



“藕断丝连”的CRISPR/Cas： 基因编辑中靶点滞留的作用与挑战*

冯依力^{1,2)*} 陈若丹^{1,2)} 谢安勇^{1,2)*}⁽¹⁾ 浙江大学医学院附属邵逸夫医院普外科, 杭州 310019; ⁽²⁾ 浙江大学转化医学研究院, 杭州 310029)

摘要 成簇规律间隔短回文重复 (clustered regulation interspaced short palindromic repeats, CRISPR) 和 CRISPR 相关蛋白质 (CRISPR-associated protein, Cas) 系统被广泛应用于基因组编辑、转录调控以及细胞实时成像等, 并已在农业、工业和医学等领域展示出巨大的应用潜力。该技术的应用取决于 CRISPR/Cas 的五大属性: 靶向、解旋、切割、滞留和旁切。本综述将主要以化脓链球菌 (*Streptococcus pyogenes*) 的 CRISPR/Cas9 为例, 聚焦于 CRISPR/Cas 的滞留属性, 梳理相关进展, 讨论其在基因编辑技术开发中的应用与挑战。

关键词 CRISPR/Cas9, 靶点滞留, 靶点解离, DNA 双链修复途径选择, 基因编辑异质性

中图分类号 Q78

DOI: 10.16476/j.pibb.2024.0274

成簇规律间隔短回文重复 (clustered regulation interspaced short palindromic repeats, CRISPR) 最初在细菌和古细菌的基因组中发现, 细菌和古细菌利用 CRISPR 及 CRISPR 相关蛋白质 (CRISPR-associated protein, Cas) 识别外源核酸物质并进行切割, 抵御外源核酸与噬菌体的入侵, 是细菌和古细菌适应性免疫的一个组成部分^[1]。2012 年, Jinek 等^[2] 将化脓链球菌 (*Streptococcus pyogenes*) 的 CRISPR/Cas9 系统的三个主要元件, 即 Cas9 核酸酶、CRISPR RNA (crRNA) 和反式激活的 crRNA (transacting crRNA, tracrRNA), 在体外组装成一个简单的、可以高效定点切割 DNA 靶点的核酸蛋白质复合物机器, 自此开启了 CRISPR 基因编辑技术的快速发展之路^[3-5]。如果没有特殊说明, 本综述所描述的 CRISPR/Cas9 主要指化脓链球菌的 CRISPR/Cas9。

CRISPR/Cas 似乎一直是基因编辑技术等待的“天选之子”, 它具有的几个内在属性, 包括靶向、解旋、切割、滞留和旁切, 为后续基因编辑技术的快速发展和应用提供了各种想象空间和可能 (图 1)。首先, CRISPR/Cas 的靶向性是 CRISPR/Cas 定点操作和应用的基础, 而且这个靶向性可控、易操作,

依赖该原理但不需要 CRISPR/Cas 切割活性的应用包括基于核酸酶活性缺失的 Cas (nuclease dead Cas, dCas) 的转录调控、表观修饰、实时影像以及联合 dCas 和转座酶的基因编辑等^[6-7]。其次, CRISPR 的解旋能力产生了单链 DNA (single-strand DNA, ssDNA), 非靶标链为碱基编辑器和引导编辑器分别提供了可靶向的 ssDNA 底物和可引导反转录的 ssDNA 引物, 这是碱基编辑器和引导编辑器工作的一个前提^[8]。而且, 加快 Cas 的解旋能力也能提高 CRISPR/Cas 的基因编辑效率^[9]。DNA 切割无疑是 CRISPR 基因编辑所依赖的一个关键属性。“不破不立”的基因编辑首先需要靶点 DNA 断裂的定点诱导, 随后依赖细胞内可操纵的 DNA 修复而产生基因编辑产物。而且, 碱基编辑器还需要靶标链的单链 DNA 断裂来定向控制修复以达到高效性, 引导编辑器则需要非靶标链 ssDNA 引物发生单链 DNA 断裂, 生成 3'-OH 来引

* 国家自然科学基金 (32371348, 32071439) 资助项目。

** 通讯联系人。

冯依力 Tel: 13968126080, E-mail: eric_feng@zju.edu.cn;

谢安勇 Tel: 13185001191, E-mail: anyongxie@zju.edu.cn

收稿日期: 2024-06-28, 接受日期: 2024-08-22

导DNA合成^[10-11]。断裂的非靶标链3'-ssDNA引物游离于CRISPR/Cas核酸复合物之外,以便与引导编辑指导RNA(prime editing guide RNA, pegRNA)的模板互补配对^[12]。

靶点滞留时间长是CRISPR的另一个独特的属性。比如,Cas9-sgRNA复合物不仅与DNA靶标结合数小时,而且在切割DNA后依旧与产物保持紧密结合^[13-16]。结合靶点后,CRISPR/Cas在靶点的滞留时间由CRISPR/Cas靶点结合与解离之间的动态平衡来调节,这个动态平衡的调节构造了CRISPR/Cas在靶点滞留这个属性,并决定了靶点滞留时间长短。这个滞留强度和时长受靶点序列、邻近的DNA代谢和染色质活动等的影响,可能影响dCas的应用效果^[13-16]。而且,CRISPR靶点滞留还发生在DNA切割后,我们的研究发现,这

种切割后的靶点滞留影响Cas9诱导的DNA双链断裂(double strand break, DSB)修复途径的选择^[17-18],操纵这个靶点滞留特性可以设计更高效、精准的基因编辑策略^[19-20]。最后需要提到CRISPR/Cas12a(也称为CRISPR/Cpf1)和CRISPR/Cas13a的旁切属性^[21]。CRISPR/Cas12a和CRISPR/Cas13a被靶点激活后,除了切割靶点,还能分别切割邻近的非靶点ssDNA和RNA。这个旁切特点被用于核酸检测技术的建立,包括DETECTR和SHERLOCK^[21-23]。除了这些属性,CRISPR演化过程中存在的微型元件也为后续微型编辑器的开发提供了可能^[24]。或许,目前还未发现的CRISPR属性将来会被挖掘,并将推动基因编辑技术的进一步发展。

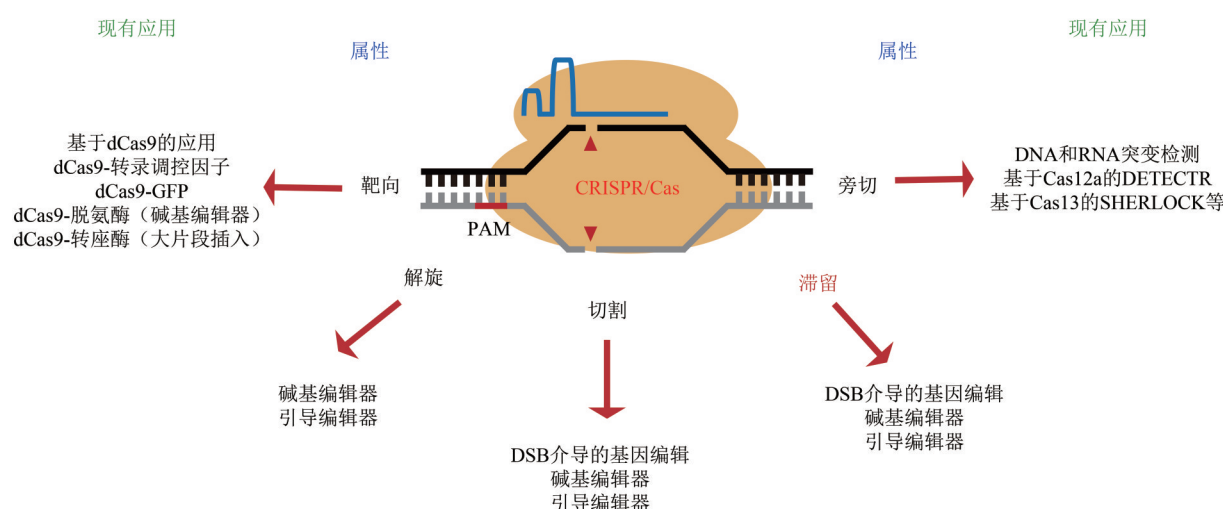


Fig. 1 Five key attributes of the CRISPR/Cas9 system and their application

图1 CRISPR/Cas9系统的五大属性及其参与的应用

CRISPR/Cas9具有五大独特属性,包括DNA靶向性、DNA解旋活性、DNA切割能力、靶点滞留特性以及旁切,这些属性在CRISPR工具应用中具有主导的或协助的作用。PAM: 间区前体邻近基序; DETECTR: DETECTR诊断系统; SHERLOCK: SHERLOCK检测技术。

1 CRISPR/Cas基因编辑技术工作原理

在众多CRISPR/Cas系统中,CRISPR/Cas9和CRISPR/Cas12a是最早开发的两个CRISPR基因编辑系统,这两个系统均包含一个Cas核酸酶和一条单指导RNA(single guide RNA, sgRNA)^[3, 25]。在广泛使用的化脓链球菌的CRISPR/Cas9系统中,sgRNA是由crRNA和tracrRNA元件融合而成^[2-4]。在CRISPR基因编辑中,还未结合sgRNA的未激活的Cas蛋白与sgRNA结合后,首先需要在庞大的基

因组中找到其靶点^[26]。CRISPR/Cas主要通过三维碰撞(3D diffusion)和一维扩散(1D diffusion)二级模式搜索靶标^[13, 27],当Cas蛋白的间区前体邻近基序(proto-spacer adjacent motif, PAM)结合结构域(PAM-interacting domain, PID)识别并结合潜在靶标的PAM时,与PAM结合的Cas蛋白将解旋PAM旁边的靶标DNA双链,允许crRNA中的间区与解开的靶标单链尝试互补配对,如果不能配对,CRISPR/Cas将脱离该位点继续搜寻,如果互补配对成功,将产生一个独特的R-loop结构,此

时, 不仅 Cas 蛋白通过氢键与 R-loop 中的 DNA 链和 RNA 链相互作用, 而且, R-loop 中的 DNA-RNA 杂交链也将帮助稳定 Cas-sgRNA-DNA 三元复合体^[26, 28]。

Cas 介导的 DNA 切割将在相对稳定的 Cas-sgRNA-DNA 三元复合体发生。以化脓链球菌 Cas9 (即 *SpCas9*) 为例, Cas9 的两个核酸酶结构域 HNH 和 RuvC 分别针对 DNA 靶标链和非靶标链切割-NGG PAM 上游第三个和第四个核苷酸之间的磷酸二酯键, 形成平末端 DSB^[2-3]。但也有数据表明, 或许受靶标核酸序列的影响, 也可能是邻近 DNA 代谢或染色质环境的干扰, RuvC 在非靶标链上的切割位点会因此而改变, 但 HNH 结构域切割位点维持不变, 从而产生 5'黏性末端^[29-32]。DNA 切割完成后, CRISPR/Cas 仍会滞留在切割的靶点上, 掩盖了 DNA 损伤。体外实验表明, Cas9-sgRNA 复合物在切割 DNA 后数小时内仍与靶 DNA 紧密结合^[13-14, 20]。细胞内, 只有在 DNA 断裂点暴露后 DNA 损伤应答途径才会激活, 而 CRISPR/Cas 的靶点解离机制以及 Cas 诱导的 DNA 断裂暴露过程均会影响后续的 DNA 修复途径选择以及最终的基因编辑产物。

一旦 CRISPR/Cas 在基因组上定点诱导 DNA DSB, 细胞内在的 DSB 修复机制将进行修复, 产生包含所需基因编辑产物的修复产物, 实现对靶标基因的编辑, 包括基因敲除 (knockout, KO) 以及基因敲入 (knock-in, KI)^[33]。在真核生物中, 细胞内 DSB 修复主要依赖两条进化保守的修复途径: 同源序列导向修复 (homology-directed repair, HDR) 或称同源重组 (homologous recombination, HR) 和非同源末端连接 (non-homologous end joining, NHEJ)。HDR 依赖同源序列模板, 在细胞周期的 S 和 G2 期发生, 通常以姊妹染色单体等位同源序列为同源模板进行 DNA 合成和修复, 恢复损伤或丢失的序列, 常被认为是一条高保真的修复途径。而 NHEJ 则是将两个 DNA 断裂末端连接起来, 不需要同源模板的帮助, 可以发生在细胞周期的各个阶段, 修复过程中会发生碱基插入或丢失, 常被认为是一条易错修复途径。NHEJ 又分为经典型 NHEJ (classical NHEJ, c-NHEJ) 和替代型末端连接 (alternative end joining, a-EJ)^[34]。作为主要的 NHEJ 方式, c-NHEJ 需要核心 NHEJ 因子组成的修复机器进行修复, 这些核心因子包括 DNA-PKcs、Ku70/Ku80 和 XRCC4/DNA 连接酶 4 等。如

果 DSB 末端是可以直接连接的互补黏性末端或平末端, c-NHEJ 的修复可以是精准高效的。我们的研究表明, 在哺乳动物细胞中, Cas9 和 Cas12a 诱导的 DSB 的 c-NHEJ 修复精准度高达 100% (平均约为 50%) (未发表数据)^[29]。NHEJ 修复中如果至少一个 c-NHEJ 核心因子没有参与, 我们称这种 NHEJ 修复为 a-EJ^[35]。a-EJ 偏好使用 2~25 个核苷酸的微同源序列来介导 DNA 末端的连接。细胞还可以通过末端的微同源序列或同源序列来选择其他两条修复途径进行修复, 即微同源序列介导的末端连接 (microhomology-mediated end joining, MMEJ) 和单链退火途径 (single strand annealing, SSA)^[36]。

因此, 基因编辑时 CRISPR 切割产生的靶点 DSB 会被不同的 DSB 修复途径竞争修复, 有的途径有利于基因编辑产物的形成, 有的途径正好相反。针对一个特定 DSB 位点, 影响修复途径选择的因素很多, 包括细胞周期、DSB 末端结构和核苷酸组成、邻近的 DNA 代谢、周围的染色质结构和活动以及细胞内特定修复途径因子的可利用性等^[20, 33]。考虑到细胞周期是 DSB 修复途径选择的一个主要决定因素, 使用细胞周期调控的小分子药物, 可以将细胞暂时停留在 S 和 G2 期, 提高 CRISPR 基因编辑中的 HDR 效率, 也可以将细胞暂时停留在 G1 期, 推动基于 NHEJ 的基因编辑^[33, 37]。甚至, CRISPR/Cas9 对靶点的切割会激活人造血干细胞和原代细胞的 p53, 诱导 G1 期的停滞, 抑制 HDR 活性^[38-39]。DSB 末端结构和核苷酸组成也可以决定 DSB 修复途径的选择以及 NHEJ 修复精准度, 比如末端微同源序列的存在和长度将影响该末端修复的 NHEJ、MMEJ 和 SSA 的参与度, 有人已尝试利用微同源序列来引导特定编辑产物的生成^[33, 37]。邻近的 DNA 复制与转录还可能影响滞留在切割靶点的 CRISPR/Cas 的解离, 并产生复制或转录偶联的 DSB, 这类 DSB 的修复也有所不同^[20]。此外, 周边染色质的结构和活动也会影响 CRISPR/Cas 的解离、修复因子 (repair factors) 的招募和 DSB 末端的加工, 从而决定 Cas 诱导的 DSB 的不同修复途径的选择^[40-42]。有证据表明, 组蛋白的伴侣蛋白复合物 FACT 帮助 CRISPR/Cas 在切割靶点的解离, 调控不同修复途径的选择^[43]。细胞本身不同修复途径的可利用性也决定了 Cas 诱导的 DSB 修复途径选择, 比如不分裂细胞不能利用 HDR 进行基因敲入的基因编辑。

即使不诱导DSB, 核酸酶失活(即dCas)或部分失活的Cas蛋白(比如包括Cas9 D10A和Cas9 H840A的Cas9 缺刻酶nCas9), 与其sgRNA伴侣, 也被用于定点靶向和滞留, 同时利用与Cas融合或sgRNA招募的效应子实现定点调控转录、实时定位、大片段靶向插入、定点碱基编辑和定点引导编辑等应用^[6-8]。目前用于与dCas融合的转录调控效应子包括转录抑制因子KRAB、转录激活因子VP64、DNA甲基转移酶DNMT3A和DNMT3L、组蛋白乙酰转移酶p300/CBP等, 而通过改造sgRNA, 联合RNA结合蛋白, 也可招募上述转录调控效应子, 调控转录^[7, 44]。类似地, 如果将转录调控效应子转换为荧光标记蛋白, 如GFP, dCas-sgRNA系统也可以用于细胞内DNA和RNA的实时影像追踪^[45]。借助dCas的靶向性, 也可将与之融合或结合的转座酶带到目标位点, 通过转座酶的作用, 最高可将10万个碱基对的DNA直接转座到基因组^[46-48]。

碱基编辑器的早期开发也是利用核酸酶完全失活的dCas9, 将胞苷脱氨酶(如APOBEC3)与之融合, 在dCas9解链时胞苷脱氨酶将R-loop中单链DNA的胞嘧啶脱氨为U, 随后的DNA合成有一定概率以U为模板与A配对, 实现C到T的转换, 即胞嘧啶碱基编辑器(cytosine base editors, CBEs)的编辑结果^[10]。进一步改良的CBEs采用Cas9 D10A突变体切割靶标链, 推动以已脱氨的非靶标链为模板进行DNA合成, 提高C到T的转换; 同时融合尿嘧啶DNA糖基化酶抑制蛋白(uracil glycosylase inhibitor, UGI), 防止靶向U的碱基切除修复, 提高碱基编辑的效率^[10]。利用改造的腺苷脱氨酶(如TadA), 后续也建立了腺嘌呤碱基编辑器(adenine base editors, ABEs), 实现腺嘌呤(adenine, A)到鸟嘌呤(guanine, G)的转换^[49]。联合DNA糖基化酶而不是DNA糖基化酶抑制蛋白也可实现碱基颠换, 即从C到G、A到C或A到T的碱基修饰^[7-8]。在此过程中, 脱氨的碱基将被DNA糖基化酶切除, 产生无碱基位点, 细胞内激活的碱基切除修复有机率产生碱基颠换。不需要脱氨酶, 直接将改造的糖基化酶与Cas9 D10A融合, 也实现了碱基转换和颠换, 拓宽了碱基修饰的范围^[50-51]。

引导编辑器则是利用nCas9(比如Cas9 H840A)切割非靶标链, 释放的非靶标链3'-ssDNA与pegRNA上的引物结合位点互补配对, 在与

nCas9融合的逆转录酶作用下, 使用pegRNA中的目标序列作为模板合成DNA^[11]。随后的DNA修复将新合成的目标序列永久整合到基因组中, 实现引导编辑。基于nCas9开发的编辑器只产生单链缺刻, 不诱导DSB, 不仅降低了脱靶效应, 同时也避免了准靶产生DSB的危害, 因此被认为安全性更高^[6-8]。然而, 如果与DNA复制偶联, 碱基编辑器和引导编辑器中的nCas9产生的单链缺刻也可能产生更具毒性的单末端DSB及其诱导的异常染色体重排^[18]。

2 CRISPR/Cas的靶点滞留及其解离

CRISPR/Cas与靶点的结合与解离是一个动态平衡, 这个动态平衡的调节构造了CRISPR/Cas靶点滞留这个属性, 并决定了靶点滞留时间长短。除了Cas-sgRNA-DNA三元复合物内在的分子相互作用及相关的构象变化参与调节这个动态平衡之外, 邻近的DNA代谢和染色质活动也会带来明显的影响。更为特别的是, 在切割靶点后CRISPR/Cas与靶点的结合与解离机制也与未切割靶点的结合与解离存在差异, 赋予了靶点滞留与解离更高的复杂性。

2.1 靶点切割后的CRISPR/Cas滞留特性

Cas9-sgRNA复合物不仅与DNA靶标结合数小时, 而且在切割DNA后依旧与产物保持紧密结合^[13-16]。Cas9-sgRNA复合物将切割后的DSB包裹在复合物内, 使其无法启动DNA损伤应答。Cas9-sgRNA复合物只有离开靶位点并暴露DSB后, 可能才启动DNA损伤应答以及修复途径选择。Cas9-sgRNA复合物长时间滞留在靶位点, 可能会增加其与DNA复制或转录机器发生碰撞的可能性, 从而造成复制叉停滞或转录相关R-loop的积累^[52]。值得注意的是, 与Cas9-sgRNA一样, 失活的Cas9(即dCas9)与靶位点结合也会产生R-loop结构。类似于转录过程中产生的R-loop, 如果大量产生, 该结构可能诱导基因组不稳定^[53]。事实上, 我们观察到dCas9在基因组上的大量结合会诱导高频的DNA损伤, 造成异常的染色体结构, 特别是在BRCA1缺陷细胞中会诱导星状染色体的形成^[18]。

在Cas9-sgRNA从靶位点释放之前, 被切割的非靶标链的3'端会提前从Cas9-sgRNA-DNA三元复合物中释放^[14, 26]。这个游离的ssDNA末端可以与外源提供的单链寡核苷酸(single-strand oligodeoxynucleotide, ssODN)模板进行部分

互补退火, 推动高效率的HDR^[14]。这种ssODN介导的HDR不依赖RAD51, 但需要范可尼贫血(Fanconi anemia, FA)途径的参与, 该途径通常与解开链间交联(interstrand cross-links, ICLs)的过程相关^[52, 54]。此外, 从Cas9-sgRNA-DNA三元复合物中释放的被切割的非靶标链的3'端可以被过表达的外源TREX2核酸外切酶或与Cas9融合的TREX2降解^[55]。游离的TREX2核酸外切酶可降解非靶标链的3'端少则5个核苷酸, 多则8个, 而与Cas9融合的TREX2可降解9个或更多的核苷酸^[55]。TREX2介导的3'-ssDNA末端的降解可用于提高Cas9介导的基因敲除效率和配对Cas9 H840A缺刻酶介导的基因敲除效率^[55-56]。

与SpCas9-sgRNA复合物不同, Cas12a-sgRNA复合物和金黄色葡萄球菌Cas9(即SaCas9)-sgRNA复合物在切割靶DNA后, 释放PAM远端的双链DNA末端, 但仍与PAM近端结合^[57-59]。尽管释放PAM远端的机制可能有所不同, Cas12a-sgRNA和SaCas9-sgRNA切割产生的两个末端的不对称释放, 将对DSB的修复和基因编辑产生一定影响。例如, Cas12a切割DNA后, DNA游离端更容易被c-NHEJ因子结合, 以保证连接的效率和精准度, 而Cas12a滞留端则因为Cas12-sgRNA的占位而阻止c-NHEJ因子的招募, 连接发生前末端易被加工, 从而产生更长的碱基丢失(数据未发表)。因此, 利用Cas12a切割产生的互补黏性末端, 联合不对称滞留而释放的PAM远端, 我们开发了CIPDEL(Cas12a-induced PAM-distal ligation)技术来实现基于c-NHEJ的、高效、精准的基因敲入(未发表数据)。该技术不仅有利于大片段插入, 而且可以在HDR活性低甚至缺失的细胞(比如神经元、骨骼肌细胞、心肌细胞、植物细胞等)中实现基因敲入(未发表数据)。

2.2 Cas9-sgRNA-DNA三元复合物中的分子相互作用

Cas9-sgRNA-DNA三元复合物的形成和稳定主要依赖于三个方面的分子相互作用: a. PID与PAM基序的分子相互作用; b. sgRNA间区序列与靶标链序列的互补配对; c. Cas9蛋白与DNA靶标链和非靶标链序列的非特异性分子相互作用。首先, SpCas9-sgRNA通过三维碰撞和一维扩散的两级搜索方式寻找PAM位点, 利用SpCas9的PID的两个保守精氨酸残基(R1333和R1335)识别标准PAM序列(5'-NGG-3'), 并与非靶标链上的GG二核苷

酸大沟相互作用^[26, 28]。识别PAM序列后, SpCas9与PAM小沟的额外相互作用帮助SpCas9解开目标DNA双链。目标DNA的定向解链促使sgRNA的间区序列与靶标单链配对, 推动R-loop结构形成和DNA切割^[26, 60]。因此, 作为Cas9靶点结合和DNA切割的先决条件, 改变PAM的识别可以扩大PAM的兼容性, 拓宽CRISPR/Cas9系统的靶向范围。的确, 通过理性设计或蛋白质进化, 直接或间接改造与PAM GG二核苷酸相互作用的Cas9关键氨基酸残基, 已成功扩展了CRISPR/Cas系统的PAM识别谱^[61-63]。

在目标DNA的定向解链中, 近PAM端的10~12个碱基对的种子区域最早解开, 并决定了靶点结合特异性^[13, 26]。R-loop结构中, sgRNA间区序列与靶标链种子序列间的完全匹配促使Cas9核酸酶构象变化, 触发靶标链和非靶标链的协同切割^[64-65]。间区序列与种子序列间的错配可能会抑制整个sgRNA间区序列与靶标链的退火和完整R-loop的形成, 显著降低甚至完全消除Cas9-sgRNA的靶点结合力以及DNA切割活性, 而种子区域之外更能容忍与sgRNA间区序列碱基的错配^[66-68]。考虑到sgRNA的间隔区和靶标链之间的碱基配对是Cas9-sgRNA-DNA复合物稳定性的一个主要决定性因素, sgRNA间区长短可能影响Cas9-sgRNA与目标DNA的结合以及Cas9的靶标特异性。但是, 与典型的20个核苷酸间区相比, 延长间区序列并没有提高Cas9的靶标亲和力和特异性, 相反, 间区截短到17~19个核苷酸可以最大限度地减少脱靶效应^[69]。这表明20个核苷酸的间区可能已经为Cas9-sgRNA的靶点结合和Cas9核酸内切酶活性提供了足够的亲和力。

除了PID与PAM的相互作用和目标双链DNA解旋后靶标单链与sgRNA间区的互补配对, Cas9和目标DNA之间的非特异性结合也增加Cas9-sgRNA-DNA三元复合物的稳定性, 并可额外调控Cas9-sgRNA-DNA复合物的构象改变和Cas9核酸酶活性。不同构象状态的Cas9-sgRNA-DNA复合物发挥不同的功能, 如靶点识别、R-loop形成、切割DNA链时的催化反应和靶点解离。每种构象状态都需要依赖Cas9与DNA之间不同的结合方式才能发挥各自的功能^[26, 64-65]。Cas9-sgRNA与靶点结合导致催化前构象变化, 并打开Cas9的中央通道, 以便容纳新形成的R-loop结构。在这个中央通道结构中, REC叶中REC3结构域和靶标链的PAM

远端DNA之间相互作用^[28]。在另一种催化前构象状态中,大量的氢键和静电引力将带负电荷的非靶标链稳定在HNH、RuvC和PID之间的带正电荷的凹槽中^[65, 70]。为了构建出既保留核酸酶活性,同时又能有效降低脱靶效应的Cas9变体,可以改变这个凹槽中的氨基酸残基来降低非特异性DNA结合^[71-73]。此外,对Cas9-sgRNA-DNA结构的解析揭示了Cas9与靶标链骨架之间直接的氢键和疏水作用,通过改变这些相互作用可以降低Cas9-sgRNA的脱靶^[28, 70, 74]。当Cas9切割DNA时,其HNH结构域会转变为具有催化能力的构象,其中D839、H840和N863等几个关键氨基酸残基的侧链可以与靶标链形成氢键^[65]。在催化后的构象中,先前构象的几种相互作用仍然存在,包括REC3和RNA-DNA杂交链之间的相互作用。因为REC3和RNA-DNA杂交链之间的相互作用有助于HNH核酸酶的激活,因此改变这种相互作用也可用于提高Cas9的特异性^[73]。催化后构象表现出的强相互作用表明,即使在DNA断裂后,Cas9-sgRNA与切割后的DNA紧密结合,在目标DNA上的滞留时间也很长^[65]。由于dCas9不切割DNA,因此dCas9可能没有类似于Cas9的催化后构象变化,其与DNA的解离方式也可能与Cas9不同。

2.3 CRISPR/Cas的靶点解离

在CRISPR/Cas基因编辑中,CRISPR/Cas的靶点滞留还需要与靶点解离维持一个精细的物化动态平衡,才能实现基因编辑的高效和精准(图2)。真核生物中,CRISPR/Cas基因编辑的靶标DNA被包裹在染色质的结构中,其中基因组DNA环绕“核心”组蛋白八聚体形成核小体,并进一步包装成更紧凑的结构单元。这种紧凑的结构不仅严重阻碍了Cas9对目标DNA的搜索和结合,而且结合后的染色质动态变化也可能影响Cas9-sgRNA在靶点的滞留和解离^[40]。同时,靶点滞留的Cas9-sgRNA也不可避免会遭遇邻近的DNA复制与转录机器^[20]。这些动态变化联同不同靶点序列差异,使得Cas9-sgRNA靶点滞留与解离更为复杂。但Cas9-sgRNA的解离方式可以简单归纳为如下两种方式:自发解离和被动解离(图2)。

自发解离是Cas9-sgRNA内在有序地解除与目标DNA的相互作用而脱离目标DNA(图2),但自发解离的机制目前还不清楚。一个可能的模型是,在Cas9-sgRNA-DNA三元复合物中,从PAM远端开始,R-loop内伴随RNA-DNA杂交链解旋而启动

靶标链和非靶标链重杂交,sgRNA间区序列与靶标链序列的互补配对逐渐消除,Cas9蛋白与DNA靶标链和非靶标链序列的非特异性分子相互作用也逐渐消失,于是,Cas9-sgRNA逐渐从PAM远端与目标DNA脱离,最后从R-loop消亡的不稳定的Cas9-sgRNA-DNA三元复合物中释放^[59]。在*Sp*Cas9-sgRNA-DNA三元复合物中,非靶标链游离出复合物外,这为引导编辑器的开发提供了便利^[11-12],但可以预测,可能因为非靶标链难以与靶标链重杂交,*Sp*Cas9-sgRNA的自发解离将难于*Sa*Cas9-sgRNA或Cas12a-sgRNA,这也可以帮助解释*Sp*Cas9在靶点上的周转远远慢于*Sa*Cas9^[13, 59]。

被动解离则是Cas9-sgRNA因为邻近的DNA代谢或染色质活动而被迫从目标DNA上脱离,这种解离可能是混乱无序的(图2)。一方面,DNA代谢(包括复制、转录、损伤和修复等)和染色质重塑产生的机械力会改变DNA的拓扑结构^[20, 75]。例如,研究表明,邻近的DNA扭转和拉伸影响Cas9-sgRNA-DNA三元复合物内R-loop的稳定性和Cas9的切割特异性^[60, 76-77]。PAM下游行进而来的Bloom综合征解旋酶(Bloom's syndrome helicase, BLM)比PAM上游来的BLM更容易取代dCas9^[78]。组蛋白分子伴侣FACT也可以通过破坏核小体的结构解离Cas9,而其他分子伴侣和重塑因子也可能参与这个过程^[43]。因此,对DNA施加的机械压力、邻近的DNA代谢和染色质的重塑活动,都可能会影响RNA-DNA杂交链中的碱基配对,以及Cas9与PAM基序和靶DNA的相互作用,从而破坏Cas9-sgRNA-DNA三元复合物的稳定性,甚至直接将Cas9-sgRNA从靶位点上移除。另一方面,除了产生DNA扭转或拉伸等机械力外,DNA复制叉、转录或修复机器可能直接与Cas9-sgRNA-DNA三元复合物碰撞,强制性地使Cas9-sgRNA从这个三元复合物中“撞”出靶点^[18, 79-81]。真核生物DNA复制的六聚体解旋酶和DNA聚合酶之间的协同作用,可能会从任意一个DNA复制方向与Cas9-sgRNA复合体产生强烈的碰撞^[18, 82]。然而,噬菌体Phi29 DNA聚合酶(DNA polymerase, DNAP)和T7 RNA聚合酶(RNA polymerase, RNAP)在体外将dCas9从DNA上解离时具有链偏好性,因此单靠碰撞力可能不足以将Cas9-sgRNA与靶DNA分离,可能还与DNAP或RNAP如何解开Cas9-sgRNA-DNA三元复合物内的RNA-DNA杂交链有关^[78, 81]。

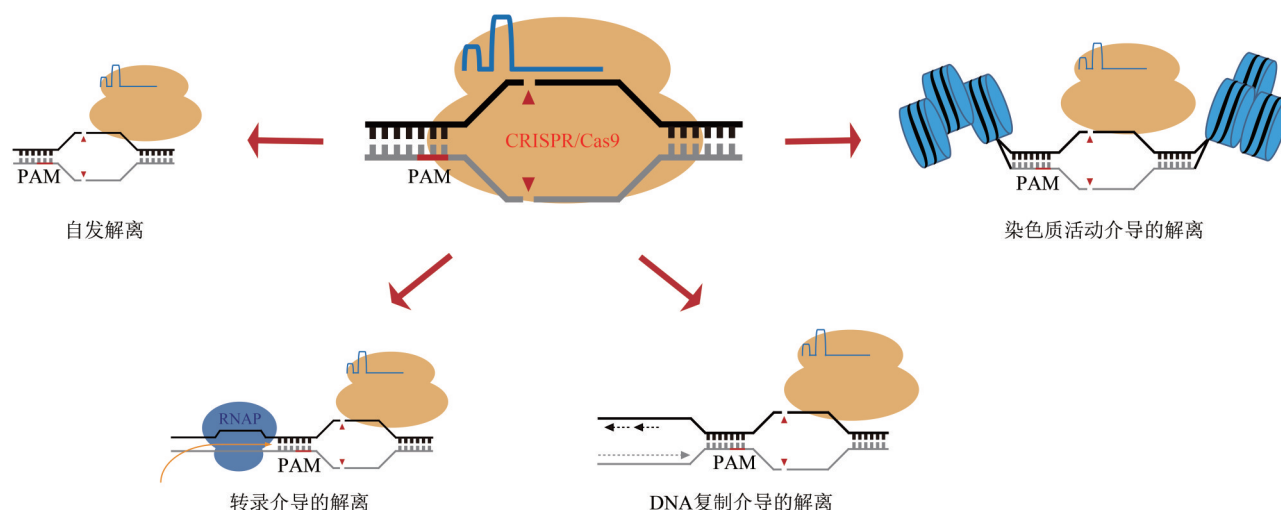


Fig. 2 Post-cleavage target residence and dissociation of CRISPR/Cas9

图2 靶点切割后CRISPR/Cas9的靶点滞留与解离

靶点滞留的Cas9-sgRNA在DNA切割后可自发解离, 也可能会遭遇邻近DNA代谢(包括DNA复制和转录)和染色质活动等事件而被动解离。PAM: 间区前体邻近基序。

3 CRISPR/Cas靶点滞留对DSB修复的影响

在CRISPR/Cas基因编辑中, 无论是DSB介导的基因编辑还是借助于单链缺刻的碱基编辑和引导编辑, CRISPR/Cas靶点滞留这个属性无疑将影响编辑产物的形成。其中, 主要的决定因素是选定位点的靶点滞留与解离的动态平衡和解离后DNA断裂的识别与修复。因为不同解离方式将产生不同构象的DNA断裂末端, 有的解离方式还限制在特定的细胞周期, 最终, 细胞将根据这些因素来决定DNA断裂修复途径的选择。每个细胞不同修复途径的选择不仅在准靶位点导致编辑产物的高度异质性, 甚至引起严重的染色体重排, 而且在不同脱靶位点影响脱靶修复产物的生成。因此, 理解CRISPR/Cas靶点滞留对DSB修复的影响有助于改进基因编辑技术, 提高该技术的效率和精准度, 拓展应用范围。

3.1 CRISPR/Cas靶点滞留对DNA损伤应答的影响

当DSB产生时, 哺乳动物细胞会采用一套进化上保守的DNA损伤应答信号通路识别并修复这些DSB。首先, DSB会被细胞内早期损伤应答因子识别, 这些早期损伤应答因子包括Ku70/Ku80/DNA-PKcs、Mre11-Rad50-NBS1 (MRN)/ATM和RPA/ATRIP/ATR三个竞争性复合物, 与断裂DNA

末端的结合将激活其中三个PI3KK的活性, 从而磷酸化下游的DNA损伤应答信号分子, 包括磷酸化组蛋白H2A变体H2AX的C端SQEY基序而形成 γ H2AX^[33, 83]。一部分信号在缺乏核小体的末端邻近区域传递和执行(因此称DNA损伤应答的“DNA域”), “DNA域”的信号对于DNA修复至关重要; 而另一部分信号则在染色质上延伸(因此称DNA损伤应答的“染色质域”), 由 γ H2AX介导, 可达兆碱基距离^[84-86]。“染色质域”招募各种损伤应答因子, 不仅帮助调控细胞DNA损伤检查点, 也协助选择修复途径^[84-86]。然而, 在CRISPR/Cas基因编辑中, DSB产生的独特性以及DSB的末端构象也决定了DNA损伤应答和DSB修复的各种变化。

通常, 在不同细胞周期产生的DSB, DSB修复途径的偏好性不同, 即使同一位点的DNA断裂, 如果出现在不同的细胞周期, 也将得到不一样的修复产物。Cas9-sgRNA结合并切割DNA靶点可以发生在细胞周期的任何时期, 然而因为切割后的靶点滞留, 断裂位点的暴露时间并不是DNA断裂时间, 而是随着滞留时间的长短变化而改变。细胞只有在识别到DNA断裂后才会启动DNA损伤应答, 选择修复途径修复DNA断裂。因此, CRISPR/Cas9靶点切割后的靶点滞留给CRISPR/Cas9诱导的DNA断裂的修复加了一层额外的调控, 这也是CRISPR/

Cas9 编辑产物异质性产生的主要原因之一^[17, 20]。研究表明, Cas9-sgRNA 形成 1~2 个 DNA 断裂便能有效延长细胞 G1 期, 而 1~4 个断裂便能使细胞有效阻滞在 G2 期^[87]。这些 DNA 损伤检查点的激活不仅源于 DNA 断裂本身, 而且受 Cas9-sgRNA 靶点滞留的强化。靶点滞留的 Cas9-sgRNA 可能阻滞或放缓邻近的转录, 导致转录 R-loop 的积累, 进而引起 DNA 复制与转录的追尾^[88]。这些事件的发生均能参与到 DNA 损伤检查点的激活, 影响 DSB 修复途径选择。

不同位点的 Cas9-sgRNA 靶点滞留能力不同, 切割 DNA 后离开靶点并暴露 DSB 所需时间也有差异, 因此 Cas9-sgRNA 在基因组不同靶点产生断裂后启动有时间差的局部 DNA 损伤应答。如果某个 Cas9-sgRNA 在切割 DNA 后迅速解离, 离开靶点并暴露 DNA 末端, 那么启动的 DNA 损伤应答可以将修复限定在 DSB 产生的细胞周期之中。若 Cas9-sgRNA 靶点切割后在靶点长久滞留, 那么细胞将无法直接检测到 DSB 的产生, 但由于邻近的 DNA 代谢或染色质活动, 将影响 DNA 损伤应答以及 DNA 损伤检查点的启动模式和调控机制。转录介导的 Cas9-sgRNA 的靶点解离可以发生在细胞周期的任何时期。在 G1 期暴露的 DSB 主要利用 NHEJ 途径修复, 而 S/G2 期暴露的 DSB 偏向 HDR 介导的修复^[33]。在 DNA 复制解离 Cas9-sgRNA 的情形下, DSB 在 S 期暴露, 并偏向于利用 HDR 途径进行修复^[17-18]。染色质活动可以在任意细胞周期解离靶点滞留的 Cas9-sgRNA, 随即启动的 DSB 修复途径选择可以保证 DSB 修复在该细胞周期内完成。

SpCas9 的靶点滞留及随后的解离同时暴露 DSB 的两个末端, 但是, *SaCas9* 和 *Cas12a* 切割 DNA 靶点后迅速释放 DSB 的 PAM 远端而滞留在 DSB 的 PAM 近端, 这种不对称解离赋予了两个 DSB 末端不同的 DNA 损伤应答信号^[13-14, 57-59]。招募的早期因子不同, 下游传递的信号也跟着改变。比如游离的 PAM 远端更容易招募 Ku70/Ku80, 偏向于激活 DNA-PKcs 介导的“DNA 域”和“染色质域”的 DNA 损伤应答 (未发表数据)。相反, Cas-sgRNA 滞留的 PAM 近端可能需要 DNA 内切酶或外切酶加工 DNA 末端, 释放 Cas-sgRNA。候选的 DNA 内切酶或外切酶包括 Mre11、CtIP、ExoI 和 DNA2 等, 而这些酶的作用将产生携带 3'-ssDNA 的 DSB 末端, 不仅偏向于激活 ATM 或 ATR 介导的 DNA 损伤应答, 也有利于 HDR 修复^[89]。这种不对

称的 DNA 损伤应答显然也增加了 CRISPR/Cas 基因编辑机制的复杂性和基因编辑产物的异质性。

3.2 CRISPR/Cas 靶点滞留对准靶位点 DSB 修复的影响及后果

Cas9 切割 DNA 后的构象变化已被解析, 然而 Cas9-sgRNA 与 DNA 解离的分子细节尚不清楚。实际上, 在哺乳动物细胞中, Cas9 所造成的损伤需要超过 15 h 的修复时间, 即使在 Cas9-sgRNA 的蛋白质核酸复合物直接递送 24 h 后仍然可以观察到编辑产物的积累^[15, 90], 这表明滞留在靶点上的 Cas9 延迟了 DNA 断裂的修复。后续的研究进一步证实, Cas9 与靶点的持久结合严重阻碍了修复机器对 DNA 断裂的识别, 并导致 DSB 修复延迟^[15, 17-18, 90]。但是在高度动态的基因组中, 滞留的 Cas9-sgRNA 有极大的几率遭遇邻近的 DNA 代谢或染色质重塑等事件, 迫使 Cas9 从靶点 DNA 上被动解离, 将暴露的 DSB 的修复限制在特定的修复环境和细胞周期^[17, 20]。

如果 Cas9-sgRNA 自发解离, 暴露出的 DSB 是典型的末端结构“干净”的平末端 DSB, 两个末端很容易被 Ku70/Ku80 或者 Mre11/Rad50/NBS1 识别和结合, 前者推动 c-NHEJ 的修复, 后者则可能进行末端切除, 产生偏向于 HDR 修复的 3'-ssDNA^[33, 37]。然而, 如果滞留在切割位点的 Cas9-sgRNA 被动解离, 暴露的 DSB 末端可能因为特殊的末端构象以及细胞周期限制而采用更为复杂的 DSB 修复途径选择, 增加基因编辑产物的异质性^[17, 20, 91]。首先, 转录进行中移动的 RNAP 一旦将 Cas9-sgRNA 从 DNA 靶点移除, 可能产生一个含有新生 mRNA-DNA 杂交链和 3'-DNA 单链的末端, 另一侧则是典型的 DSB 末端^[81]。三链末端要么直接招募特定的修复因子 (比如 ERCC6、BRCA1 和 RAD52) 进行修复, 要么通过移除新生 mRNA-DNA 杂交链中的 mRNA, 重构 DNA 末端, 启动 DSB 修复^[92]。目前并不清楚其中的调控机制和具体的修复途径偏好性, 但转录活跃区域偏向于 HDR 可能带来了一定的启示。不同于转录介导的解离, 染色质活动可能通过 DNA 扭转或拉伸等机械力移除靶点滞留的 Cas9-sgRNA, 暴露出典型的 DSB 平末端。尽管这个末端构象不会对 DSB 修复产生额外的调控, 但邻近的染色质重塑可能影响 DSB 修复步骤, 改变 DSB 修复途径选择。

DNA 复制介导的 Cas9-sgRNA 靶点解离也常有发生, 且只发生在细胞周期 S 期。一方面, 靶点切

割后滞留的Cas9-sgRNA碰巧遭遇DNA复制叉而解离。另一方面,在有的靶点,Cas9-sgRNA结合能力强、滞留时间长,转录和染色质重塑也不能迫使Cas9-sgRNA从切割的靶点解离,需要等待行进的DNA复制叉来释放靶点滞留的Cas9-sgRNA。行进复制叉上的CDC45-MCM-GINS (CMG)复合物沿着前导链移动以解开双链DNA,并将Cas9-sgRNA从靶点解离,在前导链上产生平末端的断口^[17, 93]。由于后随链上的不连续合成,在另一条姊妹染色单体上的断口并不是平末端,而是携带长3'-ssDNA的末端^[17, 93]。带有长ssDNA的末端可能不适合与c-NHEJ因子(如Ku70/Ku80)结合,但有利于RAD51重组酶的识别。而且,DNA复制介导的Cas9-sgRNA靶点解离限制在细胞周期S期,这个期间不仅HDR活性高,而且S期的姊妹染色单体的存在提供了HDR修复的序列上最佳、距离上最容易接近的同源模板。

然而,伴随着Cas9-sgRNA的靶点解离和复制叉崩塌,除了前导链产生的在一条姊妹染色单体上的平末端和后随链产生的在另一条姊妹染色单体上的带长3'单链的末端,还包括位于崩塌复制叉另一侧的平末端断口,构成了具有三末端的特殊DSB结构^[17]。目前对于如何修复这种特殊结构的认知相当缺乏。产生的三个末端两两之间也许可以通过NHEJ连接到一起,但剩余的末端因为缺乏邻近的第二末端参与修复而容易易位,甚至可能因为没有被修复,携带该末端的染色体会丢失。事实上,许多研究已经观察到Cas9诱导的染色体丢失现象。在酿酒酵母中,通过CRISPR/Cas9对染色体着丝粒附近的DNA进行特异性切割可以导致整条染色体丢失^[94]。Cas9在人的胚胎细胞内诱导的DSB可能无法修复,并在细胞进入有丝分裂后导致一条或两条染色体臂丢失^[95]。而且,三个末端的DSB也为两个姊妹染色单体末端的直接连接产生姊妹染色单体融合提供了机会,从而产生一个巨型回文染色体^[17]。这个巨型回文染色体要么携带双着丝粒,要么无着丝粒。研究表明,双着丝粒染色体在细胞的有丝分裂M期被纺锤丝拉向两极,形成染色质桥(chromatin bridge)^[96]。在细胞分裂后期,由于机械拉力使得染色质桥断裂,并导致微核的形成和染色体片段在子代细胞间不对称分配^[97]。微核的形成可能进一步诱导染色体碎裂(chromothripsis)的发生,而无着丝粒的染色体是否形成微核而诱导染色体碎裂不得而知。的确,我们的研究发现,

Cas9-sgRNA诱导的修复产物中存在姊妹染色单体融合的异常修复产物,而且这些产物的形成与DNA复制密切相关^[17]。而且,由于三个末端DSB的特殊结构及其产生的特定细胞周期,每个末端的识别和对应的DSB修复途径选择也具有其特殊性^[33, 37]。例如,携带长3'-ssDNA的末端难以结合c-NHEJ因子(如Ku70/Ku80),此时HDR可能会因为该末端的特殊性以及姊妹染色单体的存在而被优先采用,而三末端中的两个平末端却可能偏向于采用NHEJ进行修复。

当CRISPR/nCas9产生的DNA断裂是单链缺失时,DNA复制叉碰撞所产生及暴露的DSB构象与CRISPR/nCas9靶点滞留的方向密切相关^[18, 98]。CMG引导的DNA复制叉与CRISPR/nCas9产生的单链缺刻碰撞可分为四种情况(图3)。a.前导链从PAM背侧与Cas9 H840A-sgRNA诱导的单链缺刻碰撞(图3a)。碰撞后的Cas9 H840A-sgRNA可能仍与未断裂的靶标链结合在一起但释放断裂的非靶标链,前导链合成在断裂的非靶标链上进行,到达非靶标链断点时,CMG和DNA聚合酶会跑出断口而形成单末端DSB^[18, 93]。b.后随链从PAM背侧与Cas9 D10A-sgRNA诱导的单链缺刻碰撞(图3b)。碰撞后未断裂的非靶标链极有可能游离于Cas9 D10A-sgRNA复合物之外,CMG及DNA聚合酶可能依然能够在未断裂的非靶标链上引导前导链合成。如果CMG的解旋和前导链的DNA合成能通过非靶标链nCas9未切割的位点,那么在断裂的靶标链上行进的后随链DNA合成将产生一个双末端DSB^[98]。Cas9 D10A-sgRNA是否还与断裂的靶标链结合在一起目前并不清楚,然而,如果在与单链缺刻碰撞前,Cas9 D10A-sgRNA已经从靶点解离,断裂的靶标链与未切割的非靶标链已退火,CMG解旋和前导链DNA合成到达未切割的非靶标链和靶标链断口交接处,CMG复合物可能跃过缺刻口而结合下游的双链DNA,前导链的DNA合成终止,而后随链合成至靶标链的断口处产生单末端DSB^[93]。c.前导链从PAM远端与Cas9 D10A-sgRNA诱导的单链缺刻碰撞(图3c)。碰撞发生时,CMG及DNA聚合酶在遭遇靶标链上的断口前可能首先需要移除与靶标链杂交的RNA,一旦靶标链-RNA杂交链被破坏,Cas9 D10A-sgRNA在靶点的滞留难以维系而解离。而前导链合成至靶标链断口时,CMG和DNA聚合酶跑出断口,形成单末端DSB^[18]。然而,CMG也有可能跃到靶标链-RNA

杂交链, 顺利通过靶标链的切口, 继续前行, 解旋, 产生一个双末端DSB, 但目前还没有任何证据显示这种情形发生。d. 后随链从PAM远端与Cas9 H840A-sgRNA诱导的单链缺刻碰撞(图3d)。碰撞发生时, CMG及DNA聚合酶在未断裂的靶标链上首先移除与靶标链杂交的RNA, 通过破坏靶标链-RNA杂交链而解离靶点滞留的Cas9-sgRNA, 并继续在未断裂的靶标链上引导前导链合成。当CMG解旋到达未切割的靶标链与非靶标链断口的

交接处时, 跃到已退火的PAM近端3个碱基对的靶标链-非靶标链双链位置上继续前行, 解旋, 在断裂的非靶标链上的后随链DNA合成将产生一个单末端DSB^[93]。但也有研究表明, 一旦Cas9 H840A-sgRNA从靶点解离, CMG可能继续解旋, 通过靶标链的nCas9未切割位点, 引导前导链的DNA合成, 从而在断裂的非靶标链上诱导产生一个双末端DSB^[98]。

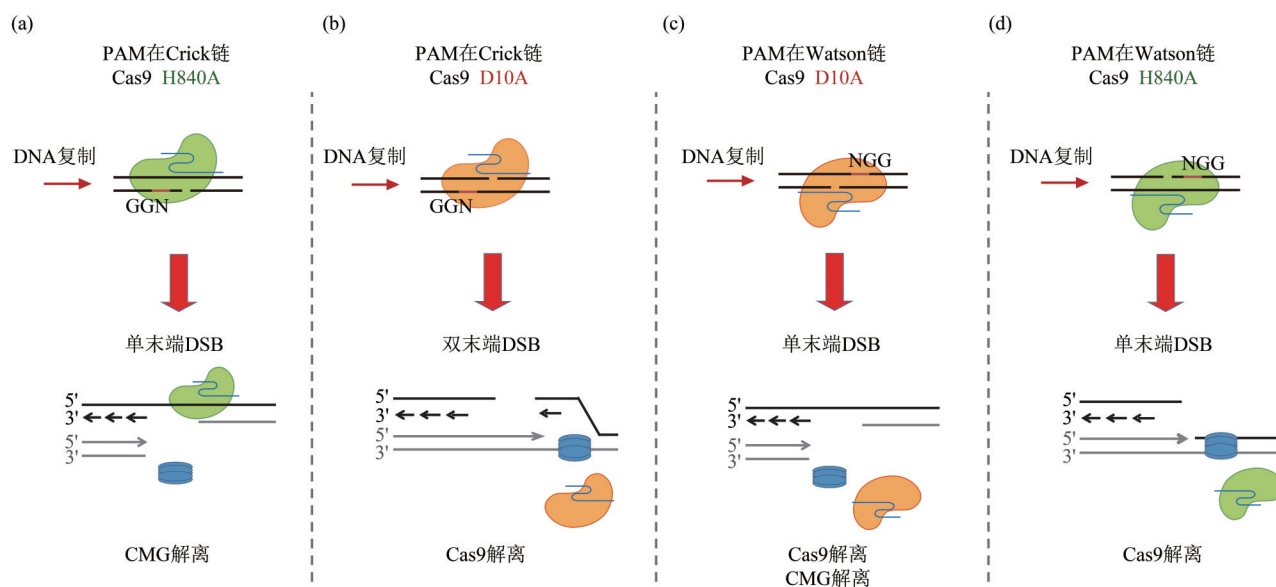


Fig. 3 Four types of collisions between DNA replication forks and nCas9-induced DNA nicks, resulting in different end configurations of replication-coupled DSBs

图3 DNA复制叉与CRISPR/nCas9诱导的单链缺刻发生碰撞的四种情形及可能产生的DSB结构特征

(a) 假定DNA复制叉从左而来, 当CMG复合物遭遇Cas9 H840A诱导的单链缺刻(PAM在Crick链上)时, 从非靶标链上移除Cas9-sgRNA, 当前导链合成至非靶标链的断点, CMG复合物离开断口产生单末端DSB, 但Cas9-sgRNA可能依旧滞留在靶标链上。(b) 遭遇Cas9 D10A诱导的单链缺刻(PAM在Crick链上)时, CMG复合物从非靶标链上移除Cas9-sgRNA, 并继续解旋、引导前导链合成通过未切割的非靶标链, 后随链合成从靶标链上解离Cas9-sgRNA, 并产生双末端DSB。(c) 遭遇Cas9 D10A诱导的单链缺刻(PAM在Watson链上)时, CMG复合物从靶标链上移除Cas9-sgRNA, 当前导链合成至靶标链的断点, CMG复合物离开断口产生单末端DSB。(d) 遭遇Cas9 H840A诱导的单链缺刻(PAM在Watson链上)时, CMG复合物从靶标链上移除Cas9-sgRNA, 并继续解旋、引导前导链合成通过未切割的靶标链, 当前导链合成至未切割靶标链和靶标链断口交接处, CMG复合物可能跃过缺刻口结合下游的双链DNA, 而后随链合成至靶标链的断口产生单末端DSB。PAM: 间隔区前体邻近基序; CMG: Cdc45-MCM-GINS复合物。

显然, 这些不同的DSB构象也需要不同的DSB修复途径选择, 目前, 领域内对此知之甚少, 但我们的研究表明, 这些DSB的修复受BRCA1的调控, 其缺陷将产生源于这些DSB的异常染色体重排, 比如高频的易位和串联倍增^[18]。因此, 从这一个角度来看, 因为单链缺刻的存在, 基于nCas9的碱基编辑器和引导编辑器并没有想象中的那样安全, DNA复制叉遭遇这两种编辑器将诱导产生不同类型的DSB构象, 给基因组稳定性带来

严重威胁。事实上, 最近研究发现, 这两种编辑器可能会在基因组上产生大片段基因删除和染色体易位, 并激活p53应答^[18, 99]。即使这两种编辑器带来的基因毒性相比于Cas9系统弱, 但依旧不可忽视。既然复制偶联DSB的产生是形成这些基因组异常修复产物的主要原因之一, 那么在进行基因编辑时, 如何避免这种DNA损伤的形成是提升编辑器安全性的重要前提, 而nCas9的滞留特性则是解决该问题的关键。

3.3 CRISPR/Cas脱靶滞留对脱靶效应的影响

脱靶效应是CRISPR基因编辑面临的一个严重问题,极大地限制了该技术的临床应用^[66]。脱靶位点的切割是由于Cas9被sgRNA引导至与间区序列不完全匹配的DNA位点,但该位点仍然可以与sgRNA退火产生R-loop结构,并随后激活Cas9介导的DNA切割^[26, 66]。然而,脱靶位点单个或多个碱基错配会降低Cas9-sgRNA在脱靶位点的亲和力和滞留时间^[66]。因此,脱靶位点的Cas9-sgRNA更容易从切割后的DNA上自发解离,而由DNA复制、转录或染色质重塑介导的解离大大降低。从脱靶位点自发解离的频率越高,在所暴露出的DSB中,越多的DSB选择c-NHEJ进行修复,而且因为Cas9诱导的平末端,c-NHEJ精准修复可以产生高频的、不能与未切割DNA底物区别的修复产物^[17, 19]。此外,Cas9在脱靶位点的再结合和再切割的能力降低,限制了脱靶效应的上调。因此,如果只从脱靶编辑产物去评估CRISPR/Cas9的脱靶效应,而不是从脱靶切割本身入手,CRISPR/Cas9的脱靶效应将被低估。特别是当脱靶编辑产物因为技术限制而难以测量时,临床应用中被低估的脱靶效应可能会带来严重后果。因此,脱靶效应的检测技术仍待进一步提高。

有效的DNA切割和高效基因编辑通常需要Cas9-sgRNA与靶DNA的紧密结合,但这也伴随着脱靶效应的提升^[26, 66]。因此,一方面,我们需要借助更强靶点结合能力或更久靶点滞留时间的Cas9变体,提高基因编辑效率,同时还要设计策略,降低这些变体的脱靶效应。基于dCas9或nCas9的基因转录调控和编辑平台也可能受益于靶点结合和靶点滞留能力的提高^[17, 19-20]。另一方面,在某些位点,对于准靶基因编辑效率而言,Cas9-sgRNA与靶DNA的分子相互作用冗余,但却导致更强的脱靶结合能力及脱靶效应。因此,消除这些冗余的分子相互作用,可以维持强大的准靶编辑活性,又能减少脱靶效应。将sgRNA的20个核苷酸的间区序列截短到17~18个,或者通过突变降低Cas9与非靶标DNA链或RNA-DNA杂交链的非特异性相互作用,成功消除Cas9-sgRNA与靶标DNA的冗余的分子相互作用,提高了Cas9-sgRNA基因编辑的特异性^[69]。

在CRISPR/Cas9基因编辑中,c-NHEJ抑制策略广泛应用于提高HDR介导的基因敲入和基因校正效率^[100-102]。然而,在许多位点,Cas9诱导的

HDR对c-NHEJ抑制策略并不敏感^[17]。此外,因为c-NHEJ的抑制将促使脱靶位点选择突变型的a-EJ进行修复,导致更强的脱靶效应^[17, 19]。因此,当使用化学或遗传学方法来抑制c-NHEJ以提高细胞中靶点基因敲入或校正效率时,脱靶效应是否加剧应当予以重点评估。当然,更好的c-NHEJ抑制方案是既能提高HDR的基因敲入效率,又不恶化脱靶效应。事实上,考虑到脱靶效应的恶化是因为目前c-NHEJ的抑制方法是对细胞全局性的影响,我们设计了一个基于dCas9的定点局部抑制c-NHEJ的方案,既提高了HDR的基因敲入效率,又没有加剧脱靶效应^[19]。

4 CRISPR/Cas靶点滞留给基因编辑带来的机遇与挑战

CRISPR/Cas在DNA上的滞留不仅影响Cas-sgRNA的解离,并且决定了特定位置处Cas诱导DSB修复途径的选择。无论是不同细胞还是同一细胞,同一靶点的断裂时间和同一时间不同靶点的断裂都会因Cas靶点滞留时间和解离形式影响修复途径的选择,这为Cas介导的基因编辑引入了一个重要变量,并增加基因编辑突变谱异质性,而且因为这个变量的不确定性,即使利用大数据分析或机器深度学习也难以预测这些异质性突变^[103-104]。因此,Cas-sgRNA的靶点滞留是CRISPR/Cas技术开发和实际应用时需要考虑的重要因素。

一方面,通过截短sgRNA或者降低Cas9与DNA之间的非特异性结合,消除Cas9-sgRNA-DNA三元复合物中冗余的分子相互作用,从而获得更高特异性的Cas9突变体,降低脱靶效应^[66]。另一方面,强劲的靶点滞留能力也是CRISPR/Cas系统高效工作的一个基础。dCas9-sgRNA通过紧密的靶点结合和靶点的稳定滞留,在转录过程中物理阻断RNAP与启动子结合以及RNAP在转录链上的移动,从而抑制基因表达^[79, 81]。而与结合到非转录链的Cas9-sgRNA相比,结合到转录链的Cas9-sgRNA更容易被行进的RNAP解离,转录抑制效果差^[79, 81]。类似地,dCas9-效应子融合在转录调控、表观修饰、靶点定位、实时影像、碱基修饰等中的工作效率也部分依赖于目标靶点结合能力和滞留时间。原则上,dCas9-效应子的靶点滞留时长决定了它们在靶点上作用的持续时间。滞留时间越长,dCas9-效应子预期工作效果越佳。甚至在nCas9碱基编辑器和引导编辑器中,nCas9-sgRNA

在靶点滞留时间越长,通过融合所携带的用于碱基修饰的脱氨酶或糖基化酶和用于引导编辑的反转录酶有更多的时间进行酶催化反应,碱基编辑器和引导编辑器的效果更好。因此,通过改造 dCas9-sgRNA 或 nCas9-sgRNA 以加强其在靶点的结合和滞留,应该是一个潜在的提高该系统工作效率的开发策略。

目前,可以采纳以下 3 种策略提高 CRISPR/Cas 靶点滞留能力。a. 基于 Cas-sgRNA-DNA 三元复合物的三维结构以及随着时间改变的动态变化,或寻找调控 Cas 与靶标链和非靶标链相互作用的 Cas 氨基酸残基,通过突变提高 Cas 蛋白与 DNA 的非特异性结合,或解析 Cas-sgRNA 靶点解离的分子细节,改造调控靶点解离的 Cas 蛋白关键氨基酸残基或 sgRNA 的关键核苷酸,延迟靶点解离^[20, 105-109]。b. 融合可以增加 Cas-sgRNA 复合物与 DNA 结合的 DNA 结合蛋白,延长滞留时间。例如,有研究表明,在 nCas9 和胞苷脱氨酶之间融合 RAD51 的单链 DNA 结合结构域,可以有效提高碱基编辑器的工作效率^[110]。最近的工作表明,通过融合双链 DNA 结合蛋白 Sso7d 或 HMG 的 DNA 结合结构域,可以提高基于 IscB 缺刻酶的碱基编辑器的效率^[111-112]。c. 通过定向进化筛选、鉴定滞留时间长的 Cas 突变体(比如 Cas9、nCas9 或 dCas9 突变体)。然而,该方向的努力仍待加强。即便如此,通过明确 Cas-sgRNA 的靶点滞留和被动解离机制及其对基因编辑的影响,不仅有助于认识和减少 CRISPR/Cas 基因编辑应用中产生的有害修复产物,而且将为建立更高效、更安全的基因编辑技术提供新的视角。

参 考 文 献

- [1] Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and Archaea. *Science*, 2010, **327**(5962): 167-170
- [2] 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
- [3] Cong L, Ran F A, Cox D, *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, **339**(6121): 819-823
- [4] Mali P, Yang L, Esvelt K M, *et al.* RNA-guided human genome engineering via Cas9. *Science*, 2013, **339**(6121): 823-826
- [5] Lander E S. The heroes of CRISPR. *Cell*, 2016, **164**(1/2): 18-28
- [6] Pacesa M, Pelea O, Jinek M. Past, present, and future of CRISPR genome editing technologies. *Cell*, 2024, **187**(5): 1076-1100
- [7] Villiger L, Joung J, Koblan L, *et al.* CRISPR technologies for genome, epigenome and transcriptome editing. *Nat Rev Mol Cell Biol*, 2024, **25**(6): 464-487
- [8] 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
- [9] Eggers A R, Chen K, Soczek K M, *et al.* Rapid DNA unwinding accelerates genome editing by engineered CRISPR-Cas9. *Cell*, 2024, **187**(13): 3249-3261.e14
- [10] 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**(7603): 420-424
- [11] 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
- [12] 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
- [13] Sternberg S H, Redding S, Jinek M, *et al.* DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, 2014, **507**(7490): 62-67
- [14] Richardson C D, Ray G J, DeWitt M A, *et al.* Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol*, 2016, **34**(3): 339-344
- [15] Kim S, Kim D, Cho S W, *et al.* Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*, 2014, **24**(6): 1012-1019
- [16] Ma H, Tu L C, Naseri A, *et al.* CRISPR-Cas9 nuclear dynamics and target recognition in living cells. *J Cell Biol*, 2016, **214**(5): 529-537
- [17] Liu S C, Feng Y L, Sun X N, *et al.* Target residence of Cas9-sgRNA influences DNA double-strand break repair pathway choices in CRISPR/Cas9 genome editing. *Genome Biol*, 2022, **23**(1): 165
- [18] Feng Y L, Liu Q, Chen R D, *et al.* DNA nicks induce mutational signatures associated with BRCA1 deficiency. *Nat Commun*, 2022, **13**(1): 4285
- [19] Feng Y L, Liu S C, Chen R D, *et al.* Proximal binding of dCas9 at a DNA double strand break stimulates homology-directed repair as a local inhibitor of classical non-homologous end joining. *Nucleic Acids Res*, 2023, **51**(6): 2740-2758
- [20] Feng Y, Liu S, Chen R, *et al.* Target binding and residence: a new determinant of DNA double-strand break repair pathway choice in CRISPR/Cas9 genome editing. *J Zhejiang Univ Sci B*, 2021, **22**(1): 73-86
- [21] Kaminski M M, Abudayyeh O O, Gootenberg J S, *et al.* CRISPR-based diagnostics. *Nat Biomed Eng*, 2021, **5**(7): 643-656
- [22] Gootenberg J S, Abudayyeh O O, Lee J W, *et al.* Nucleic acid detection with CRISPR-Cas13a/C2c2. *Science*, 2017, **356**(6336): 438-442
- [23] Chen J S, Ma E, Harrington L B, *et al.* CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity. *Science*, 2018, **360**(6387): 436-439
- [24] Tang N, Ji Q. Miniature CRISPR-Cas12 systems: mechanisms, engineering, and genome editing applications. *ACS Chem Biol*, 2024, **19**(10): 2632-2645

- 2024, **19**(7): 1399-1408
- [25] Zetsche B, Gootenberg J S, Abudayyeh O O, *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 2015, **163**(3): 759-771
- [26] Jiang F, Doudna J A. CRISPR-Cas9 structures and mechanisms. *Annu Rev Biophys*, 2017, **46**: 505-529
- [27] Globyte V, Lee S H, Bae T, *et al.* CRISPR/Cas9 searches for a protospacer adjacent motif by lateral diffusion. *EMBO J*, 2019, **38**(4): e99466
- [28] Anders C, Niewoehner O, Duerst A, *et al.* Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, 2014, **513**(7519): 569-573
- [29] Guo T, Feng Y L, Xiao J J, *et al.* Harnessing accurate non-homologous end joining for efficient precise deletion in CRISPR/Cas9-mediated genome editing. *Genome Biol*, 2018, **19**(1): 170
- [30] Shou J, Li J, Liu Y, *et al.* Precise and predictable CRISPR chromosomal rearrangements reveal principles of Cas9-mediated nucleotide insertion. *Mol Cell*, 2018, **71**(4): 498-509.e4
- [31] Lemos B R, Kaplan A C, Bae J E, *et al.* CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. *Proc Natl Acad Sci USA*, 2018, **115**(9): E2040-E2047
- [32] Shi X, Shou J, Mehryar M M, *et al.* Cas9 has no exonuclease activity resulting in staggered cleavage with overhangs and predictable di- and tri-nucleotide CRISPR insertions without template donor. *Cell Discov*, 2019, **5**: 53
- [33] Nambiar T S, Baudrier L, Billon P, *et al.* CRISPR-based genome editing through the lens of DNA repair. *Mol Cell*, 2022, **82**(2): 348-388
- [34] Chang H H Y, Pannunzio N R, Adachi N, *et al.* Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat Rev Mol Cell Biol*, 2017, **18**(8): 495-506
- [35] Ramsden D A, Carvajal-Garcia J, Gupta G P. Mechanism, cellular functions and cancer roles of polymerase-theta-mediated DNA end joining. *Nat Rev Mol Cell Biol*, 2022, **23**(2): 125-140
- [36] Jasin M, Haber J E. The democratization of gene editing: insights from site-specific cleavage and double-strand break repair. *DNA Repair*, 2016, **44**: 6-16
- [37] Xue C, Greene E C. DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. *Trends Genet*, 2021, **37**(7): 639-656
- [38] Haapaniemi E, Botla S, Persson J, *et al.* CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response. *Nat Med*, 2018, **24**(7): 927-930
- [39] Ihry R J, Worringer K A, Salick M R, *et al.* p53 inhibits CRISPR-Cas9 engineering in human pluripotent stem cells. *Nat Med*, 2018, **24**(7): 939-946
- [40] Verkuijl S A, Rots M G. The influence of eukaryotic chromatin state on CRISPR-Cas9 editing efficiencies. *Curr Opin Biotechnol*, 2019, **55**: 68-73
- [41] Hinz J M, Laughery M F, Wyrick J J. Nucleosomes inhibit Cas9 endonuclease activity *in vitro*. *Biochemistry*, 2015, **54**(48): 7063-7066
- [42] Isaac R S, Jiang F, Doudna J A, *et al.* Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *eLife*, 2016, **5**: e13450
- [43] Wang A S, Chen L C, Wu R A, *et al.* The histone chaperone FACT induces Cas9 multi-turnover behavior and modifies genome manipulation in human cells. *Mol Cell*, 2020, **79**(2): 221-233.e5
- [44] Dominguez A A, Lim W A, Qi L S. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat Rev Mol Cell Biol*, 2016, **17**(1): 5-15
- [45] Chen B, Gilbert L A, Cimini B A, *et al.* Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. *Cell*, 2013, **155**(7): 1479-1491
- [46] Klompe S E, Vo P L H, Halpin-Healy T S, *et al.* Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. *Nature*, 2019, **571**(7764): 219-225
- [47] Strecker J, Ladha A, Gardner Z, *et al.* RNA-guided DNA insertion with CRISPR-associated transposases. *Science*, 2019, **365**(6448): 48-53
- [48] Chen S P, Wang H H. An engineered cas-transposon system for programmable and site-directed DNA transpositions. *CRISPR J*, 2019, **2**(6): 376-394
- [49] 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
- [50] 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
- [51] 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
- [52] Richardson C D, Kazane K R, Feng S J, *et al.* CRISPR-Cas9 genome editing in human cells occurs *via* the Fanconi anemia pathway. *Nat Genet*, 2018, **50**(8): 1132-1139
- [53] Petermann E, Lan L, Zou L. Sources, resolution and physiological relevance of R-loops and RNA-DNA hybrids. *Nat Rev Mol Cell Biol*, 2022, **23**(8): 521-540
- [54] Ceccaldi R, Sarangi P, D'Andrea A D. The Fanconi anaemia pathway: new players and new functions. *Nat Rev Mol Cell Biol*, 2016, **17**(6): 337-349
- [55] Wang Y, Feng Y L, Liu Q, *et al.* TREX2 enables efficient genome disruption mediated by paired CRISPR-Cas9 nickases that generate 3'-overhanging ends. *Mol Ther Nucleic Acids*, 2023, **34**: 102072
- [56] Yin J, Lu R, Xin C, *et al.* Cas9 exo-endonuclease eliminates chromosomal translocations during genome editing. *Nat Commun*, 2022, **13**(1): 1204
- [57] Singh D, Mallon J, Poddar A, *et al.* Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proc Natl Acad Sci USA*, 2018, **115**(21): 5444-5449

- [58] Strohkendl I, Saifuddin F A, Rybarski J R, *et al.* Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Mol Cell*, 2018, **71**(5): 816-824.e3
- [59] Pan J, Mabuchi M, Robb G B. DNA rehybridization drives product release from Cas9 ribonucleoprotein to enable multiple-turnover cleavage. *Nucleic Acids Res*, 2023, **51**(8): 3903-3917
- [60] Szczelkun M D, Tikhomirova M S, Sinkunas T, *et al.* Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proc Natl Acad Sci USA*, 2014, **111**(27): 9798-9803
- [61] Kleinstiver B P, Prew M S, Tsai S Q, *et al.* Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, 2015, **523**(7561): 481-485
- [62] Hu J H, Miller S M, Geurts M H, *et al.* Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*, 2018, **556**(7699): 57-63
- [63] Kleinstiver B P, Prew M S, Tsai S Q, *et al.* Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol*, 2015, **33**(12): 1293-1298
- [64] Sternberg S H, LaFrance B, Kaplan M, *et al.* Conformational control of DNA target cleavage by CRISPR-Cas9. *Nature*, 2015, **527**(7576): 110-113
- [65] Zhu X, Clarke R, Puppala A K, *et al.* Cryo-EM structures reveal coordinated domain motions that govern DNA cleavage by Cas9. *Nat Struct Mol Biol*, 2019, **26**(8): 679-685
- [66] Kim D, Luk K, Wolfe S A, *et al.* Evaluating and enhancing target specificity of gene-editing nucleases and deaminases. *Annu Rev Biochem*, 2019, **88**: 191-220
- [67] Doench J G, Fusi N, Sullender M, *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol*, 2016, **34**(2): 184-191
- [68] Boyle E A, Andreasson J O L, Chircus L M, *et al.* High-throughput biochemical profiling reveals sequence determinants of dCas9 off-target binding and unbinding. *Proc Natl Acad Sci USA*, 2017, **114**(21): 5461-5466
- [69] Fu Y, Sander J D, Reyon D, *et al.* Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*, 2014, **32**(3): 279-284
- [70] Nishimasu H, Ran F A, Hsu P D, *et al.* Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, 2014, **156**(5): 935-949
- [71] Slaymaker I M, Gao L, Zetsche B, *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science*, 2016, **351**(6268): 84-88
- [72] Casini A, Olivieri M, Petris G, *et al.* A highly specific SpCas9 variant is identified by *in vivo* screening in yeast. *Nat Biotechnol*, 2018, **36**(3): 265-271
- [73] Chen J S, Dagdas Y S, Kleinstiver B P, *et al.* Enhanced proofreading governs CRISPR-Cas9 targeting accuracy. *Nature*, 2017, **550**(7676): 407-410
- [74] Kleinstiver B P, Pattanayak V, Prew M S, *et al.* High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, 2016, **529**(7587): 490-495
- [75] Bustamante C, Bryant Z, Smith S B. Ten years of tension: single-molecule DNA mechanics. *Nature*, 2003, **421**(6921): 423-427
- [76] Newton M D, Taylor B J, Driessen R P C, *et al.* DNA stretching induces Cas9 off-target activity. *Nat Struct Mol Biol*, 2019, **26**(3): 185-192
- [77] Ivanov I E, Wright A V, Cofsky J C, *et al.* Cas9 interrogates DNA in discrete steps modulated by mismatches and supercoiling. *Proc Natl Acad Sci U S A*, 2020, **117**(11): 5853-5860
- [78] Zhang Q, Wen F, Zhang S, *et al.* The post-PAM interaction of RNA-guided spCas9 with DNA dictates its target binding and dissociation. *Sci Adv*, 2019, **5**(11): eaaw9807
- [79] Qi L S, Larson M H, Gilbert L A, *et al.* Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*, 2013, **152**(5): 1173-1183
- [80] Jones D L, Leroy P, Unoson C, *et al.* Kinetics of dCas9 target search in *Escherichia coli*. *Science*, 2017, **357**(6358): 1420-1424
- [81] Clarke R, Heler R, MacDougall M S, *et al.* Enhanced bacterial immunity and mammalian genome editing *via* RNA-polymerase-mediated dislodging of Cas9 from double-strand DNA breaks. *Mol Cell*, 2018, **71**(1): 42-55.e8
- [82] Patel S S, Pandey M, Nandakumar D. Dynamic coupling between the motors of DNA replication: hexameric helicase, DNA polymerase, and primase. *Curr Opin Chem Biol*, 2011, **15**(5): 595-605
- [83] Blackford A N, Jackson S P. ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. *Mol Cell*, 2017, **66**(6): 801-817
- [84] Xie A, Kwok A, Scully R. Role of mammalian Mre11 in classical and alternative nonhomologous end joining. *Nat Struct Mol Biol*, 2009, **16**(8): 814-818
- [85] Scully R, Xie A. Double strand break repair functions of histone H2AX. *Mutat Res*, 2013, **750**(1/2): 5-14
- [86] Feng Y L, Xiang J F, Kong N, *et al.* Buried territories: heterochromatic response to DNA double-strand breaks. *Acta Biochim Biophys Sin*, 2016, **48**(7): 594-602
- [87] van den Berg J, G Manjón A, Kielbassa K, *et al.* A limited number of double-strand DNA breaks is sufficient to delay cell cycle progression. *Nucleic Acids Res*, 2018, **46**(19): 10132-10144
- [88] Hamperl S, Bocek M J, Saldivar J C, *et al.* Transcription-replication conflict orientation modulates R-loop levels and activates distinct DNA damage responses. *Cell*, 2017, **170**(4): 774-786.e19
- [89] Symington L S, Gautier J. Double-strand break end resection and repair pathway choice. *Annu Rev Genet*, 2011, **45**: 247-271
- [90] Brinkman E K, Chen T, de Haas M, *et al.* Kinetics and fidelity of the repair of Cas9-induced double-strand DNA breaks. *Mol Cell*, 2018, **70**(5): 801-813.e6
- [91] Longo G M C, Sayols S, Kotini A G, *et al.* Linking CRISPR-Cas9 double-strand break profiles to gene editing precision with BreakTag. *Nat Biotechnol*, 2024. DOI: 10.1038/s41587-024-02238-8

- [92] Puget N, Miller K M, Legube G. Non-canonical DNA/RNA structures during transcription-coupled double-strand break repair: roadblocks or bona fide repair intermediates?. *DNA Repair*, 2019, **81**: 102661
- [93] Vrtis K B, Dewar J M, Chistol G, *et al.* Single-strand DNA breaks cause replisome disassembly. *Mol Cell*, 2021, **81**(6): 1309-1318.e6
- [94] Xu H, Han M, Zhou S, *et al.* Chromosome drives *via* CRISPR-Cas9 in yeast. *Nat Commun*, 2020, **11**(1): 4344
- [95] Zuccaro M V, Xu J, Mitchell C, *et al.* Allele-specific chromosome removal after Cas9 cleavage in human embryos. *Cell*, 2020, **183**(6): 1650-1664.e15
- [96] Kagaya K, Noma-Takayasu N, Yamamoto I, *et al.* Chromosome instability induced by a single defined sister chromatid fusion. *Life Sci Alliance*, 2020, **3**(12): e202000911
- [97] Umbreit N T, Zhang C Z, Lynch L D, *et al.* Mechanisms generating cancer genome complexity from a single cell division error. *Science*, 2020, **368**(6488): eaba0712
- [98] Pavani R, Tripathi V, Vrtis K B, *et al.* Structure and repair of replication-coupled DNA breaks. *Science*, 2024, **385**(6710): eado3867
- [99] Fiumara M, Ferrari S, Omer-Javed A, *et al.* Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nat Biotechnol*, 2024, **42**(6): 877-891
- [100] Chu V T, Weber T, Wefers B, *et al.* Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells. *Nat Biotechnol*, 2015, **33**(5): 543-548
- [101] Maruyama T, Dougan S K, Truttmann M C, *et al.* Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*, 2015, **33**(5): 538-542
- [102] Yeh C D, Richardson C D, Corn J E. Advances in genome editing through control of DNA repair pathways. *Nat Cell Biol*, 2019, **21**(12): 1468-1478
- [103] Abadi S, Yan W X, Amar D, *et al.* A machine learning approach for predicting CRISPR-Cas9 cleavage efficiencies and patterns underlying its mechanism of action. *PLoS Comput Biol*, 2017, **13**(10): e1005807
- [104] Chuai G, Ma H, Yan J, *et al.* DeepCRISPR: optimized CRISPR guide RNA design by deep learning. *Genome Biol*, 2018, **19**(1): 80
- [105] Yan H, Tan X, Zou S, *et al.* Assessing and engineering the IscB- ω RNA system for programmed genome editing. *Nat Chem Biol*, 2024. DOI: 10.1038/s41589-024-01669-3
- [106] 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. *Nat Biotechnol*, 2019, **37**(3): 276-282
- [107] Kim D Y, Lee J M, Moon S B, *et al.* Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus. *Nat Biotechnol*, 2022, **40**(1): 94-102
- [108] Wu T, Liu C, Zou S, *et al.* An engineered hypercompact CRISPR-Cas12f system with boosted gene-editing activity. *Nat Chem Biol*, 2023, **19**(11): 1384-1393
- [109] Saito M, Xu P, Faure G, *et al.* Fanzor is a eukaryotic programmable RNA-guided endonuclease. *Nature*, 2023, **620**(7974): 660-668
- [110] Zhang X, Chen L, Zhu B, *et al.* Increasing the efficiency and targeting range of cytidine base editors through fusion of a single-stranded DNA-binding protein domain. *Nat Cell Biol*, 2020, **22**(6): 740-750
- [111] Han L, Hu Y, Mo Q, *et al.* Engineering miniature IscB nickase for robust base editing with broad targeting range. *Nat Chem Biol*, 2024. DOI: 10.1038/s41589-024-01670-w
- [112] Xue N, Hong D, Zhang D, *et al.* Engineering IscB to develop highly efficient miniature editing tools in mammalian cells and embryos. *Mol Cell*, 2024, **84**(16): 3128-3140.e4

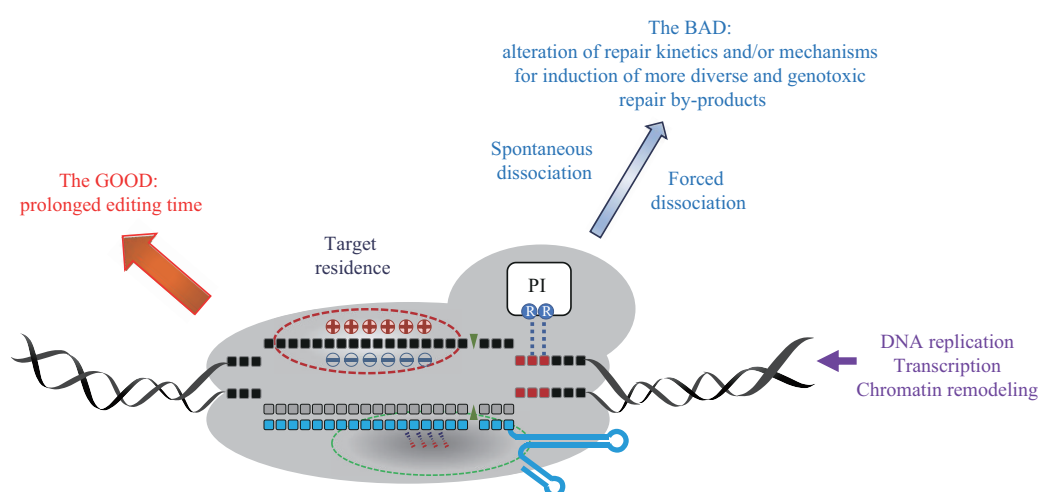
Target Residence of CRISPR/Cas in Genome Editing*

FENG Yi-Li^{1,2)**}, CHEN Ruo-Dan^{1,2)}, XIE An-Yong^{1,2)**}

⁽¹⁾Department of General Surgery, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310019, China;

⁽²⁾Institute of Translational Medicine, Zhejiang University School of Medicine, Hangzhou 310029, China)

Graphical abstract



Abstract The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) is widely used for targeted genomic and epigenomic modifications, transcriptional regulation and real-time cell imaging, and has already demonstrated great potential for applications in agriculture, industry and medicine. The promise of the technology depends upon the five intrinsic properties of CRISPR/Cas: targeting, target unwinding, target cutting, target residence, and collateral cleavage. Here, mainly using *Streptococcus pyogenes* CRISPR/Cas9 as example, we will focus on the target residence of CRISPR/Cas in applications of the CRISPR/Cas technology, summarize the recent progress, and discuss the effect of CRISPR/Cas target binding and residence on DNA double strand break repair pathway choices and the opportunities that CRISPR/Cas target residence presents to optimize the CRISPR/Cas technology.

Key words CRISPR/Cas9, target residence, target dissociation, DNA double strand break repair pathway choice, genome editing heterogeneity

DOI: 10.16476/j.pibb.2024.0274

* This work was supported by grants from The National Natural Science Foundation of China (32371348, 32071439).

** Corresponding author.

FENG Yi-Li. Tel: 86-13968126080, E-mail: eric_feng@zju.edu.cn

XIE An-Yong. Tel: 86-13185001191, E-mail: anyongxie@zju.edu.cn

Received: June 28, 2024 Accepted: August 22, 2024