Supporting Information

Kleinstiver et al. 10.1073/pnas.1117984109

SI Materials and Methods

Bacterial Strains and Plasmid Construction. *Escherichia coli* strains DH5α and ER2566 (New England Biolabs) were used for plasmid manipulations and protein expression, respectively. *E. coli* strain BW25141(λDE3) was used for genetic selection assays (1). A complete description of all plasmids used in this study are listed in Table S3, and oligonucloetides are listed in Table S4. The ryA and ryB zinc-finger genes were synthesized by Integrated DNA Technologies with 5′-BamHI and 3′-XhoI sites and a C-terminal 6-histidine tag, and cloned into pACYCDuet-1 to generate pACYCryAZf+H and pACYCryBZF+H, respectively. A stop codon was introduced at the 3′ end of the *ryAZf* gene using Quikchange (Stratagene) to generate pACYCryAZf. To create GIY-zinc finger endonucleases (GIY-ZFEs), the I-TevI and I-BmoI GIY-YIG domains were PCR-amplified from bacteriophage T4 gDNA and pACYCIBmoI, respectively, and cloned into the NcoI/BamHI sites of pACYCryAZf+H, pACYCryAZf, and pACYCryBZf+H. TevN201-ryA and TevN201R27A were subcloned into the XbaI and EcoRV sites of pTAL3 to generate the expression plasmids for the yeast reporter assay (pYTZN201 and pYTZN201R27A). To generate I-TevI LADGLIDADG homing endonucleases (Tev-LHEs), the I-OnuI E1 gene was amplified with BamHI and SalI ends to clone into the BamHI and XhoI sites of pACYCDuet-1 (PciI) to create pACYCOnuE1(+H). This vector was subsequently Quikchanged to introduce an E22Q mutation in I-OnuI E1 to create pACYCOnuE1E22Q (+H). I-TevI catalytic domains were amplified as above and cloned into PciI/BamHI of pACYCOnuE1E22Q (+H). The R27A mutants of I-TevI-zinc finger endonucleases (Tev-ZFEs) and Tev-LHEs were generated using Quikchange mutagenesis. Hybrid GIY-ZFE and Tev-LHE target sites (Figs. 1*B* and 4*B*, and Tables S1 and S2) were cloned into the toxic plasmid p11-lacY-wtx1 to generate reporter plasmids for the bacterial selection. Tev-ryA and Bmo-ryA target sites were cloned into pSP72 for in vitro cleavage assays. The Tev-ryA site hybrid homing site was also cloned into LITMUS28i using BamHI and XhoI to generate pTZHS1.35. The two-site Tev-ZF plasmids were created by subcloning the PvuII/HpaI fragment from pTZHS1.35 into the SwaI site of pTZHS1.35 to generate pTZHS2.35 and pTZHS3.35 (with the second TZHS in either orientation). The G5A or C1A/G5A mutations were introduced into pToxTZ and pTZHS plasmids by Quikchange mutagenesis. To generate the target plasmids for the yeast reporter assay, the TZ-ryA target sites from toxic plasmids containing TZ1.33, TZ1.33G5A, and TZ1.33C1A/G5A were amplified and cloned into the BglII and SpeI sites of pCP5.1. All constructs were verified by sequencing, and the amino acid sequences of all GIY-ZFEs and Tev-LHEs constructed are provided in Fig. S5.

Two-Plasmid Genetic Selection. The two plasmid genetic selection was performed as previously described (1), with toxic (reporter) plasmids containing hybrid TZ-ryA, TZ-ryB, BZ-ryA, TO target sites (Tables S1 and S2), mutant target sites (with G5A or C1A/G5A substitutions), or plasmids lacking a target site (p11-lacY-wtx1). Survival percentage was calculated by dividing the number of colonies observed on selective plates by those observed on nonselective plates.

Yeast Reporter Assay. Transformants of *Saccharomyces cerevisiae* YPH500(α) with Tev-ZFE constructs and YPH499(a) with target constructs were grown overnight (~230 rpm) at 30 °C in synthetic complete medium lacking histidine (Tev-ZFEs) or lacking tryptophan and uracil (targets). Tev-ZFEs and targets were mated by adding equal densities (~400 μL) of overnight culture to 1 mL YPD and left stationary for 5–6 h at 30 °C. Cells harvested by centrifugation were washed in 1 mL and resuspended in 4 mL of synthetic medium lacking histidine and tryptophan before shaking overnight at 30 °C. Cells were harvested by centrifugation, washed in 1 mL Z buffer (60 mM Na_2PO_4, 40 mM NaH_2PO_4, 10 mM KCl, 1 mM $MgSO_4$, pH 7.0), and suspended in 250 μL Z buffer. The suspension was diluted 20-fold into 1 mL Z buffer containing 0.27% β-mercaptoethanol, and 75 μL $CHCl_3$ and 45 μL 0.1% SDS were added before vortexing. Lysates were preincubated at 30 °C before the addition of 100 μL 4 mg/mL ortho nitrophenol (ONPG). Reactions proceeded until a yellow color developed whereby progress was stopped by the addition of 300 μL 1 M Na_2CO_3. Stopped reactions were pelleted and the absorbance of the supernatant was analyzed at 420 nm and 550 nm.

Protein Purification. Cultures overexpressing either TevN201-ZFE or BmoN221-ZFE were grown at 37 °C to an OD_{600}~0.5 and expression induced by 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG; Bio Basic) overnight at 15 °C. Cells were harvested by centrifugation at 8,983 × *g* for 12 min, resuspended in binding buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM imidazole, 5% (vol/vol) glycerol, and 1 mM DDT), and lysed by homogenization at 15,000 psi. The cell lysate was clarified by centrifugation at 20,400 × *g*, followed by sonication for 30 s, and centrifugation at 20,400 × *g* for 15 min. The clarified lysate was loaded onto a 1 mL HisTrap-HP column (GE Healthcare), washed with 15 mL binding buffer and then 10 mL wash buffer [20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 50 mM imidazole, 5% (vol/vol) glycerol, and 1 mM DDT]. Bound proteins were eluted in 1.5-mL fractions in four 5-mL step elutions with increasing concentrations of imidazole. Fractions containing GIY-ZFEs were dialyzed twice against 1 L dialysis buffer [20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 5% (vol/vol) glycerol, and 1 mM DDT] before storage at −80 °C.

Cleavage Assays. Single time-point cleavage assays to determine the $EC_{0.5max}$ of TevN201-ryA were performed in buffer containing 20 mM Tris-HCl pH 8.0, 100 mM NaCl, 10 mM $MgCl_2$, 5% (vol/vol) glycerol, 1 mM DTT and 10 nM pTZHS1.33. Reactions were incubated for 3 min at 37 °C, stopped with 5 μL stop solution (100 mM EDTA, 40% (vol/vol) glycerol, and bromophenol blue), and electrophoresed on a 1% agarose gel before staining with ethiduium bromide and analysis on an AlphaImager3400 (Alpha Innotech). The $EC_{0.5max}$ was determined by fitting the data to the equation

$$f_{([\text{endo}])} = \frac{f_{\max} * [\text{endo}]^H}{\text{EC}_{0.5\max} + [\text{endo}]^H}, \qquad \text{[S1]}$$

where $f_{([\text{endo}])}$ is the fraction of substrate cleaved at concentration of TevN201-ryA [endo], $f_{\max}$ is the maximal fraction cleavage, with 1 being the highest value, and H is the Hill constant that was set to 1. The initial reaction velocity was determined using supercoiled plasmid substrate with varying concentrations of TevN201-ryA (0.7 nM to 47 nM) and buffer as above. Aliquots were removed at various times, stopped and analyzed as above. The data for product appearance was fitted to the equation

$$\text{P} = \text{A}\left(1 - e^{-k_1 t}\right) + k_2 t, \qquad \text{[S2]}$$

where P is product (in nanomolars), A is the magnitude of the initial burst, k_1 is the rate constant (s^{-1}) of the initial burst phase,

and k_2 is the steady state rate constant (s^{-1}). The two-site plasmid cleavage assays were conducted as above, using 10 nM pTZHS2.33 or pTZHS3.33 as substrates, and ~90 nM purified TevN201-ryA. The k_{obs} rate constants were calculated from the decay of supercoiled substrate by fitting to the equation

$$[C] = [C_0]\exp(-k_1 t), \qquad \textbf{[S3]}$$

where $[C]$ (2) is the concentration (nM) of supercoiled plasmid at time t, $[C_0]$ is the initial concentration of supercoiled substrate (nM), and k_1 is the first order rate constant (in s^{-1}).

At least three independent trials were conducted for each dataset.

Cleavage Mapping. Mapping of cleavage sites was performed as previously described (3). Briefly, primers were individually end-labeled with γ -^{32}P ATP, and used in PCR reactions with pTox or pSP72 plasmids carrying TZ-ryA or BZ-ryA target sites to generate strand-specific substrates. The substrates were incubated with purified protein as above, and electrophoresed in 8% denaturing gels alongside sequencing ladders generated by cycle sequencing (US Biologicals) with the same end-labeled primers.

Bioinformatics. The distribution of the CNNNG cleavage motif was examined using a custom Perl script (Dataset S1) in a 35-bp window of ZFN sites predicted for *Dania rerio* chromosome 1 (Ensembl release 51) (4). Briefly, 40 bp flanking the downstream region of each predicted ZFN was extracted from the corresponding zebrafish chromosome 1 cDNA, and searched for position i = C and position i + 4 = G, with the occurrences of each CNNNG reported at position i (the C of the motif). The number of CNNNG motifs for unique ZFN sites was fit to a binomial distrubtion, and plotted in R.

1. Kleinstiver BP, Fernandes A, Gloor GB, Edgell DR (2010) A unified genetic, computational, and experimental framework identifies noon-conserved residues as critical for function of the homing endonuclease I-BmoI. *Nucleic Acids Res* 38:2411–2427.
2. Kleinstiver BP, Berube-Janzen W, Fernandes AM, Edgell DR (2011) Divalent metal ion differentially regulates the sequential nicking reactions of the GIY-YIG homing endonuclease I-BmoI. *PLoS ONE* 6:e23804.
3. Edgell DR, Shub DA (2001) Related homing endonucleases I-*Bmo*I and I-*Tev*I use different strategies to cleave homologous recognition sites. *Proc Natl Acad Sci USA* 98: 7898–7903.
4. Foley JE, et al. (2009) Rapid mutation of endogenous zebrafish genes using zinc finger nucleases made by Oligomerized Pool ENgineering (OPEN). *PLoS ONE* 4: e4348.

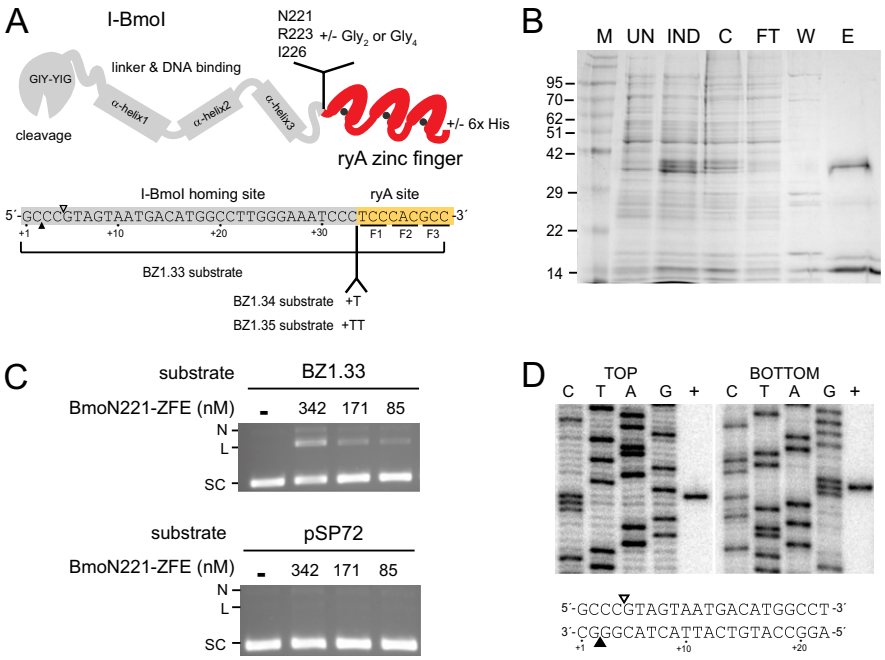

Fig. S1. Design and functionality of Bmo-ZFEs. (*A*) Schematic of Bmo-ZFE constructs, with I-BmoI protein and substrate shown in gray, and the ryA protein and binding site shown in red and yellow, respectively. (*Upper*) The fusion points for each of the Bmo-ZFEs are indicated as the last I-BmoI amino acid, with or without a 2× Glycine or 4× Glycine linker. Constructs were made with a 6xHis tag on the C-terminal end. (*Lower*) The substrate shown consists of 33-nts of the top strand I-BmoI *thyA* target site (BZ1.33), fused to the 5′ end of the ryA binding site. Substrates tested differ by the insertion of one or two T nucleotides at the junction of the *thyA*/ryA sites. (*B*) Purification of His-tagged BmoN221-ryA. Shown is a representative SDS/PAGE gel; C, crude lysate; E, elution; FT, flow-through from metal-affinity column; IND, induced culture; M, marker with molecular weights in kilodaltons indicated on the left; UN, uninduced culture; W, wash. (*C*) BmoN221-ryA cleavage specificity. Shown are representative agarose gels of cleavage assays with 10 nM pBZ1.33 or pSP72 (no target site) substrates and the indicated concentrations of BmoN221-ryA under standard assay conditions for 10 min. BmoN221-ryA cleaves the BZ-ryA target site plasmid (BZ1.33) but not the control plasmid (pSP72) lacking the target site. L, linear; N, nicked; SC, supercoiled. (*D*) Mapping of BmoN221-ryA cleavage sites on the BZ1.33 substrate, with top and bottom cleavage sites indicated by open and closed triangles, respectively.

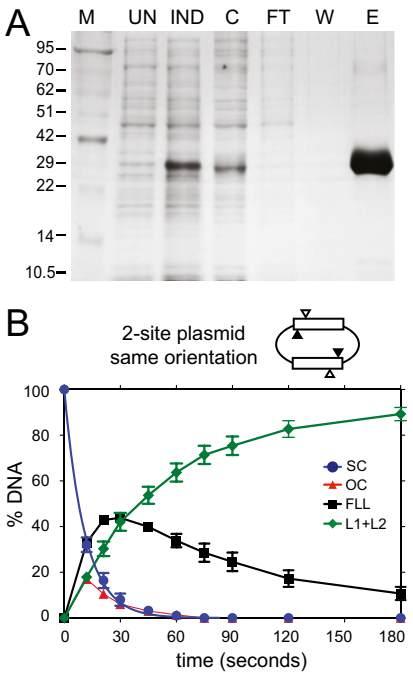

Fig. S2. TevN201-ryA purification for in vitro experiments. (*A*) Purification of TevN201-ryA. Shown is a representative SDS/PAGE gel; C, crude lysate; E, elution; FT, flow-through from metal-affinity column; IND, induced culture; M, marker with molecular weights in kilodaltons indicated on the left; UN, uninduced culture; W, wash. (*B*) Graphic representation of cleavage assays with 90 nM TevN201-ZFE and 10 nM two-site pTZ1.33 plasmid with target sites in the same orientation. Data are plotted as averages of three independent replicates with SDs; FLL, full-length linear; L1+L2, linear products; OC, open-circle (nicked); SC, supercoiled.

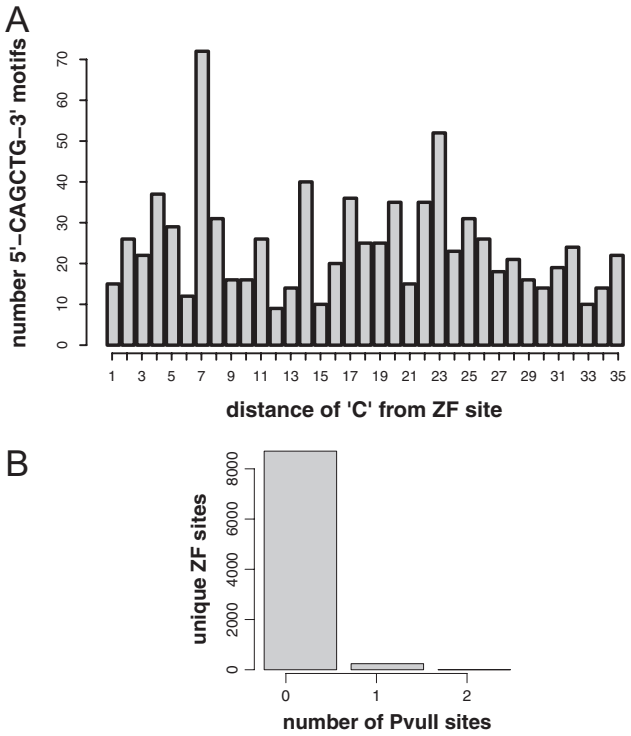

Fig. S3. PvuII site analysis. (*A*) Shown is the distribution of the 5′-CAGCTG-3′ motif in a 35-bp window flanking 8,829 predicted ZFN sites on zebrafish chromosome 1. The number of occurences of the 'C' of the motif at each distance is indicated. (*B*) Unique ZFN sites were grouped according to the number of occurences of the 5′-CAGCTG-3′ motif in the 35-bp window.

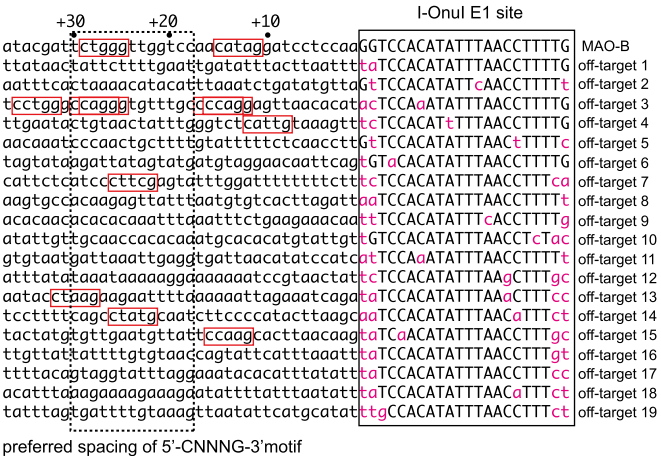

Fig. S4. Occurrence of the 5′-CNNNG-3′ motif upstream of I-OnuI E1 off-target sites. Shown is 37-nt of upstream sequence adjacent to the 22-nt I-OnuI E1 MAO-B target site, along with 19 predicted off-target sites (1). CNNNG motifs are highlighted in red, with only 3 of 19 predicted I-OnuI E1 off target sites containing a CNNNG motif at a targetable distance by Tev-LHE fusions. Nucleotide differences of the off-target sites to the I-OnuI E1 site are indicated in magenta lowercase font.

1. Takeuchi R, et al. (2011) Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proc Natl Acad Sci USA* 108:13077–13082.

>TevN201-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTGEKPF
ACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevN201G2-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTGEK
PFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevN201G4-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNGGGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTG
EKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevK203-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNIKGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTGEK
PFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevK203G2-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNIKGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTG
EKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevK203G4-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNIKGGGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTH
TGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevS206-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNIKKISGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHT
GEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevS206G2-ryA
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNIKKISGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRT
HTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>TevN201-ryA + 6xHis
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNFSRSDALTRHIRTHTGEKPF
ACDICGRKFSDPGHLVRHTKIHLRQKQLVHHHHHH

>BmoN221-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNF
SRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoN221G2-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMR
NFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoN221G4-ryA

Fig. S5. (Continued)

MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNGGGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRIC
MRNFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoR223-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNSRGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMR
NFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoR223G2-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNSRGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRIC
MRNFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoR223G4-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNSRGGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCR
ICMRNFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoI226-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNSRPVIGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRI
CMRNFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoI226G2-ryA
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNSRPVIGGGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQC
RICMRNFSRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLV

>BmoN221-ryA + 6xHis
MKSGVYKITNKNTGKFYIGSSEDCESRLKVHFRNLKNNRHINRYLNNSFNKHGEQVFIGEVIHILPIEEAIAKEQWYIDNFYEEMYNIS
KSAYHGGDLTSYHPDKRNIILKRADSLKKVYLKMTSEEKAKRWQCVQGENNPMFGRKHTETTKLKISNHNKLYYSTHKNPFKGKKHSEE
SKTKLSEYASQRVGEKNPFYGKTHSDEFKTYMSKKFKGRKPKNGSGMEPYACPVESCDRRFSQSGHLQRHIRIHTGQKPFQCRICMRNF
SRSDALTRHIRTHTGEKPFACDICGRKFSDPGHLVRHTKIHLRQKQLVHHHHHH

>TZN201-ryB + 6xHis
MGKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKING
YNIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSF
FNHKHSDITKSKISEKMKGKKPSNGSGMEPYACPVESCDRRFSQSSDLTRHIRIHTGQKPFQCRICMRNFSQSSSLVRHIRTHTGEKPF
ACDICGRKFSQSSSLVRHTKIHLRQKQLVHHHHHH

>Tev/Onu E1 E22Q S114 + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILENIQ
STWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKKAF
PEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDFVV
TKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>Tev/Onu E1 E22Q D127 + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQIC
LHIKDKSILENIQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAK
LNWGLTDELKKAFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKE
KNRSEFSWLDFVVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>Tev/Onu E1 E22Q N140 + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKR
SVGYATELGFQICLHIKDKSILENIQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLK
INGIKELVRIKAKLNWGLTDELKKAFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMN
SLITYLGCGYIKEKNRSEFSWLDFVVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRV
FHHHHHH

>Tev/Onu E1 E22Q N169 + 6xHis-Tag

Fig. S5. (Continued)

```
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNGSGSAYMSR
RESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILENIQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFE
KYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKKAFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVN
LIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDFVVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVA
KLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>Tev/Onu E1 E22Q D184 + 2xGly + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFF
NHKHSDGGGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILENIQSTWKVGVIANSGDNAVS
LRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKKAFPEIISKERSLINKNIPN
FKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDFVVTKFSDINDKIIPVFQEN
TLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>Tev/Onu E1 E22Q N201+ 4xGly
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFF
NHKHSDITKSKISEKMKGKKPSNGGGGGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILEN
IQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKK
AFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDF
VVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVF

>R27A Tev/Onu E1 E22Q N201+ 4xGly + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKAWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFF
NHKHSDITKSKISEKMKGKKPSNGGGGGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILEN
IQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKK
AFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDF
VVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>R27A Tev/Onu E1(WT) N201+ 4xGly + 6xHis-Tag
MKSGIYQIKNTLNNKVYVGSAKDFEKAWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFF
NHKHSDITKSKISEKMKGKKPSNGGGGGSGSAYMSRRESINPWILTGFADAEGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILEN
IQSTWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKK
AFPEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDF
VVTKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVFHHHHHH

>Tev/Onu E1 E22Q K203
MKSGIYQIKNTLNNKVYVGSAKDFEKRWKRHFKDLEKGCHSSIKLQRSFNKHGNVFECSILEEIPYEKDLIIERENFWIKELNSKINGY
NIADATFGDTCSTHPLKEEIIKKRSETVKAKMLKLGPDGRKALYSKPGSKNGRWNPETHKFCKCGVRIQTSAYTCSKCRNRSGENNSFF
NHKHSDITKSKISEKMKGKKPSNIKGSGSAYMSRRESINPWILTGFADAQGSFLLRIRNSNKRSVGYATELGFQICLHIKDKSILENIQ
STWKVGVIANSGDNAVSLRVTRFEDLKVIIDHFEKYPLITQKLGDYMLFKQAFCVMENKEHLKINGIKELVRIKAKLNWGLTDELKKAF
PEIISKERSLINKNIPNFKWLAGFTSGEGCFFVNLIKSNSKLGVQVQLVFSITQHIKDKNLMNSLITYLGCGYIKEKNRSEFSWLDFVV
TKFSDINDKIIPVFQENTLIGVKLEDFEDWCKVAKLIEEKKHLTESGLDEIKKIKLNMNKGRVF
```

Fig. S5. Amino acid sequences of GIY-ZFEs and Tev-LHEs.

Table S1. Tev-ZFE selection data

Table S1

Table S2. Tev-LHE selection data

Table S2

Table S3. Strains and plasmids used in this study

Table S3

1. Kleinstiver BP, Fernandes A, Gloor GB, Edgell DR (2010) A unified genetic, computational, and experimental framework identifies noon-conserved residues as critical for function of the homing endonuclease I-Bmol. *Nucleic Acids Res* 38:2411–2427.
3. Takeuchi R, et al. (2011) Tapping natural reservoirs of homing endonucleases for targeted gene modification. *Proc Natl Acad Sci USA* 108:13077–13082.
4. Chen Z, Zhao H (2005) A highly sensitive selection method for directed evolution of homing endonucleases. *Nucleic Acids Res* 33:e154.
5. Cermak T, et al. (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* 39:e82.

Table S4. Oligonucleotides used in this study

Table S4

Other Supporting Information Files

Dataset S1 (DOC)