# 类转录激活因子效应物核酸酶(TALENs)的 构建及其在基因组定点修饰中的应用

周金伟<sup>1</sup> 王灵慧<sup>2</sup> 申义君<sup>1</sup> 余树民<sup>1</sup> 曹随忠<sup>1\*</sup> ('四川农业大学动物医学院, 雅安 625014; <sup>2</sup>四川大学生命科学学院生物资源与生态环境 教育部重点实验室, 成都 610064)

摘要 类转录激活因子效应物核酸酶(transcription activator-like effector nucleases, TALENs)是 一种新发现的基因组定点修饰工具。TALENs由TALE蛋白及II型核酸内切酶Fok I组成,其中TALE 由多个重复的氨基酸序列构成,对DNA序列的识别达到一个重复单元结合一个碱基的程度,它在结合Fok I内切酶后就形成了具有DNA定点修饰功能的TALENs。该文主要介绍TALENs的构建及其 在基因组定点修饰中的应用。

关键词 类转录激活因子效应物(TALE); TALENs; 基因敲除; 基因组定点修饰; TALENs构建

# The Construction of Transcription Activator-like Effector Nucleases (TALENs) and Its Application in the Genome Fixed-point Modification

Zhou Jinwei<sup>1</sup>, Wang Linghui<sup>2</sup>, Shen Yijun<sup>1</sup>, Yu Shumin<sup>1</sup>, Cao Suizhong<sup>1\*</sup>

(<sup>1</sup>College of Veterinary Medicine, Sichuan Agricultural University, Ya'an 625000, China; <sup>2</sup>College of Life Science, Sichuan University, Key Laboratory of Biological Resources and Ecological Environment, the Ministry of Education, Chengdu 610064, China)

**Abstract** Transcription activator effector nucleases (TALENs) is a novel DNA editing technology which can induce precise site-specific modification in genome. TALENs consist of two distinct domains, an engineered DNA-binding domain derived from the transcription activator-like effector (TALE) that provides the capability and specificity for binding target DNA sequence, and a non-specific restriction endonuclease Fok I domain fused to the C-terminal of TALE that confers the nuclease activity of TALENs. Here, we summarize the technological back-ground and provide an overview of development and applications of TALENs in genome engineering.

Key words TALE; TALENs; gene knockout; targeted genome modification; TALENs construction

# 1 引言

基因组定点修饰技术是对DNA进行改造的重要技术,这项技术从出现至今都是分子生物学研究的热点,并且已在农业和生物医药领域得到广泛应用。基因修饰的方法主要包括:传统的基因打靶技术<sup>III</sup>、Cre/loxP重组酶系统、慢病毒载体技

术、锌指核酸酶技术(zine-finger nucleases, ZNFs)、 CRISPR/Cas技术(clustered regularly interspaced short palindromic repeats/cas, CRISPR/Cas)和最新研究的 TALENs技术<sup>[2]</sup>。这些技术都有各自的优缺点。Cre/ loxP重组酶系统可以进行目的性的基因敲除或者敲 进, 广泛地用于制备实验动物模型, 此系统首先要向

收稿日期: 2013-07-02 接受日期: 2013-08-07

国家自然科学基金(批准号: 31172379)资助的课题

<sup>\*</sup>通讯作者。Tel: 0835-2885312, E-mail: suizhongcao@126.com

Received: July 2, 2013 Accepted: August 7, 2013

This work was supported by the National Natural Science Foundation of China (Grant No.31172379)

<sup>\*</sup>Corresponding author. Tel: +86-835-2885312, E-mail: suizhongcao@126.com

网络出版时间: 2013-10-16 16:58 URL: http://www.cnki.net/kcms/detail/31.2035.Q.20131016.1658.001.html

基因组中插入loxP序列,因此不适合于基因治疗和 制备转基因大动物[3]。慢病毒载体技术具有整合时间 短、整合效率高的优点,但其随机整合、病毒基因的 安全性在操作过程中对目的基因的修饰都有很大的 影响<sup>[4]</sup>。ZFNs的出现被认为是基因改造技术中的一 次飞跃。ZFNs介导的基因治疗方法已经进入了临床 试验阶段,其优点在于实现了基因的靶向操作<sup>[5]</sup>,大 大提高基因组靶向修饰的效率,但是目前设计和筛 选高效率、高特异性的ZFNs仍然很难实现。制备 ZFNs所需的工作量大、周期长、成本高,而且很难 降低ZNFs的细胞毒性<sup>60</sup>,所以需要更加高效、简单 的基因组定点修饰技术。CRISPR/Cas技术能够在 DNA链特异位点进行基因的敲入、敲除,并且能同 时导入多种遗传改变,同时具有构建简单、周期短、 成本低等优点。2013年Fu等<sup>[7]</sup>指出, CRISPR/Cas进 行DNA链特异性切割的过程中, 会在预期靶点以外 的位点上生成多余的DNA突变,并且脱靶位点的突 变率有可能和靶位点一样高,或甚至更高;而最近 Hsu等<sup>[8]</sup>验证, 通过优化RNA片段的结构和调整导向 RNA序列的量可以大大降低CRISPR/Cas的脱靶效 应。2009年, Moscou等<sup>[9]</sup>和Boch等<sup>[10]</sup>同时研究发现, 类转录激活因子效应物(transcription activator-like effector, TALE)具有一个重复单元识别一个DNA 碱基的特异性,并且这种蛋白便于预测和构建。在 随后的几年里科学家们从ZNFs中受到启发,构建 出基于TALE的人工核酸内切酶并且广泛用于斑马 鱼[11]、水稻[12]、线虫[13]、果蝇[14]、青枯雷尔氏菌 (Ralstonia solanacearum)<sup>[15]</sup>、热带爪蟾<sup>[16]</sup>、猪和牛<sup>[17]</sup> 等多个物种,以及用于体外培养的人体细胞基因组 定点修饰<sup>[18-20]</sup>。TALENs技术具有和目的DNA特异 性结合、脱靶率和细胞毒性比ZFNs低以及比ZFNs 更加容易设计的优点,对于基因组的定点修饰技术 和遗传学的研究来说,又是一个重大的飞跃。

## 2 TALEs的结构

1989年, Bonas等<sup>[21]</sup>首先在一种植物病原菌黄单 胞菌(*Xanthomonas*)中分离得到TALE家族中的第一 个成员AvrBs3,并成功克隆其基因。Kay等<sup>[22]</sup>在研 究黄单胞菌致病机制时发现黄单胞菌合成的TALE 蛋白AvrBs3能够进入辣椒的细胞核,然后结合到 *upa20*基因的启动子上激活*upa20*基因从而控制细胞 的大小。在随后的几年中,相继有人报道PthXo1、 AvrXa27等TALE蛋白能够结合到宿主基因的启动子上,从而调控相应基因的表达<sup>[10]</sup>。TALE对DNA的识别方式类似真核生物的转录因子,通过识别并结合特异的DNA序列调控植物内源基因的表达,使得宿主植物对病原体的敏感性提高<sup>[23]</sup>。

TALE主要由3个部分组成: C-端含有一个核定 位信号(nuclear localization signal, NLS)和转录激活 结构域(activation domain, AD)、N-端一般含有转运 结构域(translocation domain, TD)、而中间是能够 和DNA进行特异性结合的结构域(图1A)。中间的 DNA结合结构域是一段特别长的重复氨基酸序列, 不同种类的TALE蛋白的中间结构域是由1.5~33.5 个TALE单元组成的重复氨基酸序列组成,每个 TALE单元由33~35个氨基酸残基组成[23]。在一个重 复的氨基酸序列中每个重复单元和末尾的半个重复 单元均可特异性地识别并结合DNA的一个特定核 苷酸位点<sup>[9-10]</sup>。例如TALE-AvrXs10,它结合水稻细 胞内19 bp的DNA序列只需要17.5个重复单元,在 TALE蛋白识别的靶序列中5′端第一个碱基为T,不 需要TALE蛋白单元的结合, 而最后一个碱基则由 TALE最后的0.5个重复单元约20氨基酸结合[10,24]。

在TALE蛋白的每个重复单元中,除了12和13 位的氨基酸残基可变以外,其他的氨基酸残基都是 一样的,这两个可变氨基酸残基是TALE蛋白特异性 识别DNA碱基的关键位点, 被称为重复可变双残基 (repeat variable di-residue, RVD)(图1A)。TALE蛋 白中的RVD决定了其对DNA序列的特异性识别, 不同的RVD能够特异性地识别碱基A、T、C、G 中的一种或者多种。目前发现的RVD种类有5种, 分别是HD(氨基酸名称)特异识别C、NI(氨基酸名 称)识别A、NN(氨基酸名称)识别G或A、NG(氨 基酸名称)识别T、NK(氨基酸名称)可以识别A、 T、C、G中的任一种<sup>[9-10,25]</sup>(图1B)。2013年, Meckler 等<sup>[26]</sup>报道了这5种RVD对识别DNA序列的贡献依次 为NG>HD~NN>>NI>NK。NN对G的亲和性大于 A,并且NN识别G的能力是NK的1000倍。由此表 明, TALE蛋白的重复单元跟DNA碱基之间有更好的 一一对应性,因此在识别靶点的特异性方面比ZFN 更有优势。

#### 3 TALENs技术原理

2011年, Li等<sup>[27]</sup>将TALEs与核酸酶结合构建

出能够对目的基因进行定点修饰的TALENs,首次展示了这种蛋白在基因组工程学方面的应用潜力。TALE核酸酶的构建是在ZNFs的基础上产生的,TALENs由TALEs的DNA结合结构域和来自细菌的II S型核酸内切酶Fok I的切割域融合而成(图2)。TALEs的DNA结合结构域识别DNA链上的目标

靶点, Fok I形成二聚体后才能发挥其核酸内切酶的活性。因此, TALENs和ZNFs一样通常都是成对使用的。成对的TALENs同时作用就可以对细胞内的DNA双链进行切割,造成DNA双链的断裂(double-strand break, DSB)。DSB能够激活细胞内固有的同源重组(homologous recombination, HR)或非同源末



A: structure of TALE protein. TD: translocation domain; NLS: nuclear localization signal; AD: activation domain; B: types of RVD.

ructure of TALE protein. TD. transformula domain, NES. nuclear localization signal, AD. activation domain, B. types of KVF

图1 TALE的结构和基本的识别序列









图3 TALENs进行基因敲除的原理(根据参考文献[30]修改) Fig.3 TALENs principle of gene knockout (modified from reference [30])

端连接(non-homologous end-joining, NHEJ)机制,进行DSB的修复<sup>[28]</sup>(图3)。当发生NHEJ时,会引入或者缺失一些DNA片段,从而引起新的突变,达到基因敲除的目的。当存在同源序列时,就会进行HR修复,不过这种修复的效率很低<sup>[29]</sup>。因此,TALENs在理论上能够用于对任何物种DNA链上的任何位点进行定点突变,实现物种的遗传突变。

### **4 TALENs**的构建

TALENs的构建过程主要包括三个步骤:(1)从 选定的目的基因序列中找出适合TALE蛋白结合的 目标靶点;(2)制备针对目的靶点的特异TALE蛋白, 这也是TALENs构建过程中最复杂和重要的环节; (3)在TALE蛋白上选择合适的Fok I结构域,利用分 子生物学的方法把Fok I和TALE组装在一起构建完 整的TALE人工核酸酶。

#### 4.1 TALENs靶位点的选择

TALENs靶点的选择是进行TALE蛋白设计的 关键。2011年, Cermak等<sup>[31]</sup>提出了选择TALENs靶 点的原则: (1)TALENs靶位点5′端的第一位碱基不能 是T; (2)靶位点的第二位碱基不能是腺嘌呤(A); (3) 靶序列的前一位碱基(5′端碱基)应该为T; (4)靶序列 的最后一个碱基(3′端)不是G(对应于最后的0.5个 TALE重复单元); (5)在TALENs靶位点中4种碱基各 占一定比例, 在选择的潜在靶序列中尽量遵从A占 31%±16%, C占37%±13%, G占9%±8%, T占22%±10% 的原则。2012年, Reyon等<sup>[32]</sup>指出TALENs靶点的选 择只需要遵从上述第三条原则。除了这些原则以外, 为了保证靶点在整个基因组中存在唯一性, 靶序列 一般会选择大于10 bp, 并且两个靶序列之间的间距 长度也会显著地影响TALENs活性。2010年, Christian等<sup>[33]</sup>还曾经提出两个靶位点之间的间隔应该控 制在13~30 bp之间, 其中15 bp时TALENs的活性最 高。

在TALENs潜在靶点的选择过程中通常可以利 用其他实验室建立的网站或软件进行预测或设计。 现在使用的设计网站主要有Bogdanove和Voytas实 验室最早建立的TALE-NT(https://tale-nt.cac.cornell. edu/),此网站可以选择两侧靶位点碱基的数量和两 侧位点之间的间隔长度,并且还能预测出间隔中存 在的限制性酶切位点。此外,还有Joung实验室建立 的ZiFiT网站(http://zifit.partners.org/ZiFiT/)<sup>[34]</sup>、Zhu 实验室建立的idTALE网站(http://idtale.kaust.edu. sa/)<sup>[35]</sup>、张博实验室建立的人工核酸内切酶(engineered endonuclease, EEN)综合数据库与知识库网站 EENdb(http://eendb.zfgenetics.org/)<sup>[36]</sup>以及Doyle设计 的TALE-NT 2.0软件<sup>[37]</sup>。

#### 4.2 TALE蛋白重复序列的构建

在靶点的设计过程中,为了保证设计的靶位点 在整个基因组中的唯一性,选择的靶序列一般会大 于10 bp, 这样就要求TALE至少需要含有9.5个重复 单元,编码这一重复结构的重复核苷酸序列大于1Kb。 因此、构建能够识别特定DNA序列的TALE是TALENs 技术中最关键也是最复杂的一步。围绕这一问题科 学家们发明了多种解决方法,最初人们应用全序列人 工合成这一最昂贵的方法,除了这一方法外还可以通 过分子生物学的方法进行构建。这些分子生物学的 方法主要有:限制性酶切连接法(restriction enzyme and ligation, REAL)<sup>[34]</sup>、idTALE一步酶切次序连接法<sup>[35]</sup>、 基于PCR的GG法(GG-PCR)<sup>[38-39]</sup>、传统的基于质粒 载体的GG法(GG-Vector)<sup>[27]</sup>、高通量TALENs合成技 术FLASH(fast ligation-based automatable solid-phase high-throughput)<sup>[32]</sup>、重复加帽ICA(iterative capped assembly)组装法<sup>[40]</sup>、长黏末端的LIC(ligation-independent cloning)组装法<sup>[41]</sup>以及有我国学者张博报道 的单元组装法(unit assembly, UA)<sup>[42]</sup>。

REAL法是Sander等<sup>[34]</sup>建立的,他们在构建 TALE的过程中不但要用PCR克隆出能够识别目的碱 基的TALE单元,而且使每个TALE基本单元质粒的5′ 端都带有II S型内切酶Bbs I的识别位点,3′端都带有 另一个II S型内切酶Bsa I和普通内切酶BamH I的识 别位点。在进行组装的时候分别用Bsa I和BamH I对 前一个TALE单元进行切割,Bbs I和BamH I对后一 个TALE单元进行切割,形成一对前后互补的4 bp黏 性末端和一对序列不变的BamH I黏性末端,然后连 接成为双单元的TALE骨架。这些双单元质粒中仍 然具有基本单元中的3个酶切位点,这样就可以进行 重复的酶切-连接,构建目的长度目的序列的TALE 蛋白表达载体。这种方法只需要三种酶就可以完成 组装,一周之内能完成目的TALE的构建,但是这个 方法酶切-连接的步骤非常多。

Li等<sup>[35]</sup>介绍的idTALE一步酶切次序连接法,组 装dTALE主要通过利用II型限制性内切酶PpuMI、 BsmAI、BsmBI、XhoI和SacI,这些限制性内切酶 用于目的序列的酶切-连接反应。构建过程中首先 要选择含有dHax3的pENTR221质粒载体,根据设计 的TALE重复序列选择含有7种片段(fragment)的质 粒,每种片段受到如下的限制:片段1(F1)192 bp,并 且含有PpuMI、BsmAI酶切位点;F2、F3、F4、F5 都含有BsmAI切割位点,长度分别为234 bp、161 bp、 193 bp和204 bp;F6长度为207 bp,含有XhoI、BsmAI 酶切位点;F7长度为541 bp,含有XhoI、SacI切割位 点,把含有相同切割位点的TALE序列与骨架质粒进 行混合后切割,就会形成互补的黏性末端,然后用连 接酶连接就构成了dTALE序列。

基于PCR的GG法(GG-PCR法)<sup>[38-39]</sup>和传统的基 于质粒载体的GG法(GG-Vector法)<sup>[27]</sup>都是在Golden Gate方法的基础上建立起来的, 主要是通过利用II S 型限制性核酸内切酶(例如 Bsa I、BsmB I和Bbs I)的 识别位点与切割位点相分离的特点,通过使用同一 个II S型内切酶、设计不同的黏性末端来实现一次 酶切--连接反应再按照设定的顺序组装TALE重复单 元。GG-PCR法主要是利用了TALE的两个相邻重复 单元交界处的Gly-Leu双氨基酸残基密码子的简并 性来设计不同TALE重复单元的交界序列,作为设计 特异引物和特异连接的黏末端位点<sup>[39]</sup>。GG-Vector 法基于质粒载体,用II S型内切酶产生多种一一对 应、特异匹配的黏性末端对,一次可以连接8~10个 TALE重复单元。例如在实验中可以利用TALE重复 单元中18、19位Ala-Leu氨基酸密码子的简并性(由 于编码这两种氨基酸的密码子有8个)就可以构建出 含有上一个重复单元18位到下一个重复单元19位氨 基酸残基序列的质粒,并在紧邻这一编码区外侧设 计一个BsmB I的酶切位点。再经过BsmB I酶切在每 个质粒上产生一个4 bp的5′黏性末端与含目的碱基 的模版匹配,并加入与第一个模块前端及第八个模 块后端的黏性末端相匹配的质粒骨架得到8个单元 的重复序列。经过两次或三次这样的组装可以获得 能够特异性识别16或24个碱基的TALE重复序列<sup>[27]</sup>。

FLASH组装法是在己有的376个载体基础上, 把这些载体放在96孔板的固相基质中,首先将第一 个TALE蛋白单元生物素化,然后再将生物素化的 TALE蛋白单元与磁珠相连,这样就可以实现TALE 蛋白单元的固定化,随后的酶切酶连反应就可以在 自动化仪器中快速完成。FLASH组装法优点在于省 略每一次酶切-连接之后的细菌转化和生长这一耗 时的中间步骤,从而实现快速、高通量地构建TALE 重复序列<sup>[32]</sup>。

George实验室建立的ICA法是将TALE的单体 连接在固体支持物上,并且用帽子结构阻止不完整 链的产生,从而提高dTALE的组装成功率<sup>[40]</sup>。在实 际操作中,每个选定的TALE单体都被BsmB I切割产 生前后互补的黏性末端,用于连接出任意长度、任 意序列的TALE序列。以A-B-C链构建为例,首先将 第一个起始单元(initiator)生物素化,然后与链霉素 亲和包被的磁珠相连,在这个起始单元的右端设计 出一个和C单元的右端相同的黏性末端,并且这个 末端可以和A单元的左端相连,构建出initiator-A;第 二,为了保证所有的起始单元都会连接A单元,进行 第二步连接时在加入B单元之前添加一个能够和起 始单元末端相匹配的帽子结构,这样就能保证B单 元连接到A单元的右端,形成initiator-A-B中间产物, 而剩余的未连接的单体被帽子结构封闭;第三,做 同样的处理后再添加C单元,保证连接成initiator-A-B-C,重复上面的步骤形成A-B-C-A-B-C的序列,如 果随意地选择TALE单元进行连接就可以构建出任 意长度、任意TALE单元组成的序列。ICA法能够 快速地组装TALE序列,并且能够提高TALE的成功 率,实现高通量的组装。

LIC组装法<sup>[41]</sup>首先设计出多种TALE单元的两端都含有10~30 bp的长末端突出碱基,利用能识别不同碱基的TALE单元构建出多种组合的2单元中间结构,分别标记为1/2、2/3、3/4、4/1(如图4);第二,我们可以根据需要识别的序列从2单元结构中选择合适的组合在DNA聚合酶的作用下进行退火连接,然后与含有卡拉霉素基因的骨架相连,最后在将组装好的TALE序列与哺乳动物的表达启动子相连就构建好了TALE表达载体。例如,构建一个含有18个

TALE单元的识别序列,需要进行两级组装。第一级 组装,从2单元结构中选择3个2单元结构组成6单元 结构,在构建的6单元结构上添加一个含有卡拉霉 素基因的骨架。第二级组装,将第一级组装中构建 好的3个6单元组装进行连接,组成18单元的TALE序 列,然后再把这个序列构建在含有巨噬细胞启动子 (pCMV)、卡拉霉素抗性基因(ampR)、核定位信号 (nuclear localization signal, NLS)、起始密码子ATG、 终止密码子和T7(促进序列与T7 RNA聚合酶识别), 以及*Xho* I、*Xba* I和*Not* I酶切识别位点的质粒上,这 样就成功构建了TALE核酸酶编码序列。利用这种 方法可以构建任意长度和识别任意DNA碱基序列 的TALE核酸酶(图5)。

单元组装法属于连续克隆组装法,是2011年由 我国科学家张博等<sup>[11]</sup>建立的。该方法的起始材料只 需要分别对应于识别4个碱基的4个TALE基本单元 模块,采用3个常见的限制性内切酶*Spe* I、*Nhe* I和 *Hind* III,通过简单的酶切-连接反应,即可串联组装 出任意长度的TALE重复序列。由于TALE重复序列 是首尾相接的,因此,他们利用了这一特性可以将重 复单元中任意两个相邻氨基酸残基作为串联重复序 列的起点和终点进行组装,最后再将构建的TALE序 列的首尾重复单位的两端补齐,成功构建了完整的 TALE重复序列。根据这一原理能构建任意的替代



图5 含有表达元件的TALEN序列(根据参考文献[41]修改) Fig.5 Sequence of TALEN containing the expression unit (modified from reference [41])

重复单元,进而利用氨基酸密码子的简并性在替代 重复单元的C-端、N-端设计一对同尾酶(*Nhe* I、*Spe* I) 的切割位点,再加上DNA序列中普遍存在的*Hind* III 位点,利用同尾酶可以产生相同的黏性末端,当连接 在一起后又会失去原有的酶切位点这一特性。通过 反复的酶切-连接反应,可以得到目的TALE蛋白的 DNA序列。

在上述的这些方法中,有的需要合成大量相似的质粒和相对应的引物,操作起来非常的繁琐,并且容易出错,而且需要大量价格高昂的II S类酶。虽然FLASH和ICA能够高通量和快速合成目的TALE编码序列,但是该项技术对设备的要求很高,FLASH高速合成法不但需要准备大量的质粒片段,而且还要对合成的模块进行测序才能找到正确的连接终产物<sup>[43]</sup>。单元组装法具有成本低、操作简单、技术难度低、精确度和保真度高、所需的起始载体量少、载体有累加效应等优点,当前成为一般实验室首选的方法。

## 4.3 TALE蛋白序列与Fok I的组装

用构建的TALEs序列和能够以二聚体形式起作 用的Fok I内切酶序列进行组装,就可以构建出完整 的能够特异性识别目的基因并进行切割的TALENs 序列。Miller等<sup>[44]</sup>指出重复序列跟Fok I之间的距离 (即保留来自TALE蛋白C-末端区域的肽段长度)范 围可以为28~95个氨基酸残基,其中相距63个氨基 酸残基的效率较高。2013年Joshua等<sup>[26]</sup>撰文指出, TALENs在识别目的序列的时候存在极性效应, N-端重复序列的识别亲和力比C-端强,这也是在构建 TALENs时应该注意的。

# 5 TALENs诱变效率的检测方法

应用TALENs进行基因组定点修饰后的诱变效 率检测方法主要有限制性内切酶酶切检测法[11]、基 于可以识别错配双链的错配内切酶检测法[44]、直接 克隆测序法[34]、体内或体外引入报告基因法以及体 外的SSA(single strand annealing)分析法[2]。其中限 制性内切酶酶切检测法是TALENs诱变效率检测中 最简单易行的方法,其作用原理是根据TALENs造 成的突变会导致靶位点之间的Spacer中原有酶切位 点被破坏。使用这种方法就必须在设计靶点时要 求选择的Spacer中存在一个独有的限制性内切酶位 点。检测时,通过提取用TALENs处理后细胞的基 因组DNA, 然后设计一对引物, 运用PCR扩增出包含 TALENs靶点区域的DNA片段, 然后用内切酶切割 扩增产物并进行电泳。若是野生型DNA Spacer中原 有酶切位点未被破坏, DNA链会被完全切割, 产生1-2 条带;突变型DNA由于Spacer中原有酶切位点被破 坏,只会产生一条带,并且通过计算未切开条带的亮 度占总条带的比值,就能够反映TALENs的活性<sup>[11]</sup>。

# 6 TALENs技术在基因组定点修饰中的应用

从2007年Boch发现TALE能够进入细胞核识别 upa20基因的启动子进而调控upa20基因的表达开 始,到2009年破译了TALE的一个重复单元一个碱基

| 物种                       | 靶标基因                            | 突变率           | 参考文献       |
|--------------------------|---------------------------------|---------------|------------|
| Species                  | Target gene                     | Mutation rate | References |
| Homo sapiens             | NTF3, CCR5                      | 25%           | [44]       |
|                          | VKORC1L1                        | /             | [20]       |
|                          | OCT4                            | 70%~100%      | [45]       |
| Sus scrofa               | LDLR, GHRHR                     | 40%           | [17]       |
| Bos taurus               | ACAN, GDF8                      | 25%           | [17]       |
| Rabbits                  | RAG                             | 90%           | [46]       |
| Rattus norvegicus        | IgM                             | 59%           | [47]       |
|                          | BMPR2                           | /             | [48]       |
|                          | Y chromosome                    | /             | [49]       |
| Danio rerio              | TIA1, GSK3B, RGS4, TPH1A, DRD3R | 70%           | [43]       |
|                          | TH                              | 40%           | [50]       |
| Arabidopsis thaliana     | RD29A                           | /             | [51]       |
| Saccharomyces cerevisiae | URA3, LYS2 and ADE2             | 4.5%~27%      | [27]       |

表1 TALEN技术在基因组定点修饰中的应用 Table 1 TALEN application in the targeted genome modification

的"密码"以来, TALENs技术已经应用到人体细胞和 拟南芥、酵母、斑马鱼、小鼠和大鼠等模式生物, 以及家兔、猪、牛等大动物的基因修饰(表1)。

# 7 展望

TALENs技术作为一种新兴的基因打靶技术, 从一出现就受到极大的关注。TALENs神奇之处在 于其一个重复单元识别一个碱基,这是人类迄今为 止发现的最为简单的一种核酸与蛋白相互作用方 式。利用这一特性构建的TALENs已经应用到多种 生物体内,并且有非常好的基因组定点修饰效果。 但是, TALENs的运用依然存在一些问题, 例如: (1) 如果需要设计识别20个以上碱基的TALENs, 就需要 设计出至少2 Kb的序列,并且这些序列有高度的重 复性,因此在设计的时候存在非常大的难度,并且这 么长的TALENs导入细胞以后是否会引起细胞产生 免疫反应,以及影响TALENs在细胞中的作用还不是 很清晰; (2)在目前的报道中, TALENs主要应用于基 因的敲除,但是能否进行大片段基因的插入仍然需 要进一步的研究; (3)TALE重复序列在什么长度下最 适合用于DNA序列的识别, 以及组装成的TALENs 是否具有最优的活性仍然不清楚; (4)如何更好地解 决TALENs脱靶的问题; (5)为什么TALENs在识别 DNA序列的时候存在极性以及为什么TALENs N-端 识别DNA序列的能力会比C-端强? (6)虽然说TAL-ENs的毒性比ZNFs小,但是TALENs对细胞和个体会 产生什么样的毒性仍然不清楚。

这些问题都制约着TALENs技术的应用,但是 TALENs和ZFNs相比有明显的优势:(1)TALENs制 备和筛选更加的简单,只需要预先设计出靶位点,然 后根据靶位点利用普通的分子生物学的方法就能 快速地构建出高效的TALENs,而ZFNs则需要花费 大量的精力和财力进行制备;(2)由于TALENs是一 个重复单元一个碱基的模式进行识别,因此它的识 别率更高,脱靶率更低;(3)由于ZNFs进入细胞以后 *Fox*I形成的同源二聚体就会产生细胞毒性,影响基 因组定点修饰的成功率,而TALENs对细胞的毒性 比ZNFs小;(4)TALENs结合DNA序列更为严格,因 此效率比ZNFs高<sup>[47]</sup>。因此,我们有理由相信利用 TELANs的这些优点在基础理论研究、基因治疗以 及农业转基因动植物等领域能大大提高研究的效率 并产生深远的影响。

#### 参考文献 (References)

- Capecchi MR. Altering the genome by homologous recombination. Science 1989; 244(4910): 1288-92.
- 2 Li T, Huang S, Jiang WZ, Wright D, Spalding MH, Weeks DP, et al. TAL nucleases (TALNs): Hybrid proteins composed of TALeffectors and FokI DNA-cleavage domain. Nucleic Acids Res 2011; 39(1): 359-72.
- 3 Van Duyne GD. A structural view of cre-loxp site-specific recombination. Annu Rev Biophys Biomol Struct 2001; 30: 87-104.
- 4 Cao F, Xie X, Gollan T, Zhao L, Narsinh K, Lee RJ, *et al.* Comparison of gene-transfer efficiency in human embryonic stem cells. Mol Imaging Biol 2010; 12(1): 15-24.
- 5 Whyte JJ, Zhao J, Wells KD, Samuel MS, Whitworth KM, Walters EM, *et al.* Gene targeting with zinc finger nucleases to produce cloned eGFP knockout pigs. Mol Reprod Dev 2011; 78(1): 2.
- 6 Gaj T, Gersbach CA, Barbas CR. ZFN, TALEN, and CRISPR/ Cas-based methods for genome engineering. Trends Biotechnol 2013; 31(7): 397-405.
- 7 Fu Y, Foden JA, Khayter C, Maeder ML, Reyon D, Joung JK, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat Biotechnol 2013; 31(9): 822-6.
- 8 Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol 2013; 31(9): 827-32.
- 9 Moscou MJ, Bogdanove AJ. A simple cipher governs DNA recognition by TAL effectors. Science 2009; 326(5959): 1501.
- 10 Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, *et al*. Breaking the code of DNA binding specificity of TAL-type III effectors. Science 2009; 326(5959): 1509-12.
- 11 Huang P, Xiao A, Zhou M, Zhu Z, Lin S, Zhang B. Heritable gene targeting in zebrafish using customized TALENs. Nat Biotechnol 2011; 29(8): 699-700.
- 12 Li T, Liu B, Spalding MH, Weeks DP, Yang B. High-efficiency TALEN-based gene editing produces disease-resistant rice. Nat Biotechnol 2012; 30(5): 390-2.
- 13 Wood AJ, Lo TW, Zeitler B, Pickle CS, Ralston EJ, Lee AH, et al. Targeted genome editing across species using ZFNs and TALENs. Science 2011; 333(6040): 307.
- 14 Liu J, Li C, Yu Z, Huang P, Wu H, Wei C, *et al.* Efficient and specific modifications of the *Drosophila* genome by means of an easy TALEN strategy. J Genet Genomics 2012; 39(5): 209-15.
- 15 de Lange O, Schreiber T, Schandry N, Radeck J, Braun KH, Koszinowski J, *et al.* Breaking the DNA-binding code of Ralstonia solanacearum TAL effectors provides new possibilities to generate plant resistance genes against bacterial wilt disease. New Phytol 2013; 199(3): 773-86.
- 16 Nakajima K, Nakai Y, Okada M, Yaoita Y. Targeted gene disruption in the *Xenopus tropicalis* genome using designed TALE nucleases. Zoolog Sci 2013; 30(6): 455-60.
- 17 Carlson DF, Tan W, Lillico SG, Stverakova D, Proudfoot C, Christian M, *et al*. Efficient TALEN-mediated gene knockout in livestock. Proc Natl Acad Sci USA 2012; 109(43): 17382-7.
- 18 Bultmann S, Morbitzer R, Schmidt CS, Thanisch K, Spada F, Elsaesser J, *et al.* Targeted transcriptional activation of silent oct4 pluripotency gene by combining designer TALEs and inhibition of epigenetic modifiers. Nucleic Acids Res 2012; 40(12): 5368-77.

- 19 SunN, Liang J, Abil Z, Zhao H. Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. Mol Biosyst 2012; 8(4): 1255-63.
- 20 Tie JK, Jin DY, Tie K, Stafford DW. Evaluation of warfarin resistance using TALENs-mediated vitamin K epoxide reductase knockout HEK293 cells. J Thromb Haemost 2013; 11(8): 1556-64.
- 21 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): 127-36.
- 22 Kay S, Hahn S, Marois E, Hause G, Bonas U. A bacterial effector acts as a plant transcription factor and induces a cell size regulator. Science 2007; 318(5850): 648-51.
- 23 Boch J, Bonas U. Xanthomonas AvrBs3 family-type III effectors: Discovery and function. Annu Rev Phytopathol 2010; 48: 419-36.
- 24 Romer P, Hahn S, Jordan T, Strauss T, Bonas U, Lahaye T. Plant pathogen recognition mediated by promoter activation of the pepper Bs3 resistance gene. Science 2007; 318(5850): 645-8.
- 25 Morbitzer R, Römer P, Boch J, Lahaye T. Regulation of selected genome loci using denovo-engineered transcription activator-like effector (TALE)-type transcription factors. Proc Natl Acad Sci USA 2010; 107(50): 21617-22.
- 26 Meckler JF, Bhakta MS, Kim MS, Ovadia R, Habrian CH, Zykovich A, *et al.* Quantitative analysis of TALE-DNA interactions suggests polarity effects. Nucleic Acids Res 2013; 41(7): 4118-28.
- 27 Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, et al. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Res 2011; 39(14): 6315-25.
- 28 Allen C, Kurimasa A, Brenneman MA, Chen DJ, Nickoloff JA. DNA-dependent protein kinase suppresses double-strand breakinduced and spontaneous homologous recombination. Proc Natl Acad Sci USA 2002; 99(6): 3758-63.
- 29 Rouet P, Smih F, Jasin M. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci USA 1994; 91(13): 6064-8.
- 30 张金脉, 任兆瑞. TALENs: 一种新的基因定点修饰技术. 生命 科学(Zhang Jinmai, Ren Zhaorui. TALENs: A new genome sitespecific modification technology. Chinese Bulletin of Life Sciences) 2013; 25(12): 126-31.
- 31 Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, 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.
- 32 Reyon D, Tsai SQ, Khayter C, Foden JA, Sander JD, Joung JK. FLASH assembly of TALENs for high-throughput genome editing. Nat Biotechnol 2012; 30(5): 460-5.
- 33 Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, *et al.* Targeting DNA double-strand breaks with TAL effector nucleases. Genetics 2010; 186(2): 757-61.
- 34 Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, et al. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. Nat Biotechnol 2011; 29(8): 697-8.
- 35 Li L, Piatek MJ, Atef A, Piatek A, Wibowo A, Fang X, *et al.* Rapid and highly efficient construction of TALE-based transcriptional regulators and nucleases for genome modification. Plant Mol Biol 2012; 78(4): 407-16.

- 36 Xiao A, Wu YD, Yang Z, Hu Y, Wang W, Zhang Y, et al. EENdb: A database and knowledge base of ZFNs and TALENs for endonuclease engineering. Nucleic Acids Res 2013; 41: D415-22.
- Doyle EL, Booher NJ, Standage DS, Voytas DF, Brendel VP, Vandyk JK, *et al.* TAL Effector-Nucleotide Targeter (TALE-NT)
  2.0: Tools for TAL effector design and target prediction. Nucleic Acids Res 2012; 40: W117-22.
- 38 Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P. Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. Nat Biotechnol 2011; 29(2): 149-53.
- 39 Sanjana NE, Cong L, Zhou Y, Cunniff MM, Feng G, Zhang F. A transcription activator-like effector tool box for genome engineering. Nat Protoc 2012; 7(1): 171-92.
- 40 Briggs AW, Rios X, Chari R, Yang L, Zhang F, Mali P, 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.
- 41 Schmid-Burgk JL, Schmidt T, Kaiser V, Honing K, Hornung V. A ligation-independent cloning technique for high-throughput assembly of transcription activator-like effector genes. Nat Biotechnol 2013; 31(1): 76-81.
- 42 沈 延, 黄 鹏, 张 博. TALEN 构建与斑马鱼基因组定点突变 的实验方法与流程. 遗传(Shen Yan, Huang Peng, Zhang Bo. A protocol for TALEN construction and gene targeting in zebrafish. HEREDITAS) 2013; 35(4): 533-44.
- 43 沈 延,肖 安,黄 鹏, 王唯晔, 朱作言,张 博. 类转录激活因 子效应物核酸酶(TALEN)介导的基因组定点修饰技术. 遗传 (Shen Yan, Xiao An, Huang Peng, Wang Weiye, Zhu Zuoyan, Zhang Bo. TALE nuclease engineering and targeted genome modification. HEREDITAS) 2013; 35(4): 395-409.
- Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, *et al.* A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 2011; 29(2): 143-8.
- 45 Hockemeyer D, Wang H, Kiani S, Lai CS, Gao Q, Cassady JP, *et al*. Genetic engineering of human pluripotent cells using TALE nucleases. Nat Biotechnol 2011; 29(8): 731-4.
- 46 Song J, Zhong J, Guo X, Chen Y, Zou Q, Huang J, *et al.* Generation of RAG 1- and 2-deficient rabbits by embryo microinjection of TALENs. Cell Res 2013; 23(8): 1059-62.
- 47 Tesson L, Usal C, Ménoret S, Leung E, Niles BJ, Remy S, *et al.* Knockout rats generated by embryo microinjection of TALENs. Nat Biotechnol 2011; 29(8): 695-6.
- 48 Tong C, Huang G, Ashton C, Wu H, Yan H, Ying QL. Rapid and cost-effective gene targeting in rat embryonic stem cells by TALENs. J Genet Genomics 2012; 39(6): 275-80.
- 49 Wang H, Hu YC, Markoulaki S, Welstead GG, Cheng AW, Shivalila C. TALEN-mediated editing of the mouse Y chromosome. Nat Biotechnol 2013; 31(6): 530-2.
- 50 Zu Y, Tong X, Wang Z, Liu D, Pan R, Li Z, *et al.* TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nat Methods 2013; 10(4): 329-31.
- 51 Mahfouz MM, Li L, Piatek M, Fang X, Mansour H, Bangarusamy DK, et al. Targeted transcriptional repression using a chimeric TALE-SRDX repressor protein. Plant Mol Biol 2012; 78(3): 311-21.