Stem Cell Reports, Volume *17*

Supplemental Information

TGFβ superfamily signaling regulates the state of human stem cell pluripotency and capacity to create well-structured telencephalic organoids

Momoko Watanabe, Jessie E. Buth, Jillian R. Haney, Neda Vishlaghi, Felix Turcios, Lubayna S. Elahi, Wen Gu, Caroline A. Pearson, Arinnae Kurdian, Natella V. Baliaouri, Amanda J. Collier, Osvaldo A. Miranda, Natassia Dunn, Di Chen, Shan Sabri, Luis de la Torre-Ubieta, Amander T. Clark, Kathrin Plath, Heather R. Christofk, Harley I. Kornblum, Michael J. Gandal, and Bennett G. Novitch

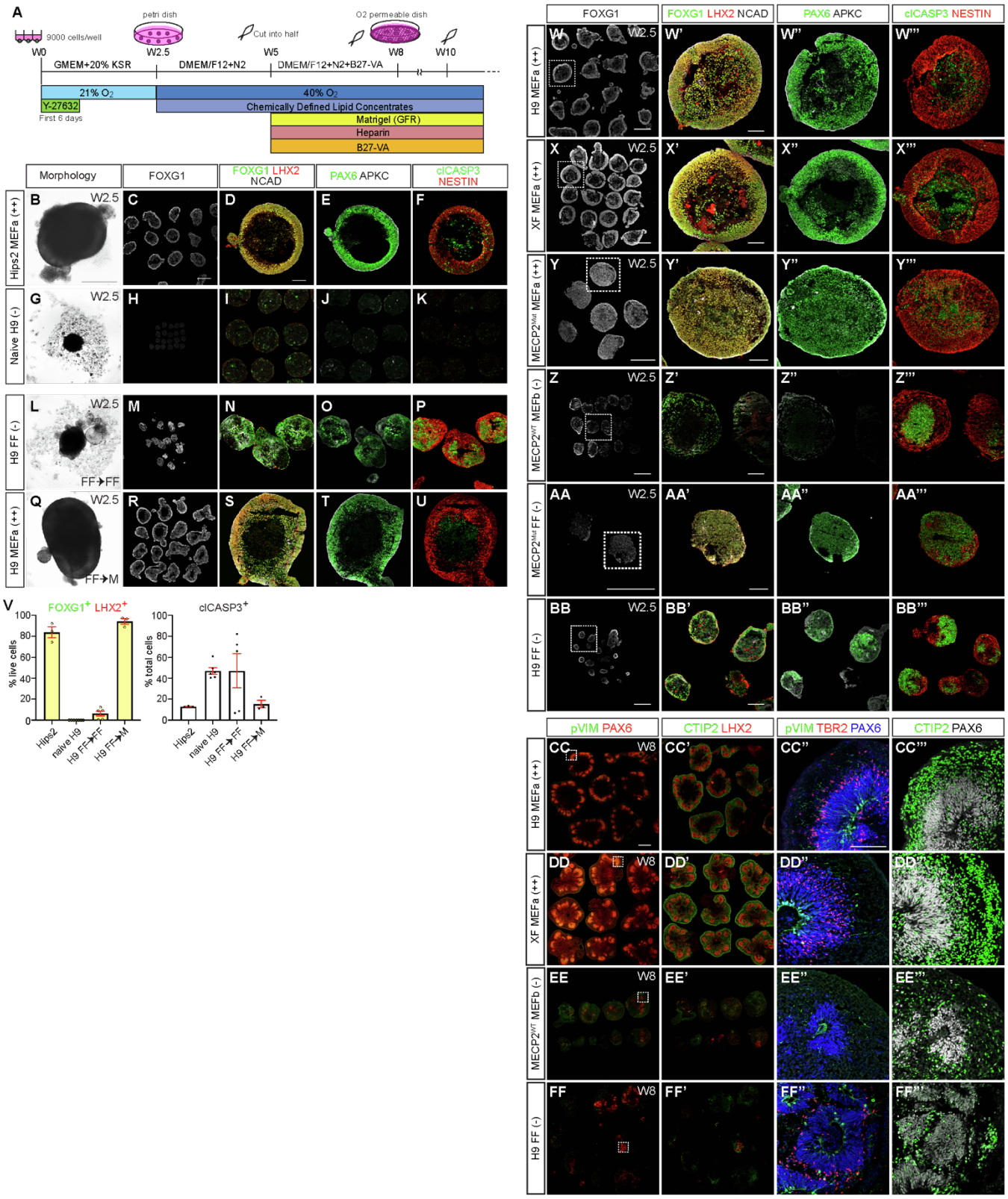

Figure S1 (related to Figures 1 and 2). The quality of cortical organoids can markedly vary according to how hPSC are maintained, and irregularities seen at early stages are associated with poorly structured organoids at later time points.

(A) Outline of the cortical organoid differentiation protocol used in most experiments, as previously described (Watanabe et al., 2017). (B-F) Hips2 hiPSCs cultured under MEFa and KSRa (MEFa) conditions successfully differentiated into cortical organoids similar to those derived from H9 hESCs maintained under the culture

conditions shown in Figure 1C-1G. (G-K) Naïve H9 cells displayed extremely poor differentiation into cortical organoids. (L-U) H9 hESCs maintained under feeder-free conditions were adapted to MEFa-supported conditions. MEFa-adaptation enabled the formerly non-competent H9 to efficiently create high quality organoids. The effects were similar to the positive effects of MEFa adaptation for feeder-free XFiPSCs as shown in Fig. 2. (V) Quantification of the percentage of $FOXG1^+$ $LHX2^+$ cortical progenitor cells out of total live cells per organoid (n = 3-7 per condition, over 1,200-3,000 cells counted per condition, from 3 different batches of organoids, 1-3 organoids per batch) and $clCASP3^+$ cells out of total cells (n = 3-6 organoids per condition, over 3,000 cells counted per condition, from 3 different batches of organoids, 1-2 organoids per batch). Data are represented as mean ± SEM. (W-Y'') Additional examples of organoids derived from different hPSC maintained under MEFa conditions showing robust expression of forebrain/cortical progenitor markers (FOXG1, LHX2, PAX6, NESTIN), consistent size, and well-formed and polarized neuroepithelium reflected by the circumferential expression of the apical membrane proteins NCAD and APKC (PKCζ). Organoids generated from MEFa-supported H9, XFiPSC, and MECP2 mutant Rett syndrome patient iPSC (Samarasinghe et al., 2021) typically showed over 80% positivity for both FOXG1 and LHX2 as well as other cortical markers such as PAX6. (Z-BB") By contrast, cortical organoids derived from H9, XFiPSC, Rett Syndrome iPSC maintained under either mTeSR1 feeder-free or suboptimal feeder-supported conditions (referred to as MEFb), fewer than 30% of cells were positive for FOXG1 and LHX2 or PAX6. Results for XFiPSC are presented in Figure 2. (CC-DD''') At W8, cortical organoids that came from batches that displayed high quality features at the W2.5 time point continued to exhibit a high percentage of cells positive for cortical progenitor markers including PAX6 and LHX2, well-defined VZ, SVZ, IZ, CP layers demarcated by PAX6, TBR2, and CTIP2 staining, and abundant formation of dividing bRG cells in the SVZ region marked by pVIM staining. (EE-FF''') In comparison, batches of cortical organoids that appeared of subpar quality at W2.5 typically displayed patchy expression of cortical progenitor markers, poorly defined and thinner SVZ, IZ, and CP layers, and reduced numbers of pVIM positive cells in the SVZ regions at W8. Scale bars: 500 μm (B, C, W, X, Y, Z, AA, BB, CC), 100 μm (D, W', X', Y', Z', AA', BB', CC"). Dotted boxes indicate regions in the lower magnification images that are enlarged in the right-hand panels. Designation of (++. +, and -) next to cell line names labels indicates the assessed organoid quality according to the rubric in Figure 1S.

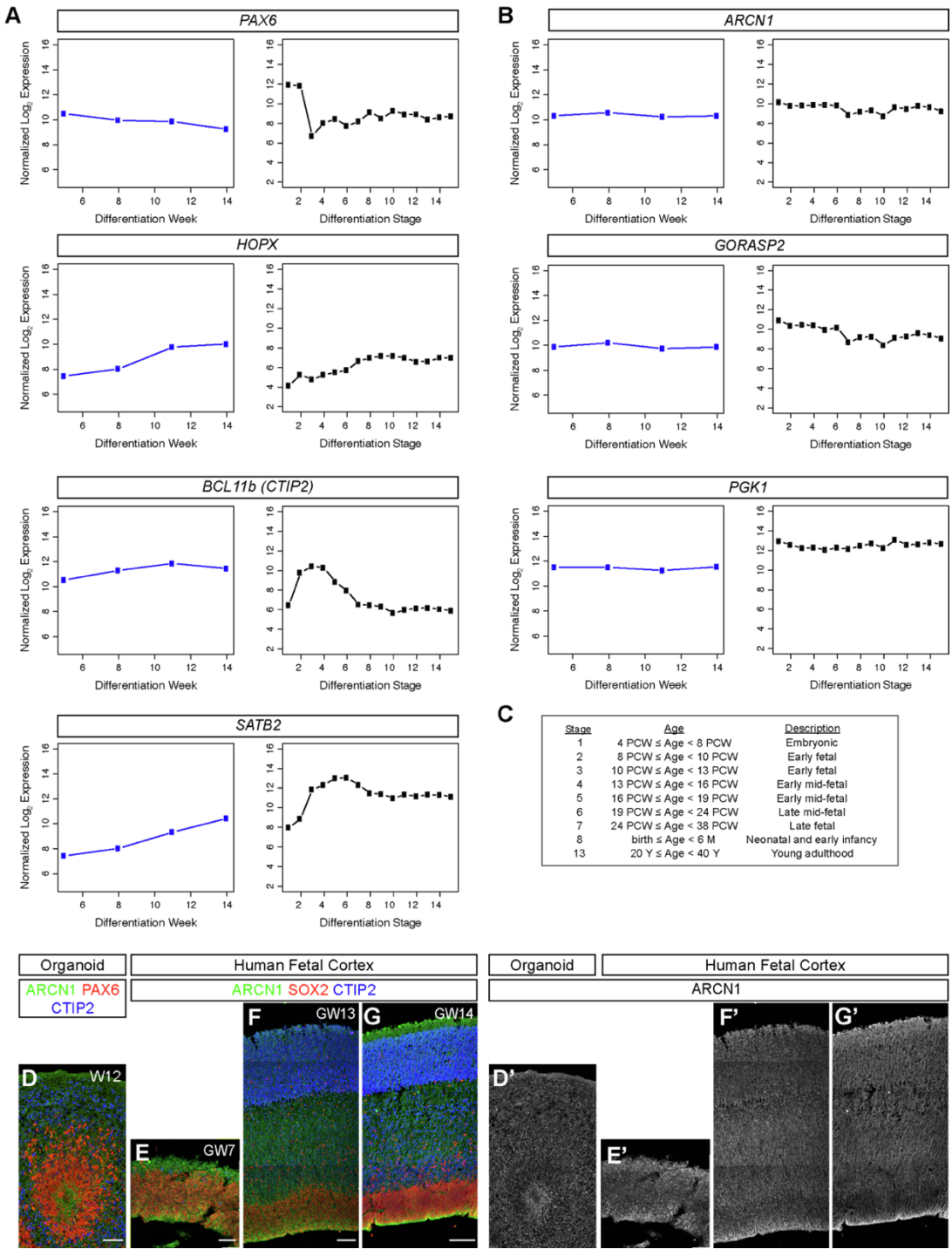

Figure S2 (related to Figures 1 and S1). Cortical organoids derived from MEFa-supported hESC show comparable expression of genes associated with metabolic stress to the developing human brain in vivo. (A-B) Comparative microarray analysis of gene expression across developmental time in brain organoids (represented by weeks in culture) and reference datasets for the developing human brain (represented as developmental stage as outlined in panel C. Each datapoint in the organoids represents the average from triplicate groups of organoids collected at the indicated time points. Note the similarity in both the expression levels of the selected genes and trajectory of expression across development in both the organoid and fetal brain reference sets. (D-G") Immunohistochemical staining shows comparable expression of the metabolic stress-associated marker ARCN1 in a representative cortical organoid (W12 example shown) and the developing human fetal cortex at GW7, 13, and 14. ARCN1 staining is present in multiple cell types including ventricular and outer radial glial cells (demarcated by PAX6 or SOX2 staining) and differentiated neurons (deep layer subset marked by CTIP2 staining). Scale bars: 50 μm (D, E) and 100 μm (F-G).

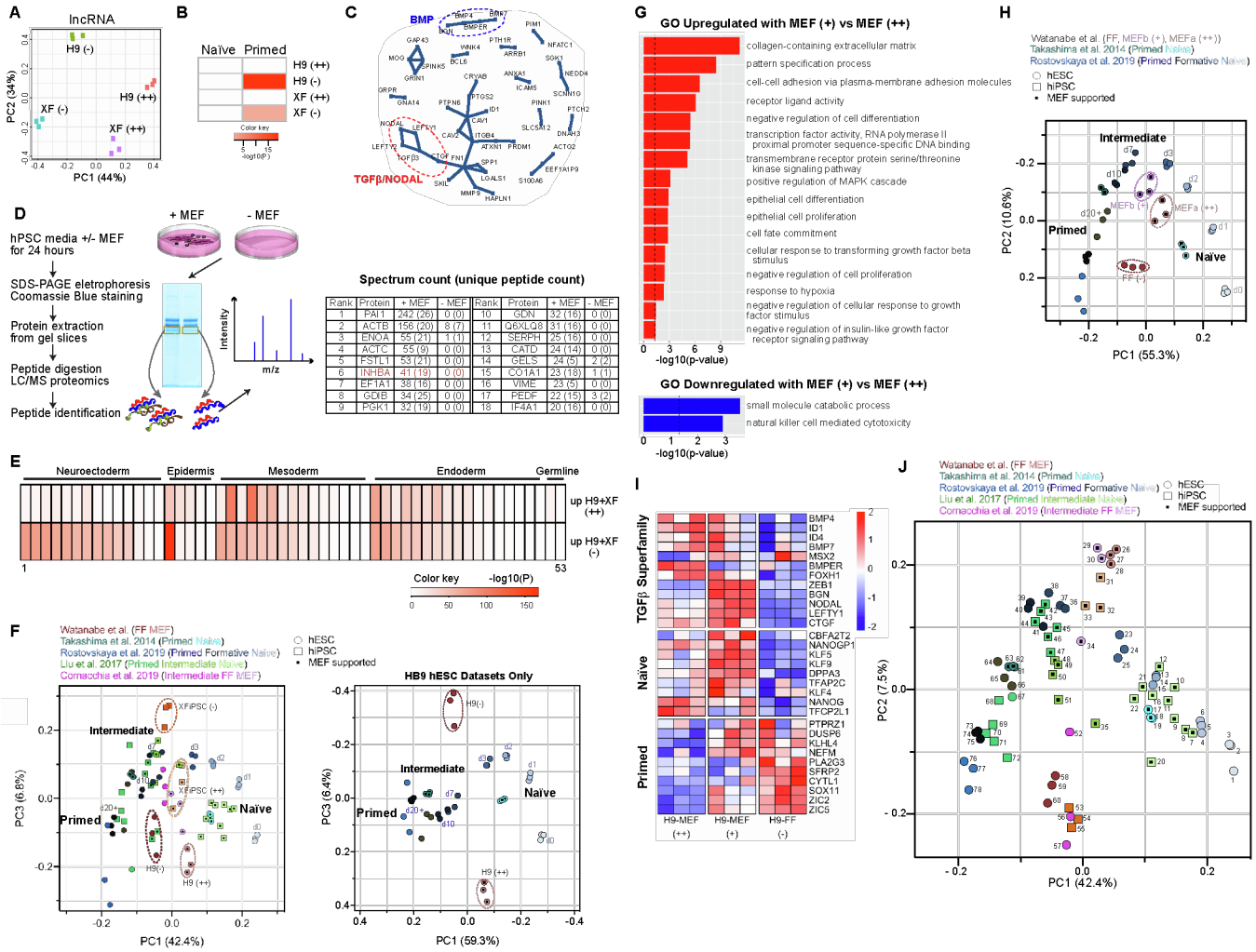

Figure S3 (related to Figure 3). Transcriptomic analysis of hPSCs under different culture conditions, MEF-secreted proteins identified by mass spectrometry, and sample key for principal component analyses. (A) Principal component analysis of lncRNAs expressed by H9 hESCs and XFiPSCs maintained under MEFa-supported conditions associated with effective cortical organoid formation (++) or feeder-free conditions associated with poor organoid formation (-). Triplicate samples are displayed. (B) Cell Marker enrichment analysis ($p<0.001$) using genes associated with naïve or primed hPSCs. (C) Protein-protein interaction analysis using Disease Association Protein-Protein Link Evaluator (DAPPLE). (D) Schematic of proteomic analysis and list of MEF-secreted proteins identified by mass spectrometry. (E) Cell marker enrichment analysis ($p<0.001$) using genes associated with mature cell lineages. A list of the cell lineages displayed (1-53) is provided in Table S5. (F) Principal component analysis integrating previously published data sets with primed, formative/intermediate, and naïve hiPSCs (Takashima et al., 2014; Liu et al., 2017; Cornacchia et al., 2019; Rostovskaya et al., 2019), showing PC1 vs PC3 related to the PC1 vs PC2 plot shown in Fig. 3F. Both x (PC1) and y (PC3) axes were reversed in the right plot to best align datapoint to the left plot. (G) Gene ontology analysis of the upregulated and downregulated transcripts in MEFa -supported competent hPSCs, H9 (++), compared to MEFb-supported semi-competent hPSCs, H9 (+). (C) Principal component analysis illustrating the differences in the overall transcriptional state of MEFb- vs. MEFa-supported H9 cells and other reference datasets used in Figure 3F. The y (PC2) axis was reversed to best match the representation of the datasets shown in panel J and other PCA plots in the study. (I) Heat map representation of genes associated with TGFβ superfamily signaling, naïve pluripotency, and primed pluripotency. The color of the tiles indicates scaled regressed gene expression data. (J) Annotated version of the plot shown in Figure 3F revealing the identity of individual hPSC samples according to the list provided in Table S4.

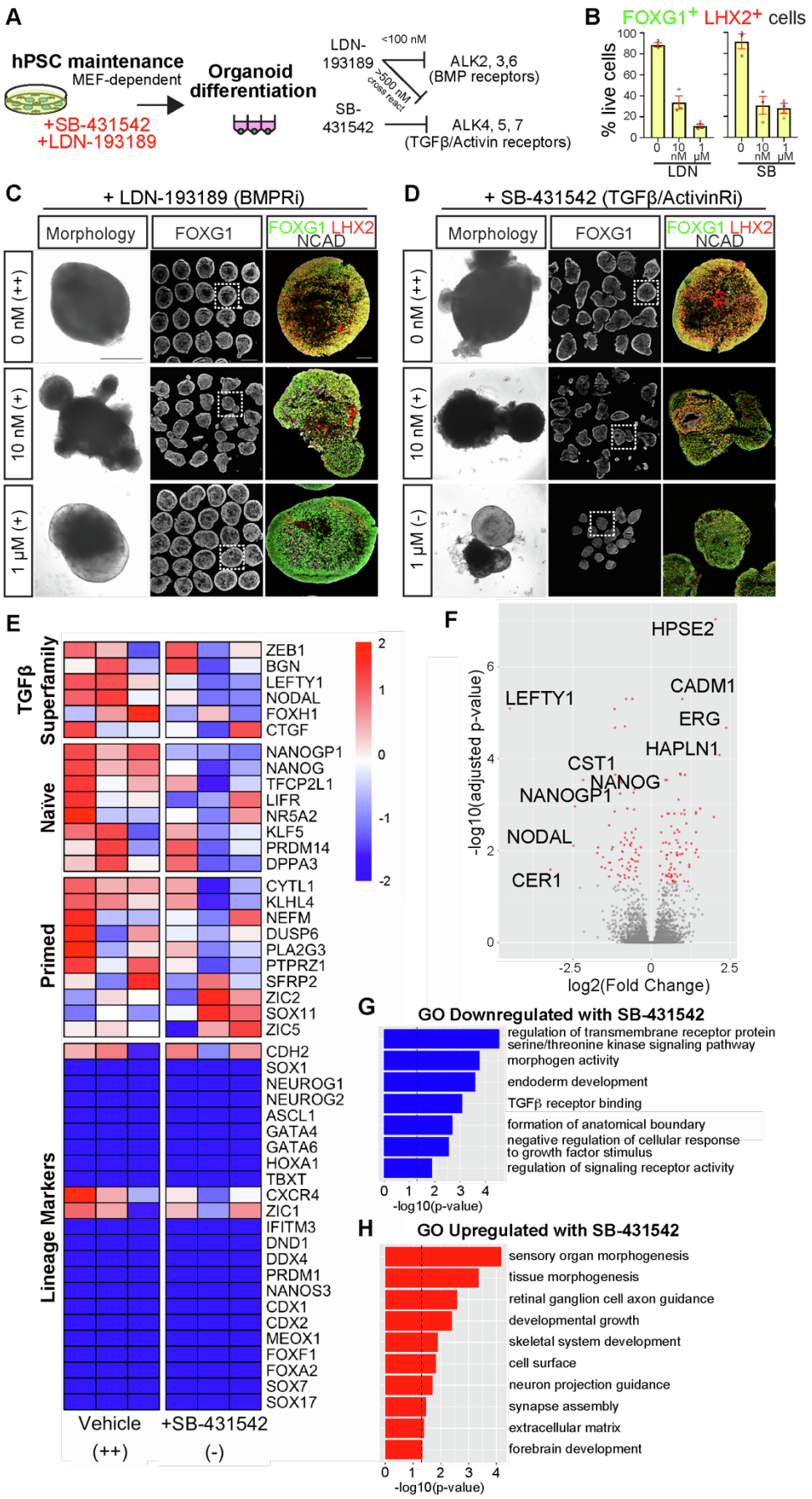
A
hPSC maintenance
MEF-dependent
+SB-431542
+LDN-193189
Organoid differentiation
LDN-193189
<100 nM
>500 nM cross react
ALK2, 3,6 (BMP receptors)
SB-431542
ALK4, 5, 7 (TGFβ/Activin receptors)
B
FOXG1+ LHX2+ cells
% live cells
LDN
SB
C
+ LDN-193189 (BMPRi)
Morphology
FOXG1
FOXG1 LHX2 NCAD
0 nM (++)
10 nM (+)
1 μM (+)
D
+ SB-431542 (TGFβ/ActivinRi)
0 nM (++)
10 nM (+)
1 μM (-)
E
TGFβ Superfamily
ZEB1
BGN
LEFTY1
NODAL
FOXH1
CTGF
Naive
NANOGP1
NANOG
TFCP2L1
LIFR
NR5A2
KLF5
PRDM14
DPPA3
Primed
CYTL1
KLHL4
NEFM
DUSP6
PLA2G3
PTPRZ1
SFRP2
ZIC2
SOX11
ZIC5
Lineage Markers
CDH2
SOX1
NEUROG1
NEUROG2
ASCL1
GATA4
GATA6
HOXA1
TBXT
CXCR4
ZIC1
IFITM3
DND1
DDX4
PRDM1
NANOS3
CDX1
CDX2
MEOX1
FOXF1
FOXA2
SOX7
SOX17
Vehicle (++)
+SB-431542 (-)
F
HPSE2
CADM1
LEFTY1
ERG
HAPLN1
CST1
NANOG
NANOGP1
NODAL
CER1
-log10(adjusted p-value)
log2(Fold Change)
G
GO Downregulated with SB-431542
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway
morphogen activity
endoderm development
TGFβ receptor binding
formation of anatomical boundary
negative regulation of cellular response to growth factor stimulus
regulation of signaling receptor activity
-log10(p-value)
H
GO Upregulated with SB-431542
sensory organ morphogenesis
tissue morphogenesis
retinal ganglion cell axon guidance
developmental growth
skeletal system development
cell surface
neuron projection guidance
synapse assembly
extracellular matrix
forebrain development
-log10(p-value)

Figure S4 (related to Figure 3). TGFβ signaling in hPSCs is required for optimal cortical organoid formation.

(A) Schematic of the experimental design. LDN-193189 or SB-431542 were added to hPSC cultures for five days before the start of the organoid differentiation procedure. LDN-193189 selectively inhibits BMP receptors when used at concentrations below 100 nM, but becomes less selective and also inhibits TGFβ and ACTIVIN receptors when used above 500 nM (Yu et al., 2008). SB-431542 is a selective inhibitor for TGFβ and ACTIVIN receptors (Inman et al., 2002). (B-D) Immunohistochemical assessment of cortical organoid formation at W2.5 from MEFa-supported organoid-competent hPSCs with and without SB-431542 (C) or LDN-193189 (D) application. Plot in (B) indicates the number of $FOXG1^{+}$ $LHX2^{+}$ cells out of total live cells. Scale bars: 500 μm, morphology and FOXG1 images; 100 μm FOXG1 LHX2 NCAD images. (E) Heat map showing the expression of genes associated with naïve pluripotency, primed pluripotency, generic pluripotency, and specific differentiated cell lineages. The color of the tiles indicates scaled regressed gene expression data. (F) Volcano plot displaying the fold change of gene expression (x-axis) and adjusted p-value (y-axis). Some naïve factors and TGFβ signaling molecules were significantly downregulated with the SB-431542 treatment. (G-H) Gene ontology analysis of the upregulated (G) and downregulated (H) transcripts in SB-431542 treated hPSCs.

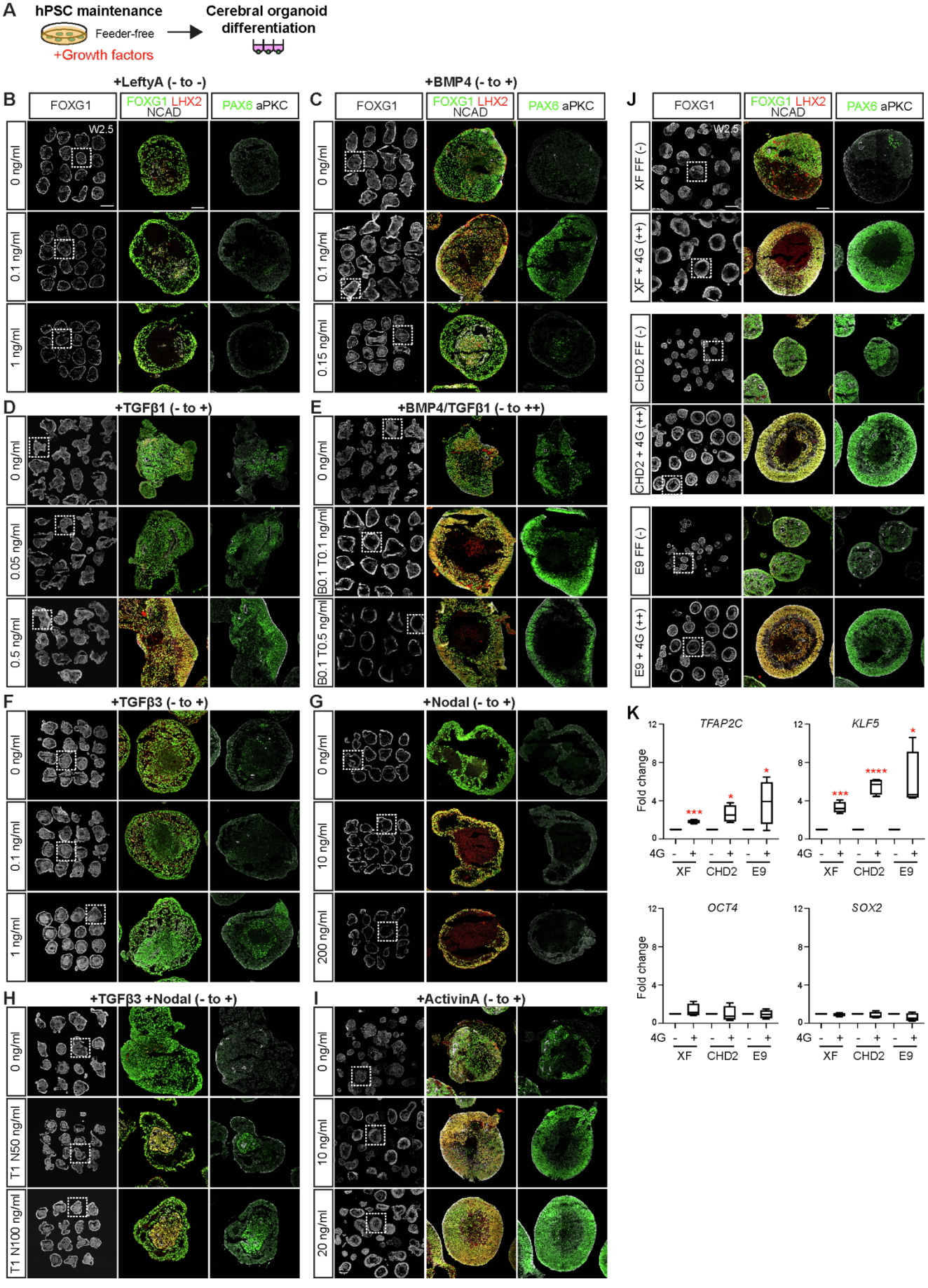
A
hPSC maintenance
Feeder-free
+Growth factors
Cerebral organoid differentiation
+LeftyA (- to -)
+BMP4 (- to +)
B
C
J
FOXG1
FOXG1 LHX2 NCAD
PAX6 aPKC
W2.5
0 ng/ml
0.1 ng/ml
1 ng/ml
0.15 ng/ml
XF FF (-)
XF + 4G (++)
CHD2 FF (-)
CHD2 + 4G (++)
E9 FF (-)
E9 + 4G (++)
D
+TGFβ1 (- to +)
E
+BMP4/TGFβ1 (- to ++)
0.05 ng/ml
0.5 ng/ml
B0.1 T0.1 ng/ml
B0.1 T0.5 ng/ml
F
+TGFβ3 (- to +)
G
+Nodal (- to +)
10 ng/ml
200 ng/ml
H
+TGFβ3 +Nodal (- to +)
T1 N50 ng/ml
T1 N100 ng/ml
I
+ActivinA (- to +)
20 ng/ml
K
TFAP2C
KLF5
OCT4
SOX2
Fold change
4G
XF
CHD2
E9

Figure S5 (related to Figures 3, 4, 5, and S3). Individual or pairs of TGFβ superfamily growth factors to feeder-free hPSCs can partially rescue cortical organoid formation, whereas full enhancement can be achieved in multiple feeder-fee hPSC lines with the 4G mixture.

(A) Schematic of the experimental design. Organoid non-competent feeder-free H9 were preconditioned for five days with TGFβ superfamily growth factors found to be associated with organoid-competent hPSCs through the transcriptomic and proteomic analyses conducted in Figures 3 and S3. (B-I) Representative examples of W2.5 organoids differentiated from H9 hESC maintained under feeder-free conditions without or with LEFTYA, BMP4, TGFβ1, BMP4/TGFβ1, TGFβ3, NODAL, TGFβ3/NODAL, or ACTIVIN A as indicated. Supplementation with most TGFβ superfamily molecules partially rescued cortical organoid production with the exception of LEFTYA. All concentrations displayed are in ng/ml. Combinations referred to as B0.1 T0.1 designates 0.1 ng/ml BMP4 and 0.1 ng/ml TGFβ. See also Table S1 for the average scores and ratings for the organoids produced in these experiments. (J) By contrast, supplementation with a four-growth factor mixture (4G)- BMP4, TGFβ1, ACTIVIN A, and TGFβ3, enhanced cortical organoid formation from diverse sets of feeder-free hPSC: XFiPSCs, *CHD2* mutant hiPSCs, and E9 healthy patient hiPSCs (Tidball et al., 2017). (K) RT-qPCR analyses showing the upregulation of *TFAP2C* and *KLF5* with the 4G application. The general pluripotency markers *OCT4* and *SOX2* were unchanged. Expression levels are normalized to feeder-free hiPSCs without 4G addition. Data are represented as mean ± SEM. n ≥ 4 independent experimental batches, each experimental replicate is an average of 3 technical replicates. Scale bars: 500 μm FOXG1 images; 100 μm FOXG1 LHX2 NCAD images.

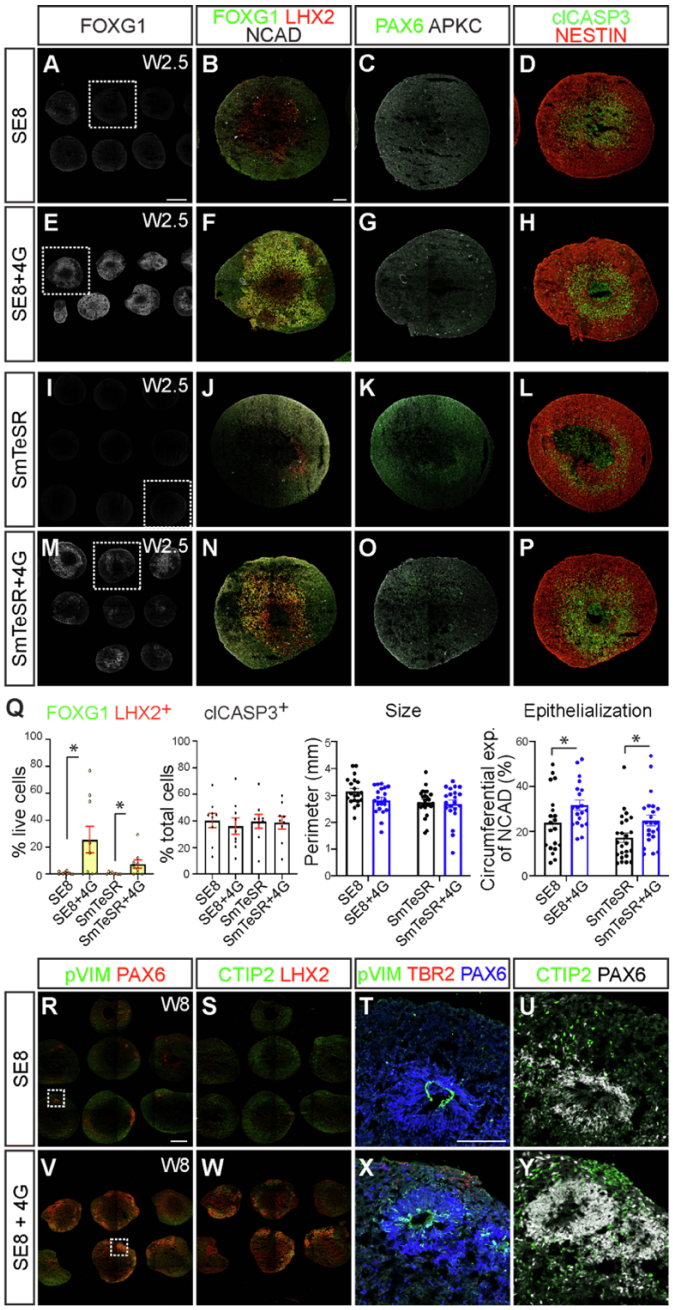

Figure S6 (related to Figure 5). 4G supplementation can enhance organoid quality using different protocols.

The effects of 4G supplementation on organoid formation were tested using the widely used Spheroid protocol (Yoon et al., 2019), which utilizes feeder-free hPSCs. Two hPSC culture conditions were utilized to form cortical spheroids: (A-H) E8-based media with vitronectin substrate (SE8) and (I-P) mTeSR1-based media with ES-qualified Matrigel substrate (SmTeSR). The formation of forebrain and cortical progenitor formation was assessed using immunostaining for FOXG1, LHX2, PAX6, and NESTIN and integrity of the apical membrane as a gauge of epithelialization using NCAD and aPKC staining. (Q) Quantification of $FOXG1^{+}LHX2^{+}$ cells out of total live cells (n = 9 organoids total per condition, 3 independent experimental batches, 3 organoids per batch, over 4000 cells per condition were counted), clCASP3 as a measure of cell death (n = 9 organoids total, 3 independent experimental batches, 3 organoids for each condition per batch, over 4000 cells/condition were counted), organoid epithelialization (n = 18 organoids, 3 independent experimental batches, 6 organoids for each condition per batch), and organoid perimeter (n = 18 organoids, 3 independent experimental batches, 6 organoids for each condition per batch). Scale bars: 500 μm (A, R), 100 μm (B, T). (R-Y) Representative examples of W8 cortical spheroids generated in the batches that were analyzed at W2.5 as described above. Results here should be compared to those analyzed in Figures 5E and S1CC-FF'''.

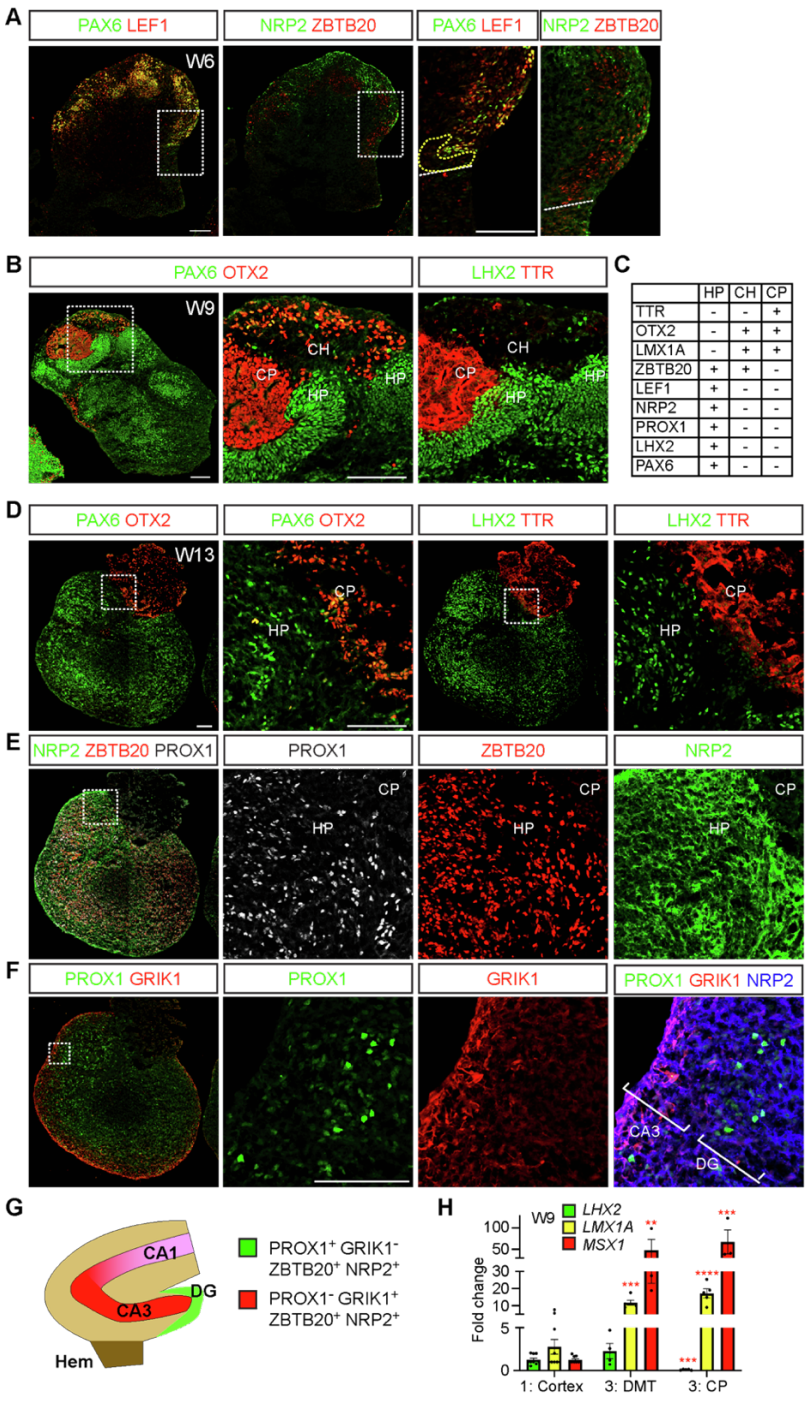

Figure S7 (related to Figure 7). hPSCs cultured under 4G feeder-free conditions can form dorsomedial telencephalic organoids.

(A) Representative examples of W6 dorsomedial telencephalic (DMT) organoids stained for hippocampal primordium (HP) markers such as PAX6, LEF1, NRP2, and ZTBT20. Images shown are adjacent sections to Figure 7F. (B) W9 DMT organoids stained for HP (PAX6 and LHX2), cortical hem (CH, OTX2), and choroid plexus (CP, OTX2 and TTR) markers. (C) List of HP, CH, and CP markers present in the organoids. (D-F) Representative examples of W13 DMT organoids. All images are from adjacent sections immunostained for CH (OTX2), CP (OTX2 and TTR), and HP (PAX6, LHX2, NRP2, ZBTB20, PROX1, and Kainate Receptor GluK1 (GRIK1) markers. GRIK1 staining, which is associated with CA3 axons in vivo, was routinely found adjacent to dentate gyrus (DG)-like regions demarcated by PROX1 expression. (G) Schematic of the organization of the fetal hippocampus in vivo, illustrating the relative positions of the DG, CA3, and CA1 regions. (H) RT-qPCR analyses of DMT markers. Expression levels are normalized to organoids without BMP4 and CHIR application. Data are represented as mean ± SEM. n ≥ 3 independent experimental replicates. All scale bars: 100 μm.

Table S1: Average scores and ratings for different culture conditions and cell lines, related to Figures 1, 2, 4, 5, S1, S4, and S5							
Cell Line	hESC/ hiPSC?	Culture Condition	Experimental Replicates	Total # Organoids	Passage Numbers	Avg. Score	Avg. rating
H9	hESC	MEFa	7	96	p50~65	8.86	++
H9	hESC	MEFb	3	58	p39-41	6.00	+
H9	hESC	FF	17	264	p50~65	3.59	-
H9	hESC	Naïve MEFc	3	78	p48-50	3.00	-
H9	hESC	LEFTYA (1 ng/ml)	3	39	p50~65	4.00	-
H9	hESC	BMP4 (0.1 ng/ml)	3	44	p50~65	6.67	+
H9	hESC	TGFβ1 (0.5 ng/ml)	3	40	p50~65	6.67	-
H9	hESC	BMP4 (0.1 ng/ml)/TGFβ1 (0.1 ng/ml)	6	85	p50~65	5.83	+
H9	hESC	TGFβ3 (1 ng/ml)	3	44	p50~65	5.00	+
H9	hESC	Nodal (200 ng/ml)	3	45	p50~65	5.00	+
H9	hESC	TGFβ1 (1 ng/ml)/Nodal (50 ng/ml)	3	48	p50~65	6.00	+
H9	hESC	ActivinA (20 ng/ml)	3	71	p50~65	7.33	+
H9	hESC	FF + #4 (+3G)	5	91	p50~65	5.80	+
H9	hESC	FF + #7 (+4G)	5	59	p50~65	8.40	++
H9	hESC	FF + #8 (+4G)	3	35	p50~65	8.00	++
H9	hESC	FF + #9 (+5G)	4	46	p50~65	7.50	+
Hips2	hiPSC	MEFa	8	96	p55~65	8.25	++
XF	hiPSC	FF	5	88	p9~25	3.80	-
XF	hiPSC	MEFa	12	212	p9~25	8.25	++
XF	hiPSC	LDN (10 nM)	3	65	p9~25	5.33	+
XF	hiPSC	LDN (1 μM)	3	65	p9~25	7.00	+
XF	hiPSC	SB (10 nM)	3	56	p9~25	5.66	+
XF	hiPSC	SB (1 μM)	3	37	p9~25	3.33	-
XF	hiPSC	FF E8	3	46	p9~25	3.33	-
XF	hiPSC	FF + 4G (#7)	4	51	p9~25	8.00	++
E9	hiPSC	FF	6	54	p19~25	3.00	-
E9	hiPSC	FF + 4G (#7)	4	62	p19~25	8.00	++
CHD2	hiPSC	FF	7	110	p17~25	3.14	-
CHD2	hiPSC	FF + 4G (#7)	4	61	p17~25	8.25	++

Table S2. Top 300 genes associated with cerebral organoid competency, related to Figure 3.
(see Excel file)

Table S3. Top 300 genes associated with cerebral organoid non-competency, related to Figure 3.
(see Excel file)

Table S4. List of the hPSC transcriptome data sets plotted in Figure 3F and S3J, also related to Figures 6E, S3F, and S3H.
(see Excel file)

Table S5: List of cell types, related to Figure S3E			
#	Description	Origin	Other origin
1	Brain - amygdala	Neuroectoderm	
2	Brain - anterior cingulate cortex	Neuroectoderm	
3	Brain - frontal cortex	Neuroectoderm	
4	Brain - substantia nigra	Neuroectoderm	
5	Brain - hippocampus	Neuroectoderm	
6	Brain - putamen/basal ganglia	Neuroectoderm	
7	Brain - spinal cord cervical	Neuroectoderm	
8	Brain - cortex	Neuroectoderm	
9	Brain - hypothalamus	Neuroectoderm	
10	Brain - nucleus accumbens/basal ganglia	Neuroectoderm	
11	Brain - caudate basal ganglia	Neuroectoderm	
12	Brain - cerebellar hemisphere	Neuroectoderm	
13	Brain - cerebellum	Neuroectoderm	
14	Nerve - tibial	Neural crest	Neuroectoderm
15	Cells transformed fibroblasts	Epidermis	
16	Skin - sun exposed lower leg	Epidermis	
17	Skin - not sun exposed suprapubic	Epidermis	
18	Breast - mammary tissue	Epidermis	
19	Pituitary	Ectoderm	
20	Adrenal gland	Mesoderm	Ectoderm
21	Artery - tibial	Mesoderm	
22	Cells EBV transformed lymphocytes	Mesoderm	
23	Muscle - skeletal	Mesoderm	
24	Artery - aorta	Mesoderm	
25	Whole blood	Mesoderm	
26	Adipose - visceral/omentum	Mesoderm	
27	Artery - coronary	Mesoderm	
28	Adipose - subcutaneous	Mesoderm	
29	Kidney - cortex	Mesoderm	
30	Uterus	Mesoderm	
31	Vagina	Mesoderm	
32	Cervix - endocervix	Mesoderm	
33	Cervix - ectocervix	Mesoderm	
34	Fallopian tube	Mesoderm	
35	Heart - left ventricle	Endoderm	
36	Esophagus - mucosa	Endoderm	
37	Liver	Endoderm	
38	Esophagus - muscularis	Endoderm	
39	Heart - artial appendage	Endoderm	
40	Esophagus - gastroesophageal junction	Endoderm	
41	Colon - sigmold	Endoderm	
42	Stomach	Endoderm	
43	Pancreas	Endoderm	
44	Colon - transverse	Endoderm	
45	Lung	Endoderm	
46	Bladder	Endoderm	Mesoderm
47	Spleen	Endoderm	
48	Minor salivary gland	Endoderm	
49	Small intestin - terminal ileum	Endoderm	
50	Thyroid	Endoderm	Neural crest
51	Prostate	Endoderm	
52	Ovary	Germline	
53	Testis	Germline	

Table S6. Key Resources, related to all Figures and Tables.
(see Excel File)

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

hPSC maintenance culture

hPSC experiments were conducted with prior approval from the University of California Los Angeles (UCLA). H9 hESC (Thomson et al., 1998), XFiPSC (line XiPSC2; Karumbayaram et al., 2012), and Hips2 (Lowry et al., 2008) lines were obtained from the UCLA Broad Stem Cell Research Center Core. UCLA1 hESCs harboring homozygous mutation of *TFAP2C* and a dox-inducible TFAP2C expression cassette were used as previously described (Diaz Perez et al., 2012; Chen et al., 2018; Pastor et al., 2018). hiPSC lines E9 (WT control) and *CHD2* heterozygous indel mutant (Tidball et al., 2017) were obtained by Dr. Jack Parent at the University of Michigan. WT and MECP2 mutant hiPSCs from a Rett syndrome patient were described in (Ohashi et al., 2018; Samarasinghe et al., 2021). hPSCs under MEF-supported conditions were maintained as previously described (Watanabe et al., 2017). 11 different MEF batches (8 purchased from Millipore, PMEF-CF, and 3 homemade batches) were used in our experiments. 5/11 (45.4%) of the MEF batches showed good performance, 3/11 (27.3%) were intermediate, and 3/11 (27.3%) were poor. Representative examples of good and intermediate MEF batches from Millipore were used in the experiments shown and designated MEFa for good, and MEFb for intermediate.

hPSCs under feeder-free conditions were maintained with mTeSR™1 (Stemcell Technologies, 85850) or Essential 8 medium (E8, ThermoFisher, A15117001) and where specified on hESC-qualified Matrigel substrate (Fisher Scientific, 08-774-552). Every 3-5 days, hPSCs were passaged at a 1:10-1:15 dilution with partial dissociation using ReLeSR (Stemcell Technologies, 5872). For the 4G method, hPSCs were preconditioned with growth factors for 3-4 days one day after passage. Growth factor concentrations were as follows: BMP4 (0.1 ng/ml, Invitrogen, PHC9534), TGFβ1 (0.1 ng/ml, R&D Systems, 240-B), ACTIVIN A (10-20 ng/ml, Peprotech, 120-14P), and NODAL (50 ng/ml, R&D Systems, 3218-ND) or TGFβ3 (1 ng/ml, R&D Systems, 8420-B3).

Primed hPSC maintenance, primed to naïve conversion, and naïve hPSC maintenance

Primed hPSCs were maintained in primed media consisting of 20% KSR in DMEM/F12 supplemented with 1x nonessential amino acids, 2 mM L-Glutamine, 0.5x Penicillin/Streptomycin (all from Invitrogen), 0.1 mM β-mercaptoethanol (Sigma-Aldrich), and 4 ng/ml FGF2 (Peprotech). Primed hPSCs were cultured on CF-1 irradiated MEFs and passaged every 4-5 days with collagenase IV (ThermoFisher Scientific) at 37°C for 5 minutes, followed by manual dissociation by pipetting.

Primed hPSCs were converted to naïve H9 as previously described (Guo et al., 2017). Briefly, primed hPSCs were dissociated into single cells with Accutase and $2x10^5$ cells per 6-well were plated in primed media with 10 μM Y-27632 onto MEFs seeded at a density of $2x10^6$ cells per 6-well plate. The following day (day 1), media was changed to cRM-1, which consists of N2B27 basal media supplemented with 1 μM PD0325901 (Cell Guidance Systems), 20 ng/ml human LIF (Millipore) and 1 mM Valproic Acid (Sigma-Aldrich). On day 4, the media was switched to cRM-2 consisting of N2B27 basal media supplemented with 1 μM PD0325901 (Cell Guidance Systems), 20 ng/ml human LIF (Millipore), 2 μM Gö6983 (Tocris) and 2 μM XAV939. From day 11 onwards, converted naïve cells were cultured in t2iLGö media. Cells were passaged on day 5, day 10, and every 4-5 days subsequently. Homogenous naïve hPSC lines were obtained after 4 passages in t2iLGö media.

Naïve hPSCs were subsequently maintained in t2iLGö media (Takashima et al., 2014) consisting of a 1:1 mixture of DMEM/F12 and Neurobasal, 0.5% N2 supplement, 1% B27 supplement, 1x nonessential amino acids, 2 mM L-Glutamine, 0.5x Penicillin/Streptomycin (all from ThermoFisher Scientific), 0.1 mM β-mercaptoethanol (Sigma-Aldrich) (N2B27 basal media) supplemented with 1 μM PD0325901 (Cell Guidance Systems), 1 μM CHIR99021 (Cell Guidance Systems), 20 ng/ml human LIF (Millipore) and 2 μM Gö6983 (Tocris) on CF-1 irradiated MEFs. Naïve hPSCs were passaged every 4 days with Accutase (ThermoFisher Scientific) at 37° C for 5 minutes. Naïve hPSCs between passages 5-7 were used for organoid experiments.

All hPSC were cultured in 21% O_2, 5% CO_2 at 37°C and subject to daily media changes.

Human fetal tissue

Experiments were performed with prior approval from the research ethics committees at the UCLA office of the Human Research Protection Program and the University of Tübingen (institutional review board [IRB] #323/2017BO2) and Novogenix Laboratories. Embryonic tissues were obtained with informed consent as discarded materials resulting from elective, legal terminations. Samples were de-identified in accordance with institutional guidelines. Specimen ages for this study are denoted as gestational weeks, as determined by the date of the last menstrual period or ultrasound and confirmed by analysis of developmental characteristics.

Tissue processing and immunohistochemistry
Brain organoids were fixed, cryoprotected, embedded, frozen, and cryosectioned as previously described (Watanabe et al., 2017). Sectioned tissues were collected onto Superfrost Plus slides (Fisher Scientific) and blocked for 30 minutes in PBS with 1% heat inactivated equine serum (Hyclone), 0.1% Triton X-100, and 0.01% sodium azide and incubated in primary antibodies (see Key Resources Table) in the blocking solution overnight at 4°C. After three washes in PBST (0.1% Triton X-100), tissue was incubated with secondary antibodies for one hour at room temperature. After three washes, tissue was mounted in ProLong® Diamond (Invitrogen) with coverslips and stored in the dark at 4°C prior to imaging. GW13-14 human fetal brain tissue was fixed in 4% PFA in PBS for 3-4 days at 4 °C. GW 7 human fetal brain tissue was fixed overnight at 4 °C. Brain tissue was then washed twice in PBS and immersed in 30% sucrose in PBS overnight at 4 °C. The tissue was lastly frozen in O.C.T. and cryosectioned at 20 µm thickness.

Microscopic imaging
Confocal images were acquired using a Zeiss LSM 780 or 800 microscope equipped with a motorized stage and Zen black or blue software. Tiled images were assembled using the Zen Tiles with the multi-focus function. For brightfield imaging, a Zeiss Axio Observer D1 microscope was used. All images were compiled in Adobe Photoshop or Fiji (Schindelin et al., 2012), with image adjustment applied to the entire image and restricted to brightness, contrast, and levels.

RNA-sequencing data analyses
STAR (version 2.4.0j; (Dobin et al., 2013)) was used to align RNA reads to the human genome (GRCh37/hg19). All samples had greater than 80% read alignment. This genome version was also used for subsequent read quantification (exon counts) with HTSeq (version 0.6.1p1; (Anders et al., 2015)). QC statistics were obtained per sample using Picard tools (http://broadinstitute.github.io/picard) to be used in downstream analysis to account for technical variation in gene expression data. QC statistics collected included: AT/GC dropout, read duplication rate, GC bias, read depth, percentage of different genomic regions covered (exons, introns, UTRs, etc.), and 5' end sequencing bias.

Differential gene expression was performed in R (https://www.r-project.org/) as follows using HTSeq exon and lncRNA counts. First, genes were filtered such that only genes with a count greater than 10 in at least 80% of samples were retained. The DESeq2 package (Love et al., 2014) was then used to both obtain normalized gene expression data (using varianceStabilizingTransform) and to calculate differentially expressed genes between conditions of interest. Based on the association of Picard QC statistics and other covariates with top normalized gene expression principal components, we included the covariates of condition (hPSC type and feeder type), RIN, RNA concentration, and the first principal component of the Picard sequencing statistics as linear model covariates during differential gene expression calculation. Groups of differentially expressed genes were identified as either significantly up- or down-regulated between conditions with a false discovery rate (FDR) < 5%.

Groups of differentially expressed genes between conditions were subjected to the following enrichment analyses. Gene ontology analysis was performed using Metascape online software (http://metascape.org) (Tripathi et al., 2015). We also used cell type markers of the naïve and primed states (Sahakyan et al., 2017) to find any enrichment in groups of differentially expressed genes using the pSI R package (Xu et al., 2014). Protein-protein interaction enrichment was established using DAPPLE (Rossin et al., 2011). For Figure S3E, GTEx RNA-seq data (gene median RPKM data version 6) was utilized. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal on 10/26/17.

Several additional analyses were conducted with a regressed dataset, in which the effects of the following covariates were removed in a linear model of the normalized gene expression data: RIN, RNA concentration, and the first principal component of the Picard QC statistics. With this approach, we could explore a dataset that only retained the effects of the conditions (hPSC type and feeder type) and the linear model residual. Comparison of different sample conditions was achieved through principal component analysis of this regressed dataset. Next, to visualize how our samples related to different states of cell maturity (naïve, intermediate, and primed), we also combined our data with that of (Takashima et al., 2014; Liu et al., 2017; Cornacchia et al., 2019; Rostovskaya et al., 2019) using ComBat in the sva R package (Leek et al., 2012) and then conducted principal component analysis with this merged dataset. We also used hierarchical clustering to group samples based on the scaled expression of the top 150 genes up-regulated in the feeder condition and the top 150 genes up-regulated in the

feeder-free condition and visualized this clustering with a heatmap (Figure 3B). We performed similar clustering analyses and heatmap visualizations with SMAD signaling players, naïve, and primed factors (Figure 3C). Clustering and heatmap visualization was conducted with the NMF (Gaujoux and Seoighe, 2010) and the ggplot2 (Wickham, 2016) packages in R.

Quantitative PCR

Reverse Transcriptase qPCR (RT-qPCR) was performed as previously described (Watanabe et al., 2017). Briefly, total RNA was extracted using a RNeasy Mini or miRNeasy Mini Kit (Qiagen) and >500 ng of total RNA was used for cDNA synthesis for each sample, using the SuperScript IV First-Strand Synthesis System (Invitrogen). For RT-qPCR reaction, LightCycler 480 SYBR Green I Master Mix and exon-spanning primer pairs listed below were used with synthesized cDNA. All primer pairs were validated for 1.8 amplification efficiency as described (Watanabe et al., 2012). All reactions were performed using a Roche LightCycler 480 real-time PCR system in triplicates, and relative expression levels were determined by normalizing the crossing points to the internal reference gene β-ACTIN. Primers used are listed in the Key Resources Table.

Immunoblotting

For the positive and negative controls, UCLA1 hESCs harboring a homozygous *TFAP2C* deletion or a doxycycline-inducible *TFAP2C* expression cassette were used as previously described (Pastor et al., 2018). In all cases, hPSCs were first washed with cold PBS and then collected for protein in RIPA lysis buffer with protease inhibitors (ThermoFisher, 78425) and phosphatase inhibitors (ThermoFisher, 78420). RIPA buffer was added to the cells followed by 5 minutes incubation on ice. Cells were then scraped off from the culture dishes, transferred to a microcentrifuge tube, and rotated for 10 minutes at 4°C. If a lysate was too viscous, it was sonicated. Lysates were centrifuged at maximum speed for 15 minutes at 4°C and the supernatant collected and snap frozen in liquid nitrogen followed by storage at -80°C. Protein lysates were then mixed with loading dye containing β-mercaptoethanol and placed in a 95°C shaking heating block followed by centrifugation at maximum speed for 2 minutes. The heated lysates were run on SurePage gels (GenScript, M00652) in MOP buffer (GenScript, M00138) for 45 minutes to 1 hour at 120 V, and proteins then transferred onto 0.45 µm nitrocellulose membranes in Tris Glycine buffer for 2 hours on ice at 300 mA. Protein transfer membranes were blocked in 5% skim milk (Bio-Rad, 170-6404) for 1 hour and incubated with primary antibodies (rabbit anti TFAP2C/AP2 Abcam GR59885-7 at 1:1000 and goat anti GAPDH Abcam ab94583 at 1:1000) on a shaker at 4°C overnight. After three 5-minute washes, membranes were incubated with secondary antibodies conjugated with HRP in 5% skim milk on the shaker for 2 hours at room temperature. We used SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher, 34095) or Pierce ECL 2 Western Blotting Substrate (ThermoFisher, 80197) for chemiluminescent detection. Membranes were scanned using a Sapphire RGBNIR^{TM} Biomolecular Imager (Azure Biosystems, Inc.), and acquired digital images quantified using the ImageJ gel function. TFAP2C protein levels were normalized to GAPDH and reported as mean ± SEM for at least four biological replicates.

Mass spectrometry

Previous work demonstrated that protein factors secreted by MEFs might be responsible for reprogramming hPSC metabolism (Gu et al., 2016). To understand which protein factor(s) in the MEF-conditioned medium might impact the ability of hPSCs to effectively form cortical organoids, a mass spectrometry-based protein identification experiment was conducted as follows: Conditioned media was collected after 24 hours incubation with or without MEFs. The protein fractions of the samples were enriched by using an Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Millipore, UFC901024). Concentrated protein samples were then separated by SDS-PAGE and the gel was processed for Coomassie Brilliant Blue staining (Thermo Scientific) according to the manufacturer's instructions. Coomassie Blue-stained bands were cut from the gels, washed twice with 50% acetonitrile, and processed for liquid chromatography tandem mass spectrometry analysis at the mass spectrometry core facility at Beth Israel Deaconess Medical Center (Boston, MA).

QUANTIFICATION AND STATISTICAL ANALYSIS

The number and description of experimental replicates and statistical analyses performed are indicated in each figure legend. In this study, three or more independent experiments were performed for all analyses unless otherwise indicated. The average scores and ratings for different cell lines and conditions are shown in Table S1. The normality and equality of variance of each data set was determined using GraphPad Prism software (D'Agostino-Pearson and Anderson-Darling tests). Appropriate parametric or non-parametric were accordingly applied. Paired and unpaid Student's two tailed t tests with or without Welch's correction and two tailed Mann-Whitney tests were calculated using Prism software. Signifiers used are as follows: no significance (n.s.) $p \geq 0.05$,

* p < 0.05, ** p < 0.01, *** p < 0.001, **** < 0.0001. All data are represented as mean ± SEM unless otherwise specified in the figure legends.

RESOURCE AVAILABILITY

Lead contact

All requests for resources, reagents, and protocols should be addressed to and will be fulfilled by the lead contact, Bennett Novitch (bnovitch@ucla.edu).

Material availability

Unique materials and reagents generated in this study are available upon request from the Lead Contact.

SUPPLEMENTAL REFERENCES

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166-169.

Chen, D., Liu, W., Zimmerman, J., Pastor, W.A., Kim, R., Hosohama, L., Ho, J., Aslanyan, M., Gell, J.J., Jacobsen, S.E.*, et al.* (2018). The TFAP2C-Regulated OCT4 Naive Enhancer Is Involved in Human Germline Formation. Cell Rep *25*, 3591-3602 e3595.

Cornacchia, D., Zhang, C., Zimmer, B., Chung, S.Y., Fan, Y., Soliman, M.A., Tchieu, J., Chambers, S.M., Shah, H., Paull, D.*, et al.* (2019). Lipid Deprivation Induces a Stable, Naive-to-Primed Intermediate State of Pluripotency in Human PSCs. Cell Stem Cell *25*, 120-136 e110.

Diaz Perez, S.V., Kim, R., Li, Z., Marquez, V.E., Patel, S., Plath, K., and Clark, A.T. (2012). Derivation of new human embryonic stem cell lines reveals rapid epigenetic progression in vitro that can be prevented by chemical modification of chromatin. Hum Mol Genet *21*, 751-764.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21.

Gaujoux, R., and Seoighe, C. (2010). Algorithms and framework for nonnegative matrix factorization (NMF).

Gu, W., Gaeta, X., Sahakyan, A., Chan, A.B., Hong, C.S., Kim, R., Braas, D., Plath, K., Lowry, W.E., and Christofk, H.R. (2016). Glycolytic metabolism plays a functional role in regulating human pluripotent stem cell state. Cell Stem Cell *19*, 476-490.

Guo, G., von Meyenn, F., Rostovskaya, M., Clarke, J., Dietmann, S., Baker, D., Sahakyan, A., Myers, S., Bertone, P., Reik, W.*, et al.* (2017). Epigenetic resetting of human pluripotency. Development *144*, 2748-2763.

Inman, G.J., Nicolas, F.J., Callahan, J.F., Harling, J.D., Gaster, L.M., Reith, A.D., Laping, N.J., and Hill, C.S. (2002). SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. Mol Pharmacol *62*, 65-74.

Karumbayaram, S., Lee, P., Azghadi, S.F., Cooper, A.R., Patterson, M., Kohn, D.B., Pyle, A., Clark, A., Byrne, J., Zack, J.A.*, et al.* (2012). From skin biopsy to neurons through a pluripotent intermediate under Good Manufacturing Practice protocols. Stem Cells Transl Med *1*, 36-43.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882-883.

Liu, X., Nefzger, C.M., Rossello, F.J., Chen, J., Knaupp, A.S., Firas, J., Ford, E., Pflueger, J., Paynter, J.M., Chy, H.S.*, et al.* (2017). Comprehensive characterization of distinct states of human naive pluripotency generated by reprogramming. Nat Methods *14*, 1055-1062.

Love, M., Anders, S., and Huber, W. (2014). Differential analysis of count data–the DESeq2 package. Genome Biol *15*, 10.1186.

Lowry, W.E., Richter, L., Yachechko, R., Pyle, A.D., Tchieu, J., Sridharan, R., Clark, A.T., and Plath, K. (2008). Generation of human induced pluripotent stem cells from dermal fibroblasts. Proc Natl Acad Sci U S A *105*, 2883-2888.

Ohashi, M., Korsakova, E., Allen, D., Lee, P., Fu, K., Vargas, B.S., Cinkornpumin, J., Salas, C., Park, J.C., Germanguz, I.*, et al.* (2018). Loss of MECP2 Leads to Activation of P53 and Neuronal Senescence. Stem Cell Reports *10*, 1453-1463.

Pastor, W.A., Liu, W., Chen, D., Ho, J., Kim, R., Hunt, T.J., Lukianchikov, A., Liu, X., Polo, J.M., Jacobsen, S.E.*, et al.* (2018). TFAP2C regulates transcription in human naive pluripotency by opening enhancers. Nat Cell Biol *20*, 553-564.

Rossin, E.J., Lage, K., Raychaudhuri, S., Xavier, R.J., Tatar, D., Benita, Y., International Inflammatory Bowel Disease Genetics, C., Cotsapas, C., and Daly, M.J. (2011). Proteins encoded in genomic regions associated with immune-mediated disease physically interact and suggest underlying biology. PLoS Genet *7*, e1001273.

Rostovskaya, M., Stirparo, G.G., and Smith, A. (2019). Capacitation of human naive pluripotent stem cells for multi-lineage differentiation. Development *146*, dev172916.

Sahakyan, A., Kim, R., Chronis, C., Sabri, S., Bonora, G., Theunissen, T.W., Kuoy, E., Langerman, J., Clark, A.T., and Jaenisch, R. (2017). Human naive pluripotent stem cells model X chromosome dampening and X inactivation. Cell Stem Cell *20*, 87-101.

Samarasinghe, R.A., Miranda, O.A., Buth, J.E., Mitchell, S., Ferando, I., Watanabe, M., Allison, T.F., Kurdian, A., Fotion, N.N., Gandal, M.J.*, et al.* (2021). Identification of neural oscillations and epileptiform changes in human brain organoids. Nat Neurosci *24*, 1488-1500.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B.*, et al.* (2012). Fiji: an open-source platform for biological-image analysis. Nat Methods *9*, 676-682.

Takashima, Y., Guo, G., Loos, R., Nichols, J., Ficz, G., Krueger, F., Oxley, D., Santos, F., Clarke, J., and Mansfield, W. (2014). Resetting transcription factor control circuitry toward ground-state pluripotency in human. Cell *158*, 1254-1269.

Thomson, J.A., Itskovitz-Eldor, J., Shapiro, S.S., Waknitz, M.A., Swiergiel, J.J., Marshall, V.S., and Jones, J.M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Tidball, A.M., Dang, L.T., Glenn, T.W., Kilbane, E.G., Klarr, D.J., Margolis, J.L., Uhler, M.D., and Parent, J.M. (2017). Rapid Generation of Human Genetic Loss-of-Function iPSC Lines by Simultaneous Reprogramming and Gene Editing. Stem Cell Reports *9*, 725-731.

Tripathi, S., Pohl, M.O., Zhou, Y., Rodriguez-Frandsen, A., Wang, G., Stein, D.A., Moulton, H.M., DeJesus, P., Che, J., Mulder, L.C.*, et al.* (2015). Meta- and Orthogonal Integration of Influenza "OMICs" Data Defines a Role for UBR4 in Virus Budding. Cell Host Microbe *18*, 723-735.

Watanabe, M., Buth, J.E., Vishlaghi, N., de la Torre-Ubieta, L., Taxidis, J., Khakh, B.S., Coppola, G., Pearson, C.A., Yamauchi, K., Gong, D.*, et al.* (2017). Self-Organized Cerebral Organoids with Human-Specific Features Predict Effective Drugs to Combat Zika Virus Infection. Cell Rep *21*, 517-532.

Watanabe, M., Kang, Y.J., Davies, L.M., Meghpara, S., Lau, K., Chung, C.Y., Kathiriya, J., Hadjantonakis, A.K., and Monuki, E.S. (2012). BMP4 sufficiency to induce choroid plexus epithelial fate from embryonic stem cell-derived neuroepithelial progenitors. J Neurosci *32*, 15934-15945.

Wickham, H. (2016). ggplot2: elegant graphics for data analysis (Springer).

Xu, X., Wells, A.B., O'Brien, D.R., Nehorai, A., and Dougherty, J.D. (2014). Cell type-specific expression analysis to identify putative cellular mechanisms for neurogenetic disorders. J Neurosci *34*, 1420-1431.

Yoon, S.J., Elahi, L.S., Pasca, A.M., Marton, R.M., Gordon, A., Revah, O., Miura, Y., Walczak, E.M., Holdgate, G.M., Fan, H.C.*, et al.* (2019). Reliability of human cortical organoid generation. Nat Methods *16*, 75-78.

Yu, P.B., Deng, D.Y., Lai, C.S., Hong, C.C., Cuny, G.D., Bouxsein, M.L., Hong, D.W., McManus, P.M., Katagiri, T., Sachidanandan, C.*, et al.* (2008). BMP type I receptor inhibition reduces heterotopic ossification. Nat Med *14*, 1363-1369.