



CRISPR-Cas 起源: TnpB 与 IscB 核酸酶的研究与应用进展

史光亮^{1,2}, 李维^{1*}, 向华², 龚路遥^{2*}

1 四川师范大学 生命科学学院, 四川 成都 610101

2 中国科学院微生物研究所, 微生物资源前期开发国家重点实验室, 北京 100101

史光亮, 李维, 向华, 龚路遥. CRISPR-Cas 起源: TnpB 与 IscB 核酸酶的研究与应用进展[J]. 微生物学报, 2024, 64(9): 3091-3104.

SHI Guangliang, LI Wei, XIANG Hua, GONG Luyao. Origin of CRISPR-Cas: progress in research and applications of TnpB and IscB[J]. Acta Microbiologica Sinica, 2024, 64(9): 3091-3104.

摘要: CRISPR-Cas 系统是广泛存在于细菌和古菌中的防御系统。基于该系统开发的基因组编辑工具已在大量物种中实现靶向编辑。目前应用最多的是 CRISPR-Cas9 和 CRISPR-Cas12a 基因组编辑工具, 但它们的蛋白大小均超过 1 000 个氨基酸, 不利于递送。来自转座子家族的 TnpB 和 IscB 蛋白(大小约 400 个氨基酸)分别被认为是 Cas12 和 Cas9 的祖先蛋白, 但其功能直到最近才被解析。它们被统称为专性移动元件引导活性(obligate mobile element-guided activity, OMEGA)蛋白, 其引导 RNA 被称为 ω RNA。此后, OMEGA 系统成为了基因编辑领域的研究热点之一。OMEGA 系统在三域生物中都有广泛分布, 而且种类多样。对 OMEGA 系统的深入研究, 将有助于开发精简、高效、安全的新型基因组编辑工具。本文围绕 OMEGA 系统的发现历程、结构特点、作用机制和在基因组编辑中的应用展开介绍, 为新型基因组编辑工具的开发和优化提供参考。

关键词: CRISPR; IscB; TnpB; 专性移动元件引导活性(OMEGA); Fanzor; 基因组编辑

资助项目: 国家自然科学基金(32230061)

This work was supported by the National Natural Science Foundation of China (32230061).

*Corresponding authors. E-mail: GONG Luyao, gongly@im.ac.cn; LI Wei, liwei001@sicnu.edu.cn

Received: 2024-02-07; Accepted: 2024-05-31; Published online: 2024-06-03

Origin of CRISPR-Cas: progress in research and applications of TnpB and IscB

SHI Guangliang^{1,2}, LI Wei^{1*}, XIANG Hua², GONG Luyao^{2*}

1 College of Life Science, Sichuan Normal University, Chengdu 610101, Sichuan, China

2 State Key Laboratory of Microbial Resources, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

Abstract: CRISPR-Cas is a defense system ubiquitous in bacteria and archaea. It has been successfully applied in genome editing in a variety of organisms. At present, CRISPR-Cas9 and CRISPR-Cas12a are the most widely used genome editing tools. However, the large protein sizes of Cas9 and Cas12a (more than 1 000 amino acids (aa)) hinder their delivery. TnpB and IscB (about 400 aa) encoded by the transposon family are considered ancestors of Cas12 and Cas9, respectively, whereas their functions are revealed just recently. They are named as obligate mobile element-guided activity (OMEGA), with the associated RNA named ω RNA. Since then, the OMEGA system has become one of the research hotspots in genome editing. OMEGA systems are diverse, with wide distribution in all the three domains of life. The in-depth research on the OMEGA system will aid in the development of new genome editing tools that are streamlined, efficient, and safe. Here, we reviewed the discovery history, structural characteristics, mechanisms of cleavage, and genome editing applications of OMEGA systems, aiming to lay a foundation for the development and optimization of genome editing tools.

Keywords: CRISPR; IscB; TnpB; obligate mobile element-guided activity (OMEGA); Fanzor; genome editing

CRISPR-Cas 系统由成簇规律间隔短回文重复序列 (clustered regularly interspaced short palindromic repeats, CRISPR) 及其关联蛋白 (CRISPR-associated proteins, Cas) 组成。其中 CRISPR 结构由高度保守的重复序列 (repeat) 和来自病毒等可移动元件的间隔序列 (spacer) 组成。CRISPR-Cas 系统的发现始于 1987 年^[1]。2007 年, 该系统被证明具有抗病毒防御功能, 获得了更多关注^[2]。紧接着, 在随后不到 5 年时间里, CRISPR-Cas 系统的基本防御机制被解析^[3-5]。2012 年, 基于 CRISPR-Cas9 的基因组编辑技术诞生^[6]。2013 年, CRISPR-Cas9 技术被成功应用于人体细胞的基因组编辑^[7]。该技术由引导 RNA (guide RNA, gRNA) 引导 Cas 核酸酶对基因组上

的靶位点进行切割, 激活胞内的修复途径从而实现基因组编辑, 具有易编程 (gRNA 序列易改变, 利用 gRNA 与 DNA 的碱基匹配实现特异性靶向)、组分简单、高效、特异等优点。随后多样化的 CRISPR-Cas 基因组编辑技术被应用于多种生物体的基因组编辑^[8], 如古菌^[9-10]、病原细菌^[11]、农作物^[12]、灵长目动物^[13], 加速了生命科学领域的发展, 并在农业育种、基因治疗等方面展现了巨大的应用前景。2020 年, 诺贝尔化学奖授予了 CRISPR-Cas9 技术的两位发明者, 表彰她们在基因组编辑领域的重大发现^[14]。当前, CRISPR-Cas 编辑技术及其衍生的单碱基编辑技术和先导编辑技术已成为基因组编辑领域的主流技术^[8]。

目前广泛应用的 CRISPR-Cas9 和 CRISPR-Cas12a 编辑技术所使用的核酸酶大小超过 1 000 个氨基酸, 并且在使用载体递送时还需整合启动子、终止信号等调控序列, 不利于递送。此外, 基于 CRISPR-Cas 系统开发单碱基编辑或者先导编辑工具, 需要再融合脱氨酶、逆转录酶或糖基化酶等不同的功能元件, 使编辑工具进一步增大^[15-18]。常用的腺相关病毒(adenovirus-associated virus, AAV)递送工具载量约 4.7 kb, 无法采用单 AAV 递送大型的编辑工具, 而采用双 AAV 递送策略时, 通常编辑效率低于单 AAV 递送^[19-20]。对于基因编辑工具, 除了小型化的需求, 还存在降低脱靶效率、提高精确性、减小原间隔序列邻近基序(protospacer adjacent motif, PAM)限制等需求^[8,21]。例如疾病治疗要求最小化的脱靶效率和高度的精确性, 而 CRISPR-Cas9 等系统会在非靶标位点发生切割并引起染色体异位等问题^[22]。此外, PAM 基序是 CRISPR-Cas 系统靶向识别的重要元件, CRISPR-Cas 系统会先识别 PAM (通常 2-5 nt, 紧邻 spacer 匹配区域)进行靶标和非靶标的区分, 之后若 gRNA 与靶标形成完全的碱基匹配, 则 Cas 蛋白可对靶标进行切割^[23]。然而, 通过工程化改造拓展 PAM 识别谱的编辑工具通常编辑效率也会降低^[24]。

近几年, 来自转座子家族、高度丰富的 OMEGA 系统(TnpB、IscB 等)^[25]获得了广泛的关注。它们被认为是 Cas9 和 Cas12 的祖先蛋白, 具有紧凑的蛋白大小(约 400 个氨基酸)和多样化的类型, 并被证明具有 RNA 引导的核酸酶活性^[25-26]。值得注意的是, 在真核生物中广泛存在的 Fanzor 蛋白(TnpB 同源蛋白)也被证明具有 RNA 引导的核酸酶活性, 表明 RNA 引导的核酸酶系统广泛存在于三域生物中^[27]。OMEGA 系统的作用机制和空间结构与已报道的 CRISPR-Cas 系统存在许多相同点, 同时存在许多差异。

OMEGA 系统不但蛋白体积小, 还具有开发成编辑工具的其他优势, 如部分系统具有识别单碱基 PAM、无旁系切割活性或脱靶效率低等特点^[25,27-28]。本文对 OMEGA 系统的发现历程、作用机制和应用进展进行介绍, 希望对基因组编辑工具的开发和优化提供参考。

1 OMEGA 的发现

CRISPR-Cas 系统分为 2 个类群, 并进一步分为 6 种类型和超过 30 种亚型。类群 1 中结合 gRNA 并参与靶标识别和干扰的效应物(effector)是由多种蛋白形成的复合体, 类群 2 的效应物是由多个结构域组成的单一蛋白; 类群 2 包括 II 型、V 型和 VI 型系统, 其中 II 型和 V 型系统通常由 gRNA 引导效应蛋白靶向切割 DNA, VI 型系统由 gRNA 引导效应蛋白靶向切割 RNA; CRISPR 结构转录后经过 Cas 等蛋白的加工, 可以得到成熟的短 CRISPR RNA (crRNA)^[29]。此外, II 型系统和部分 V 型系统会编码一个与 repeat 部分互补配对的反式激活 RNA (trans-activating CRISPR-associated RNA, tracrRNA), 参与 crRNA 的成熟过程, 并与 crRNA 形成杂交链共同指导 Cas 蛋白识别靶标^[23]。Jinek 等将 CRISPR-Cas9 系统中的 tracrRNA 与 crRNA 人工嵌合成一条单链引导 RNA (single guide RNA, sgRNA), 因此, 只需要 sgRNA 和 Cas9 蛋白 2 个组分即可进行靶向切割^[6]。

目前开发的基因组编辑工具主要基于 II 型和 V 型系统, 它们的效应蛋白分别为 Cas9 和 Cas12。来自转座子家族的 TnpB 和 IscB 蛋白可能分别是 Cas12 和 Cas9 的祖先蛋白^[29-32]。值得注意的是, 过去认为 CRISPR 样系统只存在于原核生物中, 但是最近在真核生物中也挖掘出 TnpB 和 IscB 的同源蛋白, 并证明它们具有 RNA 引导的核酸酶活性^[25,27]。对 OMEGA 的发现过程

进行梳理,将有助于挖掘更多的同源蛋白用于开发新型基因组编辑工具。

1.1 原核生物中的 OMEGA

1.1.1 Cas12 的祖先 TnpB

转座子在生物体的进化过程中起着重要的作用,参与了许多生命过程。其中 IS200/IS605 和 IS607 家族均属于最简单、最古老的转座子家族之一,在原核生物中广泛存在^[33]。这些转座子由 TnpA 转座酶催化转座,转座子末端通常存在左端(left end, LE)和右端(right end, RE)回文结构,可以被转座酶识别并切割^[34-36]。此外,这些转座子通常还编码 TnpB 蛋白^[33]。早期的研究发现 TnpB 在 TnpA 介导的转座过程中不是必需的,认为其可能参与转座调控^[37]。

TnpB 与 Cas12 蛋白具有同源性,同源区域包括 RuvC 样核酸酶结构域和富含精氨酸的桥式螺旋(bridge helix, BH)结构;由于不同 Cas12 亚型的效应蛋白与不同的 TnpB 家族蛋白接近,不同的 Cas12 亚型效应蛋白可能是由不同的 TnpB 家族蛋白独立进化而来^[31-32]。另一方面,TnpB 的结构解析结果显示 TnpB 具有 Cas12 家族蛋白的最小结构和功能核心,进一步证明了 TnpB 是 Cas12 的祖先^[38-39]。目前已鉴定出十余种 Cas12 亚型,使 V 型系统成为了亚型种类最多的 CRISPR-Cas 类型。Cas12 家族蛋白中 Cas12f 的大小与 TnpB 最接近,约 400–700 个氨基酸,前期的研究通过系统发育和功能分析提出 Cas12f 与 TnpB 关系最接近^[25-26,29,32]。最近,Chen 等对 Cas12n (过去称为 V-U4,由于缺少功能研究未被定义为 Cas12 亚型^[32])进行研究(大小约 400–700 个氨基酸),证明其具有 RNA 引导的核酸酶活性;他们结合系统发育和功能分析证明,Cas12n 是目前发现最接近 TnpB 的 Cas12 亚型^[40]。鉴于 TnpB 的多样性和分布广泛性,可能还有许多新的 Cas12 亚型等待挖掘。

1.1.2 Cas9 的祖先 IscB

Cas9 也存在大量同源蛋白,同源区域包括 BH 结构、HNH 核酸酶结构域和被它们插入的 RuvC 样核酸酶结构域;这些同源蛋白也属于 IS200/IS605 家族,其所属转座子被命名为 Cas9 样插入序列(insertion sequences Cas9-like, ISC);这些同源蛋白被命名为 IscB,被认为是 Cas9 的祖先^[30]。由于 TnpB 在原核生物中的分布远超过 IscB,并且具有最简单的结构域,IsCB 可能也起源于 TnpB^[30-32]。Altae-Tran 等进一步挖掘 Cas9 同源蛋白,发现了许多蛋白家族^[25]。其中一类蛋白包含被 BH 结构插入的 RuvC 样结构域,但缺少 HNH 结构域,蛋白大小约 350 个氨基酸;其 N 端和 IscB 相似,具有 PLMP 结构域(根据其 PLMP 保守基序命名),但其仅具有单链切口酶活性^[25]。这类蛋白同样属于 IS200/IS605 家族,之前被命名为 IscB1^[30]。为突出其结构域组成,Altae-Tran 等将其重新命名为 RuvC 样插入序列 B (insertion sequence RuvC-like OrfB, IsrB);他们还发现了一个更小的蛋白家族,大小仅约 180 个氨基酸,只包含 PLMP 结构域和 HNH 结构域,缺少 RuvC 样结构域,被命名为 HNH 样插入序列 B (insertion sequence HNH-like OrfB, IshB);另一方面,他们发现了 Cas9 的 2 个新分支:一个分支被定义为新的 II-D 亚型,具有较小的 Cas9 尺寸(约 700 个氨基酸),另一个分支属于 II-C 亚型,具有较大的 Cas9 尺寸(>1 700 个氨基酸)^[25]。

系统发育分析结果表明,所有的 Cas9 蛋白均起源于同一个 IscB 进化支,crRNA 和 tracrRNA 可能均由 ω RNA 进化产生;结合这些蛋白上下游序列中 TnpA、CRISPR 序列、Cas 适应元件(参与 spacer 获取的蛋白,包括 Cas1、Cas2、Cas4)等出现的情况,Altae-Tran 等提出 IsrB 可能是由紧凑的 RuvC 核酸酶进化而来,之后 IsrB 获得了 HNH 结构域(可能是插入了 IshB

样的蛋白), 形成 IscB 家族^[25]。结构解析的结果进一步支持以上结论^[41-42]。对于 TnpB 与 IscB 的进化关系及 TnpB 如何从 RuvC 核酸酶进化而来, 还有待深入研究。

1.2 真核生物中的 OMEGA

TnpB 和 IscB 的同源蛋白也出现在真核细胞中^[25,43], 并且同样具有 RNA 引导的核酸酶活性^[25,27,44-45]。2013 年, 真核细胞中 TnpB 的同源蛋白首次被报道, 被称为 Fanzor, 广泛存在于不同的真核物种及其病毒中^[43]。Fanzor 蛋白分为两类, 其中 Fanzor1 通常存在于不同的真核转座子中, Fanzor2 只来自 IS607 样的转座子^[43-45]。Yoon 等对 TnpB 和 Fanzor 进行系统发育分析发现, Fanzor 均起源于 IS607 家族的 TnpB; 他们提出 Fanzor 的祖先蛋白可能通过水平转移从原核生物进入真核生物, 产生 Fanzor2, 随后 Fanzor2 在真核生物中通过病毒传递等方式扩散, 并被真核转座子整合; 在这个过程中, 蛋白体积逐渐变大, 最后产生 Fanzor1^[45]。此外, 编码 Fanzor 的基因具有内含子和核定位信号, 也表明其在真核细胞中经过了长期的适应过程^[44]。另一方面, Altae-Tran 等在一种陆生绿藻的叶绿体基因组中发现了多个 IscB, 并证明其中 1 个完整的 IscB 具有核酸酶活性^[25]。目前, 尚未在高等动物中发现 Fanzor 系统。真核生物中 IscB 的分布情况也有待进一步研究。

1.3 OMEGA 的生理功能

在 OMEGA 的核酸酶活性被证明后, 这些广泛存在的 OMEGA 的生理功能被进一步解析。Karvelis 等在研究来自耐辐射奇球菌 (*Deinococcus radiodurans*) ISDra2 的 TnpB (ISDra2 TnpB) 时, 发现其识别的 PAM 序列 (5'-TTGAT-3') 正好与该转座子中 TnpA 转座酶切割和插入的识别序列一致^[26]。该序列也由此被称为转座子相关基序(transposon-associated motif,

TAM)^[25-26]。TnpA 转座酶和 TnpB 核酸酶识别序列的一致性, 暗示了转座和靶向切割之间的紧密联系; Karvelis 等基于该现象提出 TnpB 的功能可能是在 TnpA 切除转座子序列后, 对转座子两侧序列的连接处进行特异切割产生双链断裂, 引起胞内的重组修复, 从而使子代细胞中都保留原位的转座子^[26]。Xiang 等对大量 TnpB 系统进行测试, 进一步证明 TnpB 识别的 TAM 与 TnpA 识别序列相同, 即与 LE 上游 4-5 nt 序列一致^[46]。Meers 等对来自嗜热嗜脂肪地芽孢杆菌 (*Geobacillus stearothermophilus*) 的多个 TnpB 和 IscB 进行研究, 同样发现这些蛋白可以特异切割转座子两侧序列的连接产物; 此外, 与单独表达 TnpA 相比, 共表达 TnpA 和 TnpB 显著提高了转座子原位保留的数量; 以上结果证明 TnpB 和 IscB 的功能是避免转座子丢失, 促进转座子保留和扩散^[47]。真核生物中 OMEGA 的生理功能是否与原核生物中相同还有待实验验证。

2 OMEGA 系统的特征

TnpB 与 IscB 系统都是由核酸酶和 ω RNA 组成的(图 1), ω RNA 包括形成复杂三维结构的支架区和与靶标 DNA 互补配对的引导区; 虽然 OMEGA 系统中核酸酶尺寸很小, 但与 crRNA 相比, ω RNA 通常更长并且是高度结构化的^[27,38-39,41-42](图 2)。这些系统具有独特的结构和靶向识别与切割机制, 为基因组编辑工具的开发提供多样化的选择。

2.1 TnpB 的特征

TnpB 核酸酶来自 IS200/IS605 或 IS607 转座子家族, 是原核生物中分布最广泛的蛋白之一, 远超过 IscB、Cas9 或 Cas12 的分布^[25,30]。其 ω RNA 与 TnpB 编码区的 3'端有重叠, 并包含了该转座子的 RE 序列, 因此也被称为右末端元件 RNA (right end element RNA, reRNA)^[26](图 1)。

2023年4月,2个实验室分别通过冷冻电镜解析了ISDra2 TnpB- ω RNA-靶标DNA三元复合物的结构^[38-39]。ISDra2 TnpB蛋白大小为408个氨基酸,也是首个被成功用于真核细胞基因组编辑的TnpB蛋白^[26]。冷冻电镜结果显示,ISDra2 TnpB具有Cas12家族蛋白的最小结构和功能核心,也能形成二裂片结构,包括N末端的核酸识别裂片(recognition, REC)和C末端的核酸酶裂片(nuclease, NUC);其 ω RNA长度约150 nt,由116 nt的支架区和16 nt的引导区组成,包括4个茎和1个假结(茎环结构中的环上序列与其他部位的序列碱基配对,形成类似“打结”形状),假结

也在所有Cas12系统中保守存在^[38-39]。然而TnpB与Cas12系统也存在显著差别。结构比较结果显示,部分 ω RNA区域发挥了Cas12蛋白某些结构域的功能^[38-39]。此外,单个ISDra2 TnpB蛋白会被单一的 ω RNA引导,而某些Cas12蛋白会形成不对称的二聚体(如Cas12f),并且需要双RNA引导(tracrRNA和crRNA)^[48,50](表1)。结构分析结果表明,进化过程中Cas12可能通过形成不对称的二聚体或者插入多样化的REC2结构域来识别PAM远端的引导序列,从而获得参与CRISPR-Cas适应过程的结构基础;而双RNA引导序列可能是由 ω RNA进化而来^[38]。

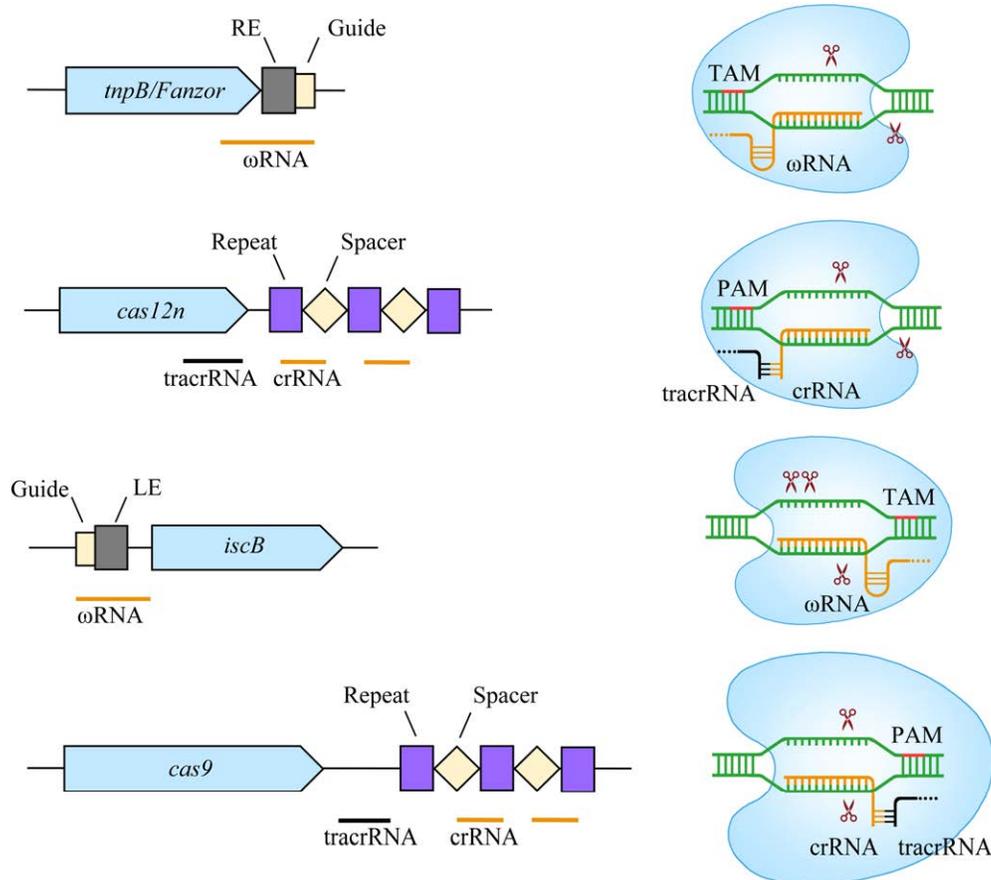


图1 OMEGA和CRISPR-Cas系统示意图

Figure 1 Schematic representation of OMEGA and CRISPR-Cas systems. This figure was drawn by the authors based on literature information to compare TnpB/Fanzor, Cas12n, IscB, and Cas9 systems. Left: Gene architecture of different systems; Right: Schematic of effector-gRNA-target DNA ternary complexes. Scissors indicate cleavage sites.

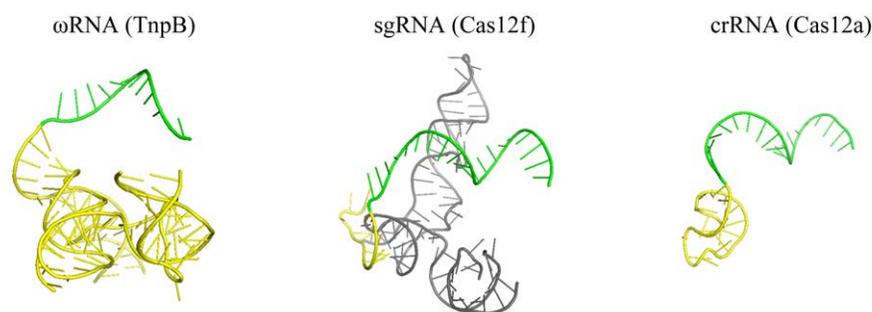


图 2 不同类型的 gRNA 结构

Figure 2 Structures of different kinds of gRNA. This figure was drawn by the authors based on literature information to compare scaffold of ω RNA from ISDra2 TnpB (PDB ID: 8H1J)^[38], sgRNA from an uncultured archaeon Cas12f (PDB ID: 7C7L)^[48], and crRNA from *Lachnospiraceae bacterium* Cas12a (PDB ID: 5XUS)^[49] ternary complexes. Green molecules indicate spacer region (guide); Grey molecules indicate tracrRNA-derived region.

表 1 OMEGA 和 CRISPR-Cas 系统特征分析

Table 1 Characteristics of OMEGA and CRISPR-Cas systems

Effector	Protein length (aa)	TracrRNA	Cleavage pattern	Collateral activity	Reference
TnpB (monomer)	~400	No	5' overhang	Yes/No	[25-26,38-39,44]
Fanzor (monomer)	400–800	No	5' overhang, 3' overhang, blunt end	No	[27,44-45]
Cas12n (likely monomer)	400–700	Yes	5' overhang	Yes	[32,40]
Cas12f (asymmetric dimer)	400–700	Yes	5' overhang (three cuts)	Yes	[29,32,48,50-51]
Cas12a (monomer)	1 200–1 500	No	5' overhang	Yes	[52-54]
IscB (monomer)	~400	No	5' overhang (three cuts)	Not tested	[25,41-42]
Cas9 (monomer)	>700	Yes	Blunt end	No	[6,25,54-56]

Xiang 等对 TnpB 系统进行深入分析, 结合其生理功能, 发现其 ω RNA 支架区和 TAM 序列均可通过生物信息分析预测^[46]。TnpB 系统的 TAM 序列通常为 4–5 nt, 大多有 T 碱基偏好^[46,57]。体外生化实验结果表明, TnpB 可以加工其自身 mRNA 产生 ω RNA^[57]。TnpB 具有依赖 TAM 的双链 DNA 切割活性, 产生黏性末端, 形成 5' 端突出; 此外, TnpB 还具有不依赖 TAM 的单链

DNA 切割活性^[25-26]。在识别单链或双链 DNA 靶标底物后, 其旁系切割活性会被激活, 可以对单链 DNA 进行非特异的切割^[25,39]。许多 Cas12 蛋白也具有旁系切割活性, 这种特性已被用于开发核酸检测工具^[40,54,58]。此外, Xu 等对来自极端嗜热古菌的 TnpB 进行了生化特征研究, 该蛋白适用温度范围广, 在 37–85 °C 条件下都有切割活性^[59]。

2.2 Fanzor 的特征

Fanzor 系统广泛存在于真核生物及其病毒中,如真菌、藻类和软体动物^[27,43-45]。Fanzor 起源于 TnpB 系统,其 ω RNA 也位于 TnpB 编码区的下游,有时会与编码区存在部分重合^[27](图 1)。Saito 等通过冷冻电镜方法对来自土壤真菌 (*Spizellomyces punctatus*) 的 Fanzor1 (SpuFz1) 进行研究,解析了 SpuFz1- ω RNA-靶标 DNA 三元复合物的结构;SpuFz1 具有与 TnpB 相似的核心区,由 REC 和 NUC 裂片形成二裂片结构;其 ω RNA 包括 75 nt 的支架区和 15 nt 的引导区,形成 2 个茎和 1 个连接区;值得注意的是,与 TnpB 系统的 ω RNA 相比,SpuFz1 系统的 ω RNA 缺少假结^[27]。对 Fanzor 系统的 ω RNA 进行二级结构预测,结果显示 Fanzor2 的 ω RNA 能形成假结,但是 Fanzor1 的 ω RNA 不能形成假结^[45]。目前,还未见关于 Fanzor2 结构的报道,对其结构的解析将有助于深入比较不同 Fanzor 的差异及 Fanzor 与 TnpB 的差异。

Fanzor 蛋白识别的 TAM 碱基数通常为 5 nt 以内,但不局限于 4-5 nt,不同的 Fanzor 蛋白偏好的碱基类型多样;Fanzor 具有多样化的切割模式,能产生 5'端突出、3'端突出或钝末端;Fanzor 仅切割具有 TAM 的靶标双链 DNA,对单链 DNA 或 RNA 均不切割^[27]。Fanzor 与 TnpB 蛋白的序列比对结果显示,Fanzor 和部分 TnpB 的 RuvC 活性位点发生了重排^[44-45]。值得注意的是,这些 RuvC 活性位点发生了重排的系统均无显著的旁系切割活性^[27,44](表 1)。这一特征为开发更安全的小型化基因组编辑工具提供可能。

2.3 IscB 的特征

IscB 核酸酶来自 IS200/IS650 转座子家族,其 ω RNA 转录自 IscB 上游序列,结构具有多样性,通常超过 200 nt^[25](图 1)。Schuler 等对来自肠道宏基因组数据的 OgeuIscB 蛋白进行研究,

通过冷冻电镜解析了 OgeuIscB- ω RNA-靶标 DNA 三元复合物的结构,其中 OgeuIscB 核酸酶的大小为 496 个氨基酸; ω RNA 为 222 nt,包括 16 nt 的引导区^[42]。OgeuIscB 也是首个被成功用于真核细胞基因组编辑的 IscB 蛋白^[25]。OgeuIscB 与 Cas9 系统在结构形成、与核酸的结合方式等方面都具有相似性;OgeuIscB 与 Cas9 蛋白最大的结构差异是 OgeuIscB 缺少 REC 裂片;不过,OgeuIscB 系统中的 ω RNA 形成了复杂的三级结构(包括 5 个茎环和 1 个假结),发挥了 Cas9 的 REC 裂片功能;因此,OgeuIscB 系统形成了更扁平的二裂片结构;Cas9 的 gRNA 也保留了部分与 ω RNA 相似的结构,包括假结;此外,IscB 蛋白的 N 端具有 PLMP 结构域,可能参与稳定 IscB- ω RNA 复合物,而 Cas9 蛋白中不包含该结构域^[42]。Kato 等也对 OgeuIscB 三元复合物进行了结构解析,进一步验证了以上结果^[41]。比较 OMEGA 系统与 CRISPR-Cas 系统的结构特点可以发现,进化过程中 RNA 的功能逐渐被蛋白取代,这些结果为 RNA 世界假说提供了支持,即 RNA 是生命的源头,生命进化的早期没有蛋白质,是 RNA 发挥催化功能^[60-61]。

IscB 识别的 PAM 也被称靶标邻近基序 (target-adjacent motif, TAM),其识别的 TAM 碱基数通常为 5 nt 以内,不同的 IscB 蛋白偏好的碱基类型多样^[25]。Altae-Tran 等对来自瓦氏别样着色菌 (*Allochromatium warmingii*) 的 IscB (AwaIscB) 进行研究,结果显示其在靶标链上的切割位点位于 TAM 上游 3 nt 处,与 Cas9 一致;然而其在非靶标链上的切割位点位于 TAM 上游 8 nt 或 12 nt 处,最终形成 5'端突出,不同于 Cas9 切割双链主要产生钝末端^[25,56](表 1)。Altae-Tran 等突变 AwaIscB 的 RuvC 活性位点后,其无法切割非靶标链;突变 HNH 活性位点后,其无法切割靶标链;突变以上 2 个结构域的活性位点后,

其无法切割任一条链^[25]。值得注意的是,来自真核生物绿藻叶绿体基因组的 IscB 蛋白具有简单的 NNG-TAM 偏好^[25]。这种具有简单 TAM 偏好的系统,有望构建 TAM 限制小的基因组编辑工具。

3 OMEGA 系统的应用

3.1 TnpB 的应用

Xiang 等对 TnpB 进行了大规模挖掘和分析:第一轮筛选出 64 个 TnpB,其中 25 个在大肠杆菌中有活性,3 个在人体细胞中有活性;结合 TnpB 的生理功能对这些系统的特点进行分析发现,其 ω RNA 支架区和 TAM 序列均可通过生物信息分析准确预测;对可能影响 TnpB 活性的因素进行分析,提出来自多拷贝转座子的 TnpB 系统可能具有更高的活性;对有活性和无活性的 TnpB 进行比较,鉴定了 10 个可能影响活性的保守氨基酸位点^[46]。以上分析为大规模筛选高活性 TnpB 核酸酶奠定了基础。他们基于以上结果重新设定筛选标准,进行了第二轮筛选,获得 14 个 TnpB:其中 8 个在大肠杆菌中有活性,2 个在人体细胞中有活性;相比于第一轮,他们得到的在大肠杆菌中有活性的 TnpB 比例从 39.1% (25/64) 提高到 57.1% (8/14),在人体细胞中有活性的 TnpB 比例从 4.7% (3/64) 提高到 14.3% (2/14);第二轮筛选出的 ISAam1 和 ISYmu1 在人体细胞中展现出最好的编辑效果,蛋白大小分别为 369 个氨基酸和 382 个氨基酸^[46]。他们将这 2 个系统和 SaCas9^[62]及 2 个优化后的 Cas12f 突变体(Un1Cas12f1^[63]和 CasMINI^[64])进行比较,分别在人和小鼠的细胞系及小鼠体内进行基因组编辑测试,结果显示这 2 个 TnpB 系统与 SaCas9 编辑效率相当,比 Cas12f 突变体效率高,特异性也不低;在目前已知的基因组编辑工具中,ISAam1 具有最小的 CRISPR 样效应

蛋白^[46]。此外, Li 等对 ISDra2 TnpB 系统的 ω RNA 进行多轮工程化改造,获得了高效的变体;将其用单 AAV 递送至小鼠肝脏,实现了酪氨酸血症的治疗;全基因组脱靶分析未检测到显著的脱靶现象,表明 TnpB 编辑工具有较高的安全性^[28]。另一方面, Xu 等利用来自极端嗜热古菌的 TnpB 开发了可以用于古菌和细菌的基因组编辑工具,其蛋白的高稳定性具有递送优势^[59]。该工作也提示,广泛存在的 OMEGA 系统为大量生物体开发内源性的基因组编辑工具提供可能。

3.2 Fanzor 的应用

Saito 等选择了 4 个在体外实验中验证有切割活性的 Fanzor 系统在人体细胞中进行基因组编辑测试,包括 2 个 Fanzor1 系统和 2 个 Fanzor2 系统;他们在人体细胞中选择了 8 个靶点进行测试,其中 3 个系统能检测到编辑活性,包括 SpuFz1,但效率均不高;于是他们进一步对 SpuFz1 蛋白及其 ω RNA 进行改造,改造后 SpuFz1 编辑效率显著提升,可达 18.4%^[27]。Jiang 等测试了 4 个 Fanzor2 系统和 1 个 Fanzor1 系统,它们在人体细胞中的编辑效率可达 15%^[44]。目前 Fanzor 系统在人体细胞中的编辑效率仍较低,需要挖掘更多的 Fanzor 系统进行测试或进一步改造已有系统。

3.3 IscB 的应用

Han 等对 OgeuIscB 蛋白及其 ω RNA 进行工程化改造,获得了可以对哺乳动物进行高效基因组编辑的工具;结合结构数据,他们对 ω RNA 进行了部分删除和突变,也对蛋白进行了选择性突变和扫描突变,最终获得的系统被命名为工程化改造的 IscB (engineered IscB, enIscB)^[65]。Han 等将 T5 外切酶与 enIscB 蛋白融合表达后,其编辑效率可以与 SpG^[66] (SpCas9 突变体,识别简单的 NGN-PAM)相当,并且具有更少的染色

体易位情况^[65]。Han 等还将突变后具有切口酶活性的 enIscB 分别与腺苷脱氨酶和胞嘧啶脱氨酶融合，开发了小型的碱基编辑器，最高的编辑效率可达 92%^[65]。

4 展望

OMEGA 系统(TnpB、Fanzor 和 IscB)来自自私的转座元件，在三域生命中具有广泛的分布，为开发大量内源性基因编辑工具提供了可能，也极大地丰富了基因组编辑工具资源。其中 TnpB 是原核生物中分布最广泛的蛋白之一，暗示可能还有许多 Cas12 新亚型等待挖掘和开发。OMEGA 蛋白具有 ω RNA 引导的核酸酶活性，原核生物中的 OMEGA 蛋白有助于转座子的保留和传播^[26,46-47]。对 TnpB 生理功能的研究，加速了高活性 TnpB 的大规模筛选^[46]。真核生物中 OMEGA 是否具有同样的生理功能仍不清楚。TnpB 和 IscB 都具有 RuvC 样结构域和 BH 结构域，它们分别被认为是 Cas12 和 Cas9 的祖先；然而 Cas12 不同亚型可能是由不同的 TnpB 独立进化而来，而 Cas9 可能起源于一个特定的 IscB 进化支^[25,29-32]。IscB 相较于 TnpB 多了 1 个 HNH 核酸酶结构域，它们之间的进化关系以及 RuvC 核酸酶进化为 OMEGA 蛋白的过程仍需要进一步解析。

目前对 OMEGA 编辑工具的开发才刚起步，有很多工作需要完成，同时也面临很多挑战。OMEGA 系统分布十分广泛，然而，想获得高效的基因组编辑工具，通常需要合成和测试大量的候选系统，其中可能只有少数具有高编辑活性。因此深入研究 OMEGA 系统的特点，建立核酸酶活性的预测机制是编辑工具开发的重点和难点。如通过建立序列、结构和功能的联系来筛选具有更高编辑潜力的系统，将加速高效编辑工具

的开发，并降低系统合成和测试的成本。Xiang 等的工作提供了很好的范例：他们对可能影响 TnpB 活性的因素进行分析，建立了大规模挖掘高活性 TnpB 核酸酶的流程，并成功获得了在真核细胞中具有高编辑活性的 TnpB^[46]。基于 IscB 的编辑工具仍较少，对来自真核细胞的 IscB 的挖掘和研究也有待进一步展开。TnpB 识别较复杂的 TAM (4–5 nt)，限制了其靶向范围，需要通过工程化改造拓展其 TAM 识别谱。然而，要在拓展编辑工具 TAM 识别谱的同时保持高编辑效率是改造的难点。目前测试的 Fanzor 在真核细胞中的编辑活性均较低。若要开发高效的编辑工具，还需进行系统性的优化改造或者筛选更高效的 Fanzor 蛋白。另一方面，最近 Liang 等通过引入突变获得了具有切口酶活性的 Cas12a 蛋白，基于此构建的先导编辑工具效率显著高于基于 Cas12a 核酸酶的先导编辑工具^[67]。TnpB 和 Fanzor 也有望通过引入合适的位点突变获得切口酶，从而构建精简的碱基编辑或先导编辑工具。

综上所述，未来对于 OMEGA 系统的研究可以从以下 3 个方面展开。第一，进一步对 OMEGA 系统进行同源搜索。通过深入分析 RuvC 核酸酶、原核生物中的 OMEGA 及真核中的 OMEGA 的进化关系，丰富编辑工具的资源库。第二，对 OMEGA 系统的作用机制和生理功能进行深入解析。更多的结构和功能解析将为工具的优化打下基础，也有助于鉴定影响 OMEGA 活性的关键因素，从而加快筛选流程、节约成本。第三，对更多 OMEGA 系统进行测试和优化。对 OMEGA 蛋白和 ω RNA 进行理性设计与定向进化，尝试提高编辑效率、减少脱靶或简化 TAM。还可以通过融合其他功能元件开发小型化的碱基编辑或先导编辑工具。

参考文献

- [1] ISHINO Y, SHINAGAWA H, MAKINO K, AMEMURA M, NAKATA A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product[J]. Journal of Bacteriology, 1987, 169(12): 5429-5433.
- [2] BARRANGOU R, FREMAUX C, DEVEAU H, RICHARDS M, BOYAVAL P, MOINEAU S, ROMERO DA, HORVATH P. CRISPR provides acquired resistance against viruses in prokaryotes[J]. Science, 2007, 315(5819): 1709-1712.
- [3] DELTCHEVA E, CHYLINSKI K, SHARMA CM, GONZALES K, CHAO YJ, PIRZADA ZA, ECKERT MR, VOGEL J, CHARPENTIER E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III[J]. Nature, 2011, 471: 602-607.
- [4] GARNEAU JE, DUPUIS MÈ, VILLION M, ROMERO DA, BARRANGOU R, BOYAVAL P, FREMAUX C, HORVATH P, MAGADÁN AH, MOINEAU S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA[J]. Nature, 2010, 468: 67-71.
- [5] BROUNS SJJ, JORE MM, LUNDRGREN M, WESTRA ER, SLIJKHUIS RJH, SNIJDERS APL, DICKMAN MJ, MAKAROVA KS, KOONIN EV, van der OOST J. Small CRISPR RNAs guide antiviral defense in prokaryotes[J]. Science, 2008, 321(5891): 960-964.
- [6] JINEK M, CHYLINSKI K, FONFARA I, HAUER M, DOUDNA JA, CHARPENTIER E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity[J]. Science, 2012, 337(6096): 816-821.
- [7] CONG L, RAN FA, COX D, LIN SL, BARRETTO R, HABIB N, HSU PD, WU XB, JIANG WY, MARRAFFINI LA, ZHANG F. Multiplex genome engineering using CRISPR/Cas systems[J]. Science, 2013, 339(6121): 819-823.
- [8] WANG JY, DOUDNA JA. CRISPR technology: a decade of genome editing is only the beginning[J]. Science, 2023, 379(6629): eadd8643.
- [9] CHENG FY, GONG LY, ZHAO DH, YANG HB, ZHOU J, LI M, XIANG H. Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon[J]. Journal of Genetics and Genomics, 2017, 44(11): 541-548.
- [10] DU KX, GONG LY, LI M, YU HY, XIANG H. Reprogramming the endogenous type I CRISPR-Cas system for simultaneous gene regulation and editing in *Haloarcula hispanica*[J]. mLife, 2022, 1(1): 40-50.
- [11] XU ZL, LI M, LI YR, CAO HL, MIAO L, XU ZC, HIGUCHI Y, YAMASAKI S, NISHINO K, WOO PCY, XIANG H, YAN AX. Native CRISPR-Cas-mediated genome editing enables dissecting and sensitizing clinical multidrug-resistant *P. aeruginosa*[J]. Cell Reports, 2019, 29(6): 1707-1717.e3.
- [12] LI SN, LIN DX, ZHANG YW, DENG M, CHEN YX, LV B, LI BS, LEI Y, WANG YP, ZHAO L, LIANG YT, LIU JX, CHEN KL, LIU ZY, XIAO J, QIU JL, GAO CX. Genome-edited powdery mildew resistance in wheat without growth penalties[J]. Nature, 2022, 602: 455-460.
- [13] KANG Y, CHU C, WANG F, NIU YY. CRISPR/Cas9-mediated genome editing in nonhuman primates[J]. Disease Models & Mechanisms, 2019, 12(10): dmm039982.
- [14] 龚路遥, 向华. 2020 年度诺贝尔化学奖: 源自微生物学前沿研究的重大突破[J]. 科学通报, 2020, 65(36): 4171-4176.
GONG LY, XIANG H. Nobel Prize in Chemistry 2020: the great breakthrough from frontier research in microbiology[J]. Chinese Science Bulletin, 2020, 65(36): 4171-4176 (in Chinese).
- [15] YE LJ, ZHAO DD, LI J, WANG YR, LI B, YANG YZ, HOU XT, WANG HB, WEI ZD, LIU XQ, LI YQ, LI SW, LIU YJ, ZHANG XL, BI CH. Glycosylase-based base editors for efficient T-to-G and C-to-G editing in mammalian cells[J]. Nature Biotechnology, 2024. DOI: <https://doi.org/10.1038/s41587-023-02050-w>.
- [16] TONG HW, WANG XC, LIU YH, LIU NN, LI Y, LUO JM, MA Q, WU DN, LI JY, XU CL, YANG H. Programmable A-to-Y base editing by fusing an adenine base editor with an N-methylpurine DNA glycosylase[J]. Nature Biotechnology, 2023, 41: 1080-1084.
- [17] ANZALONE AV, RANDOLPH PB, DAVIS JR, SOUSA AA, KOBLAN LW, LEVY JM, CHEN PJ, WILSON C, NEWBY GA, RAGURAM A, LIU DR. Search-and-replace genome editing without double-strand breaks or donor DNA[J]. Nature, 2019, 576: 149-157.
- [18] KOMOR AC, KIM YB, PACKER MS, ZURIS JA, LIU DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage[J]. Nature, 2016, 533: 420-424.

- [19] WANG D, ZHANG F, GAO GP. CRISPR-based therapeutic genome editing: strategies and *in vivo* delivery by AAV vectors[J]. *Cell*, 2020, 181(1): 136-150.
- [20] 范杏飞, 顾晶雯, 顾晓燕, 杜丽晶, 蒋俊锋, 王越. CRISPR-Cas 系统在动物及植物中不同的递送方式研究进展[J]. *解剖学杂志*, 2022, 45(2): 153-156.
FAN XF, GU JW, GU XY, DU LJ, JIANG JF, WANG Y. Progress in the study of different delivery modes of CRISPR-Cas system in animals and plants[J]. *Chinese Journal of Anatomy*, 2022, 45(2): 153-156 (in Chinese).
- [21] 于宗菲, 翁丽涵, 孙诚诚, 曹晓钰, 叶振. CRISPR/Cas9 系统技术难关: 脱靶效应及其优化方法[J]. *山东第一医科大学(山东省医学科学院)学报*, 2023, 44(1): 74-80.
YU ZF, WENG LH, SUN CC, CAO XY, YE Z. Technical difficulties of the CRISPR/Cas9 System: off-target effects and its optimization methods[J]. *Journal of Shandong First Medical University & Shandong Academy of Medical Sciences*, 2023, 44(1): 74-80 (in Chinese).
- [22] YU ZX, LU ZK, LI JJ, WANG YY, WU PF, LI YN, ZHOU YF, LI BL, ZHANG H, LIU YZ, MA LJ. PEAC-seq adopts prime editor to detect CRISPR off-target and DNA translocation[J]. *Nature Communications*, 2022, 13: 7545.
- [23] HILLE F, RICHTER H, WONG SP, BRATOVIČ M, RESSEL S, CHARPENTIER E. The biology of CRISPR-Cas: backward and forward[J]. *Cell*, 2018, 172(6): 1239-1259.
- [24] NISHIMASU H, SHI X, ISHIGURO S, GAO LY, HIRANO S, OKAZAKI S, NODA T, ABUDAYYEH OO, GOOTENBERG JS, MORI H, OURA S, HOLMES B, TANAKA M, SEKI M, HIRANO H, ABURATANI H, ISHITANI R, IKAWA M, YACHIE N, ZHANG F, et al. Engineered CRISPR-Cas9 nuclease with expanded targeting space[J]. *Science*, 2018, 361(6408): 1259-1262.
- [25] ALTAE-TRAN H, KANNAN S, DEMIRCIOGLU FE, OSHIRO R, NETY SP, MCKAY LJ, DLAKIĆ M, INSKEEP WP, MAKAROVA KS, MACRAE RK, KOONIN EV, ZHANG F. The widespread IS200/IS605 transposon family encodes diverse programmable RNA-guided endonucleases[J]. *Science*, 2021, 374(6563): 57-65.
- [26] KARVELIS T, DRUTEIKA G, BIGELYTE G, BUDRE K, ZEDAIVEINYTE R, SILANSKAS A, KAZLAUSKAS D, VENCLOVAS Č, SIKSNYS V. Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease[J]. *Nature*, 2021, 599: 692-696.
- [27] SAITO M, XU PY, FAURE G, MAGUIRE S, KANNAN S, ALTAE-TRAN H, VO S, DESIMONE A, MACRAE RK, ZHANG F. Fanzor is a eukaryotic programmable RNA-guided endonuclease[J]. *Nature*, 2023, 620: 660-668.
- [28] LI ZF, GUO RC, SUN XZ, LI GL, SHAO Z, HUO XN, YANG RR, LIU XY, CAO X, ZHANG HN, ZHANG WH, ZHANG XY, MA SY, ZHANG ML, LIU YH, YAO YN, SHI JQ, YANG H, HU CY, ZHOU YS, et al. Engineering a transposon-associated TnpB- ω RNA system for efficient gene editing and phenotypic correction of a tyrosinaemia mouse model[J]. *Nature Communications*, 2024, 15: 831.
- [29] MAKAROVA KS, WOLF YI, IRANZO J, SHMAKOV SA, ALKHNABASHI OS, BROUNS SJJ, CHARPENTIER E, CHENG D, HAFT DH, HORVATH P, MOINEAU S, MOJICA FJM, SCOTT D, SHAH SA, SIKSNYS V, TERNS MP, VENCLOVAS Č, WHITE MF, YAKUNIN AF, YAN W, et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants[J]. *Nature Reviews Microbiology*, 2020, 18: 67-83.
- [30] KAPITONOV VV, MAKAROVA KS, KOONIN EV. ISC, a novel group of bacterial and archaeal DNA transposons that encode Cas9 homologs[J]. *Journal of Bacteriology*, 2015, 198(5): 797-807.
- [31] SHMAKOV S, ABUDAYYEH OO, MAKAROVA KS, WOLF YI, GOOTENBERG JS, SEMENOVA E, MINAKHIN L, JOUNG J, KONERMANN S, SEVERINOV K, ZHANG F, KOONIN EV. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems[J]. *Molecular Cell*, 2015, 60(3): 385-397.
- [32] SHMAKOV S, SMARGON A, SCOTT D, COX D, PYZOSHA N, YAN W, ABUDAYYEH OO, GOOTENBERG JS, MAKAROVA KS, WOLF YI, SEVERINOV K, ZHANG F, KOONIN EV. Diversity and evolution of class 2 CRISPR-Cas systems[J]. *Nature Reviews Microbiology*, 2017, 15: 169-182.
- [33] SIGUIER P, GOURBEYRE E, CHANDLER M. Bacterial insertion sequences: their genomic impact and diversity[J]. *FEMS Microbiology Reviews*, 2014, 38(5): 865-891.
- [34] HE S, CORNELOUP A, GUYNET C, LAVATINE L, CAUMONT-SARCOS A, SIGUIER P, MARTY B,

- DYDA F, CHANDLER M, HOANG BT. The IS200/IS605 family and “peel and paste” single-strand transposition mechanism[J]. *Microbiology Spectrum*, 2015, 3(4): 609-630.
- [35] BOOCOCK MR, RICE PA. A proposed mechanism for IS607-family serine transposases[J]. *Mobile DNA*, 2013, 4(1): 24.
- [36] CHEN WY, MANDALI S, HANCOCK SP, KUMAR P, COLLAZO M, CASCIO D, JOHNSON RC. Multiple serine transposase dimers assemble the transposon-end synaptic complex during IS 607-family transposition[J]. *eLife*, 2018, 7: e39611.
- [37] PASTERNAK C, DULERMO R, TON-HOANG B, DEBUCHY R, SIGUIER P, COSTE G, CHANDLER M, SOMMER S. IS_{Dra2} transposition in *Deinococcus radiodurans* is downregulated by TnpB[J]. *Molecular Microbiology*, 2013, 88(2): 443-455.
- [38] NAKAGAWA R, HIRANO H, OMURA SN, NETY S, KANNAN S, ALTAE-TRAN H, YAO X, SAKAGUCHI Y, OHIRA T, WU WY, NAKAYAMA H, SHUTO Y, TANAKA T, SANO FK, KUSAKIZAKO T, KISE Y, ITOH Y, DOHMAE N, van der OOST J, SUZUKI T, et al. Cryo-EM structure of the transposon-associated TnpB enzyme[J]. *Nature*, 2023, 616: 390-397.
- [39] SASNAUSKAS G, TAMULAITIENE G, DRUTEIKA G, CARABIAS A, SILANSKAS A, KAZLAUSKAS D, VENCLOVAS Č, MONTOYA G, KARVELIS T, SIKSNYS V. TnpB structure reveals minimal functional core of Cas12 nuclease family[J]. *Nature*, 2023, 616: 384-389.
- [40] CHEN WZ, MA JC, WU ZW, WANG ZP, ZHANG HY, FU WH, PAN D, SHI J, JI QJ. Cas12n nucleases, early evolutionary intermediates of type V CRISPR, comprise a distinct family of miniature genome editors[J]. *Molecular Cell*, 2023, 83(15): 2768-2780.e6.
- [41] KATO K, OKAZAKI S, KANNAN S, ALTAE-TRAN H, ESRA DEMIRCIOLU F, ISAYAMA Y, ISHIKAWA J, FUKUDA M, MACRAE RK, NISHIZAWA T, MAKAROVA KS, KOONIN EV, ZHANG F, NISHIMASU H. Structure of the IscB- ω RNA ribonucleoprotein complex, the likely ancestor of CRISPR-Cas9[J]. *Nature Communications*, 2022, 13(1): 6719.
- [42] SCHULER G, HU CY, KE AL. Structural basis for RNA-guided DNA cleavage by IscB- ω RNA and mechanistic comparison with Cas9[J]. *Science*, 2022, 376(6600): 1476-1481.
- [43] BAO WD, JURKA J. Homologues of bacterial TnpB_IS605 are widespread in diverse eukaryotic transposable elements[J]. *Mobile DNA*, 2013, 4(1): 12.
- [44] JIANG KY, LIM J, SGRIZZI S, TRINH M, KAYABOLEN A, YUTIN N, BAO WD, KATO K, KOONIN EV, GOOTENBERG JS, ABUDAYYEH OO. Programmable RNA-guided DNA endonucleases are widespread in eukaryotes and their viruses[J]. *Science Advances*, 2023, 9(39): eadk0171.
- [45] YOON PH, SKOPINTSEV P, SHI HL, CHEN LX, ADLER BA, AL-SHIMARY M, CRAIG RJ, LOI KJ, DeTURK EC, LI Z, AMERASEKERA J, TRINIDAD M, NISONOFF H, CHEN K, LAHIRI A, BOGER R, JACOBSEN S, BANFIELD JF, DOUDNA JA. Eukaryotic RNA-guided endonucleases evolved from a unique clade of bacterial enzymes[J]. *Nucleic Acids Research*, 2023, 51(22): 12414-12427.
- [46] XIANG GH, LI YQ, SUN J, HUO YY, CAO SW, CAO YW, GUO YY, YANG L, CAI YJ, ZHANG YE, WANG HY. Evolutionary mining and functional characterization of TnpB nucleases identify efficient miniature genome editors[J]. *Nature Biotechnology*, 2023.
- [47] MEERS C, LE HC, PESARI SR, HOFFMANN FT, WALKER MWG, GEZELLE J, TANG S, STERNBERG SH. Transposon-encoded nucleases use guide RNAs to promote their selfish spread[J]. *Nature*, 2023, 622: 863-871.
- [48] TAKEDA SN, NAKAGAWA R, OKAZAKI S, HIRANO H, KOBAYASHI K, KUSAKIZAKO T, NISHIZAWA T, YAMASHITA K, NISHIMASU H, NUREKI O. Structure of the miniature type V-F CRISPR-Cas effector enzyme[J]. *Molecular Cell*, 2021, 81(3): 558-570.e3.
- [49] YAMANO T, ZETSCHKE B, ISHITANI R, ZHANG F, NISHIMASU H, NUREKI O. Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1[J]. *Molecular Cell*, 2017, 67(4): 633-645.e3.
- [50] XIAO RJ, LI Z, WANG SK, HAN RJ, CHANG LF. Structural basis for substrate recognition and cleavage by the dimerization-dependent CRISPR-Cas12f nuclease[J]. *Nucleic Acids Research*, 2021, 49(7): 4120-4128.
- [51] WU ZW, ZHANG YF, YU HP, PAN D, WANG YJ, WANG YN, LI F, LIU C, NAN H, CHEN WZ, JI QJ. Programmed genome editing by a miniature

- CRISPR-Cas12f nuclease[J]. *Nature Chemical Biology*, 2021, 17: 1132-1138.
- [52] ZETSCHKE B, GOOTENBERG JS, ABUDAYYEH OO, SLAYMAKER IM, MAKAROVA KS, ESSLETZBICHLER P, VOLZ SE, JOUNG J, van der OOST J, REGEV A, KOONIN EV, ZHANG F. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system[J]. *Cell*, 2015, 163(3): 759-771.
- [53] YAMANO T, NISHIMASU H, ZETSCHKE B, HIRANO H, SLAYMAKER IM, LI YQ, FEDOROVA I, NAKANE T, MAKAROVA KS, KOONIN EV, ISHITANI R, ZHANG F, NUREKI O. Crystal structure of Cpf1 in complex with guide RNA and target DNA[J]. *Cell*, 2016, 165(4): 949-962.
- [54] CHEN JS, MA EB, HARRINGTON LB, da COSTA M, TIAN XR, PALEFSKY JM, DOUDNA JA. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity[J]. *Science*, 2018, 360(6387): 436-439.
- [55] NISHIMASU H, RAN FA, HSU PD, KONERMANN S, SHEHATA SI, DOHMAE N, ISHITANI R, ZHANG F, NUREKI O. Crystal structure of Cas9 in complex with guide RNA and target DNA[J]. *Cell*, 2014, 156(5): 935-949.
- [56] GASIUNAS G, YOUNG JK, KARVELIS T, KAZLAUSKAS D, URBAITIS T, JASNAUSKAITE M, GRUSYTE MM, PAULRAJ S, WANG PH, HOU ZL, DOOLEY SK, CIGAN M, ALARCON C, CHILCOAT ND, BIGELYTE G, CURCURU JL, MABUCHI M, SUN ZY, FUCHS RT, et al. A catalogue of biochemically diverse CRISPR-Cas9 orthologs[J]. *Nature Communications*, 2020, 11: 5512.
- [57] NETY SP, ALTAE-TRAN H, KANNAN S, DEMIRCIOGLU FE, FAURE G, HIRANO S, MEARS K, ZHANG YG, MACRAE RK, ZHANG F. The transposon-encoded protein TnpB processes its own mRNA into ω RNA for guided nuclease activity[J]. *The CRISPR Journal*, 2023, 6(3): 232-242.
- [58] LU SH, TONG XH, HAN Y, ZHANG K, ZHANG YZ, CHEN QB, DUAN JY, LEI XL, HUANG MH, QIU Y, ZHANG DY, ZHOU X, ZHANG Y, YIN H. Fast and sensitive detection of SARS-CoV-2 RNA using suboptimal protospacer adjacent motifs for Cas12a[J]. *Nature Biomedical Engineering*, 2022, 6: 286-297.
- [59] XU Y, LIU T, WANG J, XIONG BY, LIU L, PENG N. Reprogramming an RNA-guided archaeal TnpB endonuclease for genome editing[J]. *Cell Discovery*, 2023, 9: 112.
- [60] ROBERTSON MP, JOYCE GF. The origins of the RNA world[J]. *Cold Spring Harbor Perspectives in Biology*, 2012, 4(5): a003608.
- [61] SUN A, LI CP, CHEN ZH, ZHANG SY, LI DY, YANG Y, LI LQ, ZHAO YQ, WANG KC, LI ZF, LIU JX, LIU ST, WANG J, LIU JJ G. The compact Cas π (Cas12l) 'bracelet' provides a unique structural platform for DNA manipulation[J]. *Cell Research*, 2023, 33: 229-244.
- [62] RAN FA, CONG L, YAN WX, SCOTT DA, GOOTENBERG JS, KRIZ AJ, ZETSCHKE B, SHALEM O, WU XB, MAKAROVA KS, KOONIN EV, SHARP PA, ZHANG F. *In vivo* genome editing using *Staphylococcus aureus* Cas9[J]. *Nature*, 2015, 520: 186-191.
- [63] KIM DY, LEE JM, BIN MOON S, CHIN HJ, PARK S, LIM Y, KIM D, KOO T, KO JH, KIM YS. Efficient CRISPR editing with a hypercompact Cas12f1 and engineered guide RNAs delivered by adeno-associated virus[J]. *Nature Biotechnology*, 2022, 40: 94-102.
- [64] XU XS, CHEMPARATHY A, ZENG LP, KEMPTON HR, SHANG S, NAKAMURA M, QI LS. Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing[J]. *Molecular Cell*, 2021, 81(20): 4333-4345.e4.
- [65] HAN DY, XIAO QQ, WANG YF, ZHANG HN, DONG X, LI GL, KONG XF, WANG SH, SONG JH, ZHANG WH, ZHOU JX, BI LT, YUAN Y, SHI LY, ZHONG N, YANG H, ZHOU YS. Development of miniature base editors using engineered IscB nickase[J]. *Nature Methods*, 2023, 20: 1029-1036.
- [66] WALTON RT, CHRISTIE KA, WHITTAKER MN, KLEINSTIVER BP. Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants[J]. *Science*, 2020, 368(6488): 290-296.
- [67] LIANG RH, HE ZX, ZHAO KT, ZHU HC, HU JC, LIU GW, GAO Q, LIU MY, ZHANG R, QIU JL, GAO CX. Prime editing using CRISPR-Cas12a and circular RNAs in human cells[J]. *Nature Biotechnology*, 2024. DOI: <https://doi.org/10.1038/s41587-023-02095-x>.