

CRISPR/Cas9介导的斑马鱼基因敲入技术研究进展

刘思奇, 艾怡瑞, 李世峰*, 李逸平*

(中国科学院分子细胞科学卓越创新中心/生物化学与细胞生物学研究所, 细胞生物学国家重点实验室,
上海市分子男科学重点实验室, 上海 200031)

摘要: 规律性成簇间隔的短回文重复序列(clustered regularly interspaced palindromic repeats, CRISPR)及相关蛋白组成的CRISPR/Cas9系统作为细菌和古细菌一种适应性免疫防御体系, 近年被用于多个物种的精准基因编辑。作为重要的脊椎动物发育生物学模式生物, 斑马鱼具有发育快、易饲养、繁殖力强和胚胎透明易观察等众多优点, 因此以斑马鱼为模型也开展了许多基于基因编辑的相关研究。相较于基因敲除高效的随机突变, 精准基因敲入(knock-in, KI)的低效率一直是斑马鱼基因编辑领域的短板。本文综述了在斑马鱼中使用CRISPR/Cas9系统进行基因敲入的相关研究进展, 为优化精准敲入效率以及建立斑马鱼疾病模型等方面提供借鉴。

关键词: 斑马鱼; CRISPR/Cas9; 基因编辑; 基因敲入

Advances in CRISPR/Cas9 mediated gene knock-in techniques of zebrafish

LIU Siqu, AI Yirui, LI Shifeng*, LI Yiping*

(State Key Laboratory of Cell Biology, Shanghai Key Laboratory of Molecular Andrology,
Shanghai Institute of Biochemistry and Cell Biology, Center for Excellence in Molecular Cell Science,
Chinese Academy of Sciences, Shanghai 200031, China)

Abstract: As an adaptive immune defense system in bacteria and archaea, clustered regularly interspaced short palindromic repeats and CRISPR-associated protein 9 (CRISPR/Cas9) has been applied to precise gene editing in many species. Given the zebrafish has a lot of advantages including short spawning cycle, easy feeding, high fecundity and convenient observation of transparent embryo, many studies based on gene editing have been carried out using zebrafish as a significant vertebrate developmental biological model organism. Compared with the high efficiency of random mutation by gene knock-out (KO), the low efficiency of precise gene knock-in (KI) has always been a shortcoming in the field of zebrafish gene editing. This paper reviewed the research of CRISPR/Cas9 system for targeted KI in zebrafish, and provides reference for the optimization of precise KI efficiency and the establishment of zebrafish disease models.

Key Words: zebrafish; CRISPR/Cas9; gene editing; gene knock-in

近年来, 基因编辑工具包括锌指核酸酶、类 转录激活因子效应因子核酸酶(transcription

收稿日期: 2022-10-20

基金项目: 国家重点研发计划专项项目(2019YFA0109900)

第一作者: E-mail: liusiqi2019@sibcb.ac.cn

*通信作者: 李世峰, E-mail: sfl01@sibcb.ac.cn; 李逸平, E-mail: yipingli@sibcb.ac.cn

activator-like effector nucleases, TALEN)以及规律性成簇间隔的短回文重复序列(clustered regularly interspaced palindromic repeats, CRISPR)及相关蛋白组成的CRISPR/Cas9系统均能对基因组DNA进行切割, 利用生物体自身基因修复方式使得研究人员获得基因突变个体, 极大地推进了基因功能相关研究。鉴于第三代基因编辑技术CRISPR/Cas9的简便性和特异性, 辅以修复模板, 使得外源片段插入特定基因组位点的基因敲入得以实现。这极大地推进了斑马鱼作为模式动物在生命科学中的应用。

CRISPR/Cas系统是细菌和古细菌的一种适应性免疫防御^[1]。当外来核酸初次侵入时, 细菌会将其序列整合在自身基因组中, 当再次侵入时细菌便能快速识别, 将其进行切割并降解, 从而抵抗外来入侵。许多Cas蛋白需要互相协作才能发挥切割功能, 而Cas9蛋白只需要单个蛋白即可完成RNA介导的DNA识别和切割, 这极大地提高了基因编辑的简便性^[2,3]。Jinek等^[1]保留CRISPR RNA (crRNA)与反式激活crRNA(trans-activating crRNA, tracrRNA)的重要特征, 人工合成了单链向导RNA(single guide RNA, sgRNA), 替代tracrRNA:crRNA与Cas9蛋白结合完成基因编辑过程。至此, CRISPR/Cas9系统结合人工改造sgRNA实现基因编辑的体系的建立, 被广泛应用于多种生物体系中^[4-6]。

除了通过基因敲除进行突变相关的基因功能研究外, 精准基因敲入的实现使得构建疾病模型、谱系示踪和药物筛选等研究得到了进一步的

发展。尽管在一些基因位点基因敲除(knock-out, KO)效率高, 但敲入(knock-in, KI)效率低一直是斑马鱼基因编辑领域内亟待解决的问题。目前, 在斑马鱼中已经进行了多种类型的精准敲入, 主要包括碱基替换、短标签序列插入和长荧光蛋白序列插入等。本文综述了近年来在斑马鱼中利用CRISPR/Cas9体系进行基因敲入的相关研究, 探讨了影响敲入效率的因素, 以便为优化斑马鱼基因敲入效率提供参考。

1 斑马鱼基因敲入修复途径对敲入效率的影响

外部(如射线)或内部(如化学药物)等因素的影响, 导致基因组发生双链断裂(double strand break, DSB)。正确的DSB修复对于维持基因组完整性和细胞生存都十分重要。利用CRISPR/Cas9技术在基因组特定位点产生DSB, 当细胞以某种方式进行DSB修复时, 加入合适的模板便可实现基因的敲入。

DSB修复主要有四种方式(图1): 非同源末端连接修复(classic non-homologous end joining, cNHEJ)、微同源末端修复(microhomology-mediated end joining, MMEJ)、同源重组修复(homologous recombination, HR)和单链退火修复(single-strand annealing, SSA)^[7]。其中, HR包括产生双Holliday交叉(double Holliday junction, dHJ)的经典DNA双链断裂修复方式HDR(homology-directed repair)、不产生dHJ的合成依赖单链退火修复(synthesis-dependent DNA strand annealing, SDSA)和断裂诱

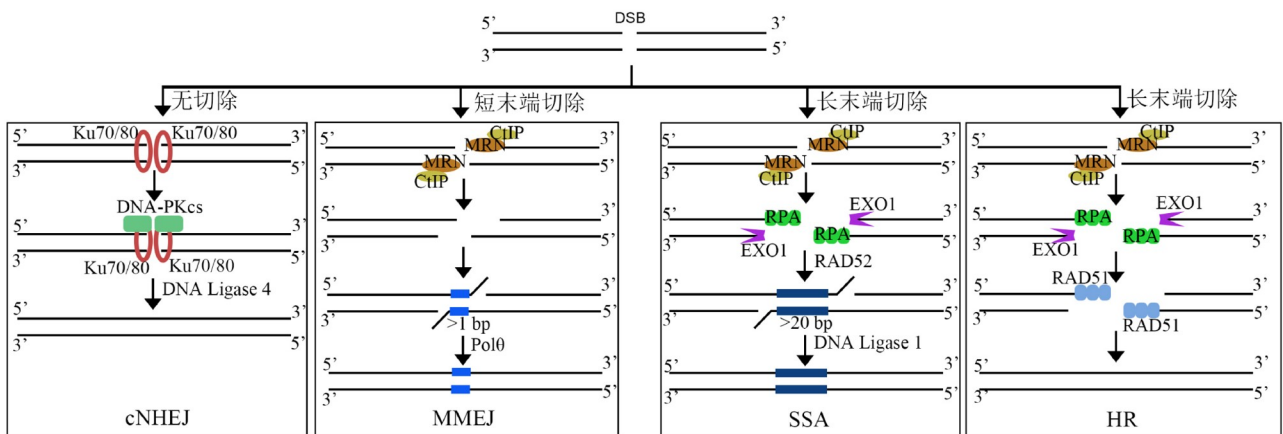


图1 DNA双链断裂主要修复机制模式图

导复制(break-induced replication, BIR)^[8,9]。在产生DSB之后,如果不发生末端切除,则招募cNHEJ相关DNA依赖蛋白激酶催化亚基(DNA-dependent protein kinase, DNA-PKcs)和DNA连接酶4(DNA Ligase 4)等进行cNHEJ修复。末端在MRN复合物[减数分裂重组11同源物(meiotic recombination 11 homolog, MRE11)、辐射敏感蛋白50(radiation sensitive protein 50, RAD50)和奈梅亨断裂综合征蛋白1(Nijmegen breakage syndrome protein 1, NBS1)组成]和C末端结合蛋白相互作用蛋白[C-terminal binding protein (CtBP)-interacting protein, CtIP]介导的切除过程中形成单链DNA时,则会抑制cNHEJ。切除单链较短则可能依赖短同源臂发生MMEJ。也可由外切酶1(exonuclease 1, EXO1)切割成长末端,形成的单链DNA则由复制蛋白A(replication protein A, RPA)包裹,然后依赖辐射敏感蛋白52(RAD52)进行SSA或在RAD51等的协助下进行HR途径修复。这几种修复途径发挥功能的细胞周期也各有不同,HDR途径通常使用姐妹染色体作为模板进行修复,只能限制在S后期/G₂期,MMEJ主要在G₁/S前期激活,而cNHEJ则在细胞周期的任意时期均能发生^[9-12]。

1.1 同源重组修复(HDR)介导的基因敲入

早在二十世纪八十年代,有研究人员将带有

同源臂的DNA序列转入小鼠胚胎干细胞中,利用细胞自身的同源重组过程进行基因敲入,自然条件下重组效率极低^[13]。基因编辑技术问世后,研究人员便尝试通过核酸内切酶来提高基因重组效率^[9,14]。直到2013年,Zu等^[15]在斑马鱼中实现了TALEN介导的同源重组,攻克了斑马鱼中无法进行同源重组操作这一难题,并证明在斑马鱼中可以通过高效的TALEN切割对(98%插入缺失效率)进行精准插入,而且达到了1.5%的生殖传递效率。

在斑马鱼胚胎中进行基因编辑时,常用的检测手段主要是PCR测序分析和荧光显微观察。表1整理了以HDR为敲入修复途径的相关案例,可以发现大多数案例均在突变位点进行敲入,从而进行碱基替换获得疾病相关模型。为了获得精准的碱基替换,可以选择HDR为修复方式进行基因敲入,也可通过不产生DSB的碱基编辑器(base editor)或先导编辑(prime editor)进行基因替换^[16-18]。

以HDR为修复方式进行斑马鱼基因敲入,实现碱基替换将点突变引入斑马鱼基因组,可以用于构建人类疾病相关模型,是人类疾病机制研究的重要手段之一。早在2014年,Irion等^[26]利用CRISPR/Cas9编辑系统尝试在斑马鱼中用野生型albino基因替换突变基因来恢复albino突变斑马鱼的无色表型,研究者使用带有sgRNA切割位点

表1 斑马鱼中以HDR实现基因敲入相关案例

靶点	敲入效果及序列	切割位点(切割效率)	供体模板(同源序列长度)	F0敲入效率	种系传递效率	参考文献
<i>ush2a; ripor2</i>	点突变C771F(126 nt); 敲除等位12 nt删除(114 nt)	突变位点附近(85.7%; 80.5%)	ssDNA	3.4%~18.0%	3/10(30%); 5/11(45%)	[19]
<i>sox3; pax6a; sox11a</i>	报告基因Flag×3(或Pax3), Bio tag, 以及HiBiT tag(或Histag)(~200 nt)	C端终止密码子上游或N端起始密码子下游(53%; 70%; 86%)	ssDNA(50-327 nt)	22%~32%	4%~21%	[20]
<i>pcna</i>	荧光报告基因LR-mNG ₁₁ -RA(117 bp)	C端终止密码子上游	ssDNA(15-25 nt)	不详	5%(只有一个胚胎)	[21]
<i>nefma</i>	荧光报告基因2A-mCherry-CaaX和KalTA4(~3 kb)	C端终止密码子上游	体内线性化环状质粒(820bp\1550bp)	不详	8%~13.3%	[22]
<i>th; nop56; rps14; twist2</i>	插入酶切位点或氨基酸突变(~300 nt)	突变位点附近(73%~80%)	ssDNA(~150nt)	25%~45%	9.5%~31.8%	[23]
<i>smad6a; tprkb; pls3; slc2a10</i>	点突变(60~180 nt)	突变位点附近	ssDNA(60\120\180 nt)	1%~4%	不详	[24]
<i>tp53; cdh5; lmna</i>	氨基酸突变(~130 nt)	突变位点附近	ssDNA(30-97nt)	~2%	4.5%~21.4%	[25]
<i>albino</i>	点突变(~1 kb)	外显子(95%)	体内线性化环状质粒	46%	~10%	[26]

的环状质粒为供体模板进行敲入, 约10%的注射F0胚胎可以发育至成年并产生黑色素, 其生殖系传递效率为3/28(10.7%); 当使用模板长度为206 bp的dsDNA作为供体模板时, 可能由于同源序列过短无法敲入。当使用环状质粒作为模板时, 在敲入序列两端加入sgRNA切割位点可以提高敲入效率。Boel等^[24]使用单链DNA(single-stranded DNA, ssDNA)为修复模板, 将与疾病相关的点突变引入斑马鱼基因组, 分析不同长度的单链修复模板对敲入效率的影响, 测序检测敲入效率, 结果显示不同基因位点敲入效率为4%~8%, 并发现非sgRNA靶向链作为修复模板有一定的敲入效率优势。Prykhodzhiy等^[25]在斑马鱼中以HDR介导的敲入体系中发现, 使用非靶向链的非对称ssDNA修复模板进行基因敲入效率较高。在构建李-弗劳梅尼癌症易感性综合征常见突变类型*tp53* R143H敲入品系时, 使用非靶向链的敲入效率是使用靶向链敲入效率的2.72倍。在构建*tp53* R217H敲入品系时, 同样发现, 非靶向链模板敲入效率更高。在*lmna*位点进行R471W敲入时发现, 硫代磷酸酯修饰的非对称靶向链模板敲入效率要高于无修饰敲入模板。这意味着在以ssDNA为模板进行不同基因位点敲入的过程中, 可以尝试不同同源臂的长度、同源臂的对称性、靶向链与非靶向链修复模板的选择、修复模板链的长度和修复模板链的化学修饰等方面来提高基因敲入效率。Bai等^[23]也通过依赖HDR的修复方式, 以长单链DNA(long single-stranded DNA, lssDNA)为修复模板, 进行斑马鱼基因敲入修复*tyr*基因的点突变。他们的结果发现, 98.5%胚胎色素恢复, 但每个胚胎恢复色素细胞数较少, 而且毒性较大, 胚胎难以发育至成年。在其他位点*th*、*nop56*和*twist2*获得了9.5%~31.8%的生殖传递效率。这与人T细胞研究中发现lssDNA为修复模板较双链修复模板的毒性更低有一定不同^[27]。

在短序列标签敲入品系构建相关研究中, Ranawakage等^[20]选择lssDNA为HDR修复模板成功在斑马鱼中进行了敲入, 效率达到了21%, 但尝试敲入更长的荧光蛋白时, 未能获得敲入品系。由于HDR修复方式插入长片段的效率偏低, 在斑马鱼中一直未能获得CRISPR/Cas9介导HDR修复方式的荧光蛋白敲入模型。直到2020年, Eschstruth

等^[22]通过HDR修复方式在斑马鱼基因组中引入长片段精准插入, 其DSB位点3'端为cNHEJ修复, 引入部分供体质粒骨架序列, 但敲入后不影响内源基因和荧光标记的表达。荧光检测发现, 约13.3%的*nefma:Cherry-CaaX*敲入胚胎可以进行生殖传递, *nefma:KalTA4*敲入胚胎的生殖传递效率约为8%, 在斑马鱼体系利用CRISPR/Cas9系统和HDR修复方式进行了荧光蛋白长片段的精准插入。后来, Kesavan等^[21]还通过改进荧光报告基因序列长度, 成功构建表达内源基因表达模式的荧光报告斑马鱼品系, 只需要插入16个氨基酸的绿色荧光蛋白(green fluorescent protein, GFP)片段GFP₁₁后, 即可在利用巨细胞病毒(cytomegalovirus, CMV)启动子等过表达GFP₁₋₁₀的条件下与其互补, 从而表达特异性绿色荧光。这一方法解决了敲入长荧光蛋白编码序列效率低的问题, 同时保证了荧光表达的特异性。最近, Carrington等^[28]通过荧光PCR和毛细管电泳预筛的方法, 以ssDNA为修复模板, 对*tcnba*位点进行了Flag标签敲入、*gata2b*位点进行了HA标签敲入和人戈谢病(Gaucher disease)相关基因*gba*进行了单核苷酸的替换, 种系传递效率在1%~5%。

1.2 微同源末端修复(MMEJ)介导的基因敲入

MMEJ修复过程不需要Ku70/80以及Ligase 4等的参与, 而是依赖于同源重组起始复合物MRN和DNA聚合酶polθ等^[29-31]。MMEJ修复常常导致DSB两侧的碱基删除, 以及染色体易位和重排, 因此是一种与“无错”的HR修复途径相反的“倾向错误”修复。研究人员通过对修复模板改造和敲入靶点的选择进行优化, 使得通过MMEJ修复方式也能获得想要的基因敲入类型^[32,33]。从表2列举的相关MMEJ介导基因敲入案例来看, 目前使用此方法进行斑马鱼基因敲入的研究并不多。

Hisano等^[35]使用MMEJ修复方式在*tyr*和*krttlc119e*基因位点进行荧光蛋白的敲入, 分析结果发现, 20 bp~40 bp长度的同源臂插入效率最高; PCR检测结果发现, F0代基因敲入效率达到38%, 能产生精准生殖传递的效率约为4.5%。但是并未观察到引入mCherry蛋白的红色荧光信号, 可能是因为插入位点在基因编码区域, 嵌合蛋白没有功能或不稳定很快被降解。在另一个基因

表2 斑马鱼中以MMEJ实现基因敲入相关案例

靶点	敲入效果及序列	切割位点(切割效率)	供体模板(同源序列长度)	F0敲入效率	种系传递效率	参考文献
<i>noto</i>	荧光报告基因2A-TagRFP-CAAX-SV40(~1 kb)	外显子(95%)	体内线性化环状质粒(12\24\48 bp)	56%~95%(5'端精准)	3/5(60%)	[34]
<i>gfap</i> ; <i>flila</i> ; <i>gatala</i>	荧光报告基因T2A-mCherry(~1 kb)	3'非翻译区(gfap gRNA效率为54%)	环状质粒(25 bp)	69%~77%	不详(F1中21%~24%)	[32]
<i>krttlc119e</i>	荧光报告基因cGFP-polyA(~1 kb)	终止密码子附近(73%)	体内线性化环状质粒(40 bp)	38%(不同程度的绿色荧光)	2条广泛表达GFP均能生殖传递	[35]

*krttlc119e*中,选择在其C端进行基因敲入,依赖修复模板40 bp同源臂进行修复,F0代可以观察到绿色荧光蛋白信号,表达出不同强度的绿色荧光,共有三条能产生GFP生殖传递的F0代。2020年,Wierson等^[34]发现,同源臂两端添加通用sgRNA来帮助修复模板释放后,短同源臂(24 bp/48 bp)与长同源臂效率相似。F0代检测结果表明,有49%胚胎表达报告基因,并有约17.4%的F1代胚胎含有敲入整合,其利用短同源臂进行敲入的效率较之前利用环状模板进行同源重组的效率提高了约10倍。在*tyr*等位点种系传递效率达到了22%~100%。但由于选择的敲入位点均在外显子区域,这可能影响内源基因的表达。去年,有研究团队又利用48 bp同源臂在内含子位置插入携带荧光报告基因的基因捕获元件,通过Cre/lox系统调控和指示*rbbp4*和*rbl*的正常表达和敲除的转变^[36]。

Luo等^[32]通过构建25 bp长度同源臂的修复模板,在不破坏内源基因表达的情况下,在斑马鱼中通过MMEJ修复方式成功构建了三个荧光蛋白精准敲入的品系,并能产生生殖传递。这意味着在斑马鱼中通过仅含有25 bp同源臂的MMEJ修复方式即可获得精准敲入的斑马鱼品系。相较于通常需要超过800 bp同源臂的HDR修复方式,短同源臂敲入方法降低了品系构建的繁琐程度。

1.3 非同源末端连接修复(cNHEJ)介导的基因敲入

细胞以cNHEJ方式进行DSB修复过程中容易发生插入或缺失,常用于基因敲除体系中。其易错的修复方式,使得利用cNHEJ修复方式难以进行精准的基因敲入。表3整理了cNHEJ介导的敲入案例,从表中可以看出,敲入位点通常选择在非编码区上,利用cNHEJ的高效性以及非编码区不破坏本底基因表达的方式来实现长片段的插入。

Auer等^[37]通过CRISPR/Cas9系统成功将5.7 kb片段插入斑马鱼基因组中,在F0代中观察到约75%的斑马鱼胚胎表达了敲入片段激活的UAS:RFP红色荧光。

为了获得内源基因顺式作用元件调控下的荧光转基因品系,Kimura等^[40]通过在N端顺式作用元件位置插入启动子*Hsp70*来介导荧光蛋白表达。Abdelmoneim等^[44]也通过相同的体系,在受雌激素调节的分子标志物*vtg1*基因位点进行敲入,敲入效率为0~10%,其中能产生生殖嵌合效率的雌性斑马鱼比例是4.8%,而雄性斑马鱼比例是7.7%。但是*Vtg1*的表达量变化远高于EGFP的变化,可能是因为增强子捕获效率不足等原因。由于有些基因的增强子在基因位点的远端,这种方式进行敲入相较于直接利用内源基因位点进行荧光标记的方式仍显不足,难以真正概括内源基因的表达情况。

Kesavan等^[39]通过CRISPR/Cas9基因编辑技术依赖的cNHEJ修复方式在斑马鱼中脑后脑边界(midbrain-hindbrain boundary, MHB)发育相关的重要基因敲入荧光蛋白,研究人员选择的插入位点位于ATG上游,尽可能使得荧光蛋白在内源启动子或增强子的调控下表达,且这种插入方式并未影响到内源基因的表达。Kesavan等^[45]通过同样的方法将*CreER-T2*插入*otx2*位点ATG上游500 bp处来产生Cre诱导品系,PCR扩增发现在5'端连接处有4 bp插入,说明虽然修复模板含有较长的同源臂,但敲入方式可能是cNHEJ。这些方案均利用了cNHEJ高效敲入的特点,又通过选择在内源基因N端进行敲入,使报告基因与内源基因表达模式尽可能一致,又不破坏内源基因的表达,但也可能无法真正模拟内源基因的表达模式。

Li等^[33]在2015年开发了一种内含子靶向cNHEJ介导的斑马鱼敲入体系,其内源基因的编码以及

表3 斑马鱼中以cNHEJ实现基因敲入相关案例

靶点	敲入效果及序列	切割位点(切割效率)	供体模板(同源序列长度)	F0 敲入效率	种系传递效率	参考文献
<i>tjpla</i> ; <i>Cldn15la</i>	荧光报告基因eGFP-polyA(~700 bp); tdTomato-polyA(~1400 bp)	最后一个外显子上游的内含子	dsDNA(~500 bp)	1%~5%	不详 (F1 中 17.1%~26.9%)	[38]
<i>Rab11a</i>	荧光报告基因eGFP-polyA(~700 bp)	起始密码子上游非编码区域	dsDNA ; ssDNA (~500bp)	1%~5%	不详 (F1 中 21.2%~22.4%)	[38]
<i>otx2</i> ; <i>pax2a</i>	荧光报告基因venus-polyA; tRFP-polyA	起始密码子上游(>50%)	体内线性化环状质粒 (500 bp)	不详	4%~7.6% ; 2.8%~20%	[39]
<i>th</i>	荧光报告基因P2A-EGFP-外显子13 (~2.1kb)	最后两个外显子中间的内含子区域(83%)	体内线性化环状质粒	23.7%	3/25(12%)	[33]
<i>eng1b</i> ; <i>evx2</i> ; <i>glyt2</i> ; <i>vglut2a</i>	荧光报告基因hs70:IR1-GFPTx ; GFPbait-hsp70:Gal4-PolyA; Tbait-hs70- IR1-GFPTx; Mbait-hs70-IR1-GFPTx	基因上游约 200-600 bp	体内线性化环状质粒	5%~10%	2.5%~66.7%	[40]
<i>neurod:eGFP</i> ; <i>vsx2:eGFP</i>	荧光报告基因eGFPbait-E2AKalTA4	eGFP编码区(66.7%)	体内线性化环状质粒	不详	75.6%; 57.6%	[37]
<i>tbx5a</i> ; <i>kctd10</i>	荧光报告基因Lox序列-切割位点后剩余 CDS-2A-tdTomato-Lox序列-PolyA-无意 突变外显子	内含子(98%)	体内线性化环状质粒	15.1%	5/9(56%)	[41]
<i>hey2</i>	荧光报告基因Lox序列-polyA-TagRFP- P2A-剪接受体-lox序列-外显子5-P2A- EGFP-polyA	最后一个外显子上游的内含子(50%)	体内线性化环状质粒	不详	2/21(9.5%)	[42]
<i>sox10</i> ; <i>isl1</i>	荧光报告基因Lox序列-切割位点后剩余 CDS-tdTomato-polyA-PolyA-EGFP-2A- 部分外显子-Lox序列	内含子	体内线性化环状质粒	64.6%	8/18(44%)	[43]

调控元件均未遭到破坏。为了选择比HDR效率更高的cNHEJ修复方式来进行整合,其sgRNA切割位点在*th*最后一个外显子前的内含子区域,这样即使5'端发生部分碱基缺失或插入也不影响内源基因的表达。将最后一个外显子E13构建到修复模板中,从而保证了内源基因表达的完整性。F0代可以观察到荧光嵌合的效率约为23.7%,其中能进行生殖传递的胚胎效率也达到了12%。这种方法使得研究人员不再依赖精确度更高的HDR修复方式就能拿到不影响内源基因功能的荧光蛋白敲入品系。

2021年,Gu等^[46]在*th-P2A-EGFP*敲入品系的*th*第8号外显子两侧引入同向*loxP*序列,当与表达Cre重组酶的转基因品系交配时可进行基因敲除,从而完成条件性敲除品系。

近年,许多研究人员依赖cNHEJ修复方式,在斑马鱼内含子位点进行双报告基因的敲入,通过Cre重组酶自由调节基因敲入或基因敲除的荧光标记。2019年,Li等^[41]在*tbx5a*位点外显子E2和E3中间的内含子进行敲入构建,载体包含E3以及剩余的蛋白编码序列,后接有2A和*tdTomato*序列,两侧为同向的*loxP*位点,再接*polyA*和无意突变的外

显子。在不存在Cre重组酶的条件下,表达*tbx5a*和*tdTomato*,*tdTomato*指示表达正确的阳性敲入。在存在Cre重组酶的条件下,*loxP*序列中间的阳性元件会被切除,从而表达提前终止的无意突变和2A-*tdGFP*,绿色荧光指示完全敲除品系。去年,Li等^[42]在*hey2*最后一个内含子敲入最后一个外显子接GFP编码序列的同时,其敲入序列前端还含方向相反的*loxP*序列,中间为颠倒的TagRFP编码序列。在Cre重组酶存在的条件下,*loxP*中间序列颠倒,使得RFP表达,并终止后续表达,从而获得表达RFP条件性敲除品系;若无重组酶Cre则表达为基因敲入荧光报告基因品系*hey2-P2A-EGFP*。Han等^[43]通过CRISPR/Cas9体系与Cre重组酶结合,利用cNHEJ修复方式在*sox10*和*isl1*位点进行了双报告基因的敲入,其通过优化载体设计使得任一方向插入均能实现双报告基因敲入。Levic等^[38]也通过靶向内含子或5'UTR区域进行荧光蛋白敲入,将一部分内源基因表达序列与外源荧光蛋白连接来保证内源基因完整性,敲入位点选择在非编码区域来避免内源基因表达的破坏。通过这一方法,研究人员在*tjpla*、*cldn15la*和*rab11a*多个基因的N端

或C端完成了荧光蛋白敲入。

2 其他敲入效率影响因素

2.1 CRISPR/Cas9和sgRNA类型对编辑效率的影响

Albadri等^[47]发现,在斑马鱼*kif5aa*位点,Cas9蛋白较Cas9信使RNA具有更高的编辑效率和更低的细胞毒性.sgRNA的切割效率是影响基因敲入效率的最主要因素之一.sgRNA选择转录起始位点TSS作为靶点的效率较内含子区域更高,sgRNA结合位点应该尽可能接近(小于10 bp)敲入或碱基替换位点^[48-50]。通常sgRNA由体外转录获得,在斑马鱼胚胎中注射化学合成tracrRNA:crRNA对较转录生成的sgRNA可能具有更高的编辑效率^[51]。在3'端增加G减少A可以提高sgRNA的稳定性和活性,但如果G重复太多可能导致形成PAM富集区域的Cas结合异常^[48]。目前所知,20个nt长度是最特异和高效率的长度,超过20 nt或短于17 nt都可能降低切割效率^[52,53]。化学修饰后的sgRNA往往更具稳定性,更有利于提高敲入效率,但也可能带来更高的脱靶效应;未经过化学修饰的sgRNA更容易降解,当然也可能降低激活免疫系统的风险^[54]。

2.2 修复模板类型对编辑效率的影响

供体修复模板的不同构象(线性化或环状质粒)对敲入效率也有一定影响。Suzuki等^[55]通过使用迷你环状质粒的修复模板进行了高效的、非同源依赖的体内基因敲入,避免了将质粒骨架整合到基因组中。在环状质粒导入细胞后利用核酸内切酶Sce I或CRISPR/Cas9对其进行线性化后的KI效率显著高于环状质粒为模板的效率^[26,56]。Aird等^[57]将HUH内切酶融合PCV序列,使得其与修复模板ssDNA可以共价结合,从而在DSB产生时达到切割酶与供体修复模板时空共定位,最终在293T细胞系中提高了敲入效率。Gu等^[58]利用链霉亲和素和生物素(monomeric streptavidin, mSA/biotin)的非特异性来提高CRISPR/Cas9组分与修复模板的亲合性,选择在可能提高重组效率具有较长G₂期的2-细胞小鼠胚胎进行注射,提高了近10倍的敲入效率。Seleit等^[59]将mSA/biotin体系运用在青鳉鱼中,成功构建荧光标记内源基因的敲入,其F0代出现嵌合体表达的效率达到了11%~59%,F0代的生殖

传输效率在不同的位点中达到了25%~100%,在F0可以产生生殖传输的胚胎中,阳性F1胚胎比例根据不同位点达到了6.6%~50%。迄今为止,在斑马鱼精准基因敲入研究中还未见使用mSA/biotin体系的报道。

2.3 修复途径调控因子对编辑效率的影响

Aida等^[60]发现,通过调控修复过程中的关键因子可提高敲入效率,他们表达DNA修复关键组分核酸外切酶Exo I以及部分其他测试蛋白可以提高人细胞中微同源臂介导基因敲入的效率。在斑马鱼中,He等^[61]发现,早期胚胎中广泛存在微同源末端介导的DNA双链断裂修复机制,通过抑制内源Ligase 4显著提高MMEJ效率,而抑制Ligase 3活性则显著降低MMEJ活性。de Vrieze等^[19]利用吗啉环反义核酸类似物(morpholino)抑制cNHEJ修复通路关键蛋白Ku70,提高了其中一个位点的HDR敲入效率。

除此之外,参与抑制或激活DNA修复过程中关键因子的小分子也是提高敲入效率的重要手段。SCR7是cNHEJ修复最后连接过程中关键因子Ligase 4的抑制剂,使用该小分子可以提高人细胞系和小鼠胚胎HDR效率^[62]。Yu等^[63]通过小分子抑制剂Brefeldin A和L755507提高小鼠胚胎干细胞中的HDR效率2~3倍。Song等^[64]发现,HDR增强因子RS-1刺激hRAD51与ssDNA结合,提高体外敲入效率2~5倍,但通过cNHEJ抑制因子SCR7则收效甚微。Boel等^[24]在斑马鱼中使用多种cNHEJ抑制剂(SCR7、NU7441、KU0060648)和HDR促进因子(RS1、L755507)在多个位点进行了敲入效率比较,在一些位点敲入效率确有提高,还有一些位点可能因为抑制剂毒性等原因使得敲入效率降低。DNA损伤修复过程中的相关蛋白过表达或抑制能否提高斑马鱼基因敲入效率呢?由于物种间演化上的差异,需要在斑马鱼中进一步的验证,与哺乳动物中的作用机制的异同也需要进一步的比较研究。

3 结论与展望

每种敲入修复方式都有其特点,根据不同的实验目的来选择敲入修复方式。HDR敲入通常需要较长的同源臂、步骤繁琐、发生细胞周期短等

因素使得HDR敲入效率低。但在诸如构建点突变或短片段的精准替换模型实验时, 利用HDR可进行精准基因敲入。选择长单链DNA为修复模板时, 需要考虑模板长度、同源臂长度、对称性和模板链的选择等条件。

针对单个目的基因进行精准敲入构建荧光标记敲入品系时, 可以选择以内含子靶向cNHEJ介导的敲入方式进行构建。与HDR相比, 虽然构建过程繁琐程度相似, 但敲入效率可能更高, 而且可以通过荧光表达时空检测是否获得目的敲入品系。若需要大量制备敲入模板进行大批量敲入构建时, 选择同源臂较长的HDR可能实验过程较为繁琐, 且准确率也相应降低, 以内含子靶向cNHEJ介导的精准敲入构建也同样较为繁琐。可选择MMEJ为修复方式, 通过在引物合成时添加短同源臂的方式来减少实验步骤, 是提高实验效率的方案之一。除了修复机制对敲入效率的影响, 切割效率对敲入效率也有极大的影响。在进行敲入实验前需要对切割效率进行优化, 具体优化方式包括Cas9的类型、sgRNA的类型、sgRNA切割位点以及递送方式等。

综上所述, 大量研究对CRISPR/Cas9系统作用过程中的多个因素如修复类型、Cas9类型、sgRNA类型、CRISPR/Cas9与修复模板亲和性等进行了优化, 建立了提高敲入效率的方法。由于影响因素众多且相关分子机制仍不清楚等, 斑马鱼体系中进行基因精准敲入仍比较困难。对多个物种的相关研究进行分析和借鉴, 可能提高斑马鱼基因敲入效率, 同时将有利于构建人类疾病相关的斑马鱼模型, 推动临床医学以及生命科学基础研究的发展。

参考文献

- [1] Jinek M, Chylinski K, Fonfara I, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, 2012, 337(6096): 816-821
- [2] Deltcheva E, Chylinski K, Sharma CM, et al. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature*, 2011, 471(7340): 602-607
- [3] 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
- [4] Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*, 2013, 339(6121): 819-823
- [5] Jinek M, East A, Cheng A, et al. RNA-programmed genome editing in human cells. *eLife*, 2013, 2: e00471
- [6] Mali P, Yang L, Esvelt KM, et al. RNA-guided human genome engineering via Cas9. *Science*, 2013, 339(6121): 823-826
- [7] Xue C, Greene EC. DNA repair pathway choices in CRISPR-Cas9-mediated genome editing. *Trends Genet*, 2021, 37(7): 639-656
- [8] Scully R, Panday A, Elango R, et al. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat Rev Mol Cell Biol*, 2019, 20(11): 698-714
- [9] Symington LS, Gautier J. Double-strand break end resection and repair pathway choice. *Annu Rev Genet*, 2011, 45(1): 247-271
- [10] Truong LN, Li Y, Shi LZ, et al. Microhomology-mediated end joining and homologous recombination share the initial end resection step to repair DNA double-strand breaks in mammalian cells. *Proc Natl Acad Sci USA*, 2013, 110(19): 7720-7725
- [11] Seol JH, Shim EY, Lee SE. Microhomology-mediated end joining: good, bad and ugly. *Mutat Res Fundamental Mol Mech Mutagenesis*, 2018, 809: 81-87
- [12] Taleei R, Nikjoo H. Biochemical DSB-repair model for mammalian cells in G₁ and early S phases of the cell cycle. *Mutat Res Genet Toxicol Environ Mutagenesis*, 2013, 756(1-2): 206-212
- [13] Thomas KR, Capecchi MR. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. *Cell*, 1987, 51(3): 503-512
- [14] Urnov FD, Miller JC, Lee YL, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*, 2005, 435(7042): 646-651
- [15] Zu Y, Tong X, Wang Z, et al. TALEN-mediated precise genome modification by homologous recombination in zebrafish. *Nat Methods*, 2013, 10(4): 329-331
- [16] Gaudelli NM, Komor AC, Rees HA, et al. Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage. *Nature*, 2017, 551(7681): 464-471
- [17] Anzalone AV, Randolph PB, Davis JR, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*, 2019, 576(7785): 149-157
- [18] Petri K, Zhang W, Ma J, et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat Biotechnol*, 2022, 40(2): 189-193
- [19] de Vrieze E, de Bruijn SE, Reurink J, et al. Efficient generation of knock-in zebrafish models for inherited

- disorders using CRISPR-Cas9 ribonucleoprotein complexes. *Int J Mol Sci*, 2021, 22(17): 9429
- [20] Ranawakage DC, Okada K, Sugio K, et al. Efficient CRISPR-Cas9-mediated knock-in of composite tags in zebrafish using long ssDNA as a donor. *Front Cell Dev Biol*, 2020, 8: 598634
- [21] Kesavan G, Machate A, Brand M. CRISPR/Cas9-based split fluorescent protein tagging. *Zebrafish*, 2021, 18(6): 369-373
- [22] Eschstruth A, Schneider-Maunoury S, Giudicelli F. Creation of zebrafish knock-in reporter lines in the *nefma* gene by Cas9-mediated homologous recombination. *Genesis*, 2020, 58(1): e23340
- [23] Bai H, Liu L, An K, et al. CRISPR/Cas9-mediated precise genome modification by a long ssDNA template in zebrafish. *BMC Genomics*, 2020, 21(1): 67
- [24] Boel A, De Saffel H, Steyaert W, et al. CRISPR/Cas9-mediated homology-directed repair by ssODNs in zebrafish induces complex mutational patterns resulting from genomic integration of repair-template fragments. *Dis Model Mech*, 2018, 11(10): dmm035352
- [25] Prykhodzhiy SV, Fuller C, Steele SL, et al. Optimized knock-in of point mutations in zebrafish using CRISPR/Cas9. *Nucleic Acids Res*, 2018, 46(17): e102
- [26] Irion U, Krauss J, Nüsslein-Volhard C. Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. *Development*, 2014, 141(24): 4827-4830
- [27] Roth TL, Puig-Saus C, Yu R, et al. Reprogramming human T cell function and specificity with non-viral genome targeting. *Nature*, 2018, 559(7714): 405-409
- [28] Carrington B, Ramanagoudr-Bhojappa R, Bresciani E, et al. A robust pipeline for efficient knock-in of point mutations and epitope tags in zebrafish using fluorescent PCR based screening. *BMC Genomics*, 2022, 23(1): 810
- [29] Bennardo N, Cheng A, Huang N, et al. Alternative-NHEJ is a mechanistically distinct pathway of mammalian chromosome break repair. *PLoS Genet*, 2008, 4(6): e1000110
- [30] Cannavo E, Cejka P. Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature*, 2014, 514(7520): 122-125
- [31] Arana ME, Seki M, Wood RD, et al. Low-fidelity DNA synthesis by human DNA polymerase theta. *Nucleic Acids Res*, 2008, 36(11): 3847-3856
- [32] Luo JJ, Bian WP, Liu Y, et al. CRISPR/Cas9-based genome engineering of zebrafish using a seamless integration strategy. *FASEB J*, 2018, 32(9): 5132-5142
- [33] Li J, Zhang B, Ren Y, et al. Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system. *Cell Res*, 2015, 25(5): 634-637
- [34] Wiersen WA, Welker JM, Almeida MP, et al. Efficient targeted integration directed by short homology in zebrafish and mammalian cells. *Elife*, 2020, 9: e53968
- [35] Hisano Y, Sakuma T, Nakade S, et al. Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. *Sci Rep*, 2015, 5(1): 8841
- [36] Liu F, Kambakam S, Almeida MP, et al. Cre/lox regulated conditional rescue and inactivation with zebrafish UFlip alleles generated by CRISPR-Cas9 targeted integration. *Elife*, 2022, 11: e71478
- [37] Auer TO, Durore K, De Cian A, et al. Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. *Genome Res*, 2014, 24(1): 142-153
- [38] Levic DS, Yamaguchi N, Wang S, et al. Knock-in tagging in zebrafish facilitated by insertion into non-coding regions. *Development*, 2021, 148(19): dev199994
- [39] Kesavan G, Chekuru A, Machate A, et al. CRISPR/Cas9-mediated zebrafish knock-in as a novel strategy to study midbrain-hindbrain boundary development. *Front Neuroanat*, 2017, 11: 52
- [40] Kimura Y, Hisano Y, Kawahara A, et al. Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. *Sci Rep*, 2014, 4(1): 6545
- [41] Li W, Zhang Y, Han B, et al. One-step efficient generation of dual-function conditional knockout and geno-tagging alleles in zebrafish. *Elife*, 2019, 8: e48081
- [42] Li J, Li HY, Gu SY, et al. One-step generation of zebrafish carrying a conditional knockout-knockin visible switch via CRISPR/Cas9-mediated intron targeting. *Sci China Life Sci*, 2020, 63(1): 59-67
- [43] Han B, Zhang Y, Bi X, et al. Bi-FoRe: an efficient bidirectional knockin strategy to generate pairwise conditional alleles with fluorescent indicators. *Protein Cell*, 2021, 12(1): 39-56
- [44] Abdelmoneim A, Clark CL, Mukai M. Fluorescent reporter zebrafish line for estrogenic compound screening generated using a CRISPR/Cas9-mediated knock-in system. *Toxicol Sci*, 2020, 173(2): 336-346
- [45] Kesavan G, Hammer J, Hans S, et al. Targeted knock-in of CreER T2 in zebrafish using CRISPR/Cas9. *Cell Tissue Res*, 2018, 372(1): 41-50
- [46] Gu S, Li J, Li S, et al. Efficient replacement of long DNA fragments via non-homologous end joining at non-coding regions. *J Mol Cell Biol*, 2021, 13(1): 75-77
- [47] Albadri S, Del Bene F, Revenu C. Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish. *Methods*, 2017, 121-122: 77-85
- [48] Liu X, Homma A, Sayadi J, et al. Sequence features

- associated with the cleavage efficiency of CRISPR/Cas9 system. *Sci Rep*, 2016, 6(1): 19675
- [49] Zhang JP, Li XL, Li GH, et al. Efficient precise knockin with a double cut HDR donor after CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol*, 2017, 18(1): 35
- [50] Okamoto S, Amaishi Y, Maki I, et al. Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. *Sci Rep*, 2019, 9(1): 4811
- [51] Hoshijima K, Jurynek MJ, Klatt Shaw D, et al. Highly efficient CRISPR-Cas9-based methods for generating deletion mutations and F0 embryos that lack gene function in zebrafish. *Dev Cell*, 2019, 51(5): 645-657.e4
- [52] Cromwell CR, Sung K, Park J, et al. Incorporation of bridged nucleic acids into CRISPR RNAs improves Cas9 endonuclease specificity. *Nat Commun*, 2018, 9(1): 1448
- [53] Fu Y, Sander JD, Reyon D, et al. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*, 2014, 32(3): 279-284
- [54] Rahdar M, McMahon MA, Prakash TP, et al. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc Natl Acad Sci USA*, 2015, 112(51): 7110-7117
- [55] Suzuki K, Tsunekawa Y, Hernandez-Benitez R, et al. *In vivo* genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature*, 2016, 540(7631): 144-149
- [56] Hoshijima K, Jurynek MJ and Grunwald DJ. Precise genome editing by homologous recombination. *Methods Cell Biol*, 2016, 135: 121-147
- [57] Aird EJ, Lovendahl KN, St. Martin A, et al. Increasing Cas9-mediated homology-directed repair efficiency through covalent tethering of DNA repair template. *Commun Biol*, 2018, 1(1): 54
- [58] Gu B, Posfai E, Rossant J. Efficient generation of targeted large insertions by microinjection into two-cell-stage mouse embryos. *Nat Biotechnol*, 2018, 36(7): 632-637
- [59] Seleit A, Aulehla A, Paix A. Endogenous protein tagging in medaka using a simplified CRISPR/Cas9 knock-in approach. *Elife*, 2021, 10: e75050
- [60] Aida T, Nakade S, Sakuma T, et al. Gene cassette knock-in in mammalian cells and zygotes by enhanced MMEJ. *BMC Genomics*, 2016, 17(1): 979
- [61] He MD, Zhang FH, Wang HL, et al. Efficient ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in zebrafish embryos. *Mutat Res Fundamental Mol Mech Mutagenesis*, 2015, 780: 86-96
- [62] Maruyama T, Dougan SK, Truttmann MC, et al. Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining. *Nat Biotechnol*, 2015, 33(5): 538-542
- [63] Yu C, Liu Y, Ma T, et al. Small molecules enhance CRISPR genome editing in pluripotent stem cells. *Cell Stem Cell*, 2015, 16(2): 142-147
- [64] Song J, Yang D, Xu J, et al. RS-1 enhances CRISPR/Cas9- and TALEN-mediated knock-in efficiency. *Nat Commun*, 2016, 7(1): 10548