

综述

TALENs 技术的研究进展及其应用前景

唐宇¹, 李丽莎^{1,2}, 林峻^{1,2}

1 福州大学生物科学与工程学院, 福建 福州 350108

2 福州大学应用基因组学研究所, 福建 福州 350108

唐宇, 李丽莎, 林峻. TALENs 技术的研究进展及其应用前景. 生物工程学报, 2015, 31(7): 1024–1038.

Tang Y, Li LS, Lin J. Advances in transcription activator-like effectors - a review. Chin J Biotech, 2015, 31(7): 1024–1038.

摘要: 转录激活因子样效应物核酸酶 (Transcription activator-like effector nucleases, TALENs), 是一种将最初发现于植物病原菌的 TALEs 蛋白 (Transcription activator-like effectors, TALEs) 同限制性核酸内切酶的切割结构域 (例如 *Fok I*) 融合而成的人工酶。TALENs 的特异性取决于数量可变的串联重复单元, 这也是 TALEs 特异性识别 DNA 的功能结构单元。理论上, 人工构建的 TALENs 可以特异性识别任意 DNA 位点并使之断裂, 进行基因敲除、插入或修饰。TALENs 技术发展至今已经在多层次、多物种中获得成功。本文总结了 TALENs 技术的研究进展, 及其与 ZFNs、CRISPR/Cas 等技术相比较的优劣之处。TALENs 在工业微生物育种领域具有特殊的应用优势, 其载体构建便利、靶向精准、编辑效率高, 以及拥有良好的生物安全性, 具备推广应用的前景。

关键词: 转录激活因子样效应物核酸酶, 基因组编辑, 工业微生物

Advances in transcription activator-like effectors - a review

Yu Tang¹, Lisha Li^{1,2}, and Jun Lin^{1,2}

1 College of Biological Science and Engineering, Fuzhou University, Fuzhou 350108, Fujian, China

2 Institute of Applied Genomics, Fuzhou University, Fuzhou 350108, Fujian, China

Abstract: As a protein originally found in plant pathogenic bacteria, transcription activator-like effectors (TALEs) can be fused with the cleaving domain of restriction endonuclease (For example *Fok I*) to form artificial nucleases named

Received: December 3, 2014; **Accepted:** January 19, 2015

Supported by: National Natural Science Foundation of China (No. 31301537), National College Students' innovation and entrepreneurship training programs (No. 201410386023), Science Development Foundation of Fuzhou University (No. 2013-XY-17).

Corresponding author: Jun Lin. Tel: +86-591-22866273; E-mail: jun@fzu.edu.cn

国家自然科学基金 (No. 31301537), 国家大学生创新创业训练计划 (No. 201410386023), 福州大学科技发展基金 (No. 2013-XY-17) 资助。

网络出版时间: 2015-02-27

网络出版地址: <http://www.cnki.net/kcms/detail/11.1998.Q.20150227.1107.001.html>

TALENs. These proteins are dependent on variable numbers of tandem Repeats of TALEs to recognize and bind DNA sequences. Each of these repeats consists of a set of approximately 34 amino acids, composed of about 32 conserved amino acids and 2 highly variable amino acids called repeat variant di-residues (RVDs). RVDs distinguish one TALE from another and can make TALEs have a simple cipher for the one-to-one recognition for proteins and DNA bases. Based on this, in theory, artificially constructed TALENs could recognize and break DNA sites specifically and arbitrarily to perform gene knockout, insertion or modification. We reviewed the development of this technology in multi-level and multi species, and its advantages and disadvantages compared with ZFNs and CRISPR/Cas technology. We also address its special advantages in industrial microbe breeding, vector construction, targeting precision, high efficiency of editing and biological safety.

Keywords: TALENs, genome editing, industrial microorganism

基因组的定点修饰, 即针对性地对遗传信息进行改造, 包括三种基本形式: 敲除、插入和替换。实际上, “替换”过程也可以理解为先“敲除”内源序列, 再于同一位点上“插入”外源序列。发展自上世纪 80 年代的经典基因打靶技术 (Gene targeting), 通过同源重组 (Homologous recombination) 解决了上述的“敲除”和“插入”难题。经典基因打靶技术在过去的 30 年里为生物技术的发展作出了贡献, 但也存在许多不足之处, 包括: 命中率低, 随机插入时有发生, 打靶精确度有待提高; 阳性克隆筛选、验证工作量巨大、效率较低; 打靶载体的设计和构建较为繁琐。

随着全基因组测序技术 (Whole-genome sequencing) 的发展以及个体化医疗需求的扩大, 在获取海量基因信息的同时, 如何分析并对这些信息加以利用, 开始成为基因组研究中亟待解决的重要课题。无论是功能性基因组研究或是临床医学应用, 基因组编辑技术都是极为重要的组成部分。理想的基因组编辑技术应当具有靶标特异性、高效且经济^[1]。而过去常用的经典基因打靶技术效率低下且费用高昂^[2-3]。因此, 人工核酸酶技术一经出现便得到迅速发

展。其原理是将非特异性核酸酶与 DNA 识别结构域人工连接形成融合工具酶。锌指核酸酶 (Zinc-finger nucleases, ZFNs) 便是这样一类工程酶。它由两个不同结构域组成: 锌指结构域 (Zinc finger), 能够识别 DNA 序列并与之结合; 核酸内切酶 *Fok I* 酶切结构域, 可以对 DNA 进行切割。ZFNs 技术能显著提高基因组编辑效率, 并成功应用于多种生物^[4-12], 基于这一技术开发 HIV 治疗药物已进入二期临床实验阶段^[13]。2009 年, 随着重复可变双残基 (Repeat variant di-residues, RVDs) 与单核苷酸对应关系的破解, 转录激活因子样效应物核酸酶 (Transcription activator-like effector nucleases, TALENs) 的研究应用进入高速发展阶段, 并有望成为替代 ZFNs 成为新的主流基因组编辑工具。TALENs 技术实现了单个蛋白质与单核苷酸的对应, 使得模块化定制的转录激活因子样效应物 (Transcription activator-like effector, TALE) 结构域可以同任意目标位点 DNA 序列识别结合, 因此 TALENs 具有比 ZFNs 更易设计、组装以及拥有更高的打靶效率。目前, TALENs 已在植物、斑马鱼、青蛙、蝙蝠、猪以及人类的体细胞和多能干细胞中得到成功应用。

近年来,基因组编辑技术得到了飞速发展,ZFNs、TALENs 技术在不断完善成熟,与此同时也有新的技术出现,例如成簇规律间隔短回文重复序列/CRISPR 相关蛋白 (Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein, CRISPR/Cas) 技术。在此篇综述中,我们将着重介绍 TALENs 技术,并就其与 ZFNs、CRISPR/Cas 之间的差异做简要介绍。

1 TALENs: 特异性识别与切割

TALEN 是 TALE 结构域与 *Fok I* 核酸内切酶结构域组合而成的融合蛋白。因 TALE 有多种变体,所以也有多种 TALEN。根据功能不同,TALENs 的结构分为 TALEs 组成的 DNA 识别结合结构域,以及 *Fok I* 核酸内切酶功能域组成的酶促切割结构域。

1.1 TALENs 的识别结合结构域

TALEs 是一类由植物病原菌黄单胞杆菌 *Xanthomonas sp.* 产生的蛋白质,可以通过 III 型分泌系统进入宿主细胞,因此也被称为 III 型效应物^[14]。一旦进入宿主细胞,一些 TALEs 就会进入细胞核,并与相对应的 DNA 序列结合,转录并激活其表达。由于遗传背景不同,一般情况下,激活某些基因会增加宿主对病原菌的增殖敏感度^[15],但某些情况下也会引起宿主的防御反应^[16]。

从结构上来说,TALEs 蛋白的 N 端一般含有 III 型分泌信号肽,C 端则包含有一个功能性核定位信号 (Nuclear localization signal, NLS) 以及一个具有真核转录激活因子特征的高效转录激活结构域 (Activation domain, AD)^[17] (图 1A)。两者之间则是决定 TALENs 特异性的 DNA

识别结合结构域,由若干 (13-29) 个重复氨基酸单元组成,每个重复氨基酸单元含有 34 个呈串联排列的氨基酸^[18]。每个重复单元的氨基酸序列几乎一致,仅 12、13 号位点的氨基酸残基是高度可变的^[19],并且是特异性识别核苷酸的关键部件 (图 1B)。因此,这两个变化的特殊氨基酸残基被合称为 RVD。其中,第 13 位残基负责识别特定核苷酸,而 12 位残基则承担了稳定 RVD 结构,并与 DNA 蛋白质骨架结合的功能^[20]。相同的发现也在 TALE-DNA 结合模拟中被提及^[21]。TALEs 中串联排列的 RVDs 与靶标 DNA 序列的核苷酸近乎一一对应。2009 年,这一对应关系被两个研究团队分别独立破解,并刊登在同一期 *Science* 杂志上。其中, Moscou 和 Bogdanove^[22]通过计算机扫描比对 RVDs 与靶标的启动子序列并分析其频率,发现 RVDs 与它们靶标位点的核苷酸直接对应,且表现为一个 RVD 对应一个核苷酸。这些对应关系存在部分的简并性,但没有明显的上下游序列依赖 (Context dependence) 现象。例如在统计了 383 个 RVD-核苷酸的对应关系后,他们发现 HD 是出现最为频繁的 RVD 并且与 C 存在很强的对应关系。其次是识别 T 的 NG 和识别 A 的 NI。再次则是识别 G 或是 A 的 NN。另一个小组的 Boch 等^[23]则通过实验分析 TALEs 与靶标 DNA 分子的对应关系进而得出了与 Moscou 等类似的结论。这些 RVD-核苷酸的对应关系被人们归纳出来并用于设计能识别特定 DNA 序列的 TALEs^[18,24-26]。

1.2 TALENs 的酶促切割结构域

近年来,许多研究中应用了人工设计的核酸酶来进行基因组编辑^[27]。这些编辑工具通常是由可定制的 DNA 识别结合域与非特异性的核

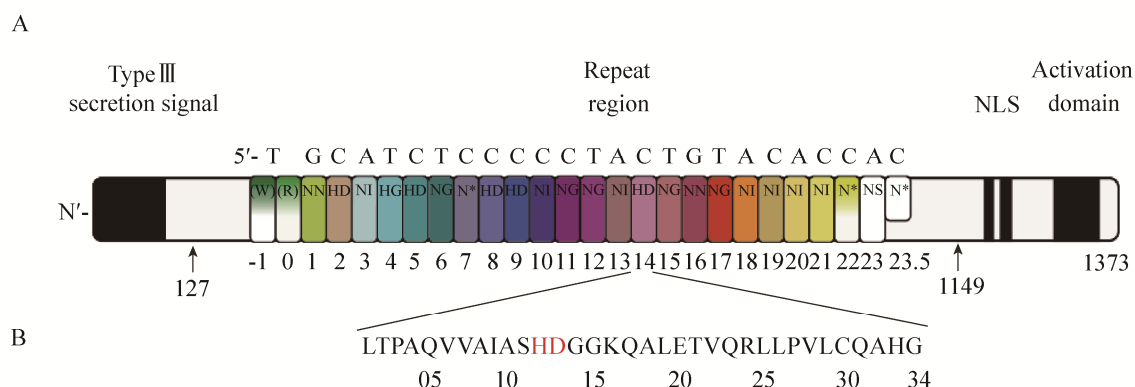


图 1 TALE 组成及 RVD 结构^[17]

Fig. 1 The composition of TALEs and structure of RVD. (A) The general construction of TALEs. (B) The general construction of repeats^[17].

酸内切酶融合而成^[28] (如 ZFNs, 也如 TALENs)。ZFNs 与 TALENs 通常连接的是 *Fok* I 核酸内切酶。*Fok* I 是最初分离于细菌海床黄杆菌 *Flavobacterium okeanoikoites* 的一类 II S 型核酸酶^[29]。在有底物 DNA 存在时用胰蛋白酶水解可以得到 41 kDa 的 N 端 DNA 结合结构域与 25 kDa 的 C 端 DNA 酶切结构域^[30]。通常, II 型限制性内切酶的切割位点与其结合位点相同或邻近, II S 型核酸内切酶则在距离结合位点固定距离处切割 DNA 双链。对于 *Fok* I 核酸酶, 其 DNA 结合结构域识别结合一段非回文序列 5'-GGATG-3'/5'-CATCC-3', 然后酶切结构域在结合位点下游距离 9 个和 13 个核苷酸的位点非特异性地切割 DNA 双链^[30-31]。*Fok* I 单体并不具有活性, 只有当它与靶标 DNA 结合并且形成二聚体, 在二价金属离子辅助下才具有酶切活性^[32]。根据此特性, 人们设想通过使两分子 *Fok* I 分别在两个相邻识别位点结合, 形成具有核酸酶功能的非同源二聚体, 切割两位点之间的 DNA 双链^[28,33-34]。

2 TALENs 的设计与构建

2010 年 6 月, Christian 等^[24]首次报道了人工构建 TALENs 技术。他们将 AvrBs3 和 PthXoI 中天然存在的 TALEs 序列与 *Fok* I 融合形成融合蛋白后对目标 DNA 进行打靶, 并对结合位点间距不同的 TALEN 打靶效率进行比较, 分析 TALEN 二聚体最佳的打靶位点距离。实验结果显示, AvrBs3 和 PthXoI 的最佳打靶位点间距分别是 21 bp 和 24 bp, 并且它们同时在 15 bp 处有相对较大的活性。TAL 结构域和 *Fok* I 结构域之间的连接区 (Linkers) 序列长度对于 TALEN 的活性影响巨大, 而两个 TALEN 结合位点的间隔区 (Spacers) 序列长度则与 *Fok* I 二聚体的形成效率息息相关。2011 年, Li 等用来自 AvrXa7 和 PthX01 中天然存在的 TALE 重复序列进行了类似的实验, 并比较了将 *Fok* I 结构域连接到 TALE 结构域 N 端和 C 端构成的 TALEN 融合蛋白的活性, 结果发现 *Fok* I 连接到 TALE 的 C 端时效率更高^[18]。对于 TALEs 的活性影响, 尽管不同 TALEs 对 N 端、C 端的依赖程度不同,

但对于绝大多数 TALEs 来说, N 端对于 TALE 与 DNA 的相互作用会更为重要^[35]。

植物中的天然 TALEs 的结合序列都是从一个 T (与 TALE 的 N 端的 0 位重复单元相对应) 开始的^[22]。尽管这个 T 碱基同 RVD 没有关联, 但却能参与引起本氏烟草 *Nicotiana benthamiana* 中的 *hax3* 的激活反应^[23]。虽然 Sun 等^[36]使用以 A、C 或 G 起始的 TALENs 也能识别并敲除导致镰刀型贫血病的人类血红蛋白 (HBB) 基因。但这个 T 碱基一般情况下仍是 TALE 活性所必需的^[37]。Scholze 等^[38]的研究显示至少有 4 个因素会影响重复单元与 DNA 的识别结合: 启动基因表达的最小重复单元数量; 重复单元与 DNA 错配引起的副作用; 不同重复单元特异性不同; 位点依赖和上下游序列依赖型错配。

而 Jankele 和 Svoboda 则从特殊重复单元类型和 TALE-DNA 结合模式中推测出来一些合理设计、组装 TALE 重复结构域的要求^[39]:

1) 选择带有 5'-T⁰ 碱基的中央重复结构域 (Central repeat domain, CRD) 识别序列。在无法满足时, 可以选用专一性较低的 5'-N⁰ 来进行调整^[40]。

2) 对待选的靶位点的唯一性进行验证 (例如, 该序列没有因存在高度重复片段而产生的多态性)。

3) 尽管一般情况下重复单元的最适长度会因不同个体的同源序列变化而存在区别^[20,41-43], 但组装 TALENs 时最好不少于 14 个重复单元, 而 TALE 转录激活因子则需要 18-20 个的重复单元。

4) 应该含有至少 4 个位置均衡的高效 RVDs (例如, HD>C, NG>T 或者 NN>G/A), 特

别是在 CRD 的末端, 高效的 RVDs 可以起到稳定 TALE-DNA 结合的作用^[44-45]。

5) 避免出现 3 个相同 RVDs 连续延伸的情况, 特别是 NG。研究显示即使只有 3 个 NG 连接成串也会导致 TALE 蛋白折叠畸形。

6) 如果需要区别对待 A、G, 则最好使用 NH 而非 NN 去识别 G^[46]。

7) 使用 NI 特异性识别 A 时需要搭配使用足够高效的 RVDs^[46]。

8) 使用含有整个 N 端区域 (N-terminal region, NTR) (约 150 个氨基酸) 以及 C 端到效应子距离合适并经过验证的 TALE 骨架。目前, 在多种生物中使用最为频繁的骨架是 Miller 等^[25]建立的。当然, Mussolino 等^[35]和 Zhang 等^[47]建立的骨架也是可靠并被广泛使用的。

9) 最后, 可以通过越来越可靠的在线工具来进行 TALEN 设计与脱靶分析。

针对由于高度重复与高度特异性而导致的 TALENs 组装困难, 研究者开发出多种方案用以克服这些困难, 从而快速、高效地组装 TALENs。Zhang 等^[47]发明了一种逐层连接策略。其原理是先组装 3 个四聚体然后再根据每个四聚体两端的独特粘性末端组装为线性串联复合物。Weber 等^[48]通过 Golden Gate 克隆技术来合成 TALEs。这一克隆技术需要 II S 型核酸酶并包含两个连续步骤: 分别是基于 *Bsa* I 序列的重复单元预装配以及基于 *Bpi* I 序列的预装配模块的合成反应。而 Cermak 等^[49]则通过设计了一种不依赖于 PCR 的策略使得 TALEs 的实用性和灵活性都得到了很大扩展。2012 年, Li 等^[50]描述了一种通过 II 型限制酶进行快速、高效地构建 TALE 转录激活因子 (TALE transcription factors,

TALE-TFs) 和 TALENs 骨架的一步限制酶连续连接反应。而 Sanjana 等^[51]改良了 TALE-TFs 用于人类基因的激活调控和 TALENs 用于调整基因的敲除或插入。还有一种固相高通量自动连接 (Fast ligation-based automatable solid-phase high-throughput, FLASH) 技术可以实现 TALENs 的大规模组装^[52]。这一方法是通过液相操作装置操作载有预组装四聚体的磁珠进行连续连接来构建 TALENs 的。而 Briggs 等^[53]设计了一种称为迭代帽组装 (Active capped assembly, ICA) 的方法, 通过连续连接重复单体从而相当成功且快速地构建定制 TALENs。这一方法与 FLASH 的不同之处在于它是基于固相支撑的, 可以使最终产物达到 21 个单体的概率大大增加。近期, Schmid-Burgk 等^[54]设计了一种非连接酶依赖型克隆 (Ligation-independent cloning, LIC) 技术。他们构建了一个五聚体 TALE 重复片段文库用以合成 TALEs。这一方法使得 TALENs 的构建更加简单快速、高通量以及自动化。尽管脱靶的影响仍是制约 TALENs 技术广泛应用的主要问题^[35,55], 但随着技术的发展^[45], 在可以预见的将来, 减小因非特异性切割而带来的细胞毒性、减小上下游序列依赖现象^[56]、增大 TALEs 的识别特异性将会使得 TALENs 的应用更为广泛和实用。

3 TALENs 的应用

随着 TALENs 组装技术的进步, TALENs 技术已经成功应用于多个领域、多个物种。包括: 线虫^[57]、斑马鱼^[58-60]、鼠^[55,61]、蚕^[62]、哺乳动物^[25]、蟋蟀^[63]、果蝇^[64]、人类^[25,36,65-67]以及多能干细胞^[68-69]。2011 年, Sander 等^[58]利用

Miller 等^[25]发明的方法构建了 TALENs, 并获得了 *hey2* 和 *gria3a* 突变的斑马鱼体细胞, 其突变效率为 11%–33%。Huang 等^[70]则利用一种单元组装法构建的 TALENs 对 *mikb* 和 *dip2a* 两个基因成功打靶并获得了可以稳定遗传的斑马鱼突变体。而到了 2012 年, Bedell 等^[71]借助于 TALENs 切割 DNA 后产生的双链断裂 (Double strand break, DSB), 把短的单链 DNA (Single-stranded DNA, ssDNA, 包含 loxP 序列) 精确地重组到斑马鱼的基因组中。这一突破性研究为条件基因打靶奠定了必要的技术基础。而 Reyon 等^[52]使用 FLASH 系统组装了 48 个 TALENs 并通过基于增强型绿色荧光蛋白 (Enhanced green fluorescent protein, eGFP) 的荧光结果去检测这些蛋白质的活性, 为 TALENs 活性检测提供了直观方法。除了应用 TALENs 进行基础研究, TALENs 也被用于治疗性应用, 包括镰刀型贫血病的治疗^[36]、慢性病毒感染^[72]以及神经性家族遗传病^[73]以及其他疾病的治疗、预防与矫正^[74-79]。此外, Kim 等^[80]使用 TALENs 技术成功获得了具有表面抗原 H-2K (K) 和潮霉素抗性蛋白的突变体, 然后分别使用磁分离和潮霉素抗性培养处理进行细胞分离。结果发现, TALENs 技术可以有效地富集突变体。这也意味着 TALEN 技术可以更广泛地应用于生物医学研究, 使核酸重组效率得到很大提高。

随着 TALEs 技术的研究深入与应用开展, 许多研究者给出了适当的建议。Zhang 等^[1]认为无论是在 TALEs 的基础研究还是临床应用中, 许多因素都应该审慎地进行考虑。这些因素包括但不限于: RVDs 的特异性与效率、连接区 (Linkers) 和间隔区 (Spacers) 的长度。此外,

TALEs 的上下游序列依赖现象与细胞毒性也都需要进行细致的验证, 特别需要注意的是对慢性病毒感染进行治疗时应考虑到可能会对宿主的基因组造成影响^[72]。另外, 由于天然 TALEs 仅在植物病原体中发现, 因此临床治疗中很有可能发生免疫应答反应。

4 TALENs 与 ZFNs、CRISPR/Cas

随着时间的推移、研究的深入, 出现了多种各具特点的基因组编辑技术。

TALENs 是以替代 ZFNs 为目标迅速发展起来的新基因组编辑技术, 并且研究、应用中同 ZFNs 存在许多相似之处。锌指结构域是目前真核生物中最常见的 DNA 结合序列, 也是人类基因组中最常使用的编码蛋白。单个锌指模块大约由 30 个氨基酸组成保守的 $\beta\beta\alpha$ 结构, 其中 α 表面的几个氨基酸会与 DNA 大沟中的三个碱基对相连^[81]。而不同的锌指模块对 DNA 的结合特异性也存在程度差异。因此, 锌指模块使得 DNA 结合蛋白的定制化成为可能。其后, 人工锌指蛋白的出现, 使得锌指蛋白可以识别结合 9-18 个碱基对的 DNA 序列。而 18 个碱基对的识别能力则意味着能在 680 亿个碱基中形成特异性, 是第一个能对人类基因组的特定序列进行特异性结合的技术^[82-83]。很多研究显示, 在相同基因打靶时 TALENs 和 ZFNs 表现出了不相上下的打靶效率^[52,55,58,68]。而 Mussolino 等^[35]比较了 CCR5 特异的 TALENs 与已经验证过的 ZFNs 之间的细胞毒性与特异性。其结果显示, TALENs 引起了约 1% 的错误突变, 而 ZFNs 的错误诱变比率则高达 11%。此外, 经过 TALENs 处理的细胞存活率是经过

ZFNs 处理细胞的两倍。

CRISPR 序列最早于 1987 年由 Ishino 等^[84]发现, Jansen 等^[85]将其正式命名。随后, 在许多细菌和古细菌中发现了 CRISPR 序列和 CRISPR 相关 (CRISPR-associated, Cas) 基因^[86]。目前已发现的 CRISPR/Cas 系统可以根据 Cas 基因核心元件的不同分为 3 种类型, I 型、III 型系统需要多个 Cas 蛋白形成复合体才能切割 DNA 双链, 而 II 型系统则只需要一个 Cas9 蛋白即可^[87]。CRISPR/Cas 系统是细菌和古细菌的一种免疫机制, 目前基因组编辑中应用的多为 II 型系统, 通过 Cas9 蛋白俘获外源 DNA 片段并整合进宿主基因组转录产生前体 CRISPR-RNA (Pre-crRNA)、反式激活 crRNA (Trans-activating crRNA, tracrRNA) 并与形成的向导 RNA (Guide RNA, gRNA) 组合成 Cas9 核酸酶复合体^[88]。在 crRNA 中的 20 bp 向导序列 (Guide sequence) 以及与向导序列结合的靶向 DNA 5' 端的 3 个碱基 (Protospacer-adjacent motif, PAM)^[89-90]介导下, Cas9 核酸酶复合体可以对外源 DNA 引入 DSBs 使之降解, 从而抵御噬菌体等外来侵入者, 进行自我保护^[91-92]。CRISPR/Cas 系统与 TALENs、ZFNs 不同之处在于, 其识别特异性是由 RNA 而非蛋白质决定, 因而设计、组装更为简便^[88,93]。一些研究显示 CRISPR/Cas 整体情况下的酶切活性比 TALENs 更为高效^[88,94]。此外, CRISPR/Cas 系统可以同时靶向基因组的多个位点进行切割或修饰^[95-97], 可以促进多位点突变疾病的治疗。

但是, 基于研究时间尚短及其结构特性, CRISPR/Cas 系统不可避免也存在许多不足: 由

于碱基识别个数的差异、PAM 序列的不严格互补^[88,97-99], CRISPR/Cas 系统存在非特异性切割现象^[98,100-101], 甚至产生较高的脱靶率^[102-103], 特异性远低于 TALENs; 而基于不同物种的遗传背景不同, 可能需要针对性地构建 Cas 基因启动子及其表达载体^[88,96,104-105]; CRISPR/Cas 系统在对不同物种以及同一物种的不同基因进行

定点编辑时, 效率存在有较大的差异^[106-107]。

比较以上 3 种技术 (表 1), ZFNs 技术积累最为丰富, 但存在脱靶率高、细胞毒性大; TALENs 技术与 ZFNs 有很大相似性, 可以借鉴 ZFNs 的研究基础, 但比 ZFNs 更有优势; CRISPR/Cas 技术最为前沿、前景更为广阔, 但受限于发展时间较短、微生物应用基础薄弱。

表 1 TALENs、ZFNs 和 CRISPR/Cas 的对比^[1]

Table 1 Comparison of TALENs, ZFNs and CRISPR/Cas^[1]

Items	TALENs	ZFNs (C ₂ H ₂)	CRISPR/Cas (Type II)
Central domain	Tandem repeats of ~34 aa	Zinc-fingers of ~30 aa	Cas protein
Determining region	RVD of 12th and 13th	Alpha-1,3,6 amino acid	20 nucleotides of crRNA and PAM sequence
Secondary structure	Two helices connected by a short RVD-containing loop	Beta folding-beta folding-alpha helix (β - β - α)	An open bilobed architecture and nucleic acid binding clefts
DNA binding pattern	One repeat for one nucleotide	One zinc finger for one trinucleotide	One nucleotide for one nucleotide
Validity	Very little apparent context dependence	Context dependence	Depend on the PAM sequence
Specificity	Relatively high	Relatively low (off-target effect)	Relatively low (off-target effect)
Degeneracy	Existing same repeat for distinct nucleotides	Having the affinity for similar but not identical sequences their intended targets	-
Toxicity	No or little (if any) detectable cytotoxicity	Yes	Yes
Optimized dimers	10–30 bp spacers with corresponding linkers	18–24 bp DNA linkers with 4–7 bp spacers	-
Applicable sites	Arbitrary theoretically	Existing modules can't target every possible sequence	Next to the PAM sequence
Design	More flexibility and repeats	Less flexibility and fingers	Simple and flexibility
Predictability	Qualified	Lack	Qualified
Intellectual property	Public cipher	Sangamo and Sigma-Aldrich	Public
Cost	Relatively low-cost and time-saving	Relatively costly and time-consuming	Very low-cost and time-saving

TALENs: transcription activator-like effectors; ZFNs: zinc-finger nucleases; CRISPR/Cas: clustered regularly interspaced short palindromic repeats/CRISPR-associated protein; C₂H₂: 2-cysteine and 2-histidine, RVD: repeat variable diresidues crRNA; CRISPR-RNA; PAM: protospacer-adjacent motif.

5 本实验室工作内容和经验

在“国家自然科学基金”等项目的资助下,本实验室在丝状真菌的基因组编辑领域已经积累有一定的实践经验。

首先,在方法选择上,纵观 ZFNs、TALENs、CRISPR/Cas 这三大技术,ZFNs 由于前文所述的种种自身限制或缺陷,目前的研究一般不会优先考虑使用 ZFNs。

而 TALENs 和 CRISPR/Cas 相比,CRISPR/Cas 技术来势汹汹,从 2013 年第一篇应用报道刊登于 *Science* 杂志起,已有数百篇相关研究报道发表于国际顶级学术期刊,其特点和优势在于操作非常简单,只需表达 Cas 酶,依靠 crRNA 就可以靶向编辑目的 DNA 序列,不需要繁琐的元件拼接过程,省时省事。

因此,在研究的早期阶段,我们优先选择 CRISPR/Cas 开展丝状真菌的基因组编辑研究。但是,由于 CRISPR/Cas 的技术缺陷所造成的问题也接踵而来:

第一,CRISPR/Cas 系统依赖 gRNA 来识别目的 DNA,因此,对 gRNA 的序列有着严格的要求,在其他物种中,gRNA 一般放置在已知的、背景非常清楚的启动子和终止子之间,例如,在人类细胞中,一般采用 U6 启动子和终止子,这样,才能保证 gRNA 在特定的碱基位置起始转录,并在特定的碱基位置终止转录。但是,丝状真菌种类繁多,对其遗传背景的研究资料较少,尤其是对启动子和终止子的研究很少,已知的启动子和终止子一般用于蛋白表达,虽然可以转录出 mRNA,但是精确定位某个启动子起始转录位点或者某个终止子终止转录位点的报道几乎没有。这样,就很难获得具有准确 5'

及 3'边界的 gRNA,在无法保证 gRNA 序列准确的情况下,精确的基因组编辑无从谈起。而且,众所周知,丝状真菌常常产生大量的 RNA 酶,而 gRNA 容易受到真菌内源 RNA 酶的攻击,一旦 gRNA 序列有所缺失,即使只缺失 1 个碱基,也可能会彻底丧失功能。

第二,前文已述,CRISPR/Cas 系统需要目标 DNA 分子上的 PAM 序列介导,例如,酿脓链球菌 *Streptococcus pyogenes* 来源的 Cas9 酶,其 PAM 序列是 NGG;脑膜炎双球菌 *Neisseria meningitidis* 来源的 Cas9 酶,其 PAM 序列是 NNNNGATT。这就意味着,CRISPR/Cas 系统只能在有 PAM 的位置处进行 DNA 编辑,这就限制了其应用,如果我们想编辑的某个位点,没有 PAM,那么就无法使用 CRISPR/Cas 系统,而这一情况在某些丝状真菌中较常遇到。

鉴于上述种种因素,我们在后期研究中,优先选择了 TALENs 系统。

TALENs 系统只需要能够表达出 TALEN 酶蛋白分子即可,因此,丝状真菌中已知的启动子和终止子均可使用,另外,TALENs 可以直接和目标 DNA 分子发生相互作用,不需要任何其他中介,不受序列特异性限制,因此靶向性更强。

其次,在实验样品的选择上,由于很多丝状真菌的菌体是多倍体,甚至有些菌体内有多个细胞核,造成了丝状真菌的遗传稳定性较差。如果基因打靶只作用于一条染色体上,或者只作用于其中一个细胞核上,那么,菌体增殖分裂的后代中,仍然会有野生型的细胞存在。这也是很多传统育种技术改造的菌株,在多次传代后出现菌株衰退的原因之一。在研究中我们发现,很多丝状真菌能产生单倍体的孢子,而这些通常被认为是休眠状态的孢子,也可以实

现外源 DNA 的转化。若能在孢子中的单倍体上成功进行基因组编辑,那么,其子代就能有非常好的遗传稳定性,避免了后期繁琐的菌株筛选和复壮步骤。

6 TALENs 技术在工业微生物领域应用前景展望

人类文明步入 21 世纪后,工业生物技术已经在世界范围内掀起了生物技术革命的第三次浪潮,为解决可持续发展中的重大问题,诸如能源危机、自然资源短缺、环境污染、医药产品与食品的生产等等做出了巨大贡献。在国家《“十二五”生物技术发展规划》中,工业生物技术亦被列为“前瞻性基础研究”的发展重点。选育和构建优良的工业微生物菌种,是工业生物技术产业化的前提条件,因此,“工业微生物育种技术”成为制约工业生物技术发展的关键技术。

丝状真菌(也俗称霉菌),广泛应用于发酵工业,在酶制剂、饲料、食品等领域有着重要用途,常见的有木霉、根霉、曲霉等等。丝状真菌与其他常见的工业微生物(诸如酵母菌、细菌)相比,丝状真菌的 DNA 转化非常困难,而且,在某些丝状真菌中,转化入细胞的外源 DNA 很容易随机整合到宿主的基因组上,此外,还有一些丝状真菌在同一个细胞中有多个细胞核,每个细胞核中的染色体又是多倍体,造成遗传稳定性极差等。上述原因,造成了丝状真菌基因工程的研究远远落后于其他物种。

相比于传统的基因打靶技术,TALENs 是革命性的,它集成了“精确”、“高效”、“简便”这三大优势,其发展与普及必是大势所趋,*Nature* 杂志社更将其评为“Method of the Year

2011”。TALENs 可以根据任意的目标 DNA 序列设计对应的结合结构域。TALENs 在工业微生物中可以有如下的应用:1) 在需要高表达生物活性物质的菌株中,TALENs 可以靶向到目标启动子或者调控区,提高目标产物的表达量;2) 在表达蛋白质的菌株中,TALENs 可以敲除菌体内源蛋白酶的基因,降低表达产物被蛋白酶降解的风险;3) 在具有反馈抑制现象的菌株中,TALENs 可以敲除导致抑制现象的基因,解除反馈抑制;4) 在抗生素生产菌株中,发酵产物中除了目标抗生素,常常还会有一些副产物,而这些副产物一般都具有毒性,且难以被纯化,TALENs 技术的出现,可以快速敲除与副产物合成有关的基因,提高发酵产物中抗生素的纯度;5) TALENs 技术还可以敲除一些菌体中冗余的、非必需的基因,这些基因常常会拖累菌体的生长速度,敲除后,可以有效地增加菌体生物量,加快发酵生产进度,节约发酵过程中所耗费的水电和人力资源等。

在工业微生物领域,制作出能够应用于人类生活的产品才是最终目的。很多丝状真菌的发酵产物,常常作为人类的食物或者药品,在日益重视食品安全和医药安全的今天,TALENs 具有比传统技术方法和 CRISPR/Cas 系统都高的靶向精确性,用 TALENs 改造的基因重组菌株,无疑具有更高的生物安全性。

7 结语

可定制的核酸酶使得研究人员能够快捷、有效地敲除、敲入目的基因。ZFNs 技术使人们对于基因与功能的关系研究进入新的局面,而更为优秀的 TALENs 的出现,则为这一进程注

入新的活力。TALENs 的核心部件 TALEs 不仅可以同限制性核酸酶结合成为可定制的高特异性核酸酶, 还能同转录激活结构域、表达结构域、染色质修饰结构域以及荧光蛋白结合, 从而形成种类繁多, 功能各异的融合蛋白。这些基于 TALEs 的新技术都具有极大的开发前景与实用价值, 但是最吸引人的还是能够任意改造基因组的 TALENs。此外, RVD 与碱基的一一对应关系是目前基因组编辑领域最为简单的蛋白质-DNA 互作机制, 在保证准确性的前提下能极大地优化基因打靶的过程与效率。

虽然 TALENs 技术已经成功应用于多种生物与领域, 但仍存在一些尚未明晰的疑问, 比如天然 TALEs 的结合位点前总有一个 T; TALEs 的末尾 RVD 是半个; RVDs 的特异性机理; TALEs 的重复数与结合效率之间的关系; TALENs 的活性影响因素; TALENs 是否会引起免疫反应等。

值得注意的是, CRISPR/Cas 技术面世后, 得到了极大的关注。鉴于 CRISPR/Cas 的便利性, 有一种主流观点认为, TALENs 和 ZFNs 技术已经过时。但是该观点有待商榷, 毕竟, 在一些特殊的应用场合或领域, TALENs 或 ZFNs 仍然有着独特的优势, 例如前文所述的丝状真菌基因改造领域, 就是一个比较典型的例子。因此, 这就要求科技工作者具体问题具体分析, 辩证地根据事物的客观属性, 有针对性地选择合适的方法技术。

REFERENCES

- [1] Zhang M, Wang F, Li S, et al. TALE: a tale of genome editing. *Prog Biophys Mol Biol*, 2014, 114(1): 25–32.
- [2] Muller U. Ten years of gene targeting: targeted mouse mutants, from vector design to phenotype analysis. *Mech Dev*, 1999, 82(1/2): 3–21.
- [3] Teng Y, Yang X. Gene targeting: the beginning of a new era in genetics. *Hereditas*, 2007, 29(11): 1291–1298 (in Chinese).
滕艳, 杨晓. 基因打靶技术: 开启遗传学新纪元. *遗传*, 2007, 29(11): 1291–1298.
- [4] Bibikova M, Carroll D, Segal DJ, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol*, 2001, 21(1): 289–297.
- [5] Bibikova M, Golic M, Golic KG, et al. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*, 2002, 161(3): 1169–1175.
- [6] Dreier B, Fuller RP, Segal DJ, et al. Development of zinc finger domains for recognition of the 5'-CNN-3' family DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem*, 2005, 280(42): 35588–35597.
- [7] Geurts AM, Cost GJ, Freyvert Y, et al. Knockout rats via embryo microinjection of zinc-finger nucleases. *Science*, 2009, 325(5939): 433.
- [8] Lloyd A, Plaisier CL, Carroll D, et al. Targeted mutagenesis using zinc-finger nucleases in *Arabidopsis*. *Proc Natl Acad Sci USA*, 2005, 102(6): 2232–2237.
- [9] Morton J, Davis MW, Jorgensen EM, et al. Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proc Natl Acad Sci USA*, 2006, 103(44): 16370–16375.
- [10] Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem*, 2001, 70: 313–340.
- [11] Wright DA, Townsend JA, Winfrey RJ, et al. High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J*. 2005, 44(4): 693–705.
- [12] Yang D, Yang H, Li W, et al. Generation of PPARgamma mono-allelic knockout pigs via zinc-finger nucleases and nuclear transfer cloning. *Cell Res*, 2011, 21(6): 979–982.
- [13] Sangamo biosciences' CEO discusses Q2 2012 results-earnings call transcript[EB/OL].

- [2014-7-16]. <http://seekingalpha.com/article/750021-sangamo-biosciences-ceo-discusses-q2-2012-results-earnings-call-transcript>.
- [14] Jiang GF, Jiang BL, Yang M, et al. Establishment of an inducing medium for type III effector secretion in *Xanthomonas campestris* pv. *campestris*. *Braz J Microbiol*, 2013, 44(3): 945–952.
- [15] Sugio A, Yang B, Zhu T, et al. Two type III effector genes of *Xanthomonas oryzae* pv. *oryzae* control the induction of the host genes OsTFIIAgamma1 and OsTFX1 during bacterial blight of rice. *Proc Natl Acad Sci USA*, 2007, 104(25): 10720–10725.
- [16] Romer P, Hahn S, Jordan T, et al. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. *Science*, 2007, 318(5850): 645–648.
- [17] Mak AN, Bradley P, Cernadas RA, et al. The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*, 2012, 335(6069): 716–719.
- [18] Li T, Huang S, Jiang WZ, et al. TAL nucleases (TALNs): hybrid proteins composed of TAL effectors and *Fok I* DNA-cleavage domain. *Nucleic Acids Res*, 2011, 39(1): 359–372.
- [19] Bonas U, Stall RE, Staskawicz B. Genetic and structural characterization of the avirulence gene *avrBs3* from *Xanthomonas campestris* pv. *vesicatoria*. *Mol Gen Genet*, 1989, 218(1): 127–136.
- [20] Boch J, Bonas U. *Xanthomonas* AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol*, 2010, 48(1): 419–436.
- [21] Bradley P. Structural modeling of TAL effector-DNA interactions. *Protein Sci*, 2012, 21(4): 471–474.
- [22] Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. *Science*, 2009, 326(5959): 1501.
- [23] Boch J, Scholze H, Schornack S, et al. Breaking the code of DNA binding specificity of TAL-type III effectors. *Science*, 2009, 326(5959): 1509–1512.
- [24] Christian M, Cermak T, Doyle EL, et al. Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics*, 2010, 186(2): 757–761.
- [25] Miller JC, Tan S, Qiao G, et al. A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*, 2011, 29(2): 143–148.
- [26] Morbitzer R, Romer P, Boch J, et al. Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE) -type transcription factors. *Proc Natl Acad Sci USA*, 2010, 107(50): 21617–21622.
- [27] Mussolino C, Cathomen T. TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol*, 2012, 23(5): 644–650.
- [28] Wah DA, Hirsch JA, Dorner LF, et al. Structure of the multimodular endonuclease *Fok I* bound to DNA. *Nature*, 1997, 388(6637): 97–100.
- [29] Sugisaki H, Kanazawa S. New restriction endonucleases from *Flavobacterium okeanokoites* (*Fok I*) and *Micrococcus luteus* (*Mlu I*). *Gene*, 1981, 16(1/3): 73–78.
- [30] Li L, Wu LP, Chandrasegaran S. Functional domains in *Fok I* restriction endonuclease. *Proc Natl Acad Sci USA*, 1992, 89(10): 4275–4279.
- [31] Szybalski W, Kim SC, Hasan N, et al. Class-II restriction enzymes—a review. *Gene*, 1991, 100: 13–26.
- [32] Bitinaite J, Wah DA, Aggarwal AK, et al. *Fok I* dimerization is required for DNA cleavage. *Proc Natl Acad Sci USA*, 1998, 95(18): 10570–10575.
- [33] Wah DA, Bitinaite J, Schildkraut I, et al. Structure of *Fok I* has implications for DNA cleavage. *Proc Natl Acad Sci USA*, 1998, 95(18): 10564–10569.
- [34] Vanamee ES, Santagata S, Aggarwal AK. *Fok I* requires two specific DNA sites for cleavage. *J Mol Biol*, 2001, 309(1): 69–78.
- [35] Mussolino C, Morbitzer R, Lutge F, et al. A novel TALE nuclease scaffold enables high genome editing activity in combination with low toxicity. *Nucleic Acids Res*, 2011, 39(21): 9283–9293.
- [36] Sun N, Liang J, Abil Z, et al. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol Biosyst*, 2012, 8(4): 1255–1263.
- [37] Wright G, Carver A, Cottom D, et al. High level expression of active human alpha-1-antitrypsin in the milk of transgenic sheep. *Biotechnology (NY)*, 1991, 9(9): 830–834.

- [38] Scholze H, Boch J. TAL effector-DNA specificity. *Virulence*, 2010, 1(5): 428–432.
- [39] Jankele R, Svoboda P. TAL effectors: tools for DNA targeting. *Brief Funct Genomics*, 2014, 13(5): 409–419.
- [40] Lamb BM, Mercer AC, Barbas CR. Directed evolution of the TALE N-terminal domain for recognition of all 5' bases. *Nucleic Acids Res*, 2013, 41(21): 9779–9785.
- [41] Joung JK, Sander JD. TALENs: a widely applicable technology for targeted genome editing. *Nat Rev Mol Cell Biol*, 2013, 14(1): 49–55.
- [42] Wefers B, Meyer M, Ortiz O, et al. Direct production of mouse disease models by embryo microinjection of TALENs and oligodeoxynucleotides. *Proc Natl Acad Sci USA*, 2013, 110(10): 3782–3787.
- [43] Guilinger JP, Pattanayak V, Reyon D, et al. Broad specificity profiling of TALENs results in engineered nucleases with improved DNA-cleavage specificity. *Nat Methods*, 2014, 11(4): 429–435.
- [44] Meckler JF, Bhakta MS, Kim MS, et al. Quantitative analysis of TALE-DNA interactions suggests polarity effects. *Nucleic Acids Res*, 2013, 41(7): 4118–4128.
- [45] Cong L, Zhou R, Kuo YC, et al. Comprehensive interrogation of natural TALE DNA-binding modules and transcriptional repressor domains. *Nat Commun*, 2012, 3: 968.
- [46] Streubel J, Blucher C, Landgraf A, et al. TAL effector RVD specificities and efficiencies. *Nat Biotechnol*, 2012, 30(7): 593–595.
- [47] Zhang F, Cong L, Lodato S, et al. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*, 2011, 29(2): 149–153.
- [48] Weber E, Engler C, Gruetzner R, et al. A modular cloning system for standardized assembly of multigene constructs. *PLoS ONE*, 2011, 6(2): e16765.
- [49] Cermak T, Doyle EL, Christian M, et al. Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res*, 2011, 39(12): e82.
- [50] Li L, Piatek MJ, Atef A, et al. Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. *Plant Mol Biol*, 2012, 78(4/5): 407–416.
- [51] Sanjana NE, Cong L, Zhou Y, et al. A transcription activator-like effector toolbox for genome engineering. *Nat Protoc*, 2012, 7(1): 171–192.
- [52] Reyon D, Tsai SQ, Khayter C, et al. FLASH assembly of TALENs for high-throughput genome editing. *Nat Biotechnol*, 2012, 30(5): 460–465.
- [53] Briggs AW, Rios X, Chari R, et al. Iterative capped assembly: rapid and scalable synthesis of repeat-module DNA such as TAL effectors from individual monomers. *Nucleic Acids Res*, 2012, 40(15): e117.
- [54] Schmid-Burgk JL, Schmidt T, Kaiser V, et al. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. *Nat Biotechnol*, 2013, 31(1): 76–81.
- [55] Tesson L, Usal C, Menoret S, et al. Knockout rats generated by embryo microinjection of TALENs. *Nat Biotechnol*, 2011, 29(8): 695–696.
- [56] Juillerat A, Dubois G, Valton J, et al. Comprehensive analysis of the specificity of transcription activator-like effector nucleases. *Nucleic Acids Res*, 2014, 42(8): 5390–5402.
- [57] Wood AJ, Lo TW, Zeitler B, et al. Targeted genome editing across species using ZFNs and TALENs. *Science*, 2011, 333(6040): 307.
- [58] Sander JD, Cade L, Khayter C, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol*, 2011, 29(8): 697–698.
- [59] Hwang WY, Peterson RT, Yeh JR. Methods for targeted mutagenesis in zebrafish using TALENs. *Methods*, 2014, 69(1): 76–84.
- [60] Cade L, Reyon D, Hwang WY, et al. Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs. *Nucleic Acids Res*, 2012, 40(16): 8001–8010.
- [61] Tong C, Huang G, Ashton C, et al. Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. *J Genet Genomics*, 2012, 39(6):

- 275–280.
- [62] Ma S, Zhang S, Wang F, et al. Highly efficient and specific genome editing in silkworm using custom TALENs. *PLoS ONE*, 2012, 7(9): e45035.
- [63] Watanabe T, Ochiai H, Sakuma T, et al. Non-transgenic genome modifications in a hemimetabolous insect using zinc-finger and TAL effector nucleases. *Nat Commun*, 2012, 3: 1017.
- [64] Liu J, Li C, Yu Z, et al. Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. *J Genet Genomics*, 2012, 39(5): 209–215.
- [65] Ding Q, Regan SN, Xia Y, et al. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell*, 2013, 12(4): 393–394.
- [66] Hu R, Wallace J, Dahlem TJ, et al. Targeting human microRNA genes using engineered Tal-effector nucleases (TALENs). *PLoS ONE*, 2013, 8(5): e63074.
- [67] Niu J, Zhang B, Chen H. Applications of TALENs and CRISPR/Cas9 in human cells and their potentials for gene therapy. *Mol Biotechnol*, 2014, 56(8): 681–688.
- [68] Hockemeyer D, Wang H, Kiani S, et al. Genetic engineering of human pluripotent cells using TALE nucleases. *Nat Biotechnol*, 2011, 29(8): 731–734.
- [69] Geissler R, Scholze H, Hahn S, et al. Transcriptional activators of human genes with programmable DNA-specificity. *PLoS ONE*, 2011, 6(5): e19509.
- [70] Huang P, Xiao A, Zhou M, et al. Heritable gene targeting in zebrafish using customized TALENs. *Nat Biotechnol*, 2011, 29(8): 699–700.
- [71] Bedell VM, Wang Y, Campbell JM, et al. *In vivo* genome editing using a high-efficiency TALEN system. *Nature*, 2012, 491(7422): 114–118.
- [72] Schiffer JT, Aubert M, Weber ND, et al. Targeted DNA mutagenesis for the cure of chronic viral infections. *J Virol*, 2012, 86(17): 8920–8936.
- [73] Hsu PD, Zhang F. Dissecting neural function using targeted genome engineering technologies. *ACS Chem Neurosci*, 2012, 3(8): 603–610.
- [74] Dupuy A, Valton J, Leduc S, et al. Targeted gene therapy of xeroderma pigmentosum cells using meganuclease and TALEN. *PLoS ONE*, 2013, 8(11): e78678.
- [75] Ma N, Liao B, Zhang H, et al. Transcription activator-like effector nuclease (TALEN)-mediated gene correction in integration-free beta-thalassemia induced pluripotent stem cells. *J Biol Chem*, 2013, 288(48): 34671–34679.
- [76] Bloom K, Ely A, Mussolino C, et al. Inactivation of hepatitis B virus replication in cultured cells and *in vivo* with engineered transcription activator-like effector nucleases. *Mol Ther*, 2013, 21(10): 1889–1897.
- [77] Ousterout DG, Perez-Pinera P, Thakore PI, et al. Reading frame correction by targeted genome editing restores dystrophin expression in cells from Duchenne muscular dystrophy patients. *Mol Ther*, 2013, 21(9): 1718–1726.
- [78] Osborn MJ, Starker CG, McElroy AN, et al. TALEN-based gene correction for epidermolysis bullosa. *Mol Ther*, 2013, 21(6): 1151–1159.
- [79] Piganeau M, Ghezraoui H, De Cian A, et al. Cancer translocations in human cells induced by zinc finger and TALE nucleases. *Genome Res*, 2013, 23(7): 1182–1193.
- [80] Kim H, Kim MS, Wee G, et al. Magnetic separation and antibiotics selection enable enrichment of cells with ZFN/TALEN-induced mutations. *PLoS ONE*, 2013, 8(2): e56476.
- [81] Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*, 1991, 252(5007): 809–817.
- [82] Beerli RR, Segal DJ, Dreier B, et al. Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci USA*, 1998, 95(25): 14628–14633.
- [83] Beerli RR, Dreier B, Barbas CR. Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci USA*, 2000, 97(4): 1495–1500.
- [84] Ishino Y, Shinagawa H, Makino K, et al. Nucleotide sequence of the *iap* gene, responsible

- for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J Bacteriol*, 1987, 169(12): 5429–5433.
- [85] Jansen R, Embden JD, Gastra W, et al. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*, 2002, 43(6): 1565–1575.
- [86] Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, 2007, 8: 172.
- [87] Sapranaukas R, Gasiunas G, Fremaux C, et al. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res*, 2011, 39(21): 9275–9282.
- [88] Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, 339(6121): 819–823.
- [89] DiCarlo JE, Norville JE, Mali P, et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res*, 2013, 41(7): 4336–4343.
- [90] Friedland AE, Tzur YB, Esvelt KM, et al. Heritable genome editing in *C. elegans* via a CRISPR-Cas9 system. *Nat Methods*, 2013, 10(8): 741–743.
- [91] Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. *Nature*, 2012, 482(7385): 331–338.
- [92] Terns MP, Terns RM. CRISPR-based adaptive immune systems. *Curr Opin Microbiol*, 2011, 14(3): 321–327.
- [93] Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*, 2013, 339(6121): 823–826.
- [94] Ding Q, Regan SN, Xia Y, et al. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell*, 2013, 12(4): 393–394.
- [95] Li W, Teng F, Li T, et al. Simultaneous generation and germline transmission of multiple gene mutations in rat using CRISPR-Cas systems. *Nat Biotechnol*, 2013, 31(8): 684–686.
- [96] Wang H, Yang H, Shivalila CS, et al. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. *Cell*, 2013, 153(4): 910–918.
- [97] Yang H, Wang H, Shivalila CS, et al. One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell*, 2013, 154(6): 1370–1379.
- [98] Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*, 2013, 31(9): 827–832.
- [99] Anders C, Niewoehner O, Duerst A, et al. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature*, 2014, 513(7519): 569–573.
- [100] Jiang W, Bikard D, Cox D, et al. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat Biotechnol*, 2013, 31(3): 233–239.
- [101] Mali P, Aach J, Stranges PB, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*, 2013, 31(9): 833–838.
- [102] Fu Y, Foden JA, Khayter C, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol*, 2013, 31(9): 822–826.
- [103] 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.
- [104] Hwang WY, Fu Y, Reyon D, et al. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*, 2013, 31(3): 227–229.
- [105] Ramalingam S, Annaluru N, Chandrasegaran S. A CRISPR way to engineer the human genome. *Genome Biol*, 2013, 14(2): 107.
- [106] Karvelis T, Gasiunas G, Miksys A, et al. crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol*, 2013, 10(5): 841–851.
- [107] Koonin EV, Makarova KS. CRISPR-Cas: evolution of an RNA-based adaptive immunity system in prokaryotes. *RNA Biol*, 2013, 10(5): 679–686.

(本文责编 陈宏宇)