



引导编辑系统的优化及在DNA大片段编辑中的应用*

焦瑶歌 姚少华**

(四川大学华西医院肿瘤中心生物治疗科, 生物治疗全国重点实验室, 成都 610041)

摘要 基因编辑技术是指利用人工核酸酶对细胞和个体中特定的基因序列进行插入、替换或删除等编辑修饰。CRISPR/Cas9 核酸酶的发现是基因编辑技术发展的一个里程碑, 但其编辑产物的精确性和脱靶效应依然是限制其应用的关键因素。近年来以引导编辑技术为代表的衍生性编辑工具因高效且精准而受到广泛关注。该系统能够以不可逆的方式在基因组中靶向引入多种类型的遗传变化, 包括 12 种可能类型的点突变, 以及片段的插入和缺失及其组合, 而无需 DNA 双链断裂 (DSB) 或者供体 DNA 模板。引导编辑技术结合了 CRISPR/Cas9 的靶向性和逆转录酶的精准编辑能力, 使得编辑产物更加精准。本综述将深入探讨引导编辑技术的发展、优化以及在 DNA 大片段编辑中的应用。

关键词 CRISPR/Cas9, 引导编辑系统, pegRNA, DNA 大片段编辑

中图分类号 Q78, Q81

DOI: 10.16476/j.pibb.2024.0266

基因编辑技术是最近二十年生命科学领域最具革命性的突破, 在基础研究和临床应用方面已展现出广阔的前景。目前, 成功用于基因组编辑的技术有归巢核酸内切酶 (homing endonuclease)^[1-2]、锌指核酸内切酶 (zinc finger endonuclease, ZFN)^[3-4]、转录激活因子类似效应物核酸酶 (transcription activator-like effector nuclease, TALEN)^[5-6] 和成簇规律间隔短回文重复 (clustered regularly interspaced short palindromic repeats, CRISPR) / CRISPR 相关蛋白 9 (CRISPR-associated protein 9, Cas9) 系统。尤其是以 Cas9 为代表的 CRISPR 核酸酶技术, 在操作上更为简便, 相比以往的归巢核酸内切酶、锌指酶、转录激活因子类似核酸酶等工具, 具有天然的优势, 是近年来分子生物学领域最重大的技术突破之一。CRISPR/Cas 系统本身是一类广泛分布于细菌和古生菌等物种的适应性免疫系统, 在 2012 年, 美国加州大学伯克利分校和瑞典马尔默大学的两女科学家 Jennifer Doudna 和 Emmanuelle Charpentier 率先证明: 源于酿脓链球菌 (*Streptococcus pyogenes*) 的 SpCas9 可在体外, 以可重编程的方式、高效而精确地对特定靶序列进

行识别和切割^[7]。优化后的 CRISPR/Cas9 系统由 Cas9 与人工设计的单指导 RNA (single-guide RNA, sgRNA) 组成, Cas9 在 sgRNA 的指引下可以精确识别并切割目标 DNA 序列, DNA 序列被切割之后, 通过细胞内的修复机制, 致使基因组序列发生改变^[8]。

除了切割 DNA, Cas9 还可以作为一个基因操作平台, 将各种功能性的蛋白质元件运载到特定 DNA 位点附近以开发衍生性的编辑工具, 实现定点的 DNA 编辑、修饰或者调控。在这些衍生性编辑工具中, 近年来发展起来的碱基编辑工具因可实现高效、精确的单碱基编辑而备受关注^[9-14]。碱基编辑工具由 Cas9 和脱氨酶融合而成, 根据脱氨酶的不同分为胞嘧啶碱基编辑器 (CBE) 和腺嘌呤碱基编辑器 (ABE), CBE 系统实现胞嘧啶 (C) 到胸腺嘧啶 (T) 的转换^[15], ABE 系统实现腺嘌呤 (A) 到

* 国家重点研发计划 (2023YFC3403200) 和四川省科技计划 (2024NSFTD0029) 资助项目。

** 通讯联系人。

Tel: 028-85503628, E-mail: shaohuayao@scu.edu.cn

收稿日期: 2024-06-25, 接受日期: 2024-08-20

鸟嘌呤(G)的转换^[16], 这对众多因为此类碱基转换导致的遗传病研究及治疗有重要意义。

之后在CBE基础上开发出的C-to-G碱基编辑器(C-to-G base editor, CGBE)和糖基化酶碱基编辑器(glycosylase base editor, GBE)能实现C-to-G和C-to-A之间的碱基颠换(transversion)^[17-18], 此技术的开发主要得益于研究者们对CBE系统编辑副产物(包括C-to-G和C-to-A的颠换)的深入研究。研究者们认为, 引入高效切除胞嘧啶脱氨产生的尿嘧啶(U)的酶——尿嘧啶DNA糖基化酶(UNG)或去掉CBE系统中的尿嘧啶糖基化酶抑制剂(UGI)组分, 可以生成无碱基(apurinic/aprimidinic site, AP)中间体, 之后实现无碱基中间体介导的高效的C-to-G和/或C-to-A碱基颠换^[17-18]。接着研究者们开发出一系列无碱基中间体依赖的碱基编辑器(gGBE、ACBE、AYBE、DAF-CBE/DAF-TBE、TSBE3等), 丰富了碱基颠换类型^[19-23]。虽然利用这些编辑系统可以实现所有12种类型的碱基转换, 但是仍然存在不足之处。

a. 旁位编辑。当编辑活性窗口内存在多个目标核苷酸时, 均会被编辑。这样可能会造成氨基酸突变或提前引入终止密码子而影响蛋白质表达。

b. 编辑纯度问题。单碱基编辑以及无碱基中间体依赖的碱基编辑器均存在编辑纯度的问题, 不能将目标核苷酸编辑为单一的指定的碱基, 不能实现精确的编辑。

c. 插入和缺失(indels)水平过高。indel会导致DNA序列发生不可预测的变化, 可能引起基因功能的改变, 影响基因表达、蛋白质结构或细胞功能等。

而2019年David R. Liu实验室^[24]研发出的基因组编辑技术——引导编辑(prime editing, PE)技术, 在一定程度上克服了这些局限性。引导编辑系统主要由两部分组成, 与逆转录酶(MLV-RT)融合表达的SpCas9切口酶(H840A)以及工程化的pegRNA(prime editing guide RNA)。该技术可以在靶位点实现精确的12种可能类型的点突变, 以及小片段插入和缺失, 而无需DNA双链断裂(double-strand breakage, DSB)或供体DNA模板^[24]。引导编辑系统被开发以来凭借其优势发展迅速, 研究者们基于该技术纠正与遗传性疾病的致病突变^[25-29], 构建疾病相关动物模型^[30-33], 助力农作物育种和改良^[34-38]等, 这些应用场景展示了引导编辑技术的多样性和潜力。除此之外, 引导编辑技术在DNA大片段编辑方面(插入、缺失、替

换、串联重复以及染色体倒位和易位等)展现出了巨大的潜力。但是编辑效率低也限制了该技术的应用, 所以如何提高引导编辑系统的编辑效率也成为研究热点, 近年来研究者在提高引导编辑效率方面的研究也取得了很大的进展。本综述详细介绍了引导编辑系统的发展, 重点介绍各种提升引导编辑效率的策略以及引导编辑技术在DNA大片段编辑中的应用。

1 引导编辑系统

1.1 引导编辑系统的概况

David R. Liu实验室提出的引导编辑系统主要由两部分组成, 一部分是与野生型逆转录酶(M-MLV-RT)融合表达的SpCas9切口酶(H840A), 其中逆转录酶来源于莫洛尼氏鼠白血病病毒(Moloney murine leukaemia virus)^[24]。另一部分是工程化的pegRNA, pegRNA是在常规sgRNA的3'端添加了引物结合序列(primer binding site, PBS)和逆转录模板(RT)^[24]。首先在pegRNA的引导下, SpCas9切口酶(H840A)会切割非靶向链——原型间隔区邻接基序(protospacer adjacent motif, PAM)所在的DNA链, 然后pegRNA 3'端的PBS可以与切口前的序列识别配对, 之后逆转录酶会以pegRNA上人工设计的逆转录模板序列为模板进行逆转录。接着在切口附近会出现包含编辑的3'侧翼序列(flap), 以及与之动态平衡的未包含编辑的5'flap, 在细胞修复过程中5'flap易被结构特异性的内切酶识别并切除, 最后通过DNA复制或修复机制将新的序列写入靶位点^[24](图1)。Osamu Nureki和张锋团队^[39]研究了该系统在多种状态下的冷冻电镜结构, 对预起始、启动转录和延伸状态的结构进行比较, 在逆转录的过程中, 逆转录酶始终相对于SpCas9保持一致的位置, 而通过逆转录酶和pegRNA合成的DNA杂合双链堆积在SpCas9的表面。对终止结构状态的功能分析显示, 逆转录酶将逆转录扩展到了预期的模板之外, 也就是将sgRNA的部分骨架序列也逆转录出来, 从而在靶基因引入了不希望的编辑^[39]。

引导编辑系统经由David R. Liu实验室开发并经过一系列的优化。首先, David R. Liu实验室^[24]将逆转录酶融合到Cas9切口酶的C端, 并将此系统称为PE1。经过测试, PE1系统可以在哺乳动物细胞基因组实现精确的点突变, 小片段的碱基插入或缺失(1~3 bp)^[24]。但是PE1的编辑效率较低,

多数位点不足5%。

考虑到逆转录酶在引导编辑系统中起着重要的作用,为了提高该系统的编辑效率,David R. Liu团队^[24]进一步优化了逆转录酶。优化从已知可以增加逆转录酶的热稳定性、逆转录连续性以及DNA:RNA的底物亲和力等突变入手。据此创建并评估了包含多种M-MLV RT突变的19种变体,最终在M-MLV逆转录酶中引入5个氨基酸突变(D200N/L603W/T330P/T306K/W313F)形成了PE2系统,与PE1相比,PE2在哺乳动物细胞基因组的编辑效率提高了1.6~5.1倍^[24]。

之后该团队对系统进行进一步的优化,引入一条靶向非编辑链的sgRNA以诱导其产生单链断裂,促使细胞以编辑链作为模板进行修复,从而提高系统的编辑效率(PE3)^[24]。与PE2相比,该策略可以将编辑效率提高1.5~4.2倍^[24]。但是PE3系统的indel水平也有所升高,这可能是由于引入sgRNA后,双链切割产生一个类似DSB的双链断裂。为了降低indel水平,David R. Liu实验室^[24]将sgRNA设计为仅可以靶向编辑链编辑后的序列(PE3b)。与PE3系统相比,PE3b系统显著降低了indel的水平,而编辑效率没有明显的降低。

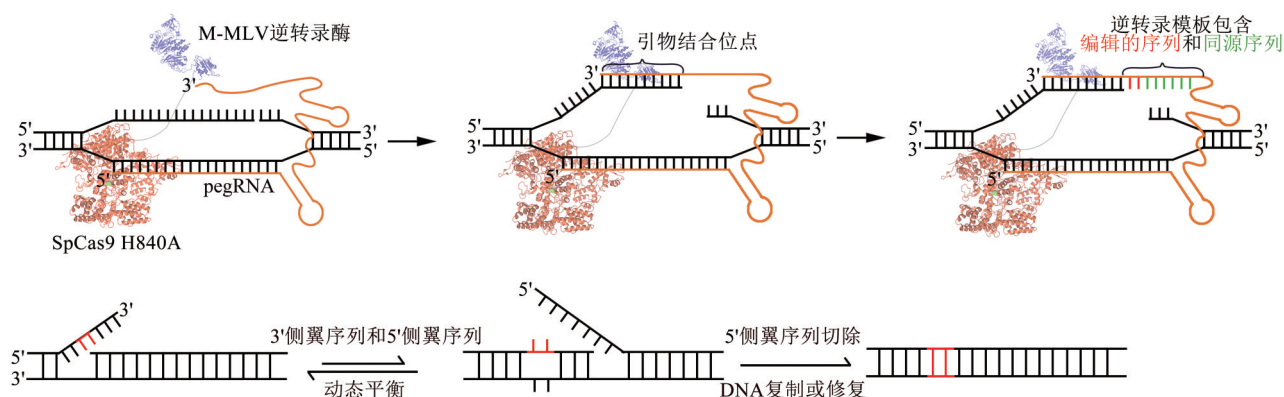


Fig. 1 The working principle diagram of prime editing system

图1 引导编辑系统的工作原理示意图

1.2 引导编辑系统的特点

1.2.1 编辑结果精确

引导编辑系统的编辑方式足够准确,可以通过逆转录模板指定任意想要写入基因组的序列,从而实现所有12种可能类型的点突变,以及小片段的插入缺失及其组合^[24]。相比之下,单碱基编辑器CBE和ABE以及新型AP位点依赖的碱基编辑器虽然可以实现所有的碱基转换类型,但是这些编辑工具编辑模式固定,无法做到编辑方案的订制,且或多或少地存在不精确性,如旁位编辑、编辑纯度等问题。

1.2.2 编辑类型多样

引导编辑系统的编辑类型多样,特别在DNA大片段编辑上面展现出无限的潜力。最初的引导编辑系统可以实现小片段的插入、缺失以及组合^[24],之后研究者们基于该系统还可以实现DNA大片段的删除、插入、替换、基因整合^[26, 40-49]以及染色体易位、倒位、串联重复^[40, 50-52]等。

1.2.3 PAM的限制小

与碱基编辑系统相比,引导编辑系统扩展了基因组编辑的范围。碱基编辑系统存在活性窗口的限制,引导编辑技术可以在靠近或远离PAM位点的位置进行编辑,编辑位置到PAM的距离可以超过30 bp^[24]。

1.2.4 较低的脱靶水平

引导编辑系统在全基因组范围具有很高的编辑特异性^[47, 53]。高彩霞团队^[53]在植物上对引导编辑系统的脱靶效应进行了研究,发现pegRNA依赖型的脱靶效应非常低,且该系统不会造成全基因组范围的不依赖于pegRNA的脱靶效应。但是在碱基编辑系统中,由于sgRNA引导碱基编辑器结合与靶向位点序列相似的位点,造成sgRNA序列依赖的脱靶^[54-55],并且碱基编辑系统特别是CBE系统会在全基因组和全转录组范围内引发严重的sgRNA非依赖性脱靶,这些突变会随机分布于整个基因组中^[56-58]。

综上所述,引导编辑技术在编辑准确性、编辑

类型、PAM限制以及脱靶水平等方面都存在一定的优势,但是由于引导编辑机制涉及到多个元件的协同作用和内源性修复或复制因子的调控,引导编辑技术的编辑效率普遍较低,这也限制了其更广泛的应用。接下来将从不同方面讨论如何优化引导编辑系统从而提高该系统的编辑效率。

2 引导编辑系统的优化

2.1 pegRNA的优化

2.1.1 RT和PBS的优化

pegRNA是引导编辑系统重要的组成部分之一,可以对其进行各方面的优化来提高引导编辑系统的编辑效率(表1)。David R.Liu团队^[24]首次提出引导编辑系统时就发现PBS和RT的参数设置与编辑效率息息相关,PBS的长度约在13 nt时效率较高,而最优的RT长度与编辑类型以及编辑位点相关^[24](表1)。此外他们发现pegRNA的3'延伸第一个碱基是C时会降低引导编辑系统的编辑效率(表1)。

pegRNA的PBS与编辑链切口前的DNA结合以启动逆转录,其熔解温度(melting temperature, T_m)决定了两条链结合的稳定性,所以PBS的 T_m 值可能影响编辑效率(表1)。高彩霞团队^[59]在植物中使用引导编辑系统时发现,具有30°C熔解温度的PBS序列使引导编辑系统的编辑效率最优。而伊成器课题组^[47]发现在哺乳动物细胞中PBS的 T_m 值范围为34~40°C时的编辑效率最优。李占军团队^[60]通过测试发现PBS的 T_m 值范围为42°C时的编辑效率最优。Scot A Wolfe团队^[61]发现对于3'端保护的pegRNA,在哺乳动物细胞中,较短的PBS长度和PBS的 T_m 值接近37°C是最佳选择。植物和哺乳动物细胞之间不同的原因可能是不同类别的细胞培养温度不同。

2.1.2 双pegRNA策略

原始的引导编辑系统中包含一个pegRNA,多个团队证实使用两个pegRNA同时靶向一个目标编辑位点可以提高引导编辑系统的编辑效率以及编辑精确度(表1)。高彩霞团队^[59]首次提出双pegRNA策略,即在DNA的正链和负链各设计一个pegRNA,通过在水稻内源位点测试,最佳 T_m 值与双pegRNA结合的策略可以将引导编辑系统编辑效率从2.9倍提高到17.4倍,大大提高了引导编辑系统在植物中的编辑效率。伊成器课题组^[47]使用配对的pegRNA编码正义和反义DNA链中的相同

编辑区域,通过与原始的引导编辑系统比较,各种编辑形式如碱基替换、插入和删除方面的效率都显著高于PE2系统,且编辑副产物远少于PE3系统。笔者所在团队使用双pegRNA系统将DNA大片段删除的效率提高了16倍,编辑精确度提高了60倍,此外该策略可以提高多碱基同时转换的效率,但是没有提高单碱基转换的效率^[49]。虽然该策略并不是对所有的编辑类型都是有效的,但是总体上从提高引导编辑系统的效率和精确度来说,使用双pegRNA策略不失为一种明智高效的选择。

2.1.3 增加pegRNA的稳定性

PegRNA是在常规sgRNA的基础上增加了PBS和RT,该序列会被外切酶降解,从而降低引导编辑系统的编辑效率。此外pegRNA 3'端的PBS与5'端部分间隔序列(spacer)互补,它们之间的退火会导致pegRNA环化,从而阻碍编辑。

为了增加pegRNA的稳定性,David R.Liu团队^[62]在pegRNA的3'端添加结构化的RNA基序(evopreQ1、mpknot),通过降低pegRNA 3'端延伸的降解来增加pegRNA的稳定性,改进之后的pegRNA(epegRNA)在多种细胞系中将引导编辑系统的编辑效率提高3~4倍(表1)。之后多个研究团队以相似的原理在pegRNA的3'端添加一段序列,如可以形成发夹结构的20 nt的csy4序列(ePE)^[63]、病毒外切核糖核酸酶抗性RNA基序Zika xr-RNA(xr-pegRNA)^[64]、G-quadruplex(G-PE)^[65],或者RNA适配体(sPEs)^[66]等来提高pegRNA的稳定性从而提高编辑效率(表1)。此外陈佳实验室团队^[67]通过改进pegRNA的小发夹结构,将其第2个茎环结构中的G/A碱基对改造为C/G碱基对开发了apegRNA,增加其稳定性将编辑效率平均提高2.77倍(表1)。

为了减少pegRNA的环化,马涵慧团队^[66]采用环状pRNA(circular pRNA, cpRNA)策略大大提高了引导编辑系统的效率,该策略将pegRNA拆分成sgRNA和pRNA(prime RNA)两部分(表1)。高彩霞团队^[68]和Erik J. Sontheimer团队^[27]采用相似的环状方案也提高了系统的编辑效率(表1)。Jing-Ruey Joanna Yeh团队^[69]将pegRNA热变性再缓慢冷却来降低pegRNA的错误折叠水平并增强其与Cas9蛋白复合的能力,该方法将核糖核蛋白介导的斑马鱼胚胎引导编辑效率提高了近25倍,并且通过引入旨在破坏pegRNA内部相互作用的点突变,将斑马鱼胚胎的编辑效率提

高6倍（表1）。Scot A Wolfe团队^[61]为了减少PBS和间隔序列之间的环化自抑制作用，在哺乳动物细胞中使用较短长度的PBS，此外引导编辑系统递送后对细胞进行瞬时冷休克处理进一步提高了编辑效率。但是其中的一些策略并不适用于体内递送。

之后 Britt Adamson 团队^[70]通过 CRISPR 干扰（CRISPR interference, CRISPRi）筛选出一个关键小RNA结合蛋白La（La protein），发现它是影响引导编辑系统编辑效率的关键调控因子。La可以结合RNA聚合酶III（Pol III）转录本3'端的多聚尿苷（polyU），保护它们免受外切核酸酶的破坏^[70]。作者将La蛋白与PEmax系统结合，增强pegRNA的稳定性，特别是当融合到C端改善编辑效果最为显著，该团队将PEmax-C-la（1-194）命名为PE7^[70]。该策略与PEmax系统强强结合，实现了更高水平的编辑效率（表1）。

2.1.4 增加pegRNA的表达水平

多个团队通过优化pegRNA的骨架来增加其表达水平，他们将骨架中连续尿嘧啶的第四个尿嘧啶突变为胞嘧啶或腺嘌呤以消除转录终止信号，以此增加了pegRNA表达，提高引导编辑系统的效率^[62-63]（表1）。此外，林照博与黄行许团队^[71]合作使用RNA聚合酶II启动子代替RNA聚合酶III启动子，适合表达含有4个以上连续U残基的复杂RNA结构（表1）。陈其军团队^[35]在玉米的研究中使用两种类型的启动子（Pol2-Pol3复合型启动子）驱动pegRNA的表达，从而提高编辑效率（表1）。杨进孝和赵久然团队^[72]通过在植物中的测试发现，在RT模板上添加多个同义突变位点能显著提升引导编辑系统的编辑效率（RT-M策略），作者猜测该策略降低了RT模板与切口之后的基因组DNA结合的概率，从而增强逆转录水平提高编辑效率（表1）。

Table 1 The strategies of pegRNA optimization
表1 pegRNA优化策略

名称	结构	参考文献
pegRNA	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—3' 逆转录模板、引物结合序列的长度；逆转录模板的第一个碱基避免是C；引物结合序列的熔解温度值；双pegRNA策略	[24, 44, 47, 49, 59-61]
epegRNA	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—evopreQ1或mpknot—3'	[62]
G-PE	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—G-quadruplex—3'	[65]
xr-pegRNA	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—Zika-xrRNA—3'	[64]
ePE	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架:连续尿嘧啶的第四个尿嘧啶突变为胞嘧啶—逆转录模板和引物结合序列—Csy4序列—间隔序列—SpCas9 sgRNA骨架—3'	[63]
sPEs	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—MS2—3'	[66]
apegRNA	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架:第2个茎环结构的G/A碱基对改造为C/G碱基对—逆转录模板和引物结合序列—3'	[67]
RT-M pegRNA	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板(含有突变)和引物结合序列—MS2—3'	[69, 72]
pegRNA F+E骨架	5'—U6启动子—间隔序列—SpCas9 sgRNA骨架:连续尿嘧啶的第四个尿嘧啶突变为胞嘧啶—逆转录模板和引物结合序列—3'	[62-63]
pegRNA Pol2启动子	5'—Pol2启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—3'	[71]
pegRNA Pol2-Pol3启动子	5'—Pol2-Pol3启动子—间隔序列—SpCas9 sgRNA骨架—逆转录模板和引物结合序列—3'	[35]
petRNA	5'—U6启动子—核糖酶1—MS2—逆转录模板和引物结合序列—核糖酶2—3'	[27]
cpRNA	5'—U6启动子—核糖酶1—5'连接序列—MS2或PP7—逆转录模板和引物结合序列—3'连接序列—核糖酶2—3'	[66]
circular RNAs	5'—U6启动子—5'R—5'L—MS2—逆转录模板和引物结合序列—3'L—3'R—3'	[68]

evopreQ1: 天然存在的最小型RNA基序之一（modified prequeosine1-1 riboswitch aptamer）；mpknot: MMLV逆转录酶的内源性模板，有助于富集MMLV-nCas9融合蛋白（frameshifting pseudoknot from Moloney murine leukemia virus）；G-quadruplex: G-四联体；Zika-xrRNA: 寨卡病毒-源自虫媒黄病毒基因组3'UTR的RNA；Csy4: 细菌CRISPR内切核糖核酸酶；MS2: MS2外壳蛋白；PP7: PP7外壳蛋白；Pol2-Pol3: RNA聚合酶II启动子-RNA聚合酶III启动子；R: 右；L: 左。

2.1.5 pegRNA的优化设计

设计高活性的pegRNA也是提高引导编辑系统效率的有效途径。相对于常规的sgRNA, pegRNA的设计更加复杂。研究人员开发出一系列可以用来设计pegRNA的在线网站, 这些网站根据需要写入基因组的序列以及靶向的基因组序列, 设计最优的PBS和RT序列^[73-79]。基于3'延伸的pegRNA策略, David R.Liu团队^[62]开发了用于识别pegRNA和3'延伸序列之间的接头序列的计算工具pegLIT, 该工具的出现使我们可以高效便捷地设计cpegRNA。

2.2 引导编辑系统蛋白的优化

Cas9蛋白和逆转录酶在引导编辑系统中发挥着重要的作用, 通过进一步优化可提高引导编辑系统的效率(表2)。

薛文团队^[80]及其他团队^[27, 81]在哺乳动物细胞上验证去除逆转录酶中的核糖核酸酶H结构域(RNase H)不影响引导编辑系统的编辑效率(表2)。高彩霞团队^[82]通过去除逆转录酶中的核糖核酸酶H结构域, 并掺入具有核酸伴侣活性的病毒核衣壳蛋白, 开发了ePPE(engineered plant prime editor)系统, 该系统在植物细胞中将编辑效率提高了1.8~3.4倍。Rasmus O. Bak团队^[83]通过对逆转录酶进行变体筛选、密码子优化和去除RNase H结构域的改进, 优化过的系统具有表达优势(1.4倍)和大小优势(短621 bp)(表2)。去除逆转录酶中核糖核酸酶H结构域的策略在不影响编辑效率的基础上更利于递送。David R.Liu团队^[84]也在PEmax系统中对逆转录酶进行了密码子优化, 同时在Cas9蛋白中引入了两个氨基酸(R221K N394K)的突变, 进一步提高了编辑效率(表2)。Gerald Schwank团队^[85]通过酵母中的蛋白质进化技术增强了引导编辑系统的活性, 得到Cas9(A259D)和逆转录酶(K445T)的突变体PE-Y18, 与PEmax相比, 编辑效率提高约3.5倍(表2)。宗媛团队^[86]在ePPE系统的基础上引入PEmax版本的Cas9蛋白的两个突变, 称为ePPEmax*系统, 之后又在逆转录酶中引入一个新的氨基酸突变位点V223A, 开发ePPE plus系统, 显著提升了引导编辑系统在小麦中的基因编辑能力(表2)。此外, 杨进孝和赵久然团队^[72]发现, 在植物中将M-MLV RT置于Cas9 N端能介导更高的编辑效率(PE-P3)(表2)。以上的优化策略大多都是对逆转录酶进行优化, 今后可以聚焦于Cas9蛋白进化, 为引导编辑系统打造最强悍的Cas9

版本。

同时研究者们也通过对引导编辑系统蛋白质优化进一步提高了编辑的精确度。Jin-Soo Kim团队^[87]发现, *SpCas9* H840A不仅可以切割non-target DNA链, 还会切割target DNA链, 只是切割效率低于野生型Cas9, 该团队使用双突变型nCas9(H840A N863A)进一步灭活HNH核酸酶结构域, 该ePE3系统使编辑的精确度提高(表2)。Gil Gregor Westmeyer团队^[88]通过将5'-3'核酸外切酶与引导编辑系统结合开发Exo-PE策略, 该策略增强了≥30 bp片段DNA插入的有效性和精确度(表2)。此外梁振团队^[89]通过将T5 5'-3'核酸外切酶融合在系统的N端, 构建了一种高效的植物引导编辑系统PE2(v2), 可以将点突变删除插入的效率提高1.7~2.9倍(表2)。我们更希望得到将上述提高编辑效率以及提高编辑精确度的策略相结合的系统, 具体的融合策略还需要进一步去探索。

除此之外, 研究者们也通过提高引导编辑系统蛋白的表达量来提高系统的编辑效率。已有文献报道CRISPR/Cas9系统中核定位信号(nuclear localization signal, NLS)序列的组成和数量会影响基因组编辑的效率, 多个团队通过优化碱基编辑系统的NLS来提高编辑效率^[90-91]。原始的引导编辑系统在N端和C端各有一个双分型核定位信号(bipartite nuclear localization signal, bpNLS)^[24], 薛文团队^[25]在该系统N端bp NLS的基础上增加一个C-myc NLS, C端改为SV40 NLS和v-bp SV40 NLS(variant bipartite SV40 NLS), 优化之后的系统PE2*更有效地在靶位点进行核苷酸转换、序列删除或插入(表2)。PEmax系统在Cas9和逆转录酶中间添加了bp SV40 NLS^[84], 进一步提高了系统蛋白质的表达量(表2)。

2.3 不同变体的引导编辑系统

2.3.1 不同Cas9变体的引导编辑系统

原始的引导编辑系统使用化脓性链球菌Cas9(*SpCas9*), 仅可以识别NGG的PAM。若将不同变体的Cas蛋白应用到引导编辑系统中, 就可以靶向非NGG序列的区域, 提高引导编辑系统的适用范围。Yongsu Kim团队^[92]生成了具有PAM灵活性的各种类型PE2变体: PE2-VQR、PE2-VRQR、PE2-VRER、PE2-NG、PE2-SpG和PE2-SpRY。此外多个研究团队还替换了其他来源的Cas9, 如*SaCas9*^[25, 84, 93]、*SaCas9KKH*^[25, 93]、*SauriCas9*^[93]、*FnCas9*^[94]和*CjCas9*^[93]。高彩霞团队^[68]还将

Cas12a 应用到引导编辑系统, 该系统可以识别富含 T 的区域, 并且同时靶向 4 个基因。但是由于 Cas12a 同时在 DNA 和 RNA 水平上具有核酸酶的活性, 这就导致 pegRNA 也会被切割, 所以该团队使用环状 RNA 来克服这一挑战。以上这些变体在增加 PAM 多样性的基础上, 还减小了引导编辑系统的尺寸, 但是与 *SpCas9* 相比, 这些 Cas 蛋白的大小并没有小很多, 之后可以将 *IsrB*^[94]、*IscB*^[95]、*TnpB*^[96] 等应用到引导编辑系统, 并通过进化得到活性更强的 Cas 蛋白。

2.3.2 不同逆转录酶变体的引导编辑系统

原始的引导编辑系统使用来自莫洛尼氏鼠白血病病毒的逆转录酶^[24], 多个团队也测试其他来源的逆转录酶。高彩霞团队^[27]使用针对植物优化的 PPE 系统, 发现原来的逆转录酶可以被 CaMV-RT (来自花椰菜花叶病毒) 和反转录子衍生性逆转录酶 (来自大肠杆菌 BL21) 代替; Erik J. Sontheimer 团队^[34]也使用其他来源的逆转录酶, 但是这些系统的编辑效率都不如原始的逆转录酶, 活性比较弱。David R.Liu 团队^[97]基于噬菌体辅助连续进化 (phage assisted continuous evolution, PACE) 技术和蛋白质工程生成了体积更小, 效率更高的 PE6a-g 系统, 其中每个系统都包含一个新的逆转录酶或 Cas9 变体, 并且测试发现 PE6a 到 PE6g 具有不同的编辑类型偏好, 可根据编辑类型选择使用不同的 PE6 系统 (表 2)。

2.3.3 不同偶联系统的引导编辑系统

常规的引导编辑系统由两部分组成, 一个用来表达 *SpCas9* 切口酶和逆转录酶, 另一个用来表达 pegRNA。Paul Q Thomas 团队^[98]研发的 Prime Editing All-in-One (PEA1) 系统以及 B. P. C. Koeleman 团队^[99]的 Prime Editing all-in-one (pAIO) 系统将表达引导编辑系统所有组分的元件全部构建在一个载体上面, 提高了编辑效率。

此外还有其他团队使用拆分的策略包括 sPE^[27]、Intein-PE^[80, 83, 100-103]、nCas9/MCP-RT 和 SunTag-PE^[27], 这些策略不仅有利于递送, 而且编辑效率提高或与全长系统相当 (表 2)。王可品团队^[104]提出了一种称为 CC-PE 的拆分系统, 其中 nCas9 和 RT 通过卷曲螺旋二聚化肽 (coiled-coil dimerization peptides) 偶联, 该肽仅包含 28 个氨基酸, 可以基于静电和疏水相互作用实现紧密和特异性的二聚化配对, 几乎不会增加载体大小的额外负担 (表 2)。

2.4 内源性因素

2.4.1 错配修复机制

虽然引导编辑系统可以在细胞中实现各种类型的 DNA 变化, 但是对相关的内源性修复机制知之甚少。David R.Liu 团队^[84]通过 CRISPRi 筛选发现错配修复 (mismatch repair, MMR) 活动强烈抑制引导编辑的效率和纯度。他们通过敲除或者降低 MMR 相关基因的表达来提高引导编辑系统的编辑效率, 通过尝试发现瞬时表达 MLH1 显性抑制 (MLH1 dominant negative, MLH1dn) 基因的效果最佳, 并由此产生了 PE4 (PE2 与 MLH1 dn ($\Delta 754-756$) 的组合) 和 PE5 (PE3 与 MLH1 dn ($\Delta 754-756$) 的组合)^[84] (表 2)。但是经过测试 MLH dn 对编辑类型有偏好性, 能增强所有 12 种点突变的编辑效率, 但对 G-C 突变的增强效果最差, 对 <15 bp 的插入缺失突变有明显的增强, 但对更长片段的插入缺失突变并无很好的增强^[84]。此外 J. I. Loizou 团队^[105]针对涵盖已知人类 DNA 修复途径的 32 个 DNA 修复因子进行了集中的遗传筛选, 他们的结果也显示, MMR 参与引导编辑的调控, 并且在不同的细胞系以及编辑类型中, 抑制 MMR 可使编辑效率提高 2~17 倍。Hyuk-Jin Cha 团队^[106]在人类胚胎干细胞中通过敲除 MMR 途径的关键蛋白质, 发现 MutS α 和 MutS β 以依赖于编辑尺寸大小的方式决定引导编辑系统的编辑效率。魏鹏程和李娟团队^[107]使用 RNA 干扰 (RNA interference, RNAi) 技术抑制了水稻中的 *OsMLH1* 基因, 在稳定转化的水稻细胞中将编辑效率提高了 1.30~2.11 倍, 并且通过条件性激活系统不影响植物的生育能力。

2.4.2 染色质环境

染色质环境会强烈影响基因编辑的编辑效率和准确性。不同的染色质状态可能导致编辑系统的访问性受限。因此, 了解和调控染色质环境对于高效的引导编辑系统至关重要。

Jay Shendure 团队^[108]发现引导编辑系统表现出强烈的染色质相关的位置效应, 对于相同的目标位点和编辑, 编辑效率从 0% 到 94% 不等。之后作者通过一种基于单细胞测序 (sci-RNA-seq3) 的技术可以将细胞受到的遗传干扰和不同染色质环境中的引导编辑系统响应联系起来, 进而发现解旋酶样转录因子 (helicase-like transcription factor, HLTF) 为引导编辑系统的依赖性抑制因子^[108]。为此他们提出可以通过分别使用 CRISPR 介导的沉默或激活

来显著降低或提高引导编辑系统的编辑效率^[108]。笔者所在团队通过药物筛选系统发现组蛋白去乙酰化酶 (histone deacetylase, HDAC) 抑制剂可以提高引导编辑系统小片段插入和缺失的编辑效率, 这可能是因为 HDAC 抑制剂促进了染色质开放状态^[109]。Kim 团队^[30]使用近端 gRNA (spacer 长度为 14 nt 或 15 nt 的 sgRNA, 可以结合到目标位点但不具有切割活性) 或染色质调节肽 (CMP-PE1-V1) 的引导编辑系统, 通过打开目标位点的染色质结构从而提高编辑效率 (表 2)。张学礼、毕昌昊团队^[110]使用 SunTag 系统和 MS2 系统招募先锋因子 P65 提升染色质可及性从而提高引导编辑系统的编辑效率 (表 2)。

2.5 其他策略

近期研究还发现, 对 p53 的短暂抑制可以提高引导编辑系统在人胚胎干细胞的编辑效率, 但是其中的机制还没有探究清楚^[111]。Kim 团队^[112]通过添加 Rad51 DNA 结合域 (DBD) 来生成更有效的

PE2 变体 (hyPE2), Rad51 结合域可以促进 DNA/RNA 杂合体的形成, 从而促进 pegRNA 与带切口的目标 ssDNA 的结合, 增强逆转录的水平, 在内源性位点进行测试时, hyPE2 比 PE2 高 1.5 或 1.4 倍 (图 3)。季泉江、黄行许课题组^[113]报道 3'-5' DNA 外切酶是阻碍细菌先导编辑系统的关键遗传因素, 他们设计了一种称为 BacPE 的引导编辑平台, 同时删除 sbcB、xseA 和 exoX DNA 外切酶, 将引导编辑效率提高了 100 倍。Richard I. Sherwood 团队^[114]提出一种肽自我编辑测序分析 (PepSeq) 的高通量方法, 以测量 12 000 个氨基酸肽的融合如何影响引导编辑效率, 开发出的 IN-PE 系统显示氨基酸肽的融合可以通过提高翻译效率从而提高系统的编辑效率 (表 2)。

这些从不同方面优化引导编辑系统的尝试, 丰富了引导编辑系统编辑效率提升的选择, 也提示引导编辑系统效率仍有进一步提升的空间。

Table 2 Different versions of the prime editing system
表2 不同版本的引导编辑系统

名称	结构							参考文献
PE1	SV40 NLS	SpCas9 H840A	33 aa	MMLV RT	SV40 NLS			[24]
PE2	SV40 NLS	SpCas9 H840A	33 aa	MMLV RT突变体	SV40 NLS			[24]
PE2*	c-Myc NLS	SV40 NLS	SpCas9 H840A	33 aa	MMLV RT突变体	v-bp SV40 NLS	SV40 NLS	[25]
PEmax	SV40 NLS	SpCas9 H840A (R221K N394K)	8 aa	SV40 NLS	8 aa	MMLV RT突变体	SV40 NLS c-Myc NLS	[84]
PE4/PE5	SV40 NLS	SpCas9 H840A	33 aa	MMLV RT突变体	SV40 NLS	MLH1 dn		[84]
PE2 ^{ΔRNaseH}	SV40 NLS	SpCas9 H840A	33 aa	MMLV RT突变体 ΔRNaseH	SV40 NLS			[27, 80-83]
PE-Y18	SV40 NLS	SpCas9 H840A (A259D)	33 aa	MMLV RT突变体 (K445T)	SV40 NLS			[85]
ePPE	SV40 NLS	SpCas9 H840A	16 aa	NC	SV40 NLS	32 aa	MMLV RT突变体 ΔRNaseH	[82]
ePPEmax*	SV40 NLS	SpCas9 H840A (R221K N394K)	16 aa	NC	SV40 NLS	32 aa	MMLV RT突变体 ΔRNaseH	[86]
ePPEplus	SV40 NLS	SpCas9 H840A (R221K N394K)	16 aa	NC	SV40 NLS	32 aa	MMLV RT突变体 ΔRNaseH V223A	[86]
PE-P3	SV40 NLS	MMLV RT突变体	33 aa	SpCas9 H840A	SV40 NLS			[72]
Exo-PE	SV40 NLS	EXO1 5'-3' 核酸外切酶	SpCas9 H840A	33 aa	MMLV RT突变体	PCP	SV40 NLS	[88]
ePE3 (H840A N863A)	SV40 NLS	SpCas9 H840A (N863A)	33 aa	MMLV RT突变体	SV40 NLS			[87]
PE2 (v2)	SV40 NLS	GGGGS 接头	TS 5'-3' 核酸外切酶	SpCas9 H840A	33 aa	MMLV RT突变体	SV40 NLS	[89]
PE6	SV40 NLS	SpCas9 H840A*	33 aa	RT突变体*	SV40 NLS			[97]
hy-PE2	SV40 NLS	SpCas9 H840A	16 aa	Rad51 DBD	33 aa	MMLV RT突变体	SV40 NLS	[112]

续表2		
名称	结构	参考文献
CMP-PE1-V1		[30]
IN-PE		[114]
PE7		[70]
sPE		[27]
Intein-PE		[80, 83, 100-103]
SunTag-PE		[27]
nCas9 and MCP-RT		[27]
CC-PE		[104]

SV40 NLS: 猴病毒40 (Simian Virus 40) 核定位信号; aa: 氨基酸; MMLV RT: 来源于莫洛尼氏鼠白血病病毒的逆转录酶; c-Myc NLS: c-Myc核定位信号; v-bp SV40 NLS: 双粒子SV40核定位信号的变体 (variant bipartite SV40 NLS); MLH1 dn: MLH1显性抑制型 (MLH1 dominant negative); RNase H: 核糖核酸酶H结构域; NC: 具有核酸伴侣活性的病毒核衣壳蛋白 (viral nucleocapsid); PCP: PP7适配体结合外壳蛋白; Rad51 DBD: Rad51结合域; HN1/H1G: HN1蛋白/H1G组蛋白; NFATC2IPp: 与NFAT (nuclear factor of activated T-cells) 转录因子家族成员相互作用的蛋白质 (nuclear factor of activated T-cells cytoplasmic 2 interacting protein); La: 关键小RNA结合蛋白; GCN4/scFv: 19个氨基酸的肽, 可被scFv抗体识别/单链可变片段 (single chain variable fragment); MCP: 构成MS2噬菌体衣壳的主要蛋白质; P3/P4: 包含28个氨基酸的肽, 基于静电和疏水相互作用实现紧密和特异性的二聚化配对。

3 引导编辑系统在DNA大片段编辑中的应用

已有研究显示, 在与人类疾病相关的75 122种遗传变异中, 约50%都是由点突变造成的, 缺失占26%, 复制占8.2%, 插入占2.5%, 插入和缺失的组合占2%^[115], 其中大部分是长片段的插入、缺失以及复制。目前实现这些DNA大片段编辑的策略多依赖CRISPR/Cas9核酸酶^[116-123], 但是基于DNA双链断裂修复的方法都容易出错, 导致基因组出现随机的插入或缺失、DNA易位和细胞凋亡途径的激活等^[124-126], 而引导编辑系统的出现给DNA大片段编辑带来了希望, 研究者们基于该系统研发出不同的策略来实现DNA大片段的编辑。

3.1 DNA大片段删除

多个研究团队基于引导编辑系统使用双pegRNA策略实现DNA大片段的删除(图2)。Jay Shendure团队^[44]使用双向pegRNA的设计方案(PRIME-Del), 使待删除片段两端携带与相反的基因组链上模板的互补序列, 从而实现目的片段的删除, 最高可实现~10 kb的删除(图2)。David.R Liu团队^[49]使用与之类似的策略twinPE可以删除高达约800个碱基对的序列(图2)。同样的, 笔者所在团队使用双pegRNA策略在也在哺乳动物细胞基因组实现大片段的删除。该团队首先尝试使用PE3系

统实现DNA大片段的删除, 通过测试发现, 只有nick sgRNA的位置在待敲除片段的3'端时, 可以实现大片段的删除。之后作者将nick sgRNA置换成pegRNA, 结果显示双pegRNA策略(Bi-PE)的编辑效率要高于PE3系统, 而产生的indels水平低于PE3系统(图2)。Bi-PE策略可以删除1 kb的DNA片段, 同时将DNA大片段删除的效率提高16倍, 编辑准确性提高60倍^[49]。之后该团队尝试利用Bi-PE系统在内源基因组上删除大片段的同时插入一小段序列^[49], 并提出两种工作方案Bi-PE-2和Bi-PE-3, 两种方案不同之处在于pegRNA的RT部分, 其中Bi-PE-2的pegRNA的RT只有edit序列(插入的小片段序列), 无同源序列(HA), Bi-PE-3 pegRNA的RT部分同时包含edit和HA序列, 内源性靶点的测试显示Bi-PE-3的策略优于Bi-PE-2^[49](图2)。总之, 这些方案与传统的Cas9介导的缺失突变方法进行比较, 引导编辑系统介导的DNA大片段删除的效率更高, 且产生错误的概率更低。

多个实验室还使用野生型Cas9代替SpCas9 H840A在基因组实现大片段的删除。薛文团队^[26]基于双pegRNA策略使用具有双链DNA切割能力的PE-Cas9实现大片段的删除, 该方法被作者称为PEDAR (PE-Cas9-based deletion and repair), 可实

现8~10 kb长片段DNA删除(图2)。最后作者将PEDAR应用于疾病模型, 将其递送至酪氨酸血症小鼠模型内, 成功实现基因治疗^[26]。类似的, 笔者所在团队研发的WT-PE技术, 使用与Bi-PE两种方案相似的bi-WT-PE class 1和bi-WT-PE class 2,

删除长度可达16.8 Mb^[40](图2)。在DNA长片段的缺失上, 与传统Cas9介导的缺失突变方法相比, 具有双链DNA切割能力的PE-Cas9策略存在效率更高、副产物更少等优点。但是不可否认的是, 这些策略仍然存在极高水平的indels。

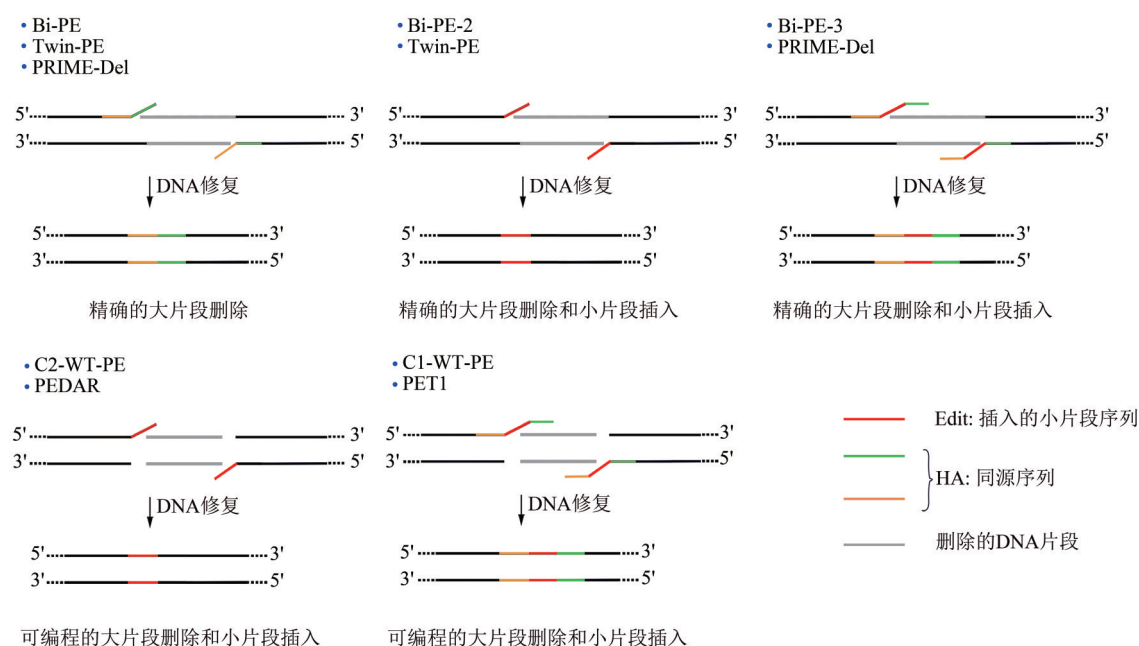


Fig. 2 Technologies of large DNA fragment deletion based on prime editing system

图2 基于引导编辑系统实现DNA大片段删除的技术

3.2 DNA大片段插入

多个研究团队基于引导编辑系统实现了DNA大片段的插入(图3)。殷昊课题组^[48]基于引导编辑系统开发了高效靶向插入的系统 GRAND editing, 该系统使用一对特殊设计的pegRNAs来实现大20~1 000 bp大小的DNA片段靶向插入(图3)。pegRNAs的逆转录模板包含需要插入到基因组的序列并且序列部分互补, 而与靶向的基因组无互补序列^[48]。尽管GRAND editing可以不依赖细胞周期实现片段的插入和置换, 但是在DNA大片段插入和置换方面, 比如在kb级别, 效率依然较低, 仍需进一步优化。

将引导编辑系统与整合酶等位点特异性重组系统结合起来可以有效地提高引导编辑系统介导的DNA大片段敲入, 其原理是先应用引导编辑系统在待编辑基因组位点上引入重组酶的作用位点, 然后利用位点特异性的重组酶将目的基因片段进一步整合进入基因组位点。目前比较常用的是来源于

Bxb1噬菌体的丝氨酸DNA重组酶, 该系统具有相对较高酶活性, 并且在人类细胞基因组序列中缺乏混杂的假位点。David R. Liu团队^[41]将双pegRNA策略twinPE与丝氨酸重组酶Bxb1联合使用时, 发现twinPE可以精确地将长数千个碱基对的基因大小的DNA序列插入到基因组中与治疗有关的位点上(图3)。之后该团队基于PACE技术对位点特异性丝氨酸重组酶Bxb1进行优化, 进化之后的系统eePASSIGE可精确整合大于10 kb的DNA片段, 整合效率比野生型Bxb1高出4.2倍^[127](图3)。同样的, Jonathan S. Gootenberg团队^[45]将引导编辑系统与精确的位点特异性整合酶(丝氨酸整合酶Bxb1)融合表达, 通过位点特异性靶向元件提供可编程插入。通过该策略(PASTE)(图3), 研究人员在三个人类细胞系中成功实现了多个基因组位点约36 kb序列的整合^[45], 此外他们从宏基因组和工程直系同源物中发现了25 614个丝氨酸整合酶, 并且对这些整合酶进行筛选, 发现来源于蜡样芽孢杆菌的

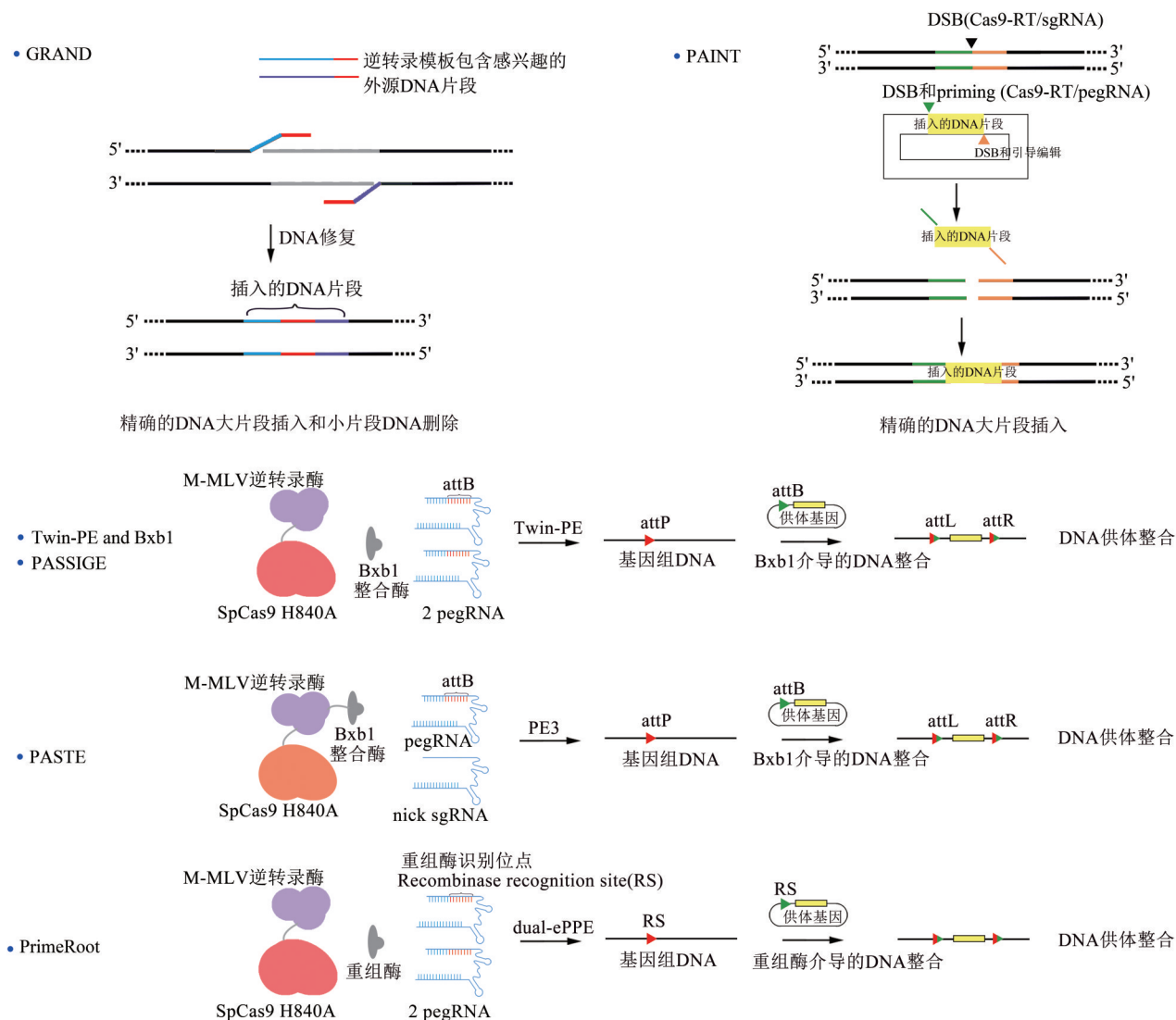


Fig. 3 Technologies of large DNA fragment insertion based on prime editing system

图3 基于引导编辑系统实现DNA大片段插入的技术

DSB: DNA双链断裂; RT: 逆转录酶; priming: 引导; attB/attP: 细菌附着位点/噬菌体的附着位点 (attachment B/ attachment P), 重组酶识别位点; attL/attR: attB和attP之间重组后形成的位点。

整合酶 BceINTa 具有更高的活性。与 David R. Liu 团队的设计类似, 该系统也是首先通过引导编辑系统将 attb 插入基因组, 从而指引整合酶发生位点特异性的 BP 重组反应, 将大片段 DNA 序列插入。高彩霞团队^[43]在植物中通过整合优化的引导编辑工具和酪氨酸家族位点特异性重组酶 Cre 以及双 pegRNA 策略, 开发出 PrimeRoot (Prime editing-mediated Recombination Of Opportune Targets) 系统, 可以在水稻和玉米中实现长达 11.1 kb 的大片段 DNA 的高效精准定点插入, 效率可达 6% (图 3)。该团队对来源于酪氨酸重组酶家族和丝氨酸重

组酶家族的整合酶进行筛选, 发现来源于酪氨酸重组酶家族的 Cre 在植物中具有更高的活性^[43]。此类策略可以将整合酶所需的附着位点插入基因组中的任何位置, 继而实现大片段的插入。这种整合方式存在一定的优势, 首先不会引起 DNA 双链断裂, 其次整合的片段较大。

李伟团队^[46]基于引导编辑系统设计了一种强大的敲入策略 PAINT (图 3)。该技术首先使用 Cas9-RT/sgRNA 在基因组产生一个切口, 接着基于 Cas9-RT/2pegRNA 在供体质载体的待敲入基因的两侧产生 flap, 该 flap 的序列由 pegRNA 决定, 且与

基因组切口两侧的序列配对, 最终基于微同源物介导的末端连接的机制将基因敲入基因组。PAINT 3.0可以在几个治疗相关的基因组位点插入2.5 kb的转基因, 敲入效率高达85%。PAINT技术是一种用于大型转基因整合的强大基因编辑工具, 但是该技术使用Cas9核酸酶亦有引入大量indels的风险。

3.3 染色体倒位和易位

染色体疾病是由染色体异常引起的一类疾病, 如染色体易位、倒位和扩增等, 由此会影响人体的生长、发育和功能^[128]。染色体疾病的建模是非常重要的, 通过研究染色体异常的模式和机制, 可以更好地理解这些疾病的发病机制, 发展更有效的治疗方法。

笔者所在团队构建的WT-PE技术可以实现染色体间易位^[40] (图4)。Yongsub Kim团队^[50]开发了基于引导编辑技术的染色体易位和倒位策略(PETI) (图4), 该技术方案与bi-WT-PE class 1设计一致。该团队基于PETI技术在人类细胞中创造了癌症相关的易位和倒位, 效率与Cas9相当, 但精确度要高于Cas9^[50]。这两种策略虽然可以实现与Cas9效率相当的编辑, 但是仍然存在大量的编辑副产物, 很有必要对精确度进行提升。

3.4 DNA大片段串联重复

基因扩增是生物进化过程的关键动力之一, 在

遗传性疾病和癌症的发病机制中起着关键的作用^[129-132]。基因扩增相关疾病的研究和治疗面临着极大的挑战, 因为对其进行建模是比较困难的, 动物模型的稀缺限制了针对此类突变的基因组编辑疗法的测试和开发。

笔者所在团队基于引导编辑系统策略开发了一种基因扩增技术TD-PE (图4), 该技术使用双pegRNA, 每种pegRNA都含有与待复制片段的5'端或3'端序列同源的RT模板。产生的ssDNA链经过重新退火, 形成一个含有部分串联重复的中间体, 然后通过内源性DNA修复机制进行修复, 实现串联重复^[51]。TD-PE可以实现50 bp~10 kb范围内精确的原位基因组片段扩增, 在测试的所有位点中, 效率最高可达28.33%, 同时通过微调pegRNA, 可以同时实现序列扩增和插入, 增加了TD-PE编辑方案的灵活性^[51]。此外, TD-PE在细胞水平实现了多种与疾病相关的串联重复, 证明TD-PE在遗传学研究中的应用前景。TD-PE的编辑精确度较高, 仅产生较低的indel水平。但是其会产生多个拷贝的串联重复, 从TD-PE的工作模式来看, 这是因为TD-PE不会破坏目标位点的序列(PAM和HA的序列), 这就允许TD-PE可以进行多次工作, 这是一把双刃剑, 一方面提供了多拷贝扩增目的序列的可能性, 另一方面也降低了编辑产物

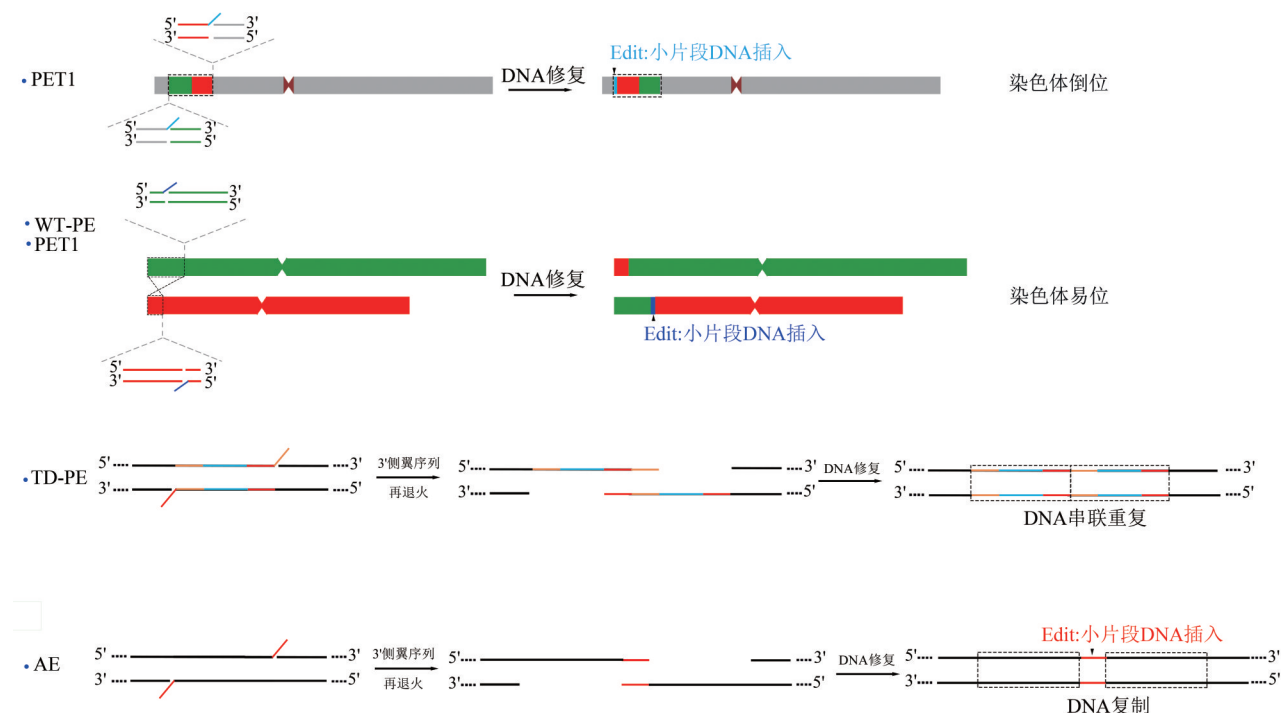


Fig. 4 Technologies of chromosome editing based on prime editing system

图4 基于引导编辑系统实现染色体编辑的技术

的特异性。

最近,殷昊团队^[52]基于类似的方案开发了一种名为 Amplification Editing (AE) 的编辑方案,可以实现 20 bp~100 Mb 的复制(图4)。AE 策略与 TD-PE 皆采用双 pegRNA 策略,但是 TD-PE 的 pegRNA 的 RT 模板是与待复制片段的 5'端或 3'端同源的序列,而 AE 的 RT 模板是新引入的与基因组无关的序列。两种编辑策略各有优缺点,首先 TD-PE 的编辑比较精确,而且通过微调 pegRNA 可以在基因组任意位置实现串联重复,而 AE 由于引入了外源序列在一定程度上破坏了原有的基因组序列。其次 AE 的编辑效率以及复制长度优于 TD-PE,但是 TD-PE 也有着实现染色体级别复制的潜力。

总的来说,两种编辑策略均可以实现 DNA 片段的精准复制,促进相关疾病模型的构建,给相关疾病的治疗带来了希望!

4 总结和展望

引导编辑系统代表了基因编辑领域的最新进展,它结合了 CRISPR/Cas9 的高效靶向性和逆转录酶的精准编辑能力,是近乎全能的编辑工具。该技术的发展为生命科学研究和医学治疗提供了巨大的潜力。

引导编辑系统的优势之一是其高度精准性。相比传统的 CRISPR/Cas9 技术,引导编辑系统能够更精确地定位到目标位点,并实现单碱基水平的编辑。这意味着科学家们可以更精细地调控基因表达,更准确地修复遗传性疾病的突变,从而为个性化医学和基因治疗提供了更多可能性。另一个引导编辑系统的优点是其对大片段基因编辑的潜力。传统的 CRISPR/Cas9 技术在 DNA 大片段编辑上存在挑战,而引导编辑系统能够通过 pegRNA 的设计,实现更大范围的 DNA 序列改变,包括基因组中的大片段插入、删除和扩增以及染色体的倒位易位等。这为研究人员提供了更多探索基因功能和基因组调控机制的机会,有望加速新药开发和疾病治疗的进展。

但是不可否认的是,引导编辑系统仍然存在一定的问题,尤其在体内基因编辑治疗的应用层面还存在编辑效率低以及递送困难。首先是编辑效率低的缺点,在本综述中已经探讨研究者们通过改造优化 pegRNA 和引导编辑蛋白以及调控内源性因素等提高编辑效率的策略,相信未来对引导编辑器的优

化和对编辑产物诱导的 DNA 损伤修复过程的深入解析必将更进一步地提高编辑效率。关于体内递送,目前已有基于腺相关病毒(adeno-associated virus, AAV)系统进行引导编辑器的局部^[101-102]或系统性递送^[93, 133]和基于类病毒颗粒局部递送^[134]等多个成功案例。然而,基于 AAV 的体内递送由于其长期表达的特性,有增加引导编辑系统脱靶编辑的风险。相信未来随着包括脂质纳米颗粒等非病毒载体以及包括慢病毒、逆转录病毒等类病毒载体技术的发展,引导编辑系统将有望更有效地应用于人类疾病的治疗。

未来,引导编辑系统还有许多发展方向和应用前景。一方面,随着技术的不断改进和优化,相信引导编辑系统的编辑效率和特异性将进一步提高,为更广泛的基因编辑应用打开更多可能性。另一方面,随着对基因组的深入理解和技术的进步,引导编辑系统有望在疾病治疗、农业改良、生物学研究等领域发挥更大的作用。特别是在个性化医学和精准基因治疗方面,引导编辑系统有望为患者提供更有效、更安全的治疗方案,成为未来生物医学领域的重要工具之一。

参 考 文 献

- [1] Jasin M. Genetic manipulation of genomes with rare-cutting endonucleases. *Trends Genet*, 1996, **12**(6): 224-228
- [2] Belfort M, Roberts R J. Homing endonucleases: keeping the house in order. *Nucleic Acids Res*, 1997, **25**(17): 3379-3388
- [3] Urnov F D, Miller J C, Lee Y L, *et al*. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 2005, **435**(7042): 646-651
- [4] Urnov F D, Rebar E J, Holmes M C, *et al*. Genome editing with engineered zinc finger nucleases. *Nat Rev Genet*, 2010, **11**(9): 636-646
- [5] Li T, Huang S, Jiang W Z, *et al*. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and FokI DNA-cleavage domain. *Nucleic Acids Res*, 2011, **39**(1): 359-372
- [6] Li T, Huang S, Zhao X, *et al*. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res*, 2011, **39**(14): 6315-6325
- [7] Jinek M, Chylinski K, Fonfara I, *et al*. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012, **337**(6096): 816-821
- [8] Cong L, Ran F A, Cox D, *et al*. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, **339**(6121): 819-823
- [9] Dever D P, Bak R O, Reinisch A, *et al*. CRISPR/Cas9 β -globin gene targeting in human haematopoietic stem cells. *Nature*, 2016, **539**(7629): 384-389

- [10] Mettananda S, Fisher C A, Hay D, *et al.* Editing an α -globin enhancer in primary human hematopoietic stem cells as a treatment for β -thalassemia. *Nat Commun*, 2017, **8**(1): 424
- [11] Nelson C E, Hakim C H, Ousterout D G, *et al.* *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science*, 2016, **351**(6271): 403-407
- [12] Rees H A, Liu D R. Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat Rev Genet*, 2018, **19**(12): 770-788
- [13] Kim K, Ryu S M, Kim S T, *et al.* Highly efficient RNA-guided base editing in mouse embryos. *Nat Biotechnol*, 2017, **35**(5): 435-437
- [14] Rossidis A C, Stratigis J D, Chadwick A C, *et al.* *In utero* CRISPR-mediated therapeutic editing of metabolic genes. *Nat Med*, 2018, **24**(10): 1513-1518
- [15] 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. *Nature*, 2016, **533**: 420-424
- [16] 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. *Nature*, 2017, **551**(7681): 464-471
- [17] Kurt I C, Zhou R, Iyer S, *et al.* CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat Biotechnol*, 2021, **39**(1): 41-46
- [18] Zhao D, Li J, Li S, *et al.* Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat Biotechnol*, 2021, **39**(1): 35-40
- [19] Tong H, Liu N, Wei Y, *et al.* Programmable deaminase-free base editors for G-to-Y conversion by engineered glycosylase. *Natl Sci Rev*, 2023, **10**(8): nwad143
- [20] Chen L, Hong M, Luan C, *et al.* Adenine transversion editors enable precise, efficient A•T-to-C•G base editing in mammalian cells and embryos. *Nat Biotechnol*, 2024, **42**(4): 638-650
- [21] Tong H, Wang X, Liu Y, *et al.* Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase. *Nat Biotechnol*, 2023, **41**(8): 1080-1084
- [22] Ye L, Zhao D, Li J, *et al.* Glycosylase-based base editors for efficient T-to-G and C-to-G editing in mammalian cells. *Nat Biotechnol*, 2024. DOI: 10.1038/s41587-023-02050-w
- [23] He Y, Zhou X, Chang C, *et al.* Protein language models-assisted optimization of a uracil-N-glycosylase variant enables programmable T-to-G and T-to-C base editing. *Mol Cell*, 2024, **84**(7): 1257-1270.e6
- [24] Anzalone A V, Randolph P B, Davis J R, *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 2019, **576**(7785): 149-157
- [25] Liu P, Liang S Q, Zheng C, *et al.* Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. *Nat Commun*, 2021, **12**(1): 2121
- [26] Jiang T, Zhang X O, Weng Z, *et al.* Deletion and replacement of long genomic sequences using prime editing. *Nat Biotechnol*, 2022, **40**(2): 227-234
- [27] Liu B, Dong X, Cheng H, *et al.* A split prime editor with untethered reverse transcriptase and circular RNA template. *Nat Biotechnol*, 2022, **40**(9): 1388-1393
- [28] Böck D, Rothgangl T, Villiger L, *et al.* Treatment of a metabolic liver disease by *in vivo* prime editing in mice. *bioRxiv*, 2021. DOI: 10.1101/2021.08.17.456632
- [29] Chemello F, Chai A C, Li H, *et al.* Precise correction of Duchenne muscular dystrophy exon deletion mutations by base and prime editing. *Sci Adv*, 2021, **7**(18): eabg4910
- [30] Park S J, Jeong T Y, Shin S K, *et al.* Targeted mutagenesis in mouse cells and embryos using an enhanced prime editor. *Genome Biol*, 2021, **22**(1): 170
- [31] Liu Y, Li X, He S, *et al.* Efficient generation of mouse models with the prime editing system. *Cell Discov*, 2020, **6**: 27
- [32] Lin J, Liu X, Lu Z, *et al.* Modeling a cataract disorder in mice with prime editing. *Mol Ther Nucleic Acids*, 2021, **25**: 494-501
- [33] Qian Y, Zhao D, Sui T, *et al.* Efficient and precise generation of Tay-Sachs disease model in rabbit by prime editing system. *Cell Discov*, 2021, **7**(1): 50
- [34] Lin Q, Zong Y, Xue C, *et al.* Prime genome editing in rice and wheat. *Nat Biotechnol*, 2020, **38**(5): 582-585
- [35] Jiang Y Y, Chai Y P, Lu M H, *et al.* Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol*, 2020, **21**(1): 257
- [36] Butt H, Rao G S, Sedeek K, *et al.* Engineering herbicide resistance *via* prime editing in rice. *Plant Biotechnol J*, 2020, **18**(12): 2370-2372
- [37] Li H, Li J, Chen J, *et al.* Precise modifications of both exogenous and endogenous genes in rice by prime editing. *Mol Plant*, 2020, **13**(5): 671-674
- [38] Xu R, Liu X, Li J, *et al.* Identification of herbicide resistance OsACC1 mutations *via* in planta prime-editing-library screening in rice. *Nat Plants*, 2021, **7**(7): 888-892
- [39] Shuto Y, Nakagawa R, Zhu S, *et al.* Structural basis for pegRNA-guided reverse transcription by a prime editor. *Nature*, 2024, **631**(8019): 224-231
- [40] Tao R, Wang Y, Hu Y, *et al.* WT-PE: prime editing with nuclease wild-type Cas9 enables versatile large-scale genome editing. *Signal Transduct Target Ther*, 2022, **7**(1): 108
- [41] Anzalone A V, Gao X D, Podracky C J, *et al.* Programmable deletion, replacement, integration and inversion of large DNA sequences with twin prime editing. *Nat Biotechnol*, 2022, **40**(5): 731-740
- [42] Zheng C, Liu B, Dong X, *et al.* Template-jumping prime editing enables large insertion and exon rewriting *in vivo*. *Nat Commun*, 2023, **14**(1): 3369
- [43] Sun C, Lei Y, Li B, *et al.* Precise integration of large DNA sequences in plant genomes using PrimeRoot editors. *Nat Biotechnol*, 2024, **42**(2): 316-327
- [44] Choi J, Chen W, Suiter C C, *et al.* Precise genomic deletions using paired prime editing. *Nat Biotechnol*, 2022, **40**(2): 218-226
- [45] Yarnall M T N, Ioannidi E I, Schmitt-Ulms C, *et al.* Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat Biotechnol*,

- 2023, **41**(4): 500-512
- [46] Wang C, Fang S, Chen Y, *et al.* High-efficiency targeted transgene integration *via* primed micro-homologues. *Cell Discov*, 2023, **9**(1): 69
- [47] Zhuang Y, Liu J, Wu H, *et al.* Increasing the efficiency and precision of prime editing with guide RNA pairs. *Nat Chem Biol*, 2022, **18**(1): 29-37
- [48] Wang J, He Z, Wang G, *et al.* Efficient targeted insertion of large DNA fragments without DNA donors. *Nat Methods*, 2022, **19**(3): 331-340
- [49] Tao R, Wang Y, Jiao Y, *et al.* Bi-PE: bi-directional priming improves CRISPR/Cas9 prime editing in mammalian cells. *Nucleic Acids Res*, 2022, **50**(11): 6423-6434
- [50] Kweon J, Hwang H Y, Ryu H, *et al.* Targeted genomic translocations and inversions generated using a paired prime editing strategy. *Mol Ther*, 2023, **31**(1): 249-259
- [51] Jiao Y, Li M, He X, *et al.* Targeted, programmable, and precise tandem duplication in the mammalian genome. *Genome Res*, 2023, **33**(5): 779-786
- [52] Zhang R, He Z, Shi Y, *et al.* Amplification editing enables efficient and precise duplication of DNA from short sequence to megabase and chromosomal scale. *Cell*, 2024, **187**(15): 3936-3952.e19
- [53] Jin S, Lin Q, Luo Y, *et al.* Genome-wide specificity of prime editors in plants. *Nat Biotechnol*, 2021, **39**(10): 1292-1299
- [54] Kim D, Kim D E, Lee G, *et al.* Genome-wide target specificity of CRISPR RNA-guided adenine base editors. *Nat Biotechnol*, 2019, **37**(4): 430-435
- [55] Kim D, Lim K, Kim S T, *et al.* Genome-wide target specificities of CRISPR RNA-guided programmable deaminases. *Nat Biotechnol*, 2017, **35**(5): 475-480
- [56] Lee H K, Willi M, Miller S M, *et al.* Targeting fidelity of adenine and cytosine base editors in mouse embryos. *Nat Commun*, 2018, **9**(1): 4804
- [57] Zuo E, Sun Y, Wei W, *et al.* Cytosine base editor generates substantial off-target single-nucleotide variants in mouse embryos. *Science*, 2019, **364**(6437): 289-292
- [58] 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
- [59] Lin Q, Jin S, Zong Y, *et al.* High-efficiency prime editing with optimized, paired pegRNAs in plants. *Nat Biotechnol*, 2021, **39**(8): 923-927
- [60] Qian Y, Wang D, Niu W, *et al.* Development of a highly efficient prime editor system in mice and rabbits. *Cell Mol Life Sci*, 2023, **80**(11): 346
- [61] Ponninselvan K, Liu P, Nyalile T, *et al.* Reducing the inherent auto-inhibitory interaction within the pegRNA enhances prime editing efficiency. *Nucleic Acids Res*, 2023, **51**(13): 6966-6980
- [62] Nelson J W, Randolph P B, Shen S P, *et al.* Engineered pegRNAs improve prime editing efficiency. *Nat Biotechnol*, 2022, **40**(3): 402-410
- [63] Liu Y, Yang G, Huang S, *et al.* Enhancing prime editing by Csy4-mediated processing of pegRNA. *Cell Res*, 2021, **31**(10): 1134-1136
- [64] Zhang G, Liu Y, Huang S, *et al.* Enhancement of prime editing *via* xrRNA motif-joined pegRNA. *Nat Commun*, 2022, **13**(1): 1856
- [65] Li X, Wang X, Sun W, *et al.* Enhancing prime editing efficiency by modified pegRNA with RNA G-quadruplexes. *J Mol Cell Biol*, 2022, **14**(4): mjac022
- [66] Feng Y, Liu S, Mo Q, *et al.* Enhancing prime editing efficiency and flexibility with tethered and split pegRNAs. *Protein Cell*, 2023, **14**(4): 304-308
- [67] Li X, Zhou L, Gao B Q, *et al.* Highly efficient prime editing by introducing same-sense mutations in pegRNA or stabilizing its structure. *Nat Commun*, 2022, **13**(1): 1669
- [68] Liang R, He Z, Zhao K T, *et al.* Prime editing using CRISPR-Cas12a and circular RNAs in human cells. *Nat Biotechnol*, 2024. DOI: 10.1038/s41587-023-02095-x
- [69] Zhang W, Petri K, Ma J, *et al.* Enhancing CRISPR prime editing by reducing misfolded pegRNA interactions. *Elife*, 2024, **12**: RP90948
- [70] Yan J, Oyler-Castrillo P, Ravisankar P, *et al.* Improving prime editing with an endogenous small RNA-binding protein. *Nature*, 2024, **628**(8008): 639-647
- [71] Huang S, Zhang Z, Tao W, *et al.* Broadening prime editing toolkits using RNA-Pol-II-driven engineered pegRNA. *Mol Ther*, 2022, **30**(9): 2923-2932
- [72] Liu M, Zhang X, Xu W, *et al.* Efficient and precise genomic deletion in rice using enhanced prime editing. *aBIOTECH*, 2024, **5**(2): 214-218
- [73] Chow R D, Chen J S, Shen J, *et al.* A web tool for the design of prime-editing guide RNAs. *Nat Biomed Eng*, 2021, **5**(2): 190-194
- [74] Li Y, Chen J, Tsai S Q, *et al.* Easy-Prime: a machine learning-based prime editor design tool. *Genome Biol*, 2021, **22**(1): 235
- [75] Hwang G H, Jeong Y K, Habib O, *et al.* PE-Designer and PE-Analyzer: web-based design and analysis tools for CRISPR prime editing. *Nucleic Acids Res*, 2021, **49**(W1): W499-W504
- [76] Anderson M V, Haldrup J, Thomsen E A, *et al.* pegIT-a web-based design tool for prime editing. *Nucleic Acids Res*, 2021, **49**(W1): W505-W509
- [77] Standage-Beier K, Tekel S J, Brafman D A, *et al.* Prime editing guide RNA design automation using PINE-CONE. *ACS Synth Biol*, 2021, **10**(2): 422-427
- [78] Hsu J Y, Grünwald J, Szalay R, *et al.* PrimeDesign software for rapid and simplified design of prime editing guide RNAs. *Nat Commun*, 2021, **12**(1): 1034
- [79] Bhagwat A M, Graumann J, Wiegandt R, *et al.* Multicrispr: gRNA design for prime editing and parallel targeting of thousands of targets. *Life Sci Alliance*, 2020, **3**(11): e202000757
- [80] Zheng C, Liang S Q, Liu B, *et al.* A flexible split prime editor using truncated reverse transcriptase improves dual-AAV delivery in mouse liver. *Mol Ther*, 2022, **30**(3): 1343-1351
- [81] Lan T, Chen H, Tang C, *et al.* Mini-PE, a prime editor with compact Cas9 and truncated reverse transcriptase. *Mol Ther Nucleic Acids*,

- 2023, **33**: 890-897
- [82] Zong Y, Liu Y, Xue C, *et al.* An engineered prime editor with enhanced editing efficiency in plants. *Nat Biotechnol*, 2022, **40**(9): 1394-1402
- [83] Gao Z, Ravendran S, Mikkelsen N S, *et al.* A truncated reverse transcriptase enhances prime editing by split AAV vectors. *Mol Ther*, 2022, **30**(9): 2942-2951
- [84] Chen P J, Hussmann J A, Yan J, *et al.* Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell*, 2021, **184**(22): 5635-5652.e29
- [85] Weber Y, Böck D, Ivaşcu A, *et al.* Enhancing prime editor activity by directed protein evolution in yeast. *Nat Commun*, 2024, **15**(1): 2092
- [86] Ni P, Zhao Y, Zhou X, *et al.* Efficient and versatile multiplex prime editing in hexaploid wheat. *Genome Biol*, 2023, **24**(1): 156
- [87] Lee J, Lim K, Kim A, *et al.* Prime editing with genuine Cas9 nickases minimizes unwanted indels. *Nat Commun*, 2023, **14**(1): 1786
- [88] Truong D J J, Geilenkeuser J, Wendel S V, *et al.* Exonuclease-enhanced prime editors. *Nat Methods*, 2024, **21**(3): 455-464
- [89] Liang Z, Wu Y, Guo Y, *et al.* Addition of the T5 exonuclease increases the prime editing efficiency in plants. *J Genet Genomics*, 2023, **50**(8): 582-588
- [90] Koblan L W, Doman J L, Wilson C, *et al.* Improving cytidine and adenine base editors by expression optimization and ancestral reconstruction. *Nat Biotechnol*, 2018, **36**(9): 843-846
- [91] Zafra M P, Schatoff E M, Katti A, *et al.* Optimized base editors enable efficient editing in cells, organoids and mice. *Nat Biotechnol*, 2018, **36**(9): 888-893
- [92] Kweon J, Yoon J K, Jang A H, *et al.* Engineered prime editors with PAM flexibility. *Mol Ther*, 2021, **29**(6): 2001-2007
- [93] Böck D, Rothgangl T, Villiger L, *et al.* *In vivo* prime editing of a metabolic liver disease in mice. *Sci Transl Med*, 2022, **14**(636): eabl9238
- [94] Oh Y, Lee W J, Hur J K, *et al.* Expansion of the prime editing modality with Cas9 from *Francisella novicida*. *Genome Biol*, 2022, **23**(1): 92
- [95] Han D, Xiao Q, Wang Y, *et al.* Development of miniature base editors using engineered IscB nickase. *Nat Methods*, 2023, **20**(7): 1029-1036
- [96] Karvelis T, Druteika G, Bigelyte G, *et al.* Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature*, 2021, **599**(7886): 692-696
- [97] Doman J L, Pandey S, Neugebauer M E, *et al.* Phage-assisted evolution and protein engineering yield compact, efficient prime editors. *Cell*, 2023, **186**(18): 3983-4002.e26
- [98] Adikusuma F, Lushington C, Arudkumar J, *et al.* Optimized nickase- and nuclease-based prime editing in human and mouse cells. *Nucleic Acids Res*, 2021, **49**(18): 10785-10795
- [99] Dirkx N, Weuring W J, De Vriendt E, *et al.* Increased prime edit rates in KCNQ2 and SCN1A via single nicking all-in-one plasmids. *BMC Biol*, 2023, **21**(1): 156
- [100] Grünewald J, Miller B R, Szalay R N, *et al.* Engineered CRISPR prime editors with compact, untethered reverse transcriptases. *Nat Biotechnol*, 2023, **41**(3): 337-343
- [101] She K, Liu Y, Zhao Q, *et al.* Dual-AAV split prime editor corrects the mutation and phenotype in mice with inherited retinal degeneration. *Signal Transduct Target Ther*, 2023, **8**(1): 57
- [102] Zhi S, Chen Y, Wu G, *et al.* Dual-AAV delivering split prime editor system for *in vivo* genome editing. *Mol Ther*, 2022, **30**(1): 283-294
- [103] Gao Z, Ravendran S, Mikkelsen N S, *et al.* A truncated reverse transcriptase enhances prime editing by split AAV vectors. *Mol Ther*, 2022, **30**(9): 2942-2951
- [104] Mu S, Chen H, Li Q, *et al.* Enhancing prime editor flexibility with coiled-coil heterodimers. *Genome Biol*, 2024, **25**(1): 108
- [105] Ferreira da Silva J, Oliveira G P, Arasa-Verge E A, *et al.* Prime editing efficiency and fidelity are enhanced in the absence of mismatch repair. *Nat Commun*, 2022, **13**(1): 760
- [106] Park J C, Kim Y J, Han J H, *et al.* MutS α and MutS β as size-dependent cellular determinants for prime editing in human embryonic stem cells. *Mol Ther Nucleic Acids*, 2023, **32**: 914-922
- [107] Liu X, Gu D, Zhang Y, *et al.* Conditional knockdown of OsMLH1 to improve plant prime editing systems without disturbing fertility in rice. *Genome Biol*, 2024, **25**(1): 131
- [108] Li X, Chen W, Martin B K, *et al.* Chromatin context-dependent regulation and epigenetic manipulation of prime editing. *Cell*, 2024, **187**(10): 2411-2427.e25
- [109] Liu N, Zhou L, Lin G, *et al.* Erratum: HDAC inhibitors improve CRISPR-Cas9 mediated prime editing and base editing. *Mol Ther Nucleic Acids*, 2022, **30**: 173
- [110] Chen R, Cao Y, Liu Y, *et al.* Enhancement of a prime editing system via optimal recruitment of the pioneer transcription factor P65. *Nat Commun*, 2023, **14**(1): 257
- [111] Li M, Zhong A, Wu Y, *et al.* Transient inhibition of p53 enhances prime editing and cytosine base-editing efficiencies in human pluripotent stem cells. *Nat Commun*, 2022, **13**(1): 6354
- [112] Song M, Lim J M, Min S, *et al.* Generation of a more efficient prime editor 2 by addition of the Rad51 DNA-binding domain. *Nat Commun*, 2021, **12**(1): 5617
- [113] Zhang H, Ma J, Wu Z, *et al.* BacPE: a versatile prime-editing platform in bacteria by inhibiting DNA exonucleases. *Nat Commun*, 2024, **15**(1): 825
- [114] Velimirovic M, Zanetti L C, Shen M W, *et al.* Peptide fusion improves prime editing efficiency. *Nat Commun*, 2022, **13**(1): 3512
- [115] Landrum M J, Lee J M, Benson M, *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*, 2016, **44**(D1): D862-D868
- [116] Hsu P D, Lander E S, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, **157**(6): 1262-1278
- [117] Anzalone A V, Koblan L W, Liu D R. Genome editing with CRISPR-Cas nucleases, base editors, transposases and prime editors. *Nat Biotechnol*, 2020, **38**(7): 824-844

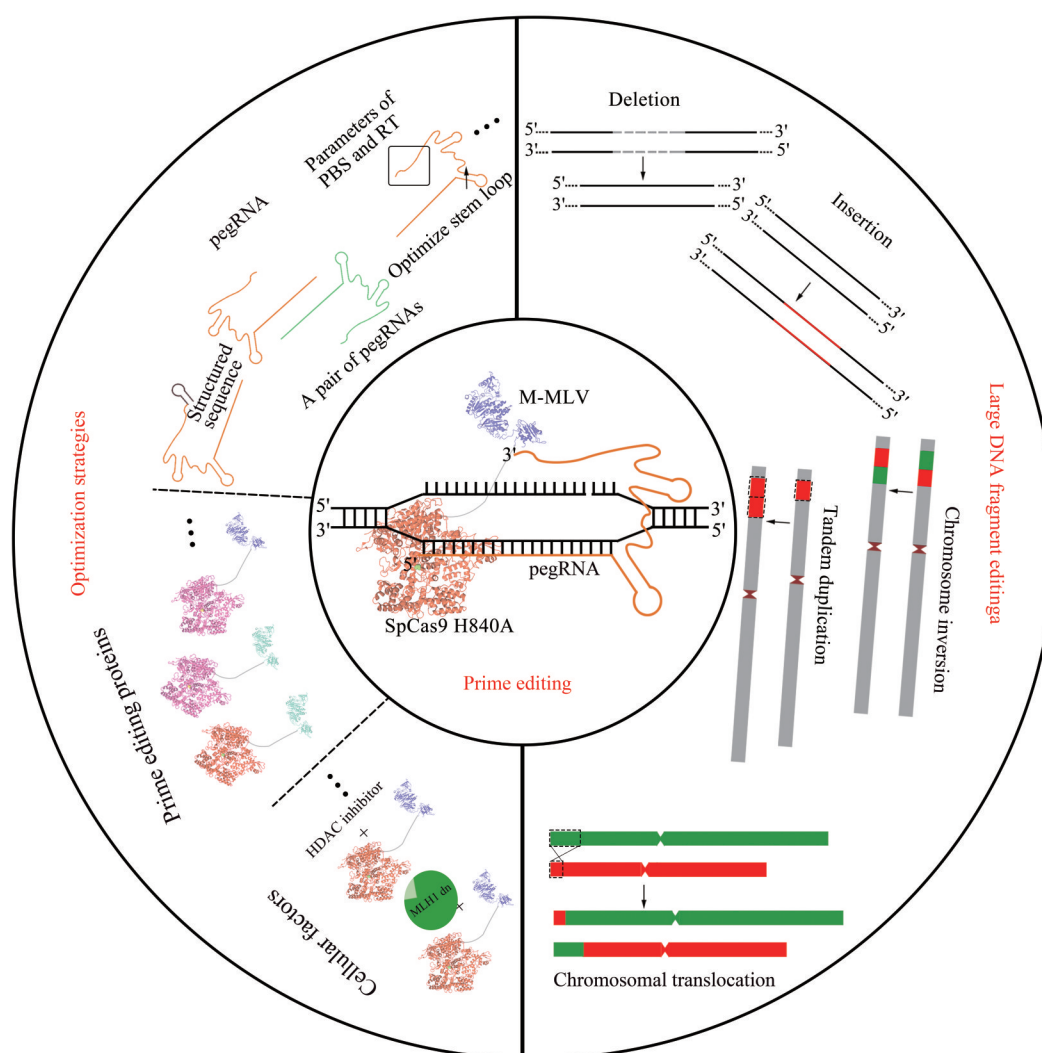
- [118] Wright A V, Nuñez J K, Doudna J A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell*, 2016, **164**(1/2): 29-44
- [119] Canver M C, Bauer D E, Dass A, *et al.* Characterization of genomic deletion efficiency mediated by clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9 nuclease system in mammalian cells. *J Biol Chem*, 2017, **292**(6): 2556
- [120] Kosicki M, Tomberg K, Bradley A. Repair of double-strand breaks induced by CRISPR-Cas9 leads to large deletions and complex rearrangements. *Nat Biotechnol*, 2018, **36**(8): 765-771
- [121] Song Y, Liu Z, Zhang Y, *et al.* Large-fragment deletions induced by Cas9 cleavage while not in the BEs system. *Mol Ther Nucleic Acids*, 2020, **21**: 523-526
- [122] Choi P S, Meyerson M. Targeted genomic rearrangements using CRISPR/Cas technology. *Nat Commun*, 2014, **5**: 3728
- [123] Guo Y, Xu Q, Canzio D, *et al.* CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell*, 2015, **162**(4): 900-910
- [124] Chapman J R, Taylor M R G, Boulton S J. Playing the end game: DNA double-strand break repair pathway choice. *Mol Cell*, 2012, **47**(4): 497-510
- [125] Geisinger J M, Stearns T. CRISPR/Cas9 treatment causes extended TP53-dependent cell cycle arrest in human cells. *Nucleic Acids Res*, 2020, **48**(16): 9067-9081
- [126] Wang H, Lu H, Lei Y S, *et al.* Development of a self-restricting CRISPR-Cas9 system to reduce off-target effects. *Mol Ther Methods Clin Dev*, 2020, **18**: 390-401
- [127] Pandey S, Gao X D, Krasnow N A, *et al.* Efficient site-specific integration of large genes in mammalian cells *via* continuously evolved recombinases and prime editing. *Nat Biomed Eng*, 2024. DOI: 10.1038/s41551-024-01227-1
- [128] Stankiewicz P, Lupski J R. Structural variation in the human genome and its role in disease. *Annu Rev Med*, 2010, **61**: 437-455
- [129] Roy A, Kumar V, Zorman B, *et al.* Recurrent internal tandem duplications of BCOR in clear cell sarcoma of the kidney. *Nat Commun*, 2015, **6**: 8891
- [130] Khater F, Langlois S, Cassart P, *et al.* Recurrent somatic BRAF insertion (p. V504_R506dup): a tumor marker and a potential therapeutic target in pilocytic astrocytoma. *Oncogene*, 2019, **38**: 2994-3002
- [131] Kottaridis P D, Gale R E, Frew M E, *et al.* The presence of a FLT3 internal tandem duplication in patients with acute myeloid leukemia (AML) adds important prognostic information to cytogenetic risk group and response to the first cycle of chemotherapy: analysis of 854 patients from the United Kingdom Medical Research Council AML 10 and 12 trials. *Blood*, 2001, **98**(6): 1752-1759
- [132] Lan X, Pritchard J K. Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals. *Science*, 2016, **352**(6288): 1009-1013
- [133] Davis J R, Banskota S, Levy J M, *et al.* Efficient prime editing in mouse brain, liver and heart with dual AAVs. *Nat Biotechnol*, 2024, **42**(2): 253-264
- [134] An M, Raguram A, Du S W, *et al.* Engineered virus-like particles for transient delivery of prime editor ribonucleoprotein complexes *in vivo*. *Nat Biotechnol*, 2024. DOI: 10.1038/s41587-023-02078-y

Optimization of Prime Editing System and Its Application in Large DNA Fragment Editing*

JIAO Yao-Ge, YAO Shao-Hua**

(Department of Biotherapy, Cancer Center and State Key Laboratory of Biotherapy, West China Hospital, Sichuan University, Chengdu 610041, China)

Graphical abstract



* This work was supported by grants from National Key Research and Development Program of China (2023YFC3403200) and Sichuan Science and Technology Program (2024NSFTD0029).

** Corresponding author.

Tel: 86-28-85503628, E-mail: shaohuayao@scu.edu.cn

Received: June 25, 2024 Accepted: August 20, 2024

Abstract Gene editing technology utilizes artificial nucleases to insert, replace, or delete specific sequences in desired genomic regions. The discovery of CRISPR/Cas9 nucleases was a milestone in the development of advanced gene editing tools, which revolutionized the field due to their simplicity and versatility. However, the limited precision of Cas9 nucleases remains a notable obstacle. Recently, derivative technologies such as prime editing have earned considerable attention for their enhanced efficiency and precision. The prime editing system consists of two components: the SpCas9 nickase (H840A) fused with reverse transcriptase (MLV-RT) and an engineered prime editing guide RNA (pegRNA). This system can irreversibly introduce various types of genetic changes into the genome, including 12 possible types of point mutations, as well as insertions, deletions and their combinations, without the need for DNA double-strand breaks (DSBs) or donor DNA templates. Prime editing offers several advantages in terms of editing accuracy, versatility, PAM constraints, and off-target effects. The editing results of prime editing system is highly accurate and can be tailored to specific needs. In addition, the system can be edited near or far from PAM sites, making it less constrained by PAM site restrictions. Moreover, it demonstrates high genome-wide specificity. The system also supports a variety of edits, demonstrating immense potential, especially in large DNA fragment editing—an area that relied heavily on CRISPR/Cas9 nucleases before. The development of prime editing, especially bi-direction prime editor, shed new light on large DNA fragment manipulations, including deletions, insertions, replacements, gene integration, as well as chromosomal translocations, inversions, and tandem duplications. Despite the significant progress made with prime editing technology, its application still faces challenges, especially low editing efficiency, which limits its potential in broader research and clinical settings. Consequently, researchers are exploring strategies to enhance the efficiency of prime editing. This review highlights several approaches to improving prime editing efficiency. These include optimizing pegRNA by refining PBS and RT parameters, increasing pegRNA stability and expression levels, and developing automated pegRNA design software. Additionally, efforts are being made to optimize the prime editing system proteins, such as screening for Cas9 and reverse transcriptase variants and performing codon optimization. The final aspect is the regulation of endogenous factors, including the inhibition of mismatch repair mechanisms and the modulation of chromatin environment. These approaches significantly enhance the practicality of prime editing in research and clinical contexts. In conclusion, prime editing represents a major advancement in the field of gene editing, offering powerful tools and methods for both basic research and clinical applications. This review will introduce the discovery, improvement and applications of prime editors, with a focus on prime editing mediated large DNA fragment manipulations. Hopefully, these insights will serve as valuable references for future research and applications of prime editing technology.

Key words CRISPR/Cas9, prime editing, pegRNA, large DNA fragment editing

DOI: 10.16476/j.pibb.2024.0266