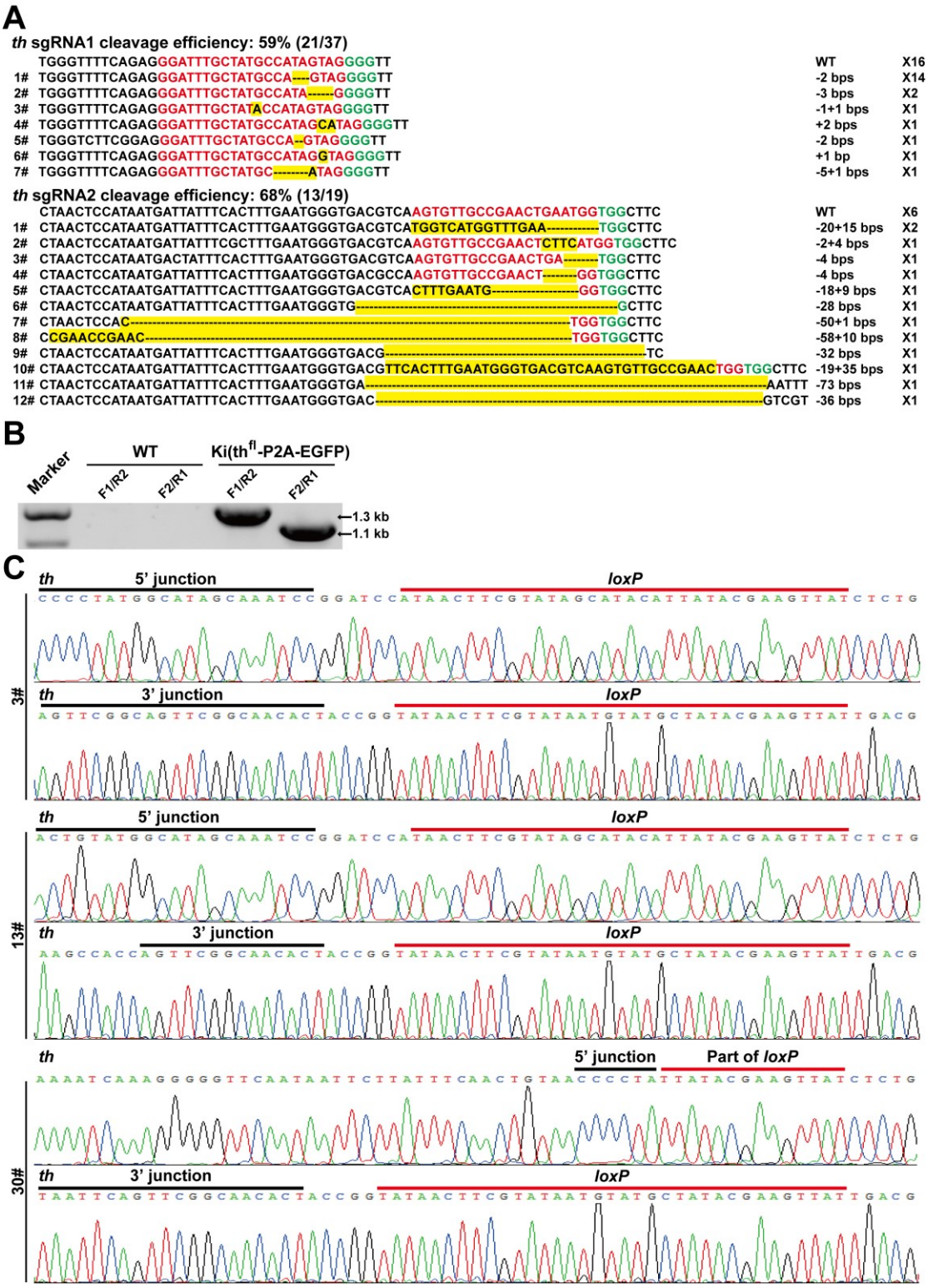
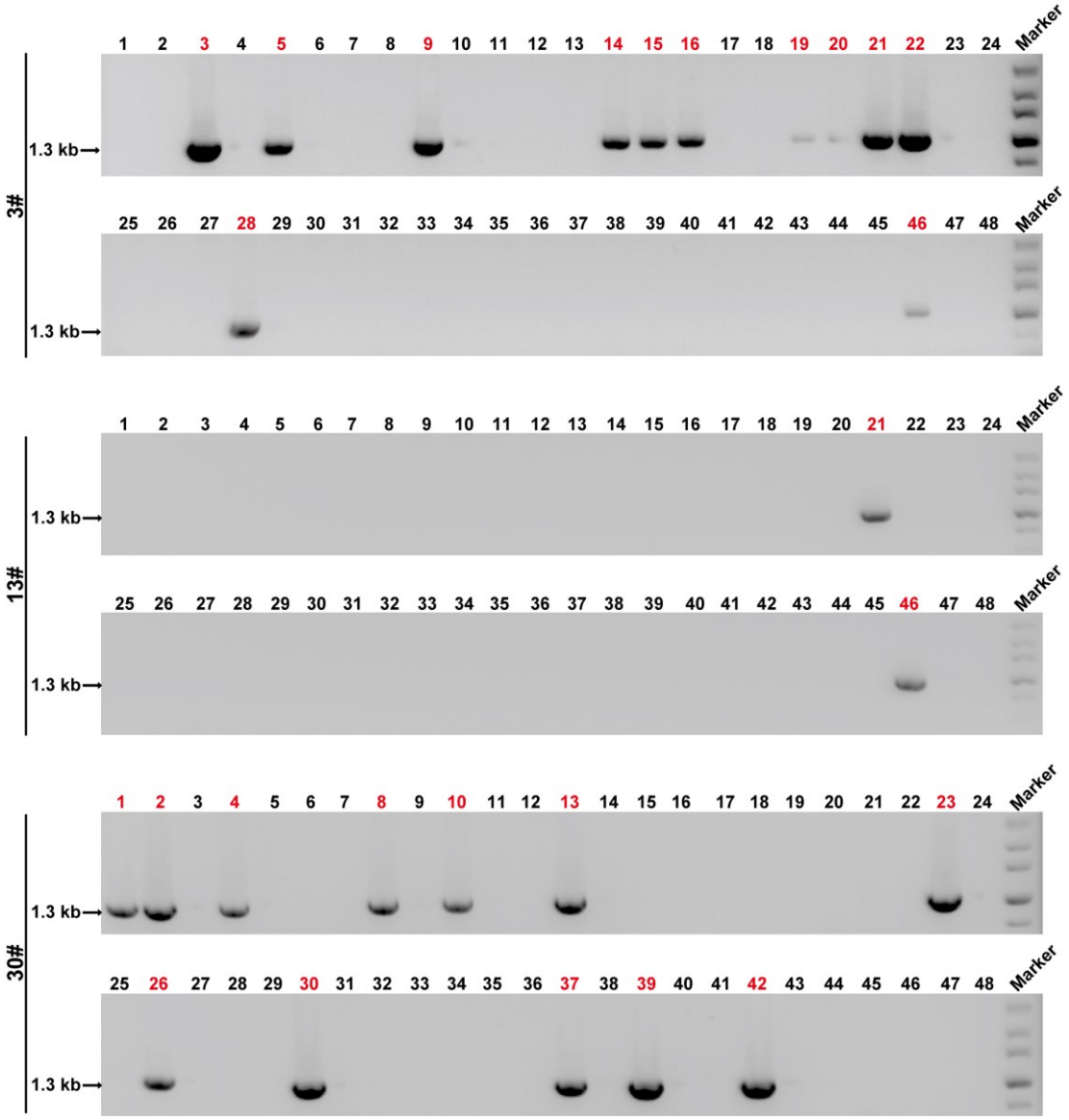
Supplementary materials

Supplementary Figures and Tables



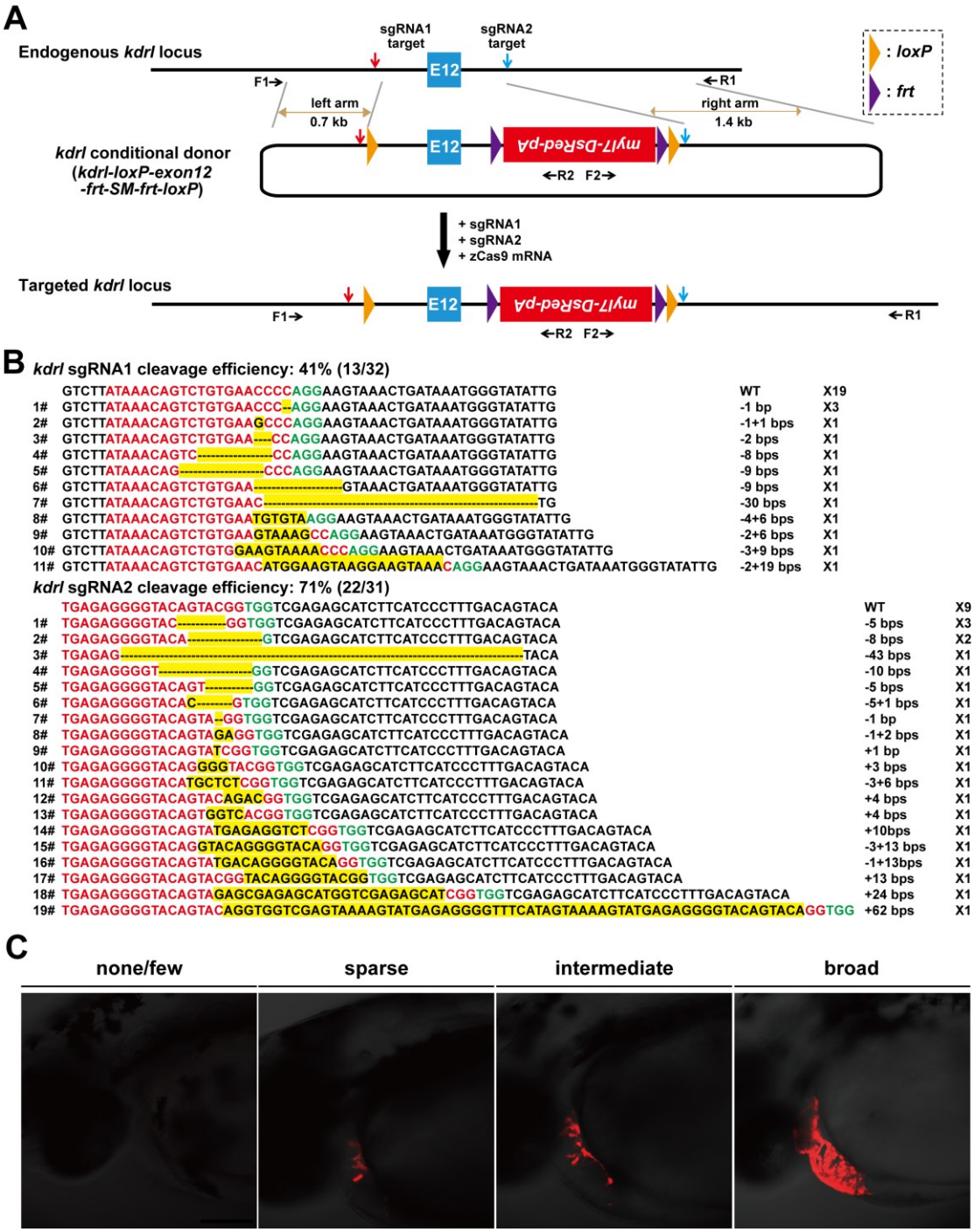
Supplementary Figure S1. Cleavage efficiency of *th* sgRNAs and verification of *th* conditional alleles.

(**A**) Cleavage efficiency of the sgRNA1 and sgRNA2 targeting the zebrafish *th* locus. The indel mutations are highlighted in yellow, and the PAM and sgRNA target sequences are shown in green and red, respectively. The numbers of insertion (+) and/or deletion (-) of base pairs (bps) are shown at the right. (**B**) PCR analysis of the 5' and 3'junctions of the targeted *th* locus. The primers of F1, R1, F2, and R2 are shown in Figure 1A. A 1.3-kb band was amplified by using the 5' junction primers F1 and R2, and a 1.1-kb band was amplified by using the 3' junction primers F2 and R1. (**C**) Sequencing reports of the 5' and 3' junction sites of the progenies of three Ki(th^{fl}-P2A-EGFP) founders.



Supplementary Figure S2. Germline transmission analysis of the three Ki(th^{fl}-P2A-EGFP) founders by PCR.

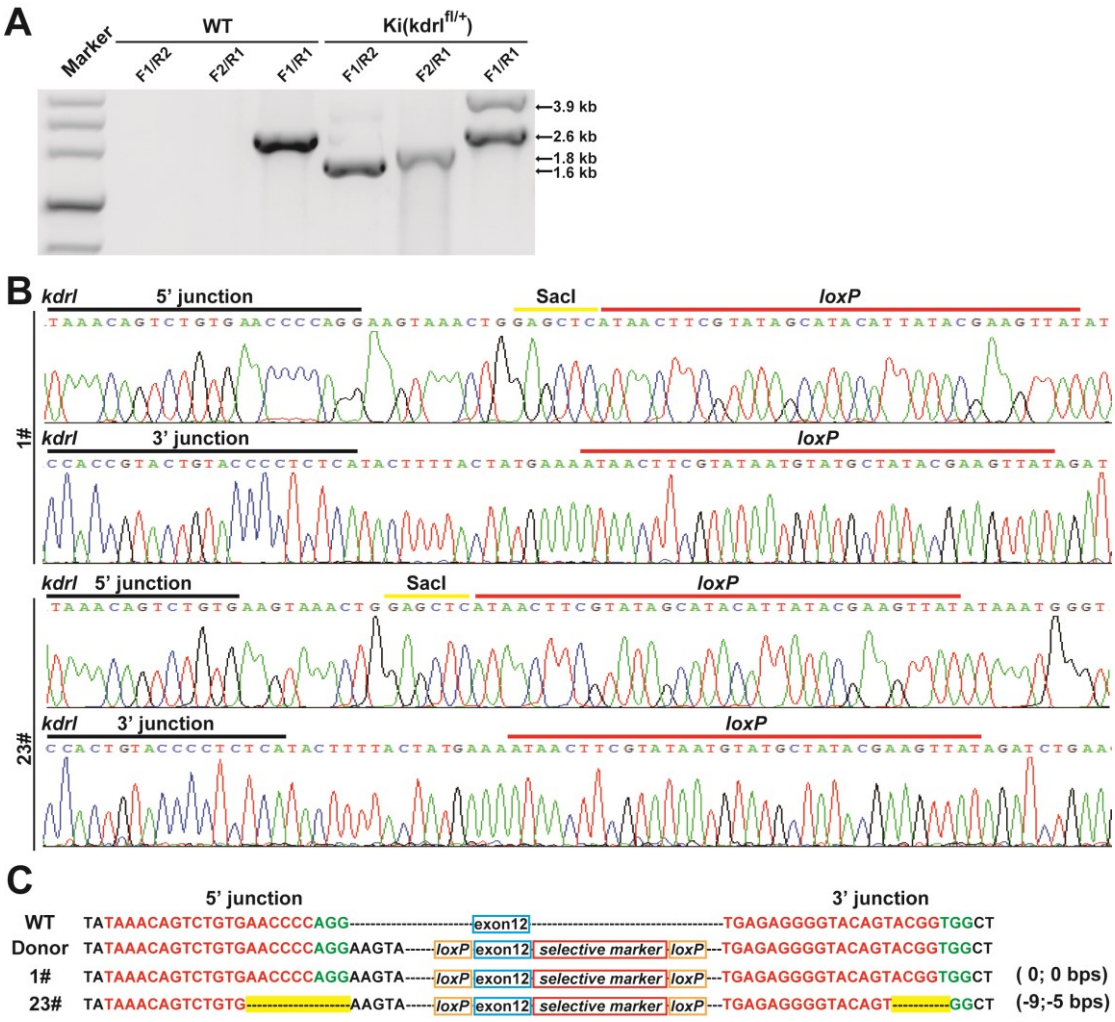
The 5' junction primers F1 and R2 were used to amplify DNA fragments from 48 individual F1 of each founder. The number in red indicates that the corresponding progeny carries a conditional allele.



Supplementary Figure S3. Design of the *kdrl* conditional allele through NHEJ-mediated DNA replacement, cleavage efficiency of *kdrl* sgRNAs and transient expression of the selective marker.

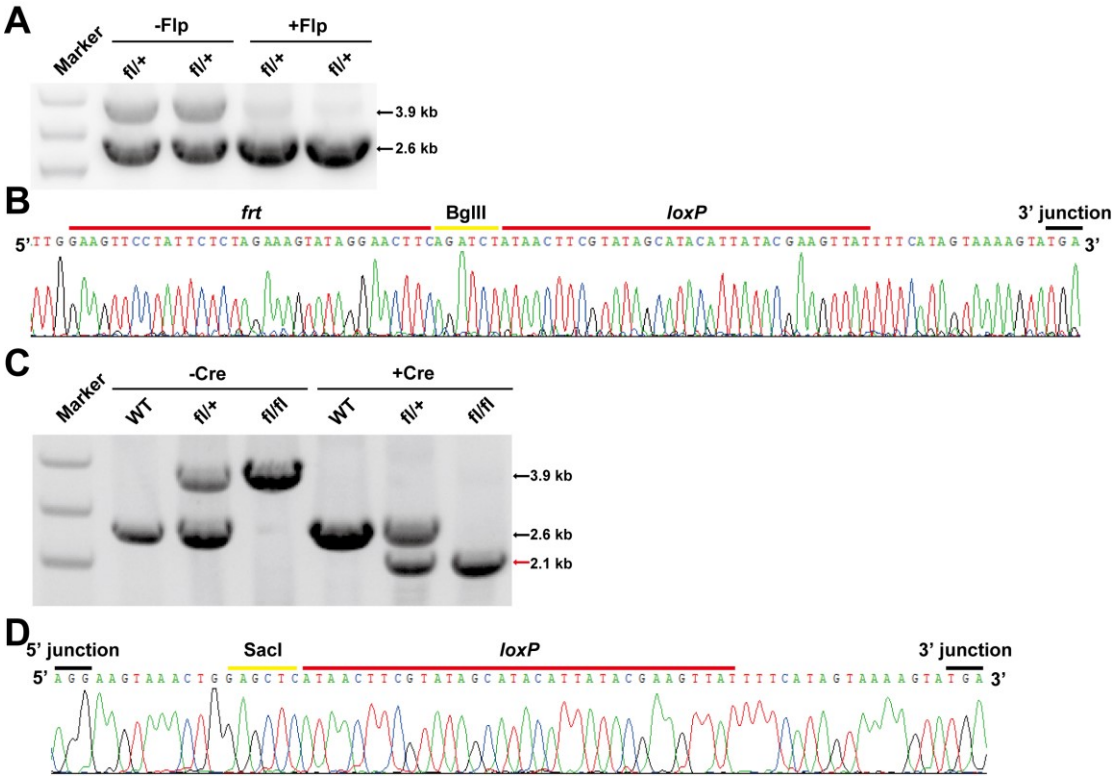
(**A**) Schematic of the strategy for NHEJ-mediated DNA replacement at the *kinase insert domain receptor like* (*kdrl*) locus of Tg(kdrl:EGFP) zebrafish, in which the vascular endothelial cells express EGFP. Zebrafish *kdrl* gene is specifically expressed in vascular endothelial cells, and its mutation causes severe defects in vascular formation at early developmental stages (Habeck et al., 2002). The sgRNA1 and sgRNA2 targets are indicated by red or blue arrows, respectively. In the *kdrl* conditional donor

(*kdrl-loxP-exon12-frt-SM-frt-loxP*), two loxP sites and homology arms (brown lines with double arrows) were added. The length of the left and right arms is 686 bps and 1297 bps, respectively. To reduce the labor of intensive screening of conditional alleles, a selective marker (SM, 1.2 kb) can expresses DsRed under the control of the *myl7* promoter, and its transcriptional direction is contrary to *kdrl* transcriptional direction. The SM is also flanked by *frt* recombination sites, thus it can be excised by Flp recombinase. The zebrafish *kdrl* has 30 exons, and E12 (12^{th} exon) contains 109 bps. The endogenous E12 is replaced by the *kdrl-loxP-exon12-frt-SM-frt-loxP* cassette at the *kdrl* locus after co-injecting the donor with the sgRNA1, sgRNA2 and zCas9 mRNA. The primers (F1, R1, F2, R2) for verification are indicated by short black arrows. (**B**) Cleavage efficiency of the sgRNA1 and sgRNA2 targeting the zebrafish *kdrl* locus. The indel mutations are highlighted in yellow, and the PAM and sgRNA target sequences are shown in green and red, respectively. The numbers of insertion (+) and/or deletion (-) of bps are shown at the right. (**C**) Differential expression of SM in the heart of F0 larvae. The embryos with broad expression of DsRed in the heart were raised to adulthood for the founder screening. Scale bars: 100 μm.



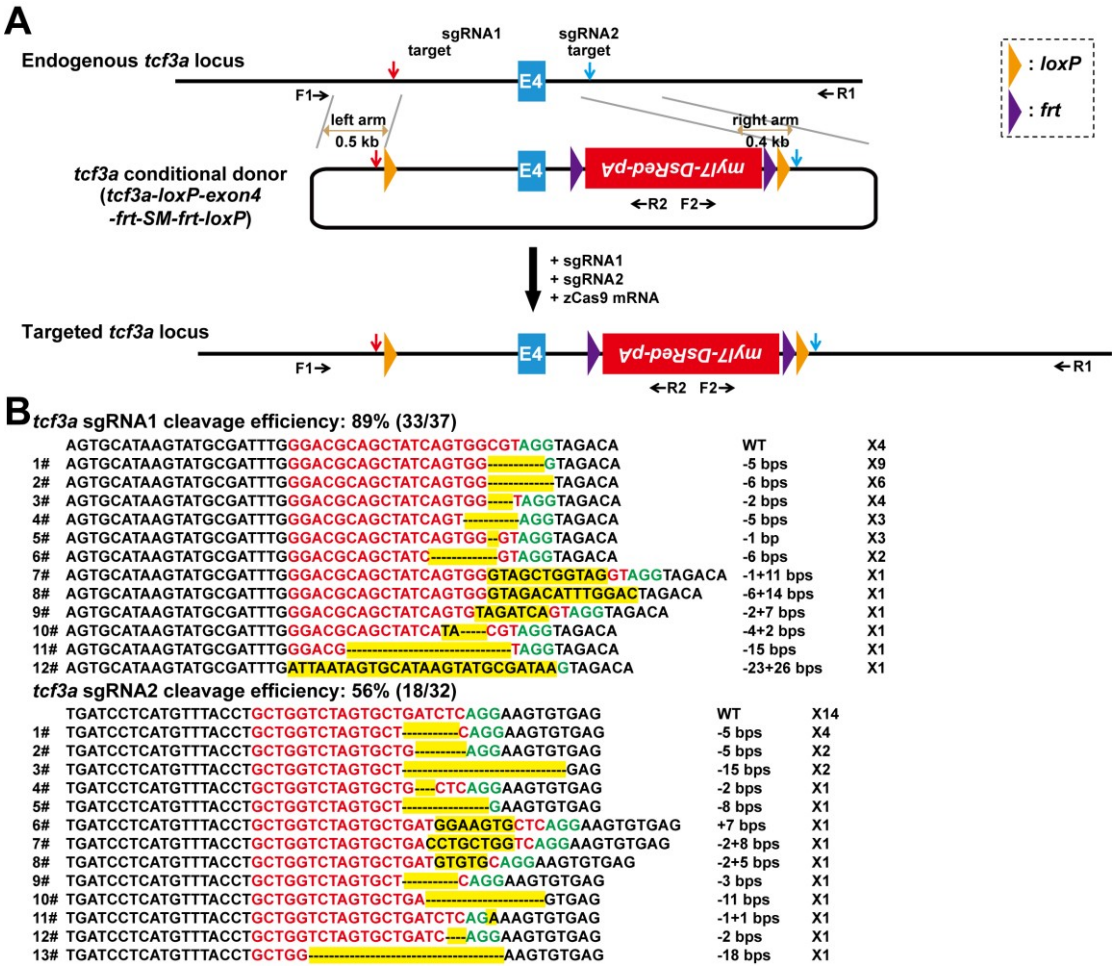
Supplementary Figure S4. Verification of *kdrl* conditional alleles.

(**A**) PCR analysis of the 5' and 3'junctions of the targeted *kdrl* locus. The primers of F1, R1, F2, and R2 are shown in Figure S3A. A 1.6-kb band was amplified by using the 5' junction primers F1 and R2, a 1.8-kb band was amplified by using the 3' junction primers F2 and R1, and a 2.6-kb band and a 3.9-kb band were amplified by using the primers F1 and R1. (**B**) Sequencing reports of the 5' and 3' junction sites of the progenies of two Ki(kdrl^{fl}) founders (1# and 23#), which are identified from 25 F0 screened (Table S1). (**C**) 5′ and 3′ junction sequences of F1 progenies of the two Ki(kdrl^{fl}) founders in comparison with the sequences of WT zebrafish and the donor vector (Table S2). The indel mutations are highlighted in yellow, and the PAM and sgRNA target sequences are shown in green and red, respectively. The numbers of insertion (+) and/or deletion (-) of bps at the 5' and 3' junctions are shown in the brackets at the right.



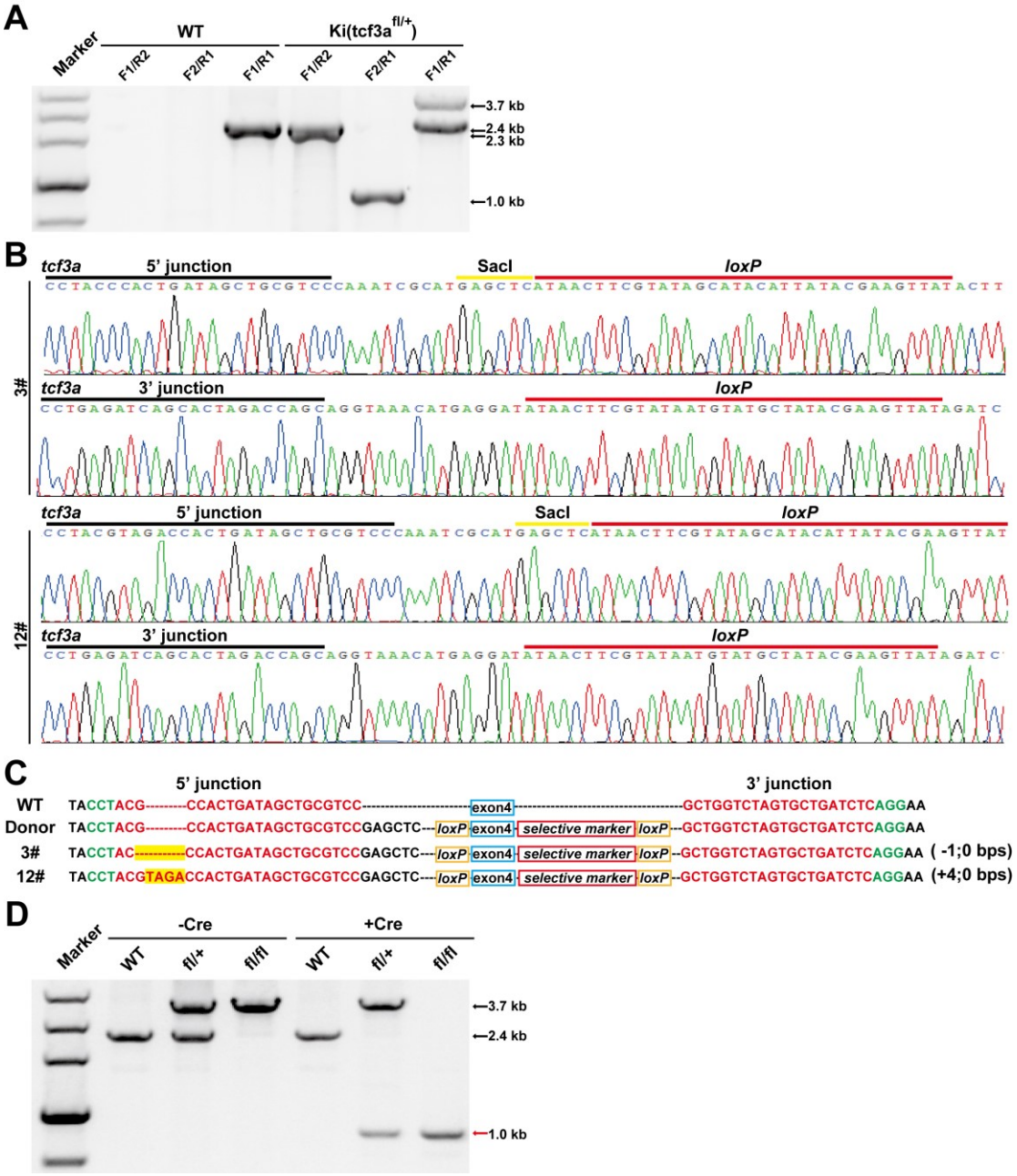
Supplementary Figure S5. Confirmation of recombinase-induced excision of the conditional Ki(kdrl^{fl}) line by PCR and sequencing.

(**A**) PCR analysis for the SM deletion in 3-dpf heterozygote Ki(kdrl^{fl}) larvae with Flp mRNA injection at one-cell stage. Flp mRNA was injected into one-cell-stage embryos generated via Ki(kdrl$^{fl/+}$) intercross for the SM deletion, and we found that 43 out of 47 injected larvae at 3 dpf did not express DsRed in the heart, indicating the occurrence of Flp-induced SM excision. (**B**) Sequencing report showing SM deletion. (**C**) PCR analysis of heterozygote Ki(kdrl^{fl}) intercross larvae without ("-Cre") or with ("+Cre") Cre mRNA injection at one-cell stage. The red arrow indicates the additional product at ~ 2.1 kb, which were in principle excised from the original ~ 3.9 kb fragment by Cre. (**D**) Sequencing report of the additional 2.1-kb product, confirming the correct deletion by Cre. Cre-induced E12 (109 bps) excision of *kdrl* caused the frame-shift mutation. In the embryos of [Tg(kdrl:EGFP);Ki(kdrl$^{fl/+}$)] intercross, one fourth (17/70) of Cre-injected larvae showed vascular defects in both the trunk and brain, consistent with the vascular phenotype of the *kdrl* mutant homozygote (Figure 1F and Ref. Habeck et al., 2002). As a control, all (68/68) intercross larvae without Cre injection displayed normal vascular patterns (Figure 1G).



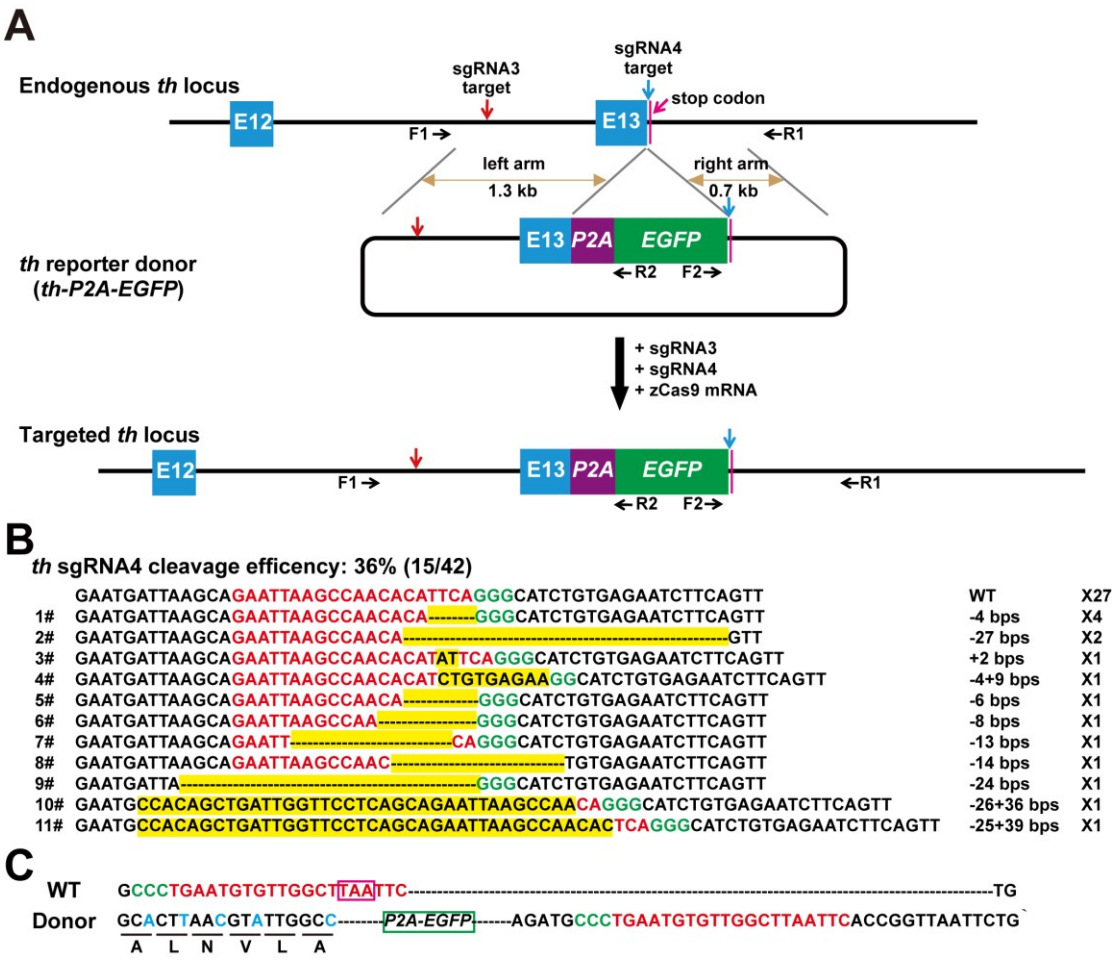
Supplementary Figure S6. Design of the *tcf3a* conditional allele through NHEJ-mediated DNA replacement approach and cleavage efficiency of *tcf3a* sgRNAs.

(**A**) Schematic of the strategy for NHEJ-mediated DNA replacement approach at the *tcf3a* locus. Zebrafish *tcf3a* gene (also named *E12*) plays an important role in muscle development at early developmental stages (Chong et al., 2007). The design of *tcf3a* conditional allele is similar to that of *kdrl* conditional allele (Figure S3A). The sgRNA1 and sgRNA2 targets are indicated by red or blue arrows, respectively. In the *tcf3a* conditional donor (*tcf3a-loxP-exon12-frt-SM-frt-loxP*), the length of the left and right arms is 530 bps and 424 bps, respectively. The length of E4 is 68 bps. (**B**) Cleavage efficiency of the sgRNA1 and sgRNA2 targeting the zebrafish *tcf3a* locus.



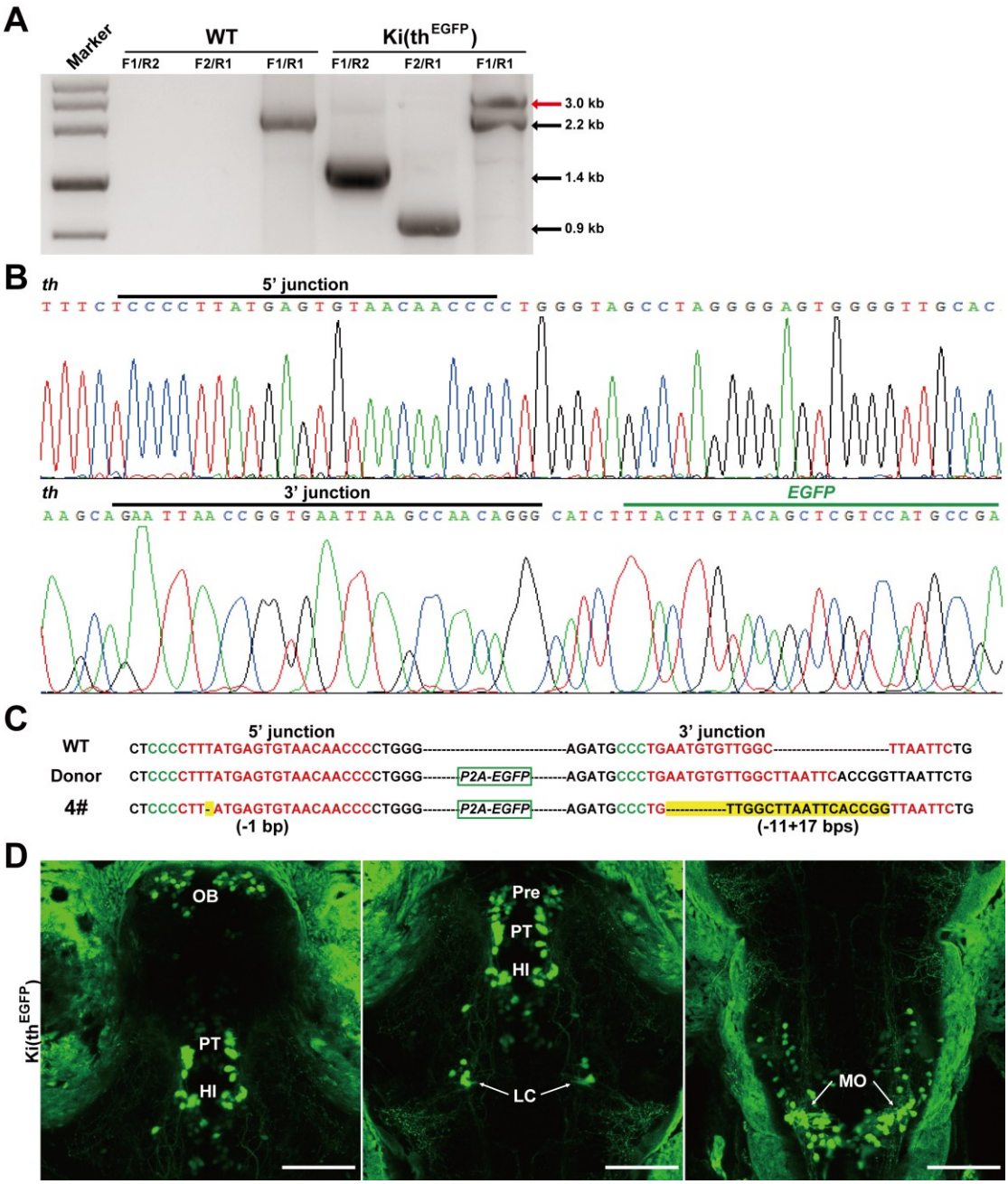
Supplementary Figure S7. Verification of *tcf3a* conditional alleles and Cre-induced excision.

(**A**) PCR analysis of the 5' and 3'junctions of the targeted *tcf3a* locus. (**B**) Sequencing reports of the 5' and 3' junction sites of the progenies of two Ki(tcf3a^{fl}) founders (3# and 12#), which are identified from 15 F0 screened (Table S1). (**C**) 5′ and 3′ junction sequences of F1 progenies of the two Ki(tcf3a^{fl}) founders in comparison with the sequences of WT zebrafish and the donor vector (Table S2). (**D**) PCR analysis of heterozygote Ki(tcf3a^{fl}) intercross larvae without ("-Cre") or with ("+Cre") Cre mRNA injection at one-cell stage. The red arrow indicates the additional product at ~ 1.0 kb, which were in principle excised from the original ~ 3.7 kb fragment by Cre. The phenotype of larvae carrying Cre-induced E4 (68 bps) excision of *tcf3a* was comparable to *tcf3a* morphants (Figure 1H, I; see also Ref. Chong et al., 2007).



Supplementary Figure S8. Design of the *th* reporter line via NHEJ-mediated DNA replacement approach and cleavage efficiency of *th* sgRNA.

(**A**) Schematic of NHEJ-mediated DNA replacement approach at the zebrafish *th* locus for generating EGFP reporter lines. The target sites of sgRNA3 and sgRNA4 are indicated by red or blue arrows, respectively. The *th* stop codon is indicated by a pink line. In the *th* reporter donor (*th-P2A-EGFP*), the homology arms are indicated by brown lines with double arrows. The length of the left and right arms is 1282 bps and 671 bps, respectively. The endogenous E13 in the *th* locus was replaced by the *th-P2A-EGFP* cassette after co-injection of the donor with the sgRNA3, sgRNA4 and zCas9 mRNA. The primers (F1, R1, F2, R2) for verification are indicated by short black arrows. (**B**) Cleavage efficiency of the sgRNA4 targeting the zebrafish *th* locus. The indel mutations are highlighted in yellow, and the PAM and sgRNA target sequences are shown in green and red, respectively. The numbers of insertion (+) and/or deletion (-) of base pairs are shown at the right. (**C**) sgRNA4 target sequences in the *th* locus and the reporter donor. As the sgRNA4 target partially overlaps with the 3' side of the *th* coding sequence (CDS), we introduced silent mutations at the 3' side of the *th* CDS in the *th-P2A-EGFP* donor vector. Therefore, the 3' CDS of the *th* in the donor cannot be excised by CRISPR/Cas9. Silent mutations are highlighted in blue. The *th* stop codon and *P2A-EGFP* are indicated by a pink or green box, respectively. Please note, the sgRNA3 with a cleavage efficiency of 83% was the same as that used in Ref. J. Li et al., 2015.

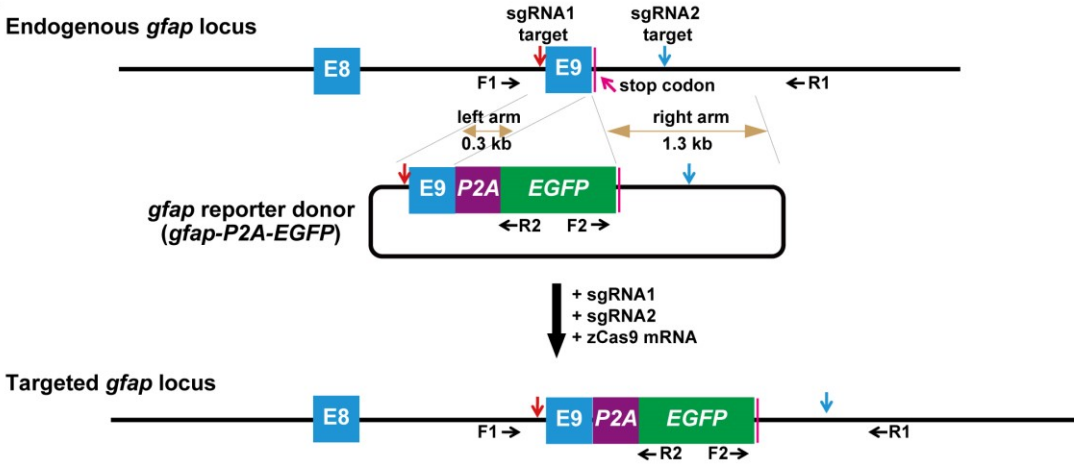


Supplementary Figure S9. Confirmation of the Ki(th^{EGFP}) reporter line.

(**A**) PCR verification of DNA replacement at the *th* locus by using the primers (F1, R1, F2 and R2) shown in Figure S8A. For the endogenous *th* locus, a 2.2-kb band was amplified by using F1 and R1. For a targeted *th* locus, 1.4-kb band was amplified by using the 5' junction primers F1 and R2, a 0.9-kb band was amplified by using the 3' junction primers F2 and R1, and a 3.0-kb band (red arrow) was amplified by using F1 and R1. We got a Ki(th^{EGFP}) founder (#4) out of 11 F0 screened (Table S1), and its F1 progenies were used for the verification. (**B**) Sequencing reports of the 5' and 3' junction sites of the progeny of the Ki(th^{EGFP}) founder. (**C**) 5′ and 3′ junction sequences of F1 progeny of the Ki(th^{EGFP}) founder in comparison with the sequence of WT zebrafish and the donor vector (Table S2). The indel mutations are highlighted in yellow, and the PAM and sgRNA target sequences are shown in green and red,

respectively. The numbers of insertion (+) and/or deletion (-) of bps at the 5' and 3' junctions are shown in the brackets. (**D**) Representative projected *in vivo* confocal images (dorsal view) of a 3-dpf Ki(th^{EGFP}) larva at different visual fields. The EGFP expression pattern was consistent with the results of whole *in situ* hybridization of *th* (Filippi et al., 2010). HI, intermediate hypothalamus; LC, locus coeruleus; MO, medulla oblongata; OB, olfactory bulb; Pre, pretectum; PT, posterior tubercular. Scale bars, 100 μm.

A



B

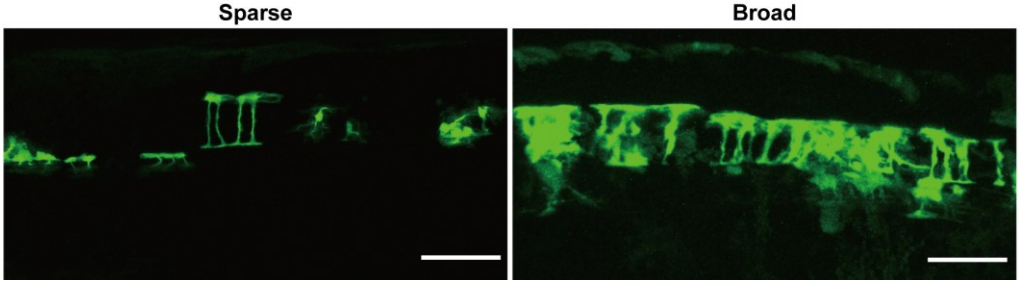
gfap sgRNA1 cleavage efficiency: 74% (26/35)

	AATTTGTTAAGTTTACTTAATCGGGCGGCGCCAAT	WT	X9
1#	AATTTGTTAAGTTTACTTAAT--GGGCGGCGCCAAT	-1 bps	X3
2#	AATTTGTTAAGTTTACTT---------GGCGGCGCCAAT	-5 bp	X2
3#	AATTTGTTAAGTTTAC----------GGCGCCAAT	-10 bps	X2
4#	AATTTGTTAAGTTTACTTACT TAT T T ACCACTAGGCTATTGGCGGCGCCAAT	-4+21 bps	X1
5#	AATTTGTTAAGTATACTTTATTATTGGCGGCGCCAAT	-5+7 bps	X1
6#	AATTTGTTAAGTTTACAAATT----GGCGGCGCCAAT	-7+5 bps	X1
7#	AATTTGTTAAGTTTACTT---------GGCGGCGCCAAT	-7+2 bps	X1
8#	AATTTGTTAAGTTTACTT--------GGGCGGCGCCAAT	-4 bps	X1
9#	AATTTGTTAAGTTTACTTAG-CGGGCGGCGCCAAT	-2+1 bps	X1
10#	AATTTGTTAAGTTTACTTAAT--------CGGCGCCAAT	-4 bps	X1
11#	AATTTGTTAAGTTTACTTAAT----------T	-13 bps	X1
12#	AATTTGTTAAGTTTACTTAATTTGTAAGTTTACTTAATTTGTAAGTAATTTGGCGGCGCCAAT	-2+30 bps	X1
13#	AATTTGTTAAGTTTACTTAAT----------CGCCAAT	-7 bps	X1
14#	AATTTGTTAAGTACTTAAGTTTACTGGGCGGCGCCAAT	-10+13 bps	X1
15#	AATTTGTTAA----------GCGCCAAT	-17 bps	X1
16#	AATTTGTTAAGTTATTTTAATTTAAGTTATTTAAGTTATTTTTAGGCGCCAAT	-13+31 bps	X1
17#	AATTTGTTAAG----------CGGCGCCAAT	-14 bps	X1
18#	AATTTGTTAAGTTTACTAGGCTATTGGCGGCGCCAAT	-6+8 bps	X1
19#	AATTTGTTAAGTTTACTTATT----GGCGGCGCCAAT	-4+2 bps	X1
20#	AATTTGTTAAGTTTACTTAATCTGGCGGCGCCAAT	-1+1 bps	X1
21#	AATTT----------GCGGCGCCAAT	-19 bps	X1
22#	AATTTGTTAAGTTTATT----------GGCGGCGCCAAT	-8+2 bps	X1

gfap sgRNA2 cleavage efficiency: 74% (29/39)

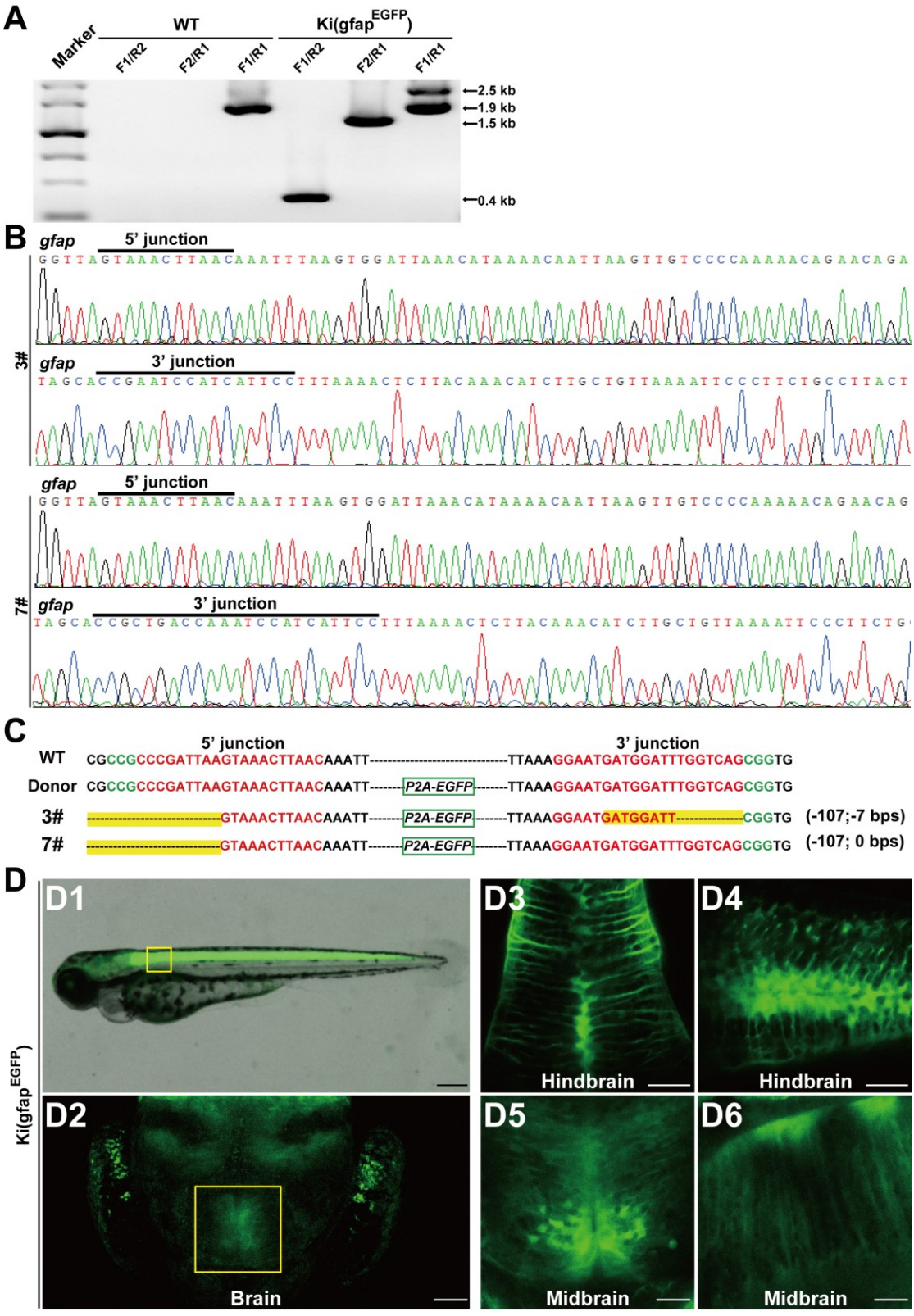
	TTAAAGGAATGATGGATTTGGTCAGCGGTGCTATTTGAACATTTTACTC	WT	X10
1#	TTAAAGGAATGATGGATTTGG----------TGCTATTTGAACATTTTACTC	-7 bps	X7
2#	TTAAAGGAATGATGGATTT---------AGCGGTGCTATTTGAACATTTTACTC	-4 bps	X2
3#	TTAAAGGAATGATGGATTTG----------CGGTGCTATTTGAACATTTTACTC	-5 bps	X2
4#	TTAAAGGAATGATGGATTTGGTT-----CGGTGCTATTTGAACATTTTACTC	-3+1 bps	X1
5#	TTAAAGGAATGATGGATTTGGTAAAGTTTGTAGCGGTGCTATTTGAACATTTTACTC	-1+9 bps	X1
6#	TTAAAGGAATGATGGATTT----------TACTC	-25 bps	X1
7#	TTAAAGGAATGATGGATTTGGTGCTATTTGCGGTGCTATTTGAACATTTTACTC	-2+7 bps	X1
8#	TTAAAGGAATGATGGATTTGGATCAGCGGTGCTATTTGAACATTTTACTC	+1 bp	X1
9#	TTAAAGGAATGATGGATTTGG--------CGGTGCTATTTGAACATTTTACTC	-4 bps	X1
10#	TTAAAGGAATGATGGATTTGGTGCTATTTAGCGGTGCTATTTGAACATTTTACTC	-1+7 bps	X1
11#	TTAAAGGAATGATGGACGGAA-CAGCGGTGCTATTTGAACATTTTACTC	-6+5 bps	X1
12#	TTAAA----------CAGCGGTGCTATTTGAACATTTTACTC	-18 bps	X1
13#	TTAAAGGAATGATGGATTTGGTGCGTTCAAACAGCGGTGCTATTTGAACATTTTACTC	+9 bps	X1
14#	TTAAAGGAATGATGGATT---------CAGCGGTGCTATTTGAACATTTTACTC	-4 bps	X1
15#	TTAAAGGAATGATGGA----------CAGCGGTGCTATTTGAACATTTTACTC	-6 bps	X1
16#	TTAAAGGAATGATGGATTTGGTCATTTAGCGGTGCTATTTGAACATTTTACTC	+4 bps	X1
17#	TTAAAGGAATGATGGATTTGGTCCCATAGCGGTGCTATTTGAACATTTTACTC	+4 bps	X1
18#	TTAAAGGAATGATGGATT----------CGGTGCTATTTGAACATTTTACTC	-7 bps	X1
19#	TTAAAGGAATGATGGATTTG----------CTATTTGAACATTTTACTC	-10 bps	X1
20#	TTAAAGGAATGATGGATTTGGTGGTGGTGGTGGTGAAATAGCGGTGCTATTTGAACATTTTACTC	-1+17 bps	X1
21#	TTAAAGGAATGATGGATTT----------ATTTGAACATTTTACTC	-13 bps	X1

C



Supplementary Figure S10. Design of *gfap* reporter line through NHEJ-mediated DNA replacement approach, cleavage efficiency of *gfap* sgRNA and verification by transient expression in F0 larvae.

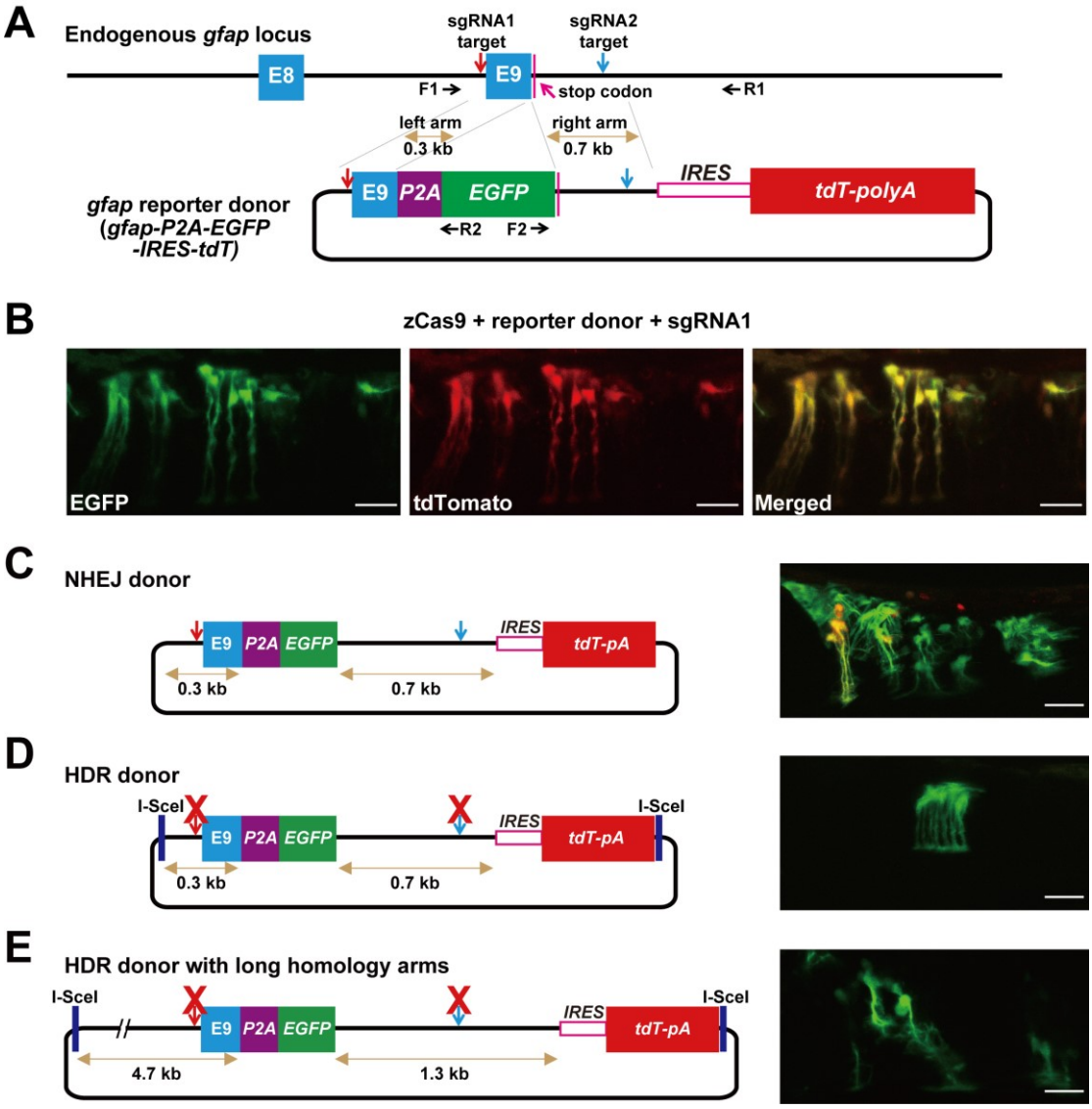
(**A**) Schematic of the strategy for NHEJ-mediated DNA replacement approach at the zebrafish *glial acidic fibrillary protein* (*gfap*) locus. *gfap* is specifically expressed in glial cells of the nervous system (Bernardos and Raymond, 2006; J. Li et al., 2015). Two sgRNAs were selected to specifically target two different sites in the *gfap*, one in the last intron (sgRNA1) and the other in 3' UTR region (sgRNA2). The *gfap* stop codon is indicated by a pink line. The homology arms are indicated by brown lines with double arrows, and the lengths of the left and right homology arms are 324 bps and 1326 bps, respectively. (**B**) Cleavage efficiency of the sgRNA1 and sgRNA2 targeting the zebrafish *gfap* locus. (**C**) Transient EGFP expression in glial cells when co-injecting the *gfap-P2A-EGFP*, zCas9 mRNA, and sgRNA1 and sgRNA2 of *gfap*. EGFP expression in the spinal cord was observed in ~ 88% (160/181) injected larvae (Table S3). Scale bars: 50 μm.



Supplementary Figure S11. Confirmation of the Ki(gfap^{EGFP}) reporter line.

(**A**) PCR verification of DNA replacement at the *gfap* locus. (**B**) Sequencing reports of the 5' and 3' junction sites of the progenies of two Ki(gfap^{EGFP}) founders (3# and 7#), which are identified from the 8 raised F0 larvae with broad EGFP expression (Table S1). (**C**) 5′ and 3′

junction sequences of F1 progenies of the two Ki(gfap^{EGFP}) founders in comparison with the sequence of WT zebrafish and the donor vector (Table S2). (**D**) Representative projected *in vivo* confocal images of 3-dpf Ki(gfap^{EGFP}) larvae at different visual fields. Lateral view: **D1**, **D4**, **D6**; dorsal view: **D2**, **D3**, **D5**. Scale bars: 250 μm (**D1**), 50 μm (**D2**), and 25 μm (**D3**-**D6**).



Supplementary Figure S12. Efficiency comparison of NHEJ- and HDR-mediated DNA replacement approach at the zebrafish *gfap* locus by using double fluorescence reporter donors.

(**A**) Schematic of the strategy for NHEJ-mediated DNA replacement approach at the zebrafish *gfap* locus by using a double fluorescence donor vector. The zebrafish *gfap* has 9 exons and E9 represents the 9th exon. The lengths of the left and right homology arms are 324 bps and 702 bps, respectively. IRES: internal ribosome entry site; tdT-polyA: tdTomato connected with a SV40 polyA sequence. In this right arm, the *gfap* transcriptional terminate signal was deleted to permit *IRES-tdT* transcription, and the target site of sgRNA2 was kept. (**B**) Co-expression of EGFP and tdTomato in glial cells when co-injecting the *gfap-P2A-EGFP-IRES-tdT*, zCas9 mRNA and *gfap* sgRNA1. (**C**) NHEJ-mediated DNA replacement at the *gfap* locus. Left, schematic of the *gfap-P2A-EGFP-IRES-tdT* donor vector for NHEJ-mediated DNA replacement. Right, representative projected *in vivo* confocal image of a 3-dpf larva co-injected with the vector, zCas9 mRNA, *gfap* sgRNA1 and *gfap* sgRNA2. In the vector, the *gfap* sgRNA1 and sgRNA2 target sites are indicated by a red or blue arrow,

respectively, and the homology arms are indicated by brown lines with double arrows. (**D**) HDR-mediated DNA replacement at the *gfap* locus by a donor vector with homology arms which are the same as in (**C**). Left, schematic of the HDR-related donor. The PAM of the sgRNA1 and sgRNA2 target sites located in the homology arms of the vector was mutated (red crosses) to prevent CRISPR/Cas9-induced excision. Right, representative projected *in vivo* confocal image of a 3-dpf larva co-injected with the vector, zCas9 mRNA, *gfap* sgRNA1 and *gfap* sgRNA2. (**E**) HDR-mediated DNA replacement at the *gfap* locus by a donor vector with homology arms longer than those used in (**D**). Left, schematic of the HDR-related donor. The lengths of the left and right arms are 4720 bps and 1326 bps, respectively. The PAM of the sgRNA1 and sgRNA2 target sites located in the homology arms of the vector was mutated (red crosses). Right, representative projected *in vivo* confocal image of a 3-dpf larva co-injected with the vector, zCas9 mRNA, *gfap* sgRNA1 and *gfap* sgRNA2. Scale bars, 20 μm (**B-E**).

Supplementary Table S1. Founder-screening rate of different edited alleles by NHEJ-mediated DNA replacement approach.

Donor	Injection	Survial	F0 screened	Founders
th-loxP-exon8-loxP	N.A.	N.A.	41	3
kdrl-loxP-exon12-frt-SM-frt-loxP	538	267	25	2
tcf3a-loxP-exon4-frt-SM-frt--loxP	511	203	15	2
th-P2A-EGFP	N.A.	166	11	1
gfap-P2A-EGFP	276	181	8	2

N.A.: not available.

Supplementary Table S2. Information of edited alleles and germline transmission rate of founders.

Donor	Founder	Gender	Intergration manner	5' junction	3' junction	ratio of germline with the knockin event
th-loxP-exon8-loxP	3#	female	Replacement	-3 bps	-7+7 bps	25% (12/48)
	13#	male	Replacement	-10 bps	-4 bps	4% (2/48)
	30#	male	Replacement	-44 bps	+3 bps	25% (12/48)
	2#	male	Insertion	-42 bps	N.A.	12% (7/59)
	26#	female	Insertion	-3 bps	N.A.	10% (5/48)
kdrl-loxP-exon12 -frt-SM-frt-loxP	1#	male	Replacement	0 bp	0 bp	28% (60/217)
	23#	female	Replacement	-9 bp	-5 bps	11% (5/44)
	16#	male	Insertion	N.A.	-5 bps	5% (8/165)
tcf3a-loxP-exon4 -frt-SM-frt-loxP	3#	male	Replacement	-1 bp	0 bp	66% (45/68)
	12#	female	Replacement	+4 bps	0 bp	8% (14/179)
	11#	male	Insertion	-2 bps	N.A.	12% (7/59)
	13#	male	Insertion	-2+19 bps	N.A.	10% (5/48)
th-P2A-EGFP	4#	male	Replacement	-1 bp	-11+17 bps	13% (16/119)
gfap-P2A-EGFP	3#	male	Replacement	-107 bps	-7 bps	12% (23/190)
	7#	male	Replacement	-107 bps	0 bp	5% (4/84)

Supplementary Table S3. Efficiency of knockin with different lengths of homology arms at the *gfap* locus.

Left arm length / upstream of sgRNA target	Right arm length / downstream of sgRNA target	sgRNAs	Donor	Normal development rate	GFP+/Total
172 bps / 5 bps	559 bps / 5 bps	1&2	circle	64.2%（79/123）	14%（11/79）
207 bps / 40 bps	594 bps / 40 bps	1&2	circle	48.6%(55/113)	40%（22/55）
324 bps / 157 bps	1326 bps / 872 bps	1&2	circle	65.6%(181/276)	88%(160/181)
4720 bps / 4553 bps	1326 bps / 872 bps	1&2	circle	53.1%(34/64)	70%(82/117)
324 bps / 157 bps	1326 bps / 872 bps	1&2	fragment	40.0%(209/523)	52%(108/209)
324 bps / 157 bps	1326 bps / 872 bps	1	circle	53.8%(56/104)	79%(44/56)

Supplementary Table S4. Target sequences of all tested sgRNAs

Donor	gRNA target sequences
th-loxP-exon8-loxP	GGATTTGCTATGCCATAGTAG GGG GCAAATCCCTCTGAAAACCC AGG GGTGGTTGCTAGAATATTCA TGG GGAATGAACTGCCAACTACTC TGG gGTGTTGCCGAACTGAATGG TGG gGAGGGCGTGTCTGCAGTTG GGG
kdrl-loxP-exon12-loxP	gTAAACAGTCTGTGAACCCC AGG GTTTCATACCTTTGATGTGA CGG GAGGAAATTACATGGGACG TGG gTGAGAGGGGTACAGTACGG TGG gCTCGGCCACCTTAGTACAG TGG GCTGCTACGCCACTGTACTA AGG gTATAAGGGTTGGAGAACAC GGG
tcf3a-loxP-exon4-loxP	gAAAATGAAAACCGAAAATT TGG gATATACATTAGGGCTGGGC TGG gAGTGGCGTAGGTAGACATT TGG GGACGCAGCTATCAGTGGCGT AGG GCATTACAGATCTGATCCTGT CGG GATCAGGAGACGTGTGTTAGT AGG GCTGGTCTAGTGCTGATCTC AGG gAGTAGGCCACAGACAGGAG AGG
gfap-P2A-EGFP	GTTAGTAAAGGACCTTGAGC AGG GGGCGGCGCCAATAGCCTAG TGG GTTAAGTTTACTTAATCGGG CGG GATCATTAAAGAGTCCACTA CGG GGTGTTAAAGTAAGGCAGAA GGG GGAATGATGGATTTGGTCAG CGG
th-P2A-EGFP	GGGTTGTTACACTCATAAAG GGG gGAATTAAGCCAACACATTCA GGG gGATACATTATTTAACGCTGA AGG

The sequences in red were used to generate DNA replacement-mediated knockin zebrafish in this study, and yellow box represents an efficient target sequence which was used in a previous study (J. Li et al., 2015). The green characters represent PAM sequences. The "g" was added for synthesizing sgRNAs.

Supplementary Table S5. Sequences of primers used.

Donor	Primers	Sequence (5' to 3')
th-loxP-exon8-loxP	F1	TACAGCCCAGATAACATCCCT
	R1	GAAACAGACTATGACTGAATGAGTT
	F2	ATGGGTGACGTCAATAACTTCGTA
	R2	TGGGTTTTCAGAGATAACTTCGT
kdrl-loxP-exon12 -frt-SM-frt-loxP	F1	TTGGCAGGGGTAAAGCACAG
	R1	AACGTGCCACTTGACATCCA
	F2	CCCTTGGTCACCTTCAGCTT
	R2	CCCACAACGAGGACTACACC
tcf3a-loxP-exon4 -frt-SM-frt-loxP	F1	TTTGCCGACCTCACCCATC
	R1	TTTAAAGGGACTTGGAGTAAAAGTG
	F2	CCCTTGGTCACCTTCAGCTT
	R2	CCCACAACGAGGACTACACC
th-P2A-EGFP	F1	ATTCCGCCGAAACGCAGTAAAGA
	R1	AGTGTACCGTAAAGGTGTTA
	F2	CATGGTCCTGCTGGAGTTCGTG
	R2	CTCGCCCTTGCTCACCATAG
gfap-P2A-EGFP	F1	AGAATTTAGGATCCCACACAACTCA
	R1	TTGCTTGAAGGCACTACTAGG
	F2	CATGGTCCTGCTGGAGTTCGTG
	R2	CTCGCCCTTGCTCACCATAG

Supplementary Discussion

Genomic DNA replacement through homogenous recombination (HR) is a common approach in generating genetically engineered animals (Capecchi, 2005). Due to its low efficiency, the HR-mediated approach usually requires both positive and negative selective systems to pre-screen gene-modified events in cultured embryonic stem cells (Downing and Battey, 2004). However, HR-mediated approach cannot be applied in zebrafish models because there is currently no the culture technology of zebrafish embryonic stem cells. Recently, with the help of different endonucleases, HDR has been used to generate knockin lines of zebrafish with various efficiencies (Zu et al., 2013; Irion et al., 2014; Shin et al., 2014; Armstrong et al., 2016; Hoshijima et al., 2016; M. Li et al., 2016; Burg et al., 2018; Tessadori et al., 2018). For most of reported zebrafish lines generated via HDR, single-stranded oligodeoxynucleotide (ssODN) was usually used as a repair template (Irion et al., 2014; Armstrong et al., 2016; Hoshijima et al., 2016; Burg et al., 2018; Tessadori et al., 2018). This strategy can only achieve genomic alteration with a few of bps such as point mutation and short tag insertion. Long DNA fragment replacement has wide application in gene editing, because it can achieve different types of genomic editing. However, long DNA fragment replacement for zebrafish genome editing is still under development.

Pre-screening via obvious phenotypes or reporter gene expression, both of which imply correct genomic replacement events, was sometimes used to increase screening efficiency in those studies (Irion et al., 2014; Shin et al., 2014; Hisano et al., 2015; Hoshijima et al., 2016). However, the requirement for pre-screening cannot be met in many cases. Without pre-screening, genomic editing mediated by HDR in zebrafish shows relatively low efficiencies (Armstrong et al., 2016; Burg et al., 2018; Tessadori et al., 2018), in particular for long DNA fragments (Zu et al., 2013; Hoshijima et al., 2016). This makes it necessary to develop more efficient methods for achieving long DNA fragments replacement.

NHEJ has been recently used to achieve targeted integration in cultured cells, mouse and zebrafish with a high efficiency through inserting a DNA fragment at a single DSB site (Maresca et al., 2013; Auer et al., 2014; Kimura et al., 2014; J. Li et al., 2015; Ota et al., 2016; Suzuki et al., 2016; Kesavan et al., 2018; W. Li et al., 2019; J. Li et al., 2020), but these strategies still cannot be readily applied for sophisticated genomic editing, in particular for

achieving traditional conditional knockout. In the present study, by taking advantage of highly active NHEJ-mediated knockin with homology arms and non-coding region targeting, we developed an efficient long DNA fragment replacement method by introducing two DSB sites at non-coding genomic regions via the CRISPR/Cas9 system. With pre-screening, the screen rate of our method for long genomic DNA replacement is 25% (*gfap* reporter allele: 2/8), which is comparable to the reported targeted integrations (Auer et al., 2014; Kimura et al., 2014; Shin et al., 2014; Hisano et al., 2015; J. Li et al., 2015; Ota et al., 2016). Without pre-screening, the screen rate of our method for long genomic DNA replacement is 9% (*th* conditional allele: 3/41, *kdrl*: 2/25, *tcf3a*: 2/15, *th* reporter allele: 1/11), which is higher than that of HDR-based method in previous studies (4%; *th*: 4/275 and *kcnh6a*: 3/43; see also Refs. Zu et al., 2013 and Hoshijima et al., 2016). In comparison with HR/HDR-based approaches in a more precise way, we did experiments at the same genomic locus *gfap* and found that the DNA replacement efficiency of NHEJ-mediated approach is much higher (see Figure 1J and K). As DSB sites were designed in non-coding regions, NHEJ-induced indels will not shift the reading frame of targeted genes, leading to an increased rate of successful DNA replacement. We also found that homology arms can increase knockin efficiency, further studies should test the feasibility of NHEJ-mediated long fragment replacement via universal construct with simply modification by just adding target sites and the short homology arms extended out of target sites. Among our data, although most of junctions in the screened replacement lines shows the NHEJ-mediated integrations, five out of twenty junctions shows the perfect repair happened without indels, implying that HDR-mediated repair cannot be excluded completely in our approach.

Materials and Methods

Zebrafish

Adult zebrafish were maintained with automatic fish housing system (ESEN, China) at 28°C. Embryos were raised under a 14h-10h light-dark cycle in 10% Hank's solution that consisted of (in mM): 140 NaCl, 5.4 KCl, 0.25 Na2HPO4, 0.44 KH2PO4, 1.3 CaCl2, 1.0 MgSO4 and 4.2 NaHCO3 (pH 7.2). The embryos for imaging experiments were treated with 0.003% 1-phenyl-2-thiourea (PTU, Sigma) from 24 hours post-fertilization to prevent pigment

formation. The transgenic line Tg(kdrl:EGFP) and knockin line of Ki(th-P2A-EGFP) were described previously (Jin et al., 2005; J. Li et al., 2015), and the transgenic line Tg(dat:mRFP-Cre) was generated by using a modified zebrafish BAC clone which contains the dat promoter. To design sgRNAs, AB/WT zebrafish were firstly screened by PCR and sequencing at sgRNA target sites. Zebrafish handling procedures were approved by the Institute of Neuroscience, Chinese Academy of Sciences.

Production of mRNA and sgRNAs *in vitro*

The zCas9 plasmid pGH-T7-zCas9 was linearized by XbaI and used as a template for Cas9 mRNA *in vitro* synthesis with the mMACHINE T7 Ultra kit (Ambion, USA) (Liu et al., 2014). Cre or Flp plasmids (in pcDNA3.1) were linearized by XmaJI and their mRNA were synthesized *in vitro* by mMACHINE T7 Ultra kit. We used the CRISPR/Cas9 design tool provided in the website (http://zifit.partners.org) to select specific targets to minimize off-target effects. The sequences of sgRNA targets tested are listed in Supplementary Table S4. A pair of oligonucleotides containing the targeting sequence of sgRNAs were annealed and cloned at the downstream of the T7 promoter in the pT7-sgRNA vector. The sgRNAs were synthesized by the MAXIscript T7 Kit (Ambion, USA) and were purified by using the mirVana™ miRNA Isolation Kit (Ambion, USA).

Replacement donor vector construction

Replacement donor plasmids were constructed by standard molecular cloning techniques with restriction enzyme cleavage and DNA ligation. The th-loxP-exon8-loxP replacement donor was constructed by ligating three fragments (a left arm, loxP-exon8-loxP and a right arm) with the pMD-19T vector via KpnI, BamHI, AgeI and SalI (Takara, Japan). The three fragments were all amplified from the genomic DNA of Ki(th-P2A-EGFP) zebrafish by using the PrimeSTAR HS DNA polymerase (Takara, Japan). The kdrl-loxP-exon12-frt-SM-frt-loxP or tcf3a-loxP-exon4-frt-SM-frt-loxP replacement donor were constructed by ligating four fragments (a left arm, loxP-exon12/4, frt-SM-frt, and a loxP-right arm) with the pMD-19T vector via KpnI, SacI, PaeI, BglII and SalI. The th-P2A-EGFP replacement donor was constructed by modifying the knockin plasmid used in our previous work (J. Li et al., 2015).

The original P2A-EGFP fragment was replaced with P2A-EGFP-sgRNA2 target by using AflII and AgeI, and then silent mutation of 5 bps in the 3' side of the th CDS was introduced by using the Fast Mutagenesis System (Transgen, China). For constructing different gfap-P2A-EGFP replacement donors, the left and right arms of gfap-P2A-EGFP were amplified from the genomic DNA of sgRNA target-screened wild-type zebrafish, and the P2A-EGFP fragment was obtained from the th-P2A-EGFP knockin plasmid. The gfap-P2A-EGFP-IRES-tdT donor was constructed by modifying gfap-P2A-EGFP plasmid, which ligated the fragment of IRES-tdT-polyA with the 3' side of the shorted right homology arm of the gfap-P2A-EGFP. The HDR-related donors were constructed by mutating the PAM regions of gfap-P2A-EGFP-IRES-tdT plasmids, and then a pair of head-to-head oriented I-SceI sites were added out of the left arm and IRES-tdT.

Microinjection of one-cell stage embryos

All donor plasmids were purified before microinjection by the Gel Extraction Kit (Qiagen, Germany). Flp mRNA and Cre mRNA were synthesized *in vitro* as zCas9 mRNA. The zCas9 mRNA, sgRNAs, and donor plasmids were co-injected into one-cell zebrafish embryos. Each embryo was injected with 1 nl solution containing 600 ng/μl zCas9 mRNA, 50 ng/μl sgRNA each, and 15 ng/μl donor plasmid into the animal pole. Only 1 U/μl I-SceI enzyme 5 X/μl buffer (NEB) were mixed when co-injecting with the HDR donor. Flp mRNA or Cre mRNA was also injected into one-cell zebrafish embryos with 1 nl containing 100 ng/μl mRNA.

PCR and sequencing for confirming sgRNA cleavage

Genomic DNA for PCR was extracted from 20 embryos at 3 dpf that were co-injected with zCas9 mRNA and sgRNAs. Fragments containing target sites were amplified by the Takara Ex Taq, and were then cloned by the TA Cloning Kit (Takara, Japan) for sequencing. Generally, we sequenced 20-50 clones for analyzing the cleavage efficiency of each sgRNA.

Founder screening and the germline mosaicism rate analysis

All F0 zebrafish were screened by out-crossing with AB/WT adults. F1 progenies of Ki(th^{EGFP}) or Ki($gfap^{EGFP}$) founder candidates were screened by EGFP signal at 3 dpf. The ratio of F1

progeny carrying the modification was calculated based on EGFP-positive larvae. 1 - 3 positive larvae were picked up to extract the genomic DNA to determine the integration manner. F1 progenies of Ki(th^{fl}-P2A-EGFP) candidates were screened by collecting ~ 30 F1 progenies of each candidate for genomic DNA extraction and PCR identification. The ratio of F1 progeny carrying the modification was analyzed by PCR analysis of 48 F1 progenies one by one. F1 progenies of Ki(kdrl^{fl}) or Ki(tcf3a^{fl}) candidates were screened first by DsRed signal in heart, then 1-3 positive larvae were picked up for genome extraction and PCR identification. The ratio of F1 progeny carrying the modification was calculated based on DsRed-positive larvae. The genomic DNA of positive F1 larvae was also used for confirmation of the integration manner.

Verification of NHEJ-mediated DNA replacement

The genomic DNA of 1 - 20 3-dpf F1 embryos of each F0 were extracted. The fragment of 5' and 3' junctions were amplified respectively with the primer pair of F1/R2 and F2/R1, whereas the full length product of the entire edited region was amplified with F1/R1 (Supplementary Table S5). The size of all PCR products was first examined by gel electrophoresis, and the full length products amplified by F1/R1 were sequenced directly or cloned and sequenced. The sequencing data were used to analyze the integration manner at junction sites or the excision at frt/loxP sites.

***In vivo* confocal imaging**

For *in vivo* confocal imaging, larvae at 3 or 5 dpf were immobilized in 1.2 % low melting-point agarose without paralysis or anesthetics. Imaging experiments were performed under a 20X water-immersion objective by using a Fluoview 1000 two-photon microscope (Olympus, Japan) (Y. Li et al., 2012). The spatial resolution of all images was 1024 × 1024 pixels.

Statistics

Statistical analysis was performed using unpaired Student's t-test. The P value less than 0.05 were considered to be statistically significant. All results are represented as mean ± SEM.

Supplementary References

Armstrong, G.A., Liao, M., You, Z., et al. (2016). Homology Directed Knockin of Point Mutations in the Zebrafish tardbp and fus Genes in ALS Using the CRISPR/Cas9 System. PLoS One *11*, e0150188.

Auer, T.O., Duroure, K., De Cian, A., et al. (2014). Highly efficient CRISPR/Cas9-mediated knock-in in zebrafish by homology-independent DNA repair. Genome research *24*, 142-153.

Bernardos, R.L., and Raymond, P.A. (2006). GFAP transgenic zebrafish. Gene Expr Patterns *6*, 1007-1013.

Burg, L., Palmer, N., Kikhi, K., et al. (2018). Conditional mutagenesis by oligonucleotide-mediated integration of loxP sites in zebrafish. PLoS genetics *14*, e1007754.

Capecchi, M.R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. Nat Rev Genet *6*, 507-512.

Chong, S.W., Nguyet, L.M., Jiang, Y.J., et al. (2007). The chemokine Sdf-1 and its receptor Cxcr4 are required for formation of muscle in zebrafish. BMC Dev Biol *7*, 54.

Downing, G.J., and Battey, J.F., Jr. (2004). Technical assessment of the first 20 years of research using mouse embryonic stem cell lines. Stem cells *22*, 1168-1180.

Filippi, A., Mahler, J., Schweitzer, J., et al. (2010). Expression of the paralogous tyrosine hydroxylase encoding genes th1 and th2 reveals the full complement of dopaminergic and noradrenergic neurons in zebrafish larval and juvenile brain. J Comp Neurol *518*, 423-438.

Habeck, H., Odenthal, J., Walderich, B., et al. (2002). Analysis of a zebrafish VEGF receptor mutant reveals specific disruption of angiogenesis. Curr Biol *12*, 1405-1412.

Hisano, Y., Sakuma, T., Nakade, S., et al. (2015). Precise in-frame integration of exogenous DNA mediated by CRISPR/Cas9 system in zebrafish. Scientific reports *5*, 8841.

Hoshijima, K., Jurynec, M.J., and Grunwald, D.J. (2016). Precise Editing of the Zebrafish Genome Made Simple and Efficient. Dev Cell *36*, 654-667.

Irion, U., Krauss, J., and Nusslein-Volhard, C. (2014). Precise and efficient genome editing in zebrafish using the CRISPR/Cas9 system. Development *141*, 4827-4830.

Jao, L.E., Wente, S.R., and Chen, W. (2013). Efficient multiplex biallelic zebrafish genome editing using a CRISPR nuclease system. Proc Natl Acad Sci U S A *110*, 13904-13909.

Jin, S.W., Beis, D., Mitchell, T., et al. (2005). Cellular and molecular analyses of vascular tube and lumen formation in zebrafish. Development *132*, 5199-5209.

Kesavan, G., Hammer, J., Hans, S., et al. (2018). Targeted knock-in of CreER (T2) in zebrafish using CRISPR/Cas9. Cell Tissue Res *372*, 41-50.

Kimura, Y., Hisano, Y., Kawahara, A., et al. (2014). Efficient generation of knock-in transgenic zebrafish carrying reporter/driver genes by CRISPR/Cas9-mediated genome engineering. Scientific reports *4*, 6545.

Li, J., Li, H.Y., Gu, S.Y., et al. (2020). One-step generation of zebrafish carrying a conditional knockout-knockin visible switch via CRISPR/Cas9-mediated intron targeting. Science China. Life sciences *63*, 59-67.

Li, J., Zhang, B.B., Ren, Y.G., et al. (2015). Intron targeting-mediated and endogenous gene integrity-maintaining knockin in zebrafish using the CRISPR/Cas9 system. Cell research *25*, 634-637.

Li, M., Zhao, L., Page-McCaw, P.S., et al. (2016). Zebrafish Genome Engineering Using the CRISPR-Cas9 System. Trends in genetics : TIG *32*, 815-827.

Li, W., Zhang, Y., Han, B., et al. (2019). One-step efficient generation of dual-function conditional knockout and geno-tagging alleles in zebrafish. eLife *8*.

Li, Y., Du, X.F., Liu, C.S., et al. (2012). Reciprocal regulation between resting microglial dynamics and neuronal activity in vivo. Dev Cell *23*, 1189-1202.

Liu, D., Wang, Z., Xiao, A., et al. (2014). Efficient gene targeting in zebrafish mediated by a zebrafish-codon-optimized cas9 and evaluation of off-targeting effect. Journal of genetics and genomics = Yi chuan xue bao *41*, 43-46.

Maresca, M., Lin, V.G., Guo, N., et al. (2013). Obligate ligation-gated recombination (ObLiGaRe): custom-designed nuclease-mediated targeted integration through nonhomologous end joining. Genome research *23*, 539-546.

Ota, S., Taimatsu, K., Yanagi, K., et al. (2016). Functional visualization and disruption of targeted genes using CRISPR/Cas9-mediated eGFP reporter integration in zebrafish. Scientific reports *6*, 34991.

Shah, A.N., Davey, C.F., Whitebirch, A.C., et al. (2015). Rapid reverse genetic screening using CRISPR in zebrafish. Nature methods *12*, 535-540.

Shin, J., Chen, J., and Solnica-Krezel, L. (2014). Efficient homologous recombination-mediated genome engineering in zebrafish using TALE nucleases. Development *141*, 3807-3818.

Suzuki, K., Tsunekawa, Y., Hernandez-Benitez, R., et al. (2016). In vivo genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. Nature *540*, 144-149.

Tessadori, F., Roessler, H.I., Savelberg, S.M.C., et al. (2018). Effective CRISPR/Cas9-based nucleotide editing in zebrafish to model human genetic cardiovascular disorders. Dis Model Mech *11*.

Zhang, Y., Huang, H., Zhang, B., et al. (2016). TALEN- and CRISPR-enhanced DNA homologous recombination for gene editing in zebrafish. Methods in cell biology *135*, 107-120.

Zu, Y., Tong, X., Wang, Z., et al. (2013). TALEN-mediated precise genome modification by homologous recombination in zebrafish. Nature methods *10*, 329-331.