Supplementary Information for Doyon et al., "Heritable Targeted Gene Disruption in Zebrafish Using Designed Zinc Finger Nucleases"

Table of Contents

Supplementary Methods

Design, assembly, and *in vitro* evaluation of ZFN constructs

The published literature has accumulated a very extensive collection of methods for developing zinc finger proteins that target a sequence of one's choosing (Pabo et al, 2001). These include the rational elaboration of the protein-DNA interface (Desjarlais & Berg, 1992), use of affinity selection methodologies such as phage display (Rebar & Pabo, 1994; Greisman & Pabo, 1997; Segal et al, 1999; Beerli et al, 2000; Dreier et al, 2001; Isalan & Choo, 2001), biological selection methods such as 1 or 2-hybrid systems in budding yeast (Bartsevich & Juliano, 2000) or bacteria (Joung et al, 2000), and the use of naturally occurring zinc fingers (Bae et al, 2003). These strategies, combined with the modular reassortment of zinc fingers with pre-characterized specificites (Wright et al, 2006), have been successfully used by a large number of laboratories to develop DNA binding domains that engage investigator-specified targets in living cells in a broad range of organisms (Choo et al, 1994; Beerli et al, 1998; Beerli et al, 2000; Zhang et al, 2000; Bibikova et al, 2001; Liu et al, 2001; Bibikova et al, 2002; Ren et al, 2002; Bartsevich et al, 2003; Porteus & Baltimore, 2003; Reynolds et al, 2003; Tan et al, 2003; Lloyd et al, 2005; Wright et al, 2005; Beumer et al, 2006; Morton et al, 2006).

The particular approach we have adopted for developing zinc finger proteins for genome editing applications relies on an archive of pre-characterized two-finger modules, each of which recognizes an experimentally validated 6 bp half-site (Isalan et al, 2001; Moore et al, 2001). To develop ZFNs directed against zebrafish *gol* and *ntl* cDNAs, these sequences were scanned for positions where modules exist in the archive that allow the fusion of two such modules to form a 4-finger protein (ZFN-R) with a composite 12 bp target site recognized on the Watson strand, and another such 4 finger protein (ZFN-L) that recognizes a 12 bp target site on the Crick strand, 5 or 6 bp away from the 5' most base pair recognized by ZFN-R.

While a detailed analysis of the relative merits of the various approaches to ZFP design is outside the scope of the present discussion, we note that the process described above reliably produces ZFPs that pass DNA binding ELISA assays (see Supplementary Figs. 2 and 5), pass functional assays for activity in yeast-based proxy systems (Fig. 2A and 3A), and, when fused to the FokI endonuclease domain, efficiently and specifically engage their intended, endogenous target loci in the context of complex genomes such as zebrafish (present work), hamster (Santiago et al, 2008), and human (Urnov et al, 2005; Lombardo et al, 2007; Miller et al, 2007; Perez et al, *Nature Biotechnology*, in press).

These ZFNs are assembled using a PCR-based procedure described in Supplementary Fig. 1. In brief, each two-finger module is amplified by PCR in a separate reaction. The two PCR products that correspond to the two-finger modules that compose each ZFN are then combined and joined by conventional restriction enzyme digestion-ligation into a ZFN expression vector to yield a gene encoding (NH_2 to COOH) a triple-FLAG tag, a nuclear localization signal, the ZFP module, and the endonuclease domain of the type IIS restriction enzyme FokI. The ZFN coding region is flanked with the cytomegalovirus (CMV) immediate-early promoter and bovine growth hormone (BGH) polyadenylation signal. The complete sequence of all ZFNs used in this work are shown in Supplementary Figs. 4 and 6. Each ZFN was evaluated for DNA binding using an ELISA assay performed as described (Bartsevich et al, 2003).

The *in vitro* consensus binding site for each ZFN was determined using a procedure described in detail elsewhere (Perez et al, *Nature Biotechnology* in press). In brief, HA-tagged ZFPs were synthesized by in vitro transcription-translation, and SELEX was performed by incubation with a pool of randomized DNA sequences and an anti-HA biotin-coupled antibody, capture of protein-bound DNA by streptavidin magnetic-coated beads, and PCR-based amplification of the bound DNA, and use of the resulting PCR pool in a second round of SELEX.

This procedure was repeated for a total of 4 rounds of selection, and DNA fragments amplified after the final round were cloned and sequenced.

Use of a Budding Yeast-Based Reporter System to Identify ZFN Pairs for Gene Disruption

Construction of the reporter plasmid

The reporter construct was targeted to the HO locus using the yeast integrating plasmid (YIp) HO-poly-KanMX-HO (Voth et al, 2001). To generate the SSA reporter (Supplementary Fig. 3), a fragment corresponding to nucleotides 1 to 750 of the *MEL1* gene (Liljestrom, 1985) (relative to the ATG) was cloned into the SalI and BamHI sites of HO-poly-KanMX-HO using the following primers: 5'-AATTGTCGACATGTTTGCTTTCTACTTTCTCACCGC-3' and 5'-AATTGGATCCCCCCATTGGAGCTGCC-3'. Then a fragment from nucleotides 299 to 2100 was cloned into the SacI and EcoRI sites using the following primers: 5'-AATTGAGCTCAGACCACCTGCATAATAACAGC-3' and 5'-AATTGAATTCGGGCAAAAATTGGTACCAATGC-3'. Finally, a 1489 bp fragment of the *PGK1* promoter was cloned into the BsiWI and SalI sites using the following primers: 5'-AATTCGTACGTCTAACTGATCTATCCAAAACTG-3' and 5'-AATTGTCGACTTGATCTTTTGGTTTTATATTTGTTG-3'.

Construction of the reporter strain

Integration of the reporter construct into the 69-1B strain (S288C background; *MAT***a** *his3 200 lys2-128δ leu2 1*) was performed as described (Voth et al, 2001). Note that the designer deletion strain BY4741, available from Open Biosystems, provides the same characteristics as the 69-1B strain used in this study. Briefly, 2 g of the reporter construct containing the target sequence was

linearized with NotI and used to transform yeast using the lithium acetate method (Gietz & Schiestl, 2007). We confirmed correct integration by colony PCR using the following primers: HO-L: 5'-TATTAGGTGTGAAACCACGAAAAGT-3' and 5'-ACTGTCATTGGGAATGTCTTATGAT-3'; HO-R: 5'-ATTACGCTCGTCATCAAAATCA-3'; 5'-CATGTCTTCTCGTTAAGACTGCAT-3'.

ZFN expression vectors

The entire coding sequence of each ZFN pair was transferred to galactose inducible expression vectors using standard cloning procedures (Mumberg et al, 1994; Urnov et al, 2005; Moehle et al, 2007). These YCp vectors, p413prom and p415prom, are available through the ATCC (Mumberg et al, 1994). Transformation of the reporter strain with the ZFN expression vectors (plasmid names listed in Supplementary Table 4) was done in deep well blocks as described (Gietz & Schiestl, 2007).

Induction of ZFN expression

To derepress the *GAL1* promoter, the pools of transformants were diluted 1:10 into 1 ml of SC His-Leu- medium containing 2% raffinose as a source of carbon and incubated overnight at 30ºC. ZFN expression was induced by diluting the raffinose cultures 1:10 into 1ml of SC His-Leu- medium containing 2% galactose. Cells were then incubated for 2 to 6 hours, before addition of 2% glucose to stop expression. Cells were then incubated overnight to allow for DSB repair and reporter expression.

Reporter assay (MEL1 assay)

The first step was to determine the cell density of the cultures in order to normalize the reporter signal to the amount of cells in the culture. This was done by a simple spectrophotometric reading

at 600 nm. The deep well block was then centrifuged at 3000g for 5 minutes to pellet yeast cells and 10 ul of the media was assayed for Mel1 activity as described (Ryan et al, 1998; Chen et al, 2004).

References for Supplementary Information

Bae KH, Kwon YD, Shin HC, Hwang MS, Ryu EH, Park KS, Yang HY, Lee DK, Lee Y, Park J, Kwon HS, Kim HW, Yeh BI, Lee HW, Sohn SH, Yoon J, Seol W, Kim JS (2003) Human zinc fingers as building blocks in the construction of artificial transcription factors. *Nat Biotechnol* **21:** 275-280

Bartsevich VV, Juliano RL (2000) Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries. *Mol Pharmacol* **58:** 1-10.

Bartsevich VV, Miller JC, Case CC, Pabo CO (2003) Engineered zinc finger proteins for controlling stem cell fate. *Stem Cells* **21:** 632-637

Beerli RR, Dreier B, Barbas CF, 3rd (2000) Positive and negative regulation of endogenous genes by designed transcription factors. *Proc Natl Acad Sci U S A* **97:** 1495-1500

Beerli RR, Segal DJ, Dreier B, Barbas CF, 3rd (1998) Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc Natl Acad Sci U S A* **95:** 14628-14633

Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D (2006) Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics* **172:** 2391-2403

Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol Cell Biol* **21:** 289-297.

Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc- finger nucleases. *Genetics* **161:** 1169-1175.

Chen Q, Chen J, Sun T, Shen J, Shen X, Jiang H (2004) A yeast two-hybrid technology-based system for the discovery of PPARgamma agonist and antagonist. *Analytical biochemistry* **335:** 253-259

Choo Y, Sanchez-Garcia I, Klug A (1994) In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **372:** 642-645.

Desjarlais JR, Berg JM (1992) Toward rules relating zinc finger protein sequences and DNA binding site preferences. *Proc Natl Acad Sci U S A* **89:** 7345-7349

Dreier B, Beerli RR, Segal DJ, Flippin JD, Barbas CF, 3rd (2001) Development of zinc finger domains for recognition of the 5'-ANN-3' family of DNA sequences and their use in the construction of artificial transcription factors. *J Biol Chem* **276:** 29466-29478.

Gietz RD, Schiestl RH (2007) Microtiter plate transformation using the LiAc/SS carrier DNA/PEG method. *Nature protocols* **2:** 5-8

Greisman HA, Pabo CO (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science* **275:** 657-661

Isalan M, Choo Y (2001) Rapid, high-throughput engineering of sequence-specific zinc finger DNA- binding proteins. *Methods Enzymol* **340:** 593-609

Isalan M, Klug A, Choo Y (2001) A rapid, generally applicable method to engineer zinc fingers illustrated by targeting the HIV-1 promoter. *Nat Biotechnol* **19:** 656-660.

Joung JK, Ramm EI, Pabo CO (2000) A bacterial two-hybrid selection system for studying protein-DNA and protein-protein interactions. *Proc Natl Acad Sci U S A* **97:** 7382-7387

Liljestrom PL (1985) The nucleotide sequence of the yeast MEL1 gene. *Nucleic Acids Res* **13:** 7257-7268

Liu PQ, Rebar EJ, Zhang L, Liu Q, Jamieson AC, Liang Y, Qi H, Li PX, Chen B, Mendel MC, Zhong X, Lee YL, Eisenberg SP, Spratt SK, Case CC, Wolffe AP (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. *J Biol Chem* **276:** 11323-11334.

Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proc Natl Acad Sci U S A* **102:** 2232-2237

Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA, Ando D, Urnov FD, Galli C, Gregory PD, Holmes MC, Naldini L (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol* **25:** 1298-1306

Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol* **25:** 778-785

Moehle EA, Rock JM, Lee YL, Jouvenot Y, Dekelver RC, Gregory PD, Urnov FD, Holmes MC (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A* **104:** 3055-3060

Moore M, Choo Y, Klug A (2001) Design of polyzinc finger peptides with structured linkers. *Proc Natl Acad Sci U S A* **98:** 1432-1436.

Morton J, Davis MW, Jorgensen EM, Carroll D (2006) Induction and repair of zinc-finger nuclease-targeted double-strand breaks in Caenorhabditis elegans somatic cells. *Proc Natl Acad Sci U S A* **103:** 16370-16375

Mumberg D, Muller R, Funk M (1994) Regulatable promoters of Saccharomyces cerevisiae: comparison of transcriptional activity and their use for heterologous expression. *Nucleic Acids Res* **22:** 5767-5768

Pabo CO, Peisach E, Grant RA (2001) Design and Selection of Novel Cys2his2 Zinc Finger Proteins. *Annu Rev Biochem* **70:** 313-340.

Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* **300:** 763.

Rebar EJ, Pabo CO (1994) Zinc finger phage: affinity selection of fingers with new DNA-binding specificities. *Science* **263:** 671-673

Ren D, Collingwood TN, Rebar EJ, Wolffe AP, Camp HS (2002) PPARγ knockdown by engineered transcription factors: exogenous PPARγ2 but not PPARγ1 reactivates adipogenesis. *Genes Dev* **16:** 27-32.

Reynolds L, Ullman C, Moore M, Isalan M, West MJ, Clapham P, Klug A, Choo Y (2003) Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proc Natl Acad Sci U S A* **100:** 1615-1620.

Ryan MP, Jones R, Morse RH (1998) SWI-SNF complex participation in transcriptional activation at a step subsequent to activator binding. *Mol Cell Biol* **18:** 1774-1782.

Santiago Y, Chan E, Liu PQ, Orlando S, Zhang L, Urnov FD, Holmes MC, Guschin D, Waite A, Miller JC, Rebar EJ, Gregory PD, Klug A, Collingwood TN (2008) Targeted gene knockout in mammalian cells using engineered zinc finger nucleases. *Proc Natl Acad Sci U S A* **105:** 5809-5814

Segal DJ, Dreier B, Beerli RR, Barbas CF, 3rd (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A* **96:** 2758-2763

Tan S, Guschin D, Davalos A, Lee YL, Snowden AW, Jouvenot Y, Zhang HS, Howes K, McNamara AR, Lai A, Ullman C, Reynolds L, Moore M, Isalan M, Berg LP, Campos B, Qi H, Spratt SK, Case CC, Pabo CO, Campisi J, Gregory PD (2003) Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity. *Proc Natl Acad Sci U S A* **100:** 11997-12002

Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature* **435:** 646-651

Voth WP, Richards JD, Shaw JM, Stillman DJ (2001) Yeast vectors for integration at the HO locus. *Nucleic Acids Res* **29:** E59-59

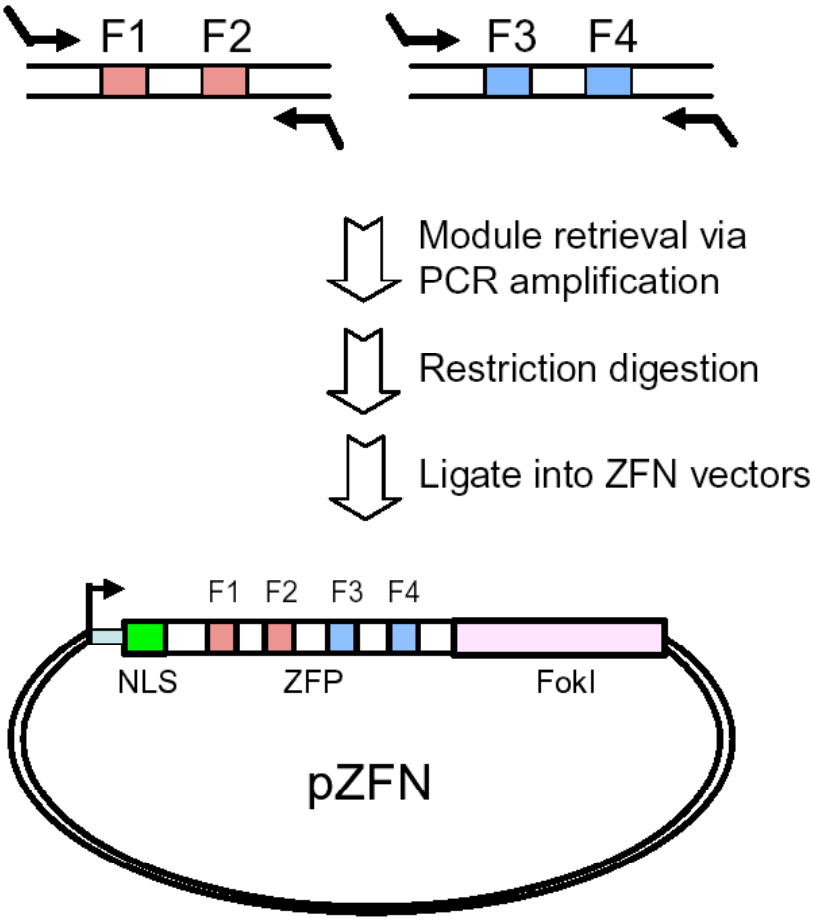
Wright DA, Thibodeau-Beganny S, Sander JD, Winfrey RJ, Hirsh AS, Eichtinger M, Fu F, Porteus MH, Dobbs D, Voytas DF, Joung JK (2006) Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. *Nature protocols* **1:** 1637-1652

Wright DA, Townsend JA, Winfrey RJ, Jr., Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J* **44:** 693-705

Zhang L, Spratt SK, Liu Q, Johnstone B, Qi H, Raschke EE, Jamieson AC, Rebar EJ, Wolffe AP, Case CC (2000) Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. *J Biol Chem* **275:** 33850-33860.

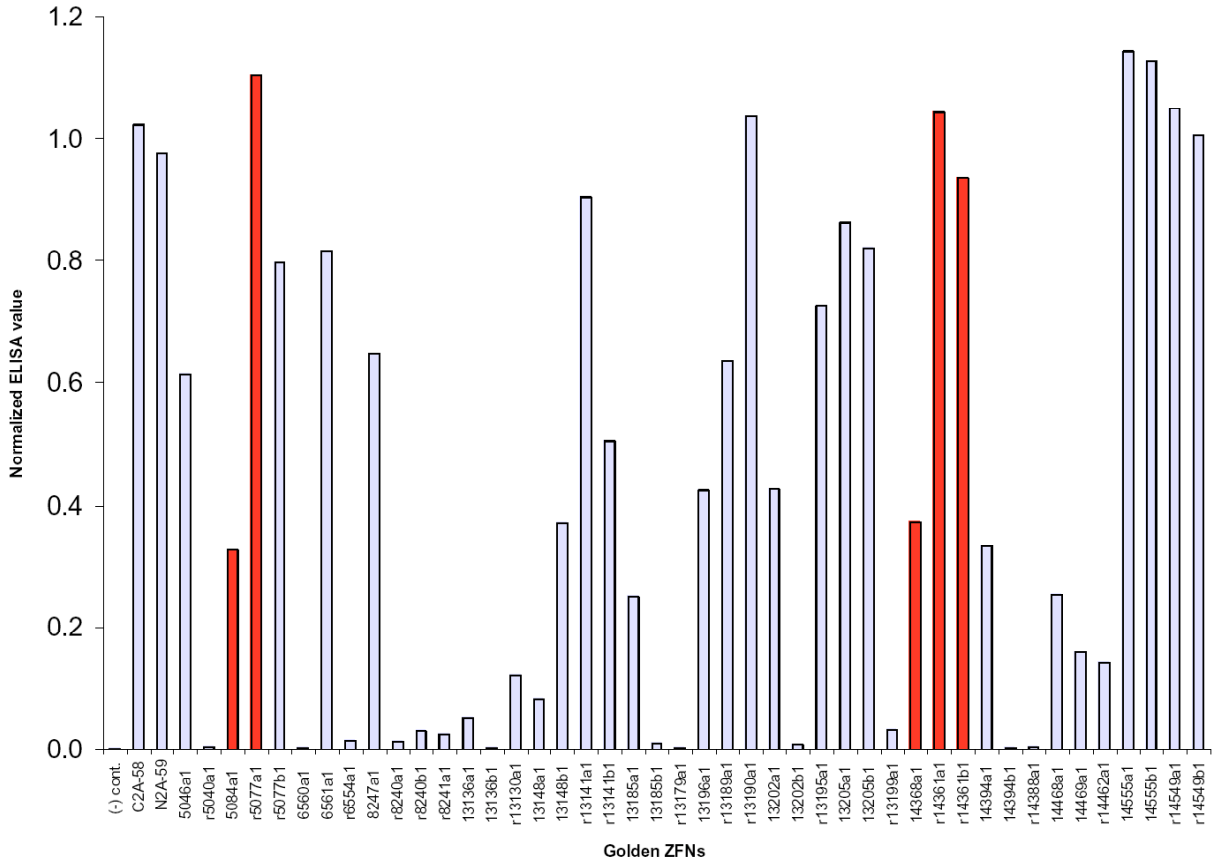
Supplementary Figure 1: Schematic of the PCR-based ZFN assembly procedure from individual 2-finger modules. Each two-finger module is amplified by PCR in a separate reaction. The two PCR products that correspond to the two-finger modules that compose each ZFN are then combined and joined by conventional restriction enzyme digestion-ligation into a ZFN expression vector to yield a gene encoding, NH_2 to COOH, a triple-FLAG tag, a nuclear localization signal, the ZFP module, and the endonuclease domain of the type IIS restriction enzyme FokI.

.



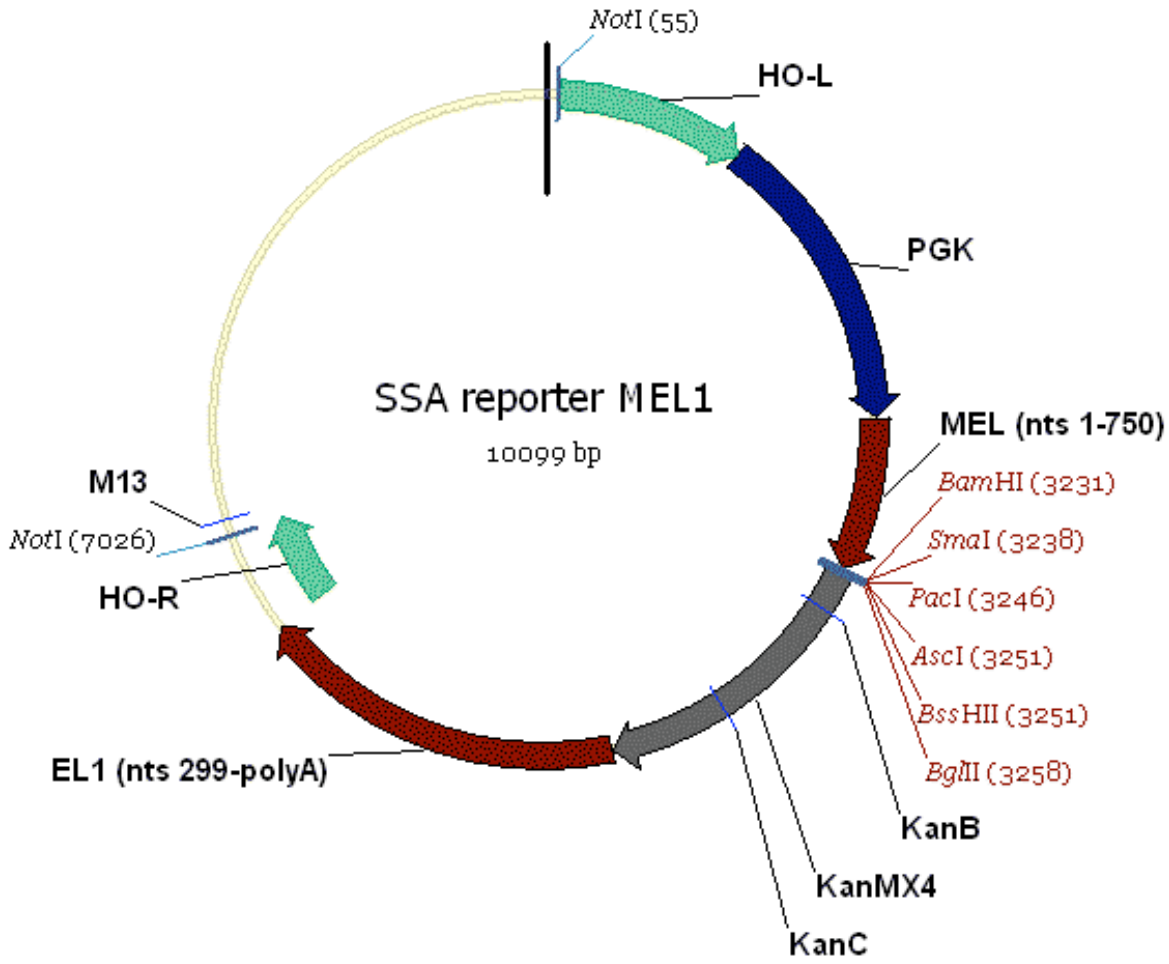
NLS: nuclear localization signal
F1-4: ZFP helices
FokI: FokI cleavage domain

Supplementary Figure 2: ELISA data for DNA binding by all the ZFPs designed and assembled against the *golden* locus. Each ZFN was evaluated for DNA binding using an ELISA assay performed as described (Bartsevich et al, 2003). The first sample is a negative control; the second and third sample, respectively, are positive controls and represent ZFN-R and ZFN-L from previous work (Urnov et al, 2005). Each ZFN is annotated by a specific base-pair position within the *golden* locus (the cDNA cannot be used as a reference in this context because some ZFNs span exon/intron boundaries). In this nomenclature, position #1 corresponds to the transcription start site of the *golden* gene as per the UCSC genome browser DanRer5 zebrafish genome annotation (July 2005); an "r" before the ZFN name indicates that its primary recognition is the Crick strand. The ZFNs highlighted in red form pairs 1 and 14/15, respectively, used in the primary text. Pair 1 ZFN-R is 5084a1 and ZFN-L is r5077a1; Pair 14/15 ZFN-R is 14368a1, Pair 14 ZFN-L is r14361b1, and Pair 15 ZFN-L is r14361a1..



Supplementary Figure 3: Schematic, restriction map, and sequence of the reporter plasmid used in the budding yeast ZFN activity screening system.

Schematic and restriction map:



Fea

tures: total length, 10099 bp
HO left homology arm: bases 67-984
PGK1 promoter: bases 985-2473
5' end of the *MEL1* gene: bases 2480-3229
MCS: bases 3231-3257
kanMX4 cassette: bases 3258-4705
3' end of the *MEL1* gene: bases 4707-6516
HO right homology arm: bases 6517-7029
pUK21 backbone: bases 7027-54

Vector sequence (also available as a Vector NTI file from furnov@sangamo.com upon request):

```
   1    ATGACCATGA TTACGCCACT AGTCCGAGGC CTCGAGATCC GATATCGCCG TGGCGGCCGC CAGCTGAAGC TTAATTATCC TGGGCACGAG
        TGAAACAAAG
        TACTGGTACT AATGCGGTGA TCAGGCTCCG GAGCTCTAGG CTATAGCGGC ACCGCCGGCG GTCGACTTCG AATTAATAGG ACCCGTGCTC
        ACTTTGTTTC
 101    CTAAAACCTT TATTTAGCAT GGCCATTGAA TGTAACAATT ATATATATCG CAAGCACAAA AAATCAAGGA GAGAGAACTA CCACTTTGTT
        CATGTGTACA
        GATTTTGGAA ATAAATCGTA CCGGTAACTT ACATTGTTAA TATATATAGC GTTCGTGTTT TTTAGTTCCT CTCTCTTGAT GGTGAAACAA
        GTACACATGT
 201    ATGTTCATTA TCTCCATAAG CAAAAAAAAA AAATAGAAAA CATATGCTAT AAGGTTGATA TTCTCACGAG TAAGCGGCAC TTGCTACTTA
        TTGACATTGC
        TACAAGTAAT AGAGGTATTC GTTTTTTTTT TTTATCTTTT GTATACGATA TTCCAACTAT AAGAGTGCTC ATTCGCCGTG AACGATGAAT
        AACTGTAACG
 301    AGATTTTTGG CTACAGAAAT AGTATATTAG AGATTATAAT TGCTAATCAA ATCAAAATAT AAAATTAGTA AACCAAACCA TTTATACCCT
        TCCTTAGTAG
        TCTAAAAACC GATGTCTTTA TCATATAATC TCTAATATTA ACGATTAGTT TAGTTTTATA TTTTAATCAT TTGGTTTGGT AAATATGGGA
        AGGAATCATC
 401    TTATGGATTG TTTTTTAATG ATATTTCTGC AAACCAAAGA AAGATTGTTA TCCAGATAGA ATTTAGTTTT GATATTCATT TTTTTGTTGA
        AGATTGAACG
        AATACCTAAC AAAAAATTAC TATAAAGACG TTTGGTTTCT TTCTAACAAT AGGTCTATCT TAAATCAAAA CTATAAGTAA AAAAACAACT
        TCTAACTTGC
 501    CCATATCTGG GCCTCATAAT TCAAAAGACG GTGCCATTAT CGGTAGCGTT TCGCATTGTA CTGGATTTCA GAAATTTCAC AGTTGATGAA
        TCGAAAAGAA
        GGTATAGACC CGGAGTATTA AGTTTTCTGC CACGGTAATA GCCATCGCAA AGCGTAACAT GACCTAAAGT CTTTAAAGTG TCAACTACTT
        AGCTTTTCTT
 601    TGGTCTCATT GCAACACGTA AGGTTAAGAT GTCCCTTTTT ACCATTATAG GCAATAAATG AATCATAAAA CGACCGTATA CTGGTGAAAT
        AGTAGGGAGA
        ACCAGAGTAA CGTTGTGCAT TCCAATTCTA CAGGGAAAAA TGGTAATATC CGTTATTTAC TTAGTATTTT GCTGGCATAT GACCACTTTA
        TCATCCCTCT
 701    ACGAGTACCT GTAGTAAAAA GTATAAATCA TAGTTAATCG GGCAATGTCC CTCGATCAAG GAGTATTGTG TCATGTTCGA GACAAACGCC
        AACATTTTTG
        TGCTCATGGA CATCATTTTT CATATTTAGT ATCAATTAGC CCGTTACAGG GAGCTAGTTC CTCATAACAC AGTACAAGCT CTGTTTGCGG
        TTGTAAAAAC
 801    TTTCTTTTGG ACAAATGTTG TTTGCATTTA TGATCCGTTA TATTTTGATC TAATGTAGAG TTGCACGTAG TTCTTACTGG CAAAGAAATC
        GATGCATACC
        AAAGAAAACC TGTTTACAAC AAACGTAAAT ACTAGGCAAT ATAAAACTAG ATTACATCTC AACGTGCATC AAGAATGACC GTTTCTTTAG
        CTACGTATGG
 901    AAAAAAGAAT AAAGGTGATA TTTGATCTTT ACCGTTTAGT TCCAACGTAA AATTGTGCCT TTGGACTTAA AATGGCGTCG TACGTCTAAC
        TGATCTATCC
        TTTTTTCTTA TTTCCACTAT AAACTAGAAA TGGCAAATCA AGGTTGCATT TTAACACGGA AACCTGAATT TTACCGCAGC ATGCAGATTG
        ACTAGATAGG
1001    AAAACTGAAA ATTACATTCT TGATTAGGTT TATCACAGGC AAATGTAATT TGTGGTATTT TGCCGTTCAA AATCTGTAGA ATTTTCTCAT
        TGGTCACATT
        TTTTGACTTT TAATGTAAGA ACTAATCCAA ATAGTGTCCG TTTACATTAA ACACCATAAA ACGGCAAGTT TTAGACATCT TAAAAGAGTA
        ACCAGTGTAA
1101    ACAACCTGAA AATACTTTAT CTACAATCAT ACCATTCTTA TAACATGTCC CCTTAATACT AGGATCAGGC ATGAACGCAT CACAGACAAA
        ATCTTCTTGA
        TGTTGGACTT TTATGAAATA GATGTTAGTA TGGTAAGAAT ATTGTACAGG GGAATTATGA TCCTAGTCCG TACTTGCGTA GTGTCTGTTT
        TAGAAGAACT
1201    CAAACGTCAC AATTGATCCC TCCCCATCCG TTATCACAAT GACAGGTGTC ATTTTGTGCT CTTATGGGAC GATCCTTATT ACCGCTTCAT
        CCGGTGATAG
        GTTTGCAGTG TTAACTAGGG AGGGGTAGGC AATAGTGTTA CTGTCCACAG TAAAACACGA GAATACCCTG CTAGGAATAA TGGCGAAGTA
        GGCCACTATC
1301    ACCGCCACAG AGGGGCAGAG AGCAATCATC ACCTGCAAAC CCTTCTATAC ACTCACATCT ACCAGTGTAC GAATTGCATT CAGAAAACTG
        TTTGCATTCA
        TGGCGGTGTC TCCCCGTCTC TCGTTAGTAG TGGACGTTTG GGAAGATATG TGAGTGTAGA TGGTCACATG CTTAACGTAA GTCTTTTGAC
        AAACGTAAGT
1401    AAAATAGGTA GCATACAATT AAAACATGGC GGGCATGTAT CATTGCCCTT ATCTTGTGCA GTTAGACGCG AATTTTTCGA AGAAGTACCT
        TCAAAGAATG
        TTTTATCCAT CGTATGTTAA TTTTGTACCG CCCGTACATA GTAACGGGAA TAGAACACGT CAATCTGCGC TTAAAAAGCT TCTTCATGGA
        AGTTTCTTAC
1501    GGGTCTTATC TTGTTTTGCA AGTACCACTG AGCAGGATAA TAATAGAAAT GATAATATAC TATAGTAGAG ATAACGTCGA TGACTTCCCA
        TACTGTAATT
        CCCAGAATAG AACAAAACGT TCATGGTGAC TCGTCCTATT ATTATCTTTA CTATTATATG ATATCATCTC TATTGCAGCT ACTGAAGGGT
        ATGACATTAA
1601    GCTTTTAGTT GTGTATTTTT AGTGTGCAAG TTTCTGTAAA TCGATTAATT TTTTTTTCTT TCCTCTTTTT ATTAACCTTA ATTTTTATTT
        TAGATTCCTG
        CGAAAATCAA CACATAAAAA TCACACGTTC AAAGACATTT AGCTAATTAA AAAAAAAGAA AGGAGAAAAA TAATTGGAAT TAAAAATAAA
        ATCTAAGGAC
1701    ACTTCAACTC AAGACGCACA GATATTATAA CATCTGCATA ATAGGCATTT GCAAGAATTA CTCGTGAGTA AGGAAAGAGT GAGGAACTAT
        CGCATACCTG
        TGAAGTTGAG TTCTGCGTGT CTATAATATT GTAGACGTAT TATCCGTAAA CGTTCTTAAT GAGCACTCAT TCCTTTCTCA CTCCTTGATA
        GCGTATGGAC
1801    CATTTAAAGA TGCCGATTTG GGCGCGAATC CTTTATTTTG GCTTCACCCT CATACTATTA TCAGGGCCAG AAAAAGGAAG TGTTTCCCTC
        CTTCTTGAAT
        GTAAATTTCT ACGGCTAAAC CCGCGCTTAG GAAATAAAAC CGAAGTGGGA GTATGATAAT AGTCCCGGTC TTTTTCCTTC ACAAAGGGAG
        GAAGAACTTA
1901    TGATGTTACC CTCATAAAGC ACGTGGCCTC TTATCGAGAA AGAAATTACC GTCGCTCGTG ATTTGTTTGC AAAAAGAACA AAACTGAAAA
        AACCCAGACA
        ACTACAATGG GAGTATTTCG TGCACCGGAG AATAGCTCTT TCTTTAATGG CAGCGAGCAC TAAACAAACG TTTTTCTTGT TTTGACTTTT
        TTGGGTCTGT
2001    CGCTCGACTT CCTGTCTTCC TATTGATTGC AGCTTCCAAT TTCGTCACAC AACAAGGTCC TAGCGACGGC TCACAGGTTT TGTAACAAGC
        AATCGAAGGT
        GCGAGCTGAA GGACAGAAGG ATAACTAACG TCGAAGGTTA AAGCAGTGTG TTGTTCCAGG ATCGCTGCCG AGTGTCCAAA ACATTGTTCG
        TTAGCTTCCA
2101    TCTGGAATGG CGGGAAAGGG TTTAGTACCA CATGCTATGA TGCCCACTGT GATCTCCAGA GCAAAGTTCG TTCGATCGTA CTGTTACTCT
        CTCTCTTTCA
        AGACCTTACC GCCCTTTCCC AAATCATGGT GTACGATACT ACGGGTGACA CTAGAGGTCT CGTTTCAAGC AAGCTAGCAT GACAATGAGA
        GAGAGAAAGT
```

```
2201        AACAGAATTG TCCGAATCGT GTGACAACAA CAGCCTGTTC TCACACACTC TTTTCTTCTA ACCAAGGGGG TGGTTTAGTT TAGTAGAACC
            TCGTGAAACT
            TTGTCTTAAC AGGCTTAGCA CACTGTTGTT GTCGGACAAG AGTGTGTGAG AAAAGAAGAT TGGTTCCCCC ACCAAATCAA ATCATCTTGG
            AGCACTTTGA
2301        TACATTTACA TATATATAAA CTTGCATAAA TTGGTCAATG CAAGAAATAC ATATTTGGTC TTTTCTAATT CGTAGTTTTT CAAGTTCTTA
            GATGCTTTCT
            ATGTAAATGT ATATATATTT GAACGTATTT AACCAGTTAC GTTCTTTATG TATAAACCAG AAAAGATTAA GCATCAAAAA GTTCAAGAAT
            CTACGAAAGA
2401        TTTTCTCTTT TTTACAGATC ATCAAGGAAG TAATTATCTA CTTTTTACAA CAAATATAAA ACCAAAAGAT CAAGTCGACA TGTTTGCTTT
            CTACTTTCTC
            AAAAGAGAAA AAATGTCTAG TAGTTCCTTC ATTAATAGAT GAAAAATGTT GTTTATATTT TGGTTTTCTA GTTCAGCTGT ACAAACGAAA
            GATGAAAGAG
2501        ACCGCATGCA TCAGTTTGAA GGGCGTTTTT GGGGTGTCTC CGAGTTACAA TGGCCTTGGT CTCACTCCAC AGATGGGTTG GGACAACTGG
            AATACGTTTG
            TGGCGTACGT AGTCAAACTT CCCGCAAAAA CCCCACAGAG GCTCAATGTT ACCGGAACCA GAGTGAGGTG TCTACCCAAC CCTGTTGACC
            TTATGCAAAC
2601        CCTGCGATGT CAGTGAACAG CTACTTCTAG ACACCGCTGA TAGAATTTCT GACTTGGGGC TAAAGGATAT GGGTTACAAG TATATCATTC
            TGGATGACTG
            GGACGCTACA GTCACTTGTC GATGAAGATC TGTGGCGACT ATCTTAAAGA CTGAACCCCG ATTTCCTATA CCCAATGTTC ATATAGTAAG
            ACCTACTGAC
2701        CTGGTCTAGC GGCAGAGATT CCGACGGTTT CCTCGTTGCA GATGAACAAA AATTTCCCAA TGGTATGGGC CATGTTGCAG ACCACCTGCA
            TAATAACAGC
            GACCAGATCG CCGTCTCTAA GGCTGCCAAA GGAGCAACGT CTACTTGTTT TTAAAGGGTT ACCATACCCG GTACAACGTC TGGTGGACGT
            ATTATTGTCG
2801        TTTCTTTTCG GTATGTATTC GTCTGCTGGT GAGTACACCT GTGCTGGATA TCCTGGGTCT CTGGGTCGTG AGGAAGAAGA TGCACAGTTC
            TTTGCAAATA
            AAAGAAAAGC CATACATAAG CAGACGACCA CTCATGTGGA CACGACCTAT AGGACCCAGA GACCCAGCAC TCCTTCTTCT ACGTGTCAAG
            AAACGTTTAT
2901        ACCGCGTTGA CTACTTGAAG TACGATAATT GTTACAATAA GGGTCAGTTT GGTACACCGG AAATTTCTTA CCACCGTTAC AAGGCCATGT
            CAGATGCTTT
            TGGCGCAACT GATGAACTTC ATGCTATTAA CAATGTTATT CCCAGTCAAA CCATGTGGCC TTTAAAGAAT GGTGGCAATG TTCCGGTACA
            GTCTACGAAA
3001        GAATAAAACT GGTAGGCCTA TATTCTATTC TCTATGTAAC TGGGGTCAGG ATTTAACATT TTACTGGGGC TCTGGTATCG CCAATTCTTG
            GAGAATGAGT
            CTTATTTTGA CCATCCGGAT ATAAGATAAG AGATACATTG ACCCCAGTCC TAAATTGTAA AATGACCCCG AGACCATAGC GGTTAAGAAC
            CTCTTACTCA
3101        GGAGATGTTA CTGCTGAGTT CACTCGTCCA GATAGCAGAT GTCCCTGTGA TGGCGATGAA TACGATTGCA AGTACGCCGG TTTCCATTGT
            TCTATTATGA
            CCTCTACAAT GACGACTCAA GTGAGCAGGT CTATCGTCTA CAGGGACACT ACCGCTACTT ATGCTAACGT TCATGCGGCC AAAGGTAACA
            AGATAATACT
3201        ATATTCTTAA CAAGGCAGCT CCAATGGGGG GATCCCCGGG TTAATTAAGG CGCGCCAGAT CTGTTTAGCT TGCCTCGTCC CCGCCGGGTC
            ACCCGGCCAG
            TATAAGAATT GTTCCGTCGA GGTTACCCCC CTAGGGGCCC AATTAATTCC GCGCGGTCTA GACAAATCGA ACGGAGCAGG GGCGGCCCAG
            TGGGCCGGTC
3301        CGACATGGAG GCCCAGAATA CCCTCCTTGA CAGTCTTGAC GTGCGCAGCT CAGGGGCATG ATGTGACTGT CGCCCGTACA TTTAGCCCAT
            ACATCCCCAT
            GCTGTACCTC CGGGTCTTAT GGGAGGAACT GTCAGAACTG CACGCGTCGA GTCCCCGTAC TACACTGACA GCGGGCATGT AAATCGGGTA
            TGTAGGGGTA
3401        GTATAATCAT TTGCATCCAT ACATTTTGAT GGCCGCACGG CGCGAAGCAA AAATTACGGC TCCTCGCTGC AGACCTGCGA GCAGGGAAAC
            GCTCCCCTCA
            CATATTAGTA AACGTAGGTA TGTAAAACTA CCGGCGTGCC GCGCTTCGTT TTTAATGCCG AGGAGCGACG TCTGGACGCT CGTCCCTTTG
            CGAGGGGAGT
3501        CAGACGCGTT GAATTGTCCC CACGCCGCGC CCCTGTAGAG AAATATAAAA GGTTAGGATT TGCCACTGAG GTTCTTCTTT CATATACTTC
            CTTTTAAAAT
            GTCTGCGCAA CTTAACAGGG GTGCGGCGCG GGGACATCTC TTTATATTTT CCAATCCTAA ACGGTGACTC CAAGAAGAAA GTATATGAAG
            GAAAATTTTA
3601        CTTGCTAGGA TACAGTTCTC ACATCACATC CGAACATAAA CAACCATGGG TAAGGAAAAG ACTCACGTTT CGAGGCCGCG ATTAAATTCC
            AACATGGATG
            GAACGATCCT ATGTCAAGAG TGTAGTGTAG GCTTGTATTT GTTGGTACCC ATTCCTTTTC TGAGTGCAAA GCTCCGGCGC TAATTTAAGG
            TTGTACCTAC
3701        CTGATTTATA TGGGTATAAA TGGGCTCGCG ATAATGTCGG GCAATCAGGT GCGACAATCT ATCGATTGTA TGGGAAGCCC GATGCGCCAG
            AGTTGTTTCT
            GACTAAATAT ACCCATATTT ACCCGAGCGC TATTACAGCC CGTTAGTCCA CGCTGTTAGA TAGCTAACAT ACCCTTCGGG CTACGCGGTC
            TCAACAAAGA
3801        GAAACATGGC AAAGGTAGCG TTGCCAATGA TGTTACAGAT GAGATGGTCA GACTAAACTG GCTGACGGAA TTTATGCCTC TTCCGACCAT
            CAAGCATTTT
            CTTTGTACCG TTTCCATCGC AACGGTTACT ACAATGTCTA CTCTACCAGT CTGATTTGAC CGACTGCCTT AAATACGGAG AAGGCTGGTA
            GTTCGTAAAA
3901        ATCCGTACTC CTGATGATGC ATGGTTACTC ACCACTGCGA TCCCCGGCAA AACAGCATTC CAGGTATTAG AAGAATATCC TGATTCAGGT
            GAAAATATTG
            TAGGCATGAG GACTACTACG TACCAATGAG TGGTGACGCT AGGGGCCGTT TTGTCGTAAG GTCCATAATC TTCTTATAGG ACTAAGTCCA
            CTTTTATAAC
4001        TTGATGCGCT GGCAGTGTTC CTGCGCCGGT TGCATTCGAT TCCTGTTTGT AATTGTCCTT TTAACAGCGA TCGCGTATTT CGTCTCGCTC
            AGGCGCAATC
            AACTACGCGA CCGTCACAAG GACGCGGCCA ACGTAAGCTA AGGACAAACA TTAACAGGAA AATTGTCGCT AGCGCATAAA GCAGAGCGAG
            TCCGCGTTAG
4101        ACGAATGAAT AACGGTTTGG TTGATGCGAG TGATTTTGAT GACGAGCGTA ATGGCTGGCC TGTTGAACAA GTCTGGAAAG AAATGCATAA
            GCTTTTGCCA
            TGCTTACTTA TTGCCAAACC AACTACGCTC ACTAAAACTA CTGCTCGCAT TACCGACCGG ACAACTTGTT CAGACCTTTC TTTACGTATT
            CGAAAACGGT
4201        TTCTCACCGG ATTCAGTCGT CACTCATGGT GATTTCTCAC TTGATAACCT TATTTTTGAC GAGGGGAAAT TAATAGGTTG TATTGATGTT
            GGACGAGTCG
            AAGAGTGGCC TAAGTCAGCA GTGAGTACCA CTAAAGAGTG AACTATTGGA ATAAAAACTG CTCCCCTTTA ATTATCCAAC ATAACTACAA
            CCTGCTCAGC
4301        GAATCGCAGA CCGATACCAG GATCTTGCCA TCCTATGGAA CTGCCTCGGT GAGTTTTCTC CTTCATTACA GAAACGGCTT TTTCAAAAAT
            ATGGTATTGA
            CTTAGCGTCT GGCTATGGTC CTAGAACGGT AGGATACCTT GACGGAGCCA CTCAAAAGAG GAAGTAATGT CTTTGCCGAA AAAGTTTTTA
            TACCATAACT
4401        TAATCCTGAT ATGAATAAAT TGCAGTTTCA TTTGATGCTC GATGAGTTTT TCTAATCAGT ACTGACAATA AAAAGATTCT TGTTTTCAAG
            AACTTGTCAT
            ATTAGGACTA TACTTATTTA ACGTCAAAGT AAACTACGAG CTACTCAAAA AGATTAGTCA TGACTGTTAT TTTTCTAAGA ACAAAAGTTC
            TTGAACAGTA
```

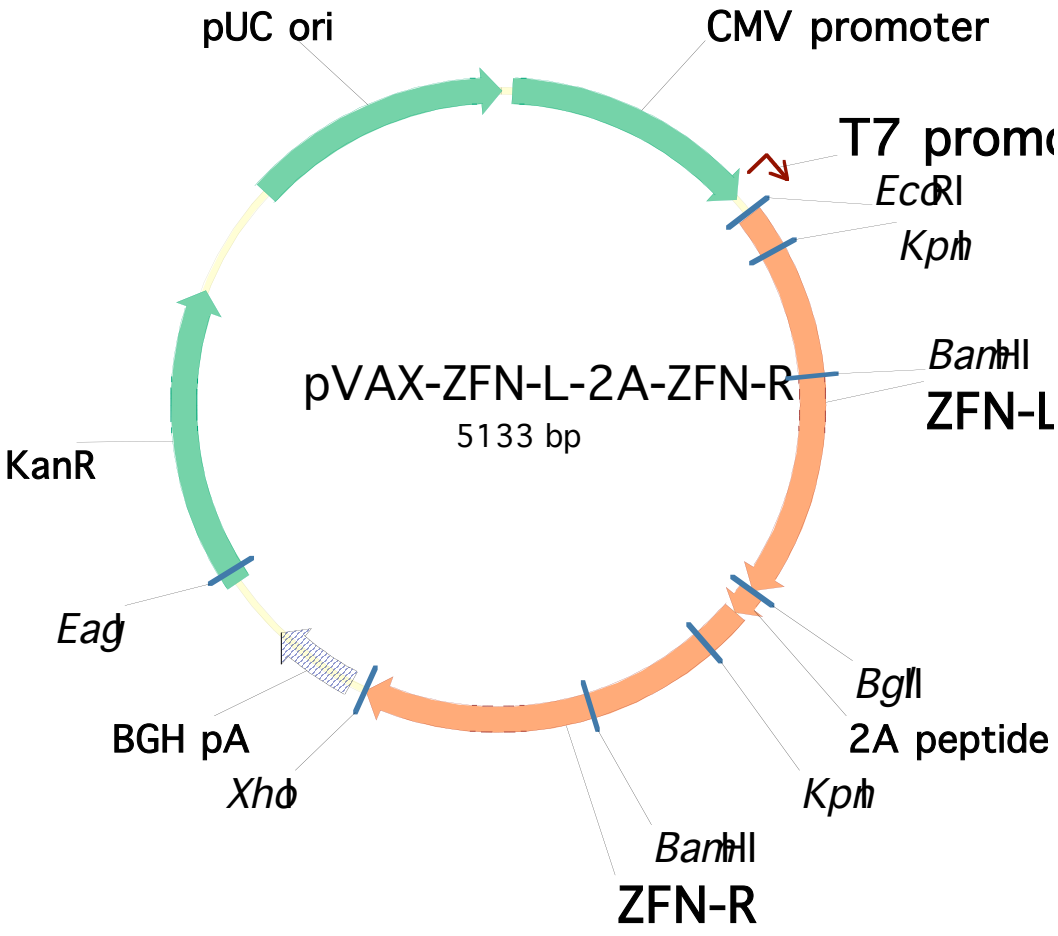
```
4501       TTGTATAGTT TTTTTATATT GTAGTTGTTC TATTTTAATC AAATGTTAGC GTGATTTATA TTTTTTTTCG CCTCGACATC ATCTGCCCAG
           ATGCGAAGTT
           AACATATCAA AAAAATATAA CATCAACAAG ATAAAATTAG TTTACAATCG CACTAAATAT AAAAAAAAGC GGAGCTGTAG TAGACGGGTC
           TACGCTTCAA
4601       AAGTGCGCAG AAAGTAATAT CATGCGTCAA TCGTATGTGA ATGCTGGTCG CTATACTGCT GTCGATTCGA TACTAACGCC GCCATCCAGT
           GTCGAAAACG
           TTCACGCGTC TTTCATTATA GTACGCAGTT AGCATACACT TACGACCAGC GATATGACGA CAGCTAAGCT ATGATTGCGG CGGTAGGTCA
           CAGCTTTTGC
4701       AGCTCAGACC ACCTGCATAA TAACAGCTTT CTTTTCGGTA TGTATTCGTC TGCTGGTGAG TACACCTGTG CTGGATATCC TGGGTCTCTG
           GGTCGTGAGG
           TCGAGTCTGG TGGACGTATT ATTGTCGAAA GAAAAGCCAT ACATAAGCAG ACGACCACTC ATGTGGACAC GACCTATAGG ACCCAGAGAC
           CCAGCACTCC
4801       AAGAAGATGC ACAGTTCTTT GCAAATAACC GCGTTGACTA CTTGAAGTAC GATAATTGTT ACAATAAGGG TCAGTTTGGT ACACCGGAAA
           TTTCTTACCA
           TTCTTCTACG TGTCAAGAAA CGTTTATTGG CGCAACTGAT GAACTTCATG CTATTAACAA TGTTATTCCC AGTCAAACCA TGTGGCCTTT
           AAAGAATGGT
4901       CCGTTACAAG GCCATGTCAG ATGCTTTGAA TAAAACTGGT AGGCCTATAT TCTATTCTCT ATGTAACTGG GGTCAGGATT TAACATTTTA
           CTGGGGCTCT
           GGCAATGTTC CGGTACAGTC TACGAAACTT ATTTTGACCA TCCGGATATA AGATAAGAGA TACATTGACC CCAGTCCTAA ATTGTAAAAT
           GACCCCGAGA
5001       GGTATCGCCA ATTCTTGGAG AATGAGTGGA GATGTTACTG CTGAGTTCAC TCGTCCAGAT AGCAGATGTC CCTGTGATGG CGATGAATAC
           GATTGCAAGT
           CCATAGCGGT TAAGAACCTC TTACTCACCT CTACAATGAC GACTCAAGTG AGCAGGTCTA TCGTCTACAG GGACACTACC GCTACTTATG
           CTAACGTTCA
5101       ACGCCGGTTT CCATTGTTCT ATTATGAATA TTCTTAACAA GGCAGCTCCA ATGGGGCAAA ATGCAGGTGT TGGTGGTTGG AATGATCTGG
           ACAATCTAGA
           TGCGGCCAAA GGTAACAAGA TAATACTTAT AAGAATTGTT CCGTCGAGGT TACCCCGTTT TACGTCCACA ACCACCAACC TTACTAGACC
           TGTTAGATCT
5201       GGTTGGTGTC GGGAATTTGA CTGACGATGA GGAAAAGGCA CATTTCTCTA TGTGGGCAAT GGTAAAGTCT CCACTTATCA TTGGTGCCAA
           TGTGAATAAC
           CCAACCACAG CCCTTAAACT GACTGCTACT CCTTTTCCGT GTAAAGAGAT ACACCCGTTA CCATTTCAGA GGTGAATAGT AACCACGGTT
           ACACTTATTG
5301       TTAAAGGCAT CTTCGTACTC AATCTATAGT CAAGCCTCTG TCATCGCAAT TAATCAAGAT TCAAATGGTA TTCCAGCAAC AAGAGTCTGG
           AGATATTATG
           AATTTCCGTA GAAGCATGAG TTAGATATCA GTTCGGAGAC AGTAGCGTTA ATTAGTTCTA AGTTTACCAT AAGGTCGTTG TTCTCAGACC
           TCTATAATAC
5401       TTTCAGACAC AGATGAATAT GGACAAGGTG AAATTCAAAT GTGGAGTGGT CCTCTTGACA ATGGTGATCA AGTGGTTGCT TTATTGAATG
           GAGGAAGCGT
           AAAGTCTGTG TCTACTTATA CCTGTTCCAC TTTAAGTTTA CACCTCACCA GGAGAACTGT TACCACTAGT TCACCAACGA AATAACTTAC
           CTCCTTCGCA
5501       ATCTAGACCA ATGAACACGA CCTTGGAAGA GATTTTTTTT GACAGCAATC TGGGTTCAAA GAAACTGACA TCGACTTGGG ATATCTACGA
           CCTATGGGCC
           TAGATCTGGT TACTTGTGCT GGAACCTTCT CTAAAAAAAA CTGTCGTTAG ACCCAAGTTT CTTTGACTGT AGCTGAACCC TATAGATGCT
           GGATACCCGG
5601       AACAGAGTTG ACAACTCGAC AGCGTCTGCT ATCCTTGGAC GGAATAAGAC AGCCACCGGT ATTCTCTACA ATGCTACGGA GCAATCCTAC
           AAAGACGGTT
           TTGTCTCAAC TGTTGAGCTG TCGCAGACGA TAGGAACCTG CCTTATTCTG TCGGTGGCCA TAAGAGATGT TACGATGCCT CGTTAGGATG
           TTTCTGCCAA
5701       TGTCTAAGAA TGATACAAGA CTGTTTGGTC AGAAAATTGG TAGTCTTTCT CCAAATGCTA TACTTAACAC GACTGTTCCA GCTCACGGTA
           TCGCCTTCTA
           ACAGATTCTT ACTATGTTCT GACAAACCAG TCTTTTAACC ATCAGAAAGA GGTTTACGAT ATGAATTGTG CTGACAAGGT CGAGTGCCAT
           AGCGGAAGAT
5801       TAGGTTGAGA CCCTCTTCTT GAGCTTATTG TTGAGCAAAG CAGGGCGAGA AGTATTGATG ATTGTTAAAA AGTTCATGAA AAAAATACTA
           CTCGAATATT
           ATCCAACTCT GGGAGAAGAA CTCGAATAAC AACTCGTTTC GTCCCGCTCT TCATAACTAC TAACAATTTT TCAAGTACTT TTTTTATGAT
           GAGCTTATAA
5901       TATTCAGAGT AACTAAATAA TAAACGACAG AATAGCCTAT CAGGTATTCC AATAGTTTTC GTTTTGTAGG TACATAATCT GAAGCCCTTG
           AACTTTTTCT
           ATAAGTCTCA TTGATTTATT ATTTGCTGTC TTATCGGATA GTCCATAAGG TTATCAAAAG CAAAACATCC ATGTATTAGA CTTCGGGAAC
           TTGAAAAAGA
6001       CGTTTACATA CTTCATTGCA TTAGCGATAT TTCACATGTG CTATACTAGT GACTTTTGTA AAATACTTTT CTGCCAATGT GATTCTAGAA
           TTGTATGACA
           GCAAATGTAT GAAGTAACGT AATCGCTATA AAGTGTACAC GATATGATCA CTGAAAACAT TTTATGAAAA GACGGTTACA CTAAGATCTT
           AACATACTGT
6101       TTTCACGAAG AGGAACAACA GCTTCAGGAG TACATACAAA CGACGAAATC ATCATATCAG CAAAACCGGA TATCGAAGAA GCAGAAGGCG
           CTAGTCTACA
           AAAGTGCTTC TCCTTGTTGT CGAAGTCCTC ATGTATGTTT GCTGCTTTAG TAGTATAGTC GTTTTGGCCT ATAGCTTCTT CGTCTTCCGC
           GATCAGATGT
6201       GGTATCTCTG TAGTTGATAC AAACAAAAAC AAAGAGTTCG AACACATTTC TGGGTCCCGA GTAGTGGCGT TGGTGTTCCA CCATTGGTCG
           AACAATTGTT
           CCATAGAGAC ATCAACTATG TTTGTTTTTG TTTCTCAAGC TTGTGTAAAG ACCCAGGGCT CATCACCGCA ACCACAAGGT GGTAACCAGC
           TTGTTAACAA
6301       TTTTGCTACC GCTCGCTCAT AGCCGGTTTC AGTAAAATTG AAGTAAGCCC AGTTTCATCT GAATTTGAGA ATAGGGTCTA AAAAGGGAGT
           ATCACCATCG
           AAAACGATGG CGAGCGAGTA TCGGCCAAAG TCATTTTAAC TTCATTCGGG TCAAAGTAGA CTTAAACTCT TATCCCAGAT TTTTCCCTCA
           TAGTGGTAGC
6401       AACGCTTCCT GATAGTCCTT AGAGAACAGT CATCAGATGT CCATATATTA CTACTACCCT TCTAGACCTG ACCACGTACT GTTGGATGCA
           TTGGTACCAA
           TTGCGAAGGA CTATCAGGAA TCTCTTGTCA GTAGTCTACA GGTATATAAT GATGATGGGA AGATCTGGAC TGGTGCATGA CAACCTACGT
           AACCATGGTT
6501       TTTTTGCCCG AATTCCTGGG GGAACAACTT CACAGAATGT TTTGTCATAT TGTCGAAGTG GTCACAAAAC AAGAGAAGTT CCGCCAATTA
           TAAAAAGGGA
           AAAAACGGGC TTAAGGACCC CCTTGTTGAA GTGTCTTACA AAACAGTATA ACAGCTTCAC CAGTGTTTTG TTCTCTTCAA GGCGGTTAAT
           ATTTTTCCCT
6601       ACCCGTATAT TTCAGCTTCA CGGATGATTT CCAGGGTGAG AGTACTGTAT ATGGGCTTAC GATAGAAGGC CATAAAAATT TCTTGCTTGG
           CAACAAAATA
           TGGGCATATA AAGTCGAAGT GCCTACTAAA GGTCCCACTC TCATGACATA TACCCGAATG CTATCTTCCG GTATTTTTAA AGAACGAACC
           GTTGTTTTAT
6701       GAAGTGAAAT CATGTCGAGG CTGCTGTGTG GGAGAACAGC TTAAAATATC ACAAAAAAAG AATCTAAAAC ACTGTGTTGC TTGTCCCAGA
           AAGGGAATCA
           CTTCACTTTA GTACAGCTCC GACGACACAC CCTCTTGTCG AATTTTATAG TGTTTTTTTC TTAGATTTTG TGACACAACG AACAGGGTCT
           TTCCCTTAGT
```

```
6801    AGTATTTTTA TAAAGATTGG AGTGGTAAAA ATCGAGTATG TGCTAGATGC TATGGAAGAT ACAAATTCAG CGGTCATCAC TGTATAAATT
        GCAAGTATGT
        TCATAAAAAT ATTTCTAACC TCACCATTTT TAGCTCATAC ACGATCTACG ATACCTTCTA TGTTTAAGTC GCCAGTAGTG ACATATTTAA
        CGTTCATACA
6901    ACCAGAAGCA CGTGAAGTGA AAAAGGCAAA AGACAAAGGC GAAAAATTGG GCATTACGCC CGAAGGTTTG CCAGTTAAAG GACCAGAGTG
        TATAAAATGT
        TGGTCTTCGT GCACTTCACT TTTTCCGTTT TCTGTTTCCG CTTTTTAACC CGTAATGCGG GCTTCCAAAC GGTCAATTTC CTGGTCTCAC
        ATATTTTACA
7001    GGCGGAATCT TACAGTGGCC TATGCGGCCG CTCTAGAACT AGTGGATCGA TCCCCAATTC GCCCTATAGT GAGTCGTATT ACAATTCACT
        GGCCGTCGTT
        CCGCCTTAGA ATGTCACCGG ATACGCCGGC GAGATCTTGA TCACCTAGCT AGGGGTTAAG CGGGATATCA CTCAGCATAA TGTTAAGTGA
        CCGGCAGCAA
7101    TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA
        GAGGCCCGCA
        AATGTTGCAG CACTGACCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA ACGTCGTGTA GGGGGAAAGC GGTCGACCGC ATTATCGCTT
        CTCCGGGCGT
7201    CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGCGCCTG ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTCAC
        ACCGCATACG
        GGCTAGCGGG AAGGGTTGTC AACGCGTCGG ACTTACCGCT TACCGCGGAC TACGCCATAA AAGAGGAATG CGTAGACACG CCATAAAGTG
        TGGCGTATGC
7301    TCAAAGCAAC CATAGTACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG TGACCGCTAC ACTTGCCAGC
        GCCCTAGCGC
        AGTTTCGTTG GTATCATGCG CGGGACATCG CCGCGTAATT CGCGCCGCCC ACACCACCAA TGCGCGTCGC ACTGGCGATG TGAACGGTCG
        CGGGATCGCG
7401    CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGTGGGCCA TCGCCCTGAT
        AGACGGTTTT
        GGCGAGGAAA GCGAAAGAAG GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA GGGGCAGTTC GAGATTTAGC CCCACCCGGT AGCGGGACTA
        TCTGCCAAAA
7501    TCGCCCTTTG ACGTTGGAGT CCACGTTCTT TAATAGTGGA CTCTTGTTCC AAACTGGAAC AACACTCAAC CCTATCTCGG GCTATTCTTT
        TGATTTATAA
        AGCGGGAAAC TGCAACCTCA GGTGCAAGAA ATTATCACCT GAGAACAAGG TTTGACCTTG TTGTGAGTTG GGATAGAGCC CGATAAGAAA
        ACTAAATATT
7601    GGGATTTTGC CGATTTCGGC CTATTGGTTA AAAAATGAGC TGATTTAACA AAAATTTAAC GCGAATTTTA ACAAAATATT AACGTTTACA
        ATTTTATGGT
        CCCTAAAACG GCTAAAGCCG GATAACCAAT TTTTTACTCG ACTAAATTGT TTTTAAATTG CGCTTAAAAT TGTTTTATAA TTGCAAATGT
        TAAAATACCA
7701    GCACTCTCAG TACAATCTGC TCTGATGCCG CATAGTTAAG CCAGCCCCGA CACCCGCCAA CACCCGCTGA CGCGCCCTGA CGGGCTTGTC
        TGCTCCCGGC
        CGTGAGAGTC ATGTTAGACG AGACTACGGC GTATCAATTC GGTCGGGGCT GTGGGCGGTT GTGGGCGACT GCGCGGGACT GCCCGAACAG
        ACGAGGGCCG
7801    ATCCGCTTAC AGACAAGCTG TGACCGTCTC CGGGAGCTGC ATGTGTCAGA GGTTTTCACC GTCATCACCG AAACGCGCGA GACGAAAGGG
        CCTCGTGATA
        TAGGCGAATG TCTGTTCGAC ACTGGCAGAG GCCCTCGACG TACACAGTCT CCAAAAGTGG CAGTAGTGGC TTTGCGCGCT CTGCTTTCCC
        GGAGCACTAT
7901    CGCCTATTTT TATAGGTTAA TGTCATGATA ATAATGGTTT CTTAGACGTC AGGTGGCACT TTTCGGGGAA ATGTGCGCGG AACCCCTATT
        TGTTTATTTT
        GCGGATAAAA ATATCCAATT ACAGTACTAT TATTACCAAA GAATCTGCAG TCCACCGTGA AAAGCCCCTT TACACGCGCC TTGGGGATAA
        ACAAATAAAA
8001    TCTAAATACA TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT AATATTGAAA AAGGAAGAGT ATGAGTATTC
        AACATTTCCG
        AGATTTATGT AAGTTTATAC ATAGGCGAGT ACTCTGTTAT TGGGACTATT TACGAAGTTA TTATAACTTT TTCCTTCTCA TACTCATAAG
        TTGTAAAGGC
8101    TGTCGCCCTT ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC GCTGGTGAAA GTAAAAGATG CTGAAGATCA
        GTTGGGTGCA
        ACAGCGGGAA TAAGGGAAAA AACGCCGTAA AACGGAAGGA CAAAAACGAG TGGGTCTTTG CGACCACTTT CATTTTCTAC GACTTCTAGT
        CAACCCACGT
8201    CGAGTGGGTT ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC GAAGAACGTT TTCCAATGAT GAGCACTTTT
        AAAGTTCTGC
        GCTCACCCAA TGTAGCTTGA CCTAGAGTTG TCGCCATTCT AGGAACTCTC AAAAGCGGGG CTTCTTGCAA AAGGTTACTA CTCGTGAAAA
        TTTCAAGACG
8301    TATGTGGCGC GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC ACTATTCTCA GAATGACTTG GTTGAGTACT
        CACCAGTCAC
        ATACACCGCG CCATAATAGG GCATAACTGC GGCCCGTTCT CGTTGAGCCA GCGGCGTATG TGATAAGAGT CTTACTGAAC CAACTCATGA
        GTGGTCAGTG
8401    AGAAAAGCAT CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT GAGTGATAAC ACTGCGGCCA ACTTACTTCT
        GACAACGATC
        TCTTTTCGTA GAATGCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC GGTATTGGTA CTCACTATTG TGACGCCGGT TGAATGAAGA
        CTGTTGCTAG
8501    GGAGGACCGA AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT GATCGTTGGG AACCGGAGCT GAATGAAGCC
        ATACCAAACG
        CCTCCTGGCT TCCTCGATTG GCGAAAAAAC GTGTTGTACC CCCTAGTACA TTGAGCGGAA CTAGCAACCC TTGGCCTCGA CTTACTTCGG
        TATGGTTTGC
8601    ACGAGCGTGA CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAACTG GCGAACTACT TACTCTAGCT TCCCGGCAAC
        AATTAATAGA
        TGCTCGCACT GTGGTGCTAC GGACATCGTT ACCGTTGTTG CAACGCGTTT GATAATTGAC CGCTTGATGA ATGAGATCGA AGGGCCGTTG
        TTAATTATCT
8701    CTGGATGGAG GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG GTTTATTGCT GATAAATCTG GAGCCGGTGA
        GCGTGGGTCT
        GACCTACCTC CGCCTATTTC AACGTCCTGG TGAAGACGCG AGCCGGGAAG GCCGACCGAC CAAATAACGA CTATTTAGAC CTCGGCCACT
        CGCACCCAGA
8801    CGCGGTATCA TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC ACGACGGGGA GTCAGGCAAC TATGGATGAA
        CGAAATAGAC
        GCGCCATAGT AACGTCGTGA CCCCGGTCTA CCATTCGGGA GGGCATAGCA TCAATAGATG TGCTGCCCCT CAGTCCGTTG ATACCTACTT
        GCTTTATCTG
8901    AGATCGCTGA GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC CAAGTTTACT CATATATACT TTAGATTGAT TTAAAACTTC
        ATTTTTAATT
        TCTAGCGACT CTATCCACGG AGTGACTAAT TCGTAACCAT TGACAGTCTG GTTCAAATGA GTATATATGA AATCTAACTA AATTTTGAAG
        TAAAAATTAA
9001    TAAAAGGATC TAGGTGAAGA TCCTTTTTGA TAATCTCATG ACCAAAATCC CTTAACGTGA GTTTTCGTTC CACTGAGCGT CAGACCCCGT
        AGAAAAGATC
        ATTTTCCTAG ATCCACTTCT AGGAAAAACT ATTAGAGTAC TGGTTTTAGG GAATTGCACT CAAAAGCAAG GTGACTCGCA GTCTGGGGCA
        TCTTTTCTAG
```

```
 9101   AAAGGATCTT CTTGAGATCC TTTTTTTCTG CGCGTAATCT GCTGCTTGCA AACAAAAAAA CCACCGCTAC CAGCGGTGGT TTGTTTGCCG
        GATCAAGAGC
        TTTCCTAGAA GAACTCTAGG AAAAAAAGAC GCGCATTAGA CGACGAACGT TTGTTTTTTT GGTGGCGATG GTCGCCACCA AACAAACGGC
        CTAGTTCTCG
 9201   TACCAACTCT TTTTCCGAAG GTAACTGGCT TCAGCAGAGC GCAGATACCA AATACTGTCC TTCTAGTGTA GCCGTAGTTA GGCCACCACT
        TCAAGAACTC
        ATGGTTGAGA AAAAGGCTTC CATTGACCGA AGTCGTCTCG CGTCTATGGT TTATGACAGG AAGATCACAT CGGCATCAAT CCGGTGGTGA
        AGTTCTTGAG
 9301   TGTAGCACCG CCTACATACC TCGCTCTGCT AATCCTGTTA CCAGTGGCTG CTGCCAGTGG CGATAAGTCG TGTCTTACCG GGTTGGACTC
        AAGACGATAG
        ACATCGTGGC GGATGTATGG AGCGAGACGA TTAGGACAAT GGTCACCGAC GACGGTCACC GCTATTCAGC ACAGAATGGC CCAACCTGAG
        TTCTGCTATC
 9401   TTACCGGATA AGGCGCAGCG GTCGGGCTGA ACGGGGGGTT CGTGCACACA GCCCAGCTTG GAGCGAACGA CCTACACCGA ACTGAGATAC
        CTACAGCGTG
        AATGGCCTAT TCCGCGTCGC CAGCCCGACT TGCCCCCCAA GCACGTGTGT CGGGTCGAAC CTCGCTTGCT GGATGTGGCT TGACTCTATG
        GATGTCGCAC
 9501   AGCATTGAGA AAGCGCCACG CTTCCCGAAG GGAGAAAGGC GGACAGGTAT CCGGTAAGCG GCAGGGTCGG AACAGGAGAG CGCACGAGGG
        AGCTTCCAGG
        TCGTAACTCT TTCGCGGTGC GAAGGGCTTC CCTCTTTCCG CCTGTCCATA GGCCATTCGC CGTCCCAGCC TTGTCCTCTC GCGTGCTCCC
        TCGAAGGTCC
 9601   GGGAAACGCC TGGTATCTTT ATAGTCCTGT CGGGTTTCGC CACCTCTGAC TTGAGCGTCG ATTTTTGTGA TGCTCGTCAG GGGGGCGGAG
        CCTATGGAAA
        CCCTTTGCGG ACCATAGAAA TATCAGGACA GCCCAAAGCG GTGGAGACTG AACTCGCAGC TAAAAACACT ACGAGCAGTC CCCCCGCCTC
        GGATACCTTT
 9701   AACGCCAGCA ACGCGGCCTT TTTACGGTTC CTGGCCTTTT GCTGGCCTTT TGCTCACATG TTCTTTCCTG CGTTATCCCC TGATTCTGTG
        GATAACCGTA
        TTGCGGTCGT TGCGCCGGAA AAATGCCAAG GACCGGAAAA CGACCGGAAA ACGAGTGTAC AAGAAAGGAC GCAATAGGGG ACTAAGACAC
        CTATTGGCAT
 9801   TTACCGCCTT TGAGTGAGCT GATACCGCTC GCCGCAGCCG AACGACCGAG CGCAGCGAGT CAGTGAGCGA GGAAGCGGAA GAGCGCCCAA
        TACGCAAACC
        AATGGCGGAA ACTCACTCGA CTATGGCGAG CGGCGTCGGC TTGCTGGCTC GCGTCGCTCA GTCACTCGCT CCTTCGCCTT CTCGCGGGTT
        ATGCGTTTGG
 9901   GCCTCTCCCC GCGCGTTGGC CGATTCATTA ATGCAGCTGG CACGACAGGT TTCCCGACTG GAAAGCGGGC AGTGAGCGCA ACGCAATTAA
        TGTGAGTTAG
        CGGAGAGGGG CGCGCAACCG GCTAAGTAAT TACGTCGACC GTGCTGTCCA AAGGGCTGAC CTTTCGCCCG TCACTCGCGT TGCGTTAATT
        ACACTCAATC
10001   CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGCGGCTCGT ATGTTGTGTG GAATTGTGAG CGGATAACAA TTTCACACAG
        GAAACAGCT
        GAGTGAGTAA TCCGTGGGGT CCGAAATGTG AAATACGAAG GCGCCGAGCA TACAACACAC CTTAACACTC GCCTATTGTT AAAGTGTGTC
        CTTTGTCGA
```

Supplementary Figure 4: Vector map and full DNA/amino acid sequence of the ZFNs used for *golden* gene disruption.

Vector map (sequence shown on next page):



ZFN expression vector sequence

1) The ZFP DNA binding helices are shown in lower case and underlined.
2) The 2A peptide sequence is double-underlined.
3) The Eag I restriction site and the T7 promoter used to linearize the plasmid and drive mRNA transcription, respectively, are indicated.

golden pair 1

```
CTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
                                                  T7 promoter

GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAA

       M  D  Y  K  D  H  D  G  D  Y  K  D  H  D  I  D  Y  K
TTCGCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

 D  D  D  D  K  M  A  P  K  K  K  R  K  V  G  I  H  G  V  P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
                                                      Kpn I
```

ZFP-L

```
 A  A  M  A  E  R  P  F  Q  C  R  I  C  M  R  N  F  S  t  s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTACCTCC

 g  s  l  s  r  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I
GGCTCCCTGTCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

 C  G  R  K  F  A  r  s  d  n  l  r  e  H  T  K  I  H  T  G
TGTGGGAGGAAGTTTGCCCGCTCCGACAACCTGCGCGAGCATACCAAGATACACACGGGA

 S  Q  K  P  F  Q  C  R  I  C  M  R  N  F  S  r  s  d  a  l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGTGACGCCCTG

 s  e  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I  C  G  R
AGCGAACACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

 K  F  A  q  n  a  t  r  t  k  H  T  K  I  H  L  R
AAATTTGCCCAGAACGCCACCCGCACAAAGCATACCAAGATACACCTGCGG
```

Fok-L

```
 G  S  Q  L  V  K  S  E  L  E  E  K  K  S  E  L  R  H  K  L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

 K  Y  V  P  H  E  Y  I  E  L  I  E  I  A  R  N  S  T  Q  D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

 R  I  L  E  M  K  V  M  E  F  F  M  K  V  Y  G  Y  R  G  K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

 H  L  G  G  S  R  K  P  D  G  A  I  Y  T  V  G  S  P  I  D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT
```

```
 Y  G  V  I  V  D  T  K  A  Y  S  G  G  Y  N  L  P  I  G  Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  E  R  Y  V  E  E  N  Q  T  R  N  K  H  L  N  P
GCCGACGAGATGGAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACCTCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  I  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  R  S
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCAGATCT
                                                    Bgl II

 G  G  G  E  G  R  G  S  L  L  T  C  G  D  V  E  E  N  P  G
GGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGC

 P  R  M  D  Y  K  D  H  D  G  D  Y  K  D  H  D  I  D  Y  K
CCTAGGATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

 D  D  D  D  K  M  A  P  K  K  K  R  K  V  G  I  H  G  V  P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
                                                   Kpn I
```

ZFP-R

```
 A  A  M  A  E  R  P  F  Q  C  R  I  C  M  R  N  F  S  d  r
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTGACCGC

 s  d  l  s  r  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I
TCCGACCTGTCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

 C  G  R  K  F  A  r  s  d  d  l  t  r  H  T  K  I  H  T  G
TGTGGGAGGAAGTTTGCCCGCTCCGACGACCTGACCCGCCATACCAAGATACACACGGGC

 G  G  G  S  Q  K  P  F  Q  C  R  I  C  M  R  N  F  S  r  s
GGAGGCGGATCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGCTCC

 d  d  l  t  r  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I
GACGACCTGACCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

 C  G  R  K  F  A  q  s  g  d  l  t  r  H  T  K  I  H  L  R
TGTGGGAGGAAGTTTGCCCAGTCCGGCGACCTGACCCGCCATACCAAGATACACCTGCGG
```

Fok-R

```
 G  S  Q  L  V  K  S  E  L  E  E  K  K  S  E  L  R  H  K  L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

 K  Y  V  P  H  E  Y  I  E  L  I  E  I  A  R  N  S  T  Q  D
```

```
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

 R  I  L  E  M  K  V  M  E  F  F  M  K  V  Y  G  Y  R  G  K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

 H  L  G  G  S  R  K  P  D  G  A  I  Y  T  V  G  S  P  I  D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

 Y  G  V  I  V  D  T  K  A  Y  S  G  G  Y  N  L  P  I  G  Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  Q  R  Y  V  K  E  N  Q  T  R  N  K  H  I  N  P
GCCGACGAGATGCAGAGATACGTGAAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  K  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACAAAACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  *
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCTGATAA


CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC
 XhoI                                                BGH poly(A) signal

CAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC

ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT

ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG

CATGCTGGGGATGCGGTGGGCTCTATGGCTTCTACTGGGCGGTTTTATGGACAGCAAGCG

AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACT

GGATGGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA

CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
                                                        EagI
```

***golden* pair 14**

```
CTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
                                                 T7 promoter

GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAA

       M  D  Y  K  D  H  D  G  D  Y  K  D  H  D  I  D  Y  K
TTCGCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

 D  D  D  D  K  M  A  P  K  K  K  R  K  V  G  I  H  G  V  P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
                                                    Kpn I
```

ZFP-L

```
 A  A  M  A  E  R  P  F  Q  C  R  I  C  M  R  N  F  S  q  s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCAGTCC

 g  n  l  a  r  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I
GGCAACCTGGCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

 C  G  R  K  F  A  t  s  a  n  l  s  r  H  T  K  I  H  T  G
TGTGGGAGGAAGTTTGCCACCTCCGCCAACCTGTCCCGCCATACCAAGATACACACGGGA

 S  Q  K  P  F  Q  C  R  I  C  M  R  N  F  S  r  s  d  t  l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGTGACACCCTG

 s  e  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I  C  G  R
AGCGAACACATCCGCACCCACACGGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

 K  F  A  r  s  q  t  r  k  t  H  T  K  I  H  L  R
AAATTTGCCCGCAGCCAGACCCGCAAAACCCATACCAAGATACACCTGCGG
```

Fok-L

```
 G  S  Q  L  V  K  S  E  L  E  E  K  K  S  E  L  R  H  K  L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

 K  Y  V  P  H  E  Y  I  E  L  I  E  I  A  R  N  S  T  Q  D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

 R  I  L  E  M  K  V  M  E  F  F  M  K  V  Y  G  Y  R  G  K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

 H  L  G  G  S  R  K  P  D  G  A  I  Y  T  V  G  S  P  I  D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

 Y  G  V  I  V  D  T  K  A  Y  S  G  G  Y  N  L  P  I  G  Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  E  R  Y  V  E  E  N  Q  T  R  N  K  H  L  N  P
GCCGACGAGATGGAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACCTCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC
```

G H F K G N Y K A Q L T R L N H I T N C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

N G A V L S V E E L L I G G E M I K A G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

T L T L E E V R R K F N N G E I N F R S
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCAGATCT
Bgl II

G G G E G R G S L L T C G D V E E N P G
GGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGC

P R M D Y K D H D G D Y K D H D I D Y K
CCTAGGATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-R

A A M A E R P F Q C R I C M R N F S d r
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTGACCGC

S h l s r H I R T H T G E K P F A C D I
TCCCACCTGTCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A r s d a l a r H T K I H T G
TGTGGGAGGAAGTTTGCCCGCTCCGACGCCCTGGCCCGCCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S d r s n l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTGACCGCTCCAACCTG

s r H I R T H T G E K P F A C D I C G R
TCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A t s g s l t r H T K I H L R
AAGTTTGCCACCTCCGGCTCCCTGACCCGCCATACCAAGATACACCTGCGG

Fok-R

G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

K Y V P H E Y I E L I E I A R N S T Q D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

R I L E M K V M E F F M K V Y G Y R G K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

Y G V I V D T K A Y S G G Y N L P I G Q

```
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  Q  R  Y  V  K  E  N  Q  T  R  N  K  H  I  N  P
GCCGACGAGATGCAGAGATACGTGAAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  K  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACAAAACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  *
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCTGATAA


CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC
 Xho I                                                BGH poly(A) signal

CAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC

ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT

ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG

CATGCTGGGGATGCGGTGGGCTCTATGGCTTCTACTGGGCGGTTTTATGGACAGCAAGCG

AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACT

GGATGGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA

CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
                                                      EagI
```

***golden* pair 15**

CTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
T7 promoter

GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAA

M D Y K D H D G D Y K D H D I D Y K
TTCGCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-L
A A M A E R P F Q C R I C M R N F S q s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCAGTCC

g n l a r H I R T H T G E K P F A C D I
GGCAACCTGGCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A t s g n l t r H T K I H T G
TGTGGGAGGAAGTTTGCCACCTCCGGCAACCTGACCCGCCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S r s d t l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGTGACACCCTG

s e H I R T H T G E K P F A C D I C G R
AGCGAACACATCCGCACCCACACGGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A r s q t r k t H T K I H L R
AAATTTGCCCGCAGCCAGACCCGCAAAACCCATACCAAGATACACCTGCGG

Fok-L
G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

K Y V P H E Y I E L I E I A R N S T Q D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

R I L E M K V M E F F M K V Y G Y R G K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

Y G V I V D T K A Y S G G Y N L P I G Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

A D E M E R Y V E E N Q T R N K H L N P
GCCGACGAGATGGAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACCTCAACCCC

N E W W K V Y P S S V T E F K F L F V S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

G H F K G N Y K A Q L T R L N H I T N C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

N G A V L S V E E L L I G G E M I K A G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

T L T L E E V R R K F N N G E I N F R S
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCAGATCT
Bgl II

G G G E G R G S L L T C G D V E E N P G
GGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGC

P R M D Y K D H D G D Y K D H D I D Y K
CCTAGGATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-R

A A M A E R P F Q C R I C M R N F S d r
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTGACCGC

s h l s r H I R T H T G E K P F A C D I
TCCCACCTGTCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A r s d a l a r H T K I H T G
TGTGGGAGGAAGTTTGCCCGCTCCGACGCCCTGGCCCGCCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S d r s n l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTGACCGCTCCAACCTG

s r H I R T H T G E K P F A C D I C G R
TCCCGCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A t s g s l t r H T K I H L R
AAGTTTGCCACCTCCGGCTCCCTGACCCGCCATACCAAGATACACCTGCGG

Fok-R

G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

K Y V P H E Y I E L I E I A R N S T Q D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

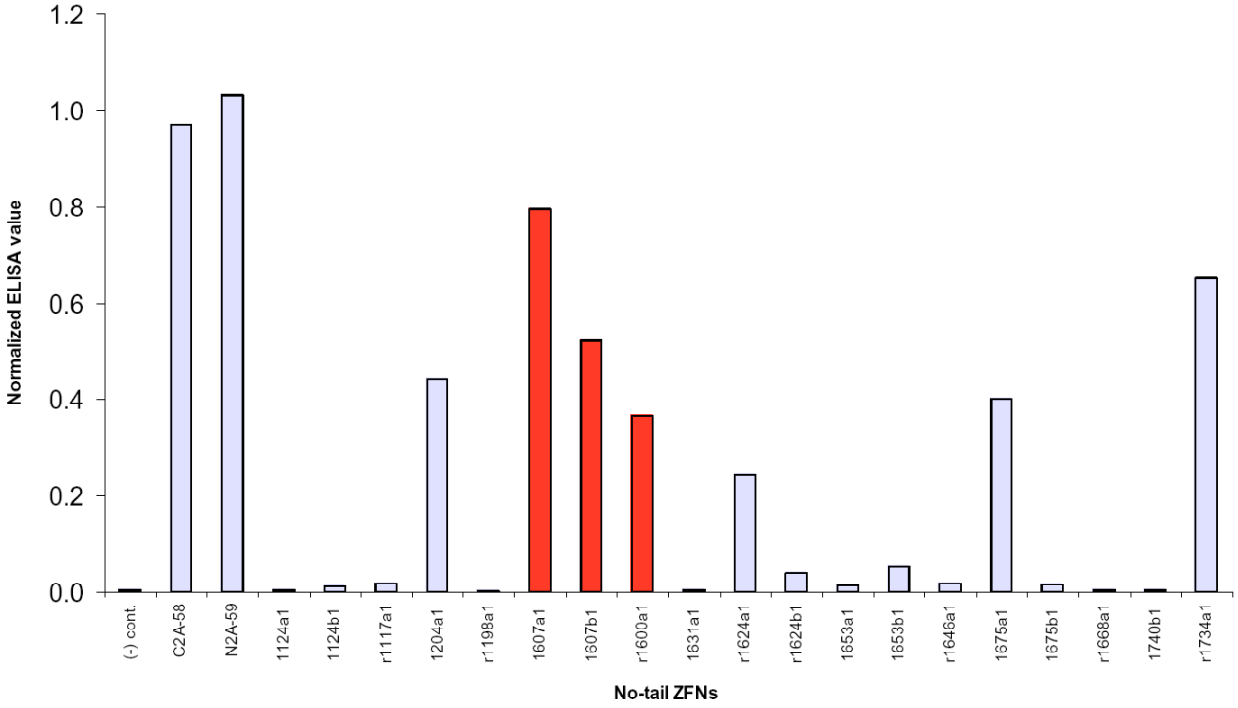
R I L E M K V M E F F M K V Y G Y R G K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

Y G V I V D T K A Y S G G Y N L P I G Q

```
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  Q  R  Y  V  K  E  N  Q  T  R  N  K  H  I  N  P
GCCGACGAGATGCAGAGATACGTGAAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  K  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACAAAACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  *
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCTGATAA


CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC
 Xho I                                                BGH poly(A) signal

CAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC

ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT

ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG

CATGCTGGGGATGCGGTGGGCTCTATGGCTTCTACTGGGCGGTTTTATGGACAGCAAGCG

AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACT

GGATGGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA

CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
                                                      Eag I
```

Supplementary Figure 5: ELISA data for DNA binding by all the ZFPs designed and assembled against the *ntl* locus. Each ZFN was evaluated for DNA binding using an ELISA assay performed as described (Bartsevich et al, 2003). The first sample is a negative control; the second and third sample, respectively, are positive controls and represent ZFN-R and ZFN-L from previous work (Urnov et al., 2005). Each ZFN is annotated by a specific base-pair position within the *ntl* locus (the cDNA cannot be used as a reference in this context because some ZFNs span exon/intron boundaries). In this nomenclature, position #1 corresponds to the transcription start site of the *ntl* gene as per the UCSC genome browser DanRer5 zebrafish genome annotation (July 2005); an "r" before the ZFN name indicates that its primary recognition is the Crick strand. The ZFNs highlighted in red form pairs 2 and 3, respectively, used in the primary text. Pair 2 ZFN-R is 1607a1, Pair 3 ZFN-L is 1607b1, and Pair 2/3 ZFN-L is r1600a1.



Supplementary Figure 6: The full DNA/amino acid sequence of the ZFNs used for *ntl* gene disruption (the vector backbone and overall arrangement is identical to that shown in Supplementary Fig. 4).

ZFN expression vector sequence

1) The ZFP DNA binding helices are shown in lower case and underlined.
2) The 2A peptide sequence is double-underlined.
3) The Eag I restriction site and the T7 promoter used to linearize the plasmid and drive mRNA transcription, respectively, are indicated.

ntl pair 2

```
CTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
                                                  T7 promoter

GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAA

       M  D  Y  K  D  H  D  G  D  Y  K  D  H  D  I  D  Y  K
TTCGCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

 D  D  D  D  K  M  A  P  K  K  K  R  K  V  G  I  H  G  V  P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
                                                    Kpn I
```

ZFP-L

```
 A  A  M  A  E  R  P  F  Q  C  R  I  C  M  R  N  F  S  r  s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGT

 d  n  l  s  r  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I
GACAACCTGAGCCGGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

 C  G  R  K  F  A  d  s  s  t  r  k  k  H  T  K  I  H  T  G
TGTGGGAGGAAATTTGCCGACAGCAGCACCCGCAAAAAGCATACCAAGATACACACGGGA

 S  Q  K  P  F  Q  C  R  I  C  M  R  N  F  S  r  s  d  h  l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGCTCCGACCACCTG

 s  a  H  I  R  T  H  T  G  E  K  P  F  A  C  D  I  C  G  R
TCCGCCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

 K  F  A  h  s  n  a  r  k  t  H  T  K  I  H  L  R
AAGTTTGCCCACTCCAACGCCCGCAAGACCCATACCAAGATACACCTGCGG
```

Fok-L

```
 G  S  Q  L  V  K  S  E  L  E  E  K  K  S  E  L  R  H  K  L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

 K  Y  V  P  H  E  Y  I  E  L  I  E  I  A  R  N  S  T  Q  D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

 R  I  L  E  M  K  V  M  E  F  F  M  K  V  Y  G  Y  R  G  K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG
```

H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

Y G V I V D T K A Y S G G Y N L P I G Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

A D E M Q R Y V E E N Q T R N K H I N P
GCCGACGAGATGCAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

N E W W K V Y P S S V T E F K F L F V S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

G H F K G N Y K A Q L T R L N H I T N C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

N G A V L S V E E L L I G G E M I K A G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

T L T L E E V R R K F N N G E I N F R S
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCAGATCT
Bgl II

G G G E G R G S L L T C G D V E E N P G
GGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGC

P R M D Y K D H D G D Y K D H D I D Y K
CCTAGGATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-R

A A M A E R P F Q C R I C M R N F S r s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGT

d t l s q H I R T H T G E K P F A C D I
GACACCCTGAGCCAGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A d r s a r t r H T K I H T G
TGTGGAAGGAAATTTGCCGACAGGAGCGCCCGCACACGGCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S r s d d l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGTGACGACCTG

s k H I R T H T G E K P F A C D I C G R
AGCAAGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A d n s n r i k H T K I H L R
AAATTTGCCGACAACAGCAACCGCATAAAGCATACCAAGATACACCTGCGG

Fok-R

G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG

BamH I

```
 K  Y  V  P  H  E  Y  I  E  L  I  E  I  A  R  N  S  T  Q  D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

 R  I  L  E  M  K  V  M  E  F  F  M  K  V  Y  G  Y  R  G  K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

 H  L  G  G  S  R  K  P  D  G  A  I  Y  T  V  G  S  P  I  D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

 Y  G  V  I  V  D  T  K  A  Y  S  G  G  Y  N  L  P  I  G  Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  Q  R  Y  V  E  E  N  Q  T  R  N  K  H  I  N  P
GCCGACGAGATGCAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  I  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  *
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCTGATAA
```

CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC
Xho I *BGH poly(A) signal*

CAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC

ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT

ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG

CATGCTGGGGATGCGGTGGGCTCTATGGCTTCTACTGGGCGGTTTTATGGACAGCAAGCG

AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACT

GGATGGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA

CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
EagI

***ntl* pair 3**

CTGGCTAACTAGAGAACCCACTGCTTACTGGCTTATCGAAATTAATACGACTCACTATAG
T7 promoter

GGAGACCCAAGCTGGCTAGCGTTTAAACTTAAGCTGATCCACTAGTCCAGTGTGGTGGAA

M D Y K D H D G D Y K D H D I D Y K
TTCGCCATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-L
A A M A E R P F Q C R I C M R N F S r s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGT

d n l s r H I R T H T G E K P F A C D I
GACAACCTGAGCCGGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A d s s t r k k H T K I H T G
TGTGGGAGGAAATTTGCCGACAGCAGCACCCGCAAAAAGCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S r s d h l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGCTCCGACCACCTG

s a H I R T H T G E K P F A C D I C G R
TCCGCCCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A h s n a r k t H T K I H L R
AAGTTTGCCCACTCCAACGCCCGCAAGACCCATACCAAGATACACCTGCGG

Fok-L
G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

K Y V P H E Y I E L I E I A R N S T Q D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

R I L E M K V M E F F M K V Y G Y R G K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

Y G V I V D T K A Y S G G Y N L P I G Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

A D E M Q R Y V E E N Q T R N K H I N P
GCCGACGAGATGCAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

N E W W K V Y P S S V T E F K F L F V S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

G H F K G N Y K A Q L T R L N H I T N C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

N G A V L S V E E L L I G G E M I K A G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

T L T L E E V R R K F N N G E I N F R S
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCAGATCT
Bgl II

G G G E G R G S L L T C G D V E E N P G
GGCGGCGGAGAGGGCAGAGGAAGTCTTCTAACATGCGGTGACGTGGAGGAGAATCCCGGC

P R M D Y K D H D G D Y K D H D I D Y K
CCTAGGATGGACTACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAG

D D D D K M A P K K K R K V G I H G V P
GATGACGATGACAAGATGGCCCCCAAGAAGAAGAGGAAGGTGGGCATTCACGGGGTACCC
Kpn I

ZFP-R

A A M A E R P F Q C R I C M R N F S r s
GCCGCTATGGCTGAGAGGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGTAGT

d t l s q H I R T H T G E K P F A C D I
GACACCCTGAGCCAGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATT

C G R K F A d r s a r t r H T K I H T G
TGTGGAAGGAAATTTGCCGACAGGAGCGCCCGCACACGGCATACCAAGATACACACGGGA

S Q K P F Q C R I C M R N F S r s d s l
TCTCAGAAGCCCTTCCAGTGTCGAATCTGCATGCGTAACTTCAGTCGCTCCGACTCCCTG

s k H I R T H T G E K P F A C D I C G R
TCCAAGCACATCCGCACCCACACAGGCGAGAAGCCTTTTGCCTGTGACATTTGTGGGAGG

K F A d n s n r i k H T K I H L R
AAGTTTGCCGACAACTCCAACCGCATCAAGCATACCAAGATACACCTGCGG

Fok-R

G S Q L V K S E L E E K K S E L R H K L
GGATCCCAGCTGGTGAAGAGCGAGCTGGAGGAGAAGAAGTCCGAGCTGCGGCACAAGCTG
BamH I

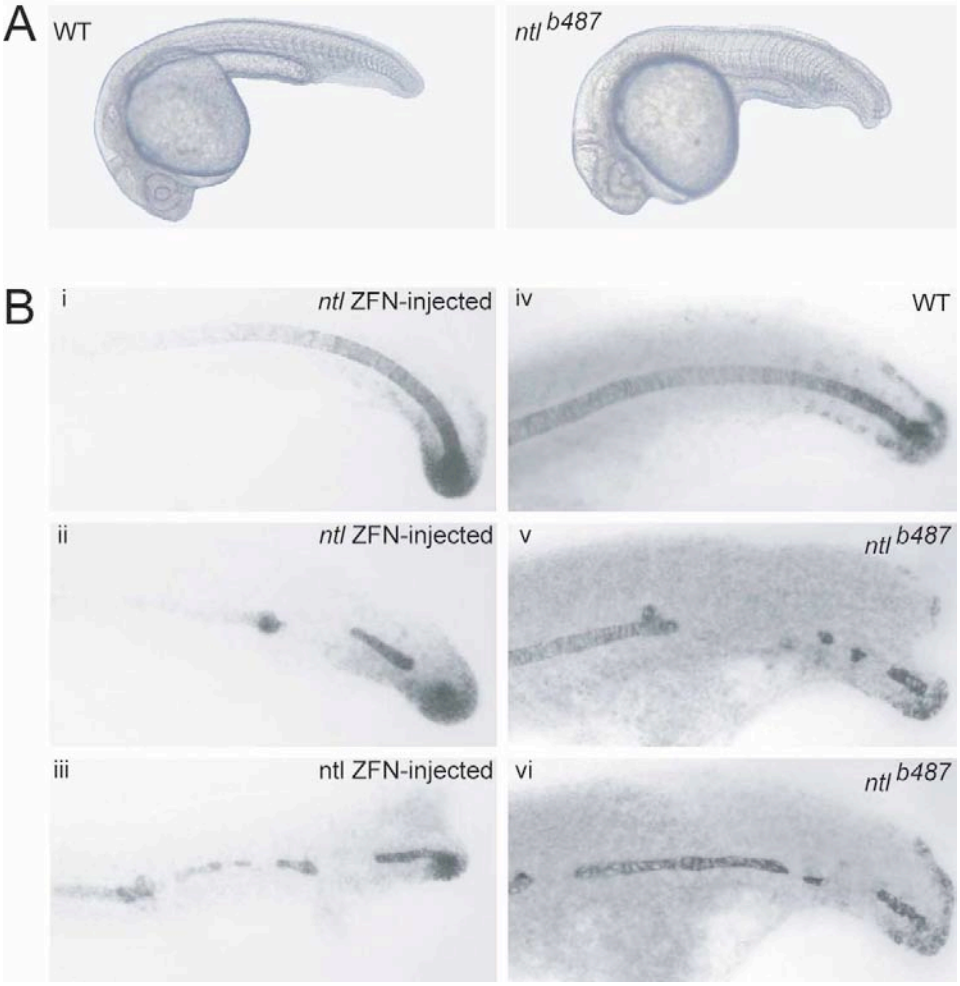
K Y V P H E Y I E L I E I A R N S T Q D
AAGTACGTGCCCCACGAGTACATCGAGCTGATCGAGATCGCCAGGAACAGCACCCAGGAC

R I L E M K V M E F F M K V Y G Y R G K
CGCATCCTGGAGATGAAGGTGATGGAGTTCTTCATGAAGGTGTACGGCTACAGGGGAAAG

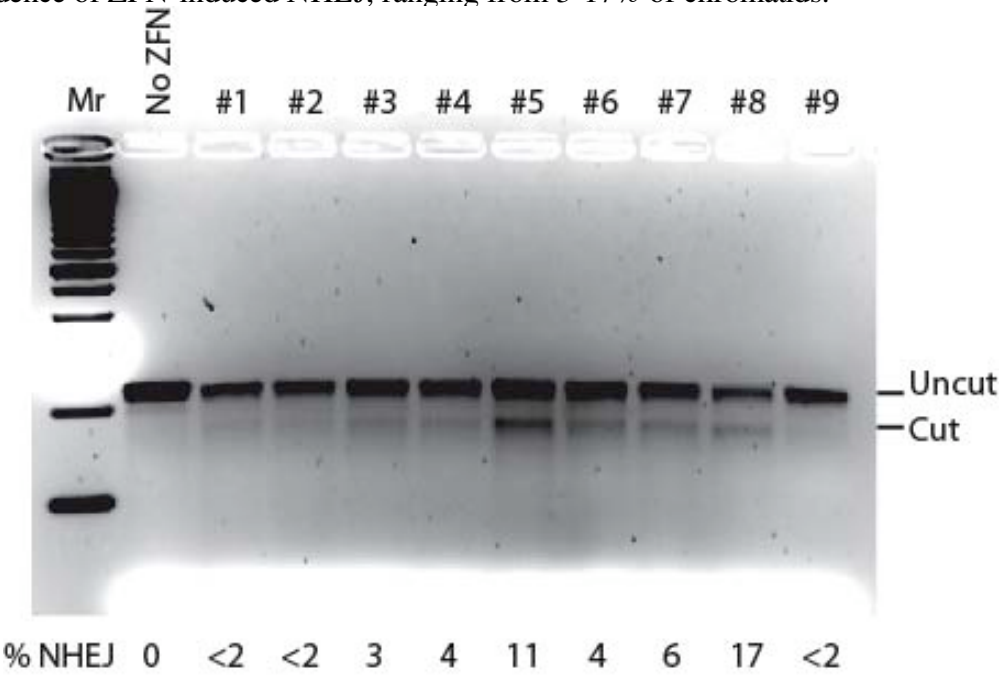
H L G G S R K P D G A I Y T V G S P I D
CACCTGGGCGGAAGCAGAAAGCCTGACGGCGCCATCTATACAGTGGGCAGCCCCATCGAT

```
 Y  G  V  I  V  D  T  K  A  Y  S  G  G  Y  N  L  P  I  G  Q
TACGGCGTGATCGTGGACACAAAGGCCTACAGCGGCGGCTACAATCTGCCTATCGGCCAG

 A  D  E  M  Q  R  Y  V  E  E  N  Q  T  R  N  K  H  I  N  P
GCCGACGAGATGCAGAGATACGTGGAGGAGAACCAGACCCGGAATAAGCACATCAACCCC

 N  E  W  W  K  V  Y  P  S  S  V  T  E  F  K  F  L  F  V  S
AACGAGTGGTGGAAGGTGTACCCTAGCAGCGTGACCGAGTTCAAGTTCCTGTTCGTGAGC

 G  H  F  K  G  N  Y  K  A  Q  L  T  R  L  N  H  I  T  N  C
GGCCACTTCAAGGGCAACTACAAGGCCCAGCTGACCAGGCTGAACCACATCACCAACTGC

 N  G  A  V  L  S  V  E  E  L  L  I  G  G  E  M  I  K  A  G
AATGGCGCCGTGCTGAGCGTGGAGGAGCTGCTGATCGGCGGCGAGATGATCAAAGCCGGC

 T  L  T  L  E  E  V  R  R  K  F  N  N  G  E  I  N  F  *
ACCCTGACACTGGAGGAGGTGCGGCGCAAGTTCAACAACGGCGAGATCAACTTCTGATAA


CTCGAGTCTAGAGGGCCCGTTTAAACCCGCTGATCAGCCTCGACTGTGCCTTCTAGTTGC
Xho I                                                BGH poly(A) signal

CAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCC

ACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCT

ATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGG

CATGCTGGGGATGCGGTGGGCTCTATGGCTTCTACTGGGCGGTTTTATGGACAGCAAGCG

AACCGGAATTGCCAGCTGGGGCGCCCTCTGGTAAGGTTGGGAAGCCCTGCAAAGTAAACT

GGATGGCTTTCTCGCCGCCAAGGATCTGATGGCGCAGGGGATCAAGCTCTGATCAAGAGA

CAGGATGAGGATCGTTTCGCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGCCG
                                                      EagI
```

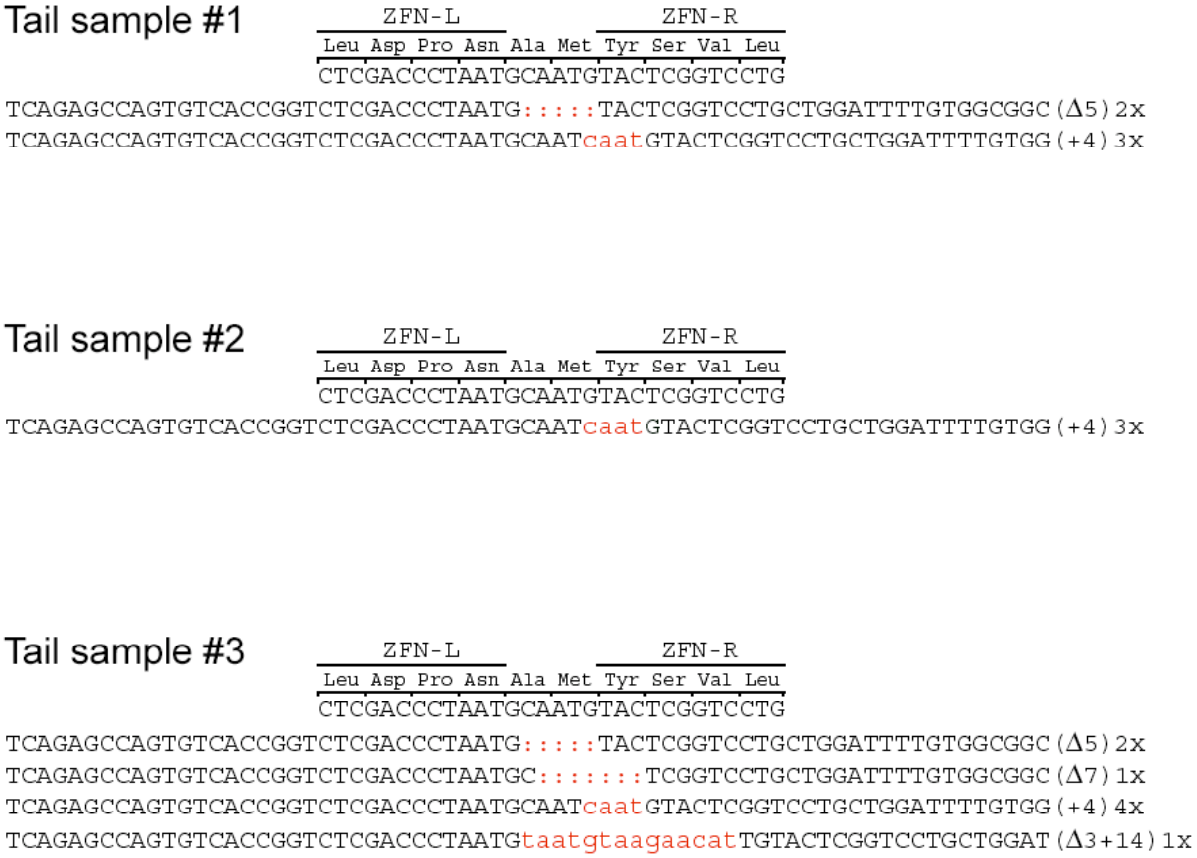
Supplementary Figure 7: ZFN injection can cause hypomorphic phenotypes.
(A) mRNA encoding high-fidelity, obligate-heterodimer (Miller et al., 2007) ZFNs was injected into 1-cell embryos from a cross between wildtype and *ntl*^{b195} heterozygous fish. ZFN-injected embryos have diminished or no notochord, misshapen somites, and reduced tail mesoderm, with phenotypes resembling both the null *ntl*^{b195} homozygous mutant (see Fig. 3C) and the hypomorphic *ntl*^{b487} homozygous mutant (top right panel). (B) When processed by in situ hybridization to detect notochordal *ntl* expression, *ntl* ZFN-injected embryos display a range of phenotypes, from normal expression (i) to disrupted and patchy expression (ii, iii), to completely absent expression (not shown). The patchy notochord phenotype is typical of that in hypomorphic *ntl*^{b487} homozygous mutant embryos (v, vi); normal wildtype siblings are shown in panel (iv).



Supplementary Figure 8: Injection of *ntl* ZFN-encoding mRNA into zebrafish embryos induces disruption of the *ntl* gene in somatic cells. Nine individual *ntl*-like embryos (#1-#9) were assayed to assess the percentage of *ntl* loci that had undergone NHEJ using an assay based on the mismatch-sensitive endonuclease ("Surveyor"; Miller et al., 2007), and most showed clear evidence of ZFN-induced NHEJ, ranging from 3-17% of chromatids.



Supplementary Figure 9: Injection of *ntl* ZFN-encoding mRNA into wildtype embryos induces tail phenotypes in juvenile and adult fish. DNA was prepared from small posterior tissue samples obtained from tailless fish like those shown in Fig. 4A. The *ntl* locus surrounding the ZFN-cleavage site was amplified and sequenced from each of three tailless embryos. In every case, *ntl* mutant-bearing amplicons represent a substantial fraction of the total (Sample 1, 5/25 (20%) *ntl*-bearing chromatids, 2 different alleles; Sample 2, 3/30 (10%) *ntl*-bearing chromatids, 1 allele; Sample 3, 8/29 (28%) *ntl*-bearing chromatids, 4 different alleles). The frequency of each allele is indicated after the allele description.



Supplementary Figure 10: Injection of *ntl* ZFN-encoding mRNA into wild-type embryos induces biallelic disruption of the *ntl* gene in somatic cells. (A) Wild-type embryos injected with mRNA encoding *ntl*-targeting ZFN Pair 2 (carrying wild-type, rather than obligate-heterodimer FokI domains) show *ntl*-like phenotypes (middle panel), and some show additional mild necrosis (right panel); uninjected embryos are shown in the left panel. In situ hybridization of representative embryos to detect notochordal *ntl* expression are inset. (B) Allelic diversity for one representative *ntl*-targeting ZFN mRNA-injected embryo from (A). In all 3 embryos, a large number of unique *ntl* alleles were observed (>20) and 64-81% of the sequenced chromatids carried an induced mutation (Embryo 1, 38/59 *ntl*-bearing chromatids, Embryo 2, 44/63 *ntl*-bearing chromatids and Embryo 3, 62/77 *ntl*-bearing chromatids).

B

ZFN-L ZFN-R

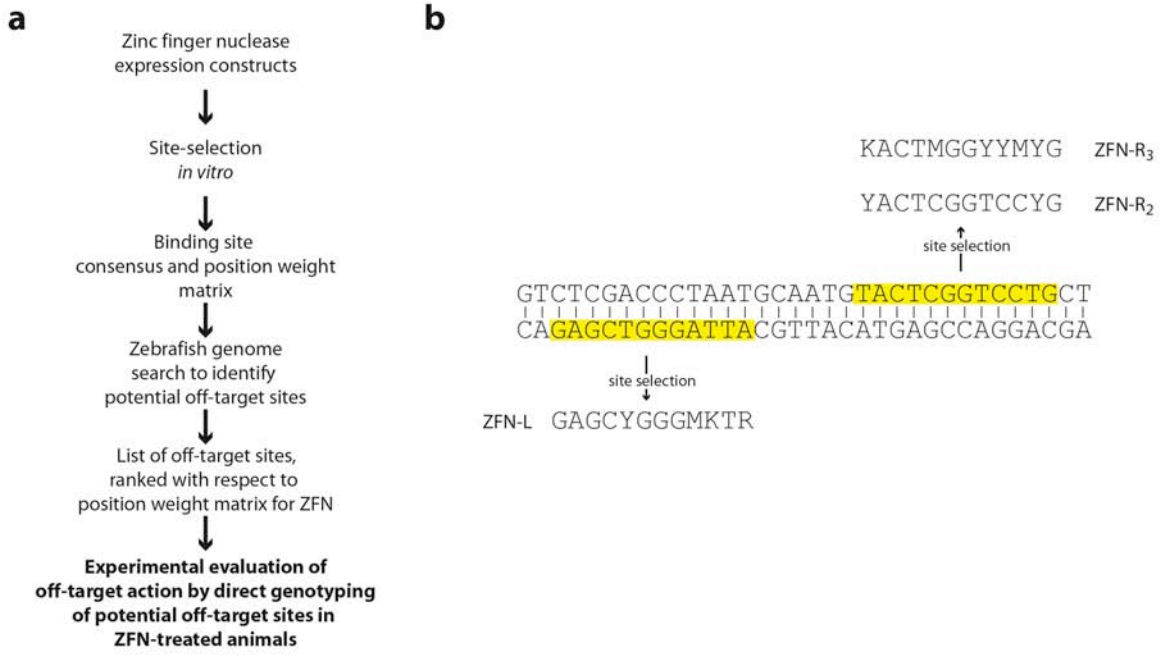
Leu Asp Pro Asn Ala Met Tyr Ser Val Leu

CTCGACCCTAATGCAATGTACTCGGTCCTG

TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGtag:GTACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ4,+3)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATaccc::TACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ6,+4)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGgattt::CTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ7,+5)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGa:::GTACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ4,+1)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGcg:::TACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ5,+2)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATG::::GTACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ4)
TCAGAGCCAGTGTCACCGGTCTCGACCC:::::CAATGTACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ5)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATG:::::TACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ5)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGtat:::::TCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ8,+3)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAA:::::::TACTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ7)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATG:::::::CTCGGTCCTGCTGGATTTTGTGGCGGCCGA(Δ7)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAAT:::::::::::::::::GCTGGATTTTGTGGCGGCCGA(Δ17)

TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGCAATcaatGTACTCGGTCCTGCTGGATTTTGTGGCGG(+4)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGgggtcctggTACTCGGTCCTGCTGGATTTTGTGGCGG(Δ5+9)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGtcgaccctagTACTCGGTCCTGCTGGATTTTGTGGCG(Δ5+10)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGCgactcaatgTACTCGGTCCTGCTGGATTTTGTGGCG(Δ4+9)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGagtgtgaatgTACTCGGTCCTGCTGGATTTTGTGGCG(Δ5+10)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGCtagaccctagTACTCGGTCCTGCTGGATTTTGTGGC(Δ4+10)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGCAggacctaatgTACTCGGTCCTGCTGGATTTTGTGG(Δ3+10)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGtagtacacaatgTACTCGGTCCTGCTGGATTTTGTGG(Δ5+12)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGtactcggtacagagTACTCGGTCCTGCTGGATTTTGT(Δ5+14)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGCAtgtactcggcatgTACTCGGTCCTGCTGGATTTTG(Δ3+13)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGgtccctgctcgacccTACTCGGTCCTGCTGGATTTTG(Δ5+15)
TCAGAGCCAGTGTCACCGGTCTCGACCCTAATGgtcctgctggtcctggtccACTCGGTCCTGCTGGATT(Δ6+19)

Supplementary Figure 11: Analysis of ZFN action at potential off-target sites.
(A) Experimental scheme to evaluate ZFN action at potential off-target sites. (B) The ZFN recognition site in the *ntl* gene, and a qualitative representation of *in vitro* site selection results showing the ZFN recognition sites determined experimentally. (C) A listing of the potential off-target sites for the ZFNs; the mismatches relative to the ZFN consensus sites are shown in lowercase.



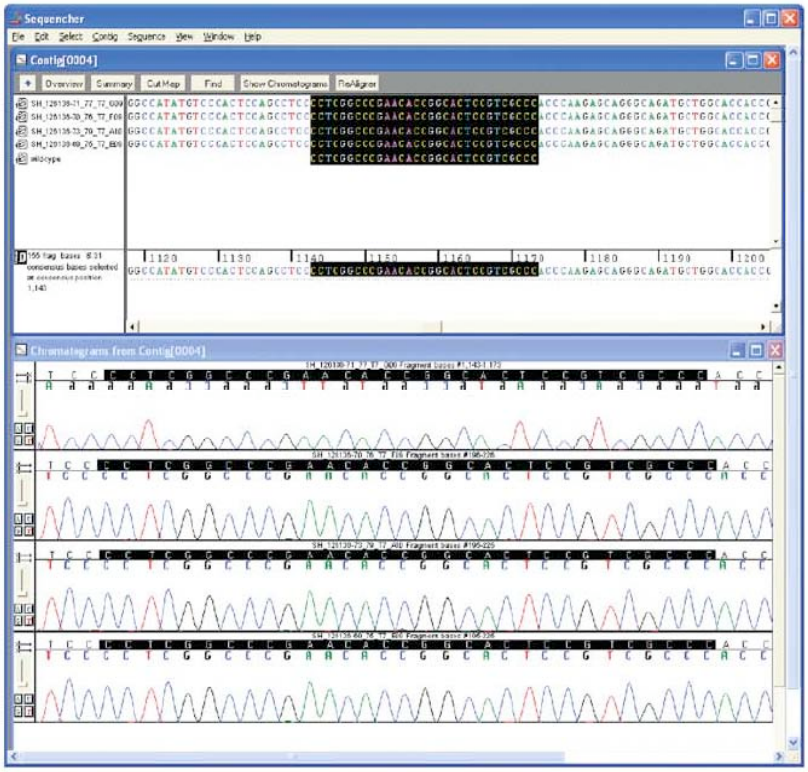
c

#	Pair	Loc	Pos	Sequence	Gene	Coding?	# of *ntl* progeny analyzed	Wild-type chromatids	Mutant chromatids
		chr19	20,038,174	TCTCGACCCTAATGCAATGTACTCGGTCCTGC	ntl	Exon 2	15	0/15	15/0
1	2/3	chr8	15,071,183	cCTCGgCCCgAAcaCc_gGcACTCcGTCgccC	prkcz	Intron 12	7	26/26	0/26
2	2/3	Unmap.	125,224,504	cCTCGcCaCacAcaCAcaccACTCGGaCCgGg		Desert	7	25/25	0/25
3	2/3	chr11	1,589,395	TCTCtgCCCgcccGCcgcGctCTCGGTCCcGg	zgc:136857	Exon 6	9	34/34	0/34
4	2	chr20	2,693,549	aCTCGgCCaactcaCtggGcACTCGGTtCaGt		Intergenic	7	28/28	0/28
5	2	chr20	285,141	ctgCGgCCCgAActCctcGTACTCGGcggcGC	si:ch211-241j12.3	Exon 7 (stop)	2	7/7	0/7
6	3	chr3	26,364,497	aCTCtGCtCTAATattgaGTACTtGGTCaTGg		Intron 1	4	15/15	0/15
7	3	chr19	25,689,552	gCTCacCCCaAAcaatcTtgACTCGGTtCaGg		Intergenic	3	12/12	0/12

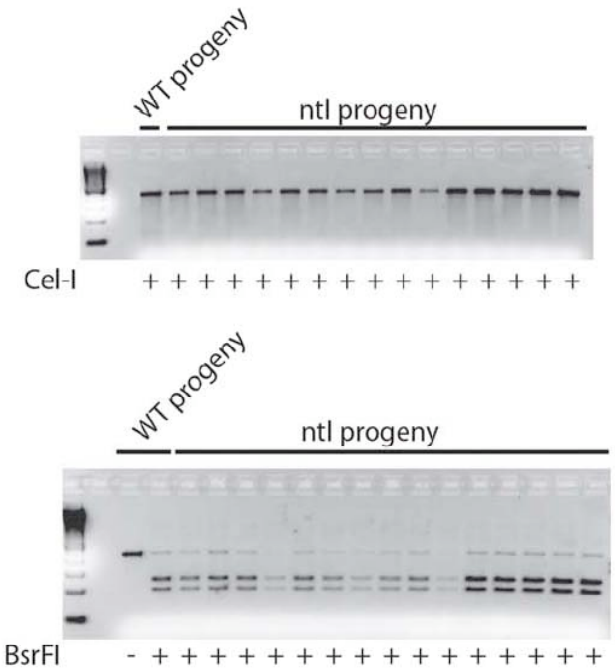
Supplementary Figure 11: Analysis of ZFN action at potential off-target sites, continued. Sequencing of the *ntl* progeny fails to identify mutant chromatids at potential off-target sites. Each of the following panels presents the sequencing and additional genotyping data for all the potential off-target sites, listed by number as in Fig. S11C. The chromatids were genotyped by amplifying the respective stretches by PCR (primers listed in Supplementary Table 3), followed by TOPO cloning and sequencing. (a) Representative chromatograms selected from *ntl* progeny of founders A-D (when applicable) are shown. In each case, the potential off-target sites are indicated with black boxes. In some cases, in addition to direct sequencing, we were able to use a mismatch sensitive endonuclease assay (unless the animal is heterozygous for a naturally occurring SNP in the region), or, if the ZFN target site overlaps with a restriction enzyme site, to do a loss-of-RFLP assay. (b) Gel-based genotyping assays using restriction enzymes or mismatch sensitive nucleases. In both cases, we used DNA extracted from wild-type progeny as a control.

Supplementary Figure 11, off-target 1

1a

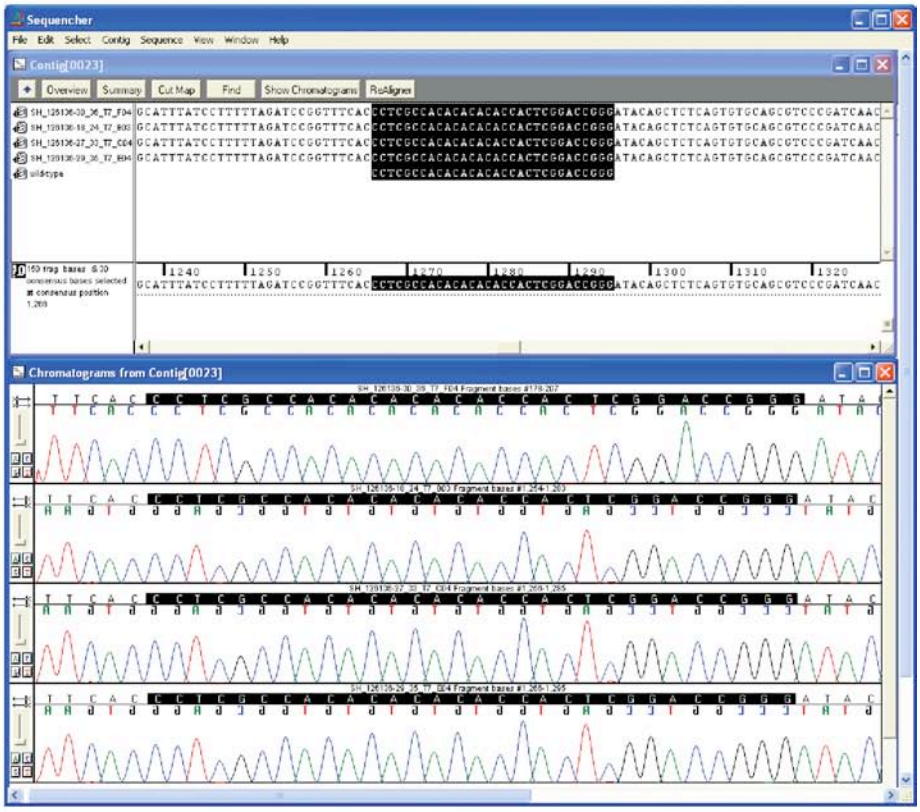


b

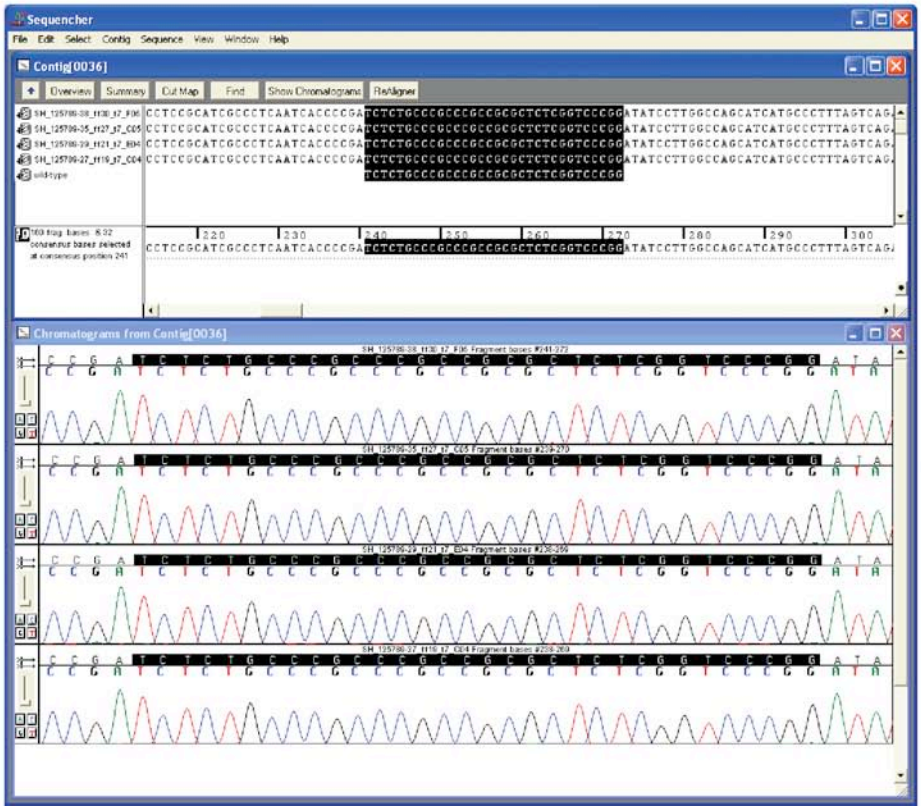


Supplementary Figure 11, off-targets 2 and 3

2a

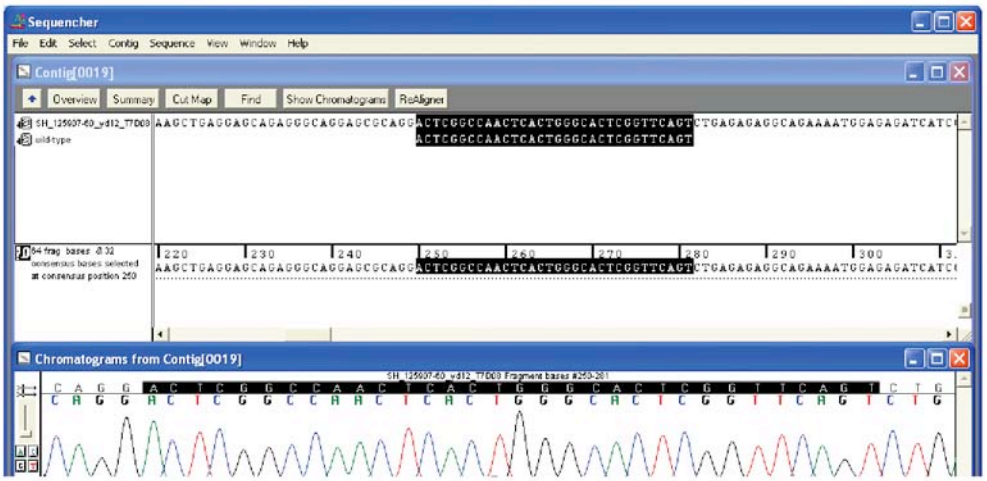


3a

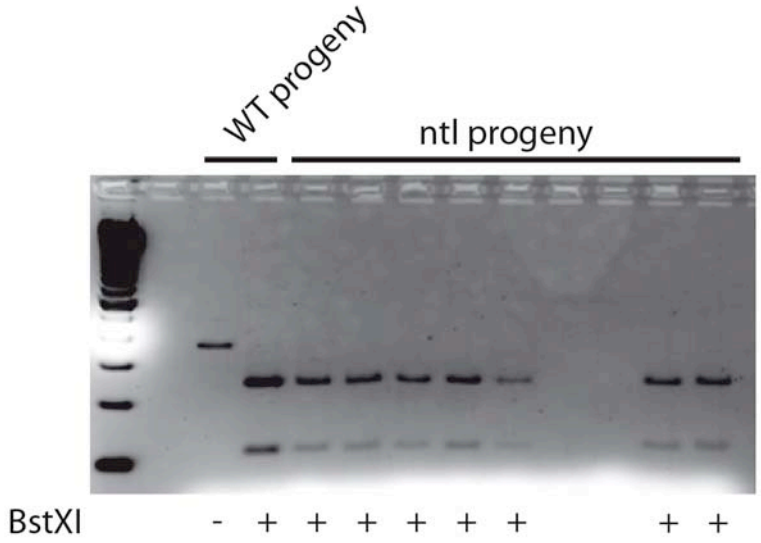


Supplementary Figure 11, off-target 4

4a

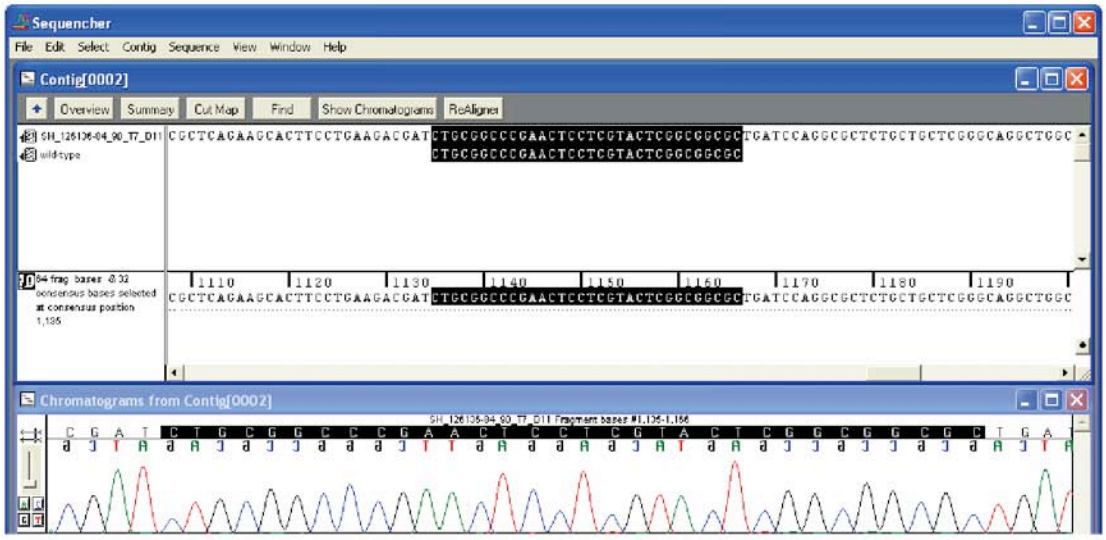


b

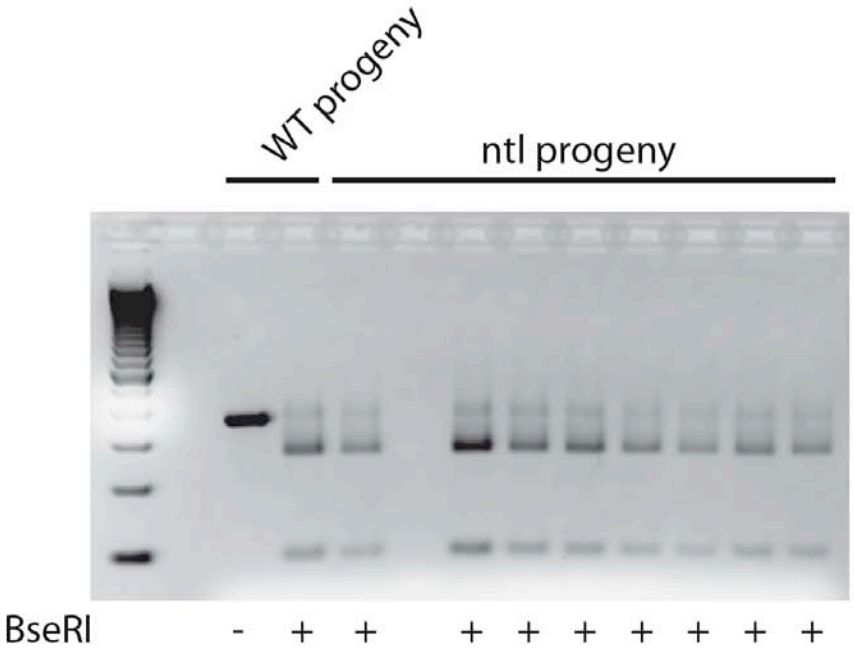


Supplementary Figure 11, off-target 5

5a

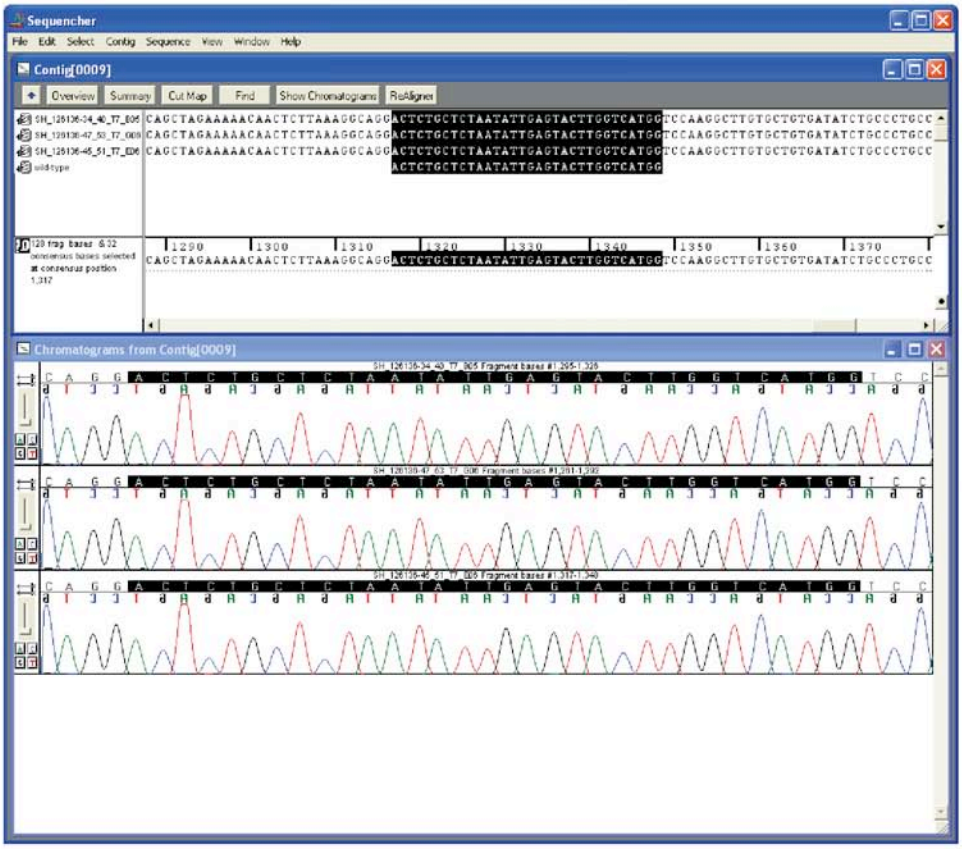


b

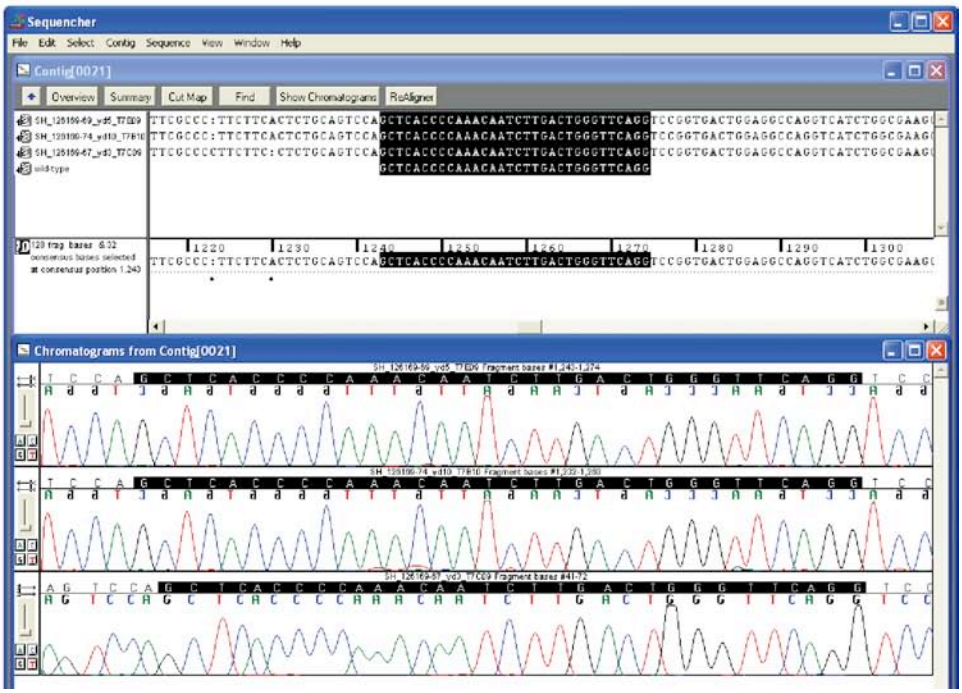


Supplementary Figure 11, off-targets 6 and 7

6a



7a



Supplementary Table 1
ZFN pairs designed against the *golden* locus (data on their testing is shown in Fig. 2A)

ZFN pair #	The codon in the *slc24a5* locus at which the DSB is induced is indicated by the cognate amino acid number in the ORF
1	Ile 166
2	Ile 166
3	Ser 355
4	Ser 355
5	Asp 381
6	Asp 381
7	Asp 381
8	Asp 381
9	Asp 397
10	Asp 397
11	Val 399
12	Ala 400
13	Ala 400
14	Val 437
15	Val 437
16	His 471
17	His 471
18	Glu 500
19	Glu 500
20	Glu 500
21	Glu 500

Supplementary Table 2
ZFN pairs designed against the *ntl* locus (data on their testing is shown in Fig. 3A)

ZFN pair #	The codon in the *ntl* locus at which the DSB is induced is indicated by the cognate amino acid number in the ORF
1	Leu 14
2	Ala 79
3	Ala 79
4	Trp 95
5	Asn 124

Supplementary Table 3
Primer sequences used to genotype potential off-target sites for ZFN action

#	Primers
1	CAGAAGTAAAGCCTGTTCAGTTTCATC TGCTTCTCTGTCTGTTCTGTGCTG
2	ACTCTCACACACACAAGATTAACGATG GCATGAAATCTCTATTCTCATCTATTCTGC
3	GTGTGTGTGGGGATGTGAAGATG ATTAGGCCGGATCTTACACCAGC
4	ACCTTGTATCCGCACACTTGATTT TTGCAACTCTAATCCAGCACCTC
5	GTGTTGTACAGGTATGACCGCTCTT GCTAAAAGGGTCCGTCAGAAGGC
6	ACATAAGACTTTTTGTGGTCC TCAGATATGCAAACCAACATGTGC
7	GGCATACTGAACTAGCGGCTACC CTTCACTCTGCAGTCCAGCTCAC

Supplementary Table 4
Plasmid designations for ZFN expression vectors

Target gene	**ZFN pair #**	**Yeast expression vectors**	**Zebrafish expression vectors**
gol	1	pSGYDLeu-12776, pSGYDHis-12775	pVAX-12776-2A-12775
gol	14	pSGYDLeu-12805, pSGYDHis-12804	pVAX-12805-2A-12804
gol	15	pSGYDLeu-12806, pSGYDHis-12804	pVAX-12806-2A-12804
ntl	2	pSGYDLeu-13370, pSGYDHis-13368	pVAX-13370-2A-13368
ntl	3	pSGYDLeu-13370, pSGYDHis-13369	pVAX-13370-2A-13369