

两种密码子优化的 *Cas9* 编码基因在斑马鱼胚胎中基因敲除效率的比较

张峰华, 王厚鹏, 黄思雨, 熊凤, 朱作言, 孙永华

中国科学院水生生物研究所, 淡水生态与生物技术国家重点实验室, 武汉 430072

摘要: 为在斑马鱼中获得特异且高效的基因敲除, 多个实验室独立人工合成了序列彼此不一的 *Cas9* cDNA 序列, 并克隆入不同的体外转录载体。本文选取两种斑马鱼密码子优化的 *Cas9* 编码序列(*zCas9_bz* 和 *zCas9_wc*), 对斑马鱼胚胎中的 7 个基因(外源 *egfp* 及内源 *chd*、*hbeffa*、*th*、*ee1a1b*、*tyr*、*tef711a*)分别进行敲除, 通过 PCR 产物测序、克隆测序和表型分析比较了两种 *Cas9* 的敲除效率。结果发现, *zCas9_wc* 在各种情况下都显现出较高的敲除效率, 而 *zCas9_bz* 的效率相对较低。

关键词: CRISPR/Cas9; 斑马鱼; 基因敲除; 突变效率

A comparison of the knockout efficiencies of two codon-optimized *Cas9* coding sequences in zebrafish embryos

Fenghua Zhang, Houpeng Wang, Siyu Huang, Feng Xiong, Zuoyan Zhu, Yonghua Sun

State Key Laboratory of Freshwater Ecology and Biotechnology, Institute of Hydrobiology, Chinese Academy of Sciences, Wuhan 430072, China

Abstract: Recent years have witnessed the rapid development of the clustered regularly interspaced short palindromic repeats/CRISPR-associated protein(CRISPR/Cas9)system. In order to realize gene knockout with high efficiency and specificity in zebrafish, several labs have synthesized distinct *Cas9* cDNA sequences which were cloned into different vectors. In this study, we chose two commonly used zebrafish-codon-optimized *Cas9* coding sequences (*zCas9_bz*, *zCas9_wc*) from two different labs, and utilized them to knockout seven genes in zebrafish embryos, including the exogenous *egfp* and six endogenous genes (*chd*, *hbeffa*, *th*, *ee1a1b*, *tyr* and *tef711a*). We compared the knockout efficiencies resulting from the two *zCas9* coding sequences, by direct sequencing of PCR products, colony sequencing and phenotypic analysis. The results showed that the knockout efficiency of *zCas9_wc* was higher than that of *zCas9_bz* in all conditions.

Keywords: CRISPR/Cas9; zebrafish; gene knockout; mutation efficiency

收稿日期: 2015-11-04; 修回日期: 2015-12-28

基金项目: 国家重大科学研究计划项目(编号: 2012CB944504), 中国科学院重点部署项目(编号: KSZD-EW-Z-001)和国家自然科学基金项目(编号: 31222052)资助[Supported by the Major National Scientific Research Projects (No. 2012CB944504), the Key Programs of the Chinese Academy of Sciences (No. KSZD-EW-Z-001) and the National Natural Science Foundation of China (No. 31222052)]

作者简介: 张峰华, 硕士生, 专业方向: 鱼类发育与生物技术。E-mail: zfh0606@gmail.com

通讯作者: 孙永华, 博士, 研究员, 研究方向: 鱼类发育与生物技术。E-mail: yhsun@ihb.ac.cn

DOI: 10.16288/j.ycz.15-452

网络出版时间: 2016-1-26 15:59:53

URL: <http://www.cnki.net/kcms/detail/11.1913.R.20160126.1559.004.html>

基因敲除是开展基因功能研究最为重要的技术手段之一。近年来, 新一代人工核酸内切酶技术——CRISPR/Cas9(Clustered regularly interspaced short palindromic repeats/CRISPR-associated protein)被广泛地应用于小鼠(*Mus musculus*)^[1]、果蝇(*Drosophila melanogaster*)^[2]、斑马鱼(*Danio rerio*)^[3,4]、线虫(*Caenorhabditis elegans*)^[5]和拟南芥(*Arabidopsis thaliana*)^[6]等各种模式生物的基因敲除和基因功能研究。CRISPR 是一类广泛分布于细菌基因组中的重复结构, 经转录并加工成短的 crRNA(CRISPR RNA)。crRNA 通过碱基配对与 tracrRNA(Trans-activating RNA)结合, 形成双链 RNA, 引导内切酶 Cas 蛋白在 crRNA 序列靶标的特定位置剪切双链 DNA^[7]。据此, 科研人员将 type II CRISPR/Cas 系统中的 crRNA 和 tracrRNA 连接起来成为一条引导 RNA(Guide-RNA, gRNA), 即可利用其对真核生物基因组进行改造, 从而在靶标 DNA 上实现特定位置的剪切, 靶标 DNA 利用非同源末端连接(Non-homologous end joining, NHEJ)机制进行双链 DNA 修复的过程中有可能引入插入或缺失(Indel)突变^[8]。

斑马鱼具有饲育容易、体外受精、体外发育、胚胎透明、组织器官再生能力强等诸多优点, 从而成为研究脊椎动物发育与人类遗传疾病的重要模式动物。由于斑马鱼产卵量较大且易于显微注射, 以 CRISPR/Cas9 基因组编辑技术为基础, 研究者先后在斑马鱼中实现了基因组水平的定点敲入^[9-11]、多基因平行敲除^[12]、同源序列介导的精确整合^[13]等。在利用 CRISPR/Cas9 技术制备斑马鱼突变品系的过程中, 往往需要保证较高的 P_0 代突变效率, 从而减少后续突变体筛选所需投入的人力和物力。目前, 在斑马鱼研究领域较为广泛使用的 *Cas9* 基因序列包括根据哺乳动物密码子优化的人源化 *hCas9*, 以及根据斑马鱼密码子优化的 *zCas9* 等。北京大学熊敬维实验室是最早在斑马鱼中实现 CRISPR/Cas9 基因敲除的实验室之一, 根据化脓性链球菌(*Streptococcus pyogenes*)*Cas9*(spCas9)的编码序列经哺乳动物密码子优化获得人源化 *hCas9* 编码基因, 克隆载体为 pXT7^[4]; 北京大学张博实验室参考 spCas9 的蛋白编码序列, 合成了斑马鱼密码子优化的 *zCas9* (简称为 *zCas9_bz*, 下同), 并克隆入 pGH 载体^[14]; 美国范德堡大学 Wenbiao Chen 实验室所使用的

Cas9 蛋白编码序列则来源于该细菌的一个新品系(MGAS1882), 其 cDNA 合成时同样依据斑马鱼的密码子进行优化(简称为 *zCas9_wc*, 下同), 克隆载体为 pT3TS^[12]。目前, 这 3 种 *Cas9* 均在斑马鱼中被广泛地使用^[11,15,16]。

本研究选取了 *zCas9_bz* 和 *zCas9_wc* 两种斑马鱼密码子优化的 *Cas9* 编码载体体外合成 mRNA, 针对斑马鱼外源 *egfp* 基因以及 *chd*、*hbegfa*、*th*、*eef1a1b*、*tyr* 和 *tcf7l1a* 等 6 个内源基因进行敲除, 并逐一比较了这两种 *Cas9* 对这 7 个基因在特定靶位的敲除效率, 以期研究者日后选用 *Cas9* 在斑马鱼中进行基因敲除提供参考。

1 材料和方法

1.1 实验材料

本研究所用 AB 品系的野生型斑马鱼以及以 AB 品系为背景的转基因斑马鱼 *Tg(kdrl:EGFP)*^[17]均来自于国家斑马鱼资源中心(武汉, <http://zfish.cn>)。斑马鱼饲养于 28.5℃ 恒温及 14 h:10 h 光暗周期条件下。用于显微注射的胚胎由雌鱼和雄鱼自然产卵而得。胚胎发育阶段定义参照文献^[18]。

1.2 方法

1.2.1 gRNA 靶位点的设计

egfp、*tyr* 和 *th* 分别参照文献中所使用的 gRNA 靶位, 具体信息如下:

egfp-target :GGCGAGGGCGATGCCACCTACGG^[9];

tyr-target :GGACTGGAGGACTTCTGGGGAGG^[12];

th-target :GGGTGATCCTGATCCAGATCCGG^[14]。

chd、*tcf7l1a*、*hbegfa* 和 *eef1a1b* 等 4 个基因的靶点为本文设计。首先从斑马鱼基因组数据库(http://www.ensembl.org/Danio_rerio/Info/Index)得到其基因序列和结构信息(ENSDARG00000006110、ENSDARG00000038159、ENSDARG00000075121 和 ENSDARG00000069951), 然后根据 gRNA 靶点设计在线网站 <http://zifit.partners.org/ZiFiT/>, 针对每个基因各设计 2 个靶位点。靶点包含 20 个碱基, 其中 5' 端为 GG, 以利于提高 T7 启动子的合成效率; 紧邻靶点 3' 端的 3 个碱基构成 PAM(Protospacer adjacent

motif)区,其序列为 NGG(N 为任意碱基)。利用 PCR 产物直接测序法初步评估靶位突变效率,若无效则重新设计靶位。经预实验最终确定有效靶位序列分别为:

chd-target :GGATTACCAGCTGCTGGTGGCGG ;

tcf7l1a-target :GGAGGAGGAGGTGATGACCTGGG ;

hbegfa-target :GGGGATTATGGGTTGCCGATAGG ;

eef1a1b-target :GGCCACCTCATTTACAAGTGTGG。

下划线表示 PAM 序列。所用靶位扩增引物见表 1。

1.2.2 gRNA 的制备及 mRNA 的合成

gRNA 的合成一般有载体克隆体外转录法^[9,12]和 PCR 产物体外转录法^[4]两种。本研究采用后一种方法,一般而言,从预订引物到合成 gRNA 可以在

2 日之内完成,这相比载体克隆体外转录法更为节省时间。具体如下:以质粒 pT7-gRNA 骨架^[4]为模板,利用携带 T7 启动子和 gRNA 靶位序列的基因特异性上游引物和通用下游引物 gRNA-RP(表 1)扩增 gRNA 模板。反应体系包含 2×Es Taq MasterMix(康为世纪,北京)25 μL,上下游引物各 2 μL,pT7-gRNA 模板 1 μL,H₂O 20 μL。反应条件:94℃ 预变性 5 min;94℃ 30 s,60℃ 30 s,72℃ 30 s,30 个循环;最后 72℃ 延伸 8 min。PCR 产物经纯化回收,再以 MAXIscript T7 Kit(Ambion,美国)试剂盒体外转录 gRNA。gRNA 经 DNase I(NEB,英国)消化 DNA 模板,然后以 Post-reaction clean up Kit(Sigma,美国)纯化回收,经 1.5%的琼脂糖凝胶电泳检测 RNA 质量。

用于制备两种 *Cas9* mRNA 的质粒 pGH-T7-

表 1 本研究所使用的引物及其序列

Table 1 Primers and their sequences used in this study

引物	序列(5'→3')	用途
egfp-cas9-F1	TGAGCAAGGGCGAGGAGC	靶位扩增
egfp-cas9-R1	CTCGATGCGGTTCCACCAG	
chd-cas9-F3	ATCCATCAATCCATTATCTT	
chd-cas9-R3	TGGTCTGTGAACACTGCC	
tcf7l1a-cas9-F1	TTCTAACCTCCACAGTCGC	
tcf7l1a-cas9-R1	GCTTCCGCAAAGTATTCC	
tyr-cas9-F1	GCGTCTCACTCTCCTCGACTCTTC	
tyr-cas9-R1	GTAGTTTCCGGCGCACTGGCAG	
th-cas9-F1	GCGTATGGAGACTTGGAGC	
th-cas9-R1	ACACTGTCTGTATTTGAAGGCA	
hbegfa-cas9-F1	ATTGGAGGCGCTGAGTGAG	
hbegfa-cas9-R1	GTTCCGTGGATGCAAAAGTC	
eef1a1b-cas9-F1	ATTCTTTAACTGGTGGTCG	
eef1a1b-cas9-R1	TCAATGGTCCTCTTGTCG	
egfp-gRNA-1	TGTAATACGACTCACTATAggcgaggcgatgccacctaGTTTGTAGAGCTAGAAAT	gRNA 模板扩增
chd-gRNA-3	TGTAATACGACTCACTATAggattaccagctgctggtggGTTTGTAGAGCTAGAAAT	
tcf7l1a-gRNA-1	TGTAATACGACTCACTATAggaggaggaggtgatgacctGTTTGTAGAGCTAGAAAT	
tyr-gRNA-1	TGTAATACGACTCACTATAggactggaggacttctggggGTTTGTAGAGCTAGAAAT	
th-gRNA-1	TGTAATACGACTCACTATAgggtgacatcgtacagatcGTTTGTAGAGCTAGAAAT	
hbegfa-gRNA-1	TGTAATACGACTCACTATAggggattatgggtgcccgatGTTTGTAGAGCTAGAAAT	
eef1a1b-gRNA-1	TGTAATACGACTCACTATAggccacatcattacaagtGTTTGTAGAGCTAGAAAT	
gRNA-RP	AAAAAAGCACCGACTCGGTGCCAC	

注:用于 gRNA 模板扩增的基因特异性上游引物序列中小写字母示 20 bp 的 gRNA 靶位序列。

zCas9_bz 及 *pT3-zCas9_wc* 载体如图 1(A 和 B), 经 *Xba* I 限制性内切酶(Fermentas, 美国)线性化后, 分别用 mMessage mMachine T7 UltraKit(Ambion, 美国)和 mMessage mMachine T3 UltraKit(Ambion, 美国)体外转录试剂盒转录 mRNA。利用微量分光光度计(Nanodrop 2000, Thermo Scientific, 美国)测定 RNA 浓度, 1.2%的琼脂糖凝胶电泳检测 RNA 质量(图 1C)。

1.2.3 显微注射

制备出合格的 gRNA 和两种 *Cas9* mRNA 后, 将每个基因所对应的 gRNA 分别与 *zCas9_bz*、*zCas9_wc* mRNA 配制注射样品。注射浓度的标准: 在不导致显著畸形率的前提下, 尽可能提高突变效

率, 具体而言, 每个样品均包含 500 ng/ μ L 的 *Cas9* mRNA 和 50 ng/ μ L 的 gRNA。对于 *egfp*, 本文将两种 *Cas9* mRNA 配制的样品注射到转基因斑马鱼 *Tg(kdrl:EGFP)*与野生型斑马鱼的杂交胚胎中。对于 *chd*、*hbeqfa*、*th*、*ee1a1b*、*tyr* 和 *tc711a* 等 6 个内源基因, 将野生型斑马鱼胚胎以链霉蛋白酶(Sigma, 美国)脱除卵膜后, 利用 FemtoJet express 微量定量注射仪(Eppendorf, 德国)分批次将样品显微注射到两胞期之前的胚胎中, 并于 0.3×Danieau's buffer 中养殖, 水温 28.5°C。为了评估两种 *Cas9* mRNA 在注射样品中的稳定性 本文设置时间梯度, 在 0 h、2 h、4 h 及 6 h 分别取样 1 μ L 电泳检测这 2 种 *Cas9* mRNA 的完整性和相对量。

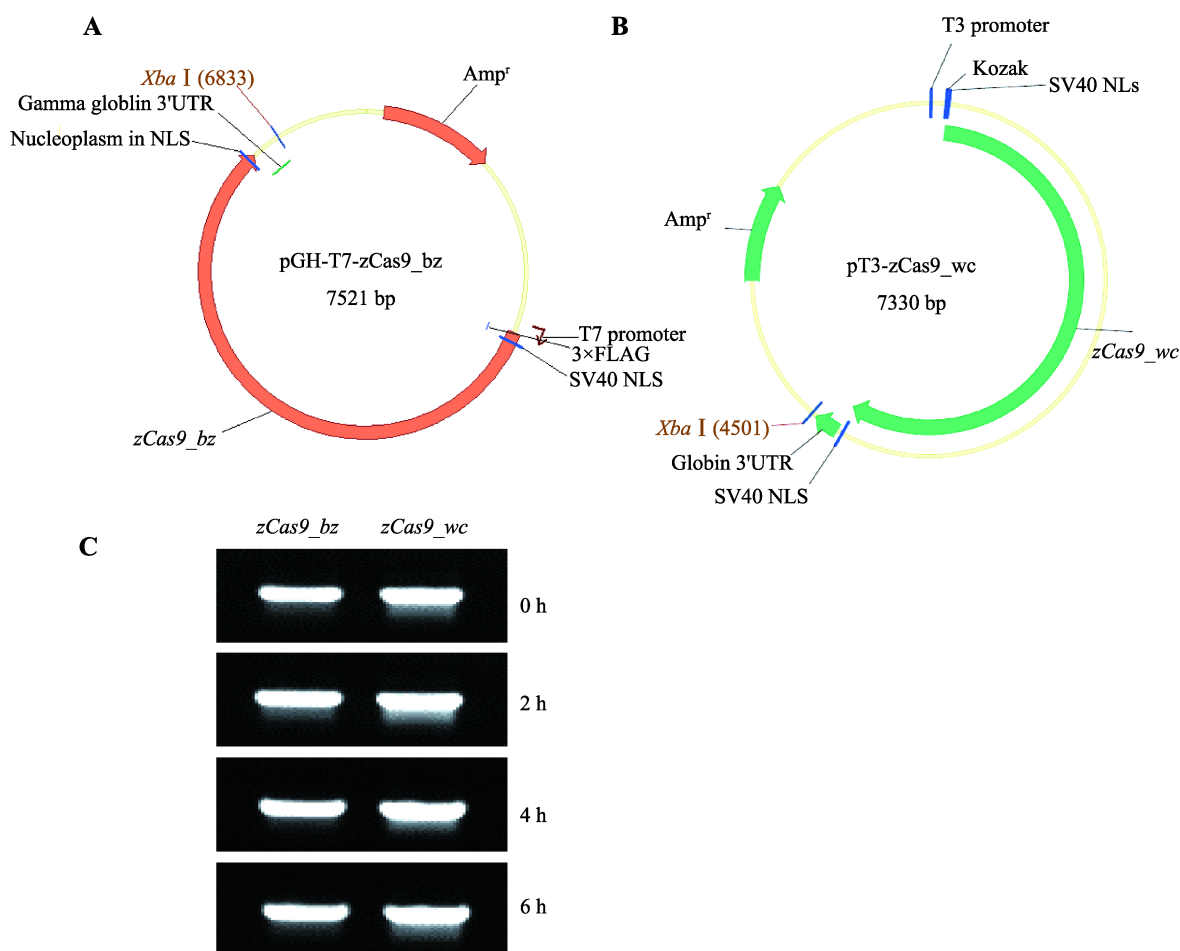


图 1 两种 *Cas9* 体外转录载体及 mRNA 稳定性比较

Fig. 1 Two *Cas9* vectors for *in vitro* transcription and comparison of the stability of their mRNAs

A、B: 分别为 *zCas9_bz* 及 *zCas9_wc* 所对应的体外转录载体示意图; C: 两种 *Cas9* mRNA 样品在配制样品后的不同时间梯度(0 h、2 h、4 h、6 h)所对应的琼脂糖凝胶电泳结果。

1.2.4 突变率检测

本文在实验中发现,发育至 10 hpf(Hours post fertilization)胚胎中基因的敲除效率即代表了注射群体的基因敲除效率,如无特殊交待,用于突变率检测的样品均取自发育至 10 hpf 的注射胚胎。每样随机取 30 颗胚胎,利用斑马鱼直接 PCR 试剂盒(成都福际生物技术有限公司)提取基因组 DNA,并利用表 1 中的靶位扩增引物从中扩增靶位序列。反应体系包括:2×Es Taq MasterMix(康为世纪,北京)25 μ L,上下游引物各 2 μ L,基因组 DNA 模板 1 μ L, H₂O 20 μ L。反应条件为:94℃预变性 5 min;然后 94℃ 30 s, *T_m*(引物复性温度)30 s, 72℃ 30 s,反应 35 个循环;最后 72℃延伸 8 min。电泳检测 PCR 产物,对于条带单一的产物直接进行 Sanger 测序;对于电泳条带不单一的 PCR 产物则采用 GeneJET Gel Extraction Kit(Thermo, 美国)回收目标条带,将纯化后的 PCR 产物进行 Sanger 测序。通过读取 PCR 混合产物的原始测序峰图,初步评价靶位 DNA 是否在 PAM 序列前后发生了 Indel 突变^[4]。

为了更加精细地评价靶位基因组的突变形式和突变效率,将回收的靶位 PCR 产物连入 pMD18-T

(TaKaRa, 中国)载体,转化 *E.Coli*, 每个样品分别挑取约 30 个阳性单克隆进行测序。分析测序结果,统计突变形式和突变比率。

2 结果与分析

2.1 两种 *Cas9* 对 *egfp* 基因敲除效率的比较

将含 *egfp* 靶位的 gRNA 分别与两种 *Cas9* mRNA 共注射到 *Tg(kdrl:EGFP)^{+/+}* 转基因斑马鱼胚胎中,30 hpf 时观察胚胎荧光消失情况并及进行统计。根据胚胎荧光消失的程度由低到高将其划分为 C1、C2、C3 3 个等级,其中 C1 为基本保持完整的荧光表达模式;C2 为存在显著的荧光消失现象;C3 为大部分荧光消失。如图 2 所示, *zCas9_bz* 及 *zCas9_wc* 所对应的胚胎荧光消失最多的级别 C3 所占的比例分别为 1.1% 及 100%,而 C2+C3 所占的比例分别为 62.0%和 100%。由此可见,对于 *egfp* 的该靶位而言, *zCas9_wc* 的敲除效率显著高于 *zCas9_bz*。

两种 *Cas9* 敲除 *egfp* 后,分别取 30 枚胚胎提取基因组,扩增包含靶位的基因组序列,利用 PCR 产物直接 Sanger 测序法初步评估 Indel 突变效率。进一步克隆测序,根据测序结果统计突变类型和比率,

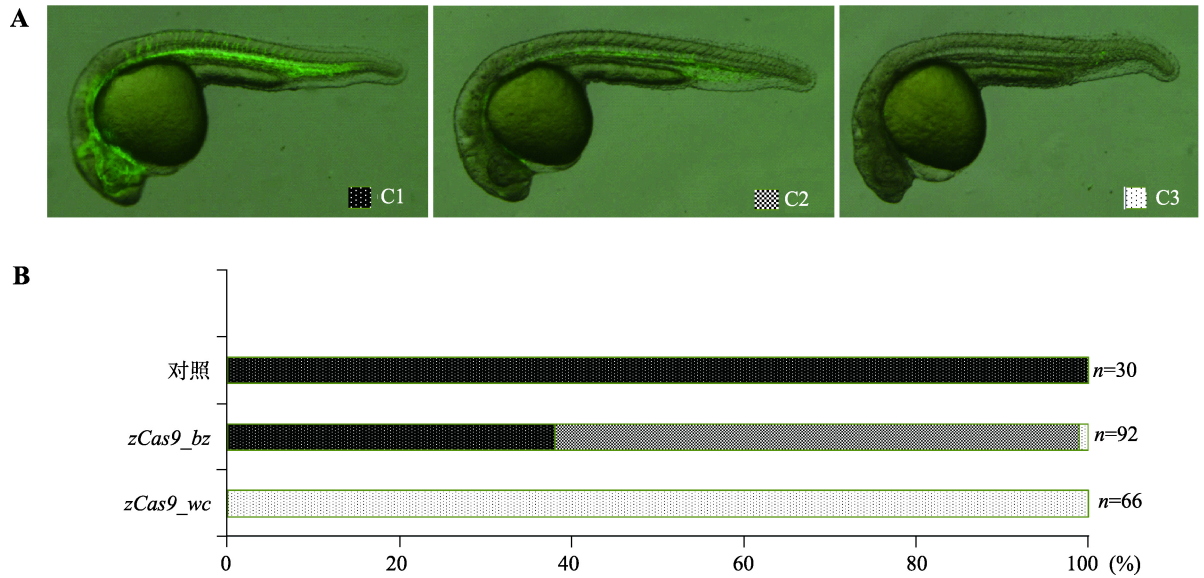


图 2 *egfp* 敲除后胚胎荧光等级划分及比例

Fig. 2 Embryonic fluorescent hierarchies and proportions after *egfp* knockout

A: *egfp* 敲除后胚胎荧光情况。根据荧光消失的程度分为 3 个等级,其中 C1 为荧光基本不消失, C2 为荧光部分消失, C3 为荧光基本全部消失; C1、C2、C3 前的方格图案与 B 图中比例统计相对应。B: *egfp* 敲除后胚胎荧光比例统计。对照为不注射的转基因胚胎情况, *n* 为每一组所统计的胚胎总数。

结果如图 3A 所示。*zCas9_bz* 及 *zCas9_wc* 所对应的突变比率分别为 8.0% 和 100.0%, 该结果所反映的两种 *Cas9* 对于 *egfp* 靶位的敲除效率与胚胎荧光消失情况基本一致, 且对于 *egfp* 靶位, *zCas9_wc* 的敲除效率甚至可达到 100%。对详细突变类型进行分析(附图 1), 未发现此两种 *Cas9* 序列在诱导 *egfp* 的突变上存在序列偏好性。

2.2 两种 *Cas9* 对斑马鱼 *chd*、*hbegfa*、*th*、*eef1a1b*、*tyr* 和 *tcf7l1a* 等 6 个基因敲除效率的比较

同样, 将两种 *Cas9* mRNA(500 pg)分别与 *chd*、*hbegfa*、*th*、*eef1a1b*、*tyr* 及 *tcf7l1a* 所对应的 gRNA(50 pg)共注射到脱除卵壳的野生型斑马鱼胚胎中。由于注射 *Cas9* 后, 在 4hpf 前即可完成基因编辑, 故在 10hpf 随机取样, 提取基因组, 从中扩增包含靶位的基因组序列, 针对 PCR 产物进行直接 Sanger 测序, 通过测序峰图中野生型序列所占的比重来初步评估突变率(图 3, B~D)。

将以上 PCR 产物进一步克隆测序, 挑取约 30 个阳性克隆测序, 统计突变类型和比率, 统计结果见图 3A。从图中可以看到, *zCas9_wc* 对 *hbegfa*、*eef1a1b*、*tyr* 以及 *tcf7l1a* 这 4 个基因在 gRNA 靶位的敲除效率均高达 95% 以上, 对 *hbegfa* 基因的敲除效率甚至可达 100%, 远高于 *zCas9_bz* 对同一基因的敲除效率。*zCas9_wc* 对 *chd* 及 *th* 在相应靶位的敲除效率略低于前面几个基因的敲除效率, 分别为 60.0% 和 78.3%, 但仍高于 *zCas9_bz* 对同一基因的敲除效率。对详细突变类型进行分析(附图 2~7), 未发现此两种 *Cas9* 在诱导突变上存在序列偏好性。

2.3 两种 *Cas9* 对斑马鱼 *chd* 和 *tyr* 基因敲除后的表型分析

为了在表型水平评估 CRISPR/Cas9 基因敲除的效率和特异性, 本文对 *chd* 和 *tyr* *Cas9* 及 gRNA 注射后的胚胎表型进行了统计。对于另外 4 个内源基因, 由于注射后未见明显表型, 故不作统计。与预期相符, *chd* gRNA 和 *Cas9* mRNA 注射后的胚胎中, 部分胚胎出现了类似于 *chd* 突变的腹部化表型^[19]。根据表型的严重程度将其划分为 C1、C2、C3 三类, 其中 C1 为野生型, C2 为出现血岛增大的表型, C3 为同时表现出血岛增大和头部及眼睛变小的表型。如

图 4A 所示, *zCas9_bz* 及 *zCas9_wc* 与 *chd* gRNA 共注射后, 表型最为严重的 C3 类胚胎所占的比例分别为 18.0% 和 48.6%, 而所有出现腹部化表型(C2+C3)胚胎的比例分别为 34.2% 和 82.6%, 这与克隆测序评估的突变效率的趋势一致。值得注意的是, 分类选取 *zCas9_wc* 注射的胚胎进行测序分析, 仅仅在表现出腹部化表型的胚胎(C2+C3)中出现了 *chd* 靶位的突变, 而所有未出现相应表型的胚胎均测序为野生型。

基因 *tyr* 突变后胚胎的色素形成会受到干扰^[10], 根据 50hpf 时胚胎色素形成的多寡将 *tyr* gRNA 和 *Cas9* mRNA 共注射的胚胎划分为 C1、C2、C3 三类, 其中 C1 为野生型, C2 为色素细胞显著减少, C3 为色素细胞基本消失。如图 4B 可见, *zCas9_bz* 和 *zCas9_wc* 敲除 *tyr* 后, C3 类所占比例分别为 10.3% 和 32.8%, C2+C3 分别占比 83.5% 和 98.4%, 与克隆测序突变率结果相符。

从以上结果可知, 注射两种 *Cas9* 和相应的早期胚胎发育相关基因的 gRNA 后, 均可以导致胚胎发育出现特异性的表型缺陷。针对以上两个靶位, *zCas9_wc* mRNA 注射后导致的表型比例均较高。由于出现特异表型的比例与序列测定结果中 Indel 突变比例存在显著关联, 因此对早期发育相关基因进行 CRISPR/Cas9 基因敲除时, 分析胚胎发育的表型比率也可以用来初步估测基因敲除的效率。

2.4 两种 *Cas9* mRNA 的稳定性及编码蛋白质序列比对分析

为探讨两种 *Cas9* 敲除效率差异的原因, 本文首先比较了这两种 *Cas9* mRNA 在不同时间梯度上是否存在量上的差异, 即它们是否存在 mRNA 稳定性上的差异。如图 1C 所示, 在 0 h、2 h、4 h 及 6 h 随着时间的增加, 两种 mRNA 电泳条带亮度相互之间并无明显差异。因此, 这两种 *Cas9* mRNA 的稳定性并不存在显著性的差异。

本文进一步对比了两种 *Cas9* 载体的 cDNA 序列及编码的氨基酸序列(图 5), 发现这两个载体至少在如下 4 个方面存在显著差异: (1)在两个体外转录载体中, 仅 *zCas9_wc* 在翻译起始位点前存在 Kozak 保守序列(GCCACCAUGG), 而已知这一序列的存在可大大提高蛋白质翻译的效率^[20]; (2)*zCas9_bz* 和 *zCas9_wc* 所编码的主体蛋白质序列存在 14 个氨基

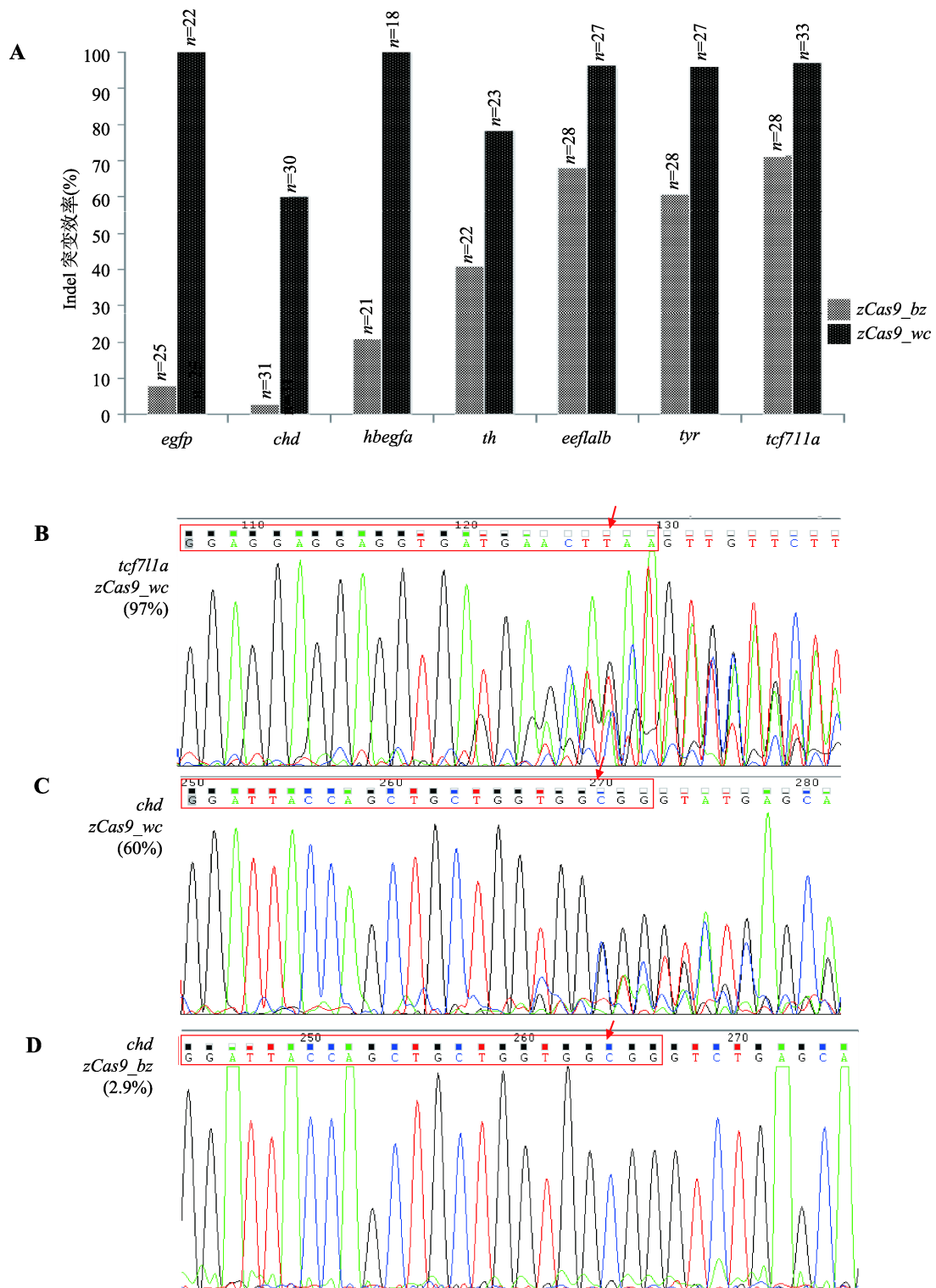
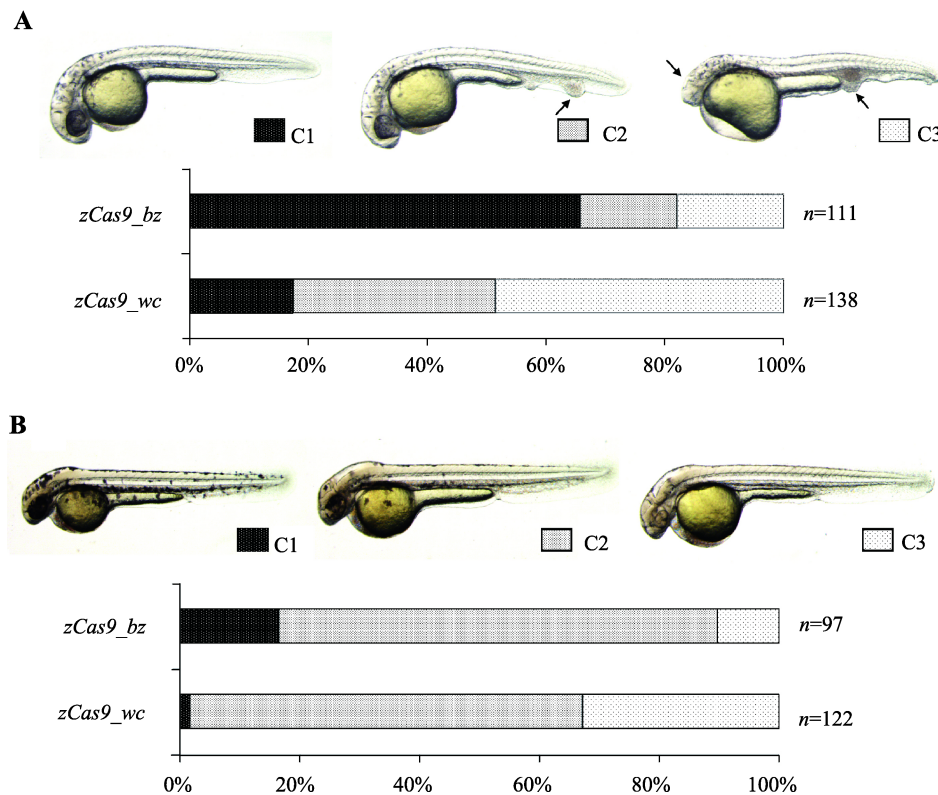


图 3 两种 *Cas9* 对 7 个靶基因敲除效率的比较

Fig. 3 Comparison of knockout efficiencies of seven target genes by the two *Cas9*

A: 两种 *Cas9* 对 7 个基因敲除后克隆测序所得突变效率, n 为测序单克隆总数; B: *zCas9_wc* 敲除 *tcf7l1a* 后胚胎 PCR 产物的直接 Sanger 测序峰图; C: *zCas9_wc* 敲除 *chd* 后胚胎 PCR 产物的直接 Sanger 测序峰图; D: *zCas9_bz* 敲除 *chd* 后胚胎 PCR 产物的直接 Sanger 测序峰图。B~D 中的红色方框所示为包含 PAM 序列的 gRNA 原始靶位位置, 红色箭头所示为 PAM 序列处的第一个测序碱基的读值, 可见敲除效率与该读值方框中的未突变碱基所占比重呈显著负相关。

图 4 *chd* 及 *tyr* 敲除后胚胎的表型统计Fig. 4 Statistics of the phenotypes after *chd* and *tyr* knockout

A: 敲除 *chd* 基因后的表型统计。C1 为野生型, C2 为出现血岛增大的表型(箭头处), C3 为同时表现出血岛增大和头部及眼睛变小的表型(箭头处); B: 敲除 *tyr* 基因后的表型统计。C1 为野生型, C2 为色素细胞有所减少, C3 为色素细胞基本消失。*n* 表示所统计的胚胎总数。

zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	120
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	94
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	240
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	214
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	360
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	334
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	480
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	454
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	600
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	574
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	720
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	694
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	840
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	813
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	960
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	933
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	1080
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	1053
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	1200
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	1173
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	1320
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	1293
zCas9_bz	MDYKDHGDYKDHDIDYKDDDDKMA	1423
zCas9_wc	MDYKDHGDYKDHDIDYKDDDDKMA	1392

图 5 两种 *Cas9* 编码蛋白的序列比对Fig. 5 Comparison of protein sequences coded by the two *Cas9*

图中虚线框分别表示 SV40 NLS(PKKKRV)、核浆素 NLS(KRPAATKKAGQAKKKK); zCas9_bz 序列中 N 端突出的序列为 3×Flag 标签(DYKDHGDYKDHDIDYKDDDDK)。

酸的差别,这与 *zCas9_wc* 来源于该菌种的一个新的品系有关;(3)两种 Cas9 蛋白仅 *zCas9_wc* 在 N 端及 C 端均有一段 SV40 大 T 抗原单分型核定位信号(PKKKRKV)^[21],而 *zCas9_bz* 仅在 N 端存在该核定位信号,而其 C 端则为双分型核浆素核定位信号(KRPAATKKAGQAKKKK),其赖氨酸残基部分类似于 SV40 核定位信号,然而其本身并非蛋白质入核的充分条件^[22]; (4)*zCas9_bz* 所编码的 Cas9 氨基酸序列的 N 端存在一 3×Flag 标签,而 Flag 标签的存在可能会在一定程度上影响到 *zCas9_bz* 蛋白产物的生物学活性^[23]。以上 4 个因素可能是 *zCas9_wc* mRNA 具有较高敲除活性,而同为斑马鱼密码子优化的 *zCas9_bz* mRNA 具有较低敲除效率的重要原因。

3 讨论

继第一代、第二代人工核酸内切酶技术 ZFNs^[24]和 TALENs^[25]之后,第三代人工核酸内切酶技术 CRISPR/Cas9 由于具有设计简单、耗时短、效率高等优点已被广泛地应用于模式动物斑马鱼的基因敲除。Jao 等^[12]利用 *zCas9_wc* 对斑马鱼(*AB* 系)外源 *egfp* 基因进行敲除,利用酶切法检测突变效率达 82%~99%,同时检测内源 *tyr* 基因的敲除效率达 93%~97%。本文中采用 *zCas9_wc* 对外源 *egfp* 和内源 *tyr* 基因进行敲除时,利用克隆测序法检测的突变率分别为 100%和 96.3%,与其结果比较一致。2014 年, Liu 等^[14]以酶切检测的方法比较了 *hCas9* 和 *zCas9_bz* 对 TU 系斑马鱼 *th* 基因在内的 11 个靶位的敲除效率,结果显示 *zCas9_bz* 的效率均高于 *hCas9*。其中以 600 pg *Cas9* mRNA 加 30 pg gRNA 的剂量注射时, *hCas9* 和 *zCas9_bz* 对 *th* 该靶位的敲除效率分别为 30%和 70%左右。本文利用 *zCas9_bz* 以 500 pg *Cas9* mRNA 加 50 pg gRNA 的剂量对 *AB* 系斑马鱼 *th* 该靶位进行敲除,克隆测序突变率达 40.9%。本文进一步选取外源 *egfp* 基因和 *chd* 等 6 个内源基因进行基因敲除,相对于 *zCas9_bz*, *zCas9_wc* 在所有分析的靶位中均表现出较高的敲除效率。因此, *zCas9_wc* 可以成为日后斑马鱼基因敲除工作中的一个好的选择。

如何科学、有效地评估 CRISPR/Cas9 基因敲除后的总体效率是一个值得探讨的问题。本研究采取

了靶点 PCR 产物直接测序法。该方法非常简单易行,仅需要开展一次 PCR 反应和一个测序反应。一般而言,从显微注射到效率评估可以在 3 日内完成。从图 3B~D 的测序峰图可以看出,通过对 PAM 序列及其下游的测序峰图进行分析,可以方便地评估野生型序列占总序列的比例(如 $n\%$),从而可以较为便捷地计算出靶位敲除的大致效率($100\% \sim n\%$)。经与图 3A 进行对比,可以发现这一评估标准获得了克隆测序结果的验证。对突变靶点的序列详细分析表明,这两种 Cas9 蛋白导致的靶位突变并不存在序列上的偏好性(附图 1~7),因此这两种 Cas9 敲除效率的差异应该不是由于 Cas9 识别靶位的特异性所导致,而可能是由于密码子优化、Kozak 序列、核定位信号、N 端标签等诸多因素的存在与否所综合导致。

本文发现一个有趣的现象:针对早期发育必需基因 *chd* 进行 Cas9 敲除时,获得了典型的腹部化表型,测序结果分析表明所有出现腹部化表型的胚胎均可获得 100%的敲除效率,而表型为野生型的胚胎的敲除效率则为 0%。后续养殖实验显示,出现腹部化表型的胚胎很难被继续饲养至性成熟,而仅有那些没有出现表型的胚胎可以被养至性成熟。这就出现了一个敲除效率和养殖存活率之间的悖论:亦即从获得基因敲除子代的角度而言,必需选取那些具有显著表型的胚胎继续饲养;而从养殖存活率的角度而言,又限制了我们选取这部分胚胎。因此,在日后的研究中,当涉及到早期发育必需基因的高效敲除时,可能需要利用原始生殖细胞替换^[26]等技术手段来解决这一问题。

总之,本研究针对斑马鱼外源 *egfp* 基因及 *chd*、*hbegfa*、*th*、*eef1a1b*、*tyr*、*tcf7l1a* 等 6 个内源基因的共计 7 个靶位,通过 PCR 产物测序、克隆测序和表型分析,系统比较了 *zCas9_bz* 及 *zCas9_wc* 对其敲除的效率。结果显示 *zCas9_wc* 效率较高,且对相当多的靶位的敲除效率接近 100%。这为研究者日后在斑马鱼中开展 CRISPR/Cas9 介导的基因敲除时选择 *Cas9* 编码序列提供了有益的参考。

近期,麻省理工学院的张峰实验室发现另一种细菌(*Staphylococcus aureus*)中存在编码一个分子量较小的 Cas9 蛋白的基因(*saCas9*),其蛋白产物不同于 *spCas9*,却一样具有高效的切割活性^[27]。Keith 实

实验室则利用斑马鱼模型, 报道了利用定点突变技术获得能够识别非经典的 NGG PAM 基序的 *Cas9* 变体^[28]。因此, 预期利用斑马鱼模型将有可能开发出靶位自由度更大、更加高效、更为特异的 CRISPR/Cas9 工具。

致 谢

感谢张博(北京大学)、陈文标(范德堡大学)教授分别提供 *zCas9_bz* 及 *zCas9_wc* 体外转录载体, 感谢熊敬维(北京大学)教授提供 gRNA 骨架载体。

附 录

附图 1~7 见电子版 www.chinagene.cn。

参考文献(References):

- [1] Shen B, Zhang J, Wu HY, Wang JY, Ma K, Li Z, Zhang XG, Zhang PM, Huang XX. Generation of gene-modified mice via Cas9/RNA-mediated gene targeting. *Cell Res*, 2013, 23(5): 720–723. [DOI]
- [2] Bassett AR, Tibbit C, Ponting CP, Liu JL. Highly efficient targeted mutagenesis of *Drosophila* with the CRISPR/Cas9 system. *Cell Rep*, 2013, 4(1): 220–228. [DOI]
- [3] Hwang WY, Fu YF, Reyon D, Maeder ML, Tsai SQ, Sander JD, Peterson RT, Yeh JRJ, Joung JK. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat Biotechnol*, 2013, 31(3): 227–229. [DOI]
- [4] Chang NN, Sun CH, Gao L, Zhu D, Xu XF, Zhu XJ, Xiong JW, Xi JJ. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res*, 2013, 23(4): 465–472. [DOI]
- [5] Tzur YB, Friedland AE, Nadarajan S, Church GM, Carlaro JA, Colaiácovo MP. Heritable custom genomic modifications in *Caenorhabditis elegans* via a CRISPR-Cas9 system. *Genetics*, 2013, 195(3): 1181–1185. [DOI]
- [6] Fauser F, Schiml S, Puchta H. Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in *Arabidopsis thaliana*. *Plant J*, 2014, 79(2): 348–359. [DOI]
- [7] Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*, 2007, 315(5819): 1709–1712. [DOI]
- [8] Cong L, Ann Ran F, Cox D, Lin SL, Barretto R, Habib N, Hsu PD, Wu XB, Jiang WY, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, 339(6121): 819–823. [DOI]
- [9] Auer TO, Duroure K, De Cian A, Concordet JP, Del Bene F. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res*, 2014, 24(1): 142–153. [DOI]
- [10] Kimura Y, Hisano Y, Kawahara A, Higashijima S. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Sci Rep*, 2014, 4: 6545. [DOI]
- [11] Li J, Zhang BB, Ren YG, Gu SY, Xiang YH, Huang C, Du JL. Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system. *Cell Res*, 2015, 25(5): 634–637. [DOI]
- [12] Jao LE, Wente SR, Chen WB. Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. *Proc Natl Acad Sci USA*, 2013, 110(34): 13904–13909. [DOI]
- [13] Hisano Y, Sakuma T, Nakade S, Ohga R, Ota S, Okamoto H, Yamamoto T, Kawahara A. Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep*, 2015, 5: 8841. [DOI]
- [14] Liu D, Wang ZX, Xiao A, Zhang YT, Li WY, Zu Y, Yao SH, Lin S, Zhang B. Efficient gene targeting in zebrafish mediated by a zebrafish-codon-optimized cas9 and evaluation of off-targeting effect. *J Genet Genomics*, 2014, 41(1): 43–46. [DOI]
- [15] Shah AN, Davey CF, Whitebitch AC, Miller AC, Moens CB. Rapid reverse genetic screening using CRISPR in zebrafish. *Nat Methods*, 2015, 12(6): 535–540. [DOI]
- [16] Dong ZJ, Dong XH, Jia WS, Cao SS, Zhao QS. Improving the efficiency for generation of genome-edited zebrafish by labeling primordial germ cells. *Int J Biochem Cell Biol*, 2014, 55: 329–334. [DOI]
- [17] Jin SW, Beis D, Mitchell T, Chen JN, Stainier DYR. Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. *Development*, 2005, 132(23): 5199–5209. [DOI]
- [18] Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. Stages of embryonic development of the zebrafish. *Dev Dyn*, 1995, 203(3): 253–310. [DOI]
- [19] Schulte-Merker S, Lee KJ, McMahon AP, Hammerschmidt M. The zebrafish organizer requires *chordino*. *Nature*, 1997,

- 387(6636): 862–863. [DOI]
- [20] Cavener DR. Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. *Nucleic Acids Res*, 1987, 15(4): 1353–1361. [DOI]
- [21] Efthymiadis A, Shao HM, Hübner S, Jans DA. Kinetic characterization of the human retinoblastoma protein bipartite nuclear localization sequence (NLS) *in vivo* and *in vitro* - A comparison with the SV40 large T-antigen NLS. *J Biol Chem*, 1997, 272(35): 22134–22139. [DOI]
- [22] Dingwall C, Robbins J, Dilworth SM, Roberts B, Richardson WD. The Nucleoplasmin nuclear location sequence is larger and more complex than that of SV-40 large T antigen. *J Cell Biology*, 1988, 107(3): 841–849. [DOI]
- [23] Terpe K. Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. *Appl Microbiol Biotechnol*, 2003, 60(5): 523–533. [DOI]
- [24] Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ, Gregory PD, Urnov FD, Amacher SL. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*, 2008, 26(6): 702–708. [DOI]
- [25] Sander JD, Cade L, Khayter C, Reyon D, Peterson RT, Joung JK, Yeh JRJ. Targeted gene disruption in somatic zebrafish cells using engineered TALENs. *Nat Biotechnol*, 2011, 29(8): 697–698. [DOI]
- [26] Ciruna B, Weidinger G, Knaut H, Thisse B, Thisse C, Raz E, Schier AF. Production of maternal-zygotic mutant zebrafish by germ-line replacement. *Proc Natl Acad Sci USA*, 2002, 99(23): 14919–14924. [DOI]
- [27] Ann Ran F, Cong L, Yan WX, Scott DA, Gootenberg JS, Kriz AJ, Zetsche B, Shalem O, Wu XB, Makarova KS, Koonin EV, Sharp PA, Zhang F. *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature*, 2015, 520(7546): 186–191. [DOI]
- [28] Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, Zheng ZL, Gonzales APW, Li ZY, Peterson RT, Yeh JRJ, Aryee MJ, Joung JK. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*, 2015, 523(7561): 481–485. [DOI]

(责任编辑: 张博)

GGCGAGGGCGAGGGCGATGCCACCTA CGG CAAGCTGAC	WT	<i>zCas9_bz</i> ×23	<i>zCas9_wc</i> ×0
GGCGAGGGCGAGGGCGAT-----ACGGCAAGCTGAC	-7bp	×1	
GGCGAGGGCGAGGGCGATGC---CTACGGCAAGCTGAC	-3bp	×1	
GGCGAGGGCGAGGGCGATGCC---CTACGGCAAGCTGAC	-2bp		×1
GGCGAGGGCGAGGGCGAT ACT ---TACGGCAAGCTGAC	-6+4bp		×1
GGCGAGGGCGAGGGCGAT GG ---CTACGGCAAGCTGAC	-4+1bp		×1
GGCGAGGGCGAGGGCGATGC-----TACGGCAAGCTGAC	-4bp		×4
GGCGAGGGCGAGGGCGATGCCA---CGGCAAGCTGAC	-4bp		×1
GGCGAGGGCGAGGGCGATGCC GG ---CGGCAAGCTGAC	-9+5bp		×1
GTCGAGGGCGAGGGCGATGCCA A -----TGAC	-12+1bp		×1
GGCGAGGGCGAGGGCGA-----GCTGAC	-15bp		×1
GGC----- TGTC ---CGGCGGCAAGTTGAC	-23+7bp		×1
GGC-----TACGGCAAGCTGAC	-21bp		×1
GGCGAGGGCGAGGGCGATG-----	-23bp		×1
GGCGAGGGCGAG-----CTGAC	-25bp		×1
GGCGAGGGCGAGGGCGA-----	-26bp		×1
-----ACGGCAAGCTGAC	-27bp		×2
GGCGAGGGCGAGGGCGATGCCAC A CTACGGCAAGCTGAC	+1bp		×1
GGCGAGGGCGAGGGCGATGCCA CGG CCTACGGCAAGCTGAC	+3bp		×1
GGCGAGGGCGAGGGCGATGC AGATGAGGG CGGCAAGCTGAC	-6+9bp		×2

附图 1 *egfp* 靶位突变类型

下划线表示靶点序列；蓝色表示 PAM 序列；红色表示缺失碱基；绿色表示碱基插入；橙色表示碱基置换；WT 为野生型，trans 为转换(transition)或颠换(transversion)。图注同附图 2~7。

CTCTCTGGATTACCAGCTGCTGCTGG CGG GTCTGAGCA	WT	<i>zCas9_bz</i> ×30	<i>zCas9_wc</i> ×12
CTCTCTGGATTACCAGCTG-----CTGAGCA	-12bp	×1	
CTCTCTGGATTACCAG T GTCTGCTGGCGGGTCTGAGCA	1bp trans		×1
CTCTCTGGATTACCAGCTGCTGG---CGGGTCTGAGCA	-3bp		×9
CTCTCTGGATTACCAGCT T ---TGGCGGGTCTGAGCA	-5+1bp		×1
CTCTCTGGATTACCAGCTGCTGG-----GTCTGAGCA	-6bp		×1
CTCTCTGGATTACCAGCTG-----GCGGGTCTGAGCA	-6bp		×2
CTCTCTGGATTACCAGC-----GGCGGGTCTGAGCA	-7bp		×1
CTCTCTGGATTACCAGCTGC-----GGGTCTGAGCA	-7bp		×1
CTCTCTGGATTACCAGCTG GGTCTGAGCAGC GCGGGTCTGAGCA	-6+12bp		×1
CTCTCTGGATTACCAGCTGCTGGT AAGGAGTCTGAGCT GCGGGTCTGAGCA	+14bp		×1

附图 2 *chd* 靶位突变类型

		<i>zCas9_bz</i>	<i>zCas9_wc</i>
TTAAACGGAGGAGGAGGTGATGACCTGGGAGCAAATGA	WT	×8	×1
TTAAACGGAGGAGGAGGTGATGA-----GGAGCAAATGA	-4bp	×1	×2
TTAAACGGAGGAGGAGGTGATG-----GGAGCAAATGA	-5bp	×2	×3
TTAAACGGAGGAGGAGGAGG-----CCTGGGAGCAAATGA	-6+1bp	×1	
TTAAACGGAGGAGGAGGTGATGA-----GCAAATGA	-7bp	×3	×3
TTAAACGGAGGAGGAGGTGA-----GGAGCACCTGA	-15+8bp	×1	
TTAAACGGAGGAGGAGGTGA-----GCAAATGA	-10bp	×1	×1
TTAAACGGAGGAGGAGGTGATGA-----	-18bp	×4	×2
-----GGAGCAAATGA	-27bp	×1	
TTAAACGGAGGAGGAGGTGATGAACCTGGGAGCAAATGA	-2+3bp	×1	
TTAAACGGAGGAGGAGGTGATGACTTAATGGGAGCAAATGA	-1+3bp	×1	
TTAAACGGAGGAGGAGGTGATGAGCAAAATGGGAGCAAATGA	-2+5bp	×1	
TTAAACGGAGGAGGAGGTGATGAACTTATTTTGCTGGGAGCAAATGA	-1+11bp	×1	
TTAAACGGAGGAGGAGGTGATGACTGAGGCAAACCTGGGAGCAAACCTGGGA	+22bp	×1	
TTAAACGGAGGAGGAGGTGATGAACCT...80bp...GGATGGGAGCAAATGA	-6+90bp	×1	
TTAAACGGAGGAGGAGGTGATG-----GGGAGCAAATGA	-4bp		×1
TTAAACGGAGGAGGAGGTGATGA-----GAGCAAATGA	-5bp		×1
TTAAACGGAGGAGGAGGTG-----GGAGCAAATGA	-8bp		×2
TTAAACGGAGGAGGAGGTGATGCA-----AAATGA	-10+2bp		×1
TTAAACGGAGGAGGAGGTGATG-----CAAATGA	-9bp		×1
TTAAACGGAGGAGGAGGTGAT-----CAAATGA	-10bp		×1
TTAAACGGAGGAGGAGGTGAAA-----AATGA	-13+3bp		×1
TTAAACGGAGGAGGAGG-----GAGCAAATGA	-11bp		×1
TTAAACGGAGGAG-----CAAATGA	-18bp		×2
TTAAAC-----TGGGAGCAAATGA	-19bp		×1
TTAAACGGAGGAGGAGGTGATGGC-----	-21+2bp		×1
TTAAACGGAGGAGGAGGTGATGAACCTGGGAGCAAATGA	+1bp		×1
TTAAACGGAGGAGGAGGTGATGGCCCTGGGAGCAAATGA	-16+17bp		×1
TTAAACGGAGGAGGAGGTGATGAGAGCAAAATGGGAGCAAATGA	-2+7bp		×1
TTAAACGGAGGAGGAGGTGATGAGAGCAAAATGGGAGCAAATGA	-3+8bp		×1
TTAAACGGAGGAGGAGGTGATGTGGGAGCAAAATGGGAGCAAATGA	-3+10bp		×1
TTAAACGGAGGAGGAGGTGAATCTTGAAAATCTTGAAAATCCTGGGAG	-3+22bp		×1
TTAAACGGAGGAGGAGGTGATACAAATGTTTCAAATGTTTTCCTGGGAG	-2+21bp		×1
TTAAACGGAGGAGGAGGTGATGAACCTCCTGAGGTGATGAACCTCCTGGG	+21bp		×1

附图 3 *tef71a* 靶位突变类型

	<i>zCas9_bz</i>	<i>zCas9_wc</i>
TTCTCTGGGGATTATGGGTTGCCGATAGGTAAATAATT WT	×16	×0
TTCTCTGGGGATTATGGGTGCCGATAGGTAAATAATT 1bp trans	×1	
TTCTCTGGGGATTATGGGTT----ATAGGTAAATAATT -4bp	×1	
TTCTCTGGGGATTATGGGTTGCCGATAGGTAAATAATT +1bp	×2	
ATGATGTTG...23bp...TCTGGGGATTATGGGTTGCCGATAGGTAAATAATT +6bp	×1	
TTCTCTGGGGATTATGGGTTAGCGATAGGTAAATAATT 3bp trans		×1
TTCTCTGGGGATTATGGGTTATGATAGGTAAATAATT 3bp trans		×1
TTCTCTGGGGATTATGGGTTG----ATAGGTAAATAATT -3bp		×2
TTCTCTGGGGATTATGGGTTGC---TAGGTAAATAATT -3bp		×2
CTCTCTGGGGATTATGGGTTA----ATAGGTAAATAATT -4+1bp		×1
TTCTCTGGGGATTATGGGTTG----ATAGGTAA-TAATT -4bp		×1
TTCTCTGGGGATTATGGG----ATAGGTAAATAATT -6bp		×1
TTCTCTGGGGATTATGGGTTGAA-----TAAATAATT -8+2bp		×1
TTCTCTGGGGATTATGG-----ATAGGTAAATAATT -7bp		×1
TTCTCTGGGGATTATGGGATT-----ATAATT -14+3bp		×1
TTCTCTGGGGATTATGGGTTGCCATAAGGTAAATAATT -1+3bp		×1
TTCTCTGGGGATTATGGGTTAATAATAATCAATAATAATT -12+14bp		×1
TTCTCTGGGGATTATGGTAATAATTGATAGGTAAATAATT -6+9bp		×2
TTCTCTGGGGATTATGGGTTGCCATGGGTTGATAGGTAAATAATT +7bp		×1
TTCTCTGGGGATTATGGGTTGTTATAATAATAGGTAGGTAAATAATT -4+13bp		×1

附图 4 *hbegfa* 靶位突变类型

	<i>zCas9_bz</i>	<i>zCas9_wc</i>
ACCACTGGCCACCTCATTTACAAGTGTGGAGGCATCGA WT	×9	×1
ACCACTGGCCACCTCATTTACGAGTGTGGAGGCATCCA 1bp trans	×1	
ACCACTGACACCTCATTTACAAGTGTGGAGGCATCGA 1bp trans	×2	
ACCACTGGCCACCTCATTTACAA-TGTGGAGGCATCGA -1bp	×1	×2
ACCACTGGCCACCTCATTTACA--TGTGGAGGCATCGA -2bp	×2	
ACCACTGGCCACCTCATTTACA---GTGGAGGCATCGA -3bp	×1	
ACCACTGGCCACCTCATTTACA----GGAGGCATCGA -5bp	×2	
ACCACTGGCCACCTCATTTACA-----GGCATCGA -8bp	×5	
A-----TCGA -33bp	×1	
ACCACTGGCCACCTCATTTACACCTTGTGTGGAGGCATCGA -1+4bp	×1	
ACCACTGGCCACCTCATTTAGGCATCGTGGAGGCATGTGGAGGCAT -4+15bp	×3	
ACCACTGGCCACCTCATTTACA-GTGTGGAGGCATCGA -1bp		×3
ACCACTGGCCACCTCATTTACT-GTGTGGAGGCATCGA -2+1bp		×1
ACCACTGGCCACCTCATTTACGA-TGTGGAGGCATCGA -3+2bp		×1
ACCACTGGCCACCTCATTTAC--GTGTGGAGGCATCGA -2bp		×1
ACCACTGGCCACCTCACCT-CA--TGTGGAGGCATCGA -8+5bp		×1
ACCACTGGCCACCCCATTTACAA-----AGGCATCGA -6bp		×1
ACCACTGGCCACCTCATTTGC-----GGAGGCATCGA -8+2bp		×1
ACCACTGGCCACCTCATT-----GTGGAGGCATCGA -7bp		×1
ACCACTGGCCACCTCATTTACAAG-----GCATCGA -7bp		×1
ACCACTGGCCACCTCATTT-----GGAGGCATCGA -8bp		×1
ACCACTGGCCACCTCATTTA-----GGCATCGA -10bp		×1
ACCACTGGCCACCTCATTTACAATGAGTGGAGGCATCGA -2+3bp		×1
ACCACTGGCCACCTCATTACCTCATGTGGAGGCATCGA -6+7bp		×1
ACCACTGGCCACCTCATTTACAAGGCATGTGGAGGCATCGA +3bp		×1
ACCACTGGCCACCTCATTTACAATGAGTGTGGAGGCATCGA +3bp		×1
ACCACTGGCCACCTCATTTACAACCTCATGTGGAGGCATCGA -2+5bp		×1
ACCACTGGCCACCTCATTTACATTACATTGTGGAGGCATCGA -2+6bp		×1
ACCACTGGCCACCTCATTTACATTACGTGTGGAGGCATCGA -2+6bp		×1
ACCACTGCCCCCTCATCTACAATTCGATGTGGAGGCATCGA -18+23bp		×1
ACCACTGGCCACCTCATTTACAAGATGAGATGTGGAGGCATCGA +6bp		×1
ACCACTGGCCACCTCATTTACAGGCATTATTGTGGAGGCATCGA -2+8bp		×1
ACCACTGGCCACCTCATTTACAATGAGGGTTGGGTAACTCAACTCAA -1+26bp		×1

附图 5 *ee1a1b* 靶位突变类型

	<i>zCas9_bz</i>	<i>zCas9_wc</i>
CTTACTGGGTGATCCTGATCCAGATC CGG ATCATATTT WT	×13	×5
CTTACTGGGTGATCCTGATCCA-----GATCATATTT -6bp	×1	
CTTACTGGGTGATCCT-----GATCCGGATCATATTT -6bp	×5	
CTTACT...111bp...AAA--TAT...94bp...CGA----- -13bp	×3	
CTTACTGGGTGATCCTGATCCAG--- T GATCATATTT -5+1bp		×1
CTTACTGGGTGATCCTGATCCAG-----GATCATATTT -5bp		×1
CTTACTGGGTGATCCTG-----ATCCGGATCATATTT -6bp		×8
CTTACTGGGTGATCCTGATCCA TATTTGA ...197bp...ACT----- -11bp&3bp trans		×1
CTTACTGGGT-----GATCCGGATCATATTT -12bp		×2
CTTACTGGGTGATCCT-----GATCATATTT -12bp		×2
CTTACTGGGTGATCCTGATCCAGATC ATCCGGATCATCC GGATCATATTT +12bp		×1
CTTACTGGGTGATCCTGATCCAG TATCATCCAGTATCGTATCCAGT +23bp		×2

附图 6 *th* 靶位突变类型

	<i>zCas9_bz</i>	<i>zCas9_wc</i>
GCGTTTGGACTGGAGGACTTCTGGGG AGG TGCAGACTC WT	×11	×1
GCGTTTGGACTGGAGGACTTCTGGG-AGGTGCAGACTC -1bp	×1	×2
GCGTTTGGACTGGAGGACTTCTGG--AGGTGCAGACTC -2bp	×1	×3
GCGTTTGGACTGGAGGACTTCTGGTG--GTGCAGACTC -2bp	×3	
GCGTTTGGACTGGAGGACT-----GTGGTGCAGACTC -6bp	×1	
GCGTTTGGACTGGAGGACT- CTG -----GTGCAGACTC -9+3bp	×1	
GCGTTTGGACTGGAGGACTTCTGGG----- GGCTCTGGAGAG -19+12bp	×1	
GCGTTTGGACTGGAGGACT-----GGTGCAGACTC -8bp	×1	×1
GCGTTTGGACTGGAGG-- T ----- G -----GTGCAGACTC -12+2bp	×1	
GCGTTTGGACTGGAGGACT-----GCAGACTC -11bp	×1	
GCGTTTGGACTGGAGGACTTCTGGGG----- -16bp	×1	
GCGTTTGGACTGGAGGACTTCTGG TGGT GCAGACTC +1bp	×1	
GCGTTTGGACTGGAGGACTTCTGG TCCAGG TGGTGCAGACTC +4bp	×1	
GCGTTTGGACTGGAGGACTTCTGG TGCAGACCT GGTGCAGACTC -3+9bp	×1	
GCGTTTGGACTGGAGGACTTCTGG ACTGGAGGGT GGTGCAGACTC +7bp	×1	
GCGTTTGGACTGGAGGACTTCTG CTGTCCAGTCTGCTGTCCAGTCT -4+23bp	×1	
GCGTTTGGACTGGAGGA GT-CGA GGGTGGTGCAGACTC -6+5bp		×1
GCGTTTGGACTGGAGGACTTCTGG-----TGCAGACTC -5bp		×3
GCGTTTGGACTGGAGGACT-----GGAGGTGCAGACTC -5bp		×1
GCGTTTGGACTGGAGGACTT-----AGGTGCAGACTC -5bp		×1
GCGTTTGGACTGGAGGACTT-----GGTGCAGACTC -7bp		×1
GCGTTTGGAC-- GAG ----- TCTGGGT -GGTGCAGACTC -17+10bp		×1
GCGTTTGGACTGGAGGACTTCTG-----CAGACTC -8bp		×2
GCGTTTGGACTGGAGGACTT----- -18bp		×1
GCGTTTGG-----TGGTGCAGACTC -18bp		×1
GCGTTTGGACTGGAGGACTTCTGG TGC AGGTGCAGACTC -2+3bp		×1
GCGTTTGGACTGGAGGACTTCT CGGGAACTGAGA AGACTC -10+13bp		×1
GCGTTTGGACTGGAGGACTTCTGG ACTG GGTGGTGCAGACTC +4bp		×1
GCGTTTGGACTGGAGGACTT GGGTGCTT GGGTGGTGCAGACTC -3+8bp		×1
GCGTTTGGACTGGAGG TGGAGCGTTTGGATCAGT GGTGCAGACTC -11+18bp		×1
GCGTTTGGACTGGAGGACTTCTGGTGA CTTCTGGTGGT GGTGCAGACTC +11bp		×1
GCGTTTGGACTGGAGGACTTCTGG TGCAGACTCGGGGAACTCG GACTC -9+20bp		×1
GACCGTTTGACGGGGGAATTCTGATTATTTGGGTGGTGCACTCT ..14bp.. AAACA +7bp&12bp trans		×1

附图 7 *tyr* 靶位突变类型