



细菌 CRISPR-Cas 系统的研究进展

胡丽, 陈实*

武汉大学药学院, 湖北 武汉 430072

摘要: 获得性免疫长期以来被视为真核生物所独有, 而 CRISPR-Cas 系统的发现则打破了这一定论。它是广泛存在于细菌和古菌中的一种获得性免疫系统, 通过捕获整合初次感染的外源核酸片段, 在 Cas 蛋白与 crRNA (CRISPR RNAs) 的共同作用下抵御相同核酸的再次入侵, 以保护宿主免受侵扰。近些年, CRISPR-Cas 系统得到广泛的关注和研究。本文主要从细菌微生物角度, 对系统分类、作用机制及原核领域应用等取得重要突破的研究进行扼要阐述, 为 CRISPR-Cas 系统的深入探究和应用拓展提供有价值的参考信息。

关键词: CRISPR-Cas 系统, 分类, 作用机制, 原核领域应用

细菌和古菌广泛存在于自然界中, 且长期以来一直受到诸如噬菌体等外源因素的侵扰。由于噬菌体的高突变率和重组能力, 极大地威胁这些微生物的生命活动。迫于生存压力它们进化出了多个防御机制来抵抗外源核酸的侵扰以保护自己^[1]。早期研究发现, 微生物主要通过多种先天性免疫系统抵御外界侵扰、比如受体遮蔽、限制修饰(R-M)系统、基因干扰、噬菌体排除、感染顿挫等^[1]。

CRISPR-Cas 系统是最近发现存在于原核生物中的一类获得性免疫系统。不同以往免疫机制, 它通过记录初次侵扰的外源核酸产生免疫记忆,

当再次入侵时, “记忆片段”的转录产物连同系统中 Cas 蛋白一起, 将核酸降解以维持宿主的正常生命活动^[2]。该系统是近几年的研究热点, 在真核及原核生物中运用极为广泛。此外, CRISPR-Cas 新系统的发现及其防御机制的研究成果颇为丰富, 下文着重概述该系统的分类和特征、防御机制及在微生物领域的前沿应用进展。

1 CRISPR-Cas 系统简述

CRISPR-Cas 系统始于 Ishino 等对大肠杆菌 K12 的研究, 无意发现一段无法确定功能的重复序列。直至 2002 年, 这段重复序列才被命名为成簇

基金项目: 国家自然科学基金(31670086)

*通信作者。Tel: +86-27-68756643; E-mail: shichen@whu.edu.cn

收稿日期: 2017-04-04; 修回日期: 2017-06-01; 网络出版日期: 2017-06-09

有规律的间隔短回文重复序列(CRISPR : Clustered Regularly Interspaced Short Palindromic Repeats)。Barrangou 等通过噬菌体感染实验证实 CRISPR 与其附近 *cas* 基因能帮助宿主抵抗噬菌体入侵^[3-5]。CRISPR-Cas 系统包含 CRISPR 基因座和 *cas* 基因两部分。其中, CRISPR 基因座主要由前导序列 (leader)、重复序列(repeat)和间隔序列(spacer)构成。前导序列富含 AT 碱基,且位于 CRISPR 基因座上游;重复序列的长度约 20-50 bp 碱基且包含 5-7 bp 回文序列,转录出的茎环结构能够稳定 RNA 的整体二级结构^[6];间隔序列为从外源捕获的核酸片段。*Cas* 基因位于基因座附近或分散于基因组其他地方,编码的蛋白在防御过程中至关重要,且作为重要依据将该系统进行科学的归类^[6]。

2 CRISPR-Cas 系统的种类和特征

在现有的数据库中,50%细菌及 90%古菌的基因组或质粒上存在至少 1 个 CRISPR 基因座^[7]。换言之,CRISPR-Cas 系统数量庞大,单一的标准难以将其归类。早期,根据 *Cas* 蛋白之间的进化关系和 *Cas* 基因组成操纵子方式,系统分为三种类型: I 型、II 型和 III 型,且分别含有 *Cas3*、*Cas9* 和 *Cas10* 特征蛋白^[8]。然而,随着对不同生物基因组的深度测序,之前的划分标准已无法将新发现的系统归至其中。依据 *Cas* 蛋白在防御过程中参与的角色, Makarova 等进一步将系统分为两大类。第一大类包括 I 型、III 型和 IV 型,它们切割外源核酸的效应因子为多个 *Cas* 蛋白形成的复合物;而第二大类系统的作用因子是比较单一的 *Cas* 蛋白,比如 II 型的 *Cas9* 蛋白和 V 型的 *Cpf* 蛋白,所以, II 型和 V 型被划为第二大类^[9]。之后, Shmakov 等人对此进行了补充,发现了 V 型的另

外 2 个亚型及 VI 型新系统^[10]。依据目前已发现的 CRISPR-Cas 系统,将其分类汇总于图 1。

2.1 I 型 CRISPR-Cas 系统

I 型系统最为复杂且 *cas* 基因最多,它由单一或多个操纵子负责。除了泛表达蛋白 *Cas1*、*Cas2* 和上述提及的特征蛋白 *Cas3* 外, I 型系统还含有 *Cas5*、*Cas6e* 和 *Cas7* 等 2 个重复系列相关蛋白 (RAMP), 它们的 RNA 酶活性在 pre-crRNA 的剪辑中起重要作用^[9]。此外, I 型系统被分为 I-A 到 I-U 七个亚型^[9-10]。*Cas8a*、*Cas8b* 和 *Cas8c* 分别为 I-A、I-B 和 I-C 的标志蛋白,它们通过与其他 *Cas* 蛋白形成 Cascade 复合物参与防御^[11]。*Cas10d* 是 I-D 亚型的标志蛋白, *Cse1* 和 *Cse2* 为 I-E 亚型标志蛋白。I-F 则含有 *Csy1*、*Csy2*、*Csy3* 和 *Csy6* 四个标志蛋白。I-U 亚型中的 U 表示未知功能的特征蛋白^[9], 比如 I-F 的一种衍生型缺失 *cas1*、*cas2* 和 *cas3* 基因,取而代之是 *csy1* 与 *csy2* 的融合基因、*csy3* 以及 *cas6f* 基因^[9]。所以,当一个 CRISPR-Cas 系统含有亚型中标志蛋白而其他基因存在差异,可将其归类为该亚型的衍生型或归为 I-U 型。

2.2 II 型 CRISPR-Cas 系统

II 型系统与 I 型相比, *Cas* 蛋白种类较简单。大蛋白 *Cas9* 是 II 型系统的特征蛋白,它能加工 pre-crRNA 促进其成熟的同时切割外源核酸^[12]。*Cas9* 与 sgRNA 和靶标 DNA 的共晶结构揭示,蛋白呈两个瓣状结构,将靶标 DNA 和 sgRNA 固定在瓣片结构的界面,瓣状间的环结构可辅助 PAM 序列的识别^[12-13]。此外, *Cas9* 除含有 *RuvC* 和 *HNH* 的核酸酶结构域,还包含部分与迁移基因同源的序列,能够编码富含精氨酸的蛋白并结合 RNA。然而,这段序列与 CRISPR 系统无直接关联性,

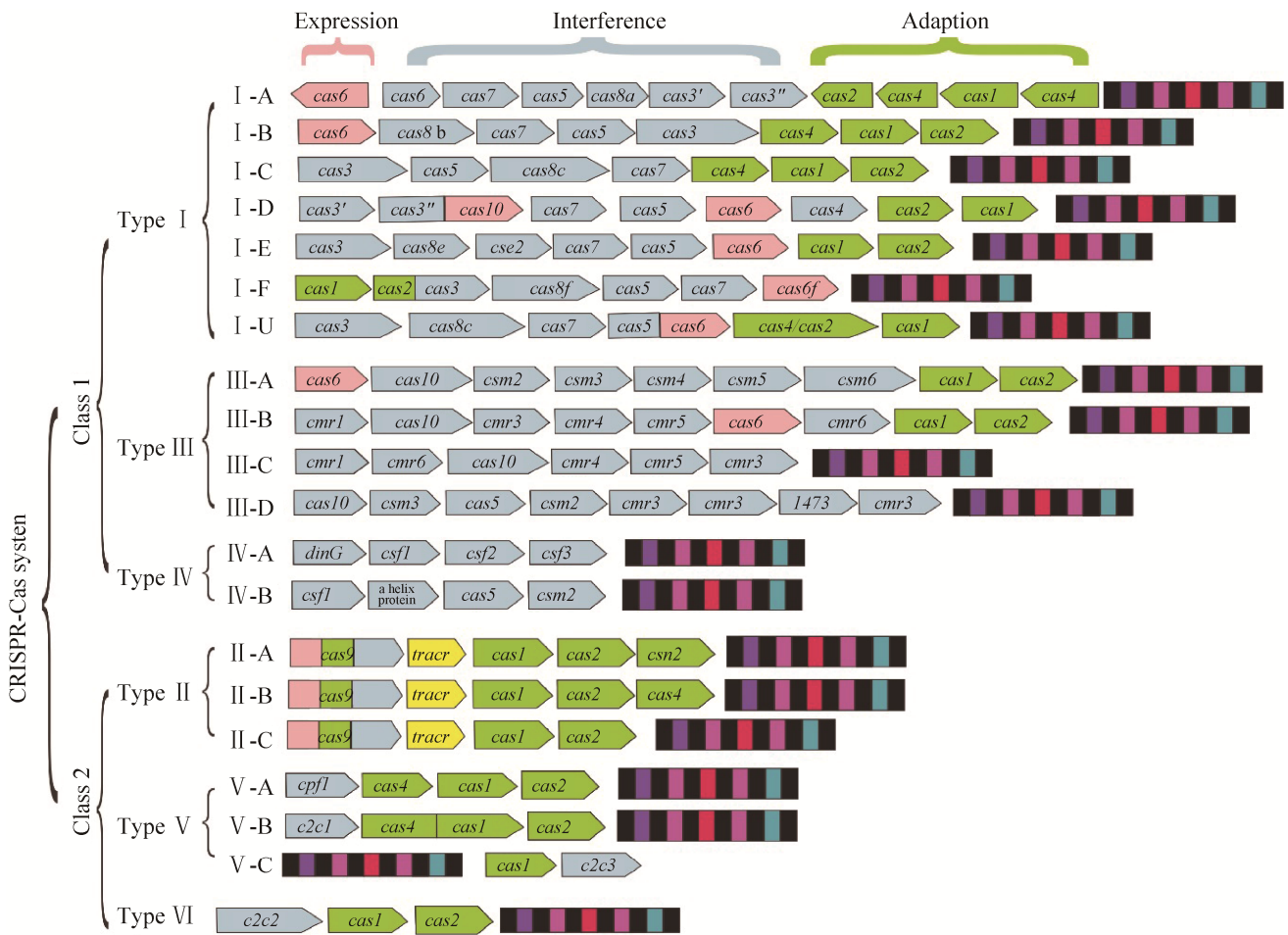


图 1. CRISPR-Cas 系统的分类、典型结构代表和功能组成元件标注示意图

Figure 1. Schematic diagram of classification, typical architecture, functional components in CRISPR-Cas system.

即暗示着 Cas9 蛋白可能不是鉴定 II 型系统的唯一特征。而 *cas9* 基因位于 *cas1* 与 *cas2* 附近可作为 II 型的一个重要标识^[14-15]。

II 型 CRISPR-Cas 系统除泛表达蛋白 Cas1、Cas2 以及特征蛋白 Cas9 外,还包含 1 个或 2 个基因,其编码的 *tracrRNA* 能辅助 RNAase III 和 Cas9 对 pre-crRNA 进行剪切,同成熟的 crRNA 部分碱基互补配对形成二聚体,进一步与 Cas9 蛋白结合形成核糖核蛋白复合物^[14]。

II 型系统目前分为: II-A、II-B 和 II-C^[8-9]。其中, II-A 系统含有 *csn2* 基因,编码的蛋白通过形

成二聚体环能结合线性双链的 DNA^[16]; II-B 型的标志蛋白 Cas4 具有 PD-EDxK 家族的 5'单链 DNA 外切核酸酶活性^[17]; II-C 系统仅有 *cas1*、*cas2* 和 *cas9* 三个基因^[14-18]。

由于 II 型系统组分简单,目前被广泛应用于基因定点编辑、基因组筛选、基因转录调控、基因组成像、基因诊疗、生态应用等研究领域^[19]。

2.3 III型 CRISPR-Cas 系统及特征

III型系统中的特征蛋白 Cas10 具有类似 PolB 家族环化酶和聚合酶功能^[20]。晶体结构揭示,其含有 RRM 折叠的掌式结构域,包含锌结合基序

的螺旋结构域,可能有活性的掌式结构域,类似 A 家族的 DNA 聚合酶和 Cmr5 的 α 螺旋结构域^[21]。该蛋白通常与含有环状循环排列的保守基序的 HD 家族核酸酶融合发挥作用^[22]。这些研究成果为 III 型系统的作用机制研究提供了重要的理论基础。

III 型 CRISPR-Cas 系统分为 4 个亚型: III-A、III-B、III-C 和 III-D^[9]。其中, III-A 亚型含有 Csm 系列蛋白、Cas1、Cas2 和 Cas6 蛋白,干扰目标为 DNA^[23-24]。III-B 亚型含有 Cmr 系列蛋白但缺乏 Cas1、Cas2 和 Cas6 蛋白,所以依赖于生物体内的其他 CRISPR 系统,干扰靶标为 RNA^[25-26]。III-C 亚型含有环化酶结构域失活的 Cas10 蛋白^[27]。III-D 亚型包含与 1473 家族同源的一个未知功能基因,同时 Cas10 蛋白缺乏 HD 结构域^[27]。

2.4 IV 型 CRISPR-Cas 系统及特征

IV 型 CRISPR-Cas 系统在基因组和质粒上均有发现,且往往仅有独立的 *cas* 基因而无相关 CRISPR 基因座。Csf1 作为特征蛋白是系统效应复合物的主要成分,该复合物还含有 RAMP 蛋白: Csf3 和 Csf2^[27]。

依据 DinG 家族解旋酶 Csf4, IV 型系统分为 IV-A 和 IV-B。其中 IV-A 型含有 *csf4* 蛋白,而 IV-B 含有一个可能组成效应复合物的 α 螺旋蛋白^[27-28]。IV 型的作用机理与 III 型相似,需要其他类型系统的辅助来发挥防御作用。

2.5 V 型和 VI 型 CRISPR-Cas 系统及特征

V 型与 II 型同属于第二大类系统。Cpf1 作为特征蛋白具有类似 RuvC 核酸酶结构域,而缺乏 HNH 核酸酶结构域,偏向于识别富含胸腺嘧啶核苷酸的 PAM 序列,并在 RNA 指导下切割磷酸二酯键产生粘性末端^[9,29-30]。近期与 crRNA 的共晶

结构解答 Cpf1 如何识别 crRNA 的问题,为 V 型的研究提供了重要依据^[29]。

根据生物基因组和宏基因组筛选, V 型系统分为 V-A、V-B 和 V-C^[9-10]。Cpf1、C2c1 和 C2c3 分别是这 3 种类型的特征蛋白。其中, C2c1 是双分子 RNA 导向的核酸酶并靶向 DNA, C2c1 和 C2c3 具有类似 RuvC 核酸酶结构域,同时与 Cpf1 结构有高度相似性^[9]。

VI 型 CRISPR-Cas 系统发现较晚, C2c2(特征蛋白)因含有 2 个 HEPN 结构域而具有核糖核酸酶活性^[9]。纤毛菌属中 C2c2 同 crRNA 的复合物和自由状态下 C2c2 的晶体结构揭示, C2c2 的两个独立活性结构域行使不同的酶切活性^[31], 表现出了蛋白的功能多样性。

3 CRISPR-Cas 系统的防御机制

作为由微生物长期进化而形成的获得性免疫系统, CRISPR-Cas 能够对各种方式的入侵产生特异的防御, 比如质粒的转化或结合、噬菌体的感染等^[32]。该防御过程包括 3 个阶段(图 1): 第一是间隔序列的获取, 称适应阶段; 第二是 crRNA 的合成, 即 CRISPR 基因座的表达; 第三是对外来核酸的抵御, 称干扰阶段^[33]。上述提及的不同系统在这 3 个阶段存在一定的差异。

3.1 CRISPR-Cas 系统适应阶段

当外源核酸入侵含有 CRISPR-Cas 系统的生物体, 宿主最先对其进行识别和扫描, 之后通过相关蛋白将其捕获并整合至先导序列下游, 形成新的间隔序列。对于大多数的 CRISPR 系统, 外源核酸的识别与选择依赖于原间隔序列邻近的 2-5 个核苷酸, 称 PAM 序列^[34], 它除了在整合阶段发挥作用, 还能够指导系统对外源核酸进行特

异切割, 从而避免宿主发生自身免疫^[34]。

I 型系统在该阶段通过两种方式捕获外源核酸。当噬菌体首次入侵生物体, 宿主采用“naive”方式整合外源 DNA^[35], 利用 RecBCD、限制核酸酶或系统相关核酸酶将外源 DNA 处理成小片段, Cas1-Cas2 复合物扫描 PAM 序列选择合适片段, 通过亲核进攻和非 Cas 蛋白修复系统将片段整合至基因座中^[35]。若系统已整合间隔序列, 噬菌体再次入侵则采用“primed”方式获取外源 DNA^[35], 该过程包含 Cas3 处理 DNA 和 Cas1-Cas2 复合物对合适片段整合等过程^[35]。II 型中 Cas9-tracrRNA 复合物同 Cas1-Cas2 复合物参与到对外源核酸整合过程, 有研究表明 II-A 亚型的 Csn2 和 II-B 亚型的 Cas4 也参与其中^[36-37]。III 型中 Cas1 蛋白能与逆转录酶融合, 将 RNA 片段直接整合至系统中^[38]。

3.2 CRISPR 基因座表达阶段

CRISPR 基因座表达的起始产物为长达 30–65 nt 的 pre-crRNA, 经过相关 Cas 蛋白和 RNA 核酸酶的加工形成成熟 crRNA, 此过程即为 CRISPR 基因座的表达^[39]。

第一大类系统中, Cas6 蛋白不依赖于金属离子识别前体 crRNA 的重复区域并进行切割, 对前体 crRNA 进行加工^[39]。I-C 亚型因缺乏 Cas6, 由 Cas5d 蛋白代替参与到该过程中, III 型往往也缺乏该类蛋白, 通过利用孤立的类似 Cas6 核酸酶来负责加工过程, 比如嗜热细菌 HB8^[40-41]。第二大类中, II 型系统的 tracrRNA 首先与前体 crRNA 部分序列互补配对形成复合物, Cas9 蛋白稳定复合物状态, 随后内源性 RNA 核酸酶对前体 crRNA 的重复序列区域剪切进行加工^[11,39-42]; Cpf1 及 C2c2 分别是 V 型和 VI 型系统的重要蛋白, 利用自身 RNA 核酸酶活性对前体 crRNA 进行加工^[31-43]。

3.3 CRISPR-Cas 系统干扰阶段

干扰阶段是防御入侵最重要的一步, 主要通过 RNA 介导准确识别外源核酸并对其进行精准剪切。第一大类系统的多亚基效应复合物和第二大类系统的单效应蛋白在此阶段发挥重要作用。

第一大类系统中, I 型系统效应复合物由 Cas 蛋白和成熟 crRNA 构成。在该过程起始, 效应复合物中 Cse1 亚基识别靶标 DNA 的 PAM 序列并与之结合, 造成 DNA 从 PAM 端逐渐解开^[43-44], 进一步 crRNA 与目标 DNA 结合形成稳定的复合物从而使效应复合物构象发生变化, 招募 Cas3 蛋白对目标 DNA 进行解旋并切割, 最后切割产物 ssDNA 作为底物被其他核酸酶彻底降解以达到防御目的^[45-46]。

同属第一类的 III 型, Csm 和 Cmr 系列蛋白为效应复合物的结构与 I 型的有一定相似性, 但作用机理存在差异, 区别在于 III 型系统依赖于 crRNA 或相关蛋白识别自身和外源核酸, 非 PAM 序列^[45-46]。此外, 上述提到 III 型系统能够干扰 RNA 是由 Cmr 效应复合物负责, 另一亚型系统 Csm 效应复合物则靶向 DNA^[24-26], 有研究表明 Csm 复合物能靶向 DNA 和 RNA^[47]以及两种亚型复合物是依赖于转录的 DNA 核酸酶^[40]。

第二大类中, II 型系统的 Cas9 蛋白与 crRNA 和 tracrRNA 形成效应复合物对外源 DNA 进行干扰, 复合物中 Cas9 蛋白通过识别原间隔序列 3' 端 PAM 序列靶向 DNA 使其解旋, 同时 crRNA 与之互补配对形成 R 环结构, 随后 Cas9 蛋白的核酸酶结构域构象发生变化并切割 DNA^[12-48]。V 型系统与 II 型相似, V-B 亚型的 C2c1 蛋白在 tracrRNA 的协助下切割外源核酸, V-A 亚型的 Cpf1 蛋白作为效应因子即可靶向外源 DNA^[33]。此外, V 型系

统通过识别位于 5'端且富含胸腺嘧啶的 PAM 序列确定目标 DNA^[33]。在原间隔序列 3'端具有特定序列条件下, VI型的效应蛋白 C2c2 在 crRNA 的介导下靶向外源 ssRNA 并进行切割^[49]。

4 CRISPR-Cas 系统在原核生物中的应用

CRISPR-Cas 系统对宿主在抵御外来核酸入侵时起到重要的作用, 其中 II 型系统是众多类型中最为简单且研究最为透彻的一种。近几年 CRISPR-Cas9 系统发展成为最热门的基因编辑工具, 在人类细胞、酵母、斑马鱼、大鼠、小鼠和玉米等真核生物得到广泛的应用。相比于真核生物, 它在原核生物中也展现出强大的编辑及基因调控等功能。

4.1 基因编辑

基因编辑技术作为研究基因功能的重要手段, 从早期的同源重组方式至锌指核酸酶和类转录激活因子效应核酸酶再到最近热门的 CRISPR-Cas9 系统, 一直在不断发展与创新。目前 CRISPR-Cas9 系统的编辑技术研究成熟, 在细菌领域得到广泛应用。多组研究表明该系统往往通过两种方式实现对细菌基因组的编辑, 一是引入外源 CRISPR-Cas9 系统, 二是利用细菌内源性 CRISPR-Cas9 进行编辑。最近曾虎、江宇等利用第一种方式分别对链霉菌及大肠杆菌等原核菌属实现了基因的定点敲除和多重基因编辑^[50]。本课题组利用同种方式, 应用外源 CRISPR-Cas9 对阿维链霉菌 M4680 的磷硫酰化修饰基因簇进行敲除, 旨在探究该种 DNA 骨架新型修饰(硫原子取代非桥连氧原子)对主要次级代谢产物阿维菌素的影响, 目前探索正在进行中。Bassalo 等改进这种方法, 将系统与噬菌体衍生型 λ Red 重组结合

以提升 CRISPR 诱发的同源直接修复效率, 实现对大肠杆菌基因组更高效的编辑^[51]。至于后者, Pyne 等通过直接引入前体 RNA 表达系统, 利用细菌自身内源性系统成功对梭状芽孢杆菌进行定点编辑^[52]。

4.2 基因表达调控

对于保守且敲除困难的基因, 应用 CRISPR-Cas9 系统也能方便地调节基因的表达水平以研究功能。当 dCas9 蛋白因核酸酶结构域失活而仅能结合目标 DNA, 即能实现对基因表达进行调控。为达到此目的, dCas9 蛋白在 crRNA 介导下主要通过阻碍 RNA 聚合酶移动和阻止 RNA 聚合酶在启动子区结合的两种方式。Cleto 等通过前者方式对基因进行抑制调控, 发现对靶向非模板链能够达到更好的抑制效果^[53]。然而更多研究成果证实第二种抑制调控方式, 不依赖于 DNA 链使目的基因表达下调更为显著^[54]。基于此, 本课题组对 DNA 磷硫酰化系统进行相似研究, 通过该系统靶向基因簇启动子区域使得修饰量下调以观察菌的形态变化及菌中其他基因表达情况, 从而探究 DNA 磷硫酰化系统的生理意义。此外, 若 dCas9 蛋白融合激活因子, 能够上调基因表达水平, Bikard 等把 dCas9 蛋白与 RNA 聚合酶的 ω 亚基融合, 在成熟 crRNA 的指导下靶向基因启动子区, 实现约 23 倍表达量的提升^[55]。这些成果体现了 CRISPR-Cas 系统对基因调控的双向性和高效性。

4.3 其他应用

微生物的分类研究是一个重要的分支, 由于 CRISPR-Cas 系统的普遍性和自身结构的特征性, 通过该系统能够简单、准确地辅助微生物菌种的分类^[56]。再者, 通过比对 CRISPR 基因座中的间隔序列, 能够对环境中的病原体进行有效追踪来

了解微生物进化史^[57]。此外,由于发生在染色体上的双链切割(DBS)损伤会导致生物体的死亡,Gomaa 等首次提出基于 CRISPR-Cas 系统筛选抗生素的理念,并在研究中得以证实^[58]。如上所述,CRISPR-Cas 系统在原核生物领域的应用,并不局限于简单的定点编辑和基因表达调控,在其他领域也得到广泛的应用。

5 CRISPR-Cas 系统的总结与展望

CRISPR-Cas 系统作为微生物重要的保护装置,自发现以来,通过科研工作者的不懈努力,不管是系统的分类、作用机制还是功能等方面逐渐被了解。对于 CRISPR-Cas 系统的分类,Koonin、Makarova 和 Shmakov 等作出了重大的贡献,合理且比较全面地概括了目前已发现的 CRISPR 种类。然而,随着研究的深入,会发现更多且有可能不符合归类要求的系统,这就需要分类依据的科学性和创新性以满足新发现系统归属。CRISPR-Cas 防御机制研究比较全面,目前公认将其分为 3 个阶段:适应、表达和干扰,但对于作用的详尽过程还有待深入研究,尤其是最近刚发现的 V 型和 VI 型系统。至于 CRISPR-Cas 系统的功能,作为最热门工具被广泛应用到基因定点编辑、基因组筛选、基因转录调控、基因组成像等领域。如本文所述,在原核生物中同是重要的研究技术。最新有研究发现应用 V-A 亚型中的 Cpf1 蛋白能够在真核细胞 HEK293T 中对基因组进行多重编辑,且操作优于 Cas9 系统^[59],这暗示着基于 Cpf1 蛋白的编辑方法也有可能应用于原核系统中。此外,VI 型系统的 C2c2 蛋白因具有 RNA 核酸酶活性,可以利用其靶向 mRNA 以调控基因表达,不管是在真核还是原核细胞中均有可能实现。总之,尽管

目前 CRISPR-Cas 系统研究比较热门,但在机制及应用方面仍存在许多问题有待深入解决与探索,在全面了解作用机制的前提下才能更好地挖掘该系统的应用潜力。

参考文献

- [1] Makarova KS, Wolf YI, Koonin EV. Comparative genomics of defense systems in archaea and bacteria. *Nucleic Acids Research*, 2013, 41(8): 4360–4377.
- [2] Heler R, Marraffini LA, Bikard D. Adapting to new threats: the generation of memory by CRISPR-Cas immune systems. *Molecular Microbiology*, 2014, 93(1): 1–9.
- [3] 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. *Journal of Bacteriology*, 1987, 169(12): 5429–5433.
- [4] Jansen R, Van Embden JDA, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology*, 2002, 43(6): 1565–1575.
- [5] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, 315(5819): 1709–1712.
- [6] Godde JS, Bickerton A. The repetitive DNA elements called CRISPRs and their associated genes: evidence of horizontal transfer among prokaryotes. *Journal of Molecular Evolution*, 2006, 62(6): 718–729.
- [7] Grissa I, Vergnaud G, Pourcel C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research*, 2007, 35(S2): W52–W57.
- [8] Makarova KS, Haft DH, Barrangou R, Brouns SJJ, Charpentier E, Horvath P, Moineau S, Mojica FJM, Wolf YI, Yakunin AF. Evolution and classification of the CRISPR-Cas systems. *Nature Reviews Microbiology*, 2011, 9(6): 467–477.
- [9] Makarova KS, Koonin EV. Annotation and classification of CRISPR-Cas systems//Lundgren M, Charpentier E, Fineran P. *Methods in Molecular Biology*. New York: Humana Press, 2015, 1311: 47–75.
- [10] 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. *Molecular Cell*, 2015, 60(3): 385–397.
- [11] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012, 337(6096): 816–821.
- [12] 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. *Cell*, 2014, 156(5): 935–949.
- [13] Jinek M, Jiang FG, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou KH, Lin S, Kaplan M, Iavarone AT, Charpentier E, Nogales E, Doudna JA. Structures of cas9 endonucleases reveal RNA-mediated conformational activation. *Science*, 2014, 343(6176): 1247997.
- [14] Chylinski K, Le Rhun A, Charpentier E. The tracrRNA and cas9 families of type II CRISPR-Cas immunity systems. *RNA Biology*, 2013, 10(5): 726–737.
- [15] Chylinski K, Makarova KS, Charpentier E, Koonin EV. Classification and evolution of type II CRISPR-Cas systems. *Nucleic Acids Research*, 2014, 42(10): 6091–6105.
- [16] Arslan Z, Wurm R, Brenner O, Ellinger P, Nagel-Steger L, Oesterhelt F, Schmitt L, Willbold D, Wagner R, Gohlke H, Smits SHJ, Pul Ü. Double-strand DNA end-binding and sliding of the toroidal CRISPR-associated protein Csn2. *Nucleic Acids Research*, 2013, 41(12): 6347–6359.
- [17] Zhang J, Kasciukovic T, White MF. The CRISPR associated protein cas4 is a 5' to 3' DNA exonuclease with an iron-sulfur cluster. *PLoS One*, 2012, 7(10): e47232.
- [18] Koonin EV, Makarova KS. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biology*, 2013, 10(5): 679–686.
- [19] Singh A, Chakraborty D, Maiti S. CRISPR/Cas9: a historical and chemical biology perspective of targeted genome engineering. *Chemical Society Reviews*, 2016, 45(24): 6666–6684.
- [20] Anantharaman V, Iyer LM, Aravind L. Presence of a classical RRM-fold palm domain in Thg1-Type 3'→5' nucleic acid polymerases and the origin of the GGDEF and CRISPR polymerase domains. *Biology Direct*, 2010, 5(1): 43.
- [21] Xing Z, Ye KQ. Crystal structure of Cmr2 suggests a nucleotide cyclase related enzyme in type III CRISPR-Cas systems. *FEBS Letter*, 2012, 586(6): 939–945.
- [22] Makarova KS, Aravind L, Grishin NV, Rogozin IB, Koonin EV. A DNA repair system specific for thermophilic archaea and bacteria predicted by genomic context analysis. *Nucleic Acids Research*, 2002, 30(2): 482–496.
- [23] Makarova KS, Wolf YI, Koonin EV. The basic building blocks and evolution of CRISPR-Cas systems. *Biochemical Society Transactions*, 2013, 41(6): 1392–1400.
- [24] Marraffini LA, Sontheimer EJ. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, 2008, 322(5909): 1843–1845.
- [25] Hale CR, Majumdar S, Elmore J, Pfister N, Compton M, Olson S, Resch AM, Glover CVC, Graveley BR, Terns RM, Terns MP. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Molecular Cell*, 2012, 45(3): 292–302.
- [26] Hale CR, Zhao P, Olson S, Duff MO, Graveley BR, Wells L, Terns RM, Terns MP. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell*, 2011, 139(5): 945–956.
- [27] Makarova KS, Aravind L, Wolf YI, Koonin EV. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biology Direct*, 2011, 6: 38.
- [28] White MF. Structure, function and evolution of the XPD family of iron-sulfur-containing 5'→3' DNA helicases. *Biochemical Society Transactions*, 2009, 37(3): 547–551.
- [29] Zetsche 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. *Cell*, 2015, 163(3): 759–771.
- [30] Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, 157(6): 1262–1278.
- [31] Liu L, Li XY, Wang JY, Wang M, Chen P, Yin ML, Li JZ, Sheng G, Wang YL. Two distant catalytic sites are responsible for C2c2 RNase activities. *Cell*, 2017, 168(1/2): 121–134.e112.
- [32] Domingo E, Holland JJ. RNA virus mutations and fitness for survival. *Annual Review of Microbiology*, 1997, 51: 151–178.
- [33] Mohanraju P, Makarova KS, Zetsche B, Zhang F, Koonin EV, van der Oost J. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science*, 2016, 353(6299): aad5147.
- [34] Mojica FJM, Diez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 2009, 155(3): 733–740.
- [35] Fineran PC, Charpentier E. Memory of viral infections by

- CRISPR-Cas adaptive immune systems: acquisition of new information. *Virology*, 2012, 434(2): 202–209.
- [36] Wei YZ, Terns RM, Terns MP. Cas9 function and host genome sampling in type II-a CRISPR-Cas adaptation. *Genes & Development*, 2015, 29 (4): 356–361.
- [37] Heler R, Samai P, Modell JW, Weiner C, Goldberg GW, Bikard D, Marraffini LA. Cas9 specifies functional viral targets during CRISPR-Cas adaptation. *Nature*, 2015, 519(7542): 199–202.
- [38] Nuñez JK, Lee ASY, Engelman A, Doudna JA. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*, 2015, 519(7542): 193–198.
- [39] Charpentier E, Richter H, van der Oost J, White MF. Biogenesis pathways of RNA guides in archaeal and bacterial CRISPR-Cas adaptive immunity. *FEMS Microbiology Reviews*, 2015, 39(3): 428–411.
- [40] Staals RHJ, Zhu YF, Taylor DW, Kornfeld JE, Sharma K, Barendregt A, Koehorst JJ, Vlot M, Neupane N, Varossieau K, Sakamoto K, Suzuki T, Dohmae N, Yokoyama S, Schaap PJ, Urlaub H, Heck AJR, Nogales E, Doudna JA, Shinkai A, van der Oost J. RNA targeting by the type iii-a CRISPR-Cas Csm complex of *Thermus thermophilus*. *Molecular Cell*, 2014, 56(4): 518–530.
- [41] Nam KH, Haitjema C, Liu XQ, Ding F, Wang HW, Delisa MP, Ke A. Cas5d protein processes Pre-crRNA and assembles into a Cascade-like interference complex in subtype I-C/Dvulg CRISPR-Cas system. *Structure*, 2012, 20(9): 1574–1584.
- [42] Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 2011, 471(7340): 602–607.
- [43] Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 2016, 532(7600): 517–521.
- [44] Rutkauskas M, Sinkunas T, Songailiene I, Tikhomirova MS, Siksnys V, Seidel R. Directional R-Loop formation by the CRISPR-Cas surveillance complex cascade provides efficient off-target site rejection. *Cell Reports*, 2015, 10(9): 1534–1543.
- [45] Van Der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nature Reviews Microbiology*, 2014, 12(7): 479–492.
- [46] Redding S, Sternberg SH, Marshall M, Gibb B, Bhat P, Guegler CK, Wiedenheft B, Doudna JA, Greene EC. Surveillance and processing of foreign DNA by the *Escherichia coli* CRISPR-Cas system. *Cell*, 2015, 163(4): 854–865.
- [47] Samai P, Pyenson N, Jiang WY, Goldberg GW, Hatoum-Aslan A, Marraffini LA. Co-transcriptional DNA and RNA cleavage during type III CRISPR-Cas immunity. *Cell*, 2015, 161(5): 1164–1174.
- [48] Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. DNA interrogation by the CRISPR RNA-guided endonuclease cas9. *Nature*, 2014, 507(7490): 62–67.
- [49] Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, Regev A, Lander ES, Koonin EV, Zhang F. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*, 2016, 353(6299): aaf5573.
- [50] Zeng H, Wen SS, Xu W, He ZR, Zhai GF, Liu YK, Deng ZX, Sun YH. Highly efficient editing of the actinorhodin polyketide chain length factor gene in *Streptomyces coelicolor* M145 using CRISPR/Cas9-CodA(sm) combined system. *Applied Microbiology and Biotechnology*, 2015, 99(24): 10575–10585.
- [51] Bassalo MC, Garst AD, Halweg-Edwards AL, Grau WC, Domaille DW, Mutalik VK, Arkin AP, Gill RT. Rapid and efficient one-step metabolic pathway integration in *E. coli*. *ACS Synthetic Biology*, 2016, 5(7): 561–568.
- [52] Pyne ME, Bruder MR, Moo-Young M, Chung DA, Chou CP. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Scientific Reports*, 2016, 6: 25666.
- [53] Cleto S, Jensen JV, Wendisch VF, Lu TK. *Corynebacterium glutamicum* metabolic engineering with CRISPR interference (CRISPRi). *ACS Synthetic Biology*, 2016, 5(5): 375–385.
- [54] Rath D, Amlinger L, Hoekzema M, Devulapally PR, Lundgren M. Efficient programmable gene silencing by Cascade. *Nucleic Acids Research*, 2014, 43(1): 237–246.
- [55] Bikard D, Jiang WY, Samai P, Hochschild A, Zhang F, Marraffini LA. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Research*, 2013, 41(15): 7429–7437.
- [56] Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley EG. Characterization and evolution of salmonella CRISPR-Cas systems. *Microbiology*, 2015, 161(2): 374–386.
- [57] Barrangou R, Dudley EG. CRISPR-based typing and next-generation tracking technologies. *Annual Review of Food Science and Technology*, 2016, 7: 395–411.

- [58] Gomaa AA, Klumpe HE, Luo ML, Selle K, Barrangou R, Beisel CL. Programmable removal of bacterial strains by use of genome-targeting CRISPR-Cas systems. *mBio*, 2013, 5(1): e00928–13.
- [59] Zetsche B, Heidenreich M, Mohanraju P, Fedorova I, Kneppers J, Degennaro EM, Winblad N, Choudhury SR, Abudayyeh OO, Gootenberg JS, Wu WY, Scott DA, Severinov K, van der Oost J, Zhang F. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nature Biotechnology*, 2017, 35(1): 31–34.

Research progress of CRISPR-Cas system in bacteria

Li Hu, Shi Chen*

School of Pharmaceutical Sciences, Wuhan University, Wuhan 430072, Hubei Province, China

Abstract: The discovery of CRISPR-Cas system breaks certain theory which adaptive immunity had been long time considered as unique characteristics of eukaryotic organism. CRISPR-Cas, a new adaptive immunity system widespread in bacteria and archaea, protects the host from invasion of exogenous nucleic acids by capturing and using the Cas protein and crRNA to resisting it when invading again. In recent years, CRISPR-Cas system has aroused extensive attention and exploration. From microbial perspective, this article introduces a brief overview on classification, mechanism and application which have achieved significant breakthrough to provide valuable reference for further research on CRISPR-Cas system and its application.

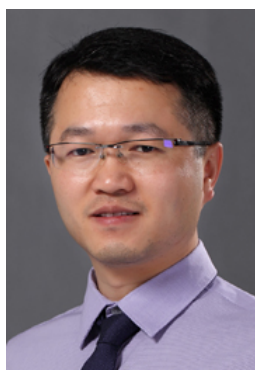
Keywords: CRISPR-Cas system, classification, mechanism, application in prokaryotic

(本文责编: 张晓丽)

Supported by the National Natural Science Foundation of China (31670086)

*Corresponding author. Tel/Fax: +86-27-68756643; E-mail: shichen@whu.edu.cn

Received: 4 April 2017; Revised: 1 June 2017; Published online: 9 June 2017



陈实, 武汉大学药学院, 楚天学者特聘教授、青年千人计划专家、湖北省特聘专家、博士生导师。入选湖北省医学领军人才培养工程。2000年毕业于华中农业大学生物技术本硕连读班。2004年博士毕业于上海交通大学; 2001–2002年赴韩国高等科技学院进行合作交流。2005–2007年在麻省理工学院进行博士后研究, 2007–2011年在哈佛大学开展研究。2011年回到武汉大学任教。任 *Microbiological Research* 和 *BMC Systems Biology* 副主编, 任 *Scientific Reports* 和 *Frontiers in Toxicogenomics* 编委。以通讯作者在 *PNAS* (2篇)、*Nature Communications*、*Nucleic Acids Research*、*Trends in Cell Biology*、*Medicinal Research Reviews*、*Chemistry & Biology*、*Protein & Cell* 等发表论文。2004年获上海市科技进步一等奖, 明治乳业生命科学奖。2007年获评《环球科学》年度全球十大科学新闻, 2008年获国家自然科学奖二等奖。3次获得新加坡陈嘉庚基金会陈嘉庚青少年发明奖(上海)。