Stem Cell Reports, Volume *18*

Supplemental Information

Organoid models of breathing disorders reveal patterning defect of hindbrain neurons caused by PHOX2B-PARMs

Kathy Nga-Chu Lui, Zhixin Li, Frank Pui-Ling Lai, Sin-Ting Lau, and Elly Sau-Wai Ngan

Description of Supplementary Files

Supplementary Data 1: Markers for 11 clusters in control HBSOs.

Supplementary Data 2: Markers of the 16 clusters identified in control and PHOX2B-7Ala HCOs.

Supplementary Data 3: Markers of the 6 cell types identified in HCOs.

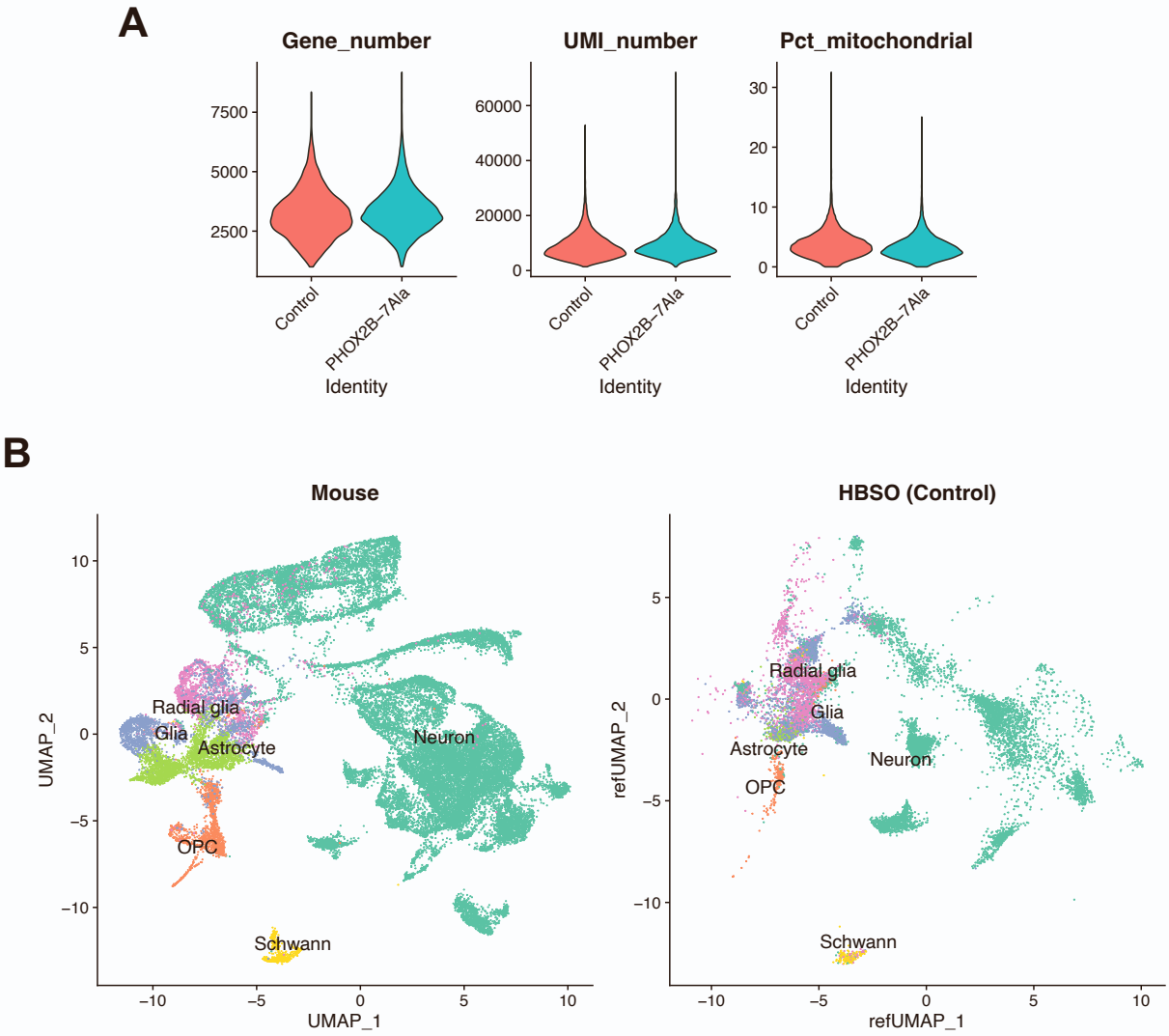
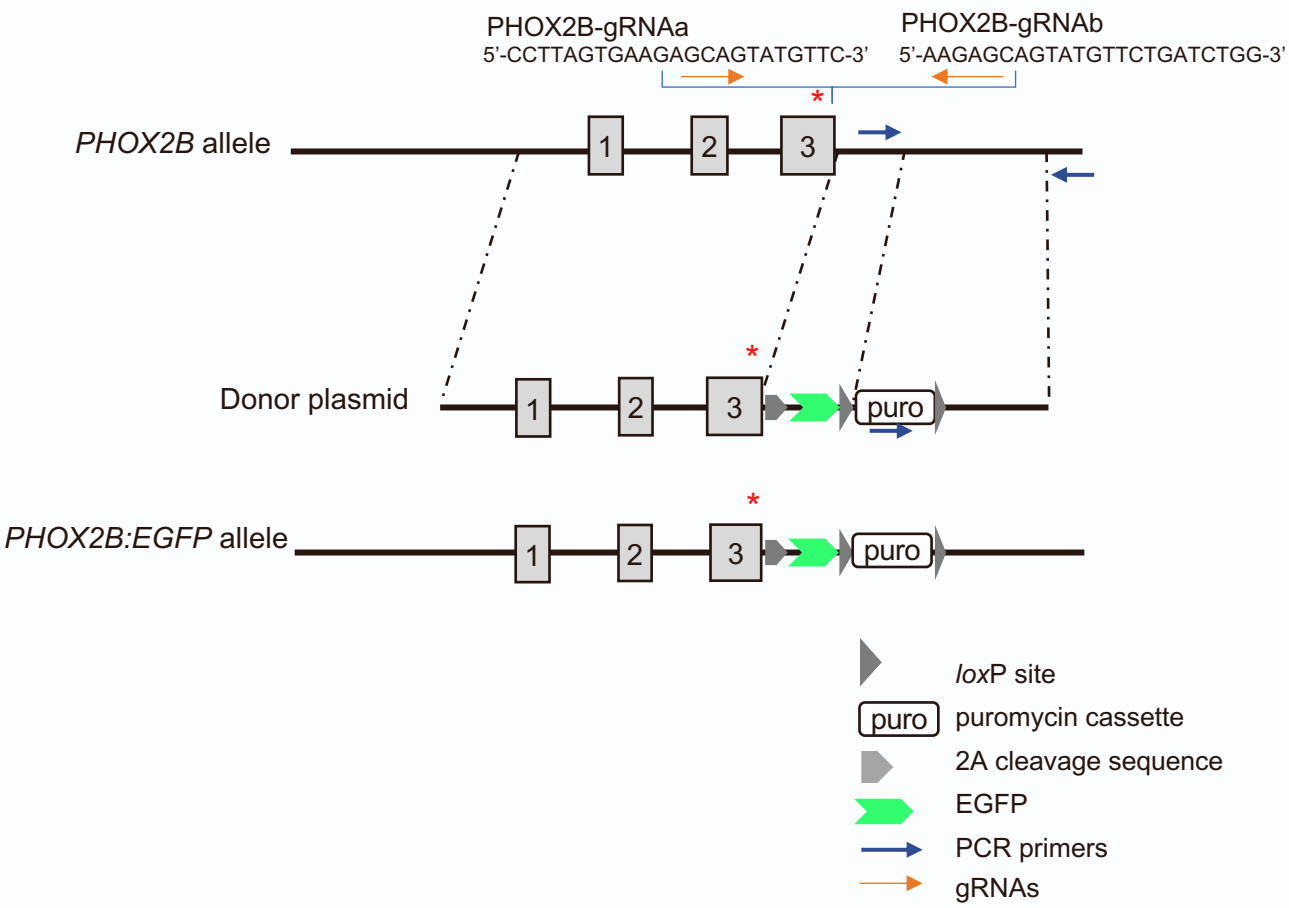


Fig. S1 Quality control and reference-based annotation of HBSO cells. (**A**) The number of detected gene and UMI number, and percentage of mitochondrial genes between control and PHOX2B-7Ala mutant HBSOs. (**B**) Distribution of the six main cell types of *in vivo* mouse hindbrain cells from E12-E16 (Manno *et al*., 2021) and our pooled HBSOs in UMAP plot (colored by cell types). Cell labels were transferred from *in vivo* mouse hindbrain to HBSOs by Seurat.



* 5 additional Ala repeats (GCA GCA GCC GCG GCA)
7 additional Ala repeats (GCG GCG GCC GCG GCA GCA GCG)
13 additional Ala repeats (GCA GCA GCC GCG GCA GCT GCT GCT GCA GCT GCA GCG GCA)

Fig. S2 Targeting strategy for introduction and correction of PARM-mutations in human *PHOX2B* allele. Additional Ala repeats were introduced to the 20-Ala polyalanine tract in exon 3 of *PHOX2B* in the control hPSC line to generate PHOX2B-PARM mutant lines by CRISPR-Cas9. The same site in the CCHS-HSCR patient hPSC line (i.e. 15C11) was corrected back to 20-Ala polyalanine tract to generate corrected hPSC line (i.e. 15C11-C).

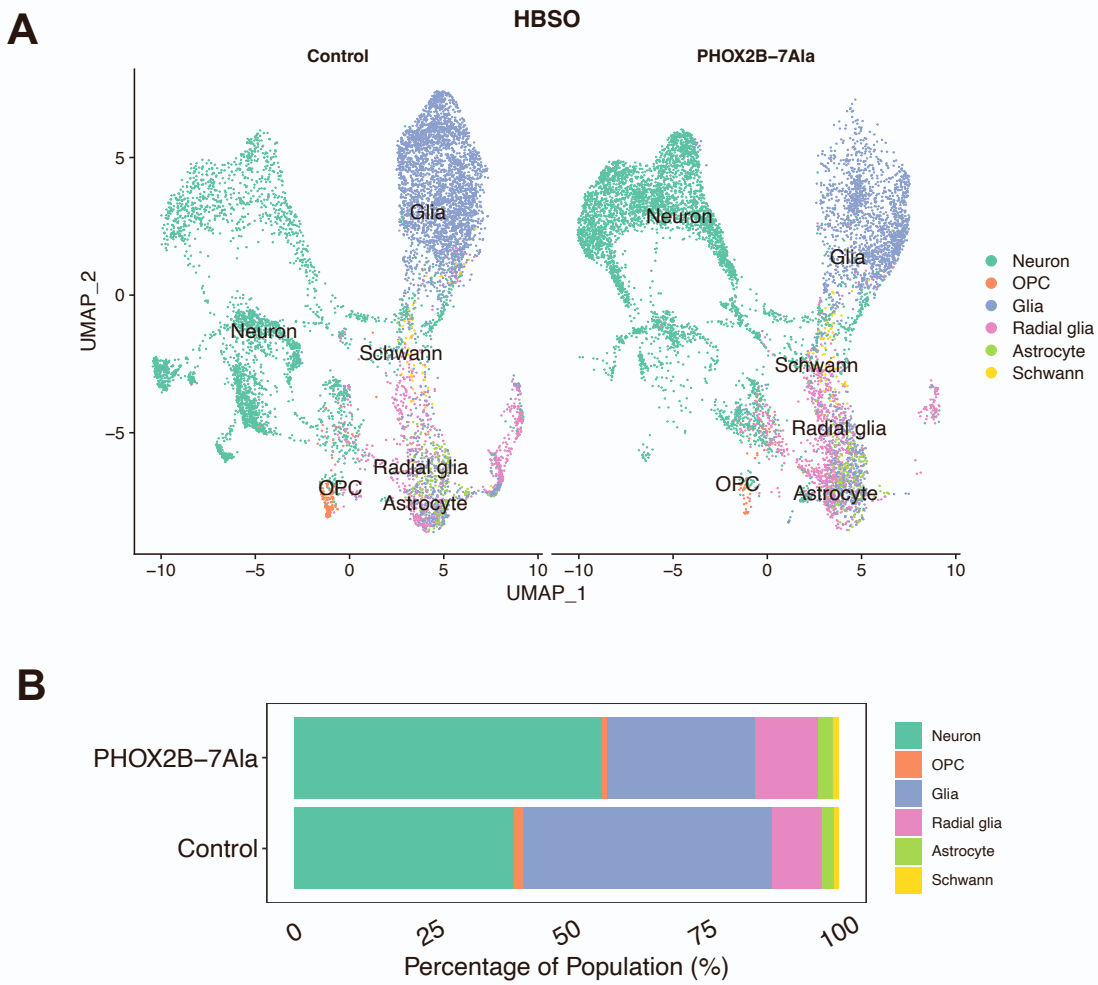


Fig. S3 Cell composition in the control and PHOX2B-7Ala mutant HBSOs as revealed by scRNA-seq. (**A**) Distribution of the six main cell types of our control and PHOX2B-7Ala mutant HBSOs in a new UMAP plot (colored by cell types). (**B**) Barplot showing the population size difference of the six main cell types between the control and PHOX2B-7Ala mutant. OPC, oligodendrocyte progenitor cell.

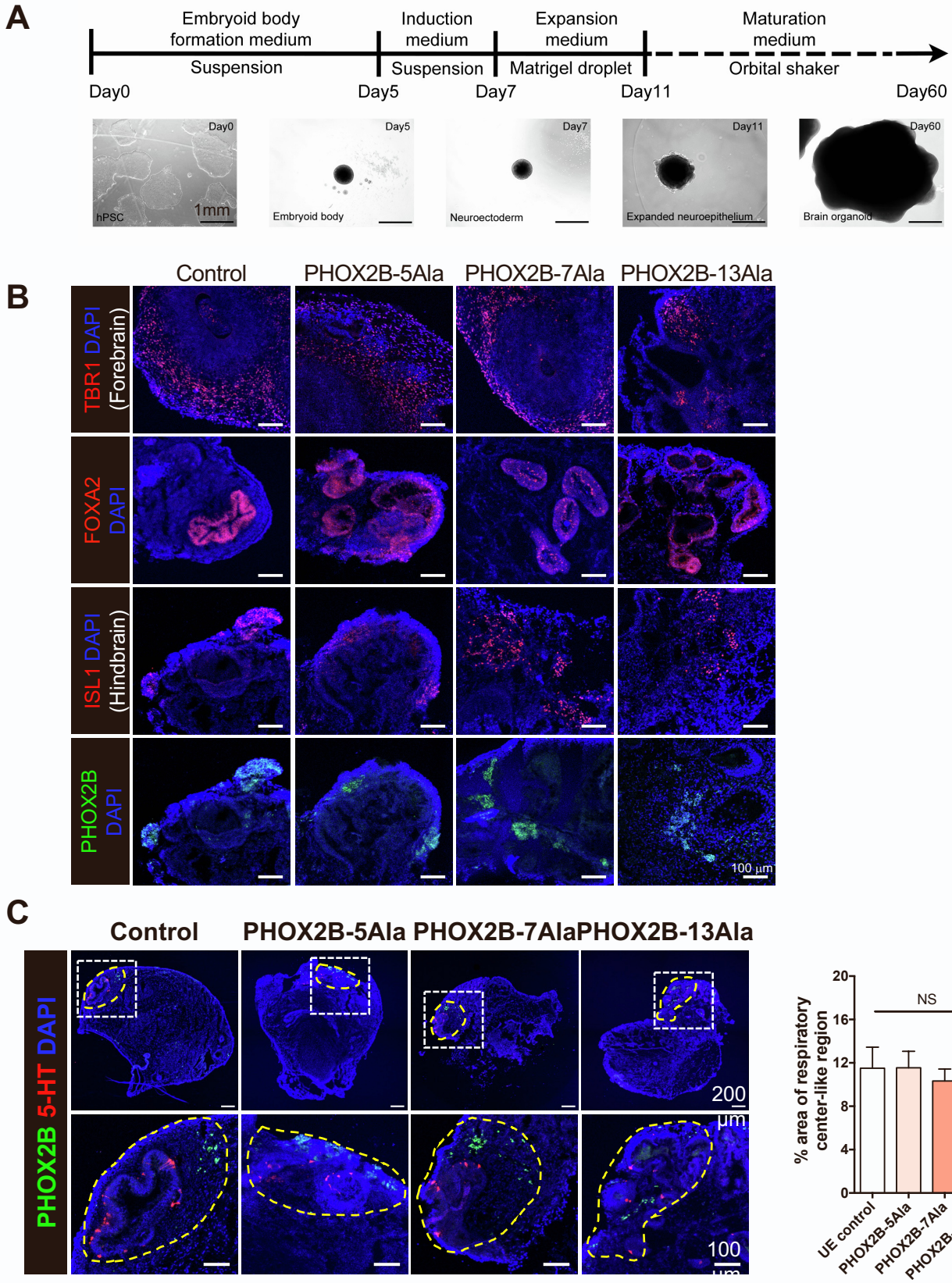


Fig. S4 Generation of unguided brain organoids (HCOs) from the control and PHOX2B-PARM mutant hPSCs. (**A**) Schematic shows the derivation of HCOs from hPSCs. Representative phase-contrast images of each stage are shown. (**B**) Immunostaining of PHOX2B and markers of forebrain (TBR1), midbrain (FOXA2) and hindbrain (ISL1) in the day-60 HCOs derived from the control and the PHOX2B-PARM mutants. (**C**) Representative images of day-60 HCOs derived from the control and PHOX2B-PARM mutants, immuno-stained with PHOX2B and 5-HT, representing the respiratory center-like regions (outlined by the yellow dashed lines). The inset in the bottom panel is an enlargement of the white squared region in the image. Bar chart showing the percentage area of respiratory center-like region in the control and PHOX2B-PARM mutant HCOs. One-way ANOVA. NS, non-significant.

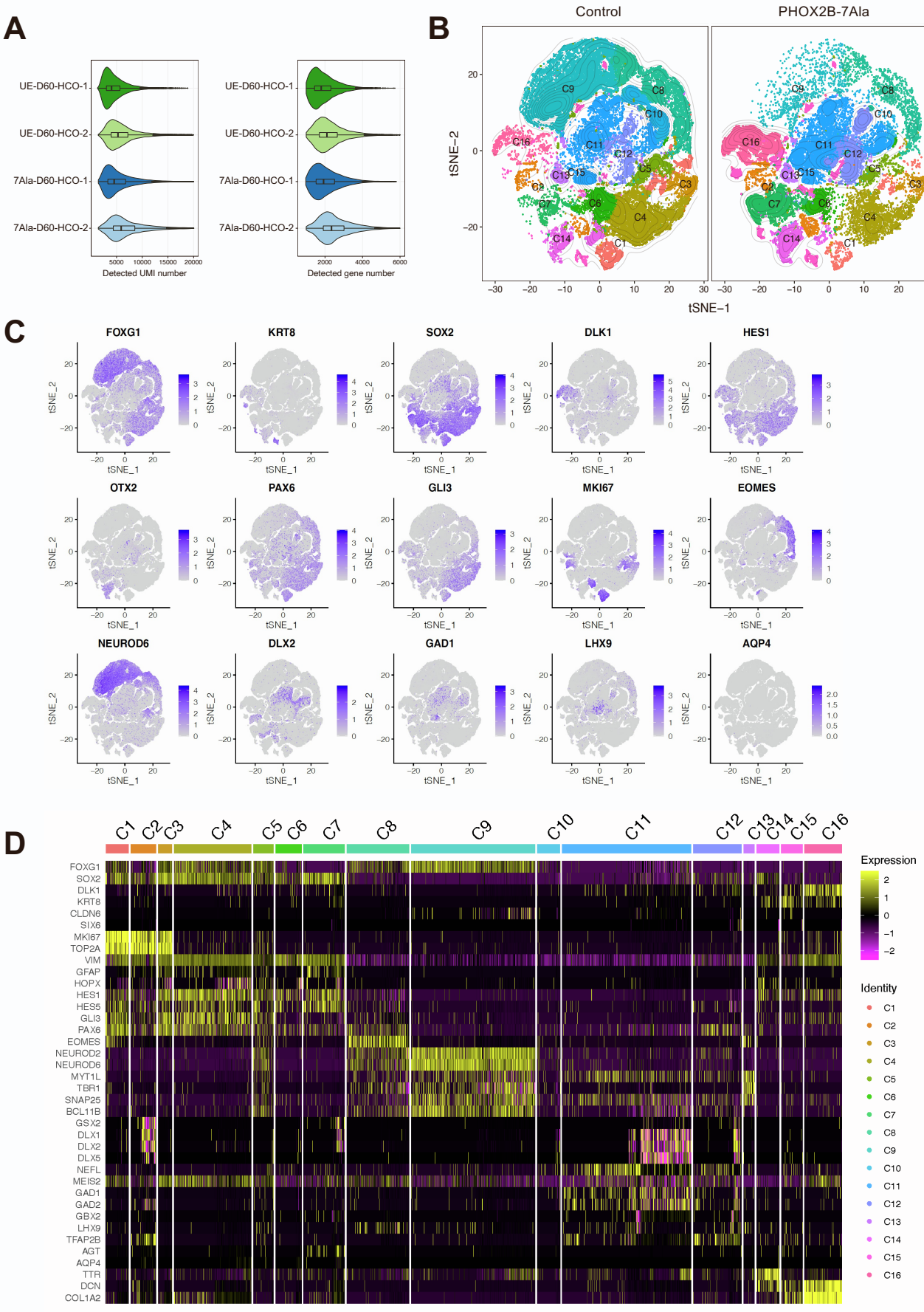


Fig. S5 Quality control and overall characteristics of the scRNA-seq dataset. (**A**) The number of detected UMI and gene numbers in four HCO samples. (**B**) Distribution of the 16 clusters from the control and PHOX2B-7Ala mutant HCOs in *t*-SNE plot (colored by clusters). (**C**) Canonical markers expressed in 16 clusters in the integrated HCO cells, colored by expression level. (**D**) Heatmap showed the unique markers of each cluster.

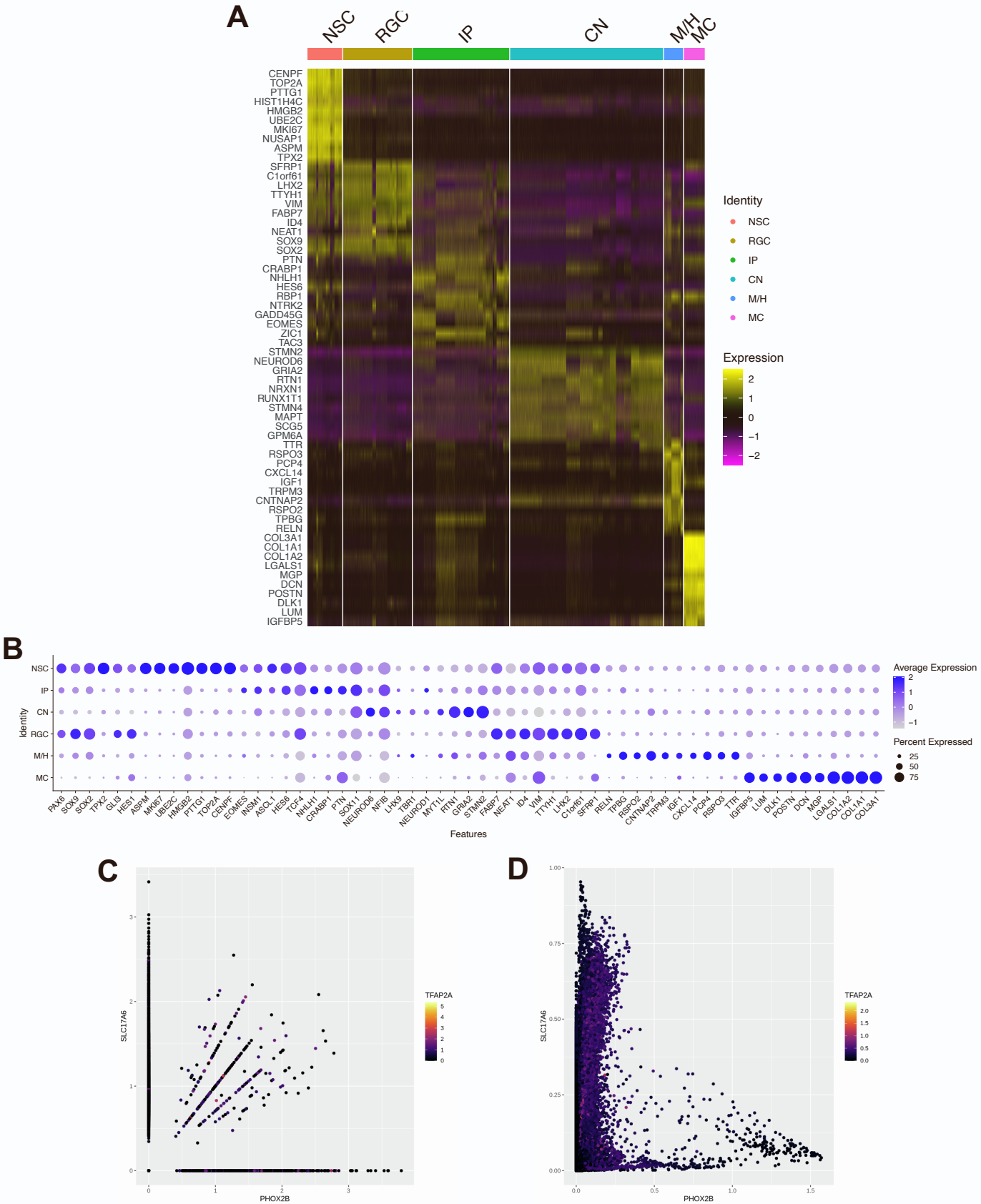
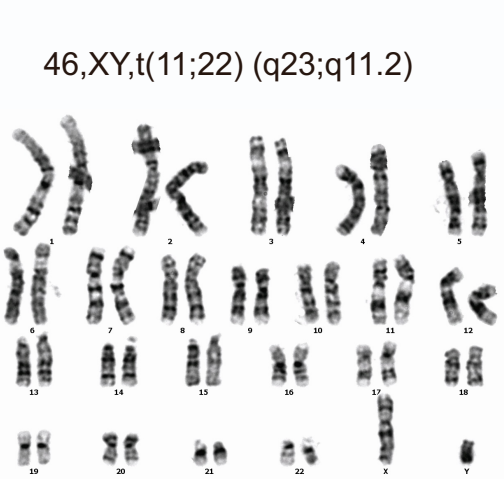


Fig. S6 Fine annotations of 16 clusters into 6 functional cell types and imputation analysis. (**A**) Heatmap showing the unique markers of 6 main cell types in the day-60 HCOs. NSC, neural stem cells; RGC, radial glial cells; IP, intermediate progenitors; CN, cortical neurons; M/H, mid/hindbrain cells; MC, mesenchymal-like cells. (**B**) Dot-plot showing the top key markers expressed in each cell type. Color of dot indicates the relative expression and size of dot indicates the percentage of cells expressing the gene. Expression of *PHOX2B* (**C**) before and (**D**) after imputation**.**

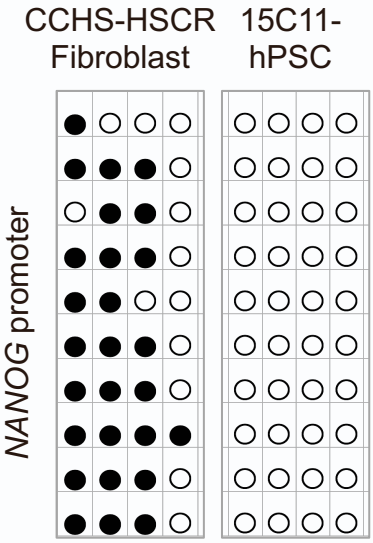
A

Clinical features:
• Central hypoventilation syndrome (CCHS) • Short Hirschsprung disease (S-HSCR) • Marcus gunn phenomenon • Global developmental delay
Genotype:
PHOX2B-7Ala-PARM t(11;22) (q23;q11.2)

B



C



D

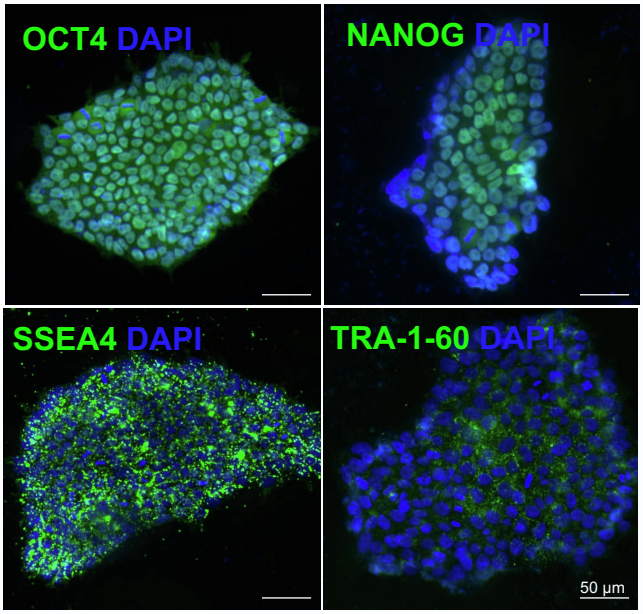


Fig. S7 Establishment and characterization of a CCHS-hPSC (15C11) line. (**A**) Table summarized the clinical features and genotype of the CCHS patient. (**B**) Karyotype of 15C11 hPSC. (**C**) Promoter of *NANOG* is unmethylated in the 15C11 hPSC. (**D**) Immunocytochemistry showing the expression of stem cell markers in the 15C11 hPSC. Scale bars: 50 μm.

Supplementary Materials and methods

Human induced pluripotent stem cells (hPSCs)

UE02302 control hPSC line was kindly provided by Dr. Guangjin Pan (Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences, China)[1]. A CCHS patient hPSC line (15C11) was generated from skin fibroblast of a CCHS patient with episomal reprogramming vector carrying the 4 reprogramming factors (*Oct4, Klf4, Sox2* and *c-Myc*) as described previously[2,3]. The study was approved by the Institutional Review Board of The University of Hong Kong and the Hospital Authority (UW 13-419). All the hPSC lines were maintained in Matrigel (Corning, 354277)-coated 6-well culture plates with mTeSR1 medium (STEMCELL Technologies, 85850) in 37 °C humidified 5% CO_2 incubator. The hPSCs were regularly passaged by treating with Dispase (STEMCELL Technologies, 07923) and the culture medium was changed daily.

Plasmid constructions

The human codon-optimized Cas9 expression plasmid was obtained from Addgene (44720). Two guide RNAs (gRNA) were designed using the CRISPR design tool (http://crispr.mit.edu) to target the 20-Ala polyalanine tract in exon 3 of *PHOX2B* (PHOX2B-gRNAa: 5'-CCT TAG TGA AGA GCA GTA TGT TC-3'; PHOX2B-gRNAb: 5'-AAG AGC AGT ATG TTC TGA TCT GG-3'). The gRNA plasmid (Addgene, 41824) was linearized with *Afl*II and the gRNA targeting sequence (see Supplementary Table 1) was cloned into gRNA plasmid using Gibson Assembly (New England Biolabs, E2611) according to Church's lab protocol[4].

The donor plasmid, pHR-PHOX2B-L-2A-EGFP, containing the homology arms spanning a region from chr4:41746593-41745806 (GRCh38.p13) and a puromycin selection cassette designed for *PHOX2B* targeting, is a generous gift from Dr. Yohan Oh (Johns Hopkins University School of Medicine, Baltimore). For construction of *PHOX2B-PARM* mutant donor plasmids containing different length of PARMs, oligos with additional Ala repeats were annealed and extended using Q5 High-fidelity DNA Polymerase (New England Biolabs, M0491). Annealed double-stranded oligos containing the desired additional Ala repeats (+5Ala, +7Ala and +13Ala) together with a 189-bp PCR fragment containing the 3'-end of the polyalanine tract region were subcloned into the ~8kb PCR fragment of the wild-type *PHOX2B* donor plasmid backbone (all designed with ~40bp overlapping sequences at each end) by

Gilson Assembly according to manufacturer's protocol. DNA sequence of all the *PHOX2B-PARM* mutant donor plasmids were confirmed by Sanger sequencing.

Generation of PHOX2B-PARM mutant hPSC cell lines

The CRISPR-Cas9 system[5] was used to target the exon3 of *PHOX2B* allele (after stop codon) in the UE02302 control hPSCs. The hPSCs were transfected with the mutant *PHOX2B-PARM* donor plasmids together with the gRNA constructs and a Cas9 nickase expression plasmid using Human Stem Cell Nucleofector Kit 2 (Lonza, VPH-5022). The transfected cells were selected for 7 days in 0.2 μg ml^{-1} Puromycin and the Puromycin-resistant clones were isolated and passaged twice before genotyping. The targeted region of *PHOX2B* gene was genotyped by using the primers listed in Supplementary Table 1 to confirm the site-specific insertion of PARMs using Sanger sequencing. The mutant hPSC clones with the confirmed PARM insertion were expanded for further assays.

Correction of PHOX2B mutation in the CCHS-hPSC line

The strategy for the correction of patient-specific *PHOX2B* mutation was similar to the protocol described above except for the use of the wild-type *PHOX2B* donor plasmid instead of the mutant *PHOX2B-PARM* donor plasmids. In brief, CCHS (15C11) hPSCs were transfected with the wild-type *PHOX2B* donor plasmid, the gRNA constructs and a Cas9-D10A nickase expression plasmid using Human Stem Cell Nucleofector Kit 2. Selection and screening of positive clones were performed as described above.

Generation of human cerebral organoids (HCOs) from hPSCs

HCOs were generated from hPSCs using a commercially optimized kit, STEMdiff Cerebral Organoid Kit (STEMCELL Technologies, 08570) following the manufacturer's protocol and previously described protocols[6,7] with modifications. First, to generate embryoid body (EB) from hPSCs for the formation of HCOs, hPSCs were washed with DPBS (Cytiva, SH30028.03) and treated with 0.5 mM EDTA solution in DPBS for 4 min at 37 °C. After removing the EDTA solution, the cells were dissociated into single cell suspension by treating with Accutase (STEMCELL Technologies, 07922) for 4 min at 37 °C. The cells were pelleted by centrifugation for 5 minutes at 270 × *g* at room temperature and then resuspended with EB

Formation Medium with ROCK inhibitor Y-27632 (10 μM). 9,000 cells were transferred to each well of 96-well round bottom ultra-low attachment plate (Corning, 7007) and kept undisturbed overnight in 37 °C incubator for the EB formation. The cells were marked as day 0 at this stage. At day 2 and 4, the EBs were fed with additional 100 μl EB Formation Medium per well. On day 5, one EB was transferred to each well of 24-well ultra-low attachment plate (Corning, 3473) with 500 μl Induction Medium to induce the formation of neuroepithelia. On day 7, the organoids were embedded in Matrigel and transferred to 6-well ultra-low attachment plate (Corning, 3471) for further expansion in Expansion Medium. On day 11, the medium was replaced with Maturation Medium and the 6-well plate (with 6-8 organoids per well) was kept shaking at 75 r.p.m. with an orbital shaker installed inside the incubator. The medium was changed every 3-4 days until day 60, when the organoids were collected.

Generation of brainstem organoids (HSBOs) from hPSCs

To generate HSBOs, hPSCs were dissociated into single cell suspension by treating with 0.5 mM EDTA solution in DPBS and then Accutase as described above. 2.7 $\times 10^6$ cells were seeded to a well of AggreWell800 24-well plate (STEMCELL Technologies, 34815) which contained 300 microwells per well (i.e. 9,000 cells per microwell). The cells were kept in mTeSR1 with 10 μM Y-27632 overnight in 37 °C incubator for the spheroid formation and they were marked as day 0 at this stage. At day 1, the spheroids from one well of AggreWell800 were split into 3 wells of 6-well ultra-low-attachment plate containing Neural Induction Medium (DMEM/F12, 1% N2 supplement (v/v), 1% GlutaMAX supplement (v/v), 1% MEM-NEAA (v/v) and 1 μg ml^{-1} Heparin) supplemented with BMP inhibitor LDN-193189 (1 μM) and TGF-β inhibitor SB431542 (10 μM). From day 3 to 6, the organoids were incubated in Neural Differentiation Medium (NDM) containing DMEM/F12 and Neurobasal Medium in 1:1 ratio with 0.5% N2 supplement (v/v), 1% B27 supplement without Vitamin A (v/v), 1% GlutaMAX (v/v), 0.5% MEM non-essential amino acids (NEAA), 2.5 μg ml^{-1} insulin, 49.5 μM β-mercaptoethanol and 1% Penicillin-Streptomycin (Pen-Strep) (v/v). LDN-193189 (1 μM), SB431542 (10 μM) and SHH agonist SAG (1 μM) were added to NVM at day 3 to 6. Starting from day 3, the organoids were kept shaking at 90 r.p.m. with an orbital shaker installed inside the incubator. From day 6 to 12, the organoids were maintained in NDM with SAG (1 μM), WNT agonist CHIR99021 (3 μM) and retinoic acid (RA) (1 μM) and the medium was changed every 2-3 days. From day 12 onwards, the organoids were maintained in Neural Maturation Medium (NMM) (same recipe as NDM but with B27 containing Vitamin A) supplemented

with 20 ng ml^{-1} BDNF, 20 ng ml^{-1} GDNF, 50 nM dibutyryl-cAMP and 200 μM ascorbic acid. The medium was changed every 3-4 days until day 60, when the organoids were collected. The suppliers and catalog numbers of the media and supplements used for the preparation of culture media were listed in Supplementary Table 1.

Multielectrode array (MEA) analysis

Day 60 HBSOs were dissociated with Papain Dissociation System (Worthington, LK003150) following the manufacturer's protocol. In brief, the HBSOs were digested with the papain solution in 37 °C incubator for 30 min with continuous shaking at 90 r.p.m. and then triturated with P1000 pipette. The triturated HBSOs were further digested in 37 °C incubator for an additional 10 min and then filtered through 40 μm cell strainer (Corning, 352340) to remove undigested cell debris. The dissociated cells were collected by centrifuging at 300 × *g* for 5 min at room temperature. The cells were then resuspended with DNase/ovomucoid inhibitor solution to stop the digestion. The cells were centrifuged at 300 × *g* for 5 min at room temperature and resuspended in BrainPhys Neuronal Medium (STEMCELL Technologies, #05790) supplemented with 0.5% N-2 supplement (v/v), 1% B-27 supplement (v/v), 1% GlutaMAX (v/v), 0.5% NEAA, 2.5 μg ml^{-1} insulin, 20 ng ml^{-1} BDNF, 20 ng ml^{-1} GDNF, 50 nM dibutyryl-cAMP, 200 μM ascorbic acid, 49.5 μM β-mercaptoethanol and 1% (v/v) Pen-Strep. Five-microliter cell suspension with 1×10^5 cells and 10 μg ml^{-1} laminin was seeded into each well of 24-well CytoView MEA plate (Axion Biosystems, M384-tMEA-24W) pre-coated with 0.1% polyethylenimine (PEI) (Sigma-Aldrich, P3143) and maintained in 37 °C humidified 5% CO_2 incubator for at least 2 weeks before MEA analysis. The culture medium was changed every 3-4 days. MEA recordings were measured at 37 °C with 5-8% CO_2 in a Maestro Edge MEA System (Axion Biosystems) to detect spontaneous firing of spikes. The MEA plate was pre-equilibrate in the MEA system for at least 5 min before measurement and each treatment was recorded for 5-10 min to calculate the mean spike firing rate. For the pharmacological experiment, retigabine (Selleckchem, S4733) (10 μM) was applied to plate immediately before recording. The MEA recordings were analyzed using the Neural Module of AxIS Navigator 1.5 software and the Neural Metric Tool software (Axion Biosystems). The mean spike firing rate of each MEA recording was normalized with the first MEA recording undertaken in 5% CO_2 or 8% CO_2 condition to calculate the fold change.

Quantitative RT-PCR (qPCR)

Total RNA from HCOs and HBSOs were extracted by RNeasy Mini Kit (Qiagen) and the RNA concentrations were determined by Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific). The RNA was then reverse-transcribed to cDNA using HiScript III All-in-one RT SuperMix (Vazyme, R333-01). For real-time qPCR, cDNA samples were amplified by Luna Universal qPCR Master Mix (New England BioLabs, M3003) using specific primer pairs with PCR profiles of 95 ºC (1 min) followed by 45 cycles of 95 ºC (15 s) and 60 ºC (30 s). Fluorescence was measured by ViiA 7 Real-Time PCR System (Thermo Fisher Scientific) at the end of each cycle. Each individual sample was assayed in triplicate and gene expression was normalized with *GAPDH* expression. The primers for qPCR analysis were listed in Supplementary Table 2.

Western blot

To extract proteins from HBSOs, HBSOs were lysed with Cell Lysis Buffer (1X) (Cell Signaling Technology, #9803) with PhosSTOP (Roche), Complete protease inhibitor cocktails (Roche) and 0.1 mM phenylmethylsulfonyl fluoride (PMSF) for 30 min on ice with vortex at intervals. The lysates were centrifuged at 10,000 × *g* at 4 ºC and the supernatant was collected. The protein extracts were separated by 12% SDS-PAGE and transferred to Immobilon-P PVDF membrane (Millipore, IPVH00010) using Trans-Blot® SD Semi-Dry Transfer Cell (Bio-Rad). The membrane was briefly washed with TBST and then blocked with blocking solution for 1 h at room temperature. After blocking, the membrane was incubated with 1:1000 rabbit anti-PHOX2B antibody (Abcam, ab) or 1:20,000 mouse anti-Actin antibody (Millipore, MAB1501) diluted in blocking solution at 4 ºC overnight. The membrane was washed with TBST thrice for 10 min and then incubated with 1:2000 goat anti-rabbit immunoglobulins-HRP (Dako, P044801) or 1:5000 anti-mouse immunoglobulins-HRP (Dako, P044701) diluted in blocking solution for 1 h at room temperature. After washing thrice with TBST for 10 min, the HRP signals were detected using WesternBright ECL (Advansta, K-12045-D50).

Immunofluorescence

The HCOs and HSBOs were fixed with 4% (w/v) paraformaldehyde (PFA) in PBS at room temperature for 1-2 h or 4 °C overnight. After washing with PBS for 5 min thrice, the samples were dehydrated with 15% (w/v) sucrose in PBS and then with 30% (w/v) sucrose in PBS at 4 °C overnight. The dehydrated samples were embedded in Tissue-Tek O.C.T. Compound (Sakura, 4583) and frozen on dry ice. The frozen blocks were stored at -80 °C until sectioning. Sections with 10-20 μm-thickness were cut using Leica CM3050 S cryostat and mounted on

Superfrost Plus Adhesion Slides (Thermo Fisher Scientific, J1800AMNZ). The frozen sections were kept at -20 °C for storage. The frozen sections were warmed at 37 °C for 10-30 min before IF procedures and washed with PBS for 5 min twice with gentle shaking. The sections were blocked with 5% (v/v) donkey serum in PBS with 0.1% (v/v) Triton X for 1 h at room temperature and then incubated with primary antibodies at 4 °C overnight. After washing with PBS for 10 min thrice, the sections were incubated with respective secondary antibodies for 2 h at room temperature. After washing with PBS for 10 min thrice, the stained cells were mounted with ProLong Diamond Antifade Mountant with DAPI (Thermo Fisher Scientific, P36971). The primary and secondary antibodies used for IF and ICC were listed in Supplementary Table 3. All the fluorescence images were acquired by Carl Zeiss LSM800 or LSM900 confocal microscope.

Droplet-based single-cell RNA-sequencing (scRNA-seq)

HBSOs and HCOs were harvested at day 60 and HBSOs were digested with Papain Dissociation System (Worthington) as described above while HCOs were digested with Accutase for 30 min at 37 °C with agitation. The HCOs were dissociated by pipetting every 10 min during digestion. All the dissociated cells were then filtered with 40 μm cell strainer and resuspended in 2% (v/v) FBS in DPBS in a density of 1×10^7 cells ml^{-1}. The viable cells were labeled with 7-AAD (5 μl per 10^6 cells; BD Biosciences, 559925) for at least 10 min on ice. After filtering through 35 μm cell strainer, the cells negative for 7-AAD were gated and sorted with BD FACSAria III Cell Sorter using purity mode with purity ≥99%. The FACS-sorted cells were then subjected to droplet-based scRNA-seq using Chromium Single Cell platform and Single Cell 3' Library Kits (10x Genomics) in Centre of PanorOmic Science (CPOS), The University of Hong Kong. In brief, cells were encapsulated into Gel Beads-in-emulsion (GEMs) by the 10X Chromium Single Cell Controller, followed by reverse transcription and library preparation to become a pool of cDNA libraries. Libraries were then purified and sequenced on an Illumina™ NextSeq 500 according to the manufacturer's protocol.

Pre-processing of droplet-based scRNA-seq data

Cellranger toolkit (version 3.0) provided by 10X Genomics were used to perform the reads de-multiplexing and alignment. unique molecular identifier (UMI) were counted by aligning to the human hg38 transcriptome. For the quality control, only cells with more than 800 detected genes were retained. Genes expressed in more than 10 cells were kept for further analysis.

Clustering and marker identification were performed by the Seurat R package[11]. The integrative analysis of the public datasets was also performed by Seurat.

Dimensionality reduction and cell clustering

The R package Seurat (version 3.1.4)[12] implemented in R (version 3.6.2) were used to perform dimensional reduction analysis. The "NormalizeData()" function from Seurat was used to normalize the raw counts, and the scale factor was set to 200,000, then followed by "FindVariableFeatures()" with default parameters to calculate highly variable genes for each sample. After performing "JackStraw()", which returned the statistical significance of PCA (Principal component analysis) scores, we selected 20 significant PCs to conduct dimension reduction and cell clustering. Then, cells were projected in 2D space using *t*-SNE (*t*-distributed Stochastic Neighbor Embedding) with default parameters.

Single cell data integration and batch effect correction

To account for batch effect among different samples, we used "FindIntegrationAnchors()" in the Seurat package to remove batch effect and merge samples to one object. In detail, the top 4,500 genes with the highest expression and dispersion from each sample were used to find the integration anchors, and then the computed anchoret was applied to perform dataset integration.

Organoid comparison

Human brainstem organoids (HBSOs) and human cerebral organoids (HCOs) were derived from the sequencing data in this article, while human cortical organoids (hFBCOs, GSE137941), human midbrain organoids (hMBOs, GSE133894), human thalamic organoids (hThaOs, GSE122342) and human arcuate organoids (hARCOs, GSE164102) were downloaded from Gene Expression Omnibus (GEO) database. Raw count matric of UMI were downloaded and normalized by NormalizeData() function in Seurat R package. Cells from each organoid models were pooled together to construct "pseudo-bulk" expression for each organoid. Similarities between different organoids were estimated by Spearman correlation.

Reference-based cell annotation

Developing [13] and adolescent [14] mouse brain atlas were downloaded from http://mousebrain.org/. Developing and adolescent neuron, glia, OPC, radial glia, Schwann cell and astrocyte located at hindbrain were extracted for comparison with our HBSO scRNA-seq dataset. Reference-based cell annotation was performed by Seurat R package. In brief, raw

count matric were normalized and scaled first. Then, FindTransferAnchors() function was applied to find shared anchors between datasets. Finally, MapQuery() was used to project query dataset (HBSOs) to reference datasets (developing and adolescent mouse brain cells). Thus, original cell type labels were transferred to HBSOs datasets based on the overall expression patterns of genes and cells in HBSOs were classified to six main cell types (neuron, glia, OPC, radial glia, Schwann cell and astrocyte).

Human cerebral organoids from public scRNA-seq datasets [15] were downloaded from https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-7552/. 2-month-old human cerebral organoids were extracted and overlaid with our HCO cells for comparison analysis. Potential technical and batch effects were removed by Seurat software[12]. The sixteen unbiased clusters were further classified into six cell types, including neural stem cells (NSC), radial glia cells (RGC), intermediate progenitors (IP), cortical neurons (CN), midbrain/hindbrain (M/H) and mesenchymal-like cells (MC) based on the expression of canonical markers and the cell annotation from the public datasets.

Marker-based neuronal subtype annotation

Neuronal cells in HBSOs were further classified into four main neuron subtypes (Cholinergic, Serotonergic, GABAergic and Glutamatergic) and two novel subtypes ($PHOX2B^{Low}$ and $PHOX2B^{High}$) based on the canonical marker genes as stated in previous literature [16]. Gene signature expression of each subtype was calculated to confirm the identity of neuronal cells.

Lineage cladogram construction

Pseudo-bulk expression (average expression) of each cell type was calculated. Pairwise Euclidean distances were calculated among cell types. Neighbor-joining (NJ) tree was estimated based on the distance matrix. ggtree R package [17] was used to visualized the lineage cladogram.

RNA velocity analysis

Velocyto (http://velocyto.org/velocyto.py/index.html) was used to calculate spliced and unspliced matric from bam files that derived from cellranger pipeline. Then, scVelo [18] was applied for RNA velocity analysis in single cells.

Single cell trajectory analysis

Single cell trajectory analysis was performed by PCA algorithm using *PHOX2B*$^{+}$ cells.

Identification of differentially expressed genes and pathway enrichment analysis

To identify unique differentially expressed genes (DEGs) among each cluster, the "FindAllMarkers()" function from Seurat was used and non-parametric Wilcoxon rank sum tests were set to evaluate the significance of each individual DEG. The DEG analysis between mutant and control based on single-cell expression data was performed using monocle (version 2.14.0) R package[19]. The DEGs with adjusted *P*-value less than 0.05 and Log_2 Fold Change (log2FC) larger than 0.5 were thought to be significant and used in downstream analysis. Gene ontology (GO) term enrichment analyses were performed using clusterProfiler (version 3.14.3) R package[20]. Terms that had an adjusted *P*-value < 0.05 was defined as significantly enriched.

Pathway score analysis

In order to measure the activity of whole pathways between samples at transcriptome level, we applied a simple additive model that ignore the interactions between genes to estimate the overall expression level of the pathways. The genes involved in GO:0007389 (pattern specification process) in GO database were used to evaluate the overall changes of the key biological processes.

Gene regulatory network (GRN) analysis

Single cell regulatory network inference and clustering (SCENIC, version 1.1.2)[21] was used to infer transcription factor networks active using scRNA-Seq data. Analysis was performed using default and recommended parameters as directed on the SCENIC vignette (https://github.com/aertslab/SCENIC) using the hg38 RcisTarget database. Kernel density line histograms showing differential AUC score distribution across conditions were plotted with ggplot2 v.3.1.1 using the regulon activity matrix ('3.4_regulonAUC.Rds', an output of the SCENIC workflow) in which columns represent cells and rows the AUC regulon activity. Fold-change (FC) difference between median AUC values was calculated and the highest changed TFs were plotted.

Differential co-expression analysis

To predict PHOX2B activated and repressed target genes, we applied a three-step strategy to obtain the target genes of PHOX2B during progenitor-to-neuron transition. In the first step, we identified the significantly correlated genes to PHOX2B which represented the genes that had

strong regulatory relationships with PHOX2B. In the second step, we utilized the differential expression profile between control and mutant to filter out the genes whose expression levels were not significantly disrupted by PHOX2B. Finally, to confine the genes to be the actual targets of PHOX2B, a public ChIP-seq dataset designed for PHOX2B from a similar cell state (neuronal progenitor) in Cistrome database [22] was utilized to obtain the direct/indirect binding genes of PHOX2B. Activated and repressed target genes were determined by the positive/negative correlation and up/down-regulation of the differential expression, respectively. In addition, TF motif enrichment analysis by FIMO tool [23] of MEME suite [24] was used to cross-validate the predicted activated targets of PHOX2B.

Imputation

Markov affinity-based graph imputation of cells (MAGIC)[25] was used to denoise the cell count matrix and fill in missing transcripts. The expression of PHOX2B could be well-recovered and most of the PHOX2B+ cells were found in cell clusters inferred to be less mature progenitors (IP) or more mature neurons (CN), and few fell into NSC in control HCOs.

Image analysis and statistical analysis

Cell numbers in immunostaining or immunocytochemistry were quantified using "Cell Counter" or "Analyze Particles" modules in Fiji ImageJ software [26,27]. Statistical analysis was performed with the software Prism 9.0 (GraphPad Software). Unless otherwise noted, the data were expressed as mean ± SEM. Two-tailed unpaired Student's *t*-test was used for comparisons involving two groups. One-way ANOVA, followed by Turkey's multiple comparison test, was used for comparisons involving more than two groups. Probability values $P < 0.05$ were considered as statistically significant.

Supplementary Table 1. List of media and supplements used for culture of hPSCs, HBSOs, HCOs, hENCCs and ENS neurons.

Media and supplements	**Applications**	**Suppliers**	**Catalog#**
mTesR1 medium	hPSCs, HBSOs	STEMCELL Technologies	85850
STEMdiff Cerebral Organoid Kit	HCOs	STEMCELL Technologies	08570
BrainPhys Neuronal Medium	HBSOs (MEA)	STEMCELL Technologies	05790
DMEM/F-12	hiPSCs, HBSOs	Thermo Fisher Scientific	11330032
DMEM/F-12 (with GlutaMAX supplement)	hENCCs and ENS neurons	Thermo Fisher Scientific	10565018
KnockOut DMEM	hENCCs	Thermo Fisher Scientific	10829018
Neurobasal Medium	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	21103049
KnockOut Serum Replacement (KOSR)	hENCCs	Thermo Fisher Scientific	10828028
B-27 Supplement	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	17504044
B-27 supplement without Vitamin A	HBSOs	Thermo Fisher Scientific	12587010
N-2 Supplement	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	17502048
Insulin	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	12585014
GlutaMAX supplement	HBSOs	Thermo Fisher Scientific	35050061
L-Glutamine	hENCCs	Thermo Fisher Scientific	25030081
MEM Non-Essential Amino Acids (NEAA)	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	11140050
β-mercaptoethanol	HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	21985023
Penicillin-Streptomycin (Pen-Strep)	HCOs, HBSOs, hENCCs and ENS neurons	Thermo Fisher Scientific	15140122
Heparin	HBSOs	Sigma-Aldrich	H3149
FGF-2	hENCCs and ENS neurons	Peprotech	100-18B
Y-27632	hPSCs, HCOs, HBSOs, hENCCs and ENS neurons	Tocris	1254
LDN-193189	HBSOs, hENCCs and ENS neurons	Reprocell	04-0074
SB431542	HBSOs, hENCCs and ENS neurons	Abcam	ab120163
SAG	HBSOs	Sigma-Aldrich	SML1314
IWR-1-endo	HBSOs	Abmole	M2782

CHIR99021	HBSOs, hENCCs and ENS neurons	Reprocell	04-0004
Retinoic acid (RA)	HBSOs, hENCCs and ENS neurons	Abcam	ab120728
BDNF	HBSOs and ENS neurons	Peprotech / Qkine	450-02 / Qk050
GDNF	HBSOs and ENS neurons	Peprotech / Shenandoah	450-10 / 100-02
NT-3	hENCCs and ENS neurons	Peprotech	450-03
β-NGF	hENCCs and ENS neurons	Peprotech	450-01
Dibutyryl-cAMP	HBSOs, hENCCs and ENS neurons	Sigma-Aldrich	D0260
Ascorbic acid	HBSOs, hENCCs and ENS neurons	Sigma-Aldrich	A4034

Supplementary Table 2. List of RT-PCR primers used in this study.

Gene	Oligo sequence (5' to 3')	Applications	Tm (°C)	Product size (bp)
HOXA2	Forward: CGTCGCTCGCTGAGTGCCTG Reverse: TGTCGAGTGTGAAAGCGTCGAGG		63.3 61.1	92
HOXB1	Forward: GAGCTTTGCACCGGCCTAT Reverse: CTTCATCCAGTCGAAGGTCCG		58.1 57.4	103
HOXB2	Forward: CCTAGCCTACAGGGTTCTCTC Reverse: CACAGAGCGTACTGGTGAAAAA		56.1 55.4	79
GBX2	Forward: ACGTCAGCAGGTTCGCTATC Reverse: AGCTGGGCTGTGACTTTGTT		57.2 57.3	166
OTX2	Forward: AGTCGAGGGTGCAGGTATG Reverse: GGCCACTTGTTCCACTCTCT		56.6 57.1	153
PAX7	Forward: CAAACACAGCATCGACGGC Reverse: CTTCAGTGGGAGGTCAGGTTC		57.3 57.2	91
OLIG3	Forward: TCATGCTCACCAGCTCCCT Reverse: CCCCCATAGATCTCGCCAAC		58.7 57.6	58
NGN1	Forward: AAGACTTCACCTACCGCCC Reverse: GGCGTTGTGTGGAGCAAGT		56.9 58.6	76
PHOX2B	Forward: AAACTCTTCACGGACCACGG Reverse: CTCCTGCTTGCGAAACTTGG		57.4 56.8	222
GAPDH	Forward: CAAGAAGGTGGTGAAGCAGGC Reverse: GCCAAATTCGTTGTCATACCAGGA		58.7 57.6	184

Supplementary Table 3. List of antibodies used in this study.

Target antigens	**Antibodies**	**Dilutions**	**Applications**	**Suppliers**	**Catalog#**	**RRID**
Actin	Mouse monoclonal anti-Actin	1:20,000	WB	Sigma-Aldrich	MAB1501	AB_2223041
CD271 (LNGFR) ($p75^{NTR}$)	Mouse anti-human CD271 (LNGFR), FITC	2μl per 10^7 cells	FACS	Miltenyi Biotec	130-113-420	AB_2733631
CD57 (HNK1)	Mouse anti-human CD57, APC	1μl per 10^7 cells	FACS	BD Biosciences	560845	AB_10563760
Cleaved Caspase-3 (C-Cas3)	Rabbit anti-Cleaved Caspase-3	1:200	IF	Cell Signaling Technology	9661	AB_2341188
Forkhead box A2 (FOXA2)	Mouse monoclonal anti-FOXA2 [7E6]	1:200	IF	Abcam	ab60721	AB_941632
Green fluorescent protein (GFP)	Sheep polyclonal anti-GFP	1:500	IF, ICC	Bio-Rad	4745-1051	AB_619712
Islet 1 (ISL1)	Rabbit monoclonal anti-ISL1 [EP4182]	1:200	IF	Abcam	ab109517	AB_10866454
Nanog Homeobox (NANOG)	Goat polyclonal anti-NANOG	1:500	ICC	R&D Systems	AF1997	AB_355097
Neurofilament (NF)	Chicken polyclonal anti-NF-L	1:500	ICC	Neuromics	CH22105	AB_2737102
Neuronal differentiation 6 (NEUORD6)	Rabbit polyclonal anti-MATH2/NEUROD6	1:200	IF	Abcam	ab85824	AB_10671651
NK2 Homeobox 2 (NKX2.2)	Mouse monoclonal anti-Nkx2.2	1:100	IF	DSHB	74.5A5	AB_531794
Octamer-binding transcription factor 4 (OCT4)	Goat polyclonal anti-Oct-3/4	1:100	ICC	Santa Cruz	sc-8629	AB_2167705
Orthodenticle homeobox 2 (OTX2)	Rabbit polyclonal anti-Otx1+Otx2	1:200	IF	Abcam	ab21990	AB_776930
Paired like homeobox 2B (PHOX2B)	Goat polyclonal anti-PHOX2B	1:200	IF, ICC	R&D Systems	AF4940	AB_10889846
Paired like homeobox 2B (PHOX2B)	Rabbit monoclonal anti-PHOX2B [EPR14423] – C-terminal	1:200	ICC, WB	Abcam	ab183741	AB_2857845
Protein gene product 9.5 (PGP9.5)	Rabbit monoclonal anti-PGP9.5 [EPR4118]	1:500	ICC	Abcam	ab108986	AB_10891773
Serotonin (5-HT)	Rat monoclonal anti-Serotonin [YC5/45]	1:200	IF	Abcam	ab6336	AB_449517
Sex-determining region Y-box 10 (SOX10)	Mouse monoclonal anti-SOX10	1:500	ICC	R&D Systems	MAB2864	AB_2195180
Sex-determining region Y-box 2 (SOX2)	Mouse monoclonal anti-SOX2	1:200	IF	R&D Systems	MAB2018	AB_358009

Stage-Specific Embryonic Antigen-4 (SSEA4)	Mouse monoclonal anti-SSEA-4	1:200	ICC	Sigma-Aldrich	MAB4304	AB_177629
T cell receptor alpha locus (TRA-1-60)	Mouse monoclonal anti-TRA-1-60	1:100	ICC	Sigma-Aldrich	MAB4360	AB_2119183
T-Box brain transcription factor 2 (TBR2)	Rabbit polyclonal anti-TBR2	1:200	IF	Abcam	ab23345	AB_778267
Tubulin beta 3 class III (TUJ1 or TUBB3)	Rabbit polyclonal anti-TUJ1	1:1,000	ICC	Abcam	ab18207	AB_444319
Tubulin beta 3 class III (TUJ1 or TUBB3)	Mouse monoclonal anti-TUJ1	1:1,000	IF, ICC	BioLegend	801202	AB_10063408
Tyrosine hydroxylase (TH)	Mouse monoclonal anti-TH	1:500	IF, ICC	Sigma-Aldrich	MAB318	AB_2201528
Vesicular glutamate transporter 2 (VGLUT2)	Guinea pig polyclonal anti-VGLUT2	1:1,000	IF	Sigma-Aldrich	AB2251-I	AB_2665454
Alexa Fluor® 488 Donkey-anti-goat IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A11055	AB_2534102
Alexa Fluor® 488 Donkey-anti-mouse IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A21202	AB_141607
Alexa Fluor® 488 Donkey-anti-rabbit IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A21206	AB_2535792
Alexa Fluor® 488 Donkey-anti-sheep IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A11015	AB_2534082
Alexa Fluor® 594 Donkey-anti-mouse IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A21203	AB_141633
Alexa Fluor® 594 Donkey-anti-rabbit IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A21207	AB_141637
Alexa Fluor® 594 Donkey-anti-rat IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A21209	AB_2535795
Alexa Fluor® 594 Goat-anti-chicken IgY (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A11042	AB_2534099
Alexa Fluor® 647 Donkey-anti-mouse IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A31571	AB_162542
Alexa Fluor® 647 Donkey-anti-rabbit IgG (H+L)		1:200	IF, ICC	Thermo Fisher Scientific	A31573	AB_2536183
Donkey anti-guinea pig IgG (H+L) highly cross-adsorbed, CF594		1:200	IF, ICC	Sigma-Aldrich	SAB4600096	AB_2728625
Goat anti-rabbit immunoglobulins-HRP		1:2,000	WB	Dako	P044801	AB_2617138
Rabbit anti-mouse immunoglobulins-HRP		1:5,000	WB	Dako	P044701	AB_2617137

Supplementary References

1. Xue Y, Cai X, Wang L, *et al*. Generating a non-integrating human induced pluripotent stem cell bank from urine-derived cells. *PLoS One*. 2013;8(8):e70573. doi:10.1371/journal.pone.0070573
2. Takahashi K, Okita K, Nakagawa M, Yamanaka S. Induction of pluripotent stem cells from fibroblast cultures. *Nat Protoc*. 2007-12-01 2007;2(12):3081-9. doi:10.1038/nprot.2007.418
3. Lai FP, Lau ST, Wong JK, *et al*. Correction of Hirschsprung-Associated Mutations in Human Induced Pluripotent Stem Cells Via Clustered Regularly Interspaced Short Palindromic Repeats/Cas9, Restores Neural Crest Cell Function. *Gastroenterology*. Jul 2017;153(1):139-153 e8. doi:10.1053/j.gastro.2017.03.014
4. Mali P, Yang L, Esvelt KM, *et al*. RNA-guided human genome engineering via Cas9. *Science*. Feb 15 2013;339(6121):823-6. doi:10.1126/science.1232033
5. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. Nov 2013;8(11):2281-2308. doi:10.1038/nprot.2013.143
6. Lancaster MA, Renner M, Martin CA, *et al*. Cerebral organoids model human brain development and microcephaly. *Nature*. Sep 19 2013;501(7467):373-9. doi:10.1038/nature12517
7. Lancaster MA, Knoblich JA. Generation of cerebral organoids from human pluripotent stem cells. *Nat Protoc*. Oct 2014;9(10):2329-40. doi:10.1038/nprot.2014.158
8. Lee G, Chambers SM, Tomishima MJ, Studer L. Derivation of neural crest cells from human pluripotent stem cells. *Nat Protoc*. Apr 2010;5(4):688-701. doi:10.1038/nprot.2010.35
9. Tang CS, Li P, Lai FP, *et al*. Identification of Genes Associated With Hirschsprung Disease, Based on Whole-Genome Sequence Analysis, and Potential Effects on Enteric Nervous System Development. *Gastroenterology*. Dec 2018;155(6):1908-1922 e5. doi:10.1053/j.gastro.2018.09.012
10. Lau ST, Li Z, Lai FPL, *et al*. Activation of Hedgehog Signaling Promotes Development of Mouse and Human Enteric Neural Crest Cells, Based on Single-Cell Transcriptome Analyses. *Gastroenterology*. Dec 2019;157(6):1556-1571 e5. doi:10.1053/j.gastro.2019.08.019
11. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. Jun 2018;36(5):411-420. doi:10.1038/nbt.4096
12. Hao Y, Hao S, Andersen-Nissen E, *et al*. Integrated analysis of multimodal single-cell data. *Cell*. Jun 24 2021;184(13):3573-3587 e29. doi:10.1016/j.cell.2021.04.048
13. La Manno G, Siletti K, Furlan A, *et al*. Molecular architecture of the developing mouse brain. *Nature*. 2021-08-05 2021;596(7870):92-96. doi:10.1038/s41586-021-03775-x
14. Zeisel A, Hochgerner H, Lönnerberg P, *et al*. Molecular Architecture of the Mouse Nervous System. *Cell*. 2018;174(4):999-1014.e22. doi:10.1016/j.cell.2018.06.021
15. Kanton S, Boyle MJ, He Z, *et al*. Organoid single-cell genomic atlas uncovers human-specific features of brain development. *Nature*. Oct 2019;574(7778):418-422. doi:10.1038/s41586-019-1654-9
16. Liu W, Venugopal S, Majid S, *et al*. Single-cell RNA-seq analysis of the brainstem of mutant SOD1 mice reveals perturbed cell types and pathways of amyotrophic lateral sclerosis. *Neurobiol Dis*. Jul 2020;141:104877. doi:10.1016/j.nbd.2020.104877
17. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2017-01-01 2016;8(1):28-36. doi:10.1111/2041-210x.12628

18. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nature Biotechnology*. 2020/12/01 2020;38(12):1408-1414. doi:10.1038/s41587-020-0591-3
19. Qiu X, Mao Q, Tang Y, *et al*. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. Oct 2017;14(10):979-982. doi:10.1038/nmeth.4402
20. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. May 2012;16(5):284-7. doi:10.1089/omi.2011.0118
21. Aibar S, Gonzalez-Blas CB, Moerman T, *et al*. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. Nov 2017;14(11):1083-1086. doi:10.1038/nmeth.4463
22. Liu T, Ortiz JA, Taing L, *et al*. Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol*. Aug 22 2011;12(8):R83. doi:10.1186/gb-2011-12-8-r83
23. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. *Bioinformatics*. Apr 1 2011;27(7):1017-8. doi:10.1093/bioinformatics/btr064
24. Bailey TL, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res*. Jul 1 2015;43(W1):W39-49. doi:10.1093/nar/gkv416
25. van Dijk D, Sharma R, Nainys J, *et al*. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell*. Jul 26 2018;174(3):716-729 e27. doi:10.1016/j.cell.2018.05.061
26. Schindelin J, Arganda-Carreras I, Frise E, *et al*. Fiji: an open-source platform for biological-image analysis. *Nature Methods*. 2012-07-01 2012;9(7):676-682. doi:10.1038/nmeth.2019
27. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*. 2012-07-01 2012;9(7):671-675. doi:10.1038/nmeth.2089