

(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
2019 年 3 月 7 日 (07.03.2019)



(10) 国际公布号
WO 2019/041296 A1

(51) 国际专利分类号:
C12N 15/79 (2006.01)

(21) 国际申请号: PCT/CN2017/100131

(22) 国际申请日: 2017 年 9 月 1 日 (01.09.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 上海科技大学 (SHANGHAITECH UNIVERSITY) [CN/CN]; 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。

(72) 发明人: 陈佳(CHEN, Jia); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。黄行许(HUANG, Xingxu); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。杨力(YANG, Li); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。杨贝(YANG, Bei); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。李潇飒(LI, Xiaosa); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。王滢(WANG, Ying); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。刘亚京

(LIU, Yajing); 中国上海市浦东新区海科路100号11号楼411A, Shanghai 201210 (CN)。

(74) 代理人: 上海申汇专利代理有限公司(SHANGHAI SHENHUI PATENT AGENT CO., LTD.); 中国上海市徐汇区漕宝路80号405室, Shanghai 200235 (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,

(54) Title: BASE EDITING SYSTEM AND METHOD

(54) 发明名称: 一种碱基编辑系统及方法

(57) Abstract: A base editing system and method constructed based on a CRISPR/Cpf1 system are provided. The base editing system utilizes Cpf1 to recognize features of a 5'-TTTN-3' (PAM) sequence, and carries out base editing in regions rich in adenine and thymine in genomic DNA. The base editing system comprises at least one of a Cpf1-BE0 expression vector, a Cpf1-BE expression vector, a Cpf1-BE-YE expression vector, a Cpf1-BE-UGI expression vector, an RNA product expressed by the Cpf1-BE0 expression vector, an RNA product expressed by the Cpf1-BE expression vector, an RNA product expressed by the Cpf1-BE-YE expression vector, an RNA product expressed by the Cpf1-BE-UGI expression vector, a protein product expressed by the Cpf1-BE0 expression vector, a protein product expressed by the Cpf1-BE expression vector, a protein product expressed by the Cpf1-BE-YE expression vector, and a protein product expressed by the Cpf1-BE-UGI expression vector. A method for carrying out base editing in the genome of a eukaryote is also provided.

(57) 摘要: 提供了一种基于CRISPR/Cpf1系统构造的碱基编辑系统及方法, 该碱基编辑系统利用Cpf1识别5'-TTTN-3'(PAM)序列的特点, 在基因组DNA中腺嘌呤、胸腺嘧啶富集区域中实现碱基编辑。所述碱基编辑系统包括Cpf1-BE0表达载体、Cpf1-BE表达载体、Cpf1-BE-YE表达载体、Cpf1-BE-UGI表达载体、Cpf1-BE0表达载体表达出的RNA产物、Cpf1-BE表达载体表达出的RNA产物、Cpf1-BE-YE表达载体表达出的RNA产物、Cpf1-BE-UGI表达载体表达出的RNA产物、Cpf1-BE0表达载体表达出的蛋白产物、Cpf1-BE表达载体表达出的蛋白产物、Cpf1-BE-YE表达载体表达出的蛋白产物、Cpf1-BE-UGI表达载体表达出的蛋白产物中的至少一种。还提供了一种可以在真核生物的基因组中实施碱基编辑的方法。

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)。

根据细则4.17的声明:

- 发明人资格(细则4.17(iv))

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))。

一种碱基编辑系统及方法

技术领域

本发明涉及基因组编辑领域，具体涉及一种基于 CRISPR/Cpf1 的碱基编辑系统实现高效率和高保真度的碱基水平基因组编辑方法及其应用。

背景技术

基因组编辑技术指利用可设计的核酸酶（分子剪刀）通过碱基插入、缺失或置换等方式，对生物体基因组 DNA 特定片段进行改造从而达到对目标基因进行编辑的一种基因工程技术。利用基因组编辑技术对细胞进行遗传学操纵，可广泛的应用于生命科学基础研究、生物技术开发、农业技术开发以及医药研发领域。例如：直接在体内校正引起遗传疾病的基因突变，将能够从根本上治疗遗传疾病；对农作物进行精确的基因工程改造，使其产量提高或能够抵抗环境污染或病原体的感染；对微生物基因组进行精准改造，从而促进可再生生物能源的开发等。

CRISPR/Cpf1 (Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1) 核酸酶可以识别基因组中的 5'-TTTN-3' (N 为任意碱基) Protospacer Adjacent Motif (PAM) 序列，并利用引导 RNA (guide RNA, gRNA) 可在多种细胞的基因组特定靶点定位，对其进行切割从而产生 DNA 双链断裂 (double strand breaks, DSBs)，然后利用细胞内源的 DNA 修复机制来实现编辑。根据不同 DNA 修复通路的激活，基因组编辑将会导致基因的失活或者突变的校正。通常来说，有两种主要的修复机制会被 DSB 所激活，一种是非同源末端连接 (non-homologous end joining, NHEJ)，另一种是源性介导修复 (homology-directed repair, HDR)。作为 DNA 双链断裂最主要的修复通路，NHEJ 在修复过程中能在 DSB 附近的基因组位点引入随机性碱基插入或缺失，从而导致基因的失活。与 NHEJ 相反，当 HDR 被激活时能够以外源性供体 DNA 作为模版，利用同源重组机制将外源性供体 DNA 的序列替换靶点基因组 DNA 的序列，从而完成基因突变的校正。但实际上，由于同源重组机制本身的局限性，HDR 介导的基因校正效率一直很低（通常小于 5%）。因此，极大限制了 CRISPR/Cpf1 基因组编辑工具从科研向应用的转化，尤其是在精准基因治疗方面的应用，这也

是基因编辑领域的一大难题。

为了提高基因突变的校正效率，碱基编辑器（base editor, BE）在近期内应运而生。这种碱基编辑器将 CRISPR 系统和 APOBEC 胞嘧啶脱氨酶家族成员整合在一起，可执行将胞嘧啶（cytosine, C）编辑为胸腺嘧啶（thymine, T）的功能。因此 CRISPR 系统与 DNA 结合的特异性决定了碱基编辑器对基因组 DNA 进行编辑的特异性。目前已有报道利用 *Streptococcus pyogenes* CRISPR associated protein 9（SpCas9）、*Staphylococcus aureus* CRISPR associated protein 9（SaCas9）成功构建剪辑编辑器，但是 SpCas9 和 SaCas9 在基因组中所识别的 PAM 序列都至少含有一个鸟嘌呤（guanosine, G）；并且在产生碱基编辑的同时也能引发一定程度的随机性碱基插入或缺失，从而降低了碱基编辑的保真度。

综上所述，创造能够与基因组 DNA 中腺嘌呤（adenosine, A）、胸腺嘧啶富集区域相结合并执行高保真度的新型碱基编辑器，有助于在各物种基因组更为广泛的区域中实现高精度碱基编辑，能极大地扩充我们对于碱基编辑器的应用，特别是在医疗领域进行精准基因治疗方面。

发明内容

本发明的目的在于克服现有碱基编辑器无法识别腺嘌呤、胸腺嘧啶富集区域，并且碱基编辑保真度不够高等缺点，利用 CRISPR/Cpf1 编辑系统发展出新型碱基编辑器，利用 Cpf1 识别 5'-TTTN-3'PAM 序列的特点，实现了在腺嘌呤、胸腺嘧啶富集区域进行高精确性，高保真度的定向碱基编辑。

为了达到上述目的，本发明提供了如下技术方案：

一种碱基编辑系统，其特征在于，包括 Cpf1-BE0 表达载体、Cpf1-BE 表达载体、Cpf1-BE-YE 表达载体、Cpf1-BE-UGI 表达载体、Cpf1-BE0 表达载体表达出的 RNA 产物、Cpf1-BE 表达载体表达出的 RNA 产物、Cpf1-BE-YE 表达载体表达出的 RNA 产物、Cpf1-BE-UGI 表达载体表达出的 RNA 产物、Cpf1-BE0 表达载体表达出的蛋白产物、Cpf1-BE 表达载体表达出的蛋白产物、Cpf1-BE-YE 表达载体表达出的蛋白产物、Cpf1-BE-UGI 表达载体表达出的蛋白产物中的至少一种。

优选地，所述的碱基编辑系统还包含 sgRNA 表达载体、sgRNA 表达载体表

达出的 RNA、和 sgRNA 表达载体表达出的蛋白中的至少一种。

优选地，所述的 Cpf1-BE0 表达载体的序列为 SEQ ID NO: 35，Cpf1-BE 表达载体的序列为 SEQ ID NO: 36，Cpf1-BE-YE 表达载体的序列为 SEQ ID NO: 37，Cpf1-BE-UGI 表达载体的序列为 SEQ ID NO: 39。

优选地，所述的 Cpf1-BE 表达载体的制备方法包括：制备含有 D832A、E925A 和 D1148A 突变的 dCpf1 片段，通过质粒重组至 pST1374-LbCpf1 载体中，得到 dCpf1 表达载体 pST1374-dCpf1，再与 APOBEC1 和 UGI 连接，得到 Cpf1-BE0 表达载体；然后在 APOBEC1 的 N 端引入 SV40 核定位信号序列，得到 Cpf1-BE 表达载体。

优选地，所述的 Cpf1-BE-YE 表达载体的制备方法包括：制备含有 W90Y 和 R126E 突变的 Cpf1-BE-YE 片段，连入 Cpf1-BE 表达载体中，得到 Cpf1-BE-YE 表达载体。

优选地，所述的 Cpf1-BE-UGI 表达载体的制备方法包括：利用 PCR 方法扩增 3×2A-UGI 编码序列，并连入 Cpf1-BE 表达载体，得到 Cpf1-BE-UGI 表达载体。

优选地，所述的 sgRNA 表达载体包括但不限于针对人基因组 DNA 靶点 DYRK1A 和 RUNX1 的 sgRNA。

本发明还提供了一种碱基编辑方法，其特征在于，包括采用上述的碱基编辑系统，在真核细胞中共转染 Cpf1-BE0 表达载体、Cpf1-BE 表达载体、Cpf1-BE-YE 表达载体、Cpf1-BE-UGI 表达载体、Cpf1-BE0 表达载体表达出的 RNA 产物、Cpf1-BE 表达载体表达出的 RNA 产物、Cpf1-BE-YE 表达载体表达出的 RNA 产物、Cpf1-BE-UGI 表达载体表达出的 RNA 产物、Cpf1-BE0 表达载体表达出的蛋白产物、Cpf1-BE 表达载体表达出的蛋白产物、Cpf1-BE-YE 表达载体表达出的蛋白产物、Cpf1-BE-UGI 表达载体表达出的蛋白产物中的至少一种；抽提转染后的真核细胞基因组 DNA。

优选地，所述的碱基编辑方法还包括利用针对靶点的特异性引物进行 PCR 扩增。

本发明还提供了上述的碱基编辑系统在真核生物的基因组中实施 C-to-T 碱基编辑的应用。

本发明首先公开了一系列新的基于 APOBEC1-dCpf1-UGI 融合蛋白的碱基编辑器(Cpf1-BEs), 同时公开了一种利用共表达 Cpf1-BEs 与 sgRNA 进行高效、高保真度碱基编辑的方法: 通过将 Cpf1-BE0 表达载体或 Cpf1-BE 表达载体和 sgRNA 表达载体共转染真核生物细胞(图 1A), 使其在靶点基因组 DNA 发生 C-T 碱基编辑, 并在 PCR 扩增靶点基因组 DNA 之后, 利用高通量二代 DNA 测序检测靶点基因组发生的 C-T 碱基编辑率和碱基插入或缺失率。

与现有技术相比, 本发明的有益效果:

本发明在国际上首次利用 CRISPR/Cpf1 系统构造了碱基编辑器, 在基因组 DNA 中腺嘌呤(adenosine, A)、胸腺嘧啶富集区域中实现了高精度和高保真度的碱基编辑, 从而有助于在各物种基因组更为广泛的区域中实现碱基编辑, 能极大地扩充我们对于碱基编辑器的应用, 特别是在医疗领域进行精准基因治疗方面。

本发明的碱基编辑方法, 可以在各种真核生物的基因组中实施高精度和高保真度的碱基编辑。

附图说明

图 1A: 实施例 1 共表达 sgRNA 与 Cpf1-BE0 或 Cpf1-BE 的方法示意图。

图 1B: 实施例 1 共表达 sgRNA 与 Cpf1-BE 比共表达 sgRNA 与 Cpf1-BE0 的方法在靶点序列中实现了更加高效的碱基编辑。C7、C10 代表其靶点序列中的第 7、10 位的胞嘧啶, NT (Non-transfected) 代表未转染样品。

图 1C: 实施例 1 共表达 sgRNA 与 Cpf1-BE0 或 Cpf1-BE 的方法在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。

图 1D: 实施例 1 共表达 sgRNA 与 Cpf1-BE 或 Cpf1-BE-YE 的方法示意图。

图 1E: 实施例 1 中比较 Cpf1-BE 和 Cpf1-BE-YE, Cpf1-BE-YE 在保持主位点 (C10) 编辑效率的同时, 极大降低了次位点 (C7) 的编辑效率。NT (Non-transfected) 代表未转染样品。

图 1F: 实施例 1 中比较 Cpf1-BE 和 Cpf1-BE-YE, Cpf1-BE-YE 介导的主位点与次位点编辑效率的比值明显增高, 从而增加了碱基编辑的精度。

图 1G: 实施例 1 中 Cpf1-BE 和 Cpf1-BE-YE 在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。

图 1H: 实施例 1 共表达 sgRNA 与 Cpf1-BE 或 Cpf1-BE-UGI 的方法示意图。

图 1I: 实施例 1 中比较 Cpf1-BE 和 Cpf1-BE-UGI, Cpf1-BE-UGI 介导的目的碱基 T 相对于非目的碱基 A 和 G 的比率明显增高,从而增加了碱基编辑的保真度。

图 1J: 实施例 1 中比较 Cpf1-BE 和 Cpf1-BE-UGI, Cpf1-BE-UGI 介导的碱基编辑率没有显著性下降。NT (Non-transfected) 代表未转染样品。

图 1K: 实施例 1 中 Cpf1-BE 和 Cpf1-BE-UGI 在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。

图 2A: 实施例 2 共表达 sgRNA 与 Cpf1-BE0 或 Cpf1-BE 的方法示意图。

图 2B: 实施例 2 共表达 sgRNA 与 Cpf1-BE 比共表达 sgRNA 与 Cpf1-BE0 的方法在靶点序列中实现了更加高效的碱基编辑。C7、C8、C10 代表其靶点序列中的第 7、8、10 位的胞嘧啶, NT (Non-transfected) 代表未转染样品。

图 2C: 实施例 2 共表达 sgRNA 与 Cpf1-BE0 或 Cpf1-BE 的方法在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。

*代表为转染样品中的本底性碱基插入或缺失。

图 2D: 实施例 2 共表达 sgRNA 与 Cpf1-BE 或 Cpf1-BE-YE 的方法示意图。

图 2E: 实施例 2 中比较 Cpf1-BE 和 Cpf1-BE-YE, Cpf1-BE-YE 在保持主位点 (C10) 编辑效率的同时, 极大降低了次位点 (C7、C8) 的编辑效率。NT (Non-transfected) 代表未转染样品。

图 2F: 实施例 2 中比较 Cpf1-BE 和 Cpf1-BE-YE, Cpf1-BE-YE 介导的主位点与次位点编辑效率的比值明显增高, 从而增加了碱基编辑的精度。

图 2G: 实施例 2 中 Cpf1-BE 和 Cpf1-BE-YE 在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。*代表为转染样品中的本底性碱基插入或缺失。

图 2H: 实施例 2 共表达 sgRNA 与 Cpf1-BE 或 Cpf1-BE-UGI 的方法示意图。

图 2I: 实施例 2 中比较 Cpf1-BE 和 Cpf1-BE-UGI, Cpf1-BE-UGI 介导的目的碱基 T 相对于非目的碱基 A 和 G 的比率明显增高,从而增加了碱基编辑的保真度。

图 2J: 实施例 2 中比较 Cpf1-BE 和 Cpf1-BE-UGI, Cpf1-BE-UGI 介导的碱基编辑率没有显著性下降。NT (Non-transfected) 代表未转染样品。

图 2K: 实施例 2 中 Cpf1-BE 和 Cpf1-BE-UGI 在靶点序列附近并没有引起随机性碱基插入或缺失。NT (Non-transfected) 代表未转染样品。*代表为转染样品中

的本底性碱基插入或缺失。

具体实施方式

下面将结合实施例对本发明的实施方案进行清楚、完整的描述，显然，所描述的实施例仅用于说明本发明的一部分实施例，而不应视为限制本发明的范围。实施例中未注明具体条件者，按照常规条件或制造商建议的条件进行。所用试剂或仪器未注明生产厂商者，视为可以通过市售购买获得的常规产品。

本发明的融合表达 APOBEC1, dCpf1 和 UGI 的方法，首先通过在载体 PST1374-LbCpf1-NLS 上引入 D832A, E925A 和 D1148A 突变，使之变成有结合但无切割活性的 dCpf1-NLS，再与 APOBEC1 和 UGI 以 XTEN 为 linker 连接，得到 Cpf1-BE0；然后在 Cpf1-BE0 基础上在 N 端引入了 SV40 核定位信号序列，得到 Cpf1-BE。并将其与 sgRNA 表达质粒共转染真核细胞，从而实现了高效率的碱基编辑。接着，为缩小碱基编辑窗口，实现高精度的碱基编辑，本发明在 APOBEC1 上引入 W90Y, R126E 突变得到 Cpf1-BE-YE，并将其与 sgRNA 表达质粒共转染真核细胞，从而实现了更高精度的碱基编辑。进一步，在 Cpf1-BE 编码框的下游利用 T2A 自剪切肽段同框连接 3×UGI 基因，得到 Cpf1-BE-UGI，并将其与 sgRNA 表达质粒共转染真核细胞，实现了高保真度的碱基编辑。

本发明以人基因组 DNA 靶点 DYRK1A（实施例 1）和 RUNX1（实施例 2）位点为例，实施了高功效高精度的 C-T 碱基编辑。

具体的，实验方法包括下列步骤：

1) 构建表达质粒。

首先通过在载体 PST1374-LbCpf1-NLS 上引入 D832A, E925A, 和 D1148A，使之变成有结合但无切割活性的 dCpf1-NLS，再与 APOBEC1 和 UGI 以 XTEN 为 linker 连接，得到 Cpf1-BE0；然后在 Cpf1-BE0 的 N 端引入 SV40 核定位信号序列，构建三者的融合表达载体，得到 Cpf1-BE。

在 APOBEC1 引入 W90Y, R126E 突变，得到 Cpf1-BE-YE。

利用 PCR 方法扩增 T2A-UGI 编码序列，并同框连入 Cpf1-BE，得到 Cpf1-BE-UGI。

构建针对人基因组 DNA 靶点 DYRK1A 和 RUNX1 的 sgRNA。

2) 转染真核细胞。

在真核细胞中共转染 sgRNA 和 Cpf1-BE0、Cpf1-BE、Cpf1-BE-YE、Cpf1-BE-UGI。

3) 基因组 DNA 抽提扩增。

抽提转染后的真核细胞基因组 DNA, 然后利用针对靶点 DYRK1A 和 RUNX1 的特异性引物进行 PCR 扩增。

4) 基因组 DNA PCR 产物的高通量二代 DNA 测序。

利用基因组 DNA 的 PCR 产物构建高通量二代 DNA 测序文库, 并进行高通量二代 DNA 测序。

本发明所述的基于 CRISPR/Cpf1 的碱基编辑方法, 可以在各种真核生物的基因组中实施高精度和高保真度的 C-to-T 碱基编辑。

在进一步描述本发明具体实施方式之前, 应理解, 本发明的保护范围不局限于下述特定的具体实施方案; 还应当理解, 本发明实施例中使用的术语是为了描述特定的具体实施方案, 而不是为了限制本发明的保护范围。

除非另外说明, 本发明中所公开的实验方法、检测方法、制备方法均采用本技术领域常规的分子生物学、生物化学、细胞生物学及相关领域的常规技术, 这些技术在现有文献中已有完善说明。

实施例 1: 人基因组 DYRK1A 位点利用 dCpf1 碱基编辑系统实现高精度和高保真度的碱基编辑

一种碱基编辑系统, 包括 Cpf1-BE0 表达载体、Cpf1-BE 表达载体、Cpf1-BE-YE 表达载体以及 Cpf1-BE-UGI 表达载体和 sgRNA 表达载体。所述的 Cpf1-BE0 表达载体的序列为 SEQ ID NO: 35, Cpf1-BE 表达载体的序列为 SEQ ID NO: 36, Cpf1-BE-YE 表达载体的序列为 SEQ ID NO: 37, Cpf1-BE-UGI 表达载体的序列为 SEQ ID NO: 39。

上述的碱基编辑系统的制备和碱基编辑方法为:

1.1 实验材料

1.1.1 试剂及质粒

引物按照常规方法合成自上海博尚生物有限公司；限制性内切酶、DNA 连接酶、高保真 DNA 聚合酶 Phusion®购自 NEB 公司；质粒重组试剂盒 Clone Express® 购自 Vazyme 公司；pCMV-BE3 购自 addgene 网站；Axygen®DNA 胶回收试剂盒购自 Corning 公司；转染试剂 Lipofectamine® LTX, Opti-MEM®购自 Thermo Fisher 公司；QuickExtract™基因组 DNA 抽提试剂、高通量 DNA 二代测序文库构建试剂盒 TruSeqChIPLibrary Preparation Kit 购自 Illumina 公司。

1.1.2 细胞株

人胎肾细胞 HEK293FT (ATCC, PTA-5077™) 培养在加入了 10% 胎牛血清(Gbico), 1%双抗的 DMEM 培养基 (Gibco)中贴壁培养。

1.2 实验方法

1.2.1 Cpf1-BE0 和 Cpf1-BE 表达载体的构建

以 pST1374-LbCpf1 质粒（按照常规方法化学合成，SEQ ID NO: 33）为模板，利用引物对（1 号和 8 号引物）和引物对（2 号和 7 号引物）进行 PCR；再将两段 PCR 产物为共同模版，利用引物对（7 号和 8 号引物）进行融合 PCR，得到含有 D832A 突变的 Cpf1-D832A 片段，并进行凝胶电泳纯化。

以 Cpf1-D832A 为模板，利用引物对（3 号和 8 号引物）和引物对（4 号和 7 号引物）进行 PCR；再将两段 PCR 产物为共同模版，利用引物对（7 号和 8 号引物）进行融合 PCR，得到含有 D832A, E925A 突变的 Cpf1-D832A-E925A 片段，并进行凝胶电泳纯化。

以 Cpf1-D832A-E925A 为模板，利用引物对（5 号和 8 号引物）和引物对（6 号和 7 号引物）进行 PCR；再将两段 PCR 产物为共同模版，利用引物对（7 号和 8 号引物）进行融合 PCR，得到含有 D832A, E925A 和 D1148A 突变的 dCpf1 片段，经凝胶电泳纯化后，利用质粒重组试剂盒 Clone Express®重组至 PstI 和 ApaI 处理产生的线性化 pST1374-LbCpf1 载体中，得到 dCpf1 表达载体 pST1374-dCpf1(SEQ ID NO: 34)。

以 pST1374-dCpf1 为模板，利用引物对（9 号和 10 号引物）扩增 dCpf1；以 pCMV-BE3 (addgene, #73021) 为模板，利用引物对（11 号和 12 号引物）扩增 UGI；再以 dCpf1 和 UGI 为共同模版，利用引物对（12 号和 13 号引物）进行融合 PCR，扩增得到 dCpf1-UGI 片段，经凝胶电泳纯化后，利用质粒重组试剂盒

Clone Express®重组连接至 SmaI 和 PmeI 处理产生的线性化 pCMV-BE3 载体中, 得到 APOBEC-dCpf1-UGI 表达载体 pCMV-Cpf1-BE0(SEQ ID NO: 35)。

以 pCMV-BE3 为模板, 利用引物对(14 号和 15 号引物)扩增 NLS-APOBEC1, 在 APOBEC1 的 N 端引入 SV40 核定位信号序列, 经 NotI 和 SmaI 酶切、凝胶电泳纯化后, 利用 T4 DNA 连接酶连入经 NotI 和 SmaI 处理产生的线性化 pCMV-Cpf1-BE0 载体中, 得到 Cpf1-BE 表达载体 pCMV-Cpf1-BE (SEQ ID NO: 36)。

1.2.2 Cpf1-BE-YE 表达载体的构建

以 pCMV-Cpf1-BE 为模板, 利用引物对(15 号和 16 号引物)和引物对(14 号和 17 号引物)进行 PCR; 再将两段 PCR 产物为共同模版, 利用引物对(14 号和 15 号引物)进行融合 PCR, 得到含有 W90Y 突变的 Cpf1-BE-Y 片段, 并进行凝胶电泳纯化。

以 Cpf1-BE-Y 为模板, 利用引物对(15 号和 18 号引物)和引物对(14 号和 19 号引物)进行 PCR; 再将两段 PCR 产物为共同模版, 利用引物对(14 号和 15 号引物)进行融合 PCR, 得到含有 W90Y, R126E 突变的 Cpf1-BE-YE 片段。经 NotI 和 SmaI 酶切、凝胶电泳纯化后, 利用 T4 DNA 连接酶连入经 NotI 和 SmaI 处理产生的线性化 pCMV-Cpf1-BE 载体中, 得到 Cpf1-BE-YE 表达载体 pCMV-Cpf1-BE-YE(SEQ ID NO: 37)。

1.2.3 Cpf1-BE-UGI 表达载体的构建

以 pCMV-Cpf1-BE 为模板, 利用引物对(7 号和 20 号引物)进行 PCR; 扩增得到含有 ApaI 酶切位点的片段, 经凝胶电泳纯化后, 利用质粒重组试剂盒 Clone Express®重组至 PmeI 和 PstI 处理产生的线性化 pCMV-Cpf1-BE 载体中, 得到含有 ApaI 酶切位点的表达载体 pCMV-Cpf1-BE-ApaI。

以化学合成的 DNA 片段 3×2A-UGI(SEQ ID NO: 38)为模板, 利用引物对(21 号和 22 号引物)进行 PCR 扩增得到含有 3×2A-UGI 片段, 经 PmeI 和 ApaI 酶切、凝胶电泳纯化后, 利用 T4 DNA 连接酶连入经 PmeI 和 ApaI 处理产生的线性化 pCMV-Cpf1-BE-ApaI 载体中, 得到 Cpf1-BE-UGI 表达载体 pCMV-Cpf1-BE-UGI(SEQ ID NO: 39)。

1.2.4 sgRNA 表达质粒的构建:

将表 1 中 23 和 24 号引物退火,将退火产物利用 T4 DNA 连接酶连入经 BsaI 和 EcoRI 线性化的 pGL3-U6-sgRNA-PGK-puromycin (addgene, 51133), 得到 LbCpf1-pGL3-U6-sgRNA (SEQ ID NO: 40)。

将 25 号和 26 号引物退火,将退火产物利用 T4 DNA 连接酶连入经限制性内切酶 BsaI 消化后的 sgRNA 表达载体 LbCpf1-pGL3-U6-sgRNA, 得到靶向人基因组 DYRK1A 位点的 sgRNA 表达质粒 psgDYRK1A (SEQ ID NO: 41)。

表 1 引物

编号	引物名	序列
1	LB D971A F	tatgttataggcattgctcgcggagagagaac
2	LB D971A R	gtttctctctccgcgagcaatgcctataacata
3	LB E1006A F	gctgttatcgccctcgcggacctaagtagcggc
4	LB E1006A R	gccgctattcaggccgcgaggcgataacagc
5	LB D1225A F	ggaagaacagacgtggccttctgatctctct
6	LB D1225A R	aggagagatcagaaaggccacgtctgttcttcc
7	LB F2096	tcgaacagctgcaggagtacgccga
8	LB R4635	accttcgaaggccctcacacttc
9	LB BE3 F1	tccccgggacctcagagtcgccacacccgaaagtatgtcaaagctggagaaattcac
10	LB BE3 R1	agtagaaccaccagacacttccgttcttcttgggtg
11	LB BE3	aagaaacggaaagtgtctgtgtgttactaatctgtc

	F2	
12	CPF BE3 fu R2	aggcacagtcgaggctgacagcgggttaactcaatggtgat
13	CPF BE3 fu F1	tgaaaagcggcagcgagactccgggacctcagagtccgccac
14	lx NLS pcr F	gagatccgcgccgctaatacgaactcactatagggagagccgccaccatgcccagaagaagaggaaa gtcatgagctcagagactggcccagtggtgctg
15	1X NLS pcr R	ctgagggtccgggagtgctgctgcccgtttcaaccgggtggcccagagaatg
16	APOBEC W90Y F1	gcattacctggtttctcagctacagcccatgcgccgaatgtag
17	APOBEC W90Y R1	ctacattcgccgcatgggctgtagctgagaaaccaggtaatgc
18	APOBEC R126E F	gtaccaccacgctgaccccgaaaatcgacaaggcctgCGGga
19	APOBEC R126E R	tcccgaggccttgctgattttcggggtcagcgtggtgtac
20	BE8.1- PmeI- ApaI-R	gaggctgacagcgggttaactcaatggtgaggcccgactttccttcttcttgg
21	ApaI- 1T2AUGI -F	aaagtcgggcccaggggcagaggaagtctgct
22	PmeI- 3T2AUGI -R	cagcgggttaactcacaccttcttttttttagg
23	L079- LbCpf1sc affold-for	accgaatttctactaagtgtagatcgagaccgagagagggtcgggtcctcttttttg

24	L080- LbCpf1sc affold-rev	aattcaaaaaaggagaccgacctctctcgggtctcgatctacacttagtagaaatt
25	hDYRK1 A cpf1 sg1 FOR	agatgaagcacatcaaggacattctaa
26	hDYRK1 A cpf1 sg1 REV	aaaattagaatgtccttgatgtgcttc

步骤 1.2.1, 1.2.2 和 1.2.3 中质粒构建实验中扩增各种片段使用 PCR 体系如下:

Phusion (NEB) 1 μ l; 5 $\times$ buffer (NEB) 4 μ l; Forward Primer (正向引物) (5 μ M) 1 μ l; Reverse Primer (反向引物) (5 μ M) 1 μ l; Template (模版) 10ng (每种模板用量均为 10ng); RNase-free water 补至 50 μ l. PCR 程序如下: 2min 95 $^{\circ}$ C; 30 $\times$ (15 sec 95 $^{\circ}$ C, 30 sec 55 $^{\circ}$ C, 2 min 72 $^{\circ}$ C); 3 min 72 $^{\circ}$ C。

1.2.5 真核细胞转染

将 psgDYRK1A 和得到的序列正确的 Cpf1-BE0 (或 Cpf1-BE 或 Cpf1-BE-YE 或 Cpf1-BE-UGI) 以 0.68 μ g: 1 μ g 的比例混合入 500 μ l Opti-MEM 中, 加入 1.68 μ l Lipofectamine plus (Invitrogen, 15338100), 加入 5.04 μ l Lipofectamine $^{\circ}$ LTX, 室温放置 5 分钟后加入 1 ml DMEM(+10% FBS)培养基和二十万 HEK293FT 细胞的 24 孔板中进行转染。12h 后换成含 1%双抗的新鲜 DMEM(+ 10% FBS)培养基。继续培养 60h 后收细胞。

1.2.6 基因组 DNA 抽提和 PCR 扩增

利用 QuickExtract $^{\text{TM}}$ 试剂按照试剂操作说明抽提转染后 HEK293FT 细胞的基因组 DNA, 将抽提得到的基因组 DNA 利用以下引物 27 和引物 28 进行 PCR 扩增, 并利用 Axygen $^{\circ}$ DNA 胶回收试剂按照试剂操作说明对 PCR 产物进行切胶回收, 用于后续的高通量二代 DNA 测序文库的构建。

表 2 引物

编号	引物名	序列

27	hDYRK1Acpf1 G1F	CCCAAAAAGGTAGAAACAAA
28	hDYRK1A cpf1 G1R	AAACTCATCCAAAGAGTATCTG

扩增基因组 DNA 使用 PCR 体系如下：

Phusion (NEB) 1 μ l; 5 $\times$ buffer (NEB) 4 μ l; Forward Primer (正向引物) (5 μ M) 1 μ l; Reverse Primer (反向引物) (5 μ M) 1 μ l; Template (模版) 100ng (每种模版用量均为 100ng); RNase-free water 补至 50 μ l。

PCR 程序如下: 2min 95 $^{\circ}$ C; 30 $\times$ (15 sec 95 $^{\circ}$ C, 30 sec 55 $^{\circ}$ C, 30 sec 72 $^{\circ}$ C); 3 min 72 $^{\circ}$ C。

1.2.7 高通量二代 DNA 测序

对 PCR 产物进行 DNA-seq 的高通量测序文库制备。具体来说，首先分别使用 QubitFluorometer 和 AglientBioanalyzer 2100 对 PCR 产物进行浓度质检和纯度鉴定；质检合格的样品，再使用 Covaris S220 超声波 DNA 破碎仪将 PCR 产物进行片段化；然后利用 Illumina 官方试剂盒 TruSeqChIP Sample Preparation Kit，对不同的 PCR 产物加上不同生物素标记 barcode，进行 PCR 扩增富集，完成测序文库的制备，并对已经构建完成的文库进行浓度和片段大小质检；最后使用 IlluminaHiSeq 2500 测序仪进行高通量测序。

如图 1 所示，和未转染 Cpf1-BE 表达载体的对照组相比，共转染 Cpf1-BE0 或 Cpf1-BE 和 sgRNA 表达载体能够在 sgRNA 靶点序列内实施定点 C-T 碱基编辑，并且后者比前者造成了更高效率的编辑（图 1B），二者均未检测到明显的碱基随机插入或缺失（图 1C）。接着，本发明在 Cpf1-BE 的 APOBEC1 基因中引入 W90Y, R126E 两个氨基酸突变，构造了碱基编辑器 Cpf1-BE-YE（图 1D）；从而缩小了碱基编辑的窗口，实现了更为精确的碱基编辑（图 1E—图 1G）。进一步的，本发明在 Cpf1-BE 的编码区下游利用 T2A 自剪切肽段连接尿嘧啶糖苷酶抑制剂（Uracil DNA glycosylase inhibitor, UGI），构造了碱基编辑器 Cpf1-BE-UGI（图 1H）；从而提高了目的碱基 T 相对于非目的碱基 A 和 G 的比例（图 1I—图 1K），实现了更高保真度的碱基编辑。

实施例 2 人基因组 RUNX1 位点利用 dCpf1 碱基编辑系统实现高功效高精度的碱基编辑

一种碱基编辑系统，包括 Cpf1-BE0 表达载体、Cpf1-BE 表达载体、Cpf1-BE-

YE 表达载体以及 Cpf1-BE-UGI 表达载体和 sgRNA 表达载体。所述的 Cpf1-BE0 表达载体的序列为 SEQ ID NO: 35, Cpf1-BE 表达载体的序列为 SEQ ID NO: 36, Cpf1-BE-YE 表达载体的序列为 SEQ ID NO: 37, Cpf1-BE-UGI 表达载体的序列为 SEQ ID NO: 39。

上述的碱基编辑系统的制备和碱基编辑方法为:

2.1 实验材料

2.1.1 试剂及质粒:同实施例 1。

2.1.2 细胞株: 同实施例 1。

2.2 实验方法

2.2.1 Cpf1-BE0 和 Cpf1-BE 表达载体的构建: 同实施例 1。

2.2.2 Cpf1-BE-YE 表达载体的构建: 同实施例 1。

2.2.3 Cpf1-BE-UGI 表达载体的构建: 同实施例 1。

2.2.4 sgRNA 表达质粒的构建

将 23 和 24 号引物退火, 将退火产物利用 T4 DNA 连接酶连入经 BsaI 和 EcoRI 线性化的 pGL3-U6-sgRNA-PGK-puromycin(addgene, 51133), 得到 LbCpf1-pGL3-U6-sgRNA(SEQ ID NO: 40)。

将 29 号和 30 号引物退火, 将退火产物利用 T4 DNA 连接酶连入经限制性内切酶 BsaI 消化后的 sgRNA 表达载体 LbCpf1-pGL3-U6-sgRNA, 得到靶向人基因组 RUNX1 位点的 sgRNA 表达质粒 psgRUNX1 (SEQ ID NO: 42)。

表 3 引物

编号	引物名	序列
29	hRUNX1 cpf1 sg1 FOR	AGATTTCTCCCCTCTGCTGGATACCTC
30	hRUNX1 cpf1 sg1 REV	AAAAGAGGTATCCAGCAGAGGGGAGAA

步骤 2.2.1, 2.2.2 和 2.2.3 中质粒构建实验中扩增各种片段使用 PCR 体系如下:

Phusion (NEB) 1 μ l; 5 $\times$ buffer (NEB) 4 μ l; Forward Primer (正向引物) (5 μ M) 1 μ l; Reverse Primer (反向引物) (5 μ M) 1 μ l; Template (模版) 10 ng (每种模板用量均为 10ng); RNase-free water 补至 50 μ l。PCR 程序如下: 2min 95°C; 30 $\times$

(15 sec 95 °C, 30 sec 55 °C, 2 min 72 °C); 3 min 72 °C。

2.2.5 真核细胞转染

将 psgRUNX1 和得到的序列正确的 Cpf1-BE0(或 Cpf1-BE 或 Cpf1-BE-YE 或 Cpf1-BE-UGI) 以 0.68μg: 1μg 的比例混合入 500μl Opti-MEM 中, 加入 1.68μl Lipofectamine plus (Invitrogen, 15338100), 加入 5.04μl Lipofectamine® LTX, 室温放置 5 分钟后加入 1 ml DMEM(+ 10% FBS)培养基和二十万 HEK293FT 细胞的 24 孔板中进行转染。12h 后换成含 1%双抗的新鲜 DMEM(+ 10% FBS)培养基。继续培养 60h 后收细胞。

2.2.6 基因组 DNA 抽提和 PCR 扩增

利用 QuickExtract™试剂按照试剂操作说明抽提转染后 HEK293FT 细胞的基因组 DNA, 将抽提得到的基因组 DNA 利用以下引物 29 和引物 30 进行 PCR 扩增, 并利用 Axygen® DNA 胶回收试剂按照试剂操作说明对 PCR 产物进行切胶回收, 用于后续的高通量二代 DNA 测序文库的构建。

表 4 引物

编号	引物名	序列
31	hRUNX1 cpf1 G1F	CTTTGGGCCTCATAACAAC
32	hRUNX1 cpf1 G1R	GGTGAAACAAGCTGCCATTT

扩增基因组 DNA 使用 PCR 体系如下:

Phusion (NEB) 1 μl; 5×buffer (NEB) 4 μl; Forward Primer (正向引物) (5 μM) 1 μl; Reverse Primer (反向引物) (5 μM) 1 μl; Template (模板) 100 ng (每种模板用量均为 100ng); RNase-free water 补至 50 μl。PCR 程序如下: 2min 95°C; 30× (15 sec 95 °C, 30 sec 55 °C, 30 sec 72 °C); 3 min 72 °C。

2.2.7 高通量二代 DNA 测序: 同实施例 1。

如图 2 所示, 和未转染 Cpf1-BE 表达载体的对照组相比, 共转染 Cpf1-BE0 或 Cpf1-BE 和 sgRNA 表达载体能够在 sgRNA 靶点序列内实施定点 C-T 碱基编辑, 并且后者比前者造成了更高效率的编辑 (图 2B), 二者均未检测到明显的碱基随机插入或缺失 (图 2C)。接着, 本发明在 Cpf1-BE 的 APOBEC1 基因中引入 W90Y, R126E 两个氨基酸突变, 构造了碱基编辑器 Cpf1-BE-YE (图 2D);

从而缩小了碱基编辑的窗口，实现了更为精确的碱基编辑（图 2E—图 2G）。进一步的，本发明在 Cpf1-BE 的编码区下游利用 T2A 自剪切肽段连接尿嘧啶糖苷酶抑制剂（Uracil DNA glycosylase inhibitor, UGI），构造了碱基编辑器 Cpf1-BE-UGI（图 2H）；从而提高了目的碱基 T 相对于非目的碱基 A 和 G 的比例（图 2I—图 2K），实现了更高保真度的碱基编辑。

以上所述仅为本发明较佳实施例而已，并不用以限制本发明，凡在本发明的精神和原则之内，所作的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

权利要求：

1、一种碱基编辑系统，其特征在于，包括 Cpf1-BE0 表达载体、Cpf1-BE 表达载体、Cpf1-BE-YE 表达载体、Cpf1-BE-UGI 表达载体、Cpf1-BE0 表达载体表达出的 RNA 产物、Cpf1-BE 表达载体表达出的 RNA 产物、Cpf1-BE-YE 表达载体表达出的 RNA 产物、Cpf1-BE-UGI 表达载体表达出的 RNA 产物、Cpf1-BE0 表达载体表达出的蛋白产物、Cpf1-BE 表达载体表达出的蛋白产物、Cpf1-BE-YE 表达载体表达出的蛋白产物、Cpf1-BE-UGI 表达载体表达出的蛋白产物中的至少一种。

2、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的碱基编辑系统还包含 sgRNA 表达载体、sgRNA 表达载体表达出的 RNA、和 sgRNA 表达载体表达出的蛋白中的至少一种。

3、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的 Cpf1-BE0 表达载体的序列为 SEQ ID NO: 35，Cpf1-BE 表达载体的序列为 SEQ ID NO: 36，Cpf1-BE-YE 表达载体的序列为 SEQ ID NO: 37，Cpf1-BE-UGI 表达载体的序列为 SEQ ID NO: 39。

4、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的 Cpf1-BE 表达载体的制备方法包括：制备含有 D832A、E925A 和 D1148A 突变的 dCpf1 片段，通过质粒重组至 pST1374-LbCpf1 载体中，得到 dCpf1 表达载体 pST1374-dCpf1，再与 APOBEC1 和 UGI 连接，得到 Cpf1-BE0 表达载体；然后在 APOBEC1 的 N 端引入 SV40 核定位信号序列，得到 Cpf1-BE 表达载体。

5、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的 Cpf1-BE-YE 表达载体的制备方法包括：制备含有 W90Y 和 R126E 突变的 Cpf1-BE-YE 片段，连入 Cpf1-BE 表达载体中，得到 Cpf1-BE-YE 表达载体。

6、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的 Cpf1-BE-UGI 表达载体的制备方法包括：利用 PCR 方法扩增 3×2A-UGI 编码序列，并连入 Cpf1-BE 表达载体，得到 Cpf1-BE-UGI 表达载体。

7、如权利要求 1 所述的碱基编辑系统，其特征在于，所述的 sgRNA 表达载体包括但不限于针对人基因组 DNA 靶点 DYRK1A 和 RUNX1 的 sgRNA。

8、一种碱基编辑方法，其特征在于，包括：采用权利要求 1-7 中任一项所述的碱基编辑系统，在真核细胞中共转染 Cpf1-BE0 表达载体、Cpf1-BE 表达载

体、Cpf1-BE-YE 表达载体、Cpf1-BE-UGI 表达载体、Cpf1-BE0 表达载体表达出的 RNA 产物、Cpf1-BE 表达载体表达出的 RNA 产物、Cpf1-BE-YE 表达载体表达出的 RNA 产物、Cpf1-BE-UGI 表达载体表达出的 RNA 产物、Cpf1-BE0 表达载体表达出的蛋白产物、Cpf1-BE 表达载体表达出的蛋白产物、Cpf1-BE-YE 表达载体表达出的蛋白产物、Cpf1-BE-UGI 表达载体表达出的蛋白产物中的至少一种；抽提转染后的真核细胞基因组 DNA。

9、如权利要求 8 所述的碱基编辑方法，其特征在于，所述的碱基编辑方法还包括利用针对靶点的特异性引物进行 PCR 扩增。

10、权利要求 1-7 中任一项所述的碱基编辑系统在真核生物的基因组中实施 C-to-T 碱基编辑的应用。

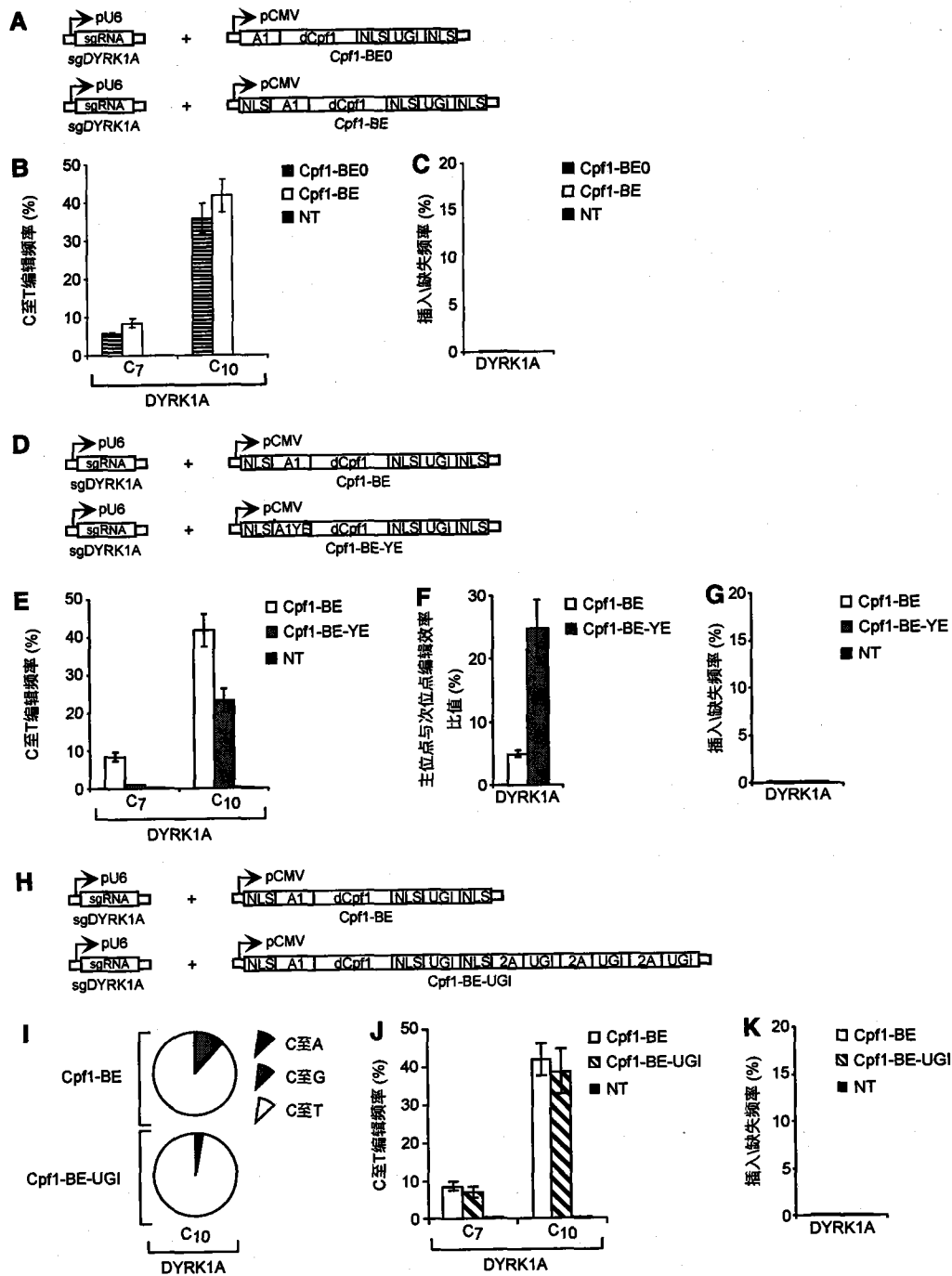


图 1

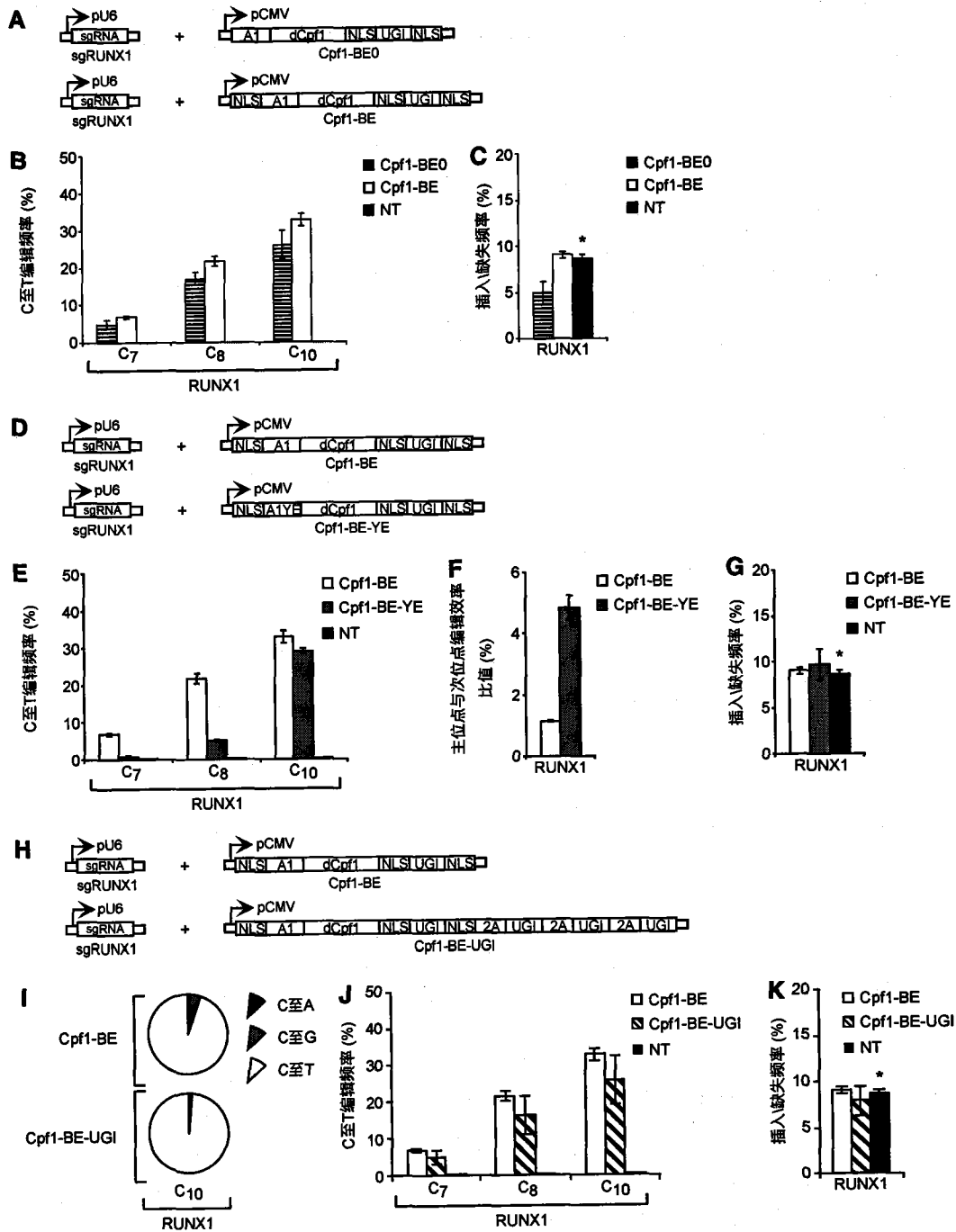


图 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/100131

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/79 (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

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)

Database: CNABS, CPRSABS, CNKI, DWPI, SIPOABS, CNTXT, WOTXT, USTXT, EPTXT, ISI Web of Knowledge, google scholar, WANFANG, CHAOXING LIBRARY, EMBL+GenBank+ddbj, Chinese Patent Biological Sequence Retrieval System; Key Words: 碱基编辑, 基因编辑, 载脂蛋白 B mRNA 编辑复合物, 胞嘧啶, 胸腺嘧啶, Cpf1, Cas12a, Clustered Regularly Interspaced Short Palindromic Repeats from Prevotella and Francisella 1, base editor, BE, nucleobase editor, cytosine, thymine, sgRNA, UGI, SV40, C-T, PAM, APOBEC1, dAsCpf1, dLbCpf1, D832A, E925A, D1148A, C-T, searching on SEQ ID Nos: 1-40

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 20170121693 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 04 May 2017 (04.05.2017), see the abstract, claims 1, 6 and 7, and description, paragraph [0209], and embodiment 9	1, 2, 5, 7, 8 (in part), 9 (in part), 10 (in part)
Y	US 20170121693 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 04 May 2017 (04.05.2017), see the abstract, claims 1, 6 and 7, and description, paragraph [0209], and embodiment 9	3, 4, 6, 8 (in part), 9 (in part), 10 (in part)

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

☒ See patent family annex.

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

Date of the actual completion of the international search 22 May 2018	Date of mailing of the international search report 04 June 2018
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LUO, Xiao Telephone No. (86-10) 62089158

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/100131

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

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

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/100131

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 20170233756 A1 (BENSON HILL BIOSYSTEMS, INC.) 17 August 2017 (17.08.2017), see the abstract, description, paragraphs [0062], [0229] and [0273], and embodiment 5	3, 4, 6, 8 (in part), 9 in part), 10 (in part)
A	CN 105907785 A (SUZHOU GENEPHARMA CO., LTD.) 31 August 2016 (31.08.2016), see entire document	1-10
A	CN 106282228 A (SUZHOU LANCEA THERAPEUTICS CO., LTD.) 04 January 2017 (04.01.2017), see entire document	1-10
A	Zetsche B et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. Cell, 25 September 2015 (25.09.2015), 163(3), ISSN: 0092-8674, see entire document	1-10
A	Wang L et al. Enhanced Base Editing by Co-expression of Free Uracil DNA Glycosylase Inhibitor. Cell Research, 29 August 2017 (29.08.2017), 27(10), ISSN: 1748-7838, see entire document	1-10
A	Makarova KS et al. Snapshot: Class 2 CRISPR-Cas Systems. Cell, 12 January 2017 (12.01.2017), 168(1-2), ISSN: 0092-8674, see entire document	1-10
A	Stella S et al. Structure of the Cpf1 Endonuclease R-loop Complex after Target DNA cleavage. Nature, 31 May 2017 (31.05.2017), 546(7659), ISSN: 1476-4687, see entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/100131

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 20170121693 A1	04 May 2017	AU 2016342380 A1	10 May 2018
		CA 3002827 A1	27 April 2017
		WO 2017070632 A2	27 April 2017
		WO 2017070633 A2	27 April 2017
US 20170233756 A1	17 August 2017	EP 3307884 A2	18 April 2018
		WO 2017141173 A2	24 August 2017
CN 105907785 A	31 August 2016	CN 106244591 A	21 December 2016
		WO 2017190664 A1	09 November 2017
CN 106282228 A	04 January 2017	WO 2018033110 A1	22 February 2018

国际检索报告

国际申请号

PCT/CN2017/100131

A. 主题的分类 C12N 15/79(2006.01)i 按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类		
B. 检索领域 检索的最低限度文献(标明分类系统和分类号) C12N 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用)) 数据库: CNABS, CPRSABS, CNKI, DWPI, SIPOABS, CNTXT, WOTXT, USTXT, EPTXT, ISI Web of Knowledge, google scholar, 万方, 超星图书, EMBL+GenBank+ddbj, 中国专利生物序列检索系统; 关键词: 碱基编辑, 基因编辑, 载脂蛋白B mRNA编辑复合物, 胞嘧啶, 胸腺嘧啶, Cpf1, Cas12a, Clustered Regularly Interspaced Short Palindromic Repeats from Prevotella and Francisella 1, base editor, BE, nucleobase editor, cytosine, thymine, sgRNA, UGI, SV40, C-T, PAM, APOBEC1, dAsCpf1, dLbCpf1, D832A, E925A, D1148A, C-T, 对SEQ ID Nos: 1-40的检索		
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	US 20170121693 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2017年 5月 4日 (2017 - 05 - 04) 参见摘要, 权利要求1, 6和7, 说明书第209段和实施例9	1, 2, 5, 7, 8(部分), 9(部分), 10(部分)
Y	US 20170121693 A1 (PRESIDENT AND FELLOWS OF HARVARD COLLEGE) 2017年 5月 4日 (2017 - 05 - 04) 参见摘要, 权利要求1, 6和7, 说明书第209段和实施例9	3, 4, 6, 8(部分), 9(部分), 10(部分)
Y	US 20170233756 A1 (BENSON HILL BIOSYSTEMS, INC.) 2017年 8月 17日 (2017 - 08 - 17) 参见摘要, 说明书第62, 229, 273段和实施例5	3, 4, 6, 8(部分), 9(部分), 10(部分)
A	CN 105907785 A (苏州吉玛基因股份有限公司) 2016年 8月 31日 (2016 - 08 - 31) 参见全文	1-10
A	CN 106282228 A (州兰希亚生物科技有限公司) 2017年 1月 4日 (2017 - 01 - 04) 参见全文	1-10
☒ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2018年 5月 22日		2018年 6月 4日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451		罗霄 电话号码 62089158

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
A	Zetsche B等. "Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System" Cell, 第163卷, 第3期, 2015年 9月 25日 (2015 - 09 - 25), ISSN: 0092-8674, 参见全文	1-10
A	Wang L等. "Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor" Cell Research, 第27卷, 第10期, 2017年 8月 29日 (2017 - 08 - 29), ISSN: 1748-7838, 参见全文	1-10
A	Makarova KS等. "SnapShot: Class 2 CRISPR-Cas Systems" Cell, 第168卷, 第1-2期, 2017年 1月 12日 (2017 - 01 - 12), ISSN: 0092-8674, 参见全文	1-10
A	Stella S等. "Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage" Nature, 第546卷, 第7659期, 2017年 5月 31日 (2017 - 05 - 31), ISSN: 1476-4687, 参见全文	1-10

第I栏

核苷酸和/或氨基酸序列(续第1页第1.c项)

1. 关于国际申请中所公开的任何核苷酸和/或氨基酸序列, 国际检索是基于下列序列表进行的:

a. ☒ 作为国际申请的一部分提交的:

☒ 附件C/ST. 25文本文件形式

☐ 纸件或图形文件形式

b. ☐ 根据细则13之三. 1(a) 仅为国际检索目的以附件C/ST. 25文本文件形式与国际申请同时提交的:

c. ☐ 仅为国际检索目的在国际申请日之后提交的:

☐ 附件C/ST. 25文本文件形式(细则13之三. 1(a))

☐ 纸件或图形文件形式(细则13之三. 1(b) 和行政规程第713段)

2. ☐ 另外, 在提交/提供了多个版本或副本的序列表的情况下, 提供了关于随后提交的或附加的副本中的信息与申请时提交的作为申请一部分的序列表的信息相同或未超出申请时提交的申请中的信息范围(如适用)的所需声明。

3. 补充意见:

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/100131

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	20170121693	A1	2017年 5月 4日	AU	2016342380	A1	2018年 5月 10日
				CA	3002827	A1	2017年 4月 27日
				WO	2017070632	A2	2017年 4月 27日
				WO	2017070633	A2	2017年 4月 27日
US	20170233756	A1	2017年 8月 17日	EP	3307884	A2	2018年 4月 18日
				WO	2017141173	A2	2017年 8月 24日
CN	105907785	A	2016年 8月 31日	CN	106244591	A	2016年 12月 21日
				WO	2017190664	A1	2017年 11月 9日
CN	106282228	A	2017年 1月 4日	WO	2018033110	A1	2018年 2月 22日