Stem Cell Reports, Volume *9*

Supplemental Information

FACS-Assisted CRISPR-Cas9 Genome Editing Facilitates Parkinson's Disease Modeling

Jonathan Arias-Fuenzalida, Javier Jarazo, Xiaobing Qing, Jonas Walter, Gemma Gomez-Giro, Sarah Louise Nickels, Holm Zaehres, Hans Robert Schöler, and Jens Christian Schwamborn

FACS assisted CRISPR-Cas9 genome editing facilitates Parkinson's disease modeling

- Supplemental Information -

Jonathan Arias-Fuenzalida [1,2,7], Javier Jarazo [1,7], Xiaobing Qing [1], Jonas Walter [1] Gemma Gomez-Giro [1,4], Sarah Louise Nickels [1,3], Holm Zaehres [4,5], Hans Robert Schöler [4,6], Jens Christian Schwamborn [1]

[1] Luxembourg Centre for Systems Biomedicine (LCSB), Developmental and Cellular Biology, University of Luxembourg, L-4362, 7 avenue des Hauts-Fourneaux, Luxembourg

[2] Graduate School of Biostudies, Kyoto University, Kyoto 606-8502, Japan

[3] Life Science Research Unit (LSRU), Systems Biology, University of Luxembourg, L-4367, 6 avenue du swing, Luxembourg

[4] Max Planck Institute for Molecular Biomedicine, Laboratory of Cell and Developmental Biology, Roentgenstrasse 20, Muenster, Germany

[5] Ruhr-University Bochum, Medical Faculty, Department of Anatomy and Molecular Embryology, 44801 Bochum, Germany

[6] Westphalian Wilhelms University Muenster, Medical Faculty, 48149 Muenster, Germany

[7] These authors equally contributed to the article

Correspondence should be addressed to J.S. (jens.schwamborn@uni.lu)

Supplemental Experimental procedures

Stem cell culture and electroporation. The following human induced pluripotent stem (iPS) cells reprogrammed with non-integrative episomal methods were used: A13777 (Gibco cat no. A13777) from female cord blood-derived $CD34^{pos}$ cells. Cell lines were cultured in Essential 8 medium (Thermo Fisher cat no. A1517001) on Geltrex (Thermo Fisher cat no. A1413301) or matrigel. Cells were normally dissociated with accutase (Thermo Fisher cat no. A1110501) and plated in media containing ROCK inhibitor Y27632 (Sigma cat no. Y0503) at 10µM for 24h after dissociation. Cells were subjected to positive selection with puromycin (Sigma cat no. P9620) at a concentration of 0.5µg/mL. Cells were electroporated using 4D-Nucleofector System (Lonza) and a 4D kit for human dermal fibroblast (Lonza cat no. V4XP). Parental pre-electroporation line presents micro-duplication 20q11.21.

Construction of sgRNA vectors and donor plasmids. Cas9 target sequences with predicted high catalytic activity were selected (Doench et al., 2014) (Table S2) and cloned into pX330 vector (Addgene 42230) as previously described (Ran et al., 2013). Primers used are indicated in Table S4. The donor vectors were pDONOR-SNCAe2-WT (Addgene 85845), pDONOR-SNCAe2-A30P (Addgene 85846), pDONOR-SNCAe3-WT (Addgene 85847), pDONOR-SNCAe3-A53T (Addgene 85848) and pDONOR-PINK1e5-I368N (Addgene 86154) in EGFP and dTOMATO containing versions. Homology arms were assembled by conventional methods (Gibson, 2011) on donor scaffolds pDONOR-tagBFP-PSM-EGFP (Addgene 100603) and pDONOR-tagBFP-PSM-dTOMATO (Addgene 100604).

In vitro RNA transcription and mRNA transfection. The coding sequence of codon-optimized hyperactive transposase Piggybac from Trichoplusia ni (Yusa et al., 2011) and the excision-only mutant (R372A/K375A) (Li et al., 2013) were amplified to incorporate the T7 promoter. Primers used are indicated in Table S4. The PCR product was used as template for in vitro transcription with an mMESSAGE mMACHINE T7 kit (Thermo Fisher cat no. AM1344) according to the manufacturer's protocol. The transcript was poly-adenylated with a Poly(A) tailing kit (Thermo Fisher cat no. AM1350) and purified with a MEGAclear transcription clean-up kit (Thermo Fisher cat no. AM1908). The transcript quality was evaluated with a Bionalayzer RNA 6000 nano (Aglient cat no. 5067-1511). Transfection was performed with Stemfect RNA transfection kit (Stemgent cat no. 00-0069) according to the manufacturer's protocol.

Fluorescent Activated Cell Sorting. FACS was conducted using sterile line sorting on a baseline and CST calibrated BD FACS ARIA III. Drop delay calibrations were ensured prior to each sample. For all human iPS cells an 85µm nozzle, a yield or purity-sorting mask and neutral density filter 2.0 were used. Cells were pre-separated with 35µm and 20µm strainers (Corning cat no. 352235 and Miltenyi cat no. 130-101-812). Sorting was conducted with single cell exclusive gating hierarchies on FSC and SSC wide and high (Figure S1A). Use of strainers and single cell gating is highly recommended (Figure S1B). For efficiency analysis, live cells were quantified by SYTOX Blue Dead Cell Stain (Thermo Fisher cat no. S34857).

Characterization of polyclones. Composition of polyclones was assessed by sub-cloning. Single cell clones were expanded and genomic DNA extracted using QuickExtract solution (Epicentre cat no. QE09050). Clones were genotyped for the left homology arm junction, right homology arm junction, and wild type junction as indicated in Figure S3, using primers in Table S4. PCR products of the left homology arm were used for Sanger sequencing of subclones of *SNCA*e2(p.A30P) polyclone 632 and *SNCA*e3(p.A53T) polyclone 636 as shown in Figure S3. The wild type junction was used for Sanger sequencing of subclones of transposed *SNCA*e2(p.A30P) polyclone 632 and transposed *SNCA*e3(p.A53T) polyclone 636 as shown in Figure S3.

Microarray Karyotype. Genomic DNA from the pre-electroporation parental, and isogenic polyclones was purified using GenElute Blood genomic DNA Kit (Sigma cat no. NA2020). Samples were processed at Bonn Univesity Life&Brain genomics facility using Illumina iScan technology (Illumina).

Immunostaining. Cells were fixed on PFA and permeabilized on PBS triton-X 0.2%. For characterizing human iPS cells, primary antibodies used were OCT4 (Santa cruz cat no. sc-5279) dilution 1:100, TRA1-81 (Millipore cat no. MAB4381) dilution 1:50, SOX2 (Abcam cat no. ab97959) dilution 1:100 and SSEA4 (Millipore cat no. MAB4304) dilution 1:50. Secondary antibodies used were donkey anti-mouse alexa fluor 488 (Thermo Fisher cat no. A-21202) and donkey anti-rabbit alexa fluor 488 (Thermo Fisher cat no. A-21206), both at dilution 1:1000. For characterizing NESCs, primary antibodies used were NESTIN (BD cat no. 611659) dilution 1:600 and SOX2 (Abcam cat no. ab97959) dilution 1:200. Secondary antibodies used were donkey anti-mouse 488 (Thermo Fisher cat no. A-21202) and donkey anti-rabbit 647 (Thermo Fisher cat no. A-31573), both at dilution 1:1000. For nuclear staining, Hoechst-33342 (Thermo Fisher cat no. 62249) was used at dilution 1:1000. Images were acquired in an inverted microscope (Zeiss Axio ObserverZ1).

NESCs differentiation and culture. Human iPS cells were clustered on aggrewell plates (Stem cell technologies cat no. 27845) for 12 hours. Embryoid bodies were transferred to ultra-low attachment plates and differentiated with the program in Figure 4A. Briefly, cells were cultured on KO-DMEM (Gibco cat no. 10829018) supplemented with 20% knock-out serum replacement (Gibco cat no. A3181501), 2mM glutamax (Gibco cat no. 35050061), 1x non-essential amino acids (Gibco cat no.11140035), 1µM dorsomorphine (Sigma cat no. P5499), 3µM CHIR99021 (Sigma cat no. SML1046) and 0.5µM purmorphamine (Sigma cat no. SML0868). From day three onwards, cells were cultured on DMEM-F12:neurobasal media (1:1) supplemented with N2 (Gibco cat no. 17502048), B27 without vitamin A (Gibco cat no. 12587001) and 2mM glutamax. For day three and four, media was supplemented with 10µM SB431542 (Sigma cat no. S4317). From day five onwards, the culture was maintained with 150µM ascorbic acid (Sigma cat no. A5960), 3µM CHIR99021 and 0.5µM purmorphamine. At day six, embryoid bodies were dissociated with accutase and plated on matrigel coated plates.

Extracellular energy flux analysis. NESCs were plated on Seahorse XFe96 assay plates (Aglient) at a density of 65k cells per well and the oxygen consumption rate was quantified in a Seahorse XFe96 Analyzer. Four baseline measurements were performed before any treatment injection. Three measurements were performed after each injection as shown in Figure 4. Final concentrations of compounds were 1µM for oligomycin (Sigma cat no. 75351), FCCP (Sigma cat no. C2920), antimycin A (Sigma cat no. A8674) and rotenone (Sigma cat no. R8875). DNA was quantified using CyQUANT kit (Thermo Fisher cat no. C7026) and normalization based on DNA content as previously described (Silva et al., 2013).

Western Blotting. For western blot analysis of NESCs total protein, an antibody against α-Synuclein (C-20)-R (Santa cruz cat no. sc-7011) was used at a dilution of 1:100, and an antibody against GAPDH (abcam cat no. ab9485) was used at a dilution of 1:1000 overnight. Blots were developed using anti-rabbit IgG HRP-linked secondary antibody (GE Healthcare Life Sciences cat no. NA934V) and west-pico chemiluminescent substrate (Thermo Fisher cat no. 34080). Membranes were imaged in a Raytest Stells system with exposure of 30s for both alpha-Synuclein and GAPDH.

Microarray. RNA was extracted from healthy control NESCs using quiazol (Qiagen cat no. 79306) and miRNeasy (Qiagen cat no. 217004). Samples were processed at the EMBL Genomics Core Facility using

Affymetrix human Gene 2.0 arrays. Results were processed using GC-RMA analysis. Gene expression omnibus accession code GSE101534.

Supplemental Figure legends

Figure S1. **Gating structures, second locus, subpopulation types and transposition optimization. Related to Figure 1, Figure 2 and Figure 3**. (A) Hierarchical single cell gating structure of SSC and FSC wide and high for single cell preparations. Representative example of *PINK1*e5(p.I368N) polyclone 517. (B) Preparation of cells with a cell strainer and single cell gating structure is essential to ensure high quality sorting. Scale bar 25µm. (C) Post-selection sorting for knock-in *PINK1*e5, using independent sgRNAs: sgRNA 517 and sgRNA 526. FACS plots are represented with 2% contour lines. (D) Purity-purity sorting allows the generation of a homogenous biallelic knock-in population. (E) Post-transposition sorting. Excision-only transposase expression removes the positive selection module for *PINK1*e5. (F) FACS analysis of parental wild type (WT) control. (G) FACS analysis for homozygous and heterozygous dTOMATO^{pos} clones for *SNCA*e2. Diagram of knock-in overlapping population types is shown (right). (H) Schematic representation of overlapping populations type U and type 2. (I) Type U (one copy integration) and type 2 (two copy integration) single cell clones present high overlap in the FSC-A dimension. Clone type U population overlaps 26.1% with the gate established by the type 2 clone. (J) Indel and wild type frequency from non-targeted allele. Sanger sequencing of the non-targeted allele from population type U presents a high frequency of indels (n=20) for *SNCA*e2 and *SNCA*e3. (K) Representative chromatogram for an indel bearing non-targeted allele of population type U. Cas9 cleavage site indicated in arrowhead. (L) Schematic representation of the transition states for PSM excision. The transition is recapitulated in E and Figure 3A-B. Puromycin resistance gene (Puro). (M) Optimization of conditions for transposase mediated excision. Representative histograms for excision of the PSM using wild type (WT) and excision-only (EO) transposase variants, and one to three transfection steps. (N) Quantification of the excision efficiency shown in M. Each condition represents three replicates. For transposase optimization assays, the EGFP^{pos} populations type 2 + type U were used. Significance determined by a one-way ANOVA. Significance level * $p<0.05$.

Figure S2. Repetitive elements decrease on-target efficiency and increase random integration events. Related to Figure 1. Flow cytometry histogram for tagBFP: (A) *SNCA* exon 2 sgRNA 630 and 632, (B) *SNCA* exon 3 sgRNA 634 and 636, (C) *PINK1* exon 5 sgRNA 517 and 526, (D) *CLN3* exon 14-15 sgRNA 788, 789 and 909, (E) *CLN3* exon 5-8 sgRNA 781 and 783, and (F) *CLN3* exon 10-13 sgRNA 561 and 563. (G) Distribution

and type of repetitive elements in the homology arms of the dsDNA donors for *SNCA*e2, *SNCA*e3, *PINK1*e5, *CLN3*e5-6, *CLN3*e10-13 and *CLN3*e14-15. (H) Predictive model for random integration. The predictive model *PR* uses the matrix of repetitive element frequency in the homology arms *A*, the repetitive elements vector *x*, and the observed incidence of tagBFP^{pos} random integration *b*. The mathematical model generates coefficients for each repetitive element and the constant of the system for random integration prediction. (I) The space of non-zero coefficients derived from H: SINE Alu and SINE Mir, allows inferring expected random integration frequencies.

Figure S3. Quantification of polyclones composition post-knock-in and post-transposition. Related to Figure 2 and 3. (A) Schematic representation of the genomic structure after knock-in and genomic structure after transposition. PSM, left homology arm (LHA) and right homology arm (RHA). The binding sites of the genotyping primers are represented (Table S4), as well as the left homology arm junction, right homology arm junction, and WT junction. (B) Genotyping PCR products of 24 clones derived from the polyclone *SNCA*e2(p.A30P) 632, and WT control. (C) Genotyping PCR products of 24 clones derived from the polyclone *SNCA*e3(A53T) 636, and WT control. (D) Genotyping PCR products of 24 clones derived from the transposed polyclone *SNCA*e2(p.A30P) 632, pre-removal polyclone, and WT control. (E) Genotyping PCR products of 24 clones derived from the transposed polyclone *SNCA*e3(p.A53T) 636, pre-removal polyclone, and WT control. (F) Representation of the left homology arm junction of *SNCA*e2 including the SNP rs104893878 and PSM interface. Sanger sequencing chromatograms of 24 clones (not shown for space limitations) derived from the polyclone *SNCA*e2(p.A30P) 632 as in B. Chromatograms show the transversion *SNCA* c.88g>c and the TTAA interface to the PSM. Knock-in (KI). (G) Representation of the left homology arm junction of *SNCA*e3 including the SNP rs104893877 and PSM interface. Sanger sequencing chromatograms of 24 clones (not shown for space limitations) derived from the polyclone *SNCA*e3(p.A53T) 636 as in C. Chromatograms show the transition *SNCA* c.209g>a and the TTAA interface to the PSM. (H) Representation of the WT junction of *SNCA*e2 including the SNP rs104893878, and the TTAA interface to the genomic region. Sanger sequencing chromatograms of 24 clones (not shown for space limitations) derived from the transposed polyclone *SNCA*e2(p.A30P) 632. Chromatograms show the transversion *SNCA* c.88g>c and the TTAA interface to the genomic region. (I) Representation of the WT junction of *SNCA*e3 including the SNP rs104893877, and the TTAA interface to the genomic region. Sanger sequencing chromatograms of 24 clones (not shown for space limitations) derived from the transposed polyclone

*SNCA*e3(p.A53T) 636. Chromatograms show the transition *SNCA* c.209g>a and the TTAA interface to the genomic region.

Figure S4. Characterization of human iPS cells and NESCs: Microarray karyotype, pluripotency, differentiation and expression levels. Related to Figure 3 and Figure 4. (A) Microarray karyotype analysis of the parental line before electroporation, (B) polyclone 6321421 *SNCA* p.A30P and (C) polyclone 6361868 *SNCA* p.A53T. (D) Immunostaining for the pluripotency markers OCT4, TRA1-81, SOX2 and SSEA4 for parental control, (E) polyclone 6321421 *SNCA* p.A30P and (F) polyclone 6361868 *SNCA* p.A53T. Scale bar 200µm. (G) Differentiation of human iPS cells to NESCs in 3D culture as shown in Figure 4A. Scale bar 500µm. (H) Immunostaining of NESCs for the neuroepithelial stem cell markers NESTIN and SOX2. Scale bar 50µm. (C) Relative expression of *SNCA* mRNA with respect to *TUBG1* and *GAPDH* transcripts in microarray expression analysis in Figure 4B. Independent samples of NESCs A13777 were used (n=3).

Supplemental Table 1. Biallelic targeting frequency

Polyclone sample	Frequency composed biallelic[a]	Frequency single channel biallelic[b]	Frequency total biallelic[c]
*SNCA*e3 636	0.032	0.179	0.390
*SNCA*e2 630	0.021	0.145	0.311
*SNCA*e2 632	0.022	0.148	0.319
*SNCA*e3 634	0.056	0.237	0.529
*SNCA*e3 636	0.012	0.110	0.231
*PINK1*e5 517	0.033	0.182	0.396
*PINK1*e5 526	0.042	0.205	0.452
Mean global			**0.375**

[a]Frequency composed biallelic is defined as the experimentally measured $EGFP^{pos}dTOMATO^{pos}$ population.
[b]Frequency of single channel biallelic represents separately the $EGFP^{pos}EGFP^{pos}$ and $dTOMATO^{pos}dTOMATO^{pos}$ population, calculated as $\sqrt{\text{frequency composed biallelic}}$.
[c]Frequency total biallelic correspond to frequency composed biallelic $+ 2 *$ frequency of single channel biallelic.

Supplemental Table 2. Genome Cas9 binding sites tested and design for PAM edited donor sequences.

sgRNA	Genomic sequence	Sequence in donor
628	gtaaaggaattcattagcca**tgg**	gtaaaggaattcattagcca**tgg**
629	ggactttcaaaggccaagga**ggg**	ggactttcaaaggccaagga**ggg**
630	gctgctgagaaaaccaaaca**ggg**	gctgctgagaaaaccaaaca**ggg**
631	agggtgttctctatgtaggt**agg**	agggtgttctctatgtaggt**agg**
632[a]	ggtgcttgttcatgagtgat**ggg**	ggtgcttgTTAAtgagtgat**g**C**g**
633	**gga**agaagatcaaaatcctatat	**gga**agaagatcaaaatcctatat
634	tgtaggctccaaaaccaagg**agg**	tgtaggctccaaaaccaagg**agg**
635	**ggt**aacacgaatataggtttcta	**ggt**aacacgaatataggtttcta
636[a]	**ggt**ttctactataaatttcatag	**g**C**t**ttctactataAATTtcatag
637	atacttgccaagaataatga**ggg**	atacttgccaagaataatga**ggg**

[a]Edited PAM sequence.

Supplemental Table 3. *SNCA* polyclones summary.

Polyclone	PAM shielded	sgRNA	FACS % non-random	FACS % composed biallelic knock-in	% correct genotype post-knock-in (n correct/total)	FACS % transposition	% correct genotype post-transposition (n correct/total)
*SNCA*e2(p.A30P) 632	YES	632	94.1	2.2	100 (24/24)	4.0	100 (24/24)
*SNCA*e3(p.A53T) 636	YES	636	56.8	1.2	100 (24/24)	1.1	100 (24/24)
*SNCA*e2(p.A30P) 630	NO	630	85.1	2.1	Not determined	3.3	Not determined
*SNCA*e3(p.A53T) 634	NO	634	34.2	5.6	Not determined	3.2	Not determined

Supplemental Table 4. Oligonucleotides used in this study.

Primer	Sequence (5' to 3')	Region (Purpose)
SNCAe2_F1 (no1615)	gaggagtcggagttgtggagaag	SNCAe2 (Genotyping)
SNCAe2_R1 (no1616)	ttcccccactgatctatgttgaagag	SNCAe2 (Genotyping)
SNCAe3_F1 (no1617)	actgaaaaatccaacattagagagg	SNCAe3 (Genotyping)
SNCAe3_R1 (no1036)	ccagaacttgccacatgctt	SNCAe3 (Genotyping)
ITR_R1 (no861)	agatgtcctaaatgcacagcg	ITR (Genotyping)
ITR_F1 (no1310)	cgtcaattttacgcatgattatctttaac	ITR (Genotyping)
SNCAe2 (no1065)	tccgtggttaggtggctaga	SNCAe2 (Sequencing)
SNCAe3 (no1034)	gggccccggtgttatctcat	SNCAe3 (Sequencing)
T7-transposase_F (no1673)	gaaattaatacgactcactataggg ccgccaccatgggcagcagcctggac	transposase CDS (T7 fusion IVT)
Transposase_R (no1693)	ggcaaacaacagatggctgg	transposase CDS (IVT)
SNCAe2_628F	caccggtaaaggaattcattagcca	synthetic (sgRNA cloning)
SNCAe2_629F	caccgggactttcaaaggccaagga	synthetic (sgRNA cloning)
SNCAe2_630F	caccggctgctgagaaaaccaaaca	synthetic (sgRNA cloning)
SNCAe2_631F	caccgagggtgttctctatgtaggt	synthetic (sgRNA cloning)
SNCAe2_632F	caccgggtgcttgttcatgagtgat	synthetic (sgRNA cloning)
SNCAe3_633F	caccgtatatcctaaaactagaaga	synthetic (sgRNA cloning)
SNCAe3_634F	caccgtgtaggctccaaaaccaagg	synthetic (sgRNA cloning)
SNCAe3_635F	caccgatctttggatataagcacaa	synthetic (sgRNA cloning)
SNCAe3_636F	caccggatactttaaatatcatctt	synthetic (sgRNA cloning)
SNCAe3_637F	caccgatacttgccaagaataatga	synthetic (sgRNA cloning)
SNCAe2_628R	aaactggctaatgaattcctttacc	synthetic (sgRNA cloning)
SNCAe2_629R	aaactccttggcctttgaaagtccc	synthetic (sgRNA cloning)
SNCAe2_630R	aaactgtttggttttctcagcagcc	synthetic (sgRNA cloning)
SNCAe2_631R	aaacacctacatagagaacaccctc	synthetic (sgRNA cloning)
SNCAe2_632R	aaacatcactcatgaacaagcaccc	synthetic (sgRNA cloning)
SNCAe3_633R	aaactcttctagttttaggatatac	synthetic (sgRNA cloning)
SNCAe3_634R	aaacccttggttttggagcctacac	synthetic (sgRNA cloning)
SNCAe3_635R	aaacttgtgcttatatccaaagatc	synthetic (sgRNA cloning)
SNCAe3_636R	aaacaagatgatatttaaagtatcc	synthetic (sgRNA cloning)
SNCAe3_637R	aaactcattattcttggcaagtatc	synthetic (sgRNA cloning)
U6_F	gagggcctatttcccatgattcc	U6 (sequencing)

Supplemental references

Doench, J.G., Hartenian, E., Graham, D.B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B.L., Xavier, R.J., and Root, D.E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. Nat Biotechnol 32, 1262-1267.

Gibson, D.G. (2011). Enzymatic assembly of overlapping DNA fragments. Methods Enzymol 498, 349-361.

Li, X., Burnight, E.R., Cooney, A.L., Malani, N., Brady, T., Sander, J.D., Staber, J., Wheelan, S.J., Joung, J.K., McCray, P.B., Jr., et al. (2013). piggyBac transposase tools for genome engineering. Proc Natl Acad Sci U S A 110, E2279-2287.

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

Silva, L.P., Lorenzi, P.L., Purwaha, P., Yong, V., Hawke, D.H., and Weinstein, J.N. (2013). Measurement of DNA concentration as a normalization strategy for metabolomic data from adherent cell lines. Anal Chem 85, 9536-9542.

Yusa, K., Zhou, L., Li, M.A., Bradley, A., and Craig, N.L. (2011). A hyperactive piggyBac transposase for mammalian applications. Proc Natl Acad Sci U S A 108, 1531-1536.

Figure S1

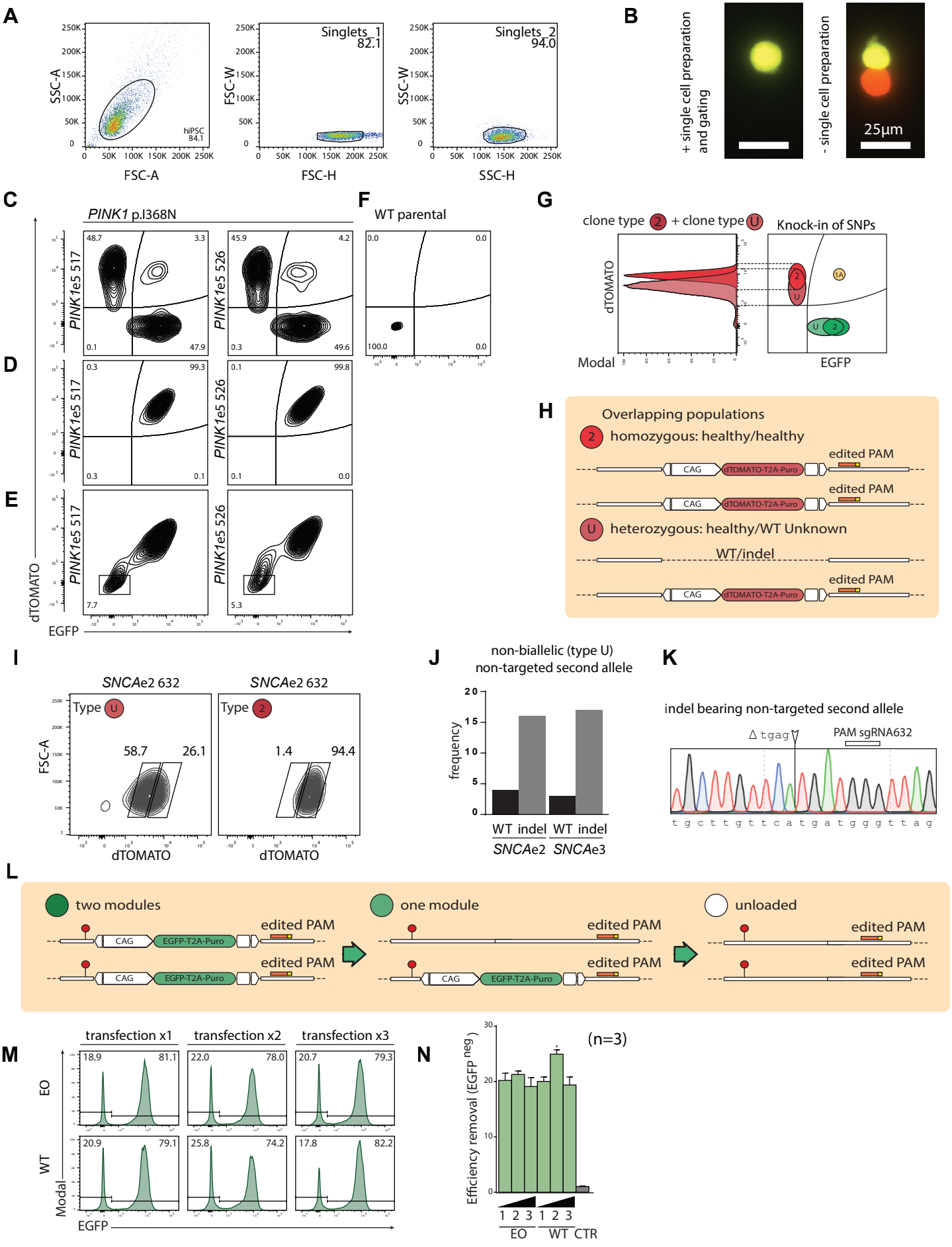
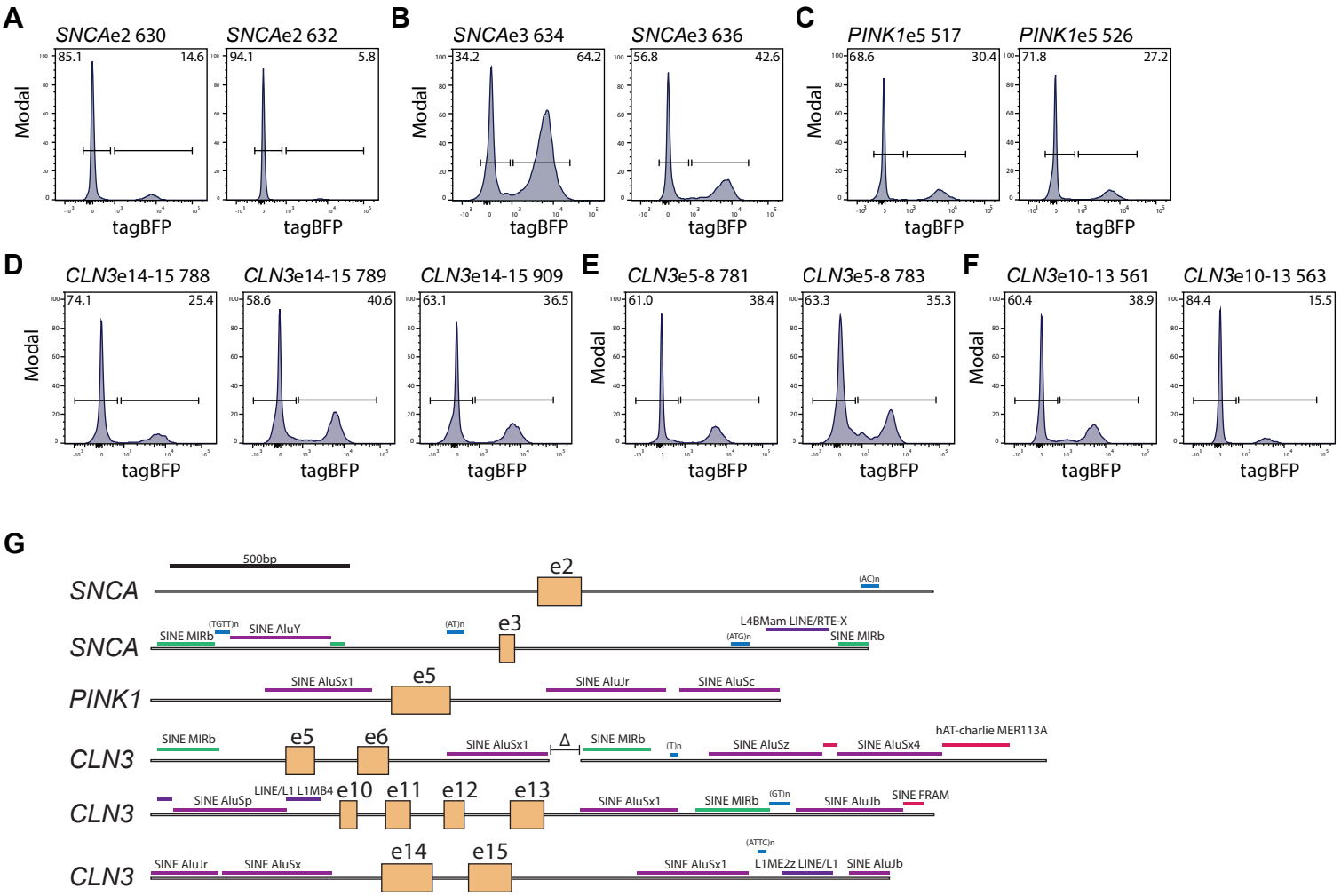

A
SSC-A
FSC-A
hiPSC
84.1
Singlets_1
82.1
FSC-W
FSC-H
Singlets_2
94.0
SSC-W
SSC-H
B
+ single cell preparation and gating
- single cell preparation
25µm
C
PINK1 p.I368N
PINK1e5 517
PINK1e5 526
48.7
3.3
0.1
47.9
45.9
4.2
0.3
49.6
D
0.3
99.3
0.3
0.1
0.1
99.8
0.1
0.0
E
7.7
5.3
dTOMATO
EGFP
F
WT parental
0.0
0.0
100.0
0.0
G
clone type 2 + clone type U
Knock-in of SNPs
dTOMATO
Modal
EGFP
1A
2
U
H
Overlapping populations
2 homozygous: healthy/healthy
edited PAM
CAG
dTOMATO-T2A-Puro
U heterozygous: healthy/WT Unknown
WT/indel
I
SNCAe2 632
Type U
Type 2
58.7
26.1
1.4
94.4
FSC-A
dTOMATO
J
non-biallelic (type U)
non-targeted second allele
frequency
WT indel WT indel
SNCAe2 SNCAe3
K
indel bearing non-targeted second allele
Δ tgag
PAM sgRNA632
t g c t t g t t c a t g a t g g g t t a g
L
two modules
one module
unloaded
edited PAM
CAG
EGFP-T2A-Puro
M
transfection x1
transfection x2
transfection x3
EO
WT
Modal
EGFP
18.9
81.1
22.0
78.0
20.7
79.3
20.9
79.1
25.8
74.2
17.8
82.2
N
(n=3)
Efficiency removal (EGFP neg)
1 2 3 1 2 3
EO WT CTR

Figure S2



H

Ax=b

frequency matrix

Loci	SINE Alu	LINE L4	SINE Mir	short repeats	Line L1	hAT charlie	υ	= measured random integration (tagBFP pos)
SNCAe3A	1	1	2	2	0	0	0	64.2
SNCAe3B	1	1	2	2	0	0	0	42.6
SNCAe2A	0	0	0	1	0	0	0	14.6
SNCAe2B	0	0	0	1	0	0	0	5.7
PINK1e5A	3	0	0	0	0	0	0	30.4
PINK1e5B	3	0	0	0	0	0	0	27.2
CLN3e14-15A	3	0	0	0	1	0	0	36.5
CLN3e14-15B	3	0	0	0	1	0	0	25.4
CLN3e14-15C	3	0	0	0	1	0	0	40.6
CLN3e10-13A	3	0	1	1	1	0	0	38.9
CLN3e10-13B	3	0	1	1	1	0	0	15.5
CLN3e5-8delA	3	0	2	0	0	1	0	38.4
CLN3e5-8delB	3	0	2	0	0	1	0	35.3
maximal-v	8.5	14.4	14.5	25	20.3	11.4	25	100.0

Predicted random

$PR = \alpha(RE1) + \beta(RE2) + \gamma(RE3) + \bullet\bullet\bullet + \delta(REn) + C$

$PR = 10.0978(\text{SINE Alu}) + (0)(\text{LINE L4}) + (1.746)(\text{SINE Mir}) + (0)(\text{short repeats}) + (0)(\text{Line L1}) + (0)(\text{hAT charlie}) + (0)(\upsilon) + C$

$PR = 10.0978(\text{SINE Alu}) + (1.746)(\text{SINE Mir}) + C \quad \text{with} \quad C=10.15$

I

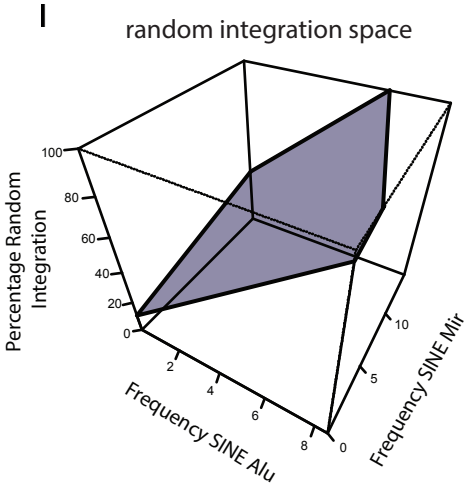


Figure S3

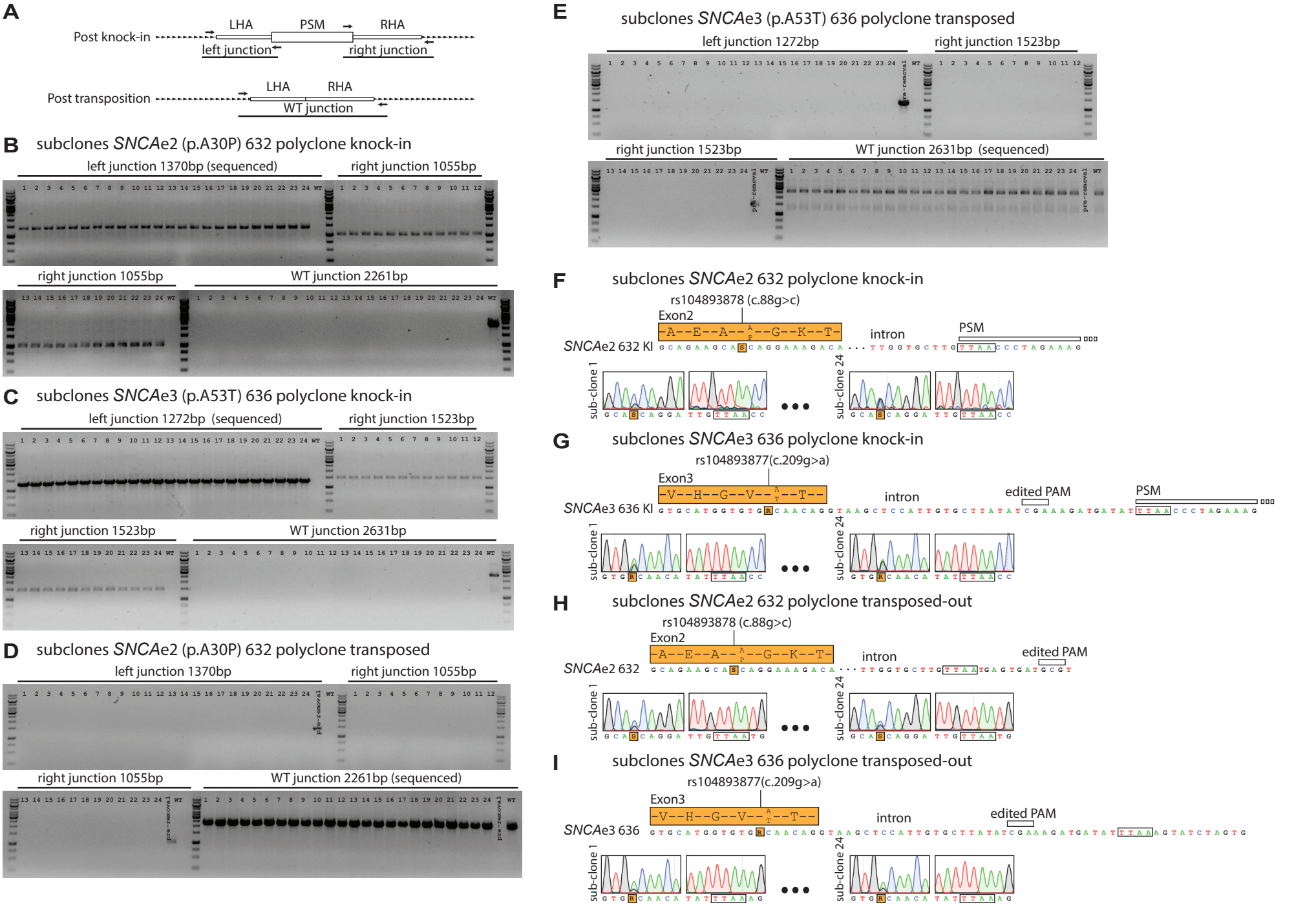
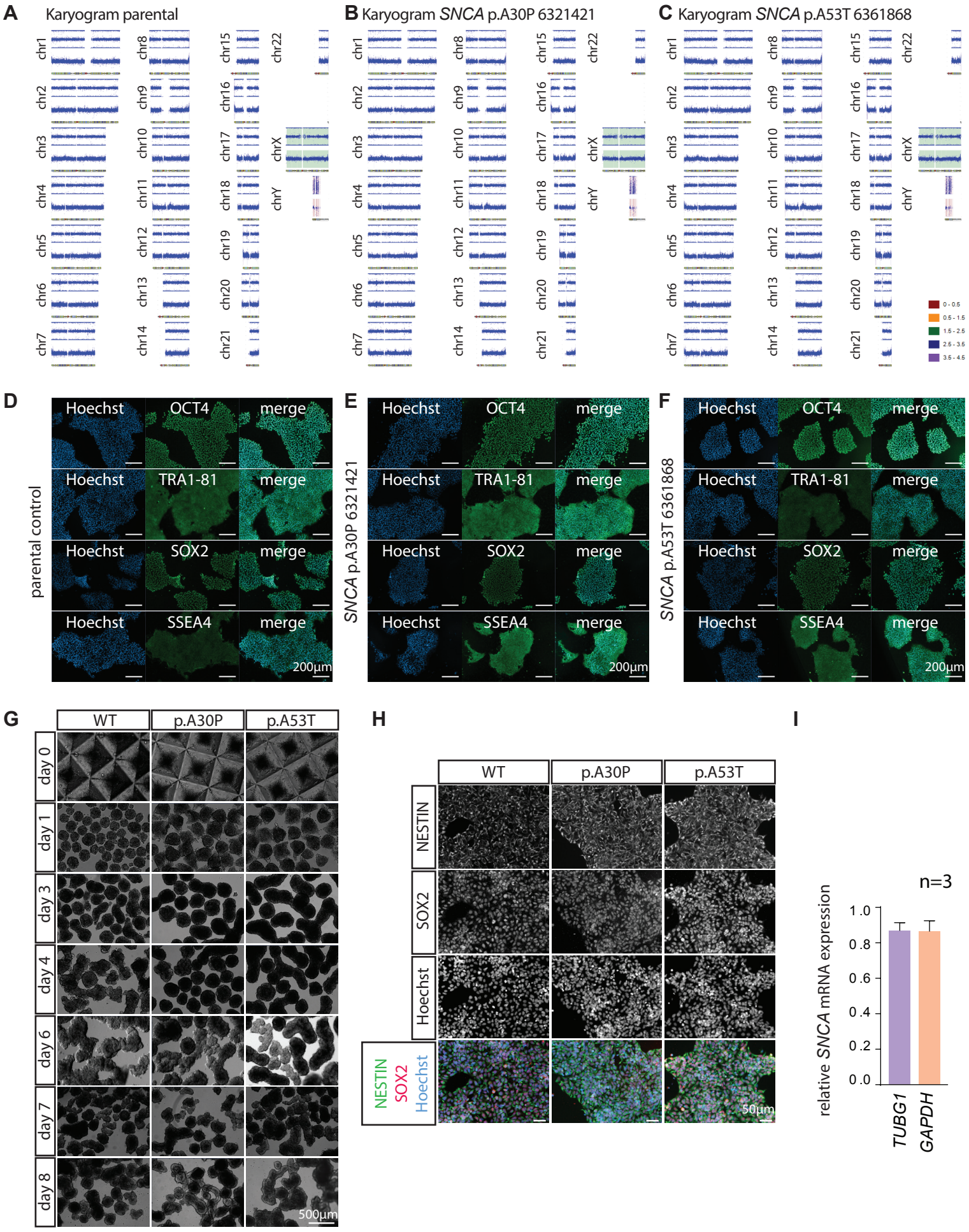

A
Post knock-in
LHA
PSM
RHA
left junction
right junction
Post transposition
LHA
RHA
WT junction
B
subclones SNCAe2 (p.A30P) 632 polyclone knock-in
left junction 1370bp (sequenced)
right junction 1055bp
right junction 1055bp
WT junction 2261bp
C
subclones SNCAe3 (p.A53T) 636 polyclone knock-in
left junction 1272bp (sequenced)
right junction 1523bp
right junction 1523bp
WT junction 2631bp
D
subclones SNCAe2 (p.A30P) 632 polyclone transposed
left junction 1370bp
right junction 1055bp
pre-removal
right junction 1055bp
WT junction 2261bp (sequenced)
pre-removal
pre-removal
E
subclones SNCAe3 (p.A53T) 636 polyclone transposed
left junction 1272bp
right junction 1523bp
pre-removal
right junction 1523bp
WT junction 2631bp (sequenced)
pre-removal
pre-removal
F
subclones SNCAe2 632 polyclone knock-in
rs104893878 (c.88g>c)
Exon2
-A--E--A--A/P--G--K--T-
intron
PSM
SNCAe2 632 KI
sub-clone 1
sub-clone 24
G
subclones SNCAe3 636 polyclone knock-in
rs104893877(c.209g>a)
Exon3
-V--H--G--V--A/T--T--
intron
edited PAM
PSM
SNCAe3 636 KI
sub-clone 1
sub-clone 24
H
subclones SNCAe2 632 polyclone transposed-out
rs104893878 (c.88g>c)
Exon2
-A--E--A--A/P--G--K--T-
intron
edited PAM
SNCAe2 632
sub-clone 1
sub-clone 24
I
subclones SNCAe3 636 polyclone transposed-out
rs104893877(c.209g>a)
Exon3
-V--H--G--V--A/T--T--
intron
edited PAM
SNCAe3 636
sub-clone 1
sub-clone 24

Figure S4


A
Karyogram parental
B Karyogram SNCA p.A30P 6321421
C Karyogram SNCA p.A53T 6361868
chr1
chr2
chr3
chr4
chr5
chr6
chr7
chr8
chr9
chr10
chr11
chr12
chr13
chr14
chr15
chr16
chr17
chr18
chr19
chr20
chr21
chr22
chrX
chrY
0 - 0.5
0.5 - 1.5
1.5 - 2.5
2.5 - 3.5
3.5 - 4.5
D
E
F
parental control
SNCA p.A30P 6321421
SNCA p.A53T 6361868
Hoechst
OCT4
merge
TRA1-81
SOX2
SSEA4
200µm
G
WT
p.A30P
p.A53T
day 0
day 1
day 3
day 4
day 6
day 7
day 8
500µm
H
NESTIN
SOX2
Hoechst
50µm
I
n=3
relative SNCA mRNA expression
1.0
0.8
0.6
0.4
0.2
0.0
TUBG1
GAPDH