

doi: 10.3969/j.issn.0490-6756.2018.02.033

# Cpf1 和 Cas9 核酸酶靶向 EGFR-L858R 突变的研究

魏恒<sup>1</sup>, 杨梅佳<sup>2</sup>, 钟坤宏<sup>2</sup>, 仝爱平<sup>2</sup>

(1. 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 610065;

2. 四川大学华西医院肿瘤中心肿瘤免疫治疗实验室, 成都 610041)

**摘要:** 表皮生长因子受体(EGFR)单核苷酸突变(2573T>G, L858R)占有EGFR突变的90%。使突变的EGFR失活对有此突变的病人非常有利。这里,应用双荧光报告分析的方法分析规律成簇间隔短回文重复(CRISPR)系统中Cpf1和Cas9在靶向EGFR-L858R突变的编辑效率。在EGFR-L858R突变位点的附近,有两个Cpf1前间区序列邻近基序(PAMs)——TTTN。并且,2573T>G突变形成了一个Cas9的PAM——NGG。因此本文通过构建两条AsCpf1的gRNAs(gRNA1和gRNA2)和一条SpCas9的gRNA(gRNA3)在体外通过双荧光蛋白分析系统去评估SpCas9和AsCpf1特异性靶向等位基因的能力。结果证实了AsCpf1和SpCas9都能够特异性的编辑突变的EGFR(2573T>G)。

**关键词:** 表皮生长因子; 规律成簇间隔短回文重复; 靶向治疗

**中图分类号:** Q291      **文献标识码:** A      **文章编号:** 0490-6756(2018)02-0401-06

## Targeting EGFR-L858R mutation by Cpf1 and Cas9 nuclease

WEI Heng, YANG Mei-Jia<sup>2</sup>, ZHONG Kun-Hong<sup>2</sup>, TONG Ai-Ping<sup>2</sup>

(1. Key Laboratory of Bio-resources and Eco-environment, Ministry of Education, College of Life Sciences, Sichuan University, Chengdu 610065, China; 2. Laboratory of Tumorimmunity therapy, Cancer Center, West China Hospital, Sichuan University and Collaborative Innovation Center, Chengdu 610041, China)

**Abstract:** Epidermal growth factor receptor(EGFR)single point mutation(2573T>G, L858R) constitute about 90% of all EGFR mutations. Selectively inactivate only mutant, not normal allele, could benefit patients with such mutations. Here, the editing efficacy and selectivity of clustered regularly interspaced short palindromic repeats(CRISPR)Cpf1 and Cas9 systems on EGFR L858R mutant allele were analyzed by dual-reporter assay in vitro. Near the mutation site, there are two TTTN protospacer adjacent motifs (PAMs) for Cpf1. 2573T>G substitution also leads to occurrence of a novel NGG PAM for Cas9. Thus we designed two AsCpf1 gRNA (gRNA1 and gRNA2) and one SpCas9 gRNA (gRNA3) and evaluated their potency and allele specificity in vitro using a dual fluorescent protein-based bioassay system. As a result, both AsCpf1 and SpCas9 demonstrated robust activities to induce specific editing of only 2573T>G mutant EGFR.

**Keywords:** Epidermal growth factor receptor; CRISPR; Targeted therapy

收稿日期: 2017-03-29

基金项目: 国家自然科学基金(31471286)

作者简介: 魏恒(1989-), 男, 广西贵港人, 硕士, 研究方向为蛋白质化学及基因工程. E-mail:476545698@qq.com.

通讯作者: 仝爱平. E-mail:aipingtong@scu.edu.cn

## 1 引言

非小细胞肺癌(Non-small-cell lung cancer, NSCLC)是所有癌症中死亡率较高的癌症<sup>[1]</sup>.在NSCLC中EGFR突变的发生率非常高,并且EGFR的突变通过影响其下游信号以及凋亡机制去调节致癌基因<sup>[2]</sup>.大约有90%的EGFR突变是由于外显子19的删除缺失和L858R的单核苷酸突变引起的,这些突变会引起EGFR的异常激活,使得在核内大量积聚.虽然对于那些EGFR突变的癌症病人来说可以用酪氨酸激酶抑制剂(TKIs)比如,吉非替尼和埃罗替尼来治疗,但是大多数的病人会因长期使用这些药物而使癌症产生抗药性并最终导致癌症的复发<sup>[3]</sup>.例如,大约50%的EGFR突变的病人是由于EGFR-T790M突变引起,这个突变会导致病人对TKIs产生抗药性<sup>[4]</sup>.

随着这几年对CRISPR/Cas9的研究和优化,使得它已经成为了用于基因编辑的一种非常有用的工具.然而,经研究发现从链球菌株(*S. pyogenes*)中分离出来的Cas9(SpCas9)能够容忍gRNAs与其互补的目标DNA之间某些碱基的错配<sup>[5-8]</sup>.但是对于SpCas9的PAM区域的序列是不允许有碱基的错配<sup>[5-8]</sup>.基于这个特点,人们可以去寻找那些在癌症中恰巧突变成为Cas9的PAM序列的位点,设计gRNA使得Cas9特异性识别基因组上的突变位点并破坏突变基因,达到特异性识别并敲除突变癌基因的目的<sup>[9-11]</sup>.

Cpf1是最近报道的依赖RNA导向的核酸酶<sup>[12,13]</sup>.不同于Cas9的PAM区域序列,Cpf1的PAM区域是富含胸腺嘧啶(T)的DNA短序列.在最近的研究中发现Cpf1具有很高的特异性,并且它的脱靶率要比Cas9的更低.此外,Cpf1的gRNAs对靠近Cpf1的PAM序列的5'端区域对单个或者连续两个核苷酸的错配非常敏感<sup>[14,15]</sup>.因此,Cpf1可以作为精准基因修正和基因组编辑的工具.

本研究的目的是检测Cpf1和Cas9这两个核酸酶靶向编辑EGFR-L858R突变的特异性和效率.

## 2 材料与方法

### 2.1 材料

2.1.1 细胞株和质粒 人胚肾细胞株HEK293T购自American Type Culture Collection(ATCC).质粒pcDNA3.1-hAsCpf1、lentiCRISPR V2、phU6-gRNA、pMD2.G、psPAX2购买于Addgene公司;质

粒PDsRed-N1和pEGFP-C1购自Clontech公司.

2.1.2 主要试剂 细胞培养基和胎牛血清购自Hyclone公司;限制性内切酶Kpn I、BamH I、Xba I、Xho I、EcoR I和T4 DNA连接酶购自NEB公司;QuickExtract™ DNA提取溶液购自Epicentre公司;PEI和Polybrene购自Sigma公司;Lipofectamine 3000购于Life Technologies公司.

### 2.2 方法

2.2.1 质粒的构建 hAsCpf1序列插入到lenti-CRISPR V2中Xba I/BamH I两个酶切位点之间构建得到pLenti-hAsCpf1.两条互补的引物退火成双链,连到phU6-gRNA质粒中用于表达AsCpf1和SpCas9的gRNAs.PDsRed-N1质粒用于扩增出RFP序列,然后链接到pEGFP-C1 Kpn I/BamH I酶切位点之间,构建成GFP-Vehicle-RFP双荧光报告质粒.包含有EGFR c.2573T>G突变位点的核酸序列和EGFR c.2573T野生型序列插入到GFP-Vehicle-RFP双荧光报告质粒的Xho I/EcoR I这两个位点之间.

2.2.2 细胞培养、慢病毒的组装和转染条件 HEK293T在含有10%胎牛血清和抗生素的DMEM培养基中培养,培养条件为37℃,5% CO<sub>2</sub>浓度.通过慢病毒的感染和嘌呤霉素筛选出稳定表达Cpf1或Cas9的细胞系(HEK293T).慢病毒的包装在HEK293T中.当10cm细胞培养皿中的HEK293T细胞密度达到80%~90%时,将1 μg pMD2.G, 3 μg psPAX2和4 μg表达Cpf1或Cas9慢病毒质粒与30 μL PEI混合均匀加到HEK293T细胞中.转染16 h之后更换培养基.转染48 h之后收集包含有病毒的上清液,用0.45 μm滤头过滤,20000g,4℃离心2h.病毒沉淀用100 μL无血清无抗生素的DMEM培养基重悬,80℃储存.20 μL的病毒重悬液中加入2 mL培养基,并加入2 μL 8 mg/mL Polybrene.用Lipofectamine 3000将1.25 μg gRNA质粒和1.25 μg双荧光蛋白报告质粒瞬时转染到AsCpf1和spCas9稳定表达的HEK293T(6孔板,细胞密度70%~90%)中,具体操作方法按照Lipofectamine 3000操作说明书.

2.2.3 DNA的提取和测序 DNA的提取使用QuickExtract™ DNA提取溶液.PCR产物的测序在擎科生物技术公司.

2.2.4 FACS分析 特异性靶向切割活性运用FACS技术分析.通过BD FACSAria™ III系统

(Becton Dickinson)统计 RFP 阳性细胞数占 GFP 阳性细胞数的比例. 每个分析重复 3 次. 统计学意义由单因素方差分析决定. P 值小于 0.05 被认为具有统计学意义.

### 3 结果与分析

#### 3.1 靶向 EGFR 2573T>G 基因编辑和双荧光报告系统的示意图

如图 Fig. 1a 所示, 在 EGFR(2573T>G)突变

位点附近, 有两个 AsCpf1 PAM 序列. 因此, 设计两条靶向 EGFR (2573T>G) gRNAs (Supplementary Table 1). EGFR 2573T>G 突变产生了一个 SpCas9 的 PAM 序列 (CGG), 基于这个 Sp-Cas9 PAM 序列设计了一条靶向 EGFR(2573T>G) gRNAs. 将表达 gRNAs 质粒和双荧光报告质粒共转到 AsCpf1 和 SpCas9 稳定表达的 HEK293T 细胞中 (Fig. 1b) 验证 AsCpf1 和 sp-Cas9 对靶向基因的切割效率.

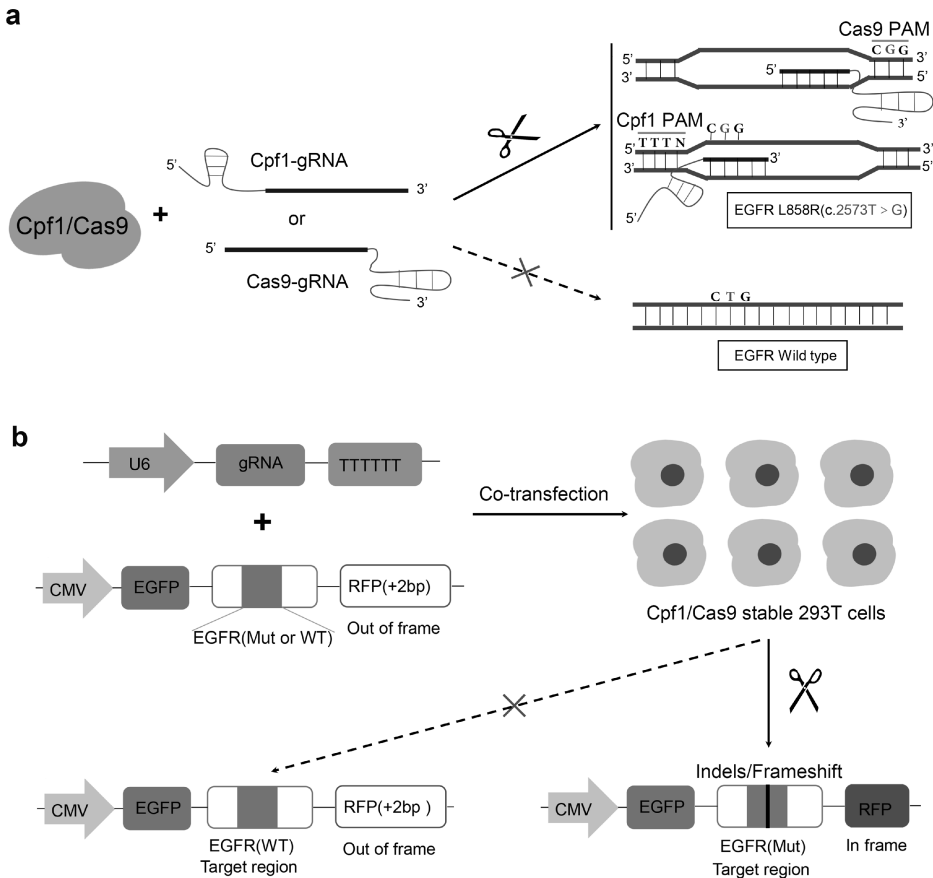


图 1 靶向编辑双荧光报告系统中 EGFR 2573T>G 位点示意图

a. 设计 AsCpf1/SpCas9 的 gRNAs 靶向 EGFR 2573T>G 突变的示意图. 设计了两条 Cpf1 的 gRNAs (gRNA1 和 gRNA2); 基于 EGFR 2573T>G 突变形成的 PAM 序列 CGG 设计了一条 SpCas9 的 gRNA (gRNA3). b. 双荧光报告质粒验证编辑效率的示意图. 双荧光报告质粒带有 CMV 启动子, 并且质粒能编码 EGFP 和 RFP. GRNAs 靶向区域和 PAM 区域序列插入到 EGFP 和 RFP 之间的多克隆位点中, 造成 RFP 移码. 切割之后由于非同源重组修复, 使得 RFP 有几率被重新编码.

Fig. 1 Schematic illustration of EGFR 2573T>G targeted gene editing and the dual fluorescence reporter systems

#### 3.2 AsCpf1 和 SpCas9 核酸酶切割效率的验证

质粒转染到 HEK293T 中 48h 之后, 可以观察到 GFP-vehicle-RFP 空白对照组中 GFP 和 RFP 同时表达在一个细胞中, 而在 GFP-WT-RFP 和 GFP-Mut-RFP 组中并没有观察到 RFP 的表达 (Fig. 2a). 当双荧光报告质粒和 gRNAs 质粒共转

染到 AsCpf1 或 SpCas9 稳定表达 HEK293T 细胞时, AsCpf1 和 SpCas9 都表现出了对突变基因序列的特异性切割 (Figs. 2b-d). FACS 统计分析显示 Cas9-gRNA3 对突变基因序列有更强的编辑效率 (Fig. 2e).

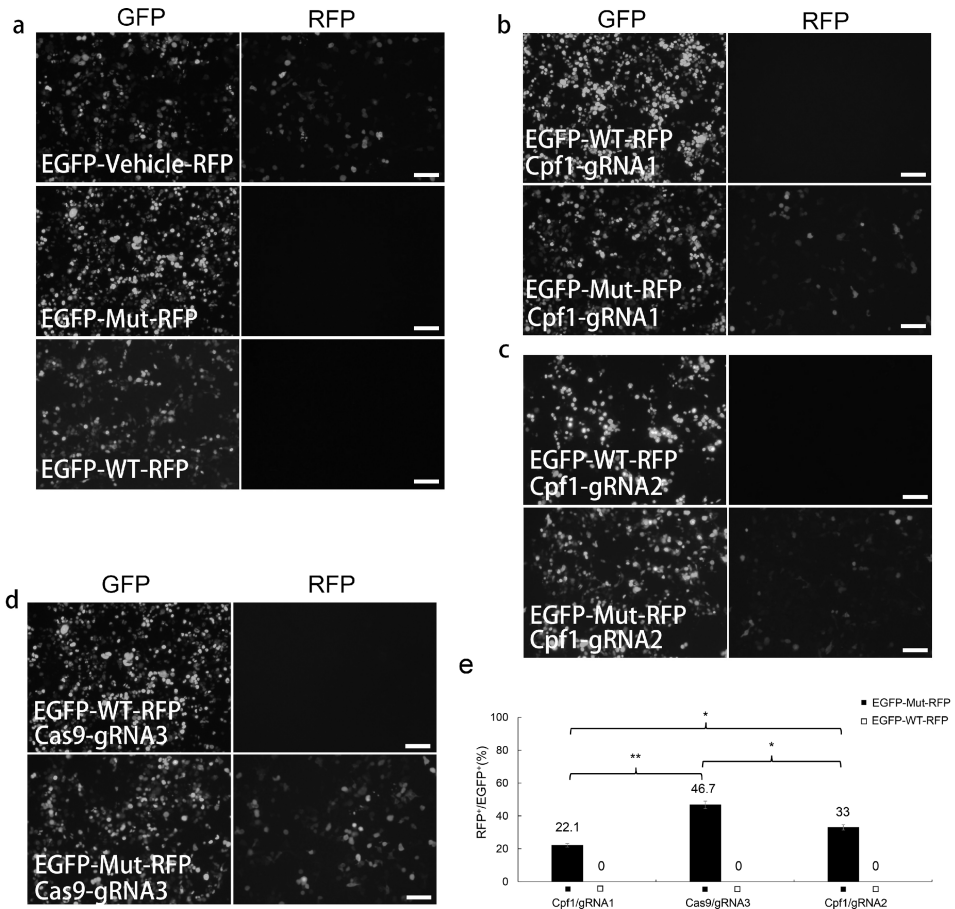


图 2 在 HEK293T 细胞中评估编辑效率

a. 瞬时转染质粒 EGFP-vehicle-RFP, EGFP-Mut-RFP 报告质粒和 EGFP-WT-RFP 报告质粒到 HEK293T 细胞中, 48 小时之后荧光显微镜观察. b-d. AsCpf1 和 SpCas9 对靶基因切割效率的验证. 所设计的 gRNAs 全部都与 EGFR 2573T>G 完全互补, 但与野生型的 EGFR 有一个碱基从错配 (AsCpf1 的 gRNAs 与野生型 EGFR 有一个碱基错配, SpCas9 与野生型 EGFR 完全互补但 PAM 序列有一个碱基错配). AsCpf1 和 SpCas9 稳定表达的 HEK293T 细胞通过慢病毒的感染和嘌呤霉素的筛选. 双荧光报告质粒和 gRNAs 质粒共转染 48 h 之后荧光显微镜成像. e. FACS 统计分析编辑效率. 误差线代表标准差 (SD), 实验经过 3 次重复. 单因素分析; \*  $P < 0.05$ ; \*\*  $P < 0.01$ . WT, 野生型; Mut, 突变型; 比例尺, 200  $\mu\text{m}$ .

Fig. 2 Evaluation of the editing efficacy and selectivity in HEK293T cells

### 3.3 测序验证靶向编辑

通过测序分析进一步验证了 AsCpf1 和 Sp-Cas9 对靶向突变序列进行了切割 (Fig. 3). 这个结果与双荧光报告系统的检测结果一致. 只有 EGFP-Mut-RFP 组的靶向切割位点下游序列中出现了重叠峰 (Figs. 3b, d & f), 而 EGFP-WT-RFP 组并没有出现重叠峰 (Figs. 3a, c & e). 结果表明 AsCpf1 和 SpCas9 均表现出了对突变 EGFR 序列的特异性编辑. 从测序峰可以明显的看出, 虽然双荧光报告系统检测和 FACS 分析都证明了 AsCpf1 的 gRNA1 和 gRNA2 都具有对突变 EGFR 编辑活性, 并且 gRNA2 相比 gRNA1 有这更好的编辑效率, 但是 SpCas9 对突变 EGFR 的编辑效率更优于 AsCpf1.

## 4 讨论

在过去的几十年间, 高通量测序作为新一代的测序技术, 已经成功鉴定出大量与肿瘤相关的体细胞突变<sup>[16,17]</sup>. 虽然 Cas9 系统能够容忍 gRNA 和其互补 DNA 之间又一个或者两个的错配, 但是, 其 PAM 序列是非常保守的<sup>[5,7,18,19]</sup>. 在最近的报道中, gRNAs 与目标序列单碱基的错配, 特别是 PAM 序列, 会限制 CRISPR/Cas 的活性<sup>[8-10,20]</sup>. 在目前的研究中发现, EGFR 2573T>G 突变恰好形成了 SpCas9 的一个 PAM 序列 (CTG>CGG), 并且我们的数据证实了 SpCas9 只对包含突变位点的序列具有切割活性.

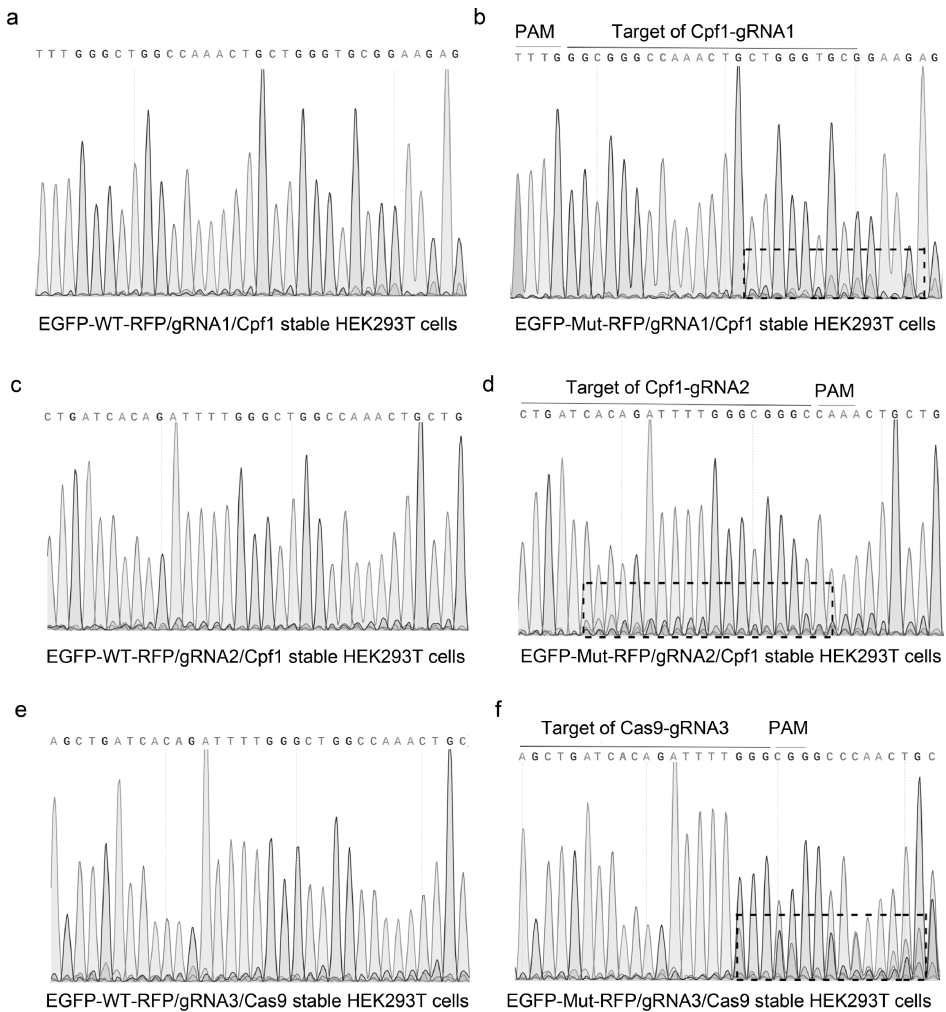


图 3 PCR 扩增靶向位置的序列, 分析

共转染 48h 之后提取基因组 DNA, a, c 和 e. 分别是 AsCpf1-gRNA1, AsCpf1-gRNA2 和 SpCas9-gRNA3 靶向野生型 EGFR 的测序结果. b, d 和 f. 分别是 AsCpf1-gRNA1, AsCpf1-gRNA2 和 SpCas9-gRNA3 靶向突变型 EGFR 切割之后的测序结果. 在黑色虚线框内出现重叠峰的区域是切割之后非同源重组修复所造成的.

Fig. 3 Representative sequencing results following co-transfection in HEK293T cells 48 h after co-transfection as described in Fig. 2

SpCas9 特异性识别 PAM 序列(5'-NGG-3'), 因此如果突变位点没有形成 PAM 序列情况下, 不能够使用 SpCas9 对突变基因进行特异性敲除, 这就是 SpCas9 使用的最大限制条件之一. LbCpf1 和 AsCpf1 靶向和切割 DNA 序列也是靠识别 PAM 序列, 但是这两个蛋白识别的 PAM 序列为 TTTN. 更重要的是, Cpf1 与 Cas9 相比具有更低的脱靶率. Cpf1 的 gRNAs 对靠近 Cpf1 的 PAM 序列的 5' 端区域对单个或者连续两个核苷酸的错配非常敏感<sup>[14,15]</sup>. 在实验中, 我们证实了 AsCpf1 对突变序列具有很强的靶向切割的活性. 因此 Cpf1 蛋白家族的出现扩大了特异性基因组编辑的 gRNAs 的可选择性.

双荧光检测实验和 FACS 分析显示 CRISPR-

Cas9 系统具有更好的基因编辑的活性, 测序结果也进一步证明了 CRISPR-Cas9 系统对 DNA 的编辑活性比 CRISPR-Cpf1 系统的强. 此外, 有报告称 Cpf1 和 Cas9 对 DNA 切割之后会引入不同的突变特征: Cpf1 主要引入核苷酸缺失突变, 很少引入插入突变; 而 Cas9 则可以引入核苷酸插入或者缺失突变<sup>[14,15]</sup>. 这可能是由于它们不同的切割机制所造成的: Cas9 对 DNA 的切割产生平末端, 而 Cpf1 切割产生粘性末端. 然而, Cpf1 对 DNA 切割之后的修复机制还未可知. 对于多数基因来说, 阅读框内 3-nt, 6-nt 或 9-nt 等短片段碱基删除不会造成基因的移码突变, 所以并不影响基因的功能. 因此, 深入了解 Cpf1 对 DNA 切割和切割之后的修复机制会更有利于我们改造 Cpf1 和运用 Cpf1 核酸酶

进行基因敲除。

在此,实验同时证明了 AsCpf1 和 SpCas9 都可以特异性靶向切割 *EGFR*2573T>G 突变基因,而对正常的 *EGFR* 不切割。因此,我们的结果显示了 AsCpf1 和 SpCas9 具有特异性靶向编辑突变基因潜在的应用价值。

#### 参考文献:

- [1] Torre L A, Bray F, Siegel R L, *et al.* Global cancer statistics[J]. *CA Cancer J Clin*, 2015, 65: 87.
- [2] Gazdar A. Activating and resistance mutations of *EGFR* in non-small-cell lung cancer: role in clinical response to *EGFR* tyrosine kinase inhibitors[J]. *Oncogene*, 2009, 28: S24.
- [3] Nguyen K S, Kobayashi S, Costa D B. Acquired Resistance to Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Non-Small-Cell Lung Cancers Dependent on the Epidermal Growth Factor Receptor Pathway[J]. *Clin Lung Cancer*, 2009, 10: 281.
- [4] Shih J Y, Gow C H, Yang P C. *EGFR* mutation conferring primary resistance to gefitinib in non-small-cell lung cancer[J]. *N Engl J Med*, 2005, 353: 207.
- [5] Fu Y, Foden J A, Khayter C, *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells [J]. *Nat Biotechnol*, 2013, 31: 822.
- [6] Lin Y, Cradick T J, Brown M T, *et al.* CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences [J]. *Nucleic Acids Res*, 2014, 42: 7473.
- [7] Pattanayak V, Lin S, Guilinger J P, *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity [J]. *Nat Biotechnol*, 2013, 31: 839.
- [8] 龙添珍,黄学琴,刘海朋,等. XIAP 在 TRAIL 抗性细胞 HCT116 *bax*<sup>-/-</sup>中的作用机制研究[J]. *四川大学学报:自然科学版*, 2016, 53: 1398.
- [9] Courtney D G, Moore J E, Atkinson S D, *et al.* CRISPR/Cas9 DNA cleavage at SNP-derived PAM enables both in vitro [J]. *Gene Ther*, 2016, 23: 108.
- [10] Li Y, Mendiratta S, Ehrhardt K, *et al.* Exploiting the CRISPR/Cas9 PAM Constraint for Single-Nucleotide Resolution Interventions [J]. *PloS one*, 2016, 11: e0144970.
- [11] Shin J W, Kim K H, Chao M J, *et al.* Permanent inactivation of Huntington's disease mutation by personalized allele-specific CRISPR/Cas9 [J]. *Hum Mol Genet*, 2016, 25: 4566.
- [12] Fonfara I, Richter H, Bratovic M, *et al.* The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA[J]. *Nature*, 2016, 532: 517.
- [13] Zetsche B, Gootenberg J S, Abudayyeh O O, *et al.* Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system[J]. *Cell*, 2015, 163: 759.
- [14] Kim D, Kim J, Hur J K, *et al.* Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells[J]. *Nat Biotechnol*, 2016, 34: 863.
- [15] Kleinstiver B P, Tsai S Q, Prew M S, *et al.* Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells [J]. *Nat Biotechnol*, 2016, 34: 869.
- [16] Gonzaga-Jauregui C, Lupski J R, Gibbs R A. Human genome sequencing in health and disease[J]. *Annu Rev Med*, 2012, 63: 35.
- [17] Vogelstein B, Papadopoulos N, Velculescu V E, *et al.* Cancer genome landscapes[J]. *Science*, 2013, 339: 1546.
- [18] Cong L, Ran F A, Cox D, *et al.* Multiplex genome engineering using CRISPR/Cas systems [J]. *Science*, 2013, 339: 819.
- [19] Jinek M, Chylinski K, Fonfara I, *et al.* A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity [J]. *Science*, 2012, 337: 816.
- [20] Gebler C, Lohoff T, Paszkowski-Rogacz M, *et al.* Inactivation of Cancer Mutations Utilizing CRISPR/Cas9[J]. *J Natl Cancer Inst*, 2016, 30: 109.