Efficient CRISPR/Cas9 genome editing with low off-target effects in zebrafish

Alexander Hruscha¹, Peter Krawitz²,³,⁴, Alexandra Rechenberg¹, Verena Heinrich², Jochen Hecht²,⁴, Christian Haass¹,⁵,⁶ and Bettina Schmid¹,⁶,*

INTRODUCTION

The induction of targeted mutations in zebrafish and other model organisms has only become possible through the use of genome editing tools, which induce mutations through DNA double-strand breaks and error-prone repair by non-homologous end joining. So far, genome editing has mostly been performed with zinc-finger nucleases (ZFNs) or transcription activation-like effector nucleases (TALENs). ZFNs can work very efficiently in zebrafish, but are expensive, difficult to design and function in a context-dependent (TALENs). ZFNs can work very efficiently in zebrafish, but are expensive, difficult to design and function in a context-dependent manner, making initial in vitro testing mandatory (Carroll, 2011). TALENs are easy to design but require laborious cloning steps (Joung and Sander, 2013). The prokaryotic clustered, regularly interspaced, short palindromic repeats (CRISPR) and the CRISPR-associated system (Cas) (CRISPR/Cas system) is a very attractive alternative since it is the only genome engineering tool described so far that relies on Watson-Crick base pairing rather than the potentially less specific protein-DNA interaction (Gaj et al., 2013).

The CRISPR/Cas9 system has been optimized and successfully used to efficiently generate insertions and deletions due to incorrect repair by non-homologous end joining of induced double-strand breaks in zebrafish (Chang et al., 2013; Hwang et al., 2013a; Hwang et al., 2013b; Jao et al., 2013). We now demonstrate that the CRISPR/Cas9 system is highly efficient and specific in generating heritable mutations and knock-in (KI) alleles in zebrafish with negligible off-target effects.

RESULTS AND DISCUSSION

We designed four single-stranded gRNAs composed of the targeting sequences of foreign DNA. Upon transcription, they serve as guide RNA (gRNA) to identify foreign DNA and target it for inactivation by nuclease-mediated cleavage. The CRISPR/Cas9 system has recently been optimized and successfully used in vertebrate cells to edit genomes (Cho et al., 2013; Horvath and Barrangou, 2013; Jinek et al., 2013; Mali et al., 2013; Shen et al., 2013; Wang et al., 2013). Recently, it has been further demonstrated to efficiently generate insertions and deletions due to incorrect repair by non-homologous end joining of induced double-strand breaks in zebrafish (Chang et al., 2013; Hwang et al., 2013a; Hwang et al., 2013b; Jao et al., 2013). We now demonstrate that the CRISPR/Cas9 system is highly efficient and specific in generating heritable mutations and knock-in (KI) alleles in zebrafish with negligible off-target effects.

CRISPR/Cas9 is a very attractive alternative since it is the only genome engineering tool described so far that relies on Watson-Crick base pairing rather than the potentially less specific protein-DNA interaction (Gaj et al., 2013). The CRISPR/Cas system is used as an innate immune strategy to inactivate foreign intruding nucleic acids (Wiedenheft et al., 2012). The CRISPR locus consists of short palindromic repeats and short interspersed sequences of foreign DNA. Upon transcription, they serve as guide RNA (gRNA) to identify foreign DNA and target it for inactivation by nuclease-mediated cleavage. The CRISPR/Cas9 system has recently been optimized and successfully used in vertebrate cells to edit genomes (Cho et al., 2013; Horvath and Barrangou, 2013; Jinek et al., 2013; Mali et al., 2013; Shen et al., 2013; Wang et al., 2013). Recently, it has been further demonstrated to efficiently generate insertions and deletions due to incorrect repair by non-homologous end joining of induced double-strand breaks in zebrafish (Chang et al., 2013; Hwang et al., 2013a; Hwang et al., 2013b; Jao et al., 2013). We now demonstrate that the CRISPR/Cas9 system is highly efficient and specific in generating heritable mutations and knock-in (KI) alleles in zebrafish with negligible off-target effects.
Fig. 1. Overview of the CRISPR/Cas9 system and mutation analysis of injected P0 founder embryos. (A) Guide oligonucleotides containing a T7 RNA polymerase binding site were annealed with T7 primer (blue). gRNA and Cas9 mRNA were co-injected into zebrafish embryos. The gRNA detects the endogenous genomic target site and can base pair via its target binding site, recruiting Cas9 protein (green) to induce a double-strand break at the target site (indicated by scissors) close to the protospacer adjacent motif (PAM). (B) PCR and restriction enzyme analysis of individual P0 embryos injected with bpt1, bpt4, C9t2 or C9t3 gRNA and Cas9 as indicated. PCR products are shown undigested (PCR) or after digestion with the appropriate restriction enzyme as indicated. Lanes 1-4 show un.injected siblings, whereas lanes 5-8 represent individual P0 founder embryos injected with the indicated gRNA/Cas9 mRNA. Red arrows indicate loss of restriction sites due to genomic mutations.

Encouraged by the efficiency and specificity of the CRISPR/Cas9 system we speculated that it might be suitable to facilitate a targeted KI approach in zebrafish. We designed a donor oligonucleotide containing an HA tag flanked by a 41 nt left and a 49 nt right homology arm to introduce the HA tag of the donor oligonucleotide into the KI site at the target region we might have even missed some mutations that do not affect the restriction enzyme recognition site. To potentially identify mutations induced at off-target sites, we performed next generation sequencing (NGS) of an amplicon of the C9 target site from pooled DNA derived from ten C9 gRNA/Cas9 mRNA-injected embryos. We found that 50.5% of the mapped reads had an indel mutation at the C9 target site (Fig. 2D; supplementary material Fig. S3A). To further determine the mutagenesis frequency at potential off-target sites we analyzed four distinct genomic regions with the highest homology to the target site (off-target sites 1-4). Only 2.2-2.5% of all reads had mutations at the off-target sites in the pool of ten injected embryos (Fig. 2D; supplementary material Fig. S3B). Mutations at off-target sites with closest homology to the target site are infrequent despite the smaller gRNA targeting sequence compared with TALEN and ZFN.

For further phenotypic characterization it is essential to generate identical mutations in every cell of the animal. To determine germline transmission rates of CRISPR/Cas9, we raised C9 gRNA/Cas9 mRNA-injected and C9 gRNA/Cas9 mRNA-injected embryos to adulthood (P0) and screened their offspring (F1) for germline transmission (Fig. 2A-C). Of the seven C9 F1 clutches analyzed, we identified three clutches (43%) with a potential mutation in one allele of the target sequence by RFLP (Fig. 2A,B). For the C9 locus we screened 20 F1 clutches and identified four that carried a mutation (20%) (Fig. 2C). This germline transmission efficiency is comparable to the 45% average germline transmission that we previously observed with 12 different ZFNs (CompoZr, Sigma) (Schmid et al., 2013; van Bebber et al., 2013) (our unpublished observations). Of the seven positive F1 clutches, on average 11% (4-20%) of the 24 tested embryos carry a mutant allele. Screening of the non-digested PCR products consistently revealed indel mutations at the target site (Fig. 2B,C). C9 founder 4 propagated a 5 bp and a 445 bp deletion through the germline, reflecting the heterogeneity of induced mutations in P0 founder fish (Fig. 2A,B). By screening for the loss of a restriction site at the target region we might have even missed some mutations that do not affect the restriction enzyme recognition site. To potentially identify mutations induced at off-target sites, we performed next generation sequencing (NGS) of an amplicon of the C9 target site from pooled DNA derived from ten C9 gRNA/Cas9 mRNA-injected embryos. We found that 50.5% of the mapped reads had an indel mutation at the C9 target site (Fig. 2D; supplementary material Fig. S3A). To further determine the mutagenesis frequency at potential off-target sites we analyzed four distinct genomic regions with the highest homology to the target site (off-target sites 1-4). Only 2.2-2.5% of all reads had mutations at the off-target sites in the pool of ten injected embryos (Fig. 2D; supplementary material Fig. S3B). Mutations at off-target sites with closest homology to the target site are infrequent despite the smaller gRNA targeting sequence compared with TALEN and ZFN.

Encouraged by the efficiency and specificity of the CRISPR/Cas9 system we speculated that it might be suitable to facilitate a targeted KI approach in zebrafish. We designed a donor oligonucleotide containing an HA tag flanked by a 41 nt left and a 49 nt right homology arm to introduce the HA tag of the donor oligonucleotide by homologous recombination into the KI approach of prescreening. A systematic analysis of KI negatively prescreened fish had a KI in the F1 generation, validating our approach of prescreening. A systematic analysis of KI parameters was performed by NGS of ten C9 KI injected embryos. pathogenesis (Gijselinck et al., 2012; Mori et al., 2013). However, knockout (KO) animal models are not yet available to test this hypothesis. C9 t2 and C9 t3 gRNA/Cas9-injected P0 embryos were efficiently mutagenized (Fig. 1B; supplementary material Fig. S1D). For further phenotypic characterization it is essential to generate identical mutations in every cell of the animal. This can only be achieved after germline transmission of mutations of the injected founder fish, which is genetically mosaic (supplementary material Fig. S2). To determine germline transmission rates of CRISPR/Cas9, we raised C9 gRNA/Cas9 mRNA-injected and C9 gRNA/Cas9 mRNA-injected embryos to adulthood (P0) and screened their offspring (F1) for germline transmission (Fig. 2A-C). Of the seven C9 F1 clutches analyzed, we identified three clutches (43%) with a potential mutation in one allele of the target sequence by RFLP (Fig. 2A,B). For the C9 locus we screened 20 F1 clutches and identified four that carried a mutation (20%) (Fig. 2C). This germline transmission efficiency is comparable to the 45% average germline transmission that we previously observed with 12 different ZFNs (CompoZr, Sigma) (Schmid et al., 2013; van Bebber et al., 2013) (our unpublished observations). Of the seven positive F1 clutches, on average 11% (4-20%) of the 24 tested embryos carry a mutant allele. Screening of the non-digested PCR products consistently revealed indel mutations at the target site (Fig. 2B,C). C9 founder 4 propagated a 5 bp and a 445 bp deletion through the germline, reflecting the heterogeneity of induced mutations in P0 founder fish (Fig. 2A,B). By screening for the loss of a restriction site at the target region we might have even missed some mutations that do not affect the restriction enzyme recognition site. To potentially identify mutations induced at off-target sites, we performed next generation sequencing (NGS) of an amplicon of the C9 target site from pooled DNA derived from ten C9 gRNA/Cas9 mRNA-injected embryos. We found that 50.5% of the mapped reads had an indel mutation at the C9 target site (Fig. 2D; supplementary material Fig. S3A). To further determine the mutagenesis frequency at potential off-target sites we analyzed four distinct genomic regions with the highest homology to the target site (off-target sites 1-4). Only 2.2-2.5% of all reads had mutations at the off-target sites in the pool of ten injected embryos (Fig. 2D; supplementary material Fig. S3B). Mutations at off-target sites with closest homology to the target site are infrequent despite the smaller gRNA targeting sequence compared with TALEN and ZFN.
A high rate of mutagenesis (50-76%) of the injected embryos was observed. Additionally, the embryos showed on average a KI rate of 3.5%, of which about half (1.7%) showed correct KI without additional mutations (Fig. 3C,D; supplementary material Fig. S5).

CRISPR/Cas9 is therefore not only efficient to generate mutations in zebrafish but also suitable for KI strategies. However, the inserted HA tag at the C9t3 locus was not detectable by western blot analysis or immunohistochemistry, most likely owing to low expression of the C13H9orf72 protein. We therefore used a similar strategy at the bpt1 site, which is actively transcribed and translated (Schmid et al., 2013). We co-injected the bpt1 gRNA, Cas9 mRNA and a donor oligonucleotide with a 33 nt left homology arm, two HA tags to improve detection sensitivity, and a 33 nt right homology arm (hereafter referred to as bpt1 KI injected) (Fig. 4A). Fourteen of the 20 (70%) injected embryos analyzed were positive by PCR analysis for the HA tag, indicating successful integration (Fig. 4B).

To determine the KI efficiency at the bpt1 locus, we subjected ten bpt1 KI injected embryos to NGS analysis. The efficiency to induce mutations varied between embryos (61-98%) and was on average 86.0% (Fig. 4C; supplementary material Fig. S6). The average HA tag integration rate was 15.6%, but only 3.5% had a correct KI without any additional mutations (Fig. 4F). The high KI efficiency at the bpt1 locus is most likely due to the highly efficient bpt1 gRNA (reaching up to 98% mutagenesis efficiency), increasing the chances of an early double-strand break and donor DNA integration. Consistent with the high mutagenesis rates in the injected founder embryos we detected reduced levels of Tardbp protein by western blot upon bpt1 KI injection (Fig. 4D). Using a sensitive anti-HA tag antibody we could further detect the Tardbp-HA fusion protein at the expected molecular weight of 44 kDa by western blot in injected founder embryos (Fig. 4D). Importantly, the nuclear localization of Tardbp (Shankaran et al., 2008) was recapitulated by immunohistochemistry with the HA-specific antibody (Fig. 4E).

In conclusion, we have adapted the protocol to efficiently synthesize gRNA without any cloning and show efficient mutagenesis, KI and germline transmission, comparable to TALEN and ZFN approaches (Bedell et al., 2012; Blackburn et al., 2013; Chang et al., 2013; Hwang et al., 2013b). Most importantly, we provide evidence of limited off-target effects and efficient KI, especially upon preselection of founder fish. In contrast to ZFN or TALEN, the design and generation of CRISPR/Cas9 RNAs is much cheaper and less time consuming, making it an excellent genome editing tool, and not only in zebrafish. This system further opens new avenues for the analysis of gene function in vivo and for the generation of new disease models in vertebrates.
MATERIALS AND METHODS

Zebrafish

Zebrafish (Danio rerio) embryos were kept at 28.5°C and were staged according to Kimmel et al. (Kimmel et al., 1995). The wild-type line AB was used. All experiments were performed in accordance with animal protection standards of the Ludwig-Maximilians University Munich and were approved by the government of Upper Bavaria (Regierung von Oberbayern, Munich, Germany).

Guide oligonucleotide

The gRNA antisense oligonucleotide sequence (5’-3’) was: AAAGCACC-GACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAA-CTTGCTATTTCTAGCTCTAAAA CNNNNNNNNNNNNNNNNNNNNNNNNNNNCT-ATAGTGAAGTCGTATTA CGC, with the T7 site shown in bold, N 20 indicating the targeting sequence, and the gRNA scaffold adopted from Mali et al. (Mali et al., 2013) shown in italics (see Fig. 1A; supplementary material Fig. S1A).

The targeting sequences (N 20) used for the respective gRNA (in parenthesis) were:

- *tardbp* targeting oligo 1 (bpt1), CCCATGTTGCTT - CGGCTGCC;
- *tardbpl* targeting oligo 4 (bplt4), AACTACTCTGAAGTGC - TGCC;
- *C13H9orf72* targeting oligo 2 (C9t2), TCACTGAGCAGCAGCA - ACCC; and
- *C13H9orf72* targeting oligo 3 (C9t3), ATCATGTCTTCA - GCCTGTCC.

For direct *in vitro* transcription, the gRNA antisense oligonucleotide was annealed to the T7 primer (TAATACGACTCACTATAG; 5 minutes at 95°C, then cooled at room temperature for 5 hours) and *in vitro* transcribed using the Ambion MEGAshortscript-T7 kit.

The gRNA antisense oligonucleotide was PCR amplified for cloning with primer 1 (GCGTAATACGACTCACTATAG) and primer 2 (which adds a DraI restriction site, CGCGTTTAAAGCACCGACTCGGTGCCAC) and the PCR product gel-purified and subcloned into pCR8/GW/TOPO/TA (Invitrogen). The insert was excised using DraI and ApaI (NEB). The excised insert and PCR product were *in vitro* transcribed using the MEGAscript-T7 kit.

Cas9

Nuclear localization sequences on both ends and one C-terminal HA tag were added to wild-type Cas9 from Streptococcus pyogenes (Addgene, 39312) by PCR (supplementary material Fig. S1C) using primer 1 (ATGGCCCCAAAAGAAGAAGCGCAAGGTAGGTGGAGGTGGAGGAG- GAGATAAGAAATACGTAAAGGCTTAG) and primer 2 (TCGCTGGGACGTCGTATGGGTAACCTCCTCCTCCCTACTCTT- CGCTTCTTCTTTGGAGCACCTCCTCCTCCGTCACCTCTT- AGCTGACT). The PCR product was subcloned into a pCR8/GW/TOPO vector (Invitrogen) and further recombined into pCS2+/GW using LR Clonase II (Invitrogen). After linearization with ApaI (NEB), Cas9 was *in vitro* transcribed using the SP6 mMESSAGE mmACHINE kit (Ambion).

Microinjection and knock in

gRNAs were microinjected (Femptojet, Eppendorf) at a final concentration of 2.4 μg/μl together with 0.5 μg/μl Cas9 mRNA into fertilized one-cell stage embryos. For KI, 100 nM donor oligonucleotide was co-injected. Uninjected siblings were used as controls. Oligonucleotides were (HA tag sequence in bold): donor oligonucleotide for C9t3 (120 nt), CTTC - TGCTTTGTGAGCCTGACGGTCTATAGGTTTGTATAAGCCCCTATC- ATG TACCCATAACGAGTCCCAAGACTACGCTTCTCTCGGCTGTAATAGCTCAGACAGCTACGCTGAGTGTGGTGTTGGGGGTTAGCTCCACCGAGCTT.
PCR analysis
DNA of 1 dpf embryos was used as a template for PCR with the following primers: C13H9orf72-for (genotyping C9t2 and C9t3), GCAACTGTT-TGTCGCCACT; C13H9orf72-rev (genotyping C9t2 and C9t3), ATCA-ACCTCCTCTGGCACAC; tardbp-for (genotyping bpt1), GAAGAAAT-TTCCAACTTCTTTC; tardbp-rev (genotyping bpt1), ACTTACAAAAAAC-ACGCTAGG; tardbpl-for (genotyping bplt4), GCCTAAGCACAATAATAGT; tardbpl-rev (genotyping bplt4), CTGACTTCTACTCTGC; C9t3-HA-tag-for (genotyping HA KI C9t3), CCCTATCATGTACCTTGCC; 2x-HA-tag-for (genotyping HA KI bpt1), CTTACCATAACCCAG-
GTCC; and tardbp-rev (genotyping HA KI in bpt1), ACTTACAAAAAAC-ACGCTAGG.

For RFLP analysis, half of the PCR product was digested with the following enzymes (NEB): tardbp (bpt1), CspCI; tardbpl (bplt4), Hpy188I; C13H9orf72 (C9t2), Ddel; C13H9orf72 (C9t3), BbsI.

Next generation sequencing
Putative off-target sites of the C9t3 gRNA were identified by BLAST search of the 13 nt 5′ of the PAM motif containing an NGG motif (previously described to be crucial for binding) (Cong et al., 2013; Jinek et al., 2012).
using Ensembl Genome assembly Zv9 (GCA_000002035.2). The DNA of ten individual embryos injected with Cas9 and C9t3 gRNA was pooled and analyzed for C9t3 gRNA target effects and off-target effects. From this pool we PCR amplified the C9t3 locus and the four potential off-target sites (supplementary material Fig. S3). Ten individual embryos injected with C9t3 gRNA, Cas9 mRNA and the donor oligonucleotide were PCR amplified and analyzed individually for KI efficiency (supplementary material Fig. S5). Adapter ligation, sequencing and tag sorting were performed by GATC Biotech (Germany). KI analysis at the bpt1 locus was performed (supplementary material Fig. S6). The length of the expected PCR products varied between 184 bp and 417 bp. We sequenced on a MiSeq platform (Illumina) in paired-end mode of appropriate length (140 bp or 250 bp). Templates of the C9t3 gRNA target sites and potential off-target sites were used as a reference sequence for determination of the off-target lesions (supplementary material Fig. S3). Templates of bpt1 and C9t3 with and without inserted HA tags and the C9t3-associated off-target sites were used as reference sequences for determination of KI efficiency (supplementary material Figs S5, S6). Reference-guided sequence assemblies were performed with Novoalign (v3, Novocraft). The frequency of insertions and deletions around the best-matching reference sequence was derived from the CIGAR code of the sequence alignment (Java and R scripts for statistical evaluation are available upon request). Primers used for barcode labeling are also available upon request.

Western blot analysis and immunohistochemistry

Western blotting (WB) and immunohistochemistry (IHC) were performed as previously described (Schmid et al., 2013). Primary antibodies used: Tardbp 4A12-111 (Tardbp epitope: TSTSGTSSSRRDAQQTY), WB 1:1; α-tubulin (Sigma, T6199), WB 1:8000; HA.11 (Covance, MMS-101IP), WB 1:2000, IHC 1:200.

Acknowledgements
We thank D. Edlbauer, P.-H. Kuhn, R. Kühn, B. Scholzenberger, S. Rothhämel and F. van Bebber for critically reading the manuscript and R. Rojas Rojas for zebrafish care.

Competing interests
The authors declare no competing financial interests.

Author contributions

Funding
The research leading to these results has received funding from the European Research Council under the European Union Seventh Framework Programme [FP7/2007-2013]ERC grant agreement no 321366-Amyloid to C.H. and the Center of Excellence in Neurodegeneration (COEN) to C.H. and B.S.

Supplementary material

Supplementary material available online at http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.099085/-/DC1

References


