Methods 121–122 (2017) 77–85

Contents lists available at ScienceDirect

Methods

journal homepage: www.elsevier.com/locate/ymeth

Genome editing using CRISPR/Cas9-based knock-in approaches in zebrafish

Shahad Albadri, Filippo Del Bene *, Céline Revenu

Institut Curie, PSL Research University, INSERM U934, CNRS UMR3215, 75248 Paris Cedex 05, France

ARTICLE INFO

Article history:
Received 15 December 2016
Received in revised form 13 February 2017
Accepted 3 March 2017
Available online 12 March 2017

Keywords:
Knock-in
CRISPR/Cas9
Genome editing
Homologous recombination
Non-homologous-end-joining
Zebrafish

ABSTRACT

With its variety of applications, the CRISPR/Cas9 genome editing technology has been rapidly evolving in the last few years. In the zebrafish community, knock-out reports are constantly increasing but insertion studies have been so far more challenging. With this review, we aim at giving an overview of the homologous directed repair (HDR)-based knock-in generation in zebrafish. We address the critical points and limitations of the procedure such as cutting efficiency of the chosen single guide RNA, use of *cas9* mRNA or Cas9 protein, homology arm size etc. but also ways to circumvent encountered issues with HDR insertions by the development of non-homologous dependent strategies. While imprecise, these homology-independent mechanisms based on non-homologous-end-joining (NHEJ) repair have been employed in zebrafish to generate reporter lines or to accurately edit an open reading frame by the use of intron-targeting modifications. Therefore, with higher efficiency and insertion rate, NHEJ-based knock-in seems to be a promising approach to target endogenous loci and to circumvent the limitations of HDR whenever it is possible and appropriate. In this perspective, we propose new strategies to generate cDNA edited or tagged insertions, which once established will constitute a new and versatile toolbox for CRISPR/Cas9-based knock-ins in zebrafish.

Contents

* Corresponding author.
E-mail address: filippo.del-bene@curie.fr (F. Del Bene).

http://dx.doi.org/10.1016/j.ymeth.2017.03.005
1046-2023/

1. Introduction

The emergence of programmable nucleases enables the targeted generation of DNA double strand breaks (DSBs) that activate cellular repair mechanisms. DSB repair can be error-prone generating small insertions and deletions (indels) but can also favour the insertion of exogenous DNA molecules either randomly or using homology arms. Programmable nucleases have thus been recognized as promising tools to knock-in genetic elements and to allow in frame insertions, the goal of which can be to fuse a tag to the open reading frame (ORF) of interest or to precisely introduce a desired mutation, for example identified in patients. The cleanest and most obvious technique for in frame knock-in (KI) is to trigger by homologous recombination (HR) the insertion of a donor DNA flanked by homologous sequences to both sides of the break in the locus of interest [1]. Short after the development of TALENS, the CRISPR/Cas9 complex has appeared as the most straightforward and simple way to induce targeted DNA DSBs and thus to activate cell repair mechanisms and stimulate KI [2]. This is routinely used for homology directed repair (HDR) in cell culture, where the possibilities of antibiotic selection of the correctly edited cells bring a considerable advantage (see Wassef et al., this issue). HDR-based integrations are however much less efficient in model organisms like mouse or zebrafish.

The main focus of this article is to analyse recent technical advances aiming at inserting with CRISPR/Cas9 genetic elements in frame in the zebrafish genome, in order to precisely modify a gene and study its function. As we will see, although feasible, HDR is however still challenging in zebrafish and requires further improvements. Consequently, other methods avoiding HDR while still using the endogenous regulatory landscape of the target gene are being developed and are promising alternatives to homology based KI.

2. The crucial step: good sgRNA to induce DSB

Highly efficient generation of DSBs stands out as the essential prerequisite to attempt KI in zebrafish. In most of the HDR success reports, sgRNAs able to induce indels in nearly every injected embryo [3–5] or resulting in above 70% of edited alleles on pools of injected embryos [6,7] have been selected to pursue further with HDR. Several detailed reviews and method papers propose protocols for sgRNA synthesis [8–10]. Here, we only will try to give an update on sgRNA selection in the perspective of optimising KI strategies.

2.1. Prediction of in vivo cutting efficiencies

Several scoring algorithms are available to predict best sgRNA and optimise cutting activity. One of them is based on a systematic analysis of 1280 sgRNAs targeting 128 zebrafish genes [11], others have been developed mainly in mammalian cells but also in other model organisms [12]. A detailed comparison of scoring algorithms with existing data sets shows highly variable correlations between scoring and actual efficiencies, and unfortunately especially low correlations for zebrafish data [12]. The Moreno-Mateos scoring however appears as the best suited for *in vitro* transcribed, injected sgRNA. From this investigation, a useful website (CRISPOR) that calculates scorings from the different algorithms but also predicts off-targets is available. In line with this report, no good correlation between scoring and *in vivo* sgRNA cleavage activity could be observed with the sgRNA sequences tested in zebrafish in our laboratory. Albeit based on zebrafish data, our results also appear uncorrelated with the Moreno-Mateos scoring (Fig. 1A).

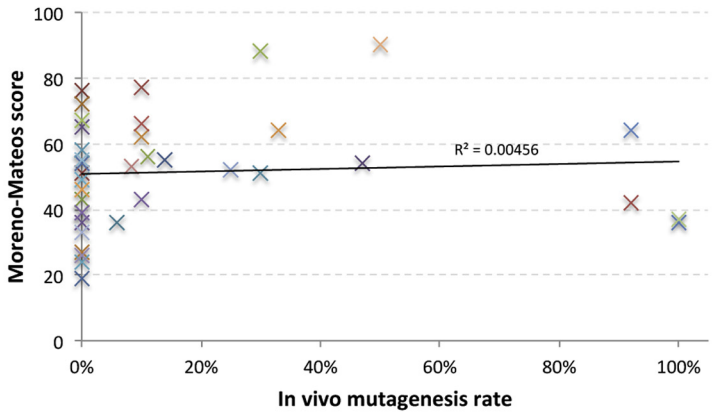
In parallel, we have tried to evaluate sgRNAs *in vitro* by cleaving PCR products of the genomic region targeted. In our hands, the efficiency of cleavage observed on gels after electrophoresis of the cleaved PCR products does not correlate with the *in vivo* efficiencies (Fig. 1B). As an example, in Fig. 1B, sgRNA #4 cuts poorly *in vitro* compared to the other tested sgRNAs, whereas it performs best *in vivo* for the analysed locus. We would thus not consider *in vitro* cutting efficiencies to rank sgRNA. We speculate that the chromatin environment of a given locus and at a given time *in vivo* can probably influence CRISPR/Cas9 cutting efficiency and this cannot be reconstituted *in vitro*. However, a sgRNA that does not cut *in vitro*, where the target site is readily available, should be excluded for further *in vivo* analysis [13]. The *in vitro* test can thus be used as a yes or no pre-screening of sgRNAs but is, according to our experience, not able to help select the best candidate among several *in vitro* cutting sgRNAs.

2.2. mRNA versus protein Cas9 injection

First studies using the CRISPR/Cas9 technique to generate DSBs and indels in zebrafish have injected sgRNAs with mRNA encoding the Cas9 protein [10]. More recently, the Cas9 protein has been directly injected in one-cell stage embryos and shown to significantly increase cleavage [8,14]. For comparison, we quantified the two techniques with a previously published sgRNA targeting the *kif5aa* locus [15]. The sgRNA was diluted at 300 ng/μL with either 150 ng/μL *Cas9* mRNA or 2 μM Cas9 protein. Approximately 1nL of this mix was injected per embryo. PCR of the locus was performed on lysates of pools of 30 embryos (24hpf) and sequenced to assess the rate of indels induced. Whereas on average 20.5% of indels were detected with *Cas9* mRNA, it reached between 90 and 100% of indels with Cas9 protein injections (either commercial from NEB or in house purified batches, kind gifts of J-P Concordet and H. Besir). Cutting efficiencies are thus largely improved by Cas9 protein injections. Moreover, survival rate is also enhanced, probably the consequence of lower toxicity of the protein *versus* mRNA. With the injection conditions described here, survival rates after *Cas9* mRNA injections vary from 70% to no survival depending on the sgRNA whereas no toxicity is detected when using Cas9 protein (meaning roughly 100% survival). Consequently, the equimolar concentrations of Cas9 protein and sgRNA injected can be largely increased, up to the mM range [14], further improving cutting efficiencies.

In line with these reports and our results, we routinely select sgRNAs exhibiting more than 80% cutting efficiency *in vivo*. sgRNAs are systematically designed against in house sequenced regions, to avoid polymorphism with the reference genome and selected on the base of low off target risk as identified using the CRISPOR website. To test sgRNA *in vivo*, one-cell stage embryos are injected with 300 ng/μL sgRNA and 2 μM Cas9 protein. Efficiency is estimated by sequencing >10 PCR fragments amplified from a pool of 30 injected embryos at approximately 24hpf (hours post fertilisation). Although weakly cutting sgRNAs are probably missed, we consider that a minimum of 10 PCR fragments sequenced is sufficient to detect highly efficient sgRNA. Based on our experience, this

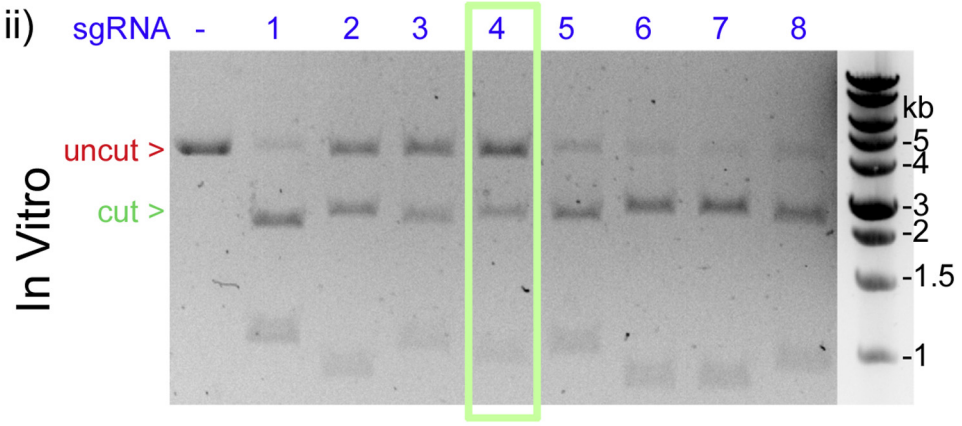
A



B i)

gtcctcaaatagaagcttctggaagTTTACTCGAATCTTTACTCCAAAATAAAAGGAAAATGCACATTATAAATGAGACTTCTGTATAGAA
TATACACCGTTTCGTTTGAGAAAAATCTTTAAGGGCACTGATTAACAAACACGTTCATGGTCATCATGACACTTTTTCTTCTCCAT
CGGTGGTTTCATGCTTTACTTTCAATGTTTGTTTGCGTGACGGCGTTGGGTTCAAATCCTGCATCGGGTGCCACTTTCTTCAGCTC
TGCATTTTCATCCTCCAGAATTTTCCTGCGATTGAGACCATTAAAGGGATAGTTTAGATTAAAATGACAATCCTTTCAGCATTTAC
CCTCCCTCACGTCATTCCAAACATGTATGATTCTTTCTTCTGTTAGTCGACTTGACGAATTTAACTCACTGACTGGGTTGGATCAG
TTCACAGATCACTGGAAAAATGTTTATAAACTATTTCAAGTTATTACAGAGTTAGTATACTGAAagtattgcatgatgcgtgtactg

ii)



iii)

In Vivo

sgRNA	TARGET pam	In vivo cut efficiency
1	GTTTGGAATGACGTGAGGGA ggg	8.3%
2	GTAAAGCATGAAACCACCGA tgg	9%
3	CTGCGATTGAGACCATTAAA ggg	0%
4	GTTGGGTTCAAATCCTGCAT cgg	36%
5	CAATCGCAGGAAAATTCTGG agg	0%
6	CGTTTGAGAAAAATCTTTAA ggg	0%
7	GTTTGAGAAAAATCTTTAA ggg	0%
8	TTCAATGTTTGTTTGCGTGA cgg	0%

Fig. 1. Predicting sgRNA cutting efficiencies remains a challenge. (A) Moreno-Mateos scoring efficiency and *in vivo* mutagenesis rate comparison reveals no correlation between the predicted and effective mutagenic potential of sgRNAs in zebrafish. 37 sgRNAs with variable scorings from the Moreno-Mateos algorithm were tested *in vivo* after injection into 1-cell stage zebrafish embryos in equimolar amounts with the Cas9 protein (NEB). The efficiency was evaluated on pools of injected embryos after DNA extraction, PCR amplification of the targeted locus, Topo-TA cloning and sequencing. For each targeted locus, 10–20 sequenced clones were aligned against the wild-type sequence of the locus and analysed for potential mutations. The predicted scores were plotted against the *in vivo* mutagenic efficiencies. The regression line and value of the coefficient of determination (R^2) indicate no correlation between Moreno-Mateos scoring values and the effective cutting efficiency of the tested sgRNAs. (B) Comparison of *in vitro versus in vivo* cutting efficiencies. (i) Sequence of the Ncam1b zebrafish locus amplified by PCR (primers are indicated in lowercase) with the sgRNA targets underlined. (ii) The 8 sgRNAs targeting the *ncam1b* gene were tested *in vitro* for their cutting potential after incubation with the amplified targeted locus, Cas9 protein and buffer (NEB) at 37 °C for 3 h. As negative control, the amplified targeted locus, Cas9 protein and buffer were incubated without sgRNA, resulting in no cutting as expected (–). All tested sgRNAs (1–8) were able to generate a cut with variable efficiencies: (1, 5–8) with high efficiency and (2–4) with lower efficiency. (iii) All sgRNAs were then injected equimolar with the Cas9 protein (1 μM) and their *in vivo* cutting efficiencies were evaluated (in the same way as in A, with 12 clones analysed by sequencing per sgRNA). The obtained results show no correlation between the *in vitro* and *in vivo* cutting efficiencies as highlighted with the sgRNA 4 (green), which was the least efficient at cutting *in vitro* but gave the best cutting results *in vivo*. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

method is the most robust, time and cost effective to routinely evaluate any sgRNA, suitable for KI applications.

3. Inserting exogenous DNA in frame with the gene of interest

3.1. Cell repair mechanisms used to KI genetic elements

As we just saw, the generation of mutations has now become pretty straightforward with the high efficiency of DSBs enabled by CRISPR/Cas9. It is mainly based on the non-homologous end joining (NHEJ) DNA damage repair mechanism, which is error prone and creates small insertions and deletions (indels) by directly ligating DNA ends. NHEJ has been used in zebrafish to knock in genetic elements at a target site with CRISPR/Cas9 [16,17] by co-injecting the CRISPR/Cas9 machinery with the genetic elements to be inserted. It however reduces the efficiency by 6 for in frame KI (1 out of 2 being correctly oriented and 1 out of 3 being in frame), which is problematic for an already challenging goal. Moreover, editing is not precise and the junctions between the donor DNA and the locus are thus unpredictable.

Homology based repair mechanisms have consequently been tested to achieve precise knock-in in zebrafish. They are of two main types. Microhomology-mediated end joining (MMEJ) uses very short homology sequences (down to 5 bases) internal to the broken DNA to join ends [18,19]. Homologous recombination is based on long homology arms (up to several kb) and results in the exchange of large nucleotide sequences between two similar DNA molecules (for review see [20]). As the precise repair molecular mechanisms that are effectively implicated in the zebrafish KI events reported in the literature have not been extensively investigated, we will refer to microhomology-directed repair (mHDR) in the case of short homology arms (up to approximately 50 nucleotides) and to large homology-directed repair (lHDR) for long homology arms (several hundreds of nucleotides). The different approaches discussed and their outcomes are summarised in Table 1.

Table 1
Different strategies used and their outcomes to KI elements in zebrafish. This table presents in a synthetic way several KI methods reported so far and their somatic efficiencies (in injected embryos) and germline targeting efficiencies (transmitting adult founder fish). It highlights the variability of the techniques and the lack of a standard methodology to generate and evaluate it. FP fluorescent protein, RFLP restriction fragment length polymorphism, NGS next generation sequencing, T7E1 T7 endonuclease 1 assay.

		locus	Strategy	sgRNA/TALEN efficiency	KI Screening	somatic efficiency (% embryos with some KI events)	precise germline targetting efficiency (% of founder fish)
mHDR, arms < 50 nucleotides	Hruscha et al. [7]	*c9orf72*	CRISPR oligonucleotides HA tag insertions	NGS 50-76% mutagenesis rate	PCR NGS	45% 1,7% precise	2/39 not precise, fin prescreening 0.0%
		tardbp	CRISPR oligonucleotides HA tag insertions	NGS 61-98% mutagenesis rate	PCR NGS	70% 3,5% precise	na
	Hwang et al. [21]	*several loci*	CRISPR oligonucleotides substitution/insertions	25-55% by T7EI assay	PCR+ seq	0 to 8,3%	na
	Armstrong et al. [22]	*tardbp*	CRISPR oligonucleotides base substitutions	na	PCR, RFLP and sequencing	na	1/46 and 3/77 2,2% and 3,9%
		fus	CRISPR oligonucleotides base substitution	na	PCR, RFLP and sequencing	na	1/47 2.1%
	Hisano et al. [6]	*keratin krtt1c119e*	CRISPR arms 40bp FP insertions	73%	GFP + cells PCR+ seq	38,0% (201/529) 100% of GFP+ 37,5% of GFP-	extrapolation: 4,5% of injected larvae 100% from broad GFP + 12,5% from intermediate GFP+
lHDR, arms > 400 nucleotides	Irion et al., [5]	*albino*	CRISPR arms 400-500bp base substitution	95% of embryos show defects in pigmentation	pigmentation (rescue of mutation) and PCR+ seq	46% of injected embryos show some pigmentation	3/28 10.7%
	Zhang et al. [23]	*twist2*	CRISPR arms 400-500bp base substitution	65% by T7EI assay	PCR	na	2/60 3.3%
	Shin et al. [24]	*sox2* *gfap*	TALEN arms>1kb FP insertion	high	FP + cells and PCR+ seq	up to 90% na	29/363 8,0% 5/44 11,3%
	Zu et al. [25]	*tyrosine hydroxylase*	TALEN arms 800-900bp FP insertion	>70%	PCR	approx 50%	4/275 1.5%
	Hoshijima et al. [3]	*golden*	TALEN arms>1kb base substitution	100% (every embryo has indels)	pigmentation and PCR	100%	7/51 13.7%
			adding fluorescent reporter (lens)		pigmentation, PCR and fluorescent lens	100%	4/39 10.3%
		kcnh6a	TALEN arms>1kb, I-sceI FP insertions	100% (every embryo has indels)	PCR	100%	2/14 14.3%
			adding fluorescent reporter (lens)		PCR and fluorescent lens	100%	3/43 7.0%

3.2. Different homology-based KI methods and their achievements

3.2.1. mHDR based KI

mHDR-based KI has first been reported using CRISPR/Cas9 in zebrafish for insertion of short nucleotide sequences that could be provided as a donor oligonucleotide [7,21]. However, in both reports, additional indels are often detected together with the integration of the intended sequence. For example, Hruscha et al. successfully knocked-in one or two HA tags (27 nucleotides) with homologous arms between 30 and 50 bp [7]. Knocked-in HA tag could be detected by PCR in 45% and 70% of injected embryos in the *C13H9orf72* and *tardbp* loci respectively. However, further investigations by sequencing revealed only 1.7% and 3.5% of perfect reads in the injected embryos. This remains low, especially considering that these are somatic modifications and might thus not be transmitted in the next generation. By injecting single stranded oligonucleotides with homology arms ranging from 7 to 53, a recent study could report the precise introduction of point mutations corresponding to identified mutations in ALS patients in the genes *tardbp* and *fus* [22]. Analysis of F1 larvae revealed that 2–4% of injected F0 transmitted the correctly edited allele to their progeny.

Larger elements, like a fluorescent protein coding sequence, have also been successfully inserted using short homology arms and a donor vector. Arms of 20 and 40 bp were giving best results on the *tyrosinase* locus, with approximately 80% of injected embryos with PCR-detected integration events of which more than 70% were precisely integrated [6]. This KI should have resulted in a Tyrosinase-mCherry fusion protein that was however undetectable by fluorescence. Using 40 bp homology arms, fusion of GFP C-terminal to the Keratin type 1 c19e protein could be generated and GFP fluorescence detected in injected embryos [6]. About a third of injected embryos showed from broad (less than 3% of reported embryos) to sparse GFP labelling in the epidermis. Embryos were sorted in three groups depending on the extent of fluorescence observed (broad, intermediate or narrow) and raised to adulthood. Few adults were screened from each group to find transmitting founders (broad 2/2-100%, intermediate 1/8-12.5% and narrow 0/10-0%) [6]. If we extrapolate from these reported percentages to the injected F0 population, an estimation of 4.5% of F1 transmitting F0 founders can be calculated, mainly isolated from the broadly GFP expressing group.

3.2.2. lHDR based KI

With the goal of making precise nucleotide substitutions, Irion et al. [5] and Zhang et al. [23] have used CRISPR/Cas9 and lHDR with a donor plasmid containing approximately 900 bp of the *albino* gene to restore pigmentation in albino zebrafish and of the *twist2* gene to mimic a human mutation found in Ablepharon macrostomia, respectively. The first study achieved nearly 10% of injected F0 able to produce pigmented larvae and the second obtained slightly more than 3% transmitting founders. Reports of successful integration of large genetic elements, as a fluorescent protein (>700 bp), with CRISPR/Cas9 using lHDR are still missing in zebrafish. Using TALENs, homologous recombinant integration of fluorescent reporters (length of 0.7–1.5 kb) was achieved with efficiencies ranging from 1.5 to 11% germ line transmission frequencies [24,25]. A recent study reported up to 14% germline transmission frequency of correctly edited alleles with homology arms of 1 kb flanking the 1.6 kb sequence to be inserted in the heart potassium channel *kcnh6a* locus [3].

This overview of the literature shows significant differences in reported efficiencies of HDR edited genomes (Table 1). This can probably largely be explained by the variability associated with each different locus that could influence both the efficiency of DSB induction and of the repair mechanisms available. It should also be noted that there is a high variability in the techniques used and no good standard currently exists, which hinders the comparison of the reported results (see Table 1). However, we can see that highest efficiencies of F1 retrieval of correctly edited alleles have been reported with the use of long homology arms, independently of the length of the sequence to insert (point mutations or large reporter sequences). Also feasible, HDR based genome editing in zebrafish remains a challenge and this highlights the need to keep on working on the technique.

3.3. Tips and tricks to improve HDR methodology

Several teams have already tested and set up methods to improve HDR efficiencies and identification of founders. Although it might still be highly locus and donor construct dependent, these reports point at a series of features to consider when designing the strategy to be used.

3.3.1. Conformation of donor DNA

Donor DNA vectors have been reported to give precisely edited alleles both when injected linearized by restriction enzymes [23–25] or as a circular plasmid. When injected as a circular plasmid, it can be designed to contain enzymatic or CRISPR/Cas9-based cleavage sites enabling linearization in the cell after injection [3,6,16]. Several studies have systematically tested the efficiencies of both methods. For the rescue of the albino locus, the toxicity of the linear donor prevented injection at high concentration resulting in low efficiencies of F0 somatic rescue. Injection of the circular vector containing CRISPR target sites drastically enhanced efficiency from 2.5% to 46% of injected larvae harbouring pigments compared to the circular vector without cleavage sites [5]. In the same direction, Hoshijima et al. reported better results with circular compared to linearized donors [3]. They also tested digestion with the meganuclease I-SceI and reported a 3 to 4-fold increase in recombination efficiency whether pre-digested *in vitro* or directly injected with the enzyme compared to circular donor injections [3].

3.3.2. Favouring HDR

Another interesting direction that has been investigated to improve HDR is to favour MMEJ and HR repair mechanisms of DSBs by blocking NHEJ. This has been reported to work in mammalian cell culture and in mouse embryos by silencing key molecular components of the NHEJ repair pathway or by treating with Scr7, an inhibitor of one of these components, the DNA ligaseIV [26,27]. The injection of a Morpholino targeting *ligase 4* was shown to also enhance MMEJ in zebrafish, using as readout the repair of an exogenously provided GFP reporter [28].

3.3.3. Screening of edited alleles

One of the bottlenecks of HDR-based KI is the identification of potential founder fish, which requires the labour-intensive screening of hundreds of animals. In all cases reported, having a visual readout to select and grow only injected embryos with the highest level of somatic editing greatly improved the identification of transmitting founders [5,6]. Best is to have a clear visual read out, as pigment rescue [3,5], but this is rarely available. A strong fluorescent signal resulting from the KI and detectable with a binocular is ideal to screen both highly edited F0 and heterozygote F1. The generation of an endogenous fusion protein however often gives too low expression to be easily detected with a fluorescent binocular [6]. Pre-screening of F0 fin biopsies by PCR has also been used as an alternative way to evaluate highly edited founders that had no visual read-out [7]. The use of one of the classical zebrafish transgenesis reporter genes, as the alpha-crystallin lens reporter, added to the donor sequence can greatly improve successful KI

detection in the F1 progeny, as has been nicely developed in a recent paper [3]. In particular, the addition of FRT Flippase sites on both sides of the reporter enables its efficient excision once the line has been selected [3,4]. In the absence of visual reporter for the expected knock-in, pools of F1 have to be screened by PCR using primers specific to the donor and to the region outside of the homology arm. For this purpose, it is advisable to create homology arms of no more than 1 kb to ease PCR and sequencing of the resulting product. For the introduction of point mutations or very short sequences, inserting a restriction site by introducing silent mutations can prevent the need for systematic sequencing [3,22]. This can be coupled to the required modifications of the sgRNA target site in the donor vector to avoid donor cleavage if needed. Finally, PCR screening of F0 frozen sperm has also been reported to give good results to identify male founder fish, the F1 generation of selected male founders being then generated by *in vitro* fertilisation [23].

4. Alternative solutions developed to circumvent HDR to generate fusion proteins

4.1. Circumventing homology for accurate targeting

While HDR-based methods are ultimately the most precise and conservative approaches for insertion, they remain challenging and are hardly efficient. To circumvent this issue, the use of imprecise homology-independent mechanisms has been so far favoured in zebrafish to insert plasmid-based donors without homology arms in order to generate reporter lines. In 2014, to convert GFP lines into Gal4 lines, Auer et al. used a strategy employing a donor vector containing an eGFP bait sequence (Gbait, approximately 200 bp) and a Gal4 transcriptional trans-activator cassette, which was co-injected with an eGFP bait targeting sgRNA and *cas9* nuclease mRNA into 1-cell stage Tg(*geneX:GFP*) transgenic embryos [16]. Therefore, using the *eGFP* coding sequence as bait, the authors allowed the conversion of existing eGFP transgenic lines into Gal4 reporters. The molecular mechanism leading to this conversion was shown to be NHEJ and not HDR mediated and it revealed that HR type of insertions are not always required for targeted in frame insertions. Using a similar approach, the strategy was improved by adding a heat shock promoter (Hsp70) upstream the *gal4* or *eGFP* cassette into the donor vector, which allowed inducing the expression of the reporter gene independently from in-frame insertions within the transcriptional starting site of the gene of interest [17,29]. To date, several new reporter lines have been generated using this strategy, conferring a powerful alternative for homology independent repair over HDR-mediated integrations. This NHEJ-based approach however only applies to the generation of reporter lines and does not enable precise editing of the endogenous ORF.

4.2. Cell cycle and repair mechanisms: a consideration for KI approaches

As discussed in the first section of this review, Cas9 protein injection gives better results than mRNA injection for the generation of DSBs. However, in our hands for NHEJ-based insertion, the injection of the sgRNA mix with *cas9* mRNA surprisingly yields better results over purified Cas9 protein. Indeed, while Hoshijima et al. successfully generated HDR-mediated insertions using the Cas9 protein [3], the co-injection of Cas9 protein along with the *hsp70:gal4* cassette developed by Kimura et al. [17] into 1-cell stage Tg(*UAS:RFP*) embryos never lead to the observation of RFP signal in the cells of interest in our conditions (n = 6 different sgRNAs injected for 3 different loci in a minimum of 200 embryos, S. Albadri unpublished). In contrast to these observations, the same injections using *cas9* mRNA lead to the detection of RFP signal where expected. In addition this also allowed us to obtain successful insertions and founders in 30% of the raised *RFP*-expressing embryos. These results suggest that NHEJ-based insertion works better with *cas9* mRNA *versus* Cas9 protein.

We speculate that the difference between the two conditions (protein *versus* mRNA) could be related to the timing of the DSB generation and, in response, the choice of the repair system. In many model organisms, it was indeed reported that NHEJ repair mostly occurs during G1 (reviewed in [30–35]). During zebrafish early development, the first 12 divisions are rapid (about 15 min each) and with no or short interphases [36,37]. NHEJ repair mechanisms could thus be disadvantaged at these stages. Microinjections into 1-cell stage zebrafish embryos using Cas9 protein will likely trigger earlier genomic cuts than mRNA. We hypothesize that at such early stages, NHEJ repair is thus less favoured than later and that alternative pathways are at work to repair DSBs. Interestingly, recent work has shown that maternally provided important components of the classical NHEJ pathway are expressed at very low levels in early zebrafish embryos (until bud stage, 10hpf) [38]. Alternative end joining pathways are crucial during these first divisions to repair DNA lesions [38]. One possible hypothesis would therefore be that injection of *cas9* mRNA will generate cuts at later developmental stages after translation, folding and assembly of the Cas9/sgRNA complex, which would allow insertion events when interphases are longer and NHEJ repair mechanisms are prevalent ([39,40], Fig. 2). While this hypothesis trying to explain our results remains to be proven, such observations may be important to take into account when aiming at the generation of knock-in.

4.3. Intron-based KIs via homology-independent repair mechanisms

An alternative method has been developed to circumvent HDR while still achieving the desired editing of the ORF of interest by combining intron targeting mediated knock-in approaches with homology independent repair mechanisms [41,42]. This approach has so far proven to be an efficient solution to conduct efficient but imprecise modifications within the targeted locus while precisely modifying the targeted ORF and maintaining its integrity. Briefly, this procedure relies on choosing a highly efficient sgRNA whose target is located in the last intron of the locus of interest before the exon containing the stop codon and the 3′UTR of the targeted gene. The donor cassette then contains the genomic region from the site targeted by the sgRNA to the 3′ end of the locus with the desired modifications of the last exon of the ORF. This strategy therefore allows the conservation of the reading frame as well as the 5′ and 3′ regulatory regions of the locus of interest by recreating in the last intron an additional modified exon with all the necessary splicing elements. It uses all the advantages conferred by NHEJ, and restricts the junction errors to intronic sequences were they are less likely to have functional consequences. In the same study, the critical role of the 3′UTR sequence is also stressed for the regulation of the exogenous gene and therefore its cloning within the donor cassette is indeed a requirement. Finally, as for other knock-in strategies, the authors point at the cleavage efficiency of the selected sgRNA and its importance to ensure the success of the procedure.

4.4. Edited cDNA insertions via homology-independent repair mechanisms

As previously discussed, higher efficiency and insertion rate of NHEJ-based knock-in can be used to circumvent the limitations of HR-dependent mechanisms, while reconstituting the integrity

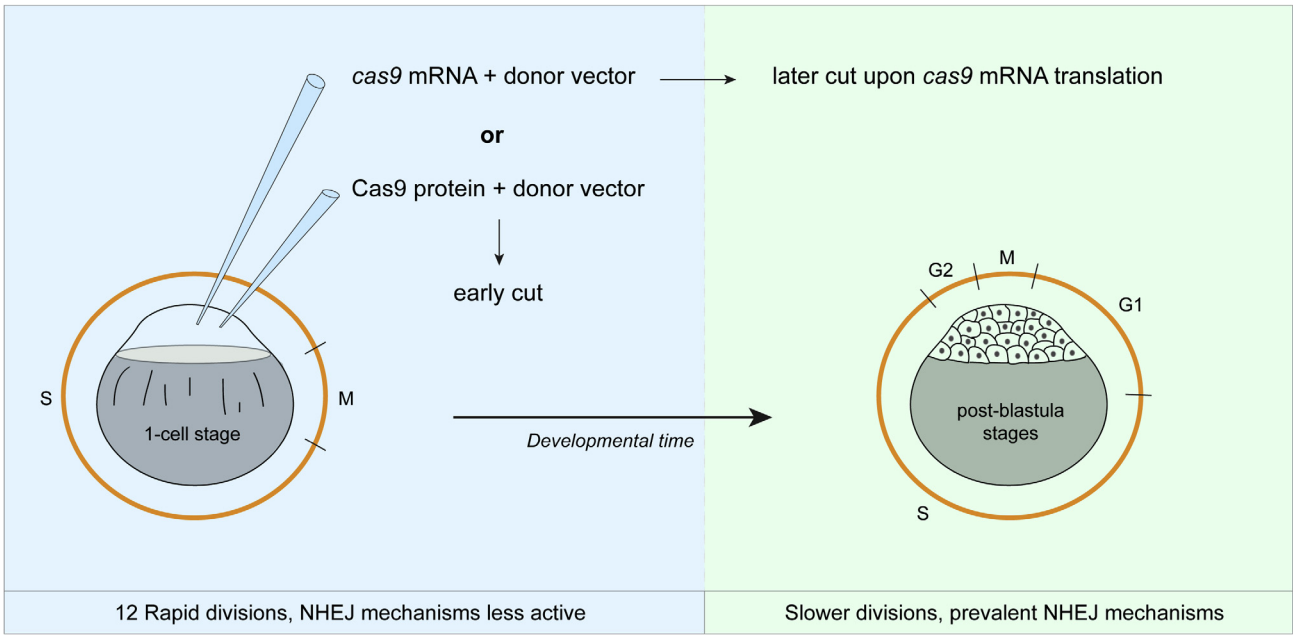


Fig. 2. Proposed model of the cell cycle during zebrafish development and repair mechanism on KI strategies. Schematic representation of two different types of KI strategies employing either *cas9* mRNA or Cas9 protein for DSB generation. The first cell divisions of the zebrafish embryos occur very rapidly (every 15 min), with little or no interphase (blue area). Consequently, during these early divisions, the classical NHEJ repair pathway would be mostly inactive. Later, as cell divisions time extend and interphases get longer, NHEJ mechanisms are preferentially used for DSB repair (green area). As a consequence, injection of Cas9 protein or *cas9* mRNA and the timing of the DSB generation could influence the type of repair mechanism used for the integration of the donor vector. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

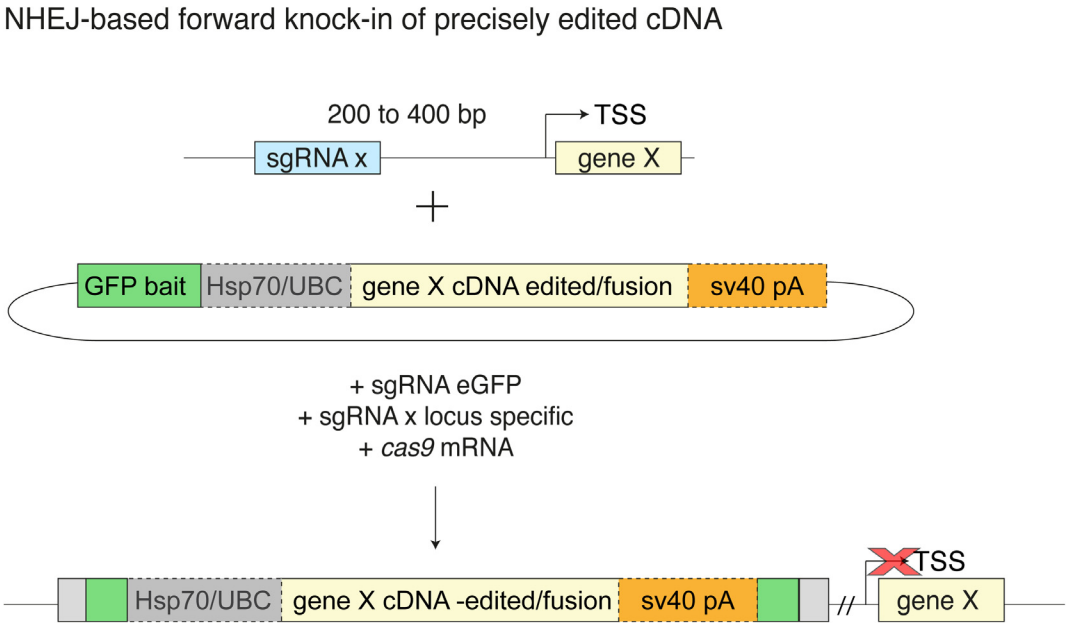


Fig. 3. Use of NHEJ mechanisms to generate accurately edited KI events. Schematic representation of a KI strategy, resulting in the expression of a precisely edited cDNA, using the NHEJ repair mechanism. This strategy aims at inserting a donor vector containing the precisely edited cDNA of the targeted gene (baring point mutations or fusion with a reporter) preceded by the eGFP bait cassette, which will be used for the linearization of the donor vector as described [16,17] and optionally a *hsp70* basal promoter or *ubc* intron cassette that will allow to boost the expression of the construct if necessary. The forward insertion of the entire construct upon linearization within the 400 first bp of the endogenous targeted locus will allow circumventing the efficiency issue linked to homology-based knock-in and allow the expression of the edited version of the gene of interest while disrupting the endogenous expression, which could be enhanced by the addition of an SV40-pA signal at the end of the tagged cDNA. Optional cassettes are depicted with dashed lines.

of the targeted locus. Combining the NHEJ KI method [16,17] with the strategy of insertion of the edited ORF, we propose to test a simplified approach aiming at inserting the modified cDNA of interest (either mutated or fused with a tag) before the ATG of the corresponding gene. In this case, the donor construct could include a similar vector containing Gbait (eGFP sgRNA target sequence) and the edited or fusion sequences of the cDNA of interest. Optionally the cDNA could be followed by an SV40 polyA (SV40pA) signal instead of the endogenous pA to prevent mRNA decay of short-lived mRNA [6]. This method has its own limitations

primarily due to the alteration of the regulation by intronic sequences and 3′UTR region (especially mRNA stability and alternative splicing). It could however still be a strategy to consider when HDR is too limiting and when the precise recapitulation of expression levels is not crucial. In the same line, the incorporation in the donor vector of enhancers (as the hsp70 promoter mentioned above) or noncoding elements of the zebrafish genome as the ubc intron that have been shown to boost expression of transgenes [43] could be considered. After NHEJ-based insertion of the edited cDNA within the first 400 bp of the transcriptional starting site of the locus of interest, the expression and localization of the protein of interest could be followed *in vivo* thanks to the in-frame fusion of a fluorescent protein or in fixed samples thanks to the use of small HA tags which would be of low interference with protein folding and localization (Fig. 3).

5. Perspectives

Although programmable nucleases and the simplicity of the CRISPR/Cas9 system have greatly advanced the field of genome editing in zebrafish, HDR-based knock-in efficiencies are still low and highly variable. This highlights the need to better understand the molecular mechanisms implicated to optimise the procedure. Both sgRNA efficiency and insertion events are still highly dependent on the locus and the developmental timing for the availability of the targeted site and the choice of the repair mechanism. However, we are at the beginning of a new era and the field is flourishing with good ideas and tricks that will end up in performing strategies for in frame insertion in zebrafish. One of them is using NHEJ-based KI to express the edited allele of choice in the endogenous regulatory genetic environment.

Recently, the first successful attempts of tissue specific or conditional CRISPR/Cas9 based knockouts have been reported in zebrafish [44–46]. The major limitation of the technique so far is the difficulty to have a definite readout of the mutated cells. Achieving highly efficient NHEJ-based KI with a fluorescent reporter (hsp70: FP-pA vector for example) could be a way to ascertain the creation of a tissue-specific null allele. Highly efficient NHEJ-based tissue-specific knock-in could also be of interest to generate a cell-type-specific reporter or to express a desired mutated allele in a tissue-specific pattern. Such techniques remain to be established but promise to constitute a versatile toolbox for CRISPR/Cas9-based KI in zebrafish.

Acknowledgements

We are grateful to J-P Concordet, Muséum National d'Histoire Naturelle Paris, and H. Besir, European Molecular Biology Laboratory Heidelberg, for providing us with Cas9 protein samples. The Del Bene team "Neural Circuit Development" is part of the Laboratoire d'Excellence (LABEX) entitled DEEP (ANR-11-LABX-0044), and of the École des Neurosciences de Paris Ile-de-France network. This work has been supported by an ATIP/AVENIR program starting grant, ERC-StG #311159-Zebratectum (F.D.B.), CNRS, INSERM, Institut Curie. CR is supported by a EU H2020 Marie Skłodowska-Curie Action fellowship (H2020-MSCA-IF-2014 #661527). SA is supported by the Fondation Berthe Fouassier (Fondation de France #2015-00060246).

References

[1] M. Jasin, J.E. Haber, The democratization of gene editing: insights from site-specific cleavage and double-strand break repair, DNA Repair (Amst) 44 (2016) 6–16.
[2] J.D. Sander, J.K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes, Nat. Biotechnol. 32 (4) (2014) 347–355.
[3] K. Hoshijima, M.J. Jurynec, D.J. Grunwald, Precise editing of the zebrafish genome made simple and efficient, Dev. Cell 36 (6) (2016) 654–667.
[4] K. Hoshijima, M.J. Jurynec, D.J. Grunwald, Precise genome editing by homologous recombination, Methods Cell Biol. 135 (2016) 121–147.
[5] U. Irion, J. Krauss, C. Nusslein-Volhard, Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system, Development 141 (24) (2014) 4827–4830.
[6] Y. Hisano et al., Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish, Sci. Rep. 5 (2015) 8841.
[7] A. Hruscha et al., Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish, Development 140 (24) (2013) 4982–4987.
[8] J.A. Gagnon et al., Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs, PLoS One 9 (5) (2014) e98186.
[9] W.Y. Hwang et al., Targeted mutagenesis in zebrafish using CRISPR RNA-guided nucleases, Methods Mol. Biol. 1311 (2015) 317–334.
[10] W.Y. Hwang et al., Efficient genome editing in zebrafish using a CRISPR-Cas system, Nat. Biotechnol. 31 (3) (2013) 227–229.
[11] M.A. Moreno-Mateos et al., CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo, Nat. Methods 12 (10) (2015) 982–988.
[12] M. Haeussler et al., Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR, Genome Biol. 17 (1) (2016) 148.
[13] Grainger S, et al. CRISPR Guide RNA Validation In Vitro. Zebrafish [Epub ahead of print] 2016, http://dx.doi.org/10.1089/zeb.2016.1358.
[14] A. Burger et al., Maximizing mutagenesis with solubilized CRISPR-Cas9 ribonucleoprotein complexes, Development 143 (11) (2016) 2025–2037.
[15] T.O. Auer et al., Deletion of a kinesin I motor unmasks a mechanism of homeostatic branching control by neurotrophin-3, Elife 4 (2015).
[16] T.O. Auer, K. Duroure, Cian A. De, J.P. Concordet, F. Del Bene, Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair, Genome Res. 24 (1) (2014) 142–153.
[17] Y. Kimura, Y. Hisano, A. Kawahara, S. Higashijima, Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering, Sci. Rep. 4 (2014) 6545.
[18] S. Nakade et al., Microhomology-mediated end-joining-dependent integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9, Nat. Commun. 5 (2014) 5560.
[19] A. Sfeir, L.S. Symington, Microhomology-mediated end joining: a back-up survival mechanism or dedicated pathway?, Trends Biochem Sci. 40 (11) (2015) 701–714.
[20] P. Sung, H. Klein, Mechanism of homologous recombination: mediators and helicases take on regulatory functions, Nat. Rev. Mol. Cell Biol. 7 (10) (2006) 739–750.
[21] W.Y. Hwang et al., Heritable and precise zebrafish genome editing using a CRISPR-Cas system, PLoS One 8 (7) (2013) e68708.
[22] G.A. Armstrong et al., Homology directed knockin of point mutations in the zebrafish tardbp and fus genes in ALS using the Crispr/Cas9 system, PLoS One 11 (3) (2016) e0150188.
[23] Y. Zhang, H. Huang, B. Zhang, S. Lin, TALEN- and CRISPR-enhanced DNA homologous recombination for gene editing in zebrafish, Methods Cell Biol. 135 (2016) 107–120.
[24] J. Shin, J. Chen, L. Solnica-Krezel, Efficient homologous recombination-mediated genome engineering in zebrafish using TALE nucleases, Development 141 (19) (2014) 3807–3818.
[25] Y. Zu et al., TALEN-mediated precise genome modification by homologous recombination in zebrafish, Nat. Methods 10 (4) (2013) 329–331.
[26] V.T. Chu et al., Increasing the efficiency of homology-directed repair for CRISPR-Cas9-induced precise gene editing in mammalian cells, Nat. Biotechnol. 33 (5) (2015) 543–548.
[27] T. Maruyama et al., Increasing the efficiency of precise genome editing with CRISPR-Cas9 by inhibition of nonhomologous end joining, Nat. Biotechnol. 33 (5) (2015) 538–542.
[28] M.D. He et al., Efficient ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in zebrafish embryos, Mutat. Res. 780 (2015) 86–96.
[29] S. Ota et al., Functional visualization and disruption of targeted genes using CRISPR/Cas9-mediated eGFP reporter integration in zebrafish, Sci. Rep. 6 (2016) 34991.
[30] K. Hiom, Coping with DNA double strand breaks, DNA Repair (Amst) 9 (12) (2010) 1256–1263.
[31] D. Branzei, M. Foiani, Regulation of DNA repair throughout the cell cycle, Nat. Rev. Mol. Cell Biol. 9 (4) (2008) 297–308.
[32] S. Panier, S.J. Boulton, Double-strand break repair: 53BP1 comes into focus, Nat. Rev. Mol. Cell Biol. 15 (1) (2014) 7–18.
[33] H. Wang et al., One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering, Cell 153 (4) (2013) 910–918.
[34] Y. Wu et al., Correction of a genetic disease in mouse via use of CRISPR-Cas9, Cell Stem Cell 13 (6) (2013) 659–662.
[35] H. Yang et al., One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering, Cell 154 (6) (2013) 1370–1379.
[36] S.F. Gilbert, Early Development in Fish, 2000.
[37] J.C. Siefert, E.A. Clowdus, C.L. Sansam, Cell cycle control in the early embryonic development of aquatic animal species, Comp. Biochem. Physiol. C Toxicol. Pharmacol. 178 (2015) 8–15.

[38] S.B. Thyme, A.F. Schier, Polq-mediated end joining is essential for surviving DNA double-strand breaks during early zebrafish development, Cell Rep. (2016).
[39] J. Dai, X. Cui, Z. Zhu, W. Hu, Non-homologous end joining plays a key role in transgene concatemer formation in transgenic zebrafish embryos, Int. J. Biol. Sci. 6 (7) (2010) 756–768.
[40] M. Hagmann et al., Homologous recombination and DNA-end joining reactions in zygotes and early embryos of zebrafish (Danio rerio) and Drosophila melanogaster, Biol. Chem. 379 (6) (1998) 673–681.
[41] J. Li, B. Zhang, J. Bu, J. Du, Intron-based genomic editing: a highly efficient method for generating knockin zebrafish, Oncotarget 6 (20) (2015) 17891–17894.
[42] J. Li et al., Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system, Cell Res. 25 (5) (2015) 634–637.
[43] E.J. Horstick et al., Increased functional protein expression using nucleotide sequence features enriched in highly expressed genes in zebrafish, Nucl. Acids Res. 43 (7) (2015) e48.
[44] J. Ablain, E.M. Durand, S. Yang, Y. Zhou, L.I.A. Zon, CRISPR/Cas9 vector system for tissue-specific gene disruption in zebrafish, Dev. Cell 32 (6) (2015) 756–764.
[45] V. Di Donato et al., 2C-Cas9: a versatile tool for clonal analysis of gene function, Genome Res. 26 (5) (2016) 681–692.
[46] L. Yin et al., Multiplex conditional mutagenesis using transgenic expression of Cas9 and sgRNAs, Genetics 200 (2) (2015) 431–441.