

## RESEARCH ARTICLE SUMMARY

## PSYCHIATRIC GENOMICS

# Transcriptome and epigenome landscape of human cortical development modeled in organoids

Anahita Amiri\*, Gianfilippo Coppola\*, Soraya Scuderi\*, Feinan Wu\*,  
Tanmoy Roychowdhury\*, Fuchen Liu, Sirisha Pochareddy, Yurae Shin, Alexias Safi,  
Lingyun Song, Ying Zhu, André M. M. Sousa, The PsychENCODE Consortium†,  
Mark Gerstein, Gregory E. Crawford, Nenad Sestan, Alexej Abyzov‡, Flora M. Vaccarino‡

**INTRODUCTION:** The human cerebral cortex has undergone an extraordinary increase in size and complexity during mammalian evolution. Cortical cell lineages are specified in the embryo, and genetic and epidemiological evidence implicates early cortical development in the etiology of neuropsychiatric disorders such as autism spectrum disorder (ASD), intellectual disabilities, and schizophrenia. Most of the disease-implicated genomic variants are located outside of genes, and the interpreta-

tion of noncoding mutations is lagging behind owing to limited annotation of functional elements in the noncoding genome.

**RATIONALE:** We set out to discover gene-regulatory elements and chart their dynamic activity during prenatal human cortical development, focusing on enhancers, which carry most of the weight upon regulation of gene expression. We longitudinally modeled human brain development using human induced pluripotent stem

cell (hiPSC)-derived cortical organoids and compared organoids to isogenic fetal brain tissue.

**RESULTS:** Fetal fibroblast-derived hiPSC lines were used to generate cortically patterned organoids and to compare organoids' epigenome and transcriptome to that of isogenic fetal brains and external datasets. Organoids model cortical development between 5 and 16 postconception weeks, thus enabling us to study transitions

## ON OUR WEBSITE

Read the full article  
at <http://dx.doi.org/10.1126/science.aat6720>

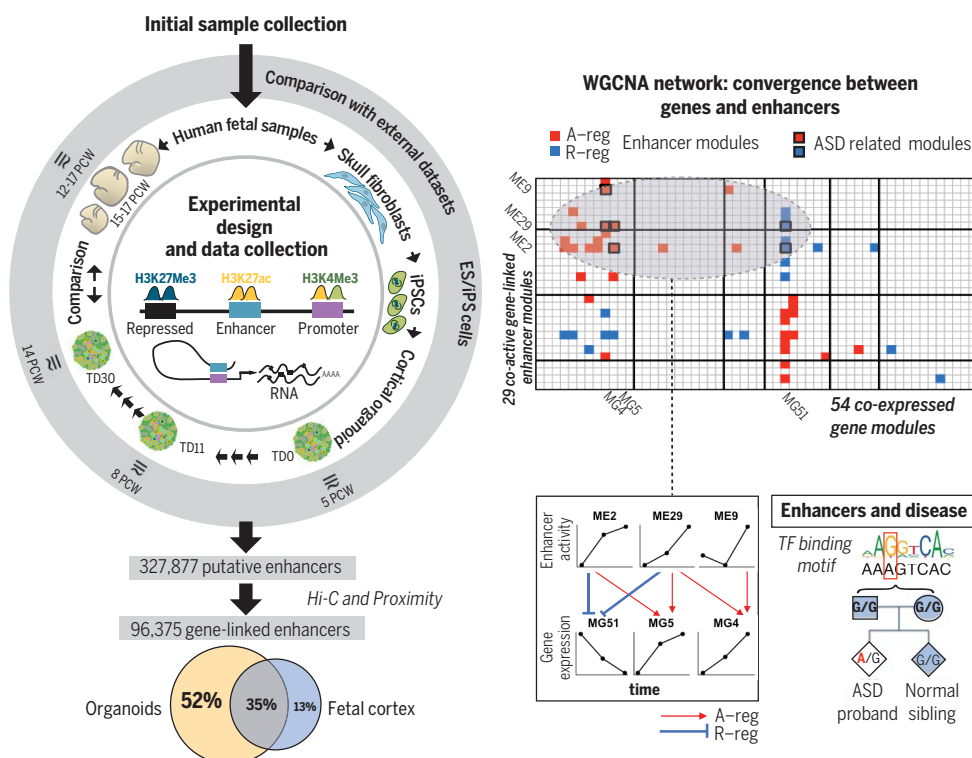
from cortical stem cells to progenitors to early neurons. The greatest changes occur at the transition from stem cells to progenitors. The regulatory landscape encompasses a total

set of 96,375 enhancers linked to target genes, with 49,640 enhancers being active in organoids but not in mid-fetal brain, suggesting major roles in cortical neuron specification. Enhancers that gained activity in the human lineage are active in the earliest stages of organoid development, when they target genes that regulate the growth of radial glial cells.

Parallel weighted gene coexpression network analysis (WGCNA) of transcriptome and enhancer activities defined a number of modules of coexpressed genes and coactive enhancers, following just six and four global temporal

patterns that we refer to as supermodules, likely reflecting fundamental programs in embryonic and fetal brain. Correlations between gene expression and enhancer activity allowed stratifying enhancers into two categories: activating regulators (A-regs) and repressive regulators (R-regs). Several enhancer modules converged with gene modules, suggesting that coexpressed genes are regulated by enhancers with correlated patterns of activity. Furthermore, enhancers active in organoids and fetal brains were enriched for ASD de novo variants that disrupt binding sites of homeodomain, Hes1, NR4A2, Sox3, and NFIX transcription factors.

**CONCLUSION:** We validated hiPSC-derived cortical organoids as a suitable model system for studying gene regulation in human embryonic brain development, evolution, and disease. Our results suggest that organoids may reveal how noncoding mutations contribute to ASD etiology. ■



**Summary of the study, analyses, and main results.** Data were generated for iPSC-derived human telencephalic organoids and isogenic fetal cortex. Organoids modeled embryonic and early fetal cortex and show a larger repertoire of enhancers. Enhancers could be divided into activators and repressors of gene expression. We derived networks of modules and supermodules with correlated gene and enhancer activities, some of which were implicated in autism spectrum disorders (ASD).

The list of author affiliations is available in the full article online.

\*These authors contributed equally to this work.

†The PsychENCODE Consortium authors and affiliations are listed in the supplementary materials.

‡Corresponding author: Email: [abyzov.alexey@mayo.edu](mailto:abyzov.alexey@mayo.edu) (A.A.); [flora.vaccarino@yale.edu](mailto:flora.vaccarino@yale.edu) (F.M.V.)  
Cite this article as Amiri et al., *Science* 362, eaat6720 (2018). DOI: 10.1126/science.aat6720

## RESEARCH ARTICLE

## PSYCHIATRIC GENOMICS

# Transcriptome and epigenome landscape of human cortical development modeled in organoids

Anahita Amiri<sup>1\*</sup>, Gianfilippo Coppola<sup>1\*</sup>, Soraya Scuderi<sup>1\*</sup>, Feinan Wu<sup>1\*</sup>,  
Tanmoy Roychowdhury<sup>2\*</sup>, Fuchen Liu<sup>3</sup>, Sirisha Pochareddy<sup>3</sup>, Yurae Shin<sup>3,4</sup>,  
Alexias Saffi<sup>5</sup>, Lingyun Song<sup>5</sup>, Ying Zhu<sup>3,6</sup>, André M. M. Sousa<sup>3</sup>,  
The PsychENCODE Consortium<sup>†</sup>, Mark Gerstein<sup>7</sup>, Gregory E. Crawford<sup>5</sup>,  
Nenad Sestan<sup>3,8</sup>, Alexej Abyzov<sup>2†</sup>, Flora M. Vaccarino<sup>1,3,8†</sup>

Genes implicated in neuropsychiatric disorders are active in human fetal brain, yet difficult to study in a longitudinal fashion. We demonstrate that organoids from human pluripotent cells model cerebral cortical development on the molecular level before 16 weeks postconception. A multiomics analysis revealed differentially active genes and enhancers, with the greatest changes occurring at the transition from stem cells to progenitors. Networks of converging gene and enhancer modules were assembled into six and four global patterns of expression and activity across time. A pattern with progressive down-regulation was enriched with human-gained enhancers, suggesting their importance in early human brain development. A few convergent gene and enhancer modules were enriched in autism-associated genes and genomic variants in autistic children. The organoid model helps identify functional elements that may drive disease onset.

**P**atterning of the mammalian brain into regions of specific size and fate, demarcated by transcription factor expression and enhancer activity, is already in progress around the time the neural tube closes in the fourth postconceptional week (PCW) in humans and fore-stalls species-specific mechanisms of neurogenesis, connectivity, and function (1–3). A growing list of genetic and epidemiological evidence implicates early neurodevelopment in the etiology of many common neuropsychiatric disorders, such as autism spectrum disorder (ASD), intellectual disabilities, and schizophrenia (4–7). Development, including cell proliferation, interaction, and differentiation, is the result of an inherent gene regulation governed by complex interactions between enhancers, promoters, noncoding RNAs, and transcription regulatory proteins. However, the understanding of epigenetic gene regulation in the

developing human brain is very limited, largely owing to the relative scarcity of available human brain tissue at early developmental time points.

The human cerebral cortex has undergone an extraordinary increase in size and complexity during mammalian evolution, in part through the symmetrical division and the exponential increase in number of radial glial (RG) cells, which are the cortical stem cells (1). The genetic and molecular underpinnings of this process are still unclear, perhaps because these events occur embryonically, before the cortical anlage is formed during the fetal period. Human induced pluripotent stem cells (hiPSCs) and hiPSC-derived organoids allow investigators to gain specific and direct insights into the genetic and molecular events that drive these very early aspects of human cortical development.

## Brain organoids match embryonic to early fetal stages of human cortical development

We produced hiPSC lines from fibroblasts isolated from human postmortem fetuses at mid-gestation, and we differentiated these lines into telencephalic organoids patterned to the dorsal forebrain; samples of cerebral cortex were collected from the same specimens for comparative analyses (fig. S1). To assess the validity of hiPSC-derived telencephalic organoids as a model of human brain development, we compared overall gene expression and regulation of organoids with isogenic cortical brain tissue. Several iPSC lines were derived from skin fibroblasts

of postmortem fetal specimens 310, 313, and 320, aged between 15 and 17 PCWs, for which cortical tissue was available (fig. S2 and table S1). The hiPSC lines derived from fetal fibroblasts were comparable to those derived from adult fibroblasts with regard to pluripotency, growth rate, and differentiation potential (figs. S3 and S4 and table S2) (8). From two hiPSC lines per each of the fetal specimens, we generated telencephalic organoids patterned to the dorsal forebrain (6), grew them under proliferative conditions for 11 days, and then moved them into a terminal differentiation (TD) medium. Organoids were randomly collected for RNA sequencing (RNA-seq) from whole cells as well as nuclear fractions and histone mark chromatin immunoprecipitation sequencing (ChIP-seq) from nuclear fractions at around day 0, day 11, and day 30 of TD in vitro (TD0, TD11, and TD30, respectively). The transcriptomes of whole cells and nuclear RNA were highly correlated (fig. S5) (8); hence, we used the cellular transcriptome for all subsequent analyses. Peaks of three histone marks [trimethylation of histone H3 on lysine 4 (H3K4me3), acetylation of histone H3 on lysine 27 (H3K27ac), and trimethylation of histone H3 on lysine 27 (H3K27me3)] were called to mark functional elements including enhancers, promoters, or polycomb-repressed regions (table S3) (8). To place organoids in a human developmental context, we then compared transcriptomes and chromatin marks from organoids with those from the corresponding isogenic cortical tissue, human embryonic stem cell (hESC) lines, and brain tissue of various ages obtained from the PsychENCODE developmental dataset (9), other PsychENCODE projects (10), and the Roadmap Epigenomics project (11) (Fig. 1A).

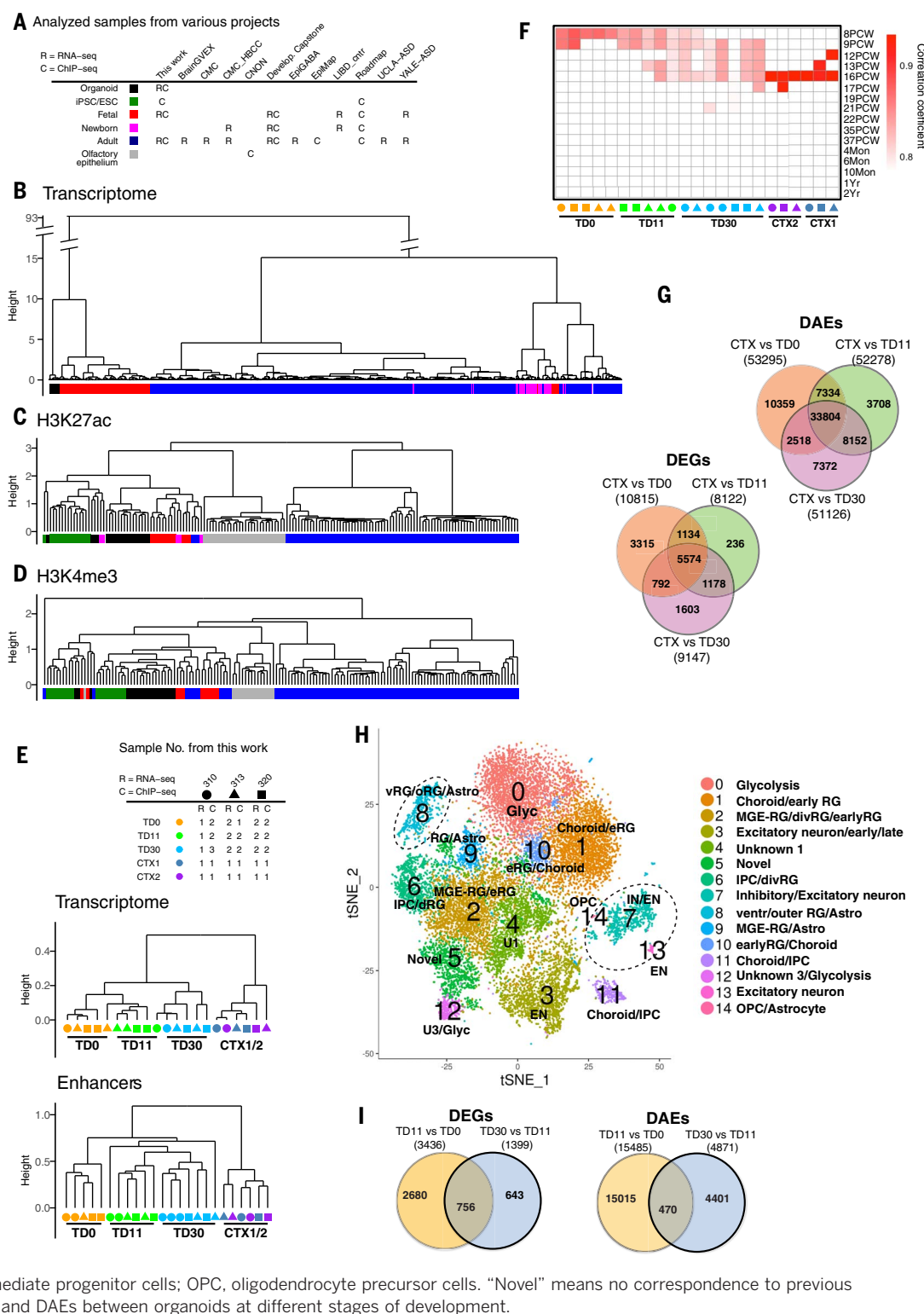
Hierarchical clustering of transcriptomes and histone marks revealed that fetal, perinatal, and adult brain samples formed separate clusters (Fig. 1, B to D), confirming fundamental differences in gene expression in prenatal versus postnatal stages of brain development (12, 13). Furthermore, hiPSC and hESC lines from different sources (including ours) and brain organoids clustered together with fetal brain tissue and separately from adult brain tissue. However, hiPSC and hESC lines formed a distinct subcluster, highlighting differences between organoids and pluripotent cells. Within each cluster, datasets for the same cell type but from different sources were highly concordant with each other (i.e., our data, those of Roadmap Epigenomics, and the PsychENCODE developmental dataset), suggesting that batch effects were not responsible for the observed clustering.

Within our datasets, organoid transcriptomes clustered by in vitro age (i.e., TD0, TD11, and TD30) irrespective of the hiPSC lines from which they were generated, suggesting that the transcriptome reveals well-defined, stage-specific cellular differentiation processes (Fig. 1E and fig. S6). Invariably, organoids clustered separately from the corresponding isogenic fetal cortex. To understand the relationships between organoids and the developing human brain, we classified the organoids against the PsychENCODE developmental dataset (9), which spans a wide range of

<sup>1</sup>Child Study Center, Yale University, New Haven, CT 06520, USA. <sup>2</sup>Department of Health Sciences Research, Center for Individualized Medicine, Mayo Clinic, Rochester, MN 55905, USA. <sup>3</sup>Department of Neuroscience, Yale University, New Haven, CT 06520, USA. <sup>4</sup>National Research Foundation of Korea, Daejeon, South Korea. <sup>5</sup>Department of Pediatrics, Division of Medical Genetics, Duke University, Durham, NC 27708, USA. <sup>6</sup>Department of Biostatistics, Yale School of Public Health, New Haven, CT, USA. <sup>7</sup>Department of Molecular Biophysics and Biochemistry, Department of Computer Science, and Department of Statistics and Data Science, Yale University, New Haven, CT 06520, USA. <sup>8</sup>Kavli Institute for Neuroscience, Yale University, New Haven, CT 06520, USA.

\*These authors contributed equally to this work. <sup>†</sup>The PsychENCODE Consortium authors and affiliations are listed in the supplementary materials. <sup>‡</sup>Corresponding author: Email: abyzov.alexey@mayo.edu (A.A.); flora.vaccarino@yale.edu (F.M.V.)

**Fig. 1. Comparison of transcriptome and epigenome of organoid and isogenic fetal**



human ages and brain regions. Organoids' transcriptomes mapped most closely to the human neocortex between 8 and 16 PCWs of development, with the isogenic fetal brain samples mapping most consistently around 16 PCW, in good agreement with their annotated age (Fig. 1F). This analysis places the organoids substantially earlier than their corresponding mid-fetal brains.

suggesting that organoids model late embryonic to early fetal stages of telencephalic development.

We next compared transcriptomes between each stage of organoid development and the postmortem fetal cortical tissue from the same individual. Overall, there was a large number of differentially expressed genes (DEGs) between each organoid stage and isogenic brain tissue, of

which roughly half was up-regulated and half down-regulated (Fig. 1G and table S4). Although some stage-specific DEGs were present, particularly at TD0 (24%), most of the differences (63%) were shared across two or more organoid stages. Top Gene Ontology (GO) terms for this common set of organoid-brain DEGs were neurogenesis and regulation of nervous system development,



whereas the TD0-specific set of organoid-brain DEGs were related to DNA replication, consistent with age and cell-type differences between fetal brain tissue and organoids (table S4). We tested this hypothesis *in silico*, by assessing for overlap between the organoid-brain DEGs and cell type-specific transcripts identified in fetal human brain (14). Genes up-regulated in the fetal cortex were consistently enriched in markers for maturing excitatory neurons, interneurons, and newborn neurons compared to all organoid stages, whereas genes up-regulated in organoids at TD0 and TD11 were enriched in markers for dividing RG (fig. S6B and table S5).

To validate bulk analyses, we performed single-nucleus RNA sequencing (snRNA-seq) (8) and analyzed the cellular composition of organoids and the fetal brain (one sample per differentiation time point and one sample for brain). We shallow-sequenced about 10,000 cells per sample and considered the top 6000 most informative cells in each sample. We retained only cells expressing at least 500 genes, resulting in a final set of 17,837 cells that were used for analysis. Batch-corrected clustering of single cell's transcriptomes by tSNE analysis from all samples identified 15 clusters (Fig. 1H), with 11 containing cells mostly from organoids and 4 containing cells mostly from the fetal cortex (fig. S6, C and D). Differential expression analysis between any individual cluster and all the others highlighted sets of marker genes for each cluster (table S6), and we used a combination of published datasets of cell markers from single-cell RNA-seq studies of fetal human brain samples (14, 15) to annotate them. The clusters largely contributed by organoid cells overlapped with those identified in human developing brains (15) (Fig. 1H and fig. S6E), and only one cluster, cluster 5, did not find any correspondence to the postmortem human dataset and was labeled "novel." These organoid-specific clusters comprised various types of RG cells including early RG (eRG), outer RG (oRG), ventricular RG (vRG), dividing RG (divRG), and truncated RG (tRG). In addition, cluster 3 expresses early- and late-born excitatory neuron (EN) markers, consistent with an organoid specification to dorsal cortex. Cell clusters specific to the fetal cortex contained inhibitory and excitatory neurons (IN/EN) (clusters 7, 13), RG cells (cluster 8), and a small oligodendrocyte precursor cell (OPC) cluster (cluster 14) (table S6). The presence of IN in the fetal cortex is expected, given that the cortex at PCW 17 is already receiving migrating interneurons from the developing basal ganglia. Timewise, our TD0 organoids (clusters 1, 2, 5, 6, and 10) containing RG and choroid cells matched with cells ranging from 6 to 9 PCWs in fetal brain samples (15). Correspondingly, our CTX1 (clusters 7, 8, 13 and 14) matched with markers (MGE-RG, RG, IN, and EN) seen in 15- to 16.5-PCW fetal brain (fig. S6, K and L). Together, the data confirmed the conclusion of bulk transcriptome analyses that organoids are younger than the fetal brain.

The fraction of cells in a cluster originating from a sample at each time point reveals some clear trends: clusters 1 (Choroid/eRG), 2 (MGE-

RG/dorsal RG/eRG), 6 (IPC/divRG), and 10 (eRG/Choroid) decreased over time, consistent with their being composed of mostly immature cells originating from organoids at TD0 (fig. S6, C and D, and table S6). By contrast, clusters 0 (Glyc) and 12 (U3/Glyc), mostly from samples at TD30, increased with time, perhaps suggesting changing metabolic requirements among neural precursors (15). The remaining clusters, in particular clusters 3 (EN), 4, and 5 (unknown), reached a maximum at TD11, consistent with findings that some newborn neurons peak at an intermediate pseudotime (15). Finally, we ordered the cells along a pseudotime (fig. S6, F to I), which revealed cell trajectories along several dimensions (8). Cells originating from TD0 samples populated the top branch and were nearly absent after the first branch point, which is consistent with the pseudotime progression (fig. S6H) from the top branch (time 0) to the left and right bottom branches (time 15). Similarly, scoring individual cells using cell cycle markers (fig. S6I) revealed a higher frequency of actively cycling cells ( $G_2$ -M or S phase) at the early pseudotimes and larger fractions of non-cycling cells ( $G_1$  phase) when moving along each path (8). In summary, from this integrated analysis emerges a highly coherent picture of organoids' temporal evolution (i.e., differentiation and maturation), representing earlier stages with respect to the corresponding 17-PCW fetal brain counterpart, and mimicking early human brain development, consistent with the classification of the bulk transcriptome with the PsychENCODE developmental Capstone dataset.

We next defined putative promoter and enhancer elements as well as repressed chromatin from histone mark data by chromatin segmentation analyses (figs. S1 and S7 and tables S7 and S8) (8). As a result, we identified 327,877 putative enhancers (H3K27ac peaks, which lack H3K4me3 and H3K27me3 signals) across organoids and fetal brains (table S9). Among these enhancers, H3K27ac signals are highly correlated with ATAC-seq (assay for transposase-accessible chromatin using sequencing) signals, confirming the open chromatin signatures and supporting the robustness of our approach (fig. S7). We further connected these enhancers to genes either by promoter-enhancer distance (within 20 kb) or by the strength of their physical interaction to gene promoters on the basis of Hi-C data for fetal brains (16). From the initial dataset of >300,000 putative enhancers, 96,375 enhancers (29.4%) were found to be associated with 22,835 protein-coding or long intergenic noncoding RNA (lincRNA) genes (out of 27,585 such genes from Gencode V25 annotation) (17) and were used for further analyses (table S10). The gene-associated enhancer dataset was corroborated by the observation of the trend that an increase in activity of enhancers or associated number of enhancers leads to higher expression of interacting genes (figs. S8 to S10).

Of the 96,375 gene-linked enhancers, 90% are concordant with those previously discovered by the ENCODE/Roadmap Consortia in various cell lines and tissues (18), and 10,243 (10%) were completely novel. Overall, 83,608 and 46,735

were active in organoids and the isogenic mid-fetal cortex, respectively. Of the former, 49,640 (59%) were active only in organoids (fig. S11E) and down-regulated in the mid-fetal brain, suggesting that organoids, and by extension, the embryonic and early fetal cortex, use roughly 1.8-fold as many enhancers as later developing cerebral cortex. Comparing enhancer numbers active in organoids across stages, an increasingly larger number became active with the progression of organoid development, with roughly 11,700 enhancers becoming active only at TD30 (fig. S11F). Furthermore, hierarchical clustering analyses based upon the degree of enhancer activity (magnitude of the H3K27ac signal) (Fig. 1E) revealed two major clusters—organoids and the fetal cortex—where organoids' enhancers clustered by *in vitro* age (i.e., TD0, TD11, and TD30) irrespective of genomic background of hiPSC lines, a pattern almost identical to that of transcriptome data (Fig. 1E and fig. S6). Finally, comparing enhancer activity between each stage of organoid development and fetal cortical tissue from the same individual showed that the three organoid stages shared a large number of differentially active enhancers (DAEs) with respect to the fetal cortex (Fig. 1G), as observed with transcriptome data. Together, these analyses reveal a close parallelism between gene expression and enhancer activities across early development and suggest that gene regulation in embryonic and early fetal development is driven by sets of early enhancers, most of which are not active in the mid-fetal cerebral cortex.

### Expression and regulatory changes defining early developmental transitions in organoids

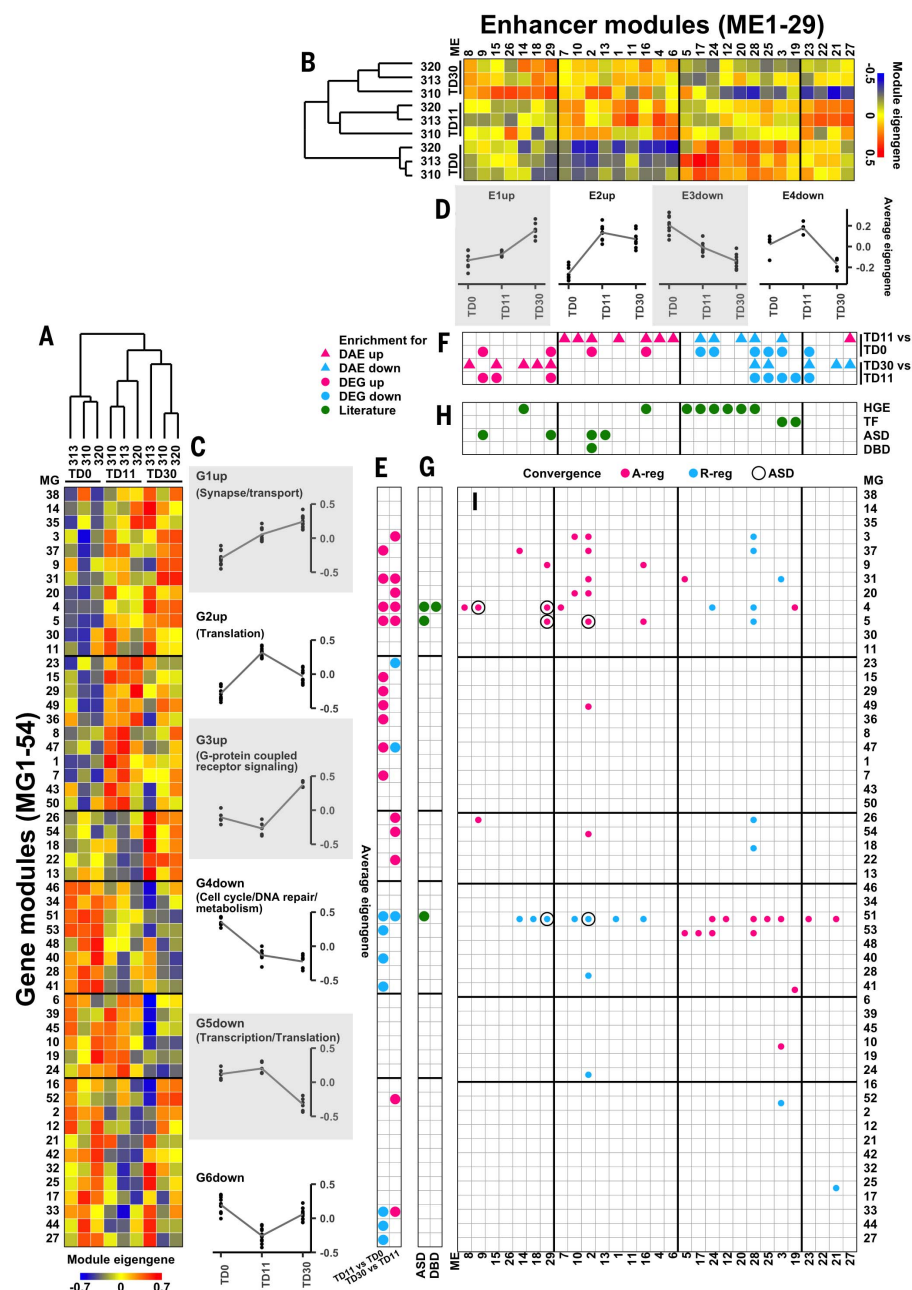
To better understand the gene-regulatory changes driving embryonic and early fetal development, we analyzed DEGs and DAEs in organoids between transitions TD0 to TD11 and TD11 to TD30. We found that the largest differences in gene expression and enhancer activity were at the first transition and that from 2% to 34% of changes were specific for this transition (Fig. 1I and tables S10 and S11), confirming that a substantial change in gene regulation must occur at the beginning of cortical stem cell differentiation. Down-regulated genes specific for the first transition were related to mitosis and regulation of the cell cycle, including cyclin-dependent kinases (CDK2, CDK4, and CDK6) and DNA repair enzymes (TP53, BRCA1/2, and PCNA), all showing a downward trend in expression, likely reflecting top proliferative activity of precursor cells at the earliest time point that decreases during differentiation (fig. S12 and table S11). Consistent with this observation, markers for cell proliferation were progressively down-regulated at the cellular level between TD0 and TD30 (fig. S3). Top functional annotations for genes down-regulated at the second transition (from TD11 to TD30) were instead related to transcriptional regulation of pluripotent and cortical precursor cells (i.e., SOX1/2, EOMES, LHX2, FOXG1, POU3F2/3, SIX3, FEZF2, EMX2, GLI1/3, NEUROD4, HES5/6, REST, and DLL3). By contrast, genes involved in the development of the

neuronal system and synaptic transmission were up-regulated at both transitions and included cell adhesion-, guidance- and synaptic molecule-related genes, including a large number of receptors, calcium and potassium channels, and synaptic membrane recycling components, as well as intellectual disability-related genes such as several CNTN family members.

Performing ChIP-seq and RNA-seq in the same samples provided an opportunity to assess the impact of enhancers on the transcription of their gene targets. We correlated enhancer activity and expression of their associated genes across the whole dataset (organoids and brain samples) to reveal that, globally, 10.6% of gene-enhancer pairs had significant positive or negative correlations, corresponding to 15,026 enhancers and 7858 genes (table S12). Observation of both positive and negative correlations is reminiscent of the finding that H3K27ac-enriched regulatory regions, commonly referred to as enhancers, can be bound by both activators and repressors of gene transcription (19). We referred to 10,192 (67.8%) enhancers with positive correlations as activating regulators (A-regs) of 5605 genes, and to 4993 (33.2%) enhancers with negative correlations as repressing regulators (R-regs) of 3251 genes. Moreover, 98.9% of enhancers are either A-regs or R-regs but not both, consistent with the notion that binding sites of activators and repressors are mutually exclusive (20). Indeed, across both transitions, we observed more pronounced correlations between expression changes of genes and activity change of linked A-regs versus linked non-A-regs; similar observations were made for R-regs (fig. S13A). Consistently, differentially active A-regs and R-regs are associated with DEGs in the expected direction, i.e., A-regs with increased activity are enriched in up-regulated DEGs, whereas R-reg with increased activity are enriched in down-regulated DEGs (Fisher's exact test,  $p < 2.2 \times 10^{-16}$  for both transitions) (fig. S13B), suggesting that differential activity of the identified enhancers is indeed driving differential gene expression across organoid development.

### Gene and enhancer network analyses

To study the temporal dynamics of gene expression and enhancer activities across the three developmental time points, we used weighted gene coexpression network analysis (WGCNA) (21). The resulting networks grouped gene transcripts in 54 coexpressed modules (MG1 to MG54) and gene-associated enhancers into 29 coactive modules (ME1 to ME29), each showing a specific trajectory along organoid differentiation (Fig. 2, A and B, and tables S12 and S14). Unsupervised hierarchical clustering of module eigengenes, which are representative of the gene expression and enhancer activity of each module, grouped samples by differentiation time point. Using k-means clustering of each module's eigengenes, we grouped the gene and enhancer modules into six and four “supermodules,” respectively, which represent higher-order clustering of the modules (Fig. 2, C and D).



**Fig. 2. Modules of coexpressed genes and coactive enhancers during organoid differentiation.** (A) Unsupervised hierarchical clustering of gene modules (1 through 54) by expression eigengenes. Rows and columns represent gene modules and samples, respectively. (B) Unsupervised hierarchical clustering of enhancer modules (1 through 29) by activity eigengenes. Rows and columns represent samples and enhancer modules, respectively. (C and D) Mean module eigengenes (lines) across differentiation times grouped by gene (C) and enhancer (D) supermodules, respectively. Dots represent values of eigengenes for individual modules. (E to H) Enrichment of gene (E and G) and enhancer (F and H) modules for DEGs and DAEs and for various enhancers and genes of interest from the literature, including HGE (human-gained enhancers) (26), TF (genes encoding transcription factors during human fetal brain development) (24), ASD (genes pertinent to autism spectrum disorder) (22), and DBD (genes pertinent to developing brain disorder) (23). (I) Correspondence between the gene and enhancer networks. The strongest A-reg (pink dots) and R-reg (cyan dots) for a subset of gene modules are overrepresented in a number of enhancer modules. Black circles emphasize converging genes and enhancer modules, both of which are ASD-associated [as shown in (G) and (H)]. Panels (E) to (I) are aligned by the gene and enhancer modules shown in (A) and (B).

Supermodules exhibit specific profiles of activities during the two transitions (8) and functional annotations (table S14). The monotonically up-regulated gene supermodule Glup comprised modules related to neurons, synapses, cell adhesion, and axon guidance and was hence dubbed as governing synapse/transport. Conversely, the down-regulated supermodule G4down comprised modules enriched in DNA repair and cell cycle-related genes and was thus dubbed as governing cell cycle/DNA repair (Fig. 2C), reflecting the cell cycle annotation of TD0-to-TD11-down-regulated DEGs (fig. S12). Other supermodules exhibited transition-specific changes. G2up, which exhibited peak up-regulated gene expression at TD11, was enriched in genes related to ribosome, translation, protein folding, and degradation. The transcription supermodule G5down, down-regulated at the second transition, included major transcription factors (TFs) expressed by cortical progenitor cells, which show down-regulation at TD11 to TD30 (fig. S12). By contrast, the G3up supermodule, up-regulated at the second transition, was enriched in G protein receptor signaling, implying a previously unknown role of these molecules in the earliest stages of cortical neuron differentiation. Patterns of gene expression and enhancer activity in the modules and supermodules were further confirmed by enrichment analysis of DEG and DAEs (Fig. 2, E and F). Specifically, gene modules and linked genes of enhancer modules were enriched with DEGs for which gene expression changes were generally in the same direction as their respective module eigengenes.

Further evidence for functional relevance of the modules and supermodules arises from intersection with genes relevant to neuropsychiatric diseases. Genes within the SFARI dataset, a curated list of genes associated with ASD, including both rare mutations and common variants (22), were significantly overrepresented in the MG4 and MG5 neuronal and synaptic modules and the MG51 cell cycle module (Fig. 2G and table S14). SFARI genes were also enriched within gene targets of four enhancer modules (ME9 and ME29 in supermodule E1up, and ME2 and ME13 in supermodule E2up) with up-regulated patterns of activity across development, one of which, the ME2 module, was also enriched in developmental brain disorder genes (23) (Fig. 2H). Enrichment analysis also showed that a set of TFs pertinent to human cortical neurogenesis (24) was preferentially associated with gene targets of two enhancer modules (ME3 and ME19, both in supermodule E3down) that have down-regulated enhancer activity across organoid development (Fig. 2H). This evidence supports the notion that organoid culture can capture dynamic gene-regulatory events present in early human brain development and that such early events are potentially involved in disease pathogenesis.

To assess the correspondence between the gene network and the enhancer network, we examined whether enhancers linked to a gene module are overrepresented in one or a small number of enhancer modules. Such convergence between a gene module and an enhancer module

would suggest that coexpressed genes are likely regulated by enhancers with correlated patterns of activity. To mitigate the ambiguity caused by multiple enhancers per gene, we focused on the strongest A-regs or R-regs of a gene, defined by the most positive or negative correlation between enhancer activity and gene expression. Indeed, we find that A-regs and R-regs of 14 and 12 gene modules, respectively, are overrepresented in a small number of enhancer modules [false discovery rate (FDR) < 0.05, Fig. 2I]. Not surprisingly, A-regs and R-regs linked to the same gene module are overrepresented in different enhancer modules with opposite trajectories over time, e.g., A-regs of MG3 in Glup converges with ME10 and ME2 in E2up, but its R-regs converges with ME28 in E3down. Such convergence between the gene network and the enhancer network suggests that coexpressed genes likely share a set of co-regulated enhancers. Moreover, enhancers discovered in organoids hint at upstream elements that regulate the expression of disease-associated genes. For example, ASD-associated MG4, MG5, and MG51 gene modules converge with ME9, ME29, and ME2, enhancer modules that are associated with ASD genes as well (Fig. 2, G to I, black circles). ME29 is particularly interesting as it contains both A-regs and R-regs for all three ASD-associated gene modules, suggesting that it may be responsible for the coordinated up- and down-regulation of genes modules involved in autism pathogenesis.

The ASD-associated gene modules—MG4, MG5, and MG51—overlapped to a significant extent with previously published ASD modules identified by in vivo analyses of differential gene expression between ASD patients and normal individuals (Fig. 3A and table S14). Our MG4 and MG5 modules were annotated by neuronal and synaptic terms (Fig. 3B) and overlapped with neuronal and synaptic modules down-regulated in the ASD postmortem cerebral cortex (25) as well as with a synapse module up-regulated in brain organoids from ASD individuals with macrocephaly (6). By contrast, our down-regulated MG51 module was annotated by cell cycle and DNA repair terms (Fig. 3B) and overlapped with M3, a module harboring protein-disrupting, rare de novo variants in ASD (4). No overlap was observed with modules related to immune dysfunction and microglia in ASD (25) (Fig. 3A). Within each ASD-associated gene module, the distribution of genes that are implicated in ASD and are targets of a member of the ME9, ME29, ME2, and ME13 ASD-associated enhancer modules appears, overall, to be skewed toward the central part of each module (i.e., the “strongest” hubs) (Fig. 3, C and D, and fig. S14). Given that hub genes are the drivers of a module, one may speculate that mutations disrupting these genes are more likely to be penetrant and/or syndromic. Looking at the first 100 hub genes (table S14), we find that the MG4 module shows two confident and two syndromic ASD-associated genes (respectively DSCAM, MYO5A, CAMK2B, and SMARCA2); the MG5 module shows three confident and three syndromic ASD-associated genes (respectively

ANK3, STXBP1, ACHE, WDR26, and ATP1A3); and the MG51 module only shows DIAPH3, a lower-confidence gene (Fig. 3C and fig. S14). Orthogonal analyses by quantitative polymerase chain reaction (qPCR) confirmed the expression level of these and other ASD genes in the organoid dataset (fig. S15). Overall, the results suggest that our organoid model may be used to unravel the roles of early prenatal neurodevelopment and genetic factors in ASD.

## Relevance of the organoid model to understanding human brain evolution

To determine whether the organoid model is useful to understanding the genetic mechanisms driving human brain evolution, we assessed the overlap of our enhancers with a list of 8996 human-gained enhancers (HGEs). These HGEs showed increased activity at very early stages of brain development (7 to 12 PCWs) in the human lineage, compared with their homologs in rhesus macaque and mouse brains at similar developmental time points (26). The majority (70%, 6295 out of 8996) of published HGEs overlapped with 9915 enhancers in our dataset, and among the latter, 3310 are associated with genes (table S15). Out of 3310 gene-associated HGEs, 2670 (85.3%) have differential activity between organoids and fetal brains, suggesting a dynamic role during brain development (fig. S16). The largest fraction of gene-associated HGEs are progressively declining in activity along organoid differentiation and from organoids to fetal brain. Among eight enhancer modules enriched with HGEs, six (all in the supermodule E3down) had decreasing activity along organoid differentiation (Fig. 2H). Genes targeted by HGEs in these six down-regulated modules were enriched in signaling pathways related to cell proliferation and cell differentiation and communication and included extracellular growth factors such as FGF7 and FGF6, FGFR1, ERBB4, IGF2, EGFL7, VEGFA, and PDGFA (table S15). Overall, among all 2908 HGE-linked genes, 824 are differentially expressed between human and macaque brain in at least one of three brain ages—438 in fetal brains, 346 in postnatal brains, and 724 in adult brains (27). Together, these findings suggest that HGEs are likely to be important regulators of genes controlling cell proliferation and cell-to-cell interactions in the human cerebral cortical primordium during the very early stages of cortical morphogenesis. These data are consistent with ATAC-seq from in vivo human brain (24), which demonstrates that HGEs are active in germinal zones and especially enriched in outer radial glia (oRG), which are expanded in humans (28).

## Gene regulation and relevance to disorders

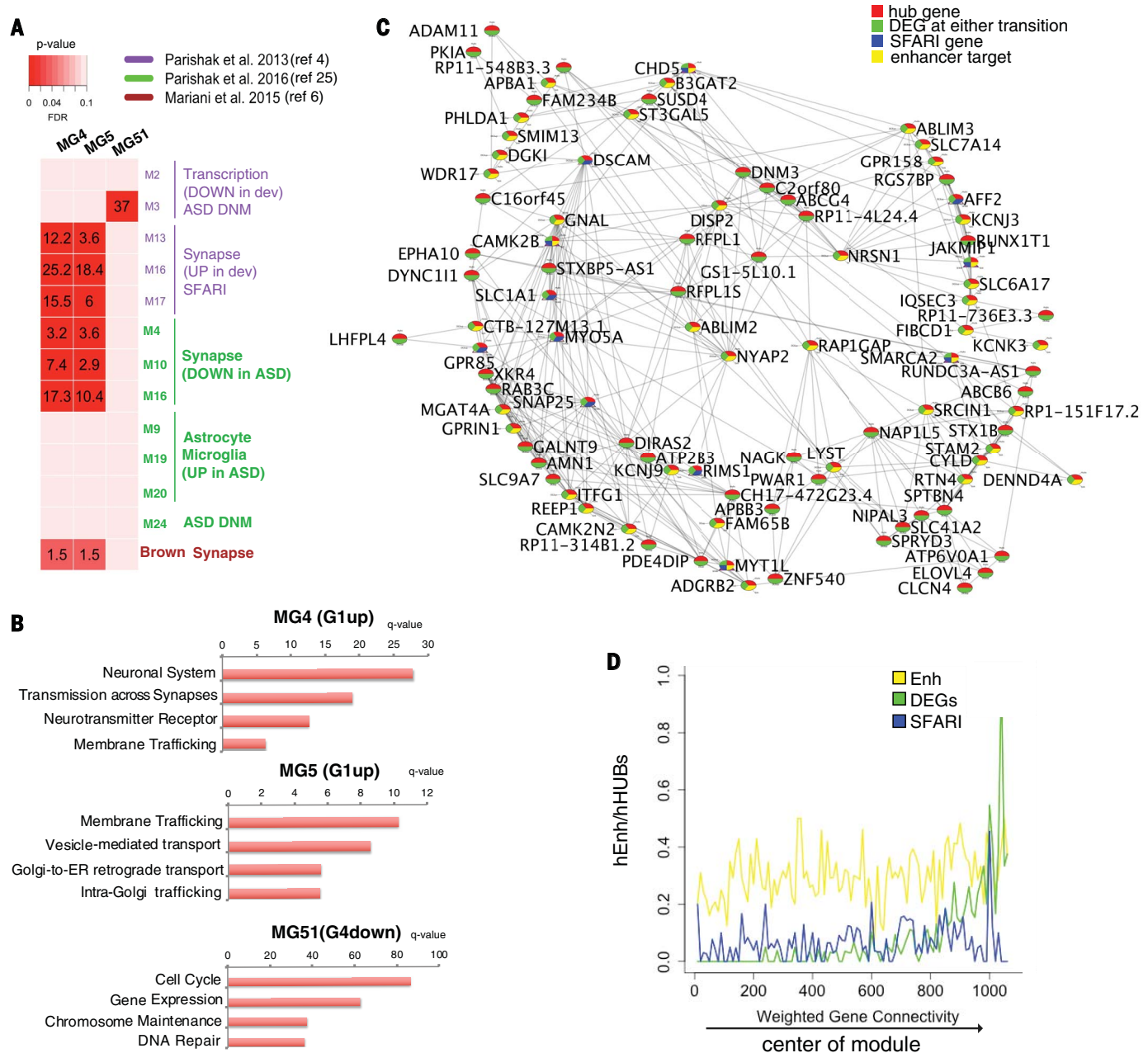
More than 24% of the ASD genes in the SFARI dataset are differentially expressed in the organoid system across time, and over 80% are linked to enhancers active in organoids or fetal brain (table S16). To understand whether enhancers active in organoids or fetal brain can inform about common and rare genetic variants that



underlie ASD, we selected three subsets from the 96,375 gene-associated enhancers: 11,448 early enhancers, only active in all organoid stages; 8999 late enhancers, only active in fetal brain; and 7865 constant enhancers, active at all stages of organoid differentiation and in fetal brain (Fig. 4A). These enhancers were analyzed for enrichment with personal variants inherited from

either parent in 540 families of the Simons Simplex Collection (SSC). Each family consisted of phenotypically normal parents, an ASD male proband, and a normal male sibling (Fig. 4A). Out of an average 3.6 million inherited single-nucleotide polymorphisms (SNPs) per person,

3327 with <5% minor allele frequency (MAF) were located within early, late, or constant enhancers (fig. S17, A to C). Among these, low-allele frequency SNPs (MAF 0.1% to 5%) were significantly enriched in probands relative to siblings in early but not in late or constant enhancers ( $p = 0.02$  by one-sample  $t$  test, Fig. 4B).



**Fig. 3. ASD-associated genes modules.** (A) Overlap of ASD gene modules MG4, MG5, and MG51 from this study with transcript modules associated with ASD from postmortem brain studies or enriched in ASD de novo mutations (DNM) (green, violet) (4, 25) and from an ASD patient-derived organoid study (brown) (6). Rows are modules from this study and columns are modules from other studies. Red shading represents the degree of enrichment between pairs of modules. Corrected  $p$  values of significant overlaps (hypergeometric test) are numerically indicated as  $-\log_{10}(p \text{ value})$ . (B) Bar plots of the top-scoring biological process terms for the ASD-

associated modules shown in (A). (C) Graphical representation of the strongest interacting hub genes in the MG4 module network. Circles: genes; lines: topological overlap above 0.95. Colors in circles annotate each gene as hub (red), DEG (green), SFARI gene (blue), and enhancer target (yellow). Enhancer target: genes targeted by enhancers in the ME9, ME29, ME13, and ME2 ASD-associated enhancer modules (Fig. 2I). (D) Frequency plots within the MG4 module showing that enhancer targets, DEGs, and SFARI genes have higher intramodular connectivity.  $x$  axis shows the weighted gene connectivity, from low (peripheral genes) to high (central hub genes).

These SNPs were also enriched in the ME2 and ME29 enhancer modules ( $p = 0.05$  and  $0.03$ , respectively, by one-sample  $t$  test) (Fig. 4B), which converge with ASD-associated gene modules (Fig. 2I). These variants are relatively common, and thus our results support the hypothesis of etiology of ASD via superposition of multiple inherited variants of low effect size (29–32).

Contrary to numerous inherited SNPs, there are only a few dozen de novo mutations (DNMs) in probands, which must have deleterious effects in order to contribute to ASD phenotypes. We compared DNMs of probands and siblings of the same family cohort (33). Out of 66,306 total DNMs, 2422 were located in our dataset of gene-associated enhancers. There was a trend of having a larger fraction of probands' DNMs in constant enhancers, which are active during a prolonged period of development (Fig. 4C and fig. S17D). We next elucidated the effect of individual DNMs in the gene-associated enhancers on TF binding. Around 24% of DNMs (out of 1240 and 1184 from proband and sibling, respectively) overlapped with at least one TF motif (figs. S17, E and F, and S18). Overall, there was a larger number of TFs with greater count of motif-breaking DNMs in probands than in siblings (more circles below the diagonal than above in Fig. 4D). A significant difference ( $p < 0.05$  by binomial test) was observed for TFs such as homeodomain, Hes1, NR4A2, Sox3, and NFIX (table S17), which are implicated in development, ASD, or mental disorders (34, 35). De novo copy-number variants at the NR4A2 gene locus at 2.q24.1, in particular, have been associated with ASD with language and cognitive impairment across multiple datasets (35). These observations provide genetic support for the relevance of enhancer elements identified in organoids in the complex etiology of ASD and link noncoding variants to ASD etiology, as previously proposed (36). Enhancers discovered in this study also inform about the possible regulatory role of SNPs that underlie the etiology of schizophrenia (37) (fig. S17G).

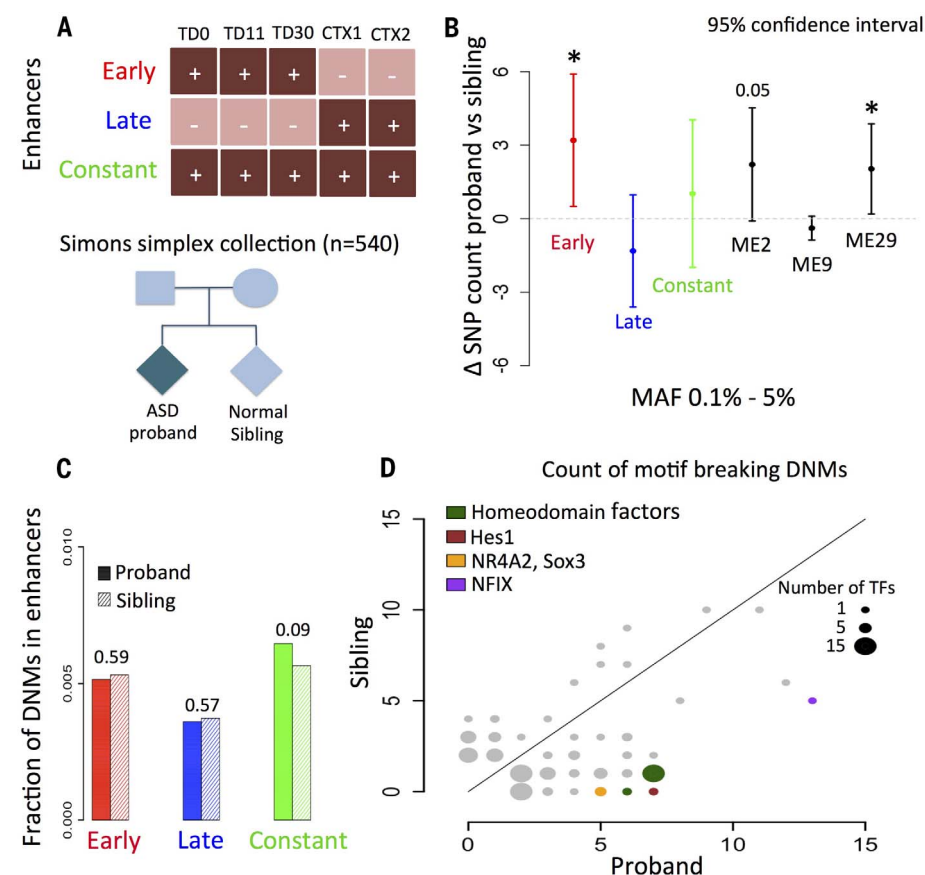
## Discussion

Using forebrain organoids, we provide an initial map of enhancer elements and corresponding transcripts that are dynamically active in the transitions between human cortical stem cells, progenitors, and early cortical neurons. Although the cataloged functional elements may require further validation of their *in vivo* activity, our findings suggest that human brain organoids provide an avenue to approach the study of the molecular and cellular events underlying brain development. Indeed, our brain organoids patterned to forebrain, on both transcriptome and regulatory levels, mimic the longitudinal development of the embryonic and early fetal cortical primordium. Because all organoid preparations (from other studies and with different protocols) patterned to the dorsal forebrain are derived from neural stem cells, it is likely that they share similar gene dynamics specific to the embryonic brain described here. Thus, our gene and enhancer analyses have wide implications, and the de-

scribed map can aid the identification of sets of genes, enhancers, and genomic variants underlying neurodevelopmental disorders and ASD in particular, because brain development is nearly complete at the time of diagnosis (38).

The majority of enhancer elements active in our organoid system are not shared with isogenic mid-fetal brain tissue, which suggests that they play a role in earlier events, i.e., progenitor proliferation and the specification of neuronal lineages. However, it remains unclear whether organoids fully recapitulate developmental processes, particularly those at later stages. Organoid preparations grown for longer periods *in vitro* may show greater overlap with mid-fetal human

brains (39, 40), although a notable aspect of the organoid system is its ability to span very early developmental transitions, which map to stages earlier than those commonly available in post-mortem human tissue. This finding is confirmed by single-cell transcriptome analyses, which revealed a wide diversity of RG and progenitor clusters throughout organoid development. All but one organoid-specific cell clusters find correspondence to cell clusters in embryonic-fetal human brain. The one that did not could be the result of *in vitro* culturing. Through longitudinal analyses, we show that many genes and their enhancer elements are differentially active in a stage-specific fashion from RG stem cells to



**Fig. 4. Enrichment of variants in gene-associated enhancers.** (A) Three subsets of enhancers were selected from all gene-associated enhancers. Early: enhancers active (denoted by +) in all organoid stages but inactive (denoted by -) in fetal brain (red); late: enhancers active in fetal brain but inactive in all organoid stages (blue); constant: enhancers active in all organoid stages and fetal brain (green). Variants in 540 families from the Simons Simplex Collection were analyzed for enrichment in these enhancer sets. (B) Comparison of inherited personal SNPs between ASD probands and normal siblings from the SSC revealed significant enrichment in probands versus siblings ( $p \leq 0.05$  by one-sample  $t$  test) of low-allele frequency SNPs (MAF 0.1 to 5%) in early enhancers (red) and enhancer modules ME2 and ME29 (black). Dashed line at value of 0 represents no difference between probands and siblings. \* $p < 0.05$ . (C) Fractions of DNMs in enhancers were compared in probands and siblings across the whole genome.  $P$  values (shown above the bars) were calculated by using the chi-square test. (D) Count of motif-breaking DNMs in all gene-associated enhancers were compared between probands and siblings. Circles represent TFs with counts of broken motifs in probands and siblings plotted on x- and y axis. The size of the circles is proportional to the number of TFs. Circles away from the diagonal represent TFs enriched with motif-breaking DNMs in probands or siblings. A few TFs in the probands (colored circles) but not in the siblings were significantly enriched ( $p < 0.05$  by binomial test) with motif-breaking DNMs.



neuronal progenitors and to young neurons. The first transition, from neural stem cells to early cortical progenitors, has the largest number of DEGs (71%) and DAEs (76%), the majority of which are specific to that step, which implies that in vivo transition from the embryonic to the fetal brain is a vulnerable step for normal brain development. Such changes reflect dynamic transitions in proliferation-related genes and transcription factors, together with the up-regulation of neuronal lineage and synaptic genes as cortical stem cells (i.e., RG cells) progressively stop dividing and acquire different neuronal identities. We found that HGEs exhibit their highest activity in RG cells, after which their activity progressively declines with differentiation. Consistent with previous findings (24), this observation implicates HGEs as regulators of the earliest phases of human brain development. Although the exact function of HGEs remains to be determined, based on enrichment for growth factors signaling pathways, their time course and the comparison with other studies, we hypothesize that they are involved in the regulation of RG cell proliferation in the cerebral cortex.

Global integrative analyses of transcriptome and enhancer elements allowed us to classify the gene-associated enhancers into elements that activate or repress gene transcription, in which activity changes in A-regs and R-regs are correlated with changes in the expression of their gene targets at each developmental transition. Because a third of those regulators likely acted as gene-repressing elements, our results point out an underappreciated layer of trans-repression during early brain development. This level of integration allows the construction of a complex regulatory network with convergent and concordant patterns of activity between gene and enhancer modules, where enhancers of coexpressed genes also exhibit correlated activity. We propose that this network portrays fundamental developmental programs in embryonic and fetal brain.

Three gene modules were enriched in genes implicated in ASD, two of which, MG4 and MG5, regulate neuron and synapses and progressively increased in expression during development; whereas the other, MG51, regulates the cell cycle, and whose expression progressively declines. Those modules overlap gene modules previously implicated in ASD based on in vivo postmortem data (25). Additionally, we found that ASD-associated gene modules converged with three ASD-associated enhancer modules, implying that other genes and enhancers in those modules may also be related to ASD by shared expression and perhaps function. This supports the validity of the organoid model for the discovery and analysis of regulatory elements whose variation may underlie the risk for neuropsychiatric disorders. Indeed, enhancers active in organoids, and, by extension, embryonic and early fetal cerebral cortices, were enriched for low population-frequency personal variants carried by ASD probands relative to unaffected siblings. Furthermore, DNMs in ASD probands more frequently disrupted binding motifs of specific transcription

factors within regulatory elements active at those stages. Those TFs, their disrupted binding motifs, and the gene targets of the enhancers with the motifs can be the subject of future functional studies on the etiology of ASD. Altogether, the evidence corroborates previous suggestions that single-nucleotide variants in noncoding regions contribute to ASD (36) and points to genes and regulatory elements underlying its onset. Thus, organoids can offer mechanistic insights into early human telencephalic development, brain evolution, and disease.

## Methods summary

Detailed materials and methods can be found in the supplementary materials. hiPSC lines were derived from skull fibroblasts of three male fetal specimens aged between 15 and 17 PCWs, from which two cerebral cortical samples each were also collected for comparative analyses. iPSCs were differentiated into telencephalic organoids patterned to the dorsal forebrain as previously described (6). Organoids were collected at three TDs for downstream analyses. Immunohistochemistry using proliferation, glutamatergic, and GABAergic neuronal markers were used for organoids' differentiation quality control. Samples from iPSCs, iPSC-derived organoids, and fetal cerebral cortical regions were used for total stranded RNA-seq (cells and nuclei), snRNA-seq (nuclei), and ChIP-seq for three histone marks (H3K4me3, H3K27ac, and H3K27me3) (nuclei). We used edgeR (41) and trended dispersion estimates to infer differentially expressed genes and differentially active enhancers. We used the Seurat pipeline (42) for single-cell RNA-seq clustering and the Monocle pipeline (43) for single-cell trajectories. ConsensusPathDB (44) and ToppGene (45) were used for functional annotation. Quantitative real-time PCR was used to cross-validate RNA-seq and DEG analyses using a random subset of the DEGs as well as DEGs implicated in ASD. ChIP-seq peaks were called by MACS2 (46), and chromatin segmentation was done by chromHMM (47). Peaks were merged into consensus peaks and annotated by the corresponding chromatin states at each TD or in the fetal cortex. We used physical proximity and published chromatin conformation (Hi-C) data (16) from the fetal brain to link enhancers to genes. Gene and enhancer modules were identified by WGCNA (21), and super-modules were defined by K-means clustering of module eigengenes. To assess the relevance of the organoid model to studying noncoding pathological mutations, personal genomic variants across the whole genome were obtained from the SFARI (Simons Simplex Collection) dataset in 540 families with ASD probands and normal siblings. We also used de novo SNPs identified in Werling *et al.* from the same cohort (33). Transcription factor binding site motifs were obtained from the JASPAR database (48).

## REFERENCES AND NOTES

1. P. Rakic, A small step for the cell, a giant leap for mankind: A hypothesis of neocortical expansion during evolution. *Trends Neurosci.* **18**, 383–388 (1995). doi: [10.1016/0166-2236\(95\)93934-P](https://doi.org/10.1016/0166-2236(95)93934-P); pmid: 7482803

2. J. L. R. Rubenstein, S. Martinez, K. Shimamura, L. Puelles, The embryonic vertebrate forebrain: The prosomeric model. *Science* **266**, 578–580 (1994). doi: [10.1126/science.7939711](https://doi.org/10.1126/science.7939711); pmid: 7939711
3. A. Visel *et al.*, A high-resolution enhancer atlas of the developing telencephalon. *Cell* **152**, 895–908 (2013). doi: [10.1016/j.cell.2012.12.041](https://doi.org/10.1016/j.cell.2012.12.041); pmid: 23375746
4. N. N. Parikshak *et al.*, Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**, 1008–1021 (2013). doi: [10.1016/j.cell.2013.10.031](https://doi.org/10.1016/j.cell.2013.10.031); pmid: 24267887
5. A. J. Willsey *et al.*, Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013). doi: [10.1016/j.cell.2013.10.020](https://doi.org/10.1016/j.cell.2013.10.020); pmid: 24267886
6. J. Mariani *et al.*, FOXG1-Dependent Dysregulation of GABA/Glutamate Neuron Differentiation in Autism Spectrum Disorders. *Cell* **162**, 375–390 (2015). doi: [10.1016/j.cell.2015.06.034](https://doi.org/10.1016/j.cell.2015.06.034); pmid: 26186191
7. S. Gulsuner *et al.*, Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* **154**, 518–529 (2013). doi: [10.1016/j.cell.2013.06.049](https://doi.org/10.1016/j.cell.2013.06.049); pmid: 23911319
8. See supplementary materials.
9. M. Li *et al.*, Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **362**, eaat7615 (2018).
10. S. Akbarian *et al.*, The PsychENCODE project. *Nat. Neurosci.* **18**, 1707–1712 (2015). doi: [10.1038/nn.4156](https://doi.org/10.1038/nn.4156); pmid: 26605881
11. A. Kundaje *et al.*, Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015). doi: [10.1038/nature14248](https://doi.org/10.1038/nature14248); pmid: 25693563
12. C. Colantoni *et al.*, Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* **478**, 519–523 (2011). doi: [10.1038/nature10524](https://doi.org/10.1038/nature10524); pmid: 22031444
13. A. T. Tebbenkamp, A. J. Willsey, M. W. State, N. Sestan, The developmental transcriptome of the human brain: Implications for neurodevelopmental disorders. *Curr. Opin. Neurol.* **27**, 149–156 (2014). doi: [10.1097/WCO.0000000000000069](https://doi.org/10.1097/WCO.0000000000000069); pmid: 24565942
14. S. J. Liu *et al.*, Single-cell analysis of long non-coding RNAs in the developing human neocortex. *Genome Biol.* **17**, 67 (2016). doi: [10.1186/s13059-016-0932-1](https://doi.org/10.1186/s13059-016-0932-1); pmid: 27081004
15. T. J. Nowakowski *et al.*, Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017). doi: [10.1126/science.aap8809](https://doi.org/10.1126/science.aap8809); pmid: 29217575
16. H. Won *et al.*, Chromosome conformation elucidates regulatory relationships in developing human brain. *Nature* **538**, 523–527 (2016). doi: [10.1038/nature19847](https://doi.org/10.1038/nature19847); pmid: 27760116
17. J. Harrow *et al.*, GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012). doi: [10.1101/gr.135350.111](https://doi.org/10.1101/gr.135350.111); pmid: 22955987
18. The ENCODE Project Consortium, An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012). doi: [10.1038/nature11247](https://doi.org/10.1038/nature11247); pmid: 22955616
19. Y. Nie, H. Liu, X. Sun, The patterns of histone modifications in the vicinity of transcription factor binding sites in human lymphoblastoid cell lines. *PLOS ONE* **8**, e60002 (2013). doi: [10.1371/journal.pone.0060002](https://doi.org/10.1371/journal.pone.0060002); pmid: 23527292
20. D. L. Daniels, W. I. Weis, Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat. Struct. Mol. Biol.* **12**, 364–371 (2005). doi: [10.1038/nsmb912](https://doi.org/10.1038/nsmb912); pmid: 15768032
21. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008). doi: [10.1186/1471-2105-9-559](https://doi.org/10.1186/1471-2105-9-559); pmid: 19114008
22. E. Larsen *et al.*, A systematic variant annotation approach for ranking genes associated with autism spectrum disorders. *Mol. Autism* **7**, 44 (2016). doi: [10.1186/s13229-016-0103-y](https://doi.org/10.1186/s13229-016-0103-y); pmid: 27790361
23. J. G. Camp *et al.*, Human cerebral organoids recapitulate gene expression programs of fetal neocortex development. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 15672–15677 (2015). doi: [10.1073/pnas.1514456112](https://doi.org/10.1073/pnas.1514456112); pmid: 26644564
24. L. de la Torre-Ubieta *et al.*, The Dynamic Landscape of Open Chromatin during Human Cortical Neurogenesis. *Cell* **172**, 289–304.e18 (2018). doi: [10.1016/j.cell.2017.12.014](https://doi.org/10.1016/j.cell.2017.12.014); pmid: 29307494
25. N. N. Parikshak *et al.*, Genome-wide changes in lncRNA, splicing, and regional gene expression patterns in autism. *Nature* **540**, 423–427 (2016). doi: [10.1038/nature20612](https://doi.org/10.1038/nature20612); pmid: 27919067

26. S. K. Reilly *et al.*, Evolutionary changes in promoter and enhancer activity during human corticogenesis. *Science* **347**, 1155–1159 (2015). doi: [10.1126/science.1260943](https://doi.org/10.1126/science.1260943); pmid: 25745175
27. Y. Zhu *et al.*, Spatiotemporal transcriptomic divergence across human and macaque brain development. *Science* **362**, eaat8077 (2018).
28. T. J. Nowakowski, A. A. Pollen, C. Sandoval-Espinosa, A. R. Kriegstein, Transformation of the Radial Glia Scaffold Demarcates Two Stages of Human Cerebral Cortex Development. *Neuron* **91**, 1219–1227 (2016). doi: [10.1016/j.neuron.2016.09.005](https://doi.org/10.1016/j.neuron.2016.09.005); pmid: 27657449
29. R. Anney *et al.*, Individual common variants exert weak effects on the risk for autism spectrum disorders. *Hum. Mol. Genet.* **21**, 4781–4792 (2012). doi: [10.1093/hmg/dds301](https://doi.org/10.1093/hmg/dds301); pmid: 22843504
30. T. Gaugler *et al.*, Most genetic risk for autism resides with common variation. *Nat. Genet.* **46**, 881–885 (2014). doi: [10.1038/ng.3039](https://doi.org/10.1038/ng.3039); pmid: 25038753
31. L. Klei *et al.*, Common genetic variants, acting additively, are a major source of risk for autism. *Mol. Autism* **3**, 9 (2012). doi: [10.1186/2040-2392-3-9](https://doi.org/10.1186/2040-2392-3-9); pmid: 23067556
32. D. J. Weiner *et al.*, Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. *Nat. Genet.* **49**, 978–985 (2017). doi: [10.1038/ng.3863](https://doi.org/10.1038/ng.3863); pmid: 28504703
33. D. M. Werling *et al.*, An analytical framework for whole-genome sequence association studies and its implications for autism spectrum disorder. *Nat. Genet.* **50**, 727–736 (2018). doi: [10.1038/s41588-018-0107-y](https://doi.org/10.1038/s41588-018-0107-y); pmid: 29700473
34. G. Stelzer *et al.*, The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinformatics* **54**, 1.30.1–1.30.33 (2016). pmid: 27322403
35. V. M. Leppa *et al.*, Rare Inherited and De Novo CNVs Reveal Complex Contributions to ASD Risk in Multiplex Families. *Am. J. Hum. Genet.* **99**, 540–554 (2016). doi: [10.1016/j.ajhg.2016.06.036](https://doi.org/10.1016/j.ajhg.2016.06.036); pmid: 27569545
36. T. N. Turner *et al.*, Genomic Patterns of De Novo Mutation in Simplex Autism. *Cell* **171**, 710–722.e12 (2017). doi: [10.1016/j.cell.2017.08.047](https://doi.org/10.1016/j.cell.2017.08.047); pmid: 28965761
37. A. F. Pardiñas *et al.*, Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018). doi: [10.1038/s41588-018-0059-2](https://doi.org/10.1038/s41588-018-0059-2); pmid: 29483656
38. S. P. Pasca, The rise of three-dimensional human brain cultures. *Nature* **553**, 437–445 (2018). pmid: 29364288
39. A. M. Pasca *et al.*, Functional cortical neurons and astrocytes from human pluripotent stem cells in 3D culture. *Nat. Methods* **12**, 671–678 (2015). doi: [10.1038/nmeth.3415](https://doi.org/10.1038/nmeth.3415); pmid: 26005811
40. X. Qian *et al.*, Brain-Region-Specific Organoids Using Mini-bioreactors for Modeling ZIKV Exposure. *Cell* **165**, 1238–1254 (2016). doi: [10.1016/j.cell.2016.04.032](https://doi.org/10.1016/j.cell.2016.04.032); pmid: 27118425
41. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010). doi: [10.1093/bioinformatics/btp616](https://doi.org/10.1093/bioinformatics/btp616); pmid: 19910308
42. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018). doi: [10.1038/nbt.4096](https://doi.org/10.1038/nbt.4096); pmid: 29608179
43. C. Trapnell *et al.*, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014). doi: [10.1038/nbt.2859](https://doi.org/10.1038/nbt.2859); pmid: 24658644
44. <https://www.doi.org/10.1093/nar/gkn698>
45. J. Chen, E. E. Bardes, B. J. Aronow, A. G. Jegga, ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res.* **37**, W305–W311 (2009). doi: [10.1093/nar/gkp427](https://doi.org/10.1093/nar/gkp427); pmid: 19465376
46. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008). doi: [10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137); pmid: 18798982
47. J. Ernst, M. Kellis, ChromHMM: Automating chromatin-state discovery and characterization. *Nat. Methods* **9**, 215–216 (2012). doi: [10.1038/nmeth.1906](https://doi.org/10.1038/nmeth.1906); pmid: 22373907
48. A. Khan *et al.*, JASPAR 2018: Update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res.* **46** (D1), D260–D266 (2018). doi: [10.1093/nar/gkx1126](https://doi.org/10.1093/nar/gkx1126); pmid: 29140473
49. PsychENCODE Consortium, Integrative Analyses of iPSC-derived Brain Organoids, Synapse (2018). doi: [10.7303/syn12033248](https://doi.org/10.7303/syn12033248)

# ACKNOWLEDGMENTS

We are grateful to the National Institute of Mental Health (NIMH), and in particular to T. Lerner and G. Senthil, for their support and guidance for this project. We thank L. Tomasini for excellent technical assistance and members of the Program in Neurodevelopment and regeneration, particularly S. Weissman, A. Szekely, J. Mariani, S. Tomasi, and others for useful comments and discussions. We thank S. Akbarian for advice on ChIP-seq. We thank G. Wang, K. Bilguvar, S. Mane, C. Castaldi, and the Yale Center for Genome Analysis for library preparation and deep sequencing. A.Ab. is also a Visiting Professor at the Child Study Center, Yale University, New Haven, CT. **Funding:** Data were generated as part of the PsychENCODE Consortium, supported by U01MH103392, U01MH103365, U01MH103346, U01MH103340, U01MH103339, R21MH109956, R21MH105881, R21MH105853, R21MH103877, R21MH102791, R01MH111721, R01MH110928, R01MH110927, R01MH110926, R01MH110921, R01MH110920, R01MH110905, R01MH109715, R01MH109677, R01MH105898, R01MH105898, R01MH094714, P50MH106934 awarded to Schahram Akbarian (Icahn School of Medicine at Mount Sinai), Gregory Crawford (Duke University), Stella Dracheva (Icahn School of Medicine at Mount Sinai), Peggy Farnham (University of Southern California), Mark Gerstein (Yale University), Daniel Geschwind (University of California, Los Angeles), Fernando Goes (Johns Hopkins University), Thomas M. Hyde (Lieber Institute for Brain Development), Andrew Jaffe (Lieber Institute for Brain Development), James A. Knowles (University of Southern California), Chunyu Liu (SUNY Upstate Medical University), Dalila Pinto (Icahn School of Medicine at Mount Sinai), Panos Roussos (Icahn School of Medicine at Mount Sinai), Stephan Sanders (University of California, San Francisco), Nenad Sestan (Yale University), Pamela Sklar (Icahn School of Medicine at Mount Sinai), Matthew State (University of California, San Francisco), Patrick Sullivan (University of North Carolina), Flora Vaccarino (Yale University), Daniel Weinberger (Lieber Institute for Brain

Development), Sherman Weissman (Yale University), Kevin White (University of Chicago and Tempus Labs, Inc.), Jeremy Willsey (University of California, San Francisco), and Peter Zandi (Johns Hopkins University). Data were generated as part of the CommonMind Consortium supported by funding from Takeda Pharmaceuticals Company Limited, F. Hoffman-La Roche Ltd, and NIH grants R01MH085542, R01MH093725, P50MH066392, P50MH080405, R01MH097276, R01-MH-075916, P50MH096891, P50MH084053S1, R37MH057881 and R37MH057881S1, HHSN271201300031C, AG02219, AG05138, and MH06692. Brain tissue for the study was obtained from the following brain bank collections: the Mount Sinai NIH Brain and Tissue Repository, the University of Pennsylvania Alzheimer's Disease Core Center, the University of Pittsburgh NeuroBioBank and Brain and Tissue Repositories and the NIMH Human Brain Collection Core. CMC Leadership: Pamela Sklar, Joseph Buxbaum (Icahn School of Medicine at Mount Sinai), Bernie Devlin, David Lewis (University of Pittsburgh), Raquel Gur, Chang-Gyu Hahn (University of Pennsylvania), Keisuke Hirai, Hiroyoshi Toyoshima (Takeda Pharmaceuticals Company Limited), Enrico Domenici, Laurent Essioux (F. Hoffman-La Roche Ltd), Lara Mangravite, Mette Peters (Sage Bionetworks), Thomas Lehner, Barbara Lipska (NIMH). **Author contributions:** The authors contributed to this study at different levels, as described in the following. Study conception: F.M.V. and A.Ab. General experimental design: F.M.V., A.Ab., G.C., and F.W. Human brain tissue procurement and processing: F.L. Human brain tissue procurement and dissection and overall project management for the PsychENCODE developmental dataset: N.S. ChIP-seq data generation from the PsychENCODE developmental dataset: S.P. and Y.S. Data generation and experimental design for iPSCs, organoid differentiation, ChIP-seq and RNA-seq assays: A.Am. and S.S. Processing of RNA-seq and ChIP-seq data: F.W., G.C., and M.G. Chromatin segmentation, Hi-C analysis and genomic variant analyses: T.R. Analysis of RNA-seq data, network analyses: G.C. Analysis of ChIP-seq data, network analyses: F.W. qPCR validation: S.S. and A.Am. Coordination of analyses, data mining and integrative analyses across datasets: F.M.V., A.Ab., F.W., and S.S. ATAC-seq assay and data analysis: A.S., L.S., and G.E.C. Rhesus interspecies DGE: A.M.M.S. and Y.Z. Display item preparation: F.M.V., A.Ab., A.Am., S.S., F.W., G.C., and T.R. Manuscript writing: F.M.V., A.Ab., A.Am., S.S., F.W., G.C., and T.R. The following authors contributed equally to the study: A.Am., S.S., F.W., G.C., T.R. All authors participated in discussion of results and manuscript editing. **Competing interests:** We declare no competing financial interests related to this article. **Data and materials availability:** Data for this manuscript will be available from Synapse ([49](https://doi.org/10.7303/syn12033248)).

# SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/362/6420/eaat6720/suppl/DC1](https://www.sciencemag.org/content/362/6420/eaat6720/suppl/DC1)  
Materials and Methods  
Figs. S1 to S18  
Tables S1 to S17  
PsychENCODE Consortium Authors and Affiliations  
References ([50–75](#))

30 March 2018; accepted 16 November 2018  
[10.1126/science.aat6720](https://doi.org/10.1126/science.aat6720)

## Transcriptome and epigenome landscape of human cortical development modeled in organoids

Anahita Amiri, Gianfilippo Coppola, Soraya Scuderi, Feinan Wu, Tanmoy Roychowdhury, Fuchen Liu, Sirisha Pochareddy, Yurae Shin, Alexias Safi, Lingyun Song, Ying Zhu, André M. M. Sousa, The PsychENCODE Consortium, Mark Gerstein, Gregory E. Crawford, Nenad Sestan, Alexej Abyzov and Flora M. Vaccarino

*Science* **362** (6420), eaat6720.  
DOI: 10.1126/science.aat6720

### ARTICLE TOOLS

<http://science.sciencemag.org/content/362/6420/eaat6720>

### SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2018/12/12/362.6420.eaat6720.DC1>

### RELATED CONTENT

<http://science.sciencemag.org/content/sci/362/6420/1262.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat7615.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat8127.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat8464.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat8077.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat4311.full>  
<http://science.sciencemag.org/content/sci/362/6420/eaat6576.full>  
<http://stm.sciencemag.org/content/scitransmed/5/168/168ra7.full>  
<http://stm.sciencemag.org/content/scitransmed/5/168/168ra8.full>  
<http://stm.sciencemag.org/content/scitransmed/8/333/333ra48.full>  
[file/content](http://science.sciencemag.org/content/sci/362/6420/1262.full#file/content)

### REFERENCES

This article cites 74 articles, 10 of which you can access for free  
<http://science.sciencemag.org/content/362/6420/eaat6720#BIBL>

### PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2018 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works