Human cerebral organoids recapitulate gene expression programs of fetal neocortex development

J. Gray Camp[a,1], Farhath Badsha[b,1], Marta Florio[b], Sabina Kanton[a], Tobias Gerber[a], Michaela Wilsch-Bräuninger[b], Eric Lewitus[c], Alex Sykes[b], Wulf Hevers[a], Madeline Lancaster[d,e], Juergen A. Knoblich[e], Robert Lachmann[f], Svante Pääbo[a,2], Wieland B. Huttner[b,2], and Barbara Treutlein[a,b,2]

[a]Max Planck Institute for Evolutionary Anthropology, Department of Evolutionary Genetics, 04103 Leipzig, Germany; [b]Max Planck Institute of Molecular Cell Biology and Genetics, 01307 Dresden, Germany; [c]Institut de Biologie, Ecole Normale Superieure, 75005 Paris, France; [d]Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom; [e]Institute of Molecular Biotechnology of the Austrian Academy of Sciences (IMBA), 1030 Vienna, Austria; and [f]Technische Universität Dresden, Universitätsklinikum Carl Gustav Carus, Klinik und Poliklinik für Frauenheilkunde und Geburtshilfe, 01307 Dresden, Germany

Contributed by Svante Pääbo, October 27, 2015 (sent for review September 22, 2015; reviewed by Colette Dehay, Yukiko Gotoh, and Rickard Sandberg)

Cerebral organoids—3D cultures of human cerebral tissue derived from pluripotent stem cells—have emerged as models of human cortical development. However, the extent to which in vitro organoid systems recapitulate neural progenitor cell proliferation and neuronal differentiation programs observed in vivo remains unclear. Here we use single-cell RNA sequencing (scRNA-seq) to dissect and compare cell composition and progenitor-to-neuron lineage relationships in human cerebral organoids and fetal neocortex. Covariation network analysis using the fetal neocortex data reveals known and previously unidentified interactions among genes central to neural progenitor proliferation and neuronal differentiation. In the organoid, we detect diverse progenitors and differentiated cell types of neuronal and mesenchymal lineages and identify cells that derived from regions resembling the fetal neocortex. We find that these organoid cortical cells use gene expression programs remarkably similar to those of the fetal tissue to organize into cerebral cortex-like regions. Our comparison of in vivo and in vitro cortical single-cell transcriptomes illuminates the genetic features underlying human cortical development that can be studied in organoid cultures.

cerebral organoid | neocortex | corticogenesis | single-cell RNA-seq | stem cells

Elucidating the cellular and molecular basis of human neocortex development and evolution has profound importance for understanding our species-specific cognitive abilities as well as our susceptibility to neurodevelopmental diseases. Neurons of the human neocortex originate during embryogenesis from cell divisions of a variety of neural progenitor cells (NPCs) located in compartmentalized germinal zones. NPC types differ in cell morphology, cell polarity, capacity to self-renew, lineage relationships, and location of mitosis (1). Apical progenitors (APs), including apical (or ventricular) radial glia (aRG), divide at the apical surface of the ventricular zone (VZ), whereas their derivative basal progenitors (BPs), including basal (or outer) radial glia (bRG) and basal intermediate progenitors (bIPs), lack apical contact and divide in the inner and outer subventricular zone (iSVZ and oSVZ) (2). In humans, both aRG and bRG are able to self-amplify by symmetric proliferative divisions. They also share the capacity to divide asymmetrically to self-renew while producing neurons either directly or via bIPs (3-6). In humans, bIPs further amplify the neuronal output of aRG and bRG by undergoing additional rounds of symmetric division before self-consuming into pairs of neurons (1, 7). Newborn neurons migrate radially from these germinal zones to finally establish in the cortical plate (CP) (8). The evolutionary expansion of the human neocortex has been linked to an increase in the proliferative potential of particular pools of NPCs, notably BPs, resulting in greater numbers of neocortical neurons (1, 2, 5, 9). Because of the challenges associated with primate experimentation, the mouse has been widely used as a model system for understanding human cortical neurogenesis. However, due to its evolutionary distance and divergent physiology, it is debatable how relevant the rodent genomic and developmental background is. Therefore, systems effectively recapitulating human cortical development are required.

Recently, self-organizing structures reminiscent of the developing human brain have been generated from pluripotent stem cells [embryonic stem cells (ESCs) or induced pluripotent stem cells (iPSCs)], presenting a unique opportunity to model human cerebral development in vitro (10). So-called cerebral organoids are generally heterogenous and allow the formation of a variety of brain-like regions including the cerebral cortex, ventral forebrain, midbrain–hindbrain boundary, and hippocampus (10). The protocol to grow cerebral organoids was designed to mimic early stages of forebrain development and then rely on the intrinsic self-organizational capacity of the cells to pattern, specify, and generate structured cerebral tissue (11). This protocol establishes cortical-like tissue with compartmentalized germinal zones including a VZ, where aRG-like cells line a lumen, express RG marker genes, undergo interkinetic nuclear migration, and divide at the apical surface, similar to their in vivo counterparts. In addition, time-lapse microscopy and immunostainings for bIP markers [e.g., TBR2;

Significance

We have used single-cell RNA sequencing to compare human cerebral organoids and fetal neocortex. We find that, with relatively few exceptions, cells in organoid cortex-like regions use genetic programs very similar to fetal tissue to generate a structured cerebral cortex. Our study is of interest, as it shows which genetic features underlying human cortical development can be accurately studied in organoid culture systems. This is important because although cerebral organoids have great promise for modeling human neurodevelopment, the extent to which organoids recapitulate neural progenitor proliferation and differentiation networks in vivo remained unclear.

Author contributions: J.G.C., F.B., S.P., W.B.H., and B.T. designed research; J.G.C., F.B., M.F., S.K., T.G., M.W.-B., E.L., A.S., and W.H. performed research; M.F., M.L., J.A.K., and R.L. contributed new reagents/analytic tools; J.G.C. and B.T. analyzed data; and J.G.C., F.B., M.F., S.P., W.B.H., and B.T. wrote the paper.

Reviewers: C.D., Stem Cell and Brain Research Institute, INSERM U846; Y.G., The University of Tokyo; and R.S., Karolinska Institutet.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the NCBI Gene Expression Omnibus (GEO) database, www.ncbi.nlm.nih.gov/geo (accession no. GSE75140).

[1]J.G.C. and F.B. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: barbara_treutlein@eva.mpg.de, paabo@eva.mpg.de, or huttner@mpi-cbg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1520760112/-/DCSupplemental.

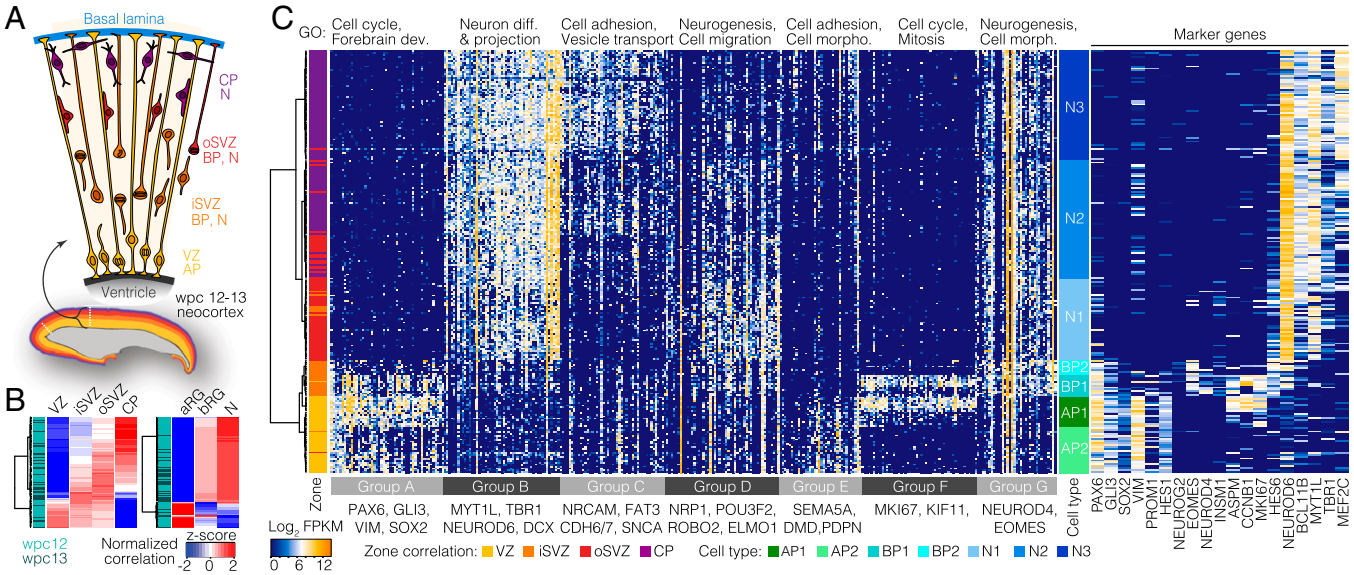

Fig. 1. Deconstructing cell composition in the fetal human neocortex. (*A*) scRNA-seq was performed on cells from two human neocortex specimens at 12–13 wpc. Schematic shows NPC types (APs, BPs) and neurons (N) enriched in zones within the human neocortex at midneurogenesis. AP, apical progenitor; BP, basal progenitor; CP, cortical plate; iSVZ, inner subventricular zone; oSVZ, outer subventricular zone; VZ, ventricular zone. (*B*) Heat maps show normalized correlation (Z-score) of single-cell transcriptomes from human wpc 12 (light green) and wpc 13 (dark green) cerebral cortex with bulk RNA-seq data from laser-microdissected zones (left, 18) or FACS-purified cell types (right, 19) from the human neocortex at the same developmental stage. (*C*) Hierarchical clustering of scRNA-seq data reveals cell types in the human fetal cortex. Each row represents a single cell and each column a gene. Genes were discovered using PCA (*SI Methods*). The maximum correlation to bulk RNA-seq data from germinal zones is shown in the left sidebar. Cell-type assignment is shown on the right sidebar. Expression of genes used to classify APs, BPs, newborn neurons (N1), and maturing neurons (N2 and N3) are show to the right of the cell-type assignment bar. Top GO enrichments are shown above the heat map, with representative genes listed below.

also known as (aka) EOMES] and neurons (TUJ1, aka TUBB3) revealed patterns of direct and indirect neurogenesis in an abventricular location, reminiscent of the SVZ. bRG-like cells, albeit more rarely than in the in vivo neocortex, were also observed. The observations of a cortical architecture and cellular behavior based largely on marker gene expression are generally consistent with in vivo patterns of human fetal cortical development. Bulk transcriptome analysis has provided a genome-wide perspective of the gene expression landscape in whole organoids (12). However, due to heterogeneity within cerebral organoids, bulk methods obscure cell-type diversity and provide coarse insight into organoid neurogenic programs. Moreover, the extent to which organoid cell composition and lineage progression are equivalent to fetal cortical tissue remains unresolved.

Single-cell RNA-seq (scRNA-seq) is a powerful method to unbiasedly identify molecularly distinct cell types in heterogeneous tissues (13–16). ScRNA-seq can also identify cells as intermediate between types, and cells can then be computationally aligned to delineate differentiation paths (15, 17). We sought to use single-cell transcriptomes to analyze the cellular composition of human cerebral organoids and fetal human cortical tissue to reconstruct the differentiation paths and gene networks underlying human corticogenesis. We find that, with some exceptions, the same genes used to build cortical tissue in vivo characterize corticogenesis in vitro. Our data thus show that genetic features underlying human cortical development can be accurately studied in organoid culture systems.

Results

Reconstructing Lineage Relationships in the Human Fetal Cerebral Cortex. To compare corticogenesis in organoids with primary fetal tissue, we first dissected the cellular composition and lineage relationships in human fetal neocortex. We analyzed 226 single-cell transcriptomes from 12- and 13-wk postconception (wpc) human neocortex specimens (Fig. 1*A*) and performed principal components analysis (PCA) on all variable genes (variance > 0.5) expressed in more than two cells to survey cell diversity in this dataset (Fig. S1*A*). The genes correlating and anticorrelating with the first principal component from the PCA distinguished NPCs and neurons, respectively, and highlighted intermediates between the two (Fig. S1*B*). We identified a single cell expressing *PECAM1* and many other endothelial markers and five interneurons derived from the ventral telencephalon (marked by *GAD1*, *DLX1/2/5/6*, and *ERBB4*; Fig. S1*C*), and we excluded these cells from further analyses. We next used all expressed genes to characterize cells based on their maximum correlation with bulk RNA-seq data from laser-dissected germinal zones, namely VZ (containing a majority of aRG), iSVZ and oSVZ (containing a majority of bIPs, bRG neurons, and migrating neurons), and CP (containing mature neurons) (18), or FACS-purified aRG, bRG, and neurons (19) (Fig. 1*B*). The transcriptome of each single fetal neocortical cell correlated maximally with one of these four different zones (CP, 100 cells; ISVZ, 24 cells; OSVZ, 62 cells; VZ, 39 cells) and with one of the purified cell populations representing diverse RG subtypes and neurons at different stages of maturation. Therefore, our scRNA-seq data likely include APs, BPs, and neurons, representing the range of neural cell types present in the human neocortex at this time point.

We performed PCA again on all fetal transcriptomes (excluding the endothelial cell and interneurons), as well as on NPCs and neurons separately, to identify genes that are most informative for defining cell subpopulations. Using these genes (Dataset S1), we hierarchically clustered cells and identified seven major cell clusters that we classified as different populations of APs, BPs, and neurons based on the expression of known cell-type marker genes and correlations with bulk data (Fig. 1*C*). Two clusters (AP1 and AP2) had a high correlation with bulk VZ, expressed genes involved in forebrain development (group A) and cell adhesion (group E), and were marked by coexpression of RG markers *PAX6, GLI3, SOX2, HES1,* and *VIM* and the AP marker *PROM1*. AP1 cells (16 cells) expressed cell-cycle genes (group F) enriched during S/G2/M phase progression (e.g., *MKI67*), whereas AP2

cells (24 cells) did not. This suggests that AP1 and AP2 may contain APs in S/G2/M and G1, respectively. We identified a cell cluster (BP1, 11 cells) that had maximum correlations with bulk iSVZ and expressed RG markers *PAX6*, *GLI3*, and *ASPM* as well as early BP markers *INSM1*, *EOMES*, and *HES6* but lacked expression of *VIM* and *SOX2*. In addition, these BP1 cells expressed a core of proneural genes, including *NEUROD4*, as well as many of the same S/G2/M cell-cycle regulators expressed in AP1 cells in the VZ (groups A, F, and G). This gene expression profile is compatible with AP-to-BP transition or an early BP. Furthermore, we identified a cluster (BP2, 10 cells) containing cells with highest correlation with bulk iSVZ and oSVZ and that was defined by coexpression of *EOMES* and early neuronal genes including *NEUROD6* (group B) but did not express *ASPM* and other S/G2/M cell-cycle regulators. This cluster is likely to contain more mature BPs that are committed to the neurogenic fate and are in the G1 phase.

In addition to NPCs, we found three clusters (neuron clusters 1–3, N1–3) distinguished by the lack of an NPC signature. Each of these clusters expressed a large group of neuronal genes (group B, including *MYT1L, BCL11B,* and *TBR1*) enriched for GO (gene ontology) terms involved in neuronal differentiation. One neuron cluster (N1, 39 cells) correlated with iSVZ and oSVZ bulk data and expressed genes involved in various aspects of neurogenesis, including neuronal migration (group D). This cluster likely contains newborn neurons en route to the CP. A second neuron cluster (N3, 67 cells) correlated with CP bulk data and expressed genes involved in cell adhesion and vesicle transport (group C), indicative of a more mature stage of neuronal differentiation. The third neuron cluster (N2, 53 cells) appears to contain cells that are transitioning between these N1 and N3 neurons. The gene expression profiles of these clusters are therefore compatible with cortical neurons at different stages of maturation.

We next inferred lineage relationships among the cells using Monocle, an algorithm that combines differential gene expression, dimension reduction, and minimal spanning tree construction to link cells along a pseudotemporally ordered path (17) (Fig. 2*A*). This revealed a cortical lineage that linked APs in the VZ, through BPs and newborn neurons in the iSVZ and oSVZ, to CP neurons. This organization was corroborated using an adjacency network based on pairwise correlations between cells (Fig. S2*A*). Notably, expression of genes known to be enriched in APs (e.g., *SOX2*), BPs (e.g., *EOMES*), and neurons (e.g., *MYT1L*) exhibited restricted expression along the lineage path (Fig. 2*B*). Heat map visualization of gene expression with cells ordered according to their pseudotemporal position along the lineage revealed a temporal sequence of gene regulatory events during differentiation and many cells at intermediate stages (Fig. S2*B*).

We constructed a transcription factor (TF) correlation network to understand the gene expression changes occurring during transitions through the AP–BP–neuron lineage (Fig. 2*C*). The network revealed two densely connected subnetworks regulating NPC proliferation and neuronal differentiation, respectively, linked by a series of TFs regulating the transition between the two major cell states. This unbiased approach highlighted known master regulators as central to each subnetwork. For example, *HES1*, *SOX9*, *PAX6*, and *SOX2* are all tightly connected and highly expressed in most APs of the VZ and are down-regulated shortly after neuronal lineage commitment occurring in the SVZ. *ASCL1*, *EOMES*, *NEUROD4*, *HES6*, and *INSM1* are all expressed concomitant with the AP-to-BP transition, consistent with their role in delamination and early neuronal specification. Another tightly connected subnetwork corresponds to genes such as *TBR1*, *MYTL1*, *BCL11A/B*, and *NEUROD6*, whose expressions are up-regulated concomitant with neuron differentiation and maintained throughout later stages of neuronal maturation. These data largely confirm the current knowledge of transcriptional regulation during cortical neurogenesis.

In conclusion, the analysis of single-cell transcriptomes in the developing human neocortex can faithfully reconstruct genetic and cellular networks involved in germinal zone organization, NPC proliferation, and NPC-to-neuron differentiation (Fig. S2*C*).

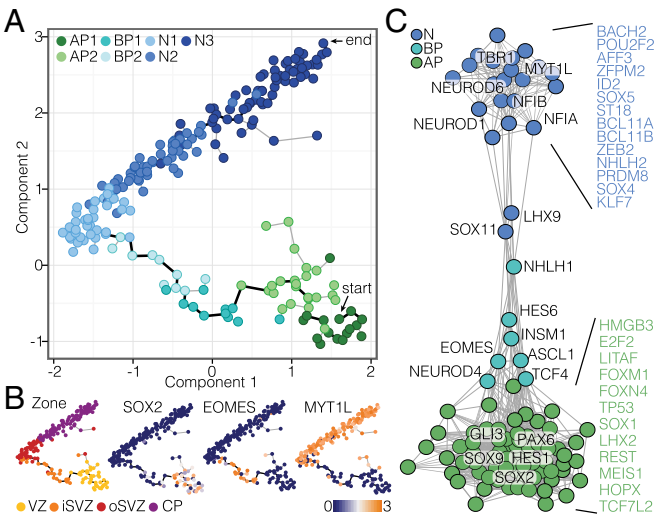

Fig. 2. Reconstructing lineage relationships in the fetal neocortex. (*A*) Monocle reveals an AP–BP–neuron lineage that correlates with the zones of the developing neocortex. Cells (circles, colored based on cell type) are arranged in the 2D independent component space based on genes used to classify cells in Fig. 1*C*. The minimal spanning tree (gray lines) connects cells, with the black line indicating the longest path. (*B*) Monocle plots with single cells are colored based on the maximum correlation with bulk RNA-seq data from cortical zones (*Far Left*) or gene expression that distinguishes the lineage transitions (*Middle Left*, *Middle Right*, and *Far Right*). (*C*) Transcription factor (TF) correlation network during lineage progression. Shown are nodes (TFs) with more than three edges, with each edge reflecting a high correlation (>0.3) between connected TFs.

Dissecting Cell Composition in Human Cerebral Organoids. To survey the cellular composition of human cerebral organoids at different stages of development, we analyzed 333 single-cell transcriptomes from five whole organoids (days 33, 35, 37, 41, and 65, iPSC-derived) and 175 single-cell transcriptomes from four microdissected cortical regions (53 d, r1, r2, ESC-derived; 58 d, r3, r4, iPSC-derived) (Fig. 3 *A–C* and Fig. S3). We combined all single-cell transcriptomes and performed PCA to identify genes most informative for defining cell populations. Using these genes (Dataset S2), we used t-distributed Stochastic Neighbor Embedding (t-SNE) to reduce the complexity of the data and visualize cell relationships in a 2D space (Fig. 3*D*). This approach organized the cells into 11 transcriptionally distinct clusters, and we performed differential expression analysis to identify genes marking each cluster (Fig. 3 *D* and *E*, Figs. S3 and S4, and Dataset S2). In this way, we identified multiple types of progenitors, neurons, and mesenchymal cells from different regions of the cerebral organoid.

We found that clusters c1, c2, and c3 are NPCs, and cluster c4 is neurons from organoid dorsal forebrain-like regions (cerebral cortex) based on enriched expression of genes expressed in nearly all fetal cerebral cortex NPCs or neurons (i.e., NPCs and Ns, *FOXG1, NFIA,* and *NFIB;* Ns, *NEUROD6*). Clusters c5 and c6 are NPCs, and c7 is neurons that lack expression of these genes that were expressed in the fetal cortex. Instead, many NPCs in c5 and c6 express *OTX2* and are likely a mixture of ventral telencephalic or hippocampal NPCs and NPCs from immature dorsal telencephalic regions. c7 is composed of neurons from ventral forebrain-like structures and includes interneuron-like cells (Fig. S4). c8 and c9 contain cycling and noncycling cells that express R-spondin genes and WNT2B and are likely from the hem signaling center in the dorsal/ventral boundary region. c10 and c11 contain cycling

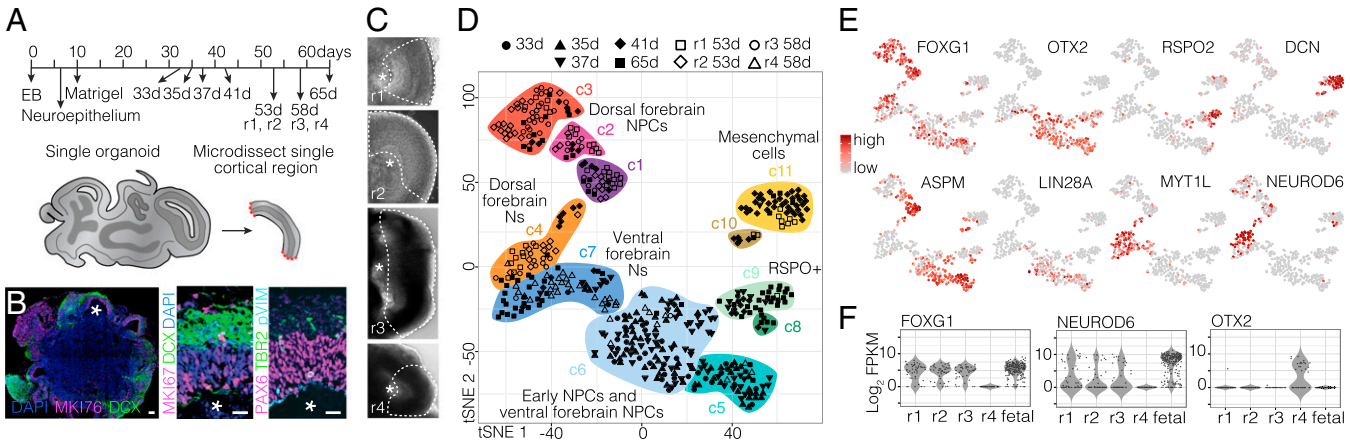

Fig. 3. Dissecting cerebral organoid cell composition using scRNA-seq. (*A*) scRNA-seq was performed on whole organoids dissociated at 33, 35, 37, 41, and 65 d after EB culture and four microdissected regions surrounding single ventricles from two organoids (day 53, r1, r2, ESC-derived; day 58, r3, r4, iPSC-derived). (*B*) The 30 d iPSC-derived organoid immunostained with proliferation marker MKI67 (magenta), neuronal marker DCX (doublecortin) (green), and DAPI (blue). (Scale bar, 100 μm.) Zoom to ventricle (asterisk) shows MKI67, DCX, and DAPI (*Left*), and NPC marker PAX6 (magenta), BP marker TBR2/EOMES (green), and mitosis marker phospo-vimentin (pVim; cyan) (*Right*). (Scale bar, 20 μm.) (*C*) Images of microdissected cortical regions (r1 and r2 from 53 d ESC-derived organoid; r3 and r4 from 58 d iPSC-derived organoid). Dotted lines show microdissection boundaries; asterisks mark ventricles. (*D*) PCA and unbiased clustering using t-SNE reveals cell populations within organoids. Shapes indicate experiments, and colors represent significant clusters. See *Dissecting Cell Composition in Human Cerebral Organoids* and *SI Results* for cluster descriptions. (*E*) Marker genes for each cluster. Cells are colored based on expression level. Cerebral cortex cells (c1, c2, c3, and c4) have high expression of *FOXG1* and low expression of *OTX2*. (*F*) Violin plot shows *FOXG1*, *NEUROD6*, or *OTX2* expression from each microdissected region compared with fetal cortex.

and noncycling mesenchymal cells that express extracellular matrix (ECM) genes and surround the periphery of cortical regions (Fig. 3*E* and Fig. S3 *G* and *H*). Please see *SI Results* for a detailed analysis and discussion of organoid cell-type composition.

We observed that each microdissected cortical-like region contained NPCs and neurons (Fig. 3*D*). However, cells from three of the four regions were *FOXG1*-positive, were *OTX2*-negative, expressed cerebral cortex markers, and were contained within the dorsal cortex clusters in the t-SNE analysis (Fig. 3 *D–F* and Fig. S3). In contrast, cells in the fourth cortical region did not express *FOXG1* or other fetal cortex markers (i.e., *NFIA*, *NFIB*, *NEUROD6*) but instead expressed *OTX2* and were contained within ventral forebrain clusters (c5, c6, and c7). Thus, individual cerebral organoids contain cortical regions with different forebrain identities, which we could discriminate due to distinct signatures of NPC and neuron populations.

Reconstructing Lineages in the Organoid Cerebral Cortex. We characterized organoid cortex-like cells (clusters 1, 2, 3, and 4; 157 cells in total) based on their maximum correlation with bulk RNA-seq data from laser-dissected germinal zones (18) or FACS-purified NPC subpopulations (Fig. S5*A*) (19). We found correlation patterns similar to the fetal tissue, suggesting that the organoid cells represent a similar range of cell types as in the fetal cortex data (CP, 27 cells; ISVZ, 3 cells; OSVZ, 19 cells; VZ, 110 cells). We next classified the organoid cells by determining the fetal cell type with which they correlate most strongly, resulting in 57 AP1, 57 AP2, 1 BP1, 6 BP2, 4 N1, 16 N2, and 16 N3 cells. Notably, we observed proportionally less BP cells (6% of NPCs) in the organoid than in the fetal tissue (34%) (Fig. S5).

As with the fetal tissue, we next performed PCA on organoid cerebral cortex-like cells and used the genes with the highest PC1–3 loadings (Dataset S3) to infer lineage relationships using Monocle as well as construct an intercellular correlation network (Fig. 4 *A* and *B* and Fig. S5*B*). Similar to the fetal tissue, the minimal spanning tree revealed a lineage path from AP through BP to the neuron, with *PAX6*, *EOMES*, and *MYT1L* exhibiting restricted expression along the lineage. Also, a side branch from the main lineage suggested rare alternative paths to neuronal fate. The adjacency network graph revealed multiple connections from AP and BP to the neuron and also highlighted AP self-renewal and proliferation in cells correlating with VZ bulk data (Fig. S5*B*). In addition, BP cells correlating with iSVZ and oSVZ were organized as intermediates between AP cells in the VZ and CP neurons. In general, this analysis confirmed that cells in the organoid cortex are organized in a zonal and cellular hierarchy consistent with what we observed in the fetal tissue.

Heat map visualization of cells ordered according to their pseudotemporal position along the lineage revealed a temporal sequence of gene regulatory events during differentiation, with many cells at intermediate stages (Fig. S5*C*). GO enrichments similar to the fetal tissue (cell cycle, mitosis, neuron projection and differentiation, forebrain development, synapse formation, and migration) were observed for gene groups that change expression along the lineage (Dataset S3). TFs involved in fetal neurogenesis showed a similar expression pattern and network topology in the organoid (Fig. 4 *B–D* and Fig. S5*D*). PCA and hierarchical clustering of organoid and fetal cells combined showed that organoid and fetal cells were distributed together within the two main subclusters representing NPCs and neurons (Fig. 4*E*). In addition, fetal and organoid cells intermix in a cell lineage network based on genes describing AP proliferation and AP–BP–neuron differentiation (Fig. 4*E*), with intercellular correlations being relatively constant along the lineage (Fig. S5). Thus, the major proportion of the variation in these data is not between in vitro and in vivo tissues but among cell states during neurogenesis.

Similarities in Neurogenic Programs Between Organoid and Fetal Cortex. We compared fetal and organoid expression of genes involved in several important cell biological processes during cortex development. Extracelluar matrix (ECM) proteins are important for NPC self-renewal (18). We analyzed the expression of genes involved in ECM production and sensing in cells from the fetal tissue and organoids (Fig. S6*A*). We find that fetal APs express collagens (*COL11A1*, *COL4A5*), numerous glycoproteins implicated in ECM signaling (*SDC4*, *LAMA1/5*, *BCAN*), integrin receptors (*ITGA6*, *ITGB8*), and glypican coreceptors (*GPC4/6*). Most of these genes (16/18, 89%) have similar AP-specific expression patterns in organoid cells. Similarly, we found strongly positive correlations between fetal and organoid cell types for genes

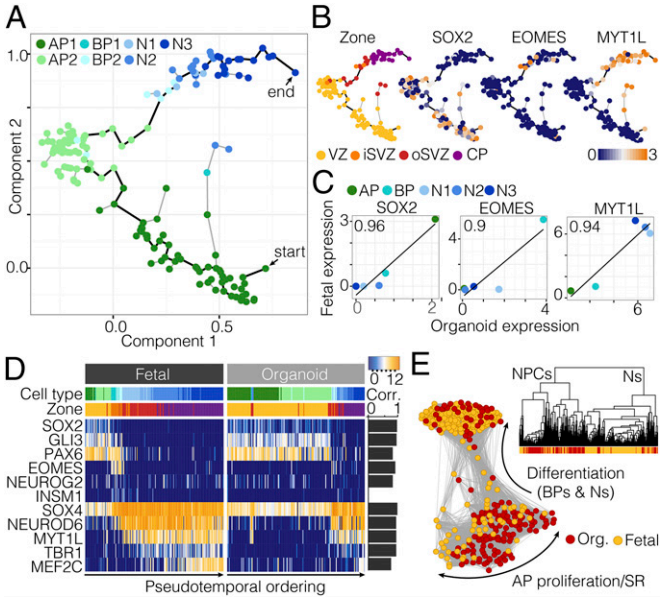

Fig. 4. Similar gene expression profiles characterize lineage progression in organoid and fetal cerebral cortex. (*A*) Organoid AP–BP–neuron lineage. PCA on organoid dorsal cortex cells identified genes used for Monocle. The minimal spanning tree (gray lines) connects cells (circles, colored by cell type). Black line indicates the longest path (Dataset S3 and Fig. S5). (*B*) Monocle plot with cells colored by maximum correlation with bulk RNA-seq data from germinal zones (*Far Left*) or marker gene expression (*Middle Left*, *Middle Right*, and *Far Right*). (*C*) Scatter plot shows correlation (Pearson) between fetal and organoid average expression per cell type for marker TFs. (*D*) Heat map shows TF expression in organoid and fetal cells ordered by pseudotime. Top bars show cell type and maximum correlation with germinal zones. Each TF's expression was averaged across cells of a given type (AP, BP, N1, N2, and N3), and the Pearson correlation between fetal and organoid cell types is shown to the right of the heat map. (*E*) Cell lineage network and dendrogram (*Top Right*) based on pairwise correlations between fetal (orange) and organoid (red) cells show that NPCs and neurons intermix.

involved in transcription regulation (10/11; 90%), RG delamination (12/15, 80%), Notch/Delta signaling (7/10, 70%), and neurite outgrowth (24/25; 96%) (Fig. 4 *D* and *E* and Fig. S6 *B–D*).

We next collected genes that (*i*) encode proteins with fixed amino acid changes in modern humans since divergence with Neanderthals (modHuman) (20), (*ii*) are mutated in human genetic disorders affecting neurogenesis (OMIM; omim.org/), (*iii*) are located nearby evolutionarily conserved sequences that have been specifically lost in the human lineage (hCondel) (21), and (*iv*) are nearby human-accelerated regions overlapping brain-accessible chromatin (haDHS) (22) (Dataset S4). We projected these changes onto a gene regulatory network inferred from correlations with the TFs that control the AP–BP–neuron lineages in the fetal cortex (Fig. 5). The expression of the majority of these genes (average, 82.5%) that may carry changes relevant to human cerebral cortex development and evolution was positively correlated between fetal and organoid cell types, suggesting these changes may be faithfully modeled in cerebral organoids.

Differences Between Organoid and Fetal Cerebral Cortex. Finally, we searched for differences in gene expression between fetal and organoid APs (AP1 and AP2 combined) and neurons (N1, N2, and N3 combined). We excluded BPs from this analysis because of the small number of BPs in the organoid dataset. We combined fold difference and the median receiver operating characteristic (ROC) test (23) to identify genes that had a high average difference (>threefold) and power to discriminate (97th percentile) between fetal and organoid cells (Dataset S4 and Fig. S7). Among genes up-regulated in the organoid, the top GO category enriched was "response to organic substance," which contained genes *FOS*, *EGR1*, *CSNK2B*, *HMGCS1*, *BCHE*, *HERPUD1*, *CLIC1*, and *ADIPOR2*. The immediate early genes *FOS* and *EGR1* are Notch signaling targets recently reported to be expressed in human but not mouse radial glia (24). Among the genes down-regulated in organoid neurons was a transporter for vitamin A (*RBP1*), which might reflect a response to the inclusion of vitamin A in the culture medium. *TUBB* (tubulin, beta class I), a structural component of microtubules, was the most differentially expressed gene with higher expression in both progenitors and neurons in organoids. Other potentially relevant differences include *PRDM8*, which has a role in assembling neuronal circuits in upper layer cortical neurons (25), and *NFIX*, which is involved in progenitor cell differentiation (26), both of which have higher expression in fetal neuronal cells than in organoid cells. Most of these genes differed in both ESC- and iPSC-derived organoids relative to the fetal tissue. Differences between fetal and organoid cells are significant compared with differences between randomly selected fetal APs (or neurons), however <5% of differentially expressed genes between fetal and organoid cells reach the average classification powers seen between APs and neurons. We conclude that the major gene expression differences between organoid and fetal tissues seem to be a response to tissue culture environment and may not reflect fundamental differences in differentiation programs.

Discussion

Single-cell gene expression analysis is a powerful technique to deconstruct tissue heterogeneity and has recently been used to characterize NPC and neurons in the fetal brain (24, 27–29). Here we applied scRNA-seq to compare cell composition and lineage relationships in fetal and organoid cerebral cortex. We find that over 80% of genes implicated in neocortex disease or evolution and are differentially expressed along the fetal cortex lineage have similar expression profiles in organoid and fetal cerebral cortex. Organoid cells thus use similar sets of genes as their fetal counterparts to perform cortical processes such as NPC proliferation and self-renewal, production of ECM, migration, adherence, delamination, and differentiation that result in structured cerebral tissue. This has important implications for using the organoid

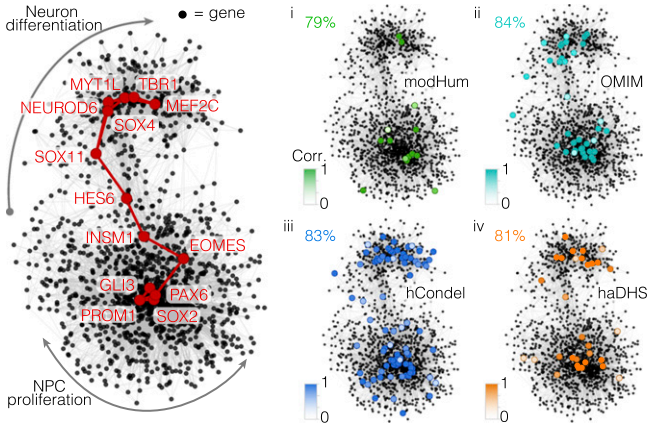

Fig. 5. Genomic scans of disease, evolutionary, and chromatin signatures highlight genetic aspects of human corticogenesis that can be modeled in vitro. Shown is the covariation network using genes that have high correlation (>0.3) with TFs controlling the AP–BP–neuron lineage from Fig. 2*C*. Select TF nodes are highlighted to delineate the path. (*i–iv*) Panels show genes that have amino acid changes that are modHum (*i*, green), OMIM (*ii*, turquoise), hCondel (*iii*, blue), or haDHS (*iv*, orange). The percentage of cells that have a positive correlation (>0.4) between fetal and organoid cells is shown, with nodes colored based on the correlation coefficient (Dataset S4).

system to model developmentally and evolutionarily relevant neocortical cell biology.

Our sampling of cells over development was relatively modest, and future higher throughput studies over an expanded time course are required. A study published while our manuscript was in review used scRNA-seq to survey the human cerebral cortex at gestational week 16–18 and reported the identification of a bRG (aka oRG) gene expression signature (29). Most of these bRG marker genes are expressed at 12–13 wpc in cells that we define as APs that highly correlate with VZ and purified aRG bulk RNA-seq data. This is consistent with the published results that the bRG signature emerges from the VZ after gestational week 13.5.

Many of the genes that differed in expression between fetal and organoid cells had relatively low expression in fetal tissue, suggesting that the differential expression observed may be due to noise in single-cell transcriptomics, whereas others seem to reflect responses to factors in the cerebral organoid culture media. One notable difference between fetal and organoid tissue was that the organoids had fewer BPs relative to APs. This might be explained by developmental time point or an underdeveloped SVZ intrinsic to the culture system. Future work will be required to understand if these differences have any bearing on corticogenesis and if culture conditions can be modified to further improve how organoid cultures reflect in vivo cerebral cortex development.

Methods

Research involving human tissue and human embryonic stem cells was approved by the Institutional Review Board of the Max Planck Institute of Molecular Cell Biology and Genetics, composed of the following people: Jussi Helppi (Chair, Head of the Animal Facility), Barbara Langen (designated veterinarian and animal welfare officer), Ronald Naumann (Head of the Transgenic Core Facility), Anke Münch-Wuttke (Animal Care Expert), and Ivan Baines (Chief Operating Officer, Max Planck Institute of Molecular Cell Biology and Genetics). In addition, the Universitätsklinikum Carl Gustav Carus of the Technische Universität Dresden Ethical Review Committee approved the research with human fetal tissue (see *SI Methods*, *Dissociating fetal tissue*).

Human fetal brain tissue (12–13 wpc) was obtained with ethical approval following elective pregnancy termination and informed written maternal consents. ESC and iPSC lines were grown in mTESR1 (Stem Cell Technologies) using standard protocols. Cerebral organoids were generated as described (10, 11), with the exception that mTESR1 was used during embryoid body (EB) formation. For immunohistochemistry and electron microscopy of organoids, see *SI Methods*. Fetal cortices were processed as described in ref. 19. Whole cerebral organoids and microdissected regions were dissociated in 2 mL of Accutase (StemPro) containing 0.2 U/μL Dnase I (Roche) for 45 min. Dissociated fetal and organoid cells were filtered through 40-, 30-, and 20-μm-diameter strainers to create a single-cell suspension. Cell viability (90–95%) was assessed using Trypan blue staining. Single-cell capture, lysis, and cDNA synthesis were performed with the Fluidigm C1 system using the SMARTer Ultra Low RNA Kit for Illumina (Clontech). cDNA size distribution was assessed by high-throughput capillary gelelectrophoresis (Advanced Analytical). Sequencing libraries were constructed in 96-well plates using the Illumina Nextera XT DNA Sample Preparation Kit (15). Up to 96 single-cell libraries were pooled, and each cell was sequenced 100 bp paired-end on Illumina HiSEq. 2500 to a depth of 2–5 million reads. A detailed description of methods and scRNAseq analyses is provided in *SI Methods*.

ACKNOWLEDGMENTS. We thank E. Taverna for comments on the manuscript, M. Renner for cotutoring organoid culture, Y. J. Chang and M. T. Garcia for help with imunohistochemistry, and B. E. Daniel for discussions. We thank K. Köhler and A. Weigert for help with iPSC culture techniques and B. Höber, A. Weihmann, J. Kelso, G. Renaud, and M. Dannemann of Max Planck Institute for Evolutionary Anthropology for DNA sequencing and bioinformatics support. J.A.K. is funded by the Austrian Academy of Sciences, Austrian Science Fund Grants I_552-B19 and Z_153_B09, and an advanced grant of the European Research Council (ERC). S.P. was supported by the Allen Family Foundation. W.B.H. was supported by Deutsche Forschungsgemeinschaft (DFG) Grant Sonderforschungsbereich 655 A2, ERC Grant 250197, the DFG-funded Center for Regenerative Therapies Dresden, and the Fonds der Chemischen Industrie.

1. Florio M, Huttner WB (2014) Neural progenitors, neurogenesis and the evolution of the neocortex. *Development* 141(11):2182–2194.
2. Dehay C, Kennedy H, Kosik KS (2015) The outer subventricular zone and primate-specific cortical complexification. *Neuron* 85(4):683–694.
3. Kriegstein A, Alvarez-Buylla A (2009) The glial nature of embryonic and adult neural stem cells. *Annu Rev Neurosci* 32:149–184.
4. Taverna E, Götz M, Huttner WB (2014) The cell biology of neurogenesis: Toward an understanding of the development and evolution of the neocortex. *Annu Rev Cell Dev Biol* 30:465–502.
5. Lui JH, Hansen DV, Kriegstein AR (2011) Development and evolution of the human neocortex. *Cell* 146(1):18–36.
6. Betizeau M, et al. (2013) Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron* 80(2):442–457.
7. Kriegstein A, Noctor S, Martínez-Cerdeño V (2006) Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat Rev Neurosci* 7(11):883–890.
8. Rakic P (2009) Evolution of the neocortex: A perspective from developmental biology. *Nat Rev Neurosci* 10(10):724–735.
9. Stahl R, et al. (2013) Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate. *Cell* 153(3):535–549.
10. Lancaster MA, et al. (2013) Cerebral organoids model human brain development and microcephaly. *Nature* 501(7467):373–379.
11. Lancaster MA, Knoblich JA (2014) Generation of cerebral organoids from human pluripotent stem cells. *Nat Protoc* 9(10):2329–2340.
12. Mariani J, et al. (2015) FOXG1-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell* 162(2):375–390.
13. Grün D, et al. (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature* 525(7568):251–255.
14. Patel AP, et al. (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* 344(6190):1396–1401.
15. Treutlein B, et al. (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature* 509(7500):371–375.
16. Zeisel A, et al. (2015) Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* 347(6226):1138–1142.
17. Trapnell C, et al. (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol* 32(4):381–386.
18. Fietz SA, et al. (2012) Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc Natl Acad Sci USA* 109(29):11836–11841.
19. Florio M, et al. (2015) Human-specific gene ARHGAP11B promotes basal progenitor amplification and neocortex expansion. *Science* 347(6229):1465–1470.
20. Prüfer K, et al. (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505(7481):43–49.
21. McLean CY, et al. (2011) Human-specific loss of regulatory DNA and the evolution of human-specific traits. *Nature* 471(7337):216–219.
22. Gittelman RM, et al. (2015) Comprehensive identification and analysis of human accelerated regulatory DNA. *Genome Res* 25(9):1245–1255.
23. Macosko EZ, et al. (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 161(5):1202–1214.
24. Pollen AA, et al. (2014) Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat Biotechnol* 32(10):1053–1058.
25. Ross SE, et al. (2012) Bhlhb5 and Prdm8 form a repressor complex involved in neuronal circuit assembly. *Neuron* 73(2):292–303.
26. Zhou B, et al. (2015) Loss of NFIX transcription factor biases postnatal neural stem/progenitor cells toward oligodendrogenesis. *Stem Cells Dev* 24(18):2114–2126.
27. Darmanis S, et al. (2015) A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci USA* 112(23):7285–7290.
28. Johnson MB, et al. (2015) Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci* 18(5):637–646.
29. Pollen AA, et al. (2015) Molecular identity of human outer radial glia during cortical development. *Cell* 163(1):55–67.
30. Rousso DL, et al. (2012) Foxp-mediated suppression of N-cadherin regulates neuroepithelial character and progenitor maintenance in the CNS. *Neuron* 74(2):314–330.
31. Koyanagi-Aoi M, et al. (2013) Differentiation-defective phenotypes revealed by large-scale analyses of human pluripotent stem cells. *Proc Natl Acad Sci USA* 110(51):20569–20574.
32. Thomson JA, et al. (1998) Embryonic stem cell lines derived from human blastocysts. *Science* 282(5391):1145–1147.
33. Renaud G, Kircher M, Stenzel U, Kelso J (2013) freelbis: An efficient basecaller with calibrated quality scores for Illumina sequencers. *Bioinformatics* 29(9):1208–1209.
34. Renaud G, Stenzel U, Maricic T, Wiebe V, Kelso J (2015) deML: Robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* 31(5):770–772.
35. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
36. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* 25(9):1105–1111.
37. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28(5):511–515.
38. Huang W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4(1):44–57.
39. Zhang HM, et al. (2012) AnimalTFDB: A comprehensive animal transcription factor database. *Nucleic Acids Res* 40(Database issue):D144–D149.