nature methods

BRIEF COMMUNICATION

https://doi.org/10.1038/s41592-018-0255-0

Reliability of human cortical organoid generation

Se-Jin Yoon[1], Lubayna S. Elahi[1], Anca M. Pașca[2], Rebecca M. Marton[1], Aaron Gordon[3], Omer Revah[1], Yuki Miura[1], Elisabeth M. Walczak[4], Gwendolyn M. Holdgate[4], H. Christina Fan[4], John R. Huguenard[5], Daniel H. Geschwind[3,6] and Sergiu P. Pașca[1,7]*

The differentiation of pluripotent stem cells in three-dimensional cultures can recapitulate key aspects of brain development, but protocols are prone to variable results. Here we differentiated multiple human pluripotent stem cell lines for over 100 d using our previously developed approach to generate brain-region-specific organoids called cortical spheroids and, using several assays, found that spheroid generation was highly reliable and consistent. We anticipate the use of this approach for large-scale differentiation experiments and disease modeling.

Recent progress in the generation of three-dimensional (3D) cultures from human pluripotent stem cells (hPSCs) promises to accelerate our understanding of human brain development and disease[1–3]. Because these 3D culture preparations are intended to model the cellular architecture of organs closely, they are known as organoids or organ spheroids[2]. For the central nervous system, organoids can display high cell diversity, recapitulate more complex cell–cell interactions among brain regions, develop to later stages than 2D cultures, and model brain disorders when patient-derived hPSCs are used[4–9]. However, one of the challenges of applying brain organoids to disease modeling and, ultimately, large-scale drug and genetic screens is the low reproducibility of differentiation[1]. Therefore, assessing the reliability of 3D neural differentiation across multiple hPSC lines and across replicate differentiations of the same lines over long periods is essential to determine what questions can be addressed with this platform.

We previously developed a directed differentiation method for specifying pyramidal cortical neurons from human induced pluripotent stem cells (hiPSCs) in a 3D culture that resembles the cerebral cortex[10,11]. These brain-region-specific organoid cultures called human cortical spheroids (hCSs) contain functional glutamatergic neurons of deep and superficial cortical layers and nonreactive astrocytes and can be maintained for very long periods (beyond 25 months)[7]. Moreover, this approach is simple and versatile: it involves no embedding in an extracellular matrix and allows other brain regions to be patterned and optionally fused into multi-region spheroids known as assembloids[12].

Here we used single-cell analyses, transcriptional profiling and immunocytochemistry during long-term in vitro differentiation to assess the reliability of hCS derivation across multiple hiPSC lines and experiments (Fig. 1a). We cultured hiPSCs in feeder-free and xeno-free conditions on human recombinant vitronectin in Essential 8 medium ($n = 15$ lines derived from 13 individuals; Supplementary Fig. 1a and Supplementary Table 1 show all hiPSC lines and assays). To derive hCSs in feeder-free conditions (hCS-FF), we then aggregated single-cell-dissociated hiPSCs in AggreWell-800 plates to obtain uniform 3D spheroids, each containing ~10,000 cells (Fig. 1b,c, Methods and Supplementary Fig. 1a). Subsequently, we applied small molecules that modulate the SMAD and Wnt pathways and the growth factors EGF and FGF2 to achieve directed differentiation. After 25 d of differentiation, hCS-FF showed strong transcriptional upregulation of the forebrain markers *FOXG1*, *SIX3* and *PAX6*, in the absence of endoderm (*SOX17*) and mesoderm (*BRACH*) markers ($n = 6$–12 hiPSC lines from 11 individuals; Fig. 1d, Supplementary Fig. 1b–d and Supplementary Table 2). Moreover, hCSs did not express hypothalamus (*RAX*) or spinal cord (*HOXB4*) markers, and expression of the ventral forebrain related genes *NKX2-1* and *GSX2* and the midbrain marker *FOXA2* was absent in 11 out of 12 differentiated hiPSC lines (Fig. 1d, Supplementary Fig. 1d).

To assess the overall success of differentiation across hiPSC lines and experiments, we carried out 4–11 differentiations with each of 12 hiPSC lines, for a total of 85 experiments. We found that in 90% of differentiations, hCS-FFs could be maintained successfully for over 100 d in vitro (Fig. 1e). Spheroids expressed cortical neural markers and showed an internal cytoarchitecture that included proliferative zones (Supplementary Fig. 2a–d). Moreover, hCS-FFs were healthy over long-term cultures, as shown by immunostaining with cleaved Caspase 3 (c-Cas3) in hCS-FFs derived from a TUBA1B-mEGFP hiPSC line (Supplementary Fig. 2e). Approximately 5% of cultures did not aggregate in the first 24 h, and another 5% of cultures disintegrated over time (most commonly in the first 20 d) (Fig. 1e and Supplementary Fig. 1a). The generation of hCS-FFs was scalable; one researcher was able to initiate the differentiation of 10–15 hiPSC lines on a weekly basis and maintain these cultures for more than 100 d.

We performed RNA-seq on hCS-FF collected at four stages of differentiation for six different hiPSC lines in at least three independent experiments (Supplementary Table 3). We used a rank–rank hypergeometric overlap (RRHO) test[13] to assess similarity between changes in hCS ($n = 6$ hiPSC lines) from an earlier stage (day 25) to a late stage (day 100) and changes between two time points in the developing human cerebral cortex[14] (Supplementary Fig. 3a,b). Similar to our previous transcriptomic data on hCSs derived from hPSCs cultured on mouse embryonic fibroblasts (hCS-MEF), we observed high overlap between hCS-FF at day 100 of differentiation and those at mid-gestation stages of development (up to post-conception week (PCW) 24). We next used principal component analysis to characterize transcriptional variance between different differentiations of six hiPSC lines over time (Fig. 1f). The predominant source of variation was related to the stage of in vitro differentiation

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, USA. [2]Department of Pediatrics, Division of Neonatology, Stanford University School of Medicine, Stanford, CA, USA. [3]Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [4]BD Biosciences, Menlo Park, CA, USA. [5]Department of Neurology and Neurological Sciences, Stanford University School of Medicine, Stanford, CA, USA. [6]Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA, USA. [7]Human Brain Organogenesis Program, Stanford University, Stanford, CA, USA. *e-mail: spasca@stanford.edu

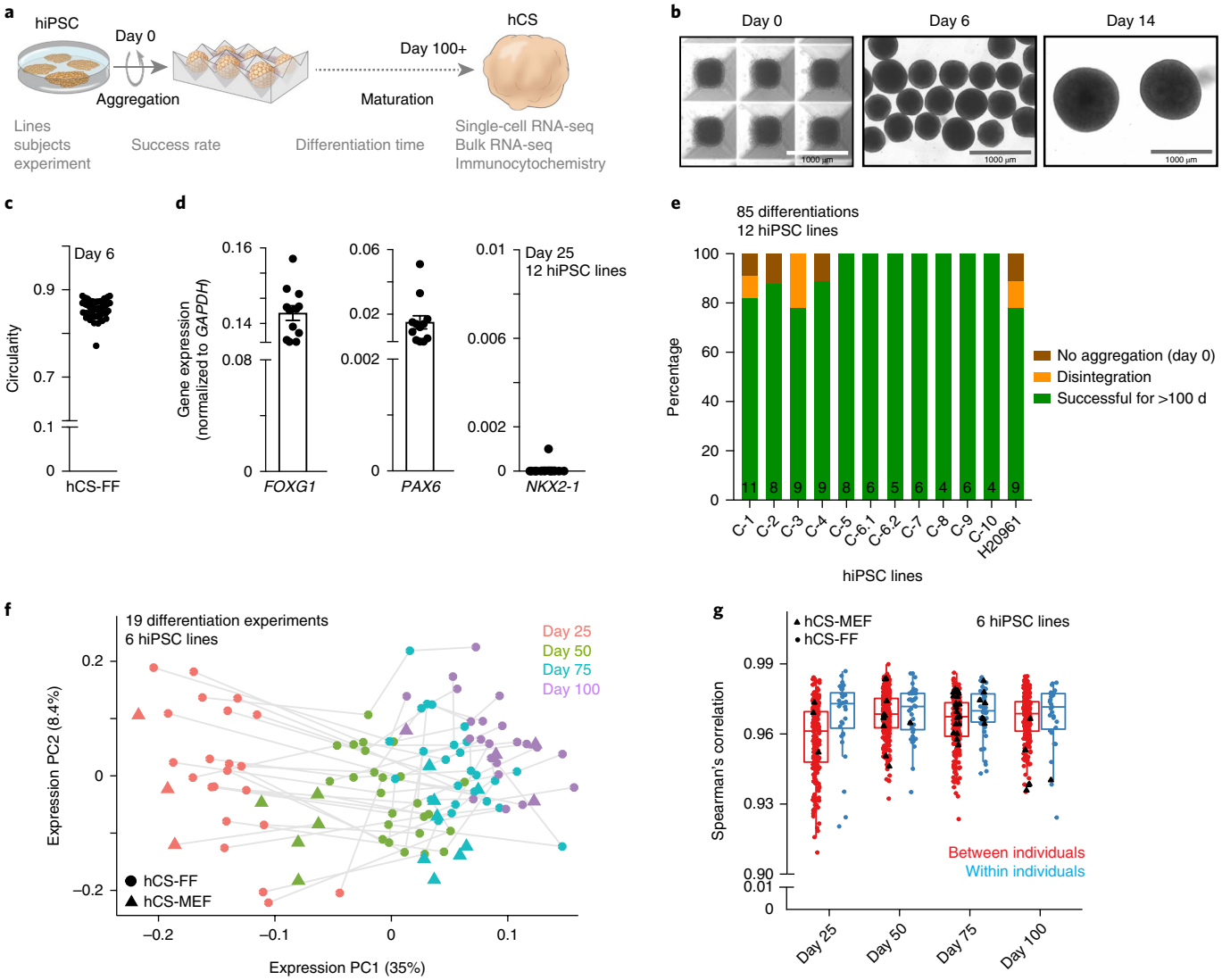

Fig. 1 | Success of differentiation and transcriptional reliability of human cortical spheroids. a, Scheme illustrating the derivation of hCS-FF from hPSCs and the assays used. **b**, Representative images of neural spheroids at days 0, 6 and 14 of differentiation. **c**, Circularity ($4\pi \times$ area/perimeter^2) of day 6 neural spheroids derived from four hiPSC lines. A value of 1.0 indicates a perfect circle. **d**, Gene expression of *FOXG1*, *PAX6* and *NKX2.1* relative to *GAPDH* in hCS-FF at day 25 of differentiation ($n = 12$ hiPSC lines from 11 subjects). Mean ± s.e.m. are shown. **e**, Percentage of successful differentiations up to 100 d for 12 hiPSC lines ($n = 85$ experiments; number per line indicated inside bars). **f**, Principal component analysis of hCS-FF and hCS-MEF at four stages of in vitro differentiation. Differentiations of the same line are indicated by a gray line (days 25, 50, 75, 100: $n = 22$, 25, 25, 22 hCS-FF and 3, 5, 8, 4 hCS-MEF samples, respectively). **g**, Spearman's correlation of samples obtained from different individuals (between individuals) or from multiple differentiations of the same hiPSC lines (within individual); two-sided Wilcoxon–Mann–Whitney test, $P < 0.03$. Day 25, 50, 75, 100: $n = 202$, 269, 281, 206 samples (between individual) and 33, 41, 47, 31 samples (within individuals), respectively. Middle hinge corresponds to median, and lower and upper hinges correspond to first and third quartiles. RNA-seq data in **f** and **g** were obtained from $n = 6$ hiPSC lines derived from six individuals and differentiated in multiple independent differentiation experiments each.

(first principal component, explaining 35% of the variance), whereas different differentiation experiments and hiPSC lines accounted for only 8.4% of the variance (principal component 2), suggesting that this protocol is reproducible across multiple individual hiPSCs, differentiation protocols (FF versus MEF) and replicate differentiations. This is also illustrated by the high Spearman's correlation coefficients (mean > 0.96) between differentiations of the same cell line (within individuals) or between different hiPSC lines (between individuals) (Wilcoxon–Mann–Whitney test, $P < 0.05$; Fig. 1g), for matched differentiation days. Previous work on forebrain organoids reported a mean Spearman's correlation coefficient in the range of 0.75–0.85 between and within individuals[5]. Interestingly, the variability was higher at earlier stages of differentiation but decreased with differentiation time (Fig. 1f).

To comprehensively characterize hCS-FF, we performed single-cell transcriptional profiling at day 105 of differentiation using stochastic barcoding[15] and compared the results to those for hCS-MEF and spheroids resembling the subpallium (hSS), as previously described[12] ($n = 24{,}237$ cells; BD Rhapsody system). As expected, clustering of cells from hCS and hSS using *t*-distributed stochastic neighbor embedding (*t*SNE) revealed a separation of dorsal and ventral differentiation conditions (Fig. 2a). Neurons expressing *STMN2* were localized in a region of the *t*SNE space distinct from progenitors and mitotically active cells (Supplementary Fig. 4a), and no cells with a mesodermal or endodermal identity were found. Cells from hCS-FF derived from three hiPSC lines, which were in turn derived from two individuals, clustered closely with dorsal forebrain (hCS-MEF)

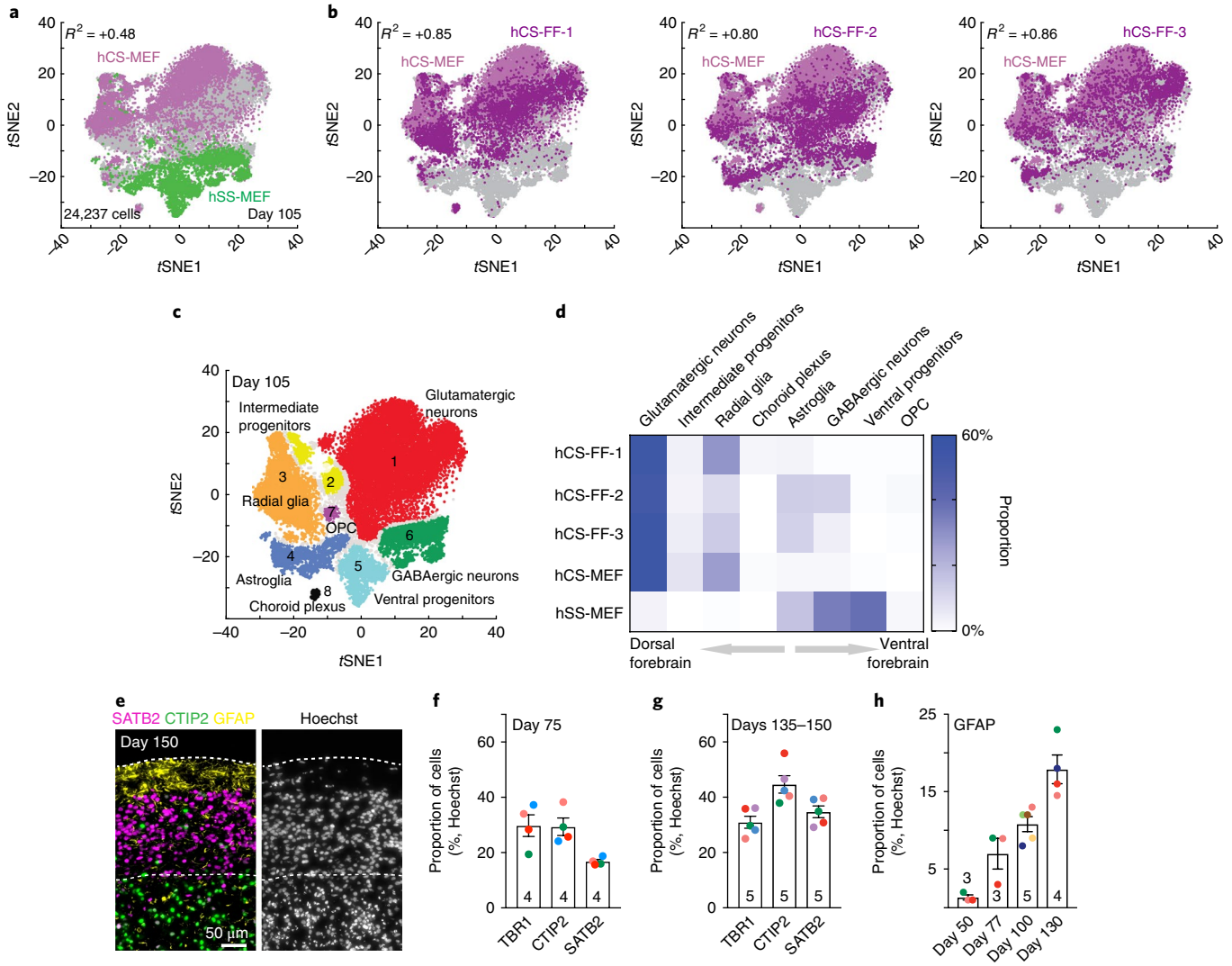

Fig. 2 | Single-cell characterization of human cortical spheroids. a,**b**, Single-cell profiling of hCS-MEF ($n = 7{,}340$ cells) and hSS-MEF ($n = 4{,}771$ cells) (**a**) versus hCS-FF derived from three different individuals (hCS-FF-1, $n = 4{,}649$ cells; hCS-FF-2, $n = 4{,}389$ cells; hCS-FF-3, $n = 3{,}088$ cells) (**b**) at day 105 of differentiation. The correlation ($\log_{10}$ transformed mean of molecules per cell per gene) between total hCS-MEF and each of the hCS-MEF cultures or hSS-MEF is indicated. **c**,**d**, Clustering (**c**) and proportions (**d**) of all single cells across conditions (glutamatergic neuron cluster 1, $n = 11{,}367$ cells; intermediate progenitor cluster 2, $n = 1{,}018$ cells; radial glia cluster 3, $n = 4{,}217$ cells; astroglia cluster 4, $n = 2{,}036$ cells; ventral progenitor cluster 5, $n = 1{,}915$ cells; GABAergic neuron cluster 6, $n = 2{,}520$ cells; OPC cluster 7, $n = 194$ cells; choroid plexus cluster 8, $n = 170$ cells; 800 cells not assigned to a cluster). **e**, Representative hCS cryosection at day 150 of differentiation stained for deep (CTIP2) and superficial (SATB2) neuronal markers and the glial marker GFAP. **f**,**g**, Proportion of cells expressing layer-specific cortical markers (TBR1, CTIP2, SATB2) at day 75 (**f**) and days 135–150 (**g**) of differentiation from four or five hiPSC lines derived from four or five individuals (two-way ANOVA, $F_{2,21} = 5.44$, $P = 0.01$ for interaction; mean ± s.e.m. are shown; sample size indicated on each column). **h**, Developmental time course for GFAP$^+$ cell generation, quantified in dissociated hCSs ($n = 3$–5 hiPSC lines per time point from seven hiPSC lines derived from six individuals; ANOVA $F_{3,11} = 21.89$, $P < 0.0001$; mean ± s.e.m.).

and separately from ventral forebrain fate (hSS-MEF). The correlation ($\log_{10}$ transformed mean of molecules per cell per gene) between hCS-MEF and each of the hCS-FF cultures was higher than $R^2 = 0.8$, whereas the correlation with hSS was $R^2 = 0.4$ (Fig. 2b and Supplementary Fig. 4b). The proportion of cortical glutamatergic neurons varied between 54% and 59% of all cells in the four hCS samples. One of the hCS samples (hCS-FF-2) included some GABAergic neurons, which is consistent with our observation that minor ventralization can sometimes occur (Fig. 1d). We have previously shown that choroid plexus cells are rare in hCS (1–2%)[12], and here we found that two of the lines did not include choroid cells.

Overall, the distribution of the cells in the eight main clusters (Fig. 2c and Supplementary Table 4) that we previously identified in forebrain spheroids was similar across hCS conditions (Fig. 2d, Spearman's $R^2 > 0.78$, $P < 0.05$ for each of the hCS-FF versus hCS-MEF and $R^2 = 0.02$ for hSS-MEF versus hCS-MEF). Moreover, temporal trajectories of gene expression for selected cortical markers in whole hCS-FF also indicated that differentiations were consistent (Supplementary Fig. 5). We complemented these analyses by immunocytochemistry and quantified the proportions of specific cell types in whole hCS-FF cryosections at two time points (day 75 and days 135–150). The proportion of layer-specific neurons changed with time, with SATB2$^+$ cells present mostly at later stages (Fig. 2e–g; two-way ANOVA, $F_{2,21} = 5.44$, $P = 0.01$ for interaction; mean ± s.e.m. are shown); s.d. values for the three cortical layer markers that we quantified at two time points were in the range of 1.4% to 7.82%. Lastly, we explored cells expressing GFAP, a marker

that is expressed early in radial glia and at later stages in astrocytes in hCS[7,10], and found a progressive increase over 6 months in vitro (Fig. 2h; ANOVA $F_{3,11} = 21.89$, $P < 0.0001$).

In slice recordings, we found that all Syn1::mCherry$^+$ neurons fired action potentials and ~70% of cells were capable of generating repetitive action potentials when depolarized (Supplementary Fig. 6a,b; $n = 13$ cells from two hiPSC lines). The firing activity was tetrodotoxin (TTX) sensitive (Supplementary Fig. 6c). Moreover, all neurons exhibited spontaneous synaptic activity, which was blocked by glutamate antagonists ($n = 13$ cells; two-sided paired t-test $^{***}P < 0.001$; Supplementary Fig. 6d,e). Lastly, we found that blue light could trigger action potentials in neurons expressing Syn1::ChR2-mCherry (Supplementary Fig. 6f,g) and that blue light pulses delivered at increasing rates (8–32 Hz) triggered frequency-dependent action potential failure (Supplementary Fig. 6h,i).

Brain organoids have the potential to recapitulate the cell diversity of the central nervous system in long-term in vitro cultures[2]. However, this has often been associated with noteworthy variability in studies involving differentiation of the same hPSC line in different experiments or across cell lines from different individuals. Because they derive a large diversity of cell types, including nonectodermal lineages, undirected organoid approaches display more variability of differentiation and bioreactor-based batch effects[16,17]. This diversity may be useful in mapping disease-associated gene expression, but it also represents a barrier to the development of disease models, which require consistent differentiation of large cohorts of hiPSC lines that often display different differentiation propensities.

Here we assessed the reliability of an hiPSC directed differentiation approach for deriving 3D cultures resembling the dorsal forebrain. This feeder-free, xeno-free approach does not involve cell (re-)plating, embedding into extracellular matrices or culture in complex environments. Using transcriptional time course, immunocytochemistry and single-cell analyses, we found that the derivation of hCS is reliable across multiple hiPSC lines and differentiation experiments. Moreover, these floating cultures are functional and can be maintained for long periods in vitro to capture human corticogenesis and astrogenesis[7]. This platform can be adapted to specify other brain regions and to derive assembloids, and it can facilitate the development of disease models using patient-derived cells, or of genetic and drug screens.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41592-018-0255-0.

Received: 14 December 2017; Accepted: 2 November 2018;
Published online: 20 December 2018

References

1. Di Lullo, E. & Kriegstein, A. R. *Nat. Rev. Neurosci.* **18**, 573–584 (2017).
2. Paşca, S. P. *Nature* **553**, 437–445 (2018).
3. Kelava, I. & Lancaster, M. A. *Cell Stem Cell* **18**, 736–748 (2016).
4. Bershteyn, M. et al. *Cell Stem Cell* **20**, 435–449 (2017).
5. Mariani, J. et al. *Cell* **162**, 375–390 (2015).
6. Lancaster, M. A. et al. *Nature* **501**, 373–379 (2013).
7. Sloan, S. A. et al. *Neuron* **95**, 779–790 (2017).
8. Qian, X. et al. *Cell* **165**, 1238–1254 (2016).
9. Blair, J. D., Hockemeyer, D. & Bateup, H. S. *Nat. Med.* **24**, 1568–1578 (2018).
10. Paşca, A. M. et al. *Nat. Methods* **12**, 671–678 (2015).
11. Sloan, S. A., Andersen, J., Paşca, A. M., Birey, F. & Paşca, S. P. *Nat. Protoc.* **13**, 2062–2085 (2018).
12. Birey, F. et al. *Nature* **545**, 54–59 (2017).
13. Stein, J. L. et al. *Neuron* **83**, 69–86 (2014).
14. BrainSpan. *BrainSpan: Atlas of the Developing Human Brain* http://www.brainspan.org (2013).
15. Fan, H. C., Fu, G. K. & Fodor, S. P. *Science* **347**, 1258367 (2015).
16. Quadrato, G. et al. *Nature* **545**, 48–53 (2017).
17. Camp, J. G. et al. *Proc. Natl Acad. Sci. USA* **112**, 15672–15677 (2015).

Acknowledgements

We acknowledge support from J. Dreux and J. Lai (at BD Genomics) for running the pipeline and Data View troubleshooting. This work was supported by the US National Institutes of Health (NIH) BRAINS Award (MH107800), the MQ Fellow Award, the NYSCF Robertson Stem Cell Investigator Award, the Donald E. and Delia B. Baxter Foundation Award, the Stanford Neurosciences Institute's Brain Rejuvenation Project, the Kwan Research Fund, a NARSAD Independent Investigator Award from the Brain & Behaviour Research Foundation (BBRF), and Stanford Start-up Funds (to S.P.P.); the California Institute of Regenerative Medicine (CIRM) and NIH U01 (MH115745) (to D.H.G. and S.P.P.); CIRM Bridges Program (to L.S.E.); a Stanford School of Medicine Dean's Postdoctoral Fellowship (to Y.M.); a National Science Foundation (NSF) Graduate Research Fellowship (to R.M.M.); NIH K12-HD000850 (Pediatric Scientist Development Program) and a Stanford Child Health Research Institute Fellowship (to A.M.P); and the Autism Science Foundation Fellowship (to A.G.).

Author contributions

S.J.Y., R.M.M. and S.P.P. developed the differentiation method. S.J.Y. and L.S.E. performed differentiation experiments. S.J.Y., L.S.E., A.M.P., R.M.M. and Y.M. characterized the protocol. E.M.W., G.M.H. and H.C.F. performed the single-cell transcriptomics experiments and analysis. A.G. and D.H.G. performed the RNA-seq and analysis. O.R., S.P.P. and J.R.H. designed, conducted or analyzed the electrophysiological experiments. S.J.Y. and S.P.P. wrote the manuscript with input from all other authors. S.P.P. supervised the work.

Competing interests

Stanford University has filed a provisional patent application that covers the generation of region-specific neural spheroids (US Application Serial No. 15/158,408). H.C.F., E.M.W. and G.M.H. were employees of BD Genomics during this study.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41592-018-0255-0.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to S.P.P.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Culture of hiPSCs. The hiPSC lines used in this study were validated using standardized methods as previously described[10,12,18] (Supplementary Table 1). Cultures were tested for mycoplasma contamination and were maintained mycoplasma free. A total of 15 hiPSC lines were collected from 13 healthy subjects. Among them, the hiPSC H20961 line was derived by the Gilad Laboratory[19], and two monoallelic mEGFP-tagged hiPSC lines, LMNB1-mEGFP (AICS-0013) and TUBA1B-mEGFP (AICS-0012), were obtained from Coriell. Approval for this study was obtained from the Stanford IRB panel, and informed consent was obtained from all subjects.

Generation of hCS-FF from hiPSC. hiPSCs were maintained on six-well plates coated with recombinant human vitronectin (VTN-N, Life Technologies, A14700) in Essential 8 medium (Life Technologies, A1517001). To coat the six-well plates, 1 ml of vitronectin (diluted at a 1:100 ratio with Dulbecco's phosphate-buffered saline (DPBS); Life Technologies, 14190) was added per well and then incubated at room temperature for 1 h. To passage hiPSCs (80–90% confluency), cells were rinsed with 3–4 ml of DPBS per well, and then 1 ml of 0.5 mM EDTA (Life Technologies, 15575) was added for 7 min at room temperature. After the EDTA was removed, 2 ml of pre-warmed complete Essential 8 medium was added to collect cells. The cell suspension was then diluted in Essential 8 medium (1:6–1:20 depending on the hiPSC line) and distributed on vitronectin-coated wells.

For the generation of 3D spheroids, hiPSCs were incubated with Accutase (Innovate Cell Technologies, AT-104) at 37 °C for 7 min and dissociated into single cells. Optionally, 1 d before spheroid formation, hiPSCs can be exposed to 1% dimethylsulfoxide (DMSO) in Essential 8 medium. To obtain uniformly sized spheroids, we used AggreWell 800 (STEMCELL Technologies, 34815) containing 300 microwells. Approximately 3×10^6 single cells were added per AggreWell 800 well in Essential 8 medium supplemented with the ROCK inhibitor Y-27632 (10 μM, Selleckchem, S1049), centrifuged at 100*g* for 3 min to capture the cells in the microwells and incubated at 37 °C with 5% CO_2. After 24 h, we collected spheroids from each microwell by firmly pipetting (with a cut end of a P1000 tip) medium in the well up and down and transferring it into ultra-low-attachment plastic dishes (Corning, 3262) in Essential 6 medium (Life Technologies, A1516401) supplemented with two SMAD pathway inhibitors—dorsomorphin (2.5 μM, Sigma-Aldrich, P5499) and SB-431542 (10 μM, Tocris, 1614). From day 2 to day 6, Essential 6 medium was changed every day and supplemented with dorsomorphin and SB-431542. Optionally, the Wnt pathway inhibitor XAV-939 (2.5 μM, Tocris, 3748) can be added with the two SMAD pathway inhibitors. On the sixth day in suspension, neural spheroids were transferred to neural medium containing Neurobasal A (Life Technologies, 10888), B-27 supplement without vitamin A (Life Technologies, 12587) and GlutaMax (1:100, Life Technologies, 35050). The neural medium was supplemented with 20 ng ml^{-1} epidermal growth factor (R&D Systems, 236-EG) and 20 ng ml^{-1} basic fibroblast growth factor (R&D Systems, 233-FB) for 19 d (until day 24), with medium changed daily in the first 10 d and every other day for the subsequent 9 d. To promote differentiation of the neural progenitors into neurons, the neural medium was supplemented with 20 ng ml^{-1} brain-derived neurotrophic factor (BDNF; Peprotech, 450-02) and 20 ng ml^{-1} NT3 (Peprotech, 450-03), with medium changes every other day. From day 43, only neural medium without growth factors was used for medium changes every 4 d.

A step-by-step protocol for the generation of cortical spheroids from hiPSCs in feeder-free, xeno-free conditions is available in *Protocol Exchange* (ref. [20]).

Cryopreservation. hCS were fixed in 4% paraformaldehyde overnight at 4 °C. They were then washed in PBS and transferred to 30% sucrose for 48–72 h. Subsequently, they were transferred into embedding medium (Tissue-Tek optimum cutting temperature (OCT) compound 4583, Sakura Finetek), snap-frozen on dry ice and stored at −80 °C. For immunohistochemistry, 10–12-μm-thick sections were cut with a cryostat (Leica).

Immunohistochemistry. Cryosections were washed with PBS to remove excess OCT compound and blocked in 10% normal goat serum (NGS), 0.3% Triton X-100 diluted in PBS for 1 h at room temperature. The sections were then incubated overnight at 4 °C with primary antibodies diluted in PBS containing 2% NGS and 0.1% Triton X-100. PBS was used to wash off the primary antibodies, and the cryosections were incubated with secondary antibodies in PBS with 2% NGS and 0.1% Triton X-100 for 1 h. The following primary antibodies were used for immunohistochemistry: anti-SATB2 (mouse, 1:50, Abcam, AB51502), anti-CTIP2 (rat, 1:300, Abcam, AB18465), anti-TBR1 (rabbit, 1:300, Abcam, AB31940), anti-TBR1 (chicken, 1:200, Millipore, AB2261), anti-TLE4 (mouse, 1:200, Santa Cruz, sc-365406), anti-HOPX (mouse, 1:100, Santa Cruz, sc-398703), anti-RELN (mouse, 1:200, MBL, D223-3), anti-BRN2 (rabbit, 1:250, GeneTex, GTX114650), anti-cCas3 (Asp175; rabbit, 1:200, CST, 9661), anti-GFAP (rabbit, 1:1,000, DAKO, Z0334) and anti-SOX2 (rabbit, 1:200, CTS, 3579).

For TBR2 staining (anti-TBR2, mouse, 1:100; R&D, MAB6166), antigen retrieval (using 10 mM Na^+ citrate and heating to 95 °C) was performed for 20 min, and sections were blocked in 0.1% Triton X-100, 10% serum and 0.2% gelatin for 1 h. The sections were incubated with primary antibodies in blocking buffer overnight at 4 °C, then washed with PBS with 0.5% Triton X-100 for 1 h and incubated with secondary antibodies diluted in blocking buffer for 1 h at room temperature.

For GFAP immunocytochemistry (rat, anti-GFAP, 1:1,000, Thermo Fisher 13-0300), hCS-FF were dissociated and plated in monolayer as described below. Plated cell coverslips were blocked in 10% NGS and 0.3% Triton X-100 diluted in PBS for 1 h at room temperature. Coverslips were incubated for 2 h at room temperature with primary antibodies diluted in PBS containing 2% NGS and 0.1% Triton X-100. PBS was used to wash the primary antibodies, and cells were then incubated with secondary antibodies in PBS with 2% NGS and 0.1% Triton X-100 for 1 h.

Images were acquired with a Zeiss Axio Imager M2 or a KEYENCE BZ-X710, and processed with ImageJ.

Dissociation of hCS and hSS. For the enzymatic dissociation of hCS-FF for culture in monolayer and immunocytochemistry (anti-GFAP), up to three spheroids were incubated with 200 μl of Accutase (Innovative Cell Technologies) for 20 min at 37 °C, washed with neural medium and gently triturated using a P200 pipette. Cells were plated on glass coverslips (15 mm, Warner Instruments) coated with poly-L-ornithine and laminin (Sigma-Aldrich) at a density of around 250,000 cells per coverslip in neural medium supplemented with BDNF, NT3 and, for the first 24 h, 10 μM Y-27632.

To dissociate hCS and hSS for single-cell profiling, we used a previously published protocol[7]. Briefly, 3–10 spheroids were chopped with a #10 blade and then incubated in 40 U ml^{-1} papain enzyme solution containing 0.46% D(+)-glucose (Sigma-Aldrich), 26 mM $NaHCO_3$ (Sigma-Aldrich), 0.5 mM EDTA (Sigma-Aldrich) in EBSS (1×, Sigma-Aldrich) at 37 °C in 5% CO_2 for 70 min. The digested spheroids were then washed and carefully triturated with a protease inhibitor stock solution containing 0.46% D(+)-glucose (Sigma-Aldrich), 26 mM $NaHCO_3$ (Sigma-Aldrich), 5 mg trypsin inhibitor (Sigma-Aldrich) in EBSS (1×, Sigma-Aldrich). After centrifugation (200*g* for 4 min), the pellet was resuspended in 0.2% bovine serum albumin diluted in PBS and supplemented with 10 μM Y-27632, and the cells were used for the single-cell RNA sequencing.

Real-time quantitative polymerase chain reaction (RT-qPCR). At least three hCS-FF were combined for RNA extraction. Total RNA was isolated using the RNeasy mini kit and RNase-Free DNase set (Qiagen), and template complementary DNA was prepared by reverse transcription using the SuperScript III First-Strand Synthesis SuperMix for RT-qPCR (Life Technologies). qPCR was performed with SYBR Green (Roche) on a ViiA7 machine (Applied Biosystems, Life Technologies). Data were processed using the QuantStudio RT-qPCR software (v1.1, Applied Biosystems). Primers and sequences are listed in Supplementary Table 2.

RNA-seq processing. RNA-seq samples are listed in Supplementary Table 3. Following extraction of total RNA, ribosomal RNA was depleted (RiboZero Gold, Illumina) and libraries were prepared using Truseq stranded RNA RiboZero Gold (Illumina). Libraries were then sequenced, generating 100-base-pair paired-end reads on an Illumina HiSeq 4000. Paired-end reads were mapped using STAR[21] to hg38 with Gencode v25 annotations. Gene expression levels were quantified using rsem (v1.3.0)[22]. Genes with fewer than ten mapped reads in over 60% of the samples were removed. Samples with standardized sample network connectivity Z scores below −2 were defined as outliers and removed[23]. To control for technical variation in the RNA sequencing (for example, read depth, % coding bases, base duplication rate), we calculated the first five principal components of the Picard sequencing metrics (http://broadinstitute.github.io/picard/) and included them in our linear model. To control for batch effects, we either included batch as a covariate in the linear model (Supplementary Fig. 3) or removed it using the ComBat function from the sva package in R (for PCA and correlation analysis see Fig. 1f,g)[24].

The GATK (v3.3) Haplotype caller was used to call SNPs from the aligned reads[25]. Resulting vcf files were converted to plink format (v1.08), and sites with more than 5% missing samples, with rare minor allele frequency (less than 0.05) and out of Hardy–Weinberg equilibrium (less than 1×10^{-6}) were filtered out[26]. These high-quality SNPs were used to run MDS together with HapMap3.3 (hg38) to infer sample race. The first two MDS values, referred to as raceMDS1/2, were then included in our linear model. For PCA and correlation analysis, raceMDS1–2 and SeqPC1–5 were regressed out before batch correction.

Single-cell gene expression (BD Rhapsody system). To capture single-cell transcriptomic information of hiPSC-derived hCS-FF samples, we used the BD Rhapsody system (formerly known as BD Resolve) (BD Biosciences) as previously reported[12,15]. Three hiPSC lines that were differentiated in parallel (that is, in the same differentiation batch) into hCS-FF were dissociated enzymatically into single cells at day 105 of differentiation and processed on the same day. Four to six spheroids from each hiPSC line were combined, the proportion of live cells was estimated using a fluorescent assay (~90%), and all cells were used for further processing. A single-cell suspension of ~10,000 cells was captured from all isolated cells, without selection, on an array of >200,000 microwells through a limited dilution approach. Beads with oligonucleotide barcodes were added to saturation so that a bead was paired with a cell in a microwell. After exposure to

cell lysis buffer, polyadenylated RNA molecules hybridized to the beads. Beads were retrieved into a single tube for reverse transcription. On cDNA synthesis, each cDNA molecule was tagged on the 5′ end (that is, the 3′ end of a messenger RNA transcript) with a molecular index and cell label indicating its cell of origin. Whole-transcriptome libraries were prepared from ~67% of the captured cells by subsampling of the Rhapsody beads that were then subject to second-strand cDNA synthesis, adaptor ligation, and universal amplification using 22 cycles of PCR. The rest of the beads were archived. Sequencing libraries were prepared using random priming PCR of the whole-transcriptome amplification products to enrich the 3′ end of the transcripts linked with the cell label and molecular indices. The libraries were sequenced on HiSeq2500 (Illumina) using 101 × 2 chemistry.

The BD Rhapsody analysis pipeline was used to process sequencing data (.fastq files). Cell labels and molecular indices were identified, and gene identity was determined by alignment against the Gencode comprehensive hg19 reference. A table containing molecule counts per gene per cell was the output.
Gene expression profiles of 4,649, 4,389 and 3,088 cells were recovered for hCS-FF-1, hCS-FF-2 and hCS-FF-3, respectively, with an average number of reads of ~38,000, ~2,000–2,300 molecules and ~1,100–1,300 genes detected per cell, with average molecular index coverage (that is, the number of times a molecule was sequenced) of 3.2–3.5.

Analysis of the single-cell transcriptome profiles was performed with BD Data View.

We compared the current datasets (hCS-FF-1–3) with single-cell data from hCS-MEF and hSS-MEF obtained on the same platform previously[12]. Cells with mitochondrial gene (with a gene symbol starting with *MT*) content > 30% were discarded, leaving a total of 24,237 cells from all samples. We extracted the expression profiles of the 338 genes (Supplementary Table 4) identified previously[12] that define the eight functional populations in forebrain spheroids, and conducted *t*SNE projection on the filtered data.

BrainSpan data. BrainSpan developmental RNA-seq data from cortical regions[14] were mapped to hg38 using Gencode v25 via STAR[21]. Gene expression levels were quantified using union exon models with featureCounts[27]. Samples with RIN scores below 8, containing less than 25% coding bases or more than 25% ribosomal bases (as called by Picard tools), were removed from the analysis. Genes with fewer than ten mapped reads in over 80% of the samples per developmental period were removed.

Transition mapping. Gene expression distributions were normalized using the trimmed mean of *M* values (TMM) method from the edgeR package[28] followed by the voom method from the limma package[29]. We calculated the log fold-change for each differentiation day or developmental period compared to every other using the limma package by applying contrasts. For the BrainSpan data in which multiple samples came from the same brain, the brain ID was used as a blocking factor in the model. Genes were ranked by log fold-change, and a hypergeometric test was used with a step size of 200 to calculate the significance of the overlap between the gene lists[13].

Electrophysiology. Sections of hCS were obtained via an approach we previously described[10]. Briefly, 3D spheroids were incubated in bicarbonate-buffered artificial cerebrospinal fluid (aCSF) at 23 °C and equilibrated with a mixture of 95% O_2 and 5% CO_2. The aCSF solution contained 126 mM NaCl, 26 mM $NaHCO_3$, 10 mM glucose, 2.5 mM KCl, 1.25 mM NaH_2PO_4, 1 mM $MgSO_4$ and 2 mM $CaCl_2$. Slicing was performed using a Leica VT1200 vibratome. Immediately after sectioning, slices were moved to a circulation chamber containing oxygenated aCSF at room temperature.

For patch-clamp recording, cells were identified by the presence of a fluorescent reporter using an upright 'slicescope' microscope (Scientifica). Recording electrodes of borosilicate glass had a resistance of 7–10 MΩ when filled with internal solution. The internal solution contained 145 mM K gluconate, 0.1 mM $CaCl_2$, 2.5 mM $MgCl_2$, 10 mM HEPES, 0.2 mM EGTA, 4 mM Na phosphocreatine. Glutamatergic currents were blocked by application of NBQX (20 µM, Tocris) and APV (100 µM, Tocris).

hCS neurons plated in monolayer and infected with the AAV-Syn1::ChR2-mCherry were activated with a 475-nm light using a 200-µm-diameter optical fiber (Thorlabs). Pulsed light (2 or 5 ms, 1.5 mW) was applied at 8, 12, 16, 24 and 32 Hz for 1 s to stimulate action-potential generation.

Data were collected using a 1550 A digitizer (Molecular Devices) and a 700B patch-clamp amplifier (Molecular Devices) and were acquired with the pClamp 10.6 software (Molecular Devices). Data were low-pass filtered at 10 kHz and digitized at 20 kHz. Data averaging, digital subtraction of null traces, and action potential and EPSC detection were obtained from Clampfit (Molecular Devices). Data were fitted and plotted using Origin (OriginLab).

Statistics. Data are presented as mean ± s.e.m., unless otherwise indicated. Distribution of the raw data was tested for normality of distribution; statistical analyses were performed using the *t*-test, Wilcoxon–Mann–Whitney test, χ^2 test or ANOVA as indicated. Sample sizes were estimated empirically or on the basis of power calculations.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Gene expression data are available in the Gene Expression Omnibus (GEO) under accession numbers GSE93811, GSE107771 and GSE120700. The data that support the findings of this study are available on request from the corresponding author.

References

18. Pașca, S. P. et al. *Nat. Med.* **17**, 1657–1662 (2011).
19. Gallego Romero, I. et al. *eLife* **4**, e07103 (2015).
20. Yoon, S. J. & Pașca, S. P. *Protoc. Exch.* https://doi.org/10.1038/protex.2018.123 (2018).
21. Dobin, A. et al. *Bioinformatics* **29**, 15–21 (2013).
22. Li, B. & Dewey, C. N. *BMC Bioinformatics* **12**, 323 (2011).
23. Oldham, M. C., Langfelder, P. & Horvath, S. *BMC Syst. Biol.* **6**, 63 (2012).
24. Johnson, W. E., Li, C. & Rabinovic, A. *Biostatistics* **8**, 118–127 (2007).
25. McKenna, A. et al. *Genome Res.* **20**, 1297–1303 (2010).
26. Purcell, S. et al. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
27. Liao, Y., Smyth, G. K. & Shi, W. *Bioinformatics* **30**, 923–930 (2014).
28. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. *Bioinformatics* **26**, 139–140 (2010).
29. Ritchie, M. E. et al. *Nucleic Acids Res.* **43**, e47 (2015).

nature research

Corresponding author(s): Sergiu Pasca

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated
☐	☒	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	The ZEN (Zeiss) and BZ-X Analyzer (KEYENCE) softwares were used for acquiring microscopy data. The BD Rhapsody software was used for single-cell gene expression. Electrophysiology data was collected using a 1550A digitizer (Molecular Devices), a 700B patch-clamp amplifier (Molecular Devices) and the pClamp 10.6 software (Molecular Devices).
Data analysis	The GraphPad Prism Version 7.0 was used for statistical analyses. ImageJ was used for image quantification. Clampfit (Molecular Devices) and Origin (OriginLab) were used for the analysis of electrophysiology data. QuantStudio Software v1.1 was used for Real-Time PCR data analysis and STAR v2.5.2b, Picard v2.5.0, GATK v3.3, featureCounts v1.5.1 and plink v1.08 were used for RNA-seq analyses. R version 3.3.2 and 3.4.0 with the following packages tidyverse v1.2.1, WGCNA v1.61, sva v3.22.0 edgeR v3.20.9, limma v3.34.9 were used for RNA-seq analyses.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Gene expression data are available in the Gene Expression Omnibus (GEO) under accession numbers GSE93811 and GSE107771. The data supporting the findings of this study are available upon request from the corresponding author.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The hiPSC lines used in each experiment are described in Supplementary Table 1. For most experiments, at least 4 hiPSC lines were used. For the electrophysiological experiments, 2 hiPSC lines were used. Sample sizes were determined empirically.
Data exclusions	Data was not excluded. However, some ventralization (forebrain) was observed in one of the 12 hiPSC lines tested in the qPCR experiments as shown in Figure 1d, and because the main goal of our analyses were on dorsal forebrain/hCS, this line was not used for immunohistochemistry analyses.
Replication	hiPSC lines were differentiated in multiple differentiation experiments to assess reliability of the methods. Success of differentiation is quantified in Figure 1e.
Randomization	The hiPSC lines used in each experiment are described in Supplementary Table 1. At least 4 hiPSC lines were used for most analyses, except for electrophysiological analysis: 2 hiPSC lines were used for the experiment. All hiPSC lines were derived from healthy individuals; no group comparisons were performed.
Blinding	Samples for RNA-seq were blinded. Most other experiments did not involve comparison accross groups, and blinding was not used.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Mouse anti-SATB2 (Abcam, AB51502), used 1:50
Rat anti-CTIP2 (Abcam, AB18465), used 1:300
Rabbit anti-TBR1 (Abcam, AB31940), used 1:300
Chicken anti-TBR1 (Millipore, AB2261, Lot. NG1820262), used 1:200
Mouse anti-TLE4 (Santa Cruz, sc-365406, Lot. C0218), used 1:200
Mouse anti-HOPX (Santa Cruz, sc-398703, Lot. F1615), used 1:100
Mouse anti-RELN (MBL, D223-3), used 1:200

Rabbit anti-BRN2 (GeneTex, GTX114650), used 1:250
Rabbit anti-cCas3 (CST, #9661, Lot. 45), used 1:200
Rabbit anti-SOX2 (CST, #3579, Lot. 5). used 1:200
Mouse anti-TBR2 (R&D, MAB6166), used 1:100
Rabbit anti-GFAP (DAKO, Z0334, Lot. 20035993), used 1:1000
Rat anti-GFAP (Thermo Fisher Scientific, 13-0300), used 1:1000

Validation

We have used or validated some of the antibodies in previous studies (Pasca et al., Nature Methods 2015; Sloan et al., Neuron 2017; Birey et al., Nature 2017; Sloan et al., Nature Protocols, 2018), such as SATB2, CTIP2, TBR1, HOPX, RELN, BRN2, SOX2, TBR2, GFAP in human cells. Moreover, most of these antibodies have been used in other studies:
The mouse anti-SATB2 (Abcam, AB51502) has been used in 87 studies according manufacturer's website, and tested for immunofluorescence staining in human fetal brains (Ozai et al., 2018).
The rat anti-CTIP2 (Abcam, AB18465) has been used in 240 studies according manufacturer's website, and tested for immunofluorescence staining in human fetal brains (Ozai et al., 2018).
The rabbit anti-TBR1 (Abcam, AB31940) has been used in 157 studies according manufacturer's website, and tested for immunofluorescence staining in human fetal brains (Ozai et al., 2018).
The chicken anti-TBR1 (Millipore, AB2261) has been tested in IHC and used in 1 study as stated on the manufacturer's website, and also tested for immunocytochemistry in human cells (Parween et al., 2017).
The mouse anti-TLE4 (Santa Cruz, sc-365406) has been used in 6 studies according manufacturer's website, and tested for immunofluorescence staining in human cells (Yang et al., 2015).
The mouse anti-HOPX (Santa Cruz, sc-398703) has been tested in IHC and used in 1 study as stated on the manufacturer's website.
The mouse anti-RELN (MBL, D223-3) has been used in 11 studies according manufacturer's website.
The rabbit anti-BRN2 (GeneTex, GTX114650) has been used in 6 studies, and tested for immunofluorescence staining in human cells according manufacturer's website.
The rabbit anti-cCas3 (CST, #9661) has been used in 3048 studies according manufacturer's website, and also tested for immunocytochemistry in human cells (Imaizumi et al., 2018).
The rabbit anti-SOX2 (CST, #3579) has been used in 93 studies according manufacturer's website, and tested for immunofluorescence staining in human cells (Kogut et al., 2018).
The mouse anti-TBR2 (R&D, MAB6166) has been used in 1 study according manufacturer's website.
The rabbit anti-GFAP (DAKO, Z0334) has been used in 8 studies according manufacturer's website, and tested for immunofluorescence staining in human cells (HD iPSX Consortium, 2017).
The rat anti-GFAP (Thermo Fisher Scientific, 13-0300) has been used in 115 studies according manufacturer's website, and tested for immunofluorescence staining in human fetal tissue (Errede et al., 2014).

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	H20961 line was obtained from the Gilad laboratory. TUBA1B-mEGFP (AICS-0012) and LMNB1-mEGFP (AICS-0013) lines were purchased from Coriell Institute.
Authentication	iPSC lines were authenticated by SNP arrays.
Mycoplasma contamination	All iPSC lines used in this study were tested for Mycoplasma free.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.