

• 新技术新方法 •

# 新一代基因组编辑系统 CRISPR/Cpf1

杨帆<sup>1,2</sup>, 李寅<sup>1</sup>

1 中国科学院微生物研究所 中国科学院微生物生理与代谢工程重点实验室, 北京 100101

2 中国科学院大学, 北京 100049

杨帆, 李寅. 新一代基因组编辑系统 CRISPR/Cpf1. 生物工程学报, 2017, 33(3): 361-371.

Yang F, Li Y. The new generation tool for CRISPR genome editing: CRISPR/Cpf1. Chin J Biotech, 2017, 33(3): 361-371

**摘要:** CRISPR/Cas 系统几乎存在于所有的细菌和古菌中, 是用来抵御外来病毒和噬菌体入侵的获得性免疫防御机制。2012 年起 CRISPR/Cas9 被改造为基因编辑工具, 并衍生出一系列高效、便捷的基因编辑工具, 迅速在基础理论、基因诊断和临床治疗等研究领域中得到广泛应用。然而, CRISPR/Cas9 也存在细胞毒性、脱靶效应和基因插入困难等一些亟待解决的问题, 在一定程度上限制了 CRISPR/Cas9 的应用。Cpf1 是 2015 年报道的一种新型 CRISPR 效应蛋白, 具有许多与 Cas9 不同的特性, 有利于克服 CRISPR/Cas9 应用中的一些限制。本文综述了近两年来对 CRISPR/Cpf1 的研究进展和应用, 并对其应用前景和发展方向进行了展望。

**关键词:** Cpf1, Cas9, CRISPR, 基因编辑

## The new generation tool for CRISPR genome editing: CRISPR/Cpf1

Fan Yang<sup>1,2</sup>, and Yin Li<sup>1</sup>

1 CAS Key Laboratory of Microbial Physiological and Metabolic Engineering, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China

2 University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract:** Almost all archaea and many bacteria achieve adaptive immunity through a diverse set of CRISPR-Cas

**Received:** January 28, 2017; **Accepted:** February 13, 2017

**Supported by:** Key Research Program of the Chinese Academy of Sciences (No. ZDRW-ZS-2016-3), National Natural Science Foundation of China (No. 31470231).

**Corresponding author:** Yin Li. Tel: +86-10-64807485; E-mail: yli@im.ac.cn

中国科学院重点部署项目 (No. ZDRW-ZS-2016-3), 国家自然科学基金 (No. 31470231) 资助。

网络出版时间: 2017-03-01

网络出版地址: <http://kns.cnki.net/kcms/detail/11.1998.Q.20170301.1110.001.html>

systems. From 2012, Cas9 has been harnessed by thousands of laboratories for genome editing applications in a variety of experimental model systems. Cas9 is driving innovative applications from basic biology to biotechnology and medicine. However, numerous challenges still remain, such as limited targeting range, toxicity, and potential for off-target mutagenesis. Cpf1 represents a class 2/type V CRISPR RNA-guided endonuclease that is distinct from the type II CRISPR Cas9 nuclease. Future exploration of strategies using Cpf1 is expected to bring solutions for some of the biggest challenges facing genome editing. Here, we review the development and applications of Cpf1 for a variety of research or applications, and highlight challenges as well as future directions.

**Keywords:** Cpf1, Cas9, CRISPR, genome editing

CRISPR/Cas 系统 (Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein) 是细菌和古菌在抵御外来病毒和噬菌体入侵的过程中不断进化而来的一种获得性免疫防御机制。CRISPR 的研究始于 1987 年, 于 2002 年被正式命名, 至 2011 年其获得性免疫防御机制基础研究清楚<sup>[1]</sup>。2012 年, Jinek 等证实了 Cas9 可以在体外条件下与 DNA 靶序列的特定位点结合并进行剪切, 由此揭开了将 CRISPR/Cas9 改造为基因编辑工具的序幕<sup>[1-2]</sup>。

人类基因组计划使人们能够读取“生命代码”, CRISPR/Cas9 及其衍生技术则是人类改写生命代码的“神笔”。近 5 年来, CRISPR/Cas9 技术为生物领域带来了巨大的冲击, 2013 年被 *Science* 杂志评为十大科学突破之一, 2015 年更是荣登榜首。利用 CRISPR/Cas9 技术, 科学家们能够高效、精确地对 DNA 序列进行修剪、替换或添加, 可以快速地实现微生物基因组编辑、动植物的品种优化、动物模型构建, 甚至推动疾病治疗的颠覆性革命<sup>[1,3-4]</sup>。

然而, 随着研究的不断深入, CRISPR/Cas9 也暴露了它的缺陷和局限性, 例如严重的脱靶效应<sup>[1,5-6]</sup>。如何改造 CRISPR/Cas9, 降低其脱靶效应, 提高基因编辑精确度, 是目前科学家们关注和研究的重点<sup>[7]</sup>。

与此同时, 科学家们也致力于寻找新的基因

编辑系统。2015 年, 麻省理工学院 (Massachusetts Institute of Technology, MIT) 的张峰小组报道了一种新的 2 类 V 型 CRISPR 效应蛋白 Cpf1, 是在单链向导 RNA 引导下与靶 DNA 特定位点结合并切割的核酸内切酶<sup>[8]</sup>。

与 Cas9 相比, Cpf1 只需一段 42-44 个核苷酸组成的单链 RNA 即可识别和剪切 DNA, 从而简化了实验设计步骤, 更有利于多基因编辑; Cpf1 能够识别富含胸腺嘧啶 (T) 的 PAM 序列 (Protospacer adjacent motif), 可以扩展 CRISPR 的编辑范围; 此外, Cpf1 剪切会产生黏性末端, 可促进目的基因通过非同源重组的方式插入靶定位点。CRISPR/Cpf1 的开发有利于突破和克服 CRISPR/Cas9 应用中的一些限制<sup>[9]</sup>, 因此被称为是新一代的 CRISPR 基因组编辑工具。

## 1 CRISPR 系统的分类和作用机制

CRISPR/Cas 系统存在于几乎所有的古菌和大多数细菌中<sup>[10]</sup>。CRISPR/Cas 基因簇由一系列 Cas 蛋白 (Cas1、Cas2、Cas4 和效应蛋白如 Cas9、Cpf1 等) 的编码基因和一段 CRISPR 序列组成, 后者由一段前导序列、许多重复序列和间隔序列顺序排列组成。

根据 Cas 基因的组成和效应蛋白的数量, CRISPR 被分为了 2 类 5 型, 共 16 种亚型<sup>[11-12]</sup>。

1 类为利用多个效应蛋白复合物干扰靶基因的 CRISPR/Cas 系统,包括 I、III 和 IV 型;2 类为利用单一的效应蛋白干扰靶基因的 CRISPR/Cas 系统,包括 II 型和 V 型。目前研究得最为清楚的 CRISPR/Cas9 为 2 类 II 型 CRISPR 系统,2015 年张锋小组新发现的 CRISPR/Cpf1 属于 2 类 V 型 CRISPR 系统。

无论哪一类型的 CRISPR 系统,其抵御外来遗传物质入侵的免疫防御过程都可分为 3 个阶段<sup>[13]</sup>:

第一为适应阶段,首次入侵的外源 DNA 的前间隔序列 (Protospacer) 被古菌或细菌中的 Cas 蛋白获取,并作为间隔序列 (Spacer) 插入 CRISPR 中的两段重复序列之间。

第二为表达阶段,外源 DNA 再次入侵时,细菌开始转录 CRISPR,形成初级转录产物 pre-crRNA,再由核糖核酸酶或 Cas 蛋白在重复序列位点内切割形成成熟的 crRNA。

第三为干扰阶段,成熟的 crRNA 与特异的 CRISPR 效应蛋白形成核糖核蛋白复合体,识别并切割能与 crRNA 互补配对的外源 DNA,造成双链断裂,激活细胞的非同源末端连接 (Non-homologous end joining, NHEJ) 或同源重组 (Homologous recombination, HR) 两种修复机制,从而实现基因的敲除、插入或修饰。

## 2 Cpf1 与 Cas9 的比较

Cpf1 首先在新凶手弗朗西斯菌 *Francisella novicida* 和普雷沃菌 *Prevotella* 等微生物中被发现和 Cas9 同时存在,根据其基因组成及作用机制,CRISPR/Cpf1 被确定为 2 类 V 型 CRISPR 系统<sup>[8]</sup>。

然而,CRISPR/Cpf1 的作用机制与同为 2

类的 CRISPR/Cas9 系统有很大的不同 (表 1)。

第一,Cpf1 的分子量比 Cas9 小。目前广泛使用的 Cas9 来源于酿脓链球菌 *Streptococcus pyogenes* (下文简称为 spCas9),含有 1 368 个氨基酸;而新凶手弗朗西斯菌 *Francisella novicida* U112、氨基酸球菌 *Acidaminococcus* sp. BV3L6 和毛螺科菌 *Lachnospiraceae bacterium* ND2006 来源的 Cpf1 (下文分别简称为 FnCpf1、AsCpf1 和 LbCpf1) 分别含有 1 300、1 307 和 1 228 个氨基酸。

第二,Cas9 和 Cpf1 均是单链 RNA 指导的核酸内切酶,但 Cas9 只具有 DNA 内切酶活性,而 Cpf1 同时具有 DNA 和 RNA 内切酶活性。

第三,CRISPR/Cas9 初级转录产物 pre-crRNA 的成熟需要反式激活 crRNA (Trans-activating crRNA, tracrRNA) 与 crRNA 互补配对,在 Cas9 存在的条件下,由 RNase III 加工完成。而 CRISPR/Cpf1 的初级转录产物 pre-crRNA 由 Cpf1 自身加工,不需要 tracrRNA 的参与<sup>[8,14]</sup>。

第四,Cas9 与由 crRNA 和 tracrRNA 配对形成的 RNA 二聚体结合,形成 Cas9-crRNA-tracrRNA 核糖核蛋白复合体,识别和切割靶基因序列。将 crRNA 和 tracrRNA 融合成一条 sgRNA (Single-guide RNA),也可以引导 Cas9 切割 DNA;sgRNA 长度一般大于 100 nt,靶定不同的基因需要设计不同的 sgRNA。而 Cpf1 与成熟 crRNA 形成 Cpf1-crRNA 核糖核蛋白二元复合体,即可实现对靶基因的识别和剪切;Cpf1 剪切 DNA 所需的 crRNA 长度仅为 42-44 nt,只需要替换 CRISPR 中的间隔序列就可以靶定不同基因,实现多个基因的同时编辑,因此设计步骤可以显著简化,实验成本也可以显著降低,同时降低了将 CRISPR/Cpf1 系统递送至编辑对

象的难度。

第五, Cas9 识别位于靶序列的 3'端的富含胞嘧啶 (G) 的 PAM 序列 (5'-NGG-3'), 而 Cpf1 识别位于靶序列 5'端富含胸腺嘧啶 (T) 的 PAM 序列。另外, 不同来源的 Cpf1 识别的 PAM 序列也有所不同: FnCpf1 识别的 PAM 序列为 5'-YTN-3', AsCpf1 和 LbCpf1 识别的 PAM 序列为 5'-TTTN-3'。Cpf1 识别的 PAM 序列与 Cas9 不同, 极大地拓宽了 CRISPR 系统基因组编辑的靶点范围, 尤其是富含 AT 的基因组。

第六, Cas9 剪切位点离 PAM 序列很近, 在其上游第 3 个核苷酸处进行切割, 若发生 NHEJ 修复, 造成的核苷酸插入或缺失会改变 PAM 临近序列, 导致 Cas9 无法再次识别和切割靶位点, 从而阻碍同源重组修复在靶位点引入正确的基因编辑; 而 Cpf1 的剪切位点离 PAM 序列较远, 在靶 DNA 链的 PAM 序列下游第 23 位核苷酸和

非靶 DNA 链的第 18 位核苷酸处进行切割, NHEJ 修复造成的核苷酸插入或缺失, 不会改变 PAM 临近序列, Cpf1 仍然可以识别和切割靶基因, 同源重组修复依然可以在靶位点引入正确的基因编辑, 从而提高了 CRISPR 系统的基因编辑效率, 也便于对同一位点进行多轮的基因编辑。

第七, Cas9 切割 DNA 形成一个平末端; 而 Cpf1 切割 DNA 形成一个有 5 个核苷酸突出的黏性末端。利用 CRISPR/Cas9 进行基因插入时, 需要在插入基因的两端添加较长的同源臂, 通过同源重组的方式将其插入靶位点; 而利用 CRISPR/Cpf1 进行基因插入时, 若在插入基因的两端添加与靶位点对应的黏性末端, 就可以使插入基因通过 NHEJ 修复以可控的方向插入靶位点, 而不必依赖于同源重组。因此, 对于同源重组发生概率较低的编辑对象, CRISPR/Cpf1 是有利的基因编辑工具。

表 1 Cpf1 和 Cas9 的比较

Table 1 The comparison between Cpf1 and Cas9

	Cpf1	Cas9
蛋白大小	含 1 200 至 1 300 个氨基酸 <sup>[8]</sup> , 分子量较小	含 1 368 个氨基酸 <sup>[15]</sup> , 分子量较大
蛋白结构	含有 RuvC 结构域和一个推定的核酸酶结构域, 不含 HNH 结构域 <sup>[18]</sup>	含有 RuvC 结构域和 HNH 结构域 <sup>[15]</sup>
性质	同时具有 DNA 和 RNA 内切酶活性 <sup>[8,14]</sup>	只具有 DNA 内切酶活性 <sup>[1]</sup>
Pre-crRNA 成熟的加工因子	不需要 tracrRNA, 由 Cpf1 加工成熟 <sup>[8,14]</sup>	crRNA 和 tracrRNA 互补形成 RNA 二聚体, 在 Cas9 蛋白存在的条件下, 由 Rnase III 加工成熟 <sup>[13]</sup>
DNA 剪切复合体	Cpf1-crRNA 核糖核蛋白复合体, crRNA 长度为 42-44 nt <sup>[8]</sup>	Cas9-crRNA-tracrRNA 核糖核蛋白复合体 (或 sgRNA-Cas9 核糖核蛋白复合体, sgRNA 长度一般大于 100 nt) <sup>[27]</sup>
DNA 识别位点	前间隔序列 5'端的 PAM 序列 <sup>[8]</sup>	前间隔序列 3'端的 PAM 序列 <sup>[1,8]</sup>
PAM 序列	T-rich 序列, 5'-YTN-3', 5'-TTTN-3' <sup>[8]</sup>	G-rich 序列, 5'-NGG-3' <sup>[1,8]</sup>
剪切位点	PAM 序列下游靶 DNA 链的第 23 位核苷酸和非靶 DNA 链的 18 位核苷酸处 <sup>[8]</sup>	PAM 序列上游第 3 位核苷酸外侧 <sup>[4]</sup>
形成末端	5 个核苷酸突出的黏性末端 <sup>[8]</sup>	平末端 <sup>[1]</sup>

### 3 Cpf1 的结构生物学研究

2014年,张峰小组解析了 Cas9-sgRNA-DNA 三元复合体的晶体结构,揭示了 Cas9 识别和切割靶向 DNA 的分子机制,并为 Cas9 的改造和优化提供了充分的依据<sup>[15]</sup>。

Cpf1 呈现与 Cas9 不同的剪切特性,如具有 RNA 内切酶活性、识别富含胸腺嘧啶 (T) 的 PAM 序列、切割产生黏性末端等。为了研究 Cpf1 识别和剪切 RNA 以及 DNA 的分子机制,研究人员对 Cpf1 核糖核蛋白的晶体结构进行了解析,并比较了 Cpf1 与 Cas9 识别和剪切目标基因机制的异同。

#### 3.1 Cpf1-crRNA 二元复合物的晶体结构解析

2016年4月,哈尔滨工业大学的黄志伟教授课题组成功解析了 LbCpf1-crRNA 核糖核蛋白复合物的晶体结构,揭示了 Cpf1 识别 crRNA 和剪切 pre-crRNA 的分子机制<sup>[16]</sup>。

与 Cas9 相似,LbCpf1 与 crRNA 结合形成的二元复合物也为二裂片结构,外观呈三角形,中心形成一个带正电荷的通道。LbCpf1 在没有 crRNA 结合的状态下处于松散的构象,crRNA 的结合并不诱导 LbCpf1 寡聚化,而是引发 Cpf1 发生显著的构象变化,转变为紧凑的三角形结构,使得 crRNA 可以和靶基因序列在蛋白中心的正电荷通道内形成异源双链核酸分子。另外,该晶体结构显示 LbCpf1 的 RNA 酶活性与其 H843、K852 以及 K869 的 3 个氨基酸有关。

然而,sgRNA 与 crRNA 长度的差异,导致 Cas9 和 sgRNA 的结合方式与 Cpf1 和 crRNA 的结合方式有很大差别。除去与靶 DNA 互补配对的向导序列,sgRNA 其余部分的长度是 crRNA 重复序列的 3.5-4 倍,并含有多个结构模块,

sgRNA 以伸展的构象与 Cas9 结合,扩大了 sgRNA 和 Cas9 的接触面积,增加了与 Cas9 之间的相互作用,以维持 Cas9 多个结构域的稳定结构。crRNA 则呈一种高度扭曲的构象,被 LbCpf1 蛋白中部的寡聚核苷酸结合结构域 (Oligonucleotide binding domain, OBD) 识别。crRNA 的长度较短,Cpf1 的多个结构域不与 crRNA 相互作用,在结合目标 DNA 时依然可以发生空间位移。

因此,Cpf1-crRNA 二元复合物结构的解析,也为研究 crRNA 如何诱导 Cpf1 识别和结合 DNA 底物提供了一定的依据。进一步对 Cpf1-crRNA-DNA 三元复合物结构的解析表明,Cpf1 和 crRNA、Cas9 和 sgRNA 不同的结合方式也导致了两者 DNA 底物识别机制的差异。

#### 3.2 Cpf1-crRNA-targeted DNA 三元复合物的晶体结构解析

为了阐述 Cpf1 如何识别和剪切 DNA,MIT 的张峰研究组和东京大学 Osamu Nureki 教授、中国科学院生物物理研究所的高璞研究组与纪念斯隆-凯特琳癌症中心的研究人员先后解析了 AsCpf1-crRNA-DNA 三元复合物晶体结构<sup>[17-18]</sup>。

Cpf1 和 Cas9 的核糖核蛋白三元复合物具有相似的整体结构。AsCpf1 与 Cas9 同样采用二裂片结构,外观呈蟹钳状,分为识别叶 (Recognition lobe, REC) 和核酸酶叶 (Nuclease lobe, NUC)。但不同的是,Cas9 核酸酶叶包含 RuvC 和 HNH 两个核酸酶结构域,分别负责切割非靶向和靶向 DNA 链,产生平末端;而 AsCpf1 的 NUC 叶没有 HNH 结构域,由 RuvC 结构域和一个推定的新的核酸酶结构域 (Nuc) 组成,分别负责切割非靶向及靶向 DNA 链,产生黏性末端。RuvC 剪切非靶向 DNA 链是 Nuc 剪切靶向 DNA 链的

先决条件。

另外, Cpf1-crRNA 二元复合物和 Cpf1-crRNA-DNA 三元复合物的结构的比较也揭示了 Cpf1 与 Cas9 不同的 DNA 识别机制。

首先, sgRNA 和 crRNA 上靠近 PAM 序列的区域均存在一段“种子序列”, 可与靶 DNA 序列互补配对, 对于 Cas9 和 Cpf1 识别和切割 DNA 至关重要。由于 sgRNA 和 crRNA 在二元复合体中呈现不同的构象, “种子序列”在 Cas9-sgRNA 二元复合体中形成有序的 A 型结构, 而在 Cpf1-crRNA 二元复合体中是无序的, 在 Cpf1-crRNA-DNA 三元复合体中才转变为有序的 A 型结构。

其次, PAM 互作裂缝在 Cas9-sgRNA 二元复合体中也已预先形成, 是 Cas9 识别目标 DNA 的另一个关键结构, 但在 Cpf1-crRNA 二元复合体中还未形成, 目标 DNA 的结合诱导 Cpf1 内部结构重排, 多个结构域发生空间位移, PAM 互作裂口经历一个“从开到关”的构象改变, 以容纳有序的 A 型种子序列和靶 DNA 在正电荷通道内形成异源双链核酸分子。

Cpf1 识别目标 DNA 时种子序列“从无序到有序”和 PAM 互作裂缝“从开到关”的构象变化, 可能有利于降低其基因编辑的脱靶效应。

## 4 CRISPR/ Cpf1 的应用研究

### 4.1 CRISPR/Cpf1 在微生物基因编辑中的应用

CRISPR/Cas9 在大肠杆菌<sup>[19-20]</sup>、酵母<sup>[21]</sup>等常用的微生物模式菌株中的应用已经趋于成熟, 通过设计一段可与靶基因区域 PAM 序列上游 20 个核苷酸互补配对的向导序列 (N20), 就可以对宿主的任意基因进行编辑。蓝细菌是一

种光合自养的原核生物, 近年来越来越多地应用于光合固碳和高附加值化学品生物合成的研究<sup>[22-24]</sup>。然而, 和大肠杆菌、酵母等微生物相比, 蓝细菌的基因编辑工具单一, 主要依靠其低概率的自发同源重组, 且由于蓝细菌的基因组拷贝数较高, 插入基因也难以稳定存在。

为了改变这一现状, 研究学者开始尝试将 CRISPR/Cas 系统改造为蓝细菌适用的精确、高效的基因编辑工具。最近, CRISPR/Cas9 被成功用于蓝细菌模式菌株细长聚球藻 *Synechococcus elongatus* PCC 7942 和 UTEX 2973 中<sup>[25-26]</sup>。然而, 研究发现 Cas9 对 *S. elongatus* PCC 7942 和 UTEX 2973 均存在不同程度的细胞毒性, 推测是由 Cas9 在非靶位点切割 DNA 导致双链断裂而引起的细胞死亡。

2016 年 12 月, Pakrasi 研究组利用 CRISPR/Cpf1 对蓝细菌进行基因编辑<sup>[27]</sup>。作者将 FnCpf1 蛋白及靶定目标基因的 crRNA 克隆在广谱宿主质粒 RSF1010 上, 成功地对 *S. elongatus* UTEX 2973、集胞藻 *Synechocystis* sp. PCC 6803 和鱼腥藻 *Anabaena* 7120 三种不同类型的蓝细菌基因组进行了基因的敲除、插入和点突变。作者发现, FnCpf1 对蓝细菌的细胞毒性比 Cas9 低得多, 因此对蓝细菌基因编辑而言, CRISPR/Cpf1 的潜力更大。

### 4.2 CRISPR/Cpf1 在植物基因编辑中的应用

目前, CRISPR/Cas9 在拟南芥、烟草、小麦和玉米等多种植物基因组编辑中都有广泛的应用<sup>[28-31]</sup>, 也逐渐应用于植物基因表达调控、表观基因组学和遗传育种等领域。尽管构建的 Cas9 突变体能够识别 5'-NGA-3'、5'-NGCG-3' 等非典型的 PAM 序列, 但是 CRISPR/Cas9 可编

辑的植物基因靶点范围,依然局限于富含鸟嘌呤(G)的基因位点<sup>[32]</sup>。由于Cpf1识别的PAM序列与Cas9不同,CRISPR/Cpf1可以进一步地扩展对植物基因组的编辑能力。

2016年12月,3个研究组先后利用CRISPR/Cpf1完成了对植物基因组的编辑<sup>[33-35]</sup>。日本学者Toki等证明了FnCpf1可成功地在烟草和水稻基因组中引入可遗传的突变,且在水稻中引发的突变效率比在烟草中要高<sup>[35]</sup>。中国农业科学院的王克剑小组和安徽省农业科学院的杨建波小组研究发现LbCpf1能够对水稻基因组进行有效编辑,但突变效率与靶基因的核苷酸组成有一定的关系,而AsCpf1则不能完成对水稻基因组编辑<sup>[33-34]</sup>。与Cas9一般造成1到2个核苷酸的插入和缺失不同,Cpf1在所有靶位点均产生6到38个核苷酸片段的缺失,因此Cpf1可能更适用于植物基因组中较长核苷酸片段的敲除。

有趣的是,Toki和杨建波小组的研究均发现,利用CRISPR/Cpf1得到的烟草和水稻突变株都存在杂合子或嵌合体,表明CRISPR/Cpf1诱导的基因突变可能只发生在1个等位基因上;或在胚胎细胞第一次分裂之后,只发生在部分体细胞中。在CRISPR/Cas9对拟南芥进行基因组编辑中也报道过类似的结果,研究者发现用组织特异性启动子表达Cas9可以减少嵌合体的产生。因此,利用诱导型或组织特异性启动子在时间和空间上控制Cpf1的表达,也许能够减少杂合子或嵌合体的产生。

另外,上述3项研究都对Cpf1在水稻中潜在的脱靶效应进行了评估。Toki等发现FnCpf1在烟草和水稻中都存在脱靶效应,但脱靶位点的基因突变频率为0-6.25%,比靶位点(21.4%-23.3%)低得多。然而,中国的两个研究组均没有检测到

LbCpf1在水稻基因组的脱靶位点处产生的突变。这可能与LbCpf1识别的PAM序列(5'-TTTN-3')比FnCpf1的PAM序列(5'-TTN-3')更长有关。PAM序列越长,基因组上潜在的脱靶位点则更少,Cpf1识别靶基因的准确度也越高。

综上所述,CRISPR/Cpf1是除CRISPR/Cas9以外可供选择的一种有效的植物基因组编辑工具,且脱靶效应低。尽管如此,CRISPR/Cpf1仍然存在突变效率较低、形成杂合子或嵌合体<sup>[35]</sup>等亟待解决的问题,还需要进一步地改良和优化,以提高CRISPR/Cpf1对植物基因组的编辑能力。

#### 4.3 CRISPR/Cpf1在动物基因编辑中的应用

CRISPR/Cas系统在动物研究中有广泛的应用。目前,研究者们利用CRISPR/Cas9成功地在人体、小鼠、果蝇和羊等动物细胞中实现了基因编辑<sup>[36-39]</sup>和基因调控<sup>[40]</sup>,在动物品种改良方面具有极大的潜力。更令人惊喜的是,越来越多的研究成果揭示了CRISPR技术在生物医学领域的广阔应用前景和无限可能性。例如,成功地构建小鼠癌症模型<sup>[41]</sup>、重现肿瘤染色体易位<sup>[42]</sup>、实现了将艾滋病毒从人类基因组中完全剔除<sup>[37]</sup>等。

CRISPR系统在动物研究中的应用仍存在一定问题。首先,CRISPR/Cas9存在严重的脱靶效应<sup>[1,6-7]</sup>,因为Cas9的特异性仅依赖于sgRNA上靠近PAM序列的10-12个核苷酸与靶基因的互补配对,而远离PAM序列的其余8-10个核苷酸发生碱基错配对靶位点的识别没有明显的影响。其次,Cas9的运送受限于病毒载体的容量大小,往往需要共转多个表达载体来实现对多个基因的编辑,增加了将Cas9导入编辑

对象的难度<sup>[1,43]</sup>。

2015年,张峰研究组在16种不同来源的Cpf1蛋白中发现AsCpf1和LbCpf1两种可在人体细胞中引入片段的插入和缺失<sup>[8]</sup>。这两种Cpf1蛋白一经发现,便被迅速地应用于动物基因编辑中<sup>[44-46]</sup>。

基因敲除小鼠模型是研究基因功能、疾病发生机理、寻找合适药物靶标的重要工具。2016年,韩国的两个研究组分别将Cpf1 mRNA及其相对应的crRNA<sup>[44]</sup>、预组装的Cpf1核糖核蛋白<sup>[45]</sup>(RNP)通过显微注射法导入小鼠胚胎中,成功获得了基因突变的小鼠,突变效率与SpCas9相当,且均没有检测到脱靶效应。

Cpf1在动物中也存在产生嵌合体的现象。嵌合体动物必须通过远交才能产生纯合的、无嵌合性的突变个体,需要花费数月或更长的时间。与植物一样,动物的嵌合性也是由Cpf1诱导的细胞突变发生在胚胎细胞第一次分裂之后导致的,因此Cpf1在动植物中发挥功能可能需要较长的时间。利用CRISPR/Cas9直接编辑无性系的供体细胞,再通过体细胞核移植到无核卵母细胞中,产生的个体嵌合现象可大幅减少;或者利用CRISPR/Cas9直接在供体精子发生干细胞中进行基因编辑,也可避免胚胎发育的全能性和多能性状态,从而消除嵌合突变体后代的产生<sup>[47]</sup>。减少CRISPR/Cpf1产生的嵌合现象也可采用与CRISPR/Cas9类似的方法。

2016年,有两项研究先后评价了Cpf1在人体细胞中的脱靶效应,并与Cas9进行了比较<sup>[48-49]</sup>。首先,Cpf1对crRNA与靶DNA之间的碱基错配比Cas9更敏感。研究发现Cpf1对crRNA的间隔序列(Spacer)中靠近PAM序列的第1至18位核苷酸的碱基错配非常敏感,发生双碱基

错配即可导致Cpf1完全失去剪切活性,对远离PAM序列的第19至23位核苷酸的碱基错配具有较高的容忍度;而Cas9与靶DNA识别的特异性依赖于靠近PAM序列的前10-12个核苷酸,而离PAM序列较远的其余8-10个核苷酸的碱基错配对靶位点的识别无显著影响。其次,Cpf1的脱靶效应比Cas9更低。研究人员利用Digenome-seq和GUIDE-seq两种方法分别在体外和体内条件下鉴定了LbCpf1和AsCpf1在人体基因组上的脱靶位点和突变频率,结果均表明LbCpf1和AsCpf1在人体基因组上的潜在脱靶位点比spCas9更少,且脱靶位点的基因突变频率也低于1%,远低于靶位点的突变频率。因此,Cpf1的脱靶效应比Cas9要小,对人类基因组的编辑具有更高的特异性。

如前所述,Cpf1在多基因编辑方面比Cas9更具优势。2016年,张峰研究组将CRISPR序列中的间隔序列分别替换成靶定不同目的基因的向导序列,与AsCpf1蛋白表达在同一个慢性病毒载体上,实现了对人体细胞中4个基因和小鼠原代皮层神经细胞中3个基因的同时编辑<sup>[43]</sup>。

综上,CRISPR/Cpf1是一种更有效的动物基因编辑工具,比CRISPR/Cas9更加适用于精确的基因组编辑。然而与CRISPR/Cas9一样,CRISPR/Cpf1在动物中也存在产生嵌合体的问题,需要进一步研究并加以解决。

## 5 展望

由于Cpf1相对于Cas9的诸多优势,CRRISPR/Cpf1的应用正在不断扩展。CRISPR/Cpf1在水稻、小鼠和人体细胞基因编辑的应用中都基本没有脱靶效应,证明了CRRISPR-Cpf1基因编辑的高度特异性,比CRISPR/Cas9更适用



于精确的基因组编辑,显示了 CRISPR/Cpf1 在疾病模型建立、抗癌药物研制、干细胞治疗等生物和医学领域的应用潜力。

CRISPR/Cpf1 依然存在诸多待解决的问题,例如,在植物和动物中会产生嵌合突变体,突变效率可能与靶基因序列的核苷酸组成有关等。如何提高 CRISPR/Cpf1 的基因编辑效率?如何扩大编辑对象和靶基因位点的范围?如何进一步提高基因编辑的精确度?如何在基因组水平上进行应用?如何将 CRISPR/Cpf1 技术应用于疾病治疗?以及由此引发的安全性和伦理问题,也都需要更加深入的研究和规则的制定加以解决。

CRISPR/Cpf1 是生物学领域非常有价值的新工具,对它的研究也才刚刚开始,仍有许多发展、改造和利用的空间。相信在不久的将来,CRISPR/Cpf1 及其衍生技术能够和 CRISPR/Cas9 一起,为科学研究和生命健康领域提供强大的和多样化的基因编辑工具。

致 谢: 感谢赵同心斧正文章手稿。

## REFERENCES

- [1] Hsu PD, Lander ES, Zhang F. Development and applications of CRISPR-Cas9 for genome engineering. *Cell*, 2014, 157(6): 1262–1278.
- [2] Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012, 337(6096): 816–821.
- [3] Ran FA. Adaptation of CRISPR nucleases for eukaryotic applications. *Analyt Biochem*, 2016, doi: 10.1016/j.ab.2016.10.018. (in Press)
- [4] Hung SSC, McCaughey T, Swann O, et al. Genome engineering in ophthalmology: application of CRISPR/Cas to the treatment of eye disease. *Prog Retin Eye Res*, 2016, 53: 1–20.
- [5] Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*, 2013, 8(11): 2281–2308.
- [6] Pattanayak V, Lin S, Guilinger JP, et al. High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat Biotechnol*, 2013, 31(9): 839–843.
- [7] Haeussler M, Concordet JP. Genome editing with CRISPR-Cas9: can it get any better?. *J Genet Genomics*, 2016, 43(5): 239–250.
- [8] Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*, 2015, 163(3): 759–771.
- [9] Fagerlund RD, Staals RHJ, Fineran PC. The Cpf1 CRISPR-Cas protein expands genome-editing tools. *Genome Biol*, 2015, 16(1): 251–253.
- [10] Mojica FJM, Díez-Villaseñor C, Soria E, et al. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol Microbiol*, 2000, 36(1): 244–246.
- [11] Makarova KS, Wolf YI, Alkhnbashi OS, et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol*, 2015, 13(11): 722–736.
- [12] Makarova KS, Haft DH, Barrangou R, et al. Evolution and classification of the CRISPR-Cas systems. *Nat Rev Microbiol*, 2011, 9(6): 467–477.
- [13] Carter J, Wiedenheft B. SnapShot: CRISPR-RNA-guided adaptive immune systems. *Cell*, 2015, 163(1): 260–260.e1.
- [14] Fonfara I, Richter H, Bratovič M, et al. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*, 2016, 532(7600): 517–521.
- [15] Nishimasu H, Ran FA, Hsu PD, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell*, 2014, 156(5): 935–949.
- [16] Dong D, Ren K, Qiu XL, et al. The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature*,

- 2016, 532(7600): 522–526.
- [17] Yamano T, Nishimasu H, Zetsche B, et al. Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell*, 2016, 165(4): 949–962.
- [18] Gao P, Yang H, Rajashankar KR, et al. Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target DNA recognition. *Cell Res*, 2016, 26(8): 901–913.
- [19] Li YF, Lin ZQ, Huang C, et al. Metabolic engineering of *Escherichia coli* using CRISPR-Cas9 mediated genome editing. *Metab Eng*, 2015, 31: 13–21.
- [20] Jiang Y, Chen B, Duan CL, et al. Multigene editing in the *Escherichia coli* genome via the CRISPR-Cas9 system. *Appl Environ Microbiol*, 2015, 81(7): 2506–2514.
- [21] Horwitz AA, Walter JM, Schubert MG, et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in Yeasts via CRISPR-Cas. *Cell Syst*, 2015, 1(1): 88–96.
- [22] Zhou J, Zhu TC, Cai Z, et al. From cyanochemicals to cyanofactories: a review and perspective. *Microb Cell Fact*, 2016, 15(1): 2–10.
- [23] Zhou J, Zhang HF, Meng HK, et al. Production of optically pure D-lactate from CO<sub>2</sub> by blocking the PHB and acetate pathways and expressing D-lactate dehydrogenase in cyanobacterium *Synechocystis* sp. PCC 6803. *Proc Biochem*, 2014, 49(12): 2071–2077.
- [24] Zhou J, Zhang FL, Meng HK, et al. Introducing extra NADPH consumption ability significantly increases the photosynthetic efficiency and biomass production of cyanobacteria. *Metab Eng*, 2016, 38: 217–227.
- [25] Wendt KE, Ungerer J, Cobb RE, et al. CRISPR/Cas9 mediated targeted mutagenesis of the fast growing cyanobacterium *Synechococcus elongatus* UTEX 2973. *Microb Cell Fact*, 2016, 15(1): 115.
- [26] Li H, Shen CR, Huang CH, et al. CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. *Metab Eng*, 2016, 38: 293–302.
- [27] Ungerer J, Pakrasi HB. Cpf1 is a versatile tool for CRISPR genome editing across diverse species of cyanobacteria. *Sci Rep*, 2016, 6: 39681–39689.
- [28] Ma XL, Zhang QY, Zhu QL, et al. A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol Plant*, 2015, 8(8): 1274–1284.
- [29] Wang YP, Cheng X, Shan QW, et al. Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat Biotechnol*, 2014, 32(9): 947–951.
- [30] Zhao YP, Zhang CS, Liu WW, et al. An alternative strategy for targeted gene replacement in plants using a dual-sgRNA/Cas9 design. *Sci Rep*, 2016, 6: 23890.
- [31] Jiang WZ, Zhou HB, Bi HH, et al. Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in *Arabidopsis*, tobacco, sorghum and rice. *Nucleic Acids Res*, 2013, 41(20): e188.
- [32] Hu XX, Wang C, Fu YP, et al. Expanding the range of CRISPR/Cas9 genome editing in rice. *Mol Plant*, 2016, 9(6): 943–945.
- [33] Xu RF, Qin RY, Li H, et al. Generation of targeted mutant rice using a CRISPR-Cpf1 system. *Plant Biotechnol J*, 2016, doi: 10.1111/pbi.12669. (in Press)
- [34] Hu XX, Wang C, Liu Q, et al. Targeted mutagenesis in rice using CRISPR-Cpf1 system. *J Genet Genomics*, 2017, 44(1): 71–73.
- [35] Endo A, Masafumi M, Kaya H, et al. Efficient targeted mutagenesis of rice and tobacco genomes using Cpf1 from *Francisella novicida*. *Sci Rep*, 2016, 6: 38169–38177.
- [36] Yin H, Xue W, Chen SD, et al. Genome editing with Cas9 in adult mice corrects a disease mutation and phenotype. *Nat Biotechnol*, 2014, 32(6): 551–553.
- [37] Liao HK, Gu Y, Diaz A, et al. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat Commun*, 2015, 6: 6413.

- [38] Ren XJ, Yang ZH, Xu J, et al. Enhanced specificity and efficiency of the CRISPR/Cas9 system with optimized sgRNA parameters in *Drosophila*. *Cell Rep*, 2014, 9(3): 1151–1162.
- [39] Wu MM, Wei CH, Lian ZX, et al. *Rosa26*-targeted sheep gene knock-in via CRISPR-Cas9 system. *Sci Rep*, 2016, 6: 24360.
- [40] Didovyk A, Borek B, Tsimring L, et al. Transcriptional regulation with CRISPR-Cas9: principles, advances, and applications. *Curr Opin Biotechnol*, 2016, 40: 177–184.
- [41] Xue W, Chen SD, Yin H, et al. CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*, 2014, 514(7522): 380–384.
- [42] Torres R, Martin MC, Garcia A, et al. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR-Cas9 system. *Nat Commun*, 2014, 5: 3964–3971.
- [43] Zetsche B, Heidenreich M, Mohanraju P, et al. Multiplex gene editing by CRISPR-Cpf1 using a single crRNA array. *Nat Biotechnol*, 2017, 35(1): 31–34.
- [44] Kim Y, Cheong SA, Lee JG, et al. Generation of knockout mice by Cpf1-mediated gene targeting. *Nat Biotechnol*, 2016, 34(8): 808–810.
- [45] Hur JK, Kim K, Been KW, et al. Targeted mutagenesis in mice by electroporation of Cpf1 ribonucleoproteins. *Nat Biotechnol*, 2016, 34(8): 807–808.
- [46] Port F, Bullock SL. Augmenting CRISPR applications in *Drosophila* with tRNA-flanked sgRNAs. *Nat Methods*, 2016, 13(10): 852–854.
- [47] Chapman KM, Medrano GA, Jaichander P, et al. Targeted germline modifications in rats using CRISPR/Cas9 and spermatogonial stem cells. *Cell Rep*, 2015, 10(11): 1828–1835.
- [48] Kleinstiver BP, Tsai SQ, Prew MS, et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat Biotechnol*, 2016, 34(8): 869–874.
- [49] Kim D, Kim J, Hur JK, et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat Biotechnol*, 2016, 34(8): 863–868.

(本文责编 陈宏宇)

**李寅** 中国科学院微生物研究所研究员，《生物工程学报》副主编。主要从事微生物分子生理学与生物技术研究，在包括 *Science*、*Proc Natl Acad Sci USA* 在内的国际期刊上发表文章 120 余篇，H 因子为 31。授权专利 30 余项，部分成果已经在工业上获得应用。担任 *Biotechnol J*、*Microb Cell Fac*、*Microbiology* 等 5 种国际期刊的编委，发展中国家科学院青年科学家网络首届执委会主席，中国科学院-发展中国家科学院生物技术卓越中心主任，中国科学院国际化战略专家咨询委员会委员，获 2012 年度发展中国家科学院地区科技发展奖。

