

CRISPR 相关转座酶及其细菌基因组编辑应用

周晓杰^{1, 2} 杨思琪^{1, 2} 张译文³ 徐佳琪⁴ 杨晟¹

(1. 中国科学院分子植物科学卓越创新中心 中国科学院合成生物学重点实验室, 上海 200032; 2. 中国科学院大学, 北京 100049; 3. 中国科学院天津工业生物技术研究所, 天津 300308; 4. 浙江大学杭州国际科创中心, 杭州 311215)

摘要: CRISPR-Cas 能够在 RNA 引导下靶向 DNA 或 RNA 的特定序列, 改变 RNA 序列即可改变靶向位点, 利用这一可重编程特性已开发出了各种强大的遗传学工具。最近发现 CRISPR 元件在进化过程中被 Tn7 转座子劫持, 由此衍生出的 CRISPR 相关转座酶 (CRISPR-associated transposases, CASTs) 系统具有 RNA 引导 DNA 整合的能力, 被部署为靶点可重编程的基因组整合工具, 在大片段和多重基因整合上具有广阔的应用前景。本文追溯了 CASTs 的发现历程, 总结了不同类型 CASTs 的基因座结构特点、介导基因整合的机制模型以及其在多种革兰氏阴性细菌中的部署和应用。

关键词: CRISPR 相关转座酶; 移动遗传元件; 基因组编辑; 多重基因整合

DOI:10.13560/j.cnki.biotech.bull.1985.2022-1163

CRISPR-associated Transposases and Their Applications in Bacterial Genome Editing

ZHOU Xiao-jie^{1, 2} YANG Si-qi^{1, 2} ZHANG Yi-wen³ XU Jia-qi⁴ YANG Sheng¹

(1. CAS Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Chinese Academy of Sciences, Shanghai 200032; 2. University of Chinese Academy of Sciences, Beijing 100049; 3. Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308; 4. ZJU-Hangzhou Global Scientific and Technological Innovation Center, Hangzhou 311215)

Abstract: CRISPR-Cas can target a specific sequence of DNA or RNA guided by a short RNA guide. The target sites can be changed with the RNA sequences. A variety of powerful genetic tools have been developed based on this reprogramming property. Recently, it has been found that some CRISPR elements are captured by Tn7 transposons during evolution, the derived CRISPR-associated transposases (CASTs) system has the ability of RNA-guided DNA integration. Some of them has been deployed as programmable genome integration tools, which have broad application prospects in large fragments and multiple gene integration. This review traces the discovery of CASTs, summarizes the characteristics of different types of CASTs loci, the mechanism models of gene integration, and their deployments and applications in a variety of Gram-negative bacteria.

Key words: CRISPR-associated transposases; mobile genetic element; genome editing; multiplexed gene integration

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and associated proteins) 是广泛存在于古菌和细菌中的适应性免疫系统, 能够提供序列特异性的保护来抵御外源 DNA 或 RNA^[1]。CRISPR-Cas 发挥免疫功能分为 3 个阶段: (1) 适

应。由适应性模块中的 Cas1 和 Cas2 组成的复合物识别外源核酸, 将其加工整合至 CRISPR 阵列中, 以提供感染的记忆^[2]; (2) crRNA 生成。CRISPR 基因座被转录成 pre-crRNA, 随后被加工为成熟 crRNA^[3]; (3) 干扰。crRNA 引导 Cas 核酸酶靶向

收稿日期: 2022-09-21

基金项目: 国家杰出青年科学基金项目 (21825804), 国家自然科学基金创新群体 (31921006)

作者简介: 周晓杰, 女, 硕士研究生, 研究方向: 基因组编辑; E-mail: zhouxiaojie@cemps.ac.cn

通讯作者: 杨晟, 男, 博士, 研究员, 研究方向: 微生物分子遗传与合成生物学; E-mail: syang@sibs.ac.cn

并切割外源入侵 DNA/RNA^[4-5]。这种由 RNA 引导的靶位点切割具有高度灵活的可编程特性,通过改变 RNA 设计靶向位点,进一步利用宿主同源重组和非同源末端连接修复机制,可以实现基因的定点编辑^[6-7]。目前,CRISPR-Cas 系统已经被开发为多种强大的工具,可满足基因插入/敲除^[6,8]、转录激活^[9]、转录抑制^[10]、碱基编辑^[11]、RNA 编辑^[12]、DNA/RNA 检测^[13-14]等多种应用场景。在细菌多重基因组编辑应用中,利用 CRISPR-Cas 产生双链断裂的编辑效率受到宿主同源重组修复能力的限制^[15],这意味着同时对工程细胞的多位点进行高效编辑仍十分困难。

转座子是能够在基因组上从一个位置移动到另一个位置的移动遗传元件^[16],最早由 Barbara McClintock 在玉米中发现,她的研究揭示玉米染色体上的激活因子能够控制解离因子从基因组上解离和在基因组随机位置上整合^[17-18]。这一发现颠覆了人们“基因在染色体上是固定的”的认知,是遗传学发展史及基因编辑工具开发进程中重要的里程碑,而转座子及相关转座酶至今仍被用作真核/原核生物的突变体筛选及基因整合工具^[16]。根据转座机制的不同,转座可分为复制型转座(copy and paste)和剪切型转座(cut and paste)^[16],少部分转座子通过滚环式复制方式^[19]和自我合成的复制方式转座^[20]。Tn7 转座子是剪切型转座的代表,能够在 5 种蛋白 TnsA、TnsB、TnsC、TnsD、TnsE 介导下进行位点特异性转座^[21],其原理是 TnsB 识别转座子两端的末端序列,并与 TnsA 结合,二者分别切割供体序列的 3' 端和 5' 端并与其形成复合物。TnsD/E 是序列特异性 DNA 结合蛋白,负责靶点选择,分别以高频率引导转位到染色体上保守的 attTn7 位点和质粒。靶位点选择和 TnsAB 复合物之间的通信由 ATP 水解蛋白 TnsC 介导,TnsC 通过与两个靶点选择器(TnsD 或 TnsE)之一的相互作用,协调转座酶募集并将供体序列插入至靶位点^[22-24]。转座子介导整合的特性使其成为天然的 DNA 载体,能够有效地进行基因组插入,已经被开发为 DNA 定点和随机整合工具^[25],但存在靶序列特异性差、插入位点不可编程的问题,极大地限制了它们在定点基因整合上的应用。

最近发现 CRISPR 元件在进化过程中被 Tn7 转座子劫持^[26],这种协同进化衍生出了 CRISPR 相关转座酶系统,能够在 RNA 的引导下介导 DNA 转座,被部署为位点可编程的基因整合工具^[27-28]。CRISPR 相关转座酶(CRISPR-associated transposases, CASTs)系统由(1)CRISPR 元件,Cas12k 或 Cascade 和 RNA;(2)转座元件,转座酶 TnsB、ATP 水解蛋白 TnsC 和 TniQ,部分系统含有 TnsA;(3)嵌在转座子末端序列(left end/right end, LE/RE)之间的供体 DNA 三部分组成^[29],但缺乏原型 CRISPR-Cas 系统中的适应性模块(Cas1、Cas2)和核酸酶活性^[30]。其介导 DNA 整合的原理是 crRNA 引导单效应元件 Cas12k 或 Cascade 以 PAM 依赖或不依赖的方式靶向靶标位点,招募转座元件 TniQ、TnsC、TnsB 或 TnsAB 复合物将带有转座酶识别末端的供体 DNA 片段整合至靶标位点下游(图 1),整个过程中不产生双链断裂,因此基于 CASTs 开发的基因整合技术同时克服了 CRISPR-Cas 系统受同源重组效率限制和转座子位点不可编程的缺点,标志着基因组编辑工具的重大进步。

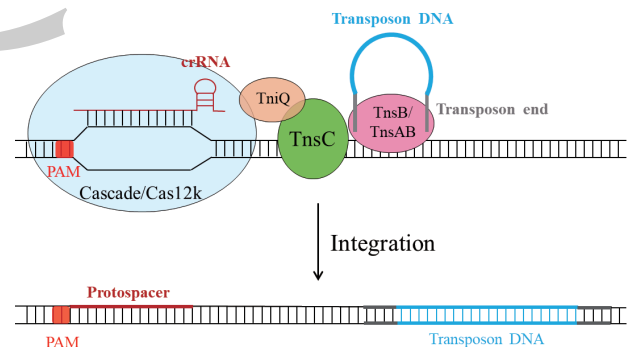


图 1 CRISPR 相关转座酶介导定点转座原理

Fig. 1 Mechanism of site-specific transposition mediated by CRISPR-associated transposases

1 CASTs 的发现

作为细菌和古菌的免疫系统,CRISPR-Cas 和移动遗传元件具有十分复杂的关系。几类移动遗传元件,病毒、质粒和转座子参与了 CRISPR-Cas 的起源和进化,而反过来 CRISPR-Cas 系统及其元件也被不同的移动遗传元件招募,形成了大量未被表征的功能^[30]。CRISPR-Cas 系统也在进化的过程中变得极

具多样性。在一些案例中, CRISPR-Cas 系统的适应性模块基因和核酸酶基因被单独编码或者完全缺失, 这些“最小” CRISPR 变体的生化和生理学功能尚未得到充分表征^[1, 31]。

2014 年, Krupovic 等^[32] 通过生物信息学分析发现了一种新型的古菌和细菌移动元件超家族, 并命名为 casposon, 它们编码与古菌和细菌的 CRISPR-Cas 适应性免疫系统相关的 Cas1 内切酶, 与 Polinton/Maverick 类的自合成真核 DNA 转座子具有几个共同特征, 包括末端倒置重复序列和 B 族 DNA 聚合酶的基因, 该研究首次发现利用 Cas1 内切酶插入宿主基因组并从中切除的移动遗传元件群体, 暗示了 CRISPR-Cas 系统和转座子的共进化过程。

2017 年, 研究人员在使用 I-F 型系统中最保守的 Cas7f 作为探针进行 PSI-BLAST 以全面鉴别 I-F 型 Cas 基因座时, 发现了含有最小 I-F 型和 I-B 型 CRISPR-Cas 系统的类 Tn7 转座子, 无核酸酶和转座子定位所需的 *tnsE*, 这种基因组成意味着它们能够进行 pre-crRNA 的处理以产生成熟 crRNA 和靶标结合, 但不发生靶向切割, 提示存在 RNA 引导的转座, 转座子编码的 CRISPR-Cas 系统在同源 DNA 位点形成 R 环, 将转座子靶向这些位点, 从而促进它们通过质粒和噬菌体传播^[26]。

2019 年, Sternberg 等^[27] 和 Zhang 等^[28] 分别表征了来自霍乱弧菌 (*Vibrio cholerae* Tn6677) 的 I-F 和霍夫曼伪枝藻 (*Scytonema hofmannii*) 的 V-K 型 CRISPR-Cas 效应子复合物指导伴随的转座酶将供体 DNA 片段包括转座子末端序列整合在基因组与引导 RNA 互补的 DNA 靶位点下游, 正式标志着可编程整合酶的发现。此后, 不断有新的 CASTs 被表征, 并被开发为高效的基因整合工具。

2 CASTs 的分类

CASTs 目前尚无全面系统的分类标准, 通常以捕获的 Cas 效应蛋白为特征, 从 CRISPR-Cas 的分类方法, 目前发现和预测的有 I-F、I-B、I-C、I-E、IV 和 V-K 共 6 种类型 (图 2)^[30, 33], 其中只有 I-F、I-B 与 V-K 型的活性已经得到表征。I-F 型 CASTs 由类 Tn7 元件捕获最小 I-F 型 CRISPR-Cas 得到, 在所有发现的亚型中具有最大的多样性。该亚型系统

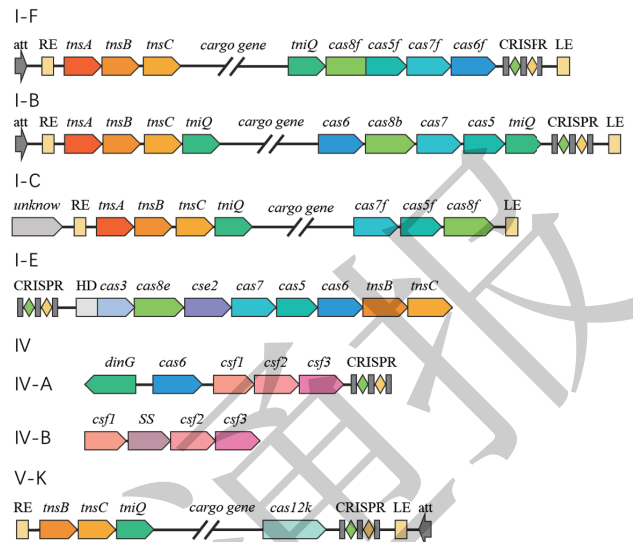


图 2 不同亚型 CASTs 的基因座结构

Fig. 2 Gene locus structure of diverse CASTs

通常由仅包含 2 段间隔序列的最小 CRISPR 阵列、Cas6f、Cas7f、Cas5f-Cas8f 融合蛋白和转座酶元件 TnsA、TnsB、TnsC 和 TniQ 组成 (图 1)^[30]。I-F 型 CASTs 目前已表征具有活性的有 21 种^[28, 34-36]。I-B 亚型系统最早在蓝藻中发现, 具有与 I-F 亚型相似的特征, 不同在于, I-B 亚型包含有 2 个 TniQ/TnsD 同源蛋白, 其中一个介导靶向基因组某些特定位点 (如 *glmS* 和 *tRNA-Val*), 这一过程称为归巢^[37]。目前已表征具有活性的 I-B 亚型 CASTs 有 3 种^[37-38]。Rybarski 等^[33] 通过宏基因组分析预测了 9 个非冗余 I-C 亚型 CASTs, 但其缺乏 CRISPR 阵列。在链霉菌科家族中发现的 I-E 亚型系统包含一个明显活跃的效应核酸酶和典型的 CRISPR 阵列, 并包含 Tn7 转座通常所需要的两个基因: *tnsB* 和 *tnsC*, 但是未识别到这些基因邻近的顺式作用末端结构, 这提示它们不是活性转座子, 相反的, 可能与 CRISPR 系统招募 Cas1 转座酶类似, 与 Tn7 相关的基因可能被 I-E 亚型 CRISPR-Cas 系统作为一个替代适应模块招募^[30]。IV 型系统是 I 类 CRISPR-Cas 系统的简化版本, 由于 IV 型系统的 Cas 蛋白序列与其他系统的同源物之间的差异巨大而被划分为一个单独的类型^[1]。IV 型系统有两种变体 IV-A 亚型和 IV-B 亚型, 都含有高度分化的 Cas5 (Csf3)、Cas7 (Csf2) 和 Cas8

样大亚基 (Csf1) 的效应模块基因, IV-A 亚型还编码 DinG 解旋酶家族^[30]。Özcan 等^[39]解析了 IV-A 亚型系统的效应复合体的结构, 发现其在结构和在所有可能的功能上都类似于由类 Tn7 转座子编码的最小的 I 型系统, 似乎是 I 型系统极其分散的衍生物。IV-B 亚型与 I-C 亚型一样, 缺乏 CRISPR 阵列, 可能通过远端来自其他活性 CRISPR-Cas 系统的 CRISPR 阵列发挥作用^[33]。除了以上含有多 Cas 效应蛋白的 CASTs 外, 含有单个 Cas 效应蛋白的 II 类 CRISPR 系统也劫持了 Tn7 转座子。V-K 型 CASTs 编码单 Cas 效应元件 Cas12k, 这是一种 TnpB 同源物, 结构预测表明这是一种失活的核酸酶, 也不具备参与干扰的各种功能^[30], 它与 I 型 CASTs 最主要的区别是缺乏 TnsA, 只能在供体片段的 3' 端发生切割, 从而介导复制型转座而非 I 型 CASTs 中的剪切型转座^[40]。目前已经表征的 V-K 型 CASTs 也有 3 种^[27, 41]。

Rybarski 等^[33]对宏基因组数据库进行生物信息学分析, 发现了 1 476 个高置信度的 CASTs, 包括含有融合 *tniQ-cas8/5*、分裂 *cas7* 和分裂 *cas5* 基因的 I-F 型系统, 含有介导归巢过程的 crRNA 的 I-B 型系统以及 Cas12 相关的重组促进核酸酶/转座酶 (recombination-promoting nuclease/transposase, Rpn) 家族转座酶, 这些发现提示了 CRISPR 相关转座酶的巨大多样性仍未被完全认识。随着演化的推进以及宏基因组数据库和生物信息学检测管道的不断改进, 会有更加多样的系统不断被发现, 这些新型的系统或许能为开发新的合成生物学使能工具提供思路。

3 CASTs 介导转座的机制模型

在所有类型的 CASTs 中, 由于组成简单, V-K 型 CASTs 是研究结构和机制的良好模型, 其介导转座机制已经得到较多解析。在 V-K 型 CASTs 中, 转座酶 TnsB 聚合成一个 C2 四聚体, 通过非特异性和序列特异性相互作用识别同源供体 DNA 末端序列, 形成配对供体复合物^[42]。单 Cas 效应蛋白 Cas12k 在向导 RNA 的引导下被定位至靶标位点, 靶标处 DNA 发生扭曲形成一个 R 环, 具有与 Tn7 中 TnsD 相同结构域的锌指蛋白 TniQ 与向导 RNA/Cas 效应

蛋白直接/间接相互作用并将靶位点信息传递给 ATP 水解蛋白 TnsC, TnsC 在 DNA 上随机聚合成螺旋细丝, 并不断延伸直到遇到 TniQ 时停止延伸, 招募前述配对供体复合物, 供体复合物中的 TnsB 触发 TnsC 的 ATP 酶活性, 在 ATP 存在下促进 TnsC 螺旋解聚至仅剩 2 圈螺旋, 并将供体 DNA 插入至靶位点下游, 在这一过程中, TnsC 和 TniQ、TnsB 的相互作用导致 TnsC 螺旋向内凹陷, 这种不对称的螺旋构象使得 TnsC 无法被 TnsB 完全解聚, 剩余的 TnsC 螺旋数决定了插入位点与靶向位点的间距及其浮动范围 (图 3)^[43-45]。

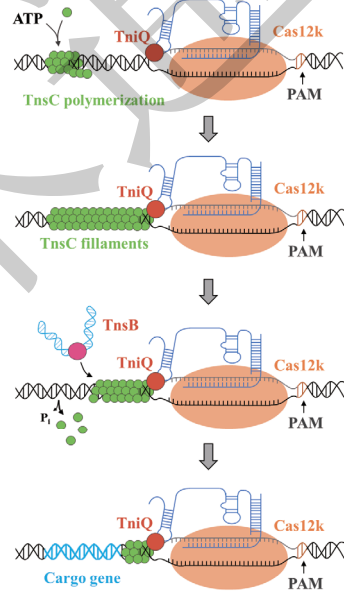


图 3 V-K 型 CASTs 介导转座机制模型

Fig. 3 Mechanistic model of V-K CASTs-mediated transposition

有趣的是, Schmitz 等^[46]在解析 ShCAST 的结构时意外发现了与 TniQ 一起纯化出来的大肠杆菌核糖体蛋白 S15 能够提升转座效率, 晶体学结构揭示 S15 和 Cas12k、sgRNA 的 tracrRNA 部分存在相互作用。大肠杆菌的 S15 蛋白通过静电相互作用和 π - π 堆积作用与 tracrRNA 的顶环产生形状互补的相互作用, 暗示 S15 能够通过稳定 tracrRNA 的顶环促进 TniQ 的招募, 这一发现为提升 CASTs 的转座活性提供了一个方向。

相比之下, I-F 型 CASTs 的机制解析较少, 目

前仅对 TniQ-Cascade 复合物的结构获得了较多的认识,而对于 CASTs 如何招募转座元件 TnsA、TnsB、TnsC 仍知之甚少。在 TniQ-Cascade 复合物中, Cas6 和 Cas8/5 分别与 crRNA 的 3' 茎环和 5' 尾部结合, Cas8/5、Cas7、Cas6 以 1:6:1 的比例组装成 Cascade 级联复合物, TniQ 以二聚体的形式,通过与 Cas6 和 Cas7.1 的相互作用被招募并锚定在级联上^[47-50]。最近的一篇研究发现,在 I-F 型 CASTs 形成转座复合物的过程中, TniQ-Cascade 复合物的 DNA 靶向是高度混杂的,能够结合广泛的脱靶位点,但所有靶向位点中只有其中的一部分能够招募 TnsC 和 TnsB。研究人员猜测在脱靶位点中, PAM 远端的错配使级联复合物的构象和 / 或 R 环结构不利于 TnsC 结合。这种 TnsC 的选择性招募增强了 CASTs 的保真性,是转座的关键检查点,但在其他 Tn 家族转座子中是否有类似的校对机制仍然未知^[51]。与 V-K 型 CASTs 不同的是, I-F 型 VchCAST 中的 TnsC 并未形成螺旋细丝,而是以单 / 双 / 三个堆叠七聚体的形态存在,说明了两种亚型 CASTs 之间转座机制的差异^[43, 51]。

此外,先前的研究表明, V-K 型 ShCAST 和 I-F 型 VchCAST 与 Tn7 转座子一样存在靶点免疫性^[27, 41],然而, Rybarski 等^[33]在分析 V 型系统基因座时发现了多个自插入和基因丢失, Yang 等^[52]通过单菌株鉴定和高通量测序证实了 I-F 型 VchCAST 系统可以在一个位点插入多个供体拷贝,这些研究结果表明 CASTs 的靶点免疫机制尚未被完全了解。

4 CASTs 的细菌基因组编辑应用

基于 CRISPR-Cas 切割 DNA 双链的基因组编辑技术,受同源重组效率的限制难以一次编辑 3 个以上靶点^[53-54]。另外也有许多研究尝试将转座子与特异性重组酶或 dCas9 进行融合,例如 mariner 转座子 Hsmar1^[55-56]、PiggyBac^[57] 和睡美人转座子^[58],以此来增强转座子系统的靶向特异性,但随机整合事件的发生率成为限制其发展的主要障碍^[57]。

CASTs 作为一种转座酶的天然变体,在被开发为基因整合工具上具有与生俱来的优势。来自霍乱弧菌的 I-F 亚型 CASTs 最早被开发为基因整合工具,

并被部署于大肠杆菌中。研究人员进一步将所有元件构建在一个 All-in-one 质粒中,得到了一个优化系统,能够在细菌中以接近 100% 的效率实现 10 k 碱基的高精度和无标记 DNA 整合^[28, 41]。V-K 和 I-B 亚型也相继被表征并开发为基因整合工具(表 1)。相比较而言, V-K 型 CASTs 比 I 型 CASTs 具有更明显的技术优势,因为它们的编码尺寸更小,具有更少的组件且单向插入^[27-28]。然而, V-K 型 CRISPR 相关转座酶系统由于缺乏 TnsA,其转座受到复制型转座机制的限制,导致生成不需要的供体质粒共整合产物。将归巢内切酶与 TnsB 融合并对供体质粒进行修饰而开发的 HELIX 技术能够使 V-K 型 CASTs 实现类似 I 型 CASTs 的剪切型转座机制^[59]。

目前, CASTs 已被部署至大肠杆菌 (*Escherichia coli*)^[28]、柠檬塔特姆氏菌 (*Tatumella citrea*)^[60]、恶臭假单胞菌 (*Pseudomonas putida*)^[41]、产酸克雷伯氏菌 (*Klebsiella oxytoca*)^[41]、希瓦氏菌 (*Shewanella oneidensis* MR-1)^[61]、苜蓿中华根瘤菌 (*Sinorhizobium meliloti*)^[62]、伯克霍尔德菌 (*Burkholderia thailandensis*)^[63]、法布氏农杆菌 (*Agrobacterium fabrum*)^[63]、蓝藻 (*Anabaena*)^[64] 中用于基因整合(表 1)。Rubin 等^[65]将基于 CASTs 的 DART (DNA-editing all-in-one RNA-guided CRISPR-Cas transposase) 系统与具有评估微生物群落中单个物种获取和整合外源 DNA 能力的 ET-Seq (environmental transformation sequencing) 结合,能够对微生物群体中的单个物种进行特定定位点的基因组编辑,研究利用该技术在合成土壤群落中鉴定出了密歇根克雷伯氏菌 (*Klebsiella michiganensis*)、猴假单胞菌 (*Pseudomonas simiae*)、罗尔斯通菌 (*Ralstonia* sp. UNC404CL21Col) 3 种遗传可及的物种。这一技术对于微生物群体中未培养微生物的定点编辑具有变革性意义。

由于不依赖于同源重组, CASTs 在大片段多重基因整合上有巨大的潜力。2020 年, Zhang 等^[60]开发了基于 CRISPR 转座酶的多拷贝基因组整合工具 MUCICAT (multicopy chromosomal integration using CRISPR-associated transposases)。其原理是使用单 crRNA 靶向基因组上重复序列,实现多拷贝位点的整合,或使用 crRNA 阵列,实现同时靶向不同位点

Table 1 Implementation of different subtype CASTs in bacteria

Note: * The transposition efficiency is not determined. ** Resistance screening is used to test transposition efficiency, thus the measured efficiency is actually on/off target ratio. *** The transposition efficiency is measured with a typical crRNA containing a typical repeats

进行基因整合（图 4-A），使用该系统生成酶基因拷贝文库，能够确定酶拷贝数的最优剂量（图 4-B）。研究人员使用该系统生成葡萄糖脱氢酶基因的拷贝文库，最终获得了比 pET-24a 表达量增加 2.6 倍的菌株^[60]。进一步使用该系统对菌株进行代谢工程改造，经过两轮整合，在多重目标位点引入终止子从而敲除了 N-乙酰葡萄糖胺降解基因簇，并对 N-乙酰葡萄糖胺合成基因簇中 *mglnS* 和 *ScGNAI* 基因的拷贝数进行了优化，整个过程仅花费 8 d^[66]。2021 年，Yang

等^[34]挖掘出来自半透明假交替单胞菌（*Pseudoalteromonas translucida* KMM520）的高活性 I-F 型 Ptr-CAST，与 VchCAST 组成正交 CASTs，分别能够以约 100% 的效率将相应的货物基因插入各自的 2 个和 5 个位点，而不会相互干扰。使用正交 CASTs，将基因靶向整合至代谢通路中竞争途径基因和其他位点，能够一步实现多 * 多拷贝整合 * 多基因中断（图 4-C），这些工具的开发大大加速了代谢工程菌株的构建进程。

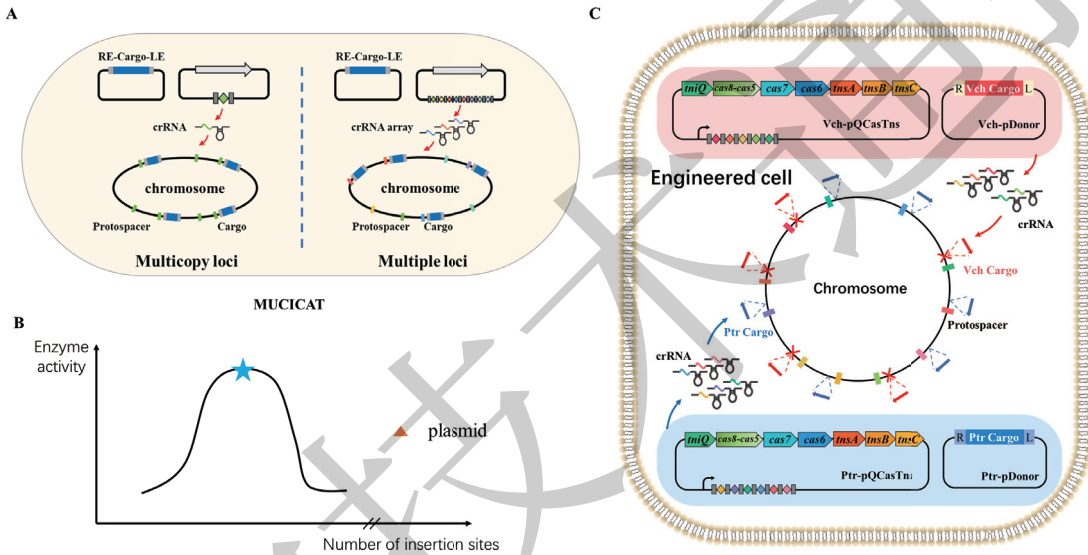


图 4 MUCICAT 的原理（A）及其在酶剂量优化（B）和代谢工程菌株改造（C）上的应用

Fig. 4 Mechanism of MUCICAT (A) and its application of enzyme dosage optimization (B) and modification (C) of metabolically engineered strain

5 总结与展望

近年来，多种 CRISPR 相关转座酶被相继挖掘，与现有基因编辑工具相比，由于不依赖于宿主 DNA 双链修复，CASTs 能够实现基因的多靶点定点插入^[60]和两正交多重编辑^[34]，是多靶整合的有力工具，并且能够高效插入大片段 DNA^[41]，为大片段基因整合工具提供了一个新的选择，是一种十分有潜力的基因编辑工具。目前，CASTs 已经被部署至多种革兰氏阴性菌和厌氧菌中，而 CASTs 能否在革兰氏阳性菌中发挥转座活性，目前尚无研究报道。尽管基于 CASTs 的基因整合工具已经能够实现大片段和多重基因整合和 / 或敲除，但其仍然存在一些问题，正交 CASTs 中两个 CAST 之间重复序列的相

似性可能导致交叉反应性^[34]以及 CASTs 系统存在类似于 CRISPR-Cas 系统中的脱靶效应，这些问题的存在提出了对 CASTs 系统进一步深入研究和优化的要求。未来，挖掘新型 CASTs，并对其进行理性或非理性改造，或许能将 CASTs 拓展至更多的细菌及真核细胞中，进一步发挥其大片段基因整合与多重编辑特性。

参考文献

[1] Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR-Cas systems [J]. Nat Rev Microbiol, 2015, 13 (11) : 722-736.

[2] Yosef I, Goren MG, Qimron U. Proteins and DNA elements essential

- for the CRISPR adaptation process in *Escherichia coli* [J]. *Nucleic Acids Res*, 2012, 40 (12) : 5569-5576.
- [3] Carte J, Wang RY, Li H, et al. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes [J]. *Genes Dev*, 2008, 22 (24) : 3489-3496.
- [4] Swarts DC, Mosterd C, van Passel MWJ, et al. CRISPR interference directs strand specific spacer acquisition [J]. *PLoS One*, 2012, 7 (4) : e35888.
- [5] Abudayyeh OO, Gootenberg JS, Konermann S, et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector [J]. *Science*, 2016, 353 (6299) : aaf5573.
- [6] Mali P, Yang LH, Esvelt KM, et al. RNA-guided human genome engineering via Cas9 [J]. *Science*, 2013, 339 (6121) : 823-826.
- [7] Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity [J]. *Science*, 2012, 337 (6096) : 816-821.
- [8] Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system [J]. *Nat Protoc*, 2013, 8 (11) : 2281-2308.
- [9] Konermann S, Brigham MD, Trevino AE, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex [J]. *Nature*, 2015, 517 (7536) : 583-588.
- [10] Qi LS, Larson MH, Gilbert LA, et al. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression [J]. *Cell*, 2013, 152 (5) : 1173-1183.
- [11] Komor AC, Kim YB, Packer MS, et al. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage [J]. *Nature*, 2016, 533 (7603) : 420-424.
- [12] Cox DBT, Gootenberg JS, Abudayyeh OO, et al. RNA editing with CRISPR-cas13 [J]. *Science*, 2017, 358 (6366) : 1019-1027.
- [13] East-Seletsky A, O'Connell MR, Knight SC, et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection [J]. *Nature*, 2016, 538 (7624) : 270-273.
- [14] Chen JS, Ma EB, Harrington LB, et al. CRISPR-Cas12a target binding unleashes indiscriminate single-stranded DNase activity [J]. *Science*, 2018, 360 (6387) : 436-439.
- [15] Luo ML, Leenay RT, Beisel CL. Current and future prospects for CRISPR-based tools in bacteria [J]. *Biotechnol Bioeng*, 2016, 113 (5) : 930-943.
- [16] Hickman AB, Dyda F. DNA transposition at work [J]. *Chem Rev*, 2016, 116 (20) : 12758-12784.
- [17] Ravindran S. Barbara McClintock and the discovery of jumping genes [J]. *Proc Natl Acad Sci USA*, 2012, 109 (50) : 20198-20199.
- [18] Fedoroff NV. The discovery and characterization of transposable elements. The collected papers of Barbara McClintock [J]. *Cell*, 1988, 53 (1) : 9-10.
- [19] Kapitonov VV, Jurka J. Rolling-circle transposons in eukaryotes [J]. *Proc Natl Acad Sci USA*, 2001, 98 (15) : 8714-8719.
- [20] Kapitonov VV, Jurka J. Self-synthesizing DNA transposons in eukaryotes [J]. *Proc Natl Acad Sci USA*, 2006, 103 (12) : 4540-4545.
- [21] Craig NL. Tn7: a target site-specific transposon [J]. *Mol Microbiol*, 1991, 5 (11) : 2569-2573.
- [22] Koch B, Jensen LE, Nybroe O. A panel of Tn7-based vectors for insertion of the gfp marker gene or for delivery of cloned DNA into Gram-negative bacteria at a neutral chromosomal site [J]. *J Microbiol Methods*, 2001, 45 (3) : 187-195.
- [23] Peters JE, Craig NL. Tn7: smarter than we thought [J]. *Nat Rev Mol Cell Biol*, 2001, 2 (11) : 806-814.
- [24] Kaczmarek Z, Czarnocki-Cieciura M, Górecka-Minakowska KM, et al. Structural basis of transposon end recognition explains central features of Tn7 transposition systems [J]. *Mol Cell*, 2022, 82(14) : 2618-2632.e7.
- [25] Morero NR, Zuliani C, Kumar B, et al. Targeting IS608 transposon integration to highly specific sequences by structure-based transposon engineering [J]. *Nucleic Acids Res*, 2018, 46 (8) : 4152-4163.
- [26] Peters JE, Makarova KS, Shmakov S, et al. Recruitment of CRISPR-Cas systems by Tn7-like transposons [J]. *Proc Natl Acad Sci USA*, 2017, 114 (35) : E7358-E7366.
- [27] Strecker J, Ladha A, Gardner Z, et al. RNA-guided DNA insertion with CRISPR-associated transposases [J]. *Science*, 2019, 365 (6448) : 48-53.
- [28] Klompe SE, Vo PLH, Halpin-Healy TS, et al. Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration [J]. *Nature*, 2019, 571 (7764) : 219-225.
- [29] Ma W, Xu YS, Sun XM, et al. Transposon-associated CRISPR-cas system: a powerful DNA insertion tool [J]. *Trends Microbiol*, 2021, 29 (7) : 565-568.

- [30] Faure G, Shmakov SA, Yan WX, et al. CRISPR-Cas in mobile genetic elements: counter-defence and beyond [J] . Nat Rev Microbiol, 2019, 17 (8) : 513-525.
- [31] Gleditsch D, Müller-Esparza H, Pausch P, et al. Modulating the Cascade architecture of a minimal Type I-F CRISPR-Cas system [J] . Nucleic Acids Res, 2016, 44 (12) : 5872-5882.
- [32] Krupovic M, Makarova KS, Forterre P, et al. Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity [J] . BMC Biol, 2014, 12: 36.
- [33] Rybarski JR, Hu K, Hill AM, et al. Metagenomic discovery of CRISPR-associated transposons [J] . Proc Natl Acad Sci USA, 2021, 118 (49) : e2112279118.
- [34] Yang SQ, Zhang YW, Xu JQ, et al. Orthogonal CRISPR-associated transposases for parallel and multiplexed chromosomal integration [J] . Nucleic Acids Res, 2021, 49 (17) : 10192-10202.
- [35] Petassi MT, Hsieh SC, Peters JE. Guide RNA categorization enables target site choice in Tn7-CRISPR-cas transposons [J] . Cell, 2020, 183 (7) : 1757-1771.e18.
- [36] Klompe SE, Jaber N, Beh LY, et al. Evolutionary and mechanistic diversity of type I-F CRISPR-associated transposons [J] . Mol Cell, 2022, 82 (3) : 616-628.e5.
- [37] Saito M, Ladha A, Strecker J, et al. Dual modes of CRISPR-associated transposon homing [J] . Cell, 2021, 184 (9) : 2441-2453.e18.
- [38] Wimmer F, Mougiakos I, Englert F, et al. Rapid cell-free characterization of multi-subunit CRISPR effectors and transposons [J] . Mol Cell, 2022, 82 (6) : 1210-1224.e6.
- [39] Özcan A, Pausch P, Linden A, et al. Type IV CRISPR RNA processing and effector complex formation in *Aromatoleum aromaticum* [J] . Nat Microbiol, 2019, 4 (1) : 89-96.
- [40] Vo PLH, Acree C, Smith ML, et al. Unbiased profiling of CRISPR RNA-guided transposition products by long-read sequencing [J] . Mob DNA, 2021, 12 (1) : 13.
- [41] Vo PLH, Ronda C, Klompe SE, et al. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering [J] . Nat Biotechnol, 2021, 39 (4) : 480-489.
- [42] Park JU, Tsai AWL, Chen TH, et al. Mechanistic details of CRISPR-associated transposon recruitment and integration revealed by cryo-EM [J] . Proc Natl Acad Sci USA, 2022, 119 (32) : e2202590119.
- [43] Park JU, Tsai AWL, Mehrotra E, et al. Structural basis for target site selection in RNA-guided DNA transposition systems [J] . Science, 2021, 373 (6556) : 768-774.
- [44] Querques I, Schmitz M, Oberli S, et al. Target site selection and remodelling by type V CRISPR-transposon systems [J] . Nature, 2021, 599 (7885) : 497-502.
- [45] Park JU, Tsai AWL, Rizo AN, et al. Structures of the holo CRISPR RNA-guided transposon integration complex [J] . bioRxiv, 2022, DOI:10.1101/2022.10.12.511933.
- [46] Schmitz M, Querques I, Oberli S, et al. Structural basis for RNA-mediated assembly of type V CRISPR-associated transposons [J] . bioRxiv, 2022. DOI:10.1101/2022.06.17.496590.
- [47] Jia N, Xie W, de la Cruz MJ, et al. Structure-function insights into the initial step of DNA integration by a CRISPR-Cas-Transposon complex [J] . Cell Res, 2020, 30 (2) : 182-184.
- [48] Wang BB, Xu WH, Yang H. Structural basis of a Tn7-like transposase recruitment and DNA loading to CRISPR-Cas surveillance complex [J] . Cell Res, 2020, 30 (2) : 185-187.
- [49] Li Z, Zhang H, Xiao RJ, et al. Cryo-EM structure of a type I-F CRISPR RNA guided surveillance complex bound to transposition protein TniQ [J] . Cell Res, 2020, 30 (2) : 179-181.
- [50] Halpin-Healy TS, Klompe SE, Sternberg SH, et al. Structural basis of DNA targeting by a transposon-encoded CRISPR-Cas system [J] . Nature, 2020, 577 (7789) : 271-274.
- [51] Hoffmann FT, Kim M, Beh LY, et al. Selective TnsC recruitment enhances the fidelity of RNA-guided transposition [J] . Nature, 2022, 609 (7926) : 384-393.
- [52] Yang JJ, Yang JW, Zhang YW, et al. CRISPR-associated transposase system can insert multiple copies of donor DNA into the same target locus [J] . CRISPR J, 2021, 4 (6) : 789-798.
- [53] Jiang Y, Chen B, Duan CL, et al. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system [J] . Appl Environ Microbiol, 2015, 81 (7) : 2506-2514.
- [54] Feng X, Zhao DD, Zhang XL, et al. CRISPR/Cas9 assisted multiplex genome editing technique in *Escherichia coli* [J] . Biotechnol J, 2018, 13 (9) : e1700604.
- [55] Bhatt S, Chalmers R. Targeted DNA transposition *in vitro* using a dCas9-transposase fusion protein [J] . Nucleic Acids Res, 2019, 47 (15) : 8126-8135.

- [56] Chen SP, Wang HH. An engineered cas-transposon system for programmable and site-directed DNA transpositions [J] . CRISPR J, 2019, 2 (6) : 376-394.
- [57] Wang G, Yang LH, Grishin D, et al. Efficient, footprint-free human iPSC genome editing by consolidation of Cas9/CRISPR and piggyBac technologies [J] . Nat Protoc, 2017, 12 (1) : 88-103.
- [58] Ma SF, Wang XL, Hu YF, et al. Enhancing site-specific DNA integration by a Cas9 nuclease fused with a DNA donor-binding domain [J] . Nucleic Acids Res, 2020, 48 (18) : 10590-10601.
- [59] Tou CJ, Orr B, Kleinstiver BP. Cut-and-paste DNA insertion with engineered type V-K CRISPR-associated transposases [J] . bioRxiv, 2022. DOI:10.1101/2022.01.07.475005.
- [60] Zhang YW, Sun XM, Wang QZ, et al. Multicopy chromosomal integration using CRISPR-associated transposases [J] . ACS Synth Biol, 2020, 9 (8) : 1998-2008.
- [61] Cheng ZH, Wu J, Liu JQ, et al. Repurposing CRISPR RNA-guided integrases system for one-step, efficient genomic integration of ultra-long DNA sequences [J] . Nucleic Acids Res, 2022, 50 (13) : 7739-7750.
- [62] Cui YL, Dong HN, Tong BS, et al. A versatile Cas12k-based genetic engineering toolkit (C12KGET) for metabolic engineering in genetic manipulation-deprived strains [J] . Nucleic Acids Res, 2022, 50 (15) : 8961-8973.
- [63] Rodríguez LT, Ellington AJ, Reisch CR. Broad-host-range mutagenesis with CRISPR-associated transposase [J] . bioRxiv, 2022. DOI:10.1101/2022.01.19.475551.
- [64] Arévalo S, Rico DP, Abarca MD, et al. Towards genome-engineering in complex cyanobacterial communities: RNA-guided transposition in *Anabaena* [J] . bioRxiv, 2022. DOI: 10.1101/2022.09.18.508393.
- [65] Rubin BE, Diamond S, Cress BF, et al. Species- and site-specific genome editing in complex bacterial communities [J] . Nat Microbiol, 2022, 7 (1) : 34-47.
- [66] Zhang YW, Yang JW, Yang SQ, et al. Programming cells by multicopy chromosomal integration using CRISPR-associated transposases [J] . CRISPR J, 2021, 4 (3) : 350-359.

(责任编辑 朱琳峰)