Stem Cell Reports
Resource
OPEN ACCESS

A Scaled Framework for CRISPR Editing of Human Pluripotent Stem Cells to Study Psychiatric Disease

Dane Z. Hazelbaker,[1] Amanda Beccard,[1] Anne M. Bara,[1] Nicole Dabkowski,[1] Angelica Messana,[1] Patrizia Mazzucato,[1] Daisy Lam,[1] Danielle Manning,[3] Kevin Eggan,[1,2,*] and Lindy E. Barrett[1,2,*]
[1]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
[2]Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA
[3]Department of Pathology, Brigham and Women's Hospital, Boston, MA 02115, USA
*Correspondence: eggan@mcb.harvard.edu (K.E.), lbarrett@broadinstitute.org (L.E.B.)
https://doi.org/10.1016/j.stemcr.2017.09.006

SUMMARY

Scaling of CRISPR-Cas9 technology in human pluripotent stem cells (hPSCs) represents an important step for modeling complex disease and developing drug screens in human cells. However, variables affecting the scaling efficiency of gene editing in hPSCs remain poorly understood. Here, we report a standardized CRISPR-Cas9 approach, with robust benchmarking at each step, to successfully target and genotype a set of psychiatric disease-implicated genes in hPSCs and provide a resource of edited hPSC lines for six of these genes. We found that transcriptional state and nucleosome positioning around targeted loci was not correlated with editing efficiency. However, editing frequencies varied between different hPSC lines and correlated with genomic stability, underscoring the need for careful cell line selection and unbiased assessments of genomic integrity. Together, our step-by-step quantification and in-depth analyses provide an experimental roadmap for scaling Cas9-mediated editing in hPSCs to study psychiatric disease, with broader applicability for other polygenic diseases.

INTRODUCTION

Combining CRISPR-Cas9 gene editing and directed differentiation of human pluripotent stem cells (hPSCs) into somatic cell types offers a powerful approach to elucidate mechanisms across the spectrum of human disease. As methodologies are fine-tuned in each of these arenas, the field is challenged to increase scale in order to keep pace with emerging human genetic data. Indeed, biobanking initiatives for hPSCs (including both human embryonic stem cells [hESCs] and human induced pluripotent stem cells [hiPSCs]) have emerged in recent years with collections aiming to include thousands of cell lines (Avior et al., 2016; McKernan and Watt, 2013). In parallel, high-throughput platforms are being developed for phenotypic characterization of hPSCs and their differentiated progeny (Leha et al., 2016; Paull et al., 2015).

These scaled initiatives hold great promise for accelerating our understanding of polygenic diseases such as diabetes (Flannick and Florez, 2016), heart disease (Kessler et al., 2016), and psychiatric disease. In the case of psychiatric disease, human genetic data have implicated hundreds of loci or genes in schizophrenia (SCZ), autism spectrum disorder (ASD), and intellectual disability (ID) (Sanders et al., 2015; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), yet only a fraction of implicated genes have well-defined neurobiological functions. Here, stem cell models are particularly relevant, as human brain tissue is largely inaccessible for *in vitro* studies, and common model organisms lack implicated brain structures (Sanders et al., 2015; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Yoon et al., 2014). Indeed, stem cell differentiation technologies now allow for the generation of many human brain cell types *in vitro* (Chambers et al., 2009; Kriks et al., 2011; Lu et al., 2016; Mertens et al., 2016; Zhang et al., 2013), facilitating the use of genetically defined hPSC lines to study disease mechanisms in relevant neuronal populations (Hendriks et al., 2016; Santos et al., 2016).

Several studies have identified disease-relevant phenotypes in cells with single-gene loss-of-function (LoF) mutations generated using CRISPR-Cas9. For example, Yi et al. (2016) generated isogenic hESCs with mutations in the ASD-associated gene *SHANK3* and derived neurons *in vitro* to identify specific channel defects, demonstrating how a LoF strategy was valuable for making discoveries of normal neurobiological function to provide insight into the disease state. Similar strategies were employed in hiPSCs to study the ASD-associated gene *CHD8* (Wang et al., 2015) and the SCZ-associated gene *DISC1* (Srikanth et al., 2015).

Despite these advances, empirical knowledge of best practices for scaling Cas9-mediated gene editing in hPSCs is lacking. Compared with more commonly used HEK293Ts or cancer cell lines (Doench et al., 2016; Hsu et al., 2013), hPSCs display poor viability at a single-cell state (Amit et al., 2000; Hendriks et al., 2016), and studies report lower frequencies of gene correction in hPSCs compared with cancer cell lines (Hockemeyer and Jaenisch, 2016). Pluripotent cells may have a lower tolerance for double-strand breaks (DSBs) compared with non-pluripotent cells (Liu et al.,

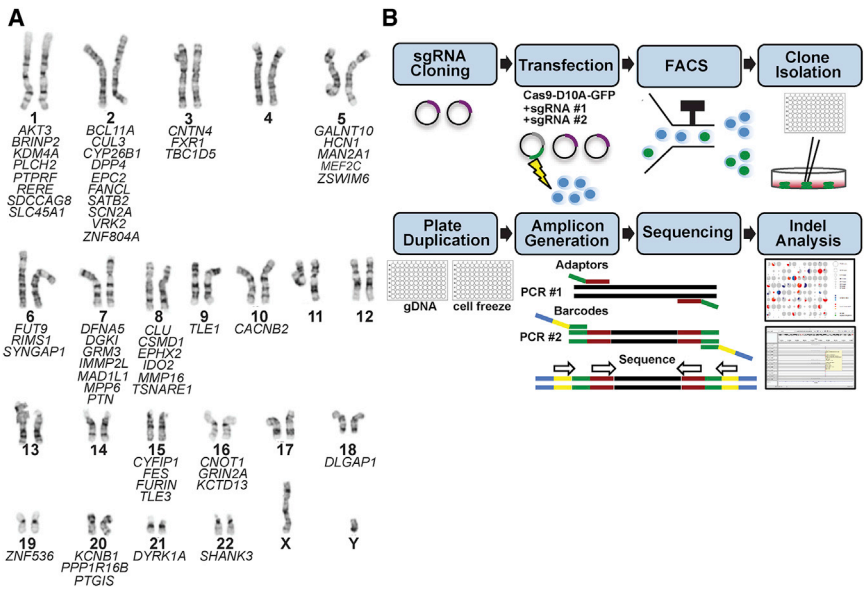

Figure 1. Selected Genes and Experimental Workflow

(A) Genes selected for editing and their chromosomal location. For gene target and sgRNA plasmid details, see Table S1.

(B) Schematic of Cas9^{D10A} nickase gene-editing workflow. For amplicon primer sequences, see Table S2.

2014) and can acquire chromosomal abnormalities with continued passage or cellular stress (Amit et al., 2000; Martins-Taylor and Xu, 2012; Taapken et al., 2011). At the level of the targeted locus, some studies report that chromatin accessibility is a significant factor in the ability of Cas9 complexes to bind DNA (Daer et al., 2016; Isaac et al., 2016; Kuscu et al., 2014; O'Geen et al., 2015; Wu et al., 2014) while others report no correlation (Perez-Pinera et al., 2013) or a decrease of Cas9 binding in heterochromatic regions (Knight et al., 2015). Thus, it is difficult to predict the outcome of an individual gene-targeting experiment in an hPSC line, a problem surmountable for single-gene studies, but one that is amplified with increased experimental scale.

Here, we applied an established Cas9-mediated gene-targeting strategy to target and genotype 58 distinct genes implicated in psychiatric disease. Our dataset, comprising nearly 5,000 individual hPSC clones, allowed us to benchmark efficiencies at each experimental step, as well as the potential influence of transcription state and nucleosome positioning on gene-editing outcome. To derive indel frequencies and features, we applied both a publicly available sequence-evaluation tool as well as an in-house tool. Finally, to address whether different hPSC lines may be more amenable to the gene-editing workflow, we targeted a subset of genes in an independent hPSC line.

RESULTS

Gene Selection, Cell-Line Selection, and Gene-Editing Workflow

We selected 58 genes implicated in SCZ, ASD, and/or ID (Basak et al., 2015; Durand et al., 2007; Golzio et al., 2012; Hamdan et al., 2009; Nishimura et al., 2007; O'Roak et al., 2012; Pickard et al., 2005; Sanders et al., 2015; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Too et al., 2016) distributed across 16 chromosomes (Figure 1A and Table S1) and applied a standardized gene-editing approach (Santos et al., 2016) (Figure 1B) to generate heterozygous and homozygous indels in hPSCs utilizing paired single-guide RNAs (sgRNAs) and Cas9^{D10A} nickase (Ran et al., 2013b). We used Cas9^{D10A} nickase based on studies suggesting equal on-target cutting efficiency and enhanced specificity with this strategy compared with single sgRNAs and Cas9 nuclease (Ran et al., 2013a). Paired sgRNAs were designed to target early constitutive exons (Table S1) in order to generate frameshift indels in the maximal number of gene isoforms.

We then selected two well-characterized hESC lines, HUES63 (XY) and WA01 (XY) (Thomson et al., 1998), with the goal of creating LoF mutations in isogenic backgrounds in order to further understand gene function and ultimately compare cellular phenotypes with those observed in patient-derived iPSCs in downstream studies. Both HUES63 and WA01 were genotyped using a high-density Illumina microarray to probe psychiatric disease risk, and we confirmed a lack of significant enrichment for marker SNPs associated with common psychiatric diseases (<2 SDs from the mean; data not shown). Both cell lines were also whole-exome sequenced at 60× coverage, confirming the absence of known structural variants predicted to have a negative impact on stem cell biology or brain development (Merkle et al., 2017).

Initially, HUES63 cells were transfected with paired sgRNA plasmids and Cas9^{D10A} nickase-GFP (PX286; kindly provided by F. Zhang). Individual GFP positive (GFP+) cells

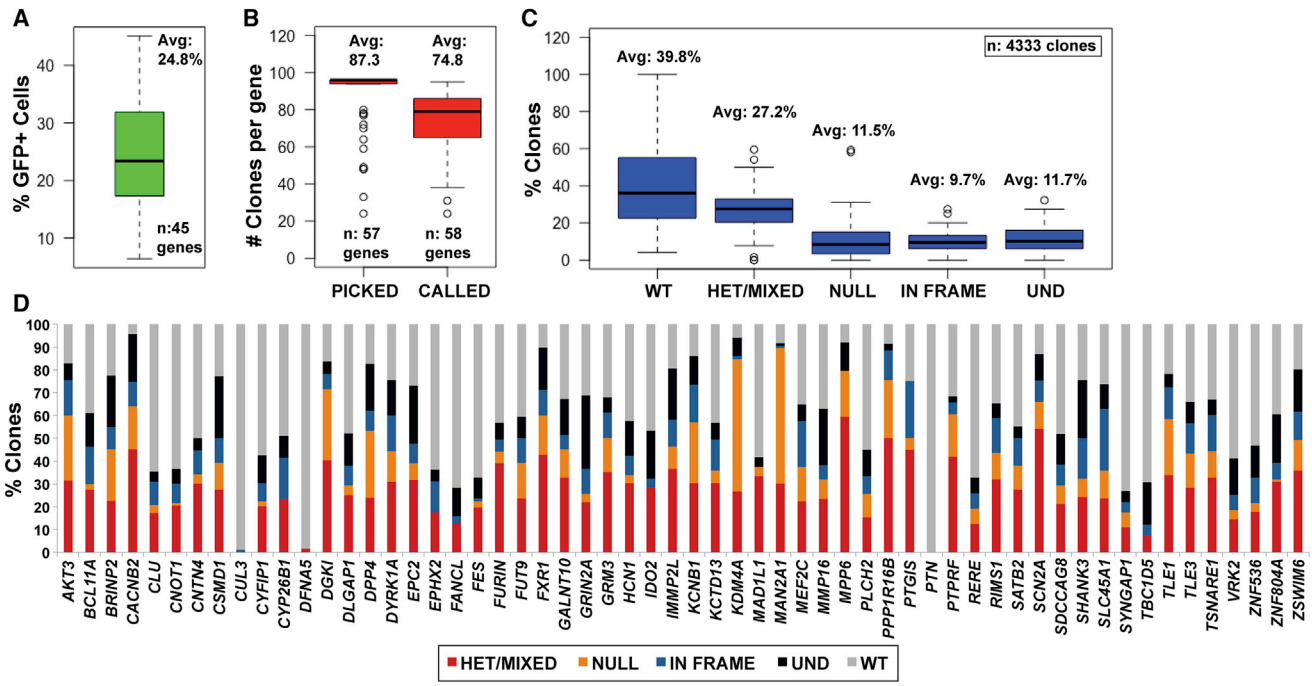

Figure 2. Indel Generation Efficiency for 58 Targeted Genes in the hPSC Line HUES63

(A) Percentage of GFP+ cells following Cas9^{D10A}-GFP transfection (n = 45 genes). Individual transfection percentages are shown in Figure S1A.

(B) Number of clones isolated ("picked"; n = 57 genes) and number of clones that yielded indel data ("called"; n = 58 genes) per gene.

(C) Percentage of clones that were putative WT, HET/MIXED, NULL, IN FRAME, or UND for 58 targeted genes (n = 4,333 clones). For examples of clone genotyping, see Figure S1B.

(D) Editing landscape across 58 genes with percentage of clones that were putative HET/MIXED (red), NULL (orange), IN FRAME (blue), UND (black), or WT (gray) for each gene.

In all boxplots, the dark horizontal line represents the median, the box represents the 25th and 75th percentiles, the whiskers represent the 5th and 95th percentiles, and the circles represent outliers.

were then isolated by fluorescence-activated cell sorting (FACS) and plated at low density to allow clonal stem cell recovery. Individual clones were isolated and expanded in 96-well plates and duplicated for cell freezing and genomic DNA (gDNA) extraction. For each clonal well, barcoded PCR amplicons spanning the Cas9 target site were generated (Figure 1B and Table S2) and pooled to create a sequencing library for each targeted gene. To assess indel generation, sequence files were input into OutKnocker (www.OutKnocker.org) (Schmid-Burgk et al., 2014).

Efficiency of Indel Generation in hPSCs

We first quantified efficiencies at key experimental steps. Considering transfection efficiency of Cas9^{D10A} nickase-GFP in hPSCs across 45 separate gene-targeting experiments, we obtained an average of 24.8% GFP+ cells (Figure 2A) with a range of 6.43%–45.1% GFP+ cells (Figure S1A). In all cases, we recovered a sufficient number of live GFP+ cells to plate at clonal density and isolate colonies. On average, we plated 10,000 cells in a 10 cm dish (range, 6,400–30,000 cells) and picked 87.3 clones per gene (considering data for 57 genes; 4,976 total clones) (Figure 2B). After analyzing sequencing data with OutKnocker, we obtained indel data on an average of 74.8 clones per gene-targeting experiment (n = 58 genes; 4,333 total clones) (Figure 2B). In total, we obtained indel data on approximately 86% of isolated clones with an average read depth of 18,878 reads per clone across all experiments. These results validated a high rate of return for each clone picked and analyzed using the above conditions.

We applied the following nomenclature in our genotype analysis to identify clones for downstream LoF analyses. Clones that had less than 25% unaligned reads and no indels were scored as putative wild-type (WT), while clones with greater than 25% unaligned reads were defined as undetermined (UND). Clones that had at least two sequence species with a minimum of one sequence species containing a frameshift indel and one sequence species not corresponding to a frameshift indel were considered putative

heterozygous or heterozygous mixed clones (HET/MIXED). The designation of heterozygous mixed clones is important as sub-cloning of these species may yield clones that are heterozygous or homozygous for frameshift indels. In cases where all sequences corresponded to identical or non-identical frameshift indels and no WT sequence, clones were considered putative frameshift null (NULL). Finally, when reads contained an in-frame indel or mix of WT and in-frame indels, but no frameshift indels, clones were designated as putative in-frame (IN FRAME). Examples of read alignments for putative WT, HET/MIXED, and NULL clones are shown in Figure S1B.

Across all experiments, we obtained an average of 39.8% WT clones, 27.2% HET/MIXED clones, 11.5% NULL clones, 9.7% IN FRAME clones, and 11.7% UND clones (Figure 2C). Individual genes showed a high degree of variability in terms of the number of clones with indels, ranging from 0% (e.g., *PTN*) to nearly 90% (e.g., *KDM4A* and *MAN2A1*) (Figure 2D). While we cannot distinguish between poor activity of sgRNAs versus indels that may be deleterious to stem cell viability and therefore selected against, 56 of 58 genes targeted resulted in clones with putative frameshift indels (Figure 2D), suggesting that the Cas9^{D10A} nickase approach was effective in hPSCs across a broad range of genes. Indeed, while we failed to genotype any clones with putative frameshift indels in *CUL3*, this gene has been shown to play an important role in mammalian cell division (Singer et al., 1999).

Influence of Gene-Specific Features on Indel Generation

To probe the contribution of gene-specific features to Cas9-mediated indel generation, we examined how transcript expression and nucleosome positioning around target genes affected outcome. By comparing editing efficiency in 53 genes with their corresponding FPKM (fragments per kilobase of transcript per million mapped reads) values in HUES63 (Cacchiarelli et al., 2015), we found that transcript levels were a poor predictor of editing efficiency (Figure 3A). We observed examples of high editing efficiency in genes expressed with FPKM <1 and low editing efficiency in genes expressed with FPKM >10 and vice versa. Indeed, editing efficiency appeared randomly distributed among high and low expressed genes, with an R^2 value of 0.01846 (Figure S2). While these analyses do not account for the influence of sgRNA expression/stability for an individual gene, our data suggest that neither high nor low transcript expression was an accurate predictor of editing outcome. Furthermore, we examined nucleosome positioning around 58 target genes using a published MNase-seq dataset (Yazdi et al., 2015) and found no correlation between editing efficiency and nucleosome dyad distance to the nearest PAM sequence, with an R^2 value of 0.00242 (Figure 3B). This suggests that efficient editing can occur at target loci in close proximity to the predicted location of nucleosomes.

Indel Properties with Cas9^{D10A} Nickase

We next analyzed the properties of indels generated in HUES63 and found that, on average, 84.72% of indels were deletions, and 15.28% of indels were insertions (Figure 4A). As expected, 33.65% of indels were in-frame, and 66.35% of indels were frameshift (Figure 4B). We observed an average deletion size of 25.7 bp (range, 1–90 bp) and an average insertion size of 5.4 bp (range, 1–10 bp) (Figures 4C–4E). There was no apparent size difference for in-frame versus frameshift deletions or in-frame versus frameshift insertions (Figure 4C) and no clear preference for indels of a specific length (Figures 4D and 4E). It is important to note that version 1.0 of OutKnocker did not call insertions greater than 10 bp, so there may be additional larger insertions present but not detected. Further, indels that made up a sizable percentage of the overall PCR amplicon (average amplicon length = 166 bp) likely failed in the library generation step, which weighted our analyses toward the detection of smaller indels. We also sought to examine the influence of 5′ overhang length generated by Cas9^{D10A} on editing efficiency by binning our sgRNA pairs into 35–44 base, 45–54 base, and 55–64 base 5′ overhang groups and determining the editing efficiency of each group. As shown in Figure 4F, 5′ overhang length had little effect on editing efficiency, in line with previous observations that optimal sgRNA cut site spacing was approximately 30–70 bases (Shen et al., 2014).

During the course of our analyses, we noted that a significant number of insertions ($\geq$3 bp) contained sequences with micro-homology to portions of the sgRNA and/or DNA target sites for Cas9^{D10A} nickase (Figure 4G). It is possible that upon generation of a DSB, the transfected sgRNA plasmids serve as readily available donor templates for homology-directed DNA repair (Cong et al., 2013), microhomology-mediated end joining (van Overbeek et al., 2016), or RNA-directed DNA repair pathways (Keskin et al., 2014). The presence of these insertions underscores the need to consider plasmid-free Cas9 delivery in hPSCs, such as RNPs, that may reduce the potential integration of unwanted sequences with high or degenerate homology to sequences adjacent to the DSB.

Utilization of a Second Sequence-Evaluation Tool

Next-generation sequencing is a powerful tool to identify indels generated by gene targeting, however, downstream analytical pipelines often use different alignment methodologies, threshold criteria, and variant effect predictors, which can substantially influence outcomes (Hasan et al., 2015). To

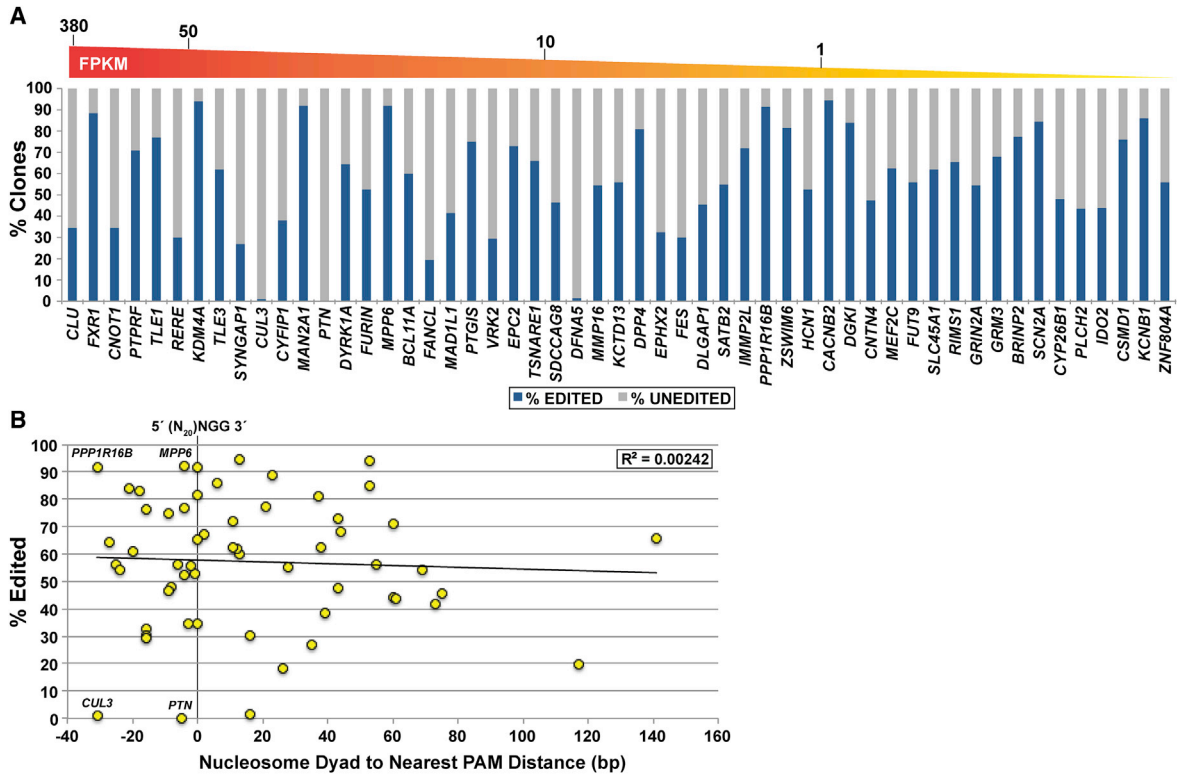

Figure 3. Overall Editing Efficiency Versus Steady-State Transcript Levels and Predicted Nucleosome Position

(A) Percentage of clones EDITED (putative HET/MIXED, NULL, IN FRAME, or UND clones with greater than 5% non-WT reads; blue) or UNEDITED (WT; gray) for each gene (n = 53) plotted with their respective FPKM values (top). For the correlation plot of gene edits relative to FPKM, see Figure S2.

(B) Correlation between editing efficiency and the distance between the closest PAM of either Cas9^{D10A} target site (n = 58 genes) to the nearest nucleosome dyad. R^2 = 0.00242. Examples of genes with a high percentage of edited clones (*PPP1R16B, MPP6*) and a low percentage of edited clones (*CUL3, PTN*) are highlighted.

further validate our findings, we developed a sequence-evaluation tool, which we called BaySnpper (Figure S3A; https://github.com/dat4git/BaySnpper). BaySnpper utilized FreeBayes (FB) (Garrison and Marth, 2012) to create composite calls from forward and reverse sequence reads and SnpEff to make predictions about the indel effect (Cingolani et al., 2012). Data were exported to generate Tab Separated Values (TSV) files as well as annotated Variant Call Format (VCF) files for visualization in Integrative Genomics Viewer (Robinson et al., 2011) or other genome browsers (Figure S3A). We then re-analyzed 30 individual gene-targeting experiments using BaySnpper and compared calls with those obtained via OutKnocker (Figures 5A, 5B, and S3B). As shown in Figure 5A, indel distribution was similar across 30 genes using data analyzed by OutKnocker and BaySnpper. Both OutKnocker and BaySnpper called an average of approximately 19–22 HET/MIXED clones, 6–7 NULL clones, and 6–8 IN FRAME clones per gene. On a gene-by-gene basis, we also observed similar proportions of HET/MIXED, NULL, and IN FRAME clones (Figure 5B). It is important to note that clone-to-clone differences did arise, presumably due to differences in alignment algorithms, read thresholds, and quality scores between the programs (Figures S3C and S3D). For example, *TSNARE1* showed highly similar proportions of HET/MIXED, NULL, and IN FRAME calls (Figure 5B), but there were discrepant indel calls between individual clones analyzed with OutKnocker and BaySnpper (Figures S3C and S3D). Thus, utilization of two analysis programs in parallel increased our confidence at the level of an individual clone.

Gene Editing in Independent hPSC Lines

To compare results in HUES63 with another hESC line, we repeated seven gene-targeting experiments in WA01 using identical sgRNAs and Cas9^{D10A} nickase-puro (PX462; Ran et al., 2013b) or Cas9 nuclease (PX459; Ran et al., 2013b) (Figure 6A and Table S3). Although seven out of seven genes targeted in WA01 resulted in clones with putative frameshift indels, average indel generation efficiency was consistently lower in WA01 compared with HUES63. To

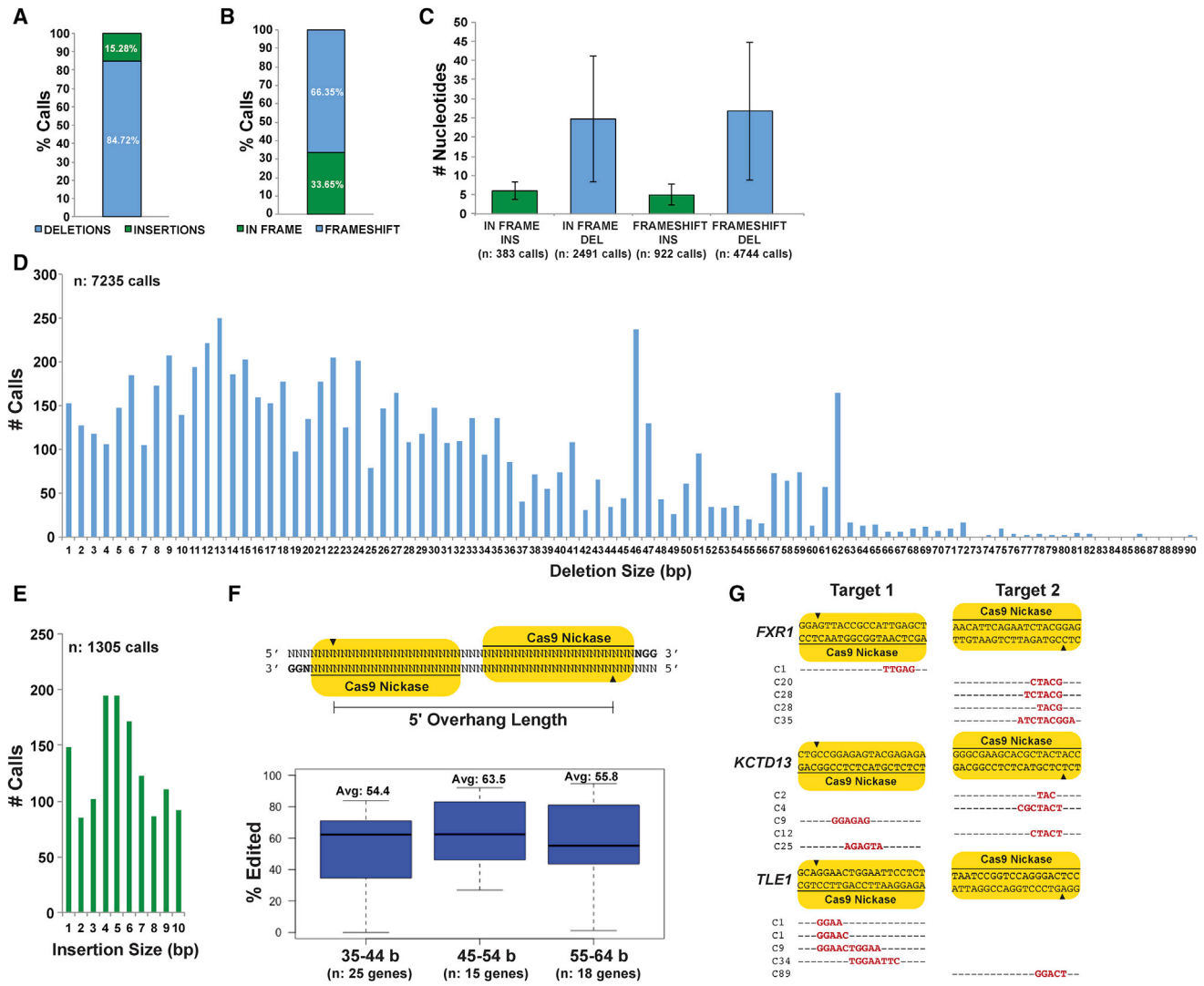

Figure 4. Indel Properties for 58 Targeted Genes in HUES63

(A) Percentage of Cas9^{D10A}-mediated indels called as deletions (blue; n = 7,235 calls) or insertions (green; n = 1,305 calls).

(B) Percentage of in-frame indels (green; n = 2,874 calls) or frameshift indels (blue; n = 5,666 calls).

(C) Average size of in-frame insertions (n = 383 calls), in-frame deletions (n = 2,491 calls), frameshift insertions (n = 922 calls), and frameshift deletions (n = 4,744 calls). Error bars represent SD.

(D) Distribution of deletion sizes (1–90 bp; n = 7,235 calls).

(E) Distribution of insertion sizes (1–10 bp; n = 1,305 calls).

(F) Schematic of Cas9^{D10A}-nickase generated 5′ overhangs at target sites (top). Boxplots depicting the influence of 5′ overhang length of paired sgRNAs on editing efficiency for 58 genes (bottom).

(G) Locations and identity of inserted DNA sequences in selected clones (C) for *FXR1*, *KCTD13*, and *TLE1* genes and their alignment to the Cas9^{D10A} target sequences. Inserted sequences for each clone are shown in red.

In all boxplots, the dark horizontal line represents the median, the box represents the 25th and 75th percentiles, the whiskers represent the 5th and 95th percentiles.

assess whether this could be attributed to the use of Cas9^{D10A} nickase specifically, we repeated four gene-targeting experiments using Cas9 nuclease (PX459 (Ran et al., 2013b)). As shown in Figure 6B, both Cas9^{D10A} nickase and Cas9 nuclease approaches yield similar outcomes, suggesting that WA01 may be less amenable to Cas9 editing. However, the properties of the indels generated remained comparable between WA01 and HUES63 (Figures 4A–4C and S4A–S4H), and there was a trend toward smaller deletions using the Cas9 nuclease (Figure S4G). These results

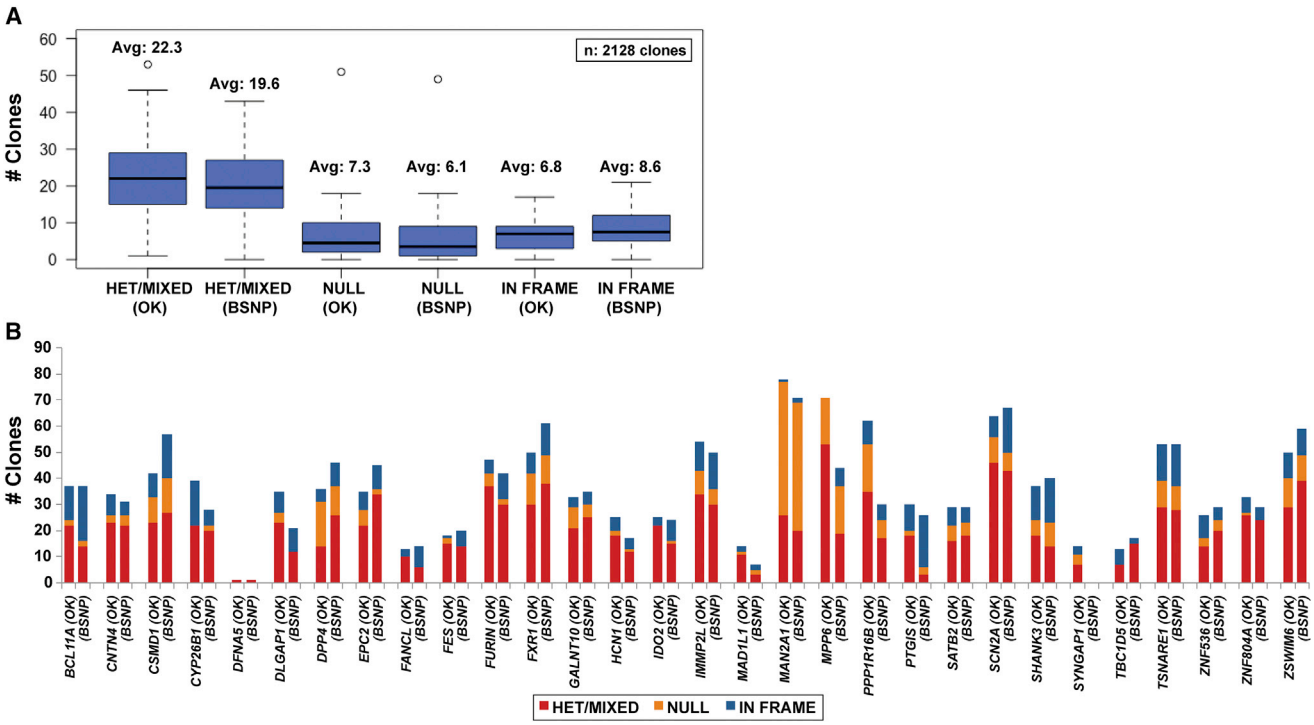

Figure 5. Comparison between OutKnocker and BaySnpper Indel Analysis Tools

(A) Average number of putative HET/MIXED, NULL, or IN FRAME clones analyzed by OutKnocker (OK) or BaySnpper (BSNP) for 30 Cas9^{D10A} targeted genes (n = 2,128 clones).

(B) Distribution of putative HET/MIXED (red), NULL (orange), or IN FRAME (blue) clones analyzed by OK or BSNP tools for 30 Cas9^{D10A} targeted genes. For examples of data visualization and call comparisons between OK and BSNP, see Figures S3A–S3D.

In all boxplots, the dark horizontal line represents the median, the box represents the 25th and 75th percentiles, the whiskers represent the 5th and 95th percentiles, and the circles represent outliers.

suggest that inherent or culture-acquired properties of individual hPSC lines may significantly affect Cas9-mediated gene-editing efficiency. From WA01 targeting experiments, clones containing putative heterozygous and homozygous frameshift indels in *CSMD1*, *CYFIP1*, *FXR1*, *SYNGAP1*, *TLE1* (Herring et al., 2016), and *TLE3* (Bara et al., 2016) are described in Table S4.

Despite reduced editing efficiency, WA01 displayed greater genomic stability compared with HUES63. Indeed, despite starting with a normal karyotype, HUES63 cells acquired chromosomal aberrations in 78 of 83 clones analyzed in the laboratory to date by SNP genotyping or G-banded karyotyping analysis (Figures 6C and 6D). We observed recurrent aberrations in HUES63 across experiments, namely trisomies 4, 11, 12, 14, 17, and X, in addition to the presence of a small number of copy number variants (CNVs; 0.27–2.4MB). Although individual clones differed in the number and specific combination of trisomies, the vast majority of clones shared trisomy 12, and many also shared trisomies 11, 14, and 17. Trisomy 4 was observed in approximately 35% of clones (data not shown). Importantly, none of the 58 genes included in this dataset were localized on chromosomes prone to trisomy. Trisomies 12, 14, 17, and X have been repeatedly identified in hPSCs *in vitro* and are thought to endow cells with a competitive growth advantage (Baker et al., 2016; International Stem Cell Initiative et al., 2011; Mayshar et al., 2010; Nguyen et al., 2013; Peterson and Loring, 2014), while trisomies 4 and 11 are less frequently reported in the literature. By contrast, WA01 acquired chromosomal aberrations in only 6 of 83 clones analyzed in the laboratory to date (Figure 6C), including one clone with trisomy 20, and five clones each harboring a single CNV (0.43–2.39 MB). The potential growth advantage conferred upon cells by chromosomal aberrations may indeed have contributed to the increased editing efficiency observed in HUES63 compared with WA01. However, it is also likely that the bottleneck of clonal selection was a significant contributor to instability rather than DNA damage caused by Cas9, as we identified similar chromosomal aberrations in HUES63 clones using the same gene-targeting protocol in the absence of Cas9 and sgRNA expression (data not shown).

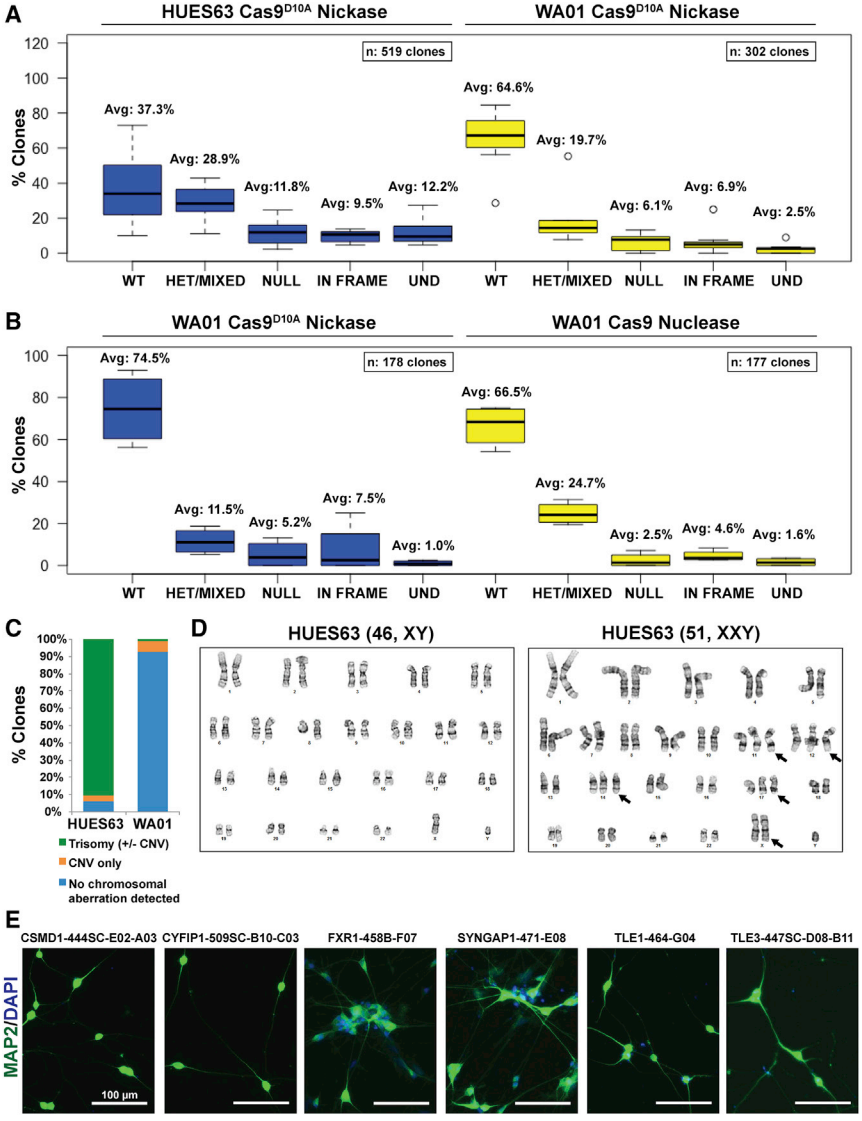

Figure 6. Gene Editing in Independent hPSC Lines

(A) Comparison of Cas9^{D10A} nickase editing efficiencies in HUES63 (blue; n = 519 clones, PX286) and WA01 (yellow; n = 302 clones, PX462) for seven selected targeting experiments using identical sgRNA pairs for *CSMD1*, *CYFIP1*, *FURIN*, *FXR1*, *SYNGAP1*, *TLE1*, and *TLE3*. Gene target and sgRNA plasmid details are in Table S3 and for indel properties, see Figures S4A–S4C.

(B) Comparison of editing efficiencies in WA01 with Cas9^{D10A} nickase (blue; n = 178 clones, PX462) and Cas9 nuclease (yellow; n = 177 clones, PX459) for four selected gene-targeting experiments with identical paired and single sgRNAs for *CSMD1*, *CYFIP1*, *FXR1*, and *TLE1*. Indel comparisons are found in Figures S4D–S4H.

(C) Percentage of clones analyzed by SNP genotyping or G-banded karyotyping analysis with no chromosomal aberration detected (blue), CNVs only (orange), or trisomies with or without CNVs (green) for HUES63 (n = 83 clones) and WA01 (n = 83 clones).

(D) Examples of G-banded metaphase analysis for parental HUES63 (46, XY) and an abnormal clone after gene targeting (51, XY +11, +12, +14, +17, +X). Arrows denote instances of aneuploidy.

(E) Generation of induced neurons from select WA01 clones with indels in *CSMD1*, *CYFIP1*, *FXR1*, *SYNGAP1*, *TLE1*, and *TLE3*. For detailed clone information, see Table S4. For pluripotency and tri-lineage differentiation potential of select WA01 clones, see Figures S5A and S5B.

In all boxplots, the dark horizontal line represents the median, the box represents the 25th and 75th percentiles, the whiskers represent the 5th and 95th percentiles, and the circles represent outliers.

Generation of Induced Neurons from Edited WA01 hPSC Lines

To further validate our hPSC lines after gene targeting, we confirmed expression of pluripotency markers as well as germ-layer differentiation capacity following embryoid body formation for a set of six WA01 clones containing putative frameshift indels in *CSMD1*, *CYFIP1*, *FXR1*, *SYNGAP1*, *TLE1*, and *TLE3* (Figures S5A and S5B) with no chromosomal aberrations detected (Table S4). To assess the relevance of these clones for the interrogation of downstream cellular phenotypes, we then generated induced neurons via ectopic expression of Ngn2 (Figure 6E) (Pak et al., 2015; Yi et al., 2016; Zhang et al., 2013).

DISCUSSION

By directly comparing editing outcomes across nearly 5,000 hPSC clones from a diverse set of 58 genes, our findings outline a readily applicable experimental and computational framework for Cas9-mediated editing in hPSCs for the study of psychiatric disease. In addition, our findings help to set expectations for large-scale studies of other complex diseases.

Without pre-testing guide pairs, we obtained putative frameshift indels in approximately 40% of clones in HUES63 and 25% of clones in WA01. The fact that the Cas9^{D10A} nickase and Cas9 nuclease strategies were equally

efficient in terms of indel generation suggests that results obtained with the Cas9^{D10A} nickase strategy are broadly representative of Cas9-mediated indel generation in these hPSCs. Given our limited ability to predict editing efficiency without pre-testing guide pairs *in vitro* and the variability in editing efficiency between different cell lines, we opted to pick a relatively large number of clones per gene-targeting experiment. A successful sequencing and analysis rate of 86% provided us with multiple clones with putative frameshift indels for the majority of genes, even in cases of lower overall editing efficiency. Based on the variability in indel generation observed at a gene-to-gene as well as cell line-to-cell line level, we recommend selecting a similarly large set of clones for initial screening until cell-line-specific data can be ascertained.

Within our gene set, we failed to see either a positive or negative correlation between steady-state transcript levels or nucleosome positioning and Cas9 editing efficiency. *A priori*, we expected highly expressed genes to be more efficiently edited due to a more open chromatin structure (Kuscu et al., 2014; O'Geen et al., 2015; Wu et al., 2014). It is possible that some discrepancies in the literature may result from applications that require Cas9 to have stable as opposed to transient binding to target DNA. For example, several studies reported a positive correlation between chromatin accessibility and Cas9 functionality using nuclease-dead Cas9 (Kuscu et al., 2014; Wu et al., 2014) as opposed to nuclease-active Cas9. As discussed by Horlbeck et al. (2016), nuclease-dead Cas9 requires persistent access to DNA for optimal functionality while nuclease-active Cas9 requires only transient access, which could perhaps be achieved following disruption of nucleosomes during replication.

Computationally, both OutKnocker and BaySnpper effectively identified indels in our dataset, and we advocate the use of a minimum of two programs simultaneously to increase confidence in individual indel calls. Indeed, similar combinatorial strategies have been suggested for evaluating indels from whole-genome sequencing data (Hasan et al., 2015). One advantage of BaySnpper is that it gave detailed information on variant effect prediction, including the level of effect (i.e., high, moderate, or low), location (i.e., coding region, splice site, intron/exon boundary), and effect (i.e., silent, frameshift, missense) as the prediction was made in the context of the whole genic structure (Cingolani et al., 2012).

Importantly, we found one of two parental cell lines, HUES63, to be chromosomally unstable, acquiring aberrations that likely favored their propagation *in vitro*. While HUES63 and WA01 were cultured under standard conditions, current gene-targeting protocols require single-cell passaging for transfection and clonal selection, which may contribute to the emergence of aberrations *in vitro*. Our analyses detected large structural aberrations, but did not address the potential for small CNVs or single nucleotide variants, which could further complicate comparisons between clonal lines assumed to be isogenic. We speculate there may be a relationship between stem cell growth properties and Cas9 editing efficiency, highlighting the importance of identifying factors that contribute to the propensity of a given hPSC line to become aneuploid, especially for applications utilizing large-scale gene editing in hPSCs.

Finally, we have validated a set of 58 sgRNAs in hPSCs for the study of diverse genes implicated in psychiatric disease as well as generated a banked resource of CRISPR-edited hPSC lines for six genes, including *CSMD1*, *CYFIP1*, *FXR1*, *SYNGAP1*, *TLE1*, and *TLE3*. It will be crucial for future studies to incorporate additional genetically diverse cell lines, including hiPSCs, which may include highly informative behavioral and developmental data from patients, to accelerate our understanding of basic underlying disease mechanisms as well as drug discovery efforts.

EXPERIMENTAL PROCEDURES

Plasmid DNA Construction

Paired sgRNAs targeting early exons of selected genes were designed using http://crispr.mit.edu. Oligonucleotides (IDT) corresponding to antisense and sense pairs of sgRNAs were cloned into a pU6-sgRNA vector (kindly provided by Neville Sanjana and Feng Zhang, Broad Institute) to generate individual sgRNA plasmids (Bara et al., 2016; Cong et al., 2013; Herring et al., 2016). The sequence and locations for each gene-specific sgRNA target site are described in Table S1.

Stem Cell Culture and Gene Targeting

The hESC line HUES63 was obtained from the Harvard Stem Cell Institute (Cambridge, MA; http://ipscore.hsci.harvard.edu) and the hESC line WA01 was obtained from WiCell Research Institute (Madison, WI) (Thomson et al., 1998). Stem cells were grown in mTeSR1 medium (Stem Cell Technologies 05850) on Geltrex-coated plates (Life Technologies A1413301) and tested to confirm the absence of mycoplasma contamination (Lonza MycoAlert) as described (Bara et al., 2016; Herring et al., 2016).

For transfection, cells were pre-incubated with 1:1 medium composed of a 1:1 mixture of mTeSR1 medium and hPSC medium (KnockOut DMEM [Gibco 10829-018] with 20% KnockOut serum replacement [Gibco 10828-028], 1% Glutamax [Gibco 35050-061], 1% non-essential amino acids [Corning 25-025-Cl], 0.1% 2-mercaptoethanol [Gibco 21985-023], and 20 ng/mL basic fibroblast growth factor [EMD Millipore GF0003AF]) supplemented with 10 μM ROCK inhibitor (Y-27632). For each transfection, 2.5×10^6 cells were electroporated at 1,050 V, 30 ms, 2 pulses (NEON, Life Technologies MPK10096), as described (Bara et al., 2016; Herring et al., 2016). The following DNA concentrations were used: 1.4 μg per sgRNA plasmid and 7 μg Cas9 plasmid. Cells were dispensed into 10 cm Geltrex-coated plates in 1:1 medium

plus 10 μM Y-27632. Twenty-four hours after transfection, GFP+/phosphotidylinositol− cells were isolated by FACS or cells were treated with 1 μM puromycin (Sigma-Aldrich P8833) for 24 hr and then maintained in 1:1 medium for 1–2 weeks to allow clonal stem cell recovery.

Individual hPSC colonies were picked and seeded into Geltrex-coated 96-well plates, expanded for 1–2 weeks, and duplicated for cell freezing and gDNA extraction. Clones were frozen in 96-well plates using 50% 1:1 medium plus 10 μM Y-27632, 40% fetal bovine serum (VWR SH30070.03), and 10% DMSO (Sigma D2650). gDNA was extracted overnight at 55°C in Tail Lysis Buffer (Viagen 102-T) with Proteinase K (Roche 03115828001) followed by a 1 hr incubation at 90°C. G-band karyotyping analysis was performed by Cell Line Genetics (Madison, WI), and SNP genotyping was performed using the Illumina Infinium PsychArray-24 Kit (Broad Institute, Cambridge, MA). SNP genotyping was also utilized for cell-line authentication.

Using WA01 for gene targeting, we observed low cell viability following FACS and therefore transfected with $Cas9^{D10A}$ nickase-puromycin (PX462) and selected for individual clones with puromycin treatment. To ensure that PX462 was competent to generate indels, we repeated four gene-targeting experiments in HUES63 using identical sgRNAs (Table S3) with PX462 instead of $Cas9^{D10A}$ nickase-GFP (PX286). We observed similar average indel frequencies overall using PX286 and PX462 in HUES63 across four genes analyzed (data not shown) indicating that GFP versus puromycin selection did not significantly alter our ability to generate indels.

DNA Sequencing

To create gene-specific libraries from extracted gDNA in 96-well format, primers containing Illumina multiplexing adapters (Ilumina) and gene-specific sequences (Table S2) were used to PCR amplify sequencing amplicons with Q5 Hot Start High-Fidelity Master Mix (NEB M04945). A second round of PCR amplification incorporated well-specific barcode IDs (Broad Institute). Barcoded PCR products were pooled and gel purified (Zymo Research D4008), run on a 2100 BioAnalyzer (Agilent Technologies) for quality assessment, and submitted for MiSeq paired-end sequencing (Broad Institute), as described (Bara et al., 2016; Hering et al., 2016).

Indel Analyses Using OutKnocker

Gene-specific FASTQ sequence files, reference amplicons, and target sites corresponding to the sense sgRNA for each gene in forward (F) and reverse (R) reads were uploaded into the OutKnocker indel analysis program (www.OutKnocker.org) (Schmid-Burgk et al., 2014). To standardize analyses, we applied a cut-off threshold of 5% read species for each individual clone. Read species that fell below the 5% threshold were not included in our analyses. To determine both indel effect on each clone (i.e., putative NULL, HET/MIXED, IN FRAME, UND) and overall editing efficiency for each gene via OutKnocker, we selected the read direction (F or R) that provided the most complete calls overall.

To determine frequencies of insertions versus deletions and putative in-frame versus frameshift indels, we tallied calls from all available OutKnocker data, including both F and R sequence reads to give an overall view of the indel landscape (with the exception of *CACNB2*, *CNOT1*, *SCN2A*, *ZSWIM6*, *CYP26B1*, *DPP4*, and *SYNGAP1* for which reads from only one direction were available). In all boxplots, the dark horizontal line represents the median, the box represents the 25th and 75th percentiles, the whiskers represent the 5th and 95th percentiles, and the circles represent outliers.

Indel Analyses Using BaySnpper

To determine indel effect and editing efficiency on selected genes with the BaySnpper indel calling tool, FASTQ files containing sequence data for F and R reads, along with the relevant barcode and gene reference amplicon sequences, were submitted to The Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Boston, MA, for processing.

Transcript and Nucleosome Position Analyses

For transcript analyses, edited and unedited percentages for each gene (n = 53) were plotted against corresponding FPKM values for HUES63 (Cacchiarelli et al., 2015). The genes *SHANK3*, *ZNF536*, *AKT3*, *GALNT10*, and *TBC1D5* were excluded from analyses as they were not present in the RNA-seq dataset. For nucleosome position analyses, edited and unedited percentages for each gene (n = 58) were plotted against the distance between the closest PAM sequence (NGG, where N = 0, bases 5′ of the PAM are negative, and bases 3′ of the N are positive) of either target site 1 or target site 2 for each gene (Table S1) to the nearest nucleosome dyad base pair as determined by MNase-seq for WA01 (Yazdi et al., 2015).

Resource Distribution

Clones derived from *CSMD1*, *CYFIP1*, *FXR1*, *SYNGAP1*, *TLE1*, and *TLE3* editing experiments performed in WA01 with no chromosomal aberrations detected are available upon request (Table S4). All sgRNA plasmids created in this study are available via Addgene (91573–91688). The plasmid pPN234 contains an additional 21 bp in the scaffold region, of which 18 bp directly match the *CYFIP1* target sequence, indicating that pPN234 contains a dimeric target sequence that will express a longer, dimeric sgRNA.

DATA AND SOFTWARE AVAILABILITY

The full BaySnpper pipeline can be accessed at https://github.com/dat4git/BaySnpper. The accession number for the next-generation sequencing data reported in this paper is NCBI SRA: SRP117350. Other data supporting the findings of this study are included within the paper and Supplemental Information.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and four tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2017.09.006.

AUTHOR CONTRIBUTIONS

K.E. and L.E.B. conceived the study and L.E.B., K.E., and D.Z.H. analyzed and interpreted data and wrote the manuscript with input from all co-authors. D.Z.H., A.B., A.M., and D.M. performed

the CRISPR-Cas9 design, cloning, and screening experiments. A.M.B., N.D., P.M., and D.L. performed the cell biology experiments.

ACKNOWLEDGMENTS

This project was funded by the Stanley Center for Psychiatric Research at the Broad Institute. We thank Yaoyu Wang and Alex Holman of the Center for Cancer Computational Biology, Dana-Farber Cancer Institute, Boston, MA, for assistance in developing the BaySnpper indel analysis pipeline. We also thank Kiki Lilliehook and members of the Barrett and Eggan labs for insightful discussions and critical reading of the manuscript, Sara Kosmaczewski and Kevin Smith for assistance with neuron generation, Feng Zhang for Cas9 plasmids, Tim Poterba and Sulagna Ghosh of the Broad Institute for assistance with SNP genotyping data, Luis Soares of Harvard Medical School for assistance with nucleosome analyses, and Florian Merkle of the University of Cambridge for technical advice and support. We also appreciate support from the Genomics Platform at the Broad Institute and the Bauer Flow Cytometry Core Facility at Harvard University.

Received: May 29, 2017
Revised: September 8, 2017
Accepted: September 11, 2017
Published: October 10, 2017

REFERENCES

Amit, M., Carpenter, M.K., Inokuma, M.S., Chiu, C.P., Harris, C.P., Waknitz, M.A., Itskovitz-Eldor, J., and Thomson, J.A. (2000). Clonally derived human embryonic stem cell lines maintain pluripotency and proliferative potential for prolonged periods of culture. Dev. Biol. *227*, 271–278.

Avior, Y., Sagi, I., and Benvenisty, N. (2016). Pluripotent stem cells in disease modelling and drug discovery. Nat. Rev. Mol. Cell Biol. *17*, 170–182.

Baker, D., Hirst, A.J., Gokhale, P.J., Juarez, M.A., Williams, S., Wheeler, M., Bean, K., Allison, T.F., Moore, H.D., Andrews, P.W., et al. (2016). Detecting genetic mosaicism in cultures of human pluripotent stem cells. Stem Cell Reports *7*, 998–1012.

Bara, A.M., Messana, A., Herring, A., Hazelbaker, D.Z., Eggan, K., and Barrett, L.E. (2016). Generation of a TLE3 heterozygous knockout human embryonic stem cell line using CRISPR-Cas9. Stem Cell Res. *17*, 441–443.

Basak, A., Hancarova, M., Ulirsch, J.C., Balci, T.B., Trkova, M., Pelisek, M., Vlckova, M., Muzikova, K., Cermak, J., Trka, J., et al. (2015). BCL11A deletions result in fetal hemoglobin persistence and neurodevelopmental alterations. J. Clin. Invest. *125*, 2363–2368.

Cacchiarelli, D., Trapnell, C., Ziller, M.J., Soumillon, M., Cesana, M., Karnik, R., Donaghey, J., Smith, Z.D., Ratanasirintrawoot, S., Zhang, X., et al. (2015). Integrative analyses of human reprogramming reveal dynamic nature of induced pluripotency. Cell *162*, 412–424.

Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat. Biotechnol. *27*, 275–280.

Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) *6*, 80–92.

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

Daer, R.M., Cutts, J.P., Brafman, D.A., and Haynes, K.A. (2016). The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. ACS Synth. Biol. *6*, 428–438.

Doench, J.G., Fusi, N., Sullender, M., Hegde, M., Vaimberg, E.W., Donovan, K.F., Smith, I., Tothova, Z., Wilen, C., Orchard, R., et al. (2016). Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat. Biotechnol. *34*, 184–191.

Durand, C.M., Betancur, C., Boeckers, T.M., Bockmann, J., Chaste, P., Fauchereau, F., Nygren, G., Rastam, M., Gillberg, I.C., Anckarsater, H., et al. (2007). Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. Nat. Genet. *39*, 25–27.

Flannick, J., and Florez, J.C. (2016). Type 2 diabetes: genetic data sharing to advance complex disease research. Nat. Rev. Genet. *17*, 535–549.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXivorg arXiv:1207.3907.

Golzio, C., Willer, J., Talkowski, M.E., Oh, E.C., Taniguchi, Y., Jacquemont, S., Reymond, A., Sun, M., Sawa, A., Gusella, J.F., et al. (2012). KCTD13 is a major driver of mirrored neuroanatomical phenotypes of the 16p11.2 copy number variant. Nature *485*, 363–367.

Hamdan, F.F., Gauthier, J., Spiegelman, D., Noreau, A., Yang, Y., Pellerin, S., Dobrzeniecka, S., Cote, M., Perreau-Linck, E., Carmant, L., et al. (2009). Mutations in SYNGAP1 in autosomal nonsyndromic mental retardation. N. Engl. J. Med. *360*, 599–605.

Hasan, M.S., Wu, X., and Zhang, L. (2015). Performance evaluation of indel calling tools using real short-read data. Hum. Genom. *9*, 20.

Hendriks, W.T., Warren, C.R., and Cowan, C.A. (2016). Genome editing in human pluripotent stem cells: approaches, pitfalls, and solutions. Cell Stem Cell *18*, 53–65.

Herring, A., Messana, A., Bara, A.M., Hazelbaker, D.Z., Eggan, K., and Barrett, L.E. (2016). Generation of a TLE1 homozygous knockout human embryonic stem cell line using CRISPR-Cas9. Stem Cell Res. *17*, 430–432.

Hockemeyer, D., and Jaenisch, R. (2016). Induced pluripotent stem cells meet genome editing. Cell Stem Cell *18*, 573–586.

Horlbeck, M.A., Witkowsky, L.B., Guglielmi, B., Replogle, J.M., Gilbert, L.A., Villalta, J.E., Torigoe, S.E., Tjian, R., and Weissman, J.S. (2016). Nucleosomes impede Cas9 access to DNA in vivo and in vitro. Elife *5*, e12677.

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.

International Stem Cell Initiative, Amps, K., Andrews, P.W., Anyfantis, G., Armstrong, L., Avery, S., Baharvand, H., Baker, J., Baker, D., Munoz, M.B., et al. (2011). Screening ethnically diverse human embryonic stem cells identifies a chromosome 20 minimal amplicon conferring growth advantage. Nat. Biotechnol. *29*, 1132–1144.

Isaac, R.S., Jiang, F., Doudna, J.A., Lim, W.A., Narlikar, G.J., and Almeida, R. (2016). Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. Elife *5*, e13450.

Keskin, H., Shen, Y., Huang, F., Patel, M., Yang, T., Ashley, K., Mazin, A.V., and Storici, F. (2014). Transcript-RNA-templated DNA recombination and repair. Nature *515*, 436–439.

Kessler, T., Vilne, B., and Schunkert, H. (2016). The impact of genome-wide association studies on the pathophysiology and therapy of cardiovascular disease. EMBO Mol. Med. *8*, 688–701.

Knight, S.C., Xie, L., Deng, W., Guglielmi, B., Witkowsky, L.B., Bosanac, L., Zhang, E.T., El Beheiry, M., Masson, J.B., Dahan, M., et al. (2015). Dynamics of CRISPR-Cas9 genome interrogation in living cells. Science *350*, 823–826.

Kriks, S., Shim, J.W., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., et al. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature *480*, 547–551.

Kuscu, C., Arslan, S., Singh, R., Thorpe, J., and Adli, M. (2014). Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat. Biotechnol. *32*, 677–683.

Leha, A., Moens, N., Meleckyte, R., Culley, O.J., Gervasio, M.K., Kerz, M., Reimer, A., Cain, S.A., Streeter, I., Folarin, A., et al. (2016). A high-content platform to characterise human induced pluripotent stem cell lines. Methods *96*, 85–96.

Liu, J.C., Lerou, P.H., and Lahav, G. (2014). Stem cells: balancing resistance and sensitivity to DNA damage. Trends Cell Biol. *24*, 268–274.

Lu, J., Zhong, X., Liu, H., Hao, L., Huang, C.T., Sherafat, M.A., Jones, J., Ayala, M., Li, L., and Zhang, S.C. (2016). Generation of serotonin neurons from human pluripotent stem cells. Nat. Biotechnol. *34*, 89–94.

Martins-Taylor, K., and Xu, R.H. (2012). Concise review: genomic stability of human induced pluripotent stem cells. Stem Cells *30*, 22–27.

Mayshar, Y., Ben-David, U., Lavon, N., Biancotti, J.C., Yakir, B., Clark, A.T., Plath, K., Lowry, W.E., and Benvenisty, N. (2010). Identification and classification of chromosomal aberrations in human induced pluripotent stem cells. Cell Stem Cell *7*, 521–531.

McKernan, R., and Watt, F.M. (2013). What is the point of large-scale collections of human induced pluripotent stem cells? Nat. Biotechnol. *31*, 875–877.

Merkle, F.T., Ghosh, S., Kamitaki, N., Mitchell, J., Avior, Y., Mello, C., Kashin, S., Mekhoubad, S., Ilic, D., Charlton, M., et al. (2017). Human pluripotent stem cells recurrently acquire and expand dominant negative P53 mutations. Nature *545*, 229–233.

Mertens, J., Marchetto, M.C., Bardy, C., and Gage, F.H. (2016). Evaluating cell reprogramming, differentiation and conversion technologies in neuroscience. Nat. Rev. Neurosci. *17*, 424–437.

Nguyen, H.T., Geens, M., and Spits, C. (2013). Genetic and epigenetic instability in human pluripotent stem cells. Hum. Reprod. Update *19*, 187–205.

Nishimura, Y., Martin, C.L., Vazquez-Lopez, A., Spence, S.J., Alvarez-Retuerto, A.I., Sigman, M., Steindler, C., Pellegrini, S., Schanen, N.C., Warren, S.T., et al. (2007). Genome-wide expression profiling of lymphoblastoid cell lines distinguishes different forms of autism and reveals shared pathways. Hum. Mol. Genet. *16*, 1682–1698.

O'Geen, H., Henry, I.M., Bhakta, M.S., Meckler, J.F., and Segal, D.J. (2015). A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Res. *43*, 3389–3404.

O'Roak, B.J., Vives, L., Fu, W., Egertson, J.D., Stanaway, I.B., Phelps, I.G., Carvill, G., Kumar, A., Lee, C., Ankenman, K., et al. (2012). Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. Science *338*, 1619–1622.

Pak, C., Danko, T., Zhang, Y., Aoto, J., Anderson, G., Maxeiner, S., Yi, F., Wernig, M., and Sudhof, T.C. (2015). Human neuropsychiatric disease modeling using conditional deletion reveals synaptic transmission defects caused by heterozygous mutations in NRXN1. Cell Stem Cell *17*, 316–328.

Paull, D., Sevilla, A., Zhou, H., Hahn, A.K., Kim, H., Napolitano, C., Tsankov, A., Shang, L., Krumholz, K., Jagadeesan, P., et al. (2015). Automated, high-throughput derivation, characterization and differentiation of induced pluripotent stem cells. Nat. Methods *12*, 885–892.

Perez-Pinera, P., Kocak, D.D., Vockley, C.M., Adler, A.F., Kabadi, A.M., Polstein, L.R., Thakore, P.I., Glass, K.A., Ousterout, D.G., Leong, K.W., et al. (2013). RNA-guided gene activation by CRISPR-Cas9-based transcription factors. Nat. Methods *10*, 973–976.

Peterson, S.E., and Loring, J.F. (2014). Genomic instability in pluripotent stem cells: implications for clinical applications. J. Biol. Chem. *289*, 4578–4584.

Pickard, B.S., Malloy, M.P., Clark, L., Lehellard, S., Ewald, H.L., Mors, O., Porteous, D.J., Blackwood, D.H., and Muir, W.J. (2005). Candidate psychiatric illness genes identified in patients with pericentric inversions of chromosome 18. Psychiatr. Genet. *15*, 37–44.

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., et al. (2013a). Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell *154*, 1380–1389.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013b). Genome engineering using the CRISPR-Cas9 system. Nat. Protocol. *8*, 2281–2308.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat. Biotechnol. *29*, 24–26.

Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., et al. (2015). Insights into autism spectrum disorder genomic

architecture and biology from 71 risk loci. Neuron *87*, 1215–1233.

Santos, D.P., Kiskinis, E., Eggan, K., and Merkle, F.T. (2016). Comprehensive protocols for CRISPR/Cas9-based gene editing in human pluripotent stem cells. Curr. Protoc. Stem Cell Biol. *38*, 5B.6.1–5B.6.60.

Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421–427.

Schmid-Burgk, J.L., Schmidt, T., Gaidt, M.M., Pelka, K., Latz, E., Ebert, T.S., and Hornung, V. (2014). OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. Genome Res. *24*, 1719–1723.

Shen, B., Zhang, W., Zhang, J., Zhou, J., Wang, J., Chen, L., Wang, L., Hodgkins, A., Iyer, V., Huang, X., et al. (2014). Efficient genome modification by CRISPR-Cas9 nickase with minimal off-target effects. Nat. Methods *11*, 399–402.

Singer, J.D., Gurian-West, M., Clurman, B., and Roberts, J.M. (1999). Cullin-3 targets cyclin E for ubiquitination and controls S phase in mammalian cells. Genes Dev. *13*, 2375–2387.

Srikanth, P., Han, K., Callahan, D.G., Makovkina, E., Muratore, C.R., Lalli, M.A., Zhou, H., Boyd, J.D., Kosik, K.S., Selkoe, D.J., et al. (2015). Genomic DISC1 disruption in hiPSCs alters Wnt signaling and neural cell fate. Cell Rep. *12*, 1414–1429.

Taapken, S.M., Nisler, B.S., Newton, M.A., Sampsell-Barron, T.L., Leonhard, K.A., McIntire, E.M., and Montgomery, K.D. (2011). Karotypic abnormalities in human induced pluripotent stem cells and embryonic stem cells. Nat. Biotechnol. *29*, 313–314.

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.

Too, L.K., Li, K.M., Suarna, C., Maghzal, G.J., Stocker, R., McGregor, I.S., and Hunt, N.H. (2016). Deletion of TDO2, IDO-1 and IDO-2 differentially affects mouse behavior and cognitive function. Behav. Brain Res. *312*, 102–117.

van Overbeek, M., Capurso, D., Carter, M.M., Thompson, M.S., Frias, E., Russ, C., Reece-Hoyes, J.S., Nye, C., Gradia, S., Vidal, B., et al. (2016). DNA repair profiling reveals nonrandom outcomes at Cas9-mediated breaks. Mol. Cell *63*, 633–646.

Wang, P., Lin, M., Pedrosa, E., Hrabovsky, A., Zhang, Z., Guo, W., Lachman, H.M., and Zheng, D. (2015). CRISPR/Cas9-mediated heterozygous knockout of the autism gene CHD8 and characterization of its transcriptional networks in neurodevelopment. Mol. Autism *6*, 55.

Wu, X., Scott, D.A., Kriz, A.J., Chiu, A.C., Hsu, P.D., Dadon, D.B., Cheng, A.W., Trevino, A.E., Konermann, S., Chen, S., et al. (2014). Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat. Biotechnol. *32*, 670–676.

Yazdi, P.G., Pedersen, B.A., Taylor, J.F., Khattab, O.S., Chen, Y.H., Chen, Y., Jacobsen, S.E., and Wang, P.H. (2015). Nucleosome organization in human embryonic stem cells. PLoS One *10*, e0136314.

Yi, F., Danko, T., Botelho, S.C., Patzke, C., Pak, C., Wernig, M., and Sudhof, T.C. (2016). Autism-associated SHANK3 haploinsufficiency causes Ih channelopathy in human neurons. Science *352*, aaf2669.

Yoon, K.J., Nguyen, H.N., Ursini, G., Zhang, F., Kim, N.S., Wen, Z., Makri, G., Nauen, D., Shin, J.H., Park, Y., et al. (2014). Modeling a genetic risk for schizophrenia in iPSCs and mice reveals neural stem cell deficits associated with adherens junctions and polarity. Cell Stem Cell *15*, 79–91.

Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron *78*, 785–798.