

碱基编辑系统研究最新进展及应用

徐鑫¹, 刘明军²

1 新疆大学 生命科学与技术学院, 新疆 乌鲁木齐 830046

2 新疆畜牧科学院生物技术研究所, 新疆 乌鲁木齐 830026

徐鑫, 刘明军. 碱基编辑系统研究最新进展及应用. 生物工程学报, 2021, 37(7): 2307-2321.

Xu X, Liu MJ. Recent advances and applications of base editing systems. Chin J Biotech, 2021, 37(7): 2307-2321.

摘要: CRISPR 系统能够在基因组 DNA 中完成精准编辑, 但依赖于细胞内的同源重组 (Homology directed recombination, HDR) 修复途径, 且效率极低。基于 CRISPR/Cas9 系统开发的碱基编辑技术 (Base editing) 通过将失去切割活性的核酸酶与不同碱基脱氨基酶融合, 构建了两套碱基编辑系统 (Base editors, BE): 胞嘧啶碱基编辑器 (Cytosine base editor, CBE) 和腺嘌呤碱基编辑器 (Adenine base editor, ABE)。这两类编辑器分别能够在不产生 DNA 双链断裂的前提下在基因靶位点完成 C>T (G>A) 或 A>G (T>C) 的替换, 最终实现精准的碱基编辑。目前碱基编辑技术已经广泛应用于基因治疗、动物模型构建、精准动物育种和基因功能分析等领域, 为基础和应用研究提供了强大的技术工具。文中概括了碱基编辑技术的研发过程、技术优势、应用现状、存在问题及改进策略, 以期对相关领域的科研人员了解和使用碱基编辑系统提供参考。

关键词: CRISPR/Cas9, 碱基编辑, 胞嘧啶碱基编辑器, 腺嘌呤碱基编辑器

Recent advances and applications of base editing systems

Xin Xu¹, and Mingjun Liu²

1 College of Life Science and Technology, Xinjiang University, Urumqi 830046, Xinjiang, China

2 Institute of Animal Biotechnology, Xinjiang Academy of Animal Science, Urumqi 830026, Xinjiang, China

Abstract: The CRISPR system is able to accomplish precise base editing in genomic DNA, but relies on the cellular homology-directed recombination repair pathway and is therefore extremely inefficient. Base editing is a new genome editing technique developed based on the CRISPR/Cas9 system. Two base editors (cytosine base editor and adenine base editor) were developed by fusing catalytically disabled nucleases with different nucleobase deaminases. These two base editors are able to perform C>T (G>A) or A>G (T>C) transition without generating DNA double-stranded breaks. The base editing technique has been widely used in gene therapy, animal models construction, precision animal breeding and gene function analysis, providing a powerful tool for basic and applied research. This review summarized the development process, technical advantages, current applications, challenges and perspectives for base editing technique, aiming to help the readers better

Received: August 3, 2020; **Accepted:** November 19, 2020

Supported by: National Major Development Program of Transgenic Breeding, China (No. 2014ZX08010-004).

Corresponding author: Mingjun Liu. Tel: +86-991-3075280; Fax: +86-991-3075270; E-mail: xjlmj2006@126.com

转基因生物新品种培育重大专项 (No. 2014ZX08010-004) 资助。

网络出版时间: 2021-03-25

网络出版地址: <https://kns.cnki.net/kcms/detail/11.1998.q.20210322.1149.005.html>

understand and use the base editing technique.

Keywords: CRISPR/Cas9, base editing, cytosine base editor (CBE), adenine base editor (ABE)

目前已知的人类致病性突变中,最大的一类是点突变,也称为单核苷酸变异 (Single-nucleotide variant, SNV)^[1-2],约 2/3 人类疾病的发生与此有关^[3]。同时,SNV 也是影响家畜性状(如生长发育、繁殖力等)的主要遗传变异^[4-5]。因此开发一种精准、高效实现碱基突变的技术对人类健康和动物遗传育种以及遗传基础研究十分重要。

基因编辑是一项功能强大的遗传工程技术,可以在生物体内和体外完成基因组 DNA 的插入、删除、替换等遗传修饰^[6-8]。CRISPR-Cas9 (Clustered regularly interspaced short palindromic repeats/CRISPR associated protein 9) 系统自 2012 年发现以来由于其高效、方便和应用广泛等特点,已经成为强大的基因编辑工具^[9-11]。Cas9 可以通过引导 RNA (Single guide RNA, sgRNA) 在目标基因位点产生 DNA 双链断裂 (Double strand break, DSB),激活细胞内的非同源末端连接 (Non-homologous end joining, NHEJ) 或同源重组 (Homology directed recombination, HDR) 修复^[12],实现对基因的各种修饰。HDR 是一种更为精确的修复方式,可以通过加入修复模板 DNA 实现单碱基突变,但受细胞周期限制,只发生在细胞的 G2 期和 S 期,因此 HDR 主要在分裂的细胞中发挥作用,甚至在体外培养的细胞系中 HDR 效率也很低^[13-15]。此外,通常情况下 NHEJ 和 HDR 两种修复途径同时存在且相互竞争,而且 NHEJ 效率比 HDR 高。因此,大多数编辑结果通常会同时发生插入或删除 (Insertions or deletions, Indels) 编辑^[15-16]。

碱基编辑 (Base editing) 是自 2016 年发展起来的基因编辑技术,已成功应用于多种生物,能够高效、精准地诱导 DNA 和 RNA 中碱基的替换。通过将不同的碱基修饰酶 (脱氨基酶) 与

CRISPR-Cas 系统结合^[17-18],构建的胞嘧啶碱基编辑器 (CBE) 和腺嘌呤碱基编辑器 (ABE),分别实现了嘧啶碱基 (C>T)^[19]和嘌呤碱基 (A>G) 的转化^[20]。

碱基编辑的原理是将失去核酸酶切割活性的 Cas9 蛋白 (Deactivated Cas9, dCas9) 或仅能切割单链产生缺口的 Cas9 蛋白 (Nickase Cas9, nCas9) 和作用于单链 DNA (Single-stranded DNA, ssDNA) 碱基的脱氨酶融合,依靠 sgRNA 将碱基编辑酶锚定到靶位点。碱基编辑系统结合 DNA 靶位点后,sgRNA 和靶位点 DNA 通过碱基序列互补配对形成“R 环”,产生局部单链 DNA (ssDNA),脱氨酶将“R 环”内 ssDNA 上一定范围内的胞嘧啶 (C) 或腺嘌呤 (A) 进行脱氨基作用 (图 1A 和图 2)^[21],经 DNA 修复或复制,实现精准的 C>T 或 A>G 的替换。

1 CBE 碱基编辑系统

APOBEC (Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like) 家族是目前报道的自然存在可作用于单链 DNA 的胞嘧啶脱氨酶^[22]。2016 年,美国哈佛大学 David R. Liu 实验室率先建立了基于融合大鼠胞嘧啶脱氨酶 rAPOBEC1 的碱基编辑器 (Base editors, BE)^[19]。胞嘧啶脱氨酶可以将 DNA 中的胞嘧啶 (C) 脱氨为尿嘧啶 (U),进而通过 DNA 修复或复制将 U 转变为胸腺嘧啶 (T),最终实现 C>T 或 G>A 的直接替换 (图 1A)。第一代碱基编辑器 (BE1) 由 dCas9 和大鼠的 APOBEC1 组成, BE1 (rAPOBEC1-XTEN-dCas9) 可以在体外有效地诱导 C>T 碱基编辑,脱氨酶的编辑窗口范围约为 5 bp,在 sgRNA 序列的第 4-8 位 (PAM 计为第 21-23 位)。在哺乳动物细胞内使用 BE1 系统需要解决的最主要问题是如何规避细胞

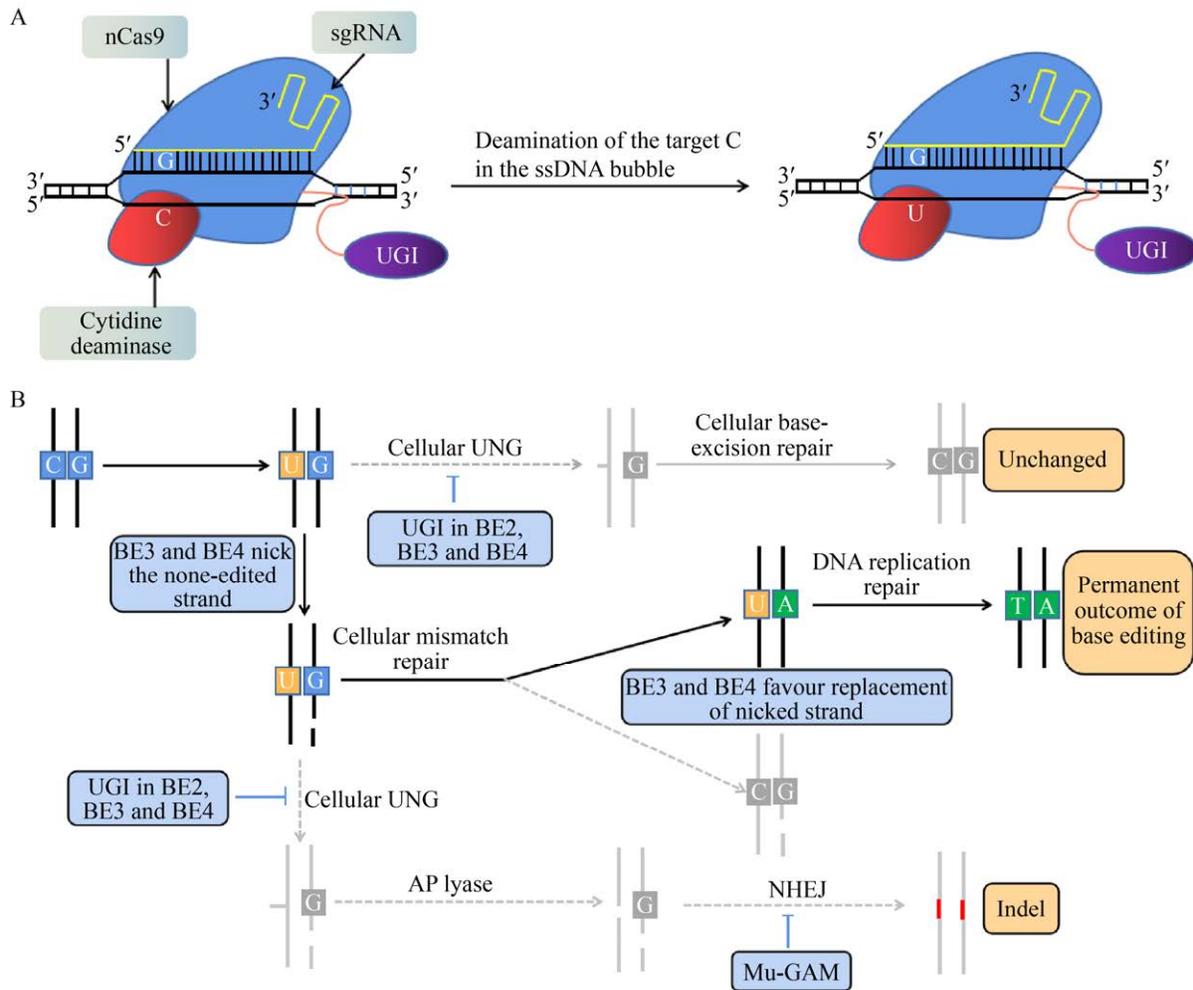


图 1 胞嘧啶碱基编辑

Fig. 1 Cytosine base editing. (A) Cytosine base-editing strategy. Cytosine deamination generates uracil, which has the same base pairing preferences as a thymine in the active site of a polymerase. (B) The cellular response to cytosine base editing. AP lyase, DNA (apyrimidinic site) lyase; indel, small insertion or deletion; Mu-GAM, bacteriophage Mu derived Gam protein; NHEJ, non-homologous end joining; UGN, Uracil DNA N-glycosylase; UGI, uracil DNA glycosylase inhibitor.

DNA 修复系统对已发生脱氨的碱基的修复。BE1 在动物细胞内脱氨作用较低（体外脱氨效率为 25%–40%，细胞内只有 0.8%–7.7%）^[19]。脱氨效率大幅度的降低与细胞修复 U•G 错配碱基对有关^[23]，细胞内的尿嘧啶 DNA 糖基化酶 (Uracil DNA N-glycosylase, UNG) 可识别错配 U•G 碱基对，通过碱基错配修复途径 (Base excision repair, BER) 将 BE1 产生的 U•G 修复恢复 C•G 配对^[24]。

为了抑制 UNG 功能，Komor 等^[19]在 BE1 的 C 端融合了来源于噬菌体 PBS 的尿嘧啶 DNA 糖基化酶抑制剂 (Uracil DNA glycosylase inhibitor, UGI)，开发了第二代碱基编辑器 BE2 (rAPOBEC1-XTEN-dCas9-UGI)，相较于 BE1 其编辑效率有所提升。为了提高 C>T 的转换效率，Komor 等又开发了第三代碱基编辑器 BE3 (rAPOBEC1-XTEN-nCas9-UGI)，将 BE2 中的 dCas9 换成 nCas9。由于 nCas9

能够在非编辑的 DNA 链上产生缺口, 细胞内的错配修复机制 (Mismatch repair, MMR) 倾向于以未产生缺口的编辑链为模板进行修复, 因此极大地增加了细胞内碱基转换效率^[19]。Komor 等在 BE3 的基础上, 在 C 端增加一个拷贝的 UGI, 同时增加 rAPOBEC1 和 nCas9 之间连接短肽的长度 (16 aa 到 32 aa), 开发了第 4 代碱基编辑器 BE4 (rAPOBEC1-XTEN-nCas9-UGI-UGI), 进一步提高了碱基的编辑效率 (图 1B)^[25]。

2016 年 Nishida 及其同事用胞苷脱氨酶 1 (Cytidine deaminase 1, CDA1) 代替 APOBEC1, 开发了一种与 BE 系统类似的称为“Target-AID (Activation-induced deaminase)”碱基编辑系统 (nCas9-CDA1-UGI)^[26], 用于嘧啶碱基替换。Target-AID 显示的编辑窗口相对于 BE3 略有偏移, 在 sgRNA 的第 2–4 位。哈佛大学 George M. Church 实验室将锌指核酸酶 (Zinc-finger nucleases, ZFN) 或转录激活因子样效应物核酸酶 (Transcription activator-like effector nucleases, TALEN) 的特异性识别单元与人类胞嘧啶脱氨酶 hAID (Human activation-induced deaminase) 融合构建了不同于 Cas9 的碱基编辑系统, 这两套系统无 PAM 结构需求, 但 C>T 转换率很低 (大肠杆菌中 13%, 人类细胞中 2.5%), 推测可能是因为 ZFN 或者 TALEN 与 DNA 结合后不能形成 ssDNA, 脱氨酶结合区域双链 DNA (Double-stranded DNA, dsDNA) 严重降低了脱氨酶的催化效率^[27]。

美国斯坦福大学 Michael C. Bassik 实验室开发的基于 MS2 招募 hAID 脱氨酶的 CRISPR-X 系统, 可以在基因组有限脱靶情况下产生多个靶位点的突变, 通过建立多位点定点突变文库可以同时靶向基因组的多个位点^[28]。中国科学院上海生命科学研究院健康科学研究所常兴实验室开发的 TAM 系统 (Targeted AID-mediated mutagenesis), 将 dCas9 和 AID 融合在 sgRNAs 的指导下可完成较大范围 (编辑范围宽达 100 nt) 的碱基编辑^[29]。

2 ABE 碱基编辑系统发展

人体有 6 种点突变形式, 这 6 种点突变形式的发生比例是不一样的^[2-3]。这种差异主要与体内发生较高频率的自发性胞嘧啶脱氨基 (估计每天每细胞大概有 100–500 次脱氨事件) 有关, 如果不进行校正, 将有大量的 G•C 碱基对突变为 A•T 碱基对^[30-31]。因此, 能够将 A•T 碱基对转变为 G•C 碱基对的分子工具备受关注, 因为它能够校正 47% 的与疾病相关的点突变^[2-3]。腺嘌呤 (A) 与胞嘧啶 (C) 一样含有一个环外的氨基, 脱氨后会改变其配对形式, 腺嘌呤 (A) 脱氨后形成肌苷 (I) (图 2), 在 DNA 水平, 肌苷会被当作鸟嘌呤 (G) 进行读码和复制, 最终实现 A>G 的转换^[32]。

开发 ABE 系统时遇到的主要障碍是已知的腺嘌呤脱氨酶都无法作用于 ssDNA, 尝试过将几种 RNA 腺嘌呤脱氨酶 (大肠杆菌 TadA、大鼠 ADA、人类 ADAR2 和人类 ADAT2) 直接替换 BE3 中的 APOBEC1 构建碱基编辑系统, 但均未检测到腺嘌呤 (A) 碱基编辑活性^[20], 说明天然存在的腺嘌呤脱氨酶需经过改造才能实现对 A 的编辑。为了克服这个问题, Gaudelli 等^[20]将大肠杆菌 tRNA 腺苷脱氨酶 TadA 改造成了一种能够作用于 ssDNA 的腺苷脱氨酶。将随机突变的 TadA 序列与编码 dCas 蛋白的序列融合构建随机突变文库, 通过恢复抗性基因的功能, 筛选获得能够作用于 ssDNA 的突变体 TadA*。将野生型的 TadA 构建到 TadA*-dCas9 中 (TadA-TadA*-nCas9), 与 TadA* 形成异源二聚体极大地提高了哺乳动物细胞中 A>G 转换效率。ABE 系统经过 7 轮改造, 筛选到了效率较高的 ABE 版本 ABE7.10, 其有效编辑窗口在 sgRNA 的第 4–7 位。不同版本的 ABE, 如 ABE7.9 或 ABE6.3, 在靠近 PAM 的位置 (第 8 位或第 9 位) 会有更高的编辑效率。ABE 的应用受到脱氧腺苷脱氨酶与 SpCas9 以外 Cas 同系物的有限兼容性限制, Richter 等使用噬菌体辅助的非连续和连续进化 (Phage-assisted non-continuous

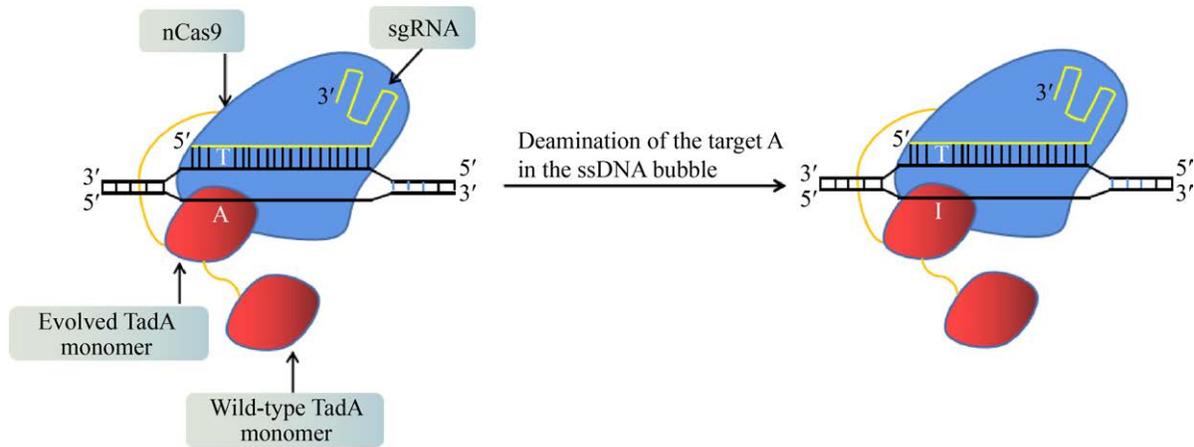


图2 腺嘌呤碱基编辑

Fig. 2 Adenine base editing. Adenosine deamination generates inosine, which has the same base pairing preferences as a guanosine in the active site of a polymerase.

and continuous evolution, PANCE 和 PACE) 继续改进 ABE7.10 的脱氨酶组分, 从而产生了目前编辑效率最高的 ABE 版本: ABE8e。与 ABE7.10 相比, ABE8e 包含 8 个额外的突变, 可将活性 (kapp) 提高 590 倍, 且与各种 Cas9 或 Cas12 同源物配对时, 可以大大提高编辑效率^[33]。Lapinaite 等通过冷冻电镜首次解析了 ABE 结合 DNA 的 3D 结构, 发现 ABE8e 比早期版本的 ABE 编辑效率高是因为 ABE8e 中的突变使得蛋白能够稳定 DNA 的构象, 加快了 A>G 转变效率^[34]。

3 碱基编辑系统存在的问题及解决办法

3.1 编辑产物的纯度

CBE 的最初几项研究结果显示, 在基因组某些位点观察到 C>R (G 或 A) 的转换, C>R 的碱基转变降低了碱基编辑产物纯度^[35-37]。Komor 等通过在缺少 UNG 基因的细胞中进行胞嘧啶碱基编辑实验, 发现 UNG 缺失的细胞中, 12 个靶位点胞嘧啶的产物纯度从平均 68% 提高到 >98%, 这表明 UNG 是影响编辑产物纯度的重要因素^[25]。BE4 在 BE3 的结构上增加了一个拷贝的 UGI, 可提高含有 UNG 细胞系的编辑产物纯度。过表达 UGI 也可以提高 BE3 编辑的产物纯度^[37], 但会导致整

个细胞内 C>T 突变率的增加^[38]。

ABE 系统的腺嘌呤碱基编辑产物纯度非常高。目前为止, 尚无报道显示有 A>Y (T 或 C) 编辑事件发生^[39-41], 可能是因为细胞从基因组 DNA 中去除肌苷 (I) 的能力比去除尿嘧啶 (U) 的能力弱得多。已知烷基腺嘌呤 DNA 糖基化酶 (Alkyl adenine DNA glycosylase, AAG; 也被称为 MPG) 可识别并去除 DNA 中的肌苷^[42], 但在 AAG 缺乏的细胞中使用 ABEs, 并不能提高其编辑效率^[20]。

3.2 产生 Indels

CBE 碱基编辑会产生低频率的 Indels, 是因为细胞内的 UNG 能够将 U 切除, 形成一个无嘧啶位点 (Apyrimidinic site, AP), AP 位点会在 AP 裂解酶的作用下在编辑链产生一个缺口^[43], 进而与 nCas9 在非编辑链产生的切口刚好形成一个 DSB, 之后进入易产生 Indels 的 NHEJ 修复途径。Komor 等发现, UNG 敲除的细胞除了编辑产物纯度提高外, Indels 的发生率也减少了。Komor 等将噬菌体 Mu 衍生的 Gam 蛋白 (Mu-GAM) 与 BE4 融合以生成 BE4-Gam。BE4-Gam 相对于 BE4 可以进一步减少 HEK293T 细胞中的 Indels 发生率, 因为 Gam 蛋白可以结合在 DSB 的末端防止其降解, 进而阻止 NHEJ 的发生^[25]。在兔子胚胎实验

中, BE4-Gam 相较于 BE3 可以显著降低 Indels 发生率并提高编辑产物纯度^[44]。

在细胞系^[45-46]、小鼠^[41]和植物^[47]的实验中, ABE 系统 Indels 的发生率很低, 普遍低于 1%, 有些实验中甚至未检测到 Indels。ABE 产生较少的 Indels 是因为缺少 DNA 修复所需的糖苷酶, 因此不会将 I 切除并在 DNA 编辑链产生切口^[20]。由于细胞内 I 被切除的效率要远远低于 U, 所以在编辑链产生的切口也更少, 因此 ABE 比 CBE 具有更低的 Indels 发生率。

3.3 编辑窗口及邻近位点编辑

BE3 系统中脱氨酶的有效编辑窗口在第 4–8 位^[19], 但位于编辑窗口外部 ssDNA R 环区域内的其他碱基仍有可能以较低的概率发生编辑。编辑窗口内存在多个可编辑的 C 或 A 时, 会导致除了目标碱基外其他碱基发生编辑。用“邻近碱基编辑 (Bystander editing)”一词来描述那些在 sgRNA 区域内除靶位点外发生的碱基编辑事件。当碱基编辑的目的是破坏启动子、mRNA 剪接位点或其他调控序列, 或者是引入提前终止密码子时, 非 CDS 区域邻近碱基编辑的发生可能是无关紧要的。但当编辑功能蛋白基因 CDS 区时, 编辑窗口内非靶标碱基的编辑, 会导致靶蛋白结构和功能改变, 尤其对基因治疗产生风险。

David R. Liu 和他的同事对 CBE 进行改造, 在 rAPOBEC1 的功能域中引入突变, 减弱了 rAPOBEC1 的脱氨活性, 导致编辑器的编辑效率降低和编辑窗口变窄 (YE1-BE3, YE2-BE3 和 YEE-BE3)。rAPOBEC1 的编辑窗口由原来的 5 nt 缩小至 1–2 nt^[48], 这些编辑窗口变窄的 CBE 编辑器可以在编辑窗口内选择靶位点的 C 进行编辑而不会编辑邻近的 C。Tan 等通过缩短 BE3 中脱氨酶与 nCas9 之间的连接短肽开发的 BE3-PAPAPAP, 以及通过截短 PmCDA1 脱氨酶羧基端序列而建立的一系列 nCDA1-BE3 版本, 在保证对靶标 C 编辑效率依然高效的同时, 可有效将编辑窗口缩

小到 1–2 nt^[49]。

但是, 有些研究需要用到编辑窗口范围较宽的碱基编辑器, 如筛选功能基因、定位蛋白结构域关键氨基酸位点等。ABE7.10 编辑窗口大约位于 sgRNA 的第 4–7 位, 而 ABE7.9 或 ABE6.3 可在第 4–9 位进行编辑。Jiang 等^[50]开发了 BE-PLUS 系统, 将 BE3 的编辑窗口由 5 nt (4–8) 拓宽至 13 nt (4–16)。基于融合 hA3A 胞嘧啶脱氨酶的碱基编辑器在人类细胞及植物中脱氨化的窗口可分别宽达 14 nt (2–13) 和 17 nt (1–17)^[51-52]。而 CP1020-CBEmax 和 CP1028-CBEmax 的编辑窗口可覆盖 sgRNA 的 8–9 nt^[53]。

3.4 PAM 序列限制

来自酿脓链球菌 *Streptococcus pyogenes* 的 Cas9 (SpCas9) 是目前应用广泛的核酸酶, 它识别的 PAM 序列为 NGG, 只有约 26% 的致病性 SNP 可以被 SpCas9 及其衍生的碱基编辑器编辑。使用来源于金黄色葡萄球菌 *Staphylococcus aureus* 的 Cas9 (SaCas9) 以及 SpCas9 和 SaCas9 的突变体构建的一系列编辑器, 如 SaBE3、Sa (KKH)-BE3、VQR-BE3、VRER-BE3 和 EQR BE3, 能够识别非 NGG 的 PAM 序列^[48]。Chen 实验室使用 Cas12a (即 Cpf1) 开发了一套 CBE 系统, 识别的 PAM 序列为 TTTV (V 可以是 A、C 或 G), 这一系统适用于富含 T 的基因组 DNA^[54]。Liu 实验室改造的 xCas9 (3.7) 替换 BE3 中的 SpCas9 构建 xBE3 系统, 使其能够在 PAM 序列为 NGN、GAA 和 GAT 目标位点发生编辑^[55]。Nureki 实验室开发的 SpCas9-NG 能够识别 PAM 序列为 NGN 的位点^[56]。最近, Russell 及其同事开发出了几乎没有 PAM 序列限制的 SpRY, 实验发现融合 SpRY 的碱基编辑器几乎能够识别所有的 PAM 位点, 并且在 PAM 序列为 NRN (R 为 A 或 G) 的靶位点编辑活性要高于 PAM 序列为 NYN (Y 为 C 或 T) 的靶位点^[57]。SpRY 的出现使消除 PAM 序列限制成为可能。

另外, 识别不同 PAM 的 Cas9 及突变体

(SaCas9^[48]、Sa(KKH)Cas9、Sp(VQR)Cas9 和 Sp(VRER)Cas9^[58-59]) 也被开发出来用于不同的 ABE 碱基编辑器。这些编辑器的产生极大地增加了 ABE 的编辑范围。

3.5 编辑序列背景依赖性

在某些包含 GC 序列的位点, rAPOBEC1 的脱氨效率不高^[19,25]。CpG 处的 DNA 甲基化可降低 rAPOBEC1 介导的碱基编辑效率,但人 APOBEC3A (hA3A) 可以更有效地编辑 CpG 中的 C, 且编辑效率比 rAPOBEC1 高^[53]。Joung 实验室利用不同胞嘧啶脱氨酶的序列偏好性, 设计了基于突变体 hA3A 的 CBE。该 CBE 对 TC 中的 C 具有偏好性^[60], 并且可减少邻近碱基编辑事件发生。进一步对 hA3A 脱氨酶的突变体进行筛选和结构改造, 产生了含有单个突变 (N57G) 的加强版的突变体 (eA3A), 能够编辑 TC 中的 C, 极大地降低了其他序列背景下 C 的编辑活性。

目前为止, 哺乳动物细胞 ABE7.10 编辑的结果显示, 在人细胞中 ABE7.10 发生编辑时没有序列背景依赖性^[20], 但是 Kim 和他同事的研究结果显示, 在拟南芥中 ABE7.10 表现出了一定的编辑偏好性, 相较于 GA、CA 或 AA 序列, ABE7.10 会优先选择编辑 TA 中的 A^[47]。

3.6 碱基编辑脱靶效应

与 CRISPR/Cas9 系统一样, BE 系统也会产生脱靶效应^[19-20]。BE 系统产生的脱靶分为可预测脱靶和不可预测脱靶, 可预测脱靶一般与以下几个原因有关: 第一, sgRNA 具有一定的容错性, 允许 sgRNA 与 ssDNA 之间存在个别碱基的错配, 脱靶效应的高低与 sgRNA 和 ssDNA 之间错配碱基数目有关。利用缩短的 sgRNA (gX18 或 gX17) 或延长的 sgRNA (gX20 或 ggX20) 均可在保持 BE 系统编辑效率的前提下降低其脱靶效应^[34,61]。第二, Cas9 核酸酶的非特异性, BE 的 DNA 靶向功能是由 Cas9 蛋白决定的, 因此在研究 BE 的脱靶编辑时利用深度测序法重点检测了那些 Cas9

能够编辑的基因位点^[19-20,48]。结果显示, Cas9 的脱靶位点如果含有 C 且在脱氨酶活性窗口内, 就会在 CBE 的作用下发生较低效率的编辑。因为并不是所有 Cas9 的脱靶位点都含有可编辑的 C, 所以使用相同的 sgRNA, CBE 的脱靶效应要比相对应的 Cas9 低。为了降低 CBE 的脱靶效应, Kleinstiver 等通过在 Cas9 中引入能够提高其保真度的突变, 构建了具有高保真度的 CBE 系统 (HF-BE3)^[62]。与原始的 CBE 相比, HF-BE3 即使与容错率很高的 sgRNA 配对也能够显著降低脱靶效率。Lee 等通过引入不同的突变建立具有高保真度的 Sniper-BE3 能够明显改善 CBE 系统的脱靶效应^[63]。Liang 等使用噬菌体来源的 anti-CRISPR 蛋白 AcrIIA5 也可以降低 BE3 和 ABE7.10 的脱靶效率^[64]。第三, 脱氨酶和 Cas9 融合蛋白在细胞内高浓度、长时间的表达也容易造成脱靶效应。可以利用碱基编辑器与 sgRNA 形成的核糖核蛋白复合物 (Ribonucleoprotein, RNP) 取代其 DNA 表达形式直接转化细胞, 减少 BE 作用时间, 从而降低脱靶效应^[61]。使用相同的 sgRNA, ABE 脱靶效应远远低于 Cas9, 甚至比 BE3 都少^[19-20]。

BE 系统不可预测脱靶效应与脱氨酶和 UGI 的过表达有关。脱氨酶作用于 ssDNA, 在 DNA 转录、复制的过程中会形成大量的 ssDNA 区域, 这些区域会成为潜在的脱靶位点。另外, 细胞内存在大量天然的 C>U 突变, UGI 的过表达阻断了 U•G 错配修复机制, 增加了 C>T 转化率。有研究表明, 在酵母及人类细胞系过表达不同胞嘧啶脱氨酶或 UGI, 会导致基因组内 C>T SNVs 的增加^[65-66]。因此传统的脱靶预测软件 (Digenome-seq、Endo VIII 或 Endo V-seq) 并不能真实有效地评估细胞内碱基脱靶水平, 全基因组测序 (WGS) 是检测细胞基因组中所有类型碱基脱靶编辑的最佳手段。杨辉实验室利用 GOT1 (Genome-wide off-target analysis by two-cell injection) 技术结合 WGS 比较突变位点, 发现 BE3 在小鼠体内引发了比自发突

变高 20 倍的 C>T 单核苷酸变异。相反, ABE 系统几乎没有脱靶的 A>G 单核苷酸变异,说明 ABE 相较于 CBE 具有很高的特异性^[67]。同一时期,高彩霞课题组发现 BE3 和 HF1-BE3 能够引起水稻全基因组水平 C>T SNVs 的增加^[68]。并且以上两项研究均显示, C>T SNVs 均匀分布在染色体间,但明显富集在转录活跃区。该现象是否会出现融合了其他脱氨酶的 CBE 系统中还有待进一步的研究。近期,高彩霞课题组改造的 BE3 突变体 A3Bctd-VHM-BE3 和 A3Bctd-KKR-BE3,经验证不仅可以提高靶位点编辑的特异性,还可以在水稻全基因组水平上消除 sgRNA 非依赖的脱靶效应^[69]。以上研究提示 WGS 应用到碱基脱靶检测中具有很大的优势,能够更加准确地评估碱基脱靶效应。出现的这些问题同时提醒现阶段将碱基编辑系统应用于疾病治疗等精准改造方面仍具有一定的风险,需要科研工作者开发出更加精准安全的 BE 系统。

早期研究报道,无论是胞嘧啶脱氨酶还是嘌呤脱氨酶都可作用于 RNA,因此 BE 系统有可能在 RNA 水平产生脱靶效应。2019 年,Grünwald 等率先证明了 BE3 和 ABE 系统在人类细胞系转录组水平存在严重的脱靶效应^[70]。其中, BE3 可在大部分表达基因的 RNA 中诱导 C>U 编辑,某些 RNA 中突变频率甚至可达 100%;而 ABEmax 在 RNA 水平对碱基 A 的脱靶编辑效率也可高达 100%^[70]。同一时期,杨辉实验室定量检测了 CBEs 和 ABEs 在 RNA 水平诱导的单核苷酸变异,结果发现 BE3 和 ABE7.10 都能产生数以万计非靶向的 RNA 突变^[71]。对于碱基编辑系统在 RNA 水平的严重脱靶现象,目前主要的解决方案是通过改造脱氨酶产生突变体,即控制脱氨酶与 RNA 的活性结合位点从而降低对 RNA 的编辑效率。基于此,Keith Joung 实验室开发的 BE3-R33A 和 BE3-R33A/K34A 突变体可将 BE3 对 RNA 的编辑分别降低 390 倍和 3 800 倍^[70]。杨辉实验室开

发的 BE3 (W90Y+R126E)、hA3A (R128A) 和 hA3A (Y130F) 可直接将 BE 脱靶率降至本底水平;针对 ABE 系统也开发了两种 ABE 突变体 ABE7.10 (D53E) 和 ABE7.10 (F148A),其中 ABE7.10 (F148A) 甚至可直接消除 ABE 系统在转录组水平的脱靶效应^[71]。David R. Liu 实验室开发的 ABEmaxAW (TadA E59A, TadA* V106W) 和 ABEmaxQW (TadA E59Q, TadA* V106W) 可在保证一定的 DNA 编辑效率的同时,降低在 RNA 水平的脱靶效应^[72]。

4 碱基编辑器应用

4.1 碱基编辑用于人类基因治疗

目前已知的人类遗传病大部分都是由点突变引起的,而 CBE 和 ABE 系统可以纠正约 60% 的致病性点突变^[2-3]。Komor 等将 BE3 以质粒的形式转染小鼠星形胶质细胞,可以将与阿尔茨海默病相关的等位基因 APOE4 转化为 APOE3r,并可以纠正乳腺癌细胞中与癌症相关的 *p53* 基因突变 (Y163C)^[19]。将密码子优化的 CBE 以质粒形式递送至患者来源的成纤维细胞中,可以纠正引起 1f 型先天性糖基化障碍的 *MPDU1* 基因中 L119P 突变^[46]。Gaudelli 等发现 ABE7.10 可以纠正永生患者来源的 LCL 细胞中遗传性血色病引起的突变 (C282Y),并且可以在成人细胞中制造增加胎儿血红蛋白 (HBG) 表达的突变^[20]。Kim 等使用 ABE7.10 纠正肌营养不良小鼠模型中肌营养不良蛋白基因 (*Dmd*) 的提前终止密码子,尽管在所有测序细胞中校正率仅为 3.3%,但肌营养不良蛋白在 17% 的肌纤维中恢复表达^[37],表明低水平的编辑通常也会产生治疗上相关的表型改变。另外,Chadwick 等利用 BE3 在鼠 *Pcsk9* 基因中产生 W159X 终止密码子突变,发现肝细胞编辑效率约 25%,并观察到 4 周后血浆 PCSK9 蛋白水平和血浆胆固醇明显降低^[73]。除了基因治疗,碱基编辑还可用于制造动物疾病模型。Lapinaite 等利用

FNLS-BE3 在 *Ctnnb1* 中产生与癌症相关的突变 (S45F), 发现在肝细胞中碱基编辑效率接近 100%, 并且与对照相比, FNLS-BE3 编辑的小鼠生长了大量可见肿瘤结节^[34]。Zhang 等利用 BE3 在斑马鱼 *tyr* 基因中产生 P302S 突变, 模拟人类常见的眼白化病, 这种方法可以研究 P302S 突变对眼部色素沉淀的影响^[74]。这些研究提示我们碱基编辑器有望成为潜在的点突变疾病治疗方法, 并且是研究细胞和体内精准遗传突变表型效应的有效方法。

4.2 CBE 引入提前终止密码子

CBE 可以通过精确的碱基编辑在需要的位置制造终止密码, 在不产生 DSB 的情况下达到敲除基因的目的。Kim 等率先使用 BE3 在小鼠胚胎的 *Dmd* 基因引入提前终止密码子建立肌营养不良小鼠疾病模型, 证明了此方法的可行性^[37]。最近建立的 CRISPR-Stop^[75]和 iSTOP^[76]方法, 实现了 BE3 介导的基因失活。Billon 等建立了一个 sgRNA 数据库, 与 BE3 结合使用, 理论上能够在人类基因组大于 98.6% 的开放阅读框中产生提前终止密码子 (参考基因组为 GRCh38); 他们公布了一个可免费访问的在线数据库, 使研究人员能够找到合适的 sgRNA, 在包括人类、小鼠和拟南芥等 8 个物种中使用 iSTOP 方法^[76]。Kuscu 等发现, 与 Cas9 敲除基因相比, 此策略可显著降低细胞凋亡^[75], 这可能是由于 DSB 可诱导较低的细胞毒性^[77-79]。尽管 NHEJ 介导的基因敲除法被广泛利用, 但此方法会产生许多其他编辑类型的细胞, 包括 DNA 易位和重排^[80], 甚至会导致细胞死亡^[77-79], 原则上通过使用碱基编辑器产生终止密码子可避免这些问题。

4.3 调节 mRNA 可变剪接

mRNA 可变剪接是转录后基因重要的调控过程, 外显子剪接过程中的重要步骤是通过剪接体机制识别外显子和内含子的高度保守的序列。几乎每个内含子都以 G 结束, 基于此, Gapinske 等

利用 CBE 将互补链的 C 转变为 T, 从而导致内含子上的 G 转变为 A, 破坏高度保守的剪接受体序列 (5'-AG-3'), 导致外显子“跨越”^[81], 此方法被称为 CRISPR-SKIP。常兴实验室利用 TAM 系统在人类细胞中实现了目前存在的 4 种 mRNA 剪接形式的改变, 并且通过编辑 *Dmd* 基因的剪接位点, 实现了目标外显子的完全跨越, 在所有表达 TAM 的细胞中修复了心肌细胞的缺陷^[82]。中山大学李剑峰实验室利用 BE3 在拟南芥和水稻中建立了有效的 mRNA 可变剪接体系^[83]。高彩霞实验室利用 BE3 在拟南芥中对 *AtAct2* 基因 5' UTR 区的内含子剪接位点处进行突变, 发现利用该方法可高效地改变 mRNA 剪接形式^[84]。

4.4 胚胎碱基编辑构建动物模型

在单细胞胚胎阶段进行基因组编辑的主要目标是产生模型生物。由于核酸酶介导的编辑策略通常无法在 F0 代中产生纯合子后代^[85-86], 因此高效的碱基编辑为解决此问题提供了一种有吸引力的选择。Kim 等利用 BE3 在单细胞阶段小鼠胚胎编辑基因 *Tyr* 的 Q68X 和 *Dmd* 的 Q871X, 建立了白化病和肌营养不良小鼠模型^[37]。Wang 等使用 BE4max 构建了早衰症 (Hutchinson-gilford progeria syndrome, HGPS) 猴子模型^[87]。Liu 等通过将 BE3、BE4-Gam 或 ABE7.10 的 mRNA 注入囊胚中建立了哈钦森-吉尔福德早衰综合征的兔子模型^[44]。Huang 及其同事通过在单细胞小鼠胚胎中共同注射 ABE7.10 和 SaBE3 mRNA 以及 sgRNAs 进行多重碱基编辑, 囊胚中同时观察到 A>G 和 C>T 的编辑^[88]。Zhang 等在小鼠胚胎共同注射两个不同的 sgRNA, 可在 F0 代中同时产生两个可遗传的 A>G 点突变^[89]。Yang 等利用 ABE7.10 建立了因 *Fah* 突变引起的 I 型酪氨酸血症小鼠模型, 并且该研究也证实 SaKKH-ABE 和 VQR-ABE 两种编辑器同样可在小鼠中产生有效的 A>G 突变^[59]。这些数据表明碱基编辑是构建动物模型的有效手段, 为进行临床疾病治疗奠定了坚实的基础。

4.5 家畜碱基编辑

家畜中许多遗传性状都是 SNP 造成的, 生物学中许多遗传疾病也是由点突变引起的, 因此碱基编辑在获得家畜优良性状以及建立人类疾病的动物模型方面有巨大优势。Li 等在猪胚胎成纤维细胞中利用 BE3 编辑 *Twist2* 基因和 *TYR* 基因, 然后通过体细胞核移植的方法建立了无脸巨口综合征和皮肤白化病的疾病模型^[90]; 赖良学实验室在猪中首次实现了多个基因靶点的同时 C>T 突变, 并建立了相应的疾病模型^[91]。Zhou 等和 Li 等利用 BE3 编辑 *SOCS2* 和 *FGF5* 基因, 成功获得了体重、体尺和毛长等生长指标增加的突变体羊^[92-93]; 笔者实验室利用 xBE4 和 xABE 系统在体外胚胎中成功编辑了与绵羊繁殖性状相关的 *FecB* (A746G)、*BMP15* (C950T) 和 *GDF9* (C1184T) 基因的 SNP 位点, 以及 *FGF5* 基因提前终止的位点 (C232T, CAG>TAG), 并获得了相应靶位点突变的囊胚。以上实验表明, 碱基编辑技术可用于家畜遗传性状的改良及大动物遗传疾病模型的建立。

4.6 植物碱基编辑

植物碱基编辑可以高效、快速产生新的植物突变体^[94]。精确的功能获得性点突变可以改良许多农艺性状, 例如植物 *ALS* 基因中的点突变能够产生对磺酰脲类和咪唑啉酮类化合物除草剂的抗性^[95]。利用 CBE 和 ABE 已在多种具有农学意义的植物中完成了碱基编辑。Gao 和他的同事证明, BE3 可以在玉米、水稻和小麦中产生有效的点突变^[96]。在另一项研究中, Kondo 及其同事发现, Target-AID 编辑器能够有效地编辑水稻和番茄^[97]。Zhou 和 Zhu 实验室的两项独立研究报告表明, ABE 在水稻中能够产生高效的碱基编辑^[39-40]。Kim 和他的同事通过农杆菌介导将 ABE7.10 转染到拟南芥和甘蓝型油菜中而产生两个表型变化^[47]; 他们使用植物优化的表达系统, 在拟南芥 FT 蛋白中产生 Y85H 突变, 从而导致迟花表型; 破坏 *PDS3* 中的剪接受体位点, 产生矮表型植株。转化后,

超过 85% 的 T1 植物显示出 >50% 的编辑效率, 并且从 T1 植物中分离的 T2 幼苗也显示出相同的表型, 表明该编辑是可遗传的^[47]。高彩霞实验室利用基于 rAPOBEC1 的 PBE 系统在水稻、小麦和玉米原生质体中实现了多个基因的编辑, 并获得了突变效率高达 45.8% 的相应突变体植株^[97]; 并利用 PBE 系统编辑小麦 *TaALS* 和 *TaACCcase* 基因, 创制了一系列抗除草剂小麦新种质^[98]。

以上研究表明, 碱基编辑是快速改造多倍体植物基因组的有效方法。从监管和消费者的角度出发, 将碱基编辑器以 RNP 的形式转入农作物将有重要意义, 因为 BE 若以质粒的形式转染存在转基因整合问题, 会给农作物贴上 GMO 的标签。碱基编辑器以 RNP 的形式转染农作物, 提供的将是无外源 DNA 的精确编辑, 从而可以避免 GMO 产生的生物安全问题^[95]。

5 总结与展望

以往在生物体单个碱基水平上高效地完成遗传信息改变似乎是不可能的, 碱基编辑器的出现使这一想法成为现实。目前为止, 已报道有两类 DNA 碱基编辑器可有效地在各种细胞和生物体的基因组中进行精准的点突变, 但 CBE 和 ABE 目前仅能够完成碱基对 6 种改变中的两种改变 (C>T, A>G)。Sakata 等将胞嘧啶脱氨酶 PmCDA1 和腺嘌呤脱氨酶 TadA 同时与 nCas9 融合, 构建了 Target-ACE 系统, 能够同时在靶位点完成 C>T 和 A>G 的改变^[99]。Kurt 等最近构建两套的 C>G 碱基编辑器 (C>G base editors, CGBE): CGBE1 和 miniCGBE1 能够在人类细胞系中完成 C>G 编辑^[100], CGBE 的出现进一步扩大了 BE 系统的应用范围。除了 BE 系统, 最新建立的 PE 基因编辑系统 (Prime editing) 能够完成全部类型的碱基转换^[101], 虽然在传统的编辑窗口内 (第 4-8 位) PE 系统的编辑效率低于 BE 系统, 但在编辑窗口外 PE 系统的编辑效率要高于 BE 系统; 而且, PE

系统碱基编辑的精准性要高于 BE, 编辑窗口内不会出现非靶碱基的编辑。BE 系统和 PE 系统互有所长, 当靶位点编辑窗口内只有一个可编辑位点时 (A 或 C), 选择 BE 系统; 当编辑窗口内有多个可编辑位点 (A 或 C) 或需要完成碱基颠换时, 选择 PE 系统。目前而言, PE 系统还无法完全替代 BE 系统完成碱基编辑, 还需要大量的实验检验 PE 系统的可靠性。

从目前 BE 系统应用于绵羊的碱基编辑工作来看, 若编辑窗口内存在多个 A 或 C, 临近位点发生脱靶编辑似乎不可避免, 且 CBE 系统还会在靶位点附近造成大量的 Indels^[92-93]。笔者实验室前期绵羊体外胚胎实验发现, xCas9 (3.7)-ABE (7.10) 系统可以将 *FecB* 基因 sgRNA 第 2-7 位 A 转变为 G, 且 A743G 的编辑效率远高于靶位点 A746G; xCas9 (3.7)-BE4 系统有机会将 *BMP15* 靶位点临近的 C 都转变为 T, 每次发生的都是多个 C 的编辑事件, 且 Indels 发生率高; 在 *GDF9* 的 C1184T 位点比较了 xCas9 (3.7)-BE4 系统和 SpCas9 介导的 HDR 方法完成精准编辑的情况, 发现 SpCas9 介导的 HDR 虽然会发生一定概率的 Indels, 但其精准编辑效率远高于 xCas9 (3.7)-BE4 系统。综上所述, 笔者认为想要完成精准碱基编辑需要关注两点: 一是 BE 系统的选择, 要选择编辑效率高、编辑窗口相对较窄、Indels 发生率低的 BE 系统; 二是不要单纯依赖 BE 系统, 针对不同的靶位点可以通过比较选择合适的编辑方式: BE 系统、Cas9 介导的 HDR 或是 PE 系统。笔者实验室的研究着重应用于生产与实践, 所以今后能够完成特异性和精准性的碱基编辑方式是实验室的关注重点。

碱基编辑技术发展的目标是在尽可能提高碱基编辑效率和靶向范围的同时, 最大程度地减少脱靶, 使其能够应用到更复杂的研究中。对于绝大多数碱基编辑而言, 目标序列都是固定的, 碱基编辑未来的发展着眼于如何提高碱基编辑的特

异性和精确性。针对 CBE 全基因组的脱靶效应, 需要开发更多的新的 CBE 系统, 在不影响靶位点编辑效率的同时, 尽可能降低全基因组脱靶编辑效率, 提高 BE 系统的安全性。

总之, 碱基编辑技术在生命科学基础研究、人类疾病治疗和生物育种等领域具有广阔的应用前景, 该技术的不断革新将会推动各领域的快速发展。

REFERENCES

- [1] Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*, 2016, 17(6): 333-351.
- [2] Landrum MJ, Lee JM, Benson M, et al. ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*, 2016, 44(D1): D862-D868.
- [3] Landrum MJ, Lee JM, Riley GR, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res*, 2014, 42(D1): D980-D985.
- [4] 田志龙, 王玉琴, 储明星. 绵羊多羔候选基因的研究及应用进展. *中国草食动物科学*, 2018, 38(3): 43-48, 57.
Tian ZL, Wang YQ, Chu MX. Review on prolific candidate genes in sheep (*Ovis aries*). *China Herbivore Science*, 2018, 38(3): 43-48, 57 (in Chinese).
- [5] 宋志芳, 石岗. 利用 GWAS 研究母猪繁殖性状的进展. *猪业科学*, 2019, 36(3): 104-106.
Song ZF, Shi G. Progress on reproductive traits of sows by using GWAS. *Swine Industry Science*, 2019, 36(3): 104-106 (in Chinese).
- [6] Kim JS. Genome editing comes of age. *Nat Protocols*, 2016, 11(9): 1573-1578.
- [7] Bak RO, Natalia GO, Porteus MH. Gene editing on center stage. *Trends Genet*, 2018, 34(8): 600-611.
- [8] Urnov FD. Genome Editing B.C. (Before CRISPR): Lasting LESSONS from the "Old Testament". *CRISPR J*, 2018, 1(1): 34-46.
- [9] Hsu PD, Lander ES, Zhang F. Development and

- applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, 157(6): 1262-1278.
- [10] Komor AC, Badran AH, Liu DR. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell*, 2017, 168(1/2): 20-36.
- [11] Knott GJ, Doudna JA. CRISPR-Cas guides the future of genetic engineering. *Science*, 2018, 361(6405): 866-869.
- [12] Ceccaldi R, Rondinelli B, D'Andrea AD. Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol*, 2016, 26(1): 52-64.
- [13] Chapman JR, Taylor MRG, Boulton SJ. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell*, 2012, 47(4): 497-510.
- [14] Cox DBT, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nat Med*, 2015, 21(2): 121-131.
- [15] Paquet D, Kwart D, Chen A, et al. Efficient introduction of specific homozygous and heterozygous mutations using CRISPR/Cas9. *Nature*, 2016, 533(7601): 125-129.
- [16] Steven L, Staahl BT, Alla RK, et al. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *eLife*, 2014, 3: e04776.
- [17] Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol*, 2004, 4(11): 868-877.
- [18] Yang B, Li XS, Lei LQ, et al. APOBEC: from mutator to editor. *J Genet Genom*, 2017, 44(9): 423-437.
- [19] Komor AC, Kim YB, Packer MS, et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*, 2016, 533(7603): 420-424.
- [20] Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A·T to G·C in genomic DNA without DNA cleavage. *Nature*, 2017, 551(7681): 464-471.
- [21] Nishimasu H, Ran FA, Hsu PD, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, 2014, 156(5): 935-949.
- [22] Harris RS, Petersen-Mahrt SK, Neuberger MS. RNA editing enzyme APOBEC1 and some of its homologs can act as DNA mutators. *Mol Cell*, 2002, 10(5): 1247-1253.
- [23] Kunz C, Saito Y, Schär P. DNA repair in mammalian cells — mismatched repair: variations on a theme. *Cell Mol Life Sci*, 2009, 66(6): 1021-1038.
- [24] Pearl LH. Structure and function in the uracil-DNA glycosylase superfamily. *Mutat Res*, 2000, 460(3/4): 165-181.
- [25] Komor AC, Zhao KT, Packer MS, et al. Improved base excision repair inhibition and bacteriophage Mu Gam protein yields C:G-to-T:A base editors with higher efficiency and product purity. *Sci Adv*, 2017, 3(8): eaao4774.
- [26] Nishida K, Arazoe T, Yachie N, et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*, 2016, 353(6305): aaf8729.
- [27] Yang LH, Briggs AW, Chew WL, et al. Engineering and optimising deaminase fusions for genome editing. *Nat Commun*, 2016, 7: 13330.
- [28] Hess GT, Frésard L, Han K, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. *Nat Methods*, 2016, 13(12): 1036-1042.
- [29] Ma YQ, Zhang JY, Yin WJ, et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells. *Nat Methods*, 2016, 13(12): 1029-1035.
- [30] Krokan HE, Drabløs F, Slupphaug G. Uracil in DNA — occurrence, consequences and repair. *Oncogene*, 2002, 21(58): 8935-8948.
- [31] Lindahl T. Instability and decay of the primary structure of DNA. *Nature*, 1993, 362(6422): 709-715.
- [32] Yasui M, Suenaga E, Koyama N, et al. Miscoding properties of 2'-deoxyinosine, a nitric oxide-derived DNA adduct, during translesion synthesis catalyzed by human DNA polymerases. *J Mol Biol*, 2008, 377(4): 1015-1023.
- [33] Richter MF, Zhao KT, Eton E, et al. Author correction: phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity. *Nat Biotechnol*, 2020, 38(7): 901.

- [34] Lapinaite A, Knott GJ, Palumbo CM, et al. DNA capture by a CRISPR-Cas9-guided adenine base editor. *Science*, 2020, 369(6503): 566-571.
- [35] Zhang H, Pan H, Zhou CY, et al. Simultaneous zygotic inactivation of multiple genes in mouse through CRISPR/Cas9-mediated base editing. *Development*, 2018, 145(20): dev168906.
- [36] Yang H, Wang HY, Shivalila CS, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, 2013, 154(6): 1370-1379.
- [37] Kim K, Ryu SM, Kim ST, et al. Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol*, 2017, 35(5): 435-437.
- [38] Wang LJ, Xue W, Yan L, et al. Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res*, 2017, 27(10): 1289-1292.
- [39] Hua K, Tao XP, Yuan FT, et al. Precise A-T to G-C base editing in the rice genome. *Mol Plant*, 2018, 11(4): 627-630.
- [40] Yan F, Kuang YJ, Ren B, et al. Highly efficient A-T to G-C base editing by Cas9n-guided tRNA adenosine deaminase in rice. *Mol Plant*, 2018, 11(4): 631-634.
- [41] Ryu SM, Koo T, Kim K, et al. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy. *Nat Biotechnol*, 2018, 36(6): 536-539.
- [42] Lau AY, Wyatt MD, Glassner BJ, et al. Molecular basis for discriminating between normal and damaged bases by the human alkyladenine glycosylase, AAG. *Proc Natl Acad Sci USA*, 2000, 97(25): 13573-13578.
- [43] Kouzminova EA, Kuzminov A. Patterns of chromosomal fragmentation due to uracil-DNA incorporation reveal a novel mechanism of replication-dependent double-stranded breaks. *Mol Microbiol*, 2008, 68(1): 202-215.
- [44] Liu ZQ, Chen M, Chen SY, et al. Highly efficient RNA-guided base editing in rabbit. *Nat Commun*, 2018, 9: 2717.
- [45] Yeh WH, Chiang H, Rees HA, et al. *In vivo* base editing of post-mitotic sensory cells. *Nat Commun*, 2018, 9: 2184.
- [46] Koblan LW, Doman JL, Wilson C, et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol*, 2018, 36(9): 843-846.
- [47] Kang BC, Yun JY, Kim ST, et al. Precision genome engineering through adenine base editing in plants. *Nat Plants*, 2018, 4(7): 427-431.
- [48] Kim YB, Komor AC, Levy JM, et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. *Nat Biotechnol*, 2017, 35(4): 371-376.
- [49] Tan JJ, Zhang F, Karcher D, et al. Engineering of high-precision base editors for site-specific single nucleotide replacement. *Nat Commun*, 2019, 10: 439.
- [50] Jiang W, Feng SJ, Huang SS, et al. BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity. *Cell Res*, 2018, 28(8): 855-861.
- [51] Zong Y, Song QN, Li C, et al. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A. *Nat Biotechnol*, 2018, 36(10): 950-953.
- [52] Wang X, Li JN, Wang Y, et al. Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion. *Nat Biotechnol*, 2018, 36(10): 946-949.
- [53] Huang TP, Zhao KT, Miller SM, et al. Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors. *Nat Biotechnol*, 2019, 37(6): 626-631.
- [54] Li XS, Wang Y, Liu YJ, et al. Base editing with a Cpf1-cytidine deaminase fusion. *Nat Biotechnol*, 2018, 36(4): 324-327.
- [55] Hu JH, Miller SM, Geurts MH, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, 2018, 556(7699): 57-63.
- [56] Nishimasu H, Shi X, Ishiguro S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*, 2018, 361(6408): 1259-1262.
- [57] Walton RT, Christie KA, Whittaker MN, et al. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, 2020, 368(6488): 290-296.
- [58] Hua K, Tao XP, Zhu JK. Expanding the base editing scope in rice by using Cas9 variants. *Plant*

- Biotechnol J, 2019, 17(2): 499-504.
- [59] Yang L, Zhang XH, Wang LR, et al. Increasing targeting scope of adenosine base editors in mouse and rat embryos through fusion of TadA deaminase with Cas9 variants. *Protein Cell*, 2018, 9(9): 814-819.
- [60] Gehrke JM, Cervantes O, Clement MK, et al. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities. *Nat Biotechnol*, 2018, 36(10): 977-982.
- [61] Kim D, Kim DE, Lee G, et al. Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat Biotechnol*, 2019, 37(4): 430-435.
- [62] Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, 2016, 529(7598): 490-495.
- [63] Lee JK, Jeong E, Lee J, et al. Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat Commun*, 2018, 9: 3048.
- [64] Liang MM, Sui TT, Liu ZQ, et al. AcrIIA5 suppresses base editors and reduces their off-target effects. *Cells*, 2020, 9(8): 1786.
- [65] Rogozin IB, Iyer LM, Liang LZ, et al. Evolution and diversification of lamprey antigen receptors: evidence for involvement of an AID-APOBEC family cytosine deaminase. *Nat Immunol*, 2007, 8(6): 647-656.
- [66] Burns MB, Lackey L, Carpenter MA, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*, 2013, 494(7437): 366-370.
- [67] Zuo EW, Sun YD, Wei W, et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science*, 2019, 364(6437): 289-292.
- [68] Jin S, Zong Y, Gao Q, et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice. *Science*, 2019, 364(6437): 292-295.
- [69] Jin S, Fei HY, Zhu ZX, et al. Rationally designed APOBEC3B cytosine base editors with improved specificity. *Mol Cell*, 2020, 79(5): 728-740.e6.
- [70] Grünwald J, Zhou RH, Garcia SP, et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors. *Nature*, 2019, 569(7756): 433-437.
- [71] Zhou CY, Sun YD, Yan R, et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis. *Nature*, 2019, 571(7764): 275-278.
- [72] Rees HA, Wilson C, Doman JL, et al. Analysis and minimization of cellular RNA editing by DNA adenine base editors. *Sci Adv*, 2019, 5(5): eaax5717.
- [73] Chadwick AC, Wang X, Musunuru K. *In vivo* base editing of PCSK9 (proprotein convertase subtilisin/kexin type 9) as a therapeutic alternative to genome editing. *Arterioscl Thrombo Vascul Biol*, 2017, 37(9): 1741-1747.
- [74] Zhang YH, Qin W, Lu XC, et al. Programmable base editing of zebrafish genome using a modified CRISPR-Cas9 system. *Nat Commun*, 2017, 8: 118.
- [75] Kuscu C, Parlak M, Tufan T, et al. CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations. *Nat Methods*, 2017, 14(7): 710-712.
- [76] Billon P, Bryant EE, Joseph SA, et al. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons. *Mol Cell*, 2017, 67(6): 1068-1079.e4.
- [77] Haapaniemi E, Botla S, Persson J, et al. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med*, 2018, 24(7): 927-930.
- [78] Ihry RJ, Worringer KA, Salick MR, et al. p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med*, 2018, 24(7): 939-946.
- [79] Aguirre AJ, Meyers RM, Weir BA, et al. Genomic copy number dictates a gene-independent cell response to CRISPR/Cas9 targeting. *Cancer Discov*, 2016, 6(8): 914-929.
- [80] Roukos V, Misteli T. The biogenesis of chromosome translocations. *Nat Cell Biol*, 2014, 16(4): 293-300.
- [81] Gapinske M, Luu A, Winter J, et al. CRISPR-SKIP: programmable gene splicing with single base editors. *Genome Biol*, 2018, 19: 107.
- [82] Yuan JJ, Ma YQ, Huang T, et al. Genetic modulation of RNA splicing with a CRISPR-guided cytidine deaminase. *Mol Cell*, 2018, 72(2): 380-394.e7.

- [83] Li ZX, Xiong XY, Wang FZ, et al. Gene disruption through base editing-induced messenger RNA missplicing in plants. *New Phytol*, 2018, 222(2): 1139-1148.
- [84] Xue CX, Zhang HW, Lin QP, et al. Manipulating mRNA splicing by base editing in plants. *Sci China Life Sci*, 2018, 61(11): 1293-1300.
- [85] Hur JK, Kim K, Been KW, et al. Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat Biotechnol*, 2016, 34(8): 807-808.
- [86] Sung YH, Baek IJ, Kim DH, et al. Knockout mice created by TALEN-mediated gene targeting. *Nat Biotechnol*, 2013, 31(1): 23-24.
- [87] Wang F, Zhang WQ, Yang QY, et al. Generation of a Hutchinson-Gilford progeria syndrome monkey model by base editing. *Prot Cell*, 2020, DOI: 10.1007/s13238-020-00740-8.
- [88] Liu Z, Lu ZY, Yang G, et al. Efficient generation of mouse models of human diseases via ABE- and BE-mediated base editing. *Nat Commun*, 2018, 9: 2338.
- [89] Ma YW, Yu L, Zhang X, et al. Highly efficient and precise base editing by engineered dCas9-guide tRNA adenosine deaminase in rats. *Cell Discov*, 2018, 4: 39.
- [90] Li ZF, Duan XY, An XM, et al. Efficient RNA-guided base editing for disease modeling in pigs. *Cell Discov*, 2018, 4: 64.
- [91] Xie JK, Ge WK, Li N, et al. Efficient base editing for multiple genes and loci in pigs using base editors. *Nat Commun*, 2019, 10: 2852.
- [92] Zhou SW, Cai B, He C, et al. Programmable base editing of the sheep genome revealed no genome-wide off-target mutations. *Front Genet*, 2019, 10: 215.
- [93] Li GW, Zhou SW, Li C, et al. Base pair editing in goat: nonsense codon introgression into *FGF5* results in longer hair. *FEBS J*, 2019, 286(23): 4675-4692.
- [94] Yin KQ, Gao CX, Qiu JL. Progress and prospects in plant genome editing. *Nat Plants*, 2017, 3: 17107.
- [95] Voytas DF, Gao CX. Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol*, 2014, 12(6): e1001877.
- [96] Zong Y, Wang YP, Li C, et al. Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat Biotechnol*, 2017, 35(5): 438-440.
- [97] Shimatani Z, Kashojiya S, Takayama M, et al. Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat Biotechnol*, 2017, 35(5): 441-443.
- [98] Zhang R, Liu JX, Chai ZZ, et al. Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nat Plants*, 2019, 5(5): 480-485.
- [99] Sakata RC, Ishiguro S, Mori H, et al. Base editors for simultaneous introduction of C-to-T and A-to-G mutations. *Nat Biotechnol*, 2020, 38(7): 865-869.
- [100] Kurt IC, Zhou RH, Iyer S, et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat Biotechnol*, 2020, DOI: 10.1038/s41587-020-0609-x.
- [101] Anzalone AV, Randolph PB, Davis JR, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 2019, 576(7785): 149-157.

(本文责编 郝丽芳)