07K		
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) DWPI, VEN, SIPOABS, CNABS, CNTXT, USTXT, EPTXT, WOTXT, 万方数据库, WANFANG DATA, CNKI, PubMed, ISI web of Knowledge, GenBank, EBI, STN: Cas9, mini, delete, deletion, RuvC1, REC1, region, 缺失, 删除, 缩小, 缩短, 接头, 结合, 化脓链球菌, Streptococcus pyogenes, SEQ ID NOs: 1-28, 申请人, 发明人		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2018163188 A1 (TSINGHUA UNIVERSITY) 14 June 2018 (2018-06-14) claims 1-20, and description, figures 1-7	1-5, 7-12, 14-68
Y	CN 110662835 A (TSINGHUA UNIVERSITY) 07 January 2020 (2020-01-07) claims 1-20	1-5, 7-12, 14-68
A	CN 106011104 A (TSINGHUA UNIVERSITY) 12 October 2016 (2016-10-12) entire document	1-68
A	WO 2020236967 A1 (THE BROAD INSTITUTE, INC. et al.) 26 November 2020 (2020-11-26) entire document	1-68
A	WO 2020243085 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK) 03 December 2020 (2020-12-03) entire document	1-68
A	WO 2020206036 A1 (THE BROAD INSTITUTE, INC. et al.) 08 October 2020 (2020-10-08) entire document	1-68
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "D" document cited by the applicant in the international application "E" earlier application or patent but published on or after the international filing date "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) "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 "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 "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 "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 "&" document member of the same patent family		
Date of the actual completion of the international search 10 May 2023		Date of mailing of the international search report 24 May 2023
Name and mailing address of the ISA/CN China National Intellectual Property Administration (ISA/ CN) China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088		Authorized officer Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	SHAMS, A. et al. "Comprehensive deletion landscape of CRISPR-Cas9 identifies minimal RNA-guided DNA-binding modules." <i>Nature Communications</i> , Vol. 12, 27 September 2021 (2021-09-27), page 3, right-hand column, last paragraph to page 4, left-hand column, paragraph 2, and figures 1-3	1-5, 7-12, 14-68
Y	MA, Dacheng et al. "Rational Design of Mini-Cas9 for Transcriptional Activation." <i>ACS Synthetic Biology</i> , Vol. 7, 21 March 2018 (2018-03-21), pages 978-985 pages 978-985	1-5, 7-12, 14-68
Y	ZHANG, Yifei et al. "Catalytic-state structure and engineering of <i>Streptococcus thermophilus</i> Cas9." <i>NATURE CATALYSIS</i> , Vol. 3, 31 October 2020 (2020-10-31), pages 813-823	1-5, 7-12, 14-68

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:
 - a. forming part of the international application as filed.
 - b. furnished subsequent to the international filing date for the purposes of international search (Rule 13ter.1(a)),
 accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2023/079545

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2018163188	A1	14 June 2018	None			
CN	110662835	A	07 January 2020	WO	2018209712	A1	22 November 2018
				EP	3625338	A1	25 March 2020
				EP	3625338	A4	20 January 2021
CN	106011104	A	12 October 2016	US	2017233703	A1	17 August 2017
WO	2020236967	A1	26 November 2020	US	2022235340	A1	28 July 2022
WO	2020243085	A1	03 December 2020	None			
WO	2020206036	A1	08 October 2020	None			

<p>A. 主题的分类</p> <p>C12N9/22(2006.01)i;C12N15/11(2006.01)i;C12N15/86(2006.01)i;C07K14/00(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. 检索领域</p> <p>检索的最低限度文献(标明分类系统和分类号)</p> <p>C12N, C07K</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>DWPI, VEN, SIPOABS, CNABS, CNTXT, USTXT, EPTXT, WOTXT, 万方数据库, CNKI, PubMed, ISI web of Knowledge, GenBank, EBI, STN: Cas9, mini, delete, deletion, RuvC1, REC1, region, 缺失, 删除, 缩小, 缩短, 接头, 结合, 化脓链球菌, Streptococcus pyogenes, SEQ ID NOs:1-28, 申请人, 发明人</p>																							
<p>C. 相关文件</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>US 2018163188 A1 (TSINGHUA UNIVERSITY) 2018年6月14日 (2018 - 06 - 14) 权利要求1-20, 说明书附图1-7</td> <td>1-5, 7-12, 14-68</td> </tr> <tr> <td>Y</td> <td>CN 110662835 A (清华大学) 2020年1月7日 (2020 - 01 - 07) 权利要求1-20</td> <td>1-5, 7-12, 14-68</td> </tr> <tr> <td>A</td> <td>CN 106011104 A (清华大学) 2016年10月12日 (2016 - 10 - 12) 全文</td> <td>1-68</td> </tr> <tr> <td>A</td> <td>WO 2020236967 A1 (THE BROAD INSTITUTE, INC. 等) 2020年11月26日 (2020 - 11 - 26) 全文</td> <td>1-68</td> </tr> <tr> <td>A</td> <td>WO 2020243085 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK.) 2020年12月3日 (2020 - 12 - 03) 全文</td> <td>1-68</td> </tr> <tr> <td>A</td> <td>WO 2020206036 A1 (THE BROAD INSTITUTE, INC. 等) 2020年10月8日 (2020 - 10 - 08) 全文</td> <td>1-68</td> </tr> </tbody> </table> <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: “A” 认为不特别相关的表示了现有技术一般状态的文件 “D” 申请人在国际申请中引证的文件 “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</p>			类型*	引用文件, 必要时, 指明相关段落	相关的权利要求	Y	US 2018163188 A1 (TSINGHUA UNIVERSITY) 2018年6月14日 (2018 - 06 - 14) 权利要求1-20, 说明书附图1-7	1-5, 7-12, 14-68	Y	CN 110662835 A (清华大学) 2020年1月7日 (2020 - 01 - 07) 权利要求1-20	1-5, 7-12, 14-68	A	CN 106011104 A (清华大学) 2016年10月12日 (2016 - 10 - 12) 全文	1-68	A	WO 2020236967 A1 (THE BROAD INSTITUTE, INC. 等) 2020年11月26日 (2020 - 11 - 26) 全文	1-68	A	WO 2020243085 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK.) 2020年12月3日 (2020 - 12 - 03) 全文	1-68	A	WO 2020206036 A1 (THE BROAD INSTITUTE, INC. 等) 2020年10月8日 (2020 - 10 - 08) 全文	1-68
类型*	引用文件, 必要时, 指明相关段落	相关的权利要求																					
Y	US 2018163188 A1 (TSINGHUA UNIVERSITY) 2018年6月14日 (2018 - 06 - 14) 权利要求1-20, 说明书附图1-7	1-5, 7-12, 14-68																					
Y	CN 110662835 A (清华大学) 2020年1月7日 (2020 - 01 - 07) 权利要求1-20	1-5, 7-12, 14-68																					
A	CN 106011104 A (清华大学) 2016年10月12日 (2016 - 10 - 12) 全文	1-68																					
A	WO 2020236967 A1 (THE BROAD INSTITUTE, INC. 等) 2020年11月26日 (2020 - 11 - 26) 全文	1-68																					
A	WO 2020243085 A1 (THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK.) 2020年12月3日 (2020 - 12 - 03) 全文	1-68																					
A	WO 2020206036 A1 (THE BROAD INSTITUTE, INC. 等) 2020年10月8日 (2020 - 10 - 08) 全文	1-68																					
国际检索实际完成的日期	2023年5月10日	国际检索报告邮寄日期	2023年5月24日																				
ISA/CN的名称和邮寄地址	中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	授权官员	黎舒婷 电话号码 (+86) 010-53961957																				

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	SHAMS, A. 等. "Comprehensive deletion landscape of CRISPR-Cas9 identifies minimal RNA-guided DNA-binding modules." NATURE COMMUNICATIONS., 第12卷, 2021年9月27日 (2021 - 09 - 27), 第3页右栏最后1段至第4页左栏第2段, 图1-3	1-5, 7-12, 14-68
Y	MA, D. C. 等. "Rational Design of Mini-Cas9 for Transcriptional Activation." ACS Synthetic Biology, 第7卷, 2018年3月21日 (2018 - 03 - 21), 第978-985页 第978-985页	1-5, 7-12, 14-68
Y	ZHANG Y. F. 等. "Catalytic-state structure and engineering of Streptococcus thermophilus Cas9." NATURE CATALYSIS., 第3卷, 2020年10月31日 (2020 - 10 - 31), 第813-823页	1-5, 7-12, 14-68

第1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核苷酸和/或氨基酸序列, 国际检索是基于下列序列列表进行的:
- a. 作为国际申请的一部分提交的;
- b. 为国际检索的目的在国际申请日之后提交(细则13之三.1(a)),
 附有说明序列列表不超出所提交国际申请公开范围的声明。
2. 本报告是在没有收到符合WIPO ST. 26标准的序列列表的情况下, 考虑了国际申请中披露的任何核苷酸和/或氨基酸序列, 在可进行有意义检索的范围内做出的。
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/CN2023/079545

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)
US	2018163188	A1	2018年6月14日	无			
CN	110662835	A	2020年1月7日	WO	2018209712	A1	2018年11月22日
				EP	3625338	A1	2020年3月25日
				EP	3625338	A4	2021年1月20日
CN	106011104	A	2016年10月12日	US	2017233703	A1	2017年8月17日
WO	2020236967	A1	2020年11月26日	US	2022235340	A1	2022年7月28日
WO	2020243085	A1	2020年12月3日	无			
WO	2020206036	A1	2020年10月8日	无			