

DOI: 10.14188/j.ajsh.2021.05.002

嗜盐古菌中 I-B 亚型 CRISPR-Cas 系统的分子机制和应用进展

张佳, 吴意, 唐兵, 唐晓峰*

(武汉大学 生命科学学院, 湖北 武汉 430072)

摘要: 嗜盐古菌是一类应用前景广阔的微生物资源,但由于缺乏高效灵活的遗传操作工具,限制了这类生物在科学研究和工业生产方面的应用。CRISPR-Cas 系统存在于绝大多数古菌中,应用其内源性 CRISPR-Cas 系统进行遗传操作是一种兼具可行性和简便性的策略。近年来,人们对 I-B 亚型 CRISPR-Cas 系统的研究已从模式古菌沃氏富盐菌(*Haloferax volcanii*)拓展到其他嗜盐古菌中。本文总结了 I-B 亚型 CRISPR-Cas 系统的分子机制,并介绍了利用这种内源性 CRISPR-Cas 系统在嗜盐古菌中进行遗传操作的研究进展。

关键词: 嗜盐古菌; CRISPR-Cas 系统; I-B 亚型

中图分类号: Q93

文献标志码: A

文章编号: 2096-3491(2021)05-0435-08

Recent progress on molecular mechanisms and application of the type I-B CRISPR-Cas system in haloarchaea

ZHANG Jia, WU Yi, TANG Bing, TANG Xiaofeng*

(College of Life Sciences, Wuhan University, Wuhan 430072, Hubei, China)

Abstract: Haloarchaea are a kind of biological resources with broad application prospects. However, due to the lack of efficient and flexible genetic manipulation tools, the application of these organisms in scientific research and industrial production is restricted. CRISPR-Cas system is present in the majority of archaea. Therefore, harnessing their endogenous CRISPR-Cas machineries for genetic manipulation is a feasible and convenient strategy. In recent years, research on the type I-B CRISPR-Cas system has expanded from the type archaeon *Haloferax volcanii* to other haloarchaea. In this paper, we summarize the recent progress on the molecular mechanisms of the type I-B CRISPR-Cas system, which has been successfully used for genetic manipulation in some haloarchaea.

Key words: haloarchaeon; CRISPR-Cas system; type I-B

0 引言

嗜盐古菌是一类存在于高盐环境中的极端微生物。由于它们能够产生在高温和高离子强度下保持催化活性的酶、可降解塑料组分聚β-羟丁酸(PHB)和聚羟基脂肪酸酯(polyhydroxyalkanoates, PHA)、

类胡萝卜素等多种能被应用在工业生物技术中的产物,一直备受科学界的关注^[1]。然而,由于嗜盐古菌细胞膜和细胞壁的特性、胞内高盐环境、染色体多倍性及其特有的核酸代谢途径,嗜盐古菌中的遗传操作一直受到限制,制约了这类生物在科学研究和工业生产方面的应用^[1]。

收稿日期: 2021-03-31 修回日期: 2021-05-09 接受日期: 2021-05-10

作者简介: 张佳(1996-),女,硕士生,主要从事极端嗜盐古菌相关研究。E-mail: 2018202040054@whu.edu.cn

* 通讯联系人: 唐晓峰(1968-),女,博士,教授,主要从事极端嗜盐、嗜热古菌适应极端环境的分子机理研究。E-mail: tangxf@whu.edu.cn

基金项目: 国家自然科学基金(31970028)

引用格式: 张佳, 吴意, 唐兵, 等. 嗜盐古菌中 I-B 亚型 CRISPR-Cas 系统的分子机制和应用进展[J]. 生物资源, 2021, 43(5): 435-442.

Zhang J, Wu Y, Tang B, et al. Recent progress on molecular mechanisms and application of the type I-B CRISPR-Cas system in haloarchaea [J]. Biotic Resources, 2021, 43(5): 435-442.

外来可移动遗传元件(mobile genetic elements, MGEs)如病毒核酸的入侵通常是有害的并且可导致细胞死亡。CRISPR (clustered regularly interspaced short palindromic repeats)是原核生物中成簇存在的具有规律间隔的短回文重复序列,Cas 蛋白(CRISPR-associated protein)能在 CRISPR RNA (crRNA)的引导下结合并切割 MGEs,二者构成了原核生物中 RNA 指导的获得性免疫系统,其中起指导作用的 crRNA 又称引导 RNA (guide RNA, gRNA)^[2]。基因组已测序的所有古菌中,约有 90% 的古菌含有 CRISPR/Cas 系统,但只有少数古菌中的 CRISPR-Cas 系统得到了充分研究^[3]。

1 嗜盐古菌简介

1.1 嗜盐古菌的分类学地位

Woese 等^[4]提出,地球上所有细胞形式的生命可以分为三个域——细菌域(Bacteria)、古菌域(Archaea)和真核生物域(Eukarya)。由于古菌经常生活在高温、高盐或厌氧等极端环境中,人们根据它们不同的生存环境和生理特征,将古菌分为极端嗜盐古菌、极端嗜热古菌和厌氧产甲烷古菌等生理类群。嗜盐古菌属于古菌域中的广古菌门(Euryarchaeota)盐古菌纲(Halobacteria)盐古杆菌目(Halobacteriales)盐古杆菌科(Halobacteriaceae)^[5]。

1.2 嗜盐古菌的生理特性

嗜盐古菌是好氧异养型微生物,大多数嗜盐古菌需要在 1.7~2.5 mol/L NaCl 浓度下生长和维持细胞结构稳定^[6]。在嗜盐古菌的细胞壁中,富含酸性氨基酸(如天冬氨酸和谷氨酸)的糖蛋白的羧基和细胞壁中的硫酸基团都能与胞外高浓度的钠离子结合,帮助维持细胞壁稳定。嗜盐古菌主要通过胞内积累与胞外 NaCl 浓度相当的 KCl,以保持胞内外渗透压平衡。

1.3 嗜盐古菌的遗传学特征

目前人们研究过的嗜盐古菌如盐沼盐古杆菌(*Halobacterium salinarum*)、地中海富盐菌(*Haloferax mediterranei*)、沃氏富盐菌(*Haloferax volcanii*)和西班牙盐盒菌(*Haloarcula hispanica*)都是多倍体^[7]。嗜盐古菌中染色体拷贝数受到生长周期和环境因素的调节,对数生长期的染色体拷贝数高于稳定期^[8]。多倍性对于嗜盐古菌生物进化方面的优势已被实验证明,包括同源重组带来的低突变率、对 X 射线辐射和干燥的高耐受性。沃氏富盐菌能够在不施加选择压力的情况下发生基因转变(gene conversion),使每个基因组拷贝保持一致性。多倍性的另

一个优势是使得嗜盐古菌在岩盐沉积物和类似的极端干燥环境中生存下来。还有研究发现,沃氏富盐菌将基因组 DNA 作为遗传物质和磷酸盐多聚物的储存形式。磷酸盐的稳定储存被认为是生物进化早期 DNA 出现的一种驱动力^[8]。

2 CRISPR-Cas 系统简介

2.1 CRISPR-Cas 系统的组成

已测序的大多数古菌(约 90%)和部分细菌(约 40%)的基因组中至少存在一个 CRISPR 基因座,詹氏甲烷热球菌(*Methanocaldococcus jannaschii*)中存在 18 个 CRISPR 基因座,是目前报道的拥有最多 CRISPR 基因座的原核生物^[9]。本实验室在鱼酱碱线菌(*Natrinema gari*)J7-2 的基因组中也发现了 3 个 CRISPR 基因座^[10]。每个 CRISPR 基因座包含一系列被可变序列间隔开的重复序列(repeat),每个可变序列对应入侵的 MGEs 上的一个片段。这些可变序列,称为间隔序列(spacer),形成了对继往感染的遗传记录,为细胞在同样的 MGEs 再次入侵时提供保护。CRISPR 序列附近通常有多个保守的蛋白质编码基因,在 CRISPR 介导的免疫过程的不同阶段发挥作用,命名时通常冠以“Cas”的前缀^[2]。

2.2 CRISPR-Cas 系统的分类

在 CRISPR-Cas 系统中, RNA 指导的 DNA 或 RNA 切割是通过效应复合体实现的,后者由一段引导 RNA (CRISPR RNA, crRNA) 和一系列 Cas 蛋白(Class 1, 1 类)或具有多结构域的单 Cas 蛋白(Class 2, 2 类)组成。在这两大类系统中,根据切割靶核酸的特征性蛋白,可将其各自分成以下类型:1 类 CRISPR-Cas 系统包括 I 型、III 型和 IV 型,2 类 CRISPR-Cas 系统包含 II 型、V 型和 VI 型。根据不同类型中 CRISPR-Cas 位点的组成,又可将其分为几种亚型。CRISPR-Cas 系统的最新分类包括两大类、6 种类型和 33 种亚型^[11]。I 型 CRISPR-Cas 系统包含最多的亚型(I-A~I-G),是在细菌和古菌中分布最广的获得性免疫系统。其中, I-B 亚型 CRISPR-Cas 系统占的比例最高。在 I 型 CRISPR-Cas 系统中,只有 I-B 亚型保留了原始的 cas 基因组成(*cas1-cas2-cas3-cas4-cas5-cas6-cas7-cas8*),其他亚型则发生了部分基因丢失和重排。根据 Cas8b 蛋白家族的种类,可以将 I-B 亚型分为三个主要分支: Cas8b1 (Hmari)、Cas8b2 (Tneap) 和 Cas8b3 (Myxan)。在 I-C~I-G 亚型中, cas 基因以单个操纵子的形式表达;而在 I-A 和 I-B 亚型中,不同的 cas 基因组成两个或两个以上的操纵子,提示其表达调控

将更为精密。根据2015年统计的结果,在含有CRISPR-Cas系统的微生物中,约有30%的古菌和15%的细菌含I-B亚型CRISPR-Cas系统^[12]。目前,全基因组经过测序的嗜盐古菌大多含I-B亚型CRISPR-Cas系统。

2.3 CRISPR-Cas系统的工作原理

CRISPR-Cas免疫系统分三个阶段起作用:适应(adaptation)、表达(expression)和干扰(interference)^[2](图1)。(1)在适应阶段,Cas整合酶复合体将一段MGEs来源的核酸片段作为新的spacer整合到CRISPR阵列中,在这个过程中repeat序列也会发生复制。(2)在表达阶段,由前导序列(leader)中的启动子元件起始CRISPR基因座的转录,产生的前体转录本(pre-crRNA)在repeat区域发生切割,形成加工后的CRISPR RNAs(crRNAs),每个成熟的crRNA包含一个完整的spacer和部分repeat序列。成熟的crRNA与Cas蛋白结合,形成效应复合体(effector complex)。(3)在干扰阶段,由效应复合体中的crRNA通过碱基互补配对识别再次入侵的MGE上的一个特定位点,触发Cas酶对外源核酸的靶向切割。

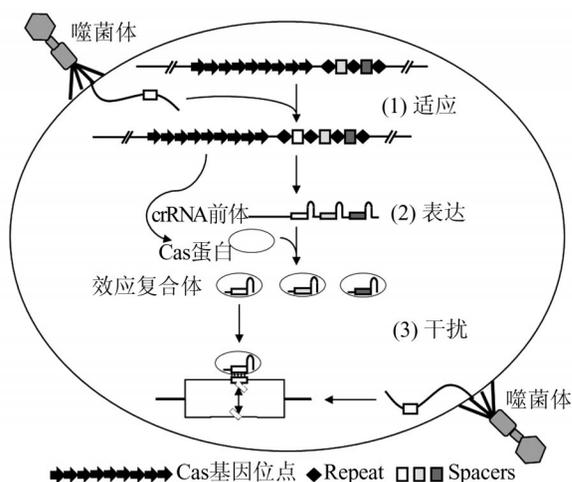


图1 CRISPR-Cas系统工作原理

Fig. 1 Mechanism of CRISPR-Cas system

I型中的效应复合体称为Cascade(CRISPR-associated complex for antiviral defense),由crRNA和多个Cas蛋白组成^[13]。Cascade定位到靶DNA上后,就会招募I型系统的特征性核酸酶Cas3进行DNA切割^[14]。Cas3蛋白的N端含有HD型磷酸水解酶/核酸酶结构域,C端延伸区中包含保守的SF2解旋酶结构域。体外实验证明Cas3具有单链DNA刺激的ATP酶活性、核酸酶活性和DNA解旋酶活性(方向为3'-5')。研究者们推测Cas3对双链DNA

分子的切割是通过先切割第一链,再切割游离出来的第二链的方式进行的^[14]。

通过储存入侵者的序列来编码产生特定的gRNA,CRISPR-Cas系统面临一个潜在的致命因素:CRISPR阵列中编码gRNA的DNA(即间隔序列),与gRNA间也存在广泛的互补配对。如果没有另外的识别机制,每个间隔序列也有可能被判定为先前的入侵者,导致基因组被攻击。此外,如果CRISPR系统错误地获取了对宿主DNA的免疫记忆,也会伴随着自身免疫的风险。因此,在编码CRISPR-Cas系统的微生物中,无论是适应阶段还是干扰阶段,都应有特定的识别机制,以区分“自我”和“非我”。

研究者将天然CRISPR阵列中的间隔序列在数据库中比对后发现,同一生物获取的不同间隔序列的原间隔序列(proto-spacer)附近存在一个相似度极高的区域(一般是2~5个核苷酸),命名为原间隔序列邻近基序(proto-spacer-adjacent motifs, PAMs)^[15]。效应复合体在检查gRNA与靶分子的互补性之前,首先扫描DNA上是否存在PAM。能与gRNA完全互补配对但缺少PAM的序列不会被核酸酶攻击。CRISPR阵列中间隔序列两侧的重复序列上缺乏可被效应复合体识别的PAM,从而有效避免自身免疫。因此,PAM是CRISPR-Cas系统靶向作用发生的“闸门”。几乎所有CRISPR系统中核酸酶功能的激活都需要某种形式的PAM。然而,不同的Cas核酸酶识别的PAM在序列、长度、复杂性、方向性、与靶点的距离等方面差别很大。而且,有研究表明,适应阶段和干扰阶段识别的PAM并不相同^[16]。

3 I-B亚型CRISPR-Cas系统相关机制研究

近年来,研究者围绕嗜盐古菌中I-B亚型CRISPR-Cas系统的工作原理,对适应、表达和干扰三个不同阶段的相关分子机制进行了探索,为后续应用奠定了理论基础。

3.1 干扰作用发生的必需条件

干扰作用,即CRISPR系统对核苷酸序列进行特异性识别和切割,是与CRISPR-Cas技术应用直接相关的一环。Fischer等^[17]发现*Natrinema gari* J7-2中的I-B亚型CRISPR-Cas系统有7个不同的PAM,当原间隔序列上游存在这样的PAM时才能触发靶向DNA切割。这是已研究的CRISPR-Cas系统中可识别PAM最多的系统。

种子序列(seed sequence)是指原间隔序列上PAM近端的序列,通常为7~12 bp,该区域内间隔

序列与原间隔序列间的碱基必须严格配对,才能触发干扰作用。Maier 等^[18]的研究表明,沃氏富盐菌中能被靶向切割的 DNA 片段需与间隔序列 5'端的前 10 个核苷酸匹配(允许在第 6 位发生错配)。

除了对原间隔序列上 PAM 和种子序列的要求,干扰作用的实现还依赖于 crRNA 的组成。嗜盐古菌中每个成熟的 crRNA 包含中间的间隔序列和两端被切割后的重复序列,分别称为 5'柄(handle)和 3'柄(handle)^[18]。研究者探究了沃氏富盐菌中不同 crRNA 变异体介导干扰反应的活性^[19],结果表明,3'柄存在缺失或突变的 crRNA 依然能够触发干扰反应,而对 5'柄进行改造则会使 crRNA 完全丧失介导干扰的能力。还有研究发现,在西班牙盐盒菌中,当间隔序列长度大于 30 bp 时,能观察到有效的干扰作用;而当间隔序列长度缩短到 30 bp 以下时,干扰作用显著减弱直至完全丧失^[20]。

在 I 型 CRISPR-Cas 系统的干扰阶段,Cascade 复合体中的 crRNA 与靶 DNA 发生碱基互补配对时形成 R 环(R-loop),进而对外源 DNA 上的靶位点进行识别。Cascade 对靶序列的识别是以方向性地(从 PAM 序列起始)形成 R 环拉链(R-loop zipping)的方式进行的^[21]。一旦遇到错配,R 环延伸终止,并以长度依赖的方式瓦解。R 环如果一直延伸到原间隔序列的末端,则会被锁定,触发 Cas3 核酸酶/解旋酶介导的 DNA 降解,而不再进行校正阅读。

综上所述,在嗜盐古菌的 CRISPR-Cas 系统中,触发成功的干扰作用需要具备如下几个要素:① 可移动遗传元件上含有能被效应复合体识别的 PAM 和种子序列;② 包含 5'柄、3'柄(部分或完全)和一定长度间隔序列的 crRNA;③ 效应复合体与原间隔序列间形成稳定的 R 环。

3.2 crRNA 的生物合成机制

在 I-B 亚型 CRISPR-Cas 系统中,由 Cas6 蛋白将 pre-crRNA 加工成单个 crRNA 分子。研究者确定了地中海富盐菌中 Cas6 的切割位点和 CRISPR 序列的转录起始点,发现 Cas6 对 pre-crRNA 的加工能力从对数早期到稳定期逐步提高,成熟 crRNA 的形成主要是在稳定期^[22]。除了 Cas6 外,其他 Cas 蛋白也会影响 crRNA 的产生和稳定:缺失 Cas1、Cas3 或 Cas4 蛋白会导致 CRISPR 前体转录本减少,进而降低成熟 crRNA 的含量;而 Cas5 和 Cas7 蛋白对于稳定成熟的 crRNA 至关重要。

根据现有研究,在嗜盐古菌的 I-B 亚型 CRISPR-Cas 系统中,CRISPR 前体转录本(pre-crRNA)是组成型表达,Cas1、Cas3 和 Cas4 蛋白与其在细胞

中稳定存在相关;而 Cas6 对 pre-crRNA 的加工呈现出生长周期依赖性,其内在的调控机制不明;加工成熟的 crRNA 需与 Cas5 和 Cas7 蛋白结合,避免被核酸酶降解。

3.3 适应发生的条件和过程

CRISPR-Cas 系统通过获取外源 DNA 片段作为新的间隔序列,建立对入侵者的适应,当外来核酸再次入侵时,通过序列配对介导靶向干扰。作为 CRISPR-Cas 系统发挥功能的基础,该系统如何获取新的间隔序列及 CRISPR 阵列的起源,一直是困扰科学家的难题。

少数 CRISPR-Cas 系统只需要 Cas1 和 Cas2 蛋白完成适应过程,但大部分系统发生适应还需要 Cas4 蛋白作为辅助因子,Cas4 蛋白在这一过程中所起的作用不明^[23]。最新研究显示,I 型系统的适应复合体是通过分阶段组装的方式发挥作用的^[24]。在第一阶段,Cas4 与 Cas1 组装形成稳定的复合体,对 protospacer 进行加工;在第二阶段,由于 Cas1-protospacer 复合体与 Cas2 结合的亲和力更强,Cas2 通过竞争性结合取代 Cas4,形成 Cas1-Cas2-protospacer 三元复合体,完成后续整合过程。

目前,在很多实验室菌株中难以观察到 CRISPR-Cas 系统对纯化病毒的有效适应现象,限制了人们对该过程发生机制的研究。西班牙盐盒菌中 I-B 亚型 CRISPR 系统是继嗜热链球菌(*Streptococcus thermophilus*)的 II-A 型系统之后第二个表现出对纯化病毒高效适应的天然系统^[25]。被病毒 HHPV-2 感染后,西班牙盐盒菌能够从病毒基因组上有选择地获取间隔序列。在 I-B 亚型 CRISPR 系统中,适应的发生除了需要 Cas1、Cas2 和 Cas4 蛋白,还需要 Cas3 和至少部分 Cascade 中的蛋白,且要求宿主的 CRISPR 阵列中存在一个能与病毒序列部分或完全匹配的间隔序列。同时,为了避免获取自身的 DNA 片段,嗜盐古菌中还存在着相应的识别机制。在完全匹配的目标序列上游,还需要一个功能性 PAM 才能启动适应过程,该机制称为引发适应(priming adaptation)^[26]。

引发适应是间隔序列获取的主要途径,至少是对于分布最广的 I 型系统而言。引发适应是由一段与外源 DNA 完全或部分配对的 crRNA 指导的,研究者对于该过程中起指导作用的 crRNA(引发 crRNA)的可塑性进行了研究^[20]。在西班牙盐盒菌中,3'柄完全去除并不影响引发适应,但严重削弱了 crRNA 的稳定性和干扰作用;而 5'柄上邻近 spacer 的 6 个核苷酸是引发适应必需的保守序列。

在验证适应发生条件的同时,研究者还对该阶段中间隔序列整合的分子机制进行了解析。在新的 spacer 整合到 CRISPR 阵列上时,邻近前导序列 (leader) 的 repeat 被准确复制。据此,人们猜测在细胞中存在控制 repeat 长度的“分子标尺”^[3]。研究人员在西班牙盐盒菌中发现,对 30 bp 的重复序列进行精确复制时需要两个位于重复序列中间的保守基序 AACCC 和 GTGGG^[27]。AACCC 基序定位在 leader-repeat 接合位点的下游~10 bp 处,这里通常是重复序列复制起始的地方。重复序列复制终止序列通常与 GTGGG 基序相隔一定的距离,可能是作为分子标尺的锚点。改变这两个基序间的距离将使复制产生的重复序列长度发生变异。获取复合体可能是通过识别这些重复序列中间的元件,来决定间隔序列整合时复制的重复序列 DNA 的长度。

尽管已有研究证实间隔序列的大小是由适应复合体的结构决定的,天然 CRISPR 阵列中的间隔序列长度并不完全相同^[28]。研究者在西班牙盐盒菌的 I-B 亚型 CRISPR 系统中探究了这种异质性的成因^[29]。在病毒感染后,菌株获取的 37,957 个新的间隔序列间存在显著的大小差异,这种差异表现出序列依赖性。spacer 3'端(PAM 远端)的第三位核苷酸偏好胞嘧啶,突变原间隔序列上的这一位点将改变最终间隔序列的大小。另外,由 PAM 不准确识别引起的 spacer 5'端(PAM 近端)滑动也会改变间隔序列的大小。上述结果表明,PAM-protospacer 序列的两端都存在不同程度的核苷酸选择性,在一定程度上对适应复合体的结构标尺进行微调,共同决定间隔序列的大小。

到目前为止,嗜盐古菌中 I-B 亚型 CRISPR-Cas 系统发生适应的条件和过程已基本研究清楚。适应的发生需要两步组装的适应复合体(Cas1-Cas4/Cas2)、Cas3 和效应复合体 Cascade(或其中部分蛋白),在 PAM 下游能与引发 crRNA 配对的原间隔序列的作用下启动间隔序列的加工、捕获和整合过程。在间隔序列整合阶段,由重复序列中的保守基序介导其自身的准确复制,而 PAM-protospacer 两端的序列将影响整合的间隔序列的长度。

在嗜盐古菌的 I-B 亚型 CRISPR-Cas 系统中,适应和干扰反应均需要成熟 crRNA 的参与,而 crRNA 的加工成熟主要是在稳定期,说明 CRISPR-Cas 系统的活性也随着稳定期的到来而提高。对此,可以作出如下猜测:在对数生长期,细胞代谢旺盛,抵抗噬菌体侵染的能力最强,而在这个时期高表达 CRISPR-Cas 系统对其而言是巨大的代谢负担,

不利于细胞快速生长分裂;而进入稳定期后,菌体密度达到最大,细胞活力下降,有害代谢产物积累,易受噬菌体侵染,此时大量成熟的 crRNA 与 Cas 蛋白结合形成效应复合体,为宿主提供有效的免疫保护。

4 I-B 亚型 CRISPR-Cas 系统在遗传操作中的应用

4.1 基因编辑

应用内源性 CRISPR-Cas 系统进行基因编辑的简便性在于,只需要在细胞中表达特定的引导 RNA,而不需要额外表达 Cas 蛋白,避免了异源表达蛋白带来的细胞毒性。图 2 显示了应用嗜盐古菌的 CRISPR-Cas 系统进行基因编辑的原理^[30]。采用的基因编辑质粒(pGE)包含一小段人造 CRISPR 序列和不含靶序列的供体 DNA。细胞接受质粒后会走向两种命运:野生型细胞的染色体会成为 CRISPR 攻击的对象,造成细胞死亡;而位点发生变异的细胞将免受 CRISPR 攻击,存活下来成为平板上的转化子。在突变体中,供体 DNA 序列替换了原来基因组位点的 DNA 序列,从而定向引入基因敲除、基因敲入和基因原位标记。

在不同的嗜盐古菌中应用上述原理进行基因组编辑时显示出不同的效果。在沃氏富盐菌中利用其 I-B 亚型 CRISPR-Cas 系统敲除一个与色素生成相关的非必需基因 *crtI* 后,CRISPR-Cas 系统介导的自靶向作用不能诱导靶位点发生预期变异^[31]。而在另一种同为多倍体嗜盐古菌西班牙盐盒菌中,由 I-B 亚型 CRISPR-Cas 系统介导的自靶向显示出较高的细胞毒性,能有效诱发同源重组,在绝大多数的转化子中实现精确的基因编辑^[32]。

两种嗜盐古菌对于 CRISPR 介导的自靶向作用显示出不同的耐受性,可能与其染色体多倍性的差异有关。在对数生长期,沃氏富盐菌中的染色体高达 18 个拷贝,而同样时期的西班牙盐盒菌中染色体只有 8 个拷贝^[32,33]。在染色体拷贝数较高的嗜盐古菌内,大量存在的自身模板将介导有效的 DNA 重组修复,和 CRISPR-Cas 系统介导的自靶向作用相互抵消,难以实现外源模板与原基因组序列间的替换。因此,在对多倍体嗜盐古菌进行基因编辑时,需结合其染色体多倍性的情况调整具体方案。

4.2 基因表达调控

CRISPR-Cas 系统已被开发为基因沉默工具,称为 CRISPR 干扰 (CRISPR interference, CRISPRi)。在这项技术中,效应复合体在 crRNA 的指导下结合到特定位点(如启动子)上,对 RNA 聚合酶形成空间位阻,从而抑制基因表达^[2]。随着古菌分子

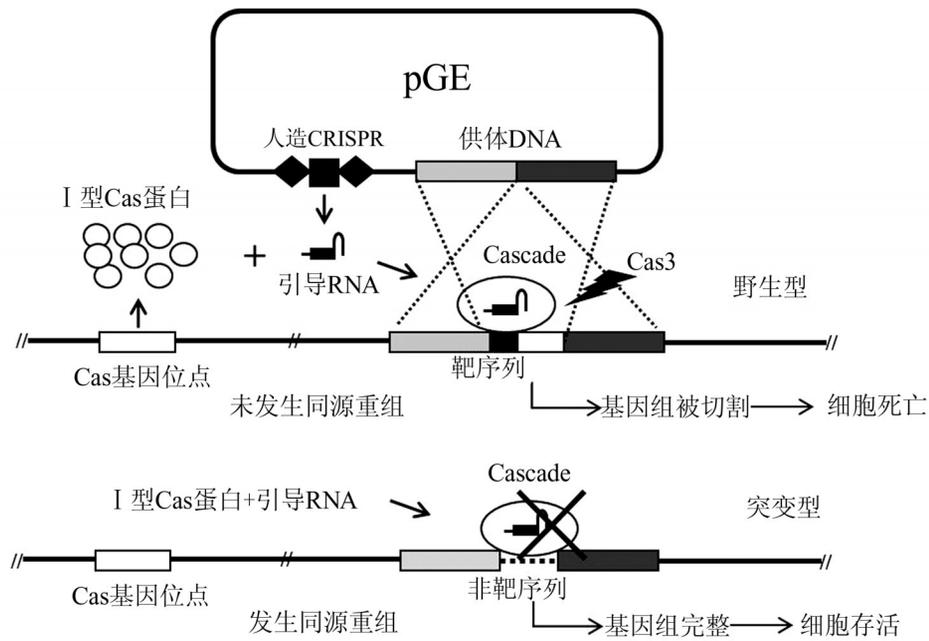


图2 I-B亚型 CRISPR-Cas 系统应用原理

Fig. 2 Application principle of type I-B CRISPR-Cas system

注:本图参照文献[30]修改绘制

Note: this figure was drawn according to the literature [30] with modifications

生物学研究的发展,迫切需要类似的转录调控工具来验证必需基因的生物功能。嗜盐古菌沃氏富盐菌中的研究表明,在敲除 *cas3* 和 *cas6b* 基因的菌株中,其 I-B 亚型 CRISPR-Cas 系统能介导有效的转录抑制^[34]。该抑制效果与靶位点的选择有关,靶向启动子区域时的抑制效果优于靶向开放阅读框,靶向模板链时的抑制效果优于靶向编码链。质粒和染色体上的基因、单顺反子或者是操纵子的一部分,都可以成为调控的对象。

5 总结和展望

目前,嗜盐古菌中 I-B 亚型 CRISPR-Cas 系统的分子机制已得到了较为全面的解析。I-B 亚型 CRISPR-Cas 系统作为自然界中分布最广的 CRISPR-Cas 系统,其分子机制的研究成果将更加具有普遍性和代表性。

I-B 亚型 CRISPR-Cas 系统作为遗传操作工具有两大突出优势:(1)更多的可识别 PAM。PAM 的要求常常限制了基因编辑位点的选择,嗜盐古菌中的 Cascade 复合体能识别多个 PAM(如沃氏富盐菌中的 Cascade 能识别 7 个 PAM),在位点的选择上有更高的灵活性。(2)更低的脱靶效率。与应用最为广泛的 CRISPR/Cas9 系统相比, I-B 亚型 CRISPR-Cas 系统的引导 RNA 与靶位点间存在更长的配对区域,从而具备更高的靶点特异性,有效降低脱靶效

率。因此,无论是应用在编码这类系统的生物(如嗜盐古菌)中,还是作为外源性 CRISPR-Cas 系统应用在其他生物中, I-B 亚型 CRISPR-Cas 系统都是极具应用潜力的遗传操作工具。

应该指出的是, I-B 亚型 CRISPR-Cas 系统的作用机制中还有一些重要问题有待解决。例如,适应阶段和干扰阶段分别有哪些 Cas 蛋白参与并采取何种方式对 PAM 进行识别? 效应复合体与靶分子形成 R 环后如何招募 Cas3 蛋白对靶 DNA 进行降解? 既然引发适应需要一段与外源 DNA 完全或部分匹配的 crRNA,那么 CRISPR 阵列中的第一个间隔序列如何被获取? 对这些问题的进一步研究及合理解释将深化人们对 CRISPR-Cas 系统的认识并促进其在基因编辑方面的应用。

参考文献

- [1] Giani M, Garbayo I, Vilchez C, *et al.* Haloarchaeal carotenoids: healthy novel compounds from extreme environments [J]. *Mar Drugs*, 2019, 17(9): 524.
- [2] Knott G J, Doudna J A. CRISPR-Cas guides the future of genetic engineering [J]. *Science*, 2018, 361(6405): 866-869.
- [3] Maier L K, Stachler A E, Brendel J, *et al.* The nuts and bolts of the *Haloferax* CRISPR-Cas system I-B [J]. *RNA Biol*, 2019, 16(4): 469-480.
- [4] Woese C R, Kandler O, Wheelis M L. Towards a natu-

- ral system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya [J]. Proc Natl Acad Sci USA, 1990, 87(12): 4576-4579.
- [5] Gupta R S, Naushad S, Fabros R, *et al.* A phylogenomic reappraisal of family-level divisions within the class Halobacteria: proposal to divide the order Halobacteriales into the families Halobacteriaceae, Haloarculaceae fam. nov., and Halococcaceae fam. nov., and the order Haloferacales into the families, Haloferacaceae and Halorubraceae fam nov [J]. Antonie Van Leeuwenhoek, 2016, 109(4): 565-587.
- [6] Eme L, Doolittle W F. Archaea [J]. Curr Biol, 2015, 25(19): R851-R855.
- [7] Zerulla K, Soppa J. Polyploidy in haloarchaea: advantages for growth and survival [J]. Front Microbiol, 2014, 5: 274.
- [8] Ludt K, Soppa J. Polyploidy in halophilic Archaea: regulation, evolutionary advantages, and gene conversion [J]. Biochem Soc Trans, 2019, 47(3): 933-944.
- [9] Marraffini L A, Sontheimer E J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea [J]. Nat Rev Genet, 2010, 11(3): 181-190.
- [10] Feng J, Liu B, Zhang Z, *et al.* The complete genome sequence of *Natrinema* sp. J7-2, a haloarchaeon capable of growth on synthetic media without amino acid supplements [J]. PLoS One, 2012, 7(7): e41621.
- [11] Makarova K S, Wolf Y I, Iranzo J, *et al.* Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants [J]. Nat Rev Microbiol, 2020, 18(2): 67-83.
- [12] Makarova K S, Wolf Y I, Alkhnbashi O S, *et al.* An updated evolutionary classification of CRISPR-Cas systems [J]. Nat Rev Microbiol, 2015, 13(11): 722-736.
- [13] Liu T Y, Doudna J A. Chemistry of Class 1 CRISPR-Cas effectors: binding, editing, and regulation [J]. J Biol Chem, 2020, 295(42): 14473-14487.
- [14] Sinkunas T, Gasiunas G, Fremaux C, *et al.* Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system [J]. EMBO J, 2011, 30(7): 1335-1342.
- [15] Leenay R T, Maksimchuk K R, Slotkowski R A, *et al.* Identifying and visualizing functional PAM diversity across CRISPR-Cas systems [J]. Mol Cell, 2016, 62(1): 137-147.
- [16] Shah S A, Erdmann S, Mojica F J, *et al.* Protospacer recognition motifs: mixed identities and functional diversity [J]. RNA Biol, 2013, 10(5): 891-899.
- [17] Fischer S, Maier L K, Stoll B, *et al.* An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA [J]. J Biol Chem, 2012, 287(40): 33351-33363.
- [18] Maier L K, Lange S J, Stoll B, *et al.* Essential requirements for the detection and degradation of invaders by the *Haloferax volcanii* CRISPR/Cas system I-B [J]. RNA Biol, 2013, 10(5): 865-874.
- [19] Maier L K, Stachler A E, Saunders S J, *et al.* An active immune defense with a minimal CRISPR (clustered regularly interspaced short palindromic repeats) RNA and without the Cas6 protein [J]. J Biol Chem, 2015, 290(7): 4192-4201.
- [20] Gong L Y, Li M, Cheng F Y, *et al.* Primed adaptation tolerates extensive structural and size variations of the CRISPR RNA guide in *Haloarcula hispanica* [J]. Nucleic Acids Res, 2019, 47(11): 5880-5891.
- [21] Rutkauskas M, Sinkunas T, Songailiene I, *et al.* Directional R-loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection [J]. Cell Rep, 2015, 10(9): 1534-1543.
- [22] Li M, Liu H, Han J, *et al.* Characterization of CRISPR RNA biogenesis and Cas6 cleavage-mediated inhibition of a provirus in the haloarchaeon *Haloferax mediterranei* [J]. J Bacteriol, 2013, 195(4): 867-875.
- [23] Zhang Z, Pan S, Liu T, *et al.* Cas4 nucleases can effect specific integration of CRISPR spacers [J]. Journal of Bacteriology, 2019, 201(12): e718-e747.
- [24] Wu C, Tang D, Cheng J, *et al.* Mechanisms of spacer acquisition by sequential assembly of the adaptation module in *Synechocystis* [J]. Nucleic Acids Res, 2021, 49(5): 2973-2984.
- [25] Li M, Wang R, Zhao D, *et al.* Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process [J]. Nucleic Acids Res, 2014, 42(4): 2483-2492.
- [26] Li M, Wang R, Xiang H. *Haloarcula hispanica* CRISPR authenticates PAM of a target sequence to prime discriminative adaptation [J]. Nucleic Acids Res, 2014, 42(11): 7226-7235.
- [27] Wang R, Li M, Gong L, *et al.* DNA motifs determining the accuracy of repeat duplication during CRISPR adaptation in *Haloarcula hispanica* [J]. Nucleic Acids Res, 2016, 44(9): 4266-4277.
- [28] Nuñez J K, Kranzusch P J, Noeske J, *et al.* Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity [J]. Nat Struct Mol Biol, 2014, 21(6): 528-534.
- [29] Li M, Gong L, Zhao D, *et al.* The spacer size of I-B CRISPR is modulated by the terminal sequence of the protospacer [J]. Nucleic Acids Res, 2017, 45(8): 4642-

- 4654.
- [30] Li Y, Pan S, Zhang Y, *et al.* Harnessing type I and type III CRISPR-Cas systems for genome editing [J]. *Nucleic Acids Res*, 2016, 44(4): e34.
- [31] Stachler A E, Turgeman-Grott I, Shtifman-Segal E, *et al.* High tolerance to self-targeting of the genome by the endogenous CRISPR-Cas system in an archaeon [J]. *Nucleic Acids Res*, 2017, 45(9): 5208-5216.
- [32] Cheng F, Gong L, Zhao D, *et al.* Harnessing the native type I-B CRISPR-Cas for genome editing in a polyploid archaeon [J]. *J Genet Genomics*, 2017, 44(11): 541-548.
- [33] Breuert S, Allers T, Spohn G, *et al.* Regulated polyploidy in halophilic Archaea [J]. *PLoS One*, 2006, 1: e92.
- [34] Stachler A E, Marchfelder A. Gene repression in haloarchaea using the CRISPR (clustered regularly interspaced short palindromic repeats)-Cas I-B system [J]. *J Biol Chem*, 2016, 291(29): 15226-15242.
- (编辑:梅楠 杨晓翠)