Stem Cell Reports
Article

OPEN ACCESS

# Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons

Eric Deneault,[1,2] Sean H. White,[3] Deivid C. Rodrigues,[4] P. Joel Ross,[4,7] Muhammad Faheem,[1,2] Kirill Zaslavsky,[4,5] Zhuozhi Wang,[2] Roumiana Alexandrova,[2] Giovanna Pellecchia,[2] Wei Wei,[4] Alina Piekna,[4] Gaganjot Kaur,[2] Jennifer L. Howe,[2] Vickie Kwan,[3] Bhooma Thiruvahindrapuram,[2] Susan Walker,[1,2] Anath C. Lionel,[1,2] Peter Pasceri,[4] Daniele Merico,[2,8] Ryan K.C. Yuen,[1,2] Karun K. Singh,[3,9,*] James Ellis,[4,5,9,*] and Stephen W. Scherer[1,2,5,6,9,*]

[1]Genetics & Genome Biology Program, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
[2]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
[3]Stem Cell and Cancer Research Institute, Department of Biochemistry and Biomedical Sciences, McMaster University, Hamilton L8S 4L8, Canada
[4]Developmental & Stem Cell Biology Program, The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
[5]Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 3H7, Canada
[6]McLaughlin Centre, University of Toronto, Toronto, ON M5S 3H7, Canada
[7]Present address: Department of Biology, University of Prince Edward Island, Charlottetown, PE C1A 4P3, Canada
[8]Present address: Deep Genomics Inc., Toronto, ON M5G 1M1, Canada
[9]Co-senior author
*Correspondence: singhk2@mcmaster.ca (K.K.S.), jellis@sickkids.ca (J.E.), stephen.scherer@sickkids.ca (S.W.S.)
https://doi.org/10.1016/j.stemcr.2018.10.003

## SUMMARY

Autism spectrum disorder (ASD) is phenotypically and genetically heterogeneous. We present a CRISPR gene editing strategy to insert a protein tag and premature termination sites creating an induced pluripotent stem cell (iPSC) knockout resource for functional studies of ten ASD-relevant genes (*AFF2/FMR2, ANOS1, ASTN2, ATRX, CACNA1C, CHD8, DLGAP2, KCNQ2, SCN2A, TENM1*). Neurogenin 2 (NGN2)-directed induction of iPSCs allowed production of excitatory neurons, and mutant proteins were not detectable. RNA sequencing revealed convergence of several neuronal networks. Using both patch-clamp and multi-electrode array approaches, the electrophysiological deficits measured were distinct for different mutations. However, they culminated in a consistent reduction in synaptic activity, including reduced spontaneous excitatory postsynaptic current frequencies in *AFF2/FMR2-*, *ASTN2-*, *ATRX-*, *KCNQ2-*, and *SCN2A*-null neurons. Despite ASD susceptibility genes belonging to different gene ontologies, isogenic stem cell resources can reveal common functional phenotypes, such as reduced functional connectivity.

## INTRODUCTION

Autism spectrum disorder (ASD) is a lifelong neurodevelopmental condition affecting reciprocal social interaction and communication, accompanied by restricted and repetitive behaviors (DSM-V, 2013). Familial clustering of ASD and related subclinical traits has been described, and with sibling recurrence risk estimates ranging from 8.1 to 18.7 (Gronborg et al., 2013; Ozonoff et al., 2011; Risch et al., 2014), a significant amount of familial liability is attributed to genetic factors (Colvert et al., 2015). Genomic microarray and sequencing studies have identified that ~10% of individuals have an identifiable genetic condition, and there are over 100 genetic disorders that can exhibit features of ASD, e.g., Fragile X and Rett syndromes (Betancur, 2011). Dozens of additional penetrant susceptibility genes have also been implicated in ASD (De Rubeis et al., 2014; Gilman et al., 2011; Pinto et al., 2014; Tammimies et al., 2015; Yuen et al., 2017), some being used in clinical testing (Carter and Scherer, 2013; Fernandez and Scherer, 2017). Genetically identified ASD-risk genes are enriched in broader functional groups consisting of synapse function, RNA processing, and transcriptional regulation (Bourgeron, 2015; De Rubeis et al., 2014; Geschwind and State, 2015; Pinto et al., 2014; Yuen et al., 2016, 2017). Importantly, so far, each risk gene or copy number variation (CNV) implicated in ASD accounts for <1% of cases, suggesting significant genetic heterogeneity (Yuen et al., 2017). Even within families, siblings can carry different penetrant mutations (Geschwind and State, 2015; Leppa et al., 2016; Yuen et al., 2015). Common genetic variants may also contribute to ASD risk (Weiner et al., 2017).

Until recently, postmortem brains were the only source of human cortical neurons to study directly the mechanistic and functional roles of ASD candidate genes *in vitro* (Varghese et al., 2017; Wintle et al., 2011). However, the still small numbers of biobanked brains, and the heterogeneous cellular content of the organ itself, as well as issues of cell viability and health are difficult to properly control in ASD (Anagnostou et al., 2014; de la Torre-Ubieta et al., 2016). The terminal differentiation status of mature neurons also precludes any potential *in vitro* studies in particular for early-onset conditions like ASD. Alternatively, somatic cells can be reprogrammed into induced pluripotent stem cells (iPSCs) that can grow indefinitely *in vitro* (Takahashi et al., 2007). Such patient-specific iPSCs provide

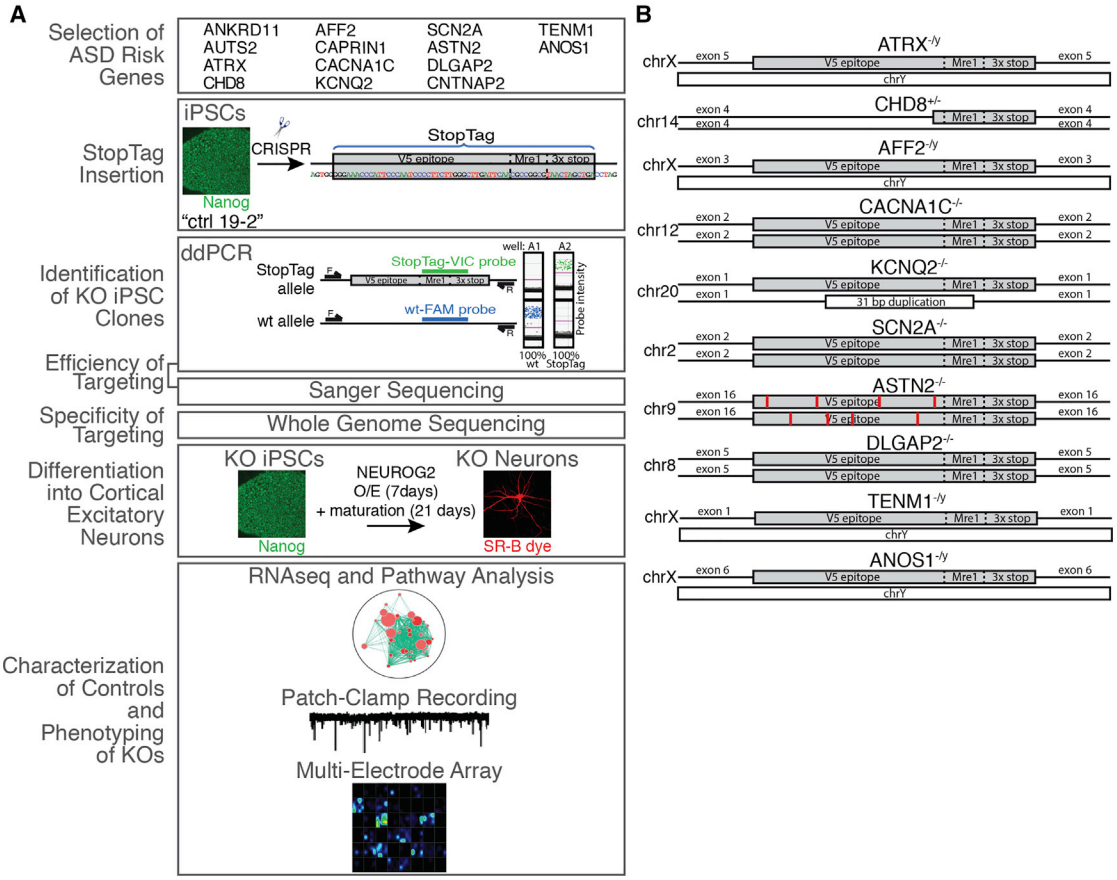


**Figure 1. Outline of the Experimental Procedure to Test the Phenotypical Consequences of Gene Repression in iPSC-Derived Glutamatergic Neurons**

(A) Unaffected human iPSC controls (ctrl) 19-2, labeled in green, were subjected to CRISPR gene editing to introduce a premature termination codon (StopTag), into a target exon of 14 ASD target genes. Knockout (KO) iPSC populations were identified by absolute quantification of StopTag versus wild-type (wt) alleles using droplet digital PCR (ddPCR). Well A1 is an example of a cell population containing 100% wt allele (FAM signal in blue) for a given target locus, while well A2 contains 100% StopTag alleles (VIC signal in green); FAM- and VIC-associated probe sequences are presented in Table S1. KO iPSCs were differentiated into glutamatergic neurons, labeled in red, by means of NGN2 transient overexpression (O/E). Neuronal phenotypes were monitored using RNA-seq, patch-clamp, and multi-electrode array recordings; F and R are ddPCR primers.

(B) Full-length integral StopTag sequence insertion was confirmed for all target genes except *CHD8*, in which the first 39 bp in 5′ of the StopTag sequence were deleted, and *ASTN2*, in which different point mutations were found (red bars); chr, chromosome; bp, base pair. See also Figures S1 and S2 and Tables S1–S3.

a newfound ability to study developmental processes, and functional characteristics, directly. Importantly, differentiation of human iPSCs into forebrain glutamatergic neurons may lead to model systems that recapitulate early molecular events in the trajectory of ASD development (Habela et al., 2016; Moretto et al., 2018). Directed induction into excitatory neurons can be achieved with high efficiency using transient ectopic expression of the transcription factor NGN2 (Zhang et al., 2013).

We devised a precise clustered regularly interspaced short palindromic repeats (CRISPR)-based strategy to efficiently generate complete knockout (KO) of any ASD-relevant gene, with all mutations made in the same "isogenic" (identical genetic background) human control iPSC line. We used the CRISPR/Cas9-mediated double-strand break mechanism coupled with error-free single-stranded template repair (SSTR) pathways (Miyaoka et al., 2014) to introduce an all-reading-frame premature termination codon (named "StopTag"; Figure 1A) into a specific exon of a target gene, designed to prevent stable RNA/protein product from being made. We hypothesized that a collection of isogenic KO lines carrying different ASD-risk null mutations would best minimize the confounding effects of genetic background. We then explored excitatory neuron

functional differences relevant to ASD for ten different successfully edited StopTag lines. Our results indicate that some ASD-risk genes display reduced synaptic activity between NGN2-derived excitatory neurons implying that ASD genes from different classes can present the same general cellular phenotype *in vitro*. We also highlight benefits and restrictions of studying ASD-risk genes using an isogenic human neural system.

## RESULTS

### Selection of ASD-Risk Genes

We selected 14 candidate ASD susceptibility genes from our ongoing whole-genome sequencing (WGS) project (the Autism Speaks MSSNG project), which aims to generate a list of penetrant genes for clinical diagnostics (Yuen et al., 2017). The evidence and priority of each gene, at the time of its selection, for having a role in ASD is described in Table 1, as is its assignment within three different functional groupings, i.e., transcriptional regulation, RNA processing, and synaptic and adhesion. We also note the importance of considering the results from the models described below to human mutation data in ASD individuals, which can present in dominant, recessive, and X-linked recessive forms, and be influenced by the sex of the carrier (Carter and Scherer, 2013; Geschwind and State, 2015).

### StopTag Insertion into ASD-Risk Genes and the Isogenic Control iPSC Line

We first used HEK293T cells to validate our SSTR-based strategy that involved introducing all-reading-frame premature termination codons (PTC; 3x stop; Figure S1A) into the DNA corresponding to an early constitutive exon for each gene (Table S1), aiming for a complete expression knockout. This insertion was delivered by a synthesized single-stranded oligodeoxynucleotide (ssODN) template (Table S1). The inserted fragment, called "StopTag" (Figures 1A and S1A), is 59 bp in length and includes a V5 epitope coding sequence. The left homology arm of ssODN was designed to insert the V5 epitope in phase with the original reading frame in order to allow the detection of truncated protein, with potential residual activity, which might have escaped non-sense mediated decay (NMD) following PTC insertion. PCR amplification confirmed the integration of the StopTag within specific target loci in *CHD8*, *DLGAP2*, and *KCNQ2* (Figure S1B).

In human iPSCs, the same StopTag insertion was used to knock out each of the 14 ASD-risk genes in a pluripotent and normal iPSC line named "Ctrl 19-2" (Figures 1A and 2A–2D). This line was reprogrammed from an unaffected father of a child with ASD who carries a *de novo* 16p11.2 microdeletion, associated with ASD (Marshall et al., 2008; Weiss et al., 2008). WGS confirmed that the unaffected father did not carry the 16p11.2 microdeletion, or any other known ASD-risk variants. Enrichment of 19-2-derived KO iPSCs was based on droplet digital PCR (ddPCR) coupled with dilution culture steps (Figure 1A and Table S2), adapted from Miyaoka et al. (2014) in order to maintain the polyclonality of cellular populations during enrichment. On average, 4.2 enrichment plates per line were necessary to purify to 100% KO (Table S2). In addition, we confirmed homozygous or hemizygous splicing of StopTag into all successful target loci (Figures 1B and S2), except for *KCNQ2* where the second allele presented a 31-bp frameshift duplication (Figure 1B and Table 1). No overt off-target mutation was found in any KO iPSC line (Table S3).

### Induction into Excitatory Neurons

Since genes involved in glutamatergic neurotransmission are associated with ASD (Autism Genome Project Consortium et al., 2007; Gilman et al., 2011), we sought to differentiate our KO iPSC lines into excitatory neurons to explore functional differences. We used the ectopic expression of NGN2 for iPSC differentiation (Figure 1A) in order to achieve homogeneous populations of neurons, and a reproducible differentiation protocol for the simultaneous assessment of multiple cell lines affected by different mutations. Using this protocol, induced neurons, when co-cultured with glial cells, display repetitive action potentials, large inward currents, and spontaneous synaptic activity by 21 days in culture (Zhang et al., 2013). We hypothesized that NGN2-induced mutant neurons could be used to monitor ASD state *in vitro*. We therefore induced NGN2 for 7 days and assessed electrophysiological properties of 19-2 control neurons using patch-clamp recordings of 21–28 days post-NGN2-induction (PNI) in the presence of cultured mouse glial cells. These 19-2 control neurons had a similar ability to fire repetitive action potentials (Figure 2E), and comparable input resistance, inward and outward currents (Figures 2F and S5A). Therefore, the NGN2 induction protocol was able to generate neurons to a similar maturation as previously reported (Yi et al., 2016; Zhang et al., 2013), enabling us to use the system to interrogate the function of ASD genes using isogenic KO iPSCs. Here, "neuron" refers to the NGN2-induced neuron.

### Transcriptional Characterization of Control 19-2 iPSCs and Neurons

We used RNA sequencing (RNA-seq) to verify the pluripotent state of the control 19-2 iPSCs, as well as the glutamatergic state of the control 19-2 neurons 4 weeks PNI cultured in the absence of glial cells. Transcript levels of 14 well-established pluripotency markers were high, i.e., from 33 to 783 reads per kilobase of transcript per million mapped reads (RPKM) in control iPSCs (Figure 2G).

**Table 1. List of 14 ASD Susceptibility Genes with Their Corresponding Functional Groupings, Transcript Levels in iPSCs, and Efficiency of Targeting**

| Gene | Chr | Molecular Function | Module | ASD Gene Reference | SFARI Score | OMIM Gene | Associated Disorders | ASD Syndrome | RPKM | Targeting Efficiency |
|---|---|---|---|---|---|---|---|---|---|---|
| *ANKRD11* | 16 | chromatin regulation | transcriptional regulation | Yuen et al., 2017 | 2 | 611192 | ADHD, EPS, ID | KBG syndrome | 26.2 | failed |
| *AUTS2* | 7 | chromatin binding | transcriptional regulation | Yuen et al., 2017 | 3 | – | ID, DD/ND | – | 16.6 | failed |
| *ATRX* | X | chromatin binding/ helicase | transcriptional regulation | Brett et al., 2014 | 4 | 300032 | EP, EPS | – | 17.7 | Hem |
| *CHD8* | 14 | chromatin binding/ helicase | transcriptional regulation | Yuen et al., 2017 | 1 | 610528 | SCZ, DD/ ND ID | – | 27.8 | Het |
| *AFF2* (FMR2) | X | RNA binding | RNA processing | Yuen et al., 2017 | 4 | 300806 | ID, ASD, EPS, EP, ADHD | fragile X syndrome | 1.7 | Hem |
| *CAPRIN1* | 11 | RNA binding | RNA processing | Jiang et al., 2013 | 3 | – | – | – | 110.9 | failed |
| *CACNA1C* | 12 | ion channel activity | synaptic and adhesion | Yuen et al., 2017 | S | 114205 | EPS, BPD, ID | Timothy syndrome | 0.2 | Hom |
| *KCNQ2* | 20 | ion channel activity | synaptic and adhesion | Yuen et al., 2017 | 3 | 602235 | ADHD, DD/ ND ID | – | 7.8 | Hom |
| *SCN2A* | 2 | ion channel activity | synaptic and adhesion | Yuen et al., 2017 | 1 | 182390 | EPS, DD/ND ADHD, ID, EP | – | 0.3 | Hom |
| *ASTN2* | 9 | calcium ion binding | synaptic and adhesion | Lionel et al., 2014 | 3 | – | ID, EPS, DD/ ND ADHD | – | 2.9 | Hom |
| *DLGAP2* | 8 | synapse | synaptic and adhesion | Marshall et al., 2008 | 4 | – | – | – | 0.2 | Hom |
| *CNTNAP2* | 7 | Cell adhesion | synaptic and adhesion | Bakkaloglu et al., 2008 | 2S | 604569 | ADHD, EP, EPS, ID | cortical dysplasia-focal epilepsy syndrome | 19.3 | failed |
| *TENM1* (ODZ1) | X | Cell adhesion/ signal transducer | synaptic and adhesion | Yuen et al., 2017 | NA | – | – | – | 0.2 | Hem |
| *ANOS1* (KAL1) | X | extracellular matrix | synaptic and adhesion | Jiang et al., 2013 | NA | 300836 | – | Kallmann syndrome | 35.8 | Hem |

Some genes like *ANOS1* and *TENM1* were considered stronger candidates for having a role in ASD early in the study, but this changed as additional genetic studies were published. SFARI, Simons Foundation Autism Research Initiative; OMIM, Online Mendelian Inheritance in Man; ADHD, attention deficit with hyperactivity disorder; EPS, extrapyramidal symptoms; ID, intellectual disability; ASD, autism spectrum disorder; EP, epilepsy; SCZ, schizophrenia; DD/NDD, developmental delay/neurodevelopmental disorder; BPD, bipolar disorder; RPKM, reads per kilobase per million mapped reads in iPSCs; Hem, hemizygous; Het, heterozygous; Hom, homozygous; NA, not available. See also Figure S2 and Table S2.

Expression of the same set of pluripotency markers was low in neurons, i.e., from 0 to 14 RPKM. We observed a similar pattern of expression to that published originally in NGN2-generated neurons (Zhang et al., 2013). For example, we recorded higher levels of the cortical markers *POU3F2* and *FOXG1* than those of *TBR1* (Figure 2G). AMPA receptor subunits *GRIA1*, *GRIA2*, and *GRIA4* were highly represented at >10 RPKM, compared with the NMDA receptor subunit *GRIN1* at 1.5 RPKM, and transcript levels for the glutamate transporter *SLC17A6* were high, while those of *SLC17A7* were much lower (Figure 2G). GABA receptor subunit *GABRA2* transcript levels were seen at higher levels compared with those of the GABA transporter *SLC32A1* (Figure 2G). We also observed high levels of the neuronal markers *MAP2*, *TUBB3*, and *NCAM1*, but near-zero levels of the astrocyte marker *GFAP* (Figure 2G). These data reflect

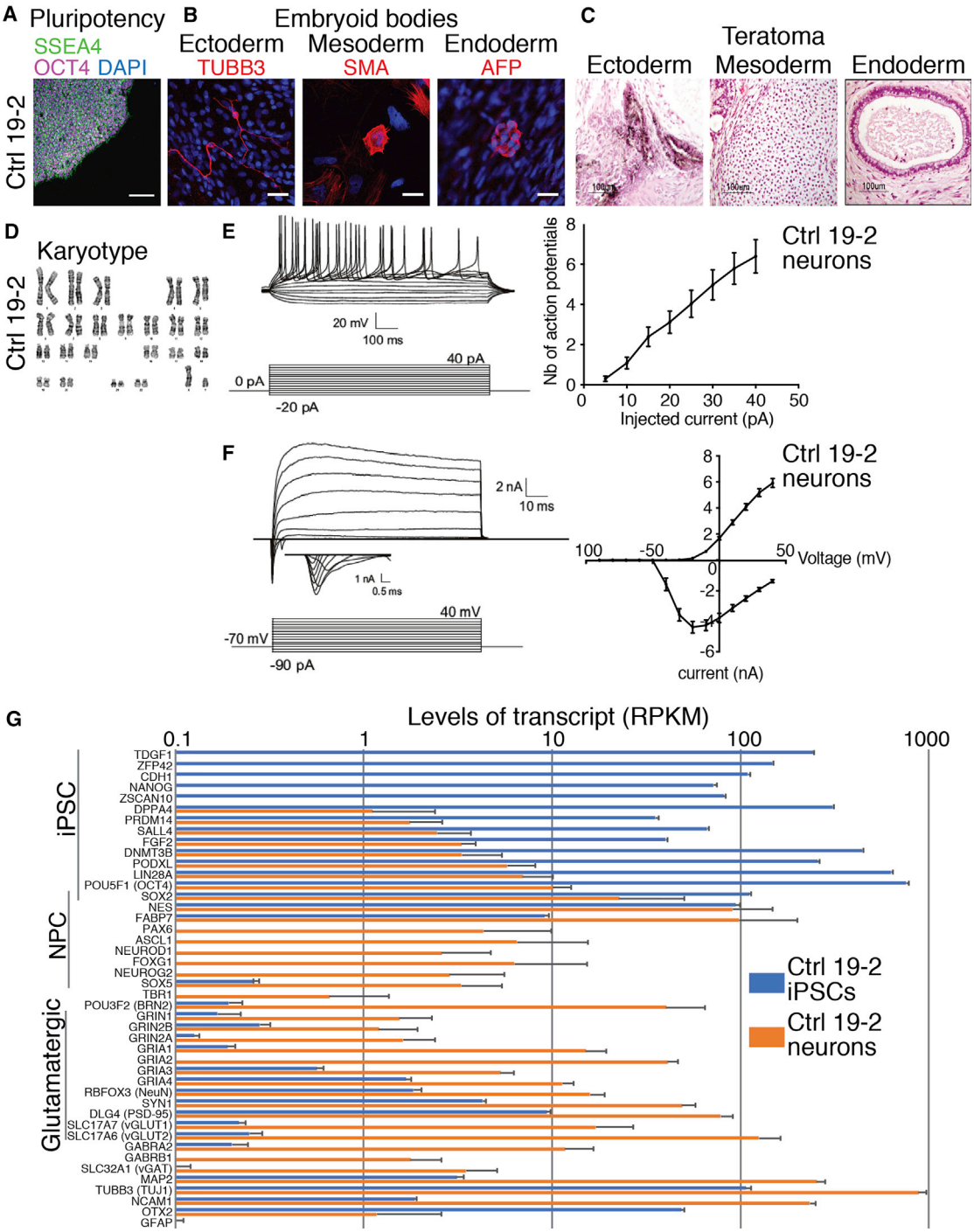


**Figure 2. Characterization of the Control 19-2 iPSCs and Neurons**

(A–D) Representative microscopic images show normal iPSC (A) pluripotency (SSEA4 and OCT4), (B) differentiation potential into the three germ layers *in vitro* (embryoid bodies: TUBB3, SMA, and AFP) or (C) *in vivo* (teratoma assays), and (D) karyotype. See also Figure S4. Scale bars: 100 μm (A), 25 μm (B), and 100 μm (C).

(E) Representative traces of action potentials recorded at different current injection on the left panel, and the number of action potentials is plotted for each step of current injection on the right panel.

*(legend continued on next page)*

a homogeneous NGN2-derived glutamatergic neuron profile, suitable for phenotyping.

### StopTag Leads to Complete KO of Target Genes

We ensured that the StopTag sequence was properly transcribed and fused to the different target transcripts. For example, a part of some reads generated from RNA-seq did not align with exon 2 of *SCN2A* on the human reference genome hg19 (Figure 3A), corresponding exactly to the StopTag sequence, which should cause premature termination of translation. We observed a significant reduction in transcript levels of five target genes (Figure 3B). For instance, transcript levels of *ATRX* were reduced by ~50% in both *ATRX*$^{-/y}$ iPSCs and neurons compared with control cells (Figure 3B). This suggests that some mutant transcripts were eliminated by NMD before complete translation due to the presence of a PTC, brought by StopTag. However, we did not observe such reduction in transcript levels for the synaptic genes *CACNA1C*, *KCNQ2*, *SCN2A*, *DLGAP2*, and *TENM1* (Figure 3B and Table S4), suggesting that these transcripts may escape NMD at least until the time of RNA extraction. Despite transcript levels, western blot analysis confirmed the absence of target proteins in mutant neurons. For example, the major protein form of *ATRX* was detected at 280 kDa in control 19-2 neurons, but was not detectable in *ATRX*$^{-/y}$ (Figure 3C).

Notwithstanding, some StopTag transcripts escaping NMD might still be translated as truncated proteins and not recognized by antibodies. Such peptides are undesirable in a KO setup since they can present some residual activity or cause other unintended damage (Kamiya et al., 2004; Luo et al., 2017). In order to reveal the presence of any truncated form of proteins, we flanked a sequence coding for a V5 epitope upstream of the 3x stop within the StopTag fragment (Figure 1A). A perfect assemblage was confirmed by Sanger sequencing for most target genes (Figure 1B). No truncated forms of protein were detectable by western blot using a V5 antibody (Figure 3D), indicating complete absence of target proteins in the testable KO neurons.

### Transcriptional Characterization of KO iPSCs and Neurons

RNA-seq profiling of the ten KO iPSC lines demonstrated their pluripotency and differentiation into glutamatergic neurons. Major iPSC markers, e.g., *NANOG* and *POU5F1*, were highly expressed exclusively in iPSCs (Figure S3). Alternatively, specific neuronal markers, e.g., *MAP2* and *SLC17A6* (vGLUT2) were found expressed only in neurons (Figure S3). Deficient *GFAP* expression indicated the absence of glial cells in our neuronal cultures (Figure S3).

### RNA-Seq and Pathway Analysis

Transcriptional co-profiling of several isogenic KO lines can reveal how different genes, belonging to different functional groupings, might regulate the expression of common gene sets or pathways. From the different lists of differentially expressed genes (DEGs) in KO iPSCs compared with controls (Table S5), we explored common pathway enrichment shared by at least three of the ten KO lines. Several different gene ontology terms and pathways associated with "neuron projection development" presented a similar profile in different KO iPSC lines (Figure 4A). For example, most of the corresponding gene sets were downregulated in *ATRX*-, *ASTN2*-, and *DLGAP2*-, while they were upregulated in *AFF2*-null iPSCs (Figure 4A). A different profile was observed with the "anchored component of membrane" pathway, i.e., predominantly upregulated in *AFF2*- and *SCN2A*-, while downregulated in *ATRX*- and *ASTN2*-null lines (Figure 4A). Another group of gene sets, associated with "negative regulation of transcription" were commonly downregulated in *ATRX*-, and upregulated in *AFF2*- and *SCN2A*-null iPSCs (Figure 4A). In general, KO of *CACNA1C* or *KCNQ2* did not have significant impact on transcriptional networks in iPSCs (Figure 4A). This may suggest that some of our KO iPSC lines already show transcriptional networks that are prone to ASD-related changes.

We also mined RNA-seq data for common mechanisms in our iPSC-derived NGN2 neurons. Since the number of DEGs passing a false discovery rate (FDR) <0.05 was substantially lower in neurons overall compared with iPSCs, we decided to lower the significance thresholds for some of the KO lines and searched for pathway enrichments, as explained in Table S6. Gene ontology terms with a Benjamini-Hochberg false discovery rate (BH-FDR) <0.05, and associated with "neurons" or "synapses," were found in high numbers only in *ATRX*$^{-/y}$ and *TENM1*$^{-/y}$ lines (Table S6). Therefore, we searched for common DEGs, instead of gene sets, found in at least two of ten mutant neuron lines.

(F) Representative traces of sodium currents on the left panel, and currents were recorded at different potentials in voltage-clamp on the right panel; 33 control 19-2 neurons were recorded from three independent differentiation experiments at day 21–28 post-NGN2-induction (PNI).

(G) Transcript levels in RPKM of a series of iPSC, neural progenitor cell (NPC), and neuron markers in control 19-2 iPSCs (blue) and control 19-2 glutamatergic neurons (orange). Values are presented as means ± SD of eight independent experiments for iPSCs and four for neurons.

See also Figure S3.

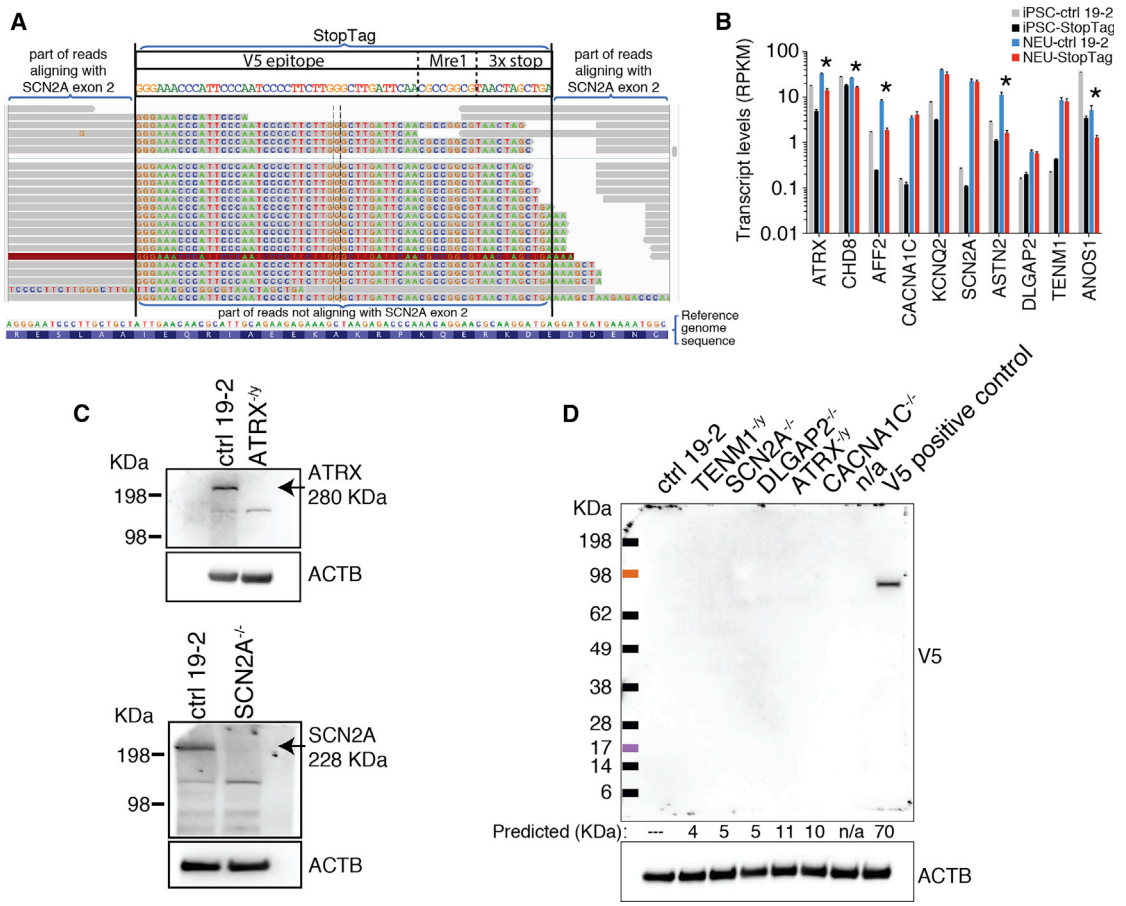


**Figure 3. Complete KO of Target Gene Expression in Neurons**

(A) Example of a target locus, i.e., exon 2 of *SCN2A*, where Spliced Transcripts Alignment to a Reference (STAR) software was used to align reads previously unmapped by TopHat. The gray part of the reads mapped to the human reference genome hg19. The colored part did not map to hg19 but aligned perfectly with the StopTag sequence, showing it is properly transcribed and fused to the target transcripts.

(B) Transcript levels in reads per kilobase per million mapped reads (RPKM; y axis) of each target gene (x axis) in their corresponding KO iPSCs (black bars) and KO neurons (red bars), as well as in control iPSCs (gray bars) and control neurons (blue bars). Values are presented as means ± SD of 2–8 independent experiments; *FDR <0.05 in neurons. See also Table S4.

(C) Western blots showing the absence of the major form of ATRX (upper panel) and SCN2A (bottom panel) proteins in their corresponding KO neurons, compared with control 19-2 neurons, 4 weeks PNI; ACTB, loading control.

(D) Western blots revealing the absence of any truncated form of proteins in different KO neuron lines, using a V5 antibody. Predicted (kDa), predicted size of potentially truncated peptides based on the insertion sites of the StopTag within each target transcript relative to the start codon position; n/a, not applicable; ACTB and TUBB3, loading controls; n/a, not available.

Several DEGs were shared between distinct KO neuron lines. For instance, *ZNF558* was downregulated in *ATRX*-, *AFF2*-, *KCNQ2*-, *SCN2A*-, *ASTN2*-, and *DLGAP2*-null neurons (Figure 4B). Conversely, *REG1A* was upregulated in *ATRX*-, *AFF2*-, and *TENM1*-null neurons (Figure 4B). Like in iPSCs, KO of *CACNA1C* did not have a major impact on the neuronal transcriptional network (Figure 4B). Interestingly, many upregulated DEGs shown in Figure 4B are also members of the cadherin superfamily PCDH involved in synapse configuration (Hirano and Takeichi, 2012). Inversely, several downregulated DEGs belong to the $C_2H_2$ zinc finger superfamily ZNF of putative transcription factors (Liu et al., 2014) (Figure 4B).

On a per line basis, we searched for enrichment in ASD-associated genes among the DEGs from each KO neuron line. For this, we used a proprietary list of 736 genes (highlighted in blue in Table S7) associated with ASD (Abrahams et al., 2013; Pinto et al., 2014; Yuen et al., 2017). The DEGs overlapping this list are highlighted in yellow for each of our KO neuron lines in Table S7. The number of ASD-associated genes that are differentially expressed is clearly higher in *ATRX*$^{-/y}$ or in *TENM1*$^{-/y}$ neurons (n > 30)

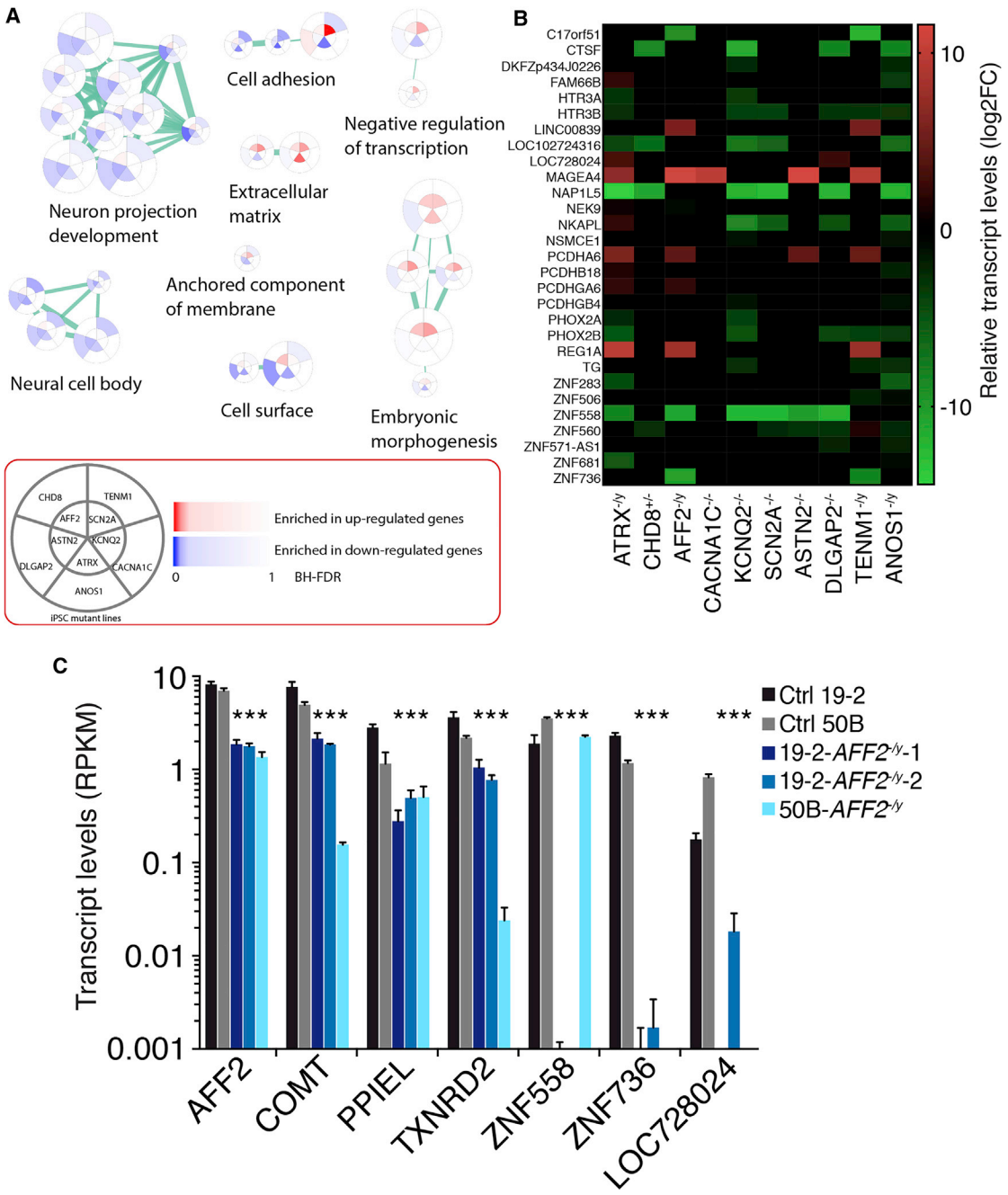


**Figure 4. Transcriptional Characterization of KO iPSCs and Neurons**

(A) Enrichment map of differentially expressed genes in different mutant iPSCs compared with the isogenic control 19-2, as revealed by RNA-seq and pathway analysis. For example, the upper right piece of each pie represents the *SCN2A*-null iPSC line (see inset), in which different gene sets associated with "embryonic morphogenesis" are upregulated (red color), with respect to the control 19-2 line. The color intensity correlates with the Benjamini-Hochberg false discovery rate (BH-FDR) values, as depicted in the inset. Values were calculated from (n) independent experiments; n = 8 for ctrl 19-2; n = 4 for $ATRX^{-/y}$, $KCNQ2^{-/-}$, $SCN2A^{-/-}$, and $ASTN2^{-/-}$; n = 2 for $AFF2^{-/y}$.

(B) Differentially expressed genes in different KO neurons compared with the isogenic control 19-2 as revealed by RNA-seq. Values are presented as log2 fold change. Values were calculated from (n) independent differentiation experiments; n = 5 for $AFF2^{-/y}$; n = 4 for ctrl 19-2, $ATRX^{-/y}$, $KCNQ2^{-/-}$, $SCN2A^{-/-}$; n = 3 for $ASTN2^{-/-}$.

*(legend continued on next page)*

compared with each of the other KO lines (n < 5; Table S7), in agreement with the number of "neuron"-associated terms in the pathway enrichment analysis (Table S6). Interestingly, among the ASD-associated genes in yellow, some are concomitantly differentially expressed in separate KO lines, e.g., *GPR139* and *HTR3A*, in *ATRX*$^{-/y}$, *KCNQ2*$^{-/-}$, and *ASTN2*$^{-/-}$ neurons (Table S7). The differential expression analyses in both iPSCs and neurons suggest common pathways and DEGs that converge toward different synaptic factors potentially affecting functional activity of neurons (Figures 4A and 4B).

### RNA-Seq Validation in a Different Genetic Background

We also suppressed the expression of *AFF2* in a different and unrelated iPSC line, namely 50B, which presented a normal pluripotency and karyotype (Figure S4). In addition to this new KO line, i.e., 50B-*AFF2*$^{-/y}$, we generated a second KO line in the 19-2 background using the same method as the first 19-2-*AFF2*$^{-/y}$ line. For optimal homogeneity of neuronal cultures, we used the same NGN2 induction protocol to obtain iPSC-derived glutamatergic neurons, on which we performed RNA-seq and differential expression analysis. We searched for significant DEGs in common with all these three *AFF2*-null lines, i.e., 19-2-*AFF2*$^{-/y}$-1, 19-2-*AFF2*$^{-/y}$-2, and 50B-*AFF2*$^{-/y}$, compared with their respective control lines 19-2 and 50B. At least seven DEGs met these criteria, i.e., *AFF2*, *COMT*, *PPIEL*, *TXNRD2*, *ZNF558*, *ZNF736*, and *LOC728024* (Figure 4C). These results confirm that at least these genes show transcript levels that are influenced by *AFF2*, and not only by the specific 19-2 genetic background. Accordingly, we found a low total number of DEGs in both genetic backgrounds (Figure S4C), which is in line with the absence of evidence for a transcriptional role of *AFF2* in the literature.

### Patch-Clamp Electrophysiological Analysis of KO Neurons

To determine if neuronal behavior was affected by any specific gene KO, we performed patch-clamp recordings on all isogenic KO neuron lines 21–28 days PNI, in co-culture with astrocytes to improve maturation. As shown in Figure S5A, overall we did not detect consistent changes in baseline properties in neurons, indicating a similar level of maturity for all KO neurons compared with controls and that individual alterations were not specific to a particular gene. *SCN2A*-deficient neurons were a notable exception since they were significantly altered in all parameters (Figure S5A), which is consistent with the known role of sodium currents in regulating neuronal excitability in mouse postnatal neurons (Planells-Cases et al., 2000).

We also measured neuronal activity using patch-clamp electrophysiology as this is perhaps more relevant to ASD. We found that overall, there was a significant reduction in neuronal activity based on the observation of significantly reduced spontaneous excitatory postsynaptic current (sEPSC) frequency in *ATRX*-, *AFF2*-, *KCNQ2*-, *SCN2A*-, and *ASTN2*-null neurons, without a corresponding change in amplitude (Figure 5A). Interestingly, *CACNA1C*$^{-/-}$, *DLGAP2*$^{-/-}$, and *ANOS1*$^{-/y}$ neurons also displayed non-significant reductions in sEPSC frequency (Figure 5A). Conversely, *TENM1*$^{-/y}$ tended toward a non-significant higher frequency of sEPSC compared with the isogenic control neurons (Figure 5A). Together, these data indicate that ASD-risk genes from different classifications can produce a similar electrophysiological phenotype.

### Multi-electrode Array Analysis of KO Neurons

The increased density of neuronal processes >28 days PNI prevented consistent clean patch-clamp recordings for longer time periods. Nonetheless, a multi-electrode array (MEA) device allows for the ability to monitor the excitability of a population of neurons in an unbiased manner by incorporating the activity of all of individual neurons in a whole well, and for long time periods in the same culture plate. We monitored the spontaneous neuronal network activity 4–8 weeks PNI using MEAs in 48-well format, in co-culture with astrocytes to improve maturation. Field recordings were acquired to estimate the activity in each well. We focused on the five genes that significantly reduced neuronal activity in Figure 5A, i.e., *ATRX*, *AFF2*, *KCNQ2*, *SCN2A*, and *ASTN2*. We sought to determine if any of these KOs would interfere with synchronized bursting events on a neuronal population level. Since the highest level of mean firing rate (MFR) was observed 8 weeks PNI in control 19-2 neurons (Figure S5B), we opted to use this time point for comparison of all KO lines. We also confirmed that most of the cultures were composed of glutamatergic neurons using different receptor inhibitors (Figure S5C).

The MFR and network burst frequency were significantly reduced in *SCN2A*-null neurons compared with control (Figure 5B). Moreover, the burst frequency was lower in *ATRX*-, *AFF2*-, *KCNQ2*-, and *SCN2A*-null neurons (Figure 5B), suggesting that loss of function of these genes

(C) RNA-seq validation of transcript levels of seven genes in three different *AFF2*-null neuron lines, i.e., 19-2-*AFF2*$^{-/y}$-1, 19-2-*AFF2*$^{-/y}$-2, and 50B-*AFF2*$^{-/y}$, compared with their respective control lines 19-2 and 50B. Values are presented as means ± SD of (n) independent differentiation experiments where n = 4 for controls 19-2 and 50B; n = 5 for 19-2-AFF2$^{-/y}$-1; n = 3 for 19-2-AFF2$^{-/y}$-2 and 50B-AFF2$^{-/y}$; *each KO line value has an FDR <0.5 compared with its respective control line.
See also Figure S4 and Tables S5–S7.

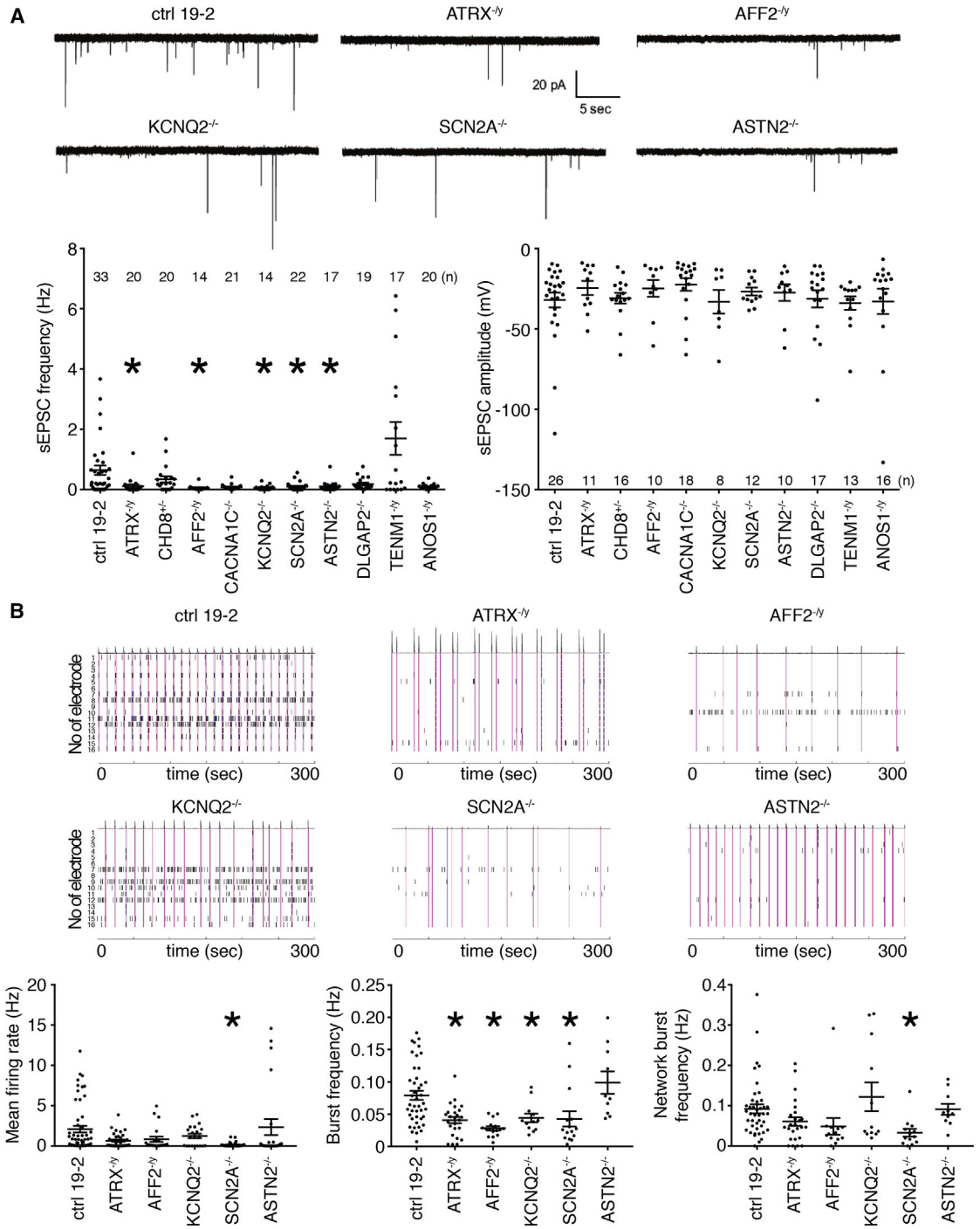


**Figure 5. Electrophysiological Phenotyping of KO iPSC-Derived Neurons**

(A) Representative traces of excitatory postsynaptic current (EPSC; top panel); spontaneous EPSC frequency (lower left panel) and amplitude (lower right panel) were recorded from different KO neurons; total number of recorded neurons is indicated on the graphs; values are presented as means ± SEM of three independent differentiation experiments for all, except two for *AFF2* and *KCNQ2*, recorded at day 21–28 PNI.

(B) Representative raster plots over a 5-min recording of multi-electrode array experiments (top panels); mean firing rate (lower left panel), burst frequency (lower middle panel), and network burst frequency (lower right panel) were recorded for the five significant genes from the sEPSC frequency graph in (A). Each spike is represented with a short black line. A burst was considered as a group of at least five

*(legend continued on next page)*

affects extracellular spontaneous network activity in glutamatergic neurons and reduces neuronal activity on a population level over a longer period of time. The absence of *AFF2* in a different genetic background, i.e., 50B-*AFF2*$^{-/y}$, also led to electrophysiological deficits (Figure S5D), confirming its role in neuronal activity. Importantly, these data are overall consistent on a per gene level compared with the sEPSC data obtained from individual neurons in Figure 5A, except *ASTN2*, and suggest that ASD-risk genes converge to disrupt excitatory neuronal activity.

## DISCUSSION

We developed a new CRISPR-based strategy, named StopTag insertion, to completely and specifically knock out the expression of ASD-relevant genes. This approach allows the systematic generation of multiple isogenic KO lines and pairs well with the highly consistent NGN2 induction to study excitatory neurons. This combination is suitable for deep phenotyping of several isogenic human glutamatergic neuron cultures, which were shown to be relevant to ASD in several studies (Habela et al., 2016; Moretto et al., 2018). The approach revealed that a common phenotype between some "functionally" diverse ASD-risk genes is reduced synaptic activity. We demonstrated that electrophysiological properties of *ATRX*-, *AFF2*-, *KCNQ2*-, *SCN2A*-, and *ASTN2*-mutant neurons were severely compromised using two complementary electrophysiology approaches, i.e., patch-clamp and MEA, which allowed for more convincing conclusions when combined together. We also showed that ASD genes from different classes display disruption of common signaling networks that are associated with neuron projection and synapse assembly. These results indicate that aberrant functional connectivity is a frequent phenotype in human neurons with ASD candidate gene null mutations.

We obtained a high rate of biallelic editing, i.e., four of the six successful autosomal genes *CHD8*, *CACNA1C*, *KCNQ2*, *SCN2A*, *ASTN2*, and *DLGAP2*. This is possibly due to the use of plasmids to express Cas9 and gRNAs, which are expected to be stable for a few days following nucleofection. We reasoned that the impact of a full KO on transcriptional networks might be more significantly detected compared with heterozygous cells, especially for ASD-relevant genes that act under a dominant model, and/or other genes exquisitely sensitive to dosage alterations. It will be worthwhile to attempt to make heterozygous mutants when possible, using different vectors such as the RNP complex, to compare the heterozygous state of ASD variants. However, many of the ASD-risk genes are located on chromosome X, and since 19-2 cells are of male origin, these genes were fully inactivated after one allele was targeted. A notable exception was *CHD8*, which was heterozygous. Despite an attempt to target our *CHD8*$^{+/-}$ line in a second round of CRISPR experiments, we failed to isolate any biallelic *CHD8*-null line. It is possible that at least one copy of *CHD8* is indispensable for survival, as proposed previously by other teams regarding mice (Nishiyama et al., 2004) and humans (Bernier et al., 2014). Since no significant results were obtained in this study with *CHD8*$^{+/-}$ cells, some specific phenotypes might be more difficult to replicate with heterozygous loss-of-function mutations, possibly due to lower penetrance.

The NGN2-neuron system may represent an advantage compared with classic dual-SMAD inhibition differentiation protocols for phenotyping experiments that require relatively high levels of cell homogeneity. Dual-SMAD inhibition neurons are more heterogeneous, including several different types of cells, they require a longer time to mature, and may not be ideal for higher-throughput studies where multiple genes need to be compared. We found that scalability and consistency are improved by the NGN2 induction protocol. Moreover, it allows the neuronal phenotype to be tested directly and eliminates any contribution from mutant glial cells. However, one caveat to the NGN2 approach is the lack of inhibitory neurons, which may be important to analyze particular ASD genes expressed in this population, e.g., *ARID1B* (Jung et al., 2017). In addition, our data suggest that a disruption in the excitation/inhibition balance contributes to ASD (Hussman, 2001; Rubenstein and Merzenich, 2003). However, to test this directly, the effect on inhibitory neurons will have to be evaluated in the future (Yang et al., 2017).

Interestingly, the five genes with a common sEPSC reduced neuronal activity phenotype (Figure 5A) fall into different molecular function groupings, i.e., *ATRX* in transcriptional regulation, *AFF2/FMR2* in RNA processing, and *KCNQ2, SCN2A, ASTN2* in synaptic and adhesion (Table 1). And KO of three of them, i.e., *ATRX*, *AFF2*, *SCN2A* (from three different groupings), converge to common transcriptional networks associated with, for

spikes, each separated by an inter-spike interval (ISI) of no more than 100 ms. A network burst (pink lines) was identified as a minimum of ten spikes with a maximum ISI of 100 ms, occurring on at least 25% of electrodes per well. From 21 to 55 different wells were recorded per line, with usually 6–8 wells per line per experiment. Values are presented as means ± SEM from (n) independent differentiation experiments, where n = 8 for ctrl 19-2; n = 6 for *ATRX*$^{-/y}$; n = 5 for *SCN2A*$^{-/-}$; n = 4 for *AFF2*$^{-/y}$, *ASTN2*$^{-/-}$, and *KCNQ2*$^{-/-}$; recorded at week 8 PNI; *p < 0.05 from one-way ANOVA (Dunnett multiple comparison test). n/d, not detected.
See also Figure S5.

example, "transcription regulation," "membrane components," and "embryonic morphogenesis" in iPSCs (Figure 4A), suggesting that early developmental events may be involved in the ASD trajectory. We also found common DEGs among the five sEPSC-featured mutant neurons, e.g., *PCDHA6*, *REG1A*, and *ZNF558* (Figure 4B), that might be responsible for this common phenotype. These results suggest that RNA-seq could be used to detect specific transcriptional signatures associated with ASD. Several DEGs were also shared between neurons mutant for genes from the same group, e.g., *KCNQ2* and *SCN2A*, which are both ion channel subunits; and between neurons mutant for *ATRX* and *AFF2*, which are both X-linked genes involved in intellectual disability and binding G4-quadruplexes associated with DNA and RNA, respectively (Bensaid et al., 2009; Law et al., 2010) (Figure 4B). Although the total number of DEGs was relatively low in our KO neurons, some were validated in a different genetic background (line 50B in Figure 4C) and were previously associated with ASD, including *COMT*, which was associated with schizophrenia (Lee et al., 2005). Different reasons to account for the low number of DEGs found overall in neurons can be posited. For instance, despite relative cell homogeneity obtained with the NGN2 induction system, we believe that our neuron cultures presented higher levels of variation in cell composition compared with our iPSC cultures, and that such variation decreased the statistical power to find DEGs. Another reason may be that the impact on transcriptional networks occurs prior to terminal differentiation of NGN2 neurons, e.g., on proliferation of neural progenitor cells (NPCs) as previously shown with *Chd8* in mice (Durak et al., 2016).

In conclusion, we have designed a CRISPR gene editing strategy for complete KO of ten ASD-risk genes of varying function and shown that they can be responsible for similar transcriptional rewiring and electrophysiological phenotypes in human iPSC-derived glutamatergic neurons. Overall, given the heterogeneity involved in ASD (Yuen et al., 2015, 2017), we believe that this type of CRISPR-isogenic KO system may be essential for stepwise controlled cellular phenotyping experiments. Whole-animal murine KO models would also be useful but the same efficiencies would not be possible. For future experiments, our isogenic KO system could also be used for the incremental creation of isogenic knockin lines where other ASD patient mutations, combinatorically, could be introduced, creating a resource for advanced functional modeling and new therapeutic testing.

## EXPERIMENTAL PROCEDURES

Reprogramming of iPSCs was performed under the approval of the Canadian Institutes of Health Research Stem Cell Oversight Committee, and the Research Ethics Board of The Hospital for Sick Children, Toronto. Reprogramming was performed from skin fibroblasts or blood cells using retrovirus, lentivirus, or Sendai virus to deliver the reprogramming factors, and iPSCs were maintained in mTeSR (STEMCELL Technologies); see Supplemental Information. The type II CRISPR/Cas9 double-nicking (Cas9D10A) (Ran et al., 2013) system with dual guide RNA (Table S1), or ribonucleoprotein (RNP complexes) was used to insert the StopTag fragment in the presence of an ssODN template, and off-target events were searched using WGS (details in the Supplemental Information). Isolation of edited iPSCs was based on ddPCR and enrichment steps and was adapted from Miyaoka et al. (2014) (see Supplemental Information). Induction of iPSCs into glutamatergic neurons was conducted using overexpression of NGN2 for a week as described in Supplemental Information. For RNA-seq, RNA libraries were prepared using NEBNext Ultra RNA Library Preparation kit for Illumina, and total RNA was used for poly(A) mRNA enrichment; pathway enrichment analysis was performed using the R package goseq version 1.28.0 and R version 3.4.1 (June 30, 2017) using a custom gene-set collection including gene ontology (GO, obtained from the R package GO.db version 3.4.1) and pathways (KEGG and Reactome collections downloaded from the respective websites on October 16, 2017); see Supplemental Information for more details. In electrophysiology, whole-cell recordings (BX51WI; Olympus) were performed at room temperature using an Axoclamp 700B amplifier (Molecular Devices) from borosilicate patch electrodes (P-97 puller; Sutter Instruments); multi-electrode array recordings were made using the Axion Maestro MEA reader (Axion Biosystems), as described in the Supplemental Information.

## ACCESSION NUMBERS

The accession number for the RNA-seq and WGS data reported in this paper is GEO: GSE107878. Specific iPSC lines are available upon request.

## SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and seven tables and can be found with this article online at https://doi.org/10.1016/j.stemcr.2018.10.003.

## AUTHOR CONTRIBUTIONS

E.D., R.K.C.Y., K.K.S., J.E., and S.W.S. designed the research project. E.D., S.H.W., M.F., P.J.R., W.W., and A.P. contributed to cell maintenance, characterization, and differentiation. E.D. conceived the StopTag KO strategy. E.D., M.F., and K.Z. contributed to the CRISPR experiments. Z.W. performed WGS off-target analyses. D.C.R. performed western blots. R.A., G.P., B.T., E.D., G.K., and D.M. participated in RNA-seq and pathway analysis. E.D. and S.H.W. performed electrophysiological experiments. J.L.H., V.K., S.W., A.C.L., and P.P. provided technical help. E.D., S.H.W., K.K.S., J.E., and S.W.S. wrote the manuscript and supervised the project. Specific contributions of the co-corresponding authors: K.K.S. lab differentiated iPSCs to neurons and conducted patch-clamp electrophysiology experiments; J.E. lab generated iPSC and differentiated to neurons for expression analyses and multi-electrode arrays;

S.W.S. lab performed the CRISPR editing, WGS, and RNA-seq experiments.

## ACKNOWLEDGMENTS

We thank the Autism Speaks MSSNG project for genomic data and linking to research ethics board-approved consented families. We also thank Drs. Melissa Carter, Wendy Roberts, Brian Chung, and Rosanna Weksberg for obtaining skin biopsies and blood work, and the families for volunteering. We also thank Sergio Pereira, Wilson Sung, Thomas Nalpathamkalam, Darwin D'Souza, Joe Whitney, Jeff MacDonald, and Liz Li for bioinformatic support; Tara Paton, Guillermo Casallo, Barbara Kellam, Ny Hoang, and Sylvia Lamoureux for technical help; T.C. Südhof for the NGN2/rtTA lentiviral constructs. Supported by the Centre for Applied Genomics, Canada, the Canadian Institutes of Health Research (CIHR), Canada, the Canadian Institute for Advanced Research (CIFAR), Canada, the McLaughlin Centre, Canada, the Canada Foundation for Innovation (CFI), Canada, the Ontario Research Fund (ORF), Autism Speaks, United States, and the Hospital for Sick Children, Canada, Foundation. Scholarships and funding was from CIHR (FDN 143295 and XGG-125818) and ORF (GL2-01-013) to S.W.S., CFI-John R. Evans Leaders Fund (JELF)/ORF and CIHR (EPS-129129) to J.E.; CIHR (148814), Ontario Brain Institute (OBI), Canada, Natural Sciences and Engineering Research Council (NSERC), Canada, and the Scottish Rite Charitable Foundation to K.K.S.; NIH (4R33MH087908), United States, to J.E. and S.W.S.; Province of Ontario Neurodevelopmental Disorders (POND) from OBI to J.E., S.W.S., and K.K.S.; E.D. was a recipient of the Banting Post-Doctoral Fellowship and the Fonds de Recherche en Santé du Québec (FRQS) Post-Doctoral Fellowship; S.H.W. was supported by a Fellowship from the Fragile X Research Foundation of Canada; D.C.R. received the International Rett Syndrome Foundation, United States, Fellowship; P.J.R. was supported by the Ontario Stem Cell Initiative Fellowship and the Ontario Ministry of Research & Innovation Fellowship; K.Z. was a recipient of the CIHR Canada Vanier Graduate Scholarship. S.W.S. holds the GlaxoSmithKline-CIHR chair in Genome Sciences at the University of Toronto and the Hospital for Sick Children.

Received: June 11, 2018
Revised: October 1, 2018
Accepted: October 2, 2018
Published: November 1, 2018

## REFERENCES

Abrahams, B.S., Arking, D.E., Campbell, D.B., Mefford, H.C., Morrow, E.M., Weiss, L.A., Menashe, I., Wadkins, T., Banerjee-Basu, S., and Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol. Autism *4*, 36.

Anagnostou, E., Zwaigenbaum, L., Szatmari, P., Fombonne, E., Fernandez, B.A., Woodbury-Smith, M., Brian, J., Bryson, S., Smith, I.M., Drmic, I., et al. (2014). Autism spectrum disorder: advances in evidence-based practice. CMAJ *186*, 509–519.

Autism Genome Project Consortium, Szatmari, P., Paterson, A.D., Zwaigenbaum, L., Roberts, W., Brian, J., Liu, X.Q., Vincent, J.B., Skaug, J.L., Thompson, A.P., et al. (2007). Mapping autism risk loci using genetic linkage and chromosomal rearrangements. Nat. Genet. *39*, 319–328.

Bakkaloglu, B., O'Roak, B.J., Louvi, A., Gupta, A.R., Abelson, J.F., Morgan, T.M., Chawarska, K., Klin, A., Ercan-Sencicek, A.G., Stillman, A.A., et al. (2008). Molecular cytogenetic analysis and resequencing of contactin associated protein-like 2 in autism spectrum disorders. Am. J. Hum. Genet. *82*, 165–173.

Bensaid, M., Melko, M., Bechara, E.G., Davidovic, L., Berretta, A., Catania, M.V., Gecz, J., Lalli, E., and Bardoni, B. (2009). FRAXE-associated mental retardation protein (FMR2) is an RNA-binding protein with high affinity for G-quartet RNA forming structure. Nucleic Acids Res. *37*, 1269–1279.

Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A.T., et al. (2014). Disruptive CHD8 mutations define a subtype of autism early in development. Cell *158*, 263–276.

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. Brain Res. *1380*, 42–77.

Bourgeron, T. (2015). From the genetic architecture to synaptic plasticity in autism spectrum disorder. Nat. Rev. Neurosci. *16*, 551–563.

Brett, M., McPherson, J., Zang, Z.J., Lai, A., Tan, E.S., Ng, I., Ong, L.C., Cham, B., Tan, P., Rozen, S., et al. (2014). Massively parallel sequencing of patients with intellectual disability, congenital anomalies and/or autism spectrum disorders with a targeted gene panel. PLoS One *9*, e93409.

Carter, M.T., and Scherer, S.W. (2013). Autism spectrum disorder in the genetics clinic: a review. Clin. Genet. *83*, 399–407.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S.R., Woodhouse, E., Gillan, N., Hallett, V., Lietz, S., Garnett, T., et al. (2015). Heritability of autism spectrum disorder in a UK population-based twin sample. JAMA Psychiatry *72*, 415–423.

de la Torre-Ubieta, L., Won, H., Stein, J.L., and Geschwind, D.H. (2016). Advancing the understanding of autism disease mechanisms through genetics. Nat. Med. *22*, 345–361.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209–215.

DSM-V. (2013). Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (American Psychiatric Association).

Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A., and Tsai, L.H. (2016). Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat. Neurosci. *19*, 1477–1488.

Fernandez, B.A., and Scherer, S.W. (2017). Syndromic autism spectrum disorders: moving from a clinically defined to a molecularly defined approach. Dialogues Clin. Neurosci. *19*, 353–371.

Geschwind, D.H., and State, M.W. (2015). Gene hunting in autism spectrum disorder: on the path to precision medicine. Lancet Neurol. *14*, 1109–1120.

Gilman, S.R., Iossifov, I., Levy, D., Ronemus, M., Wigler, M., and Vitkup, D. (2011). Rare de novo variants associated with autism

implicate a large functional network of genes involved in formation and function of synapses. Neuron *70*, 898–907.

Gronborg, T.K., Schendel, D.E., and Parner, E.T. (2013). Recurrence of autism spectrum disorders in full- and half-siblings and trends over time: a population-based cohort study. JAMA Pediatr. *167*, 947–953.

Habela, C.W., Song, H., and Ming, G.L. (2016). Modeling synaptogenesis in schizophrenia and autism using human iPSC derived neurons. Mol. Cell. Neurosci. *73*, 52–62.

Hirano, S., and Takeichi, M. (2012). Cadherins in brain morphogenesis and wiring. Physiol. Rev. *92*, 597–634.

Hussman, J.P. (2001). Suppressed GABAergic inhibition as a common factor in suspected etiologies of autism. J. Autism Dev. Disord. *31*, 247–248.

Jiang, Y.H., Yuen, R.K., Jin, X., Wang, M., Chen, N., Wu, X., Ju, J., Mei, J., Shi, Y., He, M., et al. (2013). Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. Am. J. Hum. Genet. *93*, 249–263.

Jung, E.M., Moffat, J.J., Liu, J., Dravid, S.M., Gurumurthy, C.B., and Kim, W.Y. (2017). Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. Nat. Neurosci. *20*, 1694–1707.

Kamiya, K., Kaneda, M., Sugawara, T., Mazaki, E., Okamura, N., Montal, M., Makita, N., Tanaka, M., Fukushima, K., Fujiwara, T., et al. (2004). A nonsense mutation of the sodium channel gene SCN2A in a patient with intractable epilepsy and mental decline. J. Neurosci. *24*, 2690–2698.

Law, M.J., Lower, K.M., Voon, H.P., Hughes, J.R., Garrick, D., Viprakasit, V., Mitson, M., De Gobbi, M., Marra, M., Morris, A., et al. (2010). ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell *143*, 367–378.

Lee, S.G., Joo, Y., Kim, B., Chung, S., Kim, H.L., Lee, I., Choi, B., Kim, C., and Song, K. (2005). Association of Ala72Ser polymorphism with COMT enzyme activity and the risk of schizophrenia in Koreans. Hum. Genet. *116*, 319–328.

Leppa, V.M., Kravitz, S.N., Martin, C.L., Andrieux, J., Le Caignec, C., Martin-Coignard, D., DyBuncio, C., Sanders, S.J., Lowe, J.K., Cantor, R.M., et al. (2016). Rare inherited and de novo CNVs reveal complex contributions to ASD risk in multiplex families. Am. J. Hum. Genet. *99*, 540–554.

Lionel, A.C., Tammimies, K., Vaags, A.K., Rosenfeld, J.A., Ahn, J.W., Merico, D., Noor, A., Runke, C.K., Pillalamarri, V.K., Carter, M.T., et al. (2014). Disruption of the ASTN2/TRIM32 locus at 9q33.1 is a risk factor in males for autism spectrum disorders, ADHD and other neurodevelopmental phenotypes. Hum. Mol. Genet. *23*, 2752–2768.

Liu, H., Chang, L.H., Sun, Y., Lu, X., and Stubbs, L. (2014). Deep vertebrate roots for mammalian zinc finger transcription factor subfamilies. Genome Biol. Evol. *6*, 510–525.

Luo, H., Zheng, R., Zhao, Y., Wu, J., Li, J., Jiang, F., Chen, D.N., Zhou, X.T., and Li, J.D. (2017). A dominant negative FGFR1 mutation identified in a Kallmann syndrome patient. Gene *621*, 1–4.

Marshall, C.R., Noor, A., Vincent, J.B., Lionel, A.C., Feuk, L., Skaug, J., Shago, M., Moessner, R., Pinto, D., Ren, Y., et al. (2008). Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. *82*, 477–488.

Miyaoka, Y., Chan, A.H., Judge, L.M., Yoo, J., Huang, M., Nguyen, T.D., Lizarraga, P.P., So, P.L., and Conklin, B.R. (2014). Isolation of single-base genome-edited human iPS cells without antibiotic selection. Nat. Methods *11*, 291–293.

Moretto, E., Murru, L., Martano, G., Sassone, J., and Passafaro, M. (2018). Glutamatergic synapses in neurodevelopmental disorders. Prog. Neuropsychopharmacol. Biol. Psychiatry *84*, 328–342.

Nishiyama, M., Nakayama, K., Tsunematsu, R., Tsukiyama, T., Kikuchi, A., and Nakayama, K.I. (2004). Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. Mol. Cell. Biol. *24*, 8386–8394.

Ozonoff, S., Young, G.S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., Bryson, S., Carver, L.J., Constantino, J.N., Dobkins, K., et al. (2011). Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium study. Pediatrics *128*, e488–e495.

Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., et al. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. Am. J. Hum. Genet. *94*, 677–694.

Planells-Cases, R., Caprini, M., Zhang, J., Rockenstein, E.M., Rivera, R.R., Murre, C., Masliah, E., and Montal, M. (2000). Neuronal death and perinatal lethality in voltage-gated sodium channel alpha(II)-deficient mice. Biophys. J. *78*, 2878–2891.

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

Risch, N., Hoffmann, T.J., Anderson, M., Croen, L.A., Grether, J.K., and Windham, G.C. (2014). Familial recurrence of autism spectrum disorder: evaluating genetic and environmental contributions. Am. J. Psychiatry *171*, 1206–1213.

Rubenstein, J.L., and Merzenich, M.M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. Genes Brain Behav. *2*, 255–267.

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.

Tammimies, K., Marshall, C.R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A.C., Yuen, R.K., Uddin, M., Roberts, W., Weksberg, R., et al. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-exome sequencing in children with autism spectrum disorder. JAMA *314*, 895–903.

Varghese, M., Keshav, N., Jacot-Descombes, S., Warda, T., Wicinski, B., Dickstein, D.L., Harony-Nicolas, H., De Rubeis, S., Drapeau, E., Buxbaum, J.D., et al. (2017). Autism spectrum disorder: neuropathology and animal models. Acta Neuropathol. *134*, 537–566.

Weiner, D.J., Wigdor, E.M., Ripke, S., Walters, R.K., Kosmicki, J.A., Grove, J., Samocha, K.E., Goldstein, J.I., Okbay, A., Bybjerg-Grauholm, J., et al. (2017). Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nat. Genet. *49*, 978–985.

Weiss, L.A., Shen, Y., Korn, J.M., Arking, D.E., Miller, D.T., Fossdal, R., Saemundsen, E., Stefansson, H., Ferreira, M.A., Green, T., et al. (2008). Association between microdeletion and microduplication at 16p11.2 and autism. N. Engl. J. Med. *358*, 667–675.

Wintle, R.F., Lionel, A.C., Hu, P., Ginsberg, S.D., Pinto, D., Thiruvahindrapduram, B., Wei, J., Marshall, C.R., Pickett, J., Cook, E.H., et al. (2011). A genotype resource for postmortem brain samples from the Autism Tissue Program. Autism Res. *4*, 89–97.

Yang, N., Chanda, S., Marro, S., Ng, Y.H., Janas, J.A., Haag, D., Ang, C.E., Tang, Y., Flores, Q., Mall, M., et al. (2017). Generation of pure GABAergic neurons by transcription factor programming. Nat. Methods *14*, 621–628.

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.

Yuen, R.K., Merico, D., Bookman, M., Howe, L.J., Thiruvahindrapuram, B., Patel, R.V., Whitney, J., Deflaux, N., Bingham, J., et al. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat. Neurosci. *20*, 602–611.

Yuen, R.K., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., Tong, X., Sun, Y., Cao, D., Zhang, T., et al. (2016). Genome-wide characteristics of de novo mutations in autism. NPJ Genom. Med. *1*, 160271–1602710.

Yuen, R.K., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., Chrysler, C., Nalpathamkalam, T., Pellecchia, G., Liu, Y., et al. (2015). Whole-genome sequencing of quartet families with autism spectrum disorder. Nat. Med. *21*, 185–191.

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.

**Stem Cell Reports, Volume *11***

## Supplemental Information

# Complete Disruption of Autism-Susceptibility Genes by Gene Editing Predominantly Reduces Functional Connectivity of Isogenic Human Neurons

**Eric Deneault, Sean H. White, Deivid C. Rodrigues, P. Joel Ross, Muhammad Faheem, Kirill Zaslavsky, Zhuozhi Wang, Roumiana Alexandrova, Giovanna Pellecchia, Wei Wei, Alina Piekna, Gaganjot Kaur, Jennifer L. Howe, Vickie Kwan, Bhooma Thiruvahindrapuram, Susan Walker, Anath C. Lionel, Peter Pasceri, Daniele Merico, Ryan K.C. Yuen, Karun K. Singh, James Ellis, and Stephen W. Scherer**

# Figure S1

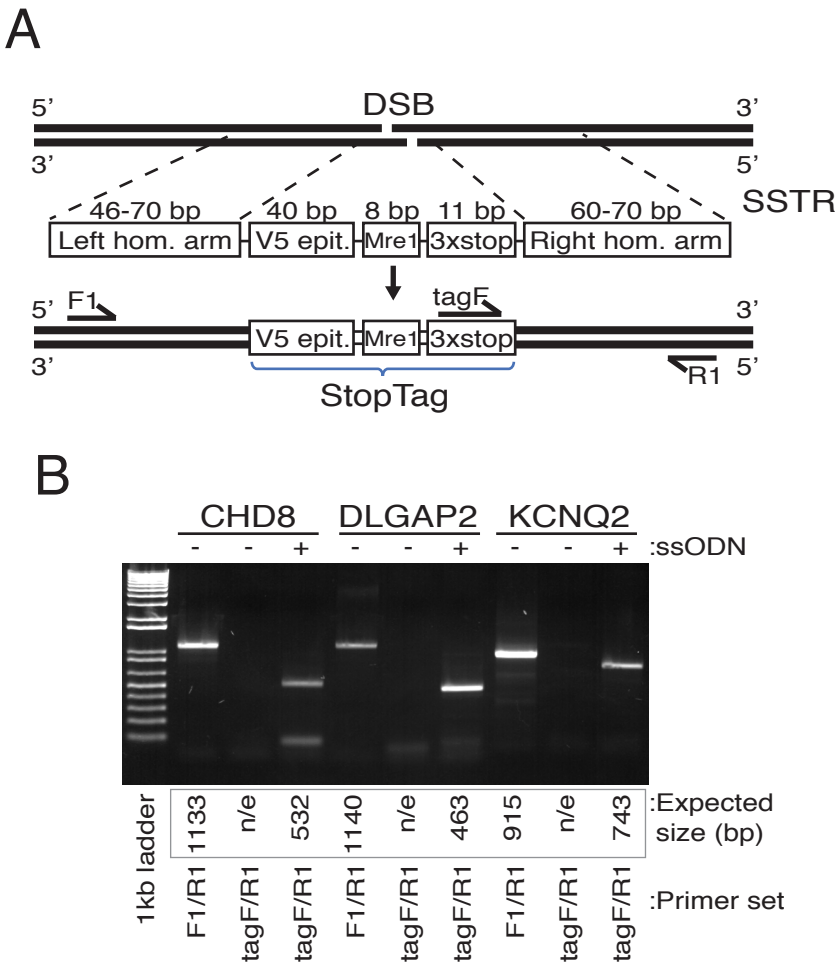


Figure S1 Validation of StopTag insertion in HEK293T cells, related to **Figure 1**. (A) Scheme of the single-stranded oligonucleotide DNA (ssODN) template used to splice the StopTag sequence into target genomic DNA; DBS = double-strand break; SSTR = single-stranded template repair; hom. = homology; epit. = epitope; F1, tagF, R1 = PCR primers. (B) Validation of StopTag insertion by PCR amplification of three selected target loci, i.e., C*HD8, DLGAP2* and *KCNQ2,* in HEK293T cells using corresponding primer sets [F1, tagF, R1 refering to (A)]; bp = base pair

# Figure S2

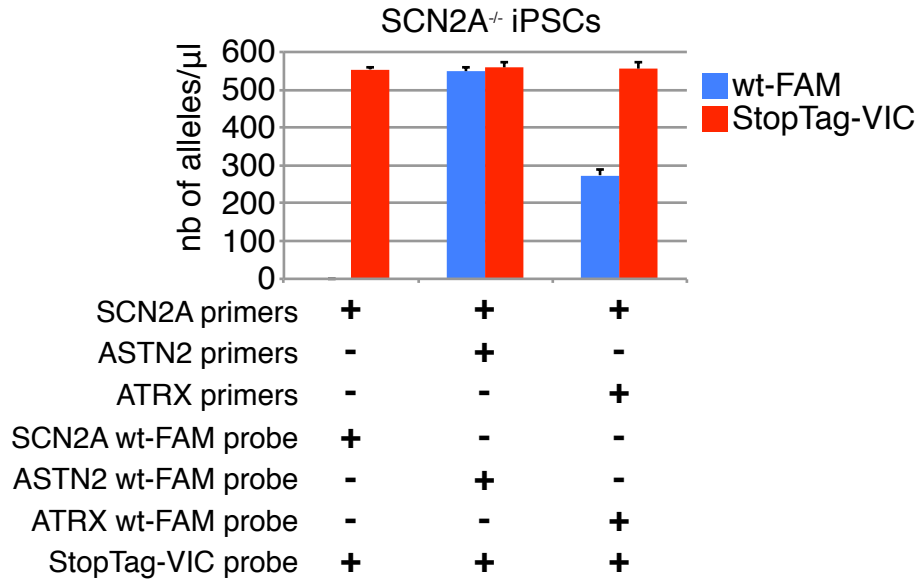


Figure S2 Confirmation of StopTag insertion and zygosity in different target loci in the human iPSC line 19-2, related to **Figure 1** and **Table 1**. Representative verification of homozygosity of StopTag insertion in SCN2A$^{StopTag/StopTag}$ (SCN2A$^{-/-}$) iPSCs using SCN2A primers/StopTag-VIC probe along with ASTN2 primers/ASTN2wt-FAM probe. No significant difference was observed between the number of SCN2A-associated StopTag-VIC alleles and the number of ASTN2-associated wt-FAM alleles, i.e., both were homozygous. The same comparison was performed with the hemizygous allele ATRX, where only half of the wt-FAM alleles were detected. Values are presented as mean±SD of three amplification reactions; wt = wild-type

Figure S3

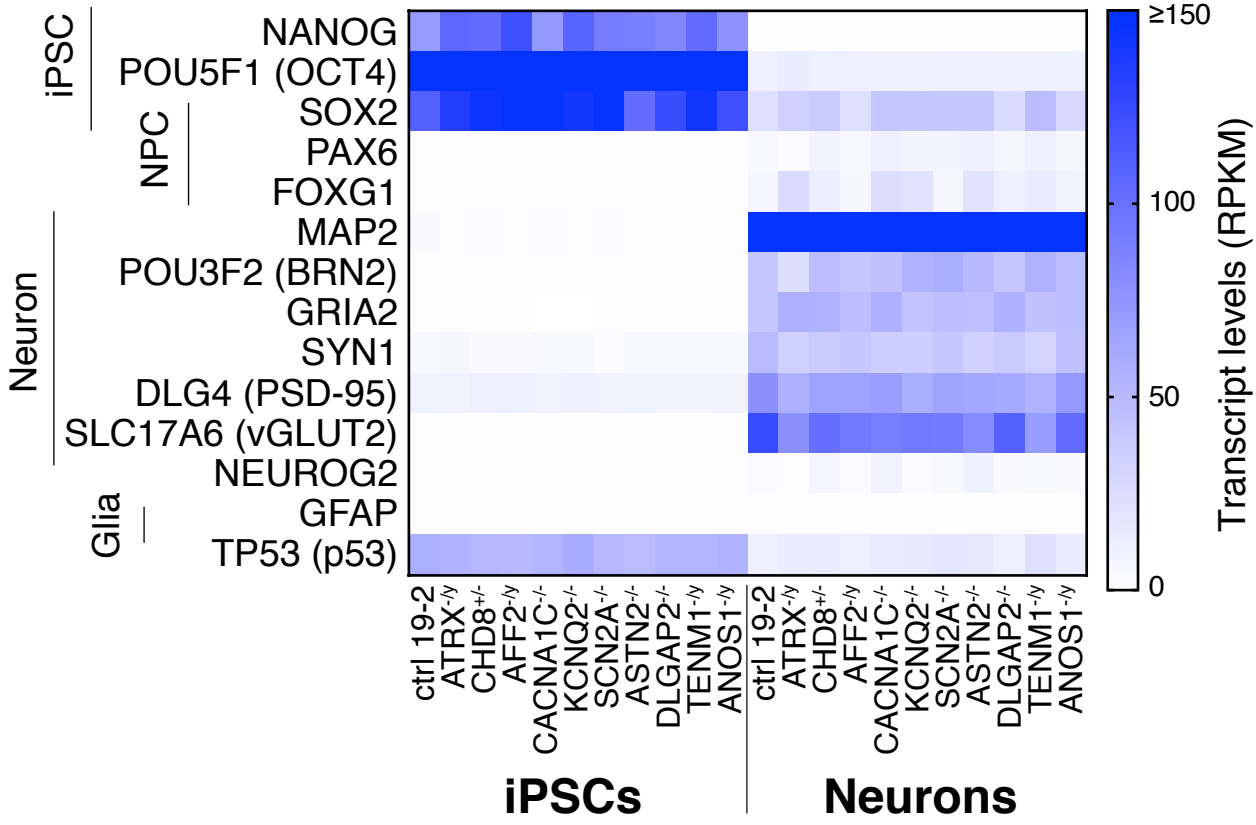


Figure S3 Transcript levels in RPKM of selected iPSC, NPC and neuronal markers in all KO cell lines, related to **Figure 2**. Values are presented as mean±SD of (n) independent experiments; n = 8 for ctrl 19-2, n = 4 for ATRX$^{-/y}$, CHD8$^{+/-}$, KCNQ2$^{-/-}$, SCN2A$^{-/-}$, ASTN2$^{-/-}$, DLGAP2$^{-/-}$, ANOS1$^{-/y}$; n = 2 for AFF2$^{-/y}$, CACNA1C$^{-/-}$ and TENM1$^{-/y}$ iPSCs; n = 5 for AFF2$^{-/y}$ and TENM1$^{-/y}$; n = 4 for ctrl 19-2, ATRX$^{-/y}$, CHD8$^{+/-}$, KCNQ2$^{-/-}$, SCN2A$^{-/-}$, DLGAP2$^{-/-}$, ANOS1$^{-/y}$; n = 3 for CACNA1C$^{-/-}$ and ASTN2$^{-/-}$ neurons. Note that color gradient ranges from 0 to 150 RPKM; all RPKM values ≥150 are presented with the same blue intensity as 150.

Figure S4

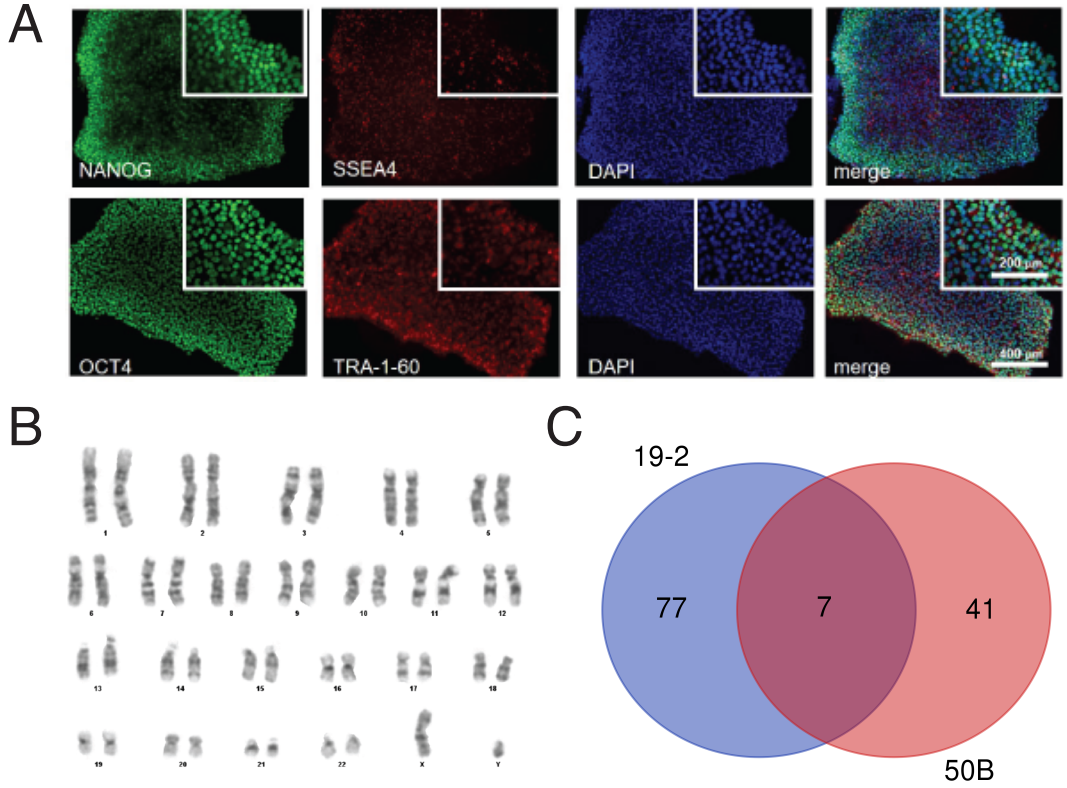


Figure S4 Characterization of the control 50B iPSC line, related to **Figures 2** and **4**. Line 50B was reprogrammed from an unaffected brother of a child with ASD, who was carrying a splice site mutation (c112+1G>C) in the gene SET (Yuen et al., 2017). Line 50B does not carry this mutation and was reprogrammed using non-integrative Sendai virus in order to minimize any artefactual genetic modification. We used ribonucleoprotein (RNP) complex as a vector to deliver the CRISPR machinery in 50B iPSCs instead of plasmids. We also used only one gRNA with the Cas9 nuclease for each target gene. We ensured that the sequence of the newly designed gRNA (Table S1) was different from that of the gRNAs used in the primary screen, further dissociating any potential off-target mutation to the observed phenotypes. (A) Immunohistochemistry revealing expression of the pluripotency markers NANOG, SSEA4, OCT4 and TRA-1-60. (B) 50B iPSCs presented a normal male karyotype; 20 cells were examined. (C) Comparison of the number of differentially expressed genes revealed by RNAseq, in AFF2-null neurons, in the two different genetic backgrounds, i.e., 19-2 and 50B.

# Figure S5

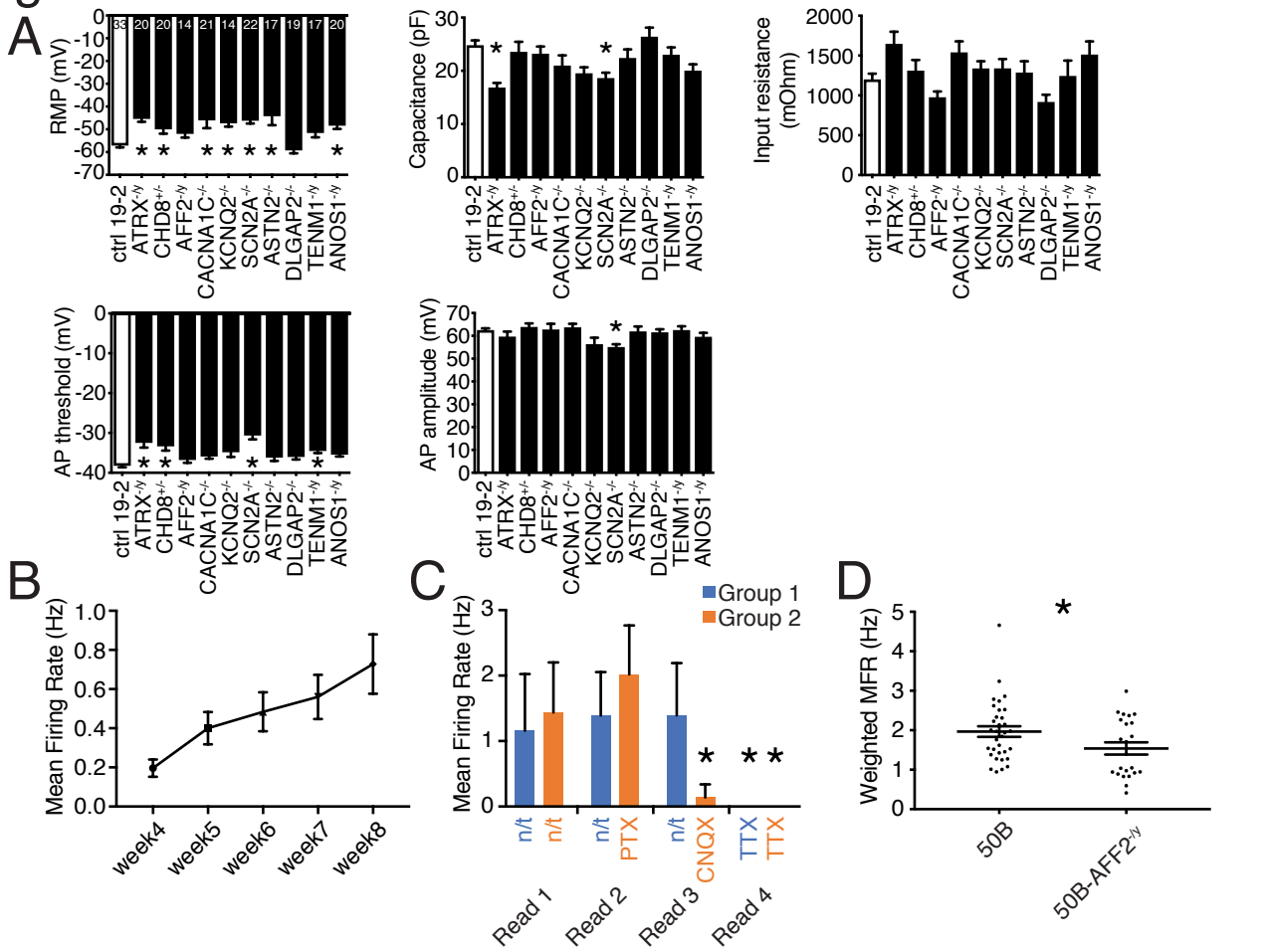


Figure S5 Electrophysiological phenotyping of KO iPSC-derived neurons, related to **Figure 5**. (A) Intrinsic properties detected by patch-clamp recordings. For example, the capacitance was similar across all lines except for ATRX- and SCN2A-null neurons where it was lower. The action potential amplitude was unchanged in all lines except for a small reduction in SCN2A-null line. While the action potential threshold was reduced in some lines, it didn't affect the intrinsic firing rate of any of the lines. The only intrinsic property to change across multiple lines was membrane potential, which was elevated in seven of ten lines. This latter finding could either be unrelated to each of the lines, or could have an impact on synaptic activity. Also, it did not correlate with the expression of maturity markers (**Figure S3**); the number of recorded neurons is indicated in the bars of the top left graph; values are presented as mean±SEM of three independent experiments, recorded at day 21-28 post-NGN2-induction (PNI); *p <0.05; n/t = not treated; mV = millivolt; pF = picofarad; mOhm = milliohm; Hz = hertz. (B) MEA recording of the mean firing rate (MFR) at one-week interval from week 4 to week 8 PNI; 43 different wells were recorded from 6 independent experiments; values are presented as mean±SEM. (C) MEA recording of the MFR of control 19-2 neurons upon treatments with different receptor inhibitors; values were acquired in four consecutive readings, i.e., before (read 1) and after addition of GABA receptor inhibitor PTX (read 2), after addition of AMPA receptor inhibitor CNQX (read 3), and after addition of sodium channel blocker TTX (read 4); no substantial change was observed after addition of PTX, confirming that GABAergic neurons were not appreciably present in our cultures; all activity was abolished after addition of TTX, indicating the human neurons were expressing functional sodium channels; our selected active wells were recorded for each group at week 8 PNI; values are presented as mean±SD of a single experiment; *p <0.05 compared to n/t. (D) MEA recording of the weighted MFR (MFR per active electrode) of AFF2 KO in a different genetic background (50B line) at week 8 PNI; values are presented as mean±SEM of 5 independent differentiation experiments (8 wells per experiment); *p <0.05

Tables S1 Description of selected target transcript and exon for each candidate gene to knockout, as well as the sequence of each guide RNA (gRNA) and template single-stranded oligonucleotide (ssODN) used for gene editing, and sequence of primers and probes used for enrichment of edited cell by ddPCR, related to **Figure 1**.

| Gene Symbol | chr | Transcript ID | Exon | Exon ID | gRNA- (5'-3') | gRNA+ (5'-3') | Offset (bp) | FAM probe (5'-3') | VIC probe (5'-3') | Forward primer ddPCR (5'-3') | Reverse primer ddPCR (5'-3') |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ANKRD11 | 16 | ENST00000378330.6 | 7 | ENSE00003504243 | TCAAAGAGCTCATCAGCGAGGGG | CGTCCCCGCGGATGGCGGCTCGG | 9 | ATCCGCGGGGACGCC | CAACGCCGGCGTAACTA | AGAAACGAGCGTGGAGAGAC | GCTCTCTGGCCCTTGCC |
| AUTS2 | 7 | ENST00000406775.6 | 1 | ENSE00001671293 | GCCGCTTTTTGCGGAGTCCATGG | GTCGCGGTCGCAGCGAGACCGGG | 1 | CGCAAAAAGCGGCGG | CAACGCCGGCGTAACTA | GAACCATGGATGGCCCGAC | CGACGACGAGGCGAGTG |
| ATRX | X | ENST00000373344.9 | 5 | ENSE00003713901 | ATCAGATGATGAAAAACCTTTGG | ACATACTTTGTTACAATTGAAGG | 3 | TGTAACAAAGTATGTAGA | CAACGCCGGCGTAACTA | TCCAGTTGTCAGAAACATTCATAATTC | TGGCAAGCTCTGCATAGTAAT |
| CHD8 | 14 | ENST00000430710.7 | 4 | ENSE00001422532 | GTTGCCCAGCGTTCGGCCAGAGG | TCGTCCTCATTTAAGACTCGAGG | 0 | TCTTAAATGAGGACGAGTTG | CAACGCCGGCGTAACTA | CGGTGAGCAGAACATACCTCG | GAGGCACCAGATGTTTTACTCTT |
| AFF2 | X | ENST00000370460.6 | 3 | ENSE00003780313 | TGACTATCACGTGACCACTCAGG | TAGCACTGTACTGGCAAGCCAGG | 6 | TCACGTGATAGTCATAAC | CAACGCCGGCGTAACTA | ACACAGCAACAGAAAATCAAAACC | AGTCTTCCAGTTTGGCTTGAGA |
| CAPRIN1 | 11 | ENST00000341394.8 | 5 | ENSE00003682460 | AACTCCTCTTCGGACAATATTGG | TAAGCTAGTAGACCCTGAACGGG | 21 | CGAAGAGGAGTTGTCAT | CAACGCCGGCGTAACTA | GAATGGAGTGCCAATATTGTC | TTGTCAACTTAAATACTCAGAAG |
| CACNA1C | 12 | ENST00000399655.5 | 2 | ENSE00000936250 | CGATGGCCGCCTGCCACGACAGG | GCCCGGCAGGCTAAGCTGATGGG | 5 | AGGCGGCCATCGACG | CAACGCCGGCGTAACTA | CTGAGCACATCCCCACCC | GGTTTCCCATATTGCTGCCG |
| KCNQ2 | 20 | ENST00000626839.2 | 1 | ENSE00003772915 | GAATTTCCTCTACAACGTGCTGG | CGGTAGAAGGCGTTGCGCTTGGG | 9 | ACGCCTTCTACCGCAA | CAACGCCGGCGTAACTA | CGGGCGCCGGGAAGCC | CCGATGGGGGTCGCC |
| SCN2A | 2 | ENST00000375437.6 | 2 | ENSE00003607689 | TGCGTTGTTCAATAGCAGCAAGG | GAGACCCAAACAGGAACGCAAGG | 19 | AACGCATTGCAGAAGA | CAACGCCGGCGTAACTA | CTTCCGCTTCTTTACCAGGGA | GGAAGAGATTTTCCTGCTTCCA |
| ASTN2 | 9 | ENST00000361209.6 | 16 | ENSE00003567870 | GCTGCTGGAAATCAACCGTGTGG | CCTTGCCTGGCCGGCGGCAAAGG | 22 | CAAGGAGCAGCTGAGC | CAACGCCGGCGTAACTA | TCCATCTACCTGTCACCTTTGC | GGTAGGAGGGCTGGTCAAGT |
| DLGAP2 | 8 | ENST00000637795.1 | 5 | ENSE00003797478 | CGGCGGCTGGGTCCGACTTCCGG | CACGTGTGGTCTGGCGCCCCCGG | 14 | AGCCGCCGCTGTGTT | CAACGCCGGCGTAACTA | CCCAGGAGCATGAAGGGC | GCTCAGCAGGTAGGGCG |
| CNTNAP2 | 7 | ENST00000361727.7 | 4 | ENSE00001369431 | TCGGATGCTGTAATTCGTGCCGG | CTATGTGCGCATAGTGCCTCTGG | 10 | TACAGCATCCGATTA | CAACGCCGGCGTAACTA | CCGGAAACATTAACTCTGACGG | ACCTGCCACAAATTTAAACAATACG |
| TENM1 | X | ENST00000371130.7 | 1 | ENSE00000857625 | GACCCTGCACGAGTATAACCAGG | CTGGAGTTGTATGACTGTCTTGG | 3 | CAAGACAGTCATACAAC | CAACGCCGGCGTAACTA | ACCAAAAGTCAAGCATGAAATGGA | AGCTCCTGGTTATACTCGTGC |
| ANOS1 | X | ENST00000262648.7 | 6 | ENSE00001090930 | GTGAATGTGCATGGAACTCGAGG | TCGAAACTGGTACCATCGGCTGG | 9 | CCAGTTTCGAGTGGCT | CAACGCCGGCGTAACTA | CTGACATAAGACCCAGCCGA | ATTCTGTGTCTGGGAAGAAAAGGA |
| AFF2 (validation in 50B line) | X | ENST00000370460.6 | 3 | ENSE00003780313 | n/a | GGCTTGCCAGTACAGTGCTAGGG | n/a | TCACGTGATAGTCATAAC | CAACGCCGGCGTAACTA | ACACAGCAACAGAAAATCAAAACC | AGTCTTCCAGTTTGGCTTGAGA |

| Gene Symbol | ssODN template sequence (5'-3') |
|---|---|
| ANKRD11 | AGATAAAGTGAACAAGAGAAACGAGCGTGGAGAGACCCGCCTGCACCGAGCCGCCATCGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACGCATCAAAGAGCTCATCAGCGAGGGGGCAGACGTCAACGTCAAGGACTTCGCAGGTGGGCACCCTGTGC |
| AUTS2 | AGCGGGGAGACCCCGGCGCAGCAGAACCATGGATGGCCCGACGCGGGGCCATGGACTCGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGATCGCGGTCGCAGCGAGACCGGGAGAGGCGCTCCCGGGGCGGGCTGGGGGCCGGCGCGGCCGGCGGCGGCG |
| ATRX | GTTGTCAGAAACATTCATAATTCTATGAACTTAATTCTTTCATAGGAAACCTTCAATTGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGATCAGATGATGAAAAACCTTTGGATGATGAAACTGTAAATGAAGATGCGTCTAATGAAAATTCAGA |
| CHD8 | AGCAGAGGCCATTGCGAGAGCCCGTGCCCGCGGTGAGCAGAACATACCTCGAGTCGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACCCAGCGTTCGGCCAGAGGAGGAAGGCGAGAAGAAACGCAGGAAGAAGAGTGCTGGGGAGAGGCTGAAAG |
| AFF2 | TGTGATACTGAATTCAACTCTAATACACAGCAACAGAAAATCAAAACCTGAGTGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACCTAGCACTGTACTGGCAAGCCAGGCCAGTGGTCAGCCAAACAAGATGCAGACTTTGACACAGGACCAGT |
| CAPRIN1 | AGTGCGGACTGACCTGAAACAAGGTTTGAATGGAGTGCCAATATTGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAGAATTCTATAAGCTAGTAGACCCTGAACGGGACATGAGCTTGAGGTATTAGTGGTATTTGATAATAGAAA |
| CACNA1C | GGCAGCGGGGCTGGCCCCTGAGCACATCCCCACCCCGGGGGCTGCCCTGTCGTGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAGCCCGGCAGGCTAAGCTGATGGGCAGCGCTGGCAATGCGACCATCTCCACAGTCAGCTCCACGCAGCGGA |
| KCNQ2 | CCTCAGCAAACCTCGCGCGGGCGGCGCGGGCGCCGGGAAGCCCCCCAAGCGCGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACAGAATTTCCTCTACAACGTGCTGGAGCGGCCGCGCGGCTGGGCGTTCATCTACCACGCCTACGTGTGAG |
| SCN2A | CACAGTCAGTGCTGGTACCGCCAGGACCTGACAGCTTCCGCTTCTTTACCAGGGAATCCCTTGCTGCTGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAAAAGCTAAGAGACCCAAACAGGAACGCAAGGATGAGGATGATGAAAATGGCCCAAAGCCAAACAGTGAC |
| ASTN2 | TCGCTGTGAGGAGAAGGGGCGCTGTCCATCTACCTGTCACCTTTGCCGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACCCACACCAGTGCTGCTGGAAATCAACCGTGTGGTGCCACTTTATACCCTCATCCAAGACAATGGCACAA |
| DLGAP2 | GGAGCGCTACTCGCCCGCGCCCAGGAGCATGAAGGGCCTTTCCGGAAGTCGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAGGGCACACGTGTGGTCTGGCGCCCCCGGAGGACTGCGAGCACCTGCACCACGGGCCCGACGCGCGGCCGC |
| CNTNAP2 | TTAAAATACTATTTCCTTTTCCAGGCATTTCCCGGAAACATTAACTCTGACGGTGTGGTCCGGCACGAAGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACGCTATGTGCGCATAGTGCCTCTGGATTGGAATGGAGAAGGTCGCATTGGACTCAGAATTGAAGTTTAT |
| TENM1 | TAGCTTACACCAGTTCTTCTGATGAGAGTGAAGATGGAAGAAAACAGTCATACAACTCCAGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAACCCTGCACGAGTATAACCAGGAGCTGAGGATGAATTACAATAGCCAGAGTAGAAAGAGGAAAGAAGT |
| ANOS1 | TCAGACCACAGACGAGCGAGTTCAACTGACTGACATAAGACCCAGCCGATGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGAGTGAATGTGCATGGAACTCGAGGCTTCACTGCCCCCAGCAAACACTTCCGTTCTTCCAAAGGTCAGTCTT |
| AFF2 (validation in 50B line) | TGTGATACTGAATTCAACTCTAATACACAGCAACAGAAAATCAAAACCTGAGTGGGGGAAACCCATTCCCAATCCCCTTCTTGGGCTTGATTCAACGCCGGCGTAACTAGCTGACAAGCCAGGCCAGTGGTCAGCCAAACAAGATGCAGACTTTGACACAGGACCAGT |

Chr = chromosome; ID = identifier; bp = base pair; n/a = not available

Table S2 Number of enrichment steps (or plates) required to isolate a 100% (homozygous or hemizygous) or 50% (heterozygous) StopTag iPSC population for each candidate genes, related to **Figure 1** and **Table 1**.

| Gene | vector | plate 1 | plate 2 | plate 3 | plate 4 | plate 5 | plate 6 | RPKM - iPSC ctrl 19-2 |
|---|---|---|---|---|---|---|---|---|
| ANKRD11 | plasmids | A4 = 0.4 | E8 = 0.3 | E10 = 2.6 | A4 = 28.2 | F9 = 4.8 | C1 = 15.1 | 26.2 |
| AUTS2 | plasmids | H3 = 0.5 | A12 = 3.1 | A10 = 0.4 | n/a | n/a | n/a | 16.6 |
| ATRX | plasmids | E11 = 1.2 | C8 = 2.4 | G2 = 13.9 | A5 = 100.0 | n/a | n/a | 17.7 |
| CHD8 | plasmids | E1 = 0.8 | B9 = 2.5 | C11 = 28.5 | E1 = 47.7 | G8 = 52.4 | G1 = 52.7 | 27.8 |
| AFF2 | plasmids | F3 = 7.0 | H2 = 100.0 | n/a | n/a | n/a | n/a | 1.7 |
| CAPRIN1 | plasmids | G1 = 0.6 | D8 = 0.2 | E1 = 0.8 | D1 = 0.06 | n/a | n/a | 110.9 |
| CACNA1C | plasmids | D9 = 1.6 | D10 = 6.7 | F10 = 15.4 | H7F3 = 44.3 | A4 = 100.0 | n/a | 0.2 |
| KCNQ2 | plasmids | H1 = 1.6 | E9 = 1.4 | G5 = 8.4 | F11 = 26.5 | E1 = 47.2 | C1 = 52.4 | 7.8 |
| SCN2A | plasmids | G6 = 2.0 | D2 = 10.9 | E8 = 48.5 | D2 = 68.9 | B6 = 86.2 | A8 = 100.0 | 0.3 |
| ASTN2 | plasmids | C4 = 3.3 | D6 = 12.1 | A2 = 43.0 | G1 = 100.0 | n/a | n/a | 2.9 |
| DLGAP2 | plasmids | C1 = 3.9 | F10 = 1.7 | B8 = 14.8 | F9 = 63.0 | C4 = 100.0 | n/a | 0.2 |
| CNTNAP2 | plasmids | G7 = 0.2 | E2 = 2.0 | F9 = 12.7 | n/a | n/a | n/a | 19.3 |
| TENM1 | plasmids | B10 = 11.3 | C9 = 7.0 | C7 = 19.1 | F1 = 100.0 | n/a | n/a | 0.2 |
| ANOS1 | plasmids | A1 = 3.0 | C8 - 4.6 | E4 = 19.0 | C6 = 88.6 | H10 = 100.0 | n/a | 35.8 |
| 50B-AFF2 | RNP complex | F4 = 7.3 | F8 = 89.0 | A1 = 100.0 | n/a | n/a | n/a | n/a |

For 10 of 14 (71%) genes, the proper targeting event was achieved (*ANKRD11, AUTS2, CAPRIN1, CNTNAP2* failed). Since our initial objective was to develop a "high-throughput" process to test mutations, here, we did not attempt additional pairs of gRNA. These four unsuccessfully-targeted genes were not necessarily less transcriptionally active than successful genes in iPSCs, as revealed by RNAseq. For example, *ANKRD11* and *CAPRIN1* presented at reads per kilobase per million mapped reads (RPKM) >25, while *DLGAP2* and *TENM1* were successfully targeted even though at RPKM 0.2. Shaded area indicates unsuccessful attempts. n/a = not available; RNP = ribonucleoprotein; RPKM = reads per kilobase per million mapped reads

Table S3 Identification of SNV/indel calls specific to the different KO line genomes with respect to the 19-2 control genome, related to **Figure 1**.

| KO Line | Chr | Strand | Variant Site | REF | ALT | Explored Locus | gRNA Site | Nb Mismatch | Mismatch Pos. | gRNA Sequence | gRNA Name | Diagnostic |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ATRX-/y | X | + | 76952179 | C | StopTag | 76951962-76952397 | 76952168 | 0 | 19 | CATACTTTGTTACAATTGA | ATRX gRNA+ | StopTag insertion on-target |
| ATRX-/y | X | - | 76952179 | G | StopTag | 76951962-76952397 | 76952144 | 0 | 19 | AAGGTTTTTCATCATCTGA | ATRX gRNA- | StopTag insertion on-target |
| ATRX-/y | 14 | + | 93462922 | TA | T | 93462722-93463122 | 93463005 | 2 | 9C4G4 | TCAGATGATCAAAAGCCTT | ATRX gRNA- | 1-bp deletion off-target, 98 bp away from cut site; intronic ITPK1 |
| CHD8+/- | 14 | + | 21896163 | A | StopTag | 21895961-21896366 | 21896168 | 0 | 19 | CGTCCTCATTTAAGACTCG | CHD8 gRNA+ | StopTag insertion on-target |
| CHD8+/- | 14 | - | 21896163 | T | StopTag | 21895961-21896366 | 21896147 | 0 | 19 | CTGGCCGAACGCTGGGCAA | CHD8 gRNA- | StopTag insertion on-target |
| AFF2-/y | X | - | 147743734 | A | StopTag | 147743525-147743943 | 147743721 | 0 | 19 | GAGTGGTCACGTGATAGTC | AFF2 gRNA- | StopTag insertion on-target |
| AFF2-/y | X | + | 147743734 | T | StopTag | 147743525-147743943 | 147743748 | 0 | 19 | AGCACTGTACTGGCAAGCC | AFF2 gRNA+ | StopTag insertion on-target |
| CACNA1C-/- | 12 | - | 2224504 | G | StopTag | 2224296-2224713 | 2224491 | 0 | 19 | GTCGTGGCAGGCGGCCATC | CACNA1C gRNA- | StopTag insertion on-target |
| CACNA1C-/- | 12 | + | 2224504 | C | StopTag | 2224296-2224713 | 2224517 | 0 | 19 | CCCGGCAGGCTAAGCTGAT | CACNA1C gRNA+ | StopTag insertion on-target |
| KCNQ2-/- | 20 | + | 62103590 | T | StopTag | 62103375-62103806 | 62103595 | 0 | 19 | GGTAGAAGGCGTTGCGCTT | KCNQ2 gRNA+ | StopTag insertion on-target |
| KCNQ2-/- | 20 | - | 62103590 | A | StopTag | 62103375-62103806 | 62103565 | 0 | 19 | GCACGTTGTAGAGGAAATT | KCNQ2 gRNA- | StopTag insertion on-target |
| SCN2A-/- | 2 | - | 166152415 | G | StopTag | 166152205-166152625 | 166152400 | 0 | 19 | TGCTGCTATTGAACAACGC | SCN2A gRNA- | StopTag insertion on-target |
| SCN2A-/- | 2 | + | 166152415 | C | StopTag | 166152205-166152625 | 166152440 | 0 | 19 | AGACCCAAACAGGAACGCA | SCN2A gRNA+ | StopTag insertion on-target |
| ASTN2-/- | 9 | + | 119413904 | C | StopTag | 119413689-119414120 | 119413895 | 0 | 19 | CTTGCCTGGCCGGCGGCAA | ASTN2 gRNA+ | StopTag insertion on-target |
| ASTN2-/- | 9 | - | 119413904 | G | StopTag | 119413689-119414120 | 119413852 | 0 | 19 | CACGGTTGATTTCCAGCAG | ASTN2 gRNA- | StopTag insertion on-target |
| DLGAP2-/- | 8 | - | 1496893 | G | StopTag | 1496683-1497104 | 1496877 | 0 | 19 | GAAGTCGGACCCAGCCGCC | DLGAP2 gRNA- | StopTag insertion on-target |
| DLGAP2-/- | 8 | + | 1496893 | C | StopTag | 1496683-1497104 | 1496912 | 0 | 19 | ACGTGTGGTCTGGCGCCCC | DLGAP2 gRNA+ | StopTag insertion on-target |
| TENM1-/y | X | + | 124097473 | C | StopTag | 124097270-124097676 | 124097477 | 0 | 19 | TGGAGTTGTATGACTGTCT | TENM1 gRNA+ | StopTag insertion on-target |
| TENM1-/y | X | - | 124097473 | G | StopTag | 124097270-124097676 | 124097453 | 0 | 19 | GGTTATACTCGTGCAGGGT | TENM1 gRNA- | StopTag insertion on-target |
| ANOS1-/y | X | + | 8553378 | T | StopTag | 8553167-8553589 | 8553380 | 0 | 19 | CGAAACTGGTACCATCGGC | ANOS1 gRNA+ | StopTag insertion on-target |
| ANOS1-/y | X | - | 8553378 | A | StopTag | 8553167-8553589 | 8553350 | 0 | 19 | CGAGTTCCATGCACATTCA | ANOS1 gRNA- | StopTag insertion on-target |

Each StopTag insertion event was found exclusively on-target as revealed by WGS. Since the 19-2 genome differs from the hg19 reference genome by probably millions of single nucleotide variants (SNVs), potential extra off-target sites might be missed if we considered only those predicted on hg19. Hence, we have listed all SNV/indel calls unique to each KO line, with respect to the 19-2 control line. Then, we mapped the corresponding gRNA sequences to 200 bp on each side of these calls. We considered as CRISPR-derived on- and off-target variants only those with a flanked matching gRNA sequence, tolerating up to five mismatches. One on-target variant was identified for each of the two gRNAs per gene, corresponding to the expected StopTag insertion. Remarkably, only one potential off-target mutation was detected within 200 bp of a sequence that had two mismatches with *ATRX* gRNA- (see row 3). This single mutation corresponds to a 1-bp deletion located in intron 4 of *IPTK1* , 98 bp away from the Cas9 cut site. Thus, we considered this deletion unlikely to be an off-target effect of CRISPR. It may have arisen as a random spontaneous mutation. No potential off-target mutation was found within any 200-bp window containing each pair of gRNAs; REF = reference; ALT = alternative; Nb = number; bp = base pair

Table S4 Distances between premature termination codon insertion (StopTag) and the next upstream and downstream exon-exon junctions (e-e) or wildtype (wt) start or stop codons, related to **Figure 3**.

| Gene | chr | exon | ins. site | dist. 5'e-e | dist. 3'e-e | decreased transcript levels (Figure 3B) | possible explanation |
|---|---|---|---|---|---|---|---|
| ATRX | X | 5/35 | 77,696,639 | 62 | 97 | yes | NMD |
| CHD8 | 14 | 4/38 | 21,428,112 | 243 | 130 | yes | NMD |
| AFF2 | X | 3/21 | 148,661,908 | 351 | 546 | yes | NMD |
| CACNA1C | 12 | 2/47 | 2,115,379 | 155 | 197 | no | higher mRNA stability |
| KCNQ2 | 20 | 1/16 | 63,472,426 | 258 | 65 | no | higher mRNA stability |
| SCN2A | 2 | 2/27 | 165,295,894 | 121 | 171 | no | higher mRNA stability |
| ASTN2 | 9 | 16/22 | 116,651,743 | 215 | 75 | yes | NMD |
| DLGAP2 | 8 | 5/15 | 1,548,767 | 141 | 945 | no | higher mRNA stability |
| TENM1 | X | 1/31 | 124,963,689 | 152 | 85 | no | StopTag <200bp downstream of wt AUG |
| ANOS1 | X | 6/14 | 8,585,364 | 97 | 61 | yes | NMD |

chr = chromosome; ins = insertion; dist = distance; NMD = nonsense-mediated decay

# Supplemental Experimental Procedures

## Skin Fibroblast Culture

The skin-punch biopsy for the control 19-2 came from the upper back by a clinician at The Hospital for Sick Children, with approval from their Research Ethics Board. Samples were immersed in 14 ml of ice-cold Alpha-MEM (Wisent Bioproducts) supplemented with penicillin 100 units/ml and streptomycin 100 µg/ml (pen/strep; ThermoFisher), and transferred immediately to the laboratory. The biopsy was cut into ~1mm$^3$ pieces in a 60-mm dish. 5 ml of collagenase 1 mg/ml (Sigma) was added and the dish was placed in 37°C incubator for 1:45 hours. Skin pieces and collagenase were then transferred to a 15-ml tube, and centrifuged at 300 g for 10 minutes. Supernatant was removed and 5 ml of trypsin 0.05%/EDTA 0.53 mM (Wisent Bioproducts) was added. The mix was pipetted up and down several times to break up tissue and placed in 37°C incubator for 30 minutes. After incubation, the mix was centrifuged at 300 g for 10 minutes, and the supernatant was removed, leaving 1 ml. The pellet was pipetted up and down vigorously to break the pieces. The mix was transferred in a T-12.5 flask along with 5 ml of Alpha-MEM, 15% Fetal Bovine Serum (FBS; Wisent Bioproducts), 1x pen/strep, and placed in a 37°C incubator. Cultured cells were fed every 5 days. At 100% confluency, the cell population was split into three 100-mm dishes for expansion and cryopreservation in liquid nitrogen.

## Skin Fibroblast Reprogramming

The iPSC line 19-2 was generated from skin fibroblasts using retroviral vectors encoding *POU5F1*, *SOX2*, *KLF4*, *MYC*, and lentiviral vectors encoding the pluripotency reporter EOS-GFP/Puro$^R$, as described previously (Hotta et al., 2009).

## Peripheral Blood Mononuclear Cells (PBMCs) Isolation and Enrichment of CD34+ Cells

Whole peripheral blood was processed at the Centre for Commercialization of Regenerative Medicine (CCRM) using Lymphoprep™ (StemCellTechnologies) in a SepMate™ tube (StemCellTechnologies) according to manufacturer's instructions. The sample was centrifuged 10 min at 1200 g. The top layer containing PBMCs was collected and mixed with 10 ml of the PBS/FBS mixture and centrifuged 8 min at 300 g. The PBMC's collected at the bottom of the tube were washed, counted and re-suspended in PBS/FBS mixture. CD34+ cells were isolated using the Human Whole Blood/Buffy Coat CD34+ Selection kit according to manufacturer's instructions (StemCellTechnologies). Isolated cells were expanded in StemSpan SFEM II media (StemCellTechnologies) and StemSpan CD34+ Expansion Supplements (StemCellTechnologies) prior to reprogramming.

## Reprogramming PBMC Using Non-Integrative Sendai Virus

The iPSC line 50B was generated from CD34+ PBMCs at CCRM using CytoTune™-iPSC 2.0 Sendai Reprogramming Kit. Expanded cells were spun down and

resuspended in StemSpan SFEM II media and StemSpan CD34+ Expansion Supplements, and placed in a single well of a 24-well dish. Virus MOI was calculated and viruses combined according to number of cells available for reprogramming and manufacturer's protocol. The virus mixture was added to cells, and washed off 24 hours after infection. 48 hours after viral delivery, cells were plated in 6-well plates in SFII and transitioned to ReproTESR for the duration of reprogramming. Once colonies were of an adequate size and morphology, individual colonies were picked and plated into E8 media (ThermoFisher). Clones were further expanded and characterized using standard assays for pluripotency, karyotyping (Medical Genetics Services, Addenbrooke's Hospital), short-tandem repeat (STR) genotyping [The Centre for Applied Genomics (TCAG)] and mycoplasma (Lonza MycoAlert). Directed differentiation was performed using kits for definitive endoderm, neural and cardiac lineages (ThermoFisher).

## iPSCs Maintenance

All iPSC lines were maintained on matrigel (Corning) coating, with complete media change every day in mTeSR™ (StemCellTechnologies). ReLeSR™ (StemCellTechnologies) was used for passaging. Accutase™ (InnovativeCellTechnologies) and 10 μM Rho-associated kinase (ROCK) inhibitor (Y-27632; StemCellTechnologies) were used for single-cell dissociation purposes.

## Gene Editing

Guide RNA sequences were devised using the CRISPR design tools (Hsu et al., 2013) or (http://www.benchling.com) in order to minimize off-target activity. We selected the highest off-target and on-target scored pair of gRNAs with minimal inclusion of single nucleotide variants with respect to the control 19-2 genome. Each ssODN sequence was designed to place the V5 fragment (red in **Table S1**) within the original translation reading frame. Two different methods were used to prepare the gRNAs. The first method involved cloning of gRNAs into expression plasmids. DNA sequences allowing fusion of crRNAs and tracrRNAs were designed and assembled as described (Yang et al., 2014) and were synthesized at TCAG. The assembled inserts were ligated into plasmid #41824 (Addgene) using Gibson Assembly® cloning kit [New England Biolabs (NEB)]. HEK293T cells (ATCC) were transfected, as 200 000 cells per well in 24-well plate, with 200 ng of each gRNA plasmids, 400 ng of pCas9D10A_GFP plasmid #44720 (Addgene), and 1 μl of 100 μM ssODN synthesized as standard desalted Ultramer® oligonucleotides [Integrated DNA Technologies (IDT)], using lipofectamine® 2000 (ThermoFisher). For human iPSCs, 3 μg of each gRNA plasmids, 5 μg of pCas9D10A_GFP plasmid, and 3 μl of 100 μM ssODN were nucleofected into $1x10^6$ iPSCs using Nucleofector™2b device (Lonza; program A-023) and Human Stem Cell Nucleofector® kit 1 (Lonza). The second method to prepare gRNAs involved ribonucleoprotein (RNP) complexes. For this, crRNA and tracrRNA were synthesized and annealed according to manufacturer's protocol (Synthego). 1.6 μl of 10 μM gRNA was incubated with 1 μl of 20 μM Cas9-2NLS Nuclease (Synthego) in 100 μl Human Stem Cell Nucleofection Solution 1 (Lonza) for 10 minutes at room temperature to form Cas9:gRNA RNP complexes. Then, 5 μl of 10 μM ssODN was added to the complexes. The mix was nucleofected into $1x10^6$ iPSCs using Nucleofector™2b (program A-023). Prior to nucleofection, iPSCs were treated with 10 μM Y-27632 for 60 minutes at 37°C.

Following nucleofection, iPSCs were evenly distributed into a flat-bottom 96-well plate in mTeSR and 10 μM Y-27632.

## Purification of Edited Cells

Near-confluent iPSCs, i.e., 10-14 days post-nucleofection or post-passaging, were treated with 10 μM Y-27632 for 60 minutes at 37°C, washed with PBS, and treated with 28 μl/well of accutase for 10 minutes at 37°C. Half of each well was harvested in a PCR 96-well plate containing 50 μl/well of DNA lysis buffer (10mM Tris pH 7.5, 10mM EDTA pH 8.0, 10mM NaCl, 0.5% N-lauroylsarcosine and freshly added 1 mg/ml proteinase K). The other half was re-seeded in a new flat-bottom matrigel-coated 96-well plate along with 250 μl/well of mTeSR™ supplemented with 10 μM Y-27632, and put at 37°C for expansion. The PCR plate containing cells and DNA lysis buffer was sealed and heated at 70°C for 10 minutes in a PCR machine, and chilled on ice for 3 minutes. 100ul/well of DNA precipitation solution (EtOH 95%, $H_2O$ 5%, NaCl 75mM), previously cooled at -80°, was added and the plate was left at room temperature for 30-60 minutes. The plate was spun at 1600 g for 4 minutes and flicked upside down quickly to remove supernatant. DNA was washed 2 consecutive times with 100ul/well EtOH 70%, and heated without seal at 70°C for 2 minutes. DNA was resuspended with 30ul/well of TE buffer, the plate was sealed and heated at 70°C for 10 minutes. The ddPCR (Bio-Rad) was performed according to manufacturer's protocol. Custom TaqMan® MGB probes (ThermoFisher) were designed according to manufacturer's recommendations. A DNA probe specific to StopTag, which encompassed an *Mre*1 restriction site and the first half of the 3x stop of StopTag, was coupled with the fluorescent dye VIC (**Table S1**). This design ensured high specificity of the StopTag-VIC probe and secured the isolation of cells that have integrated the 3x stop at the target locus. Conversely, a different probe coupled with the fluorescent dye FAM was used for each unmodified (wt) target gene. Each wt-FAM probe sequence was selected to overlap with that of reverse gRNA (gRNA-; **Table S1**), which instructs Cas9D10A to nick the plus strand. The entire wt-FAM probe sequence was removed from ssODN to avoid confusion with StopTag allele detection. The Tm of the common VIC probe was ensured to be compatible with the Tm of all other specific FAM probes. Moreover, about half of the sequence of gRNA- was removed from ssODN to prevent nicking from Cas9D10A, given that ssODN was synthesized as plus strand for each target gene. The forward primers for ddPCR were designed within the ssODN, and reverse primers outside of ssODN, to prevent PCR amplification of non-spliced ssODN sequences (**Table S1**). Primers were designed using primer blast tool (www.ncbi.nlm.nih.gov/tools/primer-blast/). The well with the highest frequency of edited alleles was identified in absolute quantification mode. When the re-seeded half of cells reached confluency, 1/10 of total cell number of the corresponding selected well were passed in a new flat-bottom matrigel-coated 96-well plate in order to enrich for edited cells while maintaining a polyclonal composition of cell population. When confluent, this secondary plate was processed for ddPCR, as described above. This procedure was repeated until a 100% edited iPSC population was obtained. For characterization of the new 100% edited lines, a PCR was performed on genomic DNA from KO cells using OneTaq® DNA polymerase (NEB) according to manufacturer's protocol in order to amplify a 500-1000 fragment encompassing the StopTag integration site. Amplified fragments were ligated into plasmids using TOPO TA cloning kit (ThermoFisher). These

plasmids were Sanger sequenced to reveal the exact sequence and zygosity of each target locus.

## Detection of CRISPR-Related Off-Target Genetic Variants

Genomic DNA was extracted from 70-80% confluent iPSCs using DNeasy® tissue kit (Qiagen), and submitted to TCAG for genomic library preparation and whole genome sequencing (WGS). DNA samples were quantified and analyzed using Qubit High Sensitivity Assay and Nanodrop OD260/280 ratio. 200 ng of DNA was used as input material for library preparation using Illumina TruSeq Nano DNA Library Prep Kit. DNA was fragmented to 550 bp on average using sonication on a Covaris LE220 instrument. Fragmented DNA was end-repaired and A-tailed. Indexed TruSeq Illumina adapters with overhang-T were added to the DNA. Libraries were loaded on a Bioanalyzer DNA High Sensitivity chip to validate size and absence of primer dimers, and were quantified by qPCR using Kapa Library Quantification Illumina/ABI Prism Kit protocol (KAPA Biosystems). Each validated library was sequenced on two lanes of a high throughput V4 flowcell on a HiSeq 2500 platform following Illumina's recommended protocol to generate paired-end reads of 125-bases in length. These reads were mapped using the bwa algorithm (0.7.12). GATK (GenotypeConcordance) (McKenna et al., 2010) was used to identify indel and SNV calls unique to the different KO line genomes with respect to the control 19-2 genome. Because of the 59-bp length of the StopTag sequence, it could have appeared as a series of GATK calls in close proximity to one another. Therefore, all calls occurring within a 200 bp window were grouped, and GATK FastaAlternateReferenceMaker was used to reconstruct the sequence within that window to identify any potential insertion. We found StopTag insertions only at the expected loci. Then, we searched for sequences matching (up to five mismatches) gRNA+ and gRNA-, separately, within a 200-bp window flanking both sides of each unique variant call, in each corresponding KO line. Since the gRNA cloning protocol that we used sets the first 5' nucleotide as G, which is dispensable for specificity (Yang et al., 2014), mismatches at that position were ignored. Likewise, mismatches at the PAM sequence were ignored.

## Lentivirus Production

$7.5x10^6$ HEK293T cells were seeded in a T-75 flask, grown in 10% fetal bovine serum in DMEM (Gibco). The next day, cells were transfected using Lipofectamine 2000 with plasmids for gag-pol (10 µg), rev (10 µg), VSV-G (5 µg), and the target constructs FUW-TetO-Ng2-P2A-EGFP-T2A-puromycin or FUW-rtTA (15 µg; gift from T.C. Südhof laboratory) (Zhang et al., 2013). Next day, the media was changed. The day after that, the media was spun down in a high-speed centrifuge at 30,000 g at 4°C for 2 hours. The supernatant was discarded and 50 µl PBS was added to the pellet and left overnight at 4°C. The next day, the solution was triturated, aliquoted and frozen at -80°C.

## Induction into Glutamatergic Neurons

$5x10^5$ iPSCs/well were seeded in a matrigel-coated 6-well plate in 2 ml of mTeSR supplemented with 10 µM Y-27632. Next day, media in each well was replaced with 2 ml fresh media plus 10 µM Y-27632, 0.8 µg/ml polybrene (Sigma), and the minimal

amount of NGN2 and rtTA lentiviruses necessary to generate 100% GFP+ cells upon doxycycline induction, depending on prior titration of a given virus batch. The day after, virus-containing media were replaced with fresh mTeSR, and cells were expanded until near-confluency. Newly generated “NGN2-iPSCs” were detached using accutase, and seeded in a new matrigel-coated 6-well plate at a density of $5x10^5$ cells per well in 2 ml of mTeSR supplemented with 10 µM Y-27632 (day 0 of differentiation). Next day (day 1), media in each well was changed for 2 ml of CM1 [DMEM-F12 (Gibco), 1x N2 (Gibco), 1x NEAA (Gibco), 1x pen/strep (Gibco), laminin (1 µg/ml; Sigma), BDNF (10 ng/µl; Peprotech) and GDNF (10 ng/µl; Peprotech) supplemented with fresh doxycycline hyclate (2 µg/ml; Sigma) and 10 µM Y-27632. The day after (day 2), media was replaced with 2 ml of CM2 [Neurobasal media (Gibco), 1x B27 (Gibco), 1x glutamax (Gibco), 1x pen/strep, laminin (1 µg/ml), BDNF (10 ng/µl) and GDNF (10 ng/µl)] supplemented with fresh doxycycline hyclate (2 µg/ml) and puromycin (5 µg/ml for 19-2-derived cells, and 2 µg/ml for 50B-derived cells; Sigma). Media was replaced with CM2 supplemented with fresh doxycycline hyclate (2 µg/ml). The same media change was repeated at day 4. At day 6, media was replaced with CM2 supplemented with fresh doxycycline hyclate (2 µg/ml) and araC (10 µM; Sigma). Two days later, these day 8 post-NGN2-induction (PNI) neurons were detached using accutase and ready to seed for subsequent experiments, as described below.

## RNAseq and q-RT-PCR

Total RNA was extracted from 70-80% confluent iPSCs using RNeasy® mini kit (Qiagen). For glutamatergic neurons, 12-well plates were coated with filter-sterilized 0.1% polyethyleneimine (PEI; Sigma) solution in borate buffer pH 8.4 for 1 hour at room temperature, washed four times with water, and dried thoroughly. $1x10^6$ neurons/well (day 8 PNI) were seeded in 2 ml CM2 media. Media was half-changed once a week with CM2 media. Four weeks later, total RNA was extracted from each well using RNeasy® mini kit (Qiagen). RNA samples were submitted to an Agilent Bioanalyzer 2100 RNA Nano chip for quality control. Concentration was determined using Qubit RNA HS Assay on a Qubit fluorometer (ThermoFisher). RNA libraries were prepared using NEBNext Ultra RNA Library Preparation kit for Illumina. Briefly, 500 ng of total RNA was used for poly-A mRNA enrichment before being split into 200-300bp fragments for 4 minutes at 94°C. Fragments were converted to double stranded cDNA, end-repaired and adenylated in 3’ to create an overhang A, allowing ligation of Illumina adapters with an overhang T. Fragments were amplified under the following conditions: initial denaturation at 98°C for 30 seconds, followed by 13 cycles of 98°C for 10 seconds, 65°C for 75 seconds, and finally an extension step of 5 minutes at 65°C. Each sample was amplified with a different barcoded adapter to allow for multiplex sequencing. RNA library integrity was verified on a Bioanalyzer 2100 DNA High Sensitivity chip (Agilent Technologies), and quantified using Kapa Library Quantification Illumina/ABI Prism Kit protocol (KAPA Biosystems). Stranded libraries were pooled in equimolar amounts and sequenced on an Illumina HiSeq 2500 platform using a High Throughput Run Mode flowcell and the V4 sequencing chemistry to generate paired-end reads of 125-bases in length, following Illumina’s recommended protocol. Data quality was assessed using FastQC v.0.11.2. Trimmed reads were screened for presence of rRNA and mtRNA sequences using FastQ-Screen v.0.4.3. RSeQC

package v.2.3.7 was used to assess read distribution, positional read duplication and confirm strandedness of alignments. Raw trimmed reads were aligned to the reference genome hg19 using Tophat v.2.0.11. Tophat alignments were processed to extract raw read counts for genes using htseq-count v.0.6.1p2. Two-condition differential expression was done with the edgeR R package, v.3.8.6. In the filtered data set were retained only genes whose reads per kilobase of expressed transcript per million reads of library (RPKM) values were >1 for the iPSCs and 0.5 for neurons in at least n-1 samples where n is the number of samples in the smaller sample group in the comparison. The method used for normalizing the data was TMM, implemented by the calcNormFactors(y) function. All samples in each pair-wise comparison were normalized together. The test for differential expression was done using the quasi-likelihood F-test functionality in edgeR. For q-RT-PCR, total RNA was reverse transcribed into cDNA using First Strand cDNA Synthesis kit according to manufacturer's protocol (Sigma), and quantitative PCR was performed using SYBR® Select Master Mix kit (LifeTechnologies) with primer sets designed using primer blast tool.

## Pathway Enrichment

Pathway enrichment analysis was performed using the R package goseq version 1.28.0 and R version 3.4.1 (2017-06-30) using a custom gene-set collection including gene ontology (GO, obtained from the R package GO.db version 3.4.1) and pathways (KEGG and Reactome collections downloaded from the respective websites on 20171016). To remove terms that were either too generic or too narrow, GO size terms were restricted to those within 100-1000 annotated genes, and pathways to 15 and 500 annotated genes, leaving 3736 total gene ontology terms and pathways including 17041 genes. Up- and down-regulated genes were separately analyzed using goseq (Wallenius' noncentral hypergeometric distribution); median transcript lengths were used to correct for count bias. To correct for multiple comparisons, Benjamini-Hochberg false discovery rate (BH-FDR) correction was applied to the p-values. The results were combined to obtain a list of pathways enriched in up-regulated and down-regulated genes.

## Patch-Clamp Recordings

Day 3 PNI neurons were replated at a density of 100 000/well of a poly-ornithin/laminin coated coverslips in a 24-well plate with CM2 media. On day 4, 50 000 mouse astrocytes were added to the plates and cultured until day 21-28 PNI for recording. At day 10, CM2 was supplemented with 2.5% FBS in accordance with (Zhang et al., 2013). Whole-cell recordings (BX51WI; Olympus) were performed at room temperature using an Axoclamp 700B amplifier (Molecular Devices) from borosilicate patch electrodes (P-97 puller; Sutter Instruments) containing a potassium-based intracellular solution (in mM): 123 K-gluconate, 10 KCL, 10 HEPES; 1 EGTA, 2 $MgCl_2$, 0.1 $CaCl_2$, 1 Mg.ATP, and 0.2 $Na_4GTP$ (pH 7.2). 0.06% sulpharhodamine dye was added to select neurons for visual confirmation of multipolar neurons. Composition of extracellular solution was (in mM): 140 NaCl, 2.5 KCl, 1 1.25 $NaH_2PO_4$, 1 $MgCl_2$, 10 glucose, and 2 $CaCl_2$ (pH 7.4). Whole cell recordings were clamped at −70 mV using Clampex 10.6 (Molecular Devices), corrected for a calculated −10 mV junction potential and analyzed

using the Template Search function from Clampfit 10.6 (Molecular Devices). Following initial breakthrough and current stabilization in voltage clamp, the cell was switched to current clamp to monitor initial spiking activity and record the membrane potential (cc=0, ~1 min post-breakthrough). Bias current was applied to bring the cell to ~70 mV whereby increasing 5 pA current steps were applied (starting at -20 pA) to generate the whole cell resistance and to elicit action potentials. Data were digitized at 10 kHz and low-pass filtered at 2 kHz. Inward and outward currents were recorded in whole-cell voltage clamp in response to consecutive 10 mV steps from -90 mV to +40 mV.

### Multi-electrode Array

48-well opaque- or clear-bottom MEA plates (Axion Biosystems) were coated with filter-sterilized 0.1% PEI solution in borate buffer pH 8.4 for 1 hour at room temperature, washed four times with water, and dried overnight. 120 000 day 8 PNI neurons (were Dox-induced for 7 days) per well were seeded in 250 µl CM2 (laminin 10 µg/ml instead of 1 µg/ml for MEA) media. The day after, 10 000 mouse astrocytes/well were seeded on top of neurons in 50 µl/well CM2 media. Astrocytes were prepared from postnatal day 1 CD-1 mice as described (Kim and Magrane, 2011). Media was half-changed once a week with CM2 media. Once a week from week 4 to 8 PNI, the electrical activity of the MEA plates was recorded using the Axion Maestro MEA reader (Axion Biosystems). The heater control was set to warm up the reader at 37°C. Each plate was first incubated for 5 minutes on the pre-warmed reader, then real-time spontaneous neural activity was recorded for 5 minutes using AxIS 2.0 software (Axion Biosystems). A bandpass filter from 200 Hz to 3 kHz was applied. Spikes were detected using a threshold of 6 times the standard deviation of noise signal on electrodes. Offline advanced metrics were re-recorded and analysed using Axion Biosystems Neural Metric Tool. An electrode was considered active if at least 5 spikes were detected per minute. A burst was considered as a group of at least 5 spikes, each separated by an inter-spike interval (ISI) of no more than 100 ms. Network bursts were identified as a minimum of 10 spikes with a maximum ISI of 100 milliseconds covered by at least 25% of electrodes in each well. No non-active well was excluded in the analysis. After the last reading at week 8, selected wells were treated with three synaptic antagonists: $GABA_A$ receptor antagonist picrotoxin (PTX; Sigma) at 100 µM, AMPA receptor antagonist 6-cyano-7-nitroquinoxaline-2,3-dion (CNQX; Sigma) at 60 µM, and sodium ion channel antagonist tetrodotoxin (TTX; Alomone labs) at 1 µM. The plates were recorded consecutively, 5-10 minutes after addition of the antagonists. A 60-minute recovery period was allowed in the incubator at 37°C between each antagonist treatment and plate recording.

## Supplemental References

Hotta, A., Cheung, A.Y., Farra, N., Vijayaragavan, K., Seguin, C.A., Draper, J.S., Pasceri, P., Maksakova, I.A., Mager, D.L., Rossant, J.*, et al.* (2009). Isolation of human iPS cells using EOS lentiviral vectors to select for pluripotency. Nat Methods *6*, 370-376.
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.
Kim, H.J., and Magrane, J. (2011). Isolation and culture of neurons and astrocytes from the mouse brain cortex. Methods Mol Biol *793*, 63-75.
McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M.*, et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297-1303.
Yang, L., Mali, P., Kim-Kiselak, C., and Church, G. (2014). CRISPR-Cas-mediated targeted genome editing in human cells. Methods Mol Biol *1114*, 245-267.
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.