

doi:10.7668/hbxb.2014.S1.023

# 植物基因功能研究新方法

彭 陈<sup>1</sup>,王 瑞<sup>2</sup>,李万昌<sup>3</sup>,郭士伟<sup>1</sup>

(1. 江苏省农业科学院 农业生物技术研究所,江苏省农业生物学重点实验室,江苏 南京 210014;  
2. 江苏省农业科技实业总公司,江苏 南京 210014;3. 河南师范大学 生命科学院,河南 新乡 453007)

**摘要:**TALEN(Transcription activator-like effector nuclease)和CRISPR/Cas系统(Clustered regularly interspaced short palindromic repeats,CRISPR. CRISPR-associated,Cas)是近年来发现的基因功能研究的新方法,且已经成功地在动物、植物及微生物等多个领域得到成功应用。它们可以对DNA双链进行切割,从而启动细胞内的DNA修复机制,以此来实现对靶位点的基因修饰。相较于传统方法,它们具有特异性高、操作简便及打靶效率高等特点。主要从结构、功能及作用机制等方面阐述这2种新基因功能的研究方法。

**关键词:**基因功能;TALEN;CRISPR/Cas系统

**中图分类号:**Q78;Q94 **文献标识码:**A **文章编号:**1000-7091(2014)增刊-0109-05

## New Method for Studying Plant Gene Function

PENG Chen<sup>1</sup>,WANG Rui<sup>2</sup>,LI Wan-chang<sup>3</sup>,GUO Shi-wei<sup>1</sup>

(1. Institute of Biotechnology,Jiangsu Academy of Agricultural Sciences,Provincial Key Laboratory of Agrobiolgy,Nanjing 210014,China;2. Jiangsu Agricultural Scientific Industry Co.,Nanjing 210014, China;3. College of Life Sciences,Henan Normal University,Xinxiang 453007,China)

**Abstract:**TALEN(Transcription activator-like effector nuclease) CRISPR/Cas system(Clustered regularly interspaced short palindromic repeats,CRISPR. CRISPR-associated,Cas) have recently emerged as new technique for targeted genome modifications. They have been successfully applied in many fields such as animals,plants and microorganism. They can generate a double-strand break(DSB). Repair of these DSBs by intracellular DNA repair mechanisms enables to achieve the target gene modification. Compared with traditional method,TALEN and CRISPR/Cas system were stronger specificity,manipulated more easily and higher targeting efficiency. Here we report the construction,function and mechanism of the TALEN and CRISPR/Cas system.

**Key words:**Gene function;TALEN;CRISPR/Cas system

随着现代测序技术和生物信息学工具的广泛应用,已经有越来越多物种的全基因组测序及功能预测工作相继完成。揭示基因在生物体生命活动中的作用,已成为生命科学领域研究的重点内容,这使得基因敲除技术得到了广泛应用,成为当前生物学研究的最热点与最前沿的工具。目前,常用的基因敲除技术主要包括同源重组、T-DNA或转座子插入、ZFN、RNAi等方法。基于同源重组原理的基因敲除技术能够准确地作用于靶位点,具有高度的特异性,但是其打靶效率较低;T-DNA或转座子插入具有随机性且筛选过程非常繁琐;ZFN技术代价昂贵,工作

繁琐,且对细胞的毒性较高;RNAi虽然是一种简单、高效、特异的基因功能分析方法,但是其作用也受到基因表达量的限制,且其干涉靶基因的同时也有可能干涉其同源基因从而造成基因沉默的表型难以鉴定。

本研究主要从结构、功能及作用机制等方面阐述了2种新的基因功能分析方法:TALENs和CRISPR/Cas系统。

## 1 TALENs

类转录激活因子效应物核酸酶(Transcription activator-like effector nuclease,TALEN)是近几年发

收稿日期:2014-10-09

基金项目:公益性行业(农业)科研专项(201303102)

作者简介:彭 陈(1988-),男,安徽合肥人,硕士,主要从事水稻真菌病害研究。

通讯作者:郭士伟(1971-),男,河南汝州人,副研究员,硕士,主要从事水稻真菌病害研究。

展起来的,可以对基因组进行定点改造的新技术,它是由植物病原菌黄单胞菌(*Xanthomonas*)分泌的一种类转录激活效应因子(Transcription activator-like effectors,TALEs)与 *Fok I* 核酸内切酶人工融合而成<sup>[1]</sup>。相对于 ZFN 技术,TALENs 具有特异性强、毒性低、打靶效率高以及操作简便等优势。自 2009 年,Moscou<sup>[2]</sup>和 Boch<sup>[3]</sup>揭示了 TALE 特异结合宿主细胞基因的机制以来,基于 TALE 蛋白的内切酶技术得到了迅猛发展,2011 年,Tesson 等<sup>[4]</sup>首先将 TALENs 成功运用于大鼠的基因敲除,从而在动物体内建立了运用 TALENs 技术进行基因敲除的新方法。此后,TALENs 技术被广泛应用于多种模式生物中。如:非洲爪蟾(*Xenopus laevis*)<sup>[5]</sup>、斑马鱼(*Danio rerio*)<sup>[6]</sup>、拟南芥(*Arabidopsis thaliana*)<sup>[7]</sup>、水稻(*Oryza sativa*)<sup>[8]</sup>、果蝇(*Drosophila*)<sup>[9]</sup>、线虫(*Caenorhabditis elegans*)<sup>[10]</sup>、裂殖酵母(*Saccharomyces cerevisiae*)等<sup>[11]</sup>。2013 年,Zhang 等<sup>[12]</sup>利用 TALENs 技术对烟草的乙酰乳酸合酶基因进行了敲除和插入试验,均成功获得了突变体,且效率分别高达 30% 和 14%,使得 TALENs 的应用更加广泛。江苏省农业生物学重点实验室也已经成功构建了稻瘟菌鞘氨醇激酶编码基因的 TALEN 载体,若能够成功获得该基因的敲除突变体,将更能丰富 TALENs 的应用范围。

### 1.1 TALEN 的结构与功能

天然的 TALE 蛋白都具有相似的结构域,主要包括:N-端的分泌信号;C-端的核定位信号(Nuclear localization signals,NLSs)和转录激活结构域(Acidic activation domain,AD);中间的 DNA 特异识别与结合结构域<sup>[13]</sup>。TALE 蛋白的 DNA 结合域都是由 33~35 个高度保守的氨基酸组成的重复单元串联而成。这些重复单元中除了第 12 和 13 位可变外,其余的氨基酸序列完全相同,这 2 个氨基酸被称为重复可变双残基(Repeat variable di-residue,RVD)。TALE 蛋白识别基因序列的机制就在于不同的 RVD 能够特异性地识别 4 种碱基中的 1 种或多种,其中最常见识别规律为:NI(Asn Ile)识别 A,NG(Asn Gly)识别 T,NN(Asn Asn)识别 G,HD(His Asp)识别 C<sup>[2-3]</sup>。Deng 等<sup>[14]</sup>通过蛋白晶体结构解析的方法进一步明确了 RVD 的功能,研究发现,只有第 13 位氨基酸具有真正的基因序列识别作用,而第 12 位主要起稳定作用。

TALE 蛋白可以和转录因子、表观遗传修饰酶等结合来实现基因转录水平的调控,而其最引人注目的应用就是用核酸内切酶 *Fok I* 替换 TALE 蛋白

C-端的 AD 形成的 TALENs。因为 *Fok I* 需要形成二聚体才能发挥作用<sup>[15]</sup>,所以 TALEN 也需要成对使用。当 TALENs 通过其 N-端的核定位信号进入细胞核内,并特异性地识别到靶位点上后,*Fok I* 即可形成有活性的二聚体并切割靶基因的 DNA 双链;DNA 双链断裂后,细胞本身将启动 2 种修复机制,即依赖同源重组的修复机制和非同源末端连接修复<sup>[16]</sup>。前者利用外源的同源序列,可以实现基因的插入,而后者在修复过程中往往导致基因突变,包括基因缺失、移码突变等,从而导致目的基因阅读框的改变,达到基因敲除的目的。

### 1.2 TALENs 构建方法

利用 TALEN 技术对基因组进行定点修饰的方法主要包括以下几个步骤。

选定目的基因上合适的 TALE 靶位点。一般选择目标基因外显子上的 2 段长度为 14~18 bp,且距离为 17~18 bp 的序列为靶位点。选择靶位点的原则主要有:①TALEN 识别位点的前 1 个碱基为 T;②如果是构建敲除载体,那么识别位点最好在目的基因的第 1 个外显子上<sup>[17]</sup>。

根据靶位点的 DNA 序列选择对应的 TALE 蛋白并与 *Fok I* 相连。这也是 TALENs 技术的最关键步骤。目前,人工组装 TALEN 的方法主要有以下 4 种:Gateway<sup>[18]</sup>、Golden Gate<sup>[19]</sup>、Toolbox<sup>[20]</sup> 以及 FLASH<sup>[17]</sup> 组装法。

导入受体细胞,筛选阳性克隆。通常可以通过借鉴制备转基因个体的方法,将 TALENs 质粒导入受体细胞,并通过 PCR 与测序相结合的方法筛选阳性克隆。

## 2 CRISPR/Cas 系统

CRISPR/Cas (Clustered regularly interspaced short palindromic repeats, CRISPR. CRISPR-associated, Cas) 系统是近年来发现的,广泛存在于细菌和古生菌中,针对噬菌体、质粒等外源 DNA 入侵的 1 种获得性免疫体系<sup>[21]</sup>。1 个完整的 CRISPR/Cas 系统主要由前导序列(Leader, L)、间隔序列(Spacer, S)与被间隔序列隔开的重复序列(Repeats, R)串联排列形成的 CRISPR 簇以及一系列 CRISPR 相关蛋白(Cas 蛋白)的编码基因构成<sup>[22]</sup>。前导序列是 1 段位于 CRISPR 簇上游,且紧邻第 1 个重复序列的不保守序列,具有启动子的功能<sup>[23]</sup>。重复序列的长度一般为 21~48 bp,并具有 5~7 个序列的回文序列,能够形成稳定的发卡结构,其重复次数可高达 250 次<sup>[24]</sup>。重复序列的 3' 末端具有高度保守的 GAAA

(C/G)基序,该结构可能是 Cas 蛋白的识别位点。每 2 个重复序列之间被间隔序列隔开,间隔序列一般来自噬菌体和质粒等外源 DNA,其长度与细菌种类和 CRISPR 位点相关。目前,已发现的 Cas 蛋白均鉴定出了核酸相关的功能域,如:核酸内切酶、核酸外切酶、RNA 和 DNA 结合域等。根据 *cas* 基因的保守程度,可将其分为核心 *cas* 基因,亚型特异性 *cas* 基因和 *RAMP* (Repeat-associated mysterious proteins) 基因<sup>[25]</sup>。

## 2.1 CRISPR/Cas 系统的作用机理

Lillestøl 等<sup>[23]</sup>研究发现,古生菌 CRISPR/Cas 系统中的间隔序列与外来噬菌体和质粒的一些基因片段具有高度同源性,Barrangou<sup>[22]</sup>和 Deveau<sup>[26]</sup>进一步研究表明该系统能够使细胞抵御外来噬菌体和质粒等的侵染。目前,虽然还不清楚 CRISPR/Cas 系统的具体作用机制,但也已明确了其作用的基本过程,主要可分为以下 3 个阶段。

新闻隔序列的获得。噬菌体或质粒等基因组侵入宿主细胞后,Cas 蛋白复合体能够识别这些外来 DNA,将其裂解得到短的原型间隔序列 (Protospacer),并加工整合形成新的 R-S 区,后者能够整合到 CRISPR/Cas 系统中的前导序列与第 1 个重复序列之间,说明前导序列可能还具有识别定位的作用。Cas 蛋白对原型间隔序列的识别位点主要位于其末端的原型间隔序列临近基序 (Protospacer adjacent motifs, PAMs)<sup>[27]</sup>。

CRISPR/Cas 的转录与加工。CRISPR 簇首先转录为长的转录前体 Pre-crRNA (Precursor CRISPR RNAs),然后通过 RNase III 或 Cas6 蛋白的作用,将 Pre-crRNA 加工成短的 crRNA。成熟的 crRNA 只含有 1 个间隔序列,并只能特异性的识别一段 DNA 序列<sup>[28]</sup>。

CRISPR/Cas 系统介导的基因修饰。crRNA 中含有的间隔序列与噬菌体等外源基因组中的原型间隔序列同源,当细胞被再次侵染时,crRNA 可以通过碱基互补配对与外源基因组结合,并利用 Cas 蛋白复合物对其进行切割,从而阻止外源 DNA 的正确表达。在此过程中,原型间隔序列临近基序和 crRNA 上的 R 区,对保护 crRNA 自身具有重要作用,它们可以帮助 Cas 蛋白复合物区分自我与非我,从而特异性地切割外源 DNA<sup>[29]</sup>。而只要外源 DNA 的原型间隔序列发生单核苷酸的突变,就能够使其逃避 CRISPR/Cas 系统的干扰。若原型间隔序列临近基序发生突变,即使 crRNA 的间隔序列与外源基因组的原型间隔序列同源,也不能发挥其免疫作用<sup>[26]</sup>。

## 2.2 CRISPR/Cas 系统的应用

CRISPR 是一种进化速度非常快的结构,其间隔序列具有多样性和特异性等特点,已被广泛应用于基因分型、流行病学以及菌落差异分析等研究。其中,利用间隔序列建立的结核分枝杆菌基因分型的标准方法就是 CRISPR 应用的典型例子<sup>[30]</sup>。CRISPR/Cas 系统作为一种抗噬菌体机制,可以为发酵工业中的噬菌体污染带来新的解决方法。同时,由于间隔序列对 DNA 的特异识别和 CRISPR/Cas 系统对基因修饰作用的可遗传性,为基因修饰特别是基因敲除和多基因敲除提供了新的有效方法。目前,基于 CRISPR/Cas 系统的基因敲除方法已经在酵母<sup>[31]</sup>、斑马鱼<sup>[32]</sup>、果蝇<sup>[33]</sup>以及人体<sup>[34]</sup>内得到成功应用。2013 年,Mali 等<sup>[34]</sup>第 1 次将 CRISPR/Cas 系统应用于人工基因修饰,并在人体细胞内得到验证。首先,他们对 *cas9* 基因进行密码子优化,使其能在人体内高效表达,并在其 C-末端增加了一段核定位信号 SV40 后,进行人工合成。其次是 gRNA (Guide RNA) 的设计,gRNA 主要由启动子、crRNA、终止子三部分构成。Mali 等<sup>[34]</sup>以 RNase III 启动子 U6 为 gRNA 的启动子,终止序列为 5'-TTTTT-3',由于 U6 是以 G 为转录起始位点,Cas9 蛋白要求 PAM 的序列为-NGG,故理论上,该 CRISPR/Cas 系统能够任意敲除一段包含有 GN<sub>20</sub>GG 序列的 DNA 片段。最后将他们转入受体细胞中,当 Cas9 蛋白将目的位点的双链切割后,细胞内将启动基因修复机制,从而产生基因缺失或插入突变,导致目的基因无法正确表达,达到基因敲除的目的。试验结果表明,基于 CRISPR/Cas 系统的基因敲除方法是一种简单、易行且高效的实验方法。此外,Mali 还利用此方法对 2 个位点同时进行了敲除操作,也得到了高的突变效率。Miao 等<sup>[35]</sup>也通过 CRISPR/Cas 系统成功得到了水稻的 *CAO1* 和 *LAZY1* 基因缺失突变体,同时,他们的试验结果还证明,即使针对 GC 含量很高的植物基因序列,CRISPR/Cas 系统也能够有较好的打靶效率。这也为水稻的分子育种提供了新方法。此前,水稻的分子育种都只能局限于对调控某个性状的单个基因进行编辑,而 CRISPR/Cas 系统能够实现对多基因位点的定点修饰,这将大大提高水稻育种的效率。而由于 CRISPR/Cas 系统的靶位点在生物中的分布频率很高,及其自身的不断被优化,可以预见,CRISPR/Cas 系统将能够成功运用于玉米、小麦等其他作物中,从而为农作物的产量及性状的提高做出重要贡献。

### 3 结论

TALENs 和 CRISPR/Cas 都是近几年新发现的基因修饰技术,虽然它们在动物领域上的应用更为广泛与成熟,但 Zhang 等<sup>[12]</sup>在烟草中成功的利用 TALENs 进行基因敲除和敲入的试验结果以及 CRISPR/Cas 成功运用于水稻基因功能研究的范例也丰富了其在植物基因功能研究方面的应用<sup>[35]</sup>。它们的发现、发展与应用为基因功能的研究提供了新方法,开拓了新视野。

虽然相较于传统的基因功能研究方法,TALENs 和 CRISPR/Cas 具有简单快速、高效及特异性高等优点,但要想能够取代基因重组或 RNAi 等传统方法而得到更广泛的应用,还有一些问题需要解决:首先,研究并评价 TALEN 和 CRISPR/Cas 系统对细胞的毒性;其次,TALEN 和 CRISPR/Cas 系统本身也存在着一定的缺陷,TALE 蛋白的序列较长且 TALEN 需要成对存在才能发挥作用,而 CRISPR/Cas 系统虽然序列较短但其打靶效应对 PAM 的依赖性较高;最后,针对不同的物种,构建载体时的启动子、终止子和入核信号的选择及载体导入受体细胞的方法还需要一定的探索工作。

#### 参考文献:

- [1] Li T, Huang S, Jiang W Z, *et al.* Hybrid proteins composed of TAL effectors and *Fok* I DNA-cleavage domain [J]. *Nucleic Acids Research*, 2011, 39(1): 359 – 372.
- [2] Moscou M J, Bogdanove A J. A simple cipher governs DNA recognition by TAL effectors [J]. *Science*, 2009, 326(5959): 1501.
- [3] Boch J, Scholze H, Schornack S, *et al.* Breaking the code of DNA binding specificity of TAL-type III effectors [J]. *Science*, 2009, 326(5959): 1509 – 1512.
- [4] Tesson L, Usal C, Ménoret S, *et al.* Knockout rats generated by embryo microinjection of TALENs [J]. *Nature Biotechnology*, 2011, 29(8): 695 – 696.
- [5] Lei Y, Guo X, Liu Y, *et al.* Efficient targeted gene disruption in *Xenopus* embryos using engineered transcription activator-like effector nucleases (TALENs) [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, 109(43): 17484 – 17489.
- [6] Huang P, Xiao A, Zhou M, *et al.* Heritable gene targeting in zebrafish using customized TALENs [J]. *Nature Biotechnology*, 2011, 29(8): 699 – 700.
- [7] Mahfouz M M, Li L, Piatek M, *et al.* Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein [J]. *Plant Molecular Biology*, 2012, 78(3): 311 – 321.
- [8] Li T, Liu B, Spalding M H, *et al.* High-efficiency TALEN-based gene editing produces disease-resistant rice [J]. *Nature Biotechnology*, 2012, 30(5): 390 – 392.
- [9] Liu J, Li C, Yu Z, *et al.* Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy [J]. *Journal of Genetics and Genomics*, 2012, 39(5): 209 – 215.
- [10] Wood A J, Lo T W, Zeitler B, *et al.* Targeted genome editing across species using ZFNs and TALENs [J]. *Science*, 2011, 333(640): 307.
- [11] Li T, Huang S, Zhao X, *et al.* Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes [J]. *Nucleic Acids Research*, 2011, 39(14): 6315 – 6325.
- [12] Zhang Y, Zhang F, Li X, *et al.* Transcription activator-like effector nucleases enable efficient plant genome engineering [J]. *Plant Physiology*, 2013, 161(1): 20 – 27.
- [13] Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: discovery and function [J]. *Annu Review Phytopathology*, 2010, 48: 419 – 436.
- [14] Deng D, Yan C, Pan X, *et al.* Structural basis for sequence-specific recognition of DNA by TAL effectors [J]. *Science*, 2012, 335(669): 720 – 723.
- [15] Kim Y G, Cha J, Chandrasegaran S. Hybrid restriction enzymes: Zinc finger fusions to *Fok* I cleavage domain [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1996, 93(3): 1156 – 1160.
- [16] Urnov F D, Rebar E J, Holmes M C, *et al.* Genome editing with engineered Zinc finger nucleases [J]. *Nature Reviews Genetics*, 2010, 11(9): 636 – 646.
- [17] Reyon D, Tsai S Q, Khayter C, *et al.* FLASH assembly of TALENs for high-throughput genome editing [J]. *Nature Biotechnology*, 2012, 30(5): 460 – 465.
- [18] Mahfouz M M, Li L, Shamimuzzaman M, *et al.* De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(6): 2623 – 2628.
- [19] Zhang F, Cong L, Lodato S, *et al.* Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription [J]. *Nature Biotechnology*, 2011, 29(2): 149 – 153.
- [20] Sanjana N E, Cong L, Zhou Y, *et al.* A transcription activator-like effector toolbox for genome engineering [J]. *Nature Protocols*, 2012, 7(1): 171 – 192.
- [21] Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea [J]. *Science*, 2010, 327

- (5962):167–170.
- [22] Barrangou R, Fremaux C, Deveau H, *et al.* CRISPR provides acquired resistance against viruses in prokaryotes [J]. *Science*, 2007, 315(5819):1709–1712.
- [23] Lillestøl R K, Redder P, Garrett R A, *et al.* A putative viral defence mechanism in archaeal cells [J]. *Archaea*, 2006, 2(1):59–72.
- [24] Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats [J]. *BMC Bioinformatics*, 2007, 8:172.
- [25] Lintner N G, Kerou M, Brumfield S K, *et al.* Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE) [J]. *The Journal of Biological Chemistry*, 2011, 286(24):21643–21656.
- [26] Deveau H, Barrangou R, Garneau J E, *et al.* Phage response to CRISPR-Encoded resistance in *Streptococcus thermophiles* [J]. *Journal of Bacteriology*, 2008, 190(4):1390–1400.
- [27] Lillestøl R K, Shah S A, Brügger K, *et al.* CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties [J]. *Molecular Microbiology*, 2009, 72(1):259–272.
- [28] Westra E R, Swarts D C, Staals R H, *et al.* The CRISPRs, they are a-changin': how prokaryotes generate adaptive immunity [J]. *Annual Review of Genetics*, 2012, 46:311–339.
- [29] Marraffini L A, Sontheimer E J. Self versus non-self discrimination during CRISPR RNA-directed immunity [J]. *Nature*, 2010, 463(7280):568–571.
- [30] Sorek R, Kunin V, Hugenholtz P. CRISPR-a widespread system that provides acquired resistance against phages in bacteria and archaea [J]. *Nature Reviews Microbiology*, 2008, 6(3):181–186.
- [31] Dicarlo J E, Norville J E, Mali P, *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems [J]. *Nucleic Acids Research*, 2013, 41(7):4336–4343.
- [32] Hwang W Y, Fu Y, Reyon D, *et al.* Efficient genome editing in zebrafish using a CRISPR-Cas system [J]. *Nature Biotechnology*, 2013, 31(3):227–229.
- [33] Bassett A R, Tibbit C, Ponting C P, *et al.* Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system [J]. *Cell Reports*, 2013, 4(1):220–228.
- [34] Mali P, Yang L H, Esvelt K M, *et al.* RNA-Guided human genome engineering via Cas9 [J]. *Science*, 2013, 339(6121):823–826.
- [35] Miao J, Guo D, Zhang J, *et al.* Targeted mutagenesis in rice using CRISPR-Cas system [J]. *Cell Research*, 2013, 23(10):1233–1236.