Published in final edited form as:

Nat Biotechnol. 2009 September ; 27(9): 851–857. doi:10.1038/nbt.1562.

Highly efficient gene targeting of expressed and silent genes in human ESCs and iPSCs using zinc finger nucleases

Dirk Hockemeyer^{1,*}, Frank Soldner^{1,*}, Caroline Beard¹, Qing Gao¹, Maisam Mitalipova¹, Russell C. DeKolver³, George E. Katibah³, Ranier Amora³, Elizabeth A. Boydston³, Bryan Zeitler³, Xiangdong Meng³, Jeffrey C. Miller³, Lei Zhang³, Edward J. Rebar³, Philip D. Gregory³, Fyodor D. Urnov³, and Rudolf Jaenisch^{1,2,4}

¹The Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, 02142 MA, USA

²Department of Biology, Massachusetts Institute of Technology, 31 Ames Street, Cambridge, MA 02139, USA

³Sangamo BioSciences, Inc., Pt. Richmond Tech Center, 501 Canal Blvd., Suite A100, Richmond, CA 94804, USA

Abstract

Human embryonic stem cells and induced pluripotent stem cells (hESCs and hiPSCs) are powerful tools for biomedical research. Realizing the full potential of these cells requires efficient genetic modification. However, techniques to generate cell type specific lineage reporters as well as reliable tools to disrupt, repair or overexpress genes by gene targeting are inefficient at best and thus are not routinely used. Here we report the highly efficient targeting of three genes in human pluripotent cells using zinc finger nuclease (ZFN) mediated genome editing. First, using ZFNs specific for the *OCT4* locus we generated OCT4-eGFP reporter cells to monitor the pluripotent state of hESCs. Secondly, we inserted a transgene into the AAVS1 locus to generate a robust drug-inducible overexpression system in hESCs. Finally, we targeted the *PITX3* gene, demonstrating that ZFNs can be used to generate reporter cells by targeting non-expressed genes in hESCs and hiPSCs.

Keywords

human embryonic stem cells; induced pluripotent stem cells; zinc finger nuclease; gene targeting; OCT4; AAVS1; PITX3; homologous recombination

In hESCs, gene targeting by homologous recombination has proven to be difficult and only few reports have described successful gene targeting since the derivation of the first hESCs more than 10 years ago^{1–9}. These studies illustrate the utility of genetically modified hESCs by gene targeting, but a general approach to manipulate the hESC genome is still lacking. Recently, a novel technique based on the introduction of DNA double-strand breaks (DSBs)

⁴Correspondence should be addressed to R.J. (jaenisch@wi.mit.edu).

*These authors contributed equally to this work

by site-specific zinc finger nucleases (ZFNs) to facilitate homologous recombination has been used to target endogenous genes in human cells^{10, 11}. To generate a ZFN the FokI nuclease domain is fused to a DNA recognition domain composed of engineered C2H2 zinc finger motifs that specify the genomic DNA binding site for the chimeric protein (Figure 1A). Upon binding of two of such fusion proteins at adjacent genomic sites, the nuclease domains dimerize, become active, and cut the genomic DNA. When provided with a donor DNA that carries homology to both sides of the DSB, the genomic site can be repaired by homology-directed repair allowing the incorporation of exogenous sequences placed between homology regions^{12, 13}. This technique, also called “genome editing”, has been previously used in systems not easily accessible to genetic modifications such as zebrafish, plants, and rats^{14–19} and has also been used to edit the *CCR5* locus in hESCs¹³.

We used the locus encoding the pluripotency gene *OCT4*, one of the few genes that has been successfully targeted in hESCs⁶, to compare the efficiency of ZFN mediated gene targeting in human ES cells with conventional homologous recombination. We designed four ZFN pairs, which recognize unique sequences in the first intron of the *OCT4* gene (Figure 1A, Supplemental Figure 1 and Supplemental Table 1) and generated targeting donor constructs with short homology arms for the three most active ZFNs pairs. Correct targeting of these donor constructs containing a splice acceptor (SA) followed by an eGFP-2A-Puromycin cassette results in the expression of two proteins: a fusion protein comprised of the first 132 amino acids of human OCT4 fused to eGFP (OCT4^{EX1}-eGFP) and the puromycin N-acetyl-transferase, both under the control of the endogenous OCT4 promoter. Southern blot analysis using external probes 3' and 5' to the donor homology regions and an internal probe against eGFP revealed that for ZFN pair#1 40 out of 42 individual cell lines established from puromycin resistant clones were correctly targeted (efficiency >94%; Figure 1B and Table 1). ZFN pair#2 had a correct targeting frequency between 36% and 53% (Table1). Of the remaining clones most were correctly targeted in the *OCT4* locus, but also carried additional non-homologous integrations (Figure 1B and Table 1). ZFN pair#3 showed the lowest targeting efficiency and generated only a few the puromycin resistant clones some of which were correctly targeted. OCT4^{EX1}-eGFP targeted cells maintained a pluripotent state, as indicated by the expression of the pluripotency markers OCT4, NANOG, SOX2, Tra-1–60 and SSEA4 (Figure 1C) and their ability to form cell types originating from all three developmental germ layers in teratoma formation assays (Figure 1D).

We detected expression of the OCT4^{EX1}-eGFP fusion protein in hESCs by western blotting with antibodies against OCT4 and eGFP (Figure 1E). When targeted hESCs were differentiated into fibroblasts, no OCT4^{EX1}-eGFP protein was detected (Figure 1E) and the cells regained puromycin sensitivity demonstrating the validity of reporter expression. Unexpectedly, clones targeted by ZFN pair #1 showed significantly lower OCT4^{EX1}-eGFP protein expression than clones targeted by ZFN pair #2 as shown by western blot and FACS analysis (Figure 1E and Supplemental Figure 2). Additionally, ZFN pair #1 targeted hESCs only tolerated puromycin concentrations of 0.5µg/ml while ZFN pair #2 targeted hESCs could be maintained in up to 2µg/mL puromycin (data not shown) suggesting that ZFN pair #1 cells expressed lower levels of puromycin N-acetyl-transferase. One possible explanation for this reduced expression in ZFN pair #1 clones is that the integration of the SA-eGFP-2A-

puromycin cassette was very close to the splice donor of the first coding exon of OCT4 (separated by 102 bp) perhaps impeding splicing efficiency and resulting in reduced OCT4^{EX1-eGFP} expression.

Approaches to express transgenes in hESCs include random integration of expression vectors and the insertion of an expression construct with isogenic vector arms into the ROSA26 locus⁴ by homologous recombination. However, unpredictable position effects or low targeting frequencies have limited the utility of these approaches. In order to develop a highly efficient and robust expression system we used ZFN technology to target the AAVS1 locus located on chromosome 19, encoding the PPP1R12C gene, which is ubiquitously expressed. This well characterized locus has been previously shown to allow stable and long-term expression of transgenes in multiple cell types including hESCs²⁰. To target the first intron of the PPP1R12C gene, we used a ZFN pair, which generates a DSB at its target locus in hESCs (Supplemental figure 1B and Supplemental Table 1) and can be used to efficiently target transgenes into transformed human cell lines (DeKolver et al., submitted).

Two constructs with identical short homology arms were used to target the AAVS1 locus: (i) A gene trap vector for the PPP1R12C promoter containing a splice acceptor- 2A- puromycin selection cassette. (ii) A puromycin selection cassette driven from the phosphoglycerol kinase (PGK) promoter (Figure 2A). As indicated by Southern blot analysis, both vectors generated approximately 50% of puromycin resistant clones that had correctly targeted insertions on one or both alleles with no additional random integrations (Figure 2B, Table 1 and Supplemental Figure 3). Targeting by vectors with an exogenous promoter-controlled selection cassette is important for targeting genes not expressed in hESCs as demonstrated below for the *PITX3* gene. All tested AAVS1-targeted hESCs, including homozygous targeted clones, retained a normal karyotype (n=4) (Supplemental Figure 4A) and remained pluripotent based on immunofluorescence staining for pluripotency markers (Supplemental Figure 4B) and teratoma formation assays (Supplemental Figure 4C). HiPSCs²¹ could be targeted with similar efficiency as hESCs (Supplemental Figure 5A, B and Table1)

To develop a transgenic overexpression system, we targeted a donor plasmid expressing eGFP under control of the constitutively active CAGGs promoter to the AAVS1 locus. HESCs targeted with this construct showed persistent and uniform GFP expression (Figure 2D). To generate an inducible expression system an AAVS1 donor plasmid containing a minimal CMV promoter and the tetracycline response element driving the eGFP cDNA (TetO-eGFP) was targeted in either the same or reverse orientation as the PPP1R12C gene (Figure 2E). ZFN-mediated targeting of BGO1 cells with these donor plasmids resulted in correctly targeted heterozygous (AAVS1-TetO-eGFP^{+/-}) and homozygous (AAVS1-TetO-eGFP^{+/+}) hESCs (Figure 2G) with similar high efficiencies as above (TetO-FW: 47%, TetO-BW 40%, Table1). Correctly targeted hESCs were transduced with a lentivirus carrying the M2rtTA reverse transactivator to render the cells doxycycline (DOX) responsive. In agreement with previously reported lentiviral transduction efficiencies of hESCs, 10% of the cells showed eGFP expression after the addition of DOX (Figure 2F). To test DOX-inducible GFP expression we withdrew DOX from the cultures for varying lengths of time and found that eGFP fluorescence became undetectable after 7 days of DOX withdrawal

(Supplemental Figure 6). We next established cell lines that showed limited silencing of the M2rtTA viral transgene by single cell subcloning (Figure 2F). FACS analysis of these cell lines, cultured under different DOX concentrations, revealed a dose-dependent relationship between DOX concentration and eGFP expression (Figure 2H). These experiments show that eGFP expression was dependent on DOX addition and on the presence of M2rtTA, indicating tight regulation of the eGFP expression cassette when integrated into the AAVS1 locus. There were no apparent differences between the two orientations of the TetO-eGFP cassette.

Finally, we tested whether ZFNs could be used to modify genes that are not expressed in hESCs and hiPSCs by targeting the first exon of *PITX3*. *PITX3* is a transcription factor expressed in some differentiated cell types such as dopaminergic neurons but not in hESCs. To generate a *PITX3* reporter, we designed a targeting construct in which the *PITX3* open reading frame was joined after amino acid 32 to eGFP followed by a polyadenylation signal and a floxed PGK-puromycin cassette (Figure 3A). Co-electroporation of the donor plasmid with the *PITX3* ZFN pair#2 (Supplemental Figure 7) into hESCs or two hiPSCs²¹ lines resulted in clones that were targeted with efficiencies of 11% and 8%, respectively, as determined by Southern blot analysis (Figure 3 B, C and Table 1). All tested hESCs and hiPSCs targeted in the *PITX3* locus maintained a normal karyotype (hESC n=3, hiPSC n=3). The PGK-puromycin cassette was subsequently removed by transient expression of Cre-recombinase (Figure 3B). The relatively high targeting efficiency in hESCs and hiPSCs demonstrates that ZFN mediated gene targeting is a robust tool to modify genes not expressed in hESCs to generate cell-type specific reporter systems. ZFN mediated gene targeting therefore has the potential to overcome one of the main obstacles in hESC and hiPSC based research.

A concern of the ZFN-targeting approach is off-target DNA breaks induced at related sequences elsewhere in the genome. This is a potential limitation of using ZFNs, as off-target cleavage may cause unpredictable genotoxic effects. While Southern blots using internal and external probes excluded additional integrations and confirmed clonality of targeted clones, these analyses do not detect ZFN-mediated double strand breaks and error prone repair elsewhere in the genome. To stringently address the frequency of off-target cleavage, we determined the DNA binding specificity for all ZFN nucleases by SELEX (Supplemental Table 3), which allowed the identification of the most probable off-target cleavage sites on a genome-wide basis (Supplemental materials and methods, and Supplemental Figures 8A–11A). Using the Cel-1 assay we quantified the actual frequency of Non-homologous end joining (NHEJ) mediated alterations in up to 10 potential off-target sites in clones generated by four different ZFNs used in this study (Table 1B and Supplemental Figure 8B and C–11B and C). In analyzing 46 genomic loci, we detected one NHEJ alteration at one genomic site in one of the 4 *PITX3* clones (Table 1 and Supplemental Figure 9 B,C and D); all the other putative off-target sites for all the ZFNs in all the clones were wild-type. Finally, we determined in heterozygous-targeted clones the frequency of NHEJ alterations on the allele that does not carry an integrated transgene. This frequency varied for the three target genes between 1/18 for *PITX3* and 1/12 for the *OCT4* clones carrying a disruption on the other allele (Table 1B); in all other cases the other allele

remained wild-type. In order to ensure a functional wild type allele in heterozygous-targeted clones, the recognition sequence of the ZFNs can be designed to recognize intron sequences as was done for the *OCT4* and AAVS1 loci.

A recent study²² described the use of ZFNs to target a site in the human PIG-A gene; together with data on targeting 5 distinct loci in 3 different genes described in the present work, these data demonstrate the utility of using ZFNs to genetically engineer hESCs and iPSCs. The possibility of off-target cleavage or of mutagenesis in the allele not carrying the insert was not addressed in that study²². It would be important interesting to determine the genome-wide spectrum of action for the 3-finger ZFNs used by Cheng and colleagues, including whether these ZFNs, which target an exon of the PIG-A gene, also cut at an identical site found in the in the PIG-A pseudogene 1 (PIGAP1) on chromosome 12.

ZFN mediated gene targeting requires only relatively short targeting arms (for AAVS1 about 500 bp for each arm) facilitating the generation of targeting constructs to achieve site-specific integration of exogenous genes whose expression can be controlled by constitutively active, inducible or tissue specific promoters. This is in contrast to targeting of genes such as the *OCT4* by conventional homologous recombination, which resulted in targeting efficiencies ranging from 20% to 40 % using isogenic targeting arms of 7.9 and 12.8kb, respectively⁶. Moreover, the use of short targeting arms combined with ZFN technology results in high targeting efficiencies even when using the same donor plasmids in genetically different cell lines, thereby eliminating the need for construction of isogenic targeting vectors. While it is possible that not all non-expressed genes can be targeted by ZFNs, the extraordinary flexibility in designing the zinc finger binding motifs^{23, 24} represents a tool that will likely allow targeting of a substantial fraction of genes. Thus, the ZFN targeting technique may help generate the genetic tools to study cell fate decisions and to use cell-type specific reporter systems to improve hESC differentiation paradigms.

EXPERIMENTAL PROCEDURES

Cell culture

Cell culture techniques have been described previously²¹. hiPSCs and the hESC lines BG01 (NIH Code: BG01; BresaGen, Inc., Athens, GA) were maintained on mitomycin C inactivated mouse embryonic fibroblast (MEF) feeder layers in hESC medium [DMEM/F12 (Invitrogen) supplemented with 15 % fetal bovine serum (FBS) (Hyclone), 5% KnockOut™ Serum Replacement (Invitrogen), 1 mM glutamine (Invitrogen), 1% nonessential amino acids (Invitrogen), 0.1 mM β -mercaptoethanol (Sigma) and 4 ng/ml FGF2 (R&D systems)]. Cultures were passaged every 5 to 7 days either manually or enzymatically with collagenase type IV (Invitrogen; 1.5 mg/ml). In order to perform FACS analysis of OCT4-eGFP clones in the absence of MEF-feeder cells, hESCs were passaged onto Matrigel coated plates in mTeSR medium (Stemcell Technologies). Karyotyping analysis was performed by Cell Line Genetics (Madison, WI, 53719).

ZFN nuclease design and ZFN expression plasmids

Zinc finger nucleases against the human *OCT4*, *AAVS1* and *PITX3* loci were designed using an archive of pre-validated 2-finger modules exactly as described^{10,16,25}; complete sequences of the ZFNs, which carried obligate heterodimer forms of the FokI endonuclease²⁶, are provided in Supplementary Table 1. The ZFNs were designed and tested at Sangamo BioSciences for the purpose of disruption of their intended target loci by transient transfection into K562, HeLa, and HEK293 cells, followed by Surveyor (Cel-1) endonuclease-based measurement of NHEJ at the target locus exactly as described^{26,25}; primers used in Cel-1 analysis are provided in the Supplementary Table 2); the ZFN expression constructs were then provided to the Jaenisch laboratory.

Targeting of hESCs and hiPSCs using ZFN mediated homologous recombination

hESCs and hiPSCs were cultured in Rho Kinase (ROCK)-inhibitor (Calbiochem; Y-27632) 24 hours prior to electroporation. Cells were harvested using 0.25% trypsin/EDTA solution (Invitrogen) and 1×10^7 cells resuspended in phosphate buffered saline (PBS) were electroporated if not otherwise indicated with 40 μ g of donor plasmids (designed and assembled by D.H. and F.S.) and 5 μ g of each ZFN encoding plasmid (Gene Pulser Xcell System, Bio-Rad: 250 V, 500 μ F, 0.4 cm cuvettes³). Cells were subsequently plated on MEF feeder layers (DR4 MEFs for puromycin selection) in hESC medium supplemented with ROCK-inhibitor for the first 24 hours. Individual colonies were picked and expanded after puromycin selection (0.5 μ g/ml) 10 to 14 days after electroporation.

Experimental genome-wide evaluation of ZFN action

A DSB persistently induced by designed ZFNs is repaired by non-homologous end joining, an error-prone process²⁷ that generates small insertions and deletions at the site of the break²⁸. This feature has been extensively used to profile the consequences of ZFN-driven editing on the target genome^{15–17, 25}. In the present work, such genotyping was performed essentially as described^{16, 25}. First, the consensus target for each ZFN was experimentally determined by SELEX as described²⁵ under conditions known to yield a biologically relevant consensus site for C2H2 ZFPs²⁹. These studies yielded the targets provided in Supplementary Table 3. Next, the human genome was searched for candidate off-target sites that provided the best match to the experimentally determined base frequency matrices obtained from the SELEX studies. In performing this step, we allowed ZFN site pairings with 5 or 6 bp between individual targets, in order to reflect the ability of our designed ZFNs to cleave equally well at these two spacings. Likewise, we also allowed an optional gap of 1 bp between the 9th and 10th bases of Oct-4 ZFN#1-R in order to reflect the binding characteristics of a longer flexible linker between the second and third fingers of this protein which allows binding to either target type. Finally, the experimentally determined base frequency matrices for each ZFN pair were used to rank the potential off-target sites. For each ZFN target, the top 10 ranked off-target sites were then genotyped as follows. Four single-cell-derived clones heterozygous for the transgene at the ZFN target site were chosen at random, genomic DNA was isolated, and every potential off-target site was amplified using 32 cycles of PCR with Accuprime Taq HiFi DNA polymerase (Invitrogen). The majority of the sites were then genotyped using the Surveyor endonuclease (“Cel-1”;

Transgenomics) assay exactly as described²⁶, with the following modification: to address the potential for biallelic homozygous disruption (which would yield no Cel-1 signal), an equal amount of PCR product amplified from control cells was added to that from the ZFN-edited clone. Following a denature-renature step and treatment by Cel-1 to cleave heteroduplexes formed from wild-type and mutated DNA strands, the reaction was resolved on a 10% nondenaturing PAGE (BioRad) in 1× TBE. One putative off-target site was found to be heterozygous for a SNP, precluding the use of Cel-1 assay; it was genotyped by cloning (TopoTA; Invitrogen) and Sanger sequencing. In addition to analyzing putative off-target sites, for each heterozygous clone, the non-transgenic allele of the ZFN target locus was genotyped by PCR (using primers shown in Supplemental Table 1), cloning, and sequencing.

Fibroblast differentiation of OCT4-eGFP hESCs

EB induced differentiation was performed as previously described³⁰. Briefly, hESC colonies were harvested using 1.5 mg/ml collagenase type IV (Invitrogen), separated from the MEF feeder cells by gravity, gently triturated and cultured for 7 days in non-adherent suspension culture dishes (Corning) in DMEM supplemented with 20% FBS. EBs were plated onto adherent tissue culture dishes and passaged according to primary fibroblast protocols using trypsin for at least four passages before the start of experiments.

Removal of PGK-Puro cassette by transient Cre-recombinase expression

HESCs targeted in the *PITX3* locus were cultured in Rho Kinase (ROCK)-inhibitor (24 hours prior to electroporation. Cells were harvested using 0.25% trypsin/EDTA solution (Invitrogen) and 1×10^7 cells resuspended in PBS were electroporated with pTurbo-Cre (40 µg; Genbank Accession Number AF334827) and pEGFP-N1 (10 µg; Clontech) as described previously³ (Gene Pulser Xcell System, Bio-Rad: 250 V, 500 µF, 0.4 cm cuvettes). Cells were subsequently plated on MEF feeder layers in hESC medium supplemented with ROCK-inhibitor. Cre-recombinase expressing cells were enriched by FACS sorting (FACS-Aria; BD-Biosciences) of a single cell suspension for eGFP expressing cells 60 hours after electroporation followed by replating at a low density in ROCK-inhibitor containing hESC medium. Individual colonies were picked 10 to 14 days after electroporation.

Lentiviral infection of hESCs

The FUW-M2rtTA lentiviral vector has been described previously³⁰. VSVG coated lentiviruses were generated in 293 cells as described previously³¹. Briefly, culture medium was changed 12 hours post-transfection and virus-containing supernatant was collected 60–72 hours post transfection. Viral supernatant was filtered through a 0.45 µm filter. This virus-containing supernatant was used to infect hESCs aggregates that were separated from feeder cells by collagenase treatment and serial washes. Two consecutive infections in the presence of 2 µg/ml of polybrene were performed over a period of 12 hours in suspension. hESC cell aggregates were replated after infection on feeder cells. Infection efficiencies were determined using FACS analysis for eGFP and SSEA4 (mouse monoclonal, Developmental Studies Hybridoma Bank) of cells cultured in the presence of doxycycline (DOX) (Sigma-Aldrich; 2 µg/ml) for two days. To enrich for eGFP expressing cells,

targeted and infected hESCs were FACS sorted as single cell suspension 2 days after doxycycline induction in the presence of ROCK-Inhibitor (FACS-Aria; BD-Biosciences) and subsequently replated in the ROCK-Inhibitor containing hESC medium. DOX responsive GFP expressing cell lines were isolated by manual picking of single colonies.

Teratoma formation and analysis

HESCs were collected by collagenase treatment (1.5 mg/ml) and separated from feeder cells by subsequent washes with medium and sedimentation by gravity. HESC aggregates were collected by centrifugation, resuspended in 250 µl of PBS and injected subcutaneously in the back of SCID mice (Taconic). Tumors generally developed within 4–8 weeks and animals were sacrificed before tumor size exceeded 1.5 cm in diameter. Teratomas were isolated and fixed in formalin. After sectioning, teratomas were diagnosed based on hematoxylin and eosin staining.

Immunocytochemistry

Cells were fixed in 4% paraformaldehyde in PBS and immunostained according to standard protocols using the following antibodies: SSEA4 (mouse monoclonal, Developmental Studies Hybridoma Bank); Tra-1–60, (mouse monoclonal, Chemicon International); hSOX2 (goat polyclonal, R&D Systems); Oct-3/4 (mouse monoclonal, Santa Cruz Biotechnology); hNANOG (goat polyclonal R&D Systems) and appropriate Molecular Probes Alexa Fluor® dye conjugated secondary antibodies (Invitrogen).

Immunoblotting

hESCs were collected by collagenase treatment (1.5 mg/ml) and separated from feeder cells by subsequent washes with medium and sedimentation by gravity. hESC derived fibroblasts were collected by trypsinization. Cells were pelleted by centrifugation and washed with PBS and again collected by centrifugation. Cells were lysed in ice-cold buffer (50 mM Tris-HCl, pH 7.4, 20% glycerol, 1 mM EDTA, 150 mM NaCl, 0.5% Triton X-100, 0.02% SDS, 1 mM dithiothreitol [DTT], 2 mM phenylmethylsulfonyl fluoride [PMSF], supplemented with proteinase inhibitor cocktail (Complete Mini, Roche). After 5 min on ice, 5 M NaCl was added to bring the final [NaCl] to 400 mM. After 5 min on ice, an equal volume of ice-cold water was added and the lysate was mixed before immediate centrifugation in a microfuge (14k rpm, 10 min). Protein concentration of the supernatant was determined by Bradford assay and 15 µg of protein was separated using 4–12% Bis-Tris gradient gels (Invitrogen). After transfer to PVDF membranes blots were probed with OCT4 (mouse monoclonal, Santa Cruz Biotechnology) or GFP (Rbt pAB to GFP, Abcam ab290-50) antibodies.

Southern blotting

Genomic DNA was separated on a 0.7% agarose gel after restriction digest with the appropriate enzymes, transferred to a nylon membrane (Amersham) and hybridized with ³²P random primer (Stratagene) labeled probes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Raaji Alagappan, Ping Xu and Elizabeth Cook for technical support and Jessica Dausman, Ruth Flannery, and Dongdong Fu for their help with animal husbandry and processing of teratomas. We thank all the members of the Jaenisch lab for helpful discussions and comments on the manuscript. D.H. is a Merck Fellow of the Life Science Research Foundation. R.J. was supported by NIH grants R37-CA084198, RO1-CA087869, and RO1-HD045022. R.J. is an adviser to Stemgen and a cofounder of Fate Therapeutics. R.C.D., G.E.K., R.A., B.Z., X.M., J.C.M., L.Z., E.J.R., P.D.G. and F.D.U. are full-time employees of Sangamo BioSciences, Inc. Requests for ZFNs should be directed to F.U. (furnov@sangamo.com).

References

1. Thomson JA, et al. Embryonic stem cell lines derived from human blastocysts. *Science*. 1998; 282:1145–1147. [PubMed: 9804556]
2. Urbach A, Schuldiner M, Benvenisty N. Modeling for Lesch-Nyhan disease by gene targeting in human embryonic stem cells. *Stem Cells*. 2004; 22:635–641. [PubMed: 15277709]
3. Costa M, et al. A method for genetic modification of human embryonic stem cells using electroporation. *Nat Protoc*. 2007; 2:792–796. [PubMed: 17446878]
4. Irion S, et al. Identification and targeting of the ROSA26 locus in human embryonic stem cells. *Nat Biotechnol*. 2007; 25:1477–1482. [PubMed: 18037879]
5. Suzuki K, et al. Highly efficient transient gene expression and gene targeting in primate embryonic stem cells with helper-dependent adenoviral vectors. *Proc Natl Acad Sci U S A*. 2008; 105:13781–13786. [PubMed: 18768795]
6. Zwaka TP, Thomson JA. Homologous recombination in human embryonic stem cells. *Nat Biotechnol*. 2003; 21:319–321. [PubMed: 12577066]
7. Davis RP, et al. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. *Blood*. 2008; 111:1876–1884. [PubMed: 18032708]
8. Xue H, et al. A Targeted Neuroglial Reporter Line Generated by Homologous Recombination in Human Embryonic Stem Cells. *Stem Cells*. 2009
9. Ruby KM, Zheng B. Gene Targeting in a HUES Line of Human Embryonic Stem Cells Via Electroporation. *Stem Cells*. 2009; 27:1496–1506. [PubMed: 19544466]
10. Urnov FD, et al. Highly efficient endogenous human gene correction using designed zinc-finger nucleases. *Nature*. 2005; 435:646–651. [PubMed: 15806097]
11. Carroll D. Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther*. 2008; 15:1463–1468. [PubMed: 18784746]
12. Moehle EA, et al. Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. *Proc Natl Acad Sci U S A*. 2007; 104:3055–3060. [PubMed: 17360608]
13. Lombardo A, et al. Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. *Nat Biotechnol*. 2007; 25:1298–1306. [PubMed: 17965707]
14. Cai CQ, et al. Targeted transgene integration in plant cells using designed zinc finger nucleases. *Plant Mol Biol*. 2009; 69:699–709. [PubMed: 19112554]
15. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA. Targeted gene inactivation in zebrafish using engineered zinc-finger nucleases. *Nat Biotechnol*. 2008; 26:695–701. [PubMed: 18500337]
16. Doyon Y, et al. Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nat Biotechnol*. 2008; 26:702–708. [PubMed: 18500334]
17. Shukla VK, et al. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. *Nature*. 2009
18. Townsend JA, et al. High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*. 2009; 459:442–445. [PubMed: 19404258]
19. Geurts AM, et al. Knockout Rats via Embryo Microinjection of Zinc-Finger Nucleases. *Science*. 2009; 325:433. [PubMed: 19628861]

20. Smith JR, et al. Robust, persistent transgene expression in human embryonic stem cells is achieved with AAVS1-targeted integration. *Stem Cells*. 2008; 26:496–504. [PubMed: 18024421]
21. Soldner F, et al. Parkinson's disease patient-derived induced pluripotent stem cells free of viral reprogramming factors. *Cell*. 2009; 136:964–977. [PubMed: 19269371]
22. Zou J, et al. Gene Targeting of a Disease-Related Gene in Human Induced Pluripotent Stem and Embryonic Stem Cells. 2009; 5:97–110.
23. Pabo CO, Peisach E, Grant RA. Design and selection of novel Cys2His2 zinc finger proteins. *Annual review of biochemistry*. 2001; 70:313–340.
24. Klug A. The discovery of zinc fingers and their development for practical applications in gene regulation. *Proc. Japan Acad.* 2005; 81:87–102.
25. Perez EE, et al. Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. *Nat Biotechnol*. 2008; 26:808–816. [PubMed: 18587387]
26. Miller JC, et al. An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol*. 2007; 25:778–785. [PubMed: 17603475]
27. Valerie K, Povirk LF. Regulation and mechanisms of mammalian double-strand break repair. *Oncogene*. 2003; 22:5792–5812. [PubMed: 12947387]
28. Bibikova M, Golic M, Golic KG, Carroll D. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. *Genetics*. 2002; 161:1169–1175. [PubMed: 12136019]
29. Phillips CM, et al. Identification of chromosome sequence motifs that mediate meiotic pairing and synapsis in *C. elegans*.
30. Hockemeyer D, et al. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell*. 2008; 3:346–353. [PubMed: 18786421]
31. Brambrink T, et al. Sequential expression of pluripotency markers during direct reprogramming of mouse somatic cells. *Cell Stem Cell*. 2008; 2:151–159. [PubMed: 18371436]

Figure 1. Targeting of *OCT4* in hESCs using ZFNs

A. Schematic overview depicting the targeting strategy for the *OCT4* locus. Probes used for Southern blot analysis are shown as red boxes, exons of the *OCT4* locus are shown as blue boxes and arrows indicate the genomic site cut by the respective ZFN pair. Donor plasmids corresponding to the cleavage location of the three ZFN pairs carried 5' and 3' homology regions covering roughly 700 bp of the *OCT4* sequence flanking the respective DSB target site are shown above; SA-GFP: splice acceptor eGFP sequence, 2A: self-cleaving peptide sequence, PURO: puromycin resistance gene, polyA: polyadenylation sequence. Inset in the

upper left depicts a cartoon of two ZFNs binding at a specific genomic site (yellow) leading to the dimerization of the FokI nuclease domains.

B. Southern blot analysis of BG01 cells targeted with the indicated ZFN pairs using the corresponding donor plasmids. Genomic DNA was digested either with EcoRI and hybridized with the external 3'-probe or digested with SacI and hybridized with the external 5'-probe or internal eGFP probe. Correctly targeted clones without additional integrations are indicated in red.

C. Immunofluorescence staining of BG01 cells targeted with the indicated ZFN pairs using the corresponding donor plasmids. Cells were stained for the pluripotency markers OCT4, NANOG, SOX2, Tra-1-60 and SSEA4.

D. Hematoxylin and eosin staining of teratoma sections generated from BG01 cells targeted with the indicated ZFN pairs and the corresponding donor plasmids.

E. Western blot analysis for the expression of OCT4 and eGFP in BG01 wild type cells and BG01 cells targeted with the indicated ZFN pairs using the corresponding donor plasmids. Cell extracts were derived from either undifferentiated cells (ES) or in vitro differentiated fibroblast-like cells (Fib.)

Figure 2. Targeting of the AAVS1 locus using ZFNs

A. Schematic overview depicting the targeting strategy for the PPP1R12C gene in the AAVS1 locus. Probes used for Southern blot analysis are shown as red boxes, the first 3 exons of PPP1R12C gene are shown as blue boxes; the arrow indicates the genomic site cut by the AAVS1-ZFNs. Donor plasmids used to target the locus are shown above; SA-Puro: splice acceptor sequence followed by a 2A self-cleaving peptide sequence and the puromycin resistance gene, pA: polyadenylation sequence, PGK: human phosphoglycerol kinase promoter, PURO: puromycin resistance gene.

- B. Southern blot analysis of BG01 cells targeted with the indicated donor plasmids using the AAVS1 ZFNs. Genomic DNA was digested with SphI and hybridized with the ³²P-labeled external 3'-probe or with the internal 5'-probe. Fragment sizes are: PGK-Puro: 5'-probe: wt=6.5 kb, targeted=4.2 kb; 3'-probe: wt=6.5 kb, targeted=3.7 kb. SA-Puro: 5'-probe: wt=6.5 kb, targeted=3.8 kb; 3'-probe: wt=6.5 kb, targeted=3.7 kb.
- C. Southern blot analysis of BG01 cells targeted with an AAVS1 donor plasmid containing a CAGGs driven eGFP cassette using the AAVS1 ZFNs. Genomic DNA was digested with SphI and hybridized with the ³²P-labeled external 3'-probe or with the internal 5'-probe. Fragment sizes are: CAGGs-GFP: 5'-probe: wt=6.5 kb, targeted=3.8 kb; 3'-probe: wt=6.5 kb, targeted=6.9 kb.
- D. Phase contrast picture and fluorescence imaging of eGFP in heterozygous or homozygous BG01 clones targeted with an AAVS1 donor plasmid containing a CAGGS driven eGFP cassette and the AAVS1 ZFNs.
- E. Schematic overview depicting the targeting strategy for the PPP1R12C gene in the AAVS1 locus with a donor construct containing a DOX-inducible TetO-eGFP. Probes used for Southern blot analysis are shown as red boxes, the first 3 exons of the PPP1R12C gene in the AAVS1 locus are shown as blue boxes and arrows indicate the genomic site cut by the ZFNs. Donor plasmids used to target the AAVS1 locus are shown above; SA-Puro: splice acceptor sequence followed by a 2A self-cleaving peptide sequence and the puromycin resistance gene, pA: polyadenylation sequence, TetO: Tetracycline response element.
- F. Phase contrast picture and fluorescence imaging of eGFP in BG01 cells either heterozygous (AAVS1 TetO-GFP^{+/-}) or homozygous (AAVS1-TetOGFP^{+/+}) for the DOX-inducible eGFP cassette targeted to the AAVS1 locus. Cells were transduced with a M2rtTA lentivirus to render the cells DOX responsive. Panel shows colonies before (top) and after FACS assisted subcloning in the presence of DOX (bottom).
- G. Southern blot analysis of BG01 cells targeted with the indicated donor plasmids using the AAVS1 ZFNs. Genomic DNA was digested with SphI and hybridized with the ³²P-labeled external 3'-probe or with the internal 5'-probe.
- H. FACS analysis of AAVS1-TetO-GFP^{+/-} and AAVS1-TetO-GFP^{+/+} subclones for GFP expression at different concentrations of DOX. BG01 cells, targeted cells prior to M2rtTA infection, and subcloned GFP responsive cell lines were cultured at different concentrations of DOX and analyzed. All cells were co-stained and analyzed for SSEA4 expression to exclude SSEA4 negative feeder cells from the analysis.

Figure 3. Targeting of *PITX3* in hESCs and hiPSCs using ZFNs

A. Schematic overview depicting the targeting strategy for the *PITX3* gene. Probes for Southern blot analysis are shown as red boxes, the first exons of the *PITX3* locus are shown as blue boxes and arrows indicate the genomic site cut by the ZFN pair #2. Donor plasmids used to target the *PITX3* locus are shown above and contained 5' and 3' homologous sequences of approximately 800 bp flanking the predicted ZFN pair #2 target site; eGFP: enhanced green fluorescent protein, PGK: human phosphoglycerol kinase promoter, PURO: puromycin resistance gene, loxP: loxP sites, pA: polyadenylation sequence. Two constructs that differed only in the orientation of this selection cassette with respect to the *PITX3* gene were successfully used to target *PITX3* (See also Table 1).

B. Southern blot analysis of BG01 cells targeted with the indicated donor plasmid using the *PITX3* ZFNs. The right panel shows Southern blot analysis of clones in which the PGK-Puro cassette was removed by transient Cre-recombinase expression. Genomic DNA was digested with HindIII and probed with ³²P-labeled external 5'-probe or with the internal 3'-probe. Fragment sizes are: 5' probe: wt=8.8 kb, targeted=7.4 kb -PGK-Puro=10.5 kb; 3'-probe: wt=8.8 kb, targeted=4.3 kb.

C. Southern blot analysis of the hiPSCs targeted with the indicated donor plasmids using the *PITX3* ZFNs. Genomic DNA was digested with HindIII and probed with ³²P-labeled external 5' probe or with the internal 3'-probe. Fragment sizes are: 5' probe: wt=8.8 kb, targeted=7.4 kb; 3'-probe: wt=8.8 kb, targeted=4.3 kb.

Table 1

Summary of targeting experiments

OCT4									
A									
cell line targeted	ZFN pair	donor	# of clones picked*	random integration*	targeted +additional integration*	correct targeted clones			
						hetero-zygous*	homo-zygous*	targeting efficiency [%]	
BGO1	control	OCT-GFP#1,2,3	2/1	2/1	0	0	0	0	
BGO1	ZFN#1 (2.5µg)	OCT-GFP#1	4/21	0/1	0	4/20	0	100/95	
BGO1	ZFN#1 (10µg)	OCT-GFP#1	17	1	0	16	0	94	
BGO1	ZFN#2 (2.5µg)	OCT-GFP#2	15/22	0/1	7/13	8/8	0	53/36	
BGO1	ZFN#2 (10µg)	OCT-GFP#2	31	1	18	12	0	39	
BGO1	ZFN#3 (2.5µg)	OCT-GFP#3	2	1	0	1	0	50	
BGO1	ZFN#3 (10µg)	OCT-GFP#3	1	0	1	0	0	0	
AAVSI									
cell line targeted	ZFN pair	donor	# of clones picked	random integration	targeted +additional integration	correct targeted clones			
						hetero-zygous	homo-zygous	targeting efficiency [%]	
BGO1	control	AAVSI/SA-Puro	10	10	0	0	0	0	
BGO1	AAVSI	AAVSI/SA-Puro	32	2	12	16	2	56	
BGO1	control	AAVSI/PGK-Puro	36	36	0	0	0	0	
BGO1	AAVSI	AAVSI/PGK-Puro	35	13	5	16	1	49	
BGO1	AAVSI	AAVSI/TetO-GFP fw	46	5	19	15	7	47	
BGO1	AAVSI	AAVSI/TetO-GFP bw	35	0	21	10	4	40	
iPS PD2 ^{1/lox} -17Puro-5	AAVSI	AAVSI/SA-Puro	23	1	8	11	3	61	
iPS PD2 ^{1/lox} -17Puro-5	AAVSI	AAVSI/PGK-Puro	15	5	5	5	0	33	
iPS PD2 ^{1/lox} -17Puro-10	AAVSI	AAVSI/PGK-Puro	37	9	9	15	4	51	

PITX3									
cell line targeted	ZFN pair	donor	# of clones picked	random integration	targeted +additional integration	correct targeted clones			targeting efficiency [%]
						hetero-zygous	homo-zygous		
BGO1**	PITX3	PITX3 GFP fw/bw	96/74	126	14/12	7/11	0		11
iPS PD2 ^{1/lox} -17Puro-10/-21Puro-20***	PITX3	PITX3 fw	30/20	23/18	4/1	3/1	0		8
B									
ZFN	NHEJ-frequency of "wt allele" in het clones		NHEJ-at the "top 10" off target sites in correctly targeted clones						
OCT4ZFN#1	1 [#] /11		0/72						
OCT4ZFN#2	0/12		0/40						
AAVS1	0/5		0/36						
PITX3 ZFN#2	1 [#] /18		1/36						

* when two numbers are show this indicates the results from two independent experiments

** first number indicates the results for targeting with the PITX3 GFP fw donor, second number indicated those with bw donor

*** first number indicates the results for targeting iPS PD2^{1/lox} -17Puro-10 cell line, the second number indicates those for iPS PD2^{1/lox} -21Puro-20 The iPSCs have previously been described in soldner et al., cell 2009

for OCT4, the mutated allele carried a deletion of 9 bp;
for PITX3, the one mutated allele carried a deletion of 8 bp.