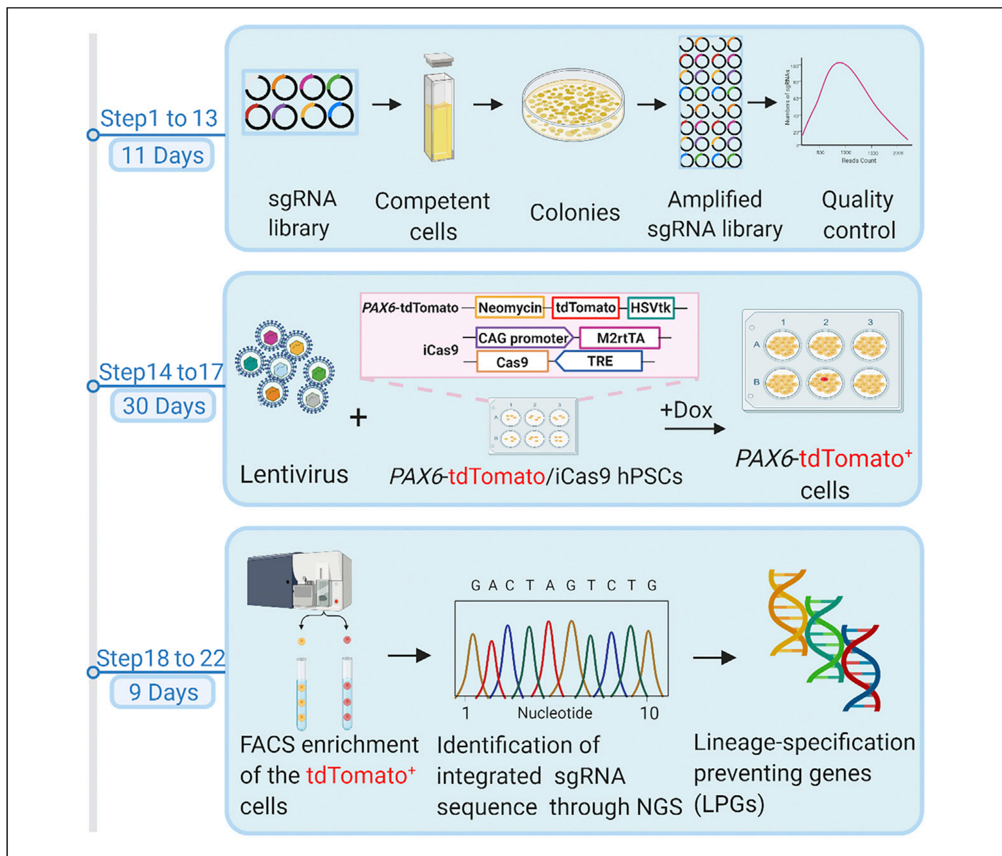
STAR Protocols

CellPress OPEN ACCESS

Protocol

Protocol for genome-scale CRISPR screening in engineered lineage reporter hPSCs to study cell fate determination



PAX6 is a key determinant of human neuroectoderm cell fate. Here, we describe a protocol for genome-scale CRISPR screening for use in genetically engineered human pluripotent stem cells (hPSCs). Using the germ layer reporter *PAX6* and an inducible CRISPR/Cas9 knockout system, we describe how to identify lineage-specific preventing genes. This protocol can be applied for use with other reporter genes to study cell fate determination in hPSCs.

Zhongshu Zhou, Lin Ma, Xiaoqing Zhang

linma@tongji.edu.cn (L.M.)
xqzhang@tongji.edu.cn (X.Z.)

Highlights

Generation of *PAX6* lineage reporter hPSCs

Detailed protocol for genome-scale CRISPR screening in hPSCs

Combining PAX6 hPSCs and CRISPR screening to study cell fate determination

Protocol allows identification of lineage specification preventing genes

Zhou et al., STAR Protocols 2, 100548
June 18, 2021 © 2021 The Authors.
https://doi.org/10.1016/j.xpro.2021.100548

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for genome-scale CRISPR screening in engineered lineage reporter hPSCs to study cell fate determination

Zhongshu Zhou,[1,3,7] Lin Ma,[1,3,5,6,7,8,9,*] and Xiaoqing Zhang[1,2,3,4,5,6,*]

[1]Translational Medical Center for Stem Cell Therapy, Shanghai East Hospital, Tongji University School of Medicine, Shanghai 200120, China

[2]Key Laboratory of Reconstruction and Regeneration of Spine and Spinal Cord Injury, Ministry of Education, Shanghai 200065, China

[3]Key Laboratory of Neuroregeneration of Shanghai Universities, Tongji University School of Medicine, Shanghai 200092, China

[4]Tsingtao Advanced Research Institute, Tongji University, Qingdao 266071, China

[5]Shanghai Institute of Stem Cell Research and Clinical Translation, Shanghai 200120, China

[6]Brain and Spinal Cord Clinical Research Center, Tongji University School of Medicine, Shanghai 200092, China

[7]These authors contributed equally

[8]Technical contact

[9]Lead contact

*Correspondence: linma@tongji.edu.cn (L.M.), xqzhang@tongji.edu.cn (X.Z.)
https://doi.org/10.1016/j.xpro.2021.100548

SUMMARY

PAX6 is a key determinant of human neuroectoderm cell fate. Here, we describe a protocol for genome-scale CRISPR screening for use in genetically engineered human pluripotent stem cells (hPSCs). Using the germ layer reporter _PAX6_ and an inducible CRISPR/Cas9 knockout system, we describe how to identify lineage-specific preventing genes. This protocol can be applied for use with other reporter genes to study cell fate determination in hPSCs.

For complete details on the use and execution of this protocol, please refer to Xu et al. (2021).

BEFORE YOU BEGIN

Preparation of sgRNA library and genetically engineered lineage reporter human pluripotent stem cells (hPSCs)

1. Choose a suitable sgRNA library for genome-scale CRISPR activation or CRISPR inhibition.

   Note: As a powerful tool for precise DNA modifications, CRISPR/Cas9 system has been widely applied in genome-scale screening. Several CRISPR/Cas9 libraries, including CRISPR activation- and CRISPR inhibition-ones (Joung et al., 2017), are commercially available. Here, we use the genome-scale CRISPR/Cas9 knockout (GeCKO) v2 library to screen genes that prevent neuroectoderm (nEc) or mesendoderm (ME) differentiation from hPSCs. To aim at a better outcome, we use one-vector format of GeCKO v2 library together with the inducible Cas9 (iCas9) system.

2. Construct the genetically engineered *PAX6*-tdTomato/iCas9 hPSCs (Figures 1A and 1B).
   a. Construct a donor plasmid harboring an in-frame coding cassette of Neo-P2A-tdTomato-T2A-HSVtk flanked by 5′- and 3′-homology arms of *PAX6*.
   b. Design a sgRNA targeting exon 4 of *PAX6* gene near the ATG start codon.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

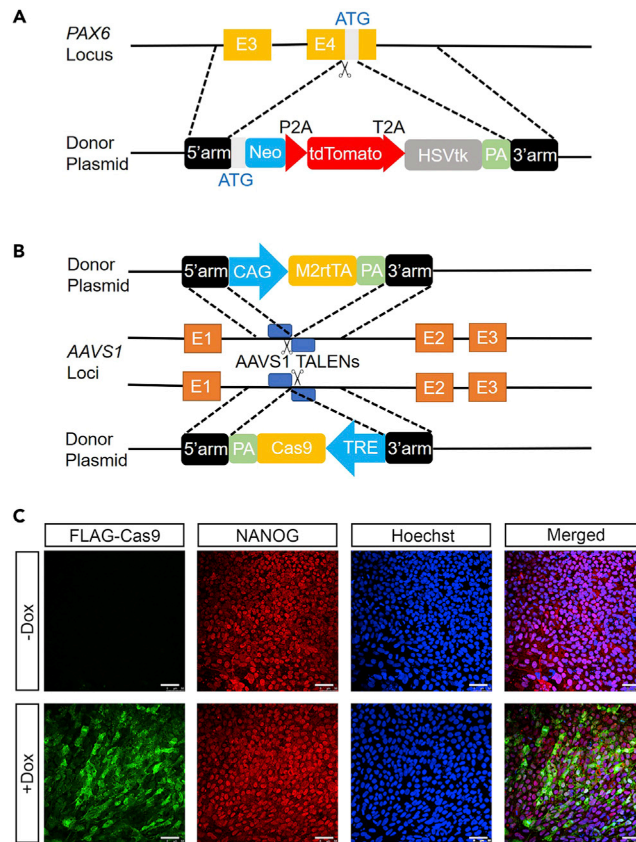


Figure 1. Construction of inducible Cas9 expression systems in *PAX6*-tdTomato reporter hPSCs

(A) Schematic view of the targeting strategy of *PAX6*-tdTomato reporter cell line through gRNA-guided CRISPR/Cas9 system. ATG, start codon; E, exon; PA, poly (A) signal; Neo, neomycin resistance gene.

(B) Schematic diagram for constructing inducible Cas9 (iCas9) expression cassette in *PAX6*-tdTomato line through TALEN (marked in blue bars)-mediated gene targeting at the *AAVS1* loci. CAG, CMV enhancer/chicken β-actin promoter; M2rtTA, reverse tetracycline transactivator; TRE, tetracycline response element, PA, poly (A) signal.

(C) Immunostaining results show that Dox treatment induces Cas9 expression, but not affects the pluripotent marker (NANOG) expression in *PAX6*-tdTomato/iCas9 hPSCs. Antibody dilutions for FLAG and NANOG are 1:1,000 and 1:500, respectively. Scale bar, 75 μm.

c. Co-transfect the donor plasmid, *PAX6* targeting sgRNA and Cas9-GFP plasmid through electroporation. After correct homologous recombination, the putative integrated hPSCs will express Neo-P2A-tdTomato-T2A-HSVtk protein under the tight control of endogenous *PAX6* gene expression regulatory machineries.
d. Construct the two-component iCas9 system (González et al., 2014) in *PAX6*-tdTomato reporter cell line. One donor plasmid contains a doxycycline (Dox)-inducible Cas9 expression cassette (3×TRE-FLAG-Cas9) and the other carries a constitutive reverse tetracycline transactivator (M2rtTA) expression cassette (CAG-M2rtTA). Both donor plasmids are flanked by 5′- and 3′-homology arms of *AAVS1* (also known as *PPP1R12C*) gene.
e. Electroporate both donor plasmids into the *PAX6*-tdTomato reporter cell line together with left and right TALEN plasmids targeting *AAVS1* loci. Lines with one allele homologously recombined with the 3×TRE-FLAG-Cas9 cassette and the other with the CAG-M2rtTA cassette is validated with genomic DNA PCR, sanger sequencing and Southern blot, and the expression pattern of Cas9 is validated via immunocytochemistry (Figure 1C).

Note: The differentiation of hPSCs into ectoderm, mesoderm and endoderm is the essential step for gastrulation and generating all functional cell types. Understanding the biological processes that determine three germ layer-entry of hPSCs is a central question in developmental and stem cell biology (Chi et al., 2017; Chi et al., 2016; Liu et al., 2019; Ma et al., 2019). Within the three germ layers, nEc specification is independent of activation of ectopic signaling (Muñoz-Sanjuán and Brivanlou, 2002). Considering the transcription factor *PAX6* is a human nEc cell fate determinant and can serve as a human nEc hallmark gene (Chen et al., 2018; Zhang et al., 2010), we engineered hPSCs with a *PAX6* reporter and the iCas9 system within the safe harbor gene *AAVS1*. Knockout of a gene in hPSCs, which leads to targeted nEc differentiation or spontaneous tri-lineage differentiation, will drive tdTomato expression in all or some of the transformed cells. For details of constructing the genetically engineered *PAX6*-tdTomato/iCas9 hPSCs, please refer to (Xu et al., 2021) and Figure 1.

Note: Given their self-renewal capacity and clonogenicity, the hPSCs are technically suitable for genetic engineering. While, the protocol described here can also be applied in other cell lines.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Goat polyclonal antibody to NANOG	R&D Systems	AF1997
Mouse monoclonal antibody to FLAG	Sigma	F3165
Hoechst 33258	Sigma	94403
Bacterial and virus strains		
DH5α Electroporation-Competent Cell	Weidi	DE1001
Chemicals, peptides, and recombinant proteins		
5×PrimeSTAR® Buffer	Takara	9158A
dNTPs (10 mM)	Takara	4019
PrimeSTAR-GXL DNA polymerase	Takara	R050A
Ampicillin	Sigma	A9518
SOC Medium	Sigma	S1797
Agarose	Amresco	9002-18-0
Puromycin	Sigma	540411
Dispase	Gibco	17105-041
Trypsin	Gibco	25200072
Doxycycline	MCE	HY-N0565B
1×phosphate buffer (PBS)	Corning	21-040-CV
FBS	ExCell Bio	FSP500
DMEM/F-12	Gibco	11330-032
DMEM (1×) for MEF growth medium	Gibco	11965-092
DMEM (1×) for HEK293FT medium	Corning	10-013-CVRC
Nonessential amino acids (NEAA)	Gibco	11140050
GlutaMAX	Gibco	35050061
KnockOut Serum Replacement	Gibco	A3181502
bFGF	PeproTech	100-18B
Glucose	Biodee	G8270
HEPES	Gibco	11344041
NaCl	SCR	10019318
Na_2HPO_4	SCR	7558-79-4
$CaCl_2$	Amresco	1631C052
NaOH	SCR	1310-73-2
Glycerol	Amresco	56-81-5

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
KCl	Sigma	V900068
Critical commercial assays		
QIAGEN Plasmid Maxi Kit	QIAGEN	12163
Universal DNA Purification Kit	TIANGEN	DP214-03
QuickExtract DNA Extraction Solution 1.0	Epicentre	QE09050
Deposited data		
Raw data	Xu et al., 2021	GEO: GSE132309
Experimental models: cell lines		
hPSCs (H9 hESCs)	WiCell	CVCL_9773
HEK293FT cells	Invitrogen	R70007
Oligonucleotides		
Primer: sgRNA-PCR-F: AATGG ACTATCATATGCTTACCGTAA CTTGAAAGTATTTCG	This paper	N/A
Primer: sgRNA-PCR-R: CTTTAGT TTGTATGTCTGTTGCTATTATG TCTACTATTCTTTCC	This paper	N/A
Primer: sgRNApool-PCR-F: AATGA TACGGCGACCACCGAGATCTAC ACTCTTTCCCTACACGACGCTCT TCCGATCTTAAGTAGAGGCTTTATAT ATCTTGTGGAAAGGACGAAACACC	This paper	N/A
Primer: sgRNApool-PCR-R: CAAGC AGAAGACGGCATACGAGATTCGC CTTGGTGACTGGAGTTCAGACGTG TGCTCTTCCGATCTCCGACTCGGT GCCCTTTTTCAA	This paper	N/A
Recombinant DNA		
CRISPR-Cas9 knockout library	Addgene	1000000049
Cas9-GFP plasmid	Addgene	44719
3×TRE-FLAG-Cas9 plasmid	Addgene	60843
psPAX2 plasmid	Addgene	12260
pMD2.G (VSVG) plasmid	Addgene	12259
AAVS1-TALEN-L plasmid	Addgene	59025
AAVS1-TALEN-R plasmid	Addgene	59026
Software and algorithms		
MAGeCK	Li et al., 2014	https://sourceforge.net/p/mageck/wiki/Home/
Cutadapt (version 1.11)	Martin, 2011	https://cutadapt.readthedocs.io/en/stable/
Bowtie2 (version 2.3.4.3)	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
ImageJ	N/A	https://imagej.nih.gov/ij/
Other		
Gene Pulser Xcell system	Bio-Rad	N/A
BD FACSVerse™ System	BD Biosciences	N/A
NanoDrop UV spectrophotometer	Thermo Fisher Scientific	N/A
Illumina HiSeq/NovaSeq system	Illumina	N/A
37°C incubator	Eppendorf	Galaxy 170 R
20-mm Cuvette	Bio-Rad	165-2082
0.22 μm Vacuum driven filter	Corning	431097
0.45 μm Filter	Millipore	HVHP02500
5 mL Round Bottom Polystyrene Test Tube with Cell Strainer Snap Cap	BD	352235
Centrifuge tube	Beckman	344058
Falcon 5 mL Round Bottom Polystyrene Test Tube	Falcon	352054
10 mL Serological pipette	Corning	4488

MATERIALS AND EQUIPMENT

HEK293FT medium, store at 4°C for 3 months		
Reagent	Final concentration	Amount
DMEM (1×)	N/A	500 mL
Fetal bovine serum (FBS)	10%	55.5 mL
Total	N/A	555.5 mL

hPSC growth medium, store at 4°C for 3 months		
Reagent	Final concentration	Amount
DMEM/F-12	N/A	500 mL
KnockOut Serum Replacement	20%	127.4 mL
GlutaMAX	0.5%	3.2 mL
Nonessential amino acids (NEAA)	1%	6.4 mL
bFGF (add before use)	4 ng/mL	
Total	N/A	637 mL
Filter through a 0.22 μm vacuum driven filter.		

2×Transfection Buffer, store at 4°C for 3 months		
Reagent	Final concentration (for 2×transfection buffer)	Amount
HEPES	0.01 g/mL	5 *g*
NaCl	0.016 g/mL	8 *g*
Na_2HPO_4	0.0002164 g/mL	0.1082 *g*
KCl	0.00076 g/mL	0.38 *g*
Glucose	0.002 g/mL	1 *g*
ddH_2O	N/A	
Total	N/A	500 mL
pH 7.04 ~7.05, adjusted with 3 M NaOH. Filter through a 0.22 μm vacuum driven filter.		

MEF growth medium, store at 4°C for 3 months		
Reagent	Final concentration	Amount
DMEM (1×)	N/A	500 mL
Fetal bovine serum (FBS)	10%	56.2 mL
Nonessential amino acids (NEAA)	1%	5.6 mL
Total	N/A	561.8 mL

STEP-BY-STEP METHOD DETAILS

Amplification of targeting library

Timing: 3 days

1. Check the efficiency of competent cells.
   a. Take out the GeCKO v2 library of plasmid and competent cells stored at −80°C and thaw them on ice. Pre-cold the 20-mm electroporation cuvette on ice. Pre-warm the SOC medium without ampicillin and the SOC agar plate (150 mm × 25 mm dish) containing 0.2 mg/mL ampicillin at a 37°C water bath and 37°C incubator for 30 min before the electroporation, respectively.
   b. Add 1 ng of plasmid into 50 μL competent cells in a 1.5 mL Eppendorf tube. Mix gently by pipetting and avoid bubbles.

c. Quickly transfer the mixture into the pre-cold 20-mm electroporation cuvette and avoid bubbles.
d. Use kimwipers to dry the surface of the cuvette. Place the cuvette in the Gene Pulser Xcell System, and electroporate the plasmid into the competent cells under the conditions as follows: Capacitance = 25 μF, Resistance = 200 Ω, Voltage = 2.5 kV.
e. Take out the cuvette from the Gene Pulser Xcell System and put it on ice for 2 min.
f. Transfer 2 μL transformation product into 500 μL pre-warmed SOC medium without ampicillin in a 1.5 mL Eppendorf tube and mix well.
g. Shake the Eppendorf tube with the transformation product in a shaker incubator at 37°C with a speed of 180 r/min for 1 h for cell recovery.
h. Spread the mixture onto the pre-warmed SOC agar plate containing ampicillin.
i. Incubate the plate in a 37°C incubator for 30 min, and continue to incubate for 14 h with the plate up-side-down.
j. Count the number of colonies generated in the SOC agar plate by ImageJ software. The efficiency of the competent cells is defined as the number of counted colonies $\times$ 2.5×10^4 colony-forming units (CFU)/μg plasmid.

Note: We use commercially available competent cells, which generate 3.614×10^4 clones in 2 μL transformation product, and the efficiency of the competent cells is 9.035×10^8 CFU/μg plasmid.

2. Calculate the volume of competent cells and the number of SOC agar plates required for sgRNA library expansion.

Note: The A pool of GeCKO v2 library contains 65,383 sgRNAs. We expect each sgRNA will be amplified for 100 times in general, and thus the A pool library should yield 6.5×10^6 colonies after transformation. During the testing, 1 ng plasmids transformed into 50 μL high-efficiency competent cells produce 9.035×10^5 colonies, and expansion of the A pool library needs a total of 400 μL competent cells (8 vials) and 180 SOC agar plates. The method for calculation is also suitable for expansion of the B pool of GeCKO v2 library or other screening libraries.

3. Electroporate the targeting sgRNA library into competent cells.
   a. Electroporate 1 ng of the targeting sgRNA library into 50 μL of competent cells at exactly the same conditions set at step 1d. Put the transformation product into 10 mL pre-warmed SOC medium without ampicillin for cell recovery as step 1 g.
   b. Perform electroporation one by one if multiple electroporation steps are required. For each electroporation, spread the transformation mixture onto 22–23 SOC agar plates containing ampicillin.
4. Collect colonies.
   a. Add 5 mL SOC medium with ampicillin onto each SOC agar plate.
   b. Gently scrape the colonies by a cell spreader, and transfer the bacteria suspension into a sterilized 2 L beaker.
   c. Wash the plate with another 5 mL SOC medium containing ampicillin, and collect the bacteria suspension into the beaker.
5. Expansion of bacteria.
   a. Separate the bacteria suspension into 10 conical flasks (500 mL) evenly.
   b. Shake the conical flasks in a shaker incubator at 37°C at a speed of 220 r/min for 8 h.
6. Collect the bacteria and extract sgRNA library using QIAGEN Plasmid Maxi Kit according to user manual.

Note: The bacteria suspension after expansion could be preserved in 30% glycerol/SOC medium at −80°C, and the preserved bacteria could be used for future library expansion with no need of extra transformation.

7. Dissolve the extracted plasmid in sterilized ddH_2O. Quantify the plasmid concentration with a NanoDrop UV spectrophotometer. Appropriate concentration of DNA plasmid is around 1 μg/μL.

⏸ **Pause point:** The DNA plasmid can be stored at −80°C for several years.

Quality control of the sgRNA library

⏱ **Timing: 8 days**

Note: A two-step PCR method is used to amplify the sgRNA sequences and add the Illumina adaptor and barcode sequences for next-generation sequencing (NGS).

8. Amplify the sgRNA library by PCR according to the following PCR program.

Component	Amount per reaction (μL)
5×Reaction Buffer	10
10 mM dNTPs	1
sgRNA-PCR-F	1
sgRNA-PCR-R	1
sgRNA library	100 ng in 10 μL
PrimeSTAR-GXL polymerase	1
ddH_2O	26
Total	50

Run the PCR according to the following program.

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial denaturation	98°C	60 s	1
Denaturation	98°C	15 s	20
Annealing	55°C	15 s	
Extension	68°C	20 s	
Final extension	68°C	5 min	1
Hold	4°C	Forever	1

9. Subject obtained PCR products to 2% (wt/vol) agarose gel, and run the gel at 120 V for 25 min at 22°C–26°C.
10. Purify the target products at around 300 bp through Universal DNA Purification Kit according to user manual.
11. Add Illumina adaptor and barcode sequences by another round of PCR for the gel-recycled products according to the following PCR program.

Component	Amount per reaction (μL)
5×Reaction Buffer	10
10 mM dNTPs	1
sgRNApool-PCR-F	1
sgRNApool-PCR-R	1
PrimeSTAR-GXL polymerase	1
Recycled products	40 ng in 4 μL
ddH_2O	32
Total	50

Run the PCR according to the following program.

PCR cycling conditions			
Steps	Temperature	Time	Cycles
Initial denaturation	98°C	60 s	1
Denaturation	98°C	15 s	20
Annealing	55°C	15 s	
Extension	68°C	20 s	
Final extension	68°C	5 min	1
Hold	4°C	Forever	1

12. Subject obtained PCR products to 2% agarose gel, and run the gel at 120 V for 25 min at 22°C–26°C. Purify the target products at around 260 bp through Universal DNA Purification Kit according to user manual.

13. Sequence the purified DNA products on the Illumina HiSeq/NovaSeq system. Check amplified library quality by MAGeCK (Li et al., 2014).
    a. Install MAGeCK with Conda/bioconda.
    b. Download the FASTQ file from Addgene.
    c. Cut adapter from the FASTQ files and just remain 20 bp sgRNA reads.
    For example, the pattern of reads1 sequence for A_pool is as follows:
    TAAGTAGA (Rd1 SP) GGCTTTATATATCTTGTGGAAAGGACGAAACACCG
    AGCTAGCTAGCTAGCTAGCT (20 bp sgRNA)
    GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAG
    TGGCACCGAGTCGGAGATCGGAAGAGCA (sgRNA scaffold)
    Reads1 sequence for A_pool should be cut 43 bp from 5' end and 87 bp from 3' end.
    ```
    $ cutadapt -u 43 -o A_pool_R1_1.fq.gz A_pool_R1.fq.gz
    $ cutadapt -u -87 -o A_pool_R1_2.fq.gz A_pool_R1_1.fq.gz
    ```
    Use reads mapping algorithms (such as Bowtie2) to map reads.
    Build bowtie2 index.
    ```
    $ bowtie2-build A_pool.fa bowtie2_ind_A_pool
    ```
    Map reads with Bowtie2.
    ```
    $ bowtie2 -x bowtie2_ind A_pool -U A_pool_R1_2.fq.gz -norc | samtools view
    -bS - > A_pool_R1.bam
    ```
    d. Use MAGeCK to count sgRNA number with default parameters and only report the reads with unique alignment.
    ```
    $ mageck count -l A_pool_library.csv -sample-label "A_pool" -fastq
    A_pool_R1.bam
    ```
    e. A count file collecting all read counts (A_pool.count.txt) and a summary file (A_pool_summary.txt) would be build by MAGeCK. Obtain the normalized counts of each sgRNA in amplified library and original sgRNA library. Qualified library will have a coverage higher than 80% and ginidex lower than 0.2. Otherwise, re-amplify the library with a larger scale.

Packaging of sgRNA library harboring lentivirus

Timing: 9 days

14. Lentivirus production
    a. Passage HEK293FT cells when reached 90% confluency. Plate HEK293FT cells into 20 100 mm culture dishes.
    b. On the next day, when the cells become 70% confluency, warm up 2×Transfection Buffer, 2 M $CaCl_2$ and sterilized ddH_2O at a 37°C water bath for 30 min before transfection.
    c. Add 5 mL ddH_2O, 737.5 µL 2 M $CaCl_2$, 100 µL sgRNA library plasmid at 1 µg/µL, 37.5 µL PAX2 at 1 µg/µL, 25 µL VSVG at 1 µg/µL in sequential in a 15 mL conical tube and mix well by

pipetting. Add 5.9 mL 2× Transfection Buffer by dropwise and mix immediately. Incubate the transfection mixture at 22°C–26°C for 1 min when it become slight turbid toward the light.

d. Evenly disperse the transfection mixture onto 5 plates of HEK29FT cells by dropwise (2.36 mL/plate). Gently swirl the dishes to ensure uniform distribution of the transfection mixture. Place the transfected dishes back into the incubator.
e. Repeat steps 14c and 14d for 3 more times to complete the transfection of the rest 15 plates.
f. Check precipitates under the microscope at 1 h after transfection. Abundant tiny calcium-phosphate-DNA precipitates with Brownian movement predicts a successful transfection.
g. Continue to culture the cells in a 37°C incubator for 16 h.
h. Aspirate the old medium and replace with 7.5 mL fresh HEK293FT medium in every dish.
i. Continue to incubate the cells for 48 h. Collect the supernatant containing viral particles in a 50 mL conical tube, and centrifuge at 3,000 × *g* for 10 min to remove cell debris. Harvest the supernatant and filter through a 0.45 μm syringe filter.
j. Sterile the ultracentrifuge tubes and buckets by soaking in 75% ethanol for 10 min. Air dry them in the biosafety cabinet.
k. Fill the filtered supernatant into 4 centrifuge tubes, and place them into the buckets. After balancing, centrifuge at 55,000 × *g* for 3 h at 16°C.
l. When ultracentrifuge completes, aspirate the supernatant carefully, and collect the virus pellet with 2 mL hPSC growth medium for each centrifuge tube. Pool all virus suspension in a 15 mL conical tube and mix well by pipetting. Keep the virus suspension at 4°C for 8–12 h, aliquot (0.5 mL/vial) and store at −80°C before use.

Pause point: The virus can be stored at −80°C for several years.

15. Titration of the packaged lentivirus solution
    a. Test the minimal killing concentration of puromycin minimum kill concentration. At the second day after passaging of hPSCs and HEK293FT cells, treat the cells with a serial of concentrations of puromycin (0, 0.25, 0.34, 0.5, 1 and 5 μg/mL) for 3 days. 3 days after treatment, observe the killing efficiency of puromycin by counting the number of living cells. The minimal killing concentration of puromycin is 0.5 μg/mL for both hPSCs and HEK293FT cells (Figure 2A).
    b. Seed HEK293FT cells in a 6-well plate at a density of 50,000 cells/well.
    c. On the next day, thaw the stored virus solution on ice and add 0, 2.5, 5, 10 and 20 μL virus solution into 5 individual wells. Mark the 6th well as control.
    d. 24 h after virus infection, add puromycin (0.5 μg/mL) into the virus-infected HEK293FT cells. Leave the 6th control well untreated and do not add puromycin into this well.
    e. Continue to culture the cells in the incubator for 3 days and replace the puromycin-containing fresh HEK293FT medium daily. For the control well, replace fresh HEK293FT medium without puromycin daily.
    f. 3 days after puromycin treatment, observe the cells of each well under a microscope (Figure 2B). The control well usually reaches 70% confluency. In puromycin treated wells, all cells in the well infected with 0 μL virus solution will be killed, and gradient increases of survival in wells infected with 2.5, 5, 10 and 20 μL virus solution will be observed.
    g. Digest the cells of each well with trypsin (0.05%) and count the cell number.
    h. Virus titer (Infectious Units (IFU)/mL) = Number of survived cells × $\frac{50{,}000}{\text{Number of control cells}}$ / Volume of virus solution added (mL).

Number of survived cells: the number of remaining cells in each well infected with various volumes of virus solution and treated with 0.5 μg/mL puromycin for 3 days;

Number of control cells: the number of cells in control well without puromycin-treatment at day 4 after seeding.

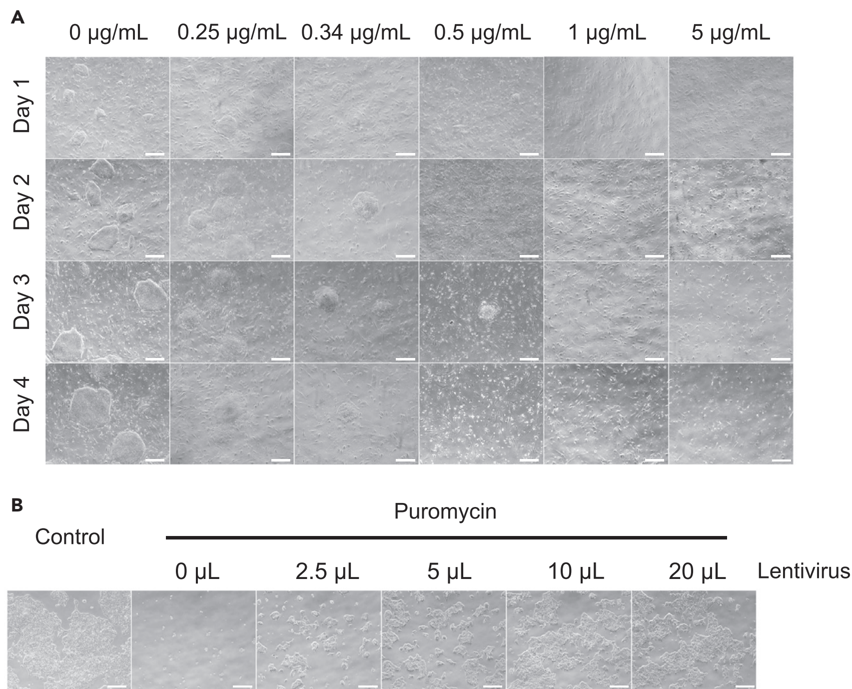


Figure 2. Titration of packaged lentivirus sgRNA library

(A) Titration the optimal concentration of puromycin for hPSCs. 0, 0.25, 0.34, 0.5, 1 and 5 µg/mL puromycin were applied in hPSCs for 3 days, and the killing efficiency was observed under the microscope. Scale bar, 100 µm.

(B) Phase contrast images show that HEK293FT cells in the control well without puromycin treatment reach 70% confluency 4 days after passaging. In puromycin treated wells, all cells in the well infected with 0 µL virus solution are killed, and the survival rates of cells infected with 2.5, 5, 10 and 20 µL virus solution show gradient increase. Scale bar, 100 µm.

Note: In our case, different dilutions of the virus end up with similar titration results, and the virus titers for GeCKO v2 library, including both A and B pools, are around 10^6 IFU/mL. If the relationship of the number of survived cells and the volume of virus solution added is not linear, only take account those wells with lower concentrations of virus treatment for calculation. Only 1 frozen-thaw cycle is acceptable, otherwise the IFU of the virus will be greatly reduced.

Note: hPSCs is not suitable for lentivirus titer test given their clonal growth behavior. Moreover, hPSCs could not survive after plating when they are digested into single cells. Here, we use HEK293FT cells to test virus titer instead of hPSCs.

Screening of lineage specification preventing genes (LPGs) in hPSCs

Timing: 30 days

16. Infect *PAX6*-tdTomato/iCas9 hPSCs with sgRNA lentiviral libraries (Figure 3A).
    a. Prepare *PAX6*-tdTomato/iCas9 hPSCs in a 6-well plate with 70% confluency.
    b. Warm up DMEM/F-12, dispase (1.5 U/mL), trypsin (0.05%) and hPSC growth medium in a 37°C water bath.
    c. Rinse one well of hPSCs with pre-warmed trypsin and digest the cells with trypsin for another 3 min in the incubator. Triturate the cell clusters into single cells with 1 mL pipette and count the total cell number of the well with a hemocytometer.
    d. Calculate the number of remaining 5 wells of cells.

CellPress
OPEN ACCESS

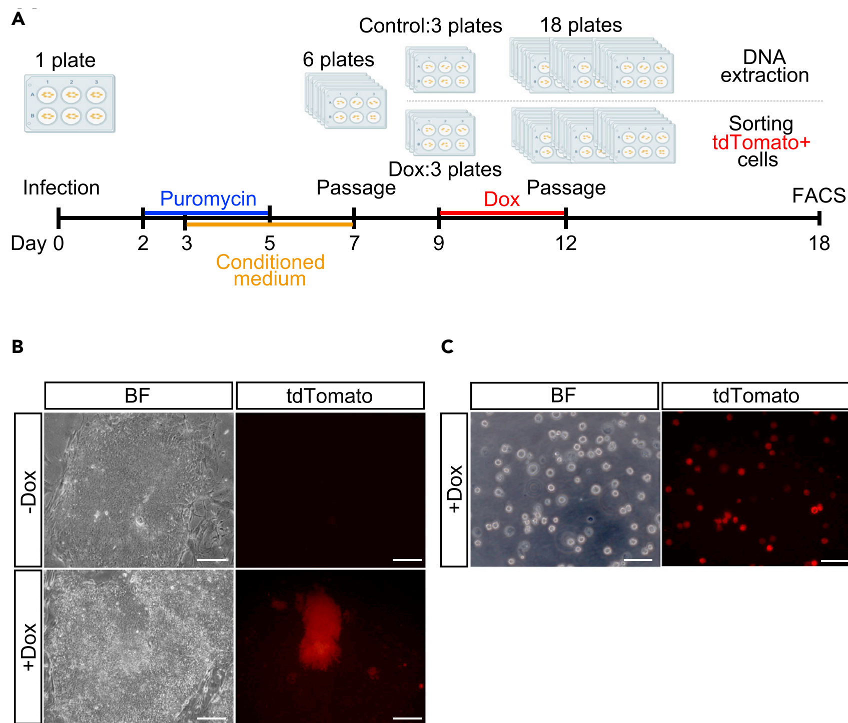


Figure 3. Genome-scale CRISPR screening in *PAX6*-tdTomato/iCas9 hPSCs

(A) Timeline of the screening procedure. 3.4 $\times 10^6$ hPSCs were infected with sgRNA lentiviral libraries of A pool or B pool and seeded in one MEF feeder-coated 6-well plate. 0.5 μg/mL puromycin was added from day 2 to day 5 to remove uninfected cells. MEF conditioned medium was supplied from day 3 to day 7. When reached 70% confluency at day 7, the library-infected hPSCs were passaged onto six 6-well plates with feeder layer. 2 days after passage, the hPSCs were treated with or without Dox (3 plates for each group) for 3 days. 5 days after passage, when the hPSCs were 70% confluent, both the control and Dox-treated cells were passaged onto eighteen 6-well plates with feeder layer. tdTomato positive cells organized into rosette-like structures were observable 6 days after the second passage, and FACS sorting and DNA extraction were performed. The whole procedure takes around 18 days.

(B) Some hPSC colones express clustered tdTomato after Dox treatment. Scale bar, 100 μm.

(C) Fluorescent images show that FACS-sorted cells have uniform tdTomato expression. Scale bar, 100 μm.

e. Digest the remaining *PAX6*-tdTomato/iCas9 hPSCs with dispase in a 37°C incubator for 3 min after brief rinse with DMEM/F-12. When the edges of hPSC colonies begin to curl, aspirate the dispase and rinse the cells with DMEM/F-12 gently.
f. Collect the hPSC colonies using a 10 mL serological pipette with 10 mL hPSC growth medium. Blow the colonies off gently and pool all the cells to a 50 mL conical tube. Pipette up and down against the bottom of the tube to break up the colonies into around 50 μm pieces.
g. Transfer the amount of cell clump suspension containing 3.4×10^6 cells into a 15 mL conical tube and centrifuge at 600 × *g* for 2 min at 22°C–26°C.
h. Aspirate the supernatant and resuspend the cell pellet in 1 mL lentivirus solution (10^6 IFU/mL) at multiplicity of infection (MOI) = 0.3 (Joung et al., 2017; Shalem et al., 2014).
i. Incubate the hPSC clumps/lentiviruses mixture in a 37°C incubator for 30 min, and gently tap the bottom of the 15 mL conical tube every 5 min. After incubation, seed the hPSC clumps/ lentiviruses mixture on feeder layer in a new 6-well plate.
j. Continue to culture the hPSCs in a 37°C incubator and replace refresh hPSC growth medium with 4 ng/mL bFGF every day.
k. 2 days after virus infection, add 0.5 μg/mL puromycin to kill uninfected cells for 3 days. Refresh the medium with puromycin daily.

l. When the library-infected *PAX6*-tdTomato/iCas9 hPSCs are 70% confluent, passage them onto 6 new 6-well plates with feeder layer.

Note: We usually maintain hPSCs on feeder layer of irradiated mouse embryonic fibroblasts (MEF) and in this way, hPSCs show no spontaneous differentiation. Ensure no occurrence of tdTomato positive cells under a fluorescence microscope every time before starting a screening experiment.

Note: If the MEF feeder layer is killed by puromycin, add MEF conditioned medium and fresh hPSC growth medium at 1:1 ratio until the cells are ready for passage. To prepare MEF conditioned medium, the MEF feeder layer is cultured with hPSC growth medium for 24 h. The conditioned medium is then collected.

Note: Before the Dox treatment, the *PAX6*-tdTomato/iCas9 hPSCs have been expanded by 25-fold for 9 days (6-fold for 5 days).
3.4×10^6 cells × 0.3 MOI × 25 = 2.55×10^7 cells
2.55×10^7 cells / 65, 383 gRNA = 390-fold representation
After Dox treatment, both the control and Dox-treated groups were further expanded for another 9 days before FACS sorting and DNA extraction, and the cells were pooled for 25 more folds at end point.
390 / 2 × 25 = 4875-fold representation
It is therefore both the control and Dox-treated cells have been expanded by 4875-fold before screening.

Note: During passage, to maintain library coverage, pool the library infected *PAX6*-tdTomato/iCas9 hPSCs together and do not discard any of the cells.

17. Induce Cas9 expression and initiate genome-scale gene knockout.
    a. At step 16l, 2 days after passaging, add 1 μg/mL Dox for 3 continuous days for 3 plates and leave the other 3 plates untreated as control.
    b. Passage and expand the control plates and Dox-treated plates onto 18 new 6-well plates, respectively.
18. FACS-enrichment of tdTomato expressing transformed cells.
    a. 6 days after passaging at step 17b, observe the tdTomato expressing pattern in control and Dox-treated cells under a fluorescence microscope. Typical tdTomato expressing cells organized into rosette like structures could be frequently seen in Dox-treated plates, but not in those control ones (Figure 3B).
    b. Digest one plate of control and one plate of Dox-treated library infected *PAX6*-tdTomato/iCas9 hPSCs into single cells with trypsin according to aforementioned method at step 16c.
    c. Pellet the single cells of both control and Dox-treated groups by centrifugation at 1,000 × *g* for 2 min.
    d. Remove the supernatant and resuspend cell pellets in 5 mL PBS containing 10% FBS.
    e. Pellet cells again at 1,000 × *g* for 2 min and resuspend them in 2 mL PBS.
    f. Pass the control and Dox-treated cell suspension through a 35 μm nylon mesh incorporated into the tube cap to remove clumps. Place them on ice.
    g. Record the percentage and collect all of the tdTomato positive cells in the Dox-treated group through FACS sorting. Also record the percentage of tdTomato positive cells in the control group by FACS, and re-confirm the percentage of tdTomato positive cells in the control group is always lower than that of the Dox-treated group.
    h. Repeat steps b-g to digest remaining plates of Dox-treated library infected *PAX6*-tdTomato/iCas9 hPSCs into single cells for further FACS sorting one plate by another.
    i. Pool all FACS enriched tdTomato cells together. Double check the tdTomato expressing cells under a fluorescence microscope (Figure 3C).

j. Pellet the tdTomato positive cells by centrifugation at a speed of 2,500 × *g* for 4 min at 22°C–26°C. Remove the supernatant and the cell pellet is ready for genomic DNA extraction. Otherwise, store at −20°C before moving to step 20.

Pause point: The collected cell pellet can be stored at −20°C for several days.

19. Digest the remaining 17 control plates into single cells with trypsin set at setp16c. Pool all the cells together and mix well by pipetting. Take 6×10^6 cells and pellet them through centrifugation at a speed of 2,500 × *g* for 4 min at 22°C–26°C. Remove the supernatant and the cell pellet is ready for genomic DNA extraction. Otherwise, store at −20°C before moving to step 20.

Pause point: The collected cell pellet can be stored at −20°C for several days.

20. Harvest the genomic DNA of the pelleted control cells with phenol/chloroform extraction, and dissolve the genomic DNA in 50 μL sterilized ddH_2O. Harvest the genomic DNA of the FACS enriched tdTomato cells by using QuickExtract DNA Extraction Solution according to the user manual, and dissolve the genomic DNA in 10 μL of sterilized ddH_2O.

Pause point: The genomic DNA can be stored at −20°C for several days.

21. Amplify the integrated sgRNAs by PCR in control and tdTomato cells, and add Illumina adaptor and barcode sequences according to aforementioned steps 8–12. For the first step of PCR amplification, add 10 μL genomic DNA templates for both the control (600 ng/μl) and tdTomato cells, although the exact amounts of DNA are variant in different conditions.

Note: In order to expand the integrated control library with ~100-fold representation, ~30 μg genomic DNA was extracted from 6×10^6 control cells and resolved in 50 μl sterilized ddH_2O (600 ng/μl). 6 μg DNA in 10 μl ddH_2O was used for PCR amplification for the integrated sgRNA library, and a total of 5 reactions were performed. All gel recycled PCR products from the 5 reactions were pooled, and 40 ng DNA products were used for the second-round PCR to add Illumina adaptor and barcode sequences.

22. Sequence the DNA products on the Illumina HiSeq/NovaSeq system, and use MAGeCK (Li et al., 2014) to identify LPGs.
    a. Cut adapter from FASTQ data and just remain 20 bp sgRNA reads.
    b. Use reads mapping algorithms (such as Bowtie2) to map reads.
    c. Use MAGeCK to count sgRNA number and compare counts of each sgRNA in different samples with default parameters and only the reads with unique alignment are reported.
    d. Get the normalized counts of each sgRNA in control sample and tdTomato cells, respectively.
    e. Estimate the statistical significance (using a negative binomial test) of enrichment for each sgRNA in the tdTomato cells compared to control cells by using the MAGeCK algorithm.
    f. Identify the LPGs by searching for genes whose sgRNAs are ranked consistently higher (by significance) using robust rank aggregation (RRA). Incorporate the negative control in MAGeCk analysis to generate null distributions and calculate the p-value and FDR for each.

Note: As a marker selection, it is also acceptable to directly compare the enriched sgRNA pools in tdTomato cells with the initial sgRNA pool obtained during the quality control step.

EXPECTED OUTCOMES

Using this genome-scale CRISPR knockout screening in *PAX6*-tdTomato/iCas9 hPSCs, we can unravel a group of LPGs, which are expressed in hPSCs and maintain their pluripotency via preventing

tri-lineage specification. According to these LPGs and their related functional modules, a full landscape of genetic wiring and biological processes that control hPSC self-renewal and tri-lineage specification can therefore be depicted.

LIMITATIONS

With the current protocol by using *PAX6*-tdTomato as a reporter, LPGs preventing nEc or tri-lineage specification are successfully unraveled. However, to elucidate LPGs specifically targeting mesoderm or endoderm specification, mesoderm or endoderm specific reporters are required. Combining of genome scale CRISPR screening and lineage reporter represents a powerful tool to study intrinsic mechanisms for governing a lineage fate. While, mechanisms related to manipulating environmental niches could not be precisely defined or even lead to misunderstanding.

TROUBLESHOOTING

Problem 1

At step 16. Spontaneous differentiation and tdTomato expressing cells always occur during normal hPSC maintenance or in the control cells without Dox treatment.

Potential solution 1

Maintaining hPSCs on MEF feeder layer will largely solve this problem. Scraping off the differentiated colonies or aberrant tdTomato expressing colonies during regular culture also helps. If a large fraction of differentiated cells frequently occur, use a new batch of cells or another hPSC cell line for the study.

Problem 2

At step 13. Amplified library has a low coverage ($\leq$80%) and high ginidex ($\geq$0.2).

Potential solution 2

We recommend scale up the competent cell volumes and transformation times. Choose one batch of competent cell with the highest transformation efficiency will also help.

Problem 3

At step 16. The screening has a high noise and non-related genes frequently appeared.

Potential solution 3

Infect the reporter cells with a lower MOI, or scale up the screening to reduce the false positives.

Problem 4

At step 2. A gene responsible for nEc commitment will fail to be screened using the KO libraries.

Potential solution 4

Activation library, such as SAM library, is recommended to screen such genes responsible for nEc commitment.

Problem 5

At steps 8–12 or 21. PCR bias may occur, which will lead to false-positive and false-negative results.

Potential solution 5

One-step PCR for both amplification and adding Illumina adaptor and barcode sequences by designing a new set of primers might minimize the PCR bias.

CellPress
OPEN ACCESS

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Lin Ma (linma@tongji.edu.cn).

Material availability

Plasmids and cell lines described in this study will be made available upon request.

Data and code availability

The data generated during this study are available at GEO: GSE132309.

ACKNOWLEDGMENTS

This work was supported by grants from the National Key Research and Development Program of China (grant no. 2018YFA0108000 and 2019YFA0110300), the National Natural Science Foundation of China (grant no. 82025020), the Science and Technology Commission of Shanghai Municipality (19JC1415100 and 21140902300), the Shanghai Municipal Education Commission (C120114), the Fundamental Research Funds for the Central Universities, and the Major Program of Development Fund for Shanghai Zhangjiang National Innovation Demonstration Zone (Stem Cell Strategic Biobank and Clinical Translation Platform of Stem Cell Technology, ZJ2018-ZD-004). Graphical abstract of this study was created with BioRender.com.

AUTHOR CONTRIBUTIONS

X.Z. supervised the study. Z.Z. performed both the experiments and analysis. X.Z., Z.Z., and L.M. wrote the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Chen, Z., Ren, X., Xu, X., Zhang, X., Hui, Y., Liu, Z., Shi, L., Fang, Y., Ma, L., Liu, Y., et al. (2018). Genetic engineering of human embryonic stem cells for precise cell fate tracing during human lineage development. Stem Cell Reports *11*, 1257–1271.

Chi, L., Fan, B., Feng, D., Chen, Z., Liu, Z., Hui, Y., Xu, X., Ma, L., Fang, Y., Zhang, Q., et al. (2017). The dorsoventral patterning of human forebrain follows an activation/transformation model. Cereb. Cortex *27*, 2941–2954.

Chi, L., Fan, B., Zhang, K., Du, Y., Liu, Z., Fang, Y., Chen, Z., Ren, X., Xu, X., Jiang, C., et al. (2016). Targeted differentiation of regional ventral neuroprogenitors and related neuronal subtypes from human pluripotent stem cells. Stem Cell Reports *7*, 941–954.

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

Joung, J., Konermann, S., Gootenberg, J.S., Abudayyeh, O.O., Platt, R.J., Brigham, M.D., Sanjana, N.E., and Zhang, F. (2017). Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat. Protoc. *12*, 828–863.

Langmead, B., and Salzberg, S., L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357–359. https://doi.org/10.1038/nmeth.1923.

Li, W., Xu, H., Xiao, T., Cong, L., Love, M.I., Zhang, F., Irizarry, R.A., Liu, J.S., Brown, M., and Liu, X.S. (2014). MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. Genome Biol. *15*, 554.

Liu, L., Chen, Z., Zhang, X., Li, S., Hui, Y., Feng, H., Du, Y., Jin, G., Zhou, X., and Zhang, X. (2019). Protection of ZIKV infection-induced neuropathy by abrogation of acute antiviral response in human neural progenitors. Cell Death Differ. *26*, 2607–2621.

Ma, L., Wang, Y., Hui, Y., Du, Y., Chen, Z., Feng, H., Zhang, S., Li, N., Song, J., Fang, Y., et al. (2019). WNT/NOTCH pathway is essential for the maintenance and expansion of human MGE progenitors. Stem Cell Reports *12*, 934–949.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet Journal *17*, 10–12. https://doi.org/10.14806/ej.17.1.200.

Muñoz-Sanjuán, I., and Brivanlou, A.H. (2002). Neural induction, the default model and embryonic stem cells. Nat. Rev.Neurosci. *3*, 271–280.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-Scale CRISPR-Cas9 Knockout Screening in Human Cells. Science *343*, 84–87. https://doi.org/10.1126/science.1247005.

Xu, X., Du, Y., Ma, L., Zhang, S., Shi, L., Chen, Z., Zhou, Z., Hui, Y., Liu, Y., Fang, Y., et al. (2021). Mapping germ-layer specification preventing genes in hPSCs via genome-scale CRISPR screening. iScience *24*, 101926.

Zhang, X., Huang, C.T., Chen, J., Pankratz, M.T., Xi, J., Li, J., Yang, Y., LaVaute, T.M., Li, X.-J., Ayala, M., et al. (2010). Pax6 Is a Human Neuroectoderm Cell Fate Determinant. Cell Stem Cell *7*, 90–100.