Stem Cell Reports, Volume *19*

Supplemental Information

Scaled and efficient derivation of loss-of-function alleles in risk genes for neurodevelopmental and psychiatric disorders in human iPSCs

Hanwen Zhang, Ada McCarroll, Lilia Peyton, Sol Díaz de León-Guerrerro, Siwei Zhang, Prarthana Gowda, David Sirkin, Mahmoud ElAchwah, Alexandra Duhe, Whitney G. Wood, Brandon Jamison, Gregory Tracy, Rebecca Pollak, Ronald P. Hart, Carlos N. Pato, Jennifer G. Mulle, Alan R. Sanders, Zhiping P. Pang, and Jubao Duan

Supplementary Information Table of Contents

Supplementary Figures

Supplementary Tables

Supplemental Experimental procedures

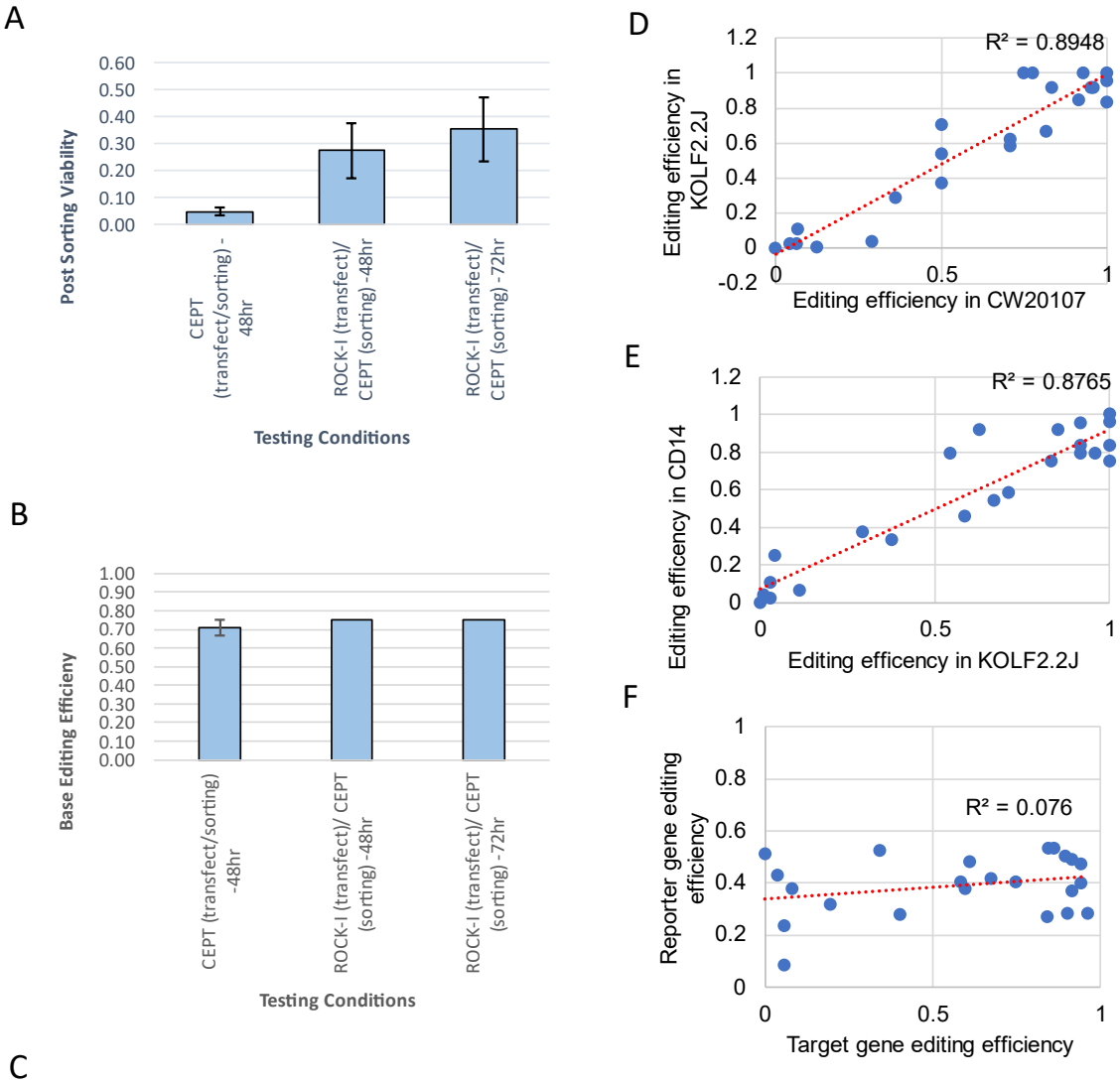
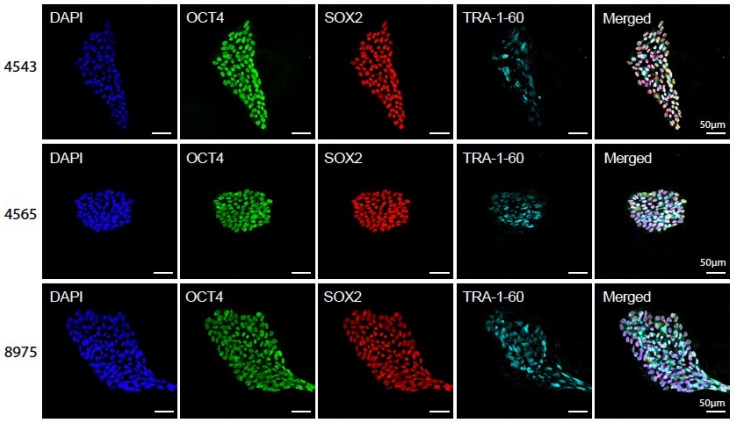
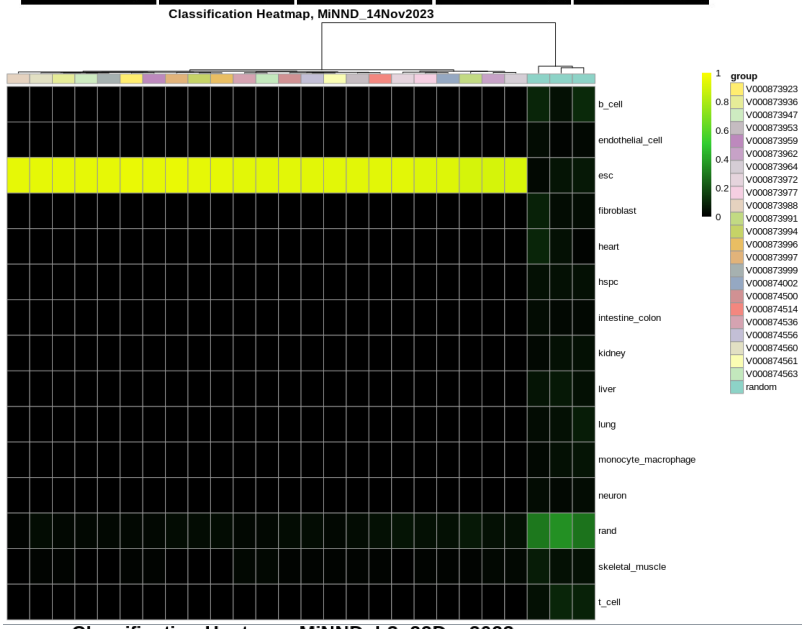

A
Post Sorting Viability
0.60
0.50
0.40
0.30
0.20
0.10
0.00
CEPT (transfect/sorting) - 48hr
ROCK-I (transfect)/ CEPT (sorting) -48hr
ROCK-I (transfect)/ CEPT (sorting) -72hr
Testing Conditions
B
Base Editing Efficieny
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
CEPT (transfect/sorting) -48hr
ROCK-I (transfect)/ CEPT (sorting) -48hr
ROCK-I (transfect)/ CEPT (sorting) -72hr
Testing Conditions
D
R² = 0.8948
Editing efficiency in KOLF2.2J
Editing efficiency in CW20107
E
R² = 0.8765
Editing efficency in CD14
Editing efficency in KOLF2.2J
F
R² = 0.076
Reporter gene editing efficiency
Target gene editing efficiency
C
Reporter Editing Efficiency
GRIA3
ASH1L
HCN4
KDM6B
SMARCC2
KMT2C
GRIN2A
TRIO
CUL1
RB1CC1
SCN2A
SP4
XPO7
ANKRD11
CACNA1G
DLL1
SETD1A
ARID1B
CHD8
AKAP11
HERC1
GABRA1
SHANK3
CW20107
KOLF2.2J
CD14

Figure S1. Reporter gene editing efficiency and optimization for increasing post-sorting single hiPSC clonal survivability (related to Figures 2 and 3). (A) Post-sorting single hiPSC clonal survivability (A) and C to T base editing efficiency (B) (assayed by Sanger sequencing of the LoF mutation site) under different transfection/cell sorting conditions. (C) Reporter gene editing efficiency for all the target genes in three hiPSC lines (CW20107, KOLF2.2J, and CD14). (D) and (E) Strong correlations of target gene editing efficiency across cell lines. (F) Weak correlation between target gene editing efficiency and reporter gene editing efficiency. Target gene editing efficiency was calculated by the genotypes of each single hiPSC clone as confirmed by Sanger sequencing. Reporter gene editing efficiency in (F) was calculated as the proportion of GFP+ cells (reporter gene edited cells) vs. BFP+ cells (cells transfected with sgRNAs) in FACS.

A


4543
4565
8975
DAPI
OCT4
SOX2
TRA-1-60
Merged
50μm

B


Classification Heatmap, MiNND_14Nov2023
b_cell
endothelial_cell
esc
fibroblast
heart
hspc
intestine_colon
kidney
liver
lung
monocyte_macrophage
neuron
rand
skeletal_muscle
t_cell
group
random

C

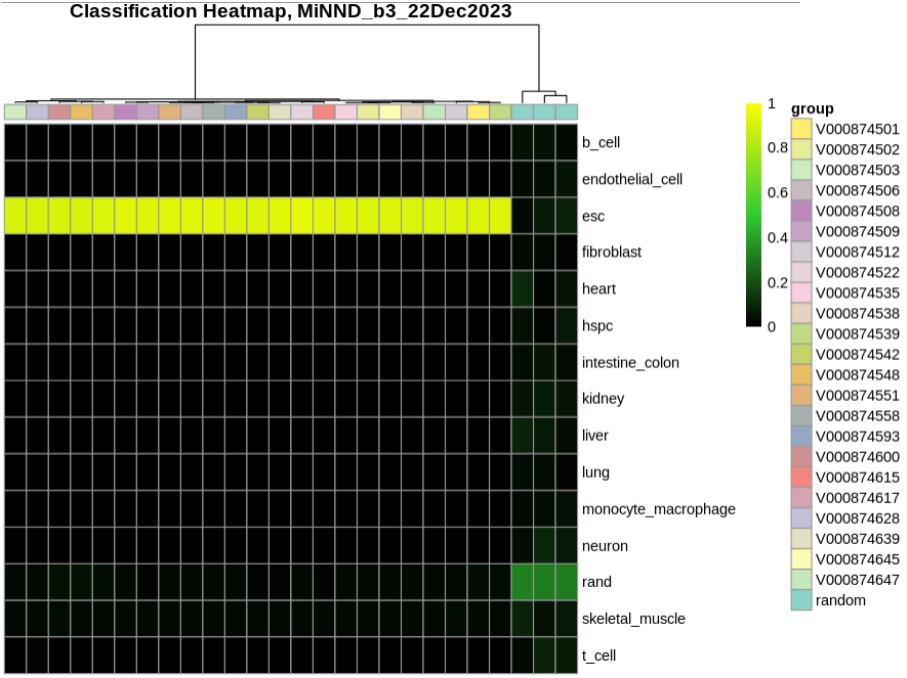

Classification Heatmap, MiNND_b3_22Dec2023
b_cell
endothelial_cell
esc
fibroblast
heart
hspc
intestine_colon
kidney
liver
lung
monocyte_macrophage
neuron
rand
skeletal_muscle
t_cell
group
random

Figure S2. Stem cell pluripotency characterization of iSTOP LoF hiPSC lines (related to Figure 4). (A) The iSTOP mutant lines were stained positive for pluripotent stem cell markers (OCT4, SOX2, TRA-1-60). Scale bar: 50μm. (B) CellNet analysis of RNA-seq data of hiPSC lines confirmed their pluripotency. Pluripotency scores showed the transcriptional similarity of the edited iSTOP LoF hiPSC lines to ESC or other non-ESC cell types. Two batches of edited hiPSC lines that were not listed in the main Figure 4 are shown.

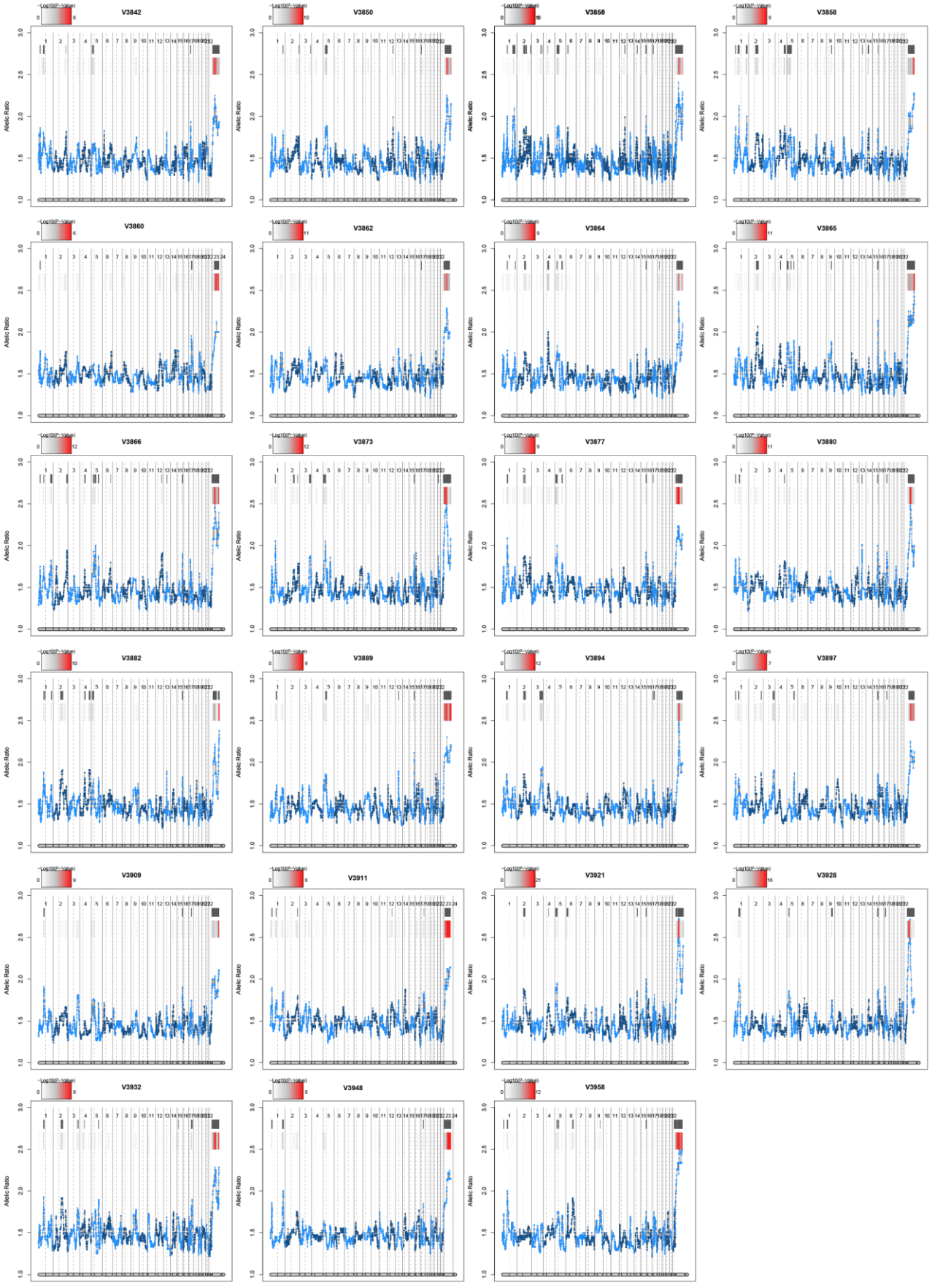
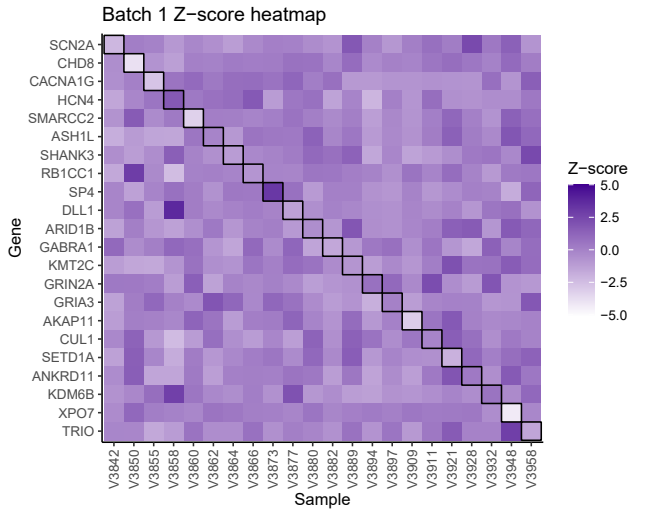
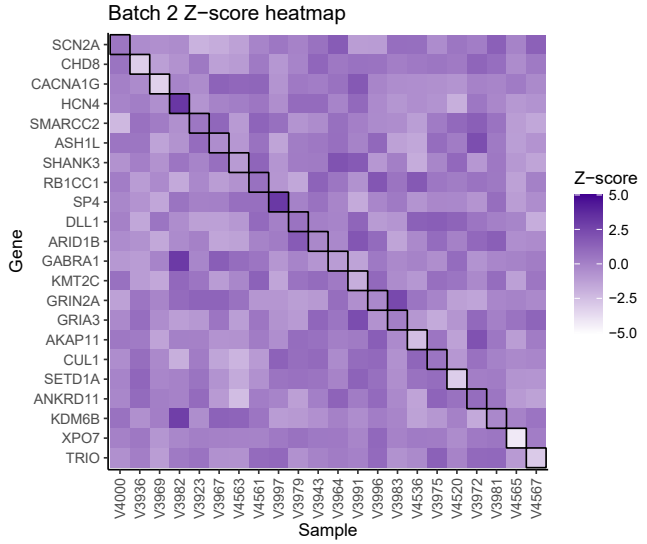

V3842
V3850
V3856
V3858
V3860
V3862
V3864
V3865
V3866
V3873
V3877
V3880
V3882
V3889
V3894
V3897
V3909
V3911
V3921
V3928
V3932
V3948
V3958
Allelic Ratio
Log10(P-Value)

Figure S3. RNA-seq-based eSNP-Karyotyping (related to Figure 4). Each panel showed the moving average of SNP intensity (RNA-seq reads) of the two alleles of heterozygous SNPs. Only the allelic ratio graphs of batch 1 edited lines are shown here; the chromosomal heterozygosity graphs and the results for all three batches are in https://zenodo.org/records/11591445. Note that only an autosomal peak interval marked with a red bar above the peak is considered to be significantly "abnormal", and no iPSC lines showed such abnormality in this batch.


Batch 1 Z−score heatmap
Gene
SCN2A
CHD8
CACNA1G
HCN4
SMARCC2
ASH1L
SHANK3
RB1CC1
SP4
DLL1
ARID1B
GABRA1
KMT2C
GRIN2A
GRIA3
AKAP11
CUL1
SETD1A
ANKRD11
KDM6B
XPO7
TRIO
V3842
V3850
V3855
V3858
V3860
V3862
V3864
V3866
V3873
V3877
V3880
V3882
V3889
V3894
V3897
V3909
V3911
V3921
V3928
V3932
V3948
V3958
Sample
Z−score
5.0
2.5
0.0
−2.5
−5.0


Batch 2 Z−score heatmap
Gene
SCN2A
CHD8
CACNA1G
HCN4
SMARCC2
ASH1L
SHANK3
RB1CC1
SP4
DLL1
ARID1B
GABRA1
KMT2C
GRIN2A
GRIA3
AKAP11
CUL1
SETD1A
ANKRD11
KDM6B
XPO7
TRIO
V4000
V3936
V3969
V3982
V3923
V3967
V4563
V4561
V3997
V3979
V3943
V3964
V3991
V3996
V3983
V4536
V3975
V4520
V3972
V3981
V4565
V4567
Sample
Z−score
5.0
2.5
0.0
−2.5
−5.0

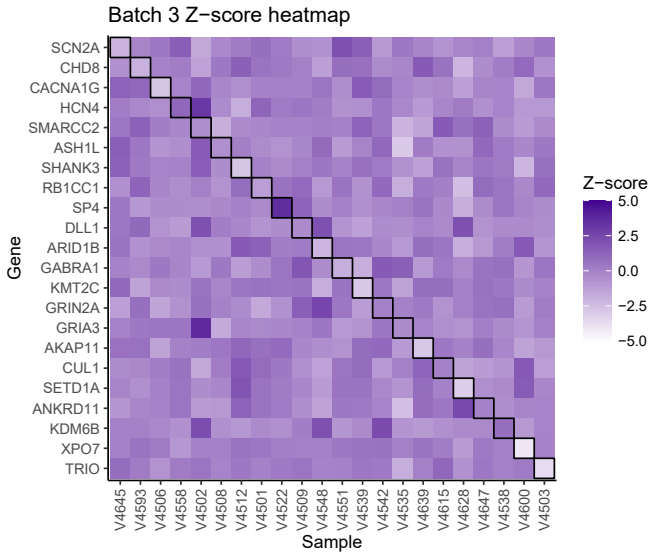
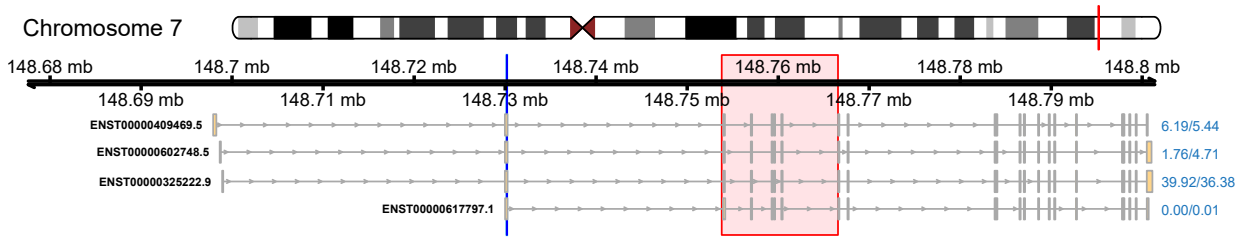
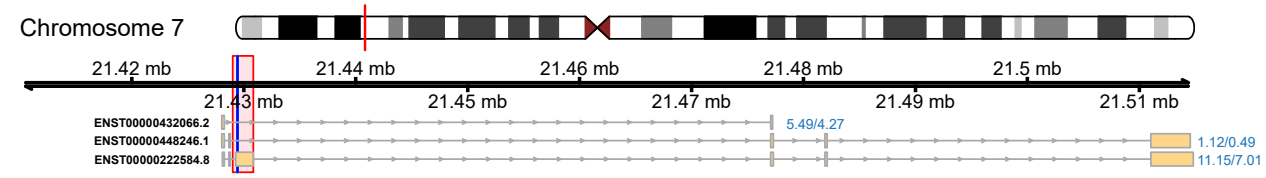

Batch 3 Z−score heatmap
Gene
SCN2A
CHD8
CACNA1G
HCN4
SMARCC2
ASH1L
SHANK3
RB1CC1
SP4
DLL1
ARID1B
GABRA1
KMT2C
GRIN2A
GRIA3
AKAP11
CUL1
SETD1A
ANKRD11
KDM6B
XPO7
TRIO
V4645
V4593
V4506
V4558
V4502
V4508
V4512
V4501
V4522
V4509
V4548
V4551
V4539
V4542
V4535
V4639
V4615
V4628
V4647
V4538
V4600
V4503
Sample
Z−score
5.0
2.5
0.0
−2.5
−5.0

Figure S4. RNA-seq-based estimation of the expected NMD for each iSTOP LoF allele derived from all three donor hiPSC lines (related to Figure 5). Shown are heat maps of Z-scored expression values (counts per million reads, CPM) of each gene in each iSTOP LoF hiPSC line for all three batches. Lined boxes indicate the edited hiPSC line that is expected to show NMD for a specific target gene.

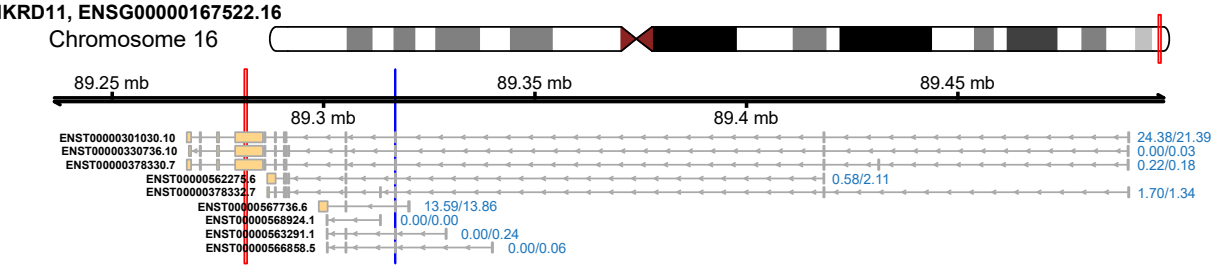
CUL1, ENSG00000055130.17



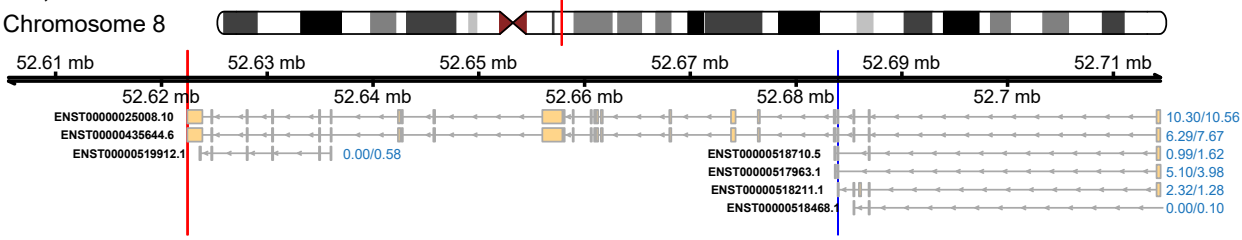
SP4, ENSG00000105866.15



ANKRD11, ENSG00000167522.16



RB1CC1, ENSG00000023287.13



ARID1B, ENSG00000049618.24

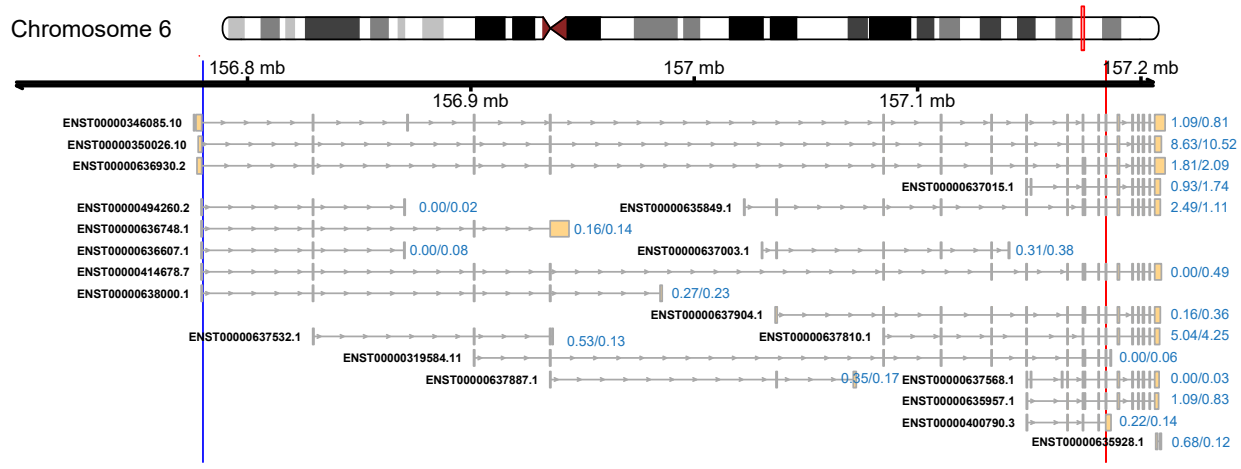


Figure S5. Transcript isoforms and the relative genomics positions of iSTOP sgRNA and antibody for each gene assayed by Western blot assay (related to Figure 5). Red boxes indicate the position of the antibody epitope or part of the protein used for generating the antibody (vendor's information), and the blue line indicates the position of sgRNA used for iSTOP mutagenesis. Only isoforms with non-zero expression (transcript per million reads, TPM) in Table S5 were included. The expression level of each transcript in the iSTOP LoF line for a specific transcript and the average expression of the transcript across all 22 iPSC lines are shown on the right side of each transcript and separated by "/".

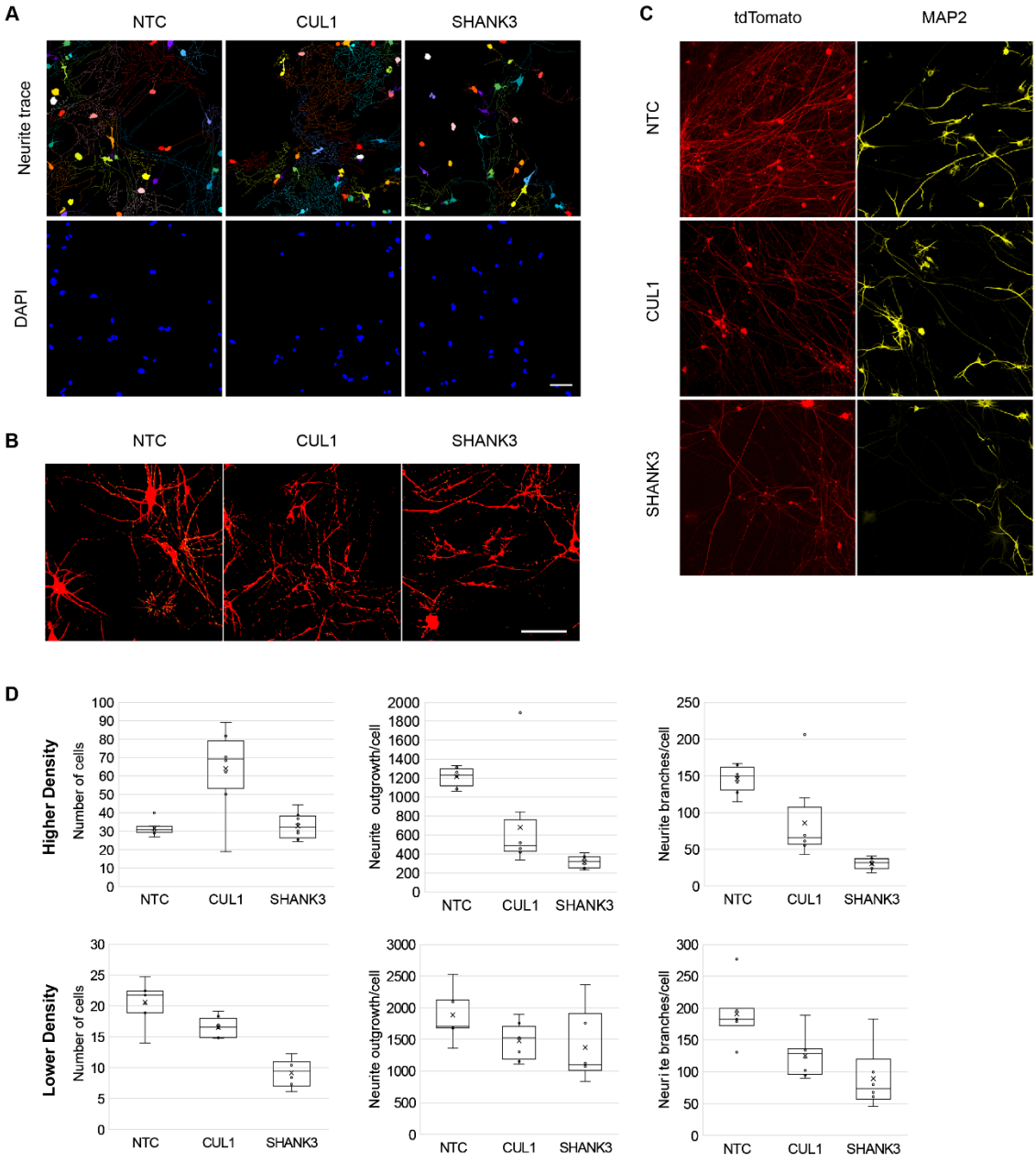

A
NTC
CUL1
SHANK3
Neurite trace
DAPI
B
NTC
CUL1
SHANK3
C
tdTomato
MAP2
NTC
CUL1
SHANK3
D
Higher Density
Number of cells
Neurite outgrowth/cell
Neurite branches/cell
Lower Density
Number of cells
Neurite outgrowth/cell
Neuri te branches/cell
NTC
CUL1
SHANK3

Figure S6. Cell segmentation and binary masks in morphometric analysis using the high content imaging and neurite analysis in low-density neuron/glia co-culture (related to Figure 6). (A) Cell segmentation for assaying neurite outgrowth and branches of LoF alleles of SHANK3 (-/-) and CUL1 (+/-) using the built-in neurite outgrowth module on the ImageXpress system. Single neurons were highlighted in rainbow color. (B) Binary masks for assaying puncta density of LoF alleles of SHANK3 (-/-) and CUL1 (+/-) using a customized synaptic module. MAP2+ and tdTomato+ mask was highlighted in red, and SYN1+ mask was highlighted in yellow. (C) Neurite traces of low-density cell culture in an independent experiment with a second parental line (KOLF2.2J) for validating the observed neurite deficits in (A) and (B). (D) The number of neurons per field of view (FOV) and the assayed neurite outgrowth/branches/cells in higher-density cultures (upper panels) or lower-density cultures (lower panels). Scale bar in all panels: 20 μm.

Supplementary Tables

Table S1. Sequences and other genomics information for iSTOP sgRNAs. Related to the main text.

iSTOP sgRNA sequences and qPCR assays for testing CBE editing efficiency												
APOE-iSTOP1-sgRNA	GGAAcAACTGACCCCGGTGG											
APOE-iSTOP2-sgRNA	GGTGcAGACACTGTCTGAGC											
iSTOP sgRNAs for Batch1-MiNND genes												
Gene name	**chr of gene**	**coordinate (hg38 bp position)**	**Strand**	**Targeted Codon**	**Number Of Isoforms**	**Relative Position in the Largest**	**NMD Prediction (%)**	**PAM: NGG**	**Off-target Prediction PAM: NGG**	**Spacing PAM: NGG**	**9 of the top 10 prioritized genes**	
ARID1B	chr6	156778374	+	CAG	2	0.066	100	GCCTcAGCCCGGCCCCGACA	0	good	ARID1B	
CACNA1G	chr17	50561655	+	CAG	32	0.027	100	TGAGCcAGGACAGCCGCCCG	0	optimal	CACNA1G	
CHD8	chr14	21429305	-	CAG	3	0.113	100	GTCCTCcAGCAGCCACAGTC	0	good	CHD8	
DLL1	chr6 KI270798v1_alt	17326	-	CAG	2	0.024	50	GTGTcAGGTAGGCGGGCAGG	0	good	DLL1	
GABRA1	chr5	161882647	+	CAG	4	0.473	100	CTAAACcAGTATGACCTTCT	0	good	GABRA1	
KMT2C	chr7	152330728	-	CAA	2	0.018	100	CAAAGAAcAATCTGCAGAAG	0	ok	KMT2C	
SCN2A	chr2	165295938	+	CAG	4	0.019	100	CAAAcAGGAACGCAAGGATG	1	good	SCN2A	
SHANK3	chr22	50679171	+	TGG	2	0.145	100	CTCCTGcCAGCCATTCTCGT	0	good	SHANK3	
SMARCC2	chr12	56189395	-	CAG	6	0.018	100	TGACCcAGTTCGACAACGTG	0	optimal	SMARCC2	
AKAP11	chr13	42299467	+	CAG	1	0.126	100	CTCTGGAcAGCAGAAGTCAT	0	ok		
ASH1L	chr1	155481573	-	CAG	3	0.146	66.67	CCCCACTcAGGAACCGCTTA	0	ok		
CUL1	chr7	148730185	+	TGG	4	0.027	100	GGTCGTCcCAGATCTGGTCC	0	ok		
GRIA3	chrX	123184560	+	CAA	6	0.009	100	GGGGcAAAGCGTGCTCCGGG	0	good		
GRIN2A	chr16	10180359	-	TGG	3	0.012	100	CCGCGCcAGACCAGAAGGGC	0	good		
HERC1	chr15	63775588	-	TGG	1	0.002	100	TCAAGcCATTTCAGCTTCAC	0	optimal		
KDM6B	chr17	7845676	+	TGG	2	0.024	100	GCAGCcATGCGCTACGAGGA	0	optimal		
RB1CC1	chr8	52683964	-	CAA	3	0.025	66.67	TATTcAACACCAGGTGCTGG	0	good		
SETD1A	chr16	30963486	+	CAG	1	0.111	100	CCCCTcAGACTGTGCCCACT	0	optimal		
SP4	chr7	21429421	+	CAA	2	0.111	50	TcAACCACAACAGCTAGAAC	0	good		
XPO7	chr8	21966923	+	CAG	2	0.027	100	ACTCcAGGCAGAGAAAGCCT	0	good		
TRIO	chr5	14280387	+	CGA	5	0.032	60	GAATAcGACAGGAGGATCTC	0	optimal		
HCN4	chr15	73368150	-	CAA	1	0.033	100	GCCGCcAAGACCCCAGCCGC	0	optimal		
ANKRD11	chr16	89316947	-	CAG	7	0.009	85.71	GGAGAAGcAGACTGGGAAAA	0	ok		

Table S2. Cell lines for generating LoF alleles (EA=European ancestry).

Cell Line	Source	Clinical Diagnosis	Age	Sex	Ethnicity
CW20107	CIRM	control	21	female	Caucasian
KOLF2.2J	Jackson Lab	control	55-59	male	EA
CD14	MGS	control	33	female	EA

Table S3. Genotypes at the iSTOP site for each LoF hiPSC line (NA=not available; no editing)

Cell line ID	Cell line ID_Alias	Batch	Source iPSC lines	Gene	Chr	Position (bp; hg38)	iSTOP genotype
g1b1c1	V3909	1	CW20107	AKAP11	chr13	42299467	C/T
g2b1c1	V3862	1	CW20107	ASH1L	chr1	155481573	A/A
g3b1c1	V3855	1	CW20107	CACNA1G	chr17	50561655	T/T
g4b1c1	V3911	1	CW20107	CUL1	chr7	148730185	G/A
g5b1c1	V3897	1	CW20107	GRIA3	chrX	123184560	T/T
g6b1c1	V3894	1	CW20107	GRIN2A	chr16	10180359	T/T
NA	NA	1	CW20107	HERC1	chr15	63775588	NA
g8b1c1	V3932	1	CW20107	KDMB6	chr17	7845676	A/A
g9b1c1	V3866	1	CW20107	RB1CC1	chr8	52683964	A/A
g10b1c1	V3921	1	CW20107	SETD1A	chr16	30963486	C/T
g11b1c1	V3873	1	CW20107	SP4	chr7	21429421	T/T
g12b1c1	V3948	1	CW20107	XPO7	chr8	21966923	T/T
g13b1c1	V3958	1	CW20107	TRIO	chr5	14280387	C/T
g14b1c1	V3858	1	CW20107	HCN4	chr15	73368150	A/A
g15b1c1	V3928	1	CW20107	ANKRD11	chr16	89316947	G/A
g16b1c1	V3864	1	CW20107	SHANK3	chr22	50679170	A/A
g17b1c1	V3850	1	CW20107	CHD8	chr14	21429305	A/A
g18b1c1	V3889	1	CW20107	KMT2C	chr7	152330728	G/A
g19b1c1	V3860	1	CW20107	SMARCC2	chr12	56189395	A/A
g20b1c1	V3842	1	CW20107	SCN2A	chr2	165295938	T/T
g21b1c1	V3877	1	CW20107	DLL1	chr6	170290088	A/A
g22b1c1	V3880	1	CW20107	ARID1B	chr6	156778374	T/T
g23b1c1	V3882	1	CW20107	GABRA1	chr5	161882647	C/T
g1b2c2	V4536	2	KOLF2.2	AKAP11	chr13	42299467	C/T
g2b2c2	V3967	2	KOLF2.2	ASH1L	chr1	155481573	A/A
g3b2c2	V3969	2	KOLF2.2	CACNA1G	chr17	50561655	T/T
g4b2c2	V3975	2	KOLF2.2	CUL1	chr7	148730185	G/A
g5b2c2	V3983	2	KOLF2.2	GRIA3	chrX	123184560	T/T
g6b2c2	V3996	2	KOLF2.2	GRIN2A	chr16	10180359	T/T
NA	NA	2	KOLF2.2	HERC1	chr15	63775588	NA
g8b2c2	V3981	2	KOLF2.2	KDMB6	chr17	7845676	A/A
g9b2c2	V4561	2	KOLF2.2	RB1CC1	chr8	52683964	A/A
g10b2c2	V4520	2	KOLF2.2	SETD1A	chr16	30963486	C/T
g11b2c2	V3997	2	KOLF2.2	SP4	chr7	21429421	T/T
g12b2c2	V4565	2	KOLF2.2	XPO7	chr8	21966923	T/T
g13b2c2	V4567	2	KOLF2.2	TRIO	chr5	14280387	C/T
g14b2c2	V3982	2	KOLF2.2	HCN4	chr15	73368150	A/A
g15b2c2	V3972	2	KOLF2.2	ANKRD11	chr16	89316947	G/A
g16b2c2	V4563	2	KOLF2.2	SHANK3	chr22	50679170	G/A

g17b2c2	V3936	2	KOLF2.2	CHD8	chr14	21429305	A/A
g18b2c2	V3991	2	KOLF2.2	KMT2C	chr7	152330728	G/A
g19b2c2	V3923	2	KOLF2.2	SMARCC2	chr12	56189395	A/A
g20b2c2	V4000	2	KOLF2.2	SCN2A	chr2	165295938	T/T
g21b2c2	V3979	2	KOLF2.2	DLL1	chr6	170290088	A/A
g22b2c2	V3943	2	KOLF2.2	ARID1B	chr6	156778374	T/T
g23b2c2	V3964	2	KOLF2.2	GABRA1	chr5	161882647	T/T
g1b3c3	V4639	3	CD14	AKAP11	chr13	42299467	C/T
g2b3c3	V4508	3	CD14	ASH1L	chr1	155481573	A/A
g3b3c3	V4506	3	CD14	CACNA1G	chr17	50561655	T/T
g4b3c3	V4615	3	CD14	CUL1	chr7	148730185	G/A
g5b3c3	V4535	3	CD14	GRIA3	chrX	123184560	T/T
g6b3c3	V4542	3	CD14	GRIN2A	chr16	10180359	T/T
NA	NA	3	CD14	HERC1	chr15	63775588	NA
g8b3c3	V4538	3	CD14	KDMB6	chr17	7845676	A/A
g9b3c3	V4501	3	CD14	RB1CC1	chr8	52683964	A/A
g10b3c3	V4628	3	CD14	SETD1A	chr16	30963486	C/T
g11b3c3	V4522	3	CD14	SP4	chr7	21429421	T/T
g12b3c3	V4600	3	CD14	XPO7	chr8	21966923	T/T
g13b3c3	V4503	3	CD14	TRIO	chr5	14280387	C/T
g14b3c3	V4558	3	CD14	HCN4	chr15	73368150	A/A
g15b3c3	V4647	3	CD14	ANKRD11	chr16	89316947	G/A
g16b3c3	V4512	3	CD14	SHANK3	chr22	50679170	A/A
g17b3c3	V4593	3	CD14	CHD8	chr14	21429305	A/A
g18b3c3	V4539	3	CD14	KMT2C	chr7	152330728	A/A
g19b3c3	V4502	3	CD14	SMARCC2	chr12	56189395	A/A
g20b3c3	V4645	3	CD14	SCN2A	chr2	165295938	T/T
g21b3c3	V4509	3	CD14	DLL1	chr6	170290088	A/A
g22b3c3	V4548	3	CD14	ARID1B	chr6	156778374	T/T
g23b3c3	V4551	3	CD14	GABRA1	chr5	161882647	C/T

Table S4. Chromosomal abnormality detected by eSNP-karyotyping in iPSC line V4167. Related to the main text, Fig. S4

chr	SNP position	-Log-*p*
chr12	57759165	3.85054963
chr12	57796735	3.85054963
chr12	59325835	3.92460732
chr12	59326071	3.89474465
chr12	59326091	4.03153324
chr12	62466937	4.20311241
chr12	62644140	4.16558805
chr12	62646492	4.24742666
chr12	62647908	4.3456919
chr12	64194285	4.45671364
chr12	64449074	4.45671364
chr12	64449075	4.36892906
chr12	64449084	4.40883795
chr12	64449140	4.48379216
chr12	64449350	4.50717128
chr12	64450281	4.48379216
chr12	64450354	4.4834706
chr12	64450369	4.50717128
chr12	64695205	4.52307202
chr12	64695505	4.52307202
chr12	64696039	4.51382747
chr12	64697060	4.52307202
chr12	64697251	4.52307202
chr12	64713715	4.64571987
chr12	64714129	4.64571987
chr12	64716472	4.64571987
chr12	65824905	4.64571987
chr12	65964288	4.64571987
chr12	65965972	4.62044753
chr12	66137398	4.52307202
chr12	66234536	4.50717128
chr12	67313337	4.51382747
chr12	67658398	4.52307202
chr12	67661419	4.52307202
chr12	67662032	4.52307202
chr12	68764959	4.48379216
chr12	68765256	4.40883795
chr12	68765355	4.44099113
chr12	68808384	4.50717128
chr12	68840106	4.50717128

chr12	68841207	4.45671364
chr12	68841626	4.40883795
chr12	68843093	4.40883795
chr12	68843224	4.45671364
chr12	68843230	4.23215593
chr12	68843233	4.38638609
chr12	68843263	4.40883795
chr12	68843698	4.40883795
chr12	68843754	4.24168138
chr12	69272155	4.24168138
chr12	69272869	4.24246392
chr12	69273496	4.24168138
chr12	69601562	3.9660467
chr12	69601565	3.86867891
chr12	70431058	3.85054963
chr12	71610121	3.81615486
chr12	71787256	3.89479184
chr12	74540073	4.06395829
chr12	75496341	4.21492583
chr12	75496553	4.2737851
chr12	75496872	4.24742666
chr12	75498671	4.23294743
chr12	75498993	4.24742666
chr12	75499601	4.40883795
chr12	75499619	4.40883795
chr12	75499706	4.48379216
chr12	76026505	4.51382747
chr12	76027761	4.51382747
chr12	76027977	4.48379216
chr12	76029027	4.48379216
chr12	76029219	4.48379216
chr12	76045879	4.48379216
chr12	76046117	4.48379216
chr12	76046488	4.48379216
chr12	76047371	4.40883795
chr12	76047374	4.2016076
chr12	76047427	3.89474465
chr12	76048101	3.85054963

Table S5. Transcript isoform expression level for all the targeted genes in LoF iPSC lines (as xlsx file). Related to the main text, Fig. 5.

Table S6. Oligos, primers, and antibodies were used for sgRNA cloning, sequencing confirmation, qPCR, and Western blot assays (as xlsx file). Related to the main text, Fig. 5, Fig. 6.

Supplemental Experimental procedures

Source hiPSC lines

The use of CW20107 (from California Institute for Regenerative Medicine - CIRM) and KOLF2.2J (from The Jackson Laboratory; an updated version of KOLF2.1J and not officially released but made available for advanced access to SSPsyGene) was part of the SSPsyGene consortium agreement on the common cell lines. KOLF2.2J, an updated version of a well-characterized KOLF2.1J, was chosen for its improved features and potential impact on the study. (Pantazis et al., 2022), with a targeted correction of a known splice-site disruption in *COL3A1* in KOLF2.1J by the MorPhic consortium (https://morphic.bio/). The original KOLF2.1J is currently listed in the catalogue of the Jackson Laboratory, but KOLF2.2J will be available soon and can already be obtained via special request. The other hiPSC line (CD14) was specific to the MiNND project and was from the Duan lab (Shi et al., 2009; Zhang et al., 2023; Zhang et al., 2020). CD14 was originally derived from the lymphocytes of a healthy donor of the Molecular Genetics of Schizophrenia (MGS) cohort (Shi *et al.*, 2009). Detailed cell line information is described in Table S2. The hiPSCs were maintained in mTeSRPlus (StemCell #100-0276) with Primocin (Invitrogen #ant-pm-1) on tissue culture plates coated with Matrigel (Fisher Scientific #08-774-552) or Geltrex (Fisher Scientific #A1413202). The study was approved by the Institutional Review Board (IRB) of NorthShore University HealthSystem, ensuring the ethical conduct of the research. More detailed information on the donor cell lines can be found in Table S2.

Chemicals and reagents

The chemicals, media, reagents used in cell culture, PCR, Sanger sequencing, and other main experiments include: BbsI-HF (NEB: R3539S), NEBuilder® HiFi DNA Assembly reagent (NEB: E2621S), mTeSR Plus (StemCell: 100-0276), Primocin (InvivoGen: ant-pm-1), Matrigel (FisherScientific: 08-774-552), Geltrex (Fisher Scientific: A1413202), DPBS (no calcium, no magnesium) (Fisher Scientific: 14-190-144), ReLeSR (StemCell: 100-0483), ROCK-Inhibitor (Tocris: 1254), mFreSR (Stem Cell: 05855), Accutase (StemCell: 07920), Trypan Blue Stain (FisherScientific: T10282), Gibco Opti-MEM I Reduced Serum Medium (Fisher Scientific: 31-985-062), Lipofectamine Stem Transfection Reagent (Fisher Scientific: STEM00003), Chroman 1 (R&D Systems: 7163/10), Emricasan (R&D Systems: 7310/5), transISRIB (R&D Systems: 5284/10), Polyamine Supplement (R&D Systems: 7739/1), QuickExtract DNA Extraction Solution 1.0 (Fisher Scientific:

NC9904870), PCRx Enhancer System (Fisher Scientific: 11-495-017), Deoxynucleoside Triphosphate Set (Sigma Aldrich: 3622614001), Shrimp Alkaline Phosphatase and buffer (Fisher Scientific: 783905000UN), E.coli exonuclease I (Fisher Scientific: 70073X5000UN), BigDye Terminator v3.1 Cycle Sequencing Kit (Fisher Scientific: 4337455), HiDi Formamide (Fisher Scientific: 4311320), RNeasy Plus Mini Kit (Qiagen: 74134), QIAshredder (Qiagen: 79654), DMSO (Sigma Aldrich: D2650-100ML), Fetal Bovine Serum (Fisher Scientific: A3160501), Mineral Oil (Sigma Aldrich: M5310), 16% Formaldehyde (Fisher Scientific: PI28908), DAPI (Fisher Scientific: EN62248). The antibodies used in immunofluorescence staining include: Synapsin I antibody (Synaptic Systems: 106-011), goat anti-tdTomato antibody (Fisher Scientific: 50-167-1115), anti-MAP2 antibody (Sigma Aldrich: AB5543), Donkey Anti-Mouse 488 (Fisher Scientific: A21202), Donkey Anti-Goat IgG (H+L) Cross Adsorbed Secondary Antibody, Alexa Fluor 568 (Fisher Scientific: A11057), Donkey Anti-Chicken IgY (H+L) Highly Cross Adsorbed Secondary Antibody, Alexa Fluor 647 (Fisher Scientfic: A78952). The main lab supplies include: 96-well non skirted PCR plate (DotScientific: 650-PCR), Corning™ Internally Threaded Cryogenic Vials 2 ML (FisherScientific: 03-374-21), 96-well plates (Corning: 353072), 24-well plates (ThermoFisher: 142475), 6-well plates (StemCell: 38016).

iSTOP gRNA design and cloning

We first retrieved the best pre-computed iSTOP-gRNA for each selected NPD gene using the iSTOP web-based tool (https://www.ciccialab-database.com/istop/#/) (Billon et al., 2017), requiring >50% NMD rate and in >50% transcript isoforms. NMD prediction was determined based on whether the targeted base was 55 nucleotides upstream of the final exon-exon junction (Billon *et al.*, 2017; Popp and Maquat, 2016). Whenever possible, the iSTOP-gRNA location was placed to the first half of the gene to ensure the resultant protein truncation (likely to be LoF) even without causing NMD. Finally, to minimize any possible off-target editing, all the selected iSTOP-gRNAs are free of any predicted off-target site via aligning to the human genome, allowing up to two mismatches in the first eight bases of the guide sequence (Billon *et al.*, 2017; Popp and Maquat, 2016). A total of 23 genes were designed for iSTOP-gRNAs in the current study (See Table S1). For cloning the designed sgRNAs, pDT-sgRNA (Addgene # 138271) vector was selected as gRNA carrier. The vector was digested with BbsI-HF. After gel purification, a single strand oligo with a prefix, gRNA of interest and a postfix was introduced into the vector backbone through Gibson assembly using NEBuilder® HiFi DNA

Assembly reagent. After cloning, miniprep plasmids were sequenced using M13 Rev primer. M13 Rev: 5' – caggaaacagctatgac – 3'. Example of a single strand oligo: 5' – atatcttgtggaaaggacgaaacaccgXXXXXXXXXXXXXXXXXXXXgttttagagctagaaatagcaagtta – 3'. After genotyping, correct clones were expanded, and transfection-grade plasmid was prepared using an endo-free plasmid kit (QIAGEN: 12362).

HEK293 culture, transfection, and editing evaluation

HEK293T cells were purchased from ATCC (Cat# CRL-3216) and maintained in DMEM with 10% FBS following the vendor's instructions. For transfection and editing efficiency evaluation, 90% confluent 293T culture was dissociated with Accutase at 37°C for 5 min. About 2×10^5 cells were replated into one well on 12 w tissue culture plates. 24 hr post replating, 2 μg of selected CBE plasmid, 1 μg of pEF-BFP (Addgene# 138272) plasmid and 1 μg of pDT-sgRNA plasmid carrying selected gRNA were transfected using Fugene HD (Promega # E2311) reagent with 1:3 DNA:Reagent ratio following vendor's instructions. 48hr post-transfection, BFP+/GFP+ and BFP+/GFP- cells were sorted through BD Aria Fusion Flow Cytometer and replated. 120 hr post-transfection, replated cells were collected after Accutase dissociation and 30-75 μl of QuickExtract DNA Extraction Solution (FisherSci # NC9904870) was added to the cell pellet for DNA extraction on a thermocycler. Extracted DNA was subsequently amplified for Sanger sequencing genotyping to evaluate editing efficiency at loci of interest. To evaluate the base editing efficiency by Sanger sequencing, we utilized an open-source tool EditR (https://moriaritylab.shinyapps.io/editr_v10/) (Kluesner et al., 2018) that takes Sanger sequencing .ab1 file and gRNA sequence as input. The percentage of each base under the sequencing peak (without sequencing noise reduction) will be calculated.

hiPSC culture and transfection

hiPSCs were maintained in mTeSRPlus (StemCell #100-0276) with Primocin (Invivogen #ant-pm-1) on tissue culture plates coated with Matrigel (Fisher Scientific #08-774-552) or Geltrex (Fisher Scientific #A1413202) throughout the mutagenesis process. The medium was changed every other day, and colonies were passaged every 4-6 days when cells reached 70% confluence. For DNA transfection, cells were plated at a density of $1.2\sim1.5\times10^5$ per well on a 24-well plate (ThermoFisher # 142475) in mTeSR Plus with 5μM ROCK-Inhibitor (Tocris #1254). The next day, antibiotics-free mTeSR Plus with 5μM ROCK-Inhibitor was changed on

the plate after ensuring appropriate cell density (60-70% confluence) and survival. Shortly after, each well was transfected with 750 ng pEF-AncBE4max (Addgene # 138270), 300 ng pEF-BFP, and 300 ng pDT-sgRNA containing the variant-specific gRNA using LipofectamineSTEM with 1:2.5 DNA:reagent ratio. In total, 23 different pDT-sgRNAs were used (one per well), and the 24th well was used as a negative control. Media was refreshed with regular mTeSR Plus containing Primocin at 24 hr and 48 hr post-transfection. At 72 hr post-transfection, cells were prepped for single-cell sorting.

Single hiPSC sorting and clonal culture

Single hiPSCs from the post-transfection culture above were sorted into 96 well plates with one cell per well using a BD FACSAria Fusion Flow Cytometer. All sorting procedures were done using mTeSR Plus with CEPT cocktail (1:10,000 chroman 1, emricasan, and transISRIB; 1:1,000 polyamine supplement) (Tristan et al., 2023). To prepare for FACS, cells were dissociated into single cells using Accutase (StemCell: 07920) for 7 min at 37°C. Cells were transferred to 15 ml tubes with 1 ml mTeSR plus to inactivate the Accutase and centrifuged at 300 × g for 3 min. The resulting cell pellets were resuspended in 700 μl media and filtered twice using 5 ml corning round bottom tubes with the blue strainer cap (Fisher Scientific: 0877123). Samples were placed on ice immediately to minimize clogging. Samples were processed and analyzed using the BD FACSAria Fusion Flow Cytometer, gating for BFP+/GFP+ cells, one single cell per well on a 96-well plate for each condition; for NTC, BFP+/GFP- cells were sorted. After sorting, the cells were not disturbed for 24 hr. 48 hr post sorting, 50 μl mTeSR plus was added to each well. 72 hr post sorting, 50 μl media refreshment for each well. 96 hr post posting, 120 μl media refreshment for each well. 144 hr post sorting, aspirated 120 μl and added 100 μl media for each well. Afterwards, 100 μl media refreshment every other day (10-14 days) until colonies appeared with an appropriate size to pick for Sanger sequencing genotyping. All media refreshments were performed using Integra's MINI 96 electronic pipette.

PCR and Sanger sequencing for LoF genotype confirmation

Once colonies reached an appropriate size and had stem cell-like morphology, 8-12 colonies were picked for each edited condition. DNA was extracted from the picked colonies using Quick Extract DNA Extraction Solution (Fisher Scientific NC9904870). Following PCR to amplify the DNA for each LoF gene, Sanger sequencing was completed to confirm that the appropriate base was changed at the desired location to create

a stop codon. The sequencing was performed on a 3730xl DNA Analyzer and the sequencing data were imported to SeqScape v2.5 for automatic analysis and genotype calling. Up to 4 colonies with confirmed homozygous editing or heterozygous editing (if there were no homozygous colonies) and good morphology were expanded for RNA isolation and cell cryopreservation. The oligos and primers for sequencing and qPCR were listed in Table S6.

RNA isolation for RNA sequencing (RNA-seq)

Based on sequencing results, two selected clones were expanded from one well on 96-well plates to one well on 6-well plates. Once reaching 70% confluency, cells were expanded a second time from one well to two wells per clone, one for cryopreservation and one for RNA extraction/RNA-seq. For RNA isolation, cells were lysed using 800 μl of Buffer RLT Plus (QIAGEN 1053393). Cell lysates in buffer RLT were stored at -80°C until ready to be isolated using the QIAGEN RNeasy Plus Mini Kit (QIAGEN 74134) following the vendor's instructions. Purified RNAs were sent to Novogene for RNA-seq.

RNA-seq data processing

Bulk RNA-seq was performed by external vendor Novogene and in 2×150 bp paired-end format of 25-30 M reads per sample. Briefly, raw FASTQ files were aligned to the human GRCh38.p14 genome by STAR 2.7.7 and subsequently counted by the built-in function of STAR at the gene level using the GTF file of GENCODE v42 with parameters `--quantMode Genecounts --alignSoftClipAtRefenceEnds No --outFilterScoreMinOverLread 0.30 --outFilterMatchNminOverLread 0.30`. Gene counts from each of the samples were collected by a customized script and collated into a single count matrix. Genes that had 0 counts in all samples were removed prior to analysis.

CellNet analysis for pluripotency

The RNA-seq data of each edited iSTOP hiPSC line were used for pluripotency evaluation by using the R package `CellNet` (Cahan et al., 2014). Briefly, the gene × sample count matrix generated in the previous step was loaded by `EdgeR` and a new count matrix contained log-transformed, library size-normalized CPM (counts-per-million) value was generated by `calcNormFactors()`, `estimateDisp()` and `cpm()`. Subsequently, the script constructed a random forest classifier using the in-built model from the `CellNet`

Package. Finally, the likelihood of each sample-cell type pair (in scores) was evaluated by passing the log-transformed gene × sample count matrix through the classifier and plotting the results in hierarchically clustered heatmaps.

Using RNA-seq data for e-SNP Karyotyping

As part of the high throughput LoF mutagenesis pipeline for evaluating possible hiPSC chromosomal abnormality due to editing or hiPSC clonal growth, we opted to use e-SNP Karyotyping rather than the classical G-band karyotyping. e-SNP Karyotyping detects any potential chromosomal aberrations, including duplications, loss of heterozygosity, and meiotic recombination. As we previously described (Zhang *et al.*, 2023; Zhang *et al.*, 2020), we used the e-Karyotyping R package developed by the Benven lab (github.io/BenvenLab/eSNPKaryotyping) (Weissbein et al., 2016) with customization to our current environment settings. Briefly, 2x150 bp of paired-end raw FASTQ files of all samples were firstly trimmed for adaptors and low-quality reads by Trimmomatic (Bolger et al., 2014). Only paired-end reads were kept post Trimmomatic processing. Trimmed read pairs were aligned to the human hg38 genome by Bowtie2 v2.5.1 to generate BAM files for eSNP Karyotyping package (Weissbein *et al.*, 2016). The original package code in R has been rewritten for GATK 4 and R 4.2 platforms. Only common SNPs (MAF > 0.05) from dbSNP 154 were retained for genome and zygosity block analysis. A rolling window of 151 bp in size was used to smooth the Allele Frequency (AF) curves when plotting the AF ratio and P values by genome coordinates. Each block represented 1.0 MB in zygosity block graphs.

Immunofluorescence staining for hiPSCs

hiPSCs were dissociated with Accutase (innovative cell technologies AT-104) and seeded into Matrigel (Corning 354234) coated round glass coverslips in a 24-well plate and kept in mTESR+ media (Stem cell technology 100-0275). Cells were kept until they formed medium-sized colonies. Cells were washed twice with 1× PBS and fixed with 4% PFA for 30 min. Samples were incubated with blocking buffer 4% BSA (A3803 Sigma), 1% Goat serum (ThermoFisher 16210072), 0.2% triton X-100 (BP151 Fisher BioReagents) in PBS for 1hr at room temperature (RT). Primary antibodies were incubated for 1 hr at RT. Samples were washed 3 times with PBS 0.2% Triton X-100, and secondary antibodies were incubated for 1 hr at RT, making sure samples were protected from light. Samples were washed 3 times with PBS 0.02% Triton X-100 and rinsed

with MiliQ water before mounting with Fluoroshield with DAPI (F6057 Sigma) and placed on a glass slide. The following antibodies were used: rabbit anti-Sox2 (Millipore AB5603), mouse IgG anti-Oct4 (Millipore MAB4401), mouse IgM anti-Tra-1-60 (Millipore MAB4360), goat anti-rabbit Alexa Fluor 546 (Invitrogen A11035), goat anti-mouse IgG Alexa Fluor 488 (Invitrogen A11001), goat anti-mouse IgM Alexa Fluor 647 (Invitrogen A21238). Confocal images were taken using a Zeiss LSM700 laser-scanning confocal microscope with a 20× objective.

Western blot

hiPSCs were grown on 6 well plates as stated above. When the cells reached ~80% confluency, they were washed twice with DPBS and lysed with 100µl RIPA buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with 0.5mM DTT, 1mM PMSF and 1× Protease inhibitor Cocktail (Sigma P8340) using a cell scraper. The lysate was transferred to a tube incubated on ice for 10 min and centrifuged at 14000 × g for 10 min at 4°C. The supernatant was transferred to a new tube and protein concentration was determined using a BCA protein assay (Thermo Scientific 23225) measuring absorbance at 562 nm in a plate reader (SpectraMax i3, Molecular Devices). For the SDS-PAGE, protein samples were prepared with 15µg of protein with 2× Laemmli sample buffer (161-0737 BIORAD) and heated for 5 min at 95°C. Samples were resolved using 7.5% or 10% acrylamide gels (4561023DC, 4561033DC BIO-RAD) and transferred into a 0.45 µm nitrocellulose membrane (1620115 BIO-RAD). Membranes were blocked for 1hr at RT using 5% non-fat milk dissolved in TBS-T (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% Tween 20). Primary antibodies were diluted in 5% non-fat milk (732-291-1940 LabScientific) or 5% BSA (A3803 Sigma-Adrich) in TBS-T and incubated overnight at 4°C. Membranes were washed 3 times with TBS-T and incubated with an HRP-conjugated secondary antibody diluted in 5% non-fat milk or BSA in TBS-T for 1 hr at RT. Membranes were washed 3 times with TBS-T and protein was visualized using the Clarity Western ECL Substrate (Bio-rad 1705060) and autoradiography films (XAR ALF 1318, LabScientific). Densitometry analysis was performed using ImageJ software, and β-actin or VCP were used as loading control. The following antibodies (Table S6) were used: ANKRD11 (1:1000 Santa Cruz sc-514916), BAF250b/ARID1B (1:3000 Cell Signaling 92964S), β-Actin (1:20000 Sigma-Aldrich A544), CUL1 (1:2000 Santa Cruz SC-17775),

FIP200/RB1CC1 (1:2000 Cell Signaling 12436S, SP4 (1:2000 Santa Cruz, SC-390124), VCP (1:20000, K331).

Lentivirus generation

Lentiviral vectors were generated by transfecting HEK293T cells with lentivirus packaging plasmids (pMDLg/pRRE, VsVG and pRSV-REV) with the desired vectors as previously described (Pang et al., 2011) using lipofectamine 3000. The following plasmids were used: pMDLg/pRRE (Addgene 12251), pRSV-Rev (Addgene #12253), pCMV-VSV-G (Addgene #8454), FUW-M2rtTA (Addgene #20342), FUW-TetO-Ngn2-P2A-puromycin (Addgene #52047), FUW-TetO-Ascl1-T2A-puromycin (Addgene #97329), FUW-TetO-Dlx2-IRES-hygromycin (Addgene #97330), TdTomato (Addgene #197033). Lentiviral particles were collected in mTeSR+ media and stored at -80°C until further use.

Neuron differentiation and coculture

hiPSCs were dissociated with Accutase (innovative cell technologies AT-104), cells were counted and 2×10^5 cells were plated per well in 6-well plates coated with Matrigel (Corning 354234) in mTESR+ (Stem cell technology 100-0275) with CETP cocktail (Chroman 1 50nM MedChemExpress HY-15392, Emricasan 5mM Selleck chem S7775, Polyamine supplement 1× Sigma Aldrich P8483, trans-ISRIB 700nM R&D systems 5284). A mixture of virus was added to the cell media before plating: i) Ngn2 + rtTA was added for excitatory neuron differentiation (Zhang et al., 2013), ii) Ascl1 + Dlx2 + rtTA was added for inhibitory neuron differentiation (Yang et al., 2017). Excitatory neurons were also transduced with a lentivirus with a plasmid expressing TdTomato on day 4. to distinguish them from inhibitory neurons. On day 1, the media was changed to Neurobasal (Gibco 21103-049) supplemented with B27 (Gibco 17504044) and GlutaMAX (Gibco 35050061), doxycycline (2µg/ml, MP biomedicals 198955) was added to the media and kept for 7 days. On days 2 and 3, infected cells were selected with Puromycin (1 µg/ml, Sigma-Aldrich P8833) for excitatory neurons or Puromycin (1µg/ml) and Hygromycin (100 µg/ml, Sigma-Aldrich H9773) for inhibitory neurons. On day 4, 8×10^3 primary mouse glia were plated into Matrigel-coated wells in a 96-well plate. On day 5, induced neurons were dissociated with Accutase and counted, and 12×10^3 excitatory and 6×10^3 inhibitory cells were seeded per well into the coverslips with mouse glia in neurobasal media with 5% FBS (R&D systems S11550) and CEPT cocktail. On day 6 media was changed with neurobasal (with B27 and GlutaMAX) 5% FBS with BDNF

(10ng/ml, PeproTech 10781-164), GDNF (10 ng/ml, PeproTech 10781-226) and NT3 (10 ng/ml, PeproTech 10781-174), Cytosine β-D-arabinofuranoside (AraC 2-4 μM Sigma-Aldrich C1768) was added to the media to stop glia proliferation. Half the media was changed every 5 days with neurobasal 5% FBS with BDNF, GDNF and NT3. On day 35 cells were washed 2 times with PBS 1× and fixed with 4% PFA for 30 min. Cells were left in PBS 0.02% sodium azide until staining.

High-content imaging

For immunofluorescence staining, **h**iPSC-derived neurons were washed twice with 1× PBS and fixed with 4% PFA for 30 min in a 96-well optical bottom plate with a polymer base (Fisher Scientific: 12-566-70) at Rutgers University (New Brunswick, NJ). Fixed neurons were stored at 4ºC in 1× PBS with 0.02% sodium azide and shipped overnight to NorthShore Research Institute (Evanston, IL). Neurons were permeabilized in 1× PBS with 0.5% Triton X-100 for 15 minutes at RT without shaking. After blocking with 3% BSA and 0.1% Triton X-100 in 1× PBS for 1hr at RT, the neurons were stained with primary antibodies, mouse anti-Synapsin 1 (1:500), goat anti-tdTomato (1 μg/ml), and chicken anti-MAP2 (1:5000), in blocking buffer for 1.5hr at RT. The samples were washed three times in 1× PBS with 0.1% Triton X-100 (0.1% PBST) for 5 min each and incubated with the secondary antibodies Donkey anti-mouse Alexa 488 (1:1000), donkey anti-goat Alexa 568 (1:1000), and donkey anti-chicken Alexa 647 (1:1000) in blocking buffer for 1 hr at RT in the dark. Next, the neurons were washed twice with 0.1% PBST for 5 min, and incubated with DAPI (0.5 μg/mL, Fisher Scientific, EN62248) at RT for 10 min. Neurons were washed with 0.05% sodium azide in PBS. The plate was stored at 4℃ and allowed to warm to RT before imaging.

For Image acquisition, the neurons were imaged using Molecular Devices (San Jose, CA) ImageXpress Micro Confocal High-Content Imaging System at both 20× and 40×. The laser wavelengths used were DAPI, FITC, Texas Red, and Cy5. Each well in the 96 well plate was imaged at 8 sites for 40× and 9 sites for 20× with 8-10 *z* stacks at 1 μm step size. For the 40× objective, the pixel size is 0.3438 μm^2 with a pinhole of 60 μm, and the 20× objective pixel size is 0.6842 μm^2 also with a pinhole of 60μm.

For image analyses, the acquired mages were analyzed as 2D maximum projection. The first two morphometrics, the mean number of neurite branches per cell and the mean length of neurite outgrowth per cell were analyzed with the built-in Neurite Outgrowth Application Module within the MetaXPress 6 software,

version 6.7.2.290. Both the mean number of neurite branches per cell and the mean length of neurite outgrowth per cell were calculated based on DAPI stain as a nuclear marker and tdTomato stain, which labels excitatory neurons (see a generation of neuron culture methods), as neurite and cell body marker. Cell bodies were defined with an approximate maximum width of 30 μm, a minimum area of 300 μm^2, and a pixel value of at least 1500 above the local background level. Nuclei were identified with an approximate minimum width of 8μm, an approximate maximum width of 20 μm, and a pixel value of at least 1500 above the local background level. Neurite outgrowths were determined with a maximum width of 2 μm, a minimum projection length of 15 μm from the cell body, and a pixel value of at least 500 above the local background level. The 20× objective images were used for neurite outgrowth analysis. For assaying the third morphometrics, excitatory synapse density, we used an in-house generated custom synaptic assay module with MetaXPress 6 software. Specifically, puncta were identified through Synapsin1 staining with an approximate minimum width of 0.5 μm, an approximate maximum width of 2 μm, and a minimum pixel value of 2500 above the local background level. The number and area of Synapsin1 positive puncta within the colocalized MAP2 and tdTomato signals were used for analysis. The puncta density was generated by the number of total area of puncta within the colocalized MAP2 and tdTomato staining divided by the area of MAP2+& tdTomato+ signal within the neurites. The 40× objective images were used for synaptic puncta density analysis.

Statistical analyses

Pearson's correlation was used to evaluate the correlations between the two groups. For high content imaging, the cellular phenotypic measurements were from 8 replicates (wells with independent cell cultures), and each well's data were averaged from 9 images. For western blot, lysates of 4 different cell cultures of 2 different passages were included. However, per the journal's policy on the type of replicate and upon the editor's request, we did not present the *P*-values from any statistical testing. To determine chromosomal abnormalities in SNP e-Karyotyping, we use the statistical cut-off as described in the method (Weissbein *et al.*, 2016).

Supplementary References

Billon, P., Bryant, E.E., Joseph, S.A., Nambiar, T.S., Hayward, S.B., Rothstein, R., and Ciccia, A. (2017). CRISPR-Mediated Base Editing Enables Efficient Disruption of Eukaryotic Genes through Induction of STOP Codons. Mol Cell *67*, 1068-1079 e1064. 10.1016/j.molcel.2017.08.008.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114-2120. 10.1093/bioinformatics/btu170.

Cahan, P., Li, H., Morris, S.A., Lummertz da Rocha, E., Daley, G.Q., and Collins, J.J. (2014). CellNet: network biology applied to stem cell engineering. Cell *158*, 903-915. 10.1016/j.cell.2014.07.020.

Kluesner, M.G., Nedveck, D.A., Lahr, W.S., Garbe, J.R., Abrahante, J.E., Webber, B.R., and Moriarity, B.S. (2018). EditR: A Method to Quantify Base Editing from Sanger Sequencing. Crispr j *1*, 239-250. 10.1089/crispr.2018.0014.

Pang, Z.P., Yang, N., Vierbuchen, T., Ostermeier, A., Fuentes, D.R., Yang, T.Q., Citri, A., Sebastiano, V., Marro, S., Sudhof, T.C., and Wernig, M. (2011). Induction of human neuronal cells by defined transcription factors. Nature *476*, 220-223. nature10202 [pii]

10.1038/nature10202.

Pantazis, C.B., Yang, A., Lara, E., McDonough, J.A., Blauwendraat, C., Peng, L., Oguro, H., Kanaujiya, J., Zou, J., Sebesta, D., et al. (2022). A reference human induced pluripotent stem cell line for large-scale collaborative studies. Cell Stem Cell *29*, 1685-1702.e1622. 10.1016/j.stem.2022.11.004.

Popp, M.W., and Maquat, L.E. (2016). Leveraging Rules of Nonsense-Mediated mRNA Decay for Genome Engineering and Personalized Medicine. Cell *165*, 1319-1322. 10.1016/j.cell.2016.05.053.

Shi, J., Levinson, D.F., Duan, J., Sanders, A.R., Zheng, Y., Pe'er, I., Dudbridge, F., Holmans, P.A., Whittemore, A.S., Mowry, B.J., et al. (2009). Common variants on chromosome 6p22.1 are associated with schizophrenia. Nature *460*, 753-757. 10.1038/nature08192.

Tristan, C.A., Hong, H., Jethmalani, Y., Chen, Y., Weber, C., Chu, P.H., Ryu, S., Jovanovic, V.M., Hur, I., Voss, T.C., et al. (2023). Efficient and safe single-cell cloning of human pluripotent stem cells using the CEPT cocktail. Nat Protoc *18*, 58-80. 10.1038/s41596-022-00753-z.

Weissbein, U., Schachter, M., Egli, D., and Benvenisty, N. (2016). Analysis of chromosomal aberrations and recombination by allelic bias in RNA-Seq. Nat Commun *7*, 12144. 10.1038/ncomms12144.

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. 10.1038/nmeth.4291.

Zhang, S., Zhang, H., Forrest, M.P., Zhou, Y., Sun, X., Bagchi, V.A., Kozlova, A., Santos, M.D., Piguel, N.H., Dionisio, L.E., et al. (2023). Multiple genes in a single GWAS risk locus synergistically mediate aberrant synaptic development and function in human neurons. Cell Genom *3*, 100399. 10.1016/j.xgen.2023.100399.

Zhang, S., Zhang, H., Zhou, Y., Qiao, M., Zhao, S., Kozlova, A., Shi, J., Sanders, A.R., Wang, G., Luo, K., et al. (2020). Allele-specific open chromatin in human iPSC neurons elucidates functional disease variants. Science *369*, 561-565. 10.1126/science.aay3983.

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. 10.1016/j.neuron.2013.05.029.