

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2018 年 2 月 8 日 (08.02.2018)



(10) 国际公布号
WO 2018/024119 A1

(51) 国际专利分类号:
C12N 15/113 (2010.01) *C12N 15/63* (2006.01)

(21) 国际申请号: PCT/CN2017/093903

(22) 国际申请日: 2017 年 7 月 21 日 (21.07.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201610627960.X 2016 年 8 月 3 日 (03.08.2016) CN

(71) 申请人: 南京大学(NANJING UNIVERSITY) [CN/
CN]; 中国江苏省南京市汉口路 22 号,
Jiangsu 210093 (CN)。

(72) 发明人: 周国华(ZHOU, Guohua); 中国江苏省南京市玄武区中山东路 305 号, Jiangsu (CN)。赵庆顺(ZHAO, Qingshun); 中国江苏省南京市浦口高新区学府路 12 号, Jiangsu (CN)。徐澍(XU, Shu); 中国江苏省南京市玄武区中山东路 305 号, Jiangsu (CN)。曹莎莎(CAO, Shasha); 中国江苏省南京市浦口高新区学府路 12 号, Jiangsu (CN)。邹秉杰(ZOU, Bingjie); 中国江苏省南京市玄武区中山东路 305 号, Jiangsu (CN)。岳芸芸(YUE, Yunyun); 中国江苏省南京市浦口高新区学府路 12 号, Jiangsu (CN)。

(74) 代理人: 北京奉思知识产权代理有限公司
(F&S INTELLECTUAL PROPERTY LTD.); 中国北京市知春路 6 号锦秋国际大厦 B 座 1403 室, Beijing 100191 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家
保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,

BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,
CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,
JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。
- 包括说明书序列表部分(细则 5.2(a))。

(54) Title: METHOD FOR EDITING TARGET POLYNUCLEOTIDE AND APPLICATION THEREOF

(54) 发明名称: 一种靶多核苷酸编辑方法及其应用

(57) Abstract: A method for editing a target polynucleotide, comprising the specific steps of: designing a pair of DNA oligonucleotide single strands for a target gene of interest, such that the pair of single strands are respectively complementary to the sense and antisense strands of the target gene; respectively generating a structure recognizable by a type of nuclease which recognizes a target sequence structure and which has a cleaving activity, and cleaving a target genomic DNA by means of the nuclease, so as to edit a genome. The structure recognized by the type of nuclease which recognizes a target sequence structure and which has a cleaving activity is a 3'-Flap structure, the 3' ends of the DNA oligonucleotide single strands are not complementary to the target gene; a recombinant structure recognizes nucleic acid endonuclease and includes a structural recognition domain, a DNA cleavage domain, a peptide linking said two domains, and a nuclear localization signal.

(57) 摘要: 一种靶多核苷酸编辑方法, 具体步骤包括针对目标靶基因, 设计一对 DNA 寡核苷酸单链, 使其与靶基因的正义和反义链分别互补, 分别产生能够被一类识别靶序列结构并具有切割活性的核酸酶识别的结构, 通过所述的核酸酶切割靶基因组 DNA, 实现对基因组的编辑。其中, 所述被一类识别靶序列结构并具有切割活性的核酸酶识别的结构为 3'-Flap 结构, DNA 寡核苷酸单链的 3' 末端不与靶基因互补; 重组的结构识别核酸内切酶, 包含结构识别功能域、DNA 切割功能域、连接二者的肽段和核定位信号。

WO 2018/024119 A1

一种靶多核苷酸编辑方法及其应用

相关申请的交叉引用

本发明要求 2016 年 8 月 3 日提交的中国专利申请号 201610627960.X
5 的优先权，其全部内容通过引用并入本文。

技术领域

本发明属于分子生物学领域，涉及一种基因组编辑方法及其应用，
具体而言涉及一种对基因组进行靶向切割的技术。

10

背景技术

DNA 编辑在许多体外或体内的分子生物学实验中都是至关重要的。
由于 II 型限制性内切酶（REases）对目标核苷酸精致而准确的切割，使
得它们成为这些实验中不可或缺的工具。迄今为止，已有 3700 种 II 型
15 REases 已经被开发，但只有 262 种不同核酸序列能够被这些酶识别
（Jurénaitė-Urbanavičienė S, et al. Proc Natl Acad Sci U S A 2007,
104:10358-10363.）。因此，这些有限的序列限制了 DNA 编辑操作中的
各种需求。为了克服此限制，已经开发了以下几种方法。第一种方法涉
及对现有 REases 的氨基酸序列进行突变，如 *Not* I 突变体，用蛋白质结
20 构信息学帮助增加 DNA 序列识别的特异性（Buchholz F. Curr Opin
Biotech 2009, 20:383-389.）。第二种方法涉及构建一种将靶向识别结构
域与 DNA 切割结构域结合融合的新型 IIS 型酶（也称为非正统酶），如
曾报道的 *Tst* I 和 *Bmr* I（Chan S, et al. Nucleic Acids Res 2007, 18,
6238-6248.）。第三种方法涉及形成一种的新型核酸酶，其融合多个能识
25 别不同 DNA 序列的限制性内切酶的切割结构域。例如，锌指核酸酶
（ZFN）（Kim Y-G, et al. Proc Natl Acad Sci U S A 1996, 93:1156-1160.）
由一个 DNA 识别域和一个非特异性核酸内切酶构成。DNA 识别域是由
一系列锌指蛋白基序（zinc-fingers）串联组成从而形成锌指阵列（Zinc
Finger Array, ZFA），并识别靶 DNA 序列；非特异性核酸内切酶如 *Fok*
30 I 切割结构域可以在靶位点切割 DNA。ZFN 已经成功地在多种生物和细

胞中被用于敲除或敲入基因（Bibikova M, et al. Genetics 2002, 161:1169-1175; Li H, et al. Nature 2011, 475:217-221; Strairmer J, et al. Nat Methods 2012, 9:993-998; Urnov FD, Rebar EJ, et al. Nat Rev Genet 2010, 11:636-646; Kim JS, et al. Nat Methods 2010, 7:91）。然而，ZFN 很难针对任何所需的 DNA 序列，因为它们偏好识别富含 G 碱基的共有序列，并且对识别靶 DNA 序列的“活性的”锌指阵列的鉴定成本很高(Isalan M. Nat Methods 2012, 9:32-34.)。另一个例子是转录激活因子样效应物核酸酶（transcription activator-like effector nucleases, TALEN），它是将多种转录激活样（TAL）基序（Li L, et al. Mol Plant 2013, 6:1318-1330.）和 *Fok I* 切割结构域融合而成。每个 TAL 基序（Repeat）可以识别一个碱基核苷酸，因而，从理论上说，有多个基序组成的这种酶可以识别靶点中相应的任何序列，且使用一对 TALENs 已成功地敲除或敲入各种生物体和细胞的基因（Bloom K, et al. Mol Ther 2013, 21:1889-1897; Christian M, et al. Genetics 2010, 186:757-761; Sung YH, et al. Nat Biotechnol 2013, 31:23-24; Tesson L, et al. Nat Biotechnol 2011, 29:695-696.; Mussolino C, et al. Nucleic Acids Res 2011, 39:9283-9293.）。但是，TALEN 需要靶序列的 5'末端有胸腺嘧啶核苷酸，其可被氨基末端隐性重复折叠（cryptic repeat fold）所识别（Mak AN, et al. Science 2012, 335:716-719.）。第四种方法是 RNA-指导的核酸内切酶（RNA-guided endonucleases, RGENs 如 CRISPR-Cas 系统），其用 RNA 替代了多肽进而识别目的序列，并将酶引导至切割靶点（Horvath P, Science 2010, 327:167-170.）。至此，使用 RGENs 也成功对人类基因组进行了基因编辑（Cho SW, et al. Nat Biotechnol 2013, 31:230-232.; Cong L, et al. Science 2013, 339:819-823.; Mali P, et al. Science 2013, 339:823-826.）。相比 ZFN 和 TALEN, RGEN 具有明显优势，它采用合成的短 RNA 分子代替了蛋白去识别序列。但是，RGEN 靶点受限于需要具有能被 Cas9 识别的 PAM 序列（Mojica FJ, et al. Microbiology 2009, 155:733-740.）。综上所述，DNA 序列的特异性限制了人工核酸内切酶在 DNA 编辑中的应用。

因此，开发一种不依赖于 DNA 序列且可切割所需序列的新型核酸内切酶具有极强的应用价值。能够识别特异 DNA 结构的侧翼核酸内切

- 酶 1 (FEN-1) 为该需求提供了可能 (Harrington JJ, et al. EMBO J 1994, 13:1235-1246.)。在 DNA 复制和修复过程中, FEN-1 参与了清除 RNA 引物或损伤 DNA 的过程 (Kaiser MW, et al. J Biol Chem. 1999, 274:21387-21394.; Kao HI, et al. J Biol Chem. 2002, 277:14379-14389.)。
- 5 新合成的 DNA 和移位区域与模板链配对碱基竞争, 从而导致双侧翼结构 (double-flap structure) 的形成 (Reynaldo LP, J Mol Biol 2000, 297:511-520.)。双侧翼结构具有一个单独的未配对的 3'核苷酸 (3' flap)。*Afu*FEN-1 在结合到 3' flap 后会催化裂解磷酸二酯键 (Chapados BR, et al. Cell 2004, 116:39-50.)。此外, *Fok* I 是 IIS 型限制性内切酶, 其由一个
- 10 N 末端的 DNA 识别结构域和 C 末端切割结构域 (Fn1) 组成。*Fok* I 的双向特性已用于新型特异性人工核酸内切酶的开发 (Li L, et al. Proc Natl Acad Sci U S A 1992, 89:4275-4279.), 如 ZFN 或 TALEN。

发明内容

- 15 因此, 本发明中, 申请人构建了由识别 3'-Flap 结构的 FEN-1 和切割 DNA 链的 *Fok* I (Fn1) 切割结构域所组成的结构指导的核酸内切酶 (Structure-guided endonuclease, SGN)。3'-Flap 结构由靶序列和人工引导 DNA (guide DNA, gDNA) 形成。根据结构引导识别, SGN 可以切割任何所需的靶向 DNA 而不需要像 type II REases、ZFN 或 TALEN 那样改变核酸内切酶或肽单元 (peptide units), 或像 RGEN 中对 RNA 分子有应用限制。
- 20

- 本发明的目的之一是提供了一种靶多核苷酸编辑方法, 其特征在于: 设计一对寡核苷酸探针, 使其与靶多核苷酸的正义链和反义链分别结合, 分别产生能够被核酸酶识别的目标序列结构, 所述核酸酶切割靶多核苷酸, 实现对靶多核苷酸的编辑。
- 25

在本发明的某个实施例中, 所述的靶多核苷酸编辑方法中, 所述靶多核苷酸是 RNA 或 DNA, 优选是基因组 DNA, 所述的基因组优选是斑马鱼基因组、哺乳动物基因组、人类基因组、或植物基因组。

- 在本发明的某个实施例中, 所述的靶多核苷酸编辑方法中, 所述一对寡核苷酸探针是 DNA, 优选所述一对寡核苷酸探针与靶多核苷酸的结
- 30

合位点间隔 0-100 bp, 优选间隔为 10-70 bp、20-60 bp、32-50 bp 或 40 bp。

在本发明的某个实施例中, 所述的靶多核苷酸编辑方法中, 所述寡核苷酸探针的长度为 20 nt 以上, 优选长度为 20-50 nt 或 25 nt。

在本发明的某个实施例中, 所述的靶多核苷酸编辑方法中, 所述的
5 目标序列结构为 5'突出核酸结构、3'突出核酸结构、切刻核酸结构、侵入核酸结构、三链核酸结构、Y 型核酸结构、3'-Flap 结构以及由错配或缺失核酸引起的鼓泡核酸结构中的至少一种的识别功能域, 优选 3'-Flap 结构, DNA 寡核苷酸单链的 3'-末端不与目标基因组 DNA 互补。

在一个具体实施例中, 所述寡核苷酸探针的 5'端对靶多核苷酸特异,
10 优选与靶多核苷酸互补; 寡核苷酸探针的 3'端不与靶多核苷酸互补, 优选 3'末端 1 个以上、1-20 个、1-10 个、1-5 个、1-4 个、1-3 个、2 个或 1 个碱基不与靶多核苷酸互补。

在本发明的某个实施例中, 所述的核酸酶是重组结构识别核酸内切酶 (SGN), 包含结构识别功能域、DNA 切割功能域和连接二者的肽段;
15 所述的结构识别功能域为能够识别结构为 5'突出核酸结构、3'突出核酸结构、切刻核酸结构、侵入核酸结构、三链核酸结构、Y 型核酸结构、3'-Flap 结构以及由错配或缺失核酸引起的鼓泡核酸结构中的至少一种的识别功能域, 优选选自 TaqPol、TthPol、TaqExo、AfuFEN、PfuFEN、MjaFEN、MthFEN、E.coliMuts、Tthmuts 和 Taqmuts 组成的组中任意一种酶的识别
20 功能域或全酶片段; 所述的切割功能域为 IIS 型核酸内切酶的切割功能域, 优选 Fok I 的部分或全部肽段; 所述的切割功能域还可选自 TaqPol、TthPol、TaqExo、AfuFEN、PfuFEN、MjaFEN、MthFEN、E.coliMuts、Tthmuts 和 Taqmuts 组成的组中任意一种酶的全酶片段的核酸内切酶结构域; 所述的
25 连接肽段为不影响结构识别与酶切功能的柔性肽段, 优选甘氨酸或丝氨酸或其组合的串联组合。

在一个具体实施例中, 所述的重组结构识别核酸内切酶含有核定位信号。

在一个具体实施例中, 所述的重组结构识别核酸内切酶的氨基酸序列选自 SEQ ID No:1、SEQ ID No:2 中任一个。

在另一个具体实施例中，所述的重组结构识别核酸内切酶的核酸序列选自 SEQ ID No:3、SEQ ID No:4、SEQ ID No:5 中任一个。

在一个具体实施例中，所述基因组 DNA 是内源 DNA 或整合到基因组的外源 DNA，优选所述靶多核苷酸编辑方法是在体实施的。

- 5 本发明的目的之二是提供一种寡核苷酸探针对，其特征在于该寡核苷酸探针对与靶多核苷酸的正义和反义链分别结合，分别产生能够被核酸酶识别的目标序列结构，所述目标序列结构能够被核酸酶识别，所述核酸酶切割靶多核苷酸，实现对靶多核苷酸的编辑。

10 在某个实施例中，所述一对寡核苷酸探针是 DNA，优选与靶多核苷酸的结合位点间隔 0-100 bp，优选间隔为 10-70 bp、20-60 bp、32-50 bp 或 40 bp。

 在一个具体实施例中，一对寡核苷酸探针与靶多核苷酸的结合位点间隔 bp 长度优选为 0 bp、1 bp、2 bp、3 bp、4 bp、5 bp、6 bp、7 bp、8 bp、9 bp、10 bp、11 bp、12 bp、13 bp、14 bp、15 bp、16 bp、17 bp、
15 18 bp、19 bp、20 bp、21 bp、22 bp、23 bp、24 bp、25 bp、26 bp、27 bp、28 bp、29 bp、30 bp、31 bp、32 bp、33 bp、34 bp、35 bp、36 bp、37 bp、38 bp、39 bp、40 bp、41 bp、42 bp、43 bp、44 bp、45 bp、46 bp、47 bp、48 bp、49 bp、50 bp、51 bp、52 bp、53 bp、54 bp、55 bp、56 bp、57 bp、58 bp、59 bp、60 bp、61 bp、62 bp、63 bp、64 bp、65 bp、66 bp、67 bp、
20 68 bp、69 bp、70 bp、71 bp、72 bp、73 bp、74 bp、75 bp、76 bp、77 bp、78 bp、79 bp、80 bp、81 bp、82 bp、83 bp、84 bp、85 bp、86 bp、87 bp、88 bp、89 bp、90 bp、91 bp、92 bp、93 bp、94 bp、95 bp、96 bp、97 bp、98 bp、99 bp、100 bp。

 在某个实施例中，所述寡核苷酸探针的长度为 20 nt 以上，优选长度为 20-50 nt 或 25 nt。
25

 在某个实施例中，所述寡核苷酸探针的 5'端对靶多核苷酸特异，优选与靶多核苷酸互补；寡核苷酸探针的 3'端不与靶多核苷酸互补，优选 3'末端 1 个以上、1-20 个、1-10 个、1-5 个、1-4 个、1-3 个、2 个或 1 个碱基不与靶多核苷酸互补。

本发明还提供了用于靶多核苷酸编辑的系统，其特征在于所述系统包括本发明中所述的寡核苷酸探针对，优选包括本发明中所述的核酸酶。

5 本发明进一步提供了用于靶多核苷酸编辑的试剂盒，其特征在于所述试剂盒包括本发明中所述的寡核苷酸探针对，优选包括本发明中所述的核酸酶。

本发明的目的之三是提供了一种在靶多核苷酸中切割大片段的方法，其特征在于采用了本发明中所述的靶多核苷酸编辑方法。

10 在某个实施例中，一种在靶多核苷酸中切割大片段的方法中所述的大片段长度大于所述寡核苷酸探针对与靶多核苷酸的结合位点的间隔，优选所述的大片段长度大于所述间隔 1-10000 bp、1-3000 bp、1-2000 bp、1-1000 bp、1-500 bp、1-200 bp 或 1-100 bp。

15 本发明还提供了一种疾病细胞模型的制作方法，其特征在于采用了本发明中所述的靶多核苷酸编辑方法或施用了本发明中所述的寡核苷酸探针对或本发明中用于靶多核苷酸编辑的系统或本发明中用于靶多核苷酸编辑的试剂盒。

本发明还提供了一种疾病动物模型的制作方法，其特征在于采用了本发明中所述的靶多核苷酸编辑方法或施用了本发明中所述的寡核苷酸探针对或本发明中用于靶多核苷酸编辑的系统或本发明中用于靶多核苷酸编辑的试剂盒。

20 本发明还提供了一种植物突变体的制作方法，其特征在于采用了本发明中所述的靶多核苷酸编辑方法或施用了本发明中所述的寡核苷酸探针对或本发明中用于靶多核苷酸编辑的系统或本发明中用于靶多核苷酸编辑的试剂盒。

25 本发明的目的之四是提供了一种疾病治疗方法，其特征在于采用了本发明中所述的靶多核苷酸编辑方法或施用了本发明中所述的寡核苷酸探针对或本发明中用于靶多核苷酸编辑的系统或本发明中用于靶多核苷酸编辑的试剂盒。

30 在某个实施例中，所述的疾病选自遗传相关疾病或非遗传相关疾病，优选所述的遗传相关疾病选自癌症、自身免疫疾病、糖尿病、血液病、心脏病、抑郁症、阿尔茨海默病、哮喘和神经疾病中的任一种。

发明的有益效果

本发明的有益效果在于所建立一种基因组编辑方法，针对目标靶基因，通过设计一对 DNA 寡核苷酸单链，使其与目标基因的正义和反义链
5 分别互补，分别产生能够被一类识别目标序列结构并具有切割活性的核酸酶识别的结构，通过所述的核酸酶切割目标基因组 DNA，最终切除大片
段 DNA，实现对基因组目的序列的靶向切割并有益于目标蛋白功能的最大程度敲除。

SGN 具有很多优势，首先在该系统中，可以较为容易地设计和合成的
10 的 gDNA 并根据需要调整其浓度；其次由 SGN 创建的大片段缺失突变的等位基因更可能产生无效的等位基因，而 ZFN、TALEN、RGEN 或新报道的 NgAgo 基因组编辑系统产生的突变是通过小的插入缺失造成的，该缺失仍然能够编码具有一些残留功能的截短蛋白质；另外在 SGN 系统中的 gDNAs 长度是可以调整的，以避免错误杂交。

15

附图说明

图 1 示出的是可用于体外实验的 SGN 编码序列和氨基酸序列的图。

图 2 示出的是 pET28a (+) -SGN 的质粒图谱图。

图 3 示出的是可用于体内斑马鱼显微注射中的携带有核定位信号的
20 SGN 编码序列和氨基酸序列（该蛋白质序列对应的编码序列是斑马鱼源化的密码子优化的）。

图 4 示出的是携带有核定位信号的 SGN 编码序列和氨基酸序列（该蛋白质序列对应的编码序列是人源化的密码子优化的）。

图 5 示出的是关于 SGN 能否切割靶 DNA 的体内和体外策略图。左
25 图是体外实验，靶 DNA 用 Cy5 基团标记修饰。设计互补于靶 DNA 的 gDNA，其 3'末端有未配对的核苷酸以形成 3'-Flap 结构。SGN 识别 3'-Flap 结构并切割靶 DNA。切割的产物通过变性聚丙烯酰胺凝胶电泳（变性的 -PAGE）和荧光成像进行分析。右图是使用的 Tg (*flk:eGFP*) 转基因斑马鱼或野生型斑马鱼胚胎研究 SGN 在体内的活性。设计一对 gDNAs，
30 其与靶转基因 GFP 或内源性基因的互补，且 gDNA 的 3'末端有未配对的

核苷酸以形成 3'-Flap 结构。将 SGN 的 mRNA（含有细胞核定位信号编码序列）和一对识别目标基因的 gDNAs 显微注射入斑马鱼胚胎。所表达的 SGN 会识别 3'-Flap 结构并切割体内靶 DNA。基因组 DNA 被消化并通过 DNA 修复途径被修复。为了检验该 DNA 编辑过程，进一步从斑马鱼胚胎中提取基因组 DNA，然后进行 GFP 或内源性基因的目的序列的 PCR 扩增、克隆和测序，分析由 SGN 引起的突变。

图 6 示出的是关于 SGN 体外切割单链 DNA 的图。变性-PAGE 示出由 SGN 切割产生的 DNA 产物。a-c：不同的单链靶 DNA（S-1，S-2 和 S-3）与不同 gDNA（gDNA-1、gDNA-2 和 gDNA-3）进行反应。泳道 1：S 加 SGN；泳道 2：S 加 gDNA；泳道 3：S 加 gDNA 和 Fok I；泳道 4：S 加 gDNA 和 FEN-1；泳道 5：S 加 gDNA 和 SGN。NC：无酶对照组。泳道 M：DNA 标准品。序列图下的箭头示意的是根据切割产物大小的可能切割位点。

图 7 示出的是不同的未配对的 3'核苷酸对 SGN 导致的 DNA 切割的效果。示意图（顶部）示出的是在单链 DNA 靶序列上 gDNA 的未配对 3'核苷酸。实线箭头指示根据切割产物大小推测的可能切割位点。变性-PAGE 结果（下部）显示了由 SGN 酶切的清晰切割产物，且在不同非配对类型的效率之间没有明显差异。(a)gDNA-1、gDNA-1-G、gDNA-1-T 靶向 S-1 ssDNA；(b)gDNA-3、gDNA-3-G、gDNA-3-A 靶向 S-3 ssDNA；(c)gDNA-4、gDNA-4-A、gDNA-4-T 靶向 S-4 ssDNA；和 (d)gDNA-5、gDNA-5-T、gDNA-5-G 靶向 S-5 的 ssDNA。

图 8 示出的是引导 DNA(gDNA)的长度影响 SGN 切割 DNA 的图。序列下的实线箭头示意的是根据切割产物大小的可能切割位点，虚线箭头表示理论切割位点。变性-PAGE 结果显示 SGN 切割后的产物。(a)gDNA-6-10 nt、gDNA-6-15 nt、gDNA-6-20 nt、gDNA-6-25 nt、gDNA-6-30 nt、gDNA-6-35 nt、gDNA-6-40 nt、gDNA-6-45 nt、gDNA-6-50 nt、gDNA-6-55 nt 和 gDNA-6-60 nt 分别靶向 S-6 ssDNA；(b)gDNA-7-10 nt、gDNA-7-15 nt、gDNA-7-20 nt、gDNA-7-25 nt、gDNA-7-30 nt、gDNA-7-35 nt、gDNA-7-40 nt、gDNA-7-45 nt、gDNA-7-50 nt、gDNA-7-55 nt 和 gDNA-7-60 nt 分别靶向 S-7 ssDNA。

图 9 示出的是关于 SGN 以二聚体形式发挥作用的实验结果图。DNA 切割率和各种 SGN 浓度之间的关系。用各种浓度的 SGN (0、0.127、0.25、0.37、0.50、0.62、1.00、2.00 nM) 切割固定浓度 100 nM 的 S-8。以切割速率 (用单位之间 t 内产物 P 的产生确定) 和对应的 SGN 浓度作图。

- 5 图中进入平台期前的黑色实线方程式是 $y=2.9684 x^{1.9255}$ (y =速率, x =SGN 浓度), 黑色的圆点表示各个 SGN 浓度下的平均速率。

图 10 示出的是关于确定 SGN 切割位点的图。(a) 确定切割位点方法的示意图。小圆代表生物素。大圆代表链霉亲和素磁珠。生物素连接的线代表靶 DNA, gDNA 中黑色部分代表 gDNA 的 3'末端的错配碱基。

- 10 (b) I 为信号强度。图中显示“GGAAGTGAC”的测序信号。S-1 (上) 测序鉴定的切割位点为实线箭头所示。(c) 图中显示“GCCCTTC”的测序信号。S-2 (上) 测序鉴定的切割位点为实线箭头所示。

图 11 示出的是体外 SGN 切割双链 DNA 的图。变性-PAGE 凝胶显示由一对 gDNA 引导 SGN 切割的产物 (标识*所示)。不同的靶 dsDNAs (S-1/S-9 (a)、S-3/S-10 (b)、S-5/S-11 (c)) 与由不同的引导 DNAs (gDNA1/gDNA9 (a)、gDNA3/gDNA10 (b)、gDNA5/gDNA11 (c)) 引导的 SGN 分别反应。泳道 1: S 加 SGN 和 gDNAs; 泳道 2: S 加 SGN; 泳道 3: S 加 gDNAs; 泳道 M: DNA 链。*: 指示切割产物。

- 20 图 12 示出的是关于 SGN 编辑斑马鱼基因组内目标基因的图。箭头指示的是大片段缺失发生的位置。(a) 上图示出靶向转基因斑马鱼报告基因 *GFP* 的 gDNAs 所在位置。下图显示的是突变分子#5-15 (表 2) 的部分测序峰图。(b) 上图示出靶向目标内源性 *znf703* 基因的 gDNAs 所在位置。下图显示的是被编辑形成的突变分子的部分测序峰图, 其中 754 bp 和其后的 11 bp 被缺失。(c) 上图示出靶向目标内源性 *cyp26b1* 基因的 gDNAs 所在位置。下图显示的是被编辑形成的突变分子的部分测序峰图, 其中 2610 bp 被缺失。

- 25 图 13 SGN 编辑内源性基因形成大片段缺失的假设机制。一个 gDNA 杂交到斑马鱼或人的基因组 DNA 的单链, 以形成 3'-Flap 结构; 与单链人工靶 DNA 的切割机理相同, SGN 结合至识别位点后切割斑马鱼基因组 DNA 的单链; 斑马鱼基因组 DNA 单链被切割后形成切刻或切口(nick)
- 30

结构并被 SGN 识别；一旦 SGN 分子结合到切刻结构，斑马鱼基因组 DNA 会被连续地切割成另一个新的具有切刻结构的切割产物。这意味着一旦切割开始时，它会一直重复；被破坏的基因组 DNA 在体内最终通过 DNA 修复途径修复从而形成突变分子。

- 5 图 14 示出的是关于 SGN 切割具有缺刻结构的双链 DNA 形成大的缺失片段。a: dsDNA 片段示意图。切刻内切酶 Nt.BstNBI 的识别位点用灰色标识，切割位点用箭头表示。b: SGN 切割产物用 2%琼脂糖凝胶分离。泳道 1: 没有 SGN 的缺刻 dsDNA 片段；泳道 2: 有 SGN 的缺刻 dsDNA 片段；泳道 M: DL-500 marker (500, 400, 300, 200, 150, 100 和 50
10 bp); 泳道 3: 没有 SGN 的 dsDNA 片段；泳道 4: 有 SGN 的 dsDNA 片段。*: 指示大片段缺失的双链 DNA。

具体实施方式

- 下面通过具体实施方式及实验数据对本发明作进一步的说明。尽管为了清楚的目的，在下文中使用了专用术语，但这些术语并不意味着定义
15 或限制本发明的范围。

如本文中所使用，术语“多核苷酸”指代是核苷酸聚合成的链状化合物。

- 如本文中所使用，术语“靶多核苷酸”是指，任何一段欲加以改造
20 或修复的目标核苷酸聚合成的链状化合物。靶多核苷酸附近的基因序列，允许外源序列在靶点处的整合，包括但不限于基因敲除（knock-out）、基因敲入（knock-in）。在具体实施方式中，靶多核苷酸是双链的 DNA 序列，包括，但不限于，细胞的染色体基因组中的 DNA 序列、细胞染色体基因组外的 DNA 序列（例如线粒体基因组）、质粒、病毒等的 DNA
25 序列。

“靶多核苷酸编辑”在某些具体实施例中指代是基因组编辑，包括由于核酸内切酶对靶基因的靶向切割导致细胞启动 DNA 损伤修复机制使得在没有外源供体 DNA 存在的情况下出现的靶基因的插入缺失突变，以及在外源供体存在的情况下外源供体靶向插入基因的基因组修饰。

- 30 “对靶多核苷酸特异”是指与靶多核苷酸结合，包括但不限于结合

可以存在一定的错配率。

如本文中所使用，术语“寡核苷酸”指代是是一类只有 60 个以下碱基的短链核苷酸的总称（包括脱氧核糖核酸 DNA 或核糖核酸 RNA 内的核苷酸）。

- 5 如本文中所使用，术语“目标序列”指代是靶序列，在具体实施例中目标基因组和靶基因组是相同含义；类似地，在一些具体实施例中目标 DNA 和靶 DNA 是相同含义，其中在一些具体实施例中使用的底物 DNA 也指代靶 DNA。

- 10 如本文中所使用，术语“核酸酶”指代是作用于水解核苷酸之间的磷酸二酯键的一种蛋白质。

引导 DNA（guide DNA）：指代是寡脱氧核糖核酸单链，亦被称作寡核苷酸探针，其中 5'序列对靶多核苷酸链特异，3'端有至少 1 个核苷酸不与靶多核苷酸单链互补，或错配。

- 15 如本文中所使用，术语“3'-Flap 结构”也称为“3'侧翼结构”，指代是由于引导 DNA（gDNA）5'端与靶多核苷酸互补形成双链但 3'端不与靶多核苷酸单链互补而在错配位置形成的一种特殊 DNA 结构。

下述实施例中的实验方法，如无特殊说明，均为常规方法。

20 具体实施例：

实施例 1 重组结构识别核酸内切酶的构建和的表达

- 25 本发明中构建了重组结构识别核酸内切酶（Structure-guided endonuclease, SGN），该内切酶由识别 3'-Flap 结构的 FEN-1 和切割 DNA 链的 *Fok* I（Fn1）切割结构域组成。其中，该序列由以下基因编码：C 端是 *Fok* I（196 个氨基酸残基），中间连接是甘氨酸-丝氨酸重复序列，和 FEN-1 酶。将该序列插入原核表达载体 pET28a(+)中，以形成 pET28a (+)-SGN。SGN 的编码序列和氨基酸序列见图 1、图 3-4 中示出，pET28a (+)-SGN 的质粒图谱见图 2 中示出。在 pET28a (+)-SGN 构建体中，SGN 基因位于 T7 启动子的下游。

30

pET28a(+)-SGN 用 CaCl_2 热休克的方法转化到宿主细菌菌株 *Arctic Express* 中。首先, 将细胞在 37°C 培养, 然后在 27°C 环境下用 IPTG (0.1 mM) 诱导 16 小时, 进而表达 SGN。收集经诱导细胞, 用超声裂解并离心。将粗提取物用镍 (Ni) 亲和色谱柱纯化 SGN。超速离心浓缩 SGN, 并用 12% 的十二烷基硫酸钠-聚丙烯酰胺凝胶电泳 (SDS-PAGE) 进行确认。获得的纯化 SGN (序列见图 1, 其中氨基酸序列为 SEQ ID No:1, 核酸序列为 SEQ ID No:3) 用于实例中的体外实验。

申请人早先假设, 一旦目标序列被 FEN-1 识别, SGN 的 Fn1 结构域应能切割 DNA 链。用于验证 SGN 能否切割靶 DNA 的体内和体外策略见图 5。在体外 (图 5, 左图), 靶 DNA 用 Cy5 基团标记修饰。设计互补于靶 DNA 的 gDNA, 其 3' 末端有未配对的核苷酸以形成 3'-Flap 结构。SGN (图 1) 识别 3'-Flap 结构并切割靶 DNA。切割的产物通过变性聚丙烯酰胺凝胶电泳 (变性的-PAGE) 和荧光成像进行分析。

然后, 使用的 *Tg (flk:eGFP)* 斑马鱼胚胎研究 SGN 在体内的活性 (图 5, 右图)。设计一对 gDNAs, 其与靶基因 GFP 正反义链分别互补, 且 gDNA 的 3' 末端有未配对的核苷酸以形成 3'-Flap 结构。将 SGN 的 mRNA (含有细胞核定位信号编码序列, 图 3, 其中核酸序列为 SEQ ID No:4, 氨基酸序列为核酸序列为 SEQ ID No:2) 和一对识别目标基因的 gDNAs 显微注射入斑马鱼胚胎。所表达的 SGN 会识别 3'-Flap 结构并切割体内靶 DNA。基因组 DNA 被消化并通过 DNA 修复途径被修复。为了检验该 DNA 编辑过程, 我们从斑马鱼胚胎中提取基因组 DNA, 然后进行 GFP 目的序列的 PCR 扩增、克隆和测序, 分析由 SGN 引起的突变。也可以采用上述类似实验方法进行野生型斑马鱼胚胎内源基因的突变。

在进行人类细胞实验时, 我们将有真核生物启动子驱动的 SGN (含核定位信号, 图 4, 核酸序列为 SEQ ID No:5, 氨基酸序列为核酸序列为 SEQ ID No:2) 表达质粒和一对识别目标基因的 gDNAs 转染人的细胞系。所表达的 SGN 会识别 3'-Flap 结构并切割体内靶 DNA。基因组 DNA 被消化并通过 DNA 修复途径被修复。为了检验该 DNA 编辑过程, 我们从人细胞系中提取基因组 DNA, 然后对目标基因组序列进行 PCR 扩增、克隆和测序, 分析由 SGN 引起的突变。

实施例 2 SGN 体外切割单链 DNA

为了测试设计 SGN 是否可以切割 DNA 链,我们将 1 ng SGN、10 pmol 底物单链 DNA (ssDNA) (S-1) 和 10 pmol gDNA-1 在 10-μL 的反应体系中进行孵育(所有的 ssDNA 和 DNA 寡核苷酸 gDNA 的序列示于表 1), 10-μL 的反应体系中还包括 MOPS (10 mM)、0.05% Tween-20、0.01% nonidet P-40 和 MgCl₂ (7.5 mM)。其中,用荧光 Cy5 基团在 S-1 的 5' 末端进行标记。且在 SGN 加入之前,先将混合物在 95℃下孵育 5 分钟, 55℃下孵育 10 分钟。然后再加入 SGN, 37℃反应 2 小时。

所述的 gDNA-1 与 S-1 形成 3'-Flap 结构。将 SGN 加入到该混合物中进行反应,然后将混合物通过变性的-PAGE 分离并通过荧光成像。具体步骤为,将反应得到的产物在变性条件下用 PAGE 进行分析。上样缓冲液含有 90%甲酰胺、0.5%EDTA, 0.1%二甲苯蓝,和 0.1%溴酚蓝。上样前,将样品 (20 μL) 孵育在沸水 5 分钟,然后在冰上冷却。然后将样品在室温上样到 20%PAGE 凝胶上,并在包含有尿素 (8.7 M) 和 Tris-硼酸盐 (89 mM) 的缓冲液中运行。电泳在 9.6 V/cm 运行 2 小时。电泳后,将凝胶浸渍在 10%乙醇固定 20 分钟。凝胶由 Tanon5200 多荧光成像仪成像 (上海,中国)。

理论上,只有经标记的目标链 (完整) 和 5'端用荧光染料 Cy5 标记的 S-1 的切割产物应被明显检测到。如图 6a 的泳道 5,通过 gDNA-1 导向作用,SGN 切割底物 S-1,产生较小分子量的条带,图中用“切割产物”表示。而仅含有 S-1 加 SGN (图 6a 的泳道 1)、S-1 加 gDNA-1 (图 6a 的泳道 2)、S-1 加 *Fok* I 和 gDNA-1 (图 6a 的泳道 3)、或 S 加 FEN-1 和 gDNA-1 (图 6a 的泳道 4) 的反应中没有发生切割现象。该结果表明, SGN 可以识别 3'-Flap 结构并切割的目的序列 DNA 链。

表 1. DNA 寡核苷酸模板、gDNA 和 PCR 引物的序列

寡核苷酸名称	序列 (5' -3')
gDNA-1	atgtcacttccccttggttctctcc
gDNA-1-G	atgtcacttccccttggttctctcg
gDNA-1-T	atgtcacttccccttggttctctct

S-1	Cy 5-ctattgcaccaggccagatgagagaaccaaggggaagtgacat
gDNA-2	ccgaagggcatgagctgcc
S-2	Cy 5-cgtgcagctcatcatgcagcagctcatgccctcgg
gDNA-3	tggttctcctgaggcctcc
gDNA-3-G	tggttctcctgaggcctcg
gDNA-3-A	tggttctcctgaggcctca
S-3	Cy 5-tctacaaccaaggagagaggcctcaggaggaacca
gDNA-4	catgtcaagatcacagatttgggcc
gDNA-4-A	catgtcaagatcacagatttgggca
gDNA-4-T	catgtcaagatcacagatttgggct
S-4	Cy 5-cagcagtttggcccgcccaaatctgtgatcttgacatg
gDNA-5	gaagcctacgtgatggccaa
gDNA-5-T	gaagcctacgtgatggccat
gDNA-5-G	gaagcctacgtgatggccag
S-5	Cy 5-acgtgggggtgtccacggaggccatcacgtaggcttc
gDNA-6-10nt	atgtcactta
gDNA-6-15nt	atgtcacttcccctg
gDNA-6-20nt	atgtcacttccccttggtta
gDNA-6-25nt (同 gDNA-1)	atgtcacttccccttggttctctcc
gDNA-6-30nt	atgtcacttccccttggttctctcatctgt
gDNA-6-35nt	atgtcacttccccttggttctctcatctggcctgt
gDNA-6-40nt	atgtcacttccccttggttctctcatctggcctggtgcaC
gDNA-6-45nt	atgtcacttccccttggttctctcatctggcctggtgcaataggga
gDNA-6-50nt	atgtcacttccccttggttctctcatctggcctggtgcaataggccctga
gDNA-6-55nt	atgtcacttccccttggttctctcatctggcctggtgcaataggccctgcatgtC
gDNA-6-60nt	atgtcacttccccttggttctctcatctggcctggtgcaataggccctgcatgtactggC
S-6	Cy 5-atagattacatccagatcatgcagggcctattgcaccaggccagatgagagaaccaaggggaagtgacat
gDNA-7-10nt	tggttctca
gDNA-7-15nt	tggttctcctgagt
gDNA-7-20nt (同 gDNA-3)	tggttctcctgaggcctcc
gDNA-7-25nt	tggttctcctgaggcctctctccg
gDNA-7-30nt	tggttctcctgaggcctctctccttgggg
gDNA-7-35nt	tggttctcctgaggcctctctccttgggtgtat
gDNA-7-40nt	tggttctcctgaggcctctctccttgggtgtagatgcg
gDNA-7-45nt	tggttctcctgaggcctctctccttgggtgtagatgctgtctg
gDNA-7-50nt	tggttctcctgaggcctctctccttgggtgtagatgctgtcttcca
gDNA-7-55nt	tggttctcctgaggcctctctccttgggtgtagatgctgtcttccctgtgg
gDNA-7-60nt	tggttctcctgaggcctctctccttgggtgtagatgctgtcttccctgtgtctcc

S-7	Cy 5-gacaacctgtgaggacacaggagaagacagcatctacaaccaaggagagaggcctcaggaggaacc
gDNA-8	ctcctgattatgac
S-8	FAM-tcagccactggcttctgtcataatcaggag-ECLIPSE
gDNA-9	ctattgcaccaggccagc
S-9	atgtcacttccccttggctctcatctggcctgggtgcaatag
gDNA-10	tctacaaccaaggat
S-10	tggttcctcctgaggcctctctccttgggtttaga
gDNA-11	acgtgggggttgcact
S-11	gaagcctacgtgatggccaccgtggacaacccccacgt
GFP F1	gtgagcaagggcgaggagctg
GFP R1	ctgtacagctcgtccatgccg
GFP F2	aagggcgaggagctgttcacc
GFP R2	agctcgtccatgccgagagtg
gDNA-gfp-F	ccgagggtgaagttcagggcgacaca
gDNA- gfp -R0	cccttcagctcgtatgcggttcacat
gDNA- gfp -R8	agtcgatgcccttcagctcgtatgcg
gDNA- gfp -R18	tcctccttgaagtcgatgcccttca
gDNA- gfp -R32	ccaggatgttgccgtcctccttgaa
gDNA- gfp -R50	tgtactccagcttgtgccccaggat
gDNA-znf703-F	cccctagcattgctggctcagaccta
gDNA-znf703-R50	ttgtctccatggctgctggaggagaa
znf703 F	tttcagtcagtcacagg
znf703 R	gactggagttggcgatttgt
gDNA-cyp26b1-F	gcagctctggcagctgaggtggacct
gDNA-cyp26b1-R32	ggggaaccccatggagcccttggcg
cyp26b1 F	atcggcacaaaaggtaaagc
cyp26b1 R	atgatcttatgtggcacagtc
cyp26b1 Fin	attcatagaaacgcgttcac
cyp26b1 Rin	agctcatattgtagctcagc
T7	taatacgactcactataggg
Sp6	atttagtgacactatag
gDNA-MES2-F	ttcaacgaggacatcgggtcttcga
gDNA-MEIS2-R	tggatcaagaaggctggacctgcttt
MEIS2-F	tgttctggcttaggctcg
MEIS2-R	agcagaaagtgggtgctt

实施例 3 SGN 切割活性不依赖于目标序列

为证明 SGN 是否有 DNA 序列的偏好性, 采用具有不同序列的 ssDNAs (S-2、S-3, 序列见表 1) 作为 SGN 的底物进行反应。10- μ L 的反应体系和条件同前述以 S-1 为底物的反应。

结果表明, 当通过 gDNA-2 或 gDNA-3 引导下, SGN 可分别将 S-2 或 S-3 切割 (见图 6b 和图 6c 的泳道 5)。而仅含有 S 加 SGN (图 6b 和图 6c 的泳道 1)、S 加 gDNA (图 6b 和图 6c 的泳道 2)、S 加 *Fok* I 和 gDNA (图 6b 和图 6c 的泳道 3)、或 S 加 FEN-1 和 gDNA (图 6b 和图 6c 的泳道 4) 的反应中没有发生切割现象。该结果表明, SGN 切割活性不依赖于目标序列但识别 3'-Flap 结构。

为了证明 gDNA 中未配对 3'核苷酸的重要性, 本发明中测试了所有类型的非配对 3'核苷酸, 包括 C-T、G-T、T-T、C-A、G-A、A-A、C-C 和 G-G。如图 7 所示, 不同非配对类型之间的效率没有明显的差异。先前报道中 (Kaiser MW, et al. J Biol Chem. 1999, 274:21387-21394.), 古细菌 FEN-1 酶中所有四个天然碱基有大致相同的效率。本发明的结果与这一发现是一致的。然后, 申请人测试了多个序列来揭示 gDNA 长度对切割效率的重要性, gDNA 长度包括 10、15、20、25、30、35、40、45、50、55、60 个核苷酸 (nt)。如图 8 所示, 当 gDNAs 的长度超过 20-nt 时, SGN 可以切割目标 DNA。但当 gDNAs 的长度为 10-nt 或 15-nt 时, 目标 DNA 没有发生裂解。根据对 *Fok* I 的结构的研究报道 (Wah DA, et al. Proc Natl Acad Sci U S A 1998, 95:10564-10569.), 当 *Fok* I 与 DNA 结合时, 切割结构域会结合在距离识别位点 13bp 的位置。这一结果表明, 该蛋白需要底物 DNA 上有足够的构象空间进行折叠和反应。然而, 如在图 6 上部中灰色箭头所示的, 当 gDNAs 的长度为 10 nt 的或 15 nt 的, 理论上的切割位点过于接近目标 DNA 的 3'末端, 在蛋白质折叠过程中所需的构象或许没有足够的空间。

实施例 4 距离非配对 3'末端核苷酸 9-10nt 处是 SGN 切割活性位点

已证明活性的 *Fok* I 是以二聚体形式发挥作用 (Wah DA, et al. Proc Natl Acad Sci U S A 1998, 95:10564-10569.)。为了研究是否 SGN 也是以二聚体形式工作, 申请人用 SGN 对 DNA 切割, 进行动力学实验。实验

前,标记底物 S-8,方法为在 5'端标记荧光染料 FAM 和在 3'端加猝灭剂,然后与 SGN 和 gDNA-8 共同孵育。通过绘制相对于各种浓度 SGN 的速率(图 9),发现当使用 0.62 nM SGN 后反应速度达到平台期。但使用小于 0.62 nM SGN 的速率/SGN 浓度曲线方程式是 $y=2.9684 x^{1.9255}$ (y=速率, x=SGN 浓度),其中 R 平方为 0.9866。这些结果证明了反应初始速度与酶浓度并不成正比,从而表明了 SGN-催化反应不是相对于 SGN 的浓度的一阶反应,这也暗示 SGN 是以二聚体发挥作用的。此外,SGN 是负责识别的 3'-Flap 结构的 FEN-1 以及负责 DNA 切割的 *Fok I* 切割结构域(Fn1)二者的融合蛋白。如先前报道(Wah DA, et al. Proc Natl Acad Sci U S A 1998, 95:10564-10569.), *Fok I* 通过其切割结构域介导的二聚体切割靶 DNA。*Fok I* 工作模型中(Wah DA, et al. Proc Natl Acad Sci U S A 1998, 95:10564-10569.), *Fok I* 分子结合在识别位点并招募另一个 *Fok I* 分子,它通过 Fn1 域提供第二催化中心。第一 *Fok I* 分子的 Fn1 在结合特定 DNA 后被激活,摆动成为一个二聚体开放构象,然后切割。因此,申请人推断,SGN 的二聚化是通过 Fn1 介导的。

为确定由 SGN 产生的切割位点的位置,申请人使用焦磷酸测序(Zhou G, et al. Anal Chem. 2006, 78:4482-4489.)对切割的链进行测序(图 10a)。

使用便携式生物发光分析仪(日立有限公司,日本)进行焦磷酸测序(Zhou G, et al. Anal Chem. 2006, 78:4482-4489.)。用链霉亲和素包被的琼脂糖珠收集生物素化的反应产物。待沉淀和洗涤后,将纯化的 dsDNAs 在碱缓冲液中变性,得到 ssDNAs。然后将固定化的生物素化的链与测序引物在 90℃的条件下退火 5 分钟。55℃反应 10 分钟,作为测序模板。焦磷酸测序混合物含有 Tris-HAC (0.1M, pH 7.7)、EDTA (2 mM)、Mg(Ac)₂ (10 mM)、0.1% BSA、DTT (1 mM)、APS (2 mM)、PVP (0.4 g/L)、D-luciferin (0.4 mM)、apyrase-VII (1.6 U/mL)、外切 Klenow 片段 (18 U/mL) 和萤火虫荧光素酶 (1 mM)。

靶序列(S-1 和 S-2)在 5'末端用生物素修饰。底物 DNAs 与 SGN 孵育后,链霉素包被的琼脂糖珠被用于捕获生物素化的反应产物。然后将固定的生物素化的链与它的 gDNA 在 85℃下退火 5 分钟,25℃下进行

10 分钟，然后用作测序模板。非配对的 3'核苷酸不会发生延伸。因此，
测序信号出现在切割位点（图 10a）。如图 10b 所示，焦磷酸测序信号
是 GGAAGTGAC。结果表明，由 SGN 介导的 S-1 切割位点位于距离
gDNA-6-20nt 的 3'末端核苷酸 9nt 的位置。针对 S-2（图 10c）的类似焦
5 磷酸测序中，发现焦磷酸测序信号是 GCCCTTC。由 SGN 介导的 S-2 切
割位点位于距离 gDNA-6-20nt 的 3'末端核苷酸 10 nt 的位置。因此，该结
果表明在切割位点距离 gDNAs 的 3'端有 9-10 个核苷酸的间隔。

实施例 5 SGN 体外部分切割双链 DNA (dsDNA)

10 既然 SGN 能够切割单链 DNA 且不依赖靶序列，而依赖由 gDNA 形
成的 3'-Flap 的方式，因此 SGN 可具备在体外酶切双链 DNA 的能力。为
证明 SGN 是否可以切割 DNA，申请人进行了三个靶 dsDNA 的检测。

首先分别孵育包含 5 pmol S-1/S-9 (S-1 的互补链) 或 S-3/S-10 (S-3
的互补链) 或 S-5/S-11 (S-5 的互补链) 的混合物以形成 dsDNA，条件
15 为 95°C 3 分钟，然后从 94°C 1 分钟开始以 1°C/分钟的速率冷却到 22°C
保持 1 分钟。在这个过程中，互补链可以相互杂交，然后分别添加 5 pmol
gDNA-1/gDNA-9 或 gDNA-3/gDNA-10 或 gDNA-5/gDNA-11，和 10 mM
MOPS、0.05% Tween-20、0.01% nonidet P-40 和 7.5 mM MgCl₂ 至混合
物中，37°C 孵育 10 分钟。最后加入 1 ng SGN 37°C 孵育 2 小时。

20 结果如图 11a、b、c 中泳道 1 所示，在包含 SGN、一对 gDNAs 和
DNA 底物 (S) 的反应中，观察到更小分子量的条带，用 “*” 标记。但
在只包含 S 和 SGN (11a、b、c 中泳道 2) 或只包含 S 和 gDNA (11a、b、
c 中泳道 3) 的反应中，没有发生切割。这些结果说明了由一对 gDNAs
引导的 SGN 可以在体外部分切割 dsDNA。

25

实施例 6 SGN 体内对基因组 DNA 的编辑

为确定 SGN 是否具有基因组编辑的活性，申请人将 1-nL 含有 200
pg 的 SGN mRNA (含有核定位信号编码序列，图 3) 和一对 gDNAs (各
50 pg，序列见表 1) 的溶液通过显微注射入 1-细胞期的 Tg (*flkl*: EGFP)
30 斑马鱼胚胎。在转基因斑马鱼的基因组中编码 GFP 的双链 DNA 的正义

链和反义链上, gDNAs 被间隔放置分别形成 3'-Flap 结构(图 12a, 顶部)。

突变的 GFP 分子用如之前文献所述的方法检测 (Dong Z et al. PLoS One 2011, 6:e28897.)。简言之, 根据说明书(南京尧顺禹, 中国)的步骤, 用 10 μ L 的 B 液孵育斑马鱼胚胎, 之后从随机挑选的 5 个胚胎中制备斑马鱼基因组 DNA 模板 (65 $^{\circ}$ C 反应 30 分钟, 95 $^{\circ}$ C 反应 10 分钟, 16 $^{\circ}$ C 反应 1 分钟)。然后将 1.0 μ L 裂解溶液作为模板, 用引物 GFP F1 和 GFP R1 (表 1) 在 20 μ L 的 PCR 混合物体系中扩增 GFP 分子。PCR 程序为: 94 $^{\circ}$ C 2 分钟, 35 个循环 (94 $^{\circ}$ C 30 秒, 60 $^{\circ}$ C 30 秒, 72 $^{\circ}$ C 30 秒), 72 $^{\circ}$ C 进行 10 分钟。扩增子在 pGEM-T 中 (Promega 公司, USA) 进行克隆。然后, 随机选择 48 个阳性转化子, 采用引物 GFP F2 和 GFP R2 (表 1) 以上面所述的方法进行 PCR 鉴定。接着, 对阳性转化子中靶分子进行测序, 以确定突变。

结果发现, 当 gDNA 间隔 0、8、18、32 或 50 个 bp, 对从显微注射了 gDNAs 与 SGN mRNA 的斑马鱼胚胎中扩增出的基因组靶序列进行测序分析, 分别揭示了 2/48、0/47、3/46、18/44、12/47 的突变比率 (表 2, 表 3)。结果表明, 基因组中, SGN 偏好于识别以 32-50 bp 的间隔序列放置的 gDNAs。

另一方面, 用 *znf703* 和 *cyp26b1* 基因作为靶基因以验证 SGN 是否可以像 ZFN 或 TALEN 那样编辑内源基因。申请人分别向 1-细胞期斑马鱼胚胎中显微注射了 1-nL 含有 200 pg 的 SGN mRNA (含有核定位信号编码序列) 和一对 gDNAs (各 50 pg, 序列见表 1) 的溶液。用上述相似方法对突变的 *znf703* 和 *cyp26b1* 基因进行检测。简言之, 1.0 μ L 来自于 5 个胚胎的裂解溶液作为模板, 分别用引物 *znf703* F1 和 *znf703* R1、*cyp26b1* F 和 *cyp26b1* R (表 1) 在 20 μ L 的 PCR 混合物体系中扩增 *znf703* 和 *cyp26b1* 分子。PCR 程序为: 95 $^{\circ}$ C 3 分钟, 30 个循环 (95 $^{\circ}$ C 15 秒, 60 $^{\circ}$ C 15 秒, 72 $^{\circ}$ C 1 分钟 30 秒), 72 $^{\circ}$ C 进行 10 分钟。为增加 *cyp26b1* 扩增的特异性, 采用了引物 *cyp26b1* Fin 和 *cyp26b1* Rin (表 1) 进行了巢式 PCR。*znf703* 扩增子和 *cyp26b1* 巢式 PCR 的扩增子分别在 pGEM-T 中 (Promega 公司, USA) 进行克隆。然后, 随机选择 96 个转化子, 采用 T7 和 Sp6 引物对 (表 1) 进行 PCR (94 $^{\circ}$ C 2 min, 30 个循环 (94 $^{\circ}$ C 30

秒, 54 °C 30 秒, 72°C 3 分钟 10 秒), 72 秒 10 分钟)。PCR 产物稀释 10 倍后, 分别用引物 *znf703* F 和 *znf703* R、*cyp26b1* Fin 和 *cyp26b1* Rin (表 1) 再进行 PCR 检测, 确定插入的片段是否来源于靶基因的突变等位基因。最后对阳性转化子靶分子进行测序鉴定突变基因型。

- 5 针对自 *znf703* 基因扩增出的等位基因片段的测序分析显示, gDNA-*znf703*-F 和 gDNA-*znf703*-R50 引导下的 SGN 诱发了比例为 1/96 的等位基因突变 (图 12b、表 2、表 4)。发生基因组编辑的突变分子中 754 个核苷酸被移除, 另外 11 个核苷酸也被删除 (图 12b, 底部)。针对自 *cyp26b1* 基因扩增出的等位基因片段的测序分析表明, 由
- 10 gDNA-*cyp26b1*-F 和 gDNA-*cyp26b1*-R32 引导的 SGN 引发了比例为 3/29 的突变 (图 12c、表 2、表 5), 发生基因组编辑的突变分子中伴随有 2610 个核苷酸大片段缺失突变 (图 12c, 底部)。结果表明, SGN 可以以低效率编辑斑马鱼基因组的内源基因。

15 表 2. 斑马鱼基因组中的目标基因经 SGN 编辑后形成的突变等位基因

靶基因	gDNAs-R	克隆号	缺失 (bp)	插入片段 (bp)	gDNA 间隔 (bp)
<i>Tg(flk1:gfp)</i> 斑马鱼中 GFP 报告基因	gDNA-gfp-R0	#1-44	659	3	0
		#1-46	659	3	0
	gDNA-gfp-R8	/	/	/	8
	gDNA-gfp-R18	#3-18	651	1	18
		#3-21	651	1	18
		#3-23	659	3	18
	gDNA-gfp-R32	#4-2	651	1	32
		#4-3	651	1	32
		#4-4	651	1	32
		#4-9	639	17	32
		#4-11	649	20	32
		#4-14	659	3	32
		#4-16	644	0	32
		#4-17	648	1	32
		#4-18	659	3	32
		#4-19	644	0	32
		#4-22	649	20	32

		#4-33	649	20	32
		#4-34	644	0	32
		#4-37	651	1	32
		#4-39	639	17	32
		#4-42	648	1	32
		#4-44	639	17	32
		#4-48	651	1	32
		#5-9	651	0	50
		#5-11	651	0	50
		#5-14	651	0	50
		#5-15	596	12	50
		#5-20	651	0	50
	gDNA-gfp-R50	#5-23	651	0	50
		#5-24	651	0	50
		#5-27	651	0	50
		#5-28	651	0	50
		#5-29	611	3	50
		#5-31	651	0	50
		#5-48	611	3	50
内 源 基 因					
<i>znf703</i>	gDNA-znf703-R50	#6	754 + 11	0	50
内 源 基 因		#7-1	2610	0	32
<i>cyp26b1</i>	gDNA-cyp26b1-R32	#7-2	2610	0	32
		#7-3	2610	0	32

发生基因靶向突变的比率：#1(2/48)、#2(0/47)、#3(3/46);、#4(18/44)、#5(12/47);、#6(1/96) 和 #7(3/29)。

表 3. *eGFP* (720 bp) 编码序列和突变序列

<i>eGFP</i> (720 bp) 编码序列	
ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACG	
TAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCT	
GAAGTTCATCTGCACCAACCGCAAGCTGCCCGTGCCCTGGCCACCCCTCGTGACCACCCCTGACCTACG	
GCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCC	
GAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGG	
TGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGG	
CAACATCCTGGGGCACAAGCTGGAGTACAACCTACAACAGCCACAACGTCTATATCATGGCCGACAAGC	
AGAAGAACGGCATCAAGGTGAAGTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGACGCTCGC	
CGACCACTACCAGCAGAACACCCCATCGGCGACGGCCCCGTGCTGCTGCCCCGACAACCACTACCTG	

AGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTTCG
TGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAA

Tg(flk:eGFP) 转基因斑马鱼的 *eGFP* 报告基因经 SGN 编辑后的突变分子的序列

1-44

GTGAGCAAGGGCGAGGAGCTGTTACC-----659bp---CTT (3bp 插入)----- CACTCTCGGCATGGACGAG
CTGTACAAG

1-46

GTGAGCAAGGGCGAGGAGCTGTTACC-----659bp---CTT (3bp 插入)----- CACTCTCGGCATGGACGAG
CTGTACAAG

3-18

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)----- ATCACTCTCGGCAT
GGACGAGCTGTACAAG

3-21

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)----- ATCACTCTCGGCAT
GGACGAGCTGTACAAG

3-23

GTGAGCAAGGGCGAGGAGCTGTTACC-----659bp---CTT (3bp 插入)----- CACTCTCGGCATGGACGAG
CTGTACAAG

4-2

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)-----ATCACTCTCGGCATG
GACGAGCTGTACAAG

4-3

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)-----ATCACTCTCGGCATG
GACGAGCTGTACAAG

4-4

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)-----ATCACTCTCGGCATG
GACGAGCTGTACAAG

4-9

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCC-----639bp-----ATCCTGTTTGTGCCCC
A (17bp 插入)-----GGATCACTCTCGGCATGGACGAGCTGTACAAG

4-11

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG-----649bp-----GCCTAACTACGGCTGCTCGT (2
0bp 插入)-----TCACTCTCGGCATGGACGAGCTGTACAAG

4-14

GTGAGCAAGGGCGAGGAGCTGTTACC-----659bp---CTT (3bp 插入)----- CACTCTCGGCATGGACGAG
CTGTACAAG

4-16

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC-----644bp-----CACTCTCGGCATGGAC
GAGCTGTACAAG

4-17

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG-----648bp-----C (1bp 插入)----- ATCACTCTCGG
CATGGACGAGCTGTACAAG

4-18

GTGAGCAAGGGCGAGGAGCTGTTACC-----659bp---CTT (3bp 插入)----- CACTCTCGGCATGGACGAG
CTGTACAAG

4-19

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC-----644bp-----CACTCTCGGCATGGAC
GAGCTGTACAAG

4-22

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG-----649bp-----GCCTAACTACGGCTGCTCGT(20
bp 插入)-----TCACTCTCGGCATGGACGAGCTGTACAAG

4-33

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG-----649bp-----GCCTAACTACGGCTGCTCGT(20
bp 插入) -----TCACTCTCGGCATGGACGAGCTGTACAAG

4-34

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATC-----644bp-----CACTCTCGGCATGGAC
GAGCTGTACAAG

4-37

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)----- ATCACTCTCGGCAT
GGACGAGCTGTACAAG

4-39

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCC-----639bp-----ATCCTGTTTGTGCCCC
A(17bp 插入)-----GGATCACTCTCGGCATGGACGAGCTGTACAAG

4-42

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTG-----648bp-----C (1bp 插入)----- ATCACTCTCGG
CATGGACGAGCTGTACAAG

4-44

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTGGTGCCCATCC-----639bp-----ATCCTGTTTGTGCCCC
A(17bp 插入)-----GGATCACTCTCGGCATGGACGAGCTGTACAAG

4-48

GTGAGCAAGGGCGAGGAGCTGTTACCGGGGTG-----651bp---A (1bp 插入)----- ATCACTCTCGGCAT
GGACGAGCTGTACAAG

5-9

GTGAGCAAGGGCGAGGAGCTGTTACCGG---651bp-----CGGGATCACTCTCGGCATGGACGAGCTGT
ACAAG

5-11	GTGAGCAAGGGCGAGGAGCTGTTACCGG----651bp-----CGGGATCACTCTCGGCATGGACGAGCTGT ACAAG
5-14	GTGAGCAAGGGCGAGGAGCTGTTACCGG----651bp-----CGGGATCACTCTCGGCATGGACGAGCTGT ACAAG
5-15	GTGAGCAAGGGCGAGGAGCTGTTACCG-----596bp-----CTGTTTTTGGCG (12bp 插入)-----AAAGACC CCAACGAGAAGCGGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATG GACGAGCTGTACAAG
5-20	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCTGT ACAAG
5-23	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCT
5-24	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCTGT ACAAG
5-27	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCTGT ACAAG
5-28	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCTGT ACAAG
5-29	GTGAGCAAGGGCGAGGAGCTGTTACCGGGGCGGTGCCCATC-----611bp-----ACC (3bp 插入)-----G GTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG
5-31	GTGAGCAAGGGCGAGGAGCTGTTACCGGCGGG----651bp----- ATCACTCTCGGCATGGACGAGCTGT ACAAG
5-48	GTGAGCAAGGGCGAGGAGCTGTTACCGGGGCGGTGCCCATC-----611bp-----ACC (3bp 插入)-----G GTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAG

表 4. *znf703* 野生序列和经 SGN 编辑后形成的突变序列

野生型 <i>znf703</i> 的部分序列
TTTGCAGTCAGTCCACGGAACCAGCAGACCTTTAAAAGTCATTACATTACCAACGATAGTTCGCCCAC

TTGTTTAGGACTCCAGGTCGACTCGATCGCCTCACTGCCGTACCGTGATCCCGCGCGTCAGGAGAAAA
GGCTGCCTATACGGATTCTCAAATGTTGACCGCTCACACAAGTCACATACTTCACCCAGAGTATCTCC
AGCCGCTCAGCTCGGCGCCCGTGAGCATTGAGGTACGACTTCTAAACCATTTGTTTCTTACTTCTGCGA
AGTAAAAATATATACATTACATTTGTATGTGTTAGAGTTGCGTGTTTGTCTAGTTTCTATAAAGCGAAA
GGGTGTTCTCTTTATTTTGTGCCCGAGTCGTGTAGTACACGCAATATTTTGCCTCAACGTTTTTTAGCTC
TAAGCTCCATGTTGCGCGGCTCGTTTAAATATGTAATTCGGATTTAGCCGAAAGATCCTGAAAGCGATT
AGCGGCAGTTTGAAGAAATAGTCCGATTTAGATGAGGCCGTGCTCTGGCAACATTTTAAGGTTTCAC
ATTCCTTAATCGCTACTGACGTGCGCCTTTACTGCTCTTAGGTGGTTAAAGCTTCCAGCATCTACGTGC
TTTTTTGGCACACAAATGGAACTTTAAAAAAGTAAATTGTGCAATCTTTTAAATAACCTGTATAAACT
CTGCATGCTCGGCACACTTTGAGTGCTGCAAACTACAACAATATGTCAAGTGATTTAATTCACCTT
GACATGCAAATTAACATCATTTAAAAATAAAATTTAATATCCATGCGAGTTTTTTATGATGATGATGTCATT
TATAAATAATCCAAGCGTAAATAGAATTATTGCATCTTCCTTGAAATATGACATATCACTGAGGGCATGC
CAAAAAATTGGTCTATAAATGAAGTGTCATTTTGTGGTGGTAAGAAAACCAAAATCATTCATTTAAAA
TTTCTTCTTTTTTTCCCATAGCTGGATGCTAAAAAGAGTCCCCTAGCATTGCTGGCTCAGACCTGCTCTC
AGATCGGAAAGCCAGATCCTCCTCCCTCCTCTAAACTGGGATCCCCTCCTCCAGCAGCCATGGAGAC
AAGGACAGCCGCTCCAGCAGCTCTAGCCTGAAGTCTGGGGAGCATCAGAACCTAGATGACAAGTCCA
GCTTTAAGCCTTACTCTAAACTGGATCAGAATGCCGAAAGGAGGGTGCTGGGATCAACAGCTCAGCA
GATAAAGCGGGGTTTCAAGTGCCAAATGGCAGTTCTTCATCTGTGACGTGCACGTCTTTACCTCCGCAT
GCCCCATCTCCACGGGCCAGCTCCCCCTCAGCAAACATCTGGACAGTCACACACACACCGGCAGTCTC
AGTCTCCTCTTTACAGAAAACAGCACATCTACAAACCACCCACATGGACTCCAAAGCTGCAGGTTCA
GACCCAGGGAATGACAGCAGCAGCAGTGGTAGTGATCGCAATGGCAAAAAAGATTTCGGACCACAATA
AATCAAGCCTGGACGTTGTACAAATCGCCAACCTCCAGTC

经 SGN 编辑后形成的突变 *znf703* 的序列

#6

TTTGCAGTCAGTCCACGGAACCAGCAGACCTTTAAAAGTCATTACATTACCAACGATAGTTCGCCAC
TTGTTTAGGACTCCAGGTCGACTCGATCGCCTCACTGCCGTACCGTGATCCCGCGCGTCAGGAGAAAA
GGCTGCCTATACGGATTCTCAAATGTTGACCGCTCACACAAGTCACATACTTCACCCAGAGTATCTCC
AGCCGCTCAGCTCGGCGCCCGTGAGCATTG-----754bp----AGCTGGATGCTAAAAAGAGTCCCCTAGCAT
TGCTGGCTCAGACCTGCTCTCAGATCGGAAAGCCAGATCCTCCTCCCTCCTCTAAACT-----11bp-----CCT
CCAGCAGCCATGGAGACAAGGACAGCCGCTCCAGCAGCTCTAGCCTGAAGTCTGGGGAGCATCAGAA
CCTAGATGACAAGTCCAGCTTTAAGCCTTACTCTAAACTGGATCAGAATGCCGAAAGGAGGGTGCTG
GGATCAACAGCTCAGCAGATAAAGCGGGGTTTCAAGTGCCAAATGGCAGTTCTTCATCTGTGACGTGC
ACGTCTTTACCTCCGCATGCCCCATCTCCACGGGCCAGCTCCCCTCAGCAAACATCTGGACAGTCACA
CACACACCGGCAGTCTCAGTCTCCTCTTTACAGAAAACAGCACATCTACAAACCACCCACATGGACT
CCAAAGCTGCAGGTTTCAAGCCAGGGAATGACAGCAGCAGCAGTGGTAGTGATCGCAATGGCAAAAA
AGATTTCGGACCACAATAAATCAAGCCTGGACGTTGTACAAATCGCCAACCTCCAGTC

表 5 *cyp26b1* 野生序列和经 SGN 编辑后形成的突变序列

cyp26b1 野生型部分序列

ATTCATAGAAACGCGCTTCACAGATCGAAATATTTGGCCATATCAATAATTTTGGAAGGACTATCTGAT
ATAATTAAGAAAACACTGCCAGTTTATGAGGGATTTATTTACATTATTTTATAGAAATCGACTTATAACAT
CAACCCCCGAAAGTTGTAAATATAAACGCAAAACAAATATGAAACGCGTAAGTCAATGAATGAATCGC
T TACTCATGCAGTGTTTACAGAGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCAACAGATGCG
CATCTGAACAATCAAAACATTTTGAAAAAGTCATGAACATATCGATTTAGCATAAGAGTCCGTAGCTTA
TATGATCCGGTAAAATGTAGAGTATTATTTACATTTTTGTTGTGATGGTTTTAAACCATCATTTGGGTAA
TCAGTGTACCTCAGTTTGAGCGCATTTTAAGTCCACAACATGAATTAACCCTGTGCGTATTCCGCGAGT
GGCAGGTGACAAGTTCCTCGACCGTAAATGAGATTGAAATATTGAAAGTGTTCGCGCCTCTCCCTCTCT
TTCTGTTTGGGCTCTCTTCCAGGTCCCGCACGCTCCAGCGCCCGTCCCGTTCACTTTTCACCCCGCCTC
AGTTCCGGTGGGGTGGGGTAAGGGGGGAGCCGCTTGACCCCTGGCCGAATGCCGCCTCACTGCCTCTG
CTATATAAACACACATTGGAGGCTCGCTGTGCGACTCGAACTACATTAAGAACTGTATGTGAAGAG
AGAGAGAGAAATAGAGAGAGAGAGAGAGAGAGAAAGAGAGAGAGAGAGAGAGCTGGAGCGCAAGAGAGAG
GTAGAGAGTGAAACCGCATCAGAAAACGGGCTAGCGAGACGCACTACAGTCATCTGAAGATTTTTACT
GCTCATAGGCATCAAACCTTCGAGGATTATTTCTTTCATTTATTTCTTCGTCTTCCATATGACCATTACTT
GATTTTCAGTGTATTTAAAGAGTTTATATCGAGCTTTAAGATTATATATAGGCCTATTTAAAAATTATACTA
AAACAATTTATTTTATATTTTCAAAAAGTAACCTTAAAACCTTTTTCTTGATTACTTATATTTGCCTTTTTTAC
TTTTGAAATTTAAAAACAACGAATCAATCTCAAGTTGGACAAGTTCACCTCCAAATTTACACTTCAAAAAT
TTGAATTTAATTGATGAATTCAGAAATATTGGTCACCGATTGGATGATTTAAATGATATACACACATT
TAAGATCGATTTGTTATATCCAGCATTCTGTTAATTAAATAGCCGACTATATTTATTTATTCATTTACTCTA
CATAAACTTTTAGGGTTGGATTTGTGCTTTTTCGAATAATTGTGCGAGAGAGTCATATCTAACAGAGTAG
GTCACTCTCCTAATTTTAGGTTTAAACCACTTTATTGCTCATCACTCCAAAGAGATATTTGAGACAAGTCC
CCGGACGTTACAGCCATGCTCTTCGAGAGTTTTGACCTTGTCTCGGCGTTGGCGACGCTGGCTGCGT
GTTTAGTGTCCATGGCACTTCTTCTGGCCGTGTCCAGCAGCTCTGGCAGCTGAGGTGGACCGCAACA
CGGGACAAGAGCTGCAAGCTGCCCATGCCAAGGGCTCCATGGGGTTCCCCATCATCGAGAAACAT
GCCACTGGTTCTTTCAGGTAAGCGCTCATTCTCCTTGTGCCTTTCCAAGTTTCCAGTCTCCCTGCTGTC
GCACGCACAGCGCTCCGGCGCGCCAGACACCGCGGGCTCGGCGCATATAGTTTTTCTGCAAGTAACT
GGGTTGTAAAGACACTTCAAGAGTAGGATAACGTCCACTAATGCTTTTTCGCACGTGTGCACTCGTGA
GCAATATTGCGCTTTAAAATAAGAACAGCTTGCAAGGTAGTTATTTCCACGTGTGAATTTCTTCCGTGC
CATATTTCCATGGTGTTAATAAGCTGTCTGTTTTAGATAGCCGTCAGTTTTGTTAATACAGCCAGTTCTC
AATCCACGTGGAACAATCTGTTTTCGGCTGCCAAGGAAGGCGCACTCCGTTCCAGAAATATTGTAGTC
CTTGGCTTCCGATGCGCGCGAGAACACAGCCACTGAGCACCGCTTGTCTAAGCGCGTCTGTTTGCGCG
CGGGGTAAAGATTGATCGCGAGCCGGGGGAGGTATAGTGTGCGCTCTTATTGAGCTTCAGTGTAATTA
TGCGCGCCGGTTTCCCAAAAAAGCCACCGTGTGTGAATGATGCCGTGCCGTTCTGTGATGATACCCA
GACAGCGCGTGGGGCGCATAAAGTGCACGCTGGAGGTCTCTCTCTCTCTCTCTCACTGTCCGCAGTAA
GAAAGTTTGTGAACGCTCAGATGAATAAAAAAAGTTATAAATTAAGATCAGACCAGCGCTATAATG
TGGTATTTTGAAAAGAGTTTATTACGCGAACCTTAGCTGTAAAATTGTCTCAAGAATTCTTTCCATATTA

AATCCACGGTGCACATTGGGATGTCTCTGGACACTGGTTTGCAGTAGCTGGAGTGGGTTTCTTTAAAG
 GGAGATTTGCGGAGATTACAGTAGGGAGGAATGTCCTGGTCTGAGTGGTGTACAGATCAAATTCTCTGG
 TCAGGAATAGATTCTCCAAACTTGCTGGAAATCCAGAAGAGAATTTCATGCAGCTGGCAGGCGCTCT
 GCAAAAACAAACAGACTTGACTGCATCTCAGAGTTATATGTGTAAAATCAATAGGATTTGGAGCGGTC
 CATTTGAGCCAAACTTTTGTCAATTTTTATTTCCACAGCTCAGTTCAGGACTGCACGGCGACTGAAAGT
 TACAGTGCAGCCTGAAATGTTATCTCTTATTTGTGCTTCTGAAAACATTCAGTGCTAATTAACATCTGCA
 TACGGAGGTTTAAGACGGGTCTTTAACTTTGCCATTTAGCTCATTTTGTGGTACTTGTCTTCTGTCTT
 CTTAAACTCAGCGGCAGATAAAACCATAAGTGGGTCAAATCTATGATTGACAGATTTATGTATTGAGA
 GGAGTCTACATTTAAGTGC GTTCTCTTGTCTGTGTCTGGCGTGATGAGTTGTGATATGATTCTGTATTT
 TGTCTGTGCATGGTTTTTTGGGAGACCACATAATATCAGATCAGAGCTTTCTTTTTTTGTGCCACTCTCAAC
 AGTCTTCCAGGTTTTTGGCTGAGCTACCAATATGAGCT

经 SGN 编辑后形成的突变 *cyp26b1* 的序列

#7-1、#7-2、#7-3（三个相同的突变序列）

ATTCATAGAAACGCGCTTCACAGATCGAAATATTTGGCCATATCAATAATTTTGGCAAGGACTA
 TCTGATATAATTAAGAAAACACTGCCAGTTTATGAGGGATTTATTTACATTATTTTATAGAAATC
 GACTTATAACATCAACCCCCGAAAGTTGTAAATATAAACGCAAAACAAATATGAAACGCGTAA
 GTCAATGAATGAATCGCTTACTCATGCAGTGTTAGAGGACTGCTGCTGCTGCTGCTGCTGCT
 GCTGCTGTGTGCAACAGATGCGCATCTGAACAATCAAAACATTTTAAAAAGTCATGAACATA
 TCGATTTAGCATAAGAGTCCGTAGCTTATATGATCCGGTAAAATGTAGAGTATTATTTACATTT
 TTGTTGTGATGGTTTTAAACCATCATTTGGGTAATCAGTGACCTCAGTTTGAGCGCATTTTAA
 GTCCACAACATGAATTAACCCTGTGCGTATTCCGCGAGTGCCAGGTGACAAGTTCCCGACCGT
 AAATGAGATTTGAAATATTGAAAGTGTTCGCCCTCTCCCTCTCTTTCTGTTTGGGCTC---2610b
 p--TCTTCCAGGTTTTTGGCTGAGCTACCAATATGAGCT

实施例 7 SGN 在细胞系对人类基因组的编辑

- 为测试 SGN 是否可对人基因组进行编辑，申请人利用完全培养基（包含高糖 DMEM（Hyclone）、10% FBS（Hyclone）和各 100 U/mL 的青霉素和链霉素（Hyclone））培养人 293T 细胞。将细胞均匀地培养在 6 孔板中，每孔 2 mL 完全培养基。待细胞密度长至 90%左右时进行细胞转染。首先，将 1 μ g 人密码子优化的 SGN 表达载体（pW1-SGN）、各 500 ng 的识别人 *MEIS2* 基因 gDNA（gDNA-MEIS2-F、gDNA-MEIS2-R，间隔 32bp 序列，表 1），与 250 μ L Opti-MEM（Gibco）培养基混匀室温放置 5 分钟。同时将 10 μ L Lipo2000（Invitrogen）与 250 μ L Opti-MEM（Gibco）培养基混匀室温放置 5 分钟。将转染载体与转染试剂共 500 μ L 混匀室温放置 15 分钟后，均匀添加到 1 个 6 孔板

的孔中。转染后 6 小时，换液一次。转染后 48 小时，收集细胞，用与前述制备斑马鱼胚胎基因组 DNA 模板相同的试剂盒（南京尧顺禹，中国）制备细胞基因组 DNA 模板，使用引物对（MEIS2-F、MEIS2-R，表 1）针对目标基因进行 PCR 扩增。PCR 反应条件为 95°C 2 分钟，30×（95°C 1 分钟，58°C 1 分钟，72°C 45 秒），最后 72°C 延伸 10 分钟。扩增完成后，按前述相同方法将扩增产物重组入 pGEM-T 载体（Promega，美国），转化产物涂板后第二天，每个平板挑选 96 个转化子，按前述同样方法进行 PCR 扩增，以鉴定插入产物是否含被编辑的基因组片段。将 PCR 产物条带大小明显小于野生型基因组片段的转化子送商业公司测序，结果有 1/96 的转化子发生了大片段缺失（表 6）。

表 6 人 *MEIS2* 野生序列和经 SGN 编辑后形成的突变等位基因序列

<i>MEIS2</i> 野生型部分序列
TGTTTCTGGCTTAGGCTCGCCTGATGGAGGAAGGGTTGTAAAGTGTGGTGCGTGCCTGTTGTATTAA AGAAGGACTCCACACCCTCCAACTTGTTGATTACAGTTCTTGCTCCTTCCCCTCTGCACCCCGTCTCT GTGTGTGTGTCTCTGCGTGGCGCTGCCCCTTAGGCACCCGTTGTTTCCTCTGTTAGCTCTGGTCTTTGA GAAGTGCGAGCTGGCGACCTGCACTCCCCGGGAACCTGGAGTGGCTGGCGGAGACGTCTGCTCCTCC GACTCCTTCAACGAGGACATCGCGGTCTTCGCCAAGCAGGTCCAGCCTTCTTGATCCACTCAACTTCG GGCAGTCCCTCACTACCCCTCCCCTTTCTCGGAGCTTTACTTGAAAGGAGGAAAGGCCAGGGACACC CTATCTGTTCTCTGGCCCAGTCCGCTCTCTACGCCTCTCCCCATCCCTGAGCCTGGGGGCGAATTGG AAGTGGAGATAAATATGGGTGTTTGTTCCTCCCCCAGCAGGAATACAGGGGTCGACCCTCGAATTAA AGTTGACCTGCTCCTAATCTGGGGCTTGGCCTTTTCCCCCGCCTCCAGCTACTTGAACAAGAATTGGG GAGAGGGGCTGTGAGGGACCCGGGAGTGCCCGGGAGGCAAGGAAGCTGGGCGCCTCTCCGAGCGGT CCCAGGCAGCAGTGGAGAGCCACAGCCTCTTGGACTGCTGGATATGGAAGTTCGAGGCCCTGAGAACG ACCCTGAGTTCCTGGCTACAGCTTAGTCGTACCAGGTCCCCCTTAATGGGGCCAGGACCAGAATGCCC TGATTCTTCTTCAGGGACTCCACTAATTTAAGGCACCAGTTTCTTCTTGTATTTTGCCAAAGCCCCCTT CCTCTTCTCCGCCATTGCCCTCCACCGTACCCATTCTCACATTTAGCTTTCTTATTTTAAATGGGGTGAT CGTTCAAGTTGACTCTCTGTTTCTTACAGGTTTCGCGCCGAAAAAGCCACTTTTTCTCAAATCCAGAGC TGGACAATTGGTAAGGCTCTCTGTTCTTGTTCCTGTTTTTCCCCAGCCTTTGCCCTCTCCCATGCCCAGAAGC ACCCACTTTCTGCT
经 SGN 编辑后形成的突变 <i>MEIS2</i> 的部分序列
TGTTTCTGGCTTAGGCTCGCCTGATGGAGGAAGGGTTGTAAAGTGTGGTGCGTGCCTGTTGTATTAA AGAAGGACTCCACACCCTCCAACTTGTTGATTACAGTTCTTGCTCCTTCCCCTCTGCACCCCGTCTCT GTGTGTGTGTCTCTGCGTGGCGCTGCCCCTTAGGCACCCGTTGTTTCCTCTGTTAGCTCTGGTCTTTGA GAAGTGCGAGCTGGCGACCTGCACTCCCCGGGAACCTGGAGTGGCTGGCGGAGACGTCTGCTCCTCC GACTCCTTCAACGAGGACATCGCGGTCTTCGCCAAGCAGGTCCAGCCTTCTTGATCCACTCAACTTCG GGCAGTCCCTCACTACCCCTCCCCTTTCTCGGAGCTTTACTTGAAAGGAGGAAAGGCCAGGGAC-----

```

-255bp-----TGGGCGCCTCTCCGAGCGGTCCCAGGCAGCAGTGGAGAGCCACAGCCTCTTGGACTGCTG
GATATGGAACCTCGAGGCCCTGAGAACGACCCTGAGTTCCTGGCTACAGCTTAGTCGTACCAGGTCCCC
CTTAATGGGGCCAGGACCAGAATGCCCTGATTCTTCTTCAGGGACTCCACTAATTTAAGGCACCAAGTTT
CTTCTTGTATTTTGCCAAAGCCCCCTTCTCTTCTCCGCCATTGCCCTCCACCGTACCCATTCTCACATT
TCAGCTTTCTTATTTTAAATGGGGTGATCGTTCAAGTTGACTCTCTGTTTCTTACAGGTTTCGCGCCGAA
AAGCCACTTTTCTCTCAAATCCAGAGCTGGACAATTTGGTAAGGCTCTCTGTTTCCTTGTTTTTCCCCA
GCCTTTGCCCTCTCCCATGCCCAGAAGCACCCACTTTCTGCT

```

实施例 8 SGN 在体内基因组编辑中的机制

已知通过 ZFN、TALEN 和 RGEN 创建的主要突变是引入小的缺失，但 SGN 却产生的是大片段的缺失。这里，申请人尝试说明这些大片段的缺失所依据的机制。如图 13 所示，1) 一个 gDNA 杂交到斑马鱼或人的基因组 DNA 的单链，以形成 3'-Flap 结构；2) 与单链人工靶 DNA 的切割机理相同，SGN 结合至识别位点后切割斑马鱼基因组 DNA 的单链；3) 斑马鱼或人基因组 DNA 单链被切割后形成有缺刻的结构；4) 据报道 (Hosfield DJ, et al. J Biol Chem. 1998, 273:27154-27161.)，FEN-1 识别缺刻结构。因为 SGN 由 FEN-1 和 Fn1 组成，申请人认为，SGN 可以识别一个带切刻的结构；5) 一种 SGN 分子结合至该带切刻的结构并切割的基因组 DNA 的单链。切割的产物还具有带切刻的结构。这意味着一旦切割开始时，它会一直重复；和 6) 被破坏的基因组 DNA 在体内通过 DNA 修复途径修复从而形成突变分子。

为了验证提出的机理的可能性，申请人测试了缺刻 dsDNA 作为底物的反应。首先，将 50 μ L 含有 dsDNA (1 μ g)、10 $\times$ NEB 缓冲液 3 (5 μ L) 和 1 μ L Nt.BstNBI (New England Biolabs, 英国) 的混合物在 37 $^{\circ}$ C 孵育 3 小时。之后，将 10 μ L 含有纯化的缺刻 dsDNA (50 ng, 来自第一步骤)、MOPS (10 mM)、0.05% Tween-20、0.01% nonidet P-40、MgCl₂ (7.5 mM) 和 SGN (1 ng) 的混合物在 37 $^{\circ}$ C 孵育 2 小时。最后，从第二步骤获得的切割产物用 2% 琼脂糖凝胶电泳进行分析。

用缺刻核酸内切酶消化 400bp 长的 dsDNA 形成单个链上均有切刻 (图 14a)，其中缺刻核酸内切酶使用的是 Nt.BstNBI，该酶的识别为点是“GAGTC”并切割后续 4 个碱基。如图 14b 的泳道 2 中所示，缺刻

dsDNA 和 SGN 共孵育的反应中，可以观察到小分子量的产物条带，相反，包含缺刻 dsDNA 但不包含 SGN 的反应（泳道 1）中，没有观察到该结果。相似地，包含 dsDNA 但不包含 SGN 的反应（泳道 3）中，包含 dsDNA 但包含 SGN 的反应（泳道 4）中，均没有观察到小分子量的产物条带。

这些提供了一些初步的证明以支持我们猜测的大片段缺失机制。但是，更深入的机制需要一些可能的结构修饰以提高 SGN 系统的精度和效率来进一步揭示。

讨论：

在这里，我们设计了一种结构导向核酸酶，在靶 DNA 和 gDNA 之间形成的 3'-Flap 结构的基础上，该酶可以识别靶 DNA，并通过 Fnl 二聚体切割靶基因。目前可用的核酸内切酶表现出对序列的偏好。而本申请中，采用结构导向方式识别和捕获靶 DNA，可以设计出满足任何靶 DNA 适合的 gDNAs，并使用 SGN 切割靶 DNA。在该系统中，可以较为容易地设计和合成 gDNA 并调整其浓度，而在 RGEN 系统中设计 gRNA 相对较为困难。日前，关于 NgAgo 基因组编辑系统一项最新研究（Gao F, et al. Nat Biotechnol 2016.）也说明 DNA 引导的优势。此外，区别于 ZFN、TALEN、RGEN 或新报道的 NgAgo 基因组编辑系统产生的突变是通过小插入缺失造成的，SGN 系统可以在斑马鱼基因组和人基因组中产生大片段缺失，因为在突变基因中一个小的插入缺失仍然能够编码具有一些残留功能的截短蛋白质，由 ZFN、TALEN、RGEN 或 NgAgo 诱导的小片段插入缺失突变等位基因并不总是无效的等位基因或敲除的等位基因。与此相反，由 SGN 创建的大片段缺失突变的等位基因更可能产生无效的等位基因。

SGN 的基因组 DNA 编辑应用中还有些问题需要进一步研究得到解决。目前，SGN 系统的效率不高，因此，我们没有从 *cyp26b1* 和 *znf703* 基因实验中观察到脱靶的情况。但根据 SGN 的工作机制，如果引导 DNA（gDNA）与基因组 DNA 产生错误杂交，那会有脱靶现象。幸运的是，相比 NgAgo 中引导 DNA 只能是 23 nt、24 nt、或 25 nt 的情况，在 SGN 系统中的 gDNAs 长度是可以调整的，以避免错误杂交。其他潜在的问题

是 FEN-1 可以针对某些天然存在的 DNA 的结构，它或许不仅会导致脱靶效应，而且还有细胞毒性。但事实上，申请人并没有在斑马鱼实验中观察到特定的细胞死亡现象。此外，与典型双链相比，产生 dsDNAs 中特定结构的挑战也是不容忽视的。目前本发明中方法的效率受限于

5 gDNA 和基因组中正确位点之间的相互作用，形成一个 3'-Flap 结构。其他策略，例如使用 PNA 或 LNA 探针作为引导 DNA 和可能的 SGN 结构修饰，应该可被用来改善和提高 SGN 的切割效率。

结论：

总之，在本发明中，申请人构建了由识别 3'-Flap 结构的 FEN-1 和切割 DNA 链的 *Fok* I (Fn1) 切割结构域所组成的结构指导的内切酶

10 (Structure-guided endonuclease, SGN)。基于靶序列和人工引导 DNA (guide DNA, gDNA) 形成的 3'-Flap 结构，SGN 可以识别靶 DNA，并通过其 Fn1 二聚体切割靶点。实验结果表明，SGN 可体外切割靶 DNA。另外使用斑马鱼胚胎和人细胞系作为孵育体系，本发明证明了 SGN 可编辑内源基因。通过结构引导识别，任何期望的靶 DNA 都可被构建的核

15 酸内切酶 SGN 所切割，而不需要改变核酸内切酶或其多肽单元（如 II 型 REases、ZFN 和 TALEN），或不受 RNA 分子使用的限制（如 RGEN 中）。它可以成为 DNA 编辑有用的替代工具。

20 以上，基于本发明的实施方式进行了说明，但本发明不限于此，本领域的技术人员应该明白，在本发明的主旨的范围内能够以进行变形和变更的方式实施，这样的变形和变更的方式，理应属于本发明的保护范围。

权利要求书

1. 一种靶多核苷酸编辑方法，其特征在于：设计一对寡核苷酸探针，使其与靶多核苷酸的正义链和反义链分别结合，分别产生能够被核酸酶识别的目标序列结构，所述核酸酶切割靶多核苷酸，实现对靶多核苷酸的编辑。
2. 如权利要求 1 所述的方法，其中，所述靶多核苷酸是 RNA 或 DNA，优选基因组 DNA，优选所述的基因组是斑马鱼基因组、哺乳动物基因组、人类基因组、或植物基因组。
3. 如权利要求 1 或 2 所述的方法，其中，所述一对寡核苷酸探针是 DNA，优选所述一对寡核苷酸探针与靶多核苷酸的结合位点间隔 0-100 bp，优选间隔为 10-70 bp、20-60 bp、32-50 bp 或 40 bp。
4. 如权利要求 1 至 3 任一项所述的方法，其中，所述寡核苷酸探针的长度为 20 nt 以上，优选长度为 20-50 nt 或 25 nt。
5. 如权利要求 1 至 4 任一项所述的方法，其中，所述的目标序列结构为 5'突出核酸结构、3'突出核酸结构、切刻核酸结构、侵入核酸结构、三链核酸结构、Y 型核酸结构、3'-Flap 结构以及由错配或缺失核酸引起的鼓泡核酸结构中的至少一种的识别功能域，优选 3'-Flap 结构，DNA 寡核苷酸单链的 3'-末端不与目标基因组 DNA 互补。
6. 如权利要求 5 所述的方法，其中，所述寡核苷酸探针的 5'端对靶多核苷酸特异，优选与靶多核苷酸互补；寡核苷酸探针的 3'端不与靶多核苷酸互补，优选 3'末端 1 个以上、1-20 个、1-10 个、1-5 个、1-4 个、1-3 个、2 个或 1 个碱基不与靶多核苷酸互补。
7. 如权利要求 1 至 6 任一项所述的方法，其中，所述的核酸酶是重组结构识别核酸内切酶（SGN），包含结构识别功能域、DNA 切割功能域和连接二者的肽段；所述的结构识别功能域为能够识别结构为 5'突出核酸结构、3'突出核酸结构、切刻核酸结构、侵入核酸结构、三链核酸结构、Y 型核酸结构、3'-Flap 结构以及由错配或缺失核酸引起的鼓泡核酸结构中的至少一种的识别功能域，优选选自 TaqPol、TthPol、TaqExo、AfuFEN、PfuFEN、MjaFEN、MthFEN、E.coliMuts、Tthmuts 和 Taqmuts 组成的组

中任意一种酶的识别功能域或全酶片段；所述的切割功能域为 IIS 型核酸内切酶的切割功能域，优选 Fok I 的部分或全部肽段；所述的切割功能域还可选自 TaqPol、TthPol、TaqExo、AfuFEN、PfuFEN、MjaFEN、MthFEN、E.coliMuts、Tthmuts 和 Taqmuts 组成的组中任意一种酶的全酶片段的核酸内切酶结构域；所述的连接肽段为不影响结构识别与酶切功能的柔性肽段，优选甘氨酸或丝氨酸或其组合的串联组合。

8. 如权利要求 7 所述的方法，其中，所述的重组结构识别核酸内切酶含有核定位信号。

9. 如权利要求 7 或 8 所述的方法，其中，所述的重组结构识别核酸内切酶的氨基酸序列选自 SEQ ID No:1、SEQ ID No:2 中任一个；所述重组结构识别核酸内切酶的核酸序列选自 SEQ ID No:3、SEQ ID No:4、SEQ ID No:5 中任一个。

10. 权利要求 1 至 9 任一项所述的方法，其中，所述基因组 DNA 是内源 DNA 或整合到基因组的外源 DNA，优选所述靶多核苷酸编辑方法是在体实施的。

11. 一种寡核苷酸探针对，其特征在于该寡核苷酸探针对与靶多核苷酸的正义和反义链分别结合，分别产生能够被核酸酶识别的目标序列结构，所述目标序列结构能够被核酸酶识别，所述核酸酶切割靶多核苷酸，实现对靶多核苷酸的编辑。

12. 如权利要求 11 所述的一种寡核苷酸探针对，其中，所述一对寡核苷酸探针对是 DNA，优选与靶多核苷酸的结合位点间隔 0-100 bp，优选间隔为 10-70 bp、20-60 bp、32-50 bp 或 40 bp。

13. 如权利要求 11 或 12 所述的一种寡核苷酸探针对，其中，所述寡核苷酸探针的长度为 20 nt 以上，优选长度为 20-50 nt 或 25 nt。

14. 如权利要求 1-13 任一项所述的一种寡核苷酸探针对，其中，所述寡核苷酸探针的 5'端对靶多核苷酸特异，优选与靶多核苷酸互补；寡核苷酸探针的 3'端不与靶多核苷酸互补，优选 3'末端 1 个以上、1-20 个、1-10 个、1-5 个、1-4 个、1-3 个、2 个或 1 个碱基不与靶多核苷酸互补。

15. 用于靶多核苷酸编辑的系统，其特征在于所述系统包括权利要求 1-14 任一项中所述的寡核苷酸探针对，优选包括权利要求 1-10 任一项中

所述的核酸酶。

16. 用于靶多核苷酸编辑的试剂盒，其特征在于所述试剂盒包括权利要求 1-14 任一项中所述的寡核苷酸探针对，优选包括权利要求 1-10 任一项中所述的核酸酶。

5 17. 一种在靶多核苷酸中切割大片段的方法，其特征在于采用了权利要求 1-10 任一项的靶多核苷酸编辑方法。

18. 如权利要求 17 所述的方法，所述的大片段长度大于所述寡核苷酸探针对与靶多核苷酸的结合位点的间隔，优选所述的大片段长度大于所述间隔 1-10000 bp、1-3000 bp、1-2000 bp、1-1000 bp、1-500 bp、1-200 bp
10 或 1-100 bp。

19. 一种疾病细胞模型的制作方法，其特征在于采用了权利要求 1-10、17、18 任一项的方法或施用了权利要求 11-14 任一项所述的寡核苷酸探针对或权利要求 15 所述的用于靶多核苷酸编辑的系统或权利要求 16 所述的用于靶多核苷酸编辑的试剂盒。

15 20. 一种疾病动物模型的制作方法，其特征在于采用了权利要求 1-10、17、18 任一项的方法或施用了权利要求 11-14 任一项所述的寡核苷酸探针对或权利要求 15 所述的用于靶多核苷酸编辑的系统或权利要求 16 所述的用于靶多核苷酸编辑的试剂盒。

21. 一种疾病治疗方法，其特征在于采用了权利要求 1-10、17、18
20 任一项的方法或施用了权利要求 11-14 任一项所述的寡核苷酸探针对或权利要求 15 所述的用于靶多核苷酸编辑的系统或权利要求 16 所述的用于靶多核苷酸编辑的试剂盒。

22. 如权利要求 21 所述的疾病治疗方法，其中所述的疾病选自遗传相关疾病或非遗传相关疾病，优选所述的遗传相关疾病选自癌症、自身免疫疾病、糖尿病、血液病、心脏病、抑郁症、阿尔茨海默病、哮喘和神经疾病中的任一种。
25

23. 一种植物突变体的制作方法，其特征在于采用了权利要求 1-10、17、18 任一项的方法或施用了权利要求 11-14 任一项所述的寡核苷酸探针对或权利要求 15 所述的用于靶多核苷酸编辑的系统或权利要求 16 所
30 述的用于靶多核苷酸编辑的试剂盒。

图 1

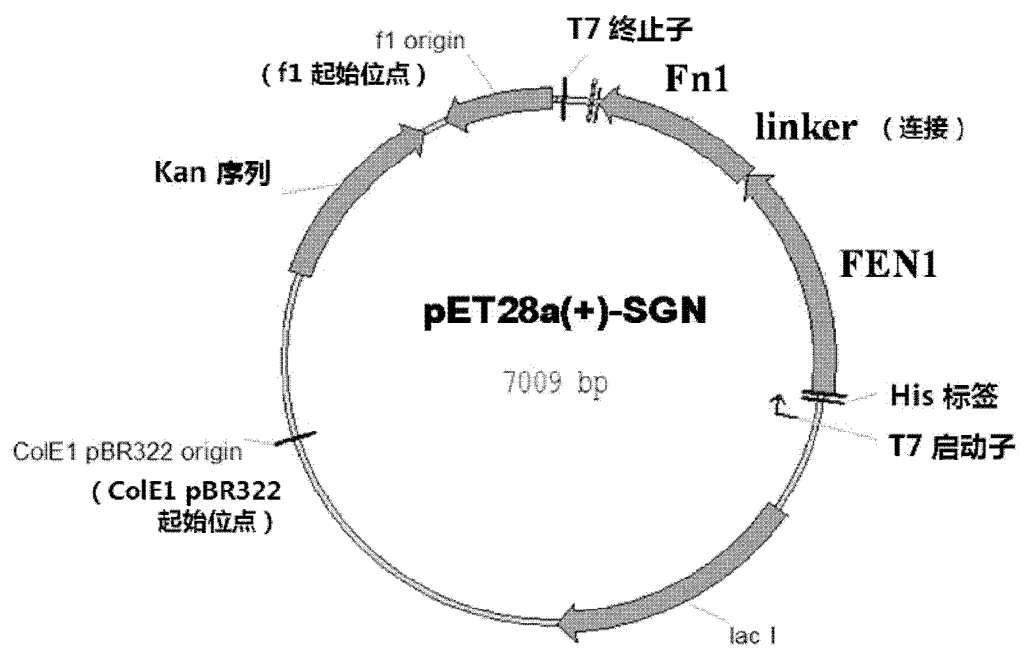


图 2

斑马鱼源密码子优化的 SGN（含核定位信号）CDS 序列

ATGGGATCTAGCCACCACCACCACCATCACTCTAGCGGACTGGTGCCTAGAGGAAGCCACATGGCTCCT
AAGAAGAAGAGAAAGGTGGGAATCCACGGAGTGCCTGCTGCTGGAGCTGATATCGGAGACCTGTTTCGAG
AGAGAGGAGGTGGAGCTGGAGTACTTCAGCGGAAAGAAGATCGCTGTGGACGCTTTCAACACACTGTAC
CAGTTCATCAGCATCATCAGACAGCCTGACGGAAACACCTCTGAAGGACAGCCAGGGAAGAATCACAAGC
CACCTGAGCGGAATCCTGTACAGAGTGAGCAACATGGTGGAGGTGGGAATCAGACCTGTGTTCTGTGTTT
GACGGAGAGCCTCCTGAGTTCAAGAAGGCTGAGATCGAGGAGAGAAAGAAGAGAAGAGCTGAGGCTGAG
GAGATGTGGATCGCTGCTCTGCAGGCTGGAGACAAGGACGCTAAGAAGTACGCTCAGGCTGCTGGAAGA
GTGGACGAGTACATCGTGGACAGCGCTAAGACACTGCTGAGCTACATGGGAATCCCTTTCTGTGGACGCT
CCTTCTGAGGGAGAAGCTCAGGCTGCTTACATGGCTGCTAAGGGAGACGTGGAGTACACAGGAAGCCAG
GACTACGACAGCCTGCTGTTTCGGAAGCCCTAGACTGGCTAGAAACCTGGCTATCACAGGAAAGAGAAAG
CTGCCTGGAAAGAACGTGTACGTGGACGTGAAGCCTGAGATCATCATCCTGGAGAGCAACCTGAAGAGA
CTGGGACTGACAAGAGAGCAGCTGATCGACATCGCTATCCTGGTGGGAACAGACTACAACGAGGGAGTG
AAGGGAGTGGGAGTGAAGAAGGCTCTGAACTACATCAAGACATACGGAGACATCTTCAGAGCTCTGAAG
GCTCTGAAGGTGAACATCGACCACGTGGAGGAGATCAGAACTTCTTCTGAACCTCCTGTGACAGAC
GACTACAGAATCGAGTTCAGAGAGCCTGACTTCGAGAAGGCTATCGAGTTCCTGTGTGAGGAGCACGAC
TTCAGCAGAGAGAGAGTGGAGAAGGCTCTGGAGAAGCTGAAGGCTCTGAAGAGCACACAGGCTACACTG
GAGAGATGGTTCGAGTTCGGAGGAGGAGGATCTGGAGGAGGAGGATCTGGAGGAGGAGGATCTCAGCTG
GTGAAGTCTGAGCTGGAGGAGAAGAAGAGCGAGCTGAGACACAAGCTGAAGTACGTGCCTCACGAGTAC
ATCGAGCTGATCGAGATCGCTAGAAACAGCACACAGGACAGAACTCCTGGAGATGAAGGTGATGGAGTTC
TTCATGAAGGTGTACGGATACAGAGGAAAGCACCTGGGAGGAAGCAGAAAGCCTGACGGAGCTATCTAC
ACAGTGGGAAGCCCTATCGACTACGGAGTGATCGTGGACACAAAGGCTTACAGCGGAGGATACAACCTG
CCTATCGGACAGGCTGACGAGATGCAGAGATACGTGGAGGAGAACCAGACAAGAAACAAGCACATCAAC
CCTAACGAGTGGTGGAAAGGTGTACCCTAGCAGCGTGACAGAGTTCAGTTCCTGTTCTGTGAGCGGACAC
TTCAAGGGAAACTACAAGGCTCAGCTGACAAGACTGAACCACATCACAACTGTAACGGAGCTGTGCTG
AGCGTGGAGGAACTGCTGATCGGAGGAGAGATGATCAAGGCTGGAACACTGACACTGGAGGAGGTGAGA
AGAAAGTTCAACAACGGAGAGATCAACTTCAAGCTGAAGAGACCTGCTGCTACAAAGAAGGCTGGACAG
GCTAAGAAGAAGAAGTAA

SEQ ID No:2

SGN 氨基酸序列（含核定位信号）

MGSSHHHHHSSGLVPRGSHMAPKKRKVGIHGVPAAADIGDLFEREEVELEYFSGKKIAVDAFNTLY
QFISIIIRQPDGTPLKDSQGRITSHLSGILYRVSNMVEVGIRPVFVFDGEPPEFKKAEIEERKKRAEAE
EMWIAALQAGDKDAKKYAQAAGRVDEYIVDSAKTLLSYMGIPFVDAPSEGEAQAAAYMAAKGDVEYTG
DYDSLFLFGSPRLARNLAITGKRKLPGKNVYVDVKPEIIILESNLRLGLTREQLIDIAILVGTDYNEGV
KGVGVKKALNYIKTYGDI FRALKALKVNIDHVEEIRNFFLNPPVTDDYRIEFREPDFEKAIEFLCEEHD
FSRERVEKALEKLKALKSTQATLERWFEFGGGGSGGGGSGGGGSQLVKSELEKKSELRHKLKYPHEY
IELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNL
PIGQADEMQRYVEENQTRNKHINPNEWVKVPSSVTEFKFLFVSGHFKGNYKAQLTRLNHIITNCNGAVL
SVEELLIGGEMIKAGTLTLEEVRKFNNGEINFKLKRPAAATKKAGQAKKKK

图 3

人源密码子优化的 SGN（含核定位信号）CDS 序列

ATGGGAAGCAGCCACCACCATCATCATCACAGCAGCGGCCCTGGTGCCTAGAGGAAGCCATATGGCCCCC
AAGAAGAAGAGAAAGGTGGGCATTACGGAGTGCTGCTGCCGGAGCTGATATTGGGGATCTGTTTCGAG
AGAGAGGAGGTGGAGCTGGAATACTTCAGCGGCAAGAAGATTGCCGTGGACGCCCTTCAATACCCGTGTAC
CAGTTCATCAGCATCATCAGACAGCCTGACGGAACCCCCCTGAAAGATAGCCAAGGCAGAATCACCAGC
CATCTGAGCGGCATCCTGTACAGAGTGAGCAACATGGTGGAGGTGGGCATCAGACCCGTGTTCTGTGTTT
GATGGAGAGCCCCCGAATTCAAGAAAAGCCGAGATCGAGGAGAGAAAAGAAGAGAAGAGCCGAGGCCGAG
GAAATGTGGATTGCCGCTCTGCAGGCTGGAGATAAAGACGCCAAGAAGTACGCTCAGGCTGCCGGAAGA
GTGGATGAGTACATCGTGGACAGCGCCAAAACCCCTGCTGAGCTACATGGGCATCCCTTTTGTGGACGCT
CCCAGCGAAGGAGAAGCCCCAAGCTGCTTACATGGCCGCCAAGGCGATGTGGAATACACAGGCAGCCAG
GACTATGACAGCCTGCTGTTTCGGAAGCCCTAGACTGGCTAGAAACCTGGCCATCACCGGCAAAAGAAAG
CTGCCCCGCAAGAACGTGTACGTGGACGTGAAGCCCCGAGATCATCATCCTGGAGAGCAACCTGAAGAGA
CTGGGCCTGACCAGAGAGCAGCTGATCGACATCGCTATCCTGGTGGGCACCGACTACAACGAAGGAGTG
AAGGGCGTGGGAGTGAAGAAAAGCCCTGAACTACATCAAGACCTACGGCGACATCTTCAGAGCTCTGAAG
GCCCTGAAGGTGAACATCGACCACGTGGAGGAGATCAGAACTTCTTCTGAACCCCCCGTGACCGAC
GATTACAGAATCGAGTTCAGAGAGCCCGACTTCGAGAAGGCCATCGAGTTCCTGTGTGAGGAGCACGAC
TTCAGCAGAGAGAGAGTGGAGAAGGCCCTGGAGAACTGAAGGCCCTGAAGAGCACACAAGCCACCCTG
GAGAGATGGTTCGAGTTTGGCGGAGGAGGAAGCGGAGGAGGAGGAAGCGGAGGAGGAGGAAGCCAGCTG
GTGAAAAGCGAGCTGGAGGAGAAGAAGAGCGAGCTGAGACACAAGCTGAAGTACGTGCCCCACGAGTAC
ATCGAGCTGATCGAGATCGCCAGAAACAGCACCCAGGACAGAATCCTGGAGATGAAGGTGATGGAGTTC
TTCATGAAGGTGTACGGCTACAGAGGCAAACATCTGGGCGGCAGCAGAAAACCTGATGGCGCCATCTAT
ACCGTGGGAAGCCCTATTGACTACGGCGTGATCGTGGACACCAAAGCTTACAGCGGCGGCTATAACCTG
CCTATCGGACAGGCTGACGAGATGCAGAGATACGTGGAGGAGAACCAGACCAGAAACAAGCACATCAAC
CCCAACGAGTGGTGAAGGTGTATCCCAGCAGCGTGACCGAGTTCAGTTCCTGTTCGTGAGCGGCCAC
TTCAAGGGCAACTACAAGGCCAGCTGACCAGACTGAACCACATCACCAACTGCAACGGAGCTGTGCTG
AGCGTGGAGGAAGTCTGATTGGCGGCGAGATGATTAAAGCCGGCACCCCTGACACTGGAGGAGGTGAGA
AGAAAGTTCAACAACGGCGAGATCAACTTCAAGCTGAAGAGACCCGCCGCTACAAAGAAAGCTGGCCAG
GCCAAGAAAAAGAAGTAA

SEQ ID No:2

SGN 氨基酸序列（含核定位信号）

MGSSHHHHHSSGLVPRGSHMAPKKRKVGIHGVPAAADIGDLFEREEVELEYFSGKKIAVDAFNTLY
QFISIIIRQPDGTPLKDSQGRITSHLSGILYRVSNMVEVGIRPVFVFDGEPPEFKKAEIEERKKRAEAE
EMWIAALQAGDKDAKKYAQAAGRVDEYIVDSAKTLLSYMGI PFVDAPSEGEAQAAAYMAAKGDVEYTGSG
DYDSLFLFGSPRLARNLAITGKRKLPGKNVYVDVKPEII ILESNLKRLGLTREQLIDIAILVGTDYNEGV
KGVGVKKALNYIKTYGDI FRALKALKVNIDHVEEIRNFFLNPPVTDDYRIEFREPDFEKAIEFLCEEHD
FSRERVEKALEKLKALKSTQATLERWFEFGGGGSGGGGSGGGGSQLVKSELEKKSELRHKLKYPHEY
IELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHLGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNL
PIGQADEMQRYVEENQTRNKHINPNEWVKVPSSVTEFKFLFVSGHFKGNYKAQLTRLNHIITNCNGAVL
SVEELLIGGEMIKAGTLTLEEVRKFNNGEINFKLKRPAAATKKAGQAKKKK

图 4

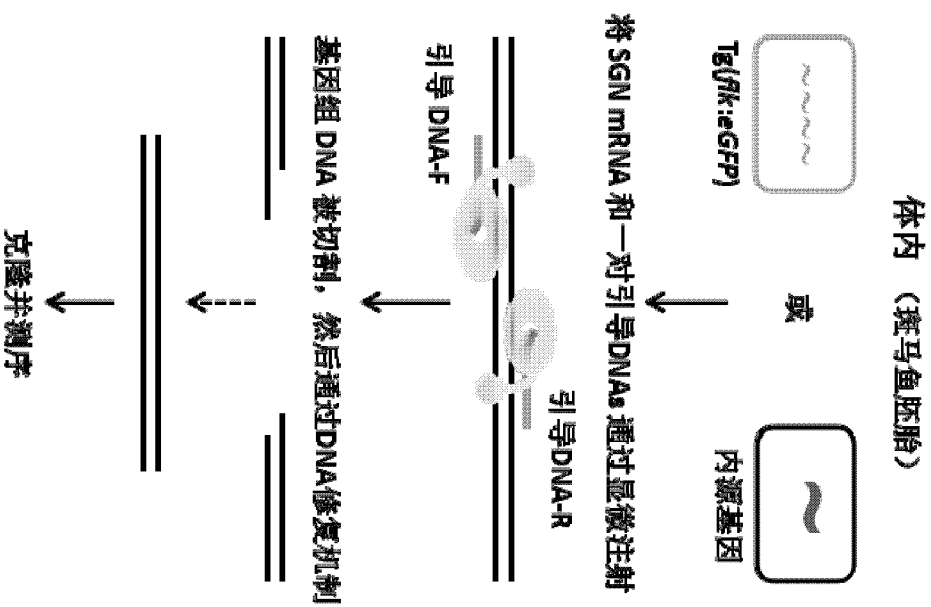
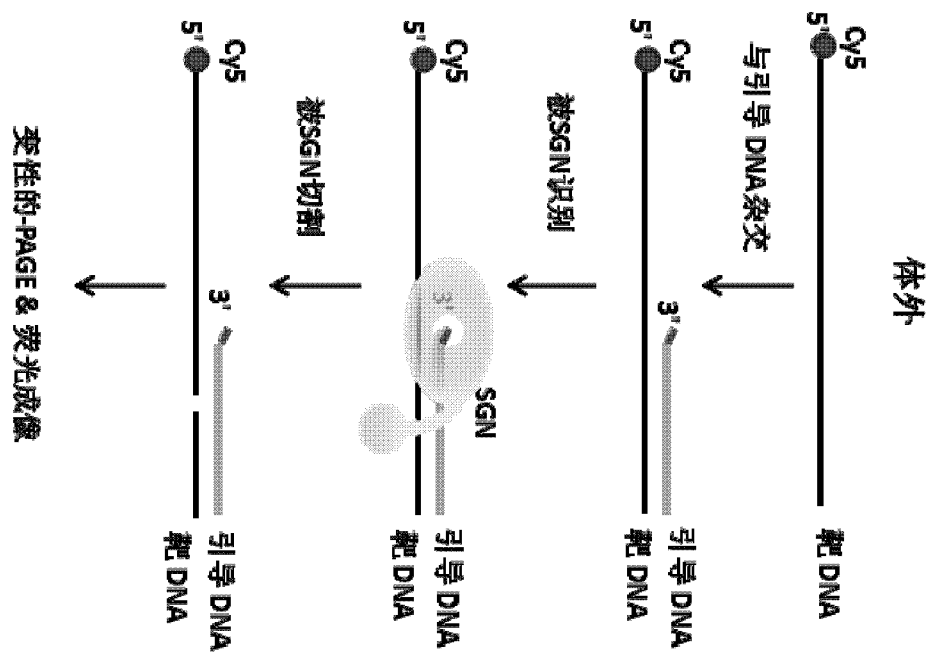


图 5

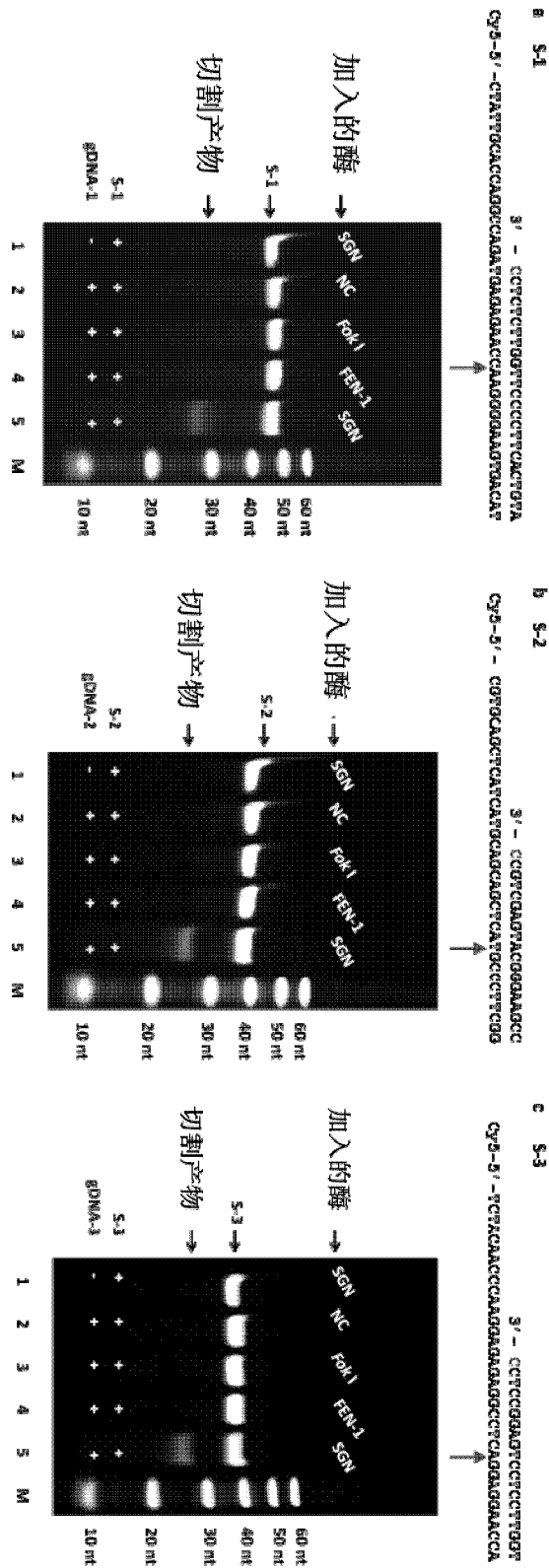


图 6

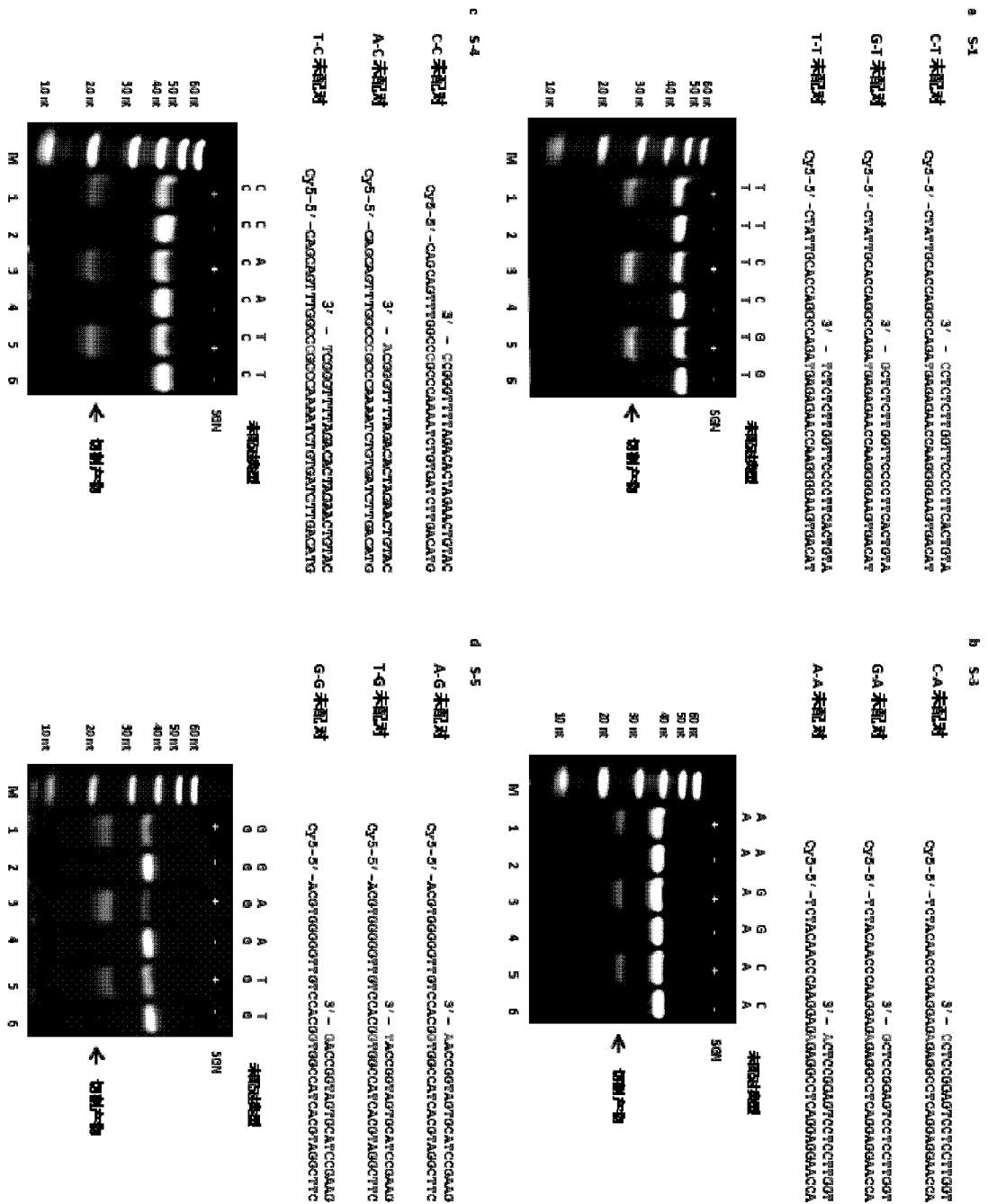


图 7



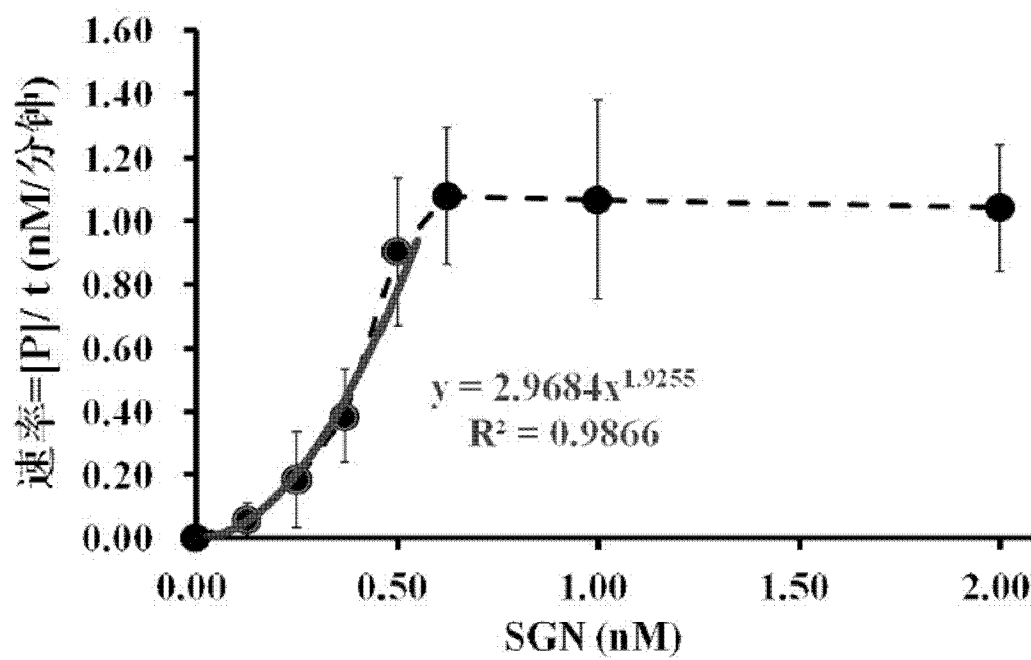


图 9

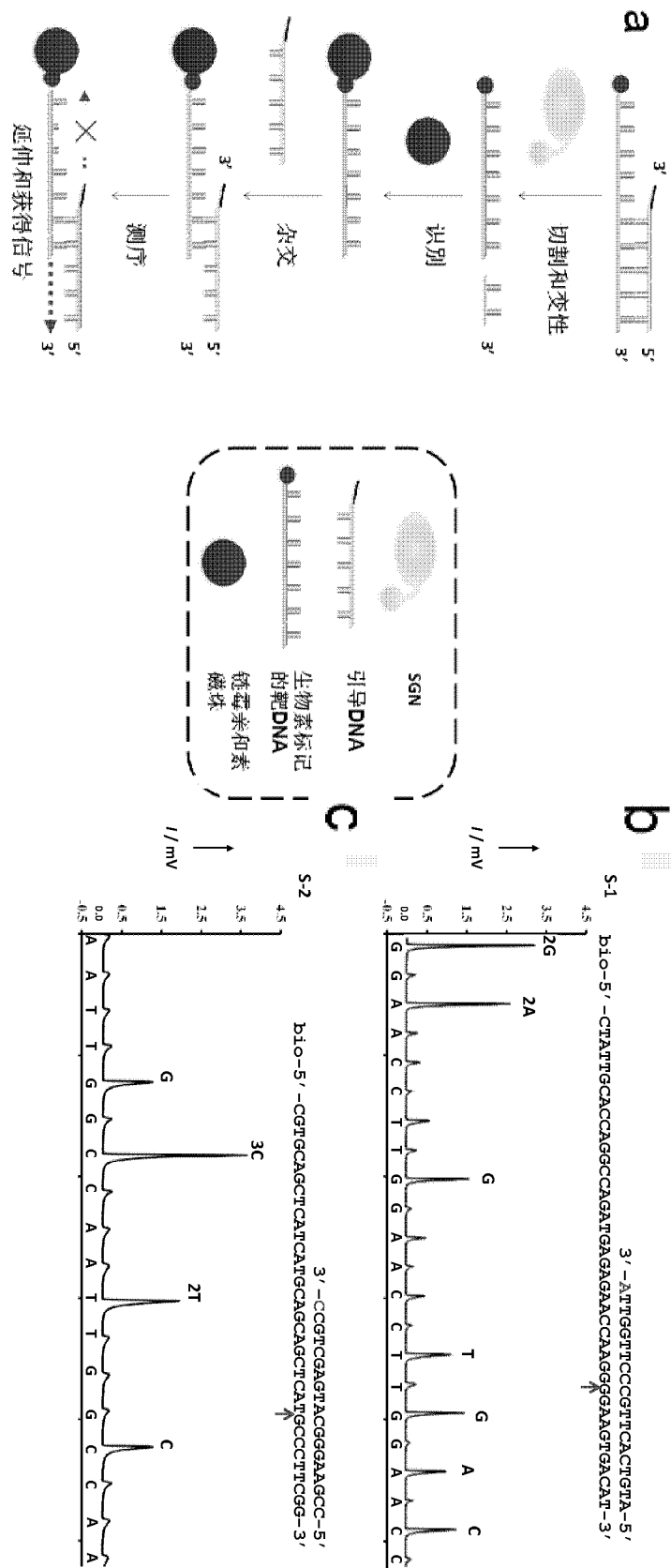


图 10

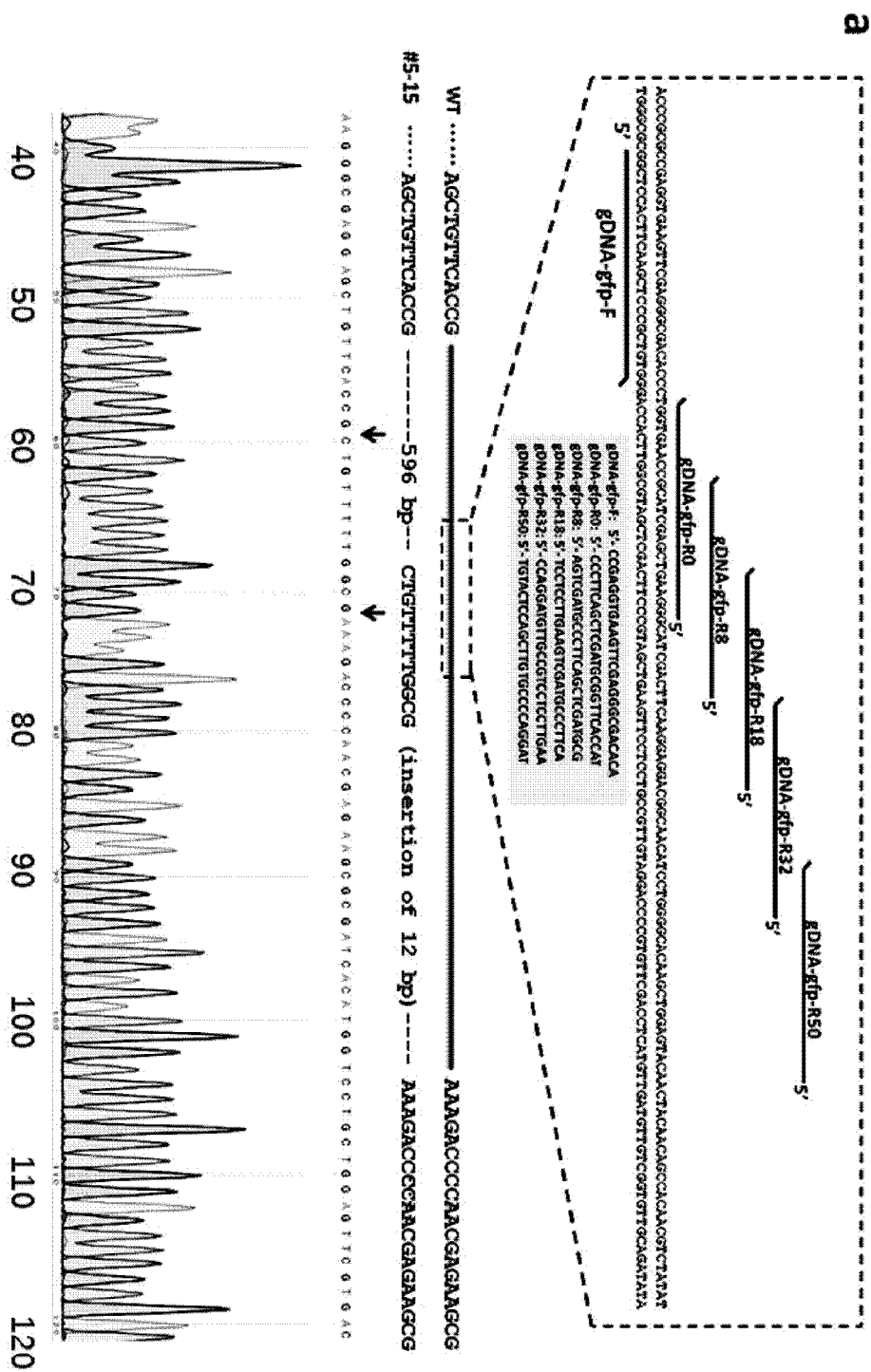
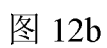


图 12a



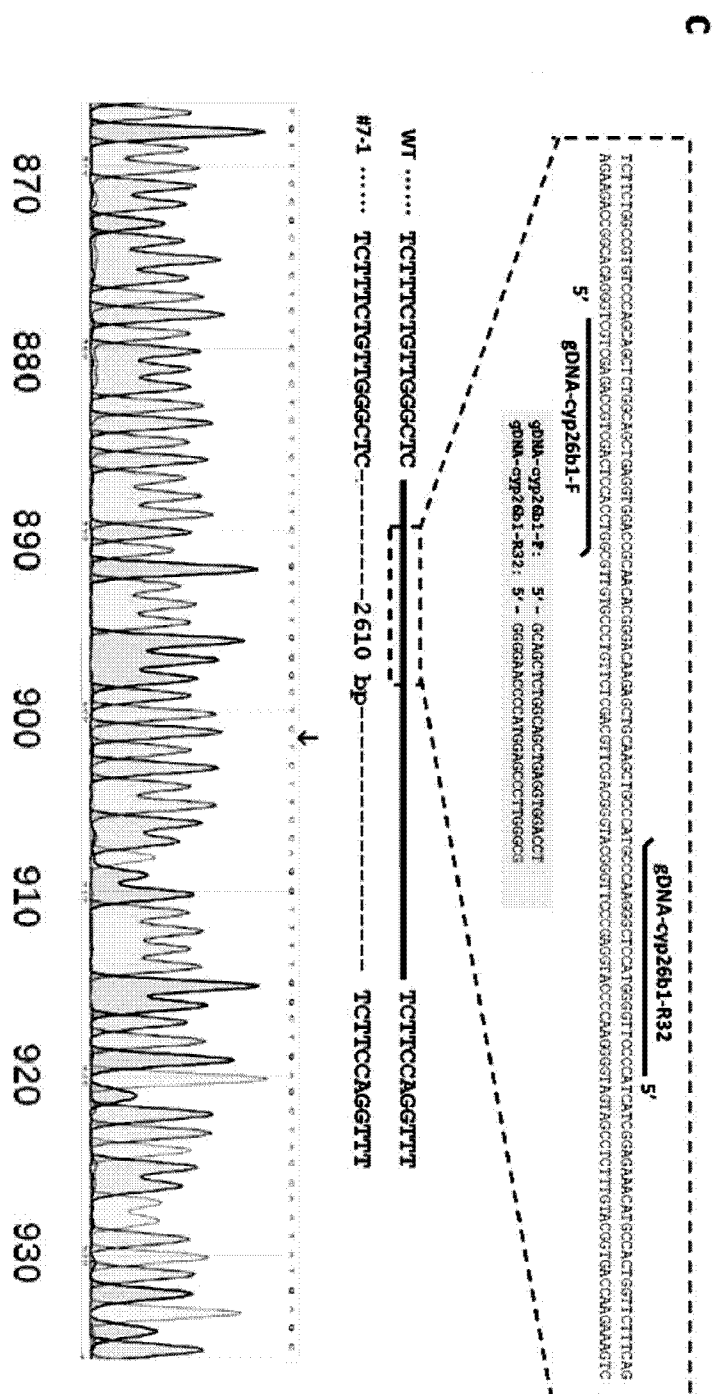


图 12c

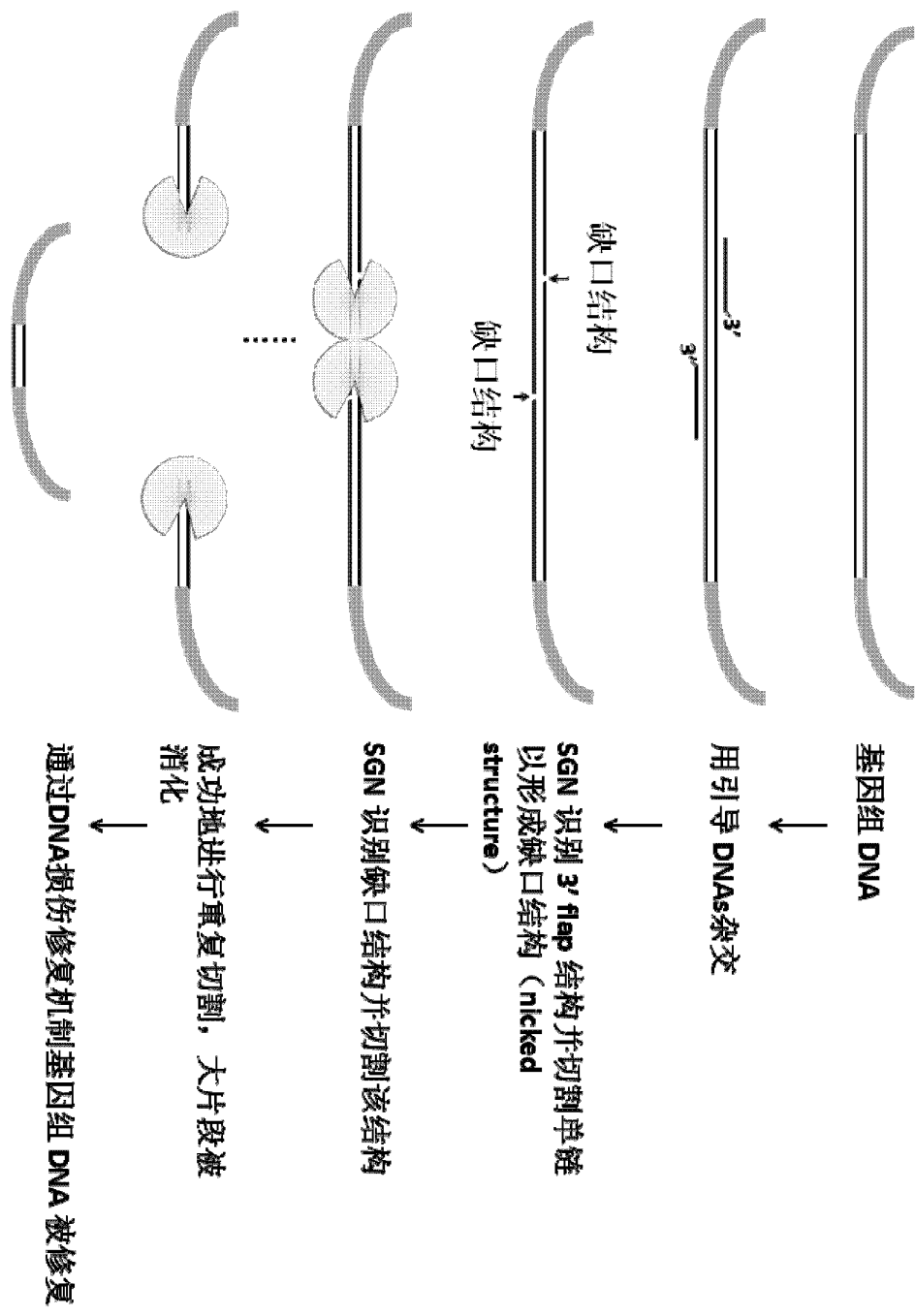


图 13

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/093903

A. CLASSIFICATION OF SUBJECT MATTER

C12N 15/113 (2010.01) i; C12N 15/63 (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 15

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)

CNABS, SIPOABS, DWPI, CNTXT, EPTXT, USTXT, WOTXT, CNKI, WANFANG, EBI, NCBI, GOOGLE, BAIDU, ISI Web of Knowledge, PUBMED: NANJING UNIVERSITY; ZHOU, Guohua; ZHAO, Qingshun; XU, Shu; CAO, Shasha; ZOU, Bingjie; YUE, Yunyun; nucleotide editing, edit, probe, primer, 3'-Flap, SGN, identification functional domain, cutting function field, FEN-1, Fok I, Fnl, Structure-guided endonuclease, Structure-guided nuclease, artificial guide DNA, guide DNA, gDNA, endonuclease, Flap endonuclease, structure-guided, gene edit+, genome edit+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	XU, S. et al. "An alternative novel tool for DNA editing without target sequence limitation: the structure-guided nuclease", GENOME BIOLOGY, vol. 17, no. 186, 16 September 2016 (16.09.2016), pages 1-12	1-23
X	CN 104745570 A (ZHOU, Guohua), 01 July 2015 (01.07.2015), claims 1-9, embodiments 4-5, and sequences 1-2	1-23
A	CN 104531633 A (LI, Yuning), 22 April 2015 (22.04.2015), the whole document	1-23
A	LI, L. et al. "Functional domains in Fok I restriction endonuclease", PROC. NATL. ACAD. SCI. USA, vol. 89, 31 May 1992 (31.05.1992), pages 4275-4279	1-23
A	CHAPADOS, B.R. et al. "Structural Basis for FEN-1 Substrate Specificity and PC-NA-Mediated Activation in DNA Replication and Repair", CELL, vol. 116, 09 January 2004 (09.01.2004), pages 39-50	1-23

☒ 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 16 October 2017 (16.10.2017)	Date of mailing of the international search report 24 October 2017 (24.10.2017)
Name and mailing address of the ISA/CN: 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 LI, Juanjuan Telephone No.: (86-10) 82246978

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/093903

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	KIM, H. et al. "A guide to genome engineering with programmable nucleases", NATURE REVIEWS, vol. 15, 02 April 2014 (02.04.2014), pages 321-334	1-23
A	HSU, P.D. et al. "Development and Applications of CRISPR-Cas9 for Genome Engineering", CELL, vol. 157, 05 June 2014 (05.06.2014), pages 1262-1277	1-23
A	HARRINGTON, J.J. et al. "The characterization of a mammalian DNA structure-specific endonuclease", THE EMBO JOURNAL, vol. 13, no. 5, 31 December 1994 (31.12.1994), pages 1235-1246	1-23

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/093903

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

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

1. ☒ Claims Nos.: 21 and 22
because they relate to subject matter not required to be searched by this Authority, namely:
[1] claims 21 and 22 relate to a method for treating diseases, and therefore do not comply with PCT Rule 39.1(iv). This report is made on the basis of amending claims 21 and 22 to be pharmaceutical use claims.
2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/CN2017/093903

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104745570 A	01 July 2015	CN 104745570 B	19 April 2017
CN 104531633 A	22 April 2015	None	

国际检索报告

国际申请号

PCT/CN2017/093903

A. 主题的分类

C12N 15/113(2010.01)i; C12N 15/63(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 15

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))

CNABS, SIPOABS, DWPI, CNTXT, EPTXT, USTXT, WOTXT, CNKI, 万方, EBI, NCBI, GOOGLE, BAIDU, ISI Web of Knowledge, PUBMED: 南京大学, 周国华, 赵庆顺, 徐澍, 曹莎莎, 邹秉杰, 岳芸芸, 核苷酸编辑, 基因编辑, 基因组编辑, 编辑, 探针, 引物, 3'-Flap, 核酸内切酶, SGN, 识别功能域, 切割功能域, FEN-1, Fok I, Fnl, Structure-guided endonuclease, Structure-guided nuclease, SGN, 人工引导DNA, guide DNA, gDNA, endonuclease, Flap endonuclease, 结构引导, 结构指导, gene edit+, genome edit+

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
PX	XU, S. 等. "An alternative novel tool for DNA editing without target sequence limitation: the structure-guided nuclease" Genome Biology, 第17卷, 第186期, 2016年 9月 16日 (2016 - 09 - 16), 第1-12页	1-23
X	CN 104745570 A (周国华) 2015年 7月 1日 (2015 - 07 - 01) 权利要求1-9、实施例4-5, 序列1-2	1-23
A	CN 104531633 A (李云英) 2015年 4月 22日 (2015 - 04 - 22) 全文	1-23
A	LI, L. 等. "Functional domains in Fok I restriction endonuclease" Proc. Natl. Acad. Sci. USA, 第89卷, 1992年 5月 31日 (1992 - 05 - 31), 第4275-4279页	1-23
A	CHAPADOS, B. R. 等. "Structural Basis for FEN-1 Substrate Specificity and PC-NA-Mediated Activation in DNA Replication and Repair" Cell, 第116卷, 2004年 1月 9日 (2004 - 01 - 09), 第39-50页	1-23

☒ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

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

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

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

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

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

“&” 同族专利的文件

国际检索实际完成的日期

2017年 10月 16日

国际检索报告邮寄日期

2017年 10月 24日

ISA/CN的名称和邮寄地址

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

传真号 (86-10) 62019451

受权官员

李娟娟

电话号码 (86-10) 82246978

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	KIM, H. 等. "A guide to genome engineering with programmable nucleases" NATURE REVIEWS, 第15卷, 2014年 4月 2日 (2014 - 04 - 02), 第321-334页	1-23
A	HSU, P.D. 等. "Development and Applications of CRISPR-Cas9 for Genome Engi- neering" Cell, 第157卷, 2014年 6月 5日 (2014 - 06 - 05), 第1262-1277页	1-23
A	HARRINGTON, J.J. 等. "The characterization of a mammalian DNA structure-spe- cific endonuclease" The EMBO Journal, 第13卷, 第5期, 1994年 12月 31日 (1994 - 12 - 31), 第1235-1246页	1-23

第II栏 某些权利要求被认为是不能检索的意见(续第1页第2项)

根据条约第17条(2)(a)，对某些权利要求未做国际检索报告的理由如下：

1. ☒ 权利要求： 21-22
因为它们涉及不要求本单位进行检索的主题，即：
[1] 权利要求21-22涉及疾病治疗的方法，因此不符合PCT第39.1(iv)的规定。本报告基于权利要求21-22修改为制药用途权利要求而作出。
2. ☐ 权利要求：
因为它们涉及国际申请中不符合规定的要求的部分，以致不能进行任何有意义的国际检索，具体地说：
3. ☐ 权利要求：
因为它们是从属权利要求，并且没有按照细则6.4(a)第2句和第3句的要求撰写。

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/093903

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	104745570	A	2015年 7月 1日	CN	104745570	B	2017年 4月 19日
CN	104531633	A	2015年 4月 22日	无			