Gene Targeting in Human Pluripotent Cells

D. Hockemeyer1 AND R. Jaenisch1,2

1Whitehead Institute and Department of Biology, 2Massachusetts Institute of Technology,
Cambridge, Massachusetts 02142

Correspondence: jaenisch@wi.mit.edu

Mouse embryonic stem cells (mESCs) have the ability to differentiate into any cell type and can generate chimeric mice when transplanted into a host blastocyst. This remarkable potential, together with the development of robust gene targeting strategies in mESCs, were essential for establishing the mouse as the most widely used model organism in biomedical research. Recent advances have allowed the isolation of human embryonic stem cells and the derivation of induced pluripotent stem cells. Genetic tools similar to those proven routine in the mouse system are needed to realize the full potential of human pluripotent cells as disease models and putative therapeutics. Gene targeting in human cells, however, has proven to be more difficult, more time-consuming, and less robust than in mESCs. In this chapter, we discuss the strategies that have been used to allow specific genetic modifications in human pluripotent cells. We focus on the novel application of custom-engineered zinc-finger nucleases for gene targeting, which has promise to become a robust tool for efficient genetic manipulation of human pluripotent cells.

The isolation of the first human embryonic stem cells (hESCs) by Thomson et al. provided hope that these cells would not only be a useful tool to study early human development but could also be used to study human genetic diseases in vitro and derive human tissue for transplantation therapies (Thomson et al. 1998). Recent advances pioneered by Yamanaka and colleagues in 2006 have allowed the generation of human pluripotent cells from adult fibroblasts (Takahashi and Yamanaka 2006; Takahashi et al. 2007), limiting the need to harvest cells from human embryos. This process, called somatic-cell reprogramming, reverts a cell’s epigenetic state to a developmentally earlier and therefore more potent state. These induced pluripotent stem cells (iPSCs) are attractive for biomedical research because they can be derived from patients who carry genetic predispositions for a given disease and from those who would benefit from cell transplantation. The realization of these applications for iPSC technology will require reliable tools for genetically modifying iPSCs. This would allow manipulation of the disease pathway of interest and potential genetic repair of disease-causing mutations in patient-derived cells. In addition, the ability to target genes in iPSCs would allow the generation of reporter cell lines that facilitate the identification and isolation of specific cell types in vitro. However, genetic manipulations of human cells by procedures that are standard in the mouse system have proven to be laborious and inefficient in hESCs and are not useful for routine applications. In this chapter, we discuss the basic principles of gene targeting and review the approaches that have been used thus far to manipulate genes in human cells. We then present a novel form of gene targeting in human pluripotent cells, using zinc-finger nucleases (ZFNs), that promises to provide a general tool to achieve gene targeting in human pluripotent cells.

GENE TARGETING IN MOUSE EMBRYONIC STEM CELLS

Before discussing gene targeting efforts in human cells, we first introduce the general principles and mechanisms of gene targeting as defined in the classic mouse system (Capecchi 1989; Koller and Smithies 1992). mESCs display high activity of the homologous recombination (HR) DNA repair pathway, a key feature that sets them apart from most other mammalian cells and allows efficient genetic manipulation by gene targeting. DNA repair by HR is one of two cellular pathways that normally repair double-strand DNA breaks (DSBs) (Helleday et al. 2007). Whereas HR repairs the DSB with high fidelity, the second pathway, the nonhomologous end joining (NHEJ) pathway, fuses the broken ends in an error-prone fashion (Valerie and Povirk 2003; Lieber 2010). The reason for the high fidelity of the HR pathway lies in its ability to use sequence information of the sister chromatid to repair DSB lesions. After the occurrence of a DSB, the DNA ends are resected, generating long single-strand 3′ DNA ends that can invade the sister-chromatid-bearing homology to the site of the DSB. The 3′ end of the broken DNA end is extended by DNA polymerases using the intact sister chromatid as a template. At this point, the sequence information of the sister chromatid is copied onto the broken strand, thereby restoring lost sequence information caused by the DSB. After this DNA synthesis step, the strand invasion is resolved and the initially broken chromosome ends are religated. In a case in which the sequence of the homologous template that is used in HR differs from the original sequence of the broken chromosome, HR can result in a gene conversion event. This gene conversion event is the underlying principle of gene targeting. During gene targeting, a cell aberrantly uses an ectopically provided DNA template.
that mimics the sister chromatid by bearing homologous sequences to the targeted genomic region but is mutated to allow modifications of the endogenous sequence. In comparison to somatic cells, mESCs exhibit a higher propensity for the HR pathway over the NHEJ pathway, thereby facilitating precise alteration of the mouse genome with high efficiency.

It was originally presumed that similar strategies could be used for genetic manipulations in human cells; however, this did not turn out to be the case due to key differences in the biology of hESCs and mESCs. The exact reasons for the low efficiency of gene targeting by HR in hESCs are not well understood but may include reduced HR activity or a higher propensity to use the NHEJ pathway. Indeed, it was reported that hESCs preferentially use NHEJ to rejoin DSBs induced by ionizing radiation, in contrast to mESCs that show a higher propensity for HR (Banuelos et al. 2008). This difference could in part be due to the reduced expression of the NHEJ proteins Ku70/80 and Lig4 in mESCs and indicates that the higher rate of HR in mESCs is partially caused by their reduced ability to use the NHEJ pathway (Banuelos et al. 2008). In addition, the substantial biological differences between mESCs and hESCs, such as the ability to grow as single cells, may profoundly affect gene targeting efficiency.

The difficulties in targeting hESCs by HR is illustrated by the observation that, so far, <20 genes have been successfully targeted in hESCs (Thomson et al. 1998; Zwaka and Thomson 2003; Urbach et al. 2004; Costa et al. 2007; Irion et al. 2007; Davis et al. 2008a,b; Suzuki et al. 2008; Ruby and Zheng 2009; Xue et al. 2009). The approach using a selection cassette without promoter that does not have a promoter, it is less likely to be expressed when aberrantly integrated into the genome. Therefore, the number of apparent random integrations is reduced and the identification of correctly targeted clones is simplified. Such gene trap approaches require active expression of the targeted gene, as is the case for OCT4 in hESCs. Therefore, this approach cannot be used to target genes specific to differentiated cells types such as lineage markers in human pluripotent cells.

Conventional gene targeting approaches have also been used to generate cell lines that report in real time the specific cell-fate decisions of hESCs during differentiation (Davis et al. 2008b; Ruby and Zheng 2009; Xue et al. 2009). Targeting of the Mixl1 (Davis et al. 2008b) and Olig2 (Xue et al. 2009) loci exemplify this experimental approach. A GFP reporter was inserted into the Mixl1 gene and used to identify and isolate cell populations enriched for primitive hematopoietic precursors. Similarly, Olig2, a transcription factor important for motor neuron and oligodendrocyte development, was targeted to generate a GFP reporter cell line. Following directed differentiation of Olig2-targeted hESCs, GFP expression was used to isolate cell types differentiating toward a neuronal or oligodendrocyte fate.

To establish a reliable system to ectopically express genes in hESCs, Irion et al. (2007) identified the human homolog of the mouse ROSA26 locus and established a gene targeting strategy for this locus. ROSA26 is frequently used as a “safe harbor” locus in the mouse because it can be targeted with high efficiency in mESCs and allows nearly ubiquitous expression without any reported detrimental effects when both alleles are disrupted (Zambrowicz et al. 1997). This locus has been successfully targeted in hESCs with an expression cassette for red fluorescent protein (RFP) (Irion et al. 2007). RFP was found in several different human cell types differentiated from these targeted hESCs, indicating that the ROSA26 locus might be suitable for transgene expression in hESCs and hESC-derived cells. However, beyond this original report, the human ROSA26 locus has thus far not been routinely used to express transgenes in hESCs.

Enhancing Gene Targeting by Alternative Forms of Homologous DNA Delivery

Gene targeting requires the delivery of ectopic DNA that can be used by cells as a template for HR. Generally, standard transfection and electroporation techniques are used to deliver targeting constructs in the form of plasmid DNA (for review, see Costa et al. 2007; Davis et al.
locus flanking the cut site can use this donor as a template to repair the DSB via a specialized form of the homologous DNA repair pathway called synthesis-dependent strand annealing (SDSA).

In SDSA (outlined in Fig. 1) (for review, see Hellday et al. 2007; Adelman and Boulton 2010), the DSB is resected by nucleases. The resulting single-strand DNA is coated by Rad51 and invades the provided donor plasmid at a homologous stretch of DNA. DNA repair polymerases then extend the single-strand DNA using the homologous strand as a template. During this step, lost DNA sequences are resynthesized or, if the exogenously applied donor contains additional sequence between the homology arms, then these get incorporated. Once DNA synthesis has generated a complementary sequence to the single-strand DNA on the other side of the DSB, the newly synthesized strand is displaced. Finally, the two strands of the DSB reanneal and synthesize DNA to fill in the gap, and DNA ligation rejoins the two strands. If the donor sequence diverges from the endogenous gene either by the addition of longer stretches of DNA or by a single point mutation, the resulting gene conversion can be detected and used to measure the frequency of gene targeting. In human cells, the rate of spontaneous gene targeting using such a construct in the absence of a DSB is low and usually occurs in less than one in one million cells. However, after SCE-I induces a DSB, targeting efficiency at the same locus and with the same donor construct increases by several orders of magnitude (Bibikova et al. 2003; Porteus and Baltimore 2003). This can result in gene targeting efficiencies of up to one in 10 cells undergoing a gene conversion (Bibikova et al. 2003; Porteus and Baltimore 2003).

After the nuclease has generated a DSB at the desired locus, the resection of the DSB is limited and results in approximately 30–100 bases of single-strand DNA (Elliott et al. 1998; Hellday et al. 2007; Moehle et al. 2007; Larocque and Jasin 2010). It is this single-strand region that is coated with RAD51 and used for homology search (Hellday et al. 2007; Adelman and Boulton 2010). As a result, only short homology arms are needed to mediate efficient gene targeting. However, the limited resection of the DSB after the nuclease cut also poses limitations because it leads to relatively short gene conversion tracts, and genetic manipulations are restricted to the region proximal to the DSB. This illustrates a fundamental difference between SDSA in human cells after a DSB and HR in mESCs, which allows genetic manipulations over long stretches of DNA.

The SCE-I nuclease has been instrumental in elucidating the details of this process; however, due its fixed 18-bp recognition site, SCE-I has limited use in gene targeting. To exploit the phenomenon that a DSB can facilitate the manipulation of genes within its proximity, site-specific nucleases called ZFNs were engineered (Bibikova et al. 2002; Porteus and Baltimore 2003). ZFNs are chimeric proteins composed of several Cys2His2 zinc fingers that form a zinc-finger DNA-binding domain, a structural motif found in a naturally occurring group of zinc-finger transcription factors (Kim and Chandrasegaran 1994; Kim et al. 1996, 1998). ZFNs are also comprised of a nuclease domain that is most commonly a

---

**ZINC-FINGER NUCLEASES**

In this section, we describe site-specific DNA nucleases (called ZFNs) that can be used to facilitate gene targeting in human cells in ways that offer several advantages over previous techniques.

An alternative strategy to enhance gene targeting efficiency came from the observation that the introduction of a DSB in a genomic region increased HR efficiency in the proximity of the break (Brenneman et al. 1996; Jasin 1996). This effect was initially described in experiments where a defined DSB was generated in cells using homing endonucleases. Homing endonucleases, such as the intron-encoded endonuclease I from *Saccharomyces cerevisiae* (SCE-I), are site-specific nucleases found in bacteria, phages, and lower eukaryotes. An important characteristic of these nucleases is that their recognition sequence is relatively large (the SCE-I site is 18 nucleotides) and therefore the mammalian genome contains only a limited number of such recognition sites. When SCE-I is expressed in cells that have been engineered to carry a single SCE-I recognition site, it binds to this site and cuts the DNA generating a DSB (Rouet et al. 1994). Cells that are provided with a “donor plasmid” that has homology to the donor id with a homology to the

---

2008a). However, alternative strategies such as the use of viruses or bacterial artificial chromosomes (BACs) to deliver targeting constructs have been developed to facilitate gene targeting in human cells.

One approach is to use adeno-associated viruses (AAVs), which can be genetically engineered to be non-integrating and maintained as episomal particles in the cell. Such AAVs have been used to deliver DNA constructs for gene targeting in human cells and can increase gene targeting efficiency to up to 1% of the infected cells being correctly targeted by HR (Russell and Hirata 1998). Examples of genes targeted with this approach in hESCs were *HPRT* and *HMGAI* (high-mobility-group AT-hook) (Khan et al. 2010). However, AAVs are commonly used to deliver targeting constructs used for a gene trap approach, because a selection cassette that carries its own promoter would be constitutively expressed from the AAV episome. Therefore, it remains to be determined whether AAVs will also prove useful for gene targeting of non-expressed genes in human pluripotent cells.

In addition to AAVs, BACs have also been used to deliver DNA to target genes in hESCs by HR (Song et al. 2010). Recombineering technology in bacteria can be used to genetically engineer BACs, which function similarly to conventional plasmids by carrying a selection cassette and homology arms. However, a key difference to regular targeting constructs is that BAC vectors can be generated with very long (>100 kb) homology arms. The increased size of the homology arms is thought to increase gene targeting efficiency in hESCs (Song et al. 2010). A BAC-based approach was used to disrupt both alleles of ATM (Khan et al. 2010), and the effect was instrumental in elucidating targeting can generally be applied to target genes in human pluripotent cells.
derivative of the FOK1 nuclease. Many aspects are considered in the design of a ZFN (Urnov et al. 2010) and different strategies have been developed (Greisman and Pabo 1997; Isalan et al. 2001; Hurt et al. 2003; Segal et al. 2003; Maeder et al. 2008; Meng et al. 2008; Foley et al. 2009; Sander et al. 2010). Briefly, each zinc finger in a zinc-finger DNA-binding domain specifically recognizes three base pairs while making only limited contact to the neighboring bases (Pavletich and Pabo 1993). This allows for the generation of combinatorial ZFN modules, a process called modular assembly (Segal et al. 2003), in which up to six zinc fingers with known binding activity are combined so that they specifically recognize a stretch of ~18 consecutive nucleotides.

The other key component of a ZFN is the nuclease domain, which is designed so that its activity is dependent on the dimerization of two half-sites (Vanamee et al. 2001; Miller et al. 2007b; Szczepak et al. 2007). This restricted enzymatic activity acts as a safeguard to prevent nonspecific DNA cleavage. Because an individual zinc finger contains only one FOK1 half-site, two ZFNs must bind with a defined spacing in order for the nuclease domain to become active. This increases the specificity of the genomic site that is recognized by a ZFN pair to up to 36 nu-
cleotides in length. After the assembly of a specific ZFN pair, its activity and cytotoxicity are generally tested by appropriate assays (Miller et al. 2007a) before it is used in gene targeting experiments.

The ZFN pair and donor plasmid containing homology to the locus of interest are introduced into the cell using standard transfection methods suitable for the individual cell type. The donor plasmid is designed to carry ~500–750 bp of homologous sequence, flanking both sides of the ZFN recognition site. If desired, a single base-pair change can be introduced around the zinc-finger recognition site in order to generate a point mutation. This could be useful to introduce or repair a disease-specific mutation in hiPSCs. Alternatively, large cassettes carrying multiple transgenes and promoters can be placed between the homology arms. Such transgene integration is a required step in the generation of reporter cell lines, where a reporter such as GFP is inserted into the genome at a specific location.

It is worth noting that ZFNs were successfully used to genetically modify model organisms that could not easily be genetically targeted with conventional techniques. Gene correction has been achieved in *Drosophila* (Bibikova et al. 2003; Beumer et al. 2006, 2008), *Arabidopsis* (Zhang et al. 2010), tobacco plants (Townsend et al. 2009), and mice (Goldberg et al. 2010). Precise gene insertion at the site of the ZFN cut was performed in corn (Shukla et al. 2009) and tobacco plants (Cai et al. 2009). Furthermore, ZFN-mediated gene disruption has been successful in zebrafish (Doyon et al. 2008; Meng et al. 2008; Foley et al. 2009), rats (Geurts et al. 2009; Mashimo et al. 2010), and *Drosophila* (Bibikova et al. 2002). In contrast to ZFN-mediated gene addition, which relies on HR with an exogenous donor, ZFN-mediated gene disruption results from DNA repair by the NHEJ pathway when no donor is supplied. As mentioned above, the NHEJ pathway is imprecise and rejoins chromosome ends without the need for homology, regenerating lost sequence information, and/or the removal of aberrantly added nucleotides. Therefore, sporadic gene disruption can occur when a ZFN cut is made in the absence of a donor plasmid, leading to loss-of-function alleles. If present in germ cells, this approach can be used to generate knockout animals (Geurts et al. 2009; Mashimo et al. 2010).

**ZFN-MEDIATED TARGETING IN HUMAN SOMATIC CELLS**

Several genes have been targeted in human somatic cells using ZFNs (Table 1). In a seminal paper, Urnov et al. (2005) described correction of a disease-relevant mutation in the *IL2Rγ* gene in human cells using ZFNs (Urnov et al. 2005). Mutations in this gene are responsible for X-linked severe combined immune deficiency (SCID). In a set of elegant experiments, these authors used ZFNs to introduce single out-of-frame point mutations into exon 5 of *IL2Rγ*. In a second targeting step, ZFNs were used to correct these point mutations and restore the protein to wild-type levels (Urnov et al. 2005).

In addition, *IL2Rγ* and other loci have been used to demonstrate that ZFNs can mediate transgene integration in a large number of human cell types, including primary cells (Lombardo et al. 2007; Moehle et al. 2007). This approach allows the standardized generation of human cell lines that express transgenes and short hairpin RNAs from defined genetic loci. ZFN-mediated transgene insertion can thereby eliminate some of the unpredictable complications generating transgenic cell lines by random integration, such as position effects, differences in copy number, or differential epigenetic silencing. We and other investigators have also recently adopted ZFN technology to genetically modify hESCs and hiPSCs (Lombardo et al. 2007; Brunet et al. 2009; Hockemeyer et al. 2009; Zou et al. 2009).

**TARGETING OCT4 AND PITX3 LOCI TO GENERATE REPORTER CELL LINES IN HUMAN PLURIPOTENT CELLS**

As mentioned above, *OCT4* had been previously targeted by conventional HR in hESCs (Zwaka and Thomson 2003). Using isogenic vector arms, targeting efficiencies were 20% with 7.9-kb homologous sequences and 40% with 12.8-kb homologous sequences. In contrast, the ZFN targeting approach we used required only short homology arms of 500–750 bp and achieved gene targeting efficiencies of up to 94% correctly targeted clones (Hockemeyer et al. 2009). In these experiments, correct targeting resulted in a reporter system that allowed the identification of *OCT4*-positive pluripotent cells by either expression of GFP or the puromycin-resistance gene. This demonstrates the advantages of ZFN-mediated targeting over conventional targeting: ZFNs require only short targeting arms but nevertheless result in substantially higher targeting efficiency. This is relevant, because a major concern and uncertainty in conventional gene targeting is the extent to which small nucleotide polymorphisms (SNPs) in targeting constructs reduce gene targeting efficiency. Examples of successful conventional targeting experiments did not explicitly use isogenic DNA to generate their targeting constructs, but a comprehensive analysis in hESCs of the impact of SNPs in targeting constructs is lacking. In addition, the optimal length of homology arms must be considered in the design of the targeting strategy. Targeting arms such as those used to conventionally target hESCs were relatively long, and this length might have been crucial to targeting success. However, to generate such long arms isogenically and isochromosomally, in order to avoid SNPs, is relatively difficult and time-consuming. ZFN-mediated targeting requires only very short targeting arms and thereby largely eliminates these concerns.

In addition to modifying the *OCT4* locus, we also used ZFNs to target *PITX3* (Hockemeyer et al. 2009), a gene encoding a transcription factor that is expressed in differentiated cells, such as dopaminergic neurons, but not in hESCs. This targeting strategy was designed to generate a GFP reporter for the transcriptional state of *PITX3* by fusing the first exon of *PITX3* with the open reading frame (ORF) of GFP. Targeting of *PITX3* is significant because it illustrates that ZFNs can be used to target genes that are not expressed in hESCs.
AAVS1 on chromosome 19 is a well-characterized locus that has been used to stably express transgenes in multiple transformed and primary cell lines without transgene silencing (Smith et al. 2008; Dekelver et al. 2010). AAVSI is the viral integration site for AAVs, which, following integration, disrupts the gene encoding the regulatory subunit 12C of protein phosphatase 1 (PPP1R12C). hESCs targeted in this locus using AAV-mediated gene integration exhibit long-term transgene expression and maintain a pluripotent state (Smith et al. 2008). To establish a robust overexpression system suitable for hESC cultures, we used a ZFN pair to target the first intron of PPP1R12C. Following the targeting of this locus with a constitutive expression cassette containing enhanced GFP, we observed long-term expression in hESCs, validating the use of ZFNs for modifying this locus. As mentioned above, an alternative approach for establishing controlled transgene expression is the isogenic targeting of the ROSA26 locus. The targeting frequency of the ROSA26 locus was relatively low, with only 2.3% of the drug-resistant clones being correctly targeted (Irion et al. 2007), whereas targeting efficiencies of up to 50% were achieved by the ZFN-mediated approach (Hockemeyer et al. 2009).

In addition to expressing enhanced GFP constitutively from the AAVSI locus, we have also introduced a doxycycline (DOX)-inducible expression cassette into the AAVSI locus using ZFNs (Hockemeyer et al. 2009). Using this system, we found that a transgene can be expressed in a quantitative and temporally controlled manner in response to different DOX concentration in the medium. This inducible system requires a DOX-responsive promoter that drives transgene expression and expression of the reverse tetracycline-controlled transactivator (rtTA). Using simultaneous targeting of both alleles of the AAVSI locus, we have established a one-step targeting strategy in which rtTA is inserted into one allele and the DOX-responsive promoter, and the transgene of interest is inserted into the

### Table 1. Gene Targeting Experiments Using ZFN in Human Cells

<table>
<thead>
<tr>
<th>Cell type</th>
<th>Genetic modification</th>
<th>Gene</th>
<th>Clonal targeting efficiency</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor cells</td>
<td>Gene disruption</td>
<td>TP73</td>
<td>Doyon et al. (2010)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MAP3K14</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EP300</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>BTK</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>TSC2</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>KDR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>NR3C1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HOXB13</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CXC4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CFXR4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCR5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL2RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene addition</td>
<td>PPP1R12C (AAVS1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene correction</td>
<td>IL2RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary human cells</td>
<td>Gene disruption, correction</td>
<td>CCR5</td>
<td>Urnov et al. (2005)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>IL2RG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene addition</td>
<td>PPP1R12C (AAVS1)</td>
<td></td>
<td>Moador et al. (2008)</td>
</tr>
<tr>
<td>hESCs/iPSCs</td>
<td>Gene disruption</td>
<td>IL2RG</td>
<td>Lombardo et al. (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCR5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIGA</td>
<td>Zou et al. (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Gene addition</td>
<td>IL2RG</td>
<td>Lombardo et al. (2007)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCR5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PIGA</td>
<td>Zou et al. (2009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>OCT4</td>
<td>up to ~90%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PITX3</td>
<td>8%–10%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PPP1R12C (AAVS1)</td>
<td>~50%</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Urnov et al. 2010.
other allele. Moreover, we have recently found that the \textit{AAVS1} locus can also be used to express short hairpin RNAs to confer stable, long-term gene knockdown in hESCs (Dekelver et al. 2010).

**OFF-TARGET SITE ALTERATIONS**

ZFN technology has the potential to be a robust and reliable tool for gene targeting in human pluripotent cells, but only a few genes have been targeted using ZFNs thus far. Nevertheless, these examples indicate that current strategies for zinc-finger design (Pabo et al. 2001; Klug 2005) might allow the generation of selective and active ZFNs for most genes in the human genome. However, the introduction of active nucleases into human cells necessitates some caution and consideration regarding potential off-target and genotoxic effects of ZFNs. For basic research, several tools are at hand to ensure specificity and quality of ZFNs. Southern blotting using internal and external probes should be used to exclude ectopic integrations and confirm clonality of targeted clones, although such analyses do not detect ZFN-mediated DSBs and subsequent error-prone repair elsewhere in the genome.

To address these concerns, we investigated the frequency of off-target DNA cleavage and subsequent NHEJ-mediated repair for all currently used ZFNs (Hockemeyer et al. 2009). We first analyzed the DNA-binding specificity for all ZFNs with SELEX (systematic evolution of ligands by exponential enrichment) (Perez et al. 2008; Phillips et al. 2009), which allowed identification of the most likely off-target cleavage sites on a genome-wide scale. Using the Cel-1 assay (Miller et al. 2007a), we then quantified the actual frequency of NHEJ-mediated alterations in identified potential off-target sites. We analyzed 46 genomic loci and detected only one NHEJ alteration at one genomic site in a \textit{PITX3}-targeted clone. All other putative off-target sites for all of the ZFNs tested were wild type. In addition, we determined the frequency of NHEJ alterations in heterozygous-targeted clones on the allele, which did not carry an integrated transgene. This frequency varied from 1 in 18 to 1 in 12 for \textit{PITX3} and \textit{OCT4} loci, respectively, whereas in all other cases, the other allele remained unaltered. To further 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 \textit{OCT4} and \textit{AAVS1} loci. In summary, the use of ZFN can in rare cases lead to off-target effects, but these can usually be detected and minimized. The use of multiple, independent ZFN-mediated manipulations to confirm cellular phenotypes should also alleviate these concerns.

**CONCLUDING REMARKS**

The combination of patient-specific hiPSCs together with ZFN technology for performing genetic modifications of the human genome will allow the generation of specific cell types carrying disease-relevant alterations. This would be a significant step toward the generation of better in vitro models for human disease and the development of potential therapeutics. Thus far, only a handful of genes in human pluripotent cells have been targeted using ZFNs. Although targeting efficiency has been very high in the genes targeted so far, it is not clear what fraction of the human genome can be engineered in this way and how the chromatin conformation of a given gene may influence targeting efficiency. It will be important to resolve these issues in future work to assess the general utility of this technology for genetically manipulating the genome of human pluripotent cells.

**ACKNOWLEDGMENTS**

We thank Helen Bateup for critical comments on this manuscript. D.H. is a Merck Fellow of the Life Science Research Foundation. R.J. was supported by National Institutes of Science grants R37-CA084198, RO1-CA087869, and RO1-HD045022. R.J. is an adviser to Stemgen and a cofounder of Fate Therapeutics.

**REFERENCES**


