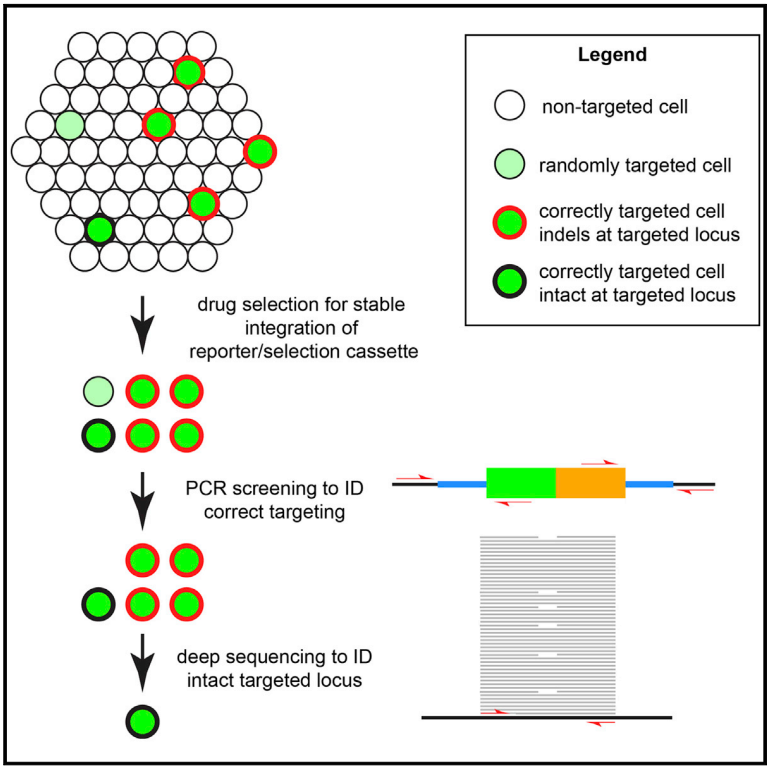
Report

Cell Reports

Efficient CRISPR-Cas9-Mediated Generation of Knockin Human Pluripotent Stem Cells Lacking Undesired Mutations at the Targeted Locus

Graphical Abstract



Authors

Florian T. Merkle,
Werner M. Neuhausser, ...,
Alexander F. Schier, Kevin Eggan

Correspondence

eggan@mcb.harvard.edu

In Brief

Gene editing with CRISPR-Cas9 holds great promise. Merkle et al. show highly efficient reporter-gene insertion at many genomic loci but observe common "on-target" mutations. These mutations can be effectively mitigated by rational targeting-strategy design and bioinformatic screening.

Highlights

- Cas9 dual nickase mediates efficient reporter-gene insertion in human stem cells
- Deleterious mutations at the targeted locus are unexpectedly common
- Rational design and bioinformatic selection enable efficient reporter generation

Merkle et al., 2015, Cell Reports *11*, 875–883
May 12, 2015 ©2015 The Authors
http://dx.doi.org/10.1016/j.celrep.2015.04.007

CellPress

Cell Reports
Report

OPEN ACCESS
CellPress

Efficient CRISPR-Cas9-Mediated Generation of Knockin Human Pluripotent Stem Cells Lacking Undesired Mutations at the Targeted Locus

Florian T. Merkle,[1,2,3,4,5,11] Werner M. Neuhausser,[1,2,5,6,7,11] David Santos,[1,2] Eivind Valen,[3,8] James A. Gagnon,[3] Kristi Maas,[1,2,5,6,7] Jackson Sandoe,[1,2,3] Alexander F. Schier,[3,5,9,10] and Kevin Eggan[1,2,3,4,5,*]
[1]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[2]The Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138, USA
[3]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA
[4]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[5]Harvard Stem Cell Institute, Cambridge, MA 02138, USA
[6]Division of Reproductive Endocrinology and Infertility, Department of Obstetrics and Gynecology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02215, USA
[7]Boston IVF, Waltham, MA 02451, USA
[8]Computational Biology Unit, Department of Informatics, University of Bergen, 5020 Bergen, Norway
[9]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA
[10]Center for Brain Science, Harvard University, Cambridge, MA 02138, USA
[11]Co-first author
*Correspondence: eggan@mcb.harvard.edu
http://dx.doi.org/10.1016/j.celrep.2015.04.007
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

SUMMARY

The CRISPR-Cas9 system has the potential to revolutionize genome editing in human pluripotent stem cells (hPSCs), but its advantages and pitfalls are still poorly understood. We systematically tested the ability of CRISPR-Cas9 to mediate reporter gene knockin at 16 distinct genomic sites in hPSCs. We observed efficient gene targeting but found that targeted clones carried an unexpectedly high frequency of insertion and deletion (indel) mutations at both alleles of the targeted gene. These indels were induced by Cas9 nuclease, as well as Cas9-D10A single or dual nickases, and often disrupted gene function. To overcome this problem, we designed strategies to physically destroy or separate CRISPR target sites at the targeted allele and developed a bioinformatic pipeline to identify and eliminate clones harboring deleterious indels at the other allele. This two-pronged approach enables the reliable generation of knockin hPSC reporter cell lines free of unwanted mutations at the targeted locus.

INTRODUCTION

Human pluripotent stem cells (hPSCs), including human embryonic and induced pluripotent stem cells (hESCs and hiPSCs), have the potential to give rise to any cell type in the body, including those affected in disease (Takahashi et al., 2007; Thomson et al., 1998). As a consequence, hPSCs can be used to generate cell-based in vitro disease models, chemical screens, and cellular therapies (Bellin et al., 2012; Egawa et al., 2012; Grskovic et al., 2011; Takahashi et al., 2007; Thomson et al., 1998). The utility of hPSCs for each of these applications can be significantly enhanced by genome engineering, the process of precisely inserting, deleting, or substituting specific genomic sequences (Merkle and Eggan, 2013; Sandoe and Eggan, 2013). In particular, the insertion of reporter genes into developmentally important loci such as *FEZF2* (Ruby and Zheng, 2009), *OLIG2* (Xue et al., 2009), or *NKX2-1* (Goulburn et al., 2011) has facilitated the identification and characterization of target cell types for disease modeling or transplantation. Similarly, disease-associated mutations have been introduced or corrected by this approach to explore the cellular phenotypes arising from these variants (Kiskinis et al., 2014; Reinhardt et al., 2013; Schwank et al., 2013).

Gene targeting has been notoriously difficult in hPSCs because these cells do not efficiently repair their DNA by homologous recombination (HR) (Liu and Rao, 2011; Zwaka and Thomson, 2003). This limitation can be addressed by the targeted introduction of DNA double-strand breaks (DSBs) at the locus of interest, which leads to an approximately 1,000-fold increase in the frequency of HR near the DSB (Rouet et al., 1994; Elliott et al., 1998). A DSB can be targeted to sites of interest in mammalian cells by designer endonucleases such as the clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 system (Barrangou et al., 2007; Jinek et al., 2012; 2013; Cho et al., 2013). Targeting of defined loci by CRISPR-Cas9 is mediated by a variable 20-base guide RNA that pairs with a cDNA sequence and by the Cas9 protein, which cleaves the DNA if the CRISPR target sequence is followed by a PAM motif (Mojica et al., 2009; Garneau et al., 2010; Wiedenheft et al., 2012).

In addition to cutting at its target site, Cas9 nuclease has been reported to cut at imperfectly matched sequences

elsewhere in the genome, leading to the formation of "off-target" insertions and deletions (indels) via the error-prone process of non-homologous end joining (NHEJ) (Fu et al., 2013). These off-target indels can be reduced by improved bioinformatic design based on analysis of CRISPR-Cas9 binding specificity (Hsu et al., 2013; Pattanayak et al., 2013; Cho et al., 2014; Fu et al., 2014; Yang et al., 2014a) and by the implementation of a mutant "nickase" variant of Cas9 (Cas9-D10A or Cas9n; Cong et al., 2013; Mali et al., 2013; Cho et al., 2014). Targeting dual Cas9 nickases (Cas9 dn) to opposite DNA stands with separate CRISPR guides leads to efficient DSB formation with 50- to 1,500-fold fewer off-target indels than Cas9 nuclease (Ran et al., 2013).

CRISPR-Cas9 has been shown to mediate reporter gene knockin at a handful of genomic loci in hPSCs (Hou et al., 2013; Mali et al., 2013; Byrne et al., 2014; González et al., 2014), but the benefits and drawbacks of genome engineering using CRISPR-Cas9, -Cas9 dn, or -Cas9n targeting strategies have not been thoroughly explored. We therefore systematically tested the ability of these three gene-targeting strategies to induce HR at 16 genomic loci in hPSCs. We report that whereas CRISPR-Cas9-mediated gene knockin was efficient, undesired indel mutations were common at both alleles of the targeted locus. To overcome this issue, we designed a Cas9-dn-targeting strategy to physically separate the two CRISPR target sites by gene insertion to mitigate indels at the targeted allele and designed a bioinformatic pipeline to identify and eliminate hPSC clones harboring on-target indels at the non-targeted allele. Together, these tools provide a framework for the efficient generation and identification of hPSC knockin clones free of on-target indels.

RESULTS

Design of CRISPR-Cas9 and Gene-Targeting Constructs

We tested the ability of distinct CRISPR-Cas9-targeting strategies to mediate the insertion of ~3-kb reporter/selection cassettes flanked by ~750-bp sequences with homology to the targeted locus (Figure S1). To ensure the concomitant delivery of CRISPR guides and Cas9 protein, we cloned CRISPR guide sequences into expression plasmids encoding both a chimeric guide/tracer sequence and either human-codon-optimized Cas9 nuclease or Cas9-D10A nickase as previously described (Cong et al., 2013). We selected CRISPR guide sequences with minimal predicted off-target binding activity (Montague et al., 2014), and in most cases, we were able to identify appropriate guides that had predicted DNA cleavage sites within 50 bp of the desired insertion (Table S1; Figure S2). We designed dual nickases to generate 5′ overhangs of 30–50 bp (Ran et al., 2013) flanking the insertion site (Figure S2).

Efficiency of CRISPR-Cas9-Mediated Gene Targeting in hPSC

We investigated the relative abilities of Cas9-, Cas9-dn-, or Cas9n-targeting strategies to mediate gene targeting in hPSCs. Because the activity of CRISPR-Cas9 may be locus and sequence dependent, we targeted 16 genomic loci. After transfection of hPSCs with CRISPR-Cas9 expression vectors and gene-targeting plasmids, they were cultured in the presence of Geneticin (G418). We then observed the emergence of Geneticin-resistant clones that maintained hPSC-like morphology, growth rates, and markers of pluripotency (Figure S1). We manually picked and assayed 1,379 clones across the 16 loci for the insertion of the reporter/selection cassette.

To screen for targeted gene insertion, we PCR amplified across the 5′ and 3′ homology arms using locus- and insertion-specific primers (Figure S1; Table S1). Targeted clones (971/1,379 clones; 70%) yielded strong, distinct amplicons of the expected size that were not seen in non-targeted control hPSCs (Figure S1). PCR analysis suggested that the remaining clones (408/1,379 clones; 30%) carried a randomly integrated drug resistance cassette and were not examined further. Most targeted clones gave both 5′ and 3′ PCR products (900/971 clones; 93%) rather than just a 5′ or 3′ amplicon (71/971 clones; 7%), suggesting insertion of both ends of the reporter/selection cassette. We confirmed the fidelity of the PCR screening method and the absence of additional genomic integrations in correctly targeted clones by Southern blotting (Figure S1).

Strikingly, all three CRISPR-Cas9 knockin strategies yielded correctly targeted clones, in some cases at nearly 100% efficiency (Figure 1A; Table 1). In contrast, we did not observe targeted clones in control experiments in which CRISPR-Cas9 was omitted (n = 4 experiments). Because a previous report failed to identify HR in hPSCs upon Cas9n administration (Hsu et al., 2013), we were surprised to observe that Cas9n mediated HR at 7/12 sites. The failure of Cas9n-mediated knockin at five other sites suggests that the activity of Cas9n may be sequence and/or locus dependent, as described in previous studies of single-strand-break-mediated HR in human cells (Metzger et al., 2011). We observed a broad range of targeting efficiencies for each Cas9-targeting strategy at different loci and a trend for average gene-targeting rates to be highest for Cas9 (78% ± 10%), intermediate for Cas9 dn (49% ± 9%), and lowest for Cas9n (15% ± 5%; Figure 1B).

The different targeting efficiencies we observed could have been due to the different Cas9-based targeting strategies employed, to the different sequences of CRISPR guides used, or to site-specific effects. To distinguish between these possibilities, we repeated the targeting experiments and held all variables constant including guide sequences, but varied the Cas9-, Cas9-dn-, or Cas9n-targeting strategy. We observed that the relative targeting efficiency of Cas9 dn was similar to that of Cas9 at the tested sites (0.96 ± 0.16) but that Cas9n produced a significantly lower percentage ($p < 0.05$) of correctly targeted clones than either Cas9 (0.43 ± 0.13) or Cas9 dn (0.22 ± 0.08; Figure 1C).

To explore the rate of CRISPR-Cas9-mediated gene knockin, we plotted the number of correctly targeted clones obtained for each targeted site and Cas9 strategy against the number of assayed drug-resistant clones (Figure 1D) and found a strong correlation between these values, with R^2 values of 0.91, 0.70, and 0.81 for Cas9, Cas9 dn, and Cas9n strategies, respectively. Linear regression curves had X-intercepts ranging from 6 to 15, consistent with a modest background level of drug-resistant

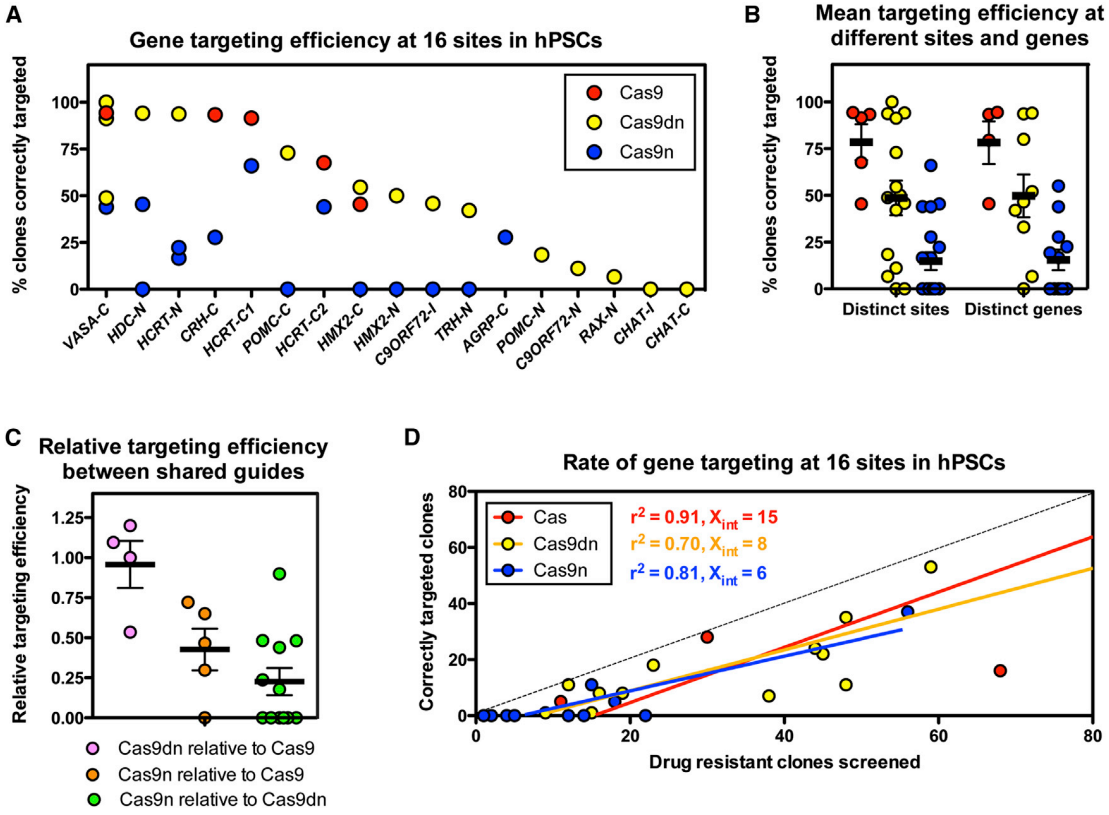


Figure 1. CRISPR-Cas9-Mediated Gene Targeting Efficiency in hPSCs

(A) Percentage of analyzed drug-resistant clones that gave PCR products of the expected size across both the 5′ and 3′ homology arms for each locus and Cas9-targeting strategy tested. Each dot represents a distinct combination of CRISPR sequence and targeted locus for a given Cas9 strategy. Clone numbers and targeting efficiencies are listed in Table 1.

(B) Summary of targeting efficiencies for each Cas9 strategy at each locus and CRISPR combination (left) or at distinct genes (right).

(C) Relative targeting efficiencies observed when using identical guide sequences but different Cas9-based strategies.

(D) Relationship between number of clones analyzed and number of clones correctly targeted for each Cas9-based strategy. Error bars in (B) and (C) denote SEM.

See also Figure S1.

colonies arising from the random integration of the targeting vector.

Correctly Targeted hPSC Clones Often Carry Undesired Mutations

CRISPR-Cas9-induced DSBs can be repaired by HR or by the error-prone NHEJ pathway, so hPSC clones that had undergone HR might also carry NHEJ-induced indels at one or both alleles of the targeted locus. Such indels are concerning, because they would likely disrupt endogenous or reporter gene function. Although CRISPR-Cas9 is known to form such "on-target" indels, their incidence in knockin reporter hPSCs is difficult to predict given the dearth of data on the insertion multi-kilobase constructs in hPSCs and the low efficiency of HDR in hPSCs relative to more extensively studied systems (Zwaka and Thomson, 2003; Liu and Rao, 2011).

To screen for on-target indels in correctly targeted hPSC clones, we PCR amplified across CRISPR target sites in both alleles and sequenced these amplicons to a median depth of 14,015 unique paired-end reads (Figures 2A, 2B, and S4A). We then calculated the indel frequency, defined as the percentage of aligned amplicons that contained indels near the CRISPR target site. This analysis showed that targeted hPSC clones carried a surprisingly high burden of indels, averaging 18% ± 3% per clone, at both the untargeted allele (Figures 2C and 2D) and the targeted allele (Figures 2E and 2F). These on-target indels were observed for all Cas9-targeting strategies (Figure 2D) and 13/13 analyzed sites. At the untargeted allele, the mean indel frequency was highest for Cas9 (37% ± 8%), intermediate for Cas9 dn (26% ± 7%), and lowest for Cas9n (4% ± 1%; Figure 2D). We found that the incidence of indels was distributed over a wide range of frequencies and that individual clones often carried multiple unique indels (alleles; Figure 2G), whose frequency within a given hPSC clone followed a roughly exponential distribution (Figure 2H). These results are unlikely to arise from the physical convergence of two or more drug-resistant clones, because the density of clones obtained in our experiments (~1 clone/cm^2) was well below commonly used clonal plating densities.

Table 1. Efficiency of hPSC Gene Knockin Mediated by Cas9, Cas9 dn, or Cas9n

	Relative Targeting Efficiency (Targeted Clones/Drug-Resistant Clones)			Absolute Targeting Efficiency (Targeted Clones/Transfected hESC)		
Locus	Cas9	Cas9 dn	Cas9n	Cas9	Cas9 dn	Cas9n
VASA-C	131/139[a] (94%)	171/199[a] (86%)	47/107[a] (44%)	-	-	-
HDC-N		32/34 (94%)	15/39 (38%)		1×10^{-5}	3×10^{-6}
HCRT-N		119/127 (94%)	3/15 (20%)		2×10^{-5}	6×10^{-7}
CRH-C	28/30 (93%)		5/18 (28%)	1×10^{-5}		1×10^{-6}
HCRT-C1	141/154 (92%)		37/56 (64%)	3×10^{-5}		7×10^{-6}
POMC-C		35/48 (73%)	0/18 (0%)		1×10^{-5}	0
HCRT-C2	16/68 (24%)		11/25 (44%)	6×10^{-6}		4×10^{-6}
HMX2-C	5/11 (45%)	24/44 (55%)	0/1 (0%)	2×10^{-6}	1×10^{-5}	0
HMX2-N		8/16 (50%)	0/5 (0%)		3×10^{-6}	0
C9ORF72-I		11/48 (23%)	0/9 (0%)		4×10^{-6}	0
TRH-N		8/19 (42%)	0/34 (0%)		3×10^{-6}	0
AGRP-C			3/18 (16%)			6×10^{-7}
POMC-N		7/38 (18%)			1×10^{-6}	
C9ORF72-N		1/9 (11%)			4×10^{-7}	
RAX-N		1/15 (7%)			4×10^{-7}	
CHAT-I		0/12 (0%)			0	
CHAT-C		0/23 (0%)			0	
Attempts	5/5 (100%)	11/13 (85%)	7/12 (58%)			
All clones	351/402 (87%)	428/632 (68%)	121/345 (35%)	1×10^{-5}	6×10^{-6}	2×10^{-6}

For each locus, the number of correctly targeted clones (numerator) and analyzed drug-resistant clones (denominator) are indicated. The ratio between these values gives the relative targeting efficiency, whereas the number of correctly targeted clones obtained per transfected hESC gives an estimate of the absolute targeting efficiency. Locus nomenclature is described in Figures S1C–S1E. In many instances, the same locus was targeted with the same CRISPR guide sequence and targeting vector but with either Cas9, Cas9 dn, or Cas9n, permitting the relative targeting efficiencies mediated by these Cas9 strategies to be compared (Figure 1C). See also Figures 1 and S1.

[a]It was not possible to calculate absolute targeting efficiencies at the *VASA* locus, because not all drug-resistant clones were screened. Mean absolute targeting efficiencies were calculated using the remaining loci.

Instead, our data support a model in which indels accumulate after gene targeting by the continued activity of plasmid-encoded CRISPR-Cas9 (Figure 2I).

Mitigation of On-Target Indels in Targeted hPSC Clones

The heterogeneous nature of the clones we observed suggested that restricting the time that CRISPR-Cas9 is active might reduce on-target indels. We therefore targeted hPSCs with a drug-inducible Cas9 plasmid and observed that indel frequencies were indeed reduced, but not eliminated (Figures S3B–S3D). We therefore asked whether targeting strategies that physically destroy or separate CRISPR target sites at the targeted allele could effectively eliminate on-target indels, as has been previously reported in other contexts (Chen et al., 2011; Yang et al., 2013, 2014b). Indeed, we observed negligible indel frequencies (0.08% ± 0.02%) at sites at which the CRISPR site was destroyed by HR, regardless of the Cas9 strategy used to mediate targeting (Figures S2 and S3E–S3G; Table S1). Next, we analyzed experiments in which gene insertion physically separated the two Cas9 dn CRISPR target sites (Figure 3A) and observed negligible indel frequencies (0.5% ± 0.4%) at eight distinct genomic sites (Figures 3B and 3C). These results indicate that undesired on-target indels can be effectively mitigated at the targeted allele in knockin hPSCs by the rational design of the targeting strategy.

Identification of Knockin hPSCs Free of Undesired On-Target Mutations

Whereas it was possible to mitigate on-target indels at the targeted allele, the untargeted allele remained vulnerable to continued CRISPR-Cas9 activity. We reasoned it should be possible to use the deep-sequencing pipeline we developed to identify and discard clones that had acquired on-target indels. We first examined the relationship between gene targeting efficiency and the frequency of indel formation at the untargeted allele and found these metrics were only weakly correlated ($R^2 < 0.5$), suggesting that gene targeting might not invariably be accompanied by on-target indels (Figure 3D). To test this hypothesis, we compared the gene targeting efficiency to the percentage of clones that had indel frequencies of 2% or less (intact) at the untargeted allele. We were encouraged to find that almost all (18/19; 95%) tested conditions yielded at least one hPSC clone that was both correctly targeted and intact at the untargeted allele (Figure 3E). When we extended this analysis to identify clones that were intact (indel frequencies of 2% or less) at both alleles, we observed such intact knockin hPSCs at 11/12 sites targeted using strategies that destroyed or physically separated CRISPR target sites at the targeted allele (Figures 3C, 3F, and S4D) but at only 5/9 sites in which the CRISPR site was left intact at both alleles (Figure 3F). Because targeting efficiencies

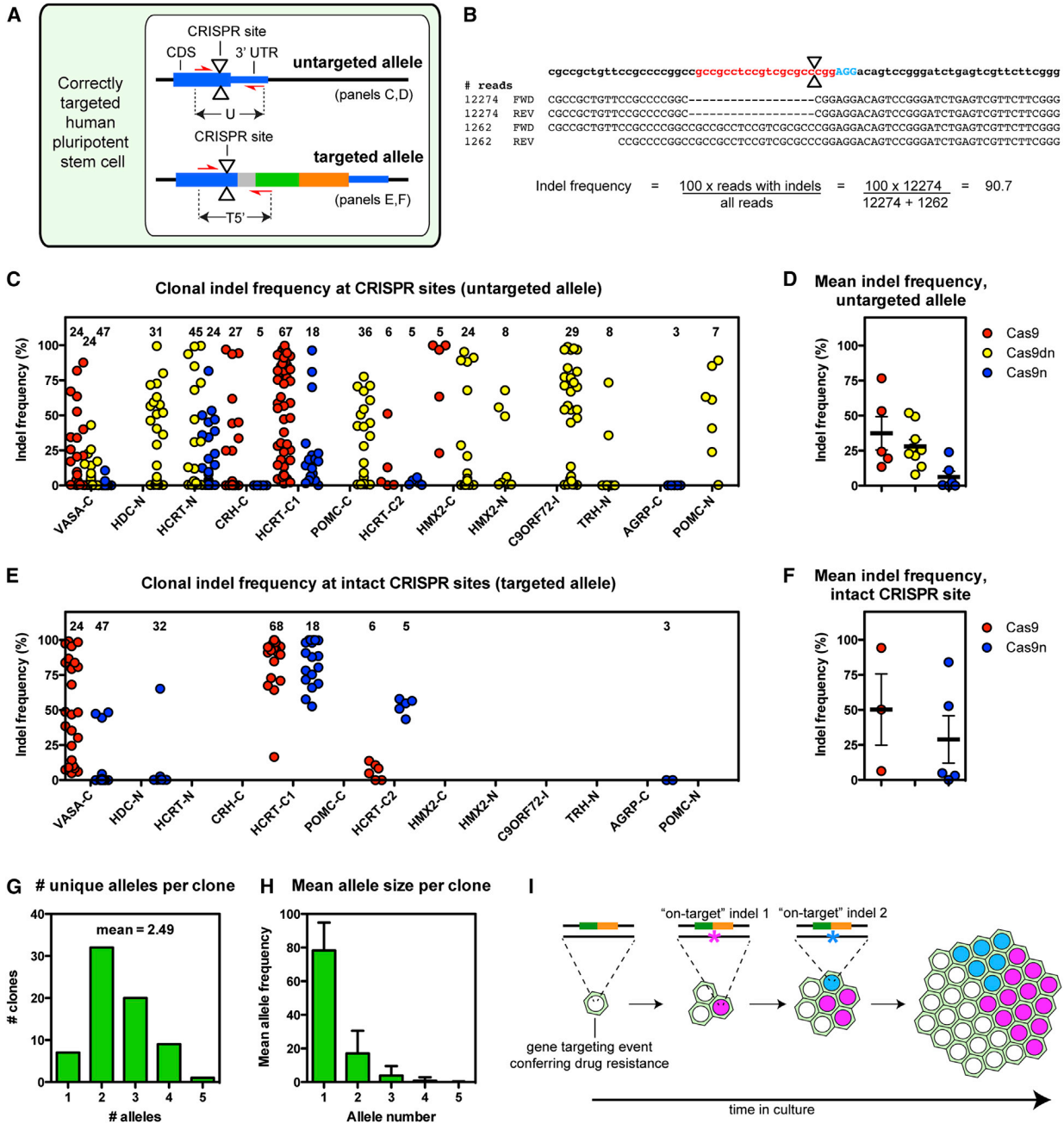


Figure 2. Frequency of CRISPR-Cas9-Induced On-Target Indels

(A) Schematic of a correctly targeted human pluripotent stem cell clone showing untargeted and targeted alleles, CRISPR-Cas9 cleavage sites (triangles), and the sequenced amplicons.

(B) Schematic of deep-sequencing data, showing the genomic locus with the CRISPR guide sequence (red), PAM motif (blue), predicted cleavage site (triangles), and the calculation of the indel frequency.

(C) The indel frequency at the untargeted allele of correctly targeted hPSC clones (number of sequenced clones indicated) is plotted for each targeted locus and Cas9 strategy.

(D) Mean indel frequencies at the untargeted allele of correctly targeted hPSC clones.

(E and F) Analysis as in (C) and (D) but performed at targeted alleles with intact CRISPR target sites. See also Figure S2.

(G) Number of intact and unique indel alleles (sub-clones) per correctly targeted hPSC clone, examined at the untargeted allele of the HCRT-C locus.

(H) Average allele frequency (sub-clone size) in hPSC clones.

(I) Schematic diagram illustrating the hypothesis of heterogeneous clone formation.

Error bars in (D), (F), and (H) denote SEM. See also Figure S2.

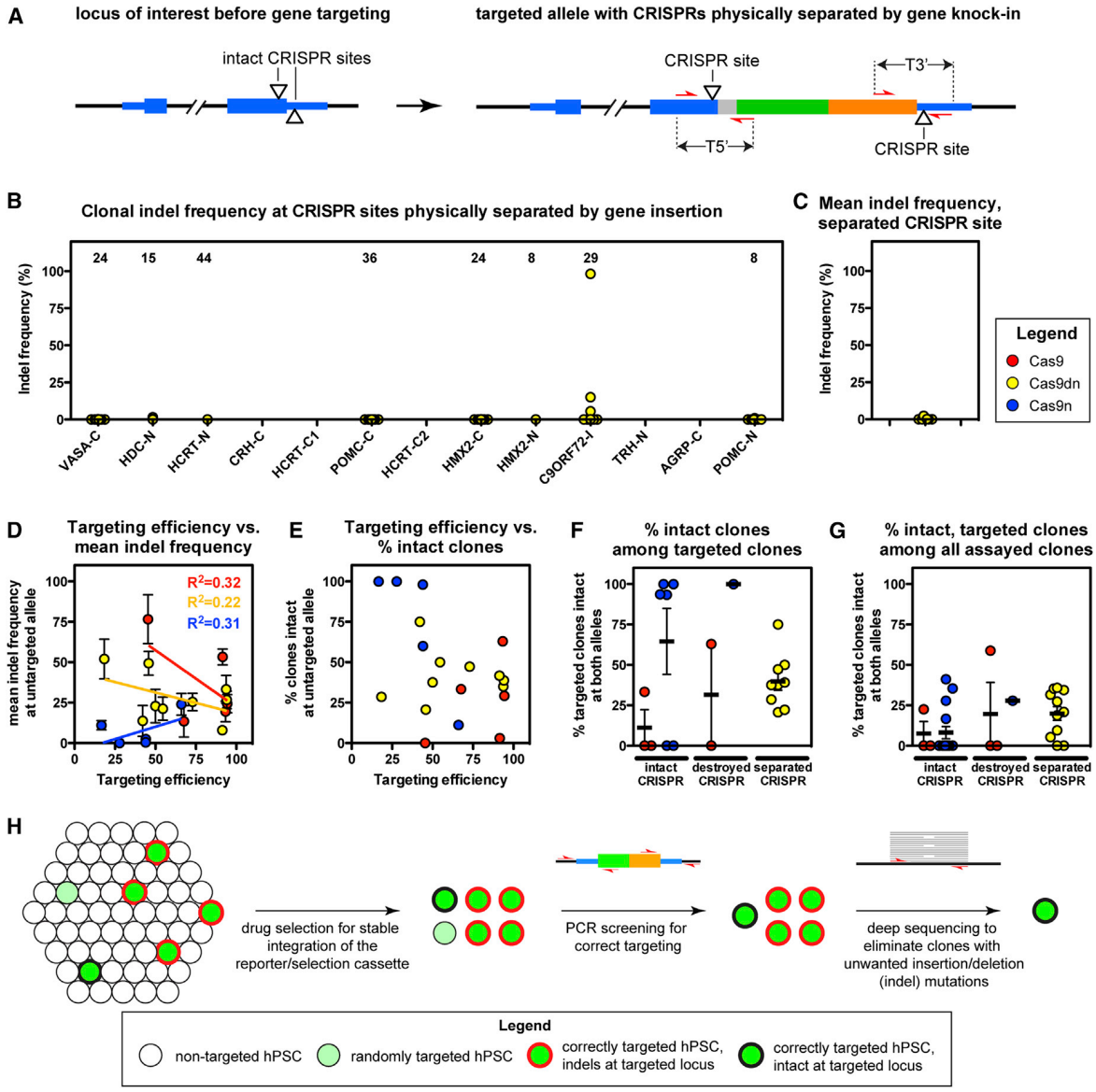


Figure 3. Strategy for Generating Intact Knockin hPSC Lines

(A) Schematic of a prototypical targeted allele with CRISPR target sites physically separated by gene insertion. Cas9 dn cleavage sites (triangles), 5′ amplicon (T5′), and 3′ amplicon (T3′) across 5′ and 3′ CRISPR target sites of the Cas9 dn strategy are illustrated.

(B and C) Indel frequencies at the targeted allele of Cas9-dn-targeted clones shown for each clone (B) or as the mean indel frequency for each locus (C).

(D) Mean indel frequencies at the untargeted allele of correctly targeted hPSC clones plotted against the targeting efficiency for each locus.

(E) Percentage of clones with indel frequencies of 2% or less at the untargeted allele plotted against targeting efficiency for each locus.

(F) Percentage of correctly targeted clones intact at both the targeted and untargeted allele.

(G) Percentage of correctly targeted and intact clones among all drug-resistant clones screened.

(H) Workflow for identifying clones of knockin hPSC clones free of unwanted mutations at the targeted locus.

Error bars in (D), (F), and (G) denote SEM. See also Figure S3.

varied over a wide range (0%–94%; Table 1), we calculated the percentage of all Geneticin-resistant clones that were correctly targeted and intact at both alleles (Figure 3G). This analysis revealed that destroying or physically separating the CRISPR target site approximately doubled the frequency of correctly targeted and intact colonies (20% ± 5% of drug-resistant colonies) relative to targeting strategies that left the CRISPR target site intact at both alleles (8% ± 3%).

Whereas it was not always possible to identify CRISPR target sites that would be destroyed by gene insertion, it was straightforward to design Cas9 dn CRISPR target sites to be physically separated by gene insertion (9/9 conditions; Figures 3F and S2).

This finding is particularly relevant in light of recent studies showing that off-target indels in human cell lines are frequently caused by CRISPR-Cas9 but almost never by Cas9 dn (Frock et al., 2015; Kim et al., 2015; Tsai et al., 2015; Wang et al., 2015). We conclude that knockin hPSCs lines can be reliably generated by combining a Cas9-dn-mediated gene-targeting strategy with bioinformatic analysis of the targeted locus to identify clones free of unwanted indels (Figure 3H).

DISCUSSION

A major advantage of hPSCs is their ability to be differentiated into disease-relevant somatic cell types for use in disease modeling or therapeutic transplantation. CRISPR-Cas9-mediated reporter knockin can enable the identification and purification of these cell types of interest. Although edited hPSC clones have been screened for off-target indels induced by CRISPR-Cas9 at genomic regions other than the targeted site (Smith et al., 2014; Suzuki et al., 2014; Veres et al., 2014), the issue of undesired on-target indels at the targeted site has not been described in detail in the context of CRISPR-Cas9-mediated gene knockin in hPSCs. Genes that are targeted for reporter knockin are often preferentially expressed and are therefore likely to participate in biological functions important for that cell type. The disruption of targeted genes by CRISPR-Cas9-induced indels might therefore fundamentally alter the biology of the cell type of interest. Furthermore, these mutations cannot be removed from the genetic background of hPSCs by the independent segregation of alleles in the germline, as would be possible in animal models.

Although we found that rational targeting strategy design could dramatically reduce the incidence of on-target indels at the targeted allele, we were surprised to find that on-target indels affected the majority of knockin hPSC clones at the other allele. We addressed this issue by developing a bioinformatic pipeline to identify and eliminate clones that might be unsuitable for downstream applications. This pipeline is inexpensive and readily accessible to the non-specialist laboratory. By combining rational targeting strategy design, efficient transfection and hPSC culture conditions, and bioinformatic screening, we generated hPSC clones that were both correctly targeted and free of on-target indels at nine out of nine analyzed genomic sites. Delivery of CRISPR-Cas9 as RNA or protein might further reduce the frequency of unwanted clones. Taken together, the tools reported here will enable hPSCs to be more fully harnessed for disease modeling, cell transplantation, and drug screening.

EXPERIMENTAL PROCEDURES

Generation of Gene Targeting and CRISPR-Cas9 Constructs

CRISPR guide RNA and Cas9 or Cas9-D10A were encoded in bicistronic expression plasmids pX330 or pX335, respectively (Cong et al., 2013; Addgene no. 42230 and no. 42335). CRISPR guide sequences were designed using the web resources available at http://www.genome-engineering.org and https://chopchop.rc.fas.harvard.edu/. Targeting constructs for genome editing were generated via Gibson assembly of four PCR products: the pDTA-TK plasmid backbone (Addgene no. 22677), the 5′ and 3′ homology arms (500–1,000 bp), and the reporter/selection cassette (Figure S1). Plasmids were isolated using an endotoxin-free Midiprep kit (QIAGEN), Sanger sequenced, and used at a concentration of >1 μg/μl.

hPSC Culture and Transfection

Human ESC and iPSC lines (hPSC) were maintained on Matrigel-coated plates in a 1:1 mix of mTeSR1 (StemCell Technologies) and mouse embryonic fibroblast (MEF)-conditioned hESC medium (MEFcm), which was changed daily. CRISPR-Cas9 and gene-targeting plasmids were introduced to hPSCs by NEON transfection (Life Technologies) following manufacturer's instructions. 120 μl of cell suspension (2.5–5 × 10^7 cells/ml) was incubated with 4 μg targeting plasmid and 1 μg CRISPR plasmid (pX330 and pX335) for Cas9 or Cas9n targeting, or 0.5 μg of each PX335 plasmid for Cas9 dn targeting, and transfected in 100 μl NEON tips (1,600 V, 20 ms, and one pulse).

Drug Selection and Clonal Expansion

hPSC clones that had stably integrated the targeting vector were identified by selection with 50 μg/ml Geneticin (G418 sulfate). Drug-resistant clones were manually picked into 96-well plates and expanded for genomic DNA extraction and continued culture. Genomic DNA was isolated by digestion with 50 μl of Direct PCR tail lysis buffer (Viagen) containing 30 μl/ml PCR-grade Proteinase K (Roche) overnight at 55°C.

Identification of Targeted Clones

Targeting efficiency was quantified by PCR screening of genomic DNA from drug-resistant clones. Primer sequences are given in Table S1. Amplicon identity and gene targeting was confirmed by deep sequencing. Southern blotting was performed on a subset of clones to confirm reporter/selection cassette insertion at the targeted locus and the absence of this cassette elsewhere in the genome.

Deep Sequencing

To assay for unwanted indels in correctly targeted hPSC clones, PCR amplicons from both the targeted and untargeted alleles of the edited locus were deep sequenced with 150-bp paired-end reads on the MiSeq Personal Sequencer (Illumina) as previously described (Gagnon et al., 2014). Indel frequencies were determined by quantifying aligned reads containing insertions or deletions 1 bp or larger.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at http://dx.doi.org/10.1016/j.celrep.2015.04.007.

AUTHOR CONTRIBUTIONS

F.T.M. and W.M.N. designed the study and wrote the manuscript under the guidance of A.F.S. and K.E. F.T.M. developed the gene-targeting and screening strategies with the help of J.S. F.T.M. and W.M.N. carried out gene-targeting experiments. D.S. assisted with construct design, PCR screening, and preparation of samples for deep sequencing. F.T.M. performed immunostaining and analyzed deep-sequencing data with the help of E.V. and J.A.G., who developed the workflow and computational tools for assessing indel frequencies and assisted with data interpretation. F.T.M., W.M.N., and K.M. performed Southern blots.

ACKNOWLEDGMENTS

We thank Feng Zhang and Neville Sanjana for their helpful advice regarding CRISPR design and Belinda von Niederhäusern for technical assistance with developing the genome editing workflow. F.T.M., A.F.S., and K.E. are supported by grants from the NIH (1R21NS071598, 5K99NS083713, HL109525, and 5P01GM099117). W.M.N. is supported by a grant from the New England Fertility Society. J.A.G. is supported by a fellowship from the American Cancer Society. E.V. is supported by a grant from the Bergen Research Foundation. K.E. is an investigator with the Howard Hughes Medical Institute.

Received: November 8, 2014
Revised: December 8, 2014
Accepted: April 1, 2015
Published: April 30, 2015

REFERENCES

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science *315*, 1709–1712.

Bellin, M., Marchetto, M.C., Gage, F.H., and Mummery, C.L. (2012). Induced pluripotent stem cells: the new patient? Nat. Rev. Mol. Cell Biol. *13*, 713–726.

Byrne, S.M., Mali, P., and Church, G.M. (2014). Genome editing in human stem cells. Methods Enzymol. *546*, 119–138.

Chen, G., Gulbranson, D.R., Hou, Z., Bolin, J.M., Ruotti, V., Probasco, M.D., Smuga-Otto, K., Howden, S.E., Diol, N.R., Propson, N.E., et al. (2011). Chemically defined conditions for human iPSC derivation and culture. Nat. Methods *8*, 424–429.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat. Biotechnol. *31*, 230–232.

Cho, S.W., Kim, S., Kim, Y., Kweon, J., Kim, H.S., Bae, S., and Kim, J.-S. (2014). Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. Genome Res. *24*, 132–141.

Cong, L., Ran, F.A., Cox, D., Lin, S., Barretto, R., Habib, N., Hsu, P.D., Wu, X., Jiang, W., Marraffini, L.A., and Zhang, F. (2013). Multiplex genome engineering using CRISPR/Cas systems. Science *339*, 819–823.

Egawa, N., Kitaoka, S., Tsukita, K., Naitoh, M., Takahashi, K., Yamamoto, T., Adachi, F., Kondo, T., Okita, K., Asaka, I., et al. (2012). Drug screening for ALS using patient-specific induced pluripotent stem cells. Sci. Transl. Med. *4*, ra104.

Elliott, B., Richardson, C., Winderbaum, J., Nickoloff, J.A., and Jasin, M. (1998). Gene conversion tracts from double-strand break repair in mammalian cells. Mol. Cell. Biol. *18*, 93–101.

Frock, R.L., Hu, J., Meyers, R.M., Ho, Y.-J., Kii, E., and Alt, F.W. (2015). Genome-wide detection of DNA double-stranded breaks induced by engineered nucleases. Nat. Biotechnol. *33*, 179–186.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. Nat. Biotechnol. *31*, 822–826.

Fu, Y., Sander, J.D., Reyon, D., Cascio, V.M., and Joung, J.K. (2014). Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nat. Biotechnol. *32*, 279–284.

Gagnon, J.A., Valen, E., Thyme, S.B., Huang, P., Ahkmetova, L., Pauli, A., Montague, T.G., Zimmerman, S., Richter, C., and Schier, A.F. (2014). Efficient mutagenesis by Cas9 protein-mediated oligonucleotide insertion and large-scale assessment of single-guide RNAs. PLoS ONE *9*, e98186.

Garneau, J.E., Dupuis, M.-È., Villion, M., Romero, D.A., Barrangou, R., Boyaval, P., Fremaux, C., Horvath, P., Magadán, A.H., and Moineau, S. (2010). The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. Nature *468*, 67–71.

González, F., Zhu, Z., Shi, Z.D., Lelli, K., Verma, N., Li, Q.V., and Huangfu, D. (2014). An iCRISPR platform for rapid, multiplexable, and inducible genome editing in human pluripotent stem cells. Cell Stem Cell *15*, 215–226.

Goulburn, A.L., Alden, D., Davis, R.P., Micallef, S.J., Ng, E.S., Yu, Q.C., Lim, S.M., Soh, C.-L., Elliott, D.A., Hatzistavrou, T., et al. (2011). A targeted NKX2.1 human embryonic stem cell reporter line enables identification of human basal forebrain derivatives. Stem Cells *29*, 462–473.

Grskovic, M., Javaherian, A., Strulovici, B., and Daley, G.Q. (2011). Induced pluripotent stem cells—opportunities for disease modelling and drug discovery. Nat. Rev. Drug Discov. *10*, 915–929.

Hou, Z., Zhang, Y., Propson, N.E., Howden, S.E., Chu, L.-F., Sontheimer, E.J., and Thomson, J.A. (2013). Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. Proc. Natl. Acad. Sci. USA *110*, 15644–15649.

Hsu, P.D., Scott, D.A., Weinstein, J.A., Ran, F.A., Konermann, S., Agarwala, V., Li, Y., Fine, E.J., Wu, X., Shalem, O., et al. (2013). DNA targeting specificity of RNA-guided Cas9 nucleases. Nat. Biotechnol. *31*, 827–832.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science *337*, 816–821.

Jinek, M., East, A., Cheng, A., Lin, S., Ma, E., and Doudna, J. (2013). RNA-programmed genome editing in human cells. eLife *2*, e00471.

Kim, D., Bae, S., Park, J., Kim, E., Kim, S., Yu, H.R., Hwang, J., Kim, J.-I., and Kim, J.-S. (2015). Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. Nat. Methods *12*, 237–243.

Kiskinis, E., Sandoe, J., Williams, L.A., Boulting, G.L., Moccia, R., Wainger, B.J., Han, S., Peng, T., Thams, S., Mikkilineni, S., et al. (2014). Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. Cell Stem Cell *14*, 781–795.

Liu, Y., and Rao, M. (2011). Gene targeting in human pluripotent stem cells. Methods Mol. Biol. *767*, 355–367.

Mali, P., Yang, L., Esvelt, K.M., Aach, J., Guell, M., DiCarlo, J.E., Norville, J.E., and Church, G.M. (2013). RNA-guided human genome engineering via Cas9. Science *339*, 823–826.

Merkle, F.T., and Eggan, K. (2013). Modeling human disease with pluripotent stem cells: from genome association to function. Cell Stem Cell *12*, 656–668.

Metzger, M.J., McConnell-Smith, A., Stoddard, B.L., and Miller, A.D. (2011). Single-strand nicks induce homologous recombination with less toxicity than double-strand breaks using an AAV vector template. Nucleic Acids Res. *39*, 926–935.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology *155*, 733–740.

Montague, T.G., Cruz, J.M., Gagnon, J.A., Church, G.M., and Valen, E. (2014). CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. *42*, W401–W407.

Pattanayak, V., Lin, S., Guilinger, J.P., Ma, E., Doudna, J.A., and Liu, D.R. (2013). High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. Nat. Biotechnol. *31*, 839–843.

Ran, F.A., Hsu, P.D., Lin, C.-Y., Gootenberg, J.S., Konermann, S., Trevino, A.E., Scott, D.A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.

Reinhardt, P., Schmid, B., Burbulla, L.F., Schöndorf, D.C., Wagner, L., Glatza, M., Höing, S., Hargus, G., Heck, S.A., Dhingra, A., et al. (2013). Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. Cell Stem Cell *12*, 354–367.

Rouet, P., Smih, F., and Jasin, M. (1994). Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc. Natl. Acad. Sci. USA *91*, 6064–6068.

Ruby, K.M., and Zheng, B. (2009). Gene targeting in a HUES line of human embryonic stem cells via electroporation. Stem Cells *27*, 1496–1506.

Sandoe, J., and Eggan, K. (2013). Opportunities and challenges of pluripotent stem cell neurodegenerative disease models. Nat. Neurosci. *16*, 780–789.

Schwank, G., Koo, B.-K., Sasselli, V., Dekkers, J.F., Heo, I., Demircan, T., Sasaki, N., Boymans, S., Cuppen, E., van der Ent, C.K., et al. (2013). Functional repair of CFTR by CRISPR/Cas9 in intestinal stem cell organoids of cystic fibrosis patients. Cell Stem Cell *13*, 653–658.

Smith, C., Gore, A., Yan, W., Abalde-Atristain, L., Li, Z., He, C., Wang, Y., Brodsky, R.A., Zhang, K., Cheng, L., and Ye, Z. (2014). Whole-genome sequencing analysis reveals high specificity of CRISPR/Cas9 and TALEN-based genome editing in human iPSCs. Cell Stem Cell *15*, 12–13.

Suzuki, K., Yu, C., Qu, J., Li, M., Yao, X., Yuan, T., Goebl, A., Tang, S., Ren, R., Aizawa, E., et al. (2014). Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones. Cell Stem Cell *15*, 31–36.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145–1147.

Tsai, S.Q., Zheng, Z., Nguyen, N.T., Liebers, M., Topkar, V.V., Thapar, V., Wyvekens, N., Khayter, C., Iafrate, A.J., Le, L.P., et al. (2015). GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. Nat. Biotechnol. *33*, 187–197.

Veres, A., Gosis, B.S., Ding, Q., Collins, R., Ragavendran, A., Brand, H., Erdin, S., Cowan, C.A., Talkowski, M.E., and Musunuru, K. (2014). Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing. Cell Stem Cell *15*, 27–30.

Wang, X., Wang, Y., Wu, X., Wang, J., Wang, Y., Qiu, Z., Chang, T., Huang, H., Lin, R.-J., and Yee, J.-K. (2015). Unbiased detection of off-target cleavage by CRISPR-Cas9 and TALENs using integrase-defective lentiviral vectors. Nat. Biotechnol. *33*, 175–178.

Wiedenheft, B., Sternberg, S.H., and Doudna, J.A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature *482*, 331–338.

Xue, H., Wu, S., Papadeas, S.T., Spusta, S., Swistowska, A.M., MacArthur, C.C., Mattson, M.P., Maragakis, N.J., Capecchi, M.R., Rao, M.S., et al. (2009). A targeted neuroglial reporter line generated by homologous recombination in human embryonic stem cells. Stem Cells *27*, 1836–1846.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. Cell *154*, 1370–1379.

Yang, L., Grishin, D., Wang, G., Aach, J., Zhang, C.-Z., Chari, R., Homsy, J., Cai, X., Zhao, Y., Fan, J.-B., et al. (2014a). Targeted and genome-wide sequencing reveal single nucleotide variations impacting specificity of Cas9 in human stem cells. Nat. Commun. *5*, 5507.

Yang, H., Wang, H., and Jaenisch, R. (2014b). Generating genetically modified mice using CRISPR/Cas-mediated genome engineering. Nat. Protoc. *9*, 1956–1968.

Zwaka, T.P., and Thomson, J.A. (2003). Homologous recombination in human embryonic stem cells. Nat. Biotechnol. *21*, 319–321.