Progenitor cell diversity in the developing mouse neocortex

Xiangbin Ruan[a,1], Bowei Kang[a,b,1], Cai Qi[a], Wenhe Lin[a], Jingshu Wang[b,2], and Xiaochang Zhang[a,c,2]

[a]Department of Human Genetics, The University of Chicago, Chicago, IL 60637; [b]Department of Statistics, The University of Chicago, Chicago, IL 60637; and [c]The Grossman Institute for Neuroscience, Quantitative Biology and Human Behavior, The University of Chicago, Chicago, IL 60637

Edited by Ulrich Mueller, Johns Hopkins University, and accepted by Editorial Board Member Jeremy Nathans January 29, 2021 (received for review September 8, 2020)

In the mammalian neocortex, projection neuron types are sequentially generated by the same pool of neural progenitors. How neuron type specification is related to developmental timing remains unclear. To determine whether temporal gene expression in neural progenitors correlates with neuron type specification, we performed single-cell RNA sequencing (scRNA-Seq) analysis of the developing mouse neocortex. We uncovered neuroepithelial cell enriched genes such as *Hmga2* and *Ccnd1* when compared to radial glial cells (RGCs). RGCs display dynamic gene expression over time; for instance, early RGCs express higher levels of *Hes5*, and late RGCs show higher expression of *Pou3f2*. Interestingly, intermediate progenitor cell marker gene *Eomes* coexpresses temporally with known neuronal identity genes at different developmental stages, though mostly in postmitotic cells. Our results delineate neural progenitor cell diversity in the developing mouse neocortex and support that neuronal identity genes are transcriptionally evident in *Eomes*-positive cells.

scRNA-Seq | neural progenitor | neuronal diversity | developmental timing | cell identity

The six-layered neocortex is evolutionarily unique in mammals and forms the physical center for the highest cognitive and information-processing functions (1). Diverse projection neuron types have been uncovered in the neocortex based on their morphology, connectivity, gene expression, and other properties (2). In rodents, cortical projection neuron types are generated sequentially by radial glial progenitor cells (RGCs) in the ventricular zone (VZ) and intermediate progenitor cells (IPCs) in the subventricular zone (SVZ) (3, 4). Neuroepithelial cells (NECs) undergo fast symmetric cell division before transitioning into RGCs which generate projection neurons in the neocortex (4). Early-born neurons are located in the deep layers V–VI and project mainly to subcortical regions, while late-born neurons populate the superficial layers II–IV and project predominantly to contra- or ipsilateral cortices (2). How neuronal birthdate is associated with its identity remains an incompletely understood question.

Dynamic gene expression is a predominant factor for cortical neuron fate specification: Genetic mapping in humans and rodents uncovered essential genes that are required for cortical neurogenesis and lamination (5, 6), large-scale RNA in situ hybridization (ISH) and expression profiling of microdissected brain tissues uncovered transcription programs that associate with cortical layers and areas (7–10), and analysis of neuronal types by retrograde labeling and knockout mice unambiguously demonstrated roles of transcription factors in neuronal fate specification (2, 11). More recently, single-cell RNA sequencing (scRNA-Seq) of adult mouse and human brains reported dozens of cortical cell types (12–15), and single-cell analyses of neocortical progenitors revealed molecular and cellular heterogeneity (16–21). scRNA-Seq of mouse cortical development reported a core RGC transcriptional program that is established at embryonic day 13.5 (E13.5) and maintained during embryonic neurogenesis (19). scRNA-Seq of pulse-tagged apical progenitors identified temporally regulated genes that switch from internally directed to more exteroceptive over time (22). Thus, cortical progenitor cells appear to encode temporal information for neuron type specification.

It has long been considered that intrinsic and sequential expression of transcription regulators in neural progenitors determines the sequential production of neuronal types. The transcription-cascade model is strongly supported by fly genetics (23), and progenitor-encoded cell lineages were demonstrated by clonal analysis in mouse neocortical cells (24). Classical heterochronic transplantation experiments supported that late neocortical progenitors had progressively restricted differentiation potential in ferrets (25, 26). Recent heterochronic transplantation experiments reported that late mouse apical progenitors retained temporal plasticity and could generate deep-layer neurons when transplanted to early brains (27). It is intriguing whether neural progenitors have temporal plasticity or progressively restricted differentiation potential. Analysis of tagged and purified apical progenitors uncovered progressive changes of transcriptional states over time (22), and it remains unclear whether neuron type-specific transcription factors are sequentially expressed in neocortical progenitors.

We investigated neural progenitor cell diversity at different developmental stages and sought to determine temporal gene expression in cortical progenitors. We performed scRNA-Seq analyses with cortical cells isolated from six developmental time points (E10.5 through E18.5) that span neuroepithelium expansion and neurogenesis. Our analysis uncovered transcriptional diversity among NECs, RGCs, and IPCs and supports that

Significance

How projection neuron types are temporally and sequentially generated in the mammalian neocortex remains unclear. We performed single-cell RNA sequencing analysis of embryonic day (E) 10.5 through E18.5 mouse neocortical cells and identified progenitor cell types across development. Our results uncovered molecular signatures for neuroepithelial cells and temporal gene expression in radial glial progenitors. Importantly, *Eomes*-positive cells display temporal expression of previously characterized neuronal identity genes. These results delineate neocortical progenitor cell diversity and the timing of neuron type specification.

Author contributions: X.Z. conceived the study; X.R., B.K., C.Q., W.L., and X.Z. performed research; X.R., B.K., W.L., J.W., and X.Z. analyzed data; J.W. and X.Z. supervised research; and B.K. and X.Z. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. U.M. is a guest editor invited by the Editorial Board.

[1]X.R. and B.K. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: jingshuw@uchicago.edu or xczhang@uchicago.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2018866118/-/DCSupplemental.

Published March 1, 2021.

previously characterized neuronal cell identity genes are transcriptionally evident in *Eomes*-positive cells.

Results

scRNA-Seq Analysis of Cell Types and Lineages in the Developing Mouse Neocortex. We performed droplet-based scRNA-Seq (Drop-Seq) with dorsal forebrain cells collected from E10.5, E12.5, E14.5, E15.5, E16.5, and E18.5 mouse embryos (28), sampling both cortical and hippocampal cells. We merged all cells with Seurat (29) and obtained 10,261 cells across six developmental stages after filtering out doublets and low-quality cells (*SI Appendix*, Fig. S1 *A* and *B*). Analysis of all filtered cells identified 19 cell clusters (Fig. 1 *A* and *B*; referred to as C0-C18 hereafter), with sampling and sequencing replicates distributed consistently across clusters (*SI Appendix*, Fig. S1*C*). We assigned cell types to individual clusters based on their marker gene expression and identified neural progenitors, neuron types, glial cells, and nonneural cells (Fig. 1 *A*–*D* and *SI Appendix*, Fig. S1*D*).

NECs from E10.5 showed high *Hmga2* expression and formed distinct clusters from later-stage neural progenitors (C6 and C14; Fig. 1 *C* and *D*). RGCs across different stages formed a coherent group of cells that showed high expression of *Pax6*, *Id4*, and mitosis genes. RGCs were further separated into three clusters (C1–C3–C10) by their mitotic phases. *Eomes* highlighted IPCs that formed two major clusters: mitotic cells in C7 and nonmitotic cells in C4. Postmitotic projection neurons expressed *Neurod1*, *Neurod6*, and subtype-specific genes: Immature neurons formed cluster C0 and lacked mature-neuron genes such as *Stmn2*; layer V–VI neurons expressing *Bcl11b* were clustered in C2 and C11, with C11 showing higher level of *Tle4*; superficial layer neurons expressing *Satb2* were clustered in C5 (Layer II–IV; Fig. 1 *A*–*D*). Hippocampal cells expressing high levels of *Crym* and *Zbtb20* were clustered in C9 (Fig. 1*C* and *SI Appendix*, Fig. S1*D*). We also observed Cajal–Retzius cells (layer I) that expressed *Reln* and *Lhx5* in C13, interneurons expressing *Dlx1* and *Gad2* in C8, and oligodendrocyte progenitors (OPCs) expressing *Tnc* and *Olig1* mostly from E18.5 and clustering in C12. Nonneural cells formed distinct clusters: endothelial cells expressing *Cldn5* in C16, microglia expressing *Tyrobp* and *Ly86* in C17, and pericytes in C18 (*SI Appendix*, Fig. S1*E*). Thus, our data identified diverse types of neural progenitors, neurons, and nonneural cells in the developing mouse neocortex.

To understand cell lineages, we removed interneurons and nonneural cells that were not derived from dorsal progenitors (*SI Appendix*, Fig. S1*F*) and performed RNA velocity and trajectory analyses on the remaining cells (Fig. 1 *E* and *F*). After regressing out cell cycle genes, three RGC clusters (C1–C3–C10) mixed well, and the two main IPC clusters also merged together (*SI Appendix*, Fig. S1*F*). Results of both RNA velocity analysis using scVelo (30) (Fig. 1*E*) and trajectory inference using Slingshot (31) (Fig. 1*F*) were projected onto a shared two-dimensional visualization of the cells using PAGA (partition-based graph abstraction) (32). scVelo and Slingshot results confirmed an NEC–RGC–IPC–neuron trajectory in the developing mouse neocortex. Interestingly, cells expressing *Eomes*, or mostly IPCs in C4 and C7, formed a distinct neck between neural progenitors and postmitotic neurons which were both spread to certain extents based on their developmental origins (Fig. 1 *E* and *F*). This is consistent with prior knowledge that *Eomes*-positive cells are derived from apical progenitors and produce over 80% of cortical projection neurons across all layers (33), suggesting a relatively convergent transcription program for IPCs at different developmental stages. To investigate neural progenitor heterogeneity and its connection to neuron types we focus on progenitor cells hereafter.

Transcription Programs Underlying Neuroepithelium-to-RGC Transition. After neural tube closure around E9.5, NECs line along the lateral ventricles and divide symmetrically to expand the progenitor pool (4). Cortical neurogenesis begins around E11.5, and concomitantly NECs extend radial fibers to become RGCs (2). NECs have a shorter cell cycle and divide symmetrically to increase their pool, while RGCs divide both asymmetrically and symmetrically for neurogenesis and self-renewal (4, 34). The transition from NECs to RGCs thus marks the elevation of progenitor cell differentiation. While it is technically challenging to isolate NECs, scRNA-Seq provides a unique way to unbiasedly investigate molecular switches mediating the NEC-to-RGC transition.

E10.5 neural progenitors showed *Sox2* expression in two distinct clusters C6 and C14 (Fig. 2*A*). C14 cells were mostly from E10.5, with few from later stages, and expressed high levels of *Wnt8b* and *Id3*, which mark cortical hem and the dorsomedial wall of the telencephalic ventricles (35). Interestingly, *Id4* and *Id3* expression marked the boundaries between dorsal neuroepithelium and the hem, suggesting their roles in brain area patterning (Fig. 2 *B* and *C*). We compared NECs (C6) and RGCs (C1–C3–C10) and identified distinct transcription programs (Fig. 2 *D*–*F* and text that follows).

C6 NECs from E10.5 showed higher expression of *Hmga2*, *Mest*, *Crabp2*, *Dlk1*, *Meg3*, and *Ccnd1* among other genes (Fig. 2 *D* and *E* and *SI Appendix*, Fig. S2 *A* and *B*). *Hmga2* was shown to promote self-renewal of neural stem cells (36), which is consistent with its high expression in NECs for expanding progenitor pools. Expression of *Ccnd1* promotes G1–S cell cycle progression and suppresses progenitor cell differentiation, which is consistent with shorter G1 length in NECs and proliferative cortical progenitors (34). *Dlk1* encodes a Delta-like noncanonical Notch ligand 1 and might play a role in NEC proliferation. *Crabp2* facilitates the binding of retinoic acid to its receptors and appears concentrated at the basal endfeet of apical progenitors (37). In contrast, RGCs in clusters C1–C3–C10 showed higher expression of *Ptn*, *Aldoc*, *Fabp7*, *Nfib*, *Tcf4*, and *Hes5* among other genes (Fig. 2 *D* and *E* and *SI Appendix*, Fig. S2*C*). *Aldoc*, *Fabp7*, and *Hes5* are known markers for RGCs, and *Hes5* is a *Notch* effector important for RGCs identity. *Tcf4* heterozygous knockout mice displayed defects in neurogenesis, neuronal migration, and differentiation, suggesting that *Tcf4* functions in RGCs and neurons (38). We further plotted the expression of NEC-enriched genes across all time points and found that the trend largely remained except for *Rpl26* (Fig. 2*F*): NEC-enriched genes were specific to E10.5 and showed decreased expression in RGCs during development. Our analyses uncovered distinct transcriptional programs that mediate the NEC-to-RGC transition.

Developmental Regulation of Radial Glial Progenitor Cells. We compared RGCs from different time points for differential gene expression. The single-cell analysis identified RGCs in clusters C1–C3–C10, representing G1–S- and G2–M-phase cells. To avoid confounding cell cycle effects, we compared E12.5 and E14.5 cells in clusters C1, C3, and C10 separately (Fig. 3*A*). E12.5 RGCs expressed higher levels of genes such as *Lix1*, *Ccnd1*, and *Gpc1*, and E14.5 RGCs showed higher expression of *Nfix*, *Ndrg2*, *Aldoc*, and other genes (Fig. 3*B* and *SI Appendix*, Fig. S3*A*). We further combined RGC from E15.5 and E16.5 (140 cells in total), compared them with RGCs from E12.5 and E14.5, and identified differentially expressed genes between any two of the three samples/time points (*SI Appendix*, Fig. S3*B*). A number of dynamically expressed RGC genes are associated with cell cycle and/or progenitor-cell identity, such as *Mki67*, *Aspm*, *Cenpe*, *Cenpf*, *Id4*, and *Hes5* (*SI Appendix*, Fig. S3*B*). We focused on transcriptional regulators and found *Pou3f2*, *Zbtb20*, and *Nfia*/*Nfib*/*Nfix* showed higher expression in late RGCs (Fig. 3 *C* and *D*). Increased expression of *Nfia*/*Nfib*/*Nfix* in RGCs during development are consistent with their roles in late-born neurons and corpus callosum formation (39, 40). The higher expression of *Pou3f2* in late RGCs is consistent with previous reports that *Pou3f2* is expressed in the ventricular zone and upper-layer

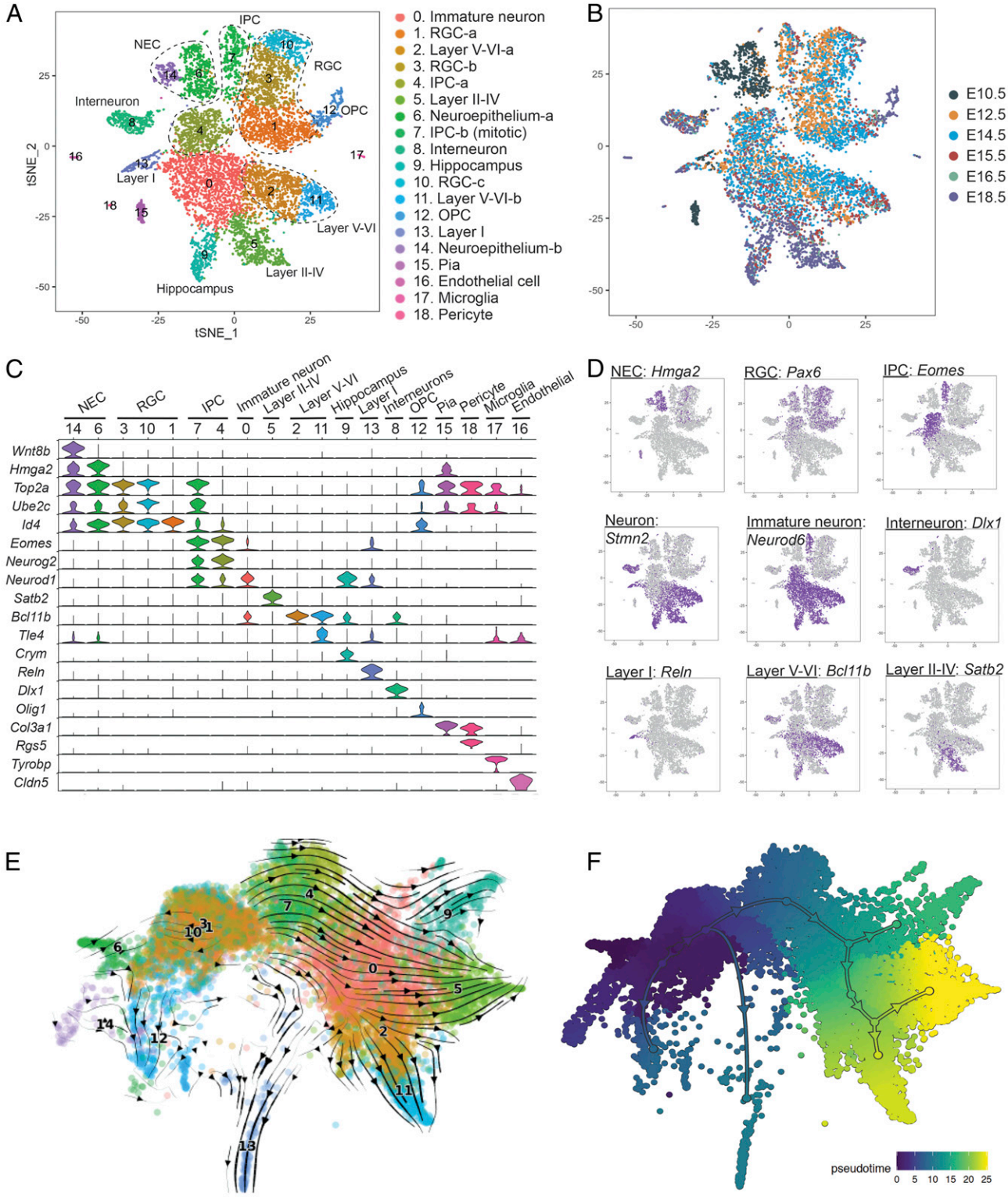

Fig. 1. scRNA-Seq uncovers cell types and lineages in the developing mouse neocortex. (*A*) tSNE plot showing 19 clusters of cells and their assigned identities from developing mouse forebrains (E10.5, E12.5, E14.5, E15.5, E16.5, and E18.5). NECs, RGCs, IPCs, OPCs, and neurons from different cortical layers and the hippocampus are highlighted. (*B*) Distribution of cells isolated from different developmental stages. (*C*) Violin plots showing marker gene expression for individual cell clusters. (*D*) Feature plots showing expression of molecular markers for NECs, RGCs, IPCs, and neuronal types. (*E*) FA plot showing RNA velocity analysis of dorsal progenitor-derived cells after adjusting for cell cycle effects. Numbers and color codes represent the same cell clusters assigned in A. (*F*) Pseudo time analysis of the cell trajectory with NECs set as the root (C6, dark blue). Colors represent the relative pseudo time of each cell.

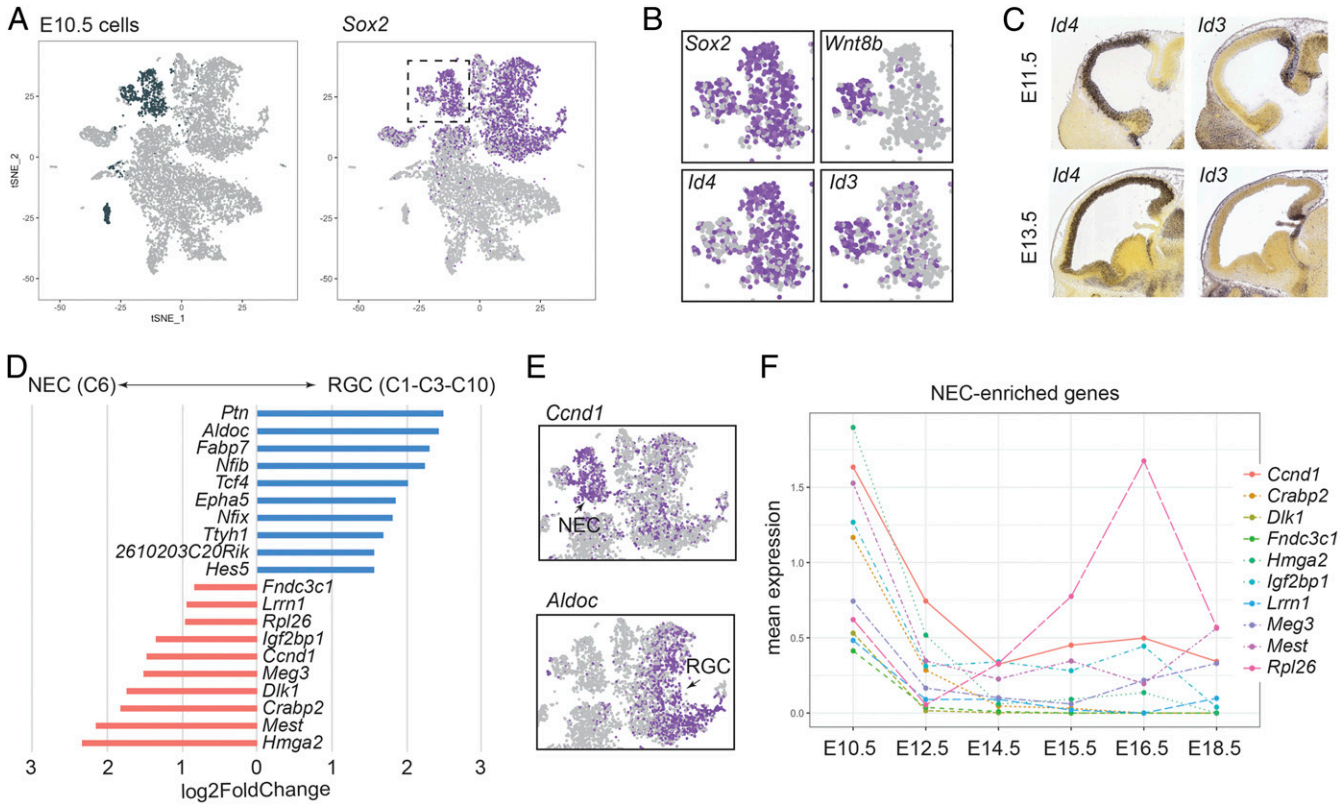

Fig. 2. Transcriptional switches underlying neuroepithelium-to-RGC transition. (*A*) E10.5 neural progenitors (*Left*) show *Sox2* expression (*Right*). (*B*) Feature plots showing that *Sox2*-positive NECs are separated into two populations marked by *Wnt8b-Id3* and *Id4* expression. (*C*) RNA ISH images showing that high *Id4* expression delineates the dorsal forebrain on sagittal sections of E11.5 and E13.5 mouse brains, in contrast to *Id3*. Images were adapted from Allen Developing Mouse Brain Atlas (2008). (*D*) Differentially expressed genes that are enriched in NECs (red, cluster 6 in Fig. 1*A*) and RGCs (blue, C1–C3–C10). (*E*) Feature plots showing that *Ccnd1* is highly expressed in NECs, and *Aldoc* is enriched in RGCs. (*F*) Expression of NEC-enriched genes in apical progenitors across different developmental stages. Shown are mean levels of scaled gene expression in C6 for E10.5 or in C1–C3–C10 for E12.5–E18.5 cells.

neurons (Fig. 3 *C* and *D*) (41). Our comparison of E12.5, E14.5, and E15.5–E16.5 RGCs identified differential gene expression that may encode temporal cues for neurogenesis.

***Eomes*-Positive Cells Temporally Express Neuronal Identity Genes.** *Eomes*-positive IPCs are derived from NECs and RGCs and generate the majority of cortical neurons directly (42). Heterochronic transplantation experiments suggested that IPCs from E15.5 might have been specified to the superficial-layer lineage (27), but it is unclear when and how IPCs adopted layer-specific identity. We sought to understand IPC diversity at the gene expression level and test the hypothesis that *Eomes*-positive cells at different developmental stages are associated with corresponding cortical neuron identity.

Eomes was predominantly expressed in clusters C4 and C7, as well as by subsets of cells in C0 and C13 (Fig. 1 *A* and *D*). We identified *Eomes*-positive cells in C0 and C13 and pooled them with cells in C4 and C7 for downstream analysis (Fig. 4*A* and *SI Appendix*, Fig. S4*A*). Subclustering of all *Eomes*-positive cells uncovered eight subgroups: mitotic cells in SC2 and SC4 (subclusters 2 and 4), *Neurog2*-positive cells in SC0 and SC3, and *Neurod1* was enriched in SC7–SC1–SC5–SC6 (Fig. 4*A* and *SI Appendix*, Fig. S4*A*).

Among *Eomes*-positive cells, individual *Neurod1*-positive and *Neurog2*-negative clusters were predominantly formed by cells from distinct developmental stages: 90% of cells in SC7 were from E10.5, 88% of SC1 cells were from E12.5, 64% of SC5 were E14.5, and all cells in SC6 were from E14.5 or later (Fig. 4*B*). Interestingly, *Eomes*-positive cells in SC7–SC1–SC5–SC6 coexpress layer-specific genes (Fig. 4*B*): *Eomes*-positive cells in SC7 (mostly from E10.5) coexpressed with Cajal–Retzius cell identity genes such as *Ebf2* and *Lhx5*, *Eomes*-positive cells in SC1 (mostly from E12.5) coexpressed deep-layer neuronal genes *Bcl11b* and *Gng3*, and *Eomes*-positive cells in SC5 and SC6 showed higher expression of superficial layer genes such as *Pou3f3* and *Bhlhe22*. Coexpression of *Eomes* messenger RNA (mRNA) with neuron type-specific transcription factors was evident in single cells (Fig. 4*C* and *SI Appendix*, Fig. S4*B*). We further confirmed coexpression of Eomes protein with lineage-specific transcription factors by immunostaining on brain sections: Eomes colocalized with Ebf2 and Tbr1 at E10.5, partially coexpressed with Bcl11b at E12.5, and costained with Pou3f3 in the SVZ at E14.5 (Fig. 4*D*). These results indicate that *Eomes* coexpresses with known neuron identity genes at the onset of cortical neuron fate specification.

To determine whether neuronal lineage gene expression in *Eomes*-positive cells occurs during or after mitosis, we ordered *Eomes*-positive cells by their estimated pseudo time (Fig. 1*F*) and calculated cell cycle scores for individual cells (*SI Appendix*, Fig. S4*C*). In the estimated pseudo time, *Neurog2*-positive cells were placed before *Neurod1* and *Neurod6* (*SI Appendix*, Fig. S4*C*), which confirms *Neurog2* as a proneural gene in the dorsal telencephalon (43). Interestingly, a significant portion of *Eomes*-positive cells expressing lineage-specific genes, such as *Tbr1-Bcl11b-Zbtb20-Bhlhe22*, showed lower cell cycle scores and were placed after mitotic cells (*SI Appendix*, Fig. S4*C*),

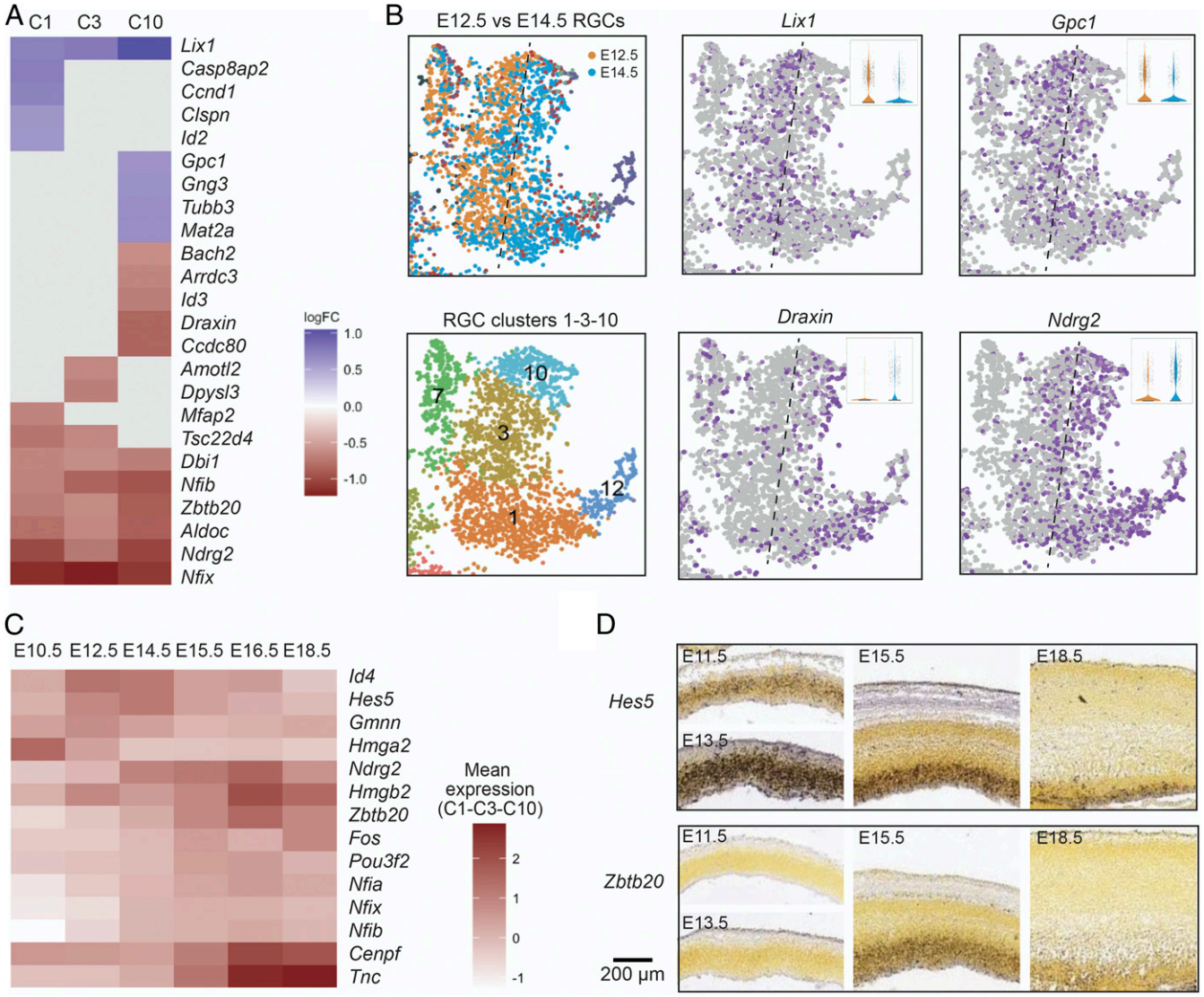

Fig. 3. Developmental regulation of gene expression in RGCs. (*A*) A heat map showing genes that are differentially expressed in E12.5 (blue) and E14.5 (red) cells in RGC clusters C1, C3, and C10. (*B*) Feature plots and violin plots (*Inserts*) showing that *Lix1* and *Gpc1* display higher expression in E12.5 RGCs, *Draxin*, and *Ndrg2* show higher expression in E14.5 RGCs. (*C*) A heat map showing genes that are predicted to regulate gene expression (GO: 165158) and are differentially expressed in RGCs across time. Averages of scaled gene expression for cells in C6 (E10.5) and C1–C3–C10 (E12.5–E18.5) are shown. Genes were identified by pairedwise comparisons of RGCs (C1–C3–C10) from E12.5, E14.5, and E15.5/E16.5. (*D*) RNA ISH images showing *Hes5* and *Zbtb20* expression at different developmental stages. *Hes5* shows higher levels of expression at E13.5, and *Zbtb20* shows higher levels of expression at E15.5. Images were adapted from Allen Developing Mouse Brain Atlas (2008).

suggesting that these cells were largely postmitotic. On the other hand, substantial amounts of *Eomes*-positive cells coexpressing *Ebf2*, *Meis2*, or *Pou3f3* showed high cell cycle scores (*SI Appendix*, Fig. S4*C*). We further costained Eomes and Ebf2 with cell cycle marker Ki67 on E10.5 dorsal brain sections and found that 31% of Eomes-Ebf2 double-positive cells were positive for Ki67 (*SI Appendix*, Fig. S4*D*). These results support that *Eomes*-positive cells from different developmental times start to express neuron type-specific genes.

Discussion

Progressive restriction of differentiation potential was considered an intrinsic program in neural progenitor cells (24, 25), and recent heterochronic transplantation experiments suggested alternative explanations (27). It remains unclear how cortical neuron fates are encoded in progenitor cells, and how RGCs and IPCs give rise to distinct neuronal subtypes. We studied the developing mouse neocortex by scRNA-Seq and identified 1) transcriptional programs that specify progenitor and neuron types in the developing neocortex, 2) dynamic expression of transcriptional regulators in RGCs, and 3) transcription of neuron type-specific genes in *Eomes*-positive cortical cells.

Our single-cell analyses identified neural progenitors, neurons, glial cells, and nonneural cells in the developing mouse neocortex. We identified marker genes specific to individual cell types (Fig. 1 and Datasets S1–S4). RNA velocity and pseudo time analyses showed the transition of NECs, RGCs, and IPCs to cortical neuron types. Importantly, we found a distinct transcription program in NECs at E10.5, where high levels of *Ccnd1* and *Hmga2* appear to promote cell cycle progression (Fig. 2 and Dataset S2). Datasets presented here may provide valuable information to guide further gain- and loss-of-function studies of

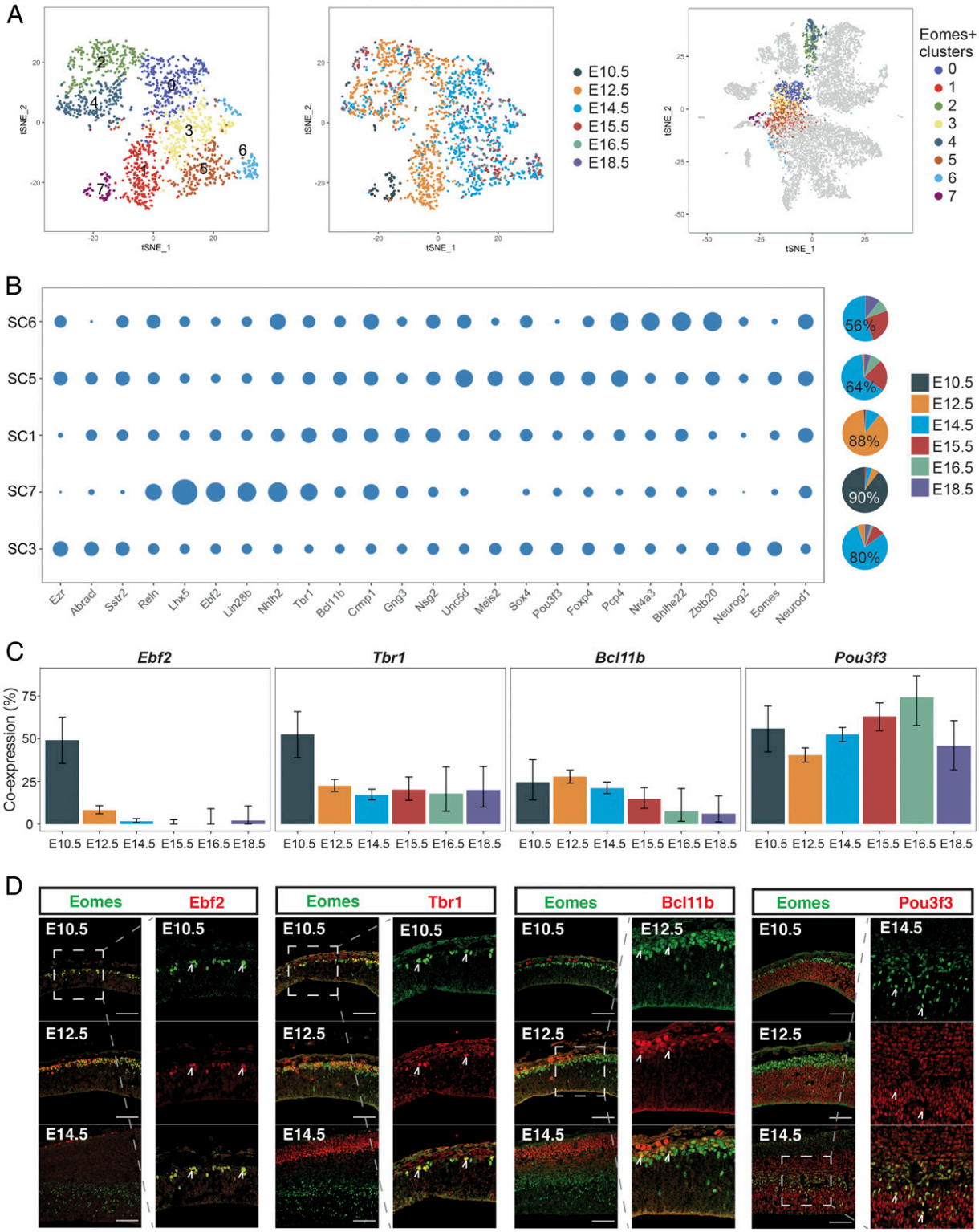

Fig. 4. *Eomes*-positive cells express neuron type-specific genes. (*A*) tSNE plot showing cellular heterogeneity of *Eomes*-positive cells (*Left*), their developmental origin (*Middle*), and distribution in Fig. 1*A* clusters (*Right*). (*B*) Developmental stages drive subclustering of *Eomes*-positive cells into clusters SC7 (90% of cells are from E10.5), SC1 (88% E12.5), and SC5 and SC6 (E14.5 and later) that cotranscribe previously characterized neuron type-specific genes. Bubble plots show the average scaled gene expressions in each subcluster. (*C*) Bar plots showing proportions of single cells that display coexpression of neuron type-specific genes with *Eomes* at different developmental stages: *Ebf2* and *Tbr1* show the highest ratio of expression in E10.5 *Eomes*-positive cells, *Bcl11b* shows the highest ratio of expression in E12.5 *Eomes*-positive cells, and *Pou3f3* shows increasing ratios of expression in E14.5–E16.5 *Eomes*-positive cells. Error bars were calculated based on the exact binomial distribution and indicate 95% confidence intervals of the estimated proportions of single cells that coexpressed neuron type-specific genes with *Eomes* at different developmental stages. (*D*) Immunostaining showing that Eomes (green) coexpresses with Ebf2 and Tbr1 at E10.5, with Bcl11b (Ctip2) at E12.5, and with Pou3f3 (Brn1) in the SVZ at E14.5. (Scale bars, 100 μm.)

cell type-specific genes in neocortex development with either in utero approaches or genetically modified animals.

The birth dates of RGC-derived projection neurons are closely associated with their layer and neuron type identities. Our analysis of E12.5, E14.5, and E15.5–E16.5 RGCs uncovered dozens of differentially expressed genes, many of which are associated with cell division and progenitor cell identity (Fig. 3 and *SI Appendix*, Fig. S3). For instance, the Notch effector *Hes5* showed higher expression in early RGCs. We also found higher expression of *Pou3f2* in late RGCs, which is consistent with the previous report that *Pou3f2* is expressed in the ventricular zone, suppresses *Hes5* expression, and marks upper-layer neurons (41). Dynamic expression of other transcription regulators such as *Nfia*/*Nfib*/*Nfix* in RGCs might interact with epigenetic regulators such as the PRC2 complex to influence neurogenesis (22). The unbiased analysis of dynamic gene expression in RGCs supports that RGCs encode temporal cues for neurogenesis.

We also sampled the dorsal medial cortex and the hippocampus and identified a cluster of hippocampal cells mainly from E18.5 (C9 in Fig. 1*A*). Interestingly, single hippocampal cells aligned in a way that closely mimicked their location in the mouse brain (*SI Appendix*, Fig. S1*D*). The *Id3*-positive cluster C14 appears to include E10.5 progenitor/primordial cells for the hippocampus, but we were unable to identify a major cluster of later hippocampal progenitors, probably because of the limited total number of cells, or to a lesser extent their molecular similarities to cortical progenitors. *Id3*- and *Wnt8b*-positive cells in C14 not only included E10.5 cells but also a few cells from E12.5 and after, which were probably progenitors from the hem and/or the hippocampus. Further studies are required to understand the bifurcation of hippocampal and cortical progenitor lineages.

The majority of cortical projection neurons are directly derived from IPCs (42), and it remains unclear how IPCs are associated with projection neuron type specification. Initial analysis of *Eomes*-positive IPCs uncovered mitotic and nonmitotic populations (Fig. 1*A*). Subclustering of *Eomes*-positive cells showed that developmental stages drove cluster formation and uncovered the temporal coexpression of *Eomes* with neuron type-specific genes in cell clusters that were correlated with developmental timing (Fig. 4). We further showed that neuron type-specific genes coexpressed with Eomes in cells that frequently had lower cell cycle scores and were negative for Ki67 (*SI Appendix*, Fig. S4 *C* and *D*). Our results suggest that *Eomes*-positive cells temporally express neuronal identity genes when IPCs are transitioning to newborn neurons.

We found that *Eomes* mRNA and protein coexpressed with *Ebf2* and *Lhx5* at E10.5, which are specifically expressed in layer I (Fig. 4 *B–D*). Interestingly, an independent study showed that Eomes is required to suppress *Ebf1*/*Ebf2*/*Ebf3* expression in mice, and loss of Eomes led to overproduction of *Ebf1*/*Ebf2*/*Ebf3*-positive cells in the dorsal cortex (44). These results indicate that Eomes can transiently coexpress with and regulate neuronal identity genes. Further studies, such as lineage tracing of *Eomes* and neuronal identity gene double-positive cells, are required to confirm whether such temporal coexpression faithfully predicts neuronal fate specification. Recessive silencing of *EOMES* was associated with microcephaly in humans (45), and deletion of *Eomes* in mice impairs IPC proliferation (42, 46). Future analysis of cell types and lineages in *Eomes* knockout mice may help to elucidate how *Eomes* is associated with temporal and sequential production of neuron types.

Methods

scRNA-Seq and Molecular Experiments. All mouse-related experiments were reviewed and approved by the Institutional Animal Care and Use Committee at the University of Chicago. CD1-timed pregnant mice were ordered from Charles River, and dorsal cortical tissues from E10.5–E18.5 embryos were dissected and dissociated with papain. Single-cell collection and library preparation followed the Drop-Seq protocol v3.1 (28). All samples were barcoded, pooled together, and sequenced in two runs on Illumina NextSeq 550. E14.5 cells were collected in two batches as described in *SI Appendix*, Fig. S1C. RNA ISH data were adapted from the Allen Developing Mouse Brain Atlas (2008). For immunostaining, embryonic brains were fixed in 4% paraformaldehyde overnight at 4 °C, cryoprotected in 25% sucrose overnight at 4 °C, embedded in Frozen Section Medium (6502; Thermo Scientific), and sectioned at 14-μm thickness in coronal direction. Slices were rinsed with 1× phosphate-buffered saline (PBS) for 5 min, incubated with blocking buffer (1× PBS containing 0.03% Triton X-100 and 5% normal donkey serum) at room temperature for 30 min, and further incubated with primary antibodies diluted in PBST buffer (1× PBS containing 0.03% Triton X-100) overnight at 4 °C. After washing three times with 1× PBS, slides were incubated for 1 h at room temperature with fluorophore-conjugated secondary antibodies in the dark. Slides were scanned with a Leica SP5 confocal microscope. The following primary antibodies were used: anti-Tbr2 (Millipore AB15894, chicken, 1:800), anti-EBF2 (R&D Systems AF7006-SP, sheep, 1:50), anti-Tbr1 (Abcam ab31940, rabbit, 1:1,000), anti-Bcl11b (Abcam ab18465, rat, 1:1,000), anti-Pou3f3 (Novus Biologicals NBP2-57011, rabbit, 1:1,000), and anti-Ki67 (Abcam ab15580, rabbit, 1:800). The secondary antibodies were all diluted at 1:2,000 in PBST buffer: donkey anti-chicken 488 (Jackson ImmunoReseach, 703-546-155), donkey anti-sheep 594 (Thermo Scientific, A11016), donkey anti-rabbit 594 (Thermo Scientific, A21207), and donkey anti-rat 594 (Thermo Scientific, A21209).

Data Processing and Cell Clustering. Sequencing reads were trimmed and processed using Drop-Seq tools (2.0.0) to obtain a unique molecular identified (UMI) count matrix. We used the Seurat package (3.1.5) (29) in R (4.0.2) and filtered out cells with total UMI counts over 6,000 or below 500 or that had over 10% mitochondrial gene counts. This resulted in a data matrix of 21,862 genes and 10,261 cells. The count matrix was then normalized by the library size, log-transformed, and scaled, following the workflow in Seurat tutorial (https://satijalab.org/seurat/) with default parameters unless otherwise stated.

Using Seurat, we performed principal component analysis dimensionality reduction using the top 500 genes with the highest expression variation across cells. Based on the Euclidean distance in the space spanned by the first 16 identified principal components, we constructed a shared nearest neighbor graph and applied the Louvain algorithm to identify clusters (47). For our initial clustering, we used default resolution = 0.8 and obtained 19 clusters (Fig. 1*A*). The *Eomes*-positive cells were defined to be the union of the original C7, a subset of C0 and C4, and a subset of C13 (1,849 cells in total), where the subsets were determined by performing subclustering and selecting *Eomes*-positive subclusters. We further identified eight subclusters in the *Eomes*-positive cells using 2,000 most variable genes and 20 principal components (Fig. 4*A* and *SI Appendix*, Fig. S4*A*).

Visualization and Differential Expression Analysis. We applied T-distributed stochastic neighbor embedding (tSNE) algorithm to visualize the raw data following the Seurat pipeline (Fig. 1*A*). Differential gene expression was analyzed using Seurat::FindMarkers with default parameters. The final marker genes we show were determined by integrating differential expression test *P* values (<0.05) and average log fold change (>0.6 or 1.8-fold unless stated otherwise). The full lists of differentially expressed genes are in Datasets S1–S4.

Trajectory Inference and RNA Velocity. We regressed out cell cycle genes before trajectory and RNA velocity analyses. The set of cell cycle genes are defined by the Gene Ontology (MGI, GO 0007049, 667 genes). We extracted the top five principal components from the cell cycle gene expression space, used them as proxies to cell cycle effects, and regressed them out with the ScaleData function in Seurat before cell trajectory was analyzed. We also calculated the cell cycle scores of each cell using the Seurat::CellCycleScoring function for *SI Appendix*, Fig. S4*C*.

We used the FA (force-directed atlas) plot (48) from PAGA (32) as the two-dimensional projection of cells in the trajectory and RNA velocity analysis in Fig. 1 *E* and *F*. After regressing out cell cycle genes, the 2,000 most highly variable genes from the residual matrix were used as the input to PAGA and initialized with the Seurat clusters in Fig. 1*A*. Interneurons and nonneural cells (clusters C8 and C15–C16–C17–C18) were not derived from dorsal progenitors and they were excluded from trajectory analyses. The trajectory inference was performed using Slingshot (v2.0.1) (31) from the Dyno

platform (49): We reclustered cells using Seurat and removed 9 outlier cells, set NECs from E10.5 as the root, and computed pseudotime using the dyno R package (Fig. 1*F*). Using velocyto (v0.6) (50), we obtained loom files of the spliced and unspliced RNA matrices. We then applied scVelo (v0.2.2) (30) to estimate transient cell states and velocities (Fig. 1*E*).

Data Availability. R codes used to replicate the figures and analysis results in this paper are available in GitHub at https://github.com/kangbw702/Progenitor-cell-diversity (51). Other packages used in this study: dplyr (1.0.0), ggplot2 (3.3.1), tidyr (1.1.0), ggsci (2.9), viridis (0.5.1), and patchwork (1.0.0). Raw sequence data and filtered gene by cell expression matrices are available on National Center for Biotechnology Information Gene Expression Omnibus (accession no. GSE161690) (52).

ACKNOWLEDGMENTS. The authors thank Elizabeth Grove, Christopher A. Walsh, Xiaoxi Zhuang, and Marcelo Nobrega for critical comments on the manuscript; Rong Zhong, Denise Fischer, Benjamin Weaver, Christina Astley, and Yu-Han Hsu for technical assistance; Pieter W. Faber and the Genomics Facility at The University of Chicago for DNA sequencing; and thank Peter Carbonetto and the Research Computing Center for hosting data analysis. This work was partially supported by grants from the National Institute of Mental Health (K01 MH109747) and the National Institute of General Medical Sciences (DP2 GM137423) to X.Z.

1. D. H. Geschwind, P. Rakic, Cortical evolution: Judge the brain by its cover. *Neuron* **80**, 633–647 (2013).
2. L. C. Greig, M. B. Woodworth, M. J. Galazo, H. Padmanabhan, J. D. Macklis, Molecular logic of neocortical projection neuron specification, development and diversity. *Nat. Rev. Neurosci.* **14**, 755–769 (2013).
3. J. H. Lui, D. V. Hansen, A. R. Kriegstein, Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).
4. M. Götz, W. B. Huttner, The cell biology of neurogenesis. *Nat. Rev. Mol. Cell Biol.* **6**, 777–788 (2005).
5. B. I. Bae, D. Jayaraman, C. A. Walsh, Genetic changes shaping the human brain. *Dev. Cell* **32**, 423–434 (2015).
6. T. J. Dixon-Salazar, J. G. Gleeson, Genetic regulation of human brain development: Lessons from mendelian diseases. *Ann. N. Y. Acad. Sci.* **1214**, 156–167 (2010).
7. J. C. Silbereis, S. Pochareddy, Y. Zhu, M. Li, N. Sestan, The cellular and molecular landscapes of the developing human central nervous system. *Neuron* **89**, 248–268 (2016).
8. A. Bernard *et al*., Transcriptional architecture of the primate neocortex. *Neuron* **73**, 1083–1099 (2012).
9. A. E. Ayoub *et al*., Transcriptional programs in transient embryonic zones of the cerebral cortex defined by high-resolution mRNA sequencing. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 14950–14955 (2011).
10. S. A. Fietz *et al*., Transcriptomes of germinal zones of human and mouse fetal neocortex suggest a role of extracellular matrix in progenitor self-renewal. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11836–11841 (2012).
11. S. Lodato, P. Arlotta, Generating neuronal diversity in the mammalian cerebral cortex. *Annu. Rev. Cell Dev. Biol.* **31**, 699–720 (2015).
12. A. Zeisel *et al*., Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
13. A. Saunders *et al*., Molecular diversity and specializations among the cells of the adult mouse brain. *Cell* **174**, 1015–1030.e16 (2018).
14. B. B. Lake *et al*., Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
15. R. D. Hodge *et al*., Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).
16. M. B. Johnson *et al*., Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015).
17. A. A. Pollen *et al*., Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nat. Biotechnol.* **32**, 1053–1058 (2014).
18. S. Zhong *et al*., A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* **555**, 524–528 (2018).
19. S. A. Yuzwa *et al*., Developmental emergence of adult neural stem cells as revealed by single-cell transcriptional profiling. *Cell Rep.* **21**, 3970–3986 (2017).
20. L. Loo *et al*., Single-cell transcriptomic analysis of mouse neocortical development. *Nat. Commun.* **10**, 134 (2019).
21. 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).
22. L. Telley *et al*., Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**, eaav2522 (2019).
23. M. Kohwi, C. Q. Doe, Temporal fate specification and neural progenitor competence during development. *Nat. Rev. Neurosci.* **14**, 823–838 (2013).
24. Q. Shen *et al*., The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat. Neurosci.* **9**, 743–751 (2006).
25. S. K. McConnell, The determination of neuronal fate in the cerebral cortex. *Trends Neurosci.* **12**, 342–349 (1989).
26. G. D. Frantz, S. K. McConnell, Restriction of late cerebral cortical progenitors to an upper-layer fate. *Neuron* **17**, 55–61 (1996).
27. P. Oberst *et al*., Temporal plasticity of apical progenitors in the developing mouse neocortex. *Nature* **573**, 370–374 (2019).
28. E. Z. Macosko *et al*., Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
29. 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).
30. V. Bergen, M. Lange, S. Peidli, F. A. Wolf, F. J. Theis, Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat. Biotechnol.* **38**, 1408–1414 (2020).
31. K. Street *et al*., Slingshot: Cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).
32. F. A. Wolf *et al*., PAGA: Graph abstraction reconciles clustering with trajectory inference through a topology preserving map of single cells. *Genome Biol.* **20**, 59 (2019).
33. T. Kowalczyk *et al*., Intermediate neuronal progenitors (basal progenitors) produce pyramidal-projection neurons for all layers of cerebral cortex. *Cereb. Cortex* **19**, 2439–2450 (2009).
34. C. Dehay, H. Kennedy, Cell-cycle control and cortical development. *Nat. Rev. Neurosci.* **8**, 438–450 (2007).
35. M. Lako *et al*., A novel mammalian wnt gene, WNT8B, shows brain-restricted expression in early development, with sharply delimited expression boundaries in the developing forebrain. *Hum. Mol. Genet.* **7**, 813–822 (1998).
36. J. Nishino, I. Kim, K. Chada, S. J. Morrison, Hmga2 promotes neural stem cell self-renewal in young but not old mice by reducing p16Ink4a and p19Arf Expression. *Cell* **135**, 227–239 (2008).
37. C. Boucherie *et al*., Neural progenitor fate decision defects, cortical hypoplasia and behavioral impairment in Celsr1-deficient mice. *Mol. Psychiatry* **23**, 723–734 (2018).
38. H. Li *et al*., Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry* **24**, 1235–1246 (2019).
39. G. Steele-Perkins *et al*., The transcription factor gene Nfib is essential for both lung maturation and brain development. *Mol. Cell. Biol.* **25**, 685–698 (2005).
40. C. E. Campbell *et al*., The transcription factor Nfix is essential for normal brain development. *BMC Dev. Biol.* **8**, 52 (2008).
41. M. H. Dominguez, A. E. Ayoub, P. Rakic, POU-III transcription factors (Brn1, Brn2, and Oct6) influence neurogenesis, molecular identity, and migratory destination of upper-layer cells of the cerebral cortex. *Cereb. Cortex* **23**, 2632–2643 (2013).
42. R. F. Hevner, Intermediate progenitors and Tbr2 in cortical development. *J. Anat.* **235**, 616–625 (2019).
43. D. J. Dennis, S. Han, C. Schuurmans, bHLH transcription factors in neural development, disease, and reprogramming. *Brain Res.* **1705**, 48–65 (2019).
44. A. Sessa *et al*., The Tbr2 molecular network controls cortical neuronal differentiation through complementary genetic and epigenetic pathways. *Cereb. Cortex* **27**, 3378–3396 (2017).
45. L. Baala *et al*., Homozygous silencing of T-box transcription factor EOMES leads to microcephaly with polymicrogyria and corpus callosum agenesis. *Nat. Genet.* **39**, 454–456 (2007).
46. S. J. Arnold *et al*., The T-box transcription factor Eomes/Tbr2 regulates neurogenesis in the cortical subventricular zone. *Genes Dev.* **22**, 2479–2484 (2008).
47. V. D. Blondel, J. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, 10008 (2008).
48. M. Jacomy, T. Venturini, S. Heymann, M. Bastian, ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PLoS One* **9**, e98679 (2014).
49. W. Saelens, R. Cannoodt, H. Todorov, Y. Saeys, A comparison of single-cell trajectory inference methods. *Nat. Biotechnol.* **37**, 547–554 (2019).
50. G. La Manno *et al*., RNA velocity of single cells. *Nature* **560**, 494–498 (2018).
51. B. Kang, Progenitor-cell-diversity. GitHub. https://github.com/kangbw702/Progenitor-cell-diversity. Deposited 15 February 2021.
52. X. Zhang, B. Kang, Progenitor cell diversity in the developing mouse neocortex. *Gene Expression Omnibus*. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1616905. Deposited 18 November 2020.