Published in final edited form as:

Nat Biotechnol. 2008 June ; 26(6): 695–701.

Targeted gene inactivation in zebrafish using engineered zinc finger nucleases

Xiangdong Meng^{1,4}, Marcus B. Noyes^{1,2}, Lihua (Julie) Zhu¹, Nathan D. Lawson^{1,3}, and Scot A. Wolfe^{1,2}

¹Program in Gene Function and Expression, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.

²Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.

³Program in Molecular Medicine, University of Massachusetts Medical School, Worcester, Massachusetts 01605, USA.

Abstract

In most vertebrate model systems direct genomic manipulation at a specific locus is still not feasible. Zinc finger nucleases (ZFNs) provide a system to introduce genomic lesions at a specific site in vertebrate cell lines^{1–3}. Here, we adapt this technology to create targeted mutations in the zebrafish germline *in vivo*. ZFNs were engineered that recognize sequences in the zebrafish ortholog of the vascular endothelial growth factor-2 receptor, *kdra*. Co-injection of mRNAs encoding these ZFNs into 1-cell stage zebrafish embryos led to mutagenic lesions at the target site that were transmitted through the germline with high frequency. Thus, engineered ZFNs can introduce heritable mutations into a vertebrate genome. Importantly, this *in vivo* gene inactivation approach obviates the need for embryonic stem cell lines and will be applicable to most animal species, especially in cases where early stage embryos are easily accessible.

The ability to perform fine genomic manipulation in the mouse has firmly established it as a central vertebrate model system. In this case, targeted gene knock-outs and knock-ins have been enabled through the establishment and use of embryonic stem cell lines in which the primary genomic manipulation is performed^{4, 5}. Similar manipulations in other vertebrate models have largely failed due to the difficulty in establishing comparable cell lines. Thus, an alternative means for generating targeted gene knockouts is essential.

ZFNs, which are a chimeric fusion between a Cys₂His₂ zinc finger protein (ZFP) and the cleavage domain of FokI endonuclease^{1, 6}, are powerful tools for genomic manipulation in plants^{7, 8}, invertebrates^{9, 10} and cell lines^{2, 3, 11}. The DNA-binding specificity is supplied by the ZFP, which can be engineered to recognize a wide variety of target sequences^{12–15}, whereas the cleavage activity is provided by the nuclease domain^{1, 6}. The nuclease domain requires dimerization for activity, and consequently two ZFNs must assemble on DNA with the appropriate geometry for efficient DNA cleavage¹⁶ (Supplementary Fig. 1a). Double-

Correspondence should be addressed to N.D.L. (nathan.lawson@umassmed.edu) or S.A.W. (scot.wolfe@umassmed.edu).

⁴Current Address: Sangamo BioSciences Inc., Point Richmond Tech Center II, 501 Canal Blvd., Suite A100, Richmond, CA 94804
Author contributions. X. M., N.D.L. and S.A.W. designed experiments and analyzed data. X.M and N.D.L carried out experiments. M.B.N. developed the omega-based bacterial one-hybrid system; S.A.W. developed the algorithm and PERL script to identify preferred recognition elements. J.Z. analyzed Solexa sequencing data. X.M., N.D.L., and S.A.W. wrote the paper.

Note: Supplementary information is available on the Nature Biotechnology website.

The authors have no declared Competing Interests.

stranded breaks that occur as a result of ZFN-mediated cleavage can lead to mutations when repaired via non-homologous end joining (NHEJ), which is an error-prone repair pathway^{10, 17}. Thus, ZFNs allow directed mutagenesis of a desired target sequence in a complex genome.

To determine the feasibility of utilizing ZFNs to induce targeted disruption of a gene within the germline of a vertebrate genome, we tested their efficacy in zebrafish. The external fertilization of zebrafish embryos makes them accessible for injection with mRNAs encoding ZFNs. Additionally, since mRNA injections can be performed directly into 1-cell stage embryos, we reasoned that co-injected ZFNs could induce lesions early enough in development that they would be transmitted through the zebrafish germline. Finally, despite its increasing use as a vertebrate model system, it is currently not possible to directly disrupt a gene of interest in the zebrafish genome.

We chose to target the zebrafish *kdra* locus due to the availability of ENU-induced mutant alleles that display vascular specific embryonic defects (Ref 18 and data not shown). We first developed an algorithm that identifies sites within a desired sequence that have favorable properties for recognition and cleavage by a pair of ZFNs (Supplementary Fig. 2). Using this approach, we identified two 9 bp zinc finger recognition elements flanking a 6 bp spacer within *kdra* exon 2 that are optimally positioned for ZFN activity (Supplementary Fig. 1c). Since this exon encodes the N-terminal region of the extracellular domain of the Kdra receptor tyrosine kinase, we would expect ZFN-induced mutations at this site that result in frameshifts to generate a null allele.

We utilized a combination of design and selection to create 3-finger ZFPs with good DNA binding specificity for each 9 bp recognition element within *kdra* exon 2 (Fig. 1). As a first step, we selected individual fingers to recognize each 3 bp subsite within the 9 bp recognition element using a bacterial one-hybrid system (see Supplementary Methods). For this purpose we constructed libraries for each finger that contain a combination of: 1) designed recognition motifs that have been previously observed in ZFP-DNA structures and 2) randomized positions where the desired binding specificity could not be predicted with high confidence (see Supplementary Methods). Individual finger libraries were first selected in the context of two anchor fingers with defined DNA binding specificity¹⁵ (Fig. 1a). Each selection yielded a restricted pool of zinc fingers with motifs suitable for recognition of each 3-bp subsite (Supplementary Fig. 3). Subsequently, the three independent finger pools that recognize each 3 bp subsite in the 9 bp recognition element were recombined via PCR to generate a new three-finger library¹⁵ (Fig. 1b). We then selected 3-finger ZFPs specific for each 9 bp recognition element from these libraries (Supplementary Fig. 4 and 5) and identified two (ZFP1 and ZFP2) that displayed specificity for at least 7 out of the 9 bp in each of the target sequences within *kdra* exon 2 (Fig. 1c and Supplementary Fig. 6). Based on the composite binding specificities of ZFP1 and ZFP2, the most suitable target site in the zebrafish genome for ZFNs containing these fingers is the desired target within *kdra* exon 2 (data not shown). Therefore, these ZFPs were incorporated into ZFNs for our studies in zebrafish.

The cassette encoding each ZFP was transferred into a pCS2 expression plasmid¹⁹ containing a nuclear localization signal, an epitope tag and one of two complementary FokI cleavage domain variants (DD and RR versions, Fig. 1d; Supplementary Fig. 1a). These modified FokI nuclease domains contain previously described mutations at the dimerization interface that reduce their propensity to homodimerize, thereby decreasing the frequency of off-target cleavage events that are due to homodimerization of an individual ZFN^{20, 21}. To determine the efficacy of the engineered ZFNs *in vivo*, we injected mRNA encoding each heterodimeric ZFN together or separately at various concentrations into 1-cell stage zebrafish embryos (Fig 1d). We then observed the injected embryos for morphological defects at 24 hours post fertilization (hpf). The majority of embryos injected with 5 pg of each exon 2 ZFN displayed

overall normal morphology at 24 hpf (Fig. 2a). Increasing amounts of both ZFNs led to non-specific developmental abnormalities (referred to as “monsters”) and death (Fig. 2a, b). We did not observe obvious cardiovascular defects known to be associated with loss of *kdra*¹⁸. This finding is consistent with previous studies that demonstrate ZFN-induced mutations usually occur in only one copy of a diploid genome leading to heterozygosity². Furthermore, embryos containing ZFN-induced mutations are mosaic (see below), which would likely preclude the appearance of vascular specific defects resulting from ZFN-induced mutations in *kdra* in injected embryos. We did not observe non-specific abnormalities or lethality in embryos injected with comparable doses of mRNA encoding either single ZFN monomer (Fig. 2a), suggesting that the observed defects are the result of excessive off-site target DNA cleavage by the heterodimeric ZFN complex.

To determine if site-specific mutations were generated by the engineered ZFNs, we performed PCR-based genotyping of injected embryos that displayed normal morphology. For this purpose, we utilized a NspI restriction site that resides in the 6 bp spacer region between each ZFN 9 bp recognition element within *kdra* exon 2 (Supplementary Fig. 1c). Since repair of ZFN-induced double strand breaks by NHEJ can lead to short deletions^{10, 17}, as well as insertions, we would expect the exon 2 NspI site to be lost in injected embryos. As shown in Fig. 2c, PCR-amplified *kdra* exon 2 fragments from individual uninjected embryos show complete digestion with NspI. By contrast, individual embryos co-injected with both ZFNs display a larger fragment consistent with the loss of the *kdra* exon 2 NspI site in a fraction of the total population (Fig. 2c). These observations imply the existence of a ZFN-directed mutagenic lesion at this site. To better characterize these lesions and quantify the mutagenesis frequency, we cloned and sequenced PCR fragments spanning *kdra* exon 2 from normal ZFN-treated embryos at 24 hpf. We observe a variety of lesions comprised of microdeletions (4, 5, and 6 bp) and small insertions (Fig. 2d). In most cases, these deletions lead to a frameshift in the *kdra* coding sequence that would result in premature truncation of the Kdra protein (data not shown). From a pool of shotgun-cloned fragments, we identified lesions in 9 out of 94 clones indicating that the ZFN-induced mutagenic frequency is approximately 10 percent in morphologically normal embryos.

We utilized Solexa sequencing technology to estimate the frequency of ZFN-induced lesions that are generated at other genomic regions with sequence similarity to the *kdra* exon 2 target site. We assayed 41 potential off-target sites that varied from the *kdra* exon 2 target site by 1 to 4 bp (heterodimeric ZFN sites) or that could be bound through homodimerization of individual monomers (homodimeric ZFN sites). We PCR-amplified each off-target site as well as the on-target site from genomic DNA isolated from embryos injected with 10 pg or 20 pg of mRNA encoding both ZFNs or from uninjected embryos. PCR products from each ZFN dose were uniquely barcoded with Solexa adapters, pooled and then analyzed in a single Solexa sequencing run. Sequencing reads unique to each experimental condition were deconvoluted using the attached barcode. No insertions or deletions were observed at any sites in the uninjected embryos. In morphologically normal embryos injected with 10 pg of both ZFNs, we detected small insertions or deletions in approximately 20 percent of on-target site fragments while only 2 off-target sites displayed similar lesions and these occurred at a frequency of ~1 percent (Fig. 2e). ZFN-induced lesions at the on-target site were 770-fold more likely to occur than in the population of off-target sites ($P\text{value} = 9.5 \times 10^{-10}$; Supplementary Table 1). In “monster” embryos from the 10 pg and 20 pg dose, we observed higher frequencies of lesions at off-target sites (Fig. 2e) although lesions are still more likely to occur at the on-target site (~170-fold, Fig 2e, Supplementary Table 1). Consistent with our use of engineered FokI heterodimers that are known to limit homodimerization^{20, 21}, we did not observe any lesions at sites expected to bind either ZFN homodimer in normal embryos (Supplementary Table 1). In summary, these results suggest that morphologically normal

embryos injected with the *kdra* exon 2 ZFNs display specific lesions at the target site with only modest collateral mutagenesis to the genome.

To determine if ZFN-induced mutations could be transmitted through the zebrafish germline, we raised embryos injected with 15 pg of mRNA encoding each ZFN. From 178 injected embryos, 69 were morphologically normal at 24 hpf and 28 of these reached adulthood (see Supplementary Fig. 7 for a summary of F1 founder analysis). We found that fertilization rates and embryo production from most ZFN-treated adults were similar to wild type adults (Supplementary Table 2, 3 and data not shown). In addition, we did not observe any increase in the appearance of non-specific developmental abnormalities in embryos from ZFN-treated adults compared to wild type (for example, see Fig 3a, c, e; Supplementary Table 2; Supplementary Fig. 8). In two cases, we identified male founders that appear to be sterile (Supplementary Table 2, 3). We identified putative founders bearing ZFN-induced mutations in *kdra* by crossing them to heterozygous carriers of the *kdra^{um6}* allele. We observed progeny embryos for presence of segmental arteries and blood circulation, both of which are absent in *kdra* mutant embryos¹⁸. To visualize segmental arteries, we utilized *kdra^{um6}* carriers bearing *Tg(fli1:egfp)^{y1}* to allow fluorescent visualization of blood vessels²². From twenty potential founders we identified 6 that gave progeny embryos that exhibited loss of segmental arteries and/or loss of circulation indicative of mutations within *kdra*. For example, we observed embryos from founder 883.2M that phenocopied those derived from an incross of *kdra^{um6}* heterozygous carriers (compare Fig. 3d, h), while embryos obtained from founder 889.3M displayed less severe segmental artery defects (Fig. 3f) similar to a previously described hypomorphic *kdra* allele¹⁸. These observations suggest a qualitative difference between the ZFN-induced lesions at the *kdra* locus in these two founders. Based on the number of mutant embryos observed in out-crosses to *kdra^{um6}* heterozygotes, we estimated that germ cell mosaicism in the identified ZFN allele-bearing founders ranged from 8 to 50 percent (Supplementary Table 2). Importantly, this degree of penetrance remained consistent upon successive crosses indicating that these mutations are stable within the zebrafish germline (Supplementary Table 3). Our results demonstrate that ZFN-induced mutations can be introduced into the zebrafish germline at high efficiency.

We genotyped individual embryos using the NspI assay described above (see Fig. 2c) to confirm that the observed vascular phenotypes were due to ZFN-induced mutations in *kdra* exon 2. For founders 883.2M, 889.3M, and 889.7M we observed a direct correlation between the appearance of an NspI resistant fragment and circulatory defects (Fig. 4a, Supplementary Fig. 9a, b). In all cases, circulatory defects correlated with trans-heterozygosity of the ZFN-induced and *kdra^{um6}* alleles (Fig. 4a, Supplementary Fig. 10 and data not shown). In the remaining founders, we observed a mixture of NspI resistant and sensitive alleles. For example, only 2 out of 8 mutant embryos derived from founder 883.7F displayed loss of NspI digestion (Supplementary Fig. 9c), suggesting the existence of multiple ZFN-induced mutant alleles. PCR analysis of mutant and wild type sibling embryos using flanking primers demonstrated the presence of a smaller fragment and indicated the existence of a large deletion at the target site (Supplementary Fig. 9d). Subsequent sequence analysis confirmed the loss of 382 bp immediately upstream of the ZFN1 recognition element (Supplementary Fig. 9d and data not shown). Thus, founder 883.7F appears to carry two distinct *kdra* exon 2 mutant alleles. Mutant embryos from founder 889.1M also displayed an apparent wild type NspI genotype (Fig. 4a). In this case we utilized a CelI nuclease assay, which cleaves DNA at sites containing base mismatches, insertions or deletions^{20, 23}, to detect potential mutations at the *kdra* exon 2 target site. As shown in Fig. 4b, PCR products amplified from wild type embryos derived from founder 889.1M fail to digest with CelI. By contrast, PCR products from mutant embryos are sensitive to CelI digestion indicating the presence of non-complementary DNA sequences within these alleles due to ZFN-induced genetic lesions at *kdra* exon 2 (Fig. 4b).

We further characterized the ZFN-induced lesions in *kdra* by cloning and sequencing PCR fragments containing exon 2 from pooled sibling mutant embryos (see Supplementary Fig. 7). Similar to somatic lesions characterized in injected embryos (see Fig. 2d), we found that ZFN-induced germline lesions included both deletions and insertions that overlapped the ZFN target site (Fig. 4c). As expected from the NspI genotyping analysis, founders 883.2M and 889.3M display deletions that eliminate the NspI site (Fig. 4c). In addition, the frequency of mutant alleles detected in a pooled embryo population (Supplementary Table 4) along with genotyping analysis (Supplementary Fig. 9a, b) suggested that these mutations were the only mutant alleles at *kdra* exon 2 carried by these respective founders. In embryos derived from founder 883.2M, we observed a 4-bp deletion that would lead to a frameshift and formation of a truncated Kdra protein (data not shown). Interestingly, the deletion observed in mutant embryos derived from founder 889.3M displayed a 24 bp deletion that would maintain the *kdra* reading frame. The milder segmental artery defect in embryos bearing this mutation is consistent with the generation of a hypomorphic *kdra* allele (see Fig. 2). In embryos derived from founders 889.1M and 889.7M, we detected insertions within the ZFN target site (Fig. 4d and Supplementary Table 4). As above, the frequency of these sequences within pooled mutant embryos suggested the existence of a single mutant allele derived from each founder (Supplementary Table 4). For founder 889.1M, a 4 bp insertion retained the NspI site, consistent with our NspI and Cell genotyping analysis (Fig. 4b, c). In embryos from founder 883.7F, we observed the occurrence of a 12 bp deletion at a low frequency within pooled mutant embryos. As described above, the 883.7F founder also bears a mutant allele characterized by a 382 bp deletion that would not be detected using our shotgun cloning strategy or preliminary NspI assay. These results indicate that all analyzed ZFN-injected founders harbor mutagenic lesions at the target site and, in all but one case, these founders bear a single mutant allele.

Using a combination of design and selection we have created a set of ZFNs that introduce genetic lesions at a desired site within the zebrafish genome. In principle, existing libraries of pre-selected zinc finger modules could be utilized to assemble ZFN targeting domains^{12-14, 24}. This type of approach, which bypasses the need for selection, would make ZFN technology even more broadly accessible to non-experts. Indeed, this strategy has been applied for generating targeted mutations in the *Drosophila* genome^{9, 10}. However, the existing databases of zinc fingers contain modules of variable quality^{10, 11, 25}, and as such, ZFNs constructed from these fingers may perform poorly, particularly in the context of most vertebrate genomes, which are considerably more complex. Until a library of *in vivo* validated finger modules is available for ZFN construction, it will likely be necessary to utilize a combination of design and selection strategies to generate ZFPs with the highest specificity for gene targeting. Using our approach, ZFNs can be generated and characterized *in vivo* in ~ 6 to 8 weeks. In principle pairs of ZFNs for different targets can be created in parallel within this timeframe with minimal expense. As more zinc finger modules are validated *in vivo*, their associated binding specificities could be utilized to generate an accessible database of potential target sites annotated throughout any genome. We envision that such a resource would ultimately allow users to generate ZFNs for a desired target gene through a streamlined selection process or, preferably, by design-only using standard molecular biology techniques.

Based on our findings, targeted gene inactivation using ZFNs is quite robust in zebrafish embryos: nearly one-third of screened founders carried ZFN-induced lesions at the target site and half of the gametes in several founders carried a single mutagenic *kdra* allele. In our current work, we crossed our founders to an existing null *kda* allele to rapidly confirm the presence of ZFN-induced lesions at the target site. However, the predominant application of this technology will likely be to generate mutations in genes for which no mutant lines exist. In these cases, founders could be identified using a PCR-based assay similar to the one described here. Subsequent generations of embryos would then be raised to adulthood and incrossed to characterize the associated phenotype. Given the high degree of germline transmission

observed in our studies, it would also be possible to perform preliminary phenotypic analysis by incrossing the initial founder fish. Therefore, generation of a family of heterozygous carriers of a desired ZFN-induced mutation would take approximately 6 to 8 months following injection to establish founders.

Our work provides a foundation to perform focused gene modification in a vertebrate genome and should allow the development of techniques for more subtle genomic manipulation. Indeed, by simply screening numerous founder individuals we show that it is possible to identify distinct lesions ranging from frameshifts to small in-frame deletions. By focusing on functional domains within a protein of interest, our approach will allow straightforward generation of a targeted allelic series of mutants. Given the success of ZFN-targeted gene modification via homologous recombination in cell culture², it will likely be feasible to achieve targeted knock-ins using this technology in zebrafish, with exciting implications for the creation of human disease models that were heretofore inaccessible. Finally, a powerful aspect of this technology is that it does not rely on the existence of species-specific ES cell lines that have proven difficult to establish for most vertebrate organisms besides mouse. Consequently, ZFN-mediated gene inactivation should translate from zebrafish to other model and non-model organisms, especially in cases where it is possible to harvest and easily manipulate fertilized embryos.

Methods

Zebrafish husbandry and genetic modification

Zebrafish adults and embryos were handled according to standard methods²⁶. These studies were approved by the UMass Medical School IACUC. The wild type line used in this study (referred to as Crawfish) was established through several incross generations of wild type fish originally obtained from Scientific Hatcheries. The *kdra^{um6}* allele was identified in an ongoing screen for vascular mutants (Lawson, *unpublished observations*). Imaging of wild type and *Tg(fli1a:egfp)^{y1}* embryos was performed as described elsewhere¹⁸. pCS2-ZFN constructs were linearized with NotI and mRNA was transcribed with SP6. We performed microinjection of ZFN mRNAs into 1-cell stage zebrafish embryos according to standard methods²⁷. Phenotypic observation of segmental arteries and circulation as well as individual embryo genotyping was performed as described previously¹⁸.

Omega-based bacterial one-hybrid selection of ZFPs

Briefly, output from the search algorithm (Supplementary Fig. 2) was manually inspected for suitable ZFP recognition sequences. One site in exon 2 (5'-ACACACCTTCAGCATGTTGGTGGGAC-3', ZFP recognition elements underlined) was deemed suitable based on the large fraction of purines within each ZFP site and the large proportion of bases that were guanines within these recognition elements. Zinc finger libraries were constructed by cassette mutagenesis and their complexity assessed as previously described²⁸ except that the annealed library and complementary oligonucleotides were cloned into the desired finger module between two unique BbsI restriction sites that generate unique, non-palindromic, 4 bp overhangs within the vector. A two-stage selection approach (adapted from Ref. 15) was used to select ZFPs from these libraries that recognize each element via an omega-based bacterial one-hybrid system²⁹.

Zinc Finger Library Construction

Individual fingers that recognize each subsite within the ZFP recognition element were selected independently from a library of partially randomized recognition helices in the context of two anchor fingers, which due to their defined specificity, ensure that the randomized finger library is positioned over the appropriate DNA-recognition element (Fig. 1A). ZFP sequences were derived from Zif268 except at the altered recognition positions. Some of the recognition

positions within each finger library were fixed based on “high confidence” interactions that were anticipated to be functional in most recognition contexts. The positions not fixed within each recognition helix (a subset of positions –1,1,2,3,5 & 6) were fully randomized (codon encoded NNK). ZFP libraries were constructed in pB1H2w2 - an ampicillin-marked vector that expresses each ZFP as a fusion to the omega-subunit of RNA polymerase. Details on the individual finger libraries are described in the Supplementary Methods.

Stage 1 B1H ZFP selections

Each ZFP library and the complementary reporter vector containing its target site were introduced into a selection strain that lacks endogenous expression of the omega subunit of RNA polymerase (*US0ΔhisBΔpyrFΔrpoZ*). For each selection 150 ng of pB1H2wZFP library DNA and 150 ng pH3U3 reporter DNA were co-transformed into the selection strain by electroporation. These cells were recovered and adapted to the NM minimal medium conditions following the previously described B1H selection protocol²⁸. About 10⁷ transformants were plated on a NM minimal medium plate (round 150×15 mm) lacking histidine and containing 0 or 10 μM IPTG and 5 or 10 mM 3-AT. The selection plates were incubated at 37°C for one or two days until a moderate number of colonies were easily discernable. Different ZFP libraries demonstrated different activities under the selection conditions. For example the F1-GGA library displayed robust growth at low stringency, so colonies were selected from the high stringency plate (10 mM 3-AT, 0 mM IPTG), whereas the F2-GGT library displayed low activity, so colonies were selected from the low stringency plate (5 mM 3-AT, 10 mM IPTG). Ideally colony numbers in the hundreds would be obtained from each selection for recombination to build the final ZFP library for each recognition element. In these selections from 35 (library F2-GGT) to 2000 (library F1-GGA) colonies were recovered on the optimal selection plate. 8 colonies from each of these plates were picked and cultured and plasmid DNA was isolated and sequenced to identify the selected fingers (sequencing primer: CAAGAGCAGGAAGCCGCTG, see Supplementary Fig. 3).

Stage 2 ZFP library assembly

Colonies from the desired selection plates were washed off the plates and their plasmid DNA was recovered as a pool as previously described²⁸. Each pool of individual fingers was amplified from the plasmid DNA by PCR: 50 μl reaction with 1 units NEB DNA Taq polymerase using 50 ng plasmid DNA template and 1 mM each primer (Supplementary Methods); denature for 1 min at 95°C, 28 cycles consisting of 94°C for 20 sec, 55°C for 20 sec and 72°C for 20 sec, followed by a final extension at 72°C for 6 min. The PCR products were run on a 1.5% agarose TAE gel, and the appropriate bands were excised and purified using a Qiagen PCR purification kit.

A library of three finger cassettes was assembled from pools of individual finger PCR products by overlapping PCR. PCR products from each individual finger (~ 50 ng each) were used as templates in a 50 μl PCR reaction containing 1 unit NEB Taq DNA polymerase and 1 × Taq buffer with dNTPs. During the first 6 reaction cycles primers were not included in the reaction, which allows the individual finger products to assemble into three-finger cassettes. The PCR reaction conditions were as follows: denature at 95°C for 1 min followed by 6 cycles consisting of 94°C for 20 sec, 60°C for 30 sec and 72°C for 30 sec. 1 mM final concentration of the F1forward and F3reverse primers was added to the reaction and then 26 additional amplification cycles were performed consisting of 94°C for 20 sec, 55 °C for 20 sec and 72 °C for 30 sec. A final extension time of 6 min at 72 °C was included before termination. The PCR products were run on a 1.5% agarose TAE gel and the appropriate band from each extension was excised and purified using a Qiagen PCR purification kit. These products were digested with NotI and BamHI and cloned into pB1H2w2 between the corresponding unique

restriction sites to build the Stage 2 ZFP libraries (as described in Ref. 28). Each three-finger ZFP library contained more than 10^6 unique clones.

Stage 2 ZFP selections and characterization

Selection of functional ZFPs that recognize the desired recognition elements were selected using the omega-based B1H system as described for the individual finger selections except that higher selection stringencies were employed: 0 to 10 μ M IPTG and 10 to 20 mM 3-AT. From 100 to 5,000 colonies survived on the various selection plates from 10^7 transformants challenged at each stringency. ZFP clones specific for the GAAGGTGTG recognition element were sequenced from the 10 mM 3-AT, 10 μ M IPTG stringency plate (about 5,000 colonies surviving, Supplementary Figure 4) and ZFP clones for the TTGGTGGGA recognition element were sequenced from the 10 and 20 mM 3-AT and 10 μ M IPTG stringency plate (about 3000 and 100 colonies survived on the two different stringencies respectively, Supplementary Fig. 5).

Two clones isolated from the GAAGGTGTG target selection (ZFP1 and ZFP3) and three clones isolated from the TTGGTGGGA target selection (ZFP2, ZFP4, ZFP5) were characterized using the omega-based B1H binding site selection system²⁹. Binding sites recognized by these clones were isolated from a 28 bp randomized pH3U3 reporter library. Individual clones were sequenced and the recognition motif was identified by analyzing the sequences for an overrepresented motif using MEME³⁰. See Supplementary Figure 6 for the list of selected sequences. The recognition motifs of each ZFP are presented as Sequence logos³¹ generated at the Weblogo site³².

Heterodimeric cleavage domains used in the ZFNs

The FokI nuclease domain was extracted by PCR from vector pST1374¹². Heterodimeric versions of this domain were generated by mutagenic PCR to create a pair of cleavage domains (D483R (RR) and R487D (DD)) in which interacting residues in the dimerization interface were swapped; these mutations have been previously described to reduce homodimeric cleavage activity and thereby reduce toxicity of ZFNs^{20, 21}.

ZFN transcription template in pCS2

Triple FLAG and HA epitopes were amplified from vectors pHWF and pHFW (The Drosophila Gateway Vector Collection) respectively by PCR and a SV40 NLS with Kozak sequence was added upstream of each epitope. The NLS-FLAG/HA fragments were inserted between BamHI and XhoI in pCS2. Following the insertion of the NLS-epitope elements into pCS2, the unique BamHI and KpnI sites were eliminated and new unique KpnI and BamHI sites were added after the NLS-FLAG/HA element to facilitate rapid sub-cloning of ZFPs from the B1H vectors. Each heterodimeric FokI cleavage domain was introduced into one NLS-epitope version of pCS2 (FokI-DD with FLAG and FokI-RR with HA) between the unique BamHI and XbaI sites and each ZFP was introduced between the unique KpnI and BamHI sites.

ZFN induced mutation detection using SURVEYOR nuclease

The SURVEYOR Mutation Detection Kit (TRANSGENOMIC) was used to detect ZFN-induced mutations within *kdra* exon 2. 25 μ l PCR reactions were carried out using 0.5 units Phusion Hot Start high fidelity DNA polymerase (FINNZYMES) in the manufacturer supplied buffer with 0.2 mM dNTPs and 1 mM primers flanking the exon 2 ZFN target site (5'-GTTTGTGTTGTTAGTGTGCCTGTGATGTAATG-3', 5'-ATAAAGTGGCCATTGAACGTAGATGCAC-3'). This amplification should produce a 230 bp DNA fragment with the ZFN target site at its center. About 30 ng genomic DNA was used

as the template in each PCR reaction. PCR reaction program: denature at 98 °C for 20 sec followed by 35 cycles consisting of 98 °C for 10 sec, 60 °C for 20 sec and 72 °C for 15 sec. 72 °C for 8 min final extension. 3 µl of each PCR products was assayed on a 1.5% TAE agarose gel to confirm successful amplification. The remaining PCR product was hybridized following the protocol in the SURVEYOR manual. 4 to 8 µl of the hybridized DNA was assayed for inhomogeneity by adding 0.5 µl SURVEYOR Enhancer S and SURVEYOR nuclease S each per 4 µl of DNA and incubating the reaction at 42 °C for 20 min. Following incubation 1 µl of loading dye was added to each reaction tube and 2.5 % agarose TAE gel was run at 90 volts for about 1 hr. The gel was subsequently stained with SYBR Gold (Invitrogen) for about 30 to 60 min and then destained with H₂O for about 20 min. DNA in the gel was visualized at 302 nm using a CYBR Gold filter.

Sequence analysis of mutations induced by ZFNs at *kdra* exon 2 in Zebrafish embryos

Genomic DNA from 24 embryos without obvious developmental defects was isolated and pooled. A 646 bp PCR fragment containing exon 2 was amplified by PCR (5'-GCAGAATTCAAACACAAGCACCGTCAGGAGCACAAGTCC-3', 5'-GCATGCGGCCGCGTTCACTGTACTCGAATGGCCTCCACAGTC-3') and purified. This PCR fragment was digested with EcoRI and NotI and cloned into pBluescript II SK+ at the corresponding sites. This library of Exon 2 segments was transformed into XL1-Blue cells (Stratagene) and cells were plated on media containing X-gal. White colonies were picked and the cloned PCR product was amplified using NEB Taq DNA polymerase and T3/T7 primers following the manufacturers protocol. 95 independent clones were sequenced with the T7 primer to determine the type and distribution of lesions at exon 2.

Off-target site identification

Off-target sites with the highest propensity to be recognized by the exon 2 ZFNs were identified by searching the genome (Zv7 repeat-masked) for matches to the consensus sequence for each ZFP based on their determined specificity ('GAXGGTGTG' and 'XXGGTGGGA', where X allows any base) with the appropriate spacing (5 or 6 bp) and orientation using a Perl algorithm. Heterodimer off-target sites were derived from sites that match 14 of 15 bp (*kdra* exon 2 is the only 15 of 15 bp match). Homodimeric sites were derived from sites that match either the 'GAXGGTGTG' composite site at 14 or 15 out of 16 bp or the 'XXGGTGGGA' composite site at 14 of 14 bp. A list of these sites is provided in Supplementary Table 5.

Off-target site lesion analysis

To assay lesion frequency at off-target sites, we performed Solexa sequencing of PCR fragments containing these sequences from zebrafish embryos. We injected wild type zebrafish embryos with mRNA encoding both *kdra* exon 2 ZFNs at a dose of 10 pg and 20 pg per embryo. At approximately 26 hours post fertilization we separated embryos based on morphology (normal or "monster") and isolated genomic DNA using the Qiagen DNeasy kit. As a control, we isolated genomic DNA from uninjected sibling embryos. We then PCR amplified 41 putative off-target sequences from uninjected, 10 pg ZFN injected normal, 10 pg injected monster and 20 pg injected monster embryos. PCR reactions were performed in 96-well plates using Advantage 2 HF polymerase (Clontech) and 20 ng of genomic DNA as template. PCR conditions were as follows: 94°C 2 min, 35 cycles: 94°C, 20 sec., 60°C, 20 sec., 65°C, 20 sec. A list of flanking primers for each site is provided in Supplementary Table 5; all primers contain an *Acl*I site. In parallel, we PCR-amplified the *kdra* exon 2 target site from genomic DNA from each set of embryos using the following primers: F, 5'-CCTGATCCACAACCTGCTTCCTGATGGATATCCAC. R 5'-ATAAAGTGGCCATTGAACGTAGATGCAC. Following amplification, the off-site PCR products were pooled together and digested with *Acl*I (NEB) to remove the terminal 16 bp of

the PCR product that is complementary to the genomic sequence, which allows sequencing to begin close to the putative lesion site. The on-site target fragment was separately digested with EcoRV (bold). Enzyme digests were heat-inactivated and a proportional amount of on-site target was mixed with the appropriate off-site pool. Each fragment pool was then gel-purified and treated with T4 DNA polymerase (NEB) with dNTPs to remove 3' overhangs created by AclI digestion. Polished fragments were purified using a Qiaquick column, treated with Klenow exo⁻ and dATP to add a 3' A overhang, and purified again using a Qiaquick column. We subsequently ligated adaptors containing a two-nucleotide "barcode" unique to each injection condition: AA – uninjected, AC – 10 pg ZFN/monster embryos, AG – 10 pg ZFN/normal embryos, AT – 20 pg/monster embryos. (Adaptor sequences are listed in the Supplementary Methods.). Adaptor-ligated PCR products were then pooled together, gel purified and PCR amplified using Illumina Genomic DNA primers (Illumina). The resulting PCR product was subjected to deep sequencing on a Solexa 1G located at the Center for AIDS Research, University of Massachusetts Medical School.

Solexa data analysis

The Firecrest and Bustard analysis module (Illumina) was used for image analysis and base calling to generate the sequence reads. This population was filtered for matches and near matches to the anticipated sequence spanning each on-target/off-target site by similarity-based alignment using phageAlign (Illumina), where the on-target site and 41 different off-target sites were examined. Sequence reads that were uniquely aligned to one of these 42 potential sequences were subsequently binned into 176 different groups based on the 2 bp barcode that proceeds each read, where four different barcodes were utilized with each representing a different ZFN dose. Thus, for each ZFN treatment there are 42 groups of sequences representing each population of the 41 off-target site sequences and the on-target site sequence group.

The presence of insertions and deletions in the sequences in each bin was identified by scanning for a 7bp "prefix" at the 5' end of each sequence prior to the spacer between the ZFP recognition sequences and a 7bp "suffix" at the 3' end of the distal ZFP binding site. The distance between the prefix and suffix pair in each sequence was determined. If the distance matched the expected sequence these sequences were binned as "correct", distances that were greater were binned as "insertions", and distances that were shorter were binned as "deletions". This analysis will miss insertions or deletions that alter either the prefix or suffix, and consequently only a subset of the potential lesions are being sampled. However, this approach is robust with regards to excluding marginal sequence reads and avoids attempts to "interpret" single base pair mutations, which could result from sequencing errors or SNPs. The number of sequence reads that contain normal sequence reads and the number of sequences with indels (insertions plus deletions) at on-target and off-target sites for each ZFN dose were computed for the subsequent statistic analysis.

The Fisher Exact Test was applied to assess whether the on-target site has a different insertion/deletion rate compared to that of the off-target sites for each ZFN dose. The odds ratio and its 95% of confidence interval was computed for each ZFN dose using fisher.test function in R v2.5 based on conditional maximum likelihood estimation. The same function was also applied to evaluate whether there is an overall significant difference among all treatments including the control group in the insertion/deletion rate of the off-target sites. When a significant overall difference was detected ($p < 0.05$), Fisher Exact Test was performed to determine whether each individual treatment affects insertion/deletion rate at the off-target sites compared with the control group.

Sequence analysis of mutant Zebrafish offspring from potential founders

Genomic DNA from each embryo that lacked circulation at 48 hpf was purified and the region surrounding Exon 2 was amplified by PCR using the primers and protocol described above for the pooled embryo analysis. 5 µl of PCR product from the embryos of each founder (6 embryos from 889.1 male founder; 6 embryos from 883.7 female founder; 3 embryos from 889.7 male founder; 6 embryos from 889.3 male founder and 6 embryos from 883.2 male founder) were pooled together, purified, digested and cloned into pBluescript II SK+ as described above. About 38 white colonies from each pool of founder DNA were amplified with T3 and T7 primers using NEB DNA Taq polymerase and these PCR products were directly sequenced using the T7 primer.

NspI analytical digestion of putative *kdra* exon 2 lesions induced by ZFNs

A 646 bp fragment containing *kdra* exon 2 was amplified using the primer set described for the SURVEYOR assay in an 11 µl PCR reaction, as described above. 1 µl of each PCR product was analyzed on 1.2% agarose gel to confirm successful amplification and the remaining 10 µl PCR product was mixed with 10 µl NspI digestion solution (1x NEB buffer 2, 1x BSA, 1.25 units of NspI, NEB) and incubated at 37°C over night. These reactions were analyzed on 2% agarose to assess cleavage of these PCR fragments.

PCR analysis of the ZFN-induced 400 bp deletion in 883.7

Primers 1 kb upstream of exon 2 target site (5'-AGCGAATTCCGAATGTCACCCATCCTGAGCAGCA-3') and 100 bp downstream of target site (5'-ATAAAGTGGCCATTGAACGTAGATGCAC-3') were used to amplify the genomic DNA of embryos displaying *kdra*^{-/-} phenotypes. These PCR products were analyzed on a 1.2% agarose gel. The majority of mutant embryos from this founder display two amplification products of different size. A representative smaller PCR from one embryo was isolated, purified and sequenced using the 3' primer.

Acknowledgements

We would like to thank Keith Joung for helpful discussions and for pST1374. We thank Charles Sagerstrom and Arndt Siekmann for critical reading of this manuscript. We are grateful to Mike Kacergis for valuable assistance in fish care and genotyping. We are grateful to Ellen Kittler and Maria Zapp for providing MPSS protocols and technical advice in support of this work. We also thank the UMass Deep Sequencing Core for MSPP analysis service and David LaPointe for providing bioinformatic and computational support. X.M., M.B.N and S.A.W. and their work were supported by 1R01GM068110 from NIGMS. N.D.L. and his work was supported by R01HL079266 from NHLBI.

References

1. Porteus MH, Carroll D. Gene targeting using zinc finger nucleases. *Nat Biotechnol* 2005;23:967–973. [PubMed: 16082368]
2. Urnov FD, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* 2005;435:646–651. [PubMed: 15806097]
3. Porteus MH, Baltimore D. Chimeric nucleases stimulate gene targeting in human cells. *Science* 2003;300:763. [PubMed: 12730593]
4. Koller BH, Smithies O. Inactivating the beta 2-microglobulin locus in mouse embryonic stem cells by homologous recombination. *Proceedings of the National Academy of Sciences of the United States of America* 1989;86:8932–8935. [PubMed: 2682666]
5. Zijlstra M, Li E, Sajjadi F, Subramani S, Jaenisch R. Germ-line transmission of a disrupted beta 2-microglobulin gene produced by homologous recombination in embryonic stem cells. *Nature* 1989;342:435–438. [PubMed: 2685607]
6. Kim YG, Cha J, Chandrasegaran S. Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences of the United States of America* 1996;93:1156–1160. [PubMed: 8577732]

7. Lloyd A, Plaisier CL, Carroll D, Drews GN. Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America* 2005;102:2232–2237. [PubMed: 15677315]
8. Wright DA, et al. High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J* 2005;44:693–705. [PubMed: 16262717]
9. Bibikova M, Beumer K, Trautman JK, Carroll D. Enhancing gene targeting with designed zinc finger nucleases. *Science* 2003;300:764. [PubMed: 12730594]
10. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D. Efficient gene targeting in *Drosophila* with zinc-finger nucleases. *Genetics* 2006;172:2391–2403. [PubMed: 16452139]
11. Alwin S, et al. Custom zinc-finger nucleases for use in human cells. *Mol Ther* 2005;12:610–617. [PubMed: 16039907]
12. Wright DA, et al. Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nat Protoc* 2006;1:1637–1652. [PubMed: 17406455]
13. Carroll D, Morton JJ, Beumer KJ, Segal DJ. Design, construction and in vitro testing of zinc finger nucleases. *Nat Protoc* 2006;1:1329–1341. [PubMed: 17406419]
14. Liu Q, Xia Z, Case CC. Validated zinc finger protein designs for all 16 GNN DNA triplet targets. *J Biol Chem* 2002;277:3850–3856. [PubMed: 11726671]
15. Hurt JA, Thibodeau SA, Hirsh AS, Pabo CO, Joung JK. Highly specific zinc finger proteins obtained by directed domain shuffling and cell-based selection. *Proc. Natl Acad. Sci. U S A* 2003;100:12271–12276. [PubMed: 14527993]
16. Bibikova M, et al. Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol* 2001;21:289–297. [PubMed: 11113203]
17. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics* 2002;161:1169–1175. [PubMed: 12136019]
18. Covassin LD, Villefranc JA, Kacergis MC, Weinstein BM, Lawson ND. Distinct genetic interactions between multiple Vegf receptors are required for development of different blood vessel types in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* 2006;103:6554–6559. [PubMed: 16617120]
19. Turner DL, Weintraub H. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev* 1994;8:1434–1447. [PubMed: 7926743]
20. Miller JC, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* 2007;25:778–785. [PubMed: 17603475]
21. Szczepek M, et al. Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat Biotechnol* 2007;25:786–793. [PubMed: 17603476]
22. Lawson ND, Weinstein BM. In vivo imaging of embryonic vascular development using transgenic zebrafish. *Dev Biol* 2002;248:307–318. [PubMed: 12167406]
23. Qiu P, et al. Mutation detection using Surveyor nuclease. *Biotechniques* 2004;36:702–707. [PubMed: 15088388]
24. Mandell JG, Barbas CF III. Zinc Finger Tools: custom DNA-binding domains for transcription factors and nucleases. *Nucl. Acids Res* 2006;34:W516–523. [PubMed: 16845061]
25. Cornu TI, et al. DNA-binding Specificity Is a Major Determinant of the Activity and Toxicity of Zinc-finger Nucleases. *Mol Ther* 2008;16:352–358. [PubMed: 18026168]
26. Westerfield, M. *The Zebrafish Book*. University of Oregon Press; Eugene, Oregon: 1993.
27. Xu, Q. *Molecular Methods in Developmental Biology*. Guille, M., editor. 127. Humana Press, Inc.; Totowa, NJ: 1999. p. 125-132.
28. Meng X, Wolfe SA. Identifying DNA sequences recognized by a transcription factor using a bacterial one-hybrid system. *Nat. protocols* 2006;1:30–45.
29. Noyes MB, et al. A systematic characterization of factors that regulate *Drosophila* segmentation via a bacterial one-hybrid system. *Nucl. Acids Res* 2008;gkn048.
30. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* 1994;2:28–36. [PubMed: 7584402]
31. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* 1990;18:6097–6100. [PubMed: 2172928]

32. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. *Genome research* 2004;14:1188–1190. [PubMed: 15173120]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Figure 1.

Engineering ZFNs that target *kdra* exon 2 using a bacterial one-hybrid system. a. First selection stage to identify single zinc fingers with optimized binding to each 3-bp subsite within a recognition element. b. Recombination of individual fingers and subsequent second selection stage against full ZFP recognition elements. c. Sequence logos are shown for the binding specificity of ZFPs incorporated into ZFNs for subsequent analysis in zebrafish embryos. d. Overview of approach for ZFN-targeted mutagenesis in zebrafish embryos.

Figure 2.

Effects of ZFN mRNA injections on zebrafish embryos. a. Graph depicting proportion of embryos with indicated phenotypes associated with increasing doses of injected ZFN mRNA. b. Examples of normal (left panel) and abnormal (“monster”, right panel) embryos injected with 15 pg of each ZFN mRNA. Left panel, arrows indicate position of the eye. Right panel, arrowhead indicates defective eye formation and severe necrosis. c. PCR genotyping to assay presence or absence of NspI site. Wild type fragments indicative of the presence of the exon 2 NspI site are indicated by red and blue arrows. The larger undigested fragment caused by loss of the NspI site is indicated by a green arrow. d. Lesions observed in morphologically normal embryos injected with 15 pg of both ZFN mRNAs. Boxes indicate wild type *kdra* sequence. Pink shaded boxes label the ZFN recognition elements. Red dashes or letters indicate the positions of deletions or insertions, respectively. e. Graph depicting lesion frequency at the *kdra* exon 2 site and a subset of off-site targets in ZFN-injected embryos. Lesion frequency is proportion of sequences containing indels at the target over total number of sequences for a given site. Data for each group of embryos were obtained using Solexa sequencing as described in Materials and Methods.

Figure 3.

Segmental artery defects in ZFN-mutation bearing embryos. a., b. Phenotypically wild type embryo from a cross between founder 883.2M and a female heterozygous for *kdra^{um6}*. b. Normal segmental artery formation; segmental arteries are indicated by white arrows. c., d. Sibling of embryo in a. displaying normal morphology, but a failure to form segmental arteries. e., f. Normal morphology but partial segmental artery formation (white arrows) in a phenotypically mutant embryo from a cross between founder 889.3M and a female heterozygous for *kdra^{um6}*. g., h. Mutant embryo from an incross of *kdra^{um6}* heterozygous parents. h. Loss of segmental arteries in a *kdra^{um6}* mutant embryo. b, d, f, h – confocal images of embryos bearing the *Tg(fli1:egfp)^{y1}* transgene.

Figure 4.

Genotypic characterization of ZFN-induced mutations. a. NspI PCR genotyping assay on individual embryos derived from founders 889.7M and 889.1M. Embryos were genotyped for *kdra^{um6}* by sequencing (see Supplementary Fig. 10 for chromatograms); +/+ indicates homozygous wild type and +/- heterozygous for the *um6* allele; M – male, F – female. b. Cell nuclease genotyping of individual embryos derived from founder 889.1M; embryos are the same as those shown in a. c. Genetic lesions identified in mutant embryos derived from individual crosses between indicated founder and a *kdra^{um6}* heterozygote. ZFN recognition elements are shaded in pink and the ZFN cleavage position on this strand is indicated by an arrow. Gray dashes or letters indicate the positions of deletions or insertions, respectively, in the progeny sequences.