

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 ***GeneWeld: a method for efficient targeted integration directed by short homology***

2

3 Wesley A. Wierson<sup>1,8</sup>, Jordan M. Welker<sup>1,8</sup>, Maira P. Almeida<sup>1,8</sup>, Carla M. Mann<sup>1</sup>, Dennis A.  
4 Webster<sup>2</sup>, Melanie E. Torrie<sup>1</sup>, Trevor J. Weiss<sup>1</sup>, Macy K. Vollbrecht<sup>2</sup>, Merrina Lan<sup>1</sup>, Kenna C.  
5 McKeighan<sup>1</sup>, Jacklyn Levey<sup>1</sup>, Zhitao Ming<sup>1</sup>, Alec Wehmeier<sup>1</sup>, Christopher S. Mikelson<sup>1</sup>, Jeffrey  
6 A. Haltom<sup>1</sup>, Kristen M. Kwan<sup>3</sup>, Chi-Bin Chien<sup>4</sup>, Darius Balciunas<sup>5</sup>, Stephen C. Ekker<sup>6</sup>, Karl J.  
7 Clark<sup>6</sup>, Beau R. Webber<sup>7</sup>, Branden Moriarity<sup>7</sup>, Staci L. Solin<sup>2</sup>, Daniel F. Carlson<sup>2</sup>, Drena L.  
8 Dobbs<sup>1</sup>, Maura McGrail<sup>1</sup>, Jeffrey J. Essner<sup>1\*</sup>

9

- 10 1. Department of Genetics, Development and Cell Biology, Iowa State University, IA, USA  
11 2. Recombinetics, Inc., St. Paul, MN, USA  
12 3. Department of Human Genetics, University of Utah School of Medicine, UT, USA  
13 4. Department of Neurobiology and Anatomy, University of Utah Medical Center, UT, USA  
14 5. Department of Biology, Temple University, PA, USA  
15 6. Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN, USA  
16 7. Department of Pediatrics, Masonic Cancer Center, University of Minnesota, Minneapolis,  
17 MN, USA  
18 8. These authors contributed equally.

19

20

21 \*Corresponding author and lead contact: [jessner@iastate.edu](mailto:jessner@iastate.edu)

22

23

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

Wierson et al. describe a targeted integration strategy, called GeneWeld, and a vector series for gene tagging, pGTag, which promote highly efficient and precise targeted integration in zebrafish, pig fibroblasts, and human cells. This approach establishes an effective genome engineering solution that is suitable for creating knock-in mutations for functional genomics and gene therapy applications. The authors describe high rates of germline transmission (50%) for targeted knock-ins at eight different zebrafish loci and efficient integration at safe harbor loci in porcine and human cells.

## **Abstract**

Choices for genome engineering and integration involve high efficiency with little or no target specificity or high specificity with low activity. Here, we describe a targeted integration strategy, called GeneWeld, and a vector series for gene tagging, pGTag (plasmids for Gene Tagging), which promote highly efficient and precise targeted integration in zebrafish embryos, pig fibroblasts, and human cells utilizing the CRISPR/Cas9 system. Our work demonstrates that *in vivo* targeting of a genomic locus of interest with CRISPR/Cas9 and a donor vector containing as little as 24 to 48 base pairs of homology directs precise and efficient knock-in when the homology arms are exposed with a double strand break *in vivo*. Our results suggest that the length of homology is not important in the design of knock-in vectors but rather how the homology is presented to a double strand break in the genome. Given our results targeting multiple loci in different species, we expect the accompanying protocols, vectors, and web interface for homology arm design to help streamline gene targeting and applications in CRISPR and TALEN compatible systems.

## **Keywords**

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1

2 CRISPR/Cas9, knock-in, homology mediated-end joining, targeted integration, zebrafish, pig  
3 fibroblasts, human K-562 cells

4

## 5 **Introduction**

6

7 Designer nucleases have rapidly expanded the way in which researchers can utilize  
8 endogenous DNA repair mechanisms for creating gene knock-outs, reporter gene knock-ins,  
9 gene deletions, single nucleotide polymorphisms, and epitope tagged alleles in diverse species  
10 (Bedell et al., 2012; Beumer et al., 2008; Carlson et al., 2012; Geurts et al., 2009; Yang et al.,  
11 2013). A single dsDNA break in the genome results in increased frequencies of recombination  
12 and promotes integration of homologous recombination (HR)-based vectors (Hasty et al., 1991;  
13 Hoshijima et al., 2016; Orr-Weaver et al., 1981; Rong and Golic, 2000; Shin et al., 2014; Zu et  
14 al., 2013). Additionally, *in vitro* or *in vivo* linearization of targeting vectors stimulates homology-  
15 directed repair (HDR) (Hasty et al., 1991; Hoshijima et al., 2016; Orr-Weaver et al., 1981; Rong  
16 and Golic, 2000; Shin et al., 2014; Zu et al., 2013). Utilizing HDR or HR at a targeted double-  
17 strand break (DSB) allows base-pair precision to directionally knock-in exogenous DNA,  
18 however, frequencies remain variable and engineering of targeting vectors is cumbersome.

19

20 Previous work has shown *Xenopus* oocytes have the ability to join or recombine linear  
21 DNA molecules that contain short regions of homology at their ends, and this activity is likely  
22 mediated by exonuclease activity allowing base pairing of the resected homology (Grzesiuk and  
23 Carroll, 1987). More recently, it was shown in *Xenopus*, silkworm, zebrafish, and mouse cells  
24 that a plasmid donor containing short ( $\leq 40$  bp) regions of homology to a genomic target site can  
25 promote precise integration at the genomic cut site when the donor plasmid is cut adjacent to

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 the homology (Aida et al., 2016; Hisano et al., 2015; Nakade et al., 2014). Gene targeting is  
2 likely mediated by the alternative-end joining/microhomology-mediated end joining (MMEJ)  
3 pathway or by a single strand annealing (SSA) mechanism (Ceccaldi et al., 2016). In contrast, in  
4 human cell culture, linear donors using a similar strategy with homologous ends have been  
5 reported to show inefficient integration until homology domains reach ~600 bp (Zhang et al.,  
6 2017), suggesting that different repair pathways may predominate depending on cell type. In the  
7 initial reports using short regions of homology for *in vivo* gene targeting in zebrafish, the level of  
8 mosaicism in F0 injected animals was high, resulting in inefficient recovery of targeted alleles  
9 through the germline (Aida et al., 2016; Hisano et al., 2015; Nakade et al., 2014).

10

11 Here, we present GeneWeld, a strategy for targeted integration directed by short  
12 homology, and demonstrate increased germline transmission rates for recovery of targeted  
13 alleles. We provide a detailed protocol and a suite of donor vectors, called pGTag, that can be  
14 easily engineered with homologous sequences (hereafter called homology arms) to a gene of  
15 interest, and a web interface for designing homology arms ([www.genesculpt.org/gtaghd/](http://www.genesculpt.org/gtaghd/)). We  
16 demonstrate that 24 or 48 base pairs of homology directly flanking cargo DNA promotes  
17 efficient gene targeting in zebrafish, pig, and human cells with frequencies up to 10-fold higher  
18 than other HR strategies. Our results also suggest that longer homology arms up to 1 kb in  
19 length provide no advantage to knock-in frequencies over short homology, and the important  
20 aspect of knock-in design is the ability to expose the homology on the knock-in cassette ends.  
21 Using short homology-arm mediated end joining, we can achieve germline transmission rates  
22 averaging approximately 50% across several zebrafish loci. Southern blot analysis in the F1  
23 generation reveals that the GeneWeld strategy can yield alleles with precise integration at both  
24 5' and 3' ends, as well as alleles that are precise on just one end. Finally, we present a strategy  
25 to delete and replace up to 48kb of genomic DNA with a donor containing homology arms  
26 flanking two distal CRISPR/Cas9 sites in a gene. These tools and methodology provide a

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 tractable solution to creating precise targeted integrations and open the door for other genome  
2 editing strategies using short homology.

3

#### 4 **Results**

5

6 The GeneWeld strategy takes advantage of two simultaneous actions to initiate targeted  
7 integration directed by short homology (Fig. 1a). First, a high efficiency nuclease introduces a  
8 DSB in the chromosomal target. Simultaneously, a second nuclease makes a DSB in the  
9 pGTag vector integration cassette exposing the short homology arms. The complementarity  
10 between the chromosomal DSB and the donor homology arms activates a MMEJ/SSA or other  
11 non-NHEJ DNA repair mechanism, together referred to as homology-mediated end joining  
12 (HMEJ). The reagents needed for this gene targeting strategy include *Cas9* mRNA to express  
13 the Cas9 nuclease, a guide RNA targeting the genomic sequence of interest, a universal gRNA  
14 (UgRNA) that targets two sites in the pGTag series donor vectors to expose the homology arms,  
15 and a pGTag/donor vector with gene specific homology arms (Fig. 1a). The universal gRNA  
16 (UgRNA) has no predicted sites in zebrafish, pig, or human genomes. Alternatively, a gene  
17 specific guide RNA can be used to expose homology arms in the donor vector. For simplicity we  
18 will refer to this set of reagents as 'GeneWeld reagents'. Using GeneWeld reagents to target  
19 various loci, we demonstrate widespread reporter gene expression in injected F0 zebrafish  
20 embryos, porcine fibroblasts, and human K-562 cells that correlates with efficient and precise in-  
21 frame integration in multiple species and cell systems.

22

#### 23 **A single 48bp short homology domain drives efficient CRISPR targeted integration**

24 To develop baseline gene targeting data, we engineered variable length homology  
25 domains to target *noto*. Homology lengths were based on observations that DNA repair  
26 enzymes bind DNA and search for homology in 3 or 4 base pair lengths (Fig. S1a) (Conway et

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 al., 2004; Singleton et al., 2002). Injection of a *noto* sgRNA that targets *noto* exon 1 and the 5'  
2 homology domain cloned into a 2A-TagRFP-CAAX-SV40 donor vector, efficient targeted  
3 integration was observed as notochord-specific RFP expression (Fig. S1a-c; Table S1-S3). The  
4 frequency of embryos with notochord-specific RFP expression increased with the length of the  
5 homology arm up to 48 bp (Fig. S1b). Somatic junction fragment analysis revealed precise  
6 integration efficiencies reaching 95% of sequenced alleles (Fig. S1d), indicating a strong  
7 correlation between expression of the reporter gene and precise somatic targeting of *noto*.

8  
9 Following these initial experiments, a 3 bp spacer sequence was included in all  
10 homology arm designs in order to separate the donor CRISPR/Cas9 target PAM and the  
11 homology domain (Fig. S2a). The spacer was included to prevent arbitrarily increasing the  
12 length of the targeting domain, since single base pair alterations in the homology region affected  
13 the tissue specific expression and presumed knock-in efficiency up to 2-fold in somatic tissue  
14 (Fig. S2b). Based on these observations the following experiments were carried out with either  
15 24 or 48 base pairs of homology to the double strand break in the targeted gene.

16

### 17 **A universal guide RNA to liberate donor homology for targeted integration**

18 To simplify donor design and liberate donor cargo *in vivo* with reproducible efficiency, a  
19 universal guide RNA sequence UgRNA, with no predicted targets in zebrafish, pig, or human  
20 genomes, was designed using optimal base composition in CRISPRScan (Fig. S3a) (Moreno-  
21 Mateos et al., 2015). To test the ability of the guide to direct a Cas9 double strand break and  
22 efficient targeted integration, the UgRNA and CGG PAM sequences were cloned 5' to the *noto*  
23 homology arm in the 2A-TagRFP-CAAX-SV40 donor vector (Fig. S3b). Zebrafish embryos  
24 injected with Cas9 mRNA, UgRNA, *noto* sgRNA, and donor plasmid resulted in 21% of injected  
25 embryos showing notochord RFP, suggesting that Cas9 with the UgRNA efficiently exposes the

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 donor vector 5' homology arm and drive precise targeted integration (Fig. S3c). The high  
2 frequency of notochord-specific RFP-positive embryos following injection suggests repair of the  
3 DSB preferentially utilizes homology in the targeting construct over the NHEJ pathway.

4

### 5 **Dual homology arm liberation directs precise 5' and 3' integration in somatic tissue**

6 We leveraged the activity of the UgrNA to develop GeneWeld, a strategy for targeted  
7 integration that results in high frequency precision repair at both 5' and 3' junctions at the target  
8 site. We built a series of vectors, pGTag, which contain sites on both sides of the cargo for  
9 cloning a short homology arm that is complementary to the 5' or 3' sequence flanking the  
10 genomic target site. The vectors also include the UgrNA sequence outside the sites for  
11 homology arm cloning (Fig. 1a, b). The final donor targeting vector will contain a cargo flanked  
12 by 5' and 3' homology arms with UgrNA sequences on both ends. Cleavage by Cas9 at the  
13 UgrNA sites liberates the DNA cargo from the plasmid backbone and exposes both 5' and 3'  
14 donor homology arms for interaction with DNA on either side of the genomic DSB (Fig. 1c).

15

16 We extended our analysis of the GeneWeld targeted integration strategy to four genes in  
17 zebrafish, *noto*, *tyrosinase (tyr)*, *esama (endothelial cell adhesion molecule a)*, and *connexin*  
18 *43.4 (cx43.4)*, and measured the frequency of precise targeted integration in somatic tissue  
19 (Fig. 2 a-d). Injection of 24 or 48 bp homology arm *noto* 2A-eGFP-SV40 donors resulted in 24%  
20 of zebrafish embryos showing extensive reporter expression in the notochord (Fig. 2 a, e),  
21 suggesting a similar precise integration efficiency compared to targeting with the single 5'  
22 homology arm 2A-TagRFP-CAAX-SV40 vector (Fig. S1, Table 1, 2). The results also suggest  
23 24 bp of homology directs targeted integration as efficiently as 48 bp, further simplifying  
24 construction of GeneWeld vectors.

25

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 Targeting exon 4 of *tyr* or exon 2 of *esama* with a 24 bp homology arm 2A-TagRFP-  
2 CAAX-SV40 donor did not result in detectable RFP signal, similar to previous reports for *tyr*  
3 (Hisano et al., 2015). However, PCR junction fragments from injected embryos showed the  
4 donor was precisely integrated in frame into *tyr* exon 4 (Fig. S4), suggesting RFP expression  
5 was below the threshold of detection. To amplify the fluorescent signal, we built pGTag 24 bp  
6 homology arm 2A-Gal4VP16- $\beta$ -actin3'UTR donors to integrate the Gal4VP16 trans-activator  
7 into the *tyr* and *esama* target sites, and injected into transgenic zebrafish embryos carrying a  
8 14xUAS-RFP reporter, *Tg(UAS:mRFP)<sup>tp12</sup>* (Balciuniene et al., 2013). This resulted in strong RFP  
9 signal in 64% of *tyr* injected animals (Fig. 2b, e), however, the embryos were highly mosaic with  
10 only 9% of RFP embryos displaying extensive expression throughout most of the pigmented  
11 cells. Targeting *esama* exon 2 with 2A-Gal4VP16- $\beta$ -actin3'UTR in the *Tg(UAS:mRFP)<sup>tp12</sup>*  
12 transgenic background resulted in 21% of embryos displaying extensive RFP expression  
13 specifically in the vasculature (Fig. 2c, e). This approach was further extended to five additional  
14 loci, targeting 2A-Gal4VP16 to *filamin a (flna)* exon 4, *moesin a (msna)* exon 2 and 6, *aquaporin*  
15 *1a1 (aqp1a1)* exon 1, *aquaporin 8a1 (aqp8a1)* exon 1, and *annexin a2a (anxa2a)* exon 3. At  
16 these loci, transient expression of RFP was observed following injection in 4-55% of  
17 *Tg(UAS:mRFP)<sup>tp12</sup>* embryos (Table S1 and S2). Taken together, these results suggest that the  
18 GeneWeld targeting method promotes high efficiency targeted integration in zebrafish embryo  
19 somatic tissue.

20

21 We next compared the frequency of GeneWeld targeted integration in zebrafish embryos  
22 with previous methods that used TALEN targeting of long homology arm (~ 1 kb) donors in  
23 combination with restriction enzyme digestion to liberate the linear donor template (Hoshijima et  
24 al., 2016; Shin et al., 2014). Targeted integration of pGTag 24 and 48 bp homology arm 2A-  
25 TagRFP-CAAX-SV40 donors into exon 2 of *connexin43.4 (cx43.4)* resulted in 38-56% and 29-



Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 47% of injected embryos showing broad RFP expression throughout the nervous system and  
2 vasculature (Fig. 2d, f). Increasing the length of the 5' and 3' homology arms to 1 kb did not  
3 significantly change the frequency of targeted integration compared to 24 bp ( $p=0.1693$ ) or 48  
4 bp ( $p=0.6520$ ) (Fig. 2 f), with 26-47% of injected embryos showing the expected neuronal and  
5 vascular RFP expression pattern (Table S1, S2). Injection without the UgrNA leaves the 1 kb  
6 homology circular donor intact and reduced targeting to 0-3% (Fig. 2f Circular HR 1 kb;  
7  $p=0.0067$ ; Table S1, S2), as expected given the low frequency of homologous recombination in  
8 embryos. In comparison to the GeneWeld method, the frequency of RFP expressing embryos  
9 after injection of linear 1 kb homology arm donor template was significantly reduced to 2-9%  
10 (Fig. 2f Linear HR 1kb;  $p=0.0111$ ; Table S1, S2). Together, these results suggest long regions  
11 of homology in the donor template do not enhance the frequency of integration at the genomic  
12 target site, compared with short 24 or 48 bp homology, when using the GeneWeld approach.  
13 The increased efficiency compared to linear donor injection may be that introduction of linear  
14 DNA into the embryo normally leads to concatemers through NHEJ. We propose that  
15 simultaneous targeting of DSBs at the genomic site and in the donor prevents early concatemer  
16 formation and favors homology directed repair.

17

18 The effect of homology length on GeneWeld integration efficiency was also tested with  
19 the 2A-Gal4VP16- $\beta$ -actin3'UTR donor targeting *esama* exon 2. The *esama* gene is expressed  
20 primarily in the vascular system. Increasing homology length from 24 bp to 1 kb dramatically  
21 increased the percentage of RFP positive embryos, from 20-23% to 82-94% ( $p=0.0001$  Fig. S5,  
22 Table S1, S2), but the majority of RFP was not vascular specific, suggesting off-target  
23 expression. A high frequency of RFP-positive embryos was also observed when the donor  
24 template was injected without UgrNA (27-53%) (Fig. S5 Circular HR 1kb) or the donor template  
25 was linearized before injection (83-94%) (Fig. S5 Linear HR 1 kb). Common repetitive elements,  
26 enhancers, or a cryptic promoter in the intronic sequence of the *esama* 1 kb homology arms

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 together may lead to off target integration and ectopic RFP expression. These results  
2 underscore the utility of the GeneWeld short homology approach for simple donor vector  
3 construction and efficient precision targeted integration.

4

### 5 **Efficient germline transmission of GeneWeld precision targeted integration events**

6 Three out of five (60%) *noto-2A-TagRFP-CAAX-SV40* injected founder fish raised to  
7 adulthood transmitted *noto-2A-TagRFP-CAAX-SV40* tagged alleles through the germline (Fig.  
8 3, Table 1, S3 and S4). Although RFP expression in *tyr-2A-Gal4VP16-β-actin* injected  
9 *Tg(UAS:mRFP)<sup>tp12</sup>* embryos was not uniform throughout melanophores, three out of eight RFP  
10 mosaic embryos raised to adulthood transmitted germline tagged alleles (37.5%) (Fig. 3, Table  
11 1, S3 and S4). Similarly, for *esama-2A-Gal4VP16-β-actin*, 12/18 (66.7%) F0s displaying  
12 widespread vasculature RFP expression were raised to adulthood transmitted *esama-2A-*  
13 *Gal4VP16-β-actin* alleles to the F1 generation. We extended the germline transmission analysis  
14 to include 6 additional loci: *flna*, two target sites in *msna* (exon 2 and 6), *aqp1a1*, *aqp8a1*, and  
15 *anxa2a*. Overall, the data reveal a combined transmission rate to the F1 generation of 49%  
16 across all GeneWeld targeted loci (Fig. 3, Table 1, S3 and S4). Taken together, these results  
17 demonstrate the GeneWeld method promotes targeted integration that is efficiently transmitted  
18 through the germline in zebrafish.

19

### 20 **Precise 5' and 3' junctions and single copy integration in F1 germline GTag alleles**

21 We performed Genomic Southern blot analyses and PCR junction fragment sequencing  
22 of F1 GTag germline alleles to determine whether GeneWeld targeting lead to precise  
23 integration at the 5' and 3' sides of the genomic target site. Southern blot analysis and  
24 sequencing of *tyr-2A-Gal4VP16-β-actin* F1 progeny demonstrated a single copy integration of  
25 the Gal4VP16 cassette (Fig. 4 a-c) with precise sequence at both 5' and 3' ends of the

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 integration site (Fig. S6). Analysis of four F1 progeny from two *noto-2A-TagRFP-CAAX-SV40*  
2 founder adults confirmed a single copy integration in *noto* exon 1 (Fig. 4, d-f). However,  
3 sequencing of PCR junction fragments in F1 progeny revealed precise 5' integration but  
4 imprecise integration at the 3' ends that could represent NHEJ repair (Fig. S6).

5  
6 Junction fragment analysis of F1s alleles from 5 additional targeted sites in *esama*, *flna*,  
7 *msna*, *aqp1a1*, and *aqp8a1* revealed precise events were primarily recovered at the 5' for all  
8 genes examined (30/31 or 97% across seven genes) (Fig. S6). This result is expected, since  
9 screening for fluorescent reporter expression from the integrated donor cargo selects for precise  
10 5' integration. For *esama*, the 3' junctions were also precise in 9/10 of the F1s examined from 6  
11 different F0s, and both *aqp1a1* and *app8a1* had precise 3' junctions. This is compared to *msna*  
12 E2 targeting with 2A-Gal4VP16- $\beta$ -actin, where only one out of the 12 F1s examined had a  
13 precise 3' junction. Together, these results indicate that GeneWeld reagents can promote with  
14 high frequency precise single copy integration at a genomic cut site without insertion of donor  
15 vector backbone sequences, although events involving NHEJ at the at the 3' end are also  
16 recovered.

17

## 18 **Homology Engineered to Distal Genomic gRNA Sites Seeds Deletion Tagging in Somatic** 19 **Tissue**

20 To further demonstrate the utility of GeneWeld targeted integration, we tested whether  
21 the pGTag donor could function to bridge two CRISPR/Cas9 genomic cuts, resulting in  
22 simultaneous deletion of endogenous sequences and integration of exogenous DNA to create a  
23 "deletion tagged" allele. The pGTag-2A-Gal4VP16 donor was cloned with 48bp homology arms  
24 to two gRNA target sites in the zebrafish *retinoblastoma1* (*rb1*) gene. Guide RNAs were  
25 designed to sites in exons 2 and 4, which are located 394 bp apart, or exons 2 and 25 which are  
26 separated by ~48.4 kb (Fig. 5a). The 5' homology arm contained sequence upstream of the cut

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 site in exon 2, while the 3' homology arm contained sequence downstream of the cut site in  
2 either exon 4 or exon 25. GeneWeld targeting with the corresponding exon 2-4 or exon 2-25  
3 pGTag-2A-Gal4VP16- $\beta$ -actin donor into Tg(*UAS:mRFP*)<sup>tpl2</sup> embryos resulted in injected  
4 embryos showing broad and ubiquitous RFP expression (Fig. 5b-c). Using the same approach,  
5 we targeted the zebrafish gene *moesina* (*msna*) at exons 2 and 6, located 7.8 kb apart, with 2A-  
6 Gal4VP16 using 48 bp of homology (Fig. 5d), and found RFP expression in a pattern consistent  
7 with the expression of *msna* (Fig. 5 e, e'). The frequency of RFP positive embryos was similar  
8 after targeting *rb1* exon 2-4 (44-78%) and *msna* exon 2-6 (50-85%), and did not seem to be  
9 affected by increasing the size of the deleted region from 394 bp to 48.4 Kb in *rb1* exon 2-25  
10 (49-70%) (Fig. 5f). Somatic junction fragment analysis detected precise integration of 2A-  
11 Gal4VP16 in both genes at the 5' upstream exon (*rb1* 97%; *msna* 85%) and 3' downstream  
12 exon (*rb1* 67%; *msna* 45%) (Fig. S7). However, only one out of 16 (6%) *rb1* e2-25-2A-  
13 Gal4VP16 targeted F0 founders transmitted a precise 5' junction through the germline, but the  
14 3' junction could not be amplified by PCR (Table S3, S4). None of the 10 *msna* e2-e6 2A-  
15 Gal4VP16 targeted F0 zebrafish transmitted a deletion tagged allele to the next generation. In  
16 contrast, targeting 2A-Gal4VP16- $\beta$ -actin to exon 2 or 6 alone resulted in 2 out of 7 F0s  
17 transmitting a targeted allele to the next generation (Table S3, S4).

18

19 Together, these results demonstrate simultaneous targeting of two distal genomic cut  
20 sites can create precise HMEJ integration at both ends of a pGTag reporter cassette in somatic  
21 tissue, but these events are not easily passed through the germline. This was reinforced by  
22 attempting deletion tagging at additional loci, including *kdrl*, *s1pr1*, and *vegfaa*, which showed  
23 32-81% expression in F0 embryos, but no germline transmission to the F1 generation (Table  
24 S2, S3).

25

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

## 1 **Integration of Exogenous DNA Using HMEJ in Porcine and Human Cells is More Efficient** 2 **than HR**

3 To determine if HMEJ integration directed by short homology functions efficiently in large  
4 animal systems, we tested the GeneWeld targeting strategy in *S. scrofa* fibroblasts. The  
5 *ROSA26* safe harbor locus was targeted with a cassette that drives ubiquitous eGFP expression  
6 from the UbC promoter (Fig. 6a-c). GeneWeld reagents, where the genomic sgRNA was  
7 replaced with mRNAs encoding a TALEN pair to generate a genomic DSB in the first intron of  
8 *ROSA26*, were delivered to pig fibroblasts by electroporation. This strategy was compared to  
9 cells electroporated with just the TALEN pair and a HR donor containing approximately 750 bp  
10 of homology flanking the genomic target site. GFP expression was observed in 23% of colonies  
11 using GeneWeld reagents, compared to 2% of colonies using the HR donor with ~750 bp  
12 homology arms. Co-occurring precise 5' and 3' junctions were observed in over 50% of the  
13 GFP+, GeneWeld engineered colonies while none of the GFP+, HR colonies contained both  
14 junctions. Sequencing of junctions from 8 GFP+, GeneWeld engineered colonies that were  
15 positive for both junctions showed precise integration in 7/8 colonies at the 5' junction and 8/8  
16 colonies at the 3' junction.

17  
18 The GeneWeld strategy was also used to target integration of a MND:GFP reporter  
19 (Halene et al., 1999) into the *AAVS1* safe harbor locus in human K-562 cells (Fig. 6d-f).  
20 Integrations were attempted with either GeneWeld reagents or an HR donor targeting the  
21 *AAVS1* cut site by electroporation of K-562 cells. Cells were FACs sorted by GFP at day 14  
22 following electroporation. With GeneWeld reagents over 50% of cells were GFP positive,  
23 compared to only 6% of cells electroporated with the HR donor. This suggests the GeneWeld  
24 strategy promoted efficient integration and stable expression of the MND:GFP cassette at the  
25 *AAVS1* locus (Fig. S8). Expression was maintained over 50 days, and 5' precise junction  
26 fragments were observed following PCR amplification in bulk cell populations (Fig. S9). The

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 results above demonstrate that the GeneWeld strategy outperforms traditional HR techniques in  
2 mammalian cell systems and is effective without antibiotic selection.

3

#### 4 **Discussion**

5

6 The results described here demonstrate the utility of short homology-based gene  
7 targeting for engineering precise integration of exogenous DNA and expand the potential of  
8 efficient tagging to diverse loci with differing endogenous expression levels. Key to our strategy  
9 was to observe reporter gene expression that correlated with precise integration (Fig. 2 and Fig.  
10 S1) and selecting these animals for examination of germline transmission to the F1 generation.  
11 However, we have observed that restricted patterns of gene expression such as *noto* do not  
12 necessarily correlate with the degree of precise integration throughout the embryo following  
13 injection (data not shown). Moreover, we show that using short homology to bridge distal ends  
14 together simultaneously creates a deletion and a reporter integration, however, these events are  
15 not easily passed through the germline. This emphasizes that the events observed in the  
16 somatic tissue of the zebrafish embryo following injection do not necessarily predict  
17 transmission through the germline. We demonstrate efficient integration of cargos up to  
18 approximately 2 kb in length in zebrafish, pig fibroblasts, and human cells, and in zebrafish,  
19 these events can be passed through the germline in nearly 50% of the injected embryos.  
20 Both CRISPR/Cas9 and TALENs are effective as GeneWeld genomic editors,  
21 providing flexibility in deployment and genome-wide accessibility.

22

23 Several components of the GeneWeld strategy may lead to enhanced somatic and  
24 germline targeting efficiencies in zebrafish as compared to previous reports (Hisano et al.,  
25 2015). Canonical NHEJ is highly active during rapid cell divisions in early zebrafish  
26 embryogenesis (Bedell et al., 2012). However, given the correct sequence context surrounding

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 the dsDNA break, MMEJ is the preferred method of non-conservative repair (Ata et al., 2018;  
2 He et al., 2015; Kent et al., 2015). GeneWeld homology arms are rationally designed based on  
3 the known homology searching activity of RAD51 and strand annealing activity of RAD52  
4 (Conway et al., 2004; Singleton et al., 2002). In our experiments at *noto*, gene targeting is  
5 significantly reduced when 48 bp homology arms are altered by 1 bp to 47 or 49 bp (Fig. S3).  
6 This suggests that optimal short homology arms should be designed in groups of 3 and/or 4 bp  
7 increments. We are currently testing this hypothesis further. Additionally, Shin et al., 2014  
8 showed the highest rates of somatic targeting when their donor was linearized *in vitro* inside a  
9 ~1 kb 5' homology arm, leaving 238 bp of homology flanking the knock-in cargo. Thus, it is  
10 tempting to speculate that gene targeting in these experiments proceeded not through HR, but  
11 through other related HMEJ DNA repair pathway more similar to the findings presented here.

12

13         Using GeneWeld reagents we observed that the length of the homology arm does not  
14 influence the observed somatic frequency of reporter gene expression when examined between  
15 24 bp and 1 kb in length. However, there are distinct disadvantages to using long homology  
16 arms in experimental design and execution. First, longer homology arms often require the use of  
17 intronic sequences which are highly enriched of repetitive sequences that can potentially  
18 mediate random integrations across the genome. Second, alternative promoters and enhancer  
19 sequences present in intronic regions can also affecting the efficiency of reporter gene  
20 expression, as we likely observed with the *esama* gene. Lastly, cloning of the homology arms  
21 can be arduous, since intronic sequences are highly variable across individuals, and therefore, it  
22 is recommended that the longer homology arms are cloned from individuals used for gene  
23 targeting. Taken together, we suggest that using the GeneWeld method with short homology  
24 arms is a preferable choice for efficient precise exogenous DNA integration.

25

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1           The dramatic shift of DNA repair at genomic DSBs from cis-NHEJ to trans-HMEJ using  
2 GeneWeld donors likely also influences enhanced editing of the germline. Across all zebrafish  
3 experiments with germline transmission, 49% of founders transmitted tagged alleles, with 17.4%  
4 of gametes carrying the edited allele of interest (Table S3, S4), demonstrating decreased  
5 germline mosaicism and increased germline transmission from previous reports. Given that our  
6 somatic knock-in and germline transmission rates are higher than most published reports, we  
7 conclude that GeneWeld is a more effective homology-based method for generating precisely  
8 targeted knock-in alleles in zebrafish.

9  
10           While targeting *noto* with 5' only homology shows an increase in targeting efficiency with  
11 longer homology, increasing homologies on both ends of the cargo DNA did not increase  
12 targeting efficiency (Fig. 2e). Positive events are selected only by fluorescently tagged alleles,  
13 indicating precise 5' integration patterns. We speculate that inclusion of homology at the 3' end  
14 of our cargo creates competition for the donor DNA ends, as not all editing events are precise at  
15 both 5' and 3' junctions (Fig. 4 and Fig. S4). Thus, it is conceivable that precise events at the 3'  
16 end could preclude precise integration at the 5' end during some editing events, and vice versa.  
17 It is tempting to speculate that this data hints at synthesis dependent strand annealing (SDSA)  
18 as a possible DNA repair mechanism for pGTag donor integration (Ceccaldi et al., 2016). After  
19 strand invasion using either of the homology domains and replication through the reporter,  
20 second DNA end capture may abort before or after replication through the opposing homology  
21 domain, resulting in imprecision, as greater than or equal to 150 bp is required for proper  
22 second end capture in yeast (Mehta et al., 2017). Experiments to address this hypothesis by  
23 varying homology arm lengths flanking the donor cassettes and including negative selection  
24 markers are of note for future work in determining the genetic mechanisms that promote  
25 efficient integration.

26



Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1           Timing and turnover of Cas9 during the genomic editing event can influence cut  
2 efficiency and somatic mosaicism/germline transmission rates (Clarke et al., 2018; Zhang et al.,  
3 2018), increasing the interest of using RNP during all precision gene editing applications.  
4 However, we were unable to observe fluorescence following injection of GeneWeld components  
5 with RNPs or detect targeted integrations at a high frequency (unpublished data). We  
6 hypothesize this is due to Cas9 and UgrRNA locating and binding to the UgrRNA sites on the  
7 pGTag donors during dilution of the injection mixture. This heteroduplex either activates DSBs  
8 on the donor *in vitro*, or directly after injection, before the genomic gRNA can locate and cut the  
9 genome. Thus, the stochastics of DNA end availability are altered using RNPs and integration  
10 activity is greatly reduced. Injection of the GeneWeld plasmid donor and RNPs in separate  
11 injection mixtures does not produce integration in zebrafish embryos (unpublished data).  
12 Further experiments could address these limitations through the use of inducible nuclease  
13 systems.

14

15           Targeting genes with lethal phenotypes, such as tumor suppressors or other genes  
16 required for embryogenesis, is of interest to the zebrafish community. However, using  
17 fluorescence to screen for targeted events can be misleading. For example, the RFP signal is  
18 dramatically reduced or lost upon biallelic inactivation of *noto*, likely when notochord cells  
19 translocate to muscle cells (Melby et al., 1996; Talbot et al., 1995). Additionally, though deletion  
20 tagging using two target sites in the genome seems to be robust in somatic tissue (Fig. 5),  
21 germline transmission of deletion tags is rare. This suggests that edited germ cells may be lost  
22 to apoptosis due to the additional cut in the genome, or that heterozygous deletion tagged  
23 alleles are recognized during homologous chromosome pairing and are repaired or lost as germ  
24 cells mature. Similar susceptibility of stem cells to apoptosis following gene editing has been  
25 previously observed (Ihry et al., 2018; Li et al., 2018). In both of these cases, it may be

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 necessary to modulate GeneWeld reagent concentrations in order to avoid biallelic inactivation  
2 of the genomic target, or to ensure homozygous deletion tagging.

3  
4       Amplification of the fluorescence signal using GAL4/VP16 allowed us to target several  
5 genes for which we did not observe a fluorescence report from integration of a fluorescent  
6 protein directly. While this approach is advantageous for selecting correctly targeted embryos to  
7 examine for germline integration, GAL4/VP16 may have toxic effects as reported previously  
8 (Ogura et al., 2009). For example, we found dominant phenotypes in the F1 generation for both  
9 *msna* and *flna* which could reflect toxicity from high levels of expression of GAL4/VP16.

10 Alternatively, these gene could also display haploinsufficiency or express a partial protein  
11 product that functions in a dominant manner. Heterozygous *msna* mutants targeting exon 5 in  
12 the F1 generation display phenotypes similar to morpholino targeting of this gene (Wang et al.,  
13 2010) (data not shown), suggesting haploinsufficiency or a dominant negative peptide is a likely  
14 explanation.

15  
16       GeneWeld is also an effective strategy to precisely control exogenous DNA integration in  
17 mammalian cell lines. While our data shows an approximate 10-fold increase in targeted  
18 integration using 48 bp of homology to drive HMEJ versus HR, Zhang et al. (2017) concluded  
19 that targeted integration did not appreciably increase until homology arms of ~600 bp were used  
20 (Zhang et al., 2017). However, this could reflect differences in the experimental design or cell  
21 types used and suggest different DNA repair pathways may be more prevalent in certain  
22 conditions. Deciphering the DNA repair pathway used for HMEJ in zebrafish and mammalian  
23 cells is paramount to increasing editing efficiencies in basic research and for gene therapy.

24  
25       Given the high efficiency and precision of GeneWeld, additional applications to efficiently  
26 introduce other gene modifications, such as single or multiple nucleotide polymorphisms, by

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 exon or gene replacement is possible using the deletion tagging method. Further, GeneWeld  
2 could be used to create conditional alleles by targeting conditional gene break systems into  
3 introns (Clark et al., 2011). In conclusion, our suite of donor vectors with validated integration  
4 efficiencies, methods, and web interface for pGTag donor engineering will serve to streamline  
5 experimental design and broaden the use of designer nucleases for homology-based gene  
6 editing at CRISPR/Cas9 and TALEN cut sites in zebrafish. We also demonstrate an advanced  
7 strategy for homology-based gene editing at CRISPR/Cas9 and TALEN cut sites in mammalian  
8 cell lines. Our results open the door for more advanced genome edits in animal agriculture and  
9 human therapeutics.

10

## 11 **Methods**

12

### 13 **Contact for reagent and resource sharing**

14 Further information and requests for resources and reagents should be directed  
15 to and will be fulfilled by the Lead Contact, Jeffrey Essner ([jessner@iastate.edu](mailto:jessner@iastate.edu)).

16

### 17 **Experimental model and subject details**

18 Zebrafish were maintained in Aquatic Habitats (Pentair) housing  
19 on a 14 hour light/10 hour dark cycle. Wild-type WIK were obtained from the Zebrafish  
20 International Resource Center. The  $Tg(miniToI2/14XUAS:mRFP, \gamma Cry:GFP)^{tp12}$ , shortened  
21 to  $Tg(UAS:mRFP)^{tp12}$ , was previously described (Balciuniene et al., 2013). All experiments were  
22 carried out under approved protocols from Iowa State University IACUC.

23

24 The human K-562 chronic myelogenous leukemia cell line (ATCC CCL-243) used in  
25 gene targeting experiments was cultured at 37°C in 5% CO<sub>2</sub> in RPMI-1640 (Thermo Fisher

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 Scientific) supplemented with 10% fetal bovine serum (FBS) and  
2 Penicillin/Streptomycin. Electroporation was conducted with  $1.5 \times 10^5$  cells in a 10  $\mu$ l tip using  
3 the Neon electroporation device (Thermo Fisher Scientific) with the following conditions: 1450V,  
4 10ms, 3x pulse. Nucleic acid dosages were as follows: 1.5  $\mu$ g Cas9 mRNA  
5 (Trilink Biotechnologies), 1  $\mu$ g each chemically modified sgRNA (Synthego), and 1  $\mu$ g donor  
6 plasmid.

7  
8 Fibroblasts were cultured in DMEM (high glucose) supplemented to 10% vol/vol FBS, 20  
9 mM L-glutamine and 1X Pen/ Strep solution and transfected using the Neon™ system  
10 (Invitrogen). Briefly,  $1 \times 10^6$  fibroblasts were transfected with 1  $\mu$ g of polyadenylated ROSA  
11 TALEN mRNA, 1  $\mu$ g of universal gRNA mRNA, 1  $\mu$ g of polyadenylated Cas9 mRNA and 1  $\mu$ g of  
12 donor plasmid. Transfected cells were cultured for 3 days at 30°C before low density plating,  
13 extended culture (10 days) and colony isolation. Individual colonies were aspirated under  
14 gentle trypsinization, replated into 96- well plates and cultured for 3-4 days.

## 16 **pGTag series vectors**

17 To build the pGTag vector series, 2A-TagRFP, 2A-eGFP, and 2A-Gal4/VP16  
18 cassettes were assembled from a 2A-TagRFP-CAAX construct, p494. To clone  
19 the eGFP cassette, the plasmid p494 was amplified with primers F-p494-XhoI and R-p494-SpeI  
20 to generate unique enzyme sites in the backbone. The eGFP coding  
21 sequence (Clontech Inc.) was amplified with the primers F-eGFP-SpeI and R-eGFP-XhoI to  
22 generate the corresponding enzyme sites on the eGFP coding sequence. Fragments  
23 were digested with SpeI-HF and XhoI (NEB) and following column purification with the Qiagen  
24 miniprep protocol, were ligated to the plasmid backbone with T4 ligase (Fisher).

25

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1           The Gal4/VP16 coding sequence and zebrafish  $\beta$ -actin 3' untranslated region was  
2 amplified from vector pDB783 (Balciuniene et al., 2013) with primers F-2A-Gal4-BamHI and R-  
3 Gal4-NcoI to add a 2A peptide to the 5' end of the Gav4Vp16 cDNA. The resulting PCR product  
4 was then cloned into the intermediate Topo Zero Blunt vector (Invitrogen) and used for  
5 mutagenesis PCR with primers F and R 'gal4-Ecofix' to disrupt the internal EcoRI restriction  
6 site. The resulting Gal4/VP16 sequence was cloned into the BamHI and NcoI sites in the p494  
7 backbone.

8  
9           The 5' universal/optimal guide site and *lacZ* cassette were added to pC-2A-TagRFP-  
10 CAAX-SV40, pC-2A-eGFP-SV40, and pC-2A-Gal4VP16- $\beta$ -actin with the following steps.  
11 The *lacZ* was first amplified with primers F-*lacZ* and R-*lacZ*, which add the type IIS enzyme  
12 sites to either end of the *lacZ*. The resulting PCR product was then cloned into an intermediate  
13 vector with the Zero Blunt® TOPO® PCR Cloning Kit (Invitrogen). This intermediate was used  
14 as a template in a nested PCR to add the Universal  
15 guide sequence GGGAGGCGTTTCGGGCCACAGCGG to the end of the *lacZ*  
16 sequence. The nested PCR used primers F-*lacZ*-universal-1 and R-*lacZ*-universal-BamHI to  
17 add the first part of the universal guide to one end and a BamHI site to the other. This was used  
18 as template for PCR with the primers F-*lacZ*-universal-EcoRI and R-*lacZ*-universal-BamHI to  
19 add the remainder of the universal guide and an EcoRI site. The fragment was column purified  
20 as above, digested with EcoRI-HF and BamHI-HF and cloned into the appropriate sites in the  
21 three vectors.

22  
23           The 3' universal guide and type 2 restriction enzyme sites were cloned into each  
24 vector in two steps. A segment from a Carp beta-actin intron containing a 99 bp spacer flanked  
25 by two BspQI sites was amplified using the primers F-3'-uni-1 and R-3'-uni-1 to add the

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 universal site to one side of the spacer. This product was column purified as above and used as  
2 template for the second amplification with primers F-3'-uniNco1 and R-3'-uniEagI to add cloning  
3 sites. This product was column purified and cloned using the Topo zero blunt kit. This  
4 intermediate was digested with NcoI-HF and EagI, and the BspQI fragment purified and cloned  
5 into the three vectors as above. Ligations were grown at 30°C to reduce the possibility of  
6 recombination between the two universal guide sites.

7  
8 Correct clones for pU-2A-TagRFP-CAAX-U, pU-2A-eGFP-U, and pU-2A-Gal4/VP16-  
9 U were selected and used as template for mutagenesis PCR with KOD to remove  
10 extra BspQI sites from the backbone with primers F/R-BBfix, digested  
11 with DpnI (NEB), and ligated with T4 ligase. A correct pU-2A-TagRFP-CAAX-U clone was used  
12 as template for PCR with F/R-TagRFPfix to interrupt the BspQI site in the TagRFP coding  
13 sequence as above. A correct clone of pU-2A-Gal4/VP16-U was selected and used as template  
14 with primers F/R-Bactfix to remove the BspQI site in the Beta-actin terminator, the product was  
15 re-cloned as above. All constructs were sequence verified.

16

### 17 **Homology arm design and donor vector construction**

18 For detailed methods, see Supplementary gene targeting protocol. In brief,  
19 homology arms of specified length directly flanking a genomic targeted double strand  
20 break were cloned into the pGTag vector, in between the UgrNA sequence and the cargo.  
21 A three nucleotide buffer sequence lacking homology to the genomic target site was engineered  
22 between the donor UgrNA PAM and the homology arms, in order to maintain the specified  
23 homology arm length. To generate 1kb homology arms for zebrafish genes *cx43.4* and *esama*,  
24 ~2kb of genomic DNA surrounding the CRISPR target site was PCR amplified from adult WIK  
25 finclips using the proofreading enzyme KOD, and then sequenced to identify polymorphisms.

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 pGTag 1kb homology arm vectors were injected into embryos from adults with the matching  
2 genomic sequence. See Table S4 for all homology arms, gRNA target sites, and spacers.

3

#### 4 **Zebrafish embryo injection**

5 pT3TS-nCas9n was a gift from Wenbiao Chen (Addgene plasmid #  
6 46757). XbaI linearized pT3TS-nCas9n was purified under RNase-free conditions with the  
7 Promega PureYield Plasmid Miniprep System. Linear, purified pT3TS-nCas9n was used as  
8 template for *in vitro* transcription of capped, polyadenylated mRNA  
9 with the Ambion T3TS mMessage mMachine Kit. mRNA was purified using  
10 Qiagen miRNeasy Kit. The genomic and universal sgRNAs were generated using cloning  
11 free sgRNA synthesis as described in (Varshney et al., 2015) and purified using  
12 Qiagen miRNeasy Kit. Donor vector plasmid DNA was purified with the  
13 Promega PureYield Plasmid Miniprep System. *noto*, *cx43.4*, *tyrosinase*, and *moesina*, were  
14 targeted by co-injection of 150 pg of nCas9n mRNA, 25 pg of genomic sgRNA,  
15 25 pg of UsgRNA (when utilized), and 10 pg of donor DNA diluted in RNase free  
16 ddH<sub>2</sub>O. The *rb1* targeting mixture contained 300 pg nCas9n mRNA. 2 nl was delivered to each  
17 embryo.

18

#### 19 **Recovery of zebrafish germline knock-in alleles**

20 Injected animals were screened for fluorescence reporter expression on a Zeiss  
21 Discovery dissection microscope and live images captured on a Zeiss LSM 700 laser scanning  
22 confocal microscope. RFP/GFP positive embryos were raised to adulthood and outcrossed to  
23 wildtype WIK adults to test for germline transmission of fluorescence in F1  
24 progeny. *tyr*, *esama*, *rb1* and *msna* embryos targeted with Gal4VP16 were crossed  
25 to *Tg(UAS:mRFP)<sup>tp12</sup>*.

26

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

## 1 **DNA isolation and PCR genotyping**

2 Genomic DNA for PCR was extracted by digestion of single embryos in 50mM NaOH at  
3 95°C for 30 minutes and neutralized by addition of 1/10<sup>th</sup> volume 1M Tris-HCl pH 8.0. Junction  
4 fragments were PCR-amplified with primers listed in Table S6 and the products TOPO-  
5 TA cloned before sequencing.

6

## 7 **Southern blot analysis**

8 Genomic Southern blot and copy number analysis was performed as described  
9 previously (McGrail et al., 2011). PCR primers used for genomic and donor specific probes are  
10 listed in Table S6.

11

## 12 **Junction fragment analysis in pig fibroblasts**

13 Individual colonies were scored for GFP expression and prepared for PCR by washing  
14 with 1X PBS and resuspension in PCR-safe lysis buffer (10 mM Tris-Cl, pH 8.0; 2 mM EDTA;  
15 2.5% (vol/vol) Tween-20; 2.5% (vol/vol) Triton-X 100; 100 µg/mL Proteinase K followed by  
16 incubation at 50°C for 60 min and 95°C for 15 min. PCR was performed using  
17 1X Accustart Supermix (Quanta) with the primers: 5' junction F-5'  
18 TAGAGTCACCCAAGTCCCGT-3', R-5'- ACTGATTGGCCGCTTCTCCT-3'; 3' junction F-5'-  
19 GGAGGTGTGGGAGGTTTTT-3', R-5'- TGATTTTCATGACTTGCTGGCT-3'. ROSA TALEN  
20 sequences are: TAL FNG NI NI HD HD NG NN NI NG NG HD NG NG NN NN; TAL RHD NN  
21 NG NI HD NI HD HD NG NN HD NG HD NI NI NG.

22

## 23 **K-592 Flow Cytometry**

24 K-562 cells were assessed for GFP expression every 7 days for 28 days following  
25 electroporation. Flow cytometry was conducted on an LSRII instrument (Becton Dickinson) and  
26 data was analyzed using FlowJo software v10 (Becton Dickinson). Dead cells were excluded



Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 from analysis by abnormal scatter profile and exclusion based on Sytox Blue viability dye  
2 (Thermo Fisher Scientific).

3  
4 Junction PCR to detect targeted integration was conducted using external genomic  
5 primers outside of the 48bp homology region and internal primers complementary to the  
6 expression cassette. PCR was conducted using Accuprime HIFI Taq (Thermo Fisher Scientific).  
7 PCR products from bulk population were sequenced directly.

8

### 9 **Quantification and statistical analysis**

10 Statistical analysis was performed using GraphPad Prism software. Data plots represent  
11 mean +/- s.e.m. of n independent experiments, indicated in the text. *p* values were calculated  
12 with two-tailed unpaired *t*-test. Statistical parameters are included in the Figure legends.

13

### 14 **Data and software availability**

15 The webtool **GTagHD** was developed to assist users in designing oligonucleotides for  
16 targeted integration using the pGTag vector suite. GTagHD guides users through  
17 entering: 1) the guide RNA for cutting their cargo-containing plasmid; 2) the guide RNA for  
18 cutting their genomic DNA sequence; (3) the genomic DNA sequence, in the form of  
19 a GenBank accession number or copy/pasted DNA sequence; and 4) the length of  
20 microhomology to be used in integrating the plasmid cargo. If the user is utilizing one of  
21 the pGTag series plasmids, GTagHD can also generate a GenBank/ApE formatted file for that  
22 plasmid, which includes the user's incorporated oligonucleotide sequences. GTagHD is freely  
23 available online at <http://genesulpt.org/gtaghd/> and for download at [https://github.com/Dobbs-](https://github.com/Dobbs-Lab/GTagHD)  
24 [Lab/GTagHD](https://github.com/Dobbs-Lab/GTagHD).

25

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 **Key resources table**

2

<b>REAGENT or RESOURCE</b>	<b>SOURCE</b>	<b>IDENTIFIER</b>
<b>Antibodies</b>		
<b>Bacterial and Virus Strains</b>		
NEB Stable Competent <i>E. coli</i>	NEB	CAT#C30401
One Shot TOP10 Chemically Competent <i>E. coli</i>	ThermoFisher	CAT#C404010
<b>Biological Samples</b>		
<b>Chemicals, Peptides, and Recombinant Proteins</b>		
<b>Critical Commercial Assays</b>		
pCR4 TOPO TA Cloning Kit for Sequencing	ThermoFisher	CAT#K457502
<b>Deposited Data</b>		
<b>Experimental Models: Cell Lines</b>		
Human: K562 cell line	ATCC	CAT# ATCC CCL-243
Porcine: Fibroblast cell line	Recombinetics, Inc.	Recombinetics, Inc.
<b>Experimental Models: Organisms/Strains</b>		
Zebrafish: WIK strain	ZIRC	CAT#ZL84
Zebrafish: pDB790	Balciunas Lab	Upon request
<b>Oligonucleotides</b>		
See Table S6 for all primers used	This paper	N/A
<b>Recombinant DNA</b>		
p494-2a-TagRFP-CAAX-SV40	This paper	Deposited at AddGene
pGTag-2A-TagRFP-CAAX-SV40	This paper	Deposited at AddGene
pGTag-2A-Gal4VP16-bactin	This paper	Deposited at AddGene
pGTag-2A-eGFP-SV40	This paper	Deposited at AddGene
<b>Software and Algorithms</b>		
GTagHD	This paper	<a href="https://github.com/Dobbs-Lab/GTagHD">https://github.com/Dobbs-Lab/GTagHD</a>
<b>Other</b>		

3

4

5

6

7

8

9

10

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22

**Acknowledgements**

This work was supported by NIH grants R24OD020166 (JJE, MM, DLD, KJC, SCE), GM088424 (JJE), and GM63904 (SCE).

**Author Contributions**

WAW, JMW, MM and JJE conceived the study; WAW, JMW, MM and JJE wrote the manuscript with input from SCE and KJC; WAW, JMW, MPA, MET, KCM, ML, ZM, AW and JAH designed and performed the zebrafish experiments; JMW and TJW designed and created the vector suite with input from WAW, KJC, SCE, KMK, C-BC, DB, MM, and JJE; CMM created the web interface with input from JMW, WAW, CSM and DLD; SLS, DAW, MKV and DFC designed and performed the pig fibroblast experiments. BSM and BRW designed and performed the human cell line experiments.

**Declaration of interests**

JJE, MM, SLS, KJC, DW, and DFC have a financial conflict of interest with Recombinetics, Inc.; JJE and SCE with Immusoft, Inc.; JJE, MM, WAW, KJC and SCE with ForgeBio and LifEngine and LifEngine Animal Technologies; BSM and BRW with B-MoGen Biotechnologies, Inc.

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

## 1 **Figure Legends**

2

3 **Figure 1. GeneWeld strategy and pGTag vector series** (a) GeneWeld reagent components  
4 are designed for simultaneous genome and donor nuclease targeting to reveal short regions of  
5 homology. Red arrowheads represent *in vivo* designer nuclease DSBs. Components include: 1 -  
6 Designer nuclease mRNA, either Cas9 to target both the genome and donor, or Cas9 to target  
7 the donor and TALEN to cut the genome; 2 - sgRNA for targeting Cas9 to genome; 3 - Universal  
8 sgRNA to liberate donor cargo and homologous ends; and 4 - pGTag donor of interest with  
9 short homology arms. (b) Type IIs restriction endonucleases *BfuAI* and *BspQI* create  
10 incompatible ends outside of their recognition sequence, allowing digestion and ligation of both  
11 homology arms into the vector in a single reaction. Homology arm fragments are formed by  
12 annealing complementary oligonucleotides to form dsDNA with sticky ends for directional  
13 cloning into the vector. XFP = Green or Red Fluorescent Protein. pA = SV40 or  $\beta$ -actin 3'  
14 untranslated region. Red and green fluorescent proteins were cloned into the pGTag vectors,  
15 and for each color, subcellular localization sequences for either nuclear localization (NLSs) and  
16 membrane localization (CAAX) are provided (c) Schematic of GeneWeld targeting *in vivo*. After  
17 designer nuclease creates targeted double-strand breaks in the genome and donor, end  
18 resection likely precedes homology recognition and strand annealing, leading to integration of  
19 the donor without vector backbone. Red arrowheads represent *in vivo* designer nuclease DSBs.

20

21 **Figure 2. HMEJ strategy promotes efficient somatic targeting of knock-in cassettes in**  
22 **zebrafish.** (a-d) Live confocal images of F0 injected embryos showing fluorescent reporter  
23 expression after GeneWeld targeted integration. *noto-2A-eGFP-SV40* at mid somite stage (a,  
24 a'). *tyr-2A-Gal4VP16- $\beta$ -actin;Tg(UAS:mRFP)<sup>tp12</sup>* at 5 days post fertilization (dpf) (b, b'). *esama-*  
25 *2A-Gal4VP16- $\beta$ -actin;Tg(UAS:mRFP)<sup>tp12</sup>* at 2 dpf (c) and 3 dpf (c'). *cx43.4-2A-tagRFP-CAAX-*  
26 *SV40* at 31 hours post fertilization (d, d'). (e) Frequency of embryos with reporter gene

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 expression following GeneWeld targeting at *noto*, *tyr* and *esama*. 5' and 3' homology lengths  
2 flanking donor cargos indicated in base pairs as 24/24 or 48/48. (f) Comparison of the frequency  
3 of RFP expressing embryos after targeting *cx43.4* exon 2 using GeneWeld 24/24 bp homology,  
4 GeneWeld 48/48 bp homology, Geneweld 1kb/1kb homology, Circular HR 1kb/1kb (injection did  
5 not include UgRNA, \* $p=0.0067$ ), Linear HR 1kb/1kb (donor was linearized before injections,  
6 \* $p=0.0111$ ). Data represents mean +/- s.e.m. of 3 independent targeting experiments.  $p$  values  
7 calculated using Students  $t$  test. Scale bars, 100  $\mu\text{m}$ .

8

9 **Figure 3. Live confocal images of F1 zebrafish with inherited germline alleles of**

10 **integrated GTag reporters.** (a, a') *Tg(noto-2A-TagRFP-CAAX-SV40)* embryo at mid somite  
11 stage showing expression in the notochord and floor plate. (b, b') *Tg(tyr-2A-Gal4Vp16- $\beta$ -actin)*  
12 displaying expression in the melanocytes in a 5 dpf larva. (c, c') *Tg(esama-2A-Gal4Vp16- $\beta$ -*  
13 *actin)* larva showing expression in the vascular system at 4 dpf. (d, d') *Tg(flna-2A-Gal4VP16- $\beta$ -*  
14 *actin)* embryo at 1 dpf showing widespread expression. (e, e' and f, f') *Tg(msna-2A-Gal4VP16-*  
15  *$\beta$ -actin)* targeted to either exon 2 or exon 6 showed expression in the central nervous system  
16 and vasculature at 2 dpf. (g, g' and h, h') *Tg(aqp1a1-2A-Gal4VP16- $\beta$ -actin)* and *Tg(aqp8a1-2A-*  
17 *Gal4VP16- $\beta$ -actin)* display expression in the trunk and tail vasculature at 2 dpf. All images are  
18 lateral views, and the *Gal4VP16* integrations have *Tg(UAS:mRFP)<sup>tp12</sup>* in the background for  
19 visualization of expression. Scale bars are 100  $\mu\text{m}$ .

20

21 **Figure 4. Molecular analysis of F1 GeneWeld GTag targeted alleles at *tyr* and *noto*.** (a-c)

22 Molecular analysis of *Tg(tyr-2A-GAL4/VP16)* F1 offspring from a single targeted F0 founder. (a)  
23 Schematic of expected integration pattern for *tyr* targeted with pGTag-2A-GAL4/VP16. 148 bp  
24 *tyr* probe in Exon 3 and 583bp probe in GAL4/VP16 are indicated. (b) GAL4/VP16 and (c) *tyr*  
25 probed Southern blots of genomic DNA from wild type (WIK) and 4 individual GAL4/VP16

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 positive F1s. The expected 7400 bp band is detected with both probes, suggesting a single  
2 copy integration. (d-f) *Tg(ното-2A-RFP)* F1 targeted integration alleles from 2 independent F0  
3 founders. (d) *ното* gene model with location of restriction enzymes used for genomic Southern  
4 blot analysis. Location of the 513 bp *ното* probe is indicated (dark lines). The predicted and  
5 recovered alleles are shown. (e) Southern blots of F1 *Tg(ното-2A-RFP)* individuals hybridized  
6 with RFP probe. F1 from founder F0#1 contain a ~2100 bp band corresponding to integration  
7 plus deletion of ~400 bp in *ното*. F1 progeny from founder F0#2 show two bands: a ~3700 bp  
8 band corresponding to integration of the reporter plus 2000 bp of vector backbone, and a ~1500  
9 bp band which may represent an off-target integration. Loading controls (10, 1) correspond to  
10 10 copies or 1 copy of RFP containing plasmid. WIK, wild type control DNA. (f) Southern blot in  
11 (d) stripped and re-hybridized with the *ното*-specific probe. A 1342 bp band representing the  
12 wild type allele was detected in all individuals. The integration allele in F1s from F0 #1 was not  
13 detected due to deletion of the region containing the probe. F1s from F0 #2 contain the ~3700  
14 bp band corresponding to the *ното-2A-RFP* integration allele.

15

16 **Figure 5. Deletion tagged alleles created with the GeneWeld strategy in zebrafish somatic**

17 **tissue.** (a) Schematic for *Gal4VP16* reporter integration to tag a deletion allele of *rb1* exons 2-4

18 (top) and *rb1* exons 2-25 (bottom). Arrowheads designate CRISPR/Cas9 DSBs. CRISPR

19 gRNAs in two exons are expected to excise the intervening genomic DNA. The targeting vector

20 contains a 5' homology arm flanking the upstream exon target site and a 3' homology arm

21 flanking the downstream exon target site. (b-c) Live confocal image of F0 *Tg(UAS:mRFP)<sup>tp12</sup>*

22 embryo after *2A-Gal4VP16* deletion tagging *rb1* exons 2-4 (b, b') and *rb1* exons 2-25 (c, c'). (d)

23 Schematic for *2A-Gal4VP16* deletion tagging of *msna* exons 2-6. (e, e') Live confocal image of

24 F0 *Tg(UAS:mRFP)<sup>tp12</sup>* embryo after *2A-Gal4VP16* deletion tagging at *msna* exons 2-6. (f)

25 Somatic reporter efficiency of targeted deletion tagging using 48 bp homology arms for *rb1*

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 exons 2-4, *rb1* exons 2-25, and *msna* exons 2-6. Data represents mean +/- s.e.m. of 4 (*rb1*) and  
2 5 (*msna*) independent targeting experiments. Scale bars 200  $\mu\text{m}$  (b, c, c', e); 100  $\mu\text{m}$  (b', e').

3

4 **Figure 6. HMEJ-based targeted integration with UgRNA-based vectors promotes efficient**

5 **knock-in in porcine fibroblasts and human K-562 cells.** (a) Strategy for integration using

6 HMEJ and HR donors into intron 1 of *S. scrofa ROSA26* locus. Arrowheads CRISPR/Cas9 (for

7 HMEJ donor) and TALEN (genome) DSBs. (b) Targeting efficiency of the HMEJ donor vs the

8 HR donor as reported by GFP positive colonies out of total colonies. (c) Percent of GFP positive

9 colonies analyzed containing properly sized junction fragments, comparing HMEJ and HR

10 donors. Data are from three independently targeted cell populations. Data represents mean +/-

11 s.e.m. of 3 independent targeting experiments. (d) Diagram of HR and HMEJ strategies for

12 targeted integration of a MND:GFP reporter cassette into the human *AAVS1* locus. (e) Flow

13 cytometry analysis of GFP expression 14 days post-electroporation for each targeting modality:

14 HR (left), HMEJ without universal sgRNA (middle), and HMEJ with universal sgRNA (right).

15 Stable gate was drawn to measure the uniformly expressing population formed by targeted

16 integration and was set based on episome only controls. (f) Quantitation of stable GFP

17 expressing population as measured by flow cytometry at day 14. Data are from three

18 independently targeted cell populations. Data represents mean +/- s.e.m. of 3 independent

19 targeting experiments. p values calculated using two-tailed unpaired t-test.

20

21

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

1 **Tables**

2 **Table 1. Germline transmission of zebrafish GeneWeld GTag integrations.**

3

Genomic target	Exon	Homology	F0 expression	F0s transmitting	Germline transmission percentage
<i>noto</i>	E1	24/24	24%	3/5	60%
<i>tyr</i>	E4	24/24	64%	3/8	38%
<i>cx43.4*</i>	E2	24/24	50%	0/1	0%
<i>cx43.4*</i>	E2	48/48	38%	0/4	0%
<i>esama</i>	E4	24/24	21%	12/18	67%
<i>flna</i>	E4	48/42	100%	3/4	75%
<i>msna</i>	E2	48/48	55%	1/4	25%
<i>msna</i>	E6	48/48	26%	1/3	33%
<i>aqp1a1</i>	E1	48/48	4%	2/9	22%
<i>aqp8a1</i>	E1	48/48	14%	1/1	100%
<i>anxa2a<sup>^</sup></i>	E3	48/48	35%	4/4	100%
			<b>Total</b>	30/61	49%

4  
5 F0's raised to adulthood were outcrossed and screened for germline transmission of properly  
6 localized fluorescent signal. F0s transmitting/F0s outcrossed x 100 = Germline transmission  
7 percentage. F0s were considered screened if 75 F1 embryos were examined for fluorescence.  
8 \*Other experiments showed *cx43.4* alleles could be transmitted through the germline in 3 out 11  
9 F0 fish or 27% with a similar vector (data not shown). TALEN indels of *cx43.4* result in sex  
10 determination defects, suggesting germline defects could contribute to variable frequencies for  
11 germline transmission (data not shown).  
12 <sup>^</sup>Transmission is based on expression in the vasculature only.  
13 *msn* and *flna* integrations display a dominant or haploinsufficient phenotype in the F0 and F1  
14 generations.  
15

16

17



Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

## 1 **References**

- 2
- 3 Aida, T., Nakade, S., Sakuma, T., Izu, Y., Oishi, A., Mochida, K., Ishikubo, H., Usami, T.,  
4 Aizawa, H., Yamamoto, T., *et al.* (2016). Gene cassette knock-in in mammalian cells and  
5 zygotes by enhanced MMEJ. *BMC Genomics* *17*, 979.
- 6 Ata, H., Ekstrom, T.L., Martinez-Galvez, G., Mann, C.M., Dvornikov, A.V., Schaeffbauer, K.J.,  
7 Ma, A.C., Dobbs, D., Clark, K.J., and Ekker, S.C. (2018). Robust activation of microhomology-  
8 mediated end joining for precision gene editing applications. *PLoS Genet* *14*, e1007652.
- 9 Balciuniene, J., Nagelberg, D., Walsh, K.T., Camerota, D., Georgette, D., Biemar, F., Bellipanni,  
10 G., and Balciunas, D. (2013). Efficient disruption of Zebrafish genes using a Gal4-containing  
11 gene trap. *BMC Genomics* *14*, 619.
- 12 Bedell, V.M., Wang, Y., Campbell, J.M., Poshusta, T.L., Starker, C.G., Krug, R.G., 2nd, Tan, W.,  
13 Penheiter, S.G., Ma, A.C., Leung, A.Y., *et al.* (2012). In vivo genome editing using a high-  
14 efficiency TALEN system. *Nature* *491*, 114-118.
- 15 Beumer, K.J., Trautman, J.K., Bozas, A., Liu, J.L., Rutter, J., Gall, J.G., and Carroll, D. (2008).  
16 Efficient gene targeting in *Drosophila* by direct embryo injection with zinc-finger nucleases. *Proc*  
17 *Natl Acad Sci U S A* *105*, 19821-19826.
- 18 Carlson, D.F., Tan, W., Lillico, S.G., Stverakova, D., Proudfoot, C., Christian, M., Voytas, D.F.,  
19 Long, C.R., Whitelaw, C.B., and Fahrenkrug, S.C. (2012). Efficient TALEN-mediated gene  
20 knockout in livestock. *Proc Natl Acad Sci U S A* *109*, 17382-17387.
- 21 Ceccaldi, R., Rondinelli, B., and D'Andrea, A.D. (2016). Repair Pathway Choices and  
22 Consequences at the Double-Strand Break. *Trends Cell Biol* *26*, 52-64.
- 23 Clark, K.J., Balciunas, D., Pogoda, H.M., Ding, Y., Westcot, S.E., Bedell, V.M., Greenwood,  
24 T.M., Urban, M.D., Skuster, K.J., Petzold, A.M., *et al.* (2011). In vivo protein trapping produces a  
25 functional expression codex of the vertebrate proteome. *Nat Methods* *8*, 506-515.
- 26 Clarke, R., Heler, R., MacDougall, M.S., Yeo, N.C., Chavez, A., Regan, M., Hanakahi, L.,  
27 Church, G.M., Marraffini, L.A., and Merrill, B.J. (2018). Enhanced Bacterial Immunity and  
28 Mammalian Genome Editing via RNA-Polymerase-Mediated Dislodging of Cas9 from Double-  
29 Strand DNA Breaks. *Mol Cell* *71*, 42-55 e48.
- 30 Conway, A.B., Lynch, T.W., Zhang, Y., Fortin, G.S., Fung, C.W., Symington, L.S., and Rice,  
31 P.A. (2004). Crystal structure of a Rad51 filament. *Nat Struct Mol Biol* *11*, 791-796.
- 32 Geurts, A.M., Cost, G.J., Freyvert, Y., Zeitler, B., Miller, J.C., Choi, V.M., Jenkins, S.S., Wood,  
33 A., Cui, X., Meng, X., *et al.* (2009). Knockout rats via embryo microinjection of zinc-finger  
34 nucleases. *Science* *325*, 433.
- 35 Grzesiuk, E., and Carroll, D. (1987). Recombination of DNAs in *Xenopus* oocytes based on  
36 short homologous overlaps. *Nucleic Acids Res* *15*, 971-985.
- 37 Halene, S., Wang, L., Cooper, R.M., Bockstoce, D.C., Robbins, P.B., and Kohn, D.B. (1999).  
38 Improved expression in hematopoietic and lymphoid cells in mice after transplantation of bone  
39 marrow transduced with a modified retroviral vector. *Blood* *94*, 3349-3357.
- 40 Hasty, P., Rivera-Perez, J., Chang, C., and Bradley, A. (1991). Target frequency and integration  
41 pattern for insertion and replacement vectors in embryonic stem cells. *Mol Cell Biol* *11*, 4509-  
42 4517.
- 43 He, M.D., Zhang, F.H., Wang, H.L., Wang, H.P., Zhu, Z.Y., and Sun, Y.H. (2015). Efficient  
44 ligase 3-dependent microhomology-mediated end joining repair of DNA double-strand breaks in  
45 zebrafish embryos. *Mutat Res* *780*, 86-96.
- 46 Hisano, Y., Sakuma, T., Nakade, S., Ohga, R., Ota, S., Okamoto, H., Yamamoto, T., and  
47 Kawahara, A. (2015). Precise in-frame integration of exogenous DNA mediated by  
48 CRISPR/Cas9 system in zebrafish. *Sci Rep* *5*, 8841.

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

- 1 Hoshijima, K., Jurynek, M.J., and Grunwald, D.J. (2016). Precise Editing of the Zebrafish
- 2 Genome Made Simple and Efficient. *Dev Cell* 36, 654-667.
- 3 Ihry, R.J., Worringer, K.A., Salick, M.R., Frias, E., Ho, D., Theriault, K., Kommineni, S., Chen, J.,
- 4 Sondey, M., Ye, C., *et al.* (2018). p53 inhibits CRISPR-Cas9 engineering in human pluripotent
- 5 stem cells. *Nat Med* 24, 939-946.
- 6 Kent, T., Chandramouly, G., McDevitt, S.M., Ozdemir, A.Y., and Pomerantz, R.T. (2015).
- 7 Mechanism of microhomology-mediated end-joining promoted by human DNA polymerase
- 8 theta. *Nat Struct Mol Biol* 22, 230-237.
- 9 Li, X.L., Li, G.H., Fu, J., Fu, Y.W., Zhang, L., Chen, W., Arakaki, C., Zhang, J.P., Wen, W.,
- 10 Zhao, M., *et al.* (2018). Highly efficient genome editing via CRISPR-Cas9 in human pluripotent
- 11 stem cells is achieved by transient BCL-XL overexpression. *Nucleic Acids Res.*
- 12 McGrail, M., Hatler, J.M., Kuang, X., Liao, H.K., Nannapaneni, K., Watt, K.E., Uhl, J.D.,
- 13 Largaespada, D.A., Vollbrecht, E., Scheetz, T.E., *et al.* (2011). Somatic mutagenesis with a
- 14 Sleeping Beauty transposon system leads to solid tumor formation in zebrafish. *PLoS One* 6,
- 15 e18826.
- 16 Mehta, A., Beach, A., and Haber, J.E. (2017). Homology Requirements and Competition
- 17 between Gene Conversion and Break-Induced Replication during Double-Strand Break Repair.
- 18 *Mol Cell* 65, 515-526 e513.
- 19 Melby, A.E., Warga, R.M., and Kimmel, C.B. (1996). Specification of cell fates at the dorsal
- 20 margin of the zebrafish gastrula. *Development* 122, 2225-2237.
- 21 Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K.,
- 22 and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9
- 23 targeting in vivo. *Nat Methods* 12, 982-988.
- 24 Nakade, S., Tsubota, T., Sakane, Y., Kume, S., Sakamoto, N., Obara, M., Daimon, T., Sezutsu,
- 25 H., Yamamoto, T., Sakuma, T., *et al.* (2014). Microhomology-mediated end-joining-dependent
- 26 integration of donor DNA in cells and animals using TALENs and CRISPR/Cas9. *Nat Commun*
- 27 5, 5560.
- 28 Ogura, E., Okuda, Y., Kondoh, H., and Kamachi, Y. (2009). Adaptation of GAL4 activators for
- 29 GAL4 enhancer trapping in zebrafish. *Dev Dyn* 238, 641-655.
- 30 Orr-Weaver, T.L., Szostak, J.W., and Rothstein, R.J. (1981). Yeast transformation: a model
- 31 system for the study of recombination. *Proc Natl Acad Sci U S A* 78, 6354-6358.
- 32 Rong, Y.S., and Golic, K.G. (2000). Gene targeting by homologous recombination in *Drosophila*.
- 33 *Science* 288, 2013-2018.
- 34 Shin, J., Chen, J., and Solnica-Krezel, L. (2014). Efficient homologous recombination-mediated
- 35 genome engineering in zebrafish using TALE nucleases. *Development* 141, 3807-3818.
- 36 Singleton, M.R., Wentzell, L.M., Liu, Y., West, S.C., and Wigley, D.B. (2002). Structure of the
- 37 single-strand annealing domain of human RAD52 protein. *Proc Natl Acad Sci U S A* 99, 13492-
- 38 13497.
- 39 Talbot, W.S., Trevarrow, B., Halpern, M.E., Melby, A.E., Farr, G., Postlethwait, J.H., Jowett, T.,
- 40 Kimmel, C.B., and Kimmel, D. (1995). A homeobox gene essential for zebrafish notochord
- 41 development. *Nature* 378, 150-157.
- 42 Varshney, G.K., Pei, W., LaFave, M.C., Idol, J., Xu, L., Gallardo, V., Carrington, B., Bishop, K.,
- 43 Jones, M., Li, M., *et al.* (2015). High-throughput gene targeting and phenotyping in zebrafish
- 44 using CRISPR/Cas9. *Genome Res* 25, 1030-1042.
- 45 Wang, Y., Kaiser, M.S., Larson, J.D., Nasevicius, A., Clark, K.J., Wadman, S.A., Roberg-Perez,
- 46 S.E., Ekker, S.C., Hackett, P.B., McGrail, M., *et al.* (2010). Moesin1 and Ve-cadherin are
- 47 required in endothelial cells during in vivo tubulogenesis. *Development* 137, 3119-3128.
- 48 Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step
- 49 generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome
- 50 engineering. *Cell* 154, 1370-1379.

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

- 1 Zhang, J.P., Li, X.L., Li, G.H., Chen, W., Arakaki, C., Botimer, G.D., Baylink, D., Zhang, L., Wen,
- 2 W., Fu, Y.W., *et al.* (2017). Efficient precise knockin with a double cut HDR donor after
- 3 CRISPR/Cas9-mediated double-stranded DNA cleavage. *Genome Biol* 18, 35.
- 4 Zhang, Y., Zhang, Z., and Ge, W. (2018). An efficient platform for generating somatic point
- 5 mutations with germline transmission in the zebrafish by CRISPR/Cas9-mediated gene editing.
- 6 *J Biol Chem* 293, 6611-6622.
- 7 Zu, Y., Tong, X., Wang, Z., Liu, D., Pan, R., Li, Z., Hu, Y., Luo, Z., Huang, P., Wu, Q., *et al.*
- 8 (2013). TALEN-mediated precise genome modification by homologous recombination in
- 9 zebrafish. *Nat Methods* 10, 329-331.
- 10

Figure 1

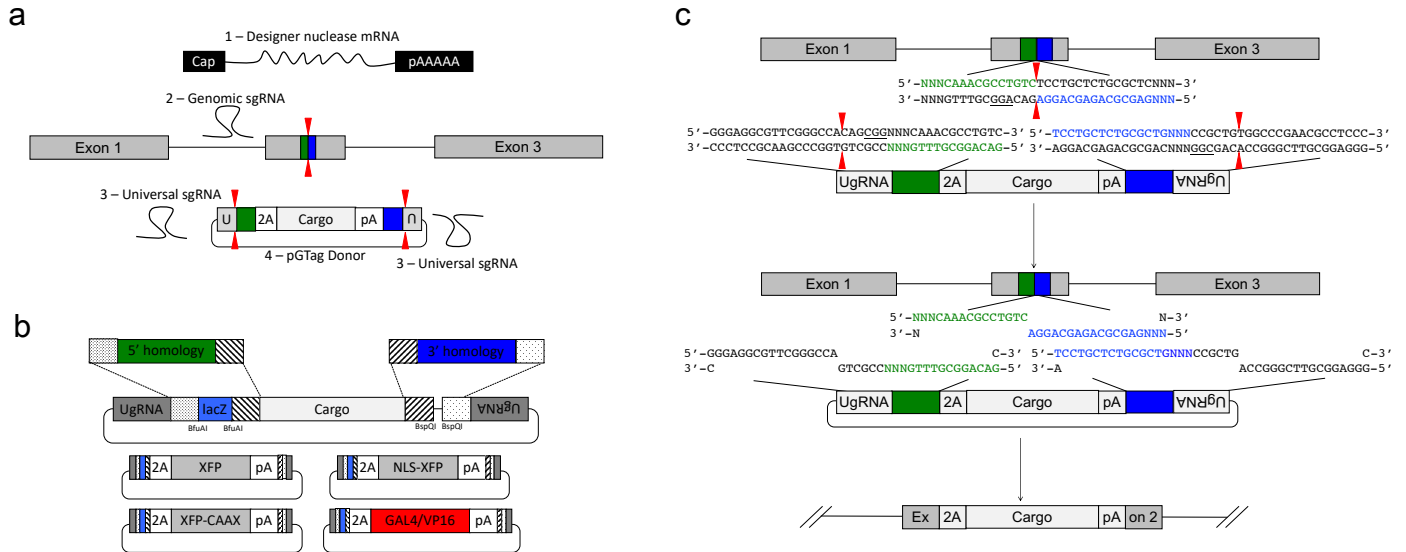


Figure 2

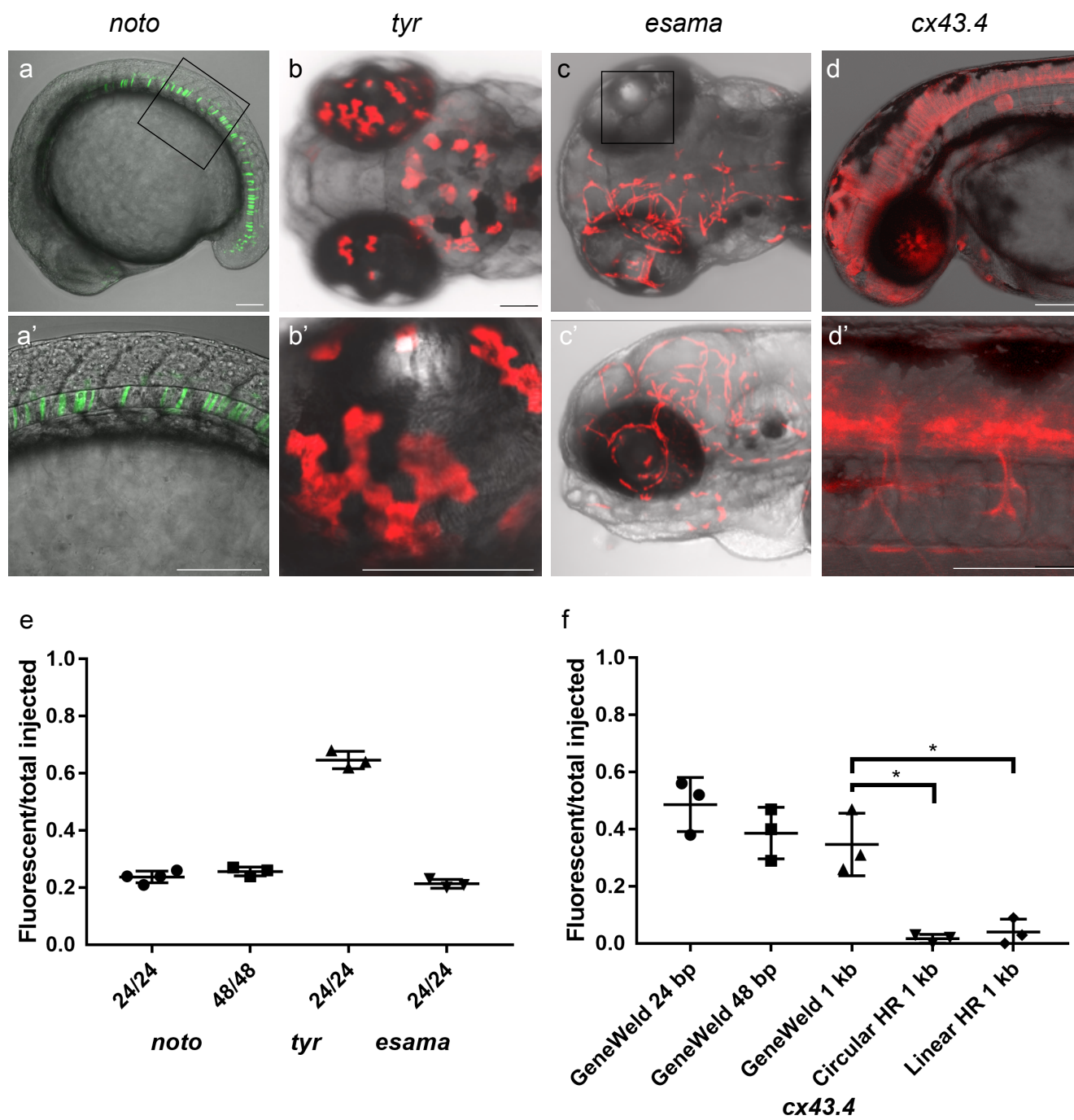


Figure 3

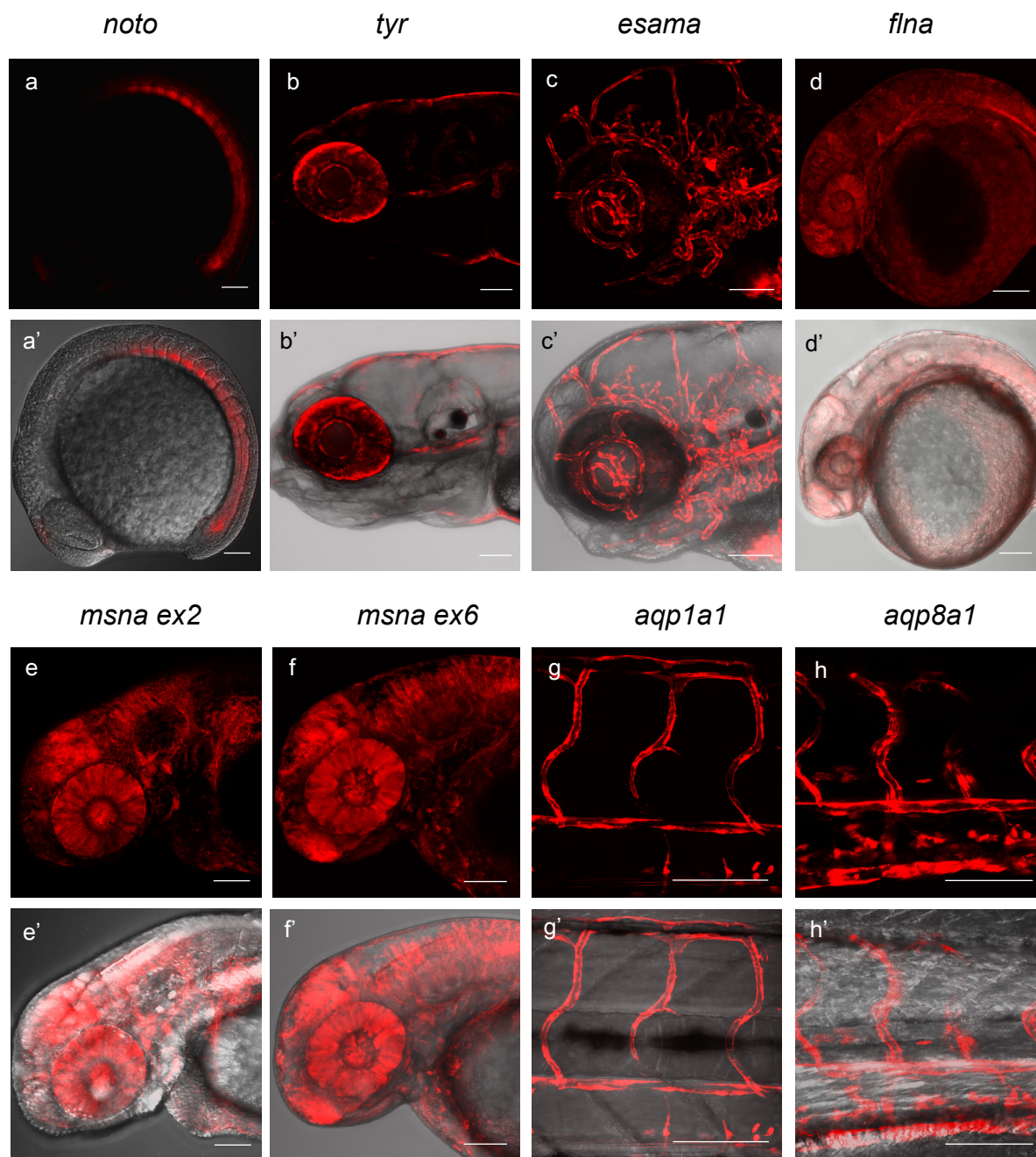


Figure 4

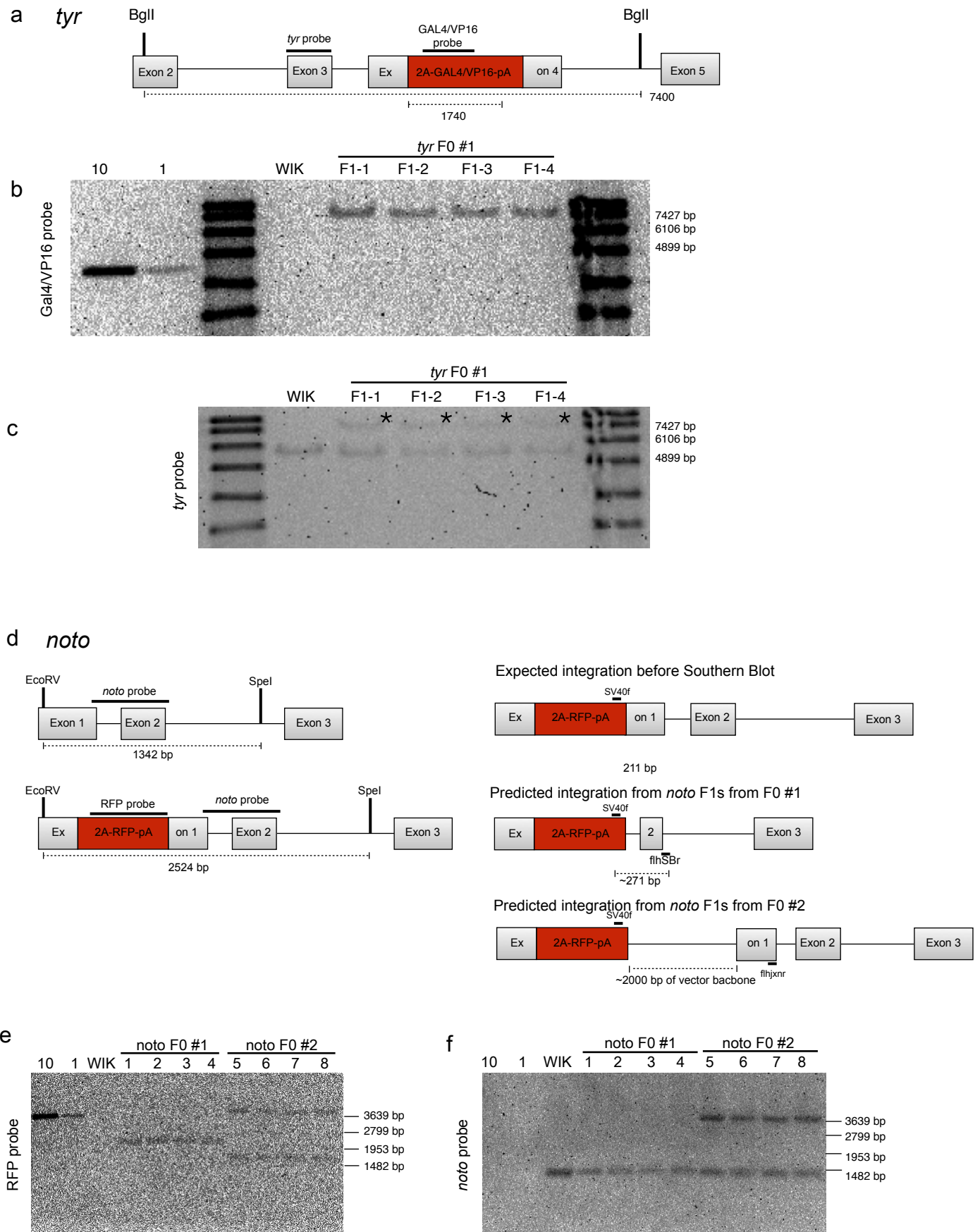
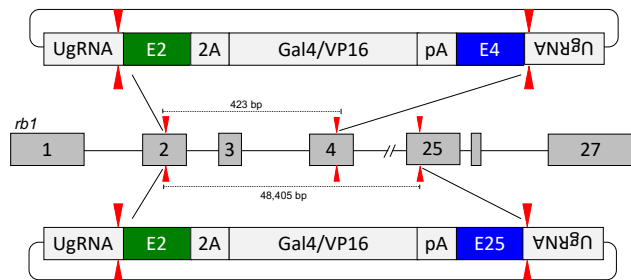
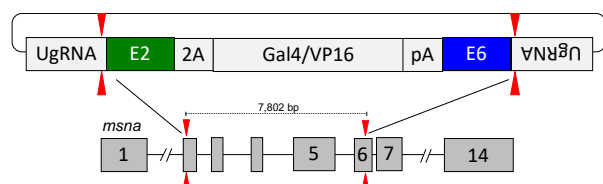


Figure 5

a *rb1*



d *msna*



f

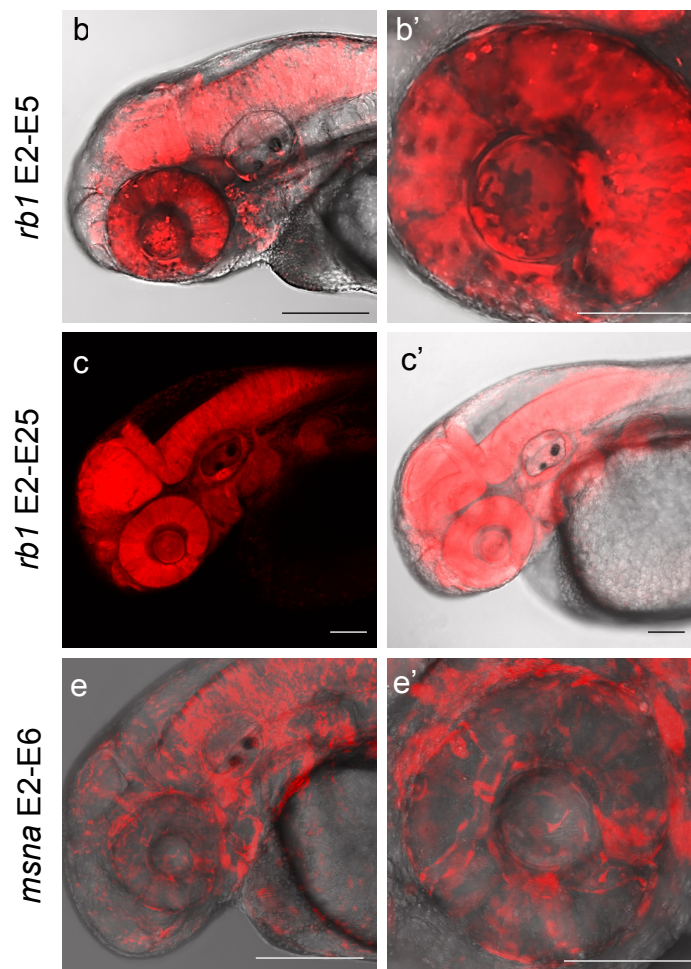
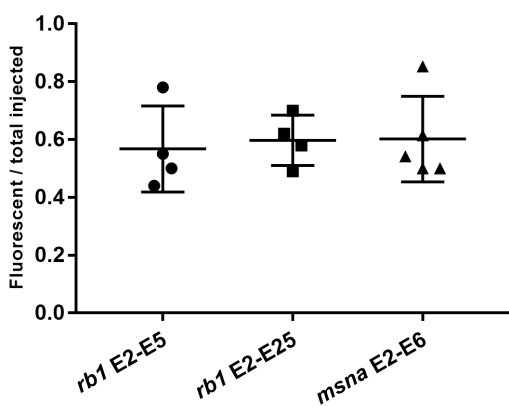
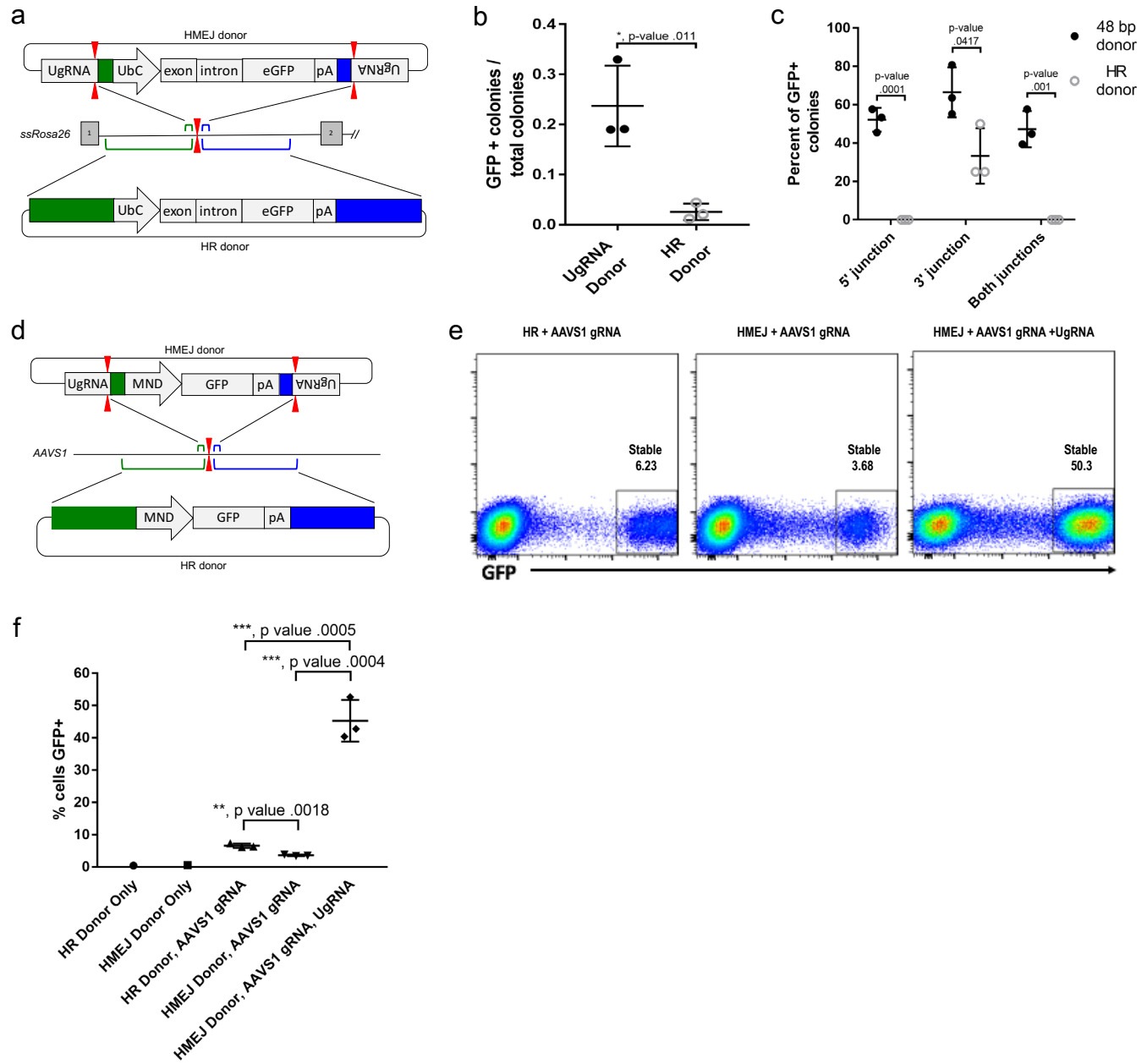




Figure 6



Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

***GeneWeld: a method for efficient targeted integration directed by short homology***

**Supplementary Figure 1. Short homology to the *noto* gene from a single homology arm 5'**

**to the gRNA target site targets integration in zebrafish embryos.** (a) Schematic for *noto*

homology arm and donor vector design. gRNA is the *noto* non-coding template strand. Black

bars represent 12, 24, and 48 bp homology arms. PAM sequences are underlined. (b) Targeting

efficiency of 12, 24, and 48 bp *noto* 5' only donors. Data represents mean +/- s.e.m. of 3

independent targeting experiments. p values calculated using two-tailed unpaired t-test. (c) Live

confocal image of *noto-2A-TagRFP-CAAX-SV40* targeted embryo showing specific RFP

expression in the notochord. Scale bar is 100  $\mu$ m. (d) Sanger sequencing of cloned 5' junction

fragments from RFP positive F0 embryos, aligned to the expected sequence from a precise

integration event.

**Supplementary Figure 2. Single base pair differences in homology arm length 5' to the**

**Cas9/gRNA cut site influence integration frequencies in zebrafish embryos.** (a) Schematic

for targeting 2A-TagRFP-CAAX-SV40 into *noto* exon 1 with 5' homology to the Cas9/gRNA cut

site containing 47, 48, or 49 bp of homology. (b) The frequency of injected zebrafish embryos

displaying notochord RFP expression after targeting *noto* exon 1 with donors containing 47, 48,

or 49 bp of 5' homology. Data represents mean +/- s.e.m. of 3 (47 bp, 49 bp) or 7(48 bp)

independent targeting experiments. p values calculated using two-tailed unpaired t-test.

**Supplementary Figure 3. The Universal gRNA (UgRNA) promotes high efficiency targeted**

**integration.** (a) Universal gRNA (UgRNA) sequence. Cas9 PAM underlined. (b) Schematic

showing the sequence of UgRNA in the targeting domain of the knock-in cassette. Sequence in

green from the *noto* gene is the engineered homology in the donor vector for HMEJ. The Cas9

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

PAM is underlined. (b) Frequency of injected embryos displaying RFP expression in the notochord compared to total injected embryos following *noto* targeting using UgRNA to liberate the homology in the donor.

**Supplementary Figure 4. Integration of Gal4/VP16 amplifies signal of targeted *tyr*.** (a) PCR of 5' junction fragments and sequencing results from junction fragments between the pGTag vector and the *tyr* locus amplified from randomly selected RFP negative embryos after injection with GeneWeld reagents for targeting *tyr* with 2A-tagRFP-CAAX-SV40pA. Most F0 injected zebrafish contain the expected 5' junction fragment (marked with an '\*'). The junction fragments from F0-1 and -2 were isolated for sequencing, and precise integrations were observed (b) Efficiency of 5' homology integration to target RFP or GAL4/VP16 into *tyr* and detect RFP expression. Data represents mean +/- s.e.m. of 3 independent targeting experiments. p values calculated using two-tailed unpaired t-test.

**Supplementary Figure 5. GeneWeld targeting efficiency comparison.** Comparison of the frequency of RFP expressing embryos after targeting *esama* exon 2 using GeneWeld 24/24 bp homology, Geneweld 1kb/1kb homology, Circular HR 1kb/1kb (injection did not include UgRNA), Linear HR 1kb/1kb (donor was linearized before injections). Increasing the length of the homology arm to 1kb significantly increased the frequency of RFP expressing embryos using GeneWeld ( $p=0.0001$ ), Circular, or Linear template. Data represents mean +/- s.e.m. of 3 independent targeting experiments.  $p$  value calculated using Students  $t$  test.

**Supplementary Figure 6. Sequence of PCR junction fragments amplified from genomic DNA from F1 transgenic zebrafish adults generated by GeneWeld targeted integration.**

Precise integration at the 5' and 3' ends in F1 progeny from F0 founder fish targeted at *tyr*,

Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

*esama*, *flna*, *msna*, *aqp1a1*, and *aqp8a1*. *noto* F1 progeny from founder #1 had a precise 5' junction and imprecise 3' junction. *noto* F1 progeny from founder #2 had 5' precise junction; no 3' junction was amplified by PCR. Lowercase letters represent "padding" nucleotides used to bring homology in frame of the coding region based on Cas9 cut site. Red letters represent mismatches unless otherwise noted below. *esama* F1 3' junctions contain a single nucleotide variant shown in red. One *esama* F1 3' junction included a 20 bp insertion (strike-through).

**Supplementary Figure 7. Sequences of 5' and 3' junction fragments from *rb1* exon 2-4, *rb1* exon 2-25, and *msna* exon 2-6 deletion tagged alleles in F0 injected embryos.**

Detection of precise and imprecise 5' and 3' junction fragments in somatic tissue of F0 embryos targeted with two guides targeting two exons and a pGTag-Gal4VP16 donor with 5' and 3' homology arms corresponding to the 5' exon and 3' exon target sites. Cloned PCR amplicons were sequenced from 3 individual embryos for each targeted deletion tagging experiment.

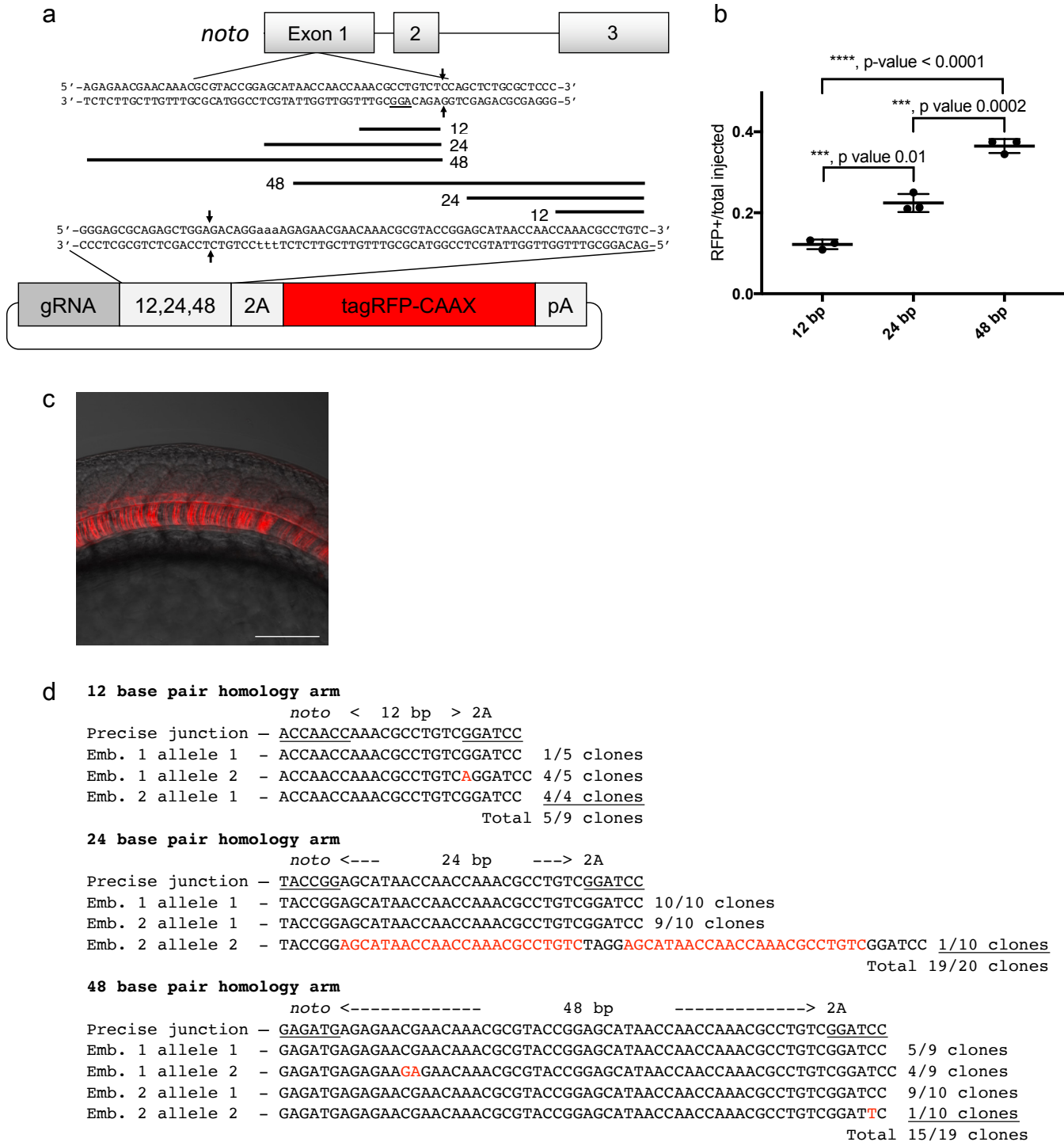
**Supplementary Figure 8. HMEJ-mediated targeted integration of an MND:GFP reporter at the *AAVS1* locus in human K-562 cells.** FACs sorted percent of GFP+ cells out of total K-562 cells at day 7, 21, 28, and 50. (b) Summary data for percent of stable GFP+ K-562 cells from day 7, 14, 21, and 28. (b') Summary data for percent of total cells GFP+ from day 7, 14, 21, 28, and 50. Data represents mean +/- s.e.m. of 3 independent targeting experiments. p values calculated using two-tailed unpaired t-test.

**Supplementary Figure 9. Direct sequencing of 5' junction PCR products derived from three independently targeted bulk cell populations.** (a) Direct sequencing of 5' junction PCR products derived from three independently targeted bulk cell populations. 48 bp HMEJ homology region and remainder of genomic *AAVS1* gRNA are indicated. Genomic sequence is

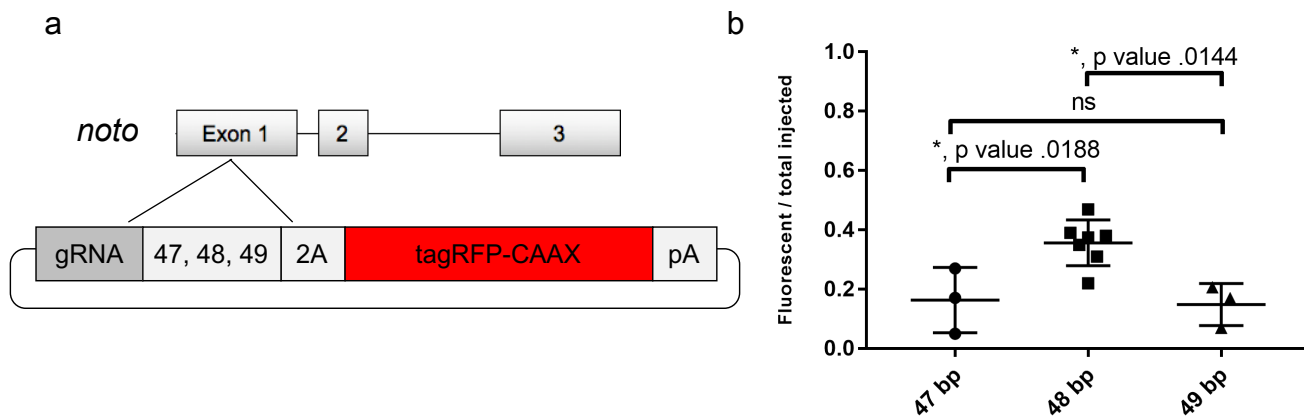
Wierson, Welker, Almeida et al. GeneWeld: a method for efficient targeted integration directed by short homology

directly left of the 48 bp HMEJ region and vector sequence is directly to the right of the AAVS1 gRNA cut site.

## Supplemental figure 1

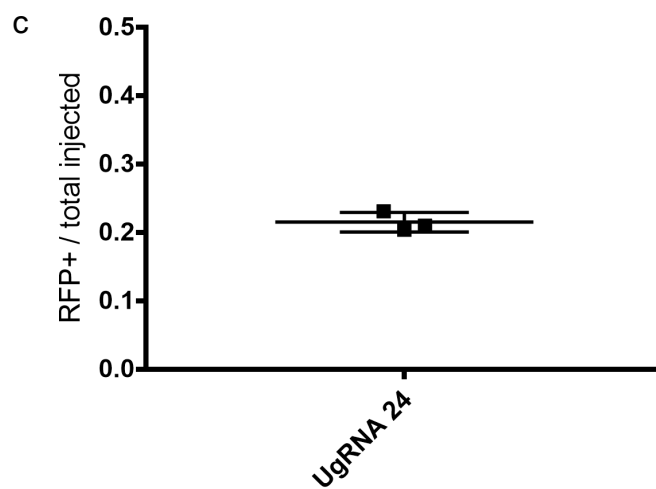
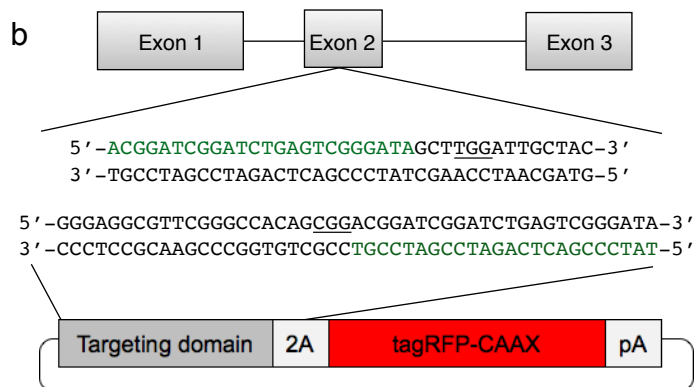


## Supplemental figure 2



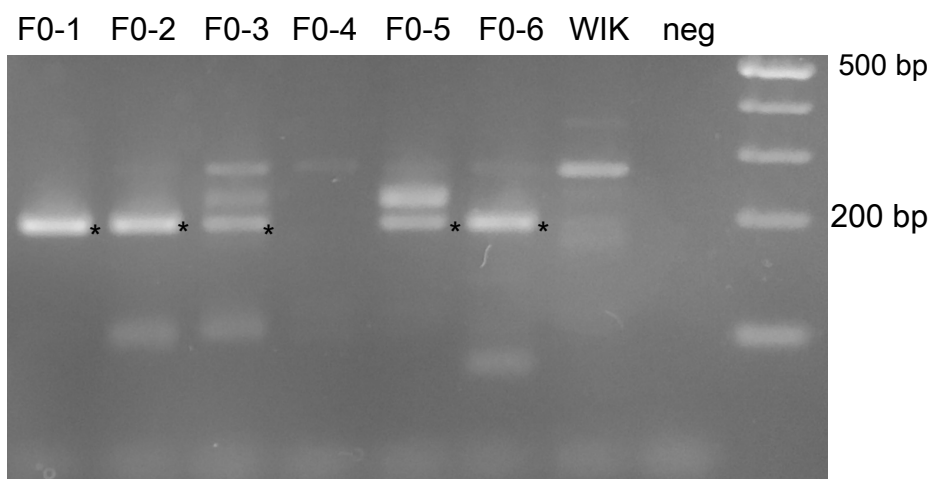
### Supplemental figure 3

a 5' - GGGAGGCGUUCGGGCCACAGCGG - 3'





## Supplemental figure 4



### 24 base pair homology arm

Tyr <--- 24 bp homology ---> 2a

Precise junction - ACAACGACGGATACTTCATGGTGCCCTTCATTGGATCC

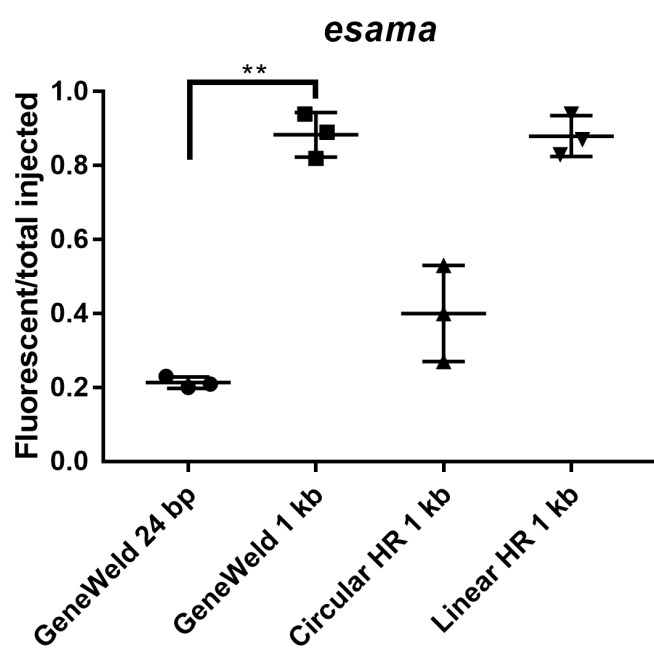
Emb. 1 allele 1 - ACAACGACGGATACTTCATGGTGCCCTTCATTGGATCC 3/4 clones

Emb. 1 allele 2 - ACAACGACGGATACTTCATGGTGCCCTTCATTGGACC 1/4 clones

Emb. 2 allele 1 - ACAACGACGGATACTTCATGGTGCCCTTCATTGGATCC 3/3 clones

Total 7/7 clones precise integration

Supplemental figure 5



## Supplemental figure 6

### **tyr F1 junction fragments - DNA from single F1** 24 base pair homology arms -

#### **5' F1 junction**

Tyr <--- 24 bp domain ---> Vector  
Precise junction - ACAACGACGGATACTTCATGGTGCCCTTCAttGGATCC  
F1 - ACAACGACGGATACTTCATGGTGCCCTTCAttGGATCC

#### **3' F1 junction**

Vector<--- 24 bp domain ---> Tyr  
Precise junction - CCATGGTCCCTCTCTACAGGAACGGAGACTATTTTC  
F1 - CCATGGTCCCTCTCTACAGGAACGGAGACTATTTTC

### **noto F1 junction fragments from DNA represented on Southern Blot** 24 base pair homology arms -

#### **5' F1 junctions from FO #1**

noto <--- 24 bp domain --->Vector  
Precise junction - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #1 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #2 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #3 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #4 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC

#### **3' F1 junctions from FO #1**

--24bp homology domain-- Alternate homology  
Knock-in alignment - TCCAGCTCTGCGCTCCCGCTTATT-----ATCTGCTCTCCAAC TCACT  
F1 #1 no sequencing performed  
F1 #2 - TCCAGCTCTGCGCTCCCGCTTATTCCGCTTATCTGCTCTCCAAC TCACT  
F1 #3 - TCCAGCTCTGCGCTCCCGCTTATTCCGCTTATCTGCTCTCCAAC TCACT  
F1 #4 - TCCAGCTCTGCGCTCCCGCTTATTCCGCTTATCTGCTCTCCAAC TCACT

#### **5' F1 junctions from FO #2**

noto <--- 24 bp domain --->Vector  
Precise junction - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #5 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #6 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #7 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC  
F1 #8 - TACCGGAGCATAACCAACCAACGcCTGTCCGGATCC

#### **3' F1 junctions from FO #2**

No junctions obtained

**esama F1 junction fragments**

24 base pair homology arm -

Supplemental figure 6 continued

**5' F1 junctions**

esama <--- 24 bp domain ---> Vector  
Precise junction - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#4, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#4, F1#2 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#5, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#5, F1#2 - GATAAGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#6, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#6, F1#2 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#7, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#7, F1#2 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#9, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC  
F0#11, F1#1 - CTTATGAAAATGTGGATGTGATCCAAGGGAttGGATCC

**3' F1 junctions**

Vector<--- 24 bp domain -> esama  
Precise junction - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#4, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#4, F1#2 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#5, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#5, F1#2 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#6, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#6, F1#2 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#7, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATGAAGGTGCTGCAGGCTTCATATTCGA  
F0#7, F1#2 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#9, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA  
F0#11, F1#1 - CCATGGAGATGGTGGTGCAGGCTTCATATTCGA

**flna F1 junction fragments**

48 base pair homology arm -

**5' F1 junctions**

flna <--- 48 bp domain --->Vector  
Precise junction - ATGACTTACCTGTCCCAGTTTCCCAAAGCCAAACTCAAGCCTGGTGCCCTCTGGGATCC  
F1#1 - ATGACTTACCTGTCCCAGTTTCCCAAAGCCAAACTCAAGCCTGGTGCCCTCTGGGATCC  
F1#2 - ATGACTTACCTGTCCCAGTTTCCCAAAGCCAAACTCAAGCCTGGTGCCCTCTGGGATCC  
F1#3 - ATGACTTACCTGTCCCAGTTTCCCAAAGCCAAACTCAAGCCTGGTGCCCTCTGGGATCC  
F1#4 - ATGACTTACCTGTCCCAGTTTCCCAAAGCCAAACTCAAGCCTGGTGCCCTCTGGGATCC

**3' F1 junctions**

not determined

**msna F1 junctions for exon 2 integration**

48 base pair homology arm -

**5' F1 junctions**

msna <--- 48 bp domain ---> Vector  
Precise junction - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#1 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#2 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#3 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#4 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#5 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#6 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#7 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#8 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#9 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#10 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#11 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC  
F1#12 - GTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATccGGAGCC

**3' F1 junction**

Vector<--- 48 bp domain ---> msna  
Precise junction - AGGAAGCAGCACCACAGGAAACAGTTATTTGACCAGGTTTGTGTGGCCTCTTTTTTTT  
F1 - AGGAAGCAGCACCACAGGAAACAGTTATTTGACCAGGTTTGTGTGGCCTCTTTTTTTT

## Supplemental figure 6 continued

### **aqp1a1 F1 junction fragments**

#### **5' F1 junction**

aqp1a1<--- 48 bp domain ---> Vector  
**Precise junction** - CCGTCAGTCAGTCAGTCATGAACGAGCTGAAGAGCAAGGCTTTCTGGCGGGCCGccGGATCC  
F2 line 1 - CCGTCAGTCAGTCAGTCATGAACGAGCTGAAGAGCAAGGCTTTCTGGCGGGCCGccGGATCC

#### **3' F1 junction**

Vector<--- 48 bp domain --->aqp1a1  
**Precise junction** - AGGAAGTCCTGGCCGAGCTGCTGGGAATGACCCTGTTTCATCTTCCTCAGCATTACAGCAG  
F2 line 1 - AGGAAGTCCTGGCCGAGCTGCTGGGAATGACCCTGTTTCATCTTCCTCAGCATTACAGCAG

### **aqp8a1 F1 junction fragments**

#### **5' F1 junction**

aqp1a1<--- 48 bp domain ---> Vector  
**Precise junction** - GAGTCGTCGGCTCTTTCTCTTCATGTTTGTGGGCTGCGTGTCGTCATGGGCAttGGATCC  
F1#1 - GAGTCGTCGGCTCTTTCTCTTCATGTTTGTGGGCTGCGTGTCGTCATGGGCAttGGATCC

#### **3' F1 junction**

Vector<--- 48 bp domain --->aqp1a1  
**Precise junction** - AGGAAGACGTGGGCATCAGCGGGAGCATCCAGCCC GCCCTGGCACACGGACTAGCACTGG  
F1#1 - AGGAAGACGTGGGCATCAGCGGGAGCATCCAGCCC GCCCTGGCACACGGACTAGCACTGG

## Supplemental Figure 7

### *rb1 e2-e4 deletion tagging*

#### 5' junction

*rb1 e2*<--- 48 bp domain ---> Vector  
Precise junction - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC  
Embryo 1 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC 5/5 clones precise  
Embryo 2 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAGAGGAAATTAGGAGCATGGacGGATCC 4/5 clones precise  
Embryo 3 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC 5/5 clones precise

#### 3' junction

Vector <--- 48 bp domain --->*rb1 e4*  
Precise junction - AGGAAGCCTAAGGTCAATAGCGCCGTACCCGCCTAGAGAACAATAACGATGTGACTTTGGCCCTCTACCAAAGATTTGTA  
Embryo 1 - AGGAAGCCTAAGGTCAATAGCGCCGTACCCGCCTAGAGAACAATAACGATGTGACTTTGGCCCTCTACCAAAGATTTGTA 1/5 clones precise  
Embryo 2 - AGGAAGCCTAAGGTCAATAGCGCCGTACCCGCCTAGAGAACAATAACGATGTGACTTTGGCCCTCTACCAAAGATTTGTA 2/5 clones precise  
Embryo 3 - AGGAAGCCTAAGGTCAATAGCGCCGTACCCGCCTAGAGAACAATAACGATGTGACTTTGGCCCTCTACCAAAGATTTGTA 5/5 clones precise

### *rb1 e2-e25 deletion tagging*

#### 5' junction

*rb1 e2*<--- 48 bp domain ---> Vector  
Precise junction - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC  
Embryo 1 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC 5/5 clones precise  
Embryo 2 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC 5/5 clones precise  
Embryo 3 - CGAGGAGCTCCAGTCCACTAACTCCATCTGTGATCATGCATGGAGAATATGGGAAAGAGAAATTAGGAGCATGGacGGATCC 5/5 clones precise

#### 3' junction

Vector <--- 48 bp domain --->*rb1 e25*  
Precise junction - AGGAAGCCCTCAAAGACTCAGATTCGATATGGACGGACAAGATGAAGCAGACGGAAGGTGGGAGTCATGATCAGTTTACTCT  
Embryo 1 - AGGAAGCCCTCAAAGACTCAGATTCGATATGGACGGACAAGATGAAGCAGACGGAAGGTGGGAGTCATGATCAGTTTACTCT 5/5 clones precise  
Embryo 2 - AGGAAGCCCTCAAAGACTCAGATTCGATATGGACGGACAAGATGAAGCAGACGGAAGGTGGGAGTCATGATCAGTTTACTCT 4/5 clones precise  
Embryo 3 - AGGAAGCCCTCAAAGACTCAGATTCGATATGGACGGACAAGATGAAGCAGACGGAAGGTGGGAGTCATGATCAGTTTACTCT 3/5 clones precise

### *msna e2-6 deletion tagging*

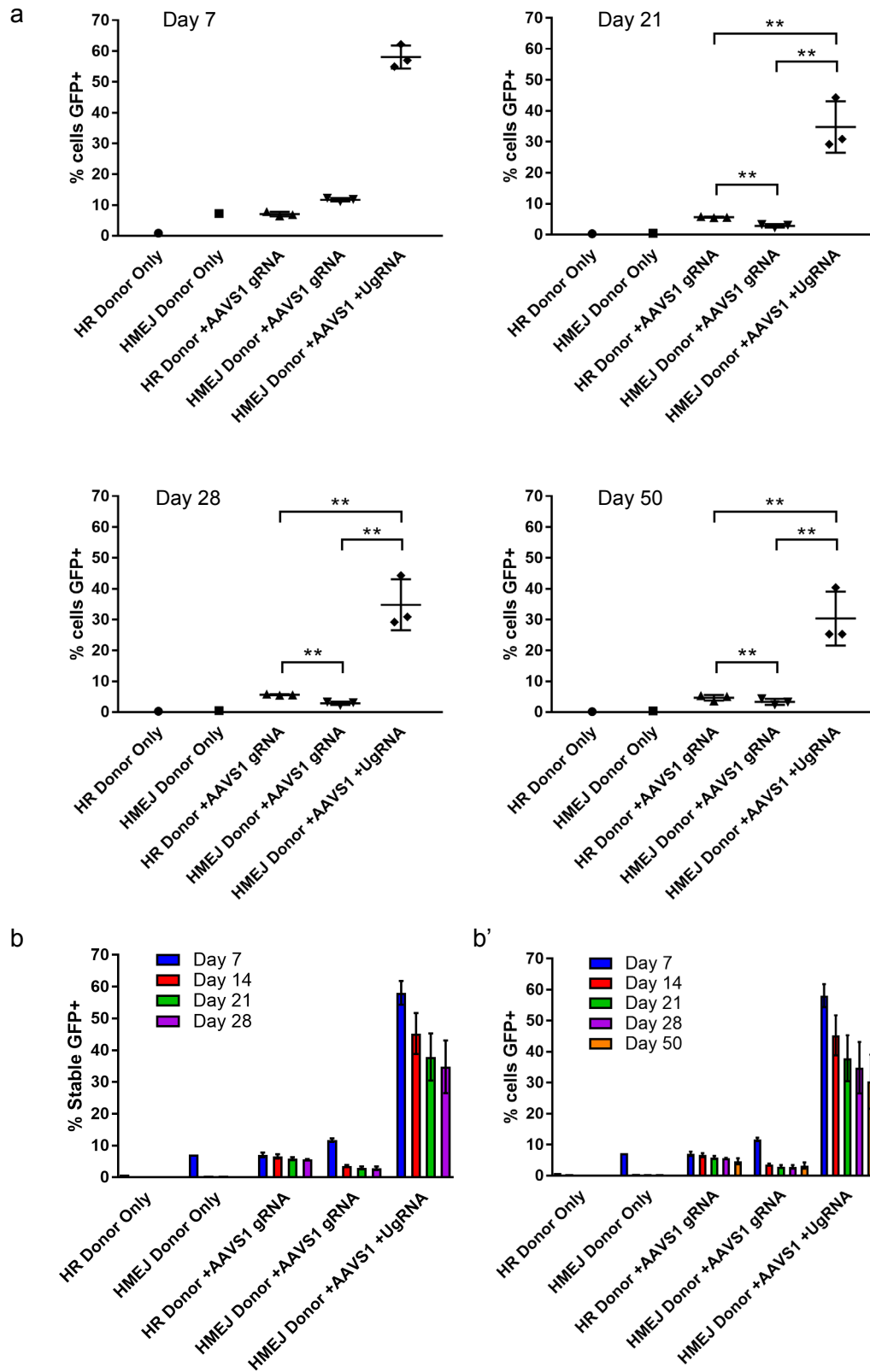
#### 5' junction

*msna* <---- 48 bp domain --->Vector  
Precise junction - GATCAGTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATCC  
Embryo 1 - GATCAGTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATCC 4/4 clones single SNP  
Embryo 2 - GATCAGTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATCC 4/5 clones precise  
Embryo 3 - GATCAGTGTTTCGTGTGACTACAATGGATGCCGAGCTGGAGTTTGCCATCCAACCCGGATCC 3/4 clones precise

#### 3' junction

vector<---- 48 bp domain ---> *msna*  
Precise junction - AGGAAGCAAGGGCATGTTGAGGTACAGACAATGGAATGTGCTCTTGCTATTTTTCTGTTT  
Embryo 1 - No junction obtained  
Embryo 2 - AGGAAGCAAGGGCATGTTGAGGTACAGACAATGGAATGTGCTCTTGCTATTTTTCTGTTT 1/5 clones single SNP  
Embryo 2 - AGGAAGCAAGGGCATGTTGAGGTACAGACAATGGAATGTGCTCTTGCTATTTTTCTGTTT 2/5 clones precise  
Embryo 2 - AGGAAGCAAGGGCATGTTGAGGTACAGACAATGGAATGTGCTCTTGCTATTTTTCTGTTT 2/5 clones single SNP  
Embryo 3 - AGGAAGCAAGGGCATGTTGAGGTACAGACAATGGAATGTGCTCTTGCTATTTTTCTGTTT 4/5 clones precise

Supplemental Figure 8



## Supplemental figure 9

### **AAVS1 5' junction**

AAVS1 <--- 48 bp domain --->Vector  
Predicted junction - GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTACGATCG  
Repetition 1 - GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTACGATCG  
Repetition 2 - GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTACGATCG  
Repetition 3 - GCTCTGGTTCTGGGTACTTTTATCTGTCCCCTCCACCCCACAGTGGGGCCACTACGATCG



Supplementary Table 1 - Gene targeting experiments and knock-in averages.								
Genomic target	Donor vector	Donor sgRNA target (genomic or UgRNA)	Homology length (5'/3')	Experiment number	Reporter positive embryos	Total embryos	Percent with positive report	Standard Error
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	12/x	Average	22	178	12.4%	0.55%
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	24/x	Average	30	131	22.9%	0.74%
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	48/x	Average	40	114	35.1%	1.01%
<i>noto</i> E2	2A-TagRFP-CAAX-SV40	UgRNA	24/x	Average	32	150	21.3%	0.66%
<i>noto</i> E1	pGTag-2A-eGFP-SV40	UgRNA	24/24	Average	45	185	24.3%	0.90%
<i>noto</i> E1	pGTag-2A-eGFP-SV40	UgRNA	48/48	Average	44	172	25.6%	0.82%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	UgRNA	24/24	Average	84	132	63.6%	1.44%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	24/24	Average	42	84	50.0%	4.39%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	48/48	Average	31	81	38.3%	4.15%
<i>esama</i> E2	ag-2A-Gal4VP16-ba	UgRNA	48/48	Average	34	162	21.0%	0.59%
<i>msna</i> E2	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	102	185	55.1%	9.87%
<i>msna</i> E6	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	25	96	26.0%	
<i>anxa2a</i> E3	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	12	34	35.3%	
<i>flna</i> E4	pGTag-2A-Gal4VP16-bactin	UgRNA	48/42	Average	9	9	100.0%	
<i>aqp1a1</i> E1	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	1	25	4.0%	
<i>aqp8a1</i> E1	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	3	21	14.3%	
<i>rb1</i> E2-E4	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	273	463	59.0%	6.44%
<i>rb1</i> E2-E25	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	292	486	60.1%	3.78%
<i>msna</i> E2-E6	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	91	151	60.3%	5.92%
<i>kdrl</i> E3-E30	pGTag-2A-Gal4VP16-bactin	UgRNA	42/48	Average	8	23	34.8%	
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	UgRNA	48/27	Average	76	94	80.9%	
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	Average	29	92	31.5%	
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	Average	60	177	33.9%	5.24%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	Average	0	211	0.0%	0.00%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	na	1Kb/1Kb	Average	2	127	1.6%	0.80%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	Average	0	81	0.0%	0.00%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	Average	5	94	5.3%	2.12%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	Average	0	82	0.0%	0.00%

Supplementary Table 2 - Gene targeting experiments and knock-in percentages.

Genomic target	Donor vector	Donor sgRNA target (genomic or UgRNA)	Homology length (5'/3')	Experiment number	Reporter positive embryos	Total embryos	Percent with positive report	Standard Error
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	12/x	1	4	32	12.5%	0.55%
		genomic sgRNA	12/x	2	12	91	13.2%	
		genomic sgRNA	12/x	3	6	55	10.9%	
		Average			22	178	12.4%	
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	24/x	1	10	47	21.3%	0.74%
		genomic sgRNA	24/x	2	10	41	24.4%	
		genomic sgRNA	24/x	3	10	43	23.3%	
		Average			30	131	22.9%	
<i>noto</i> E1	2A-TagRFP-CAAX-SV40	genomic sgRNA	48/x	1	10	29	34.5%	1.01%
		genomic sgRNA	48/x	2	15	40	37.5%	
		genomic sgRNA	48/x	3	15	45	33.3%	
		Average			40	114	35.1%	
<i>noto</i> E2	2A-TagRFP-CAAX-SV40	UgRNA	24/x	1	13	62	21.0%	0.66%
		UgRNA	24/x	2	10	49	20.4%	
		UgRNA	24/x	3	9	39	23.1%	
		Average			32	150	21.3%	
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	24/24	1	18	68	26.5%	0.90%
		UgRNA	24/24	2	9	38	23.7%	
		UgRNA	24/24	3	6	28	21.4%	
		UgRNA	24/24	4	12	51	23.5%	
<i>noto</i> E1	pGTag-2A-eGFP-SV40	UgRNA	48/48	1	9	33	27.3%	0.82%
		UgRNA	48/48	2	17	71	23.9%	
		UgRNA	48/48	3	18	68	26.5%	
		Average			44	172	25.6%	
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	UgRNA	24/24	1	21	31	67.7%	1.44%
		UgRNA	24/24	2	21	33	63.6%	
		UgRNA	24/24	3	42	68	61.8%	
	Average			84	132	63.6%		
	2A-TagRFP-CAAX-SV40	genomic sgRNA	24/x	1	0	41	0.0%	
		genomic sgRNA	24/x	2	0	22	0.0%	
Average				0	63	0.0%		
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	24/24	1	8	21	38.1%	4.39%
		UgRNA	24/24	2	15	29	51.7%	
		UgRNA	24/24	3	19	34	55.9%	
		Average			42	84	50.0%	
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	48/48	1	10	34	29.4%	4.15%
		UgRNA	48/48	2	15	32	46.9%	
		UgRNA	48/48	3	6	15	40.0%	
		Average			31	81	38.3%	
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	UgRNA	20/20	1	1	18	5.6%	0.87%
		UgRNA	20/20	2	2	32	6.3%	
		UgRNA	20/20	3	1	36	2.8%	
		Average			4	86	4.7%	
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	UgRNA	24/24	1	9	40	22.5%	0.59%
		UgRNA	24/24	2	13	65	20.0%	
		UgRNA	24/24	3	12	57	21.1%	
		Average			34	162	21.0%	
<i>msna</i> E2	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	11	34	32.4%	9.87%
		UgRNA	48/48	2	91	151	60.3%	
		Average			102	185	55.1%	
<i>msna</i> E6	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	25	96	26.0%	26.0%
<i>anxa2a</i> E3	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	12	34	35.3%	35.3%
		Average			12	34	35.3%	
<i>flna</i> E4	pGTag-2A-Gal4VP16-bactin	UgRNA	48/42	1	9	9	100.0%	100.0%
<i>aqp1a1</i> E1	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	1	25	4.0%	4.0%
		UgRNA	48/48	2	n/a	n/a	n/a	
<i>aqp8a1</i> E1	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	3	21	14.3%	14.3%
		Average			3	21	14.3%	
<i>rb1</i> E2-4	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	48	108	44.0%	6.44%
		UgRNA	48/48	2	56	111	50.0%	
		UgRNA	48/48	3	53	96	55.0%	
		UgRNA	48/48	4	116	148	78.0%	
<i>rb1</i> E2-25	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	47	76	62.0%	3.78%
		UgRNA	48/48	2	58	119	49.0%	
		UgRNA	48/48	3	87	149	58.0%	
		UgRNA	48/48	4	100	142	70.0%	
<i>msna</i> E2-E6	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	8	13	61.5%	5.92%
		UgRNA	48/48	2	29	34	85.3%	
		UgRNA	48/48	3	17	34	50.0%	
		UgRNA	48/48	4	26	48	54.2%	
		UgRNA	48/48	5	11	22	50.0%	
<i>kdrl</i> E3-E30	pGTag-2A-Gal4VP16-bactin	UgRNA	42/48	1	8	23	34.8%	34.8%
		Average			8	23	34.8%	
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	UgRNA	48/27	1	76	94	80.9%	80.9%
		Average			76	94	80.9%	
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	UgRNA	48/48	1	29	92	31.5%	31.5%
		Average			29	92	31.5%	
<i>cx43.4</i> E2 GeneWeld	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	1	22	72	30.6%	25.9%
		UgRNA	1Kb/1Kb	2	24	51	47.1%	
		UgRNA	1Kb/1Kb	3	14	54	25.9%	

				Average	60	177	33.9%	5.24%
<i>cx43.4 E2</i> Plasmid Control	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	1	0	78	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	2	0	50	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	3	0	83	0.0%	
	Average				0	211	0.0%	0.00%
<i>cx43.4 E2</i> Genomic Guide only	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	1	1	30	3.3%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	2	1	46	2.2%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/1Kb	3	0	51	0.0%	
	Average				2	127	1.6%	0.80%
<i>cx43.4 E2</i> Universal Guide only	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	1	0	44	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	2	0	24	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	UgRNA	1Kb/1Kb	3	0	13	0.0%	
	Average				0	81	0.0%	0.00%
<i>cx43.4 E2</i> Grundwald Linear Temple	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	1	1	31	3.2%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	2	4	45	8.9%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	3	0	18	0.0%	
	Average				5	94	5.3%	2.12%
<i>cx43.4 E2</i> Linerar Control	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	1	0	23	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	2	0	16	0.0%	
	pGTag-2A-TagRFP-CAAX-SV40	n/a	1Kb/700	3	0	43	0.0%	
	Average				0	82	0.0%	0.00%
<i>esama E2</i> GeneWeld	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	1	17	19	89.5%	
	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	2	18	22	81.8%	
	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	3	30	32	93.8%	
	Average				65	73	89.0%	2.87%
<i>esama E2</i> Plasmid Control	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	1	2	12	16.7%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	2	9	35	25.7%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	3	1	18	5.6%	
	Average				12	65	18.5%	4.76%
<i>esama E2</i> Genomic Guide only	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	1	8	15	53.3%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	2	3	11	27.3%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/1Kb	3	8	20	40.0%	
	Average				19	46	41.3%	6.14%
<i>esama E2</i> Universal Guide only	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	1	11	17	64.7%	
	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	2	13	18	72.2%	
	pGTag-2A-Gal4VP16-bactin	UgRNA	1Kb/1Kb	3	9	20	45.0%	
	Average				33	55	60.0%	6.63%
<i>esama E2</i> Grundwald Linear Temple	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	1	27	31	87.1%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	2	20	24	83.3%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	3	16	17	94.1%	
	Average				63	72	87.5%	2.58%
<i>esama E2</i> Linerar Control	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	1	17	18	94.4%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	2	29	31	93.5%	
	pGTag-2A-Gal4VP16-bactin	n/a	1Kb/700	3	20	21	95.2%	
	Average				66	70	94.3%	0.40%

Supplementary table 3 -  
Germline Transmission

Genomic target	F0s outcrossed	F0s transmitting	Germline transmission percentage
<i>noto</i> E1 24/24	5	3	60.0%
<i>tyr</i> E4 48/48	8	3	37.5%
<i>cx43.4</i> 24/24	1	0	0.0%
<i>cx43.3</i> 48/48	4	0	0.0%
<i>esama</i> E2 48/48	18	12	66.7%
<i>msna</i> E2 48/48	4	1	25.0%
<i>msna</i> E6 48/48	3	1	33.3%
<i>anxa2a</i> E3 48/48	4	4	100.0%
<i>flna</i> E4 48/42	3	3	100.0%
<i>aqp1a1</i> E1 48/48	9	2	22.2%
<i>aqp8a1</i> E1 48/48	1	1	100.0%
<i>rb1</i> E2-E4 48/48	10	0	0.0%
<i>rb1</i> E2-E25 48/48	16	1	6.3%
<i>msna</i> E2-E6 48/48	8	0	0.0%
<i>kdrl</i> E3-E30 44/48	3	0	0.0%
<i>s1pr1</i> E2-E2 48/48	24	0	0.0%
<i>vegfaa</i> E3-E7 48/48	16	0	0.0%
<i>mmp14a</i> E1-E10 48/48	4	0	0.0%

Supplementary Table 4  
- F0 outcrosses for  
transmission; asterisk  
denotes in-cross

Genomic target	Knock-in	Homology length	Individual F0	Crossed to	Reporter positive F1 embryos	Total F1 embryos	Percent germline transmission
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	1	Casper	4	28	14.3%
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	2	Casper	0	172	0.0%
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	3	Casper	0	15	0.0%
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	4	Casper	69	146	47.3%
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	1	Casper	11	61	18.0%
<i>noto</i> E1	pGTag-2A-TagRFP-CAAX-SV40	24/24	2	Casper	0	81	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	1	UAS:RFP	0	174	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	2	UAS:RFP	0	122	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	3	UAS:RFP	0	45	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	4	UAS:RFP	0	87	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	5	UAS:RFP	0	103	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	6	UAS:RFP	8	89	9.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	7	UAS:RFP	13	151	8.6%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	8	UAS:RFP	0	113	0.0%
<i>tyr</i> E4	pGTag-2A-Gal4VP16-bactin	24/24	9	UAS:RFP	19	155	12.3%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV42	24/24	1	flil:EGFP	0	39	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV42	24/24	2	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV43	24/24	3	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV44	24/24	4	flil:EGFP	0	15	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV45	24/24	5	flil:EGFP	0	26	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV46	24/24	6	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV47	24/24	7	flil:EGFP	0	23	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV48	24/24	8	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV49	24/24	9	flil:EGFP	0	45	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV50	24/24	10	flil:EGFP	0	82	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	1	flil:EGFP	0	130	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	2	flil:EGFP	0	100	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	3	flil:EGFP	0	75	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	4	flil:EGFP	0	52	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	5	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	6	flil:EGFP	0	12	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	7	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	8	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	9	flil:EGFP	0	34	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	10	flil:EGFP	0	86	0.0%
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	11	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	12	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	13	flil:EGFP	0	0	N/A
<i>cx43.4</i> E2	pGTag-2A-TagRFP-CAAX-SV40	48/48	14	flil:EGFP	0	0	N/A
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	1	p08790	0	21	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	2	p08790	0	212	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	3	p08790	0	31	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	4	p08790	1	4	25.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	5	p08790	1	12	8.3%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	6	p08790	14	104	13.5%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	7	p08790	0	87	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	8	p08790	0	209	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	9	p08790	0	132	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	10	p08790	4	18	22.2%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	11	p08790	0	37	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	12	p08790	11	43	25.6%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	13	p08790	14	97	14.4%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	14	p08790	0	91	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	15	p08790	7	127	5.5%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	16	p08790	0	25	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	17	p08790	30	137	21.9%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	18	p08790	8	265	3.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	19	p08790	31	227	13.7%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	20	p08790	11	146	7.5%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	21	p08790	0	188	0.0%
<i>esama</i> E2	pGTag-2A-Gal4VP16-bactin	24/24	22	p08790	3	66	4.5%
<i>msna</i> E2	pGTag-2A-Gal4VP16-bactin	48/48	1	p08790	42	144	29.2%
<i>msna</i> E2	pGTag-2A-Gal4VP16-bactin	48/48	2	p08790	0	377	0.0%

<i>msna</i> E2	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	0	202	0.0%
<i>msna</i> E2	pGTag-2A-GalMVP16-bactin	48/48	4	pDB790	0	150	0.0%
<i>msna</i> E6	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	23	311	7.4%
<i>msna</i> E6	pGTag-2A-GalMVP16-bactin	48/48	2	pDB790	0	67	0.0%
<i>msna</i> E6	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	0	54	0.0%
<i>anxa2a</i> E3	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	16	73	21.9%
<i>anxa2a</i> E3	pGTag-2A-GalMVP16-bactin	48/48	2	pDB790	10	22	45.5%
<i>anxa2a</i> E3	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	37	70	52.9%
<i>anxa2a</i> E3	pGTag-2A-GalMVP16-bactin	48/48	4	pDB790	23	36	63.9%
<i>flna</i> E4	pGTag-2A-GalMVP16-bactin	48/42	1	pDB790	21	64	32.8%
<i>flna</i> E4	pGTag-2A-GalMVP16-bactin	48/42	2	pDB790	1	52	1.9%
<i>flna</i> E4	pGTag-2A-GalMVP16-bactin	48/42	3	pDB790	4	21	19.0%
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	2	pDB790	19	281	6.8%
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	4	pDB790	n>1	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	5	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	6	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	7	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	8	pDB790	0	N/A	N/A
<i>agp1a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	9	pDB790	0	N/A	N/A
<i>agp8a1</i> E1	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	1	7	14.3%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	0	38	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	2	pDB790	0	103	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	0	61	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	4	pDB790	0	73	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	5	pDB790	0	57	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	6	pDB790	0	69	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	7	pDB790	0	37	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	8	pDB790	0	113	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	9	pDB790	0	79	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	10	pDB790	0	20	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	11	pDB790	0	136	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	12	pDB790	0	124	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	13	pDB790	0	129	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	14	pDB790	0	39	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	15	pDB790	0	123	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	16	pDB790	0	40	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	17	pDB790	0	99	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	18	pDB790	0	208	0.0%
<i>rb1</i> E2-E4	pGTag-2A-GalMVP16-bactin	48/48	19	pDB790	0	241	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	1	pDB790	0	49	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	2	pDB790	0	78	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	3	pDB790	0	162	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	4	pDB790	12	25	48.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	5	pDB790	0	79	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	6	pDB790	0	4	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	7	pDB790	0	200	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	8	pDB790	0	38	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	9	pDB790	0	128	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	10	pDB790	0	7	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	11	pDB790	0	119	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	12	pDB790	0	136	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	13	pDB790	0	76	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	14	pDB790	0	159	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	15	pDB790	0	168	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	16	pDB790	0	139	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	17	pDB790	0	25	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	18	pDB790	0	4	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	19	pDB790	0	81	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	20	pDB790	0	71	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	21	pDB790	0	75	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	22	pDB790	0	81	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	23	pDB790	0	76	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	24	pDB790	0	72	0.0%
<i>rb1</i> E2-E25	pGTag-2A-GalMVP16-bactin	48/48	25	pDB790	0	69	0.0%

<i>rb1</i> E2-E25	pGTag-2A-Gal4VP16-bactin	48/48	26	pDB790	0	33	0.0%
<i>rb1</i> E2-E25	pGTag-2A-Gal4VP16-bactin	48/48	27	pDB790	0	74	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	1	pDB790	0	180	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	2	pDB790	0	53	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	3	pDB790	0	0	N/A
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	4	pDB790	0	0	N/A
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	5	pDB790	0	93	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	6	pDB790	0	23	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	7	pDB790	0	357	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	8	pDB790	0	100	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	9	pDB790	0	235	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	10	pDB790	0	237	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	11	pDB790	0	84	0.0%
<i>msno</i> E2-E6	pGTag-2A-Gal4VP16-bactin	48/48	12	pDB790	0	117	0.0%
<i>kdrl</i> E3-E30	pGTag-2A-Gal4VP16-bactin	42/48	1	pDB790	0	250	0.0%
<i>kdrl</i> E3-E30	pGTag-2A-Gal4VP16-bactin	42/48	2	pDB790	0	250	0.0%
<i>kdrl</i> E3-E30	pGTag-2A-Gal4VP16-bactin	42/48	3	pDB790	0	250	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	1	<i>s1pr1</i> E2-E2 Animal 2	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	2	<i>s1pr1</i> E2-E2 Animal 1	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	3	<i>s1pr1</i> E2-E2 Animal 4	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	4	<i>s1pr1</i> E2-E2 Animal 3	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	5	<i>s1pr1</i> E2-E2 Animal 6	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	6	<i>s1pr1</i> E2-E2 Animal 5	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	7	<i>s1pr1</i> E2-E2 Animal 8	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	8	<i>s1pr1</i> E2-E2 Animal 7	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	9	<i>s1pr1</i> E2-E2 Animal 10	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	10	<i>s1pr1</i> E2-E2 Animal 9	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	11	<i>s1pr1</i> E2-E2 Animal 12	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	12	<i>s1pr1</i> E2-E2 Animal 11	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	13	<i>s1pr1</i> E2-E2 Animal 14	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	14	<i>s1pr1</i> E2-E2 Animal 13	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	15	<i>s1pr1</i> E2-E2 Animal 16	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	16	<i>s1pr1</i> E2-E2 Animal 15	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	17	<i>s1pr1</i> E2-E2 Animal 18	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	18	<i>s1pr1</i> E2-E2 Animal 17	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	19	<i>s1pr1</i> E2-E2 Animal 20	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	20	<i>s1pr1</i> E2-E2 Animal 19	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	21	<i>s1pr1</i> E2-E2 Animal 22	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	22	<i>s1pr1</i> E2-E2 Animal 21	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	23	<i>s1pr1</i> E2-E2 Animal 24	0	150	0.0%
<i>s1pr1</i> E2-E2	pGTag-2A-Gal4VP16-bactin	48/27	24	<i>s1pr1</i> E2-E2 Animal 23	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	1	<i>vegfaa</i> E3-E7 Animal 2	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	2	<i>vegfaa</i> E3-E7 Animal 1	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	3	<i>vegfaa</i> E3-E7 Animal 4	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	4	<i>vegfaa</i> E3-E7 Animal 3	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	5	<i>vegfaa</i> E3-E7 Animal 6	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	6	<i>vegfaa</i> E3-E7 Animal 5	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	7	<i>vegfaa</i> E3-E7 Animal 8	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	8	<i>vegfaa</i> E3-E7 Animal 7	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	9	<i>vegfaa</i> E3-E7 Animal 10	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	10	<i>vegfaa</i> E3-E7 Animal 9	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	11	<i>vegfaa</i> E3-E7 Animal 12	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	12	<i>vegfaa</i> E3-E7 Animal 11	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	13	<i>vegfaa</i> E3-E7 Animal 14	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	14	<i>vegfaa</i> E3-E7 Animal 13	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	15	<i>vegfaa</i> E3-E7 Animal 16	0	150	0.0%
<i>vegfaa</i> E3-E7	pGTag-2A-Gal4VP16-bactin	48/48	16	<i>vegfaa</i> E3-E7 Animal 15	0	150	0.0%
<i>mmp14a</i> E1-E10	pGTag-2A-Gal4VP16-bactin	48/48	1	<i>mmp14a</i> E1-E10 Animal 2	0	112	0.0%
<i>mmp14a</i> E1-E10	pGTag-2A-Gal4VP16-bactin	48/48	2	<i>mmp14a</i> E1-E10 Animal 1	0	112	0.0%
<i>mmp14a</i> E1-E10	pGTag-2A-Gal4VP16-bactin	48/48	3	<i>mmp14a</i> E1-E10 Animal 4	0	150	0.0%
<i>mmp14a</i> E1-E10	pGTag-2A-Gal4VP16-bactin	48/48	4	<i>mmp14a</i> E1-E10 Animal 3	0	150	0.0%

Supplementary table 5 - Targeting domain information for all gene targeting experiments. N/A means not applicable

Genomic target	Donor vector	Donor target sequence	5' Spacer sequence	5' Homology length	3' Spacer sequence	3' Homology	3' Homology arm
nr01 E1	2A-TgRFP-CAAX-SV40	GGGAGCGAGAGCTGGAGAGG	GGGAGCGAGAGCTGGAGAGG	12	CAAGCGCTCT	N/A	N/A
nr01 E1	2A-TgRFP-CAAX-SV40	GGGAGCGAGAGCTGGAGAGG	GGGAGCGAGAGCTGGAGAGG	24	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
nr01 E1	2A-TgRFP-CAAX-SV40	GGGAGCGAGAGCTGGAGAGG	GGGAGCGAGAGCTGGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
nr01 E2	2A-TgRFP-CAAX-SV40	GACTGGGAGAGAGCTGGAGAGG	GACTGGGAGAGAGCTGGAGAGG	24	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
nr01 E1	p0Tag-2A-TgRFP-CAAX-SV40	GGGAGCGAGAGCTGGAGAGG	GGGAGCGAGAGCTGGAGAGG	24	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
nr01 E1	p0Tag-2A-eGFP-SV40	GGGAGCGAGAGCTGGAGAGG	GGGAGCGAGAGCTGGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
nr01 E1	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	24	GGGAGCGAGAGCTGGAGAGG	N/A	N/A
ca43.4 E2	p0Tag-2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	24	AAATCTCAACCACTCAACTCTG	888	24
ca43.4 E2	p0Tag-2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	48	GGGAGCGAGAGCTGGAGAGG	ccc	48
ca43.4 E2	p0Tag-2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	20	GGGAGCGAGAGCTGGAGAGG	ccc	20
ca43.4 E2	p0Tag-2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	40	GGGAGCGAGAGCTGGAGAGG	ggg	40
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	18	GGGAGCGAGAGCTGGAGAGG	ggc	18
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	20	GGGAGCGAGAGCTGGAGAGG	aaa	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	24	GGGAGCGAGAGCTGGAGAGG	ggg	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	24	GGGAGCGAGAGCTGGAGAGG	ccc	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	36	GGGAGCGAGAGCTGGAGAGG	aaa	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	40	GGGAGCGAGAGCTGGAGAGG	ggc	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	48	GGGAGCGAGAGCTGGAGAGG	ccc	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	60	GGGAGCGAGAGCTGGAGAGG	ttt	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	72	GGGAGCGAGAGCTGGAGAGG	ccc	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	96	GGGAGCGAGAGCTGGAGAGG	888	N/A
ca43.4 E2	2A-TgRFP-CAAX-SV40	GAGCCATATCTTGCCCAAG	GAGCCATATCTTGCCCAAG	1027	GGGAGCGAGAGCTGGAGAGG	N/A	1027
esrrb E2	p0Tag-2A-GHVP18-biadin	GATGTGTATCTGAGGAGAGAGG	GATGTGTATCTGAGGAGAGAGG	24	GGGAGCGAGAGCTGGAGAGG	ttc	24
nr01 E2	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	ttt	48
nr01 E3	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	ggg	48
nr01 E4	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	ggg	48
ap0b1 E1	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	ggg	48
ap0b1 E1	p0Tag-2A-GHVP18-biadin	GTTCTGTAGAGAGGATGAGAGG	GTTCTGTAGAGAGGATGAGAGG	48	GGGAGCGAGAGCTGGAGAGG	ggg	48
rpl E2-E4	p0Tag-2A-GHVP18-biadin	E2 GTCACAGCAGAGTCACTAAAGG	E2 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	N/A	48
rpl E2-E5	p0Tag-2A-GHVP18-biadin	E2 GTCACAGCAGAGTCACTAAAGG	E2 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	N/A	48
myo E2-E6	p0Tag-2A-GHVP18-biadin	E2 GTCACAGCAGAGTCACTAAAGG	E2 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	N/A	48
klf E3-E8	p0Tag-2A-GHVP18-biadin	E3 GTCACAGCAGAGTCACTAAAGG	E3 GTCACAGCAGAGTCACTAAAGG	42	GGGAGCGAGAGCTGGAGAGG	ccc	42
slpr E2-E2	p0Tag-2A-GHVP18-biadin	E3 GTCACAGCAGAGTCACTAAAGG	E3 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	ccc	48
wdr5 E1-E7	p0Tag-2A-GHVP18-biadin	E3 GTCACAGCAGAGTCACTAAAGG	E3 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	ccc	48
mmp14 E1-E10	p0Tag-2A-GHVP18-biadin	E3 GTCACAGCAGAGTCACTAAAGG	E3 GTCACAGCAGAGTCACTAAAGG	48	GGGAGCGAGAGCTGGAGAGG	ccc	48
ROS26	HME1 donor	TAACTGATCTTGCCCAAG	TAACTGATCTTGCCCAAG	48	GGGAGCGAGAGCTGGAGAGG	ggc	48
ROS36	HME1 donor	TAACTGATCTTGCCCAAG	TAACTGATCTTGCCCAAG	760	GGGAGCGAGAGCTGGAGAGG	N/A	760
ANV1	HME1 donor	GTCACCACTCTTGCCCAAG	GTCACCACTCTTGCCCAAG	48	GGGAGCGAGAGCTGGAGAGG	ccc	48
ANV1	HR donor	GTCACCACTCTTGCCCAAG	GTCACCACTCTTGCCCAAG	804	GGGAGCGAGAGCTGGAGAGG	N/A	804



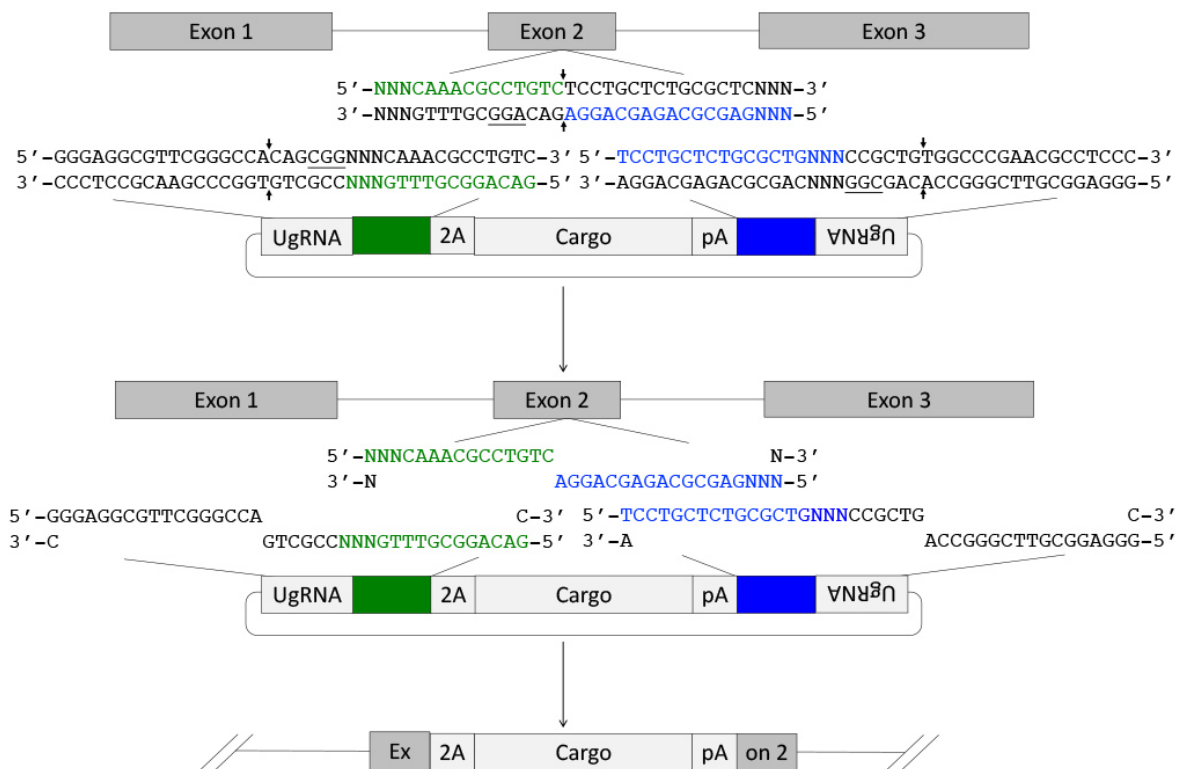
Supplementary table 6 - Primers and oligos		
Primer name	Sequence	Purpose
notojn5'f	GCTTTTGGACTAAACAGCGCCATG	Junction fragment analysis
notojn3'r	cttgctgacacagctccacg	Junction fragment analysis
notojn3'delr	CGATGTTATACTGCTCTTTTAGTTTTGTACATAT	Junction fragment analysis
tyrxjn5'f	CATCTTTGAGCAGTGGTTGAGGAGA	Junction fragment analysis
tyrxjn3'r	esam	Junction fragment analysis
esamajxn5'f	ggctcttcagtcagcgatttaagtc	Junction fragment analysis
esamajxn3'r	CATTTTCAGTCTGGTAGCAGACTG	Junction fragment analysis
cx43.4jxn5'f	GCAGGACTGAGACGGTGGTA	Junction fragment analysis
cx43.4jxn3'r	CAAAATGCATCGTAGCAAAACG	Junction fragment analysis
Rb2Jf	AAGGACAAGGATCTGAGTTTG	Junction fragment analysis
galJf	GCAAACGGCTTAACCTTCC	Junction fragment analysis
GalJR	GCCTTGATTCACTTCTGTCA	Junction fragment analysis
Rb4JR	GCTTTGATCAAACTCAA	Junction fragment analysis
Rb25JR	AGCCAGCTCTGGATCAGT	Junction fragment analysis
RFP5'jxnR	CCTTAATCAGTCTCCGCCCTTAGA	Junction fragment analysis
sv403'jxnR	GGGGAGGTGGGAGGTTTT	Junction fragment analysis
gfp5'R	GCTGAATCTGTGGCGTTTA	Junction fragment analysis
GFPS'F	ACATGCTCTGCTGAGACTT	Junction fragment analysis
F-msna-exon2	TTCTCTCAATCATTCATTGACA	Junction fragment analysis
R-msna-exon2	GGGACGGCTACAAGAAG	Junction fragment analysis
R-msna-exon6	CGTGTGATTGACAGGGTCCAC	Junction fragment analysis
F-msna-exon6	GTTCTGGAGCAGCAAGC	Junction fragment analysis
anxa2a5'Cut5'JxnPrR	CTTCAGACGAGCAGAGTGA	Junction fragment analysis
anxa2a5'Cut5'JxnPrF	TTTCTCAATGTTGGCTGT	Junction fragment analysis
R-1a1junc	agattagaggctgctcg	Junction fragment analysis
R-1a1junc	gctgagaagatgaacagg	Junction fragment analysis
KDRLEx3JxnF	TTGTTTTGTTGACTTCCAAT	Junction fragment analysis
KDRLEx3JxnR	GGGTGGTGGAGTAGCAAGA	Junction fragment analysis
S1PR1.5'JxnF	ACAAGGACCGGGACTCA	Junction fragment analysis
S1PR1.3'JxnR	CGAGACGAAAAGTTACGA	Junction fragment analysis
vegfaaEx3.5'JxnF	AACACTCTGCTTTGCTCC	Junction fragment analysis
vegfaaEx7.3'JxnR	TGCTGTCCGTTTAAAAAGTT	Junction fragment analysis
R-8a1-exon1_junc	tgattccccaaatgct	Junction fragment analysis
F-8a1-exon1_junc	gccctcttgagcactaca	Junction fragment analysis
noto5Bf	CAGATGCCACACTCGGTT	Southern blot analysis
noto5Br	CGATGTTATACTGCTCTTTTAGTTTTGTACATAT	Southern blot analysis
tyr5Be3f	GTTTTGTAATCCTGAGACGGGTTTG	Southern blot analysis
tyr5Be3r	CTGTCAATAAAGCATGATGATGATGAAATGG	Southern blot analysis
esama5Be3f	ATCATCTATCTGTAATGGAGACTTCAG	Southern blot analysis
esama5B4r	CACAGTGTGGCAGTGAAGCTTC	Southern blot analysis
gal45Br	CTGAAGAACACTGGGAGTGTGC	Southern blot analysis
gal45Bf	TTACATATCCAGAGCCGTAGGG	Southern blot analysis
rtp5Bf	ATGGTGTCTAAGGGCAAGAGC	Southern blot analysis
rtp5Br	AGCTTCAGGACCATGTGCC	Southern blot analysis
F-2A-Gal4-BamHI	GGATCCGGAGCCAGCACTCTCTCTGTTAAAGCAAGCAGGAGACGTGGAAGAAAACCCCGGTCTCTctagaAAGCTACTGCTCTTATCGAAC	pGTag vector cloning
R-Gal4-NcoI	gcaTCCATGGTAATTTATTTAGCAGTAGATAGCTATATTTGTGAAACGC	pGTag vector cloning
F-eGFP-SpeI	agacatactagtagtggtagcaaggcgaggagctg	pGTag vector cloning
R-eGFP-XhoI	tggatctcagtagggcgcctctgagactatag	pGTag vector cloning
F-p494-XhoI	cgctgctcagtagggcgcctctgagactatag	pGTag vector cloning
R-p494-SpeI	agcactagtagAGACCCGGGTTTTCTT	pGTag vector cloning
F-lacZ	tgatcacagctgctgcaacgaatgaatgagtt	pGTag vector cloning
R-lacZ	atctgtggcaggtcattgcaatgagctg	pGTag vector cloning
F-lacZ-universal-1	GGGAGGCGTTCCGGCCACAGCGGacagcaggtgcaacgaatgaatg	pGTag vector cloning
R-lacZ-universal-BamHI	ctgatcgatctctgagcagctcattg	pGTag vector cloning
F-lacZ-universal-EcoRI	gtacatgaattGGGAGGCGTTCCGGCCACAG	pGTag vector cloning
F-3'-uni-1	CGTTGCTAGCAAGAGTGAaga	pGTag vector cloning
R-3'-uni-1	GGGAGGCGTTCCGGCCACAGCGGagaagatcattatcagctg	pGTag vector cloning
F-3'-uniNco1	tgcaGccatgCGTTGCTAGCAAGAGTGAaga	pGTag vector cloning
R-3'-uniEagI	tgcaGccgcccAGATCCCATCGTACGCCGG	pGTag vector cloning
F-TagRFPfix	5'phos\cttaactcagctcctccctag	pGTag vector cloning
R-TagRFPfix	5'phos-gagaaatcagcaatgagctgac	pGTag vector cloning
F-BBfix	5'phos\AATACGCAAAACCGCTCTCC	pGTag vector cloning
R-BBfix	5'phos\GGGGCCaCaTCGGCTTC	pGTag vector cloning
F-Bactfix	5'phos\TTTAAAGTCAAACCCATGACTG	pGTag vector cloning
R-Bactfix	5'phos\CTGGCAGTCTCTCTGTTAA	pGTag vector cloning
F-Bact-Ascl	ATATGTGGCGCCAGGACTGTTACCACTCACG	pGTag vector cloning
R-Bact-NcoI	ACAACGCCATGGTAATTTATAGC	pGTag vector cloning
F-gal4-Ecofix	5'phos\AAGATCTCTGAGCCGCC	pGTag vector cloning
R-gal4-Ecofix	5'phos\ATTCGGGCTGCACTCGA	pGTag vector cloning
noto E1 guide	taatacgtactactataGGGAGCGCAGAGTGTGGAGAGCgttttagactagaa	gRNA oligo A
noto E2 guide	taatacgtactactataGACTGGAGAAAGATTCGCGgttttagactagaa	gRNA oligo A
tyr E4 guide	taatacgtactactataGTTCTCTAGAGAGGGATGAgtttagactagaa	gRNA oligo A
cx43.4 E2 guide	taatacgtactactataGGCCATATCTCCCAAGgttttagactagaa	gRNA oligo A
esama E2 guide	taatacgtactactataGGGATGATCCAGGAAAGgttttagactagaa	gRNA oligo A
msna E2 guide	taatacgtactactataGGTTCCTGTGGTGGTgttttagactagaa	gRNA oligo A
msna E6 guide	taatacgtactactataGGCTGGCAGGAGGAGCAAGgttttagactagaa	gRNA oligo A
anxa2a E3 guide	taatacgtactactataGGTGGATCAAAGTGGCTCgttttagactagaa	gRNA oligo A
flna E4 guide	taatacgtactactataGGATTGAGTTTGGGCTCAGgttttagactagaa	gRNA oligo A
aqp1a1 E1	taatacgtactactataGGCTTCTGGCGGCCCTCgttttagactagaa	gRNA oligo A
aqp8a1 E1	taatacgtactactataGGTGTCCGTCTGCGCAAGgttttagactagaa	gRNA oligo A
rb1 E2 guide	taatacgtactactataGGAGAGGAGATCAGATGAGgttttagactagaa	gRNA oligo A
rb1 E4 guide	taatacgtactactataGGCAGCAGAGATTCATCTgttttagactagaa	gRNA oligo A
rb1 E25 guide	taatacgtactactataGGCAAGCGCACTCTCgttttagactagaa	gRNA oligo A
kdr1 E3 guide	taatacgtactactataGGCTAGAGAAATGTCTAGgttttagactagaa	gRNA oligo A
kdr1 E30 guide	taatacgtactactataGGAGACTGCTCCCTGACTgttttagactagaa	gRNA oligo A
s1pr1 E2a guide	taatacgtactactataGGAGAACCAACAATTTCTgttttagactagaa	gRNA oligo A
s1pr1 E2b guide	taatacgtactactataGGGAAACATAGTCTTCTgttttagactagaa	gRNA oligo A
vegfaa E3 guide	taatacgtactactataGGGGCAAGACCCGAGAGCTGgttttagactagaa	gRNA oligo A
vegfaa E7 guide	taatacgtactactataGGGTTGCTCGATCATCTgttttagactagaa	gRNA oligo A
mmp14a E1 guide	taatacgtactactataGGACGAAACAACACTGTGgttttagactagaa	gRNA oligo A
mmp14a E10 guide	taatacgtactactataGGGAGCCGGCACAGATGAGGgttttagactagaa	gRNA oligo A
CRISPR Oligo B	AAAAGCACCGACTGGGTGCCACTTTTCAAGTTGATAAAGGACTAGCCCTTTTAACTTGCTATTCTAGCTTAAAC	Oligo B for gRNA synthesis
noto E1 5' arm bait A	12 aattcGGGAGCGCAGAGTGGAGACAGGCAACGCTGTCTGcg	Homology arm oligo
noto E1 5' arm bait B	12 gatccGACAGAGCTTTGCTGTCTCCAGCTCTGCGCTCCCg	Homology arm oligo
noto E1 5' arm bait A	24 aattcGGGAGCGCAGAGTGGAGACAGGACATAACCAACCAAAACGCCTGTCTGcg	Homology arm oligo
noto E1 5' arm bait B	24 gatccGACAGGCGTTTGGTTGGTTATGCTCTGCTCTCCAGCTCTGCGCTCCCg	Homology arm oligo
noto E1 5' arm bait A	48 aattcGGGAGCGCAGAGTGGAGACAGGAGAAAGCAAAACCGCTACCCGAGCATAAACCAACCAAAACGCCTGTG	Homology arm oligo
noto E1 5' arm bait B	48 gatccACAGGCGTTTGGTTGGTTATGCTCCGGTACGCGTTTGTCTGCTCTCTGCTGTCCAGCTCTGCGCTCCCg	Homology arm oligo
noto E1 5' arm A 24	aattcGGGAGCGCAGAGTGGAGACAGGAAAGCAAAACCAACCAAAACGCCTGTCTGcg	Homology arm oligo
noto E1 5' arm B 24	gatccGACAGGCGTTTGGTTGGTTATGCTCTCTGCTCTCCAGCTCTGCGCTCCCg	Homology arm oligo
noto E1 3' arm A 24	caatggTCCAGCTCTGCGCTCCCGCTTATTCGCCGTGGCCCGAAGCCCTCCC	Homology arm oligo
noto E1 3' arm B 24	ggcgcGGAGGCGCTCCGGCCACAGCGGAAATGAGCGGAGCGCAGAGCTGGAC	Homology arm oligo
noto E1 5' arm A 48	CGGGaaaAGAGAACGAAACAAACCGCTACCCGAGCATAAACCAACCAAAACGCCTGTCTG	Homology arm oligo
noto E1 5' arm B 48	atccGACAGGCGTTTGGTTGGTTATGCTCCGGTACCCGCTTGTCTGCTCTCTctt	Homology arm oligo
noto E1 3' arm A 48	aagTCCAGCTCTGCGCTCCCGCTTATTCAGCTCCGAGTGCACACTTCCCGggg	Homology arm oligo



## Gene Targeting Protocol for Integrations with pGTag Vectors using CRISPR/Cas9

### Targeting strategy (Figure 1):

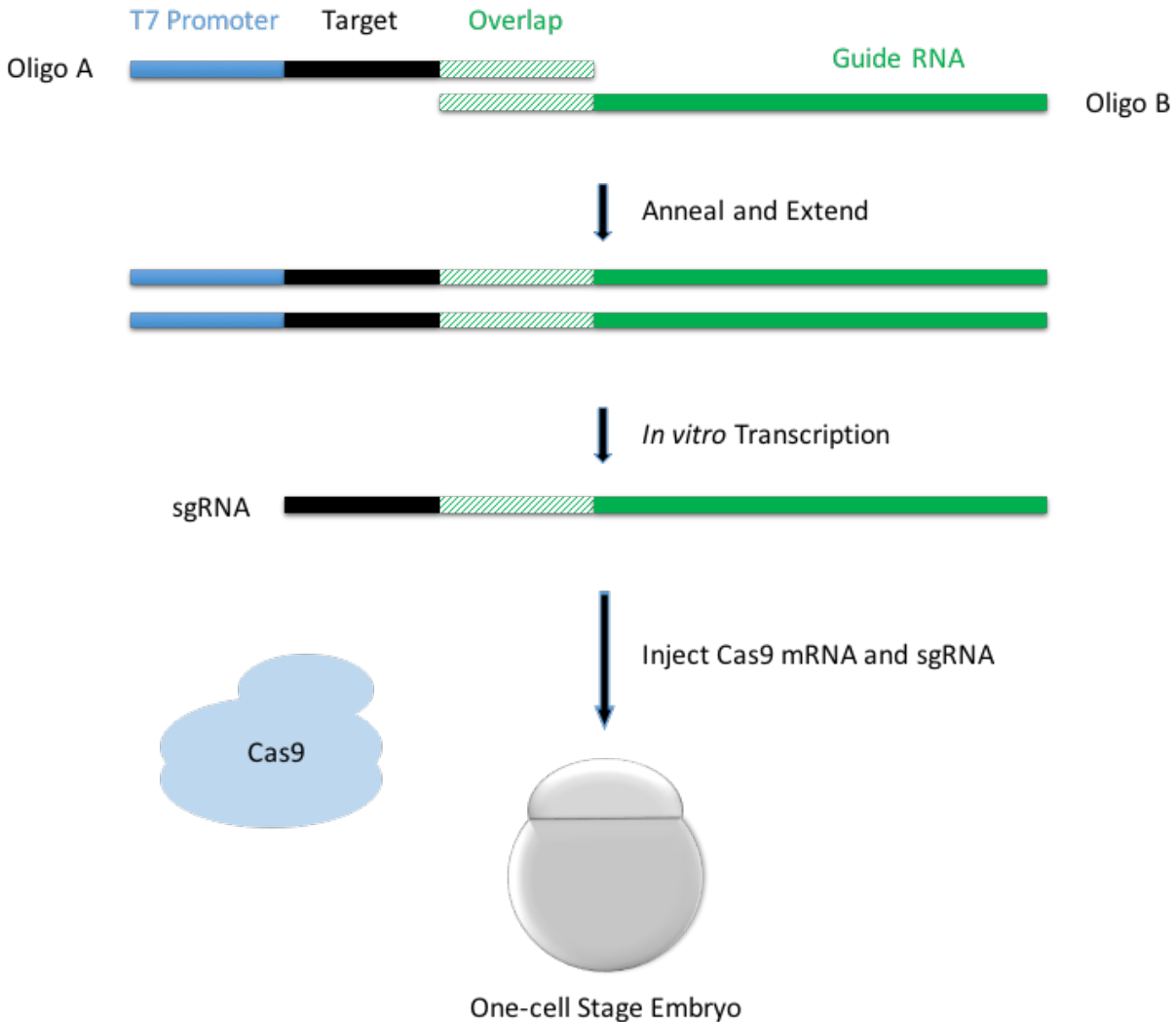
- A. Selection of a CRISPR/spCas9 target site downstream of the first AUG in the gene of interest
- B. Synthesize sgRNA and spCas9 mRNA
- C. Injection of sgRNA and spCas9 mRNA
- D. Testing for indel production/mutagenesis
- E. Design short homology arms
- F. One Pot Cloning of Homology Arms into pGTag Vectors
- G. Injection of GeneWeld reagents (spCas9 mRNA, Universal sgRNA (UgRNA), genomic sgRNA and pGTag homology vector) into 1-cell zebrafish embryos
- H. Examine embryos for fluorescence and junction fragments



**Figure 1.** Targeting integration of the pGTag vectors into the 5' region of a gene. Upon CRISPR/Cas9 targeting and cutting of both the genome (with a sgRNA) and plasmid donor (with UgRNA), the genomic and plasmid DNA likely undergo end resection mediated by the MRN complex and *Exo1*, resulting in annealing of complementary homology arms. This promotes precise homology-directed integration of cargo DNA at the CRISPR/Cas9 double-strand break.

- A. Selection of a CRISPR/spCas9 target site downstream of the first AUG in the gene of interest

1. To select a CRISPR/Cas9 target site in a 5' exon, find and download the targeted gene's genomic and coding sequences.
  - a. At <ensemble.org> Search for the gene name of interest for the species of interest and open the Transcript page.
  - b. In the left-hand side bar click on "Exons" to find the first coding exon and initiation ATG. If there are alternative transcripts for the gene, make sure there are not alternative initiation ATGs. If there are alternative start codons, target the first 5' exon that is conserved in all transcripts to generate a strong allele.
  - c. Download the transcript and 5' exon to be targeted as separate sequence files.
  - d. Using ApE: <<http://biologylabs.utah.edu/jorgensen/wayned/ape/>> annotate the coding sequence with the exons.
  
2. Use CRISPRScan (<http://www.crisprscan.org/>) (Moreno-Mateos et al., 2015) to efficiently identify target sites and generate oligos for sgRNA synthesis for the target gene.
  - a. Select the "Predict gRNAs" on the lower right-hand side of the home page of the CRISPRScan website.
  - b. Paste the 5' exon sequence into the indicated box. If the exon is very large, start with a small amount of sequence. Ideally exon sequence of ~200 bp near the desired target site. Do not design CRISPRs to intron/exon borders. If there are problems with the copy and pasting of exon sequence, first paste the sequence into a new ape file, save, then copy and paste from the new file.
  - c. Select "Zebrafish (Danio rerio)" as the species
  - d. Select "Cas9 – nGG" as the enzyme.
  - e. Select "In vitro T7 promoter".
  - f. Click on "Get sgRNAs." Examine the output. The generated targets are ranked by CRISPRScan from high to low. Select a target site (the 20 bp that are capitalized in the oligo column) from those given by CRISPRScan using the following criteria (The best gRNAs will have all of these):
    - i. An exact match to the genomic locus., When an oligo is clicked on the page will display additional information to the right. In the section called "Site Type" any mismatches in the oligo are displayed. Exact matches including 5'GG- are ideal for in vitro transcription and 100% genomic target match.
    - ii. The target is in the desired location of the gene.
    - iii. The Target is on the reverse (template) strand. Reverse strand guides are more favorable, but either will work
    - iv. A high CRISPRScan score, and a lower CFD score. However, lower score sgRNA targets may work fine.
  - g. Annotate the selected target sequence in the transcript sequence files.
  - h. For sgRNA synthesis the entire oligo sequence from CRISPRscan containing the selected target will need to be synthesized. This oligo is represented as "Oligo A" in Figure 2.



**Figure 2.** Cloning-free gRNA synthesis. Oligo A is composed of the T7 promoter at the 5' end, target sequence for gRNA, and gRNA overlap sequence for gRNA synthesis. CRISPRScan provides direct output for Oligo A. The strategy for gRNA production using Oligo A is based on (Varshney et al., 2015).

3. **Alternative to CRISPRScan:** Designing “CRISPR Oligo A” from a genomic target sequence. Skip this section if Oligo A was designed with CRISPRScan.

If the target sequence was identified using tools other than CRISPRScan, Oligo A can be designed manually. (Note: CRISPRScan will use a shorter overlap region but this does not affect template production). Add T7 and Overlap sequences (see Figure 2) to the 20 bp of target sequence without the PAM. Oligo A for the targeted gene will look like the example below:

5'-TAATACGACTCACTATAGGNNNNNNNNNNNNNNNNNNNNNNGTTTTAGAGC  
TAGAAATAGC-3'

The sequences in blue (first 17 characters) are the T7 promoter, the grey GG are part of the T7 promoter and ideally are part of the target sequence (see below), the Ns are the target

sequence, and the sequence in green (last 20 characters) are the overlap region to synthesize the non-variable part of the sgRNA. The T7 promoter works optimally with the two grey GGs, however, these GGs will be transcribed by T7 and thus become a part of the sgRNA. Target sequences that contain the GGs may work better, but there are differing reports in the literature on the importance of this (Moreno-Mateos et al., 2015). If possible, select a target that starts with GG. Refer to Moreno-Mateos et al., 2015 for other gRNA architectures with variations on the 5'GG motif.

- a. If the target sequence did not have two Gs at the beginning, additional G's will need to be added to the start of the target sequence for efficient transcription as outlined below:

\*The lower case 'g' is an extra 'G' not in the genomic sequence; the upper-case G is in the genomic. Lower case gs will not base pair with the genomic target.

- i. without GG: ggN NNN NNN NNN NNN NNN NNN N (22 bp) – 2 bases are added,
- ii. with one G: gGN NNN NNN NNN NNN NNN NNN (21 bp) – one base is added, G is part of the target sequence.
- iii. with two G: GGN NNN NNN NNN NNN NNN NN (20 bp) – no bases are added, GG is part of the target sequence.

Oligo A is made by taking this target sequence with 5'GG and pasting it into a clean file.

- b. Copy and paste the T7 promoter sequence to the 5' end of the target sequence:  
TAATACGACTCACTATA
  - c. Copy and paste the Overlap sequence to the 3' end of the target sequence:  
GTTTTAGAGCTAGAAATAGC
  - d. Check the sequences to ensure they are correct and that the PAM is NOT present in this oligo.
4. Oligo B design (Figure 2) contains the conserved guide RNA sequence: All Oligo Bs will be the same and can be ordered in large quantities.

5'-GATCCGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTT  
AACTTGCTATTTCTAGCTCTAAAAC-3'

5. To increase yield of the sgRNA synthesis the primers "T7 primer" (5'-TAATACGACTCACTATA-3') and "3'gRNA primer" (5'-GATCCGCACCGACTCGGTG-3') are also required.
6. For checking for mutagenesis at the target site, design ~20 bp DNA primers for PCR amplification to amplify at least 130 bp of DNA surrounding the target site. Mutagenesis is estimated through comparison of PCR products from injected and uninjected embryos, by visualizing small insertions and/or deletions (Indels) using electrophoresis, or by sequencing.
  - a. Primer 3 is used for primer design:  
[http://biotools.umassmed.edu/bioapps/primer3\\_www.cgi](http://biotools.umassmed.edu/bioapps/primer3_www.cgi)
  - b. Paste DNA sequence surrounding the target site into the web interface. It is recommended to use 160 – 300 bp of exon sequence centered on the cut site for

primer design. Intron sequence can be used, but this often contains polymorphisms that can lead to amplification failure.

- c. Locate the target sequence, including the PAM sequence (*italicized* below), and predict the cut site (3 bp into the target sequence from the PAM represented here by the 'x'). Mark the targeted exon sequence approximately 65-150 bp on both sides of the cut site by putting [square brackets] around it. Primer3 will design primers outside this sequence. This design allows the primers to be used for both checking of mutagenesis and for junction fragment analysis when checking for integration.

Example:

```
CGGCCTCGGGATCCACCGGCC[AGAATCGATATACTACGATGAACAGAGCAAATTT  
GTGTGTAATACGGTCGCCACCATGGCCTx CCTCGGTTTGCTACGATGCATTTGCAC  
CACTCTCATGTCCGGTTCTGGG]AGGACGTCATCAAGGAGTTCATGCGCTTCAA  
GGTGCGCATGGAGGGCTCCGTGAAC
```

- d. Set the “Primer Size” variables to Min = 130, Opt = 170, and Max = 300. Everything else can be left at the defaults.
- e. Click on “Pick Primers”
- f. Select primers from the output. Note the “product size” expected and the “tm” or melting temperature of each primer/pair. Smaller product sizes are easier to visualize mutagenesis.

## B. Synthesize the sgRNA

General guidelines and good laboratory practices for working with DNA and RNA. DNA, RNA and the enzymes are sensitive to contamination from dust and skin. Following these guidelines will prevent the degradation of the DNA and RNA you are trying to make:

- Be clean. Clean the workbench, pipetmen, racks, and centrifuges with RNase Away or something equivalent.
- Wear gloves and change when contaminated. Contamination will occur when gloves contact hair, face, skin, or the floor.
- Keep everything on ice unless the protocol indicates otherwise.
- Centrifuge components to the bottom of the tube before use, after mixing, after use, and after incubation steps.
- Do not vortex enzymes. Gently flick the tube or pipet up and down to mix samples.
- Avoid touching the walls of the tube when pipetting.
- Use a new pipette tip for each new dip.
- Dispense solutions from a pipet to the bottom of the tube, or into the liquid at the bottom of the tube when setting up reactions.
- Only remove 1.5 ml centrifuge tube and PCR tubes from their package while wearing gloves. Reseal the tube package after tubes are removed.

### ***Assembly of CRISPR Oligos A + B into a Transcription Template***

1. For synthesis of the gRNA from Oligo A and B, make a 100  $\mu$ M freezer stock and 1  $\mu$ M working stock for each oligo. All oligos are described in Section A starting on page 7.

2. Centrifuge ordered oligos briefly before opening, to move all dried DNA flakes to the bottom of the tube.
3. Add a volume (x  $\mu\text{L}$ ) of RNase-free water to make a 100  $\mu\text{M}$  stock. The tubes should be labeled with the gene name as well as the number of nmol in the tube. The amount of water to be added will need to be calculated based on the nanomoles of material contained within.
4. Vortex for 30 seconds.
5. Centrifuge briefly.
6. Make a 100-fold dilution of each 100  $\mu\text{M}$  stock Oligo A and B in separate 1.5 ml tubes.
  - a. Label one 1.5 mL centrifuge tube per Oligo A with name of oligo, date, and "1  $\mu\text{M}$ " to indicate working stocks.

1  $\mu\text{L}$  of 100  $\mu\text{M}$  Oligo A stock or Oligo B  
99  $\mu\text{L}$  of RNF-water  
100  $\mu\text{L}$  total

- b. Vortex.
  - c. Briefly centrifuge.
  - d. Store all stocks in freezer at  $-20\text{ }^{\circ}\text{C}$  for long-term storage.
7. Set up the following reaction in PCR tubes. The next two steps will generate a short segment of DNA (gDNA or guideDNA Template) which will be used as a template for synthesis of RNA:

12.5  $\mu\text{L}$  2X KOD Master Mix  
1  $\mu\text{L}$  Oligo A (1  $\mu\text{M}$ )  
1  $\mu\text{L}$  Oligo B (1  $\mu\text{M}$ )  
1  $\mu\text{L}$  T7 primer (10  $\mu\text{M}$ )  
1  $\mu\text{L}$  gRNA 3' primer (10  $\mu\text{M}$ )  
8.5  $\mu\text{L}$  RNF-water  
25  $\mu\text{L}$  total

8. Run PCR under the following conditions:

Denature at  $98^{\circ}\text{C}$  for 2 minutes  
Denature at  $98\text{ }^{\circ}\text{C}$  for 30 sec.  
Anneal at  $50\text{ }^{\circ}\text{C}$  30 sec.  
Extend at  $70\text{ }^{\circ}\text{C}$  30 sec.  
Go to (step 2) nine times.  
Extend at  $70\text{ }^{\circ}\text{C}$  2 min  
Hold  $4\text{ }^{\circ}\text{C}$  forever.

9. Run 1.2% agarose gel in 1X TAE to check that the template was synthesized:
  - a. Remove 3  $\mu\text{L}$  of the reaction and place in a 1.5 ml tube.
  - b. Mix in 1  $\mu\text{L}$  of 6x loading buffer.
  - c. Load all 4  $\mu\text{L}$  of the sample on the gel. Run the gel at 125 V for 30 minutes. Be sure



- to load a molecular weight marker.
- d. Check on the transilluminator and image the gel.
- e. A single 120 bp band should be detected when 3  $\mu$ L is loaded on gel.

### ***In vitro transcription (IVT) using the gRNA template***

1. Use the Ambion T7 Megascript Kit for transcription reagents, but follow the instructions below.
2. Thaw the T7 10X Reaction Buffer and RNF-water at room temperature, and thaw the ribonucleotides solutions on ice.
3. Vortex the T7 10X Reaction Buffer to make sure all DTT is solubilized. No white flecks should be visible.
4. Microcentrifuge all reagents briefly before opening to prevent loss of reagents and/or contamination by materials that may be present around the rim of the tube(s).
5. Keep the T7 Enzyme Mix on ice or in a -20 °C block during assembly of the reaction.
6. Make a master mix for each reaction. Assemble the reaction at room temperature on the bench. Add reagents from largest to smallest volume, adding the 10X Reaction Buffer second to last and the T7 Enzyme Mix last.

*Note:* Components in the transcription buffer can lead to precipitation of the template DNA if the reaction is assembled on ice. If the reaction precipitates, the synthesis reaction will not fully occur.

#### 7. Reagent list:

- 10  $\mu$ L of RNF-water
- 5  $\mu$ L of gDNA template (100 to 500 ng total)
- 4  $\mu$ L of NTP (1  $\mu$ L of each; A, U, C, G)
- 1  $\mu$ L of 10x transcription buffer – must be fully resuspended at room temp
- 1  $\mu$ L of T7 polymerase enzyme mix

8. Incubate at 37 °C for 4 to 16 hours. Longer incubations result in considerably better yields.
9. Add 1  $\mu$ L of Turbo DNase and incubate for 15 min at 37 °C. This will digest the template DNA in the sample.
10. Optional quality control step: Run 2  $\mu$ L of sample on a 1.2% gel in 1X TAE.
  - a. Clean the gel box, comb and tray with RNase Away, rinse with DI water.
  - b. Remove 2  $\mu$ L of sample into a clean 1.5 ml (Keep RNA on ice!)
  - c. Add 3  $\mu$ L of RNF-water and 5  $\mu$ L of Ambion RNA loading buffer with formamide.
  - d. Vortex briefly.
  - e. Spin down samples briefly.
  - f. Run all of this mixture on a 1.2% agarose gel/1X TAE, at 100 V for 1 hour.

- g. Image gel. 2 bands should be visible at ~100 and 120 bp.

### ***Purification of guide RNA***

1. Use the miRNeasy Qiagen kit for purification of gRNAs according to the manufacturer's instructions.
2. After Purification verify presence of RNA by running a 1.2% gel in 1X TAE.
3. Clean the gel box, comb and tray with RNase Away, rinse with DI water. Run on a 1.5% agarose gel/1X TAE, at 100 V for 1 hour as above.
4. Image gel. 2 bands should be visible at ~100 and 120 bp.
5. Nanodrop the RNA sample to determine the concentration.
6. Store RNA at -20 °C.

### ***Preparation of SpCas9 mRNA***

1. Digest ~5-10 µg pT3TS-nCas9n plasmid with Xba1 (plasmid Addgene #46757 (Jao et al., 2013)).
2. Purify digested DNA with Qiagen PCR cleanup kit or Promega PureYield Plasmid Miniprep System. Elute in RNF-water.
3. Run 100-500 ng on 1.2% agarose gel in 1X TAE to confirm the plasmid is linearized.
4. Use 100 ng to 1 µg DNA as template for in vitro transcription reaction.
5. Use mMMESSAGE mMACHINE T3 kit Life Technologies (AM1348) and follow the instructions in the manual.
6. Use the miRNeasy Qiagen kit for purification of nCas9n mRNA according to the manufacturer's instructions.
7. Verify mRNA integrity by mixing 1 µL of purified Cas9, 4 µL of RNF water, 5 µL glyoxyl dye (Ambion).
8. Heat mixture at 50 °C for 30 minutes, then place on ice.
9. Clean the gel box, comb and tray with RNase Away, rinse with DI water.
10. Run all 10 µL of RNA mixture on 1.2% agarose gel in 1X TAE at 100 V for 1 hour as above. One band should be visible at 4.5 kb.
11. Nanodrop the RNA sample to determine the concentration. Concentrations between 0.45 and 1 µg/µL are expected.
12. Aliquot and store RNA at -80 °C.

### **C. Injection of sgRNA and spCas9 mRNA**

The injections here are designed to deliver 25 pg of gRNA and 300 pg of Cas9 mRNA in 2 nL of

fluid to embryos at one-cell stage.

Injection trays are cast with 1.2% agarose with 1X embryo media (Zebrafish Book; [zfin.org](http://zfin.org)) in polystyrene petri dishes (Fisher No. FB0875713). Injection trays can be used multiple times and stored at 4°C for up to three weeks between use.

1. Trays are pre-warmed to 28.5 °C prior to injection by placing them in the 28.5 °C incubator. Try to mitigate tray cooling while not in use.
2. Glass needles are pulled from Kwik-Fil borosilicate glass capillaries (No. 1B100-4) on a Flaming/Brown Micropipette puller (Model P-97).

Injection samples are made to contain the following diluted in RNF water or injection buffer (final concentration: 12.5 mM HEPES pH 7.5, 25 mM Potassium Acetate, 37.5 mM Potassium Chloride, 0.0125 % glycerol, 0.025 mM DTT pH 7.5)

- a. 12.5 ng/μL of genomic gRNA
  - b. 150 ng/μL of mRNA for Cas9
3. Needles are loaded with 1.5 to 2.5 μL of sample, and then loaded onto a micro-manipulator attached to a micro injector (Harvard Apparatus PLI - 90) set to 30-40 PSI with an injection time of 200 msec.
  4. Needles are calibrated by breaking the end of the tip off with sterile tweezers, ejecting 10 times to produce a droplet of fluid, collecting the droplet into a 1 μL capillary tube (Drummond No. 1-000-0010), and measuring the distance from the end of the capillary to the meniscus of the droplet. This distance is converted to volume (where 1 mm = 30 nL) and adjusted to achieve an effective injection volume of 2 nL. Volumes are adjusted by changing the injection time. There is a linear relationship between volume and time at a set pressure. Avoid injection times less than 100msec and over 400 msec.
  5. One cell embryos that have been collected from mating cages are pipetted from collection petri dishes to the wells on the injection tray.
  6. Use the micro-manipulator and microscope to pierce the one-cell of embryos on the injection tray at an angle of 30° with the needle. Inject 2 nL of sample in the one-cell near the center of the cell-yolk boundary.
  7. After embryos have been injected, wash them from the injection tray into a clean petri dish with embryo media.
  8. Keep 20 - 40 embryos separate as uninjected controls. Treat and score the control embryos in the same way as the injected embryos.
  9. At 3 - 5 hrs post injection remove any unfertilized or dead embryos from the dishes. This will prevent death of the still developing embryos.

#### **D. Testing for indel production/mutagenesis**

##### ***Phenotypic scoring of embryos***

1. The gRNA itself may be toxic to the developing embryos. Injection toxicity can be estimated by the number dead embryos from a round of injection compared to the un-injected control dish. Count and remove any brown/dead embryos from injected and un-injected dishes. If there are significantly more dead embryos in the injected dish then the guide may be toxic, impure, or very effective at disrupting a required gene. Reducing the amount of guide or Cas9 mRNA injected may help reduce toxicity.
2. Score and document embryonic phenotypes on days 1 - 4 post fertilization (dpf). Under a dissection microscope examine the un-injected controls and injected embryos, and sort the embryos into categories.
3. Scoring categories
  - -Severe- These embryos have some parts that look like a control embryos, but are missing key features. Examples: embryos that lack their head, eyes, or tail, or embryos that have an unnaturally contorted shape or are asymmetric.
  - -Mild- These embryos appear mostly normal, but have slight defects such as small eyes, pericardial edema, shortened trunk/tail, or curled/twisted tails.
  - -Normal- appears normal and similar to controls.

### ***Digestion of embryos for isolation of genomic DNA for mutagenesis analysis***

Genomic DNA (GDNA) can be isolated from zebrafish embryos aged between 1 and 5 dpf using this protocol. Embryos can be analyzed as individuals or as pools (maximum 5) from the same injection.

1. Dechorionate embryos, if they have not emerged from the chorion.
2. It is recommended to screen a minimum of 3 embryos from each scoring category for mutagenesis. Place each embryo, including controls, into separate PCR tubes. Remove as much of the fish water as possible. If needed, spin briefly and remove additional water.
3. Add 20  $\mu$ L of 50 mM NaOH per embryo.
4. Heat the embryos at 95°C in a thermocycler for 15 minutes.
5. Vortex samples for 10 seconds. Be sure that the tubes are sealed to prevent sample loss while vortexing.
6. Spin samples down and heat for an additional 15 min at 95°C in a thermocycler.
7. Vortex samples and then spin the tubes down again. The embryos should be completely dissolved.
7. Neutralize the samples by adding 1  $\mu$ L of 1 M Tris pH 8.0 per 10  $\mu$ L NaOH. Mix by vortexing then spin down.
8. Genomic DNA should be kept at 4°C while in use and stored at -20°C.

### ***Analysis of CRISPR/Cas9 mutagenesis efficiency at targeted gene locus.***

1. Set up the following PCR reactions for each tube of digested embryos using the primers designed at the end of section A, page 10.

12.5  $\mu$ L of 2x GoTaq Mastermix  
1  $\mu$ L of Forward Primer (10  $\mu$ M)  
1  $\mu$ L of Reverse Primer (10  $\mu$ M)  
1  $\mu$ L of gDNA template (digested embryos)  
9.5  $\mu$ L of nuclease-free water  
25  $\mu$ L total

2. Vortex and briefly spin down the PCR reactions.
3. Run the following PCR program to amplify the targeted locus.

95°C 2 minutes  
95°C 30 seconds ]  
55°C\* 30 seconds ] x 35 cycles  
72°C 30 seconds ]  
72°C 5 minutes  
4°C hold

\*if the primers were designed with higher or lower  $t_m$ 's than the annealing temperature in line three, then that temperature will need to be adjusted to 2°C below the designed primer  $t_m$ .

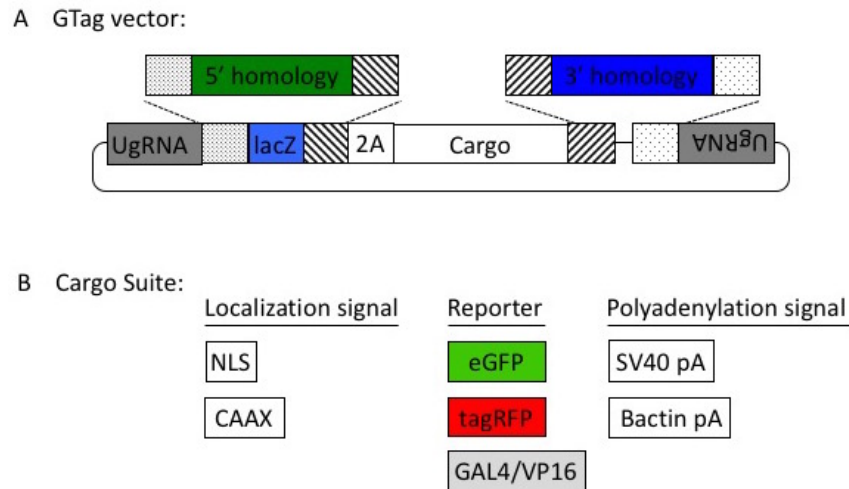
4. Run up to 7  $\mu$ L of PCR product on a 3.0% agarose gel, 1X TAE, for 1 hr at 80-100V.
5. Analyze the gel for DNA bands that appear diffuse or different in size from the control lane. This indicates that the presence of indels in the gene of interest
6. Alternatively clone and sequence PCR products or sequence them directly to verify the presence of indels.

### **E. Design short homology arms**

Homology directed gene targeting allows the integration of exogenous DNA into the genome with precision to the base pair level. However, designing and cloning individual targeting vectors and homology arms for each gene of interest can be time consuming. The pGTag vector series provides versatility for ease of generation of knockout alleles (Figure 3). The vectors contain BfuAI and BspQI type II restriction enzymes for cloning of short homology arms (24 or 48 bp) using Golden Gate cloning. The pGTag vectors require in-frame integration for proper reporter gene function. The reporter gene consists of several parts. First, a 2A peptide sequence causes translational skipping, allowing the following protein to dissociate from the locus peptide. Second and third, eGFP, TagRFP, or Gal4VP16 coding sequences for the reporter protein have a choice of sequence for localization domains, including cytosolic (no) localization, a nuclear localization signal (NLS), or a membrane localization CAAX sequence. Finally translation is terminated by one of two different polyadenylation sequences (pA); a  $\beta$ -actin pA from zebrafish or the SV40pA.

For many genes, the signal from integration of the report protein is too weak to observe. In these cases the Gal4VP16 vector allows for amplification of the report to observable expression levels in F0s and subsequent generations. A 14XUAS/RFP Tol2 plasmid is provided to make a transgenic line for use with the Gal4VP16 vector.

Sequence maps for these plasmids can be downloaded at [www.genesculpt.org/gtaghd/](http://www.genesculpt.org/gtaghd/)



**Figure 3.** The pGTag vectors allow one step cloning of homology arms.

All vectors can be obtained through Addgene ([www.addgene.org](http://www.addgene.org)). Because the pGTag plasmids contain repeated sequences, they may be subject to recombination in certain strains of bacteria. **It is strongly recommended that they are propagated at 30°C to reduce the possibility recombination.**

The web tool, GTagHD [www.genesculpt.org/gtaghd/](http://www.genesculpt.org/gtaghd/), allows for quick design of cloning ready homology arm oligos for a gene of interest. To use the tool, choose the "Submit Single Job" tab. Follow the instructions in the tab.

There should be 4 oligos (two pairs that will be annealed) generated that should be ordered for cloning. If there are any problems with the sequences and values that were entered, the web page will display the errors and give advice on how to fix them.

The following protocol describes how to design homology arm oligos manually:

*\*Note\** In the following section when orientation words are used, they are used in the context of the reading frame of the genetic locus of interest. Example: A 5' template strand CRISPR means that the target site for the CRISPR is on the template strand at the locus and is toward the 5' end of the gene. Upstream homology domains are 5' of the CRISPR cut and downstream homology domains are 3' of the cut with respect to the gene being targeted. Also note: Upper case and lower case bases are not specially modified; they are typed the way they are as a visual marker of the different parts of the homology arms.

### ***For the Upstream Homology Domain***

1) Open the sequence file for the gene of interest and identify the CRISPR site. (In this example it is a Reverse CRISPR target in Yellow, the PAM is in Orange, coding sequence is in purple)

Copy the 48 bp 5' of the CRISPR cut (the highlighted section below) into a new sequence file; this is the upstream homology.

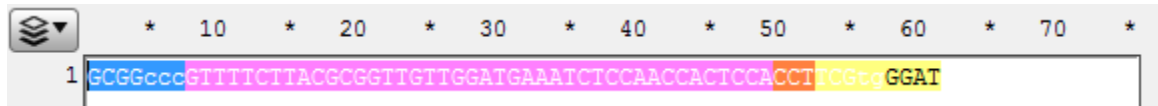
2) Observe the next three bases immediately upstream of the 48 bp of homology, and pick a base not present to be the 3 bp spacer between the homology and the Universal PAM in the vector. (Here the three bases are “GGA” so “ccc” was chosen for the spacer)

Add the spacer to the new file 5' (in front) of the homology, see below. The spacer acts a non-homologous buffer between the homology and the eventual 6 bp flap from the universal guide sequence that will occur when the cassette is liberated and may improve intended integration rates over MMEJ events.

3) Determine where the last codon is in the homology. Here the 3' G in the homology domain is the first base in the codon cut by this CRISPR target. Complete the codon by adding the remaining bases (called padding on GTagHD) for that codon from your sequence to ensure your integration event will be in frame.

4) Add the BfuAI enzyme overhang sequences for cloning, to the ends of the homology domain. 5'-GCGG and 3'-GGAT. (Here both overhangs are added to prevent errors in copying sequence for the oligos in the next two steps.)

5) The Upstream Homology Oligo A will be this sequence from the beginning to the end of the last codon (see highlighted below). Copy and paste this sequence into a new file and save it. In this example this oligo sequence is 5'-GCGGcccGTTTTCTTACGCGGTTGTTGGATGAAATCTCCAACCACTCCACCTTCGtg-3'.

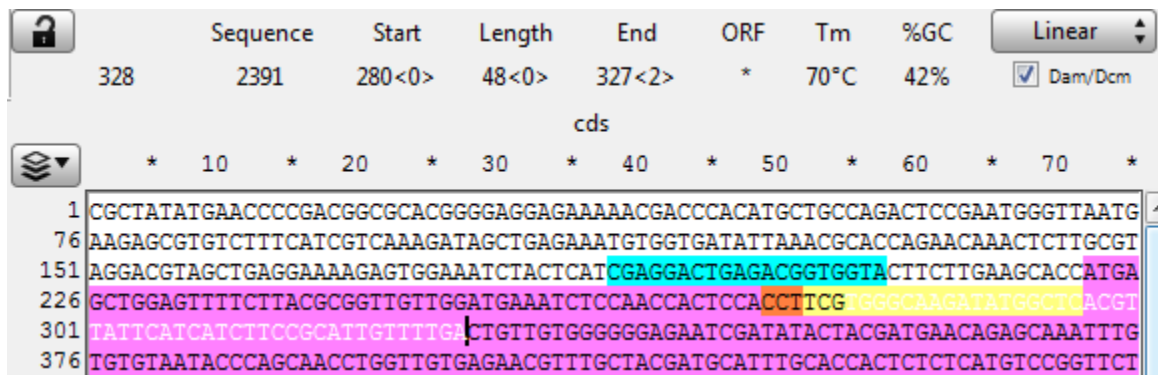


6) The Upstream Homology Oligo B will be the reverse complement of this sequence from beginning of the spacer to the end of the sequence (see highlighted below). Copy the reverse complement, paste it into a new file, and save it. In this example this oligo sequence is 5'-ATCCcaCGAAGGTGGAGTGGTTGGAGATTTTCATCCAACAACCGCGTAAGAAACggg-3'.

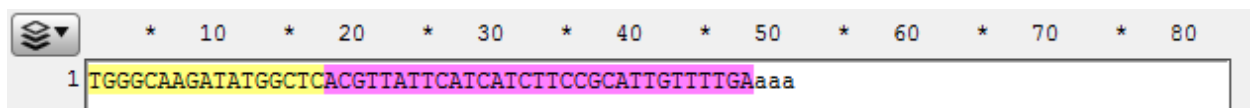


### For the Downstream Homology Domain

7) Open sequence file for the gene of interest and identify the CRISPR site. (Reverse CRISPR target in Yellow, PAM in Orange, coding sequence is in purple)  
Copy the 48 bp 3' of the CRISPR cut into a new sequence file; this is the downstream homology.

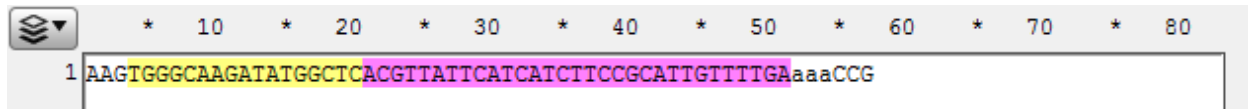


8) Observe the next three bases downstream of the 48 bp of homology, and pick a base not present to be the 3 bp spacer between the homology and the Universal PAM in the vector. (Here the bases are "CTG" so "aaa" was chosen for the spacer.)  
Add the spacer to the new file 3' of (after) the homology.



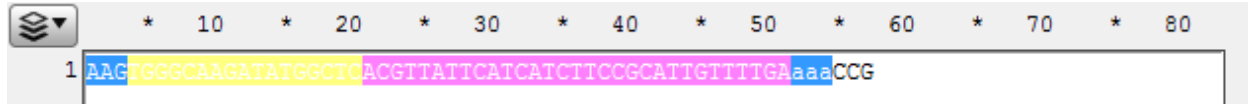
9) Add the BspQI enzyme overhang sequences for cloning, to the ends of the homology domain. 5'-AAG and 3'-CCG. (Here both overhangs are added to prevent errors in copying sequence for the oligos in the next two steps.)





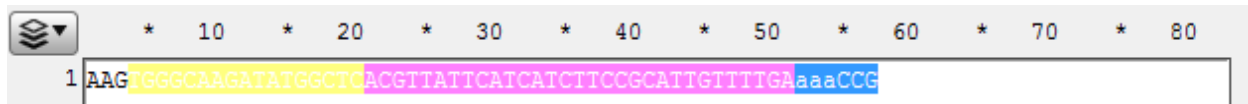
1 AAGTGGGCAAGATATGGCTCACGTTATTCATCATCTTCCGCATTGTTTTGAaaaCCG

10) The Downstream Homology Oligo A will be this sequence from the beginning of the sequence to the end of the spacer (see highlighted below). In this example this oligo sequence is 5'-AAGTGGGCAAGATATGGCTCACGTTATTCATCATCTTCCGCATTGTTTTGAaaa-3'.



1 AAGTGGGCAAGATATGGCTCACGTTATTCATCATCTTCCGCATTGTTTTGAaaaCCG

11) The Downstream Homology Oligo B (will be the reverse compliment of this sequence from the beginning of the homology to the end of the sequence (see highlighted below). In this example this oligo sequence is 5'-CGGtttCAAACAATGCGGAAGATGATGAATAACGTGAGCCATATCTTGCCCA-3'



1 AAGTGGGCAAGATATGGCTCACGTTATTCATCATCTTCCGCATTGTTTTGAaaaCCG

## F. One Pot Cloning of Homology Arms into pGTag Vectors

\*\*Note if the homology arm oligos contain either the sequence "5'-ACCTGC-3'" or "5'-GAAGAGC-3'" (or their compliments) the cloning reaction will be less efficient.

\*Note some sequences just don't work very well. **Ligation is more efficient with annealed homology arms and the purified ~1.2 kb and ~2.4kb fragments from vectors that have been digested with BfuAI and BspQI. If problems are encountered, one homology arm can also be cloned sequentially.**

### 1. Homology Arm Annealing

Anneal upstream and downstream homology oligo pairs separately:

4.5  $\mu$ L oligo A at 10 uM  
4.5  $\mu$ L oligo B at 10 uM  
4  $\mu$ L 10x Buffer 3.1 from NEB  
27  $\mu$ L dH<sub>2</sub>O  
total = 40  $\mu$ L

Incubate at 98°C for 4 min, 98°C 45 sec x 90 steps decrementing temp 1°C/cycle, 4°C hold

(Alternatively heat in 95-98°C water for 5 minutes, and then remove the boiling beaker from the heat source and allow it to cool to room temp for 2 hours, before placing samples on ice.)

### 2. 1-Pot Digest

Assemble the following:

4.0  $\mu$ L dH<sub>2</sub>O  
2  $\mu$ L Plasmid at 50 ng/ $\mu$ L  
1  $\mu$ L 10x Buffer 3.1 from NEB  
1  $\mu$ L 5' annealed homology arm  
1  $\mu$ L 3' annealed homology arm  
0.5  $\mu$ L BfuAI enzyme from NEB  
0.5  $\mu$ L BspQI enzyme from NEB  
10  $\mu$ L total

Incubate at 50°C for 1 hr, place on ice.

### 3. Ligation

Add the following:

3  $\mu$ L 5x T4 quick ligase buffer  
1.5  $\mu$ L dH<sub>2</sub>O  
0.5  $\mu$ L T4 quick ligase  
15  $\mu$ L total

Incubate 8-10 min at room temperature (to overnight). Store at -20 °C,

### 4. Transformation

- a. On ice, thaw 1 (one) vial competent cells (50  $\mu$ L) for every 2 ligation reactions. (approx. 5 min). It is recommended to use NEB Stable Competant E. coli (C3040H) cells to limit recombination.
- b. While cells are thawing, label the microcentrifuge tubes for each ligation and put on ice.
- c. Once the cells are thawed, use a pipette to transfer 25  $\mu$ L of the competent cells into each labeled tube.
- d. Add 1.5  $\mu$ L of a ligation reaction into competent cells to transform.
  - a. Amount of ligation reaction added should be less than 5% of volume of competent cells.
- e. Mix by tapping the tube several times or gently mixing with the pipet tip.
  - a. Do NOT mix by pipetting, this will lyse the cells.
- f. Incubate on ice for 5 to 20 minutes.
- g. Heat shock the cells by submerging the portion of the tube containing the cells in a 42°C water bath for 40 - 50 seconds.
- h. Incubate cells on ice for 2 minutes.
- i. Add 125  $\mu$ L of room temperature LB to each transformation.
- j. Incubate cells at 30°C for 1- 1.5 hour(s) in a shaking incubator.
- k. While the transformed cells are recovering, spread 40  $\mu$ L of X-Gal solution, and 40  $\mu$ L IPTG 0.8 M on LB Kanamycin selection plates.
  - a. X-Gal is lethal to cells while wet, it is recommended to first label the plates and then place them in a 30°C incubator to dry.
- l. After recovery and the X-Gal is dry, Plate 150  $\mu$ L of each transformation on the corresponding correctly labeled plate.
- m. Incubate plates overnight at 30°C.

## 5. Growing colonies

Pick 3 white colonies from each plate and grow in separate glass culture tubes with 3 mL LB/Kanamycin.

Or to pre-screen colonies by colony PCR:

- a. Pick up to 8 colonies with a pipet tip and resuspend them in separate aliquots of 5  $\mu$ L dH<sub>2</sub>O. Place the tip in 3 ml of LB/Kan, label, and store at 4°C.
- b. Make a master mix for your PCR reactions containing the following amounts times the number of colonies you picked.

7.5  $\mu$ L 2x Gotaq mastermix  
5.5  $\mu$ L dH<sub>2</sub>O  
0.5  $\mu$ L primer at 10 uM "F3'-check" 5'- GGCGTTGTCTAGCAAGGAAG -3'  
0.5  $\mu$ L primer at 10 uM "3'\_pgtag\_seq" 5'-ATGGCTCATAACACCCCTTG-3'  
14  $\mu$ L total

- c. Aliquot 14  $\mu$ L of mixed master mix into separate labeled PCR tubes.
- d. Add 1  $\mu$ L of colony to each reaction as template.
- e. or 20 ng purified plasmid as control.
- f. Cycle in a thermocycler

95°C 2 minutes  
95°C 30 seconds ]  
57°C 30 seconds ] x 35 cycles  
72°C 30 seconds ]  
72°C 5 minutes  
4°C hold

- g. Run 5  $\mu$ L of PCR product on a 1% agarose gel. You should get bands that are a different size than the control.

## 6. Mini Prep Cultures

Follow Qiagen Protocol

## 7. Sequencing of Plasmids

The 5' homology arm can be sequenced by the 5'\_pgtag\_seq primer:

5'-GCATGGATGTTTTCCCAGTC-3'.

The 3' homology arm can be sequenced with the "3'\_pgtag\_seq" primer:

5'-ATGGCTCATAACACCCCTTG-3'.

## G. Injection of GeneWeld Reagents (spCas9 mRNA, Universal sgRNA (UgRNA), genomic sgRNA and pGTag homology vector) into 1-cell zebrafish embryos

***Prepare and collect the following reagents for injection***

1. Prepare nCas9n mRNA from pT3TS-nCas9n (Addgene #46757 from (Jao et al., 2013)) as described above (page 14).
2. Synthesize UgRNA and purify as described above (page 11) using the following oligo A:

5'-TAATACGACTCACTATAGGGAGGCGTTTCGGGCCACAGGTTTTAGAGCTAGAAATAGC-3'

Corresponding to the universal target sequence: GGGAGGCGTTTCGGGCCACAG

Alternatively, the UgRNA can be directly ordered from IDT and resuspended in RNF water.

5'-GGGAGGCGUUCGGGCCACAGGUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAGGCUAGUCCGUUAUCAACUUGAAAAAGUGGCACCGAGUCGGUGCGGAUC-3'

3. The pGTag homology vectors should be purified a second time prior to microinjection under RNase free conditions with the Promega PureYield Plasmid Miniprep System beginning at the endotoxin removal wash and eluted in RNF water.

### ***Embryo Injections for Integration of pGTag vectors***

Injections are performed as previously described in 2 nl per embryo with the addition of the UgRNA and targeting pGTag DNA.

Final per embryo:

Injection mixture:

150 pg of nCas9n mRNA  
25 pg of genomic gRNA  
25 pg of UgRNA  
10 pg of pGTag DNA

75 pg/nl of nCas9n mRNA  
12.5 pg/nl of genomic gRNA  
12.5 pg/nl of UgRNA  
5 pg/nl of pGTag DNA

### **H. Examine embryos for fluorescence and junction fragments**

Embryos are examined for fluorescence under a Zeiss Discovery dissecting microscope with a 1X objective at 70-100X magnification. If weak signals are observed, embryos are manually dechorionated, and viewed on glass depression well slides. If no or weak signals were observed, Gal4VP16 integrations are attempted in a 14XUAS-RFP background. Embryos displaying widespread fluorescence in expression domains consistent with the targeted gene are examined for junction fragments or raised to adulthood for outcrossing.

F0 Junction fragment analysis between the genomic locus and the targeting vector is carried out by isolating DNA from embryos followed by PCR. The following primers are used for junction fragment analysis and must be paired with gene specific primers (5' to 3'):

5' pGTag junctions:

R-Gal4-5'juncM	GCCTTGATTCCACTTCTGTCA	with a gene specific forward primer
R-RFP-5'junc	CCTTAATCAGTTCCTCGCCCTTAGA	
R-eGFP-5'junc	GCTGAACTTGTGGCCGTTTA	

3' pGTag junctions:

F-Gal4-3'juncM	GCAAACGGCCTTAACTTTCC	with a gene specific reverse primer
F-Gal4-3'juncJ	CTACGGCGCTCTGGATATGT	
F-RFP-3'junc	CGACCTCCCTAGCAAACCTGGGG	
F-eGFP-3'junc	ACATGGTCTGCTGGAGTTC	

PCR amplification of junction fragments can be a result of artifacts (Won and Dawid, 2017), so it is important to carryout control amplifications with injected embryos that lack the genomic gRNA. F0 analysis by PCR of junction fragments is carried out to examine correct targeting. F-Gal4-3'juncM and F-Gal4-3'juncJ are two alternate primers for amplification of junction fragments from the Gal4 cassette due to gene specific mis-priming depending on the target loci.

7.5  $\mu$ L 2x Gotaq mastermix  
5.5  $\mu$ L dH<sub>2</sub>O  
0.5  $\mu$ L primer at 10 uM genomic primer  
0.5  $\mu$ L primer at 10 uM pGTag primer  
14  $\mu$ L total

1. Aliquot 14  $\mu$ L of mixed master mix into separate labeled PCR tubes.
2. Add 1  $\mu$ L of genomic DNA to each reaction as template.
3. Cycle in a thermocycler with the following steps:

95°C	2 minutes	
95°C	30 seconds	]
55°C	30 seconds	] x 35 cycles
72°C	30 seconds	]
72°C	5 minutes	
4°C	hold	

4. Run 5  $\mu$ L of PCR product on a 1.2 % agarose gel in 1XTAE. Putative junction fragments should give bands that are of predicted size.

F0 animals that are positive for the reporter gene are raised to adults then outcrossed and examined for fluorescence as above. The Gal4VP16 system can lead to silencing resulting in mosaic patterns in F1 embryos. F1 embryos displaying fluorescence are examined for junction fragments as above, raised to outcross to make F2 families or sacrificed at 3 weeks post fertilization for Southern-Blot analysis of integrations. F0 and F1 identified fish can be incrossed or backcrossed to get an initial impression of the homozygous phenotypes. It is recommended that lines are continuously outcrossed once established.

## References:

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.

Moreno-Mateos, M.A., Vejnar, C.E., Beaudoin, J.D., Fernandez, J.P., Mis, E.K., Khokha, M.K., and Giraldez, A.J. (2015). CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat Methods* 12, 982-988.

Varshney, G.K., Pei, W., LaFave, M.C., Idol, J., Xu, L., Gallardo, V., Carrington, B., Bishop, K., Jones, M., Li, M., *et al.* (2015). High-throughput gene targeting and phenotyping in zebrafish using CRISPR/Cas9. *Genome Res* 25, 1030-1042.

Won, M., and Dawid, I.B. (2017). PCR artifact in testing for homologous recombination in genomic editing in zebrafish. *PLoS One* 12, e0172802.