

## 特约评述

DOI: 10.12211/2096-8280.2022-053

## 碱基编辑技术及其在微生物合成生物学中的应用

王雁南, 孙宇辉

(武汉大学药学院, 组合生物合成与新药发现教育部重点实验室, 湖北 武汉 430071)

**摘要:** CRISPR/Cas 系统的发现与发展对生命科学领域产生了革命性的影响, 借助 CRISPR/Cas 系统研发出的一系列工具为相关领域的研究带来了极大的便利。碱基编辑器便是其中一类基于 CRISPR/Cas 系统开发的可实现碱基转换和颠换的基因编辑工具, 其通过将胞苷或腺苷脱氨酶以及其他功能元件与失去双链切割活性的 Cas 蛋白相融合, 由 sgRNA 引导, 实现对基因组上目标位置胞嘧啶或腺嘌呤的碱基转换。碱基编辑器一经开发便在生物学、医学等领域展现出巨大的应用潜力, 虽然经过不断优化, 但目前在使用时仍然存在着许多制约因素。本文简述了几种主要碱基编辑器的发展, 并介绍了碱基编辑器存在的靶向范围受限和脱靶编辑的问题以及现有的优化措施。同时列举了我国部分科研工作者将碱基编辑技术运用于微生物合成生物学领域的进展, 并展望了碱基编辑技术的发展及其在微生物合成生物学领域的应用前景。

**关键词:** 碱基编辑器; 特异性优化; 靶向范围拓展; 微生物; 合成生物学

**中图分类号:** Q812 **文献标志码:** A

## Base editing technology and its application in microbial synthetic biology

WANG Yannan, SUN Yuhui

(Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Ministry of Education), School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, Hubei, China)

**Abstract:** The discovery and development of the CRISPR/Cas system have a revolutionary influence on life sciences. A series of tools derived from the CRISPR/Cas system have brought great convenience to research in the field of life sciences. The base editors developed based on the CRISPR/Cas system are gene editing tools that can achieve base conversions and transversion on target. The base editors are constructed by fusing cytosine or adenosine deaminase, and other functional elements to Cas proteins with abolished double strand DNA cleavage activity to convert cytidine or adenine into other bases at genome on-target sites guided by sgRNAs. Base editors have shown great potential in biology, medicine and related fields. Although they have already been continuously optimized, there are still problems affecting further application of base editors. In this review, we briefly describe

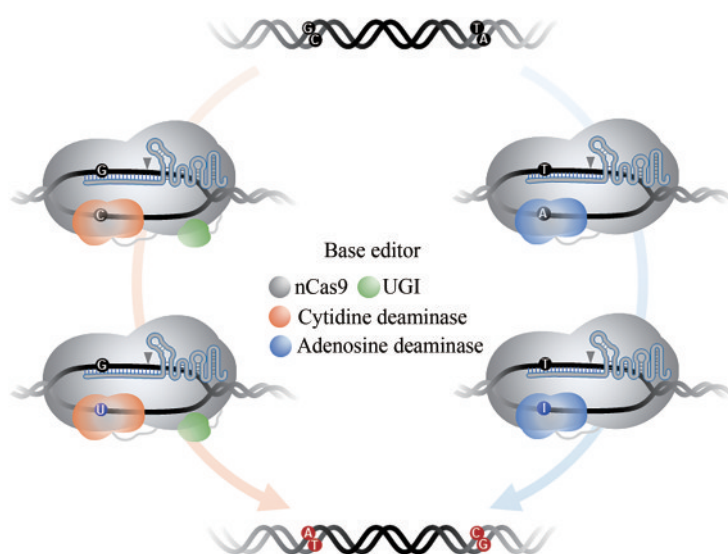
收稿日期: 2022-09-28 修回日期: 2022-11-13

基金项目: 国家重点研发计划合成生物学重点专项 (2018YFA0903200)

引用本文: 王雁南, 孙宇辉. 碱基编辑技术及其在微生物合成生物学中的应用[J]. 合成生物学, 2023, 4(4): 720-737

Citation: WANG Yannan, SUN Yuhui. Base editing technology and its application in microbial synthetic biology [J]. Synthetic Biology Journal, 2023, 4(4): 720-737

the development of DNA base editors. Furthermore, we introduce in detail the problems of the limited editing range of base editors as well as the corresponding optimization strategies by increasing the target sites recognized by the locator moiety and expanding or narrowing the editing window of the effector moiety. At the same time, we introduce several off-target editing detection methods specially developed for base editing. Based on usual and developed detection methods, multiple and frequent off-target editing caused by base editors were found at both DNA and RNA levels. We also introduced various effective optimization strategies to improve the editing specificity of the base editors in every respect. Most of these strategies are based on protein modification, but also on optimization of sgRNA and spatio-temporal regulation of base editing systems. These measures greatly enrich the application scenarios of the base editors. Then, we discuss the progress on applying base editors to the field of microbial synthetic biology, including revealing the metabolic pathway and synthesis mechanism of natural products as well as improving the production of target compounds in multiple species. Finally, we envisage the promising development of base editing in synthetic biology in the future.



**Keywords:** base editors; improvement of specificity; expansion of target range; microorganism; synthetic biology

自DNA双螺旋结构发现和重组DNA技术发明以来，人们对于如何高效地实现基因的精准分析和定向改造的探索就不曾停止过。基于CRISPR/Cas系统开发的一系列基因编辑工具，因其灵活、高效、普适的特点得到了非常广泛的应用，2020年诺贝尔化学奖授予对CRISPR/Cas系统的发现和发展做出重要贡献的Emmanuelle Charpentier和Jennifer Doudna，更使CRISPR/Cas系统及其相关技术受到了人们的瞩目和厚望。

原核生物的CRISPR/Cas系统是一类可遗传的适应性免疫系统，它可以存储原核生物曾被感染

的信息，当再次被感染时，可利用RNA引导的核酸酶破坏噬菌体遗传物质，从而发挥防御或免疫功能。人们将该系统中成簇规律间隔的短回文重复序列命名为CRISPR (clustered regularly interspaced short palindromic repeats) 序列，而将一些邻近CRISPR序列且具有特征性的相关基因命名为Cas (CRISPR-associated gene)<sup>[1]</sup>。CRISPR/Cas系统被用作基因编辑工具时，通常由单链引导RNA (single guide RNA, sgRNA) 引导Cas蛋白结合在目标位点并产生双链断裂 (double strand break, DSB)，随后经细胞的同源修复 (homology-

directed repair, HDR) 介导精准编辑或非同源末端连接修复 (nonhomologous end-joining, NHEJ) 引入插入和缺失 (insertions and deletions, Indels) 等突变, 从而实现对目标序列的靶向编辑<sup>[2-3]</sup>。尽管 CRISPR/Cas 系统相比于以往基因编辑技术, 已经展现出巨大优势和成功, 但人们发现 CRISPR/Cas 系统在使用时仍存在不少问题, 例如在对人类遗传疾病中最常见的遗传变异——点突变<sup>[4]</sup> 进行编辑修复或是构建遗传疾病模型时, 依靠 HDR 进行修复的策略编辑效率较低, 并且在应用时还存在一系列的限定条件。此外, Cas 蛋白可能会在脱靶位点产生双链断裂, 甚至目标位点 DNA 修复途径的活化都可能对细胞产生不利影响<sup>[5-6]</sup>。因此, 一种无需引入双链断裂即可实现碱基转换的基因编辑工具——碱基编辑器 (base editor, BE) 应运而生<sup>[7]</sup>。该工具通过将胞苷或腺苷脱氨酶以及其他功能元件与失去双链切割活性的 Cas 蛋白相融合, 由 sgRNA 引导, 实现对基因组上目标位置的胞嘧啶或腺嘌呤的碱基转换。碱基编辑技术一经问世, 便在生物学、医学等领域展现出巨大的应用潜力。

合成生物学是近些年发展日益蓬勃的交叉学科, 它将工程原理应用于对生命系统的研究和改造, 以期对细胞行为进行创建、编程和控制。合成生物学在过去数十年中, 在化工、能源、材料、农业、环境和医药等领域已展露出巨大的生命力。随着合成生物学的发展, 越来越多的化合物可以通过改造培养微生物获得, 人们也在不断追求高效精准地改造微生物的能力。而 CRISPR/Cas 系统的人工靶向可设计性等优点受到了合成生物学领域研究者的青睐, 基于 CRISPR/Cas 系统开发的碱基编辑技术在保留上述 CRISPR/Cas 系统优点的同时, 避免了对目标位点 DNA 造成双链断裂, 降低了基因编辑产生的细胞毒性。因此, 本文对目前已开发的几类主要的 DNA 碱基编辑器的发展和工作原理进行了阐述, 并重点介绍了碱基编辑器靶向范围受限和拓展措施, 以及造成的脱靶编辑的检测和优化两方面内容。同时, 本文还介绍了我国部分科研工作者在微生物合成生物学领域运用碱基编辑技术的一些应用, 并展望了碱基编辑技术的发展及其在微生物合成生物学领域的应用前景。

## 1 碱基编辑技术概述

### 1.1 碱基编辑器的发展

#### 1.1.1 胞嘧啶碱基编辑器和腺嘌呤碱基编辑器的发展

DNA 碱基编辑器最初是由美国哈佛大学 David R. Liu 研究团队于 2016 年开发的一种基因编辑工具<sup>[7]</sup>, 其开发的胞嘧啶碱基编辑器 (cytosine base editor, CBE) 可以实现碱基对 C-G 到 T-A 的转换。他们将大鼠来源的胞苷脱氨酶 APOBEC1 通过一段 XTEN linker 连接在 dCas9 (含 D10A 和 H840A 突变) 的 N 端, 得到的第一代碱基编辑器 BE1 通过 sgRNA 引导结合在目标位点, 由 dCas9 打开双链 DNA 形成 DNA: RNA R-loop 结构。胞苷脱氨酶对暴露出的单链 DNA 上局部的胞嘧啶脱氨, 形成尿嘧啶。U•G 碱基对随后经过体内 DNA 修复或复制依次形成 U•A、T-A 碱基对, 即在 DNA 非靶标链上 4~8 号位的窗口内 (将 NGG PAM 所在位置定义为 21~23, 下同) 实现了 C 到 T 的转换。较低的编辑效率可能是由于体内的尿嘧啶 DNA 糖基化酶 (uracil DNA glycosylase, UDG) 切除了脱氨形成的 U, 使得 DNA 被修复回 C-G 碱基对。于是该团队将抑制 UDG 活性的尿嘧啶 DNA 糖基化酶抑制剂 (uracil DNA glycosylase inhibitor, UGI) 连接于 BE1 中 dCas9 的 C 端, 得到的第二代碱基编辑器 BE2 编辑活性有了明显的提升。而第三代碱基编辑器 BE3 则是将 BE2 中的 dCas9 替换为 nCas9 (D10A), 通过在 DNA 靶标链上产生切刻, 使 DNA 以发生脱氨后的 DNA 非靶标链为模板进行修复, 进一步提升了碱基编辑的效率, 而且仅产生了 1.1% 的 Indels 突变 (表 1)。

随后, David R. Liu 研究团队<sup>[30]</sup> 通过更换 BE3 各功能元件之间的 linker, 并在 UGI 的 C 端再连接一个 UGI, 得到的第四代碱基编辑器 BE4 进一步提高了碱基编辑的效率。在 BE4 的 N 端连接来自噬菌体 Mu 的 Gam 蛋白, 利用 Gam 蛋白结合双链断裂的 DNA 末端, 使 DNA 免于降解, BE4-Gam 相比 BE4 降低了 Indels 突变的产生频率, 并且 BE4 和 BE4-Gam 均提高了 C 转换为 T 的产物纯度。此外, 他们还通过在 BE4 两端加上核定位信号, 并对

表1 代表性碱基编辑器比较

Table 1 Comparison of representative base editors

碱基编辑器	版本	组成	PAM (5'-3')	特点	参考文献
CBE	BE3	rAPOBEC1- <i>SpnCas9</i> -UGI	NGG	初代碱基编辑器, TC 基序偏好性	[7]
	Target-AID	<i>SpnCas9</i> -CDA1-UGI	NGG	初代碱基编辑器	[8]
	hA3A-BE3	hAPOBEC3A- <i>SpnCas9</i> -UGI	NGG	可编辑甲基化的 C 及 GC 基序的 C	[9]
	BE4-Max	rAPOBEC1- <i>SpnCas9</i> -UGI-UGI	NGG	密码子优化, 高活性	[10]
	dCpf1-BE	rAPOBEC1-d <i>LbCpf1</i> -UGI	TTTV	识别富含 T 的 PAM	[11]
	Target-AID-NG	<i>SpnCas9</i> (NG)-CDA1-UGI	NG	识别 NG PAM	[12]
	A3A-PBE	hAPOBEC3A- <i>SpnCas9</i> -UGI	NGG	1~17 号位编辑窗口, 几乎无基序偏好性	[13]
	BE-PLUS	SunTag- <i>SpnCas9</i> -scFv-rAPOBEC1-UGI	NGG	4~16 号位编辑窗口, 高保真	[14]
	A3G-BEs	hAPOBEC3G- <i>SpnCas9</i> -UGI-UGI	NGG	编辑 CC 基序中的第二个 C	[15]
	BE-PAPAPAP	rAPOBEC1- <i>SpnCas9</i> -UGI	NGG	5~6 号位编辑窗口, PAPAPAP 替换 XTEN linker	[16]
	HF-BE3	rAPOBEC1-H <i>FnCas9</i> -UGI	NGG	减少 sgRNA 依赖的脱靶编辑	[17]
	YE1-BE4	rAPOBEC1 (W90Y, R126E)- <i>SpnCas9</i> -UGI-UGI	NGG	缩小编辑窗口, 减少非 sgRNA 依赖的脱靶编辑	[18]
	SECURE BE3	rAPOBEC1 (R33A, K34A)- <i>SpnCas9</i> -UGI	NGG	减少 RNA 层面的脱靶编辑, 缩小编辑窗口	[19]
	DdCBEs	TALE-split DddAtox-UGI	None	同时编辑 DNA 双链, 可编辑线粒体、叶绿体基因组	[20]
	ZFDs	ZFP-split DddAtox-UGI	None	同时编辑 DNA 双链, 可编辑线粒体、叶绿体基因组	[21]
	ABE	ABE7.10	ecTadA-ecTadA*- <i>SpnCas9</i>	NGG	初代碱基编辑器
ABEmax		ecTadA-ecTadA*- <i>SpnCas9</i>	NGG	密码子优化, 高活性	[10]
ABE8e		ecTadA-ecTadA*- <i>SpnCas9</i>	NGG	高活性, TadA-8e	[23]
TaC9-ABE		ecTadA-ecTadA*-TALE, <i>SpnCas9</i>	NGG	消除 sgRNA 依赖的脱靶编辑	[24]
ABE-n <i>SpCas9</i> -DS		ecTadA*-ecTadA*- <i>SpnCas9</i> (DS)	NGG	拓宽编辑窗口, 减少 RNA 层面的脱靶编辑	[25]
ACBE	A&C-BEmax	rAPOBEC1/hAID-ecTadA-ecTadA*- <i>SpnCas9</i> -UGI-UGI	NGG	可同时编辑 A 或 C	[26]
	sgBE	<i>SpnCas9</i> , MCP-cytosine/adenosine deaminase	NGG	可同时编辑 A 或 C	[27]
GBE	CGBE1	eUNG-rAPOBEC1 (R33A)- <i>SpnCas9</i> -UGI-UGI	NGG	可实现 C 变 G 的碱基颠换	[28]
	CGBE	rAPOBEC1- <i>SpnCas9</i> -rXRCC1	NGG	利用碱基切除修复通路实现 C 变 G 的碱基颠换	[29]

BE4 进行密码子优化及祖先序列重构 (ancestral reconstruction), 由此获得的 BE4max 和 AncBE4max 成为目前编辑效率最高的 CBE 之一<sup>[10]</sup>。巧合的是, 同在 2016 年, 日本神户大学 Akihiko Kondo 研究团队<sup>[8]</sup>也报道了他们研发的 Target-AID (activation-induced cytidine deaminase) 胞嘧啶碱基编辑器, 由 nCas9 (D10A) 的 C 端连接七鳃鳗来源的胞嘧啶脱氨酶 (*Petromyzon marinus* cytidine deaminase 1, *PmCDA1*) 和 UGI 组成, 其编辑窗口相比于 BE3 更靠近 PAM 远端的 2~4 号位。

2017 年, David R. Liu 研究团队<sup>[22]</sup>进一步报道了可以实现碱基对 A-T 到 G-C 转换的腺嘌呤碱基

编辑器 (adenine base editor, ABE)。与 CBE 不同, 自然界未发现以 DNA 为底物的腺苷脱氨酶, 于是该研究团队通过定向进化和蛋白质工程使大肠杆菌 tRNA 腺苷脱氨酶 (*Escherichia coli* tRNA adenosine deaminase, *ecTadA*) 克服了这一限制, 利用二聚的 TadA 与 nCas9 (D10A) 相连接, 并通过一系列优化, 获得了腺嘌呤碱基编辑器 ABE7.10 (wtTadA-TadA\*-nCas9, \*表示蛋白变体)。ABE7.10 在 sgRNA 引导下结合在目标位点, 由 nCas9 打开双链 DNA, 二聚的 TadA 对暴露的单链 DNA 上局部的腺嘌呤脱氨, 形成次黄嘌呤 I, 经过体内的错配修复和复制过程, I•T 碱基对依次转换为 I•C,

G-C碱基对, 实现在DNA非靶标链上4~7号位编辑窗口内A到G的转换, 产物纯度极高, 且该过程中Indels突变的发生频率几乎可以忽略不计。之后, David R. Liu研究团队<sup>[10]</sup>一直在设法提高ABE的编辑效率, 通过在ABE7.10两端添加核定位信号和密码子优化, 获得了活性更强的ABEmax。进一步通过对TadA-7.10使用噬菌体辅助进化方法进行优化, 获得的ABE8e催化速率相比ABE7.10提高了590倍, 显示出编辑活性以及与各种Cas蛋白变体和同系物兼容性的大幅提升, 由此带来更多的脱靶编辑也可以通过在TadA-8e中引入V106W突变来加以改善<sup>[23]</sup>。

### 1.1.2 其他碱基编辑器的发展

随着CBE和ABE相继问世, 越来越多的碱基编辑需求寄望于依靠碱基编辑器来实现, 但是CBE和ABE仅能实现在目标位点C-G到T-A或A-T到G-C的碱基对转换, 无法实现其他碱基之间的转变。于是人们自然想到通过将胞苷脱氨酶和腺苷脱氨酶一起融合在nCas9的同侧或两侧, 获得的ACBE (fusion adenine and cytosine base editor)可以同时实现目标位点C-G到T-A和A-T到G-C碱基对的转换<sup>[26, 31-35]</sup>。人们将融合了胞苷脱氨酶和腺苷脱氨酶的版本ACBE在哺乳动物细胞和植物细胞中进行了测试, 并对系统中的其他因素, 例如蛋白之间的linker及sgRNA的长度进行了系统的考察, 发现这些ACBE都具有较低的脱靶活性。除了将两个不同的脱氨酶与nCas9融合的策略, 也有研究团队利用普通版本的CBE或ABE并通过在sgRNA上引入MS2 RNA适配体招募另一种脱氨酶的策略, 也可以实现ACBE的相似功能<sup>[27]</sup>。这些ACBE丰富了目标位点碱基转换的产物种类, 在遗传疾病的治疗、细胞谱系追踪以及基因随机突变领域都具有巨大的应用前景<sup>[26, 31-35]</sup>。

在碱基编辑器的研究和应用过程中, 人们发现CBE和ABE不仅仅介导C到T或A到G的碱基转换, 还可能会引起C到G或C到A的碱基颠换<sup>[36-38]</sup>。通过分析, 发现这一现象可能由于C脱氨后形成的U被UDG切除形成AP位点 (apurinic/aprimidinic site) 后, 经历细胞内的碱基切除修复过程而导致。于是人们利用UDG替代CBE中的UGI, 促进细胞切除C脱氨形成的U, 提高了C碱

基颠换的效率。由此开发的GBE (glycosylase base editors) 和CGBE (C-to-G base editors) 可以实现哺乳动物细胞内C到G<sup>[28, 39]</sup>, 以及大肠杆菌中C到A的碱基颠换<sup>[39]</sup>。此外, 有研究利用参与碱基切除修复途径的XRCC1蛋白, 以及改善影响C到G转变过程的其他因素, 开发了不同版本的CGBE<sup>[29, 40-41]</sup>。虽然相继有这一类碱基编辑器的开发被报道, 但是细胞内C到G的转变过程并未揭示清晰, 这也使得该碱基编辑器在研究和使用过程中遇到的一些问题还难以解答, 如GBE在哺乳动物细胞和大肠杆菌中产生了不同的编辑结果<sup>[39]</sup>。不过随着该过程的机制被逐渐阐明, 更多更高效的CGBE版本有望被开发出来。

此外, 人们还开发了可以同时编辑DNA双链的碱基编辑器: David R. Liu研究团队通过将能对DNA双链中的胞嘧啶进行脱氨的细菌间毒素DddA蛋白拆分成两个部分, 并分别与靶向定位作用的转录激活因子样效应物 (transcription activator-like effector, TALE) 或者不同的Cas蛋白以及UGI融合, 可以实现当两个定位蛋白在DNA上结合位置相距较近时, 聚合的DddA可以对两者之间的DNA双链上的C同时进行脱氨。而其他情况下, 两个拆分的DddA无法在体内聚合, 大大降低了该编辑工具的细胞毒性<sup>[20]</sup>。其中, 将拆分的DddA与两个TALE融合得到的碱基编辑器DdCBEs (RNA-free DddA-derived cytosine base editors), 因无需RNA引导便可发挥作用, 使其可以对无法通过CRISPR/Cas系统进行编辑的线粒体DNA和叶绿体DNA实现有效的编辑<sup>[42]</sup>。但是, 人们通过全基因组测序方法检测到DdCBE会在细胞核内造成大量的DNA脱靶编辑<sup>[43]</sup>。而将分裂的DddA与两个不同的Cas蛋白 [如dSpCas9和SaKKH-Cas9 (D10A)] 融合之后形成的编辑器, 由于系统过于繁复, 反而使得其应用变得有限。韩国首尔国立大学Jin-Soo Kim研究团队<sup>[21]</sup>也发表了类似的工作, 他们开发的锌指脱氨酶 (zinc finger deaminase, ZFD) 将分裂的DddA与锌指DNA结合蛋白 (zinc-finger DNA binding protein) 相融合, 也能在核DNA和线粒体DNA上实现类似的编辑。之后, Jin-Soo Kim研究团队<sup>[44]</sup>在DdCBE的基础上结合腺苷脱氨酶TadA, 经历一系列优化改造, 开发出了可以在线

粒体基因组实现A到G转化的碱基编辑器TALED (TALE-linked deaminase)。

### 1.1.3 RNA碱基编辑器的发展

除了DNA碱基编辑器，人们也开发了一系列RNA碱基编辑器来避免DNA编辑所带来的遗传信息不可逆改变以及安全与伦理问题。RNA碱基编辑使用的脱氨酶大多来自ADAR (adenosine deaminases that act on RNA) 家族的腺苷脱氨酶，ADAR包含独特的RNA结合基序，识别定位双链RNA的部分区域，ADAR脱氨酶域 (ADAR deaminase domain, ADAR<sub>DD</sub>) 将腺苷脱去氨基，实现A到I的转变<sup>[45-46]</sup>。人们利用引入发夹结构或者化学修饰的反义RNA与目标RNA互补，供融合了其他蛋白结合域或标签的ADAR<sub>DD</sub>结合<sup>[47-53]</sup>，并将需要编辑的腺苷对侧的反义RNA设计为非互补的胞苷，提高更倾向在A•C错配处脱氨的ADAR在目标腺苷处的编辑效率<sup>[50]</sup>，这一策略被应用于大多数RNA碱基编辑器的设计中。而北京大学魏文胜研究团队<sup>[54]</sup>通过设计ADAR招募RNA (ADAR-recruiting RNA, arRNA) 来招募细胞内天然的ADAR1或ADAR2蛋白，实现目标RNA从A到I的转换。美国哈佛大学张锋研究团队将ADAR2<sub>DD</sub>与dPspCas13b融合，通过RNA引导至目标位点，开发了新的RNA碱基编辑工具REPAIR (RNA Editing for Programmable A-to-I Replacement)。REPAIR对报告基因转录产物的编辑效率高于之前通过反义RNA实现的RNA碱基编辑，而在细胞内源全转录本的编辑中，其效率也与之前的技术相比更好或持平。为了提升REPAIR的编辑特异性，张锋研究团队在ADAR2<sub>DD</sub>引入了E488Q和T375G两个点突变，开发出了REPAIRv2，大幅减少了全转录组水平和目标腺苷上下游近端的脱靶编辑，但也一定程度上降低了目标位点的编辑效率<sup>[55]</sup>。他们之后通过对ADAR2<sub>DD</sub>进行定向进化，开发出了RESCUE (RNA Editing for Specific C to U Exchange) 系统，可以实现目标位点RNA碱基从C到U的转换，同时保留了A到I转换的活性，仅需通过在目标核苷酸对侧的gRNA中设计不同的错配核苷酸即可实现差异化的编辑<sup>[56]</sup>。除了ADAR家族的脱氨酶，上海科技大学池天研究团队和中国科学院植物生理生态研究所李轩研究团队基于

APOBEC3A的RNA的催化活性，构建出可实现C到U的RNA碱基编辑器，其中郝沛研究团队还基于晶体结构开发了只具有RNA催化功能的APOBEC3A (A3A) 变体<sup>[57-58]</sup>。

## 1.2 碱基编辑器编辑范围的拓展

碱基编辑工具在不引入DNA双链断裂的情况下，实现了高效率的碱基转换 (包括颠换)，其优异的性能使人们希望碱基编辑器可以适用于更多位点的编辑。碱基编辑器可编辑的位点受限于Cas蛋白的定位以及Cas蛋白打开DNA双链之后形成的R-loop与脱氨酶相互作用。人们从这两方面入手，开发了一系列具有不同靶向位点和编辑窗口的碱基编辑器。

### 1.2.1 碱基编辑器靶向位点的拓展

David R. Liu研究团队<sup>[59]</sup>通过将rAPOBEC1与Cas9的工程变体融合，获得了可以靶向非NGG PAM的碱基编辑器SaBE3、Sa(KKH)-BE3、VQR-BE3、VRER-BE3和EQR-BE3。2018年，该研究团队又开发了可识别NG、GAA和GAT PAM的xCas9，并考察了它与胞苷脱氨酶和腺苷脱氨酶融合后的编辑效果，发现xCas9与BE3、ABE7.10 (指除了Cas9蛋白的其余部分) 有很好的兼容性，可以靶向它们之前无法编辑的位点，且xCas9 (3.7)-ABE的编辑效率甚至高于ABE7.10<sup>[60]</sup>。之后，他们通过定向进化获得了可以识别NRRH、NRTH和NRCH PAM的Cas9变体，结合脱氨酶后，使碱基编辑器可以对PAM为NR的位点进行编辑<sup>[61]</sup>。SpCas9-NG和SpRY也被用于对非NGG PAM的位点进行碱基编辑，研究发现nSpCas9-NG-AID (Target-AID-NG) 的编辑效率要优于xCas9-BE4<sup>[12]</sup>；而融合SpRY的碱基编辑器几乎可以独立于PAM靶向目标位置<sup>[62]</sup>。此外，人们还开发了基于ScCas9和SmacCas9的碱基编辑器，分别识别NNG和NAAA序列的PAM<sup>[63-65]</sup>。除了利用Cas9蛋白变体拓展靶向位点，Cas12a蛋白也被用于脱氨酶的定位。上海科技大学陈佳研究团队<sup>[11]</sup>通过将rAPOBEC1与dLbCpf1 (dLbCas12a) 融合，获得了可以识别富含胸腺嘧啶核苷酸PAM (TTTV) 的dCpf1-BE。美国麻省总医院J. Keith Joung研究团

队<sup>[66]</sup>通过结构改造获得了可以识别非TTTV PAM的*AsCas12a*蛋白变体*enAsCas12a*,其可以引导胞苷脱氨酶(*enAsBE*)在非TTTV PAM的靶点进行碱基编辑。然而,与这些融合Cas9变体的CBE不同,使用Cas9蛋白变体构建的ABE与经典的ABE版本(ABE7.10)相比,大部分的碱基编辑效率较低,可能是因为*TadA*蛋白最初是基于与*SpCas9*融合后共同进化而来。此外,美国Beam Therapeutics公司Giuseppe Ciaramella研究团队<sup>[67]</sup>也通过基于ABE7.10的进化获得了*TadA\*8*,其融合*SpCas9-NG*和*SaCas9*展现了较高的编辑活性。

### 1.2.2 碱基编辑器编辑窗口的拓展

通过更换一些碱基编辑器的Cas蛋白,也可以拓宽碱基编辑器的编辑窗口。如*SaABEmax*和*SaKKH-ABEmax*编辑器,相较ABEmax(*SpCas9*)将编辑窗口扩大为4~14号位<sup>[68]</sup>。David R. Liu研究团队<sup>[69]</sup>通过将脱氨酶与不同方式循环排列的*SpCas9*(Cas9 circular permutants, Cas9-CPs)相融合,试图使脱氨酶更容易接触到单链DNA底物。融合获得的两个CP-CBEmax保持了与CBEmax一致的编辑活性,而将编辑窗口拓展为4~11号位,并且同时提高了C到T的产物纯度。这一结果印证了循环排列的*SpCas9*的新末端使融合的脱氨酶和UGI更容易接触到单链DNA底物,从而扩大编辑窗口并阻止UDG切除U的推测。而融合后的CP-ABEmax将编辑窗口扩大到了4~12号位,还有一例虽未扩大编辑窗口,但是显示出编辑窗口的平移,并且这些CP-CBEmax和CP-ABEmax造成的Indels突变比例均有所下降<sup>[68]</sup>。

除了可以通过更换Cas蛋白靶向更多的位点,也可以通过更换脱氨酶拓展碱基编辑器的编辑窗口,并改变碱基编辑的序列偏好性。中国科学院遗传与发育生物学研究所高彩霞研究团队<sup>[13]</sup>通过用APOBEC3A替换APOBEC1,将碱基编辑器的编辑窗口扩大到1~17号位,并且有效地解决了APOBEC1偏好编辑T下游的C(5'-TC-3'基序偏好性),而对G下游的C(GC基序)编辑效率低的问题,A3A-PBE碱基编辑器对C上下游序列几乎没有选择性。融合APOBEC3A的碱基编辑器也弥补了APOBEC1脱氨酶无法对甲基化的C进行编辑的缺陷<sup>[9]</sup>。David R. Liu研究团队<sup>[70]</sup>通过用CDA1替

换APOBEC1,并经过密码子优化构建了CDA1-BE4max,将编辑窗口扩大到1~9号位。将其进一步定向进化获得的evoCDA1-BE4max将编辑窗口扩大到1~13号位,其与用类似方式优化获得的另外两个碱基编辑器evoAPOBEC1-BE4max和evoFERNY-BE4max均对GC基序中的C展现良好的编辑活性。美国莱斯大学高雪研究团队<sup>[15]</sup>通过利用仅保留C端结构域的APOBEC3G(A3G)替换APOBEC1,并通过蛋白质工程改造,获得了A3G-BE4.4、A3G-BE5.13和A3G-BE5.14三个版本的碱基编辑器,将编辑窗口扩大到约4~15号位,并且在编辑底物为CC基序的情况下只编辑第二个C,大幅减少了对第一个C的编辑。上海科技大学黄行许研究团队<sup>[71]</sup>通过优化人A3G蛋白,开发出的oA3G-BE3对CC和CCC基序中下划线所在位置的C的编辑相比BE3更具选择性。吉林大学李占军研究团队<sup>[72]</sup>通过在AID脱氨酶中引入突变,融合在nCas9(D10A)的N端,获得的碱基编辑器eAID-BE4max将编辑窗口拓展为1~11号位。中国科学院脑科学与智能技术卓越创新中心仇子龙研究团队<sup>[73]</sup>通过融合nCas9(D10A)及一系列七鳃鳗胞苷脱氨酶开发了一系列编辑窗口位于PAM远端、近端以及覆盖整个Protospacer区域的碱基编辑器。

也有人通过改变脱氨酶与DNA的相互作用来拓展编辑的范围。黄行许研究团队开发了BE-PLUS编辑器,利用了包含多个GCN4多肽的SunTaq信号放大系统<sup>[74]</sup>,通过将SunTaq标签融合在nCas9(D10A)的N端,招募单独表达的scFV-APOBEC-UGI-GB1(scFV抗体特异性识别GCN4多肽,GB1为促溶标签,消除蛋白质聚集),可以将编辑窗口拓宽到4~16号位<sup>[14]</sup>。华东师范大学李大力研究团队通过在BE4max、A3A-BE4max及eA3A-BE4max<sup>[75]</sup>三个碱基编辑器的脱氨酶与Cas蛋白之间融合Rad51蛋白的单链DNA结合域(single-stranded DNA-binding domain, ssDBD),获得的hyBE4max、hyA3A-BE4max和heyA3A-BE4max三个碱基编辑器相比之前的版本大幅拓宽了编辑窗口<sup>[76]</sup>。四川大学姚少华研究团队<sup>[77]</sup>将胞苷脱氨酶嵌合在Cas9蛋白的PAM相互作用域(PAM-interacting domain, PI domain)内,获得了

编辑窗口拓展为4~14号位的碱基编辑器 BE-PIGS。复旦大学程天林研究团队<sup>[25]</sup>通过将腺苷脱氨酶嵌合于Cas9内部的特定位点,获得了多个版本的 ABE-nSpCas9-DS 碱基编辑器,这一系列的碱基编辑器将 ABE 编辑窗口拓宽为2~16号位,并且降低了RNA脱靶编辑的风险。瑞士苏黎世联邦理工学院 Gerald Schwank 研究团队<sup>[78]</sup>通过用 TadA 替换 Cas9 的 HNH 结构域,缩小了整个编辑器的大小,并获得了编辑窗口靠近 PAM 的碱基编辑器。

此外, Jin-Soo Kim 研究团队<sup>[79]</sup>通过在 sgRNA 5'端延长2~3 nt,将 ABE7.10 的编辑窗口向 PAM 远端延伸2~3 nt。类似编辑窗口延伸的现象也在 dCas-CDA-UL、nCas9 (D10A)-AID 和 CRISPR-CDA-nCas9-UGI 等碱基编辑器结合5'端延长的 sgRNA 时被发现<sup>[80-82]</sup>。

除了拓宽编辑窗口,许多的碱基编辑场景需要碱基编辑器拥有更小的编辑窗口从而产生更精准的编辑。David R. Liu 研究团队通过引入突变降低 rAPOBEC1 的编辑活性及底物结合能力,获得了四种 (YE1-BE3、EE-BE3、YE2-BE3 和 YEE-BE3) 相比于 BE3 略微降低了编辑活性但是将编辑窗口范围缩小到1~2 nt 的碱基编辑器,而通过使用截短的 sgRNA 未能很好地实现这一点<sup>[59]</sup>。之后李占军研究团队<sup>[83]</sup>在 YE1 的基础上构建了 YFE-BE4max, 同样将编辑窗口缩小到3 nt 并在兔子中实现高效率、低脱靶的编辑。德国马普分子植物生理所 Ralph Bock 研究团队<sup>[16, 84]</sup>通过用5~7个氨基酸的刚性 linker (PAPAP/PAPAPAP) 替换 BE3 中原有的 XTEN linker, 将编辑窗口缩小到了5~7号位,而在 nSpCas9 的 N 端融合截短 C 端的 CDA1 或 A3A, 分别使得碱基编辑几乎只发生在3号位或5~6号位。J. Keith Joung 研究团队<sup>[19]</sup>为了减少碱基编辑器引发的 RNA 随机突变,开发了两个在脱氨酶中引入突变的碱基编辑器变体——BE3-R33A 和 BE3-R33A/K34A, 其在大幅降低 RNA 随机突变频率的同时,也将编辑窗口缩小到了5~7号位。中国科学院天津工业生物技术研究所张学礼研究团队<sup>[85]</sup>通过设计非完全匹配的 sgRNA 增加了碱基编辑器对目标序列单个碱基的编辑比例,并且提高了碱基编辑的效率。

## 1.3 碱基编辑器编辑特异性的优化

### 1.3.1 碱基编辑器脱靶编辑的检测

在开发和应用单碱基编辑器的过程中,人们发现单碱基编辑器可以在目标位点及脱靶位点产生多种不需要的编辑副产物。因此,研究者针对性地开发出了一系列具有更高特异性的单碱基编辑器,以改善这些问题。

通用的编辑产物检测方法包括对目标位点,以及对通过软件预测或文献已发表的脱靶位点进行桑格测序 (Sanger sequencing) 或深度测序 (deep sequencing), 从而进行定点考察<sup>[7]</sup>, 或是通过全基因组测序 (whole-genome sequencing, WGS) 了解在全基因组的脱靶情况<sup>[80]</sup>, 这些方法都具有各自的局限性。

Jin-Soo Kim 研究团队<sup>[86]</sup>最早在 CBE 脱靶编辑的全基因组检测方面取得突破,他们使用 BE3ΔUGI 在体外对编辑位点脱氨并在 DNA 靶标链上产生切刻,之后利用尿嘧啶特异性切除试剂,即一种尿嘧啶 DNA 糖基化酶及 DNA 糖基化酶-裂解酶内切酶 VIII 的混合物,将 C 脱氨形成的 U 所在位置的 DNA 链切断,形成双链断裂。之后, DNA 样本在末端修复和连接后通过全基因组测序,与参考基因组序列或未经编辑处理的对照基因组测序序列比对,寻找基因组上 BE3ΔUGI 的编辑位点,该方法被称为 Digenome-seq (digested-genome sequencing)。随后,中山大学松阳洲研究团队和 Jin-Soo Kim 研究团队<sup>[87-88]</sup>分别开发了针对 ABE 脱靶编辑全基因组检测的方法,通过类似的策略,先用 ABE7.10 在体外对基因组 DNA 编辑并产生切刻,之后用内切酶 V 对 DNA 再次进行处理,切断脱氨形成的肌苷 3'下游的第二个磷酸二酯键,从而产生双链断裂,经过 DNA 末端修复后通过全基因组测序寻找编辑位点。但是这种检测 ABE 脱靶位点的方法会遗漏 ABE 导致的 C 到其他碱基的转变,并且以上几种方法都是检测体外编辑的产物,并不能客观地展示体内真实发生的编辑,于是高彩霞研究团队和中国科学院神经科学研究所杨辉研究团队分别对水稻和小鼠体内的碱基编辑进行了全基因组分析:高彩霞研究团队<sup>[89]</sup>分别用三种不同的碱基编辑器编辑水稻,并设置只转化碱基编

编辑器而不转化 sgRNA 的对照, 之后对编辑植株及野生型植株基因组 DNA 进行测序, 用野生型植株的测序结果过滤背景突变, 便可观测到碱基编辑器所引发的突变; 杨辉研究团队<sup>[90]</sup> 在小鼠受精卵分裂为两个细胞的时候, 对其中一个细胞注射碱基编辑的相关元件以及荧光蛋白基因, 在胚胎成长后的第 14.5 天, 对胚胎进行消化, 通过流式细胞仪对基于荧光蛋白表达分选出发生编辑和未发生编辑的细胞, 并对二者进行全基因组测序。对比测序结果, 过滤掉未编辑细胞中的背景突变, 便可得到发生编辑细胞的突变情况, 这种方法被命名为 GOTI (genome-wide off-target analysis by two-cell embryo injection)。David R. Liu 研究团队和美国 Beam Therapeutics 公司 Nicole M. Gaudelli 研究团队<sup>[18, 91]</sup> 还利用报告系统或者通过其他 Cas9 蛋白结合 DNA 形成 R-loop 结构暴露出单链 DNA, 分别设计了在细菌和哺乳动物细胞中不需要通过全基因组测序即可检测碱基编辑器脱靶编辑的方法。之后北京大学伊成器研究团队<sup>[92]</sup> 开发了针对胞嘧啶碱基编辑敏感度更高的检测技术 Detect-seq (dU-detection enabled by C-to-T transition during sequencing), 通过利用化学标记和生物素 Pulldown 技术, 追踪 CBE 体内编辑后的基因组产物中的中间体脱氧尿苷来揭示 CBE 的编辑组 (editome)。

### 1.3.2 碱基编辑器在目标位点产生的编辑副产物及优化策略

通过上述检测手段, 人们发现碱基编辑器造成的非预期编辑包括碱基颠换、旁观者编辑 (bystander editing) 以及 Indels 突变。产生碱基颠换的原因在前文已有介绍, 即 C 脱氨后形成的 U 被 UDG 切除, 形成 AP 位点后经历细胞内的碱基切除修复而产生 C 到 A 或 G 的颠换; 而 ABE 的编辑 A 的过程中发生碱基颠换则未见报道, 可能由于体内从 DNA 中切除肌苷的酶的活性或丰度较低, 导致 A 被编辑后经历碱基切除修复的可能性极小<sup>[22]</sup>。当编辑窗口内存在多个 C 或者 A 时, 除了目标碱基, 周围的碱基也被编辑, 这种情况被称为旁观者编辑。虽然在某些情况下旁观者编辑的影响不大, 例如由于密码子简并性, 旁观者编辑可能只会导致沉默突变, 或者导致同类型氨基酸的转变,

但是其他情况下, 旁观者编辑可能会产生非预期的严重后果。Indels 突变的产生则主要由碱基编辑器中 nCas9 产生的切刻或是碱基脱氨之后经历碱基切除修复所导致。检测结果显示, 碱基编辑后确实有少量的 Indels 突变发生, 并且 CBE 产生的频率一般要高于 ABE<sup>[22]</sup>, 但碱基编辑导致 Indels 突变产生的机理还未被完全揭示<sup>[93]</sup>。

针对上述情况, 人们采取了一系列的改进措施。为了避免产生碱基颠换, 人们在碱基编辑器中引入更多的 UGI 来降低碱基颠换的概率<sup>[14, 30]</sup>。为了减少旁观者编辑, 人们一方面通过上文介绍的缩小编辑窗口的方法, 减少脱氨酶可接触的底物数量。另一方面利用脱氨酶编辑序列的偏好性, 使得脱氨酶只编辑某些位置的 C, 例如改造的 A3G 几乎只编辑 CC 基序中的第二个 C<sup>[15]</sup>。此外, 韩国汉阳大学 Sangsu Bae 研究团队<sup>[94]</sup> 通过在 TadA7.10 中引入突变, 降低了其对旁观者胞嘧啶脱氨的活性及 RNA 编辑的活性。美国莱斯大学 Anatoly B. Kolomeisky 等研究团队<sup>[95]</sup> 通过数学建模分析预测目标碱基和旁观者碱基的编辑效率, 并据此开发了提高目标碱基与旁观者碱基编辑比的 A3A 版本。针对 Indels 突变, 通过使用 dCas9<sup>[7]</sup> 或 AsCas12a<sup>[66]</sup> 引导脱氨酶, 或者利用噬菌体 Mu 的 Gam 蛋白<sup>[30]</sup> 均可以降低 Indels 突变的发生频率。

### 1.3.3 碱基编辑器造成的 sgRNA 依赖和非 sgRNA 依赖的 DNA 脱靶编辑及优化策略

除了在目标位点产生非预期的编辑外, 碱基编辑器脱靶到其他 DNA 位点也会产生非预期的编辑产物, 包括 sgRNA 依赖的 DNA 脱靶编辑以及非 sgRNA 依赖的 DNA 脱靶编辑。

碱基编辑器所造成的 sgRNA 依赖的 DNA 脱靶编辑的发生位点并不完全与 sgRNA 引导 Cas 蛋白所造成的脱靶编辑的位点相同, 其中还受到编辑窗口内是否存在脱氨底物、底物上下游的核苷酸种类以及脱氨酶能否接触到 R-loop 结构中 DNA 链等因素影响。人们发现这一类脱靶编辑的发生相比非 sgRNA 依赖的脱靶编辑概率较小<sup>[89]</sup>, 但仍给碱基编辑器的应用带来了潜在的风险, 因此人们对该问题进行了一系列的优化。通过使用高保真 Cas 蛋白替换野生型 Cas 蛋白可以有效降低 CBE 和 ABE 这一类脱靶编辑发生频率, 例如 HF-Cas9,

xCas9 (3.7)、Sniper-Cas9、HypaCas9 等<sup>[17, 60, 75, 87, 96]</sup>。中国科学院广州生物医药与健康研究院赖良学与五邑大学邹庆剑研究团队合作将腺苷脱氨酶或胞苷脱氨酶与转录激活因子样效应子融合, 并依靠靶向附近的 dCas9 打开 DNA 双链, 开发了 TaC9-ABE 和 TaC9-CBE 碱基编辑器。该系统可以消除 sgRNA 依赖的脱靶编辑, 而不影响靶向编辑效率<sup>[24, 97]</sup>。有研究显示使用截短的 sgRNA 可以在不大幅降低目标位点编辑活性的情况下, 减少碱基编辑器 sgRNA 依赖的 DNA 脱靶编辑<sup>[96]</sup>, 但是也有研究发现使用截短的 sgRNA 会导致碱基编辑器在目标位点和脱靶位点的编辑活性均有所降低, 并且当 sgRNA 与 DNA 的错配出现在 PAM 远端时, 使用截短的 sgRNA 反而会展现出更高的脱靶活性<sup>[86]</sup>, 表明在碱基编辑器中使用截短 sgRNA 来提升特异性的方法并不如在 Cas9 中适用<sup>[98]</sup>。而使用 5'端添加 1~2 个 G 的 sgRNA 则可以在不降低目标位点编辑活性的情况下减少 sgRNA 依赖的 DNA 脱靶编辑<sup>[86]</sup>, 但也有研究显示脱靶的减少只在 sgRNA 5'端添加 2 个 G 的情况下出现<sup>[88]</sup>。武汉大学孙宇辉研究团队<sup>[99]</sup>通过在 sgRNA 的 5'端引入气泡发夹结构, 也可以有效地提升碱基编辑器的编辑特异性。

碱基编辑器所造成的非 sgRNA 依赖的 DNA 脱靶编辑通常由 CBE 引发, 即使在没有 sgRNA 的情况下, CBE 也能引发一定量的非 sgRNA 依赖的 DNA 脱靶编辑。而在 ABE 的编辑产物中几乎没有检测到此类脱靶编辑, 发生频率仅接近自发突变。这些突变主要是 C 到 T 的单核苷酸变异, 被认为是由胞苷脱氨酶和 UGI 在整个基因组中长期随机的反应而产生。而 DNA 腺苷脱氨酶本身由 RNA 腺苷脱氨酶进化而来, 与 DNA 结合能力较弱, 所以 ABE 几乎不引发非 sgRNA 依赖的 DNA 脱靶编辑<sup>[89-90, 100]</sup>。针对这一问题, David R. Liu 研究团队<sup>[18]</sup>通过测试, 发现 YE1-BE4、YE2-BE4、R33A+K34A-BE4 等 7 个碱基编辑器均减少了 DNA 层面的各种脱靶问题, 其中 YE1 与其他 Cas 变体展现了更好的兼容性, 可以高效地编辑更多的位点。Nicole M. Gaudelli 研究团队<sup>[91]</sup>从 153 个脱氨酶中筛选出 4 个脱氨酶变体——*Pp*ABOBEC1、*Rr*A3F、*Am*ABOBEC1 和 *Ss*ABOBEC3b, 由它们所构建的碱

基编辑器降低了非 sgRNA 依赖的脱靶编辑活性, 并通过对这些脱氨酶进行工程改造, 进一步降低了融合这些脱氨酶的碱基编辑器 DNA 和 RNA 的脱靶编辑频率。杨辉研究团队<sup>[101]</sup>从 23 个理性设计的 CBE 变体中筛选出在目标位点保留了高编辑效率, 同时引发极低 DNA 和 RNA 脱靶编辑和旁观者编辑的 YE1-BE3-FNLS 变体。

### 1.3.4 碱基编辑器造成的 RNA 脱靶编辑及优化策略

此外, 人们还通过 RNA 测序, 发现碱基编辑器除了造成 DNA 层面的脱靶编辑, 还会引起大量的非 sgRNA 依赖的 RNA 脱靶编辑<sup>[19, 102-104]</sup>, 脱靶的原因被认为是 CBE 和 ABE 脱氨酶均具有对 RNA 脱氨的活性。而 RNA 脱靶编辑除了低水平、广泛而随机地发生在内源 RNA 上, 也可能造成对 sgRNA 和碱基编辑器 mRNA 的编辑, 进而可能诱发 DNA 层面脱靶编辑的发生<sup>[19]</sup>。针对这些问题, 人们通过在脱氨酶中引入突变, 构建了多种碱基编辑器变体, 减少了 RNA 脱靶编辑的发生, 其中一些突变也减少了 DNA 脱靶编辑的发生<sup>[19, 23, 67, 91, 101-104]</sup>。另外, 也有研究通过将腺苷脱氨酶嵌合于 Cas9 内部的特定位点, 限制了腺苷脱氨酶的活性, 减少了 RNA 脱靶编辑的发生<sup>[25]</sup>。

### 1.3.5 碱基编辑器特异性优化的其他策略

除了蛋白层面的工程改造, 也可以通过时空控制减少碱基编辑器的剂量或作用时间来提升碱基编辑器的特异性。碱基编辑器可以通过核糖核蛋白<sup>[17-18, 23, 105-106]</sup>、mRNA<sup>[61, 67, 79, 107]</sup>或逆转录病毒颗粒<sup>[108]</sup>的方式递送到细胞内。通过这几种方式递送的碱基编辑器在体内都更容易被降解, 减少了碱基编辑器在体内的存在时间, 从而减少脱靶编辑发生的机会。利用 anti-CRISPR 蛋白 Acr II A5 可以有效地抑制 Cas9 的活性<sup>[109]</sup>, 利用 anti-deaminase (Ade) 蛋白可以有效地抑制脱氨酶 APOBEC3A 的活性<sup>[110]</sup>, 从而减少碱基编辑器的脱靶。通过在 sgRNA 中设计 RNA 适体酶 (aptamer) 来招募脱氨酶编辑可以提升编辑的特异性<sup>[111]</sup>。此外通过类似于拆分 Cas 蛋白的方式<sup>[112-113]</sup>拆分脱氨酶<sup>[114]</sup>, 可以精准地控制编辑的时间, 减少额外编辑的发生。陈佳研究团队<sup>[115]</sup>利用其首次发现的、具有胞苷脱氨酶抑制作用的胞苷脱氨酶的调节结构域调控胞

苷脱氨酶，开发了高特异性的碱基编辑器 tBEs (transformer base editors)。

## 2 碱基编辑技术在微生物合成生物学领域的应用

由于碱基编辑器可以在不引入双链 DNA 断裂的情况下实现碱基的转换，于是人们可以方便地将碱基编辑器应用于提前终止密码子的引入和修复。如 CBE 可以将 CAA、CAG、CGA 或 TGG 密码子转换为终止密码子 TAA、TAG 和 TGA 来提前终止基因的表达<sup>[116-117]</sup>，以及通过 ABE 将起始密码子 ATG 转变成 ACG 扰乱基因翻译的起始，都可以起到类似基因失活的作用<sup>[118]</sup>，用于考察基因的功能或定向改造生物代谢途径。相反，ABE 也可以通过将提前的终止密码子变成谷氨酰胺 (CAA、CAG) 或精氨酸 (CGA) 的密码子，使提前终止的基因翻译恢复正常<sup>[119]</sup>。在以代谢产物为主要研究目标的微生物中，人们对于脱靶所带来的非预期后果的容忍度远高于以人类疾病治疗为目标的研究。因此，碱基编辑技术和策略迅速被运用于微生物合成生物学研究领域，我国科研工作者在许多重要微生物天然产物的生物合成途径研究及改造中一展身手，也为基因编辑技术的发展提供了持续不断的驱动力。例如：

中国科学院分子植物卓越中心杨晟研究团队<sup>[120]</sup>利用其构建的适用于梭菌 *Clostridium beijerinckii* 的胞嘧啶碱基编辑器 pCBEclo-opt，在 *C. beijerinckii* NCIMB 8052 中使得编码木糖代谢的转录调节因子 *xyIR* 产生 C-G 至 T-A 的突变，所获得的突变株木糖消耗量比野生型菌株高 10%。

北京大学席建中研究团队通过 CBE (dCas9-CDA1-UL) 在基因中引入提前的终止密码子，逐个敲除了红细菌 *Rhodobacter sphaeroides* 2.4.1 中参与抗氧化剂辅酶 Q<sub>10</sub> 生物合成的系列基因，揭示了这些基因在合成辅酶 Q<sub>10</sub> 过程中的重要性。还通过敲除八氢番茄红素合酶基因 *crtB*，以及同时敲除 *crtB* 和参与光合氧化还原信号调节的基因 *ppsR*，将野生型中辅酶 Q<sub>10</sub> 的产量分别提高了 6% 和 39%<sup>[121]</sup>，向人们展现出了该技术和策略的实用性。

孙宇辉研究团队<sup>[122]</sup>构建了适用于链霉菌属的 CBE 和 ABE，并通过 CBE 在 *Streptomyces avermitilis* MA-4680 中 *olm*、*rpp*、*pks1*、*pks4* 和 *pks8* 五个甚至多达九个 PKS 基因簇内同时引入提前终止密码子，提高 *S. avermitilis* MA-4680 中阿维菌素的产量。之后，该研究团队将其开发的适用于链霉菌的碱基编辑器成功应用于揭示潮霉素 B 的生物合成途径。他们利用 CBE 在 *Streptomyces hygroscopicus* subsp. *hygroscopicus* DSM 40578 参与潮霉素 B 生物合成的五个候选基因内引入提前的终止密码子或突变编码区内的保守残基序列使基因失活。结合体内外的实验证据揭示了蛋白 HygJ、HygL 和 HygD 负责核苷二磷酸庚糖的连续脱氢、转氨基和转糖基化，HygY 是一种不同寻常、作用于羟基碳的自由基 S-腺苷甲硫氨酸依赖的差向异构酶，而 HygM 是一个在多平行代谢网络中发挥作用的多功能甲基转移酶。这些工作，特别在多基因同步考察上都展现出了碱基编辑器有别于传统方法和策略在天然产物生物合成机制研究领域的独特优越性<sup>[123]</sup>。

中国科技大学俞汉青研究团队<sup>[124]</sup>也在希瓦氏菌 *Shewanella oneidensis* MR-1 中开发了快速且强大的碱基编辑器 pCBEso，并应用该工具快速鉴定出 *S. oneidensis* MR-1 内参与 N-乙酰氨基葡萄糖 (GlcNAc) 或葡萄糖代谢的关键基因。此外，他们还利用 pCBEso 构建了具有扩展碳源利用谱的 *S. oneidensis* MR-1 工程菌株，当葡萄糖或 GlcNAc 成为唯一碳源时，其对多种有机污染物 (偶氮染料和有机砷化合物) 的降解率高于野生型菌株。

上海交通大学肖毅研究团队<sup>[125]</sup>针对假单胞菌属的腺嘌呤碱基编辑器开发了一种基于 ABE 的蛋白质功能障碍预测方法 DABE-CSP (dysfunction via ABE through CRISPOR-SIFT prediction)，利用该方法，他们发现了黏康酸 (muconic acid) 代谢途径中的基因 *catB* 失活或 *catB* 和 *catC* 同时失活会导致黏康酸的积累，并且后者以 100% 的产率积累黏康酸。DABE-CSP 方法的建立可以加快对假单胞菌属物种的研究。

上海师范大学芦银华研究团队<sup>[126]</sup>利用其构建的胞嘧啶碱基编辑器 dCas9-CDA-UL<sub>str</sub>，在链霉菌 *Streptomyces rapamycinicus* 2001 内抑制雷帕霉素合成的基因 *rapS* 中引入提前终止密码子，使雷帕霉

素产量提高了63.3%。

浙江大学吴坚平研究团队<sup>[127]</sup>利用其构建的适用于假单胞菌属的多重胞嘧啶碱基编辑系统提高了 *Pseudomonas putida* KT2440 中儿茶酸的产量。他们利用该系统在原儿茶酸3,4-双加氧酶的β亚基(PcaH)和丙酮酸激酶(PykA)编码基因中引入提前终止密码子,并通过该系统在3-脱氧-2-阿拉伯庚酮糖-7-磷酸合成酶的同工酶AroF-2内引入了不同的点突变(G136E、G136K和G136R),其中KT2440: : AroF2(G136E)-PJ23119ΔpcaHΔpykA菌株摇瓶发酵的儿茶酸产量较AroF-2未编辑菌株和原始菌株分别增加了69.01%和611.17%。

中国科学院天津工业生物技术研究所马延和研究团队<sup>[128]</sup>开发了基于碱基编辑器的一种多重自动化谷氨酸棒杆菌碱基编辑方法(multiplex automated *Corynebacterium glutamicum* base editing method, MACBETH),并利用MACBETH产生多重基因失活、用于提高谷氨酸产量的突变株文库,发现丙酮酸激酶编码基因pyk和乳酸脱氢酶编码基因ldhA双基因失活菌株可将谷氨酸产量相较野生型提高3倍。此外,他们通过集成机器人系统实现了MACBETH的自动化,使该方法能够每月生成数千个定向设计的*C. glutamicum*代谢工程的菌株。该研究团队利用自动化平台构建了一个包含94个转录因子基因失活文库,成功率达到100%,该工作展现了碱基编辑器快速便捷地生成突变菌株的能力,成为工业应用中多重自动化细菌基因组编辑的强大工具<sup>[128]</sup>。随后该研究团队利用拓展了PAM序列的Cas蛋白变体(nVQR-Cas9(D10A)、nVRER-Cas9(D10A)、nxCas9 3.7(D10A)和nCas9-NG(D10A)),替换了MACBETH方法中构建的CBE中的野生型nCas9,拓展了CBE在*C. glutamicum*中的编辑范围,利用nxCas9 3.7(D10A)-AID去除了lysC基因编码的天冬氨酸激酶产生的反馈抑制,使工程菌株产生1.7 g/L赖氨酸,而在野生型菌株中,赖氨酸的产量低至几乎无法检测<sup>[82]</sup>。之后,该研究团队还开发了一种基于碱基编辑器靶向且无模板的表达调控方法(base editor-targeted and template-free expression regulation, BETTER)用于在微生物中大规模微调基因的表达,它可以在原位生成大量不同的核糖体结合位点、5'非翻译

区或启动子的遗传组合,并无需合成、转化和整合DNA供体。他们应用BETTER在模式微生物*C. glutamicum*和*Bacillus subtilis*中同时调节多达十个基因的表达,以增强宿主的底物摄取或天然产物生物合成,分别获得改善了木糖分解代谢、甘油分解代谢或番茄红素生物合成的工程菌株<sup>[129]</sup>。

江南大学周哲敏研究团队<sup>[81]</sup>构建了可以实现体内原位蛋白进化的工具CRISPR-CDA-nCas9-UGI,并利用该工具在芽孢杆菌*B. subtilis*中对影响细胞功能的Sec转位酶进行进化突变,获得了提升运输能力的突变株。此外,他们还利用CRISPR-CDA-nCas9-UGI对抗菌肽抗性决定因子bceB基因进行进化,获得了对抗菌肽抗性更高的突变株。BceB蛋白在保护细胞壁生物合成靶点免受抗菌肽的抑制方面发挥重要作用,杆菌肽抗性突变体的创建对于构建产生杆菌肽的细胞工厂具有重要价值<sup>[81]</sup>。

此外,脱氨酶还可以与随机结合DNA的蛋白融合,设计成为产生DNA随机突变的工具。浙江大学连佳长研究团队<sup>[130]</sup>将不同的非特异性ssDNA结合蛋白与胞苷脱氨酶融合,理论上可以在酵母菌*Saccharomyces cerevisiae*全基因组将胞嘧啶替换成胸腺嘧啶的突变。他们利用这些rBE,提高了*S. cerevisiae*对异丁醇和乙酸盐的耐受性并增加了产β-胡萝卜素*S. cerevisiae*的β-胡萝卜素产量。他们还使用rBE对*S. cerevisiae*基因组进行持续进化,获得了对浓度为9%的异丁醇具有抗性的*S. cerevisiae*细胞。张学礼研究团队<sup>[131]</sup>将DNA解旋酶与活化诱导的胞苷脱氨酶(AID)融合形成酶复合物Helicase-AID,Helicase-AID可以在细胞内整个染色体中随机引入碱基突变。该研究团队利用这一工具分别对产β-胡萝卜素的*E. coli* KH001和*S. cerevisiae* 4742crt进行多轮随机突变,使得突变株β-胡萝卜素产量分别增加371.4%和75.4%。

### 3 展 望

碱基编辑器自问世以来,从最初可以实现目标位点编辑窗口内C到T、A到G的碱基转换,到之后开发和优化出的碱基变换种类不断丰富,编

辑窗口、靶向位点不断拓展以及特异性不断优化的新版本,极大地加速了碱基编辑器在生物学、医学等领域的应用。然而,碱基编辑器的碱基编辑偏好性,可以靶向编辑的位点和碱基受限,以及脱靶编辑的问题仍有待进一步完善。例如,虽然目前较多的碱基编辑器在编辑窗口上都比早期版本有所拓展,但是碱基编辑器对于单个或少数碱基的精确编辑仍较难实现,而精确地在基因元件中进行碱基转换对于改造微生物来说也至关重要。此外,碱基编辑器的脱靶编辑所造成的细菌毒性以及微生物的表型变化也会给微生物代谢途径改造带来障碍,虽然目前已报道了一些脱靶效应低至与自发突变水平相当的碱基编辑器,但是这些系统往往包含较多的元件,系统繁复,效能有限,因此开发或者寻找到适合微生物的简便精准的碱基编辑器还需要合成生物学研究工作者们的不断努力。同时,还可以通过对以碱基为底物的酶的发现和改造或是对生物体内核酸参与的生化反应的深入了解,利用碱基结构的变化或是体内不同的核酸修复途径,创造更多碱基变换的可能,例如由嘌呤变嘧啶等,将更能拓展碱基编辑器的应用场景。

此外,与其他在微生物中使用的基因编辑工具一样,碱基编辑器也存在着对难以建立遗传操作的微生物无法递送外源DNA的障碍。而由碱基编辑器实现的碱基转变也需要考虑到微生物产生回复突变的可能性。

碱基编辑器在微生物合成生物学中的应用虽然较CRISPR/Cas系统的基因编辑有较大优势,但是主要的问题仍旧是碱基编辑所能实现的碱基变换种类较少,虽然可以实现引入提前终止密码子等,但是对于某些氨基酸密码子的转变仍无能为力。因此创造碱基编辑器更多碱基变换的可能变得尤为重要。

可以相信随着人们研究的不断深入,碱基编辑器的性能也将日趋完善,在基础科学、农业和医学等领域将发挥更大的作用。

## 参 考 文 献

- [1] JANSEN R, VAN EMBDEN J D A, GAASTRA W, et al. Identification of genes that are associated with DNA repeats in prokaryotes[J]. *Molecular Microbiology*, 2002, 43(6): 1565-1575.
- [2] CONG L, RAN F A, COX D, et al. Multiplex genome engineering using CRISPR/Cas systems[J]. *Science*, 2013, 339(6121): 819-823.
- [3] MALI P, YANG L H, ESVELT K M, et al. RNA-guided human genome engineering via Cas9[J]. *Science*, 2013, 339(6121): 823-826.
- [4] LANDRUM M J, LEE J M, BENSON M, et al. ClinVar: public archive of interpretations of clinically relevant variants[J]. *Nucleic Acids Research*, 2016, 44(D1): D862-D868.
- [5] HAAPANIEMI E, BOTLA S, PERSSON J, et al. CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response[J]. *Nature Medicine*, 2018, 24(7): 927-930.
- [6] IHRY R J, WORRINGER K A, SALICK M R, et al. p53 Inhibits CRISPR-Cas9 engineering in human pluripotent stem cells [J]. *Nature Medicine*, 2018, 24(7): 939-946.
- [7] KOMOR A C, KIM Y B, PACKER M S, et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage[J]. *Nature*, 2016, 533(7603): 420-424.
- [8] NISHIDA K, ARAZOE T, YACHIE N, et al. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems[J]. *Science*, 2016, 353(6305): aaf8729.
- [9] WANG X, LI J N, WANG Y, et al. Efficient base editing in methylated regions with a human APOBEC3A-Cas9 fusion[J]. *Nature Biotechnology*, 2018, 36(10): 946-949.
- [10] KOBLAN L W, DOMAN J L, WILSON C, et al. Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction[J]. *Nature Biotechnology*, 2018, 36(9): 843-846.
- [11] LI X S, WANG Y, LIU Y J, et al. Base editing with a Cpf1-cytidine deaminase fusion[J]. *Nature Biotechnology*, 2018, 36(4): 324-327.
- [12] NISHIMASU H, SHI X, ISHIGURO S, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space[J]. *Science*, 2018, 361(6408): 1259-1262.
- [13] ZONG Y, SONG Q N, LI C, et al. Efficient C-to-T base editing in plants using a fusion of nCas9 and human APOBEC3A[J]. *Nature Biotechnology*, 2018, 36(10): 950 - 953.
- [14] JIANG W, FENG S J, HUANG S S, et al. BE-PLUS: a new base editing tool with broadened editing window and enhanced fidelity[J]. *Cell Research*, 2018, 28(8): 855-861.
- [15] LEE S S, DING N, SUN Y D, et al. Single C-to-T substitution using engineered APOBEC3G-nCas9 base editors with minimum genome- and transcriptome-wide off-target effects[J]. *Sci-*

- ence Advances, 2020, 6(29): eaba1773.
- [16] TAN J J, ZHANG F, KARCHER D, et al. Engineering of high-precision base editors for site-specific single nucleotide replacement[J]. *Nature Communications*, 2019, 10: 439.
- [17] REES H A, KOMOR A C, YE H W, et al. Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery[J]. *Nature Communications*, 2017, 8: 15790.
- [18] DOMAN J L, RAGURAM A, NEWBY G A, et al. Evaluation and minimization of Cas9-independent off-target DNA editing by cytosine base editors[J]. *Nature Biotechnology*, 2020, 38(5): 620-628.
- [19] GRÜNEWALD J, ZHOU R H, GARCIA S P, et al. Transcriptome-wide off-target RNA editing induced by CRISPR-guided DNA base editors[J]. *Nature*, 2019, 569(7756): 433-437.
- [20] MOK B Y, DE MORAES M H, ZENG J, et al. A bacterial cytidine deaminase toxin enables CRISPR-free mitochondrial base editing[J]. *Nature*, 2020, 583(7817): 631-637.
- [21] LIM K, CHO S I, KIM J S. Nuclear and mitochondrial DNA editing in human cells with zinc finger deaminases[J]. *Nature Communications*, 2022, 13(1): 366.
- [22] GAUDELLI N M, KOMOR A C, REES H A, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage[J]. *Nature*, 2017, 551(7681): 464-471.
- [23] RICHTER M F, ZHAO K T, ETON E, et al. Phage-assisted evolution of an adenine base editor with improved Cas domain compatibility and activity[J]. *Nature Biotechnology*, 2020, 38(7): 883-891.
- [24] LIU Y, ZHOU J Z, LAN T, et al. Elimination of Cas9-dependent off-targeting of adenine base editor by using TALE to separately guide deaminase to target sites[J]. *Cell Discovery*, 2022, 8: 28.
- [25] LI S, YUAN B, CAO J X, et al. Docking sites inside Cas9 for adenine base editing diversification and RNA off-target elimination[J]. *Nature Communications*, 2020, 11: 5827.
- [26] ZHANG X H, ZHU B Y, CHEN L, et al. Dual base editor catalyzes both cytosine and adenine base conversions in human cells[J]. *Nature Biotechnology*, 2020, 38(7): 856-860.
- [27] WANG Y H, ZHOU L F, TAO R, et al. sgBE: a structure-guided design of sgRNA architecture specifies base editing window and enables simultaneous conversion of cytosine and adenosine[J]. *Genome Biology*, 2020, 21(1): 222.
- [28] KURT I C, ZHOU R H, IYER S, et al. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells[J]. *Nature Biotechnology*, 2021, 39(1): 41-46.
- [29] CHEN L W, PARK J E, PAA P, et al. Programmable C : G to G : C genome editing with CRISPR-Cas9-directed base excision repair proteins[J]. *Nature Communications*, 2021, 12: 1384.
- [30] KOMOR A C, ZHAO K T, PACKER M S, 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[J]. *Science Advances*, 2017, 3(8): eaao4774.
- [31] GRÜNEWALD J, ZHOU R H, LAREAU C A, et al. A dual-deaminase CRISPR base editor enables concurrent adenine and cytosine editing[J]. *Nature Biotechnology*, 2020, 38(7): 861-864.
- [32] LI C, ZHANG R, MENG X B, et al. Targeted, random mutagenesis of plant genes with dual cytosine and adenine base editors[J]. *Nature Biotechnology*, 2020, 38(7): 875-882.
- [33] SAKATA R C, ISHIGURO S, MORI H, et al. Base editors for simultaneous introduction of C-to-T and A-to-G mutations[J]. *Nature Biotechnology*, 2020, 38(7): 865-869.
- [34] TAO W Y, LIU Q, HUANG S S, et al. CABE-RY: a PAM-flexible dual-mutation base editor for reliable modeling of multi-nucleotide variants[J]. *Molecular Therapy-Nucleic Acids*, 2021, 26: 114-121.
- [35] XIE J K, HUANG X Y, WANG X, et al. ACBE, a new base editor for simultaneous C-to-T and A-to-G substitutions in mammalian systems[J]. *BMC Biology*, 2020, 18(1): 131.
- [36] ARBAB M, SHEN M W, MOK B, et al. Determinants of base editing outcomes from target library analysis and machine learning[J]. *Cell*, 2020, 182(2): 463-480.e30.
- [37] KIM H S, JEONG Y K, HUR J K, et al. Adenine base editors catalyze cytosine conversions in human cells[J]. *Nature Biotechnology*, 2019, 37(10): 1145-1148.
- [38] MA Y Q, ZHANG J Y, YIN W J, et al. Targeted AID-mediated mutagenesis (TAM) enables efficient genomic diversification in mammalian cells[J]. *Nature Methods*, 2016, 13(12): 1029-1035.
- [39] ZHAO D D, LI J, LI S W, et al. Glycosylase base editors enable C-to-A and C-to-G base changes[J]. *Nature Biotechnology*, 2021, 39(1): 35-40.
- [40] KOBLAN L W, ARBAB M, SHEN M W, et al. Efficient C•G-to-G•C base editors developed using CRISPRi screens, target-library analysis, and machine learning[J]. *Nature Biotechnology*, 2021, 39(11): 1414-1425.
- [41] YUAN T L, YAN N N, FEI T Y, et al. Optimization of C-to-G base editors with sequence context preference predictable by machine learning methods[J]. *Nature Communications*, 2021,

- 12: 4902.
- [42] LI R Q, CHAR S N, LIU B, et al. High-efficiency plastome base editing in rice with TAL cytosine deaminase[J]. *Molecular Plant*, 2021, 14(9): 1412-1414.
- [43] WEI Y H, LI Z F, XU K, et al. Mitochondrial base editor DdCBE causes substantial DNA off-target editing in nuclear genome of embryos[J]. *Cell Discovery*, 2022, 8: 27.
- [44] CHO S I, LEE S H, MOK Y G, et al. Targeted A-to-G base editing in human mitochondrial DNA with programmable deaminases[J]. *Cell*, 2022, 185(10): 1764-1776.e12.
- [45] BASS B L. RNA editing by adenosine deaminases that act on RNA[J]. *Annual Review of Biochemistry*, 2002, 71: 817-846.
- [46] BAZAK L, HAVIV A, BARAK M, et al. A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes[J]. *Genome Research*, 2014, 24(3): 365-376.
- [47] MONTIEL-GONZALEZ M F, VALLECILLO-VIEJO I, YUDOWSKI G A, et al. Correction of mutations within the cystic fibrosis transmembrane conductance regulator by site-directed RNA editing[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(45): 18285-18290.
- [48] MONTIEL-GONZÁLEZ M F, VALLECILLO-VIEJO I C, ROSENTHAL J J C. An efficient system for selectively altering genetic information within mRNAs[J]. *Nucleic Acids Research*, 2016, 44(21): e157.
- [49] SCHNEIDER M F, WETTENGEL J, HOFFMANN P C, et al. Optimal guideRNAs for re-directing deaminase activity of hADAR1 and hADAR2 in trans[J]. *Nucleic Acids Research*, 2014, 42(10): e87.
- [50] STAFFORST T, SCHNEIDER M F. An RNA-deaminase conjugate selectively repairs point mutations[J]. *Angewandte Chemie International Edition*, 2012, 51(44): 11166-11169.
- [51] VOGEL P, MOSCHREF M, LI Q, et al. Efficient and precise editing of endogenous transcripts with SNAP-tagged ADARs[J]. *Nature Methods*, 2018, 15(7): 535-538.
- [52] VOGEL P, SCHNEIDER M F, WETTENGEL J, et al. Improving site-directed RNA editing *in vitro* and in cell culture by chemical modification of the guideRNA[J]. *Angewandte Chemie International Edition*, 2014, 53(24): 6267-6271.
- [53] WETTENGEL J, REAUTSCHNIG P, GEISLER S, et al. Harnessing human ADAR2 for RNA repair-recoding a PINK<sub>1</sub> mutation rescues mitophagy[J]. *Nucleic Acids Research*, 2017, 45(5): 2797-2808.
- [54] QU L, YI Z Y, ZHU S Y, et al. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs[J]. *Nature Biotechnology*, 2019, 37(9): 1059-1069.
- [55] COX D B T, GOOTENBERG J S, ABUDAYYEH O O, et al. RNA editing with CRISPR-Cas13[J]. *Science*, 2017, 358(6366): 1019-1027.
- [56] ABUDAYYEH O O, GOOTENBERG J S, FRANKLIN B, et al. A cytosine deaminase for programmable single-base RNA editing[J]. *Science*, 2019, 365(6451): 382-386.
- [57] HUANG X X, LV J J, LI Y Q, et al. Programmable C-to-U RNA editing using the human APOBEC3A deaminase [J]. *The EMBO Journal*, 2020, 39(22): e104741.
- [58] TANG G Y, XIE B R, HONG X N, et al. Creating RNA specific C-to-U editase from APOBEC3A by separation of its activities on DNA and RNA substrates[J]. *ACS Synthetic Biology*, 2021, 10(5): 1106-1115.
- [59] KIM Y B, KOMOR A C, LEVY J M, et al. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions[J]. *Nature Biotechnology*, 2017, 35(4): 371-376.
- [60] HU J H, MILLER S M, GEURTS M H, et al. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity[J]. *Nature*, 2018, 556(7699): 57-63.
- [61] MILLER S M, WANG T N, RANDOLPH P B, et al. Continuous evolution of SpCas9 variants compatible with non-G PAMs [J]. *Nature Biotechnology*, 2020, 38(4): 471-481.
- [62] WALTON R T, CHRISTIE K A, WHITTAKER M N, et al. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants[J]. *Science*, 2020, 368(6488): 290-296.
- [63] CHATTERJEE P, JAKIMO N, JACOBSON J M. Minimal PAM specificity of a highly similar SpCas9 ortholog[J]. *Science Advances*, 2018, 4(10): eaau0766.
- [64] CHATTERJEE P, LEE J, NIP L, et al. A Cas9 with PAM recognition for adenine dinucleotides[J]. *Nature Communications*, 2020, 11(1): 2474.
- [65] LIU Z Q, SHAN H H, CHEN S Y, et al. Efficient base editing with expanded targeting scope using an engineered Spy-mac Cas9 variant[J]. *Cell Discovery*, 2019, 5: 58.
- [66] KLEINSTIVER B P, SOUSA A A, WALTON R T, et al. Engineered CRISPR-Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing [J]. *Nature Biotechnology*, 2019, 37(3): 276-282.
- [67] GAUDELLI N M, LAM D K, REES H A, et al. Directed evolution of adenine base editors with increased activity and therapeutic application[J]. *Nature Biotechnology*, 2020, 38(7): 892-900.

- [68] HUANG T P, ZHAO K T, MILLER S M, et al. Circularly permuted and PAM-modified Cas9 variants broaden the targeting scope of base editors[J]. *Nature Biotechnology*, 2019, 37(6): 626-631.
- [69] OAKES B L, FELLMANN C, RISHI H, et al. CRISPR-Cas9 circular permutants as programmable scaffolds for genome modification[J]. *Cell*, 2019, 176(1/2): 254-267.e16.
- [70] THURONYI B W, KOBLAN L W, LEVY J M, et al. Continuous evolution of base editors with expanded target compatibility and improved activity[J]. *Nature Biotechnology*, 2019, 37(9): 1070-1079.
- [71] YU W X, LI J N, HUANG S S, et al. Harnessing A3G for efficient and selective C-to-T conversion at C-rich sequences[J]. *BMC Biology*, 2021, 19(1): 34.
- [72] LIU Z Q, SHAN H H, CHEN S Y, et al. Improved base editor for efficient editing in GC contexts in rabbits with an optimized AID-Cas9 fusion[J]. *FASEB Journal*, 2019, 33(8): 9210-9219.
- [73] CHENG T L, LI S, YUAN B, et al. Expanding C-T base editing toolkit with diversified cytidine deaminases[J]. *Nature Communications*, 2019, 10: 3612.
- [74] TANENBAUM M E, GILBERT L A, QI L S, et al. A protein-tagging system for signal amplification in gene expression and fluorescence imaging[J]. *Cell*, 2014, 159(3): 635-646.
- [75] GEHRKE J M, CERVANTES O, CLEMENT M K, et al. An APOBEC3A-Cas9 base editor with minimized bystander and off-target activities[J]. *Nature Biotechnology*, 2018, 36(10): 977-982.
- [76] ZHANG X H, CHEN L, ZHU B Y, et al. Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain[J]. *Nature Cell Biology*, 2020, 22(6): 740-750.
- [77] WANG Y H, ZHOU L F, LIU N, et al. BE-PIGS: a base-editing tool with deaminases inlaid into Cas9 PI domain significantly expanded the editing scope[J]. *Signal Transduction and Targeted Therapy*, 2019, 4: 36.
- [78] VILLIGER L, SCHMIDHEINI L, MATHIS N, et al. Replacing the *SpCas9* HNH domain by deaminases generates compact base editors with an alternative targeting scope[J]. *Molecular Therapy Nucleic Acids*, 2021, 26: 502-510.
- [79] RYU S M, KOO T, KIM K, et al. Adenine base editing in mouse embryos and an adult mouse model of Duchenne muscular dystrophy[J]. *Nature Biotechnology*, 2018, 36(6): 536-539.
- [80] BANNO S, NISHIDA K, ARAZOE T, et al. Deaminase-mediated multiplex genome editing in *Escherichia coli*[J]. *Nature Microbiology*, 2018, 3(4): 423-429.
- [81] HAO W L, CUI W J, CHENG Z Y, et al. Development of a base editor for protein evolution *via in situ* mutation *in vivo*[J]. *Nucleic Acids Research*, 2021, 49(16): 9594-9605.
- [82] WANG Y, LIU Y, LI J W, et al. Expanding targeting scope, editing window, and base transition capability of base editing in *Corynebacterium glutamicum*[J]. *Biotechnology and Bioengineering*, 2019, 116(11): 3016-3029.
- [83] LIU Z Q, CHEN S Y, SHAN H H, et al. Efficient base editing with high precision in rabbits using YFE-BE4max[J]. *Cell Death & Disease*, 2020, 11: 36.
- [84] TAN J J, ZHANG F, KARCHER D, et al. Expanding the genome-targeting scope and the site selectivity of high-precision base editors[J]. *Nature Communications*, 2020, 11: 629.
- [85] ZHAO D D, JIANG G, LI J, et al. Imperfect guide-RNA (igRNA) enables CRISPR single-base editing with ABE and CBE[J]. *Nucleic Acids Research*, 2022, 50(7): 4161-4170.
- [86] KIM D, LIM K, KIM S T, et al. Genome-wide target specificities of CRISPR RNA-guided programmable deaminases[J]. *Nature Biotechnology*, 2017, 35(5): 475-480.
- [87] KIM D, KIM D E, LEE G, et al. Genome-wide target specificity of CRISPR RNA-guided adenine base editors[J]. *Nature Biotechnology*, 2019, 37(4): 430-435.
- [88] LIANG P P, XIE X W, ZHI S Y, et al. Genome-wide profiling of adenine base editor specificity by EndoV-seq[J]. *Nature Communications*, 2019, 10(1): 67.
- [89] JIN S, ZONG Y, GAO Q, et al. Cytosine, but not adenine, base editors induce genome-wide off-target mutations in rice[J]. *Science*, 2019, 364(6437): 292-295.
- [90] ZUO E W, SUN Y D, WEI W, et al. Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos[J]. *Science*, 2019, 364(6437): 289-292.
- [91] YU Y, LEETE T C, BORN D A, et al. Cytosine base editors with minimized unguided DNA and RNA off-target events and high on-target activity[J]. *Nature Communications*, 2020, 11(1): 2052.
- [92] LEI Z X, MENG H W, LV Z C, et al. Detect-seq reveals out-of-protospacer editing and target-strand editing by cytosine base editors[J]. *Nature Methods*, 2021, 18(6): 643-651.
- [93] LEI L Q, CHEN H Q, XUE W, et al. APOBEC3 induces mutations during repair of CRISPR-Cas9-generated DNA breaks[J]. *Nature Structural & Molecular Biology*, 2018, 25(1): 45-52.
- [94] JEONG Y K, LEE S, HWANG G H, et al. Adenine base editor engineering reduces editing of bystander cytosines[J]. *Nature*

- Biotechnology, 2021, 39(11): 1426-1433.
- [95] WANG Q, YANG J, ZHONG Z C, et al. A general theoretical framework to design base editors with reduced bystander effects[J]. Nature Communications, 2021, 12: 6529.
- [96] LEE J K, JEONG E, LEE J, et al. Directed evolution of CRISPR-Cas9 to increase its specificity[J]. Nature Communications, 2018, 9: 3048.
- [97] ZHOU J Z, LIU Y, WEI Y H, et al. Eliminating predictable DNA off-target effects of cytosine base editor by using dual guiders including sgRNA and TALE[J]. Molecular Therapy, 2022, 30(7): 2443-2451.
- [98] FU Y F, SANDER J D, REYON D, et al. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs[J]. Nature Biotechnology, 2014, 32(3): 279-284.
- [99] HU Z W, WANG Y N, LIU Q, et al. Improving the precision of base editing by bubble hairpin single guide RNA[J]. mBio, 2021, 12(2): e00342-21.
- [100] MCGRATH E, SHIN H, ZHANG L Y, et al. Targeting specificity of APOBEC-based cytosine base editor in human iPSCs determined by whole genome sequencing[J]. Nature Communications, 2019, 10: 5353.
- [101] ZUO E W, SUN Y D, YUAN T L, et al. A rationally engineered cytosine base editor retains high on-target activity while reducing both DNA and RNA off-target effects[J]. Nature Methods, 2020, 17(6): 600-604.
- [102] GRÜNEWALD J, ZHOU R H, IYER S, et al. CRISPR DNA base editors with reduced RNA off-target and self-editing activities[J]. Nature Biotechnology, 2019, 37(9): 1041-1048.
- [103] REES H A, WILSON C, DOMAN J L, et al. Analysis and minimization of cellular RNA editing by DNA adenine base editors [J]. Science Advances, 2019, 5(5): eaax5717.
- [104] ZHOU C Y, SUN Y D, YAN R, et al. Off-target RNA mutation induced by DNA base editing and its elimination by mutagenesis[J]. Nature, 2019, 571(7764): 275-278.
- [105] JANG H K, JO D H, LEE S N, et al. High-purity production and precise editing of DNA base editing ribonucleoproteins[J]. Science Advances, 2021, 7(35): eabg2661.
- [106] YEH W H, CHIANG H, REES H A, et al. *In vivo* base editing of post-mitotic sensory cells[J]. Nature Communications, 2018, 9: 2184.
- [107] JIANG T T, HENDERSON J M, COOTE K, et al. Chemical modifications of adenine base editor mRNA and guide RNA expand its application scope[J]. Nature Communications, 2020, 11(1): 1979.
- [108] BANSKOTA S, RAGURAM A, SUH S, et al. Engineered virus-like particles for efficient *in vivo* delivery of therapeutic proteins[J]. Cell, 2022, 185(2): 250-265.
- [109] LIANG M M, SUI T T, LIU Z Q, et al. Acr II A5 suppresses base editors and reduces their off-target effects[J]. Cells, 2020, 9(8): 1786.
- [110] LIU Z Q, CHEN S Y, LAI L X, et al. Inhibition of base editors with anti-deaminases derived from viruses[J]. Nature Communications, 2022, 13: 597.
- [111] COLLANTES J C, TAN V M, XU H T, et al. Development and characterization of a modular CRISPR and RNA aptamer mediated base editing system[J]. The CRISPR Journal, 2021, 4(1): 58-68.
- [112] GANGOPADHYAY S A, COX K J, MANNA D, et al. Precision control of CRISPR-Cas9 using small molecules and light [J]. Biochemistry, 2019, 58(4): 234-244.
- [113] ZETSCHKE B, VOLZ S E, ZHANG F. A split-Cas9 architecture for inducible genome editing and transcription modulation[J]. Nature Biotechnology, 2015, 33(2): 139-142.
- [114] BERRÍOS K N, EVITT N H, DEWEERD R A, et al. Controllable genome editing with split-engineered base editors[J]. Nature Chemical Biology, 2021, 17(12): 1262-1270.
- [115] WANG L J, XUE W, ZHANG H X, et al. Eliminating base-editor-induced genome-wide and transcriptome-wide off-target mutations[J]. Nature Cell Biology, 2021, 23(5): 552-563.
- [116] BILLON P, BRYANT E E, JOSEPH S A, et al. CRISPR-mediated base editing enables efficient disruption of eukaryotic genes through induction of STOP codons[J]. Molecular Cell, 2017, 67(6): 1068-1079.
- [117] KUSCU C, PARLAK M, TUFAN T R, et al. CRISPR-STOP: gene silencing through base-editing-induced nonsense mutations[J]. Nature Methods, 2017, 14(7): 710-712.
- [118] WANG X J, LIU Z W, LI G L, et al. Efficient gene silencing by adenine base editor-mediated start codon mutation[J]. Molecular Therapy, 2020, 28(2): 431-440.
- [119] LEE C I, JO D H, HWANG G H, et al. CRISPR-Pass: gene rescue of nonsense mutations using adenine base editors[J]. Molecular Therapy, 2019, 27(8): 1364-1371.
- [120] LI Q, SEYS F M, MINTON N P, et al. CRISPR-Cas9<sup>D10A</sup> nickase-assisted base editing in the solvent producer *Clostridium beijerinckii*[J]. Biotechnology and Bioengineering, 2019, 116(6): 1475-1483.
- [121] LUO Y F, GE M, WANG B L, et al. CRISPR/Cas9-deaminase enables robust base editing in *Rhodobacter sphaeroides* 2.4.1 [J]. Microbial Cell Factories, 2020, 19(1): 93.
- [122] ZHONG Z Y, GUO J H, DENG L, et al. Base editing in *Strep-*

- tomyces* with Cas9-deaminase fusions[EB/OL]. bioRxiv, 2019 [2022-09-01]. <https://www.biorxiv.org/content/10.1101/630137v1>.
- [123] LI S C, LIU Q, ZHONG Z Y, et al. Exploration of hygromycin B biosynthesis utilizing CRISPR-Cas9-associated base editing [J]. ACS Chemical Biology, 2020, 15(6): 1417-1423.
- [124] CHENG L, MIN D, HE R L, et al. Developing a base-editing system to expand the carbon source utilization spectra of *Shewanella oneidensis* MR-1 for enhanced pollutant degradation [J]. Biotechnology and Bioengineering, 2020, 117(8): 2389-2400.
- [125] ABDULLAH, WANG P J, HAN T R, et al. Adenine base editing system for *Pseudomonas* and prediction workflow for protein dysfunction via ABE[J]. ACS Synthetic Biology, 2022, 11(4): 1650-1657.
- [126] ZHAO Y W, TIAN J Z, ZHENG G S, et al. Multiplex genome editing using a dCas9-cytidine deaminase fusion in *Streptomyces*[J]. Science China Life Sciences, 2020, 63(7): 1053-1062.
- [127] SUN J, LU L B, LIANG T X, et al. CRISPR-assisted multiplex base editing system in *Pseudomonas putida* KT2440[J]. Frontiers in Bioengineering and Biotechnology, 2020, 8: 905.
- [128] WANG Y, LIU Y, LIU J, et al. MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method[J]. Metabolic Engineering, 2018, 47: 200-210.
- [129] WANG Y, CHENG H J, LIU Y, et al. *In-situ* generation of large numbers of genetic combinations for metabolic reprogramming via CRISPR-guided base editing[J]. Nature Communications, 2021, 12: 678.
- [130] PAN Y J, XIA S Y, DONG C, et al. Random base editing for genome evolution in *Saccharomyces cerevisiae*[J]. ACS Synthetic Biology, 2021, 10(10): 2440-2446.
- [131] WANG J, ZHAO D D, LI J, et al. Helicase-AID: a novel molecular device for base editing at random genomic loci[J]. Metabolic Engineering, 2021, 67: 396-402.



**通讯作者：**孙宇辉(1967—),男,博士,教授,博士生导师。研究方向为微生物药物合成生物学及其编辑工具开发和优化。

E-mail: yhsun@whu.edu.cn



**第一作者：**王雁南(1994—),男,博士。研究方向为基因编辑工具开发和优化。

E-mail: ynw@whu.edu.cn