



Developmental programming and lineage branching of early human telencephalon

Lin Ma^{1,2,3,4,†}, Yanhua Du^{5,†}, Yi Hui^{1,2,3,†}, Nan Li^{1,2,3,†}, Beibei Fan^{1,2,3}, Xiaojie Zhang⁶, Xiaocui Li⁷, Wei Hong⁷, Zhiping Wu⁷, Shuwei Zhang^{1,2,3}, Shanshan Zhou^{1,2,3}, Xiangjie Xu^{1,2,3}, Zhongshu Zhou^{1,2,3}, Cizhong Jiang⁸ , Ling Liu^{1,2,3,4} & Xiaoqing Zhang^{1,2,3,8,9,10,*} 

Abstract

The dorsal and ventral human telencephalons contain different neuronal subtypes, including glutamatergic, GABAergic, and cholinergic neurons, and how these neurons are generated during early development is not well understood. Using scRNA-seq and stringent validations, we reveal here a developmental roadmap for human telencephalic neurons. Both dorsal and ventral telencephalic radial glial cells (RGs) differentiate into neurons via dividing intermediate progenitor cells (IPCs_{div}) and early postmitotic neuroblasts (eNBs). The transcription factor *ASCL1* plays a key role in promoting fate transition from RGs to IPCs_{div} in both regions. RGs from the regionalized neuroectoderm show heterogeneity, with restricted glutamatergic, GABAergic, and cholinergic differentiation potencies. During neurogenesis, IPCs_{div} gradually exit the cell cycle and branch into sister eNBs to generate distinct neuronal subtypes. Our findings highlight a general RGs-IPCs_{div}-eNBs developmental scheme for human telencephalic progenitors and support that the major neuronal fates of human telencephalon are predetermined during dorsoventral regionalization with neuronal diversity being further shaped during neurogenesis and neural circuit integration.

Keywords *ASCL1*; development; lineage branching; neurogenesis; telencephalon

Subject Categories Development; Neuroscience

DOI 10.15252/embj.2020107277 | Received 11 November 2020 | Revised 21 August 2021 | Accepted 24 August 2021 | Published online 24 September 2021

The EMBO Journal (2021) 40: e107277

Introduction

Generating a large while diversified neuronal pool and constructing well-organized excitatory or inhibitory neuronal networks are key features of the human telencephalon, which endows us with immense abilities for higher cognition, language, emotional control, decision-making, and tool use (Rakic, 2009; Hansen *et al*, 2010; Lui *et al*, 2011; Miller *et al*, 2014; Silbereis *et al*, 2016; Hodge *et al*, 2019). However, one unanswered question in developmental biology is how these abundant neuronal subtypes in the human brain are finely programmed and generated during early development.

When the anterior neural plate curves and ultimately meets along the dorsal midline to form the neural tube, the embryonic telencephalon is clearly demarcated into dorsal and ventral halves. The pallium, which gives rise to the cortex (CTX), and the subpallium, which develops into the basal ganglia, are two distinct regions endowed with different neurological functions. The dorsoventral (DV) regionalization of neuroectodermal cells within the neural plate is coordinated by antagonizing signaling molecules, such as BMPs and BMP antagonists, as well as WNTs and Sonic hedgehog (SHH), which emanate from the organizers located separately at the dorsal and ventral boundaries (Wilson & Rubenstein, 2000; Gunhaga *et al*, 2003; Li *et al*, 2009; Chi *et al*, 2016b). Consistent results across species suggest that these signaling molecules determine telencephalic DV regional identities by regulating master transcription factors (TFs), such as PAX6 for dorsal regionalization and NKX2.1 for ventral regionalization (Sussel *et al*, 1999; Gulacsi & Anderson, 2006; Zhang *et al*, 2010). Genetic studies have further revealed that these TFs reciprocally antagonize each other and thereby form a clear DV boundary (Chi *et al*, 2016b).

Regional neural progenitors proliferate and ultimately exit cell cycle to generate neurons. Cellular and molecular events that

1 Translational Medical Center for Stem Cell Therapy, Shanghai East Hospital, School of Medicine, Tongji University, Shanghai, China

2 Key Laboratory of Neuroregeneration of Shanghai Universities, School of Medicine, Tongji University, Shanghai, China

3 Shanghai Institute of Stem Cell Research and Clinical Translation, Shanghai, China

4 Department of Pathology and Pathophysiology, School of Medicine, Tongji University, Shanghai, China

5 Department of Immunology and Microbiology, Shanghai Institute of Immunology, Shanghai Jiao Tong University School of Medicine, Shanghai, China

6 Department of Obstetrics and Gynecology, Shanghai Baoshan Luodian Hospital, Shanghai, China

7 Shanghai First Maternity and Infant Hospital, School of Medicine, Tongji University, Shanghai, China

8 Key Laboratory of Spine and Spinal Cord Injury Repair and Regeneration of Ministry of Education, Orthopaedic Department of Tongji Hospital, Shanghai, China

9 Brain and Spinal Cord Innovative Research Center, School of Medicine, Tongji University, Shanghai, China

10 Tsingtao Advanced Research Institute, Tongji University, Qingdao, China

*Corresponding author: Tel: +86-21-65985003; E-mail: xqzhang@tongji.edu.cn

†These authors contributed equally to this work

control proliferation vs. differentiation are therefore key in setting the developmental programs that generate the desired neuronal population size (Chenn & Walsh, 2002; Fan *et al.*, 2018; Zhong *et al.*, 2018). In both dorsal and ventral embryonic telencephalons, there are two pools of amplifying progenitors, radial glial cells (RGs) located in the ventricular zone (VZ), and their downstream intermediate progenitor cells (IPCs) in the subventricular zone (SVZ) (Polen *et al.*, 2015; Nowakowski *et al.*, 2017). The continuous proliferation of RGs and IPCs results in the characteristic cortical gyrification and most possibly the unique cognitive abilities in humans (Johnson *et al.*, 2018). Indeed, human has a dramatically enlarged outer subventricular zone (oSVZ), which harbors expanded pool of cortical precursors from midcorticogenesis onward (Fietz *et al.*, 2010; Lui *et al.*, 2011; Johnson *et al.*, 2018). The patterned gyrification and the emergence of enlarged oSVZ of human telencephalon suggest that human RGs and IPCs might have adopted evolutionarily distinct molecular features, endowing them with advantageous properties in proliferation, differentiation, and fate determination (Lui *et al.*, 2011; Johnson *et al.*, 2018).

Cortex is subdivided into frontal lobe, parietal lobe, occipital lobe, and temporal lobe. Although different cortical lobes exhibit specialized functions and neuronal projections, they show common layering of excitatory glutamatergic (Glu) neurons organized into columnar microcircuits during orchestrated sequential neurogenesis and neuronal maturation (Zeng *et al.*, 2012; Lodato & Arlotta, 2015; Lake *et al.*, 2016; Fan *et al.*, 2018). In the ventral telencephalon, neural progenitors ultimately differentiate into GABAergic neurons and cholinergic neurons (ChAT), which either migrate or project to distal areas. Among these, parvalbumin (PV) and somatostatin (SST) GABAergic interneurons (INs) from the medial ganglionic eminence (MGE), and vasoactive intestinal peptide (VIP), cholecystokinin (CCK), and reelin (RELN) INs from the caudal ganglionic eminence (CGE) all migrate tangentially to populate the CTX, where they modulate the excitability of Glu neurons by local inhibition (Wamsley & Fishell, 2017; Mayer *et al.*, 2018; Mi *et al.*, 2018; Fishell & Kepecs, 2020). To date, the elaborate roadmap for the developmental trajectory and underlying contributing factors of these diversified excitatory and inhibitory neurons in human telencephalon have not been fully elucidated.

We have revealed that human telencephalon adopts evolutionarily distinct neural fate induction and DV regionalization programs, emphasizing the essentiality in exploring the conservative or unique aspects of human brain development (Li *et al.*, 2009; Zhang *et al.*,

2010; Chi *et al.*, 2016b). Indeed, pioneered single-cell RNA sequencing (scRNA-seq) studies start to elucidate neuronal compositions and temporal gene expression characteristics along with human brain development (Nowakowski *et al.*, 2017; Fan *et al.*, 2018, 2020; Zhong *et al.*, 2018). Here, we conducted scRNA-seq in human dorsal telencephalon, including prefrontal cortex (PFC) and occipital lobe (OL), as well as MGE in ventral telencephalon at early stages essential for neuronal lineage branching and neurogenesis followed by stringent validations. Our results reveal that both dorsal and ventral human telencephalons adopt similar cellular developmental programs for governing progenitor proliferation and neurogenesis. Our study also reveals that the major neuronal fates of CTX and MGE have already been predetermined during DV regionalization, and neuronal diversity is further broadened during neurogenesis, neuronal maturation, and neural circuit integration.

Results

Dorsal and ventral telencephalons harbor shared developmental programs at key points of neurogenesis

The large volume and complex features of human telencephalon originate from two key developmental events, that is, lineage branching and progenitor expansion vs. neurogenesis to generate variable neuronal subtypes (Fig 1A). To unravel these developmental processes, we dissected two CTX regions, including PFC and OL, from dorsal region, as well as the MGE from ventral region of human telencephalon at gestational week (GW) 12 and GW14 (Fig EV1A and B), key time points when progenitor proliferation and substantial neurogenesis occur (Laclef & Métin, 2018). To rule out batch effects, three biological replicates were performed for PFC and MGE at each time point (Fig EV1C and Appendix Table S1). The dissected PFC, OL, and MGE were then digested into single cells and subjected to scRNA-seq (Fig 1B). Quality screening revealed a total of 27,833 cells from the PFC^{GW12}, 25,276 cells from the MGE^{GW12}, 11,833 cells from the OL^{GW12}, 46,671 cells from the PFC^{GW14}, 24,561 cells from the MGE^{GW14}, and 14,176 cells from the OL^{GW14}, which generated strongly valid single-cell transcriptomic data after filtering the cells with low gene counts or low unique molecular identifiers (Fig EV1C and D).

The 124,341 single cells from MGE and PFC at both GW12 and GW14 were pooled, clustered based on the principal components analysis (PCA), and visualized using Uniform Manifold

Figure 1. Dorsal and ventral telencephalons share developmental programs at key points of neurogenesis.

- PAX6⁺ human neuroectoderm (NE) specifies into dorsal and ventral telencephalons during morphogen-triggered regional patterning. These regional progenitors expand, exit cell cycle, and differentiate into glutamatergic neurons, GABAergic neurons, and cholinergic neurons. Dissecting cellular and molecular mechanisms underlying progenitor expansion vs. neurogenesis and lineage branching remains a key question.
- Schematic illustrating the dissection of the prefrontal cortex (PFC) and medial ganglionic eminence (MGE) for single-cell RNA-seq.
- Visualization of MGE and PFC cells at GW12 and GW14 using Uniform Manifold Approximation and Projection (UMAP) plots. The forebrain gene *FOXP1* was uniformly expressed in both MGE and PFC cells. MGE progenitors, PFC progenitors, MGE neurons, and PFC neurons were distinguished by the combinatorial expression of marker genes, namely, *NKX2.1*, *PAX6*, *LHX6*, and *NEUROD6*, respectively. Direction of arrows represents the trend of differentiation.
- Both MGE (49,837) and PFC (74,504) cells were grouped into two categories, progenitors and neurons based on their cycling states. Details of other cell types were demonstrated in Fig EV1.
- Quantification of the percentage of progenitors and neurons in MGE and PFC at GW12 or GW14. Three biological replicates for MGE and PFC at GW12 or GW14. Central bands, median; boxes, interquartile range; whiskers, maximum/minimum.
- Heatmap displaying the differentially expressed genes (DEGs) enriched in progenitors or neurons in MGE and PFC. The enriched biological processes and the representative genes of each group are shown on the right. Fisher's exact test.

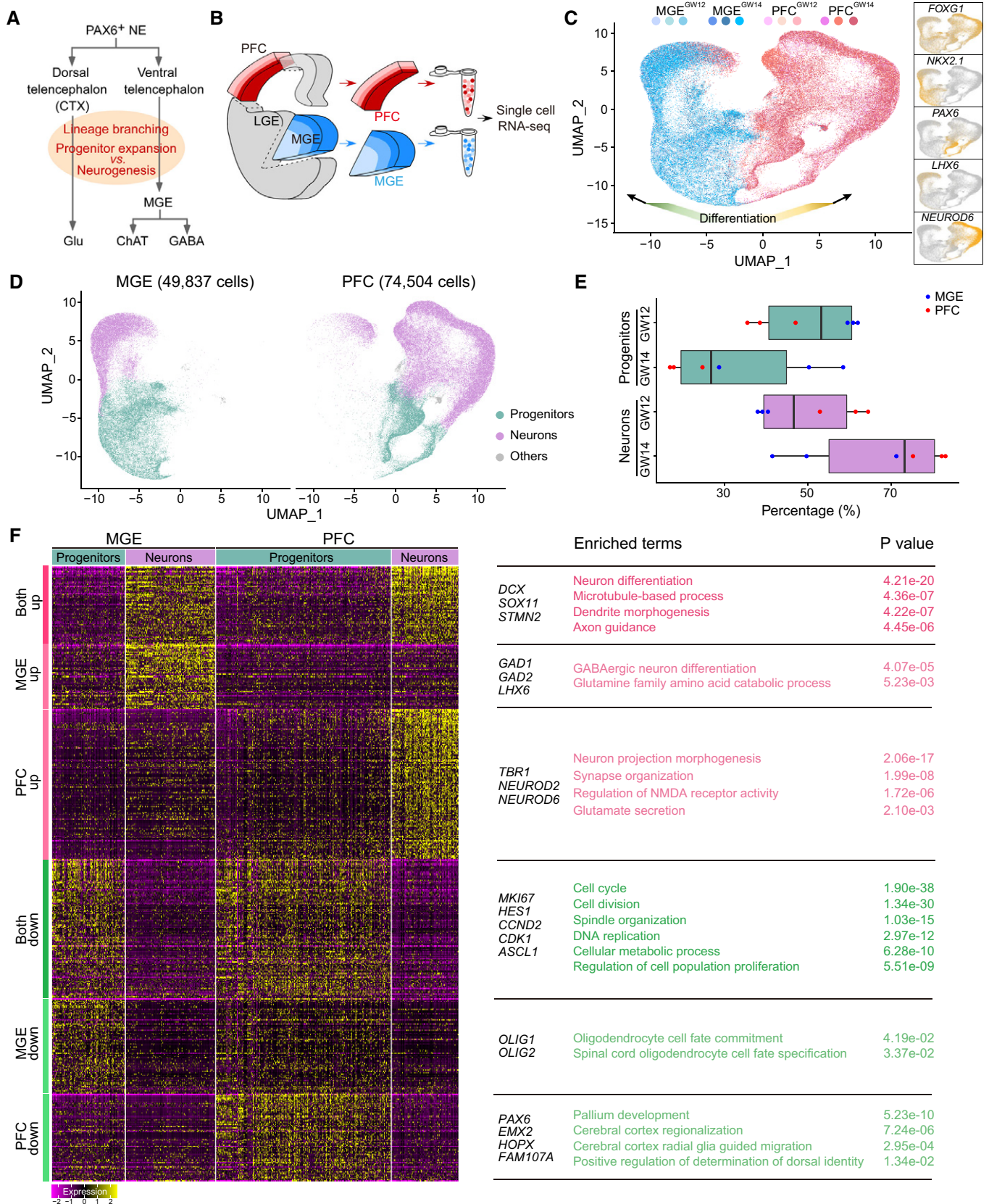


Figure 1.

Approximation and Projection (UMAP; Figs 1C and EV1E). The MGE cells were largely separated from those derived from the PFC, suggesting that their regional identities were fated once regionalized from the neuroectoderm (NE; Fig 1C). Three biological replicates showed consistent cell-type clustering, indicating minimal batch effects (Fig 1C). The forebrain gene *FOXP1* was uniformly expressed in both MGE and PFC cells (Fig 1C). DV separation was validated via dorsal vs. ventral regional identity genes. *NKX2.1* and *PAX6*, *FABP7*, *EMX2*, and *EOMES*, which are regionally expressed in MGE and PFC progenitors, respectively (Gulacsi & Anderson, 2006; Zhang et al., 2010), were expressed in corresponding populations (Figs 1C and EV1F). *LHX6*, *GAD1*, *GAD2*, *DLX2*, *DLX5*, and *NEUROD2* and *NEUROD6*, which are regulators expressed in late-stage progenitors and neurons, and are required for distinct neuronal differentiation (Bormuth et al., 2013; Kepecs & Fishell, 2014), were present in subpopulations of MGE and PFC, respectively (Figs 1C and EV1F). In contrast, other regional identity genes, such as *EN1* and *TAL2* for midbrain (Metzakopian et al., 2012), *VAX1* for hypothalamus (Hallonet et al., 1998; Wang et al., 2015), and *GSX2*, *HTR3A*, and *EBF1* for lateral/caudal ganglionic eminences (LGE/CGE) (Kepecs & Fishell, 2014; Mayer et al., 2018), were not detected (Fig EV1G).

Uniform Manifold Approximation and Projection plot revealed that both MGE and PFC populations were grouped into two major clusters (Fig 1D). The lower populations mainly expressed *NKX2.1* or *PAX6*, whereas the upper ones predominantly expressed *LHX6* and *NEUROD6* in MGE and PFC, respectively (Fig 1C). Meanwhile, the lower populations from MGE and PFC highly expressed marker genes related to G2/M or S phase, and the upper populations were enriched with G1 phase characteristic genes (Fig EV1H). The expression patterns of cell fate- and cell cycle-featured genes highlighted that the two major clusters in MGE and PFC represented either progenitor or neuronal fate, and a proliferation-to-differentiation transition in corresponding regions (Fig 1D). In addition, the proliferation-to-differentiation transition pattern was in line with the results showing that both MGE and PFC contained larger populations of neurons but smaller populations of progenitors at GW14 as compared to those at GW12 (Fig 1E).

At GW12 and GW14, microglia and oligodendrocyte progenitor cells (OPCs) started to populate both regions with more cells occurred in PFC at GW14 (Fig EV1E, I, and J). These results suggest that the developing telencephalon begins to construct its complex cell-type compositions when neural progenitors are undergoing massive expansion and neurogenesis in both dorsal and ventral regions.

To study the underlying cellular events during the transition from proliferative progenitors to postmitotic neurons of MGE and PFC, we used Seurat FindMarkers to identify the differentially expressed genes (DEGs) between progenitors and neurons. We conducted heatmap hierarchical clustering of the DEGs and highlighted their biological functions through Gene Ontology (GO) analysis (Fig 1F). In total, we identified 440 down-regulated genes and 339 up-regulated genes in neurons as compared to progenitors (Fig EV1K). 40% of the down-regulated DEGs (174/440) were shared in MGE and PFC (Figs 1F and EV1K). The common down-regulated DEGs were largely enriched in functional modules of cell cycle and cell division, whereas the unique down-regulated DEGs were mostly related to regional development (Fig 1F). Results also showed that 26% (89/339) of the up-regulated DEGs were shared by MGE and PFC, and were enriched in biological functions related to neuron differentiation, dendrite morphogenesis, and axon guidance (Figs 1F and EV1K). These data indicate that dorsal/ventral telencephalic progenitors might employ similar cell cycle exit and neurite outgrowth programs during progenitor-to-neuron transition, but, at the same time, preserve their own regional identities.

A total of 26,009 cells of OL at GW12 and GW14 were similarly clustered into progenitors, neurons, microglia, and OPCs (Fig EV1L). Similarity matrix represented high pairwise Pearson's correlations between PFC and OL in both progenitors and neurons (Fig EV1M), indicating PFC and OL have almost identical developmental processes. This was also supported by the similar expression pattern of featured genes in PFC and OL (Fig EV1N).

Gene expression trajectory identifies a common cell fate transition route applicable for both dorsal and ventral telencephalic progenitors

Progenitors from MGE (25,857) and CTX (pooled data from PFC and OL, 26,423 cells) were categorized into RGs and IPCs, and each of them could be further classified into 2 or 3 subclusters (Figs 2A and EV2A). Valid clustering of RGs and IPCs was confirmed by corresponding enrichment of RG- and IPC-featured genes for either MGE or CTX as revealed by another group (Nowakowski et al., 2017) (Fig EV2B). At GW14, the proportion of RGs decreased while that of IPCs increased in both MGE and CTX as compared to those at GW12 (Fig EV2C), suggesting progressive transition from RG to IPC.

Regional genes *NKX2.1* and *PAX6* were mutually exclusively expressed in MGE and CTX, respectively (Fig 2B). Notch signaling-

Figure 2. Gene expression trajectory of both dorsal and ventral telencephalic progenitors identifies a common cell fate transition route.

- A Progenitors in MGE (25,857) and CTX (26,423) were clustered into radial glia cells (RGs) and intermediate progenitor cells (IPCs). Dotted lines distinguish RG and IPC clusters.
- B Expression pattern of featured genes in clusters of MGE and CTX progenitors.
- C Scatter diagram showing enriched DEGs in RGs or IPCs of MGE and CTX. Red dots and blue dots represent common RG-enriched genes and IPC-enriched genes of both MGE and CTX.
- D Functional annotation of RG- and IPC-enriched genes in MGE and CTX.
- E, F Single-cell trajectories of RG and IPC cells in MGE (Panel E) and CTX (Panel F) by Monocle analysis.
- G Scatter plot showing enriched DEGs in dividing IPCs (IPCs_div) or early postmitotic neuroblasts (eNBs). Red dots and blue dots represent common IPC_div-enriched genes and eNB-enriched genes of both MGE and CTX.
- H Biological processes for DEGs enriched in either IPCs_div or eNBs of MGE and CTX.

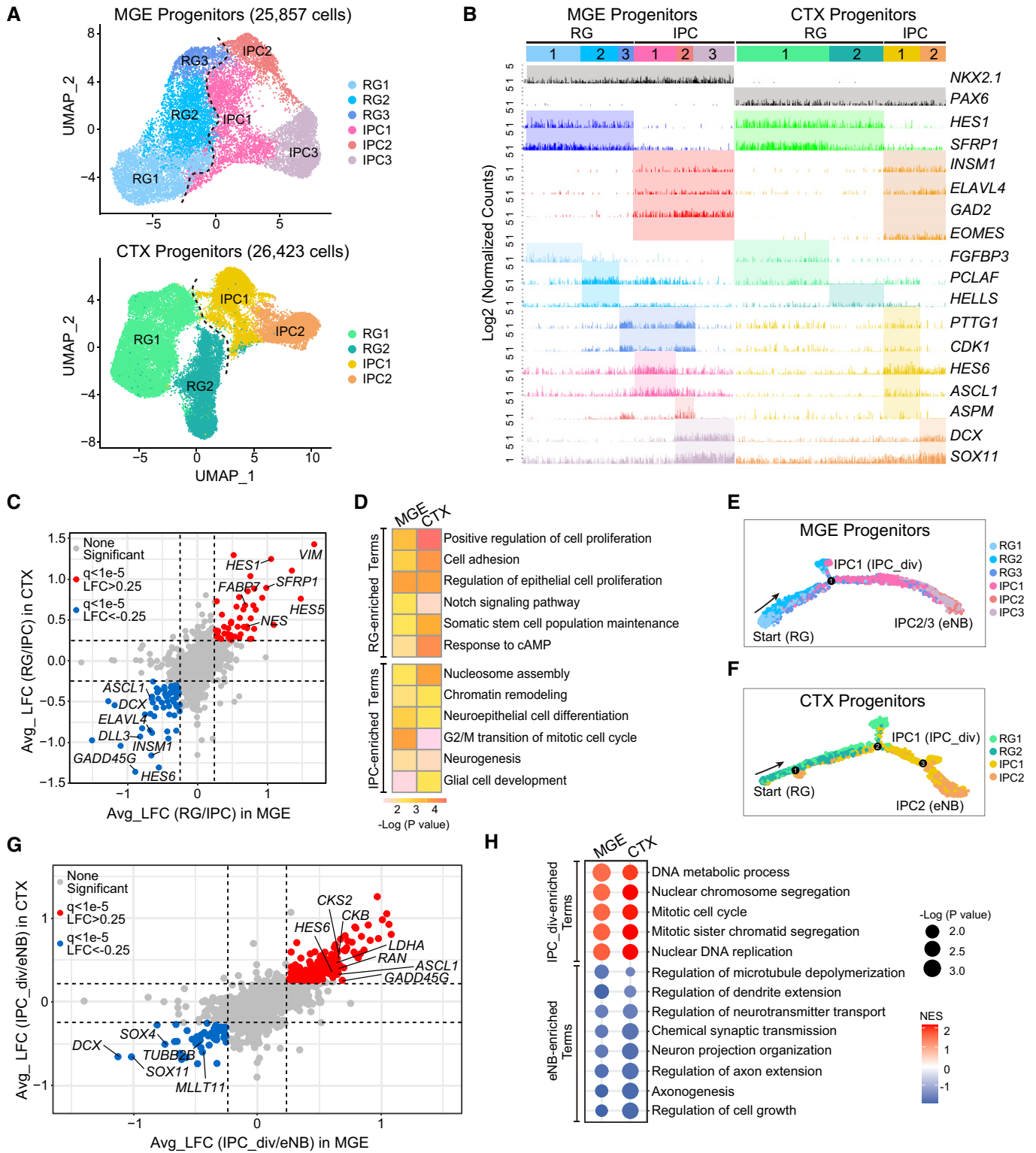


Figure 2.

regulated genes, such as *HES1* and *HES5*, were highly expressed in RGs of both MGE and CTX (Fig 2B and C), coinciding with key roles of Notch signaling pathway in the maintenance and expansion of human telencephalic progenitors (Hansen et al, 2010; Fiddes et al,

2018; Ma et al, 2019). MGE and CTX RGs also uniformly expressed *VIM*, an intermediate filament and scaffold protein that regulates epithelial mesenchymal transition and balances Notch signaling, and *SFRP1* (Fig 2B and C), a WNT signaling inhibitor (Chen et al,

2018a; Gopinathan *et al*, 2019). Both MGE and CTX IPCs showed uniform *INSM1* and *ELAVL4* expression (Fig 2B and C). *INSM1* is a transcription factor that is required for the delamination of RGs from the ventricular surface during mouse cortical development (Farkas *et al*, 2008). *ELAVL4* is an RNA binding protein, and its knockout causes anatomically disrupted Glu neurons with comorbid epileptic encephalopathy (EE)- and autism spectrum disorder (ASD)-like disease phenotypes (DeBoer *et al*, 2014). The shared expression pattern of *INSM1* and *ELAVL4* in CTX and MGE IPCs reinforced us to hypothesize that they might exert similar roles in regulating IPC development in MGE as well as they do in the CTX. As expected, *GAD2*, gene expressed in later-stage progenitors for INs (Kepecs & Fishell, 2014), and *EOMES*, IPC-featured gene for Glu neurons (Zhong *et al*, 2018), were specifically expressed in the IPCs of MGE and CTX, respectively (Fig 2B).

Differentially expressed genes between RGs and IPCs of both MGE and CTX were retrieved, with a large population of RG- and IPC-enriched DEGs conserved in MGE and CTX (Fig 2C), indicating similar programs underlying the transition from RGs to IPCs in both regions. GO analysis demonstrated that common RG-featured genes were enriched in transcription, cell proliferation, stem cell maintenance, and cell adhesion, whereas common IPC-featured genes were enriched in neuron migration, chromatin remodeling, and G2/M transition of mitotic cell cycle (Fig 2D). Together, these results suggest that both MGE and CTX have similar temporal gene expression in controlling the developmental trajectory from RGs to IPCs, although they are targeting toward distinct neuronal fates.

We also retrieved characteristic genes for each subcluster of RGs and IPCs of MGE and CTX (Figs 2B and EV2D). To study the interrelationships among these subclusters, Monocle analysis was then performed to examine their developmental processes (Fig 2E and F). In both regions, majority of subclusters were lined sequentially in pseudotime alignment, suggesting that these subclusters were grouped based largely on temporal aspects during development rather than on lineage identities. RG-featured cells in both MGE and CTX were detected at the start point, whereas IPC-featured cells appeared at the end (Fig 2E and F).

Remarkably, IPC-featured cells in both MGE and CTX were clearly demarcated into two pseudotime populations, including IPC1 and IPC2/3 in MGE, as well as IPC1 and IPC2 in CTX (Fig 2E and F). This pattern was further validated via cluster analysis using *hclust* (Fig EV2E). Furthermore, according to gene expression profiles, IPC1 from both MGE and CTX expressed hallmark genes essential for progenitor maintenance and proliferation, such as *RAN* and *CKS2* (Spruck *et al*, 2003; Zaoui *et al*, 2019) (Figs 2G and EV2D). In contrast, IPC2/3 from MGE, and IPC2 from CTX expressed early-born neuronal genes, such as *DCX* and *SOX11* (Wang *et al*, 2011; Kuwajima *et al*, 2017) (Figs 2G and EV2D). We therefore defined IPC1 in MGE and CTX as dividing IPCs (IPC_{s-div}), and IPC2/3 in MGE and IPC2 in CTX as early postmitotic neuroblasts (eNBs; Fig 2E–G). Gene Set Enrichment Analysis (GSEA) revealed that common DEGs between IPC_{s-div} and eNBs in MGE and CTX were prominent in biological processes of mitotic cell cycle for IPC_{s-div}-enriched DEGs and of neuron projection organization for eNB-enriched DEGs (Fig 2H). These results suggest that both dorsal and ventral telencephalons have conserved developmental trajectories of cell type progressing from

RGs to IPC_{s-div}, and then to eNBs. Among these, RGs and IPC_{s-div} are pivotal stages for progenitor expansion, while the transition from IPC_{s-div} to eNBs represents the initiation of cell cycle exit and neurogenesis.

IPC_{s-div} is a pivotal stage bridging progenitor proliferation and neurogenesis. *ASCL1*, *GADD45G*, *HES6*, *HMGB1*, *HMG2*, and *RNASEH2B* were specifically expressed in IPC_{s-div} of both MGE and CTX (Fig 3A). *HES6*, *HMGB1*, and *HMG2* have been reported to regulate proliferation of neuronal progenitor cells in telencephalon (Methot *et al*, 2013; Gao *et al*, 2020; Zhao *et al*, 2020). Gene regulatory network analysis of DEGs between RGs and IPCs or IPC_{s-div} and eNBs revealed that *ASCL1* is a hub gene in regulating *HES6* and a batch of eNB-featured genes (Fig 3B), highlighting its central role in controlling progenitor proliferation and differentiation. The involvement of *ASCL1* in generation of GABA INs from ventral telencephalic progenitors has been extensively studied in rodent (Casarosa *et al*, 1999; Peltopuro *et al*, 2010). Recent studies also showed that *ASCL1* was expressed in VZ and SVZ in fetal cortex of monkey and human being (Ma *et al*, 2013; Alzu'bi & Clowry, 2019), consistent with our current data (Fig 3A and C). To validate the expression pattern of *ASCL1* in the developing human brain, we performed RNAscope fluorescent *in situ* hybridization (Wang *et al*, 2012), and confirmed that *ASCL1* was widely expressed in the SVZ region of both MGE and CTX at GW13 (Fig 3D–F). The broad expression of *ASCL1* in IPC_{s-div} of human MGE and CTX drove us to hypothesize that it might exert important functional roles in both regions.

ASCL1 mediates the transition from RG to IPC in both MGE and CTX

To investigate the role of *ASCL1* in human telencephalon development, we mirrored CTX and MGE development via *in vitro*-targeted differentiation of human embryonic stem cells (hESCs) with our well-established protocol (Zhang *et al*, 2010; Chi *et al*, 2016a, 2016b; Chen *et al*, 2018b; Ma *et al*, 2019) (Fig 4A). The mutual exclusive expression pattern of *PAX6* and *NKX2.1* in the CTX and MGE progenitors indicated high patterning efficiency and pure dorsal and ventral progenitors obtained from hESCs (Fig 4B). Transient expression of *ASCL1* at day 25, a differentiation period equivalent to IPCs, was observed in both MGE and CTX (Fig EV3A), in line with its *in vivo* expression pattern.

To illustrate the role of *ASCL1*, we genetically engineered an *ASCL1*-knockout (KO) hESC line, with KO efficiency confirmed at both protein and mRNA levels (Fig 4C and D). The *ASCL1*-KO hESCs could be maintained *in vitro* and efficiently specified into neuroectoderm (NE) cells and then telencephalic neural progenitors (Fig EV3B and C). RNA-seq experiments showed that at day 25, both MGE and CTX progenitors after *ASCL1* KO showed higher expression of RG-featured genes, such as *HES1*, *VIM*, and *SFRP1*, but much lower expression of IPC/IPC_{s-div}/eNB-featured genes, such as *INSM1*, *ELAVL4*, *HES6*, *SOX11*, and *DCX* (Fig 4E). In addition, a large proportion of DEGs in the MGE and CTX progenitors after *ASCL1* KO were overlapped (Fig 4F). GO-term enrichment analysis revealed that common up-regulated genes in CTX and MGE cells after KO were clustered in biological functions of cell proliferation and cell adhesion, while common down-regulated genes in KO were clustered in neurotransmitter secretion and axon guidance (Fig 4G).

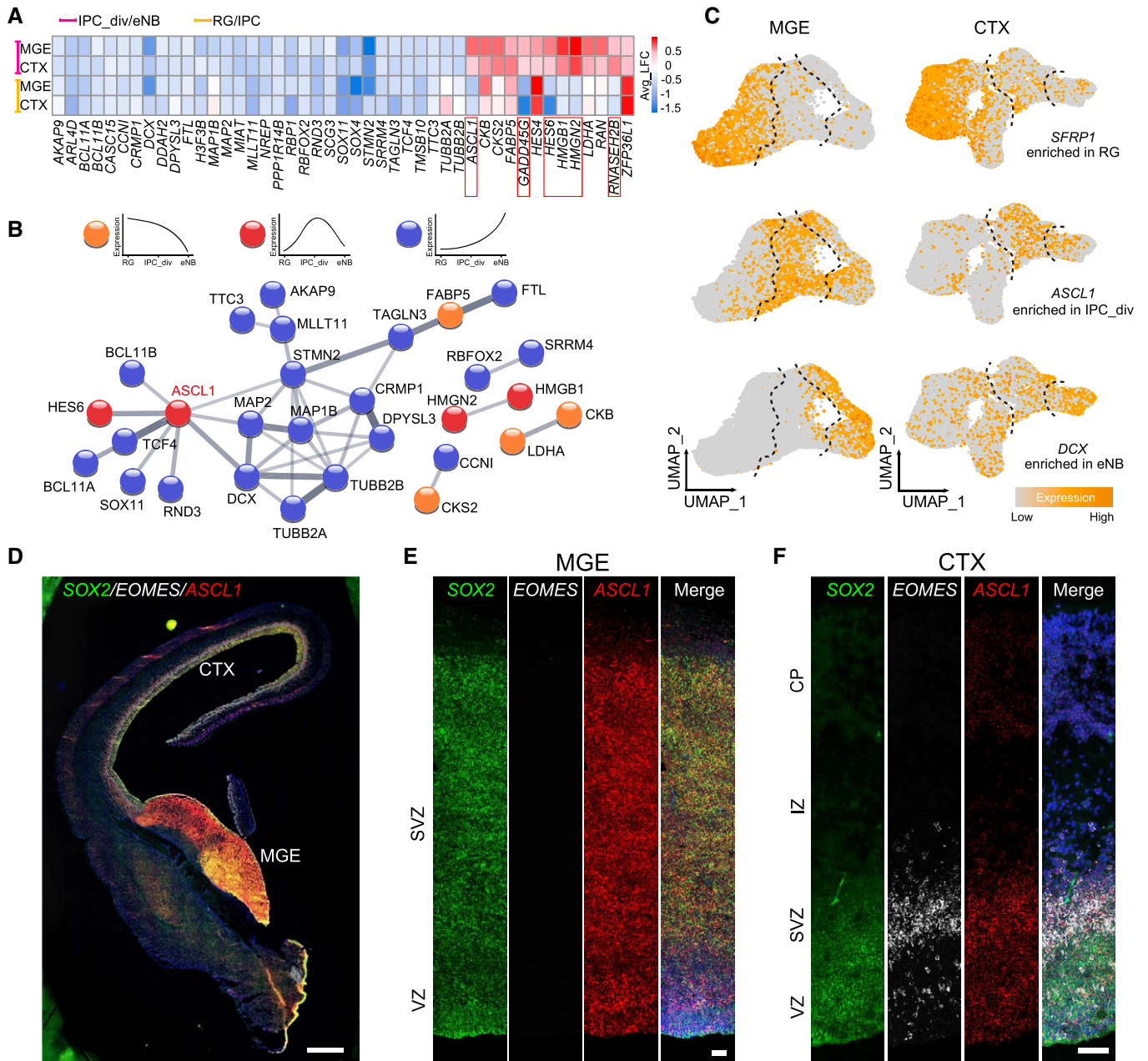


Figure 3. ASCL1 is the hub gene expressing in dividing IPCs.

- A Heatmap showing the featured DEGs between IPCs_div and eNBs, as well as RGs and IPCs in MGE and CTX. The red boxed *ASCL1*, *GADD45G*, *HES6*, *HMGB1*, *HMGN2* and *RNASEH2B* represent genes specifically enriched in IPCs_div of both MGE and CTX.
- B Protein–protein interaction (PPI) network analysis of DEGs revealed in panel A.
- C UMAP plots showing preferential expression of *SFRP1*, *ASCL1*, and *DCX* in RGs, IPCs_div, and eNBs, respectively, in both MGE and CTX.
- D RNAscope fluorescent *in situ* hybridization showing expression of *SOX2*, *EOMES*, and *ASCL1* mRNAs in coronal sections of the developing human brain at GW13. Scale bar, 1 mm.
- E, F Higher-magnification images showing *ASCL1* are predominantly expressed in the SVZ of both MGE (Panel E) and CTX (Panel F) at GW13. VZ, ventricular zone; SVZ, subventricular zone; IZ, intermediate zone; CP: cortical plate. Scale bars, 50 μ m.

Reintroducing of *ASCL1* with a doxycycline (Dox)-inducible overexpression system (rescue, Res) reversed the gene expression changes associated with *ASCL1* KO in both MGE and CTX progenitors with remarkable down-regulation of RG-feared genes and up-regulation

of IPC/IPC_div/eNB-feared genes (Figs 4H and EV3D–G). Together, these results reveal a common but pivotal role of *ASCL1* in both dorsal and ventral telencephalon developments in humans by mediating the fate transition from RGs to IPCs.

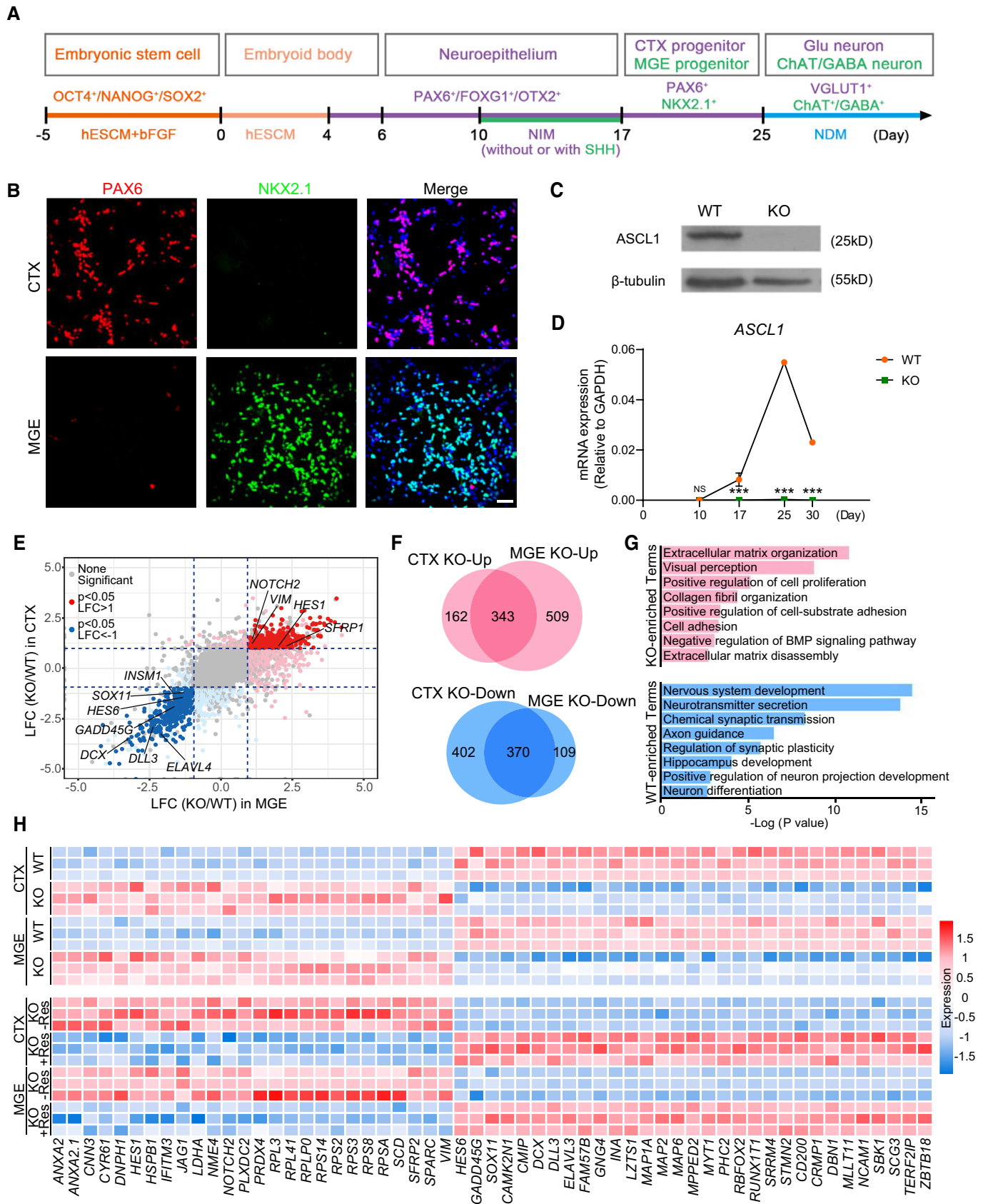


Figure 4.

Figure 4. ASCL1 mediates the transition from RG to IPC in both MGE and CTX.

- A Schematic representation of the procedure for generating MGE and CTX progenitors and neurons from human embryonic stem cells (hESCs). Without ectopic morphogens, NEs were targeted to a cortical fate from day 10 to 17. SHH was added to the NE from day 10 to 17 for ventralization. Cortical progenitors generated Glu neurons, whereas MGE progenitors produced ChAT and GABA neurons.
- B PAX6 and NKX2.1 immunostaining experiment showing CTX and MGE regional progenitors specified from hESCs with high efficiency. Scale bar, 50 μ m.
- C Western blot showing lack of ASCL1 protein expression in *ASCL1*-KO MGE progenitors differentiated from hESCs at day 25.
- D Quantitative PCR analyses showing lack of *ASCL1* mRNA expression in *ASCL1*-KO cells along with neural lineage differentiation of hESCs toward a CTX fate. Three biological replicates for wild type (WT) and KO. Data are presented as mean \pm SD. Unpaired t-test, NS, none significant; *** $P < 0.001$.
- E Scatter plot showing most of the DEGs between WT and *ASCL1*-KO progenitors at day 25 are shared in CTX and MGE. Red dots represent shared up-regulated genes, and blue dots represent shared down-regulated genes upon *ASCL1* KO.
- F Venn diagrams showing the number of genes related to panel E.
- G Functional annotations of KO- and WT-enriched DEGs showing a retarded RG to IPC transition upon *ASCL1* KO.
- H Heatmap showing the gene expression changes upon *ASCL1* KO are almost completely rescued by reintroducing of *ASCL1* (rescue, Res) at days 17–25 into hESCs differentiated toward either MGE or CTX fate.

Mapping resident and migrated neuronal populations in the developing human cortex

Neurons in PFC and OL at GW12 and GW14 were clustered into excitatory neurons expressing *SLC17A7*, *NEUROD2*, and *NEUROD6*; INs expressing *SLC32A1*, *GAD2*, *GAD1*, and *DLX5*; and Cajal–Retzius cells expressing *RELN*, *PCP4*, *LHX5*, *CALB2*, and *ZIC1* (Figs 5A and B, and EV4A). Excitatory neurons were further grouped into five subclusters, ExN1–5 (Fig 5A and B). Along with development from GW12 to GW14, the percentages of ExN2, ExN3, and ExN4 were slightly increased, whereas proportion of ExN5 was dramatically decreased (Fig 5C). These results suggest that the birth of ExN5 might process ahead of other excitatory neurons. Pseudotime analysis showed that ExN1 to ExN5 were nicely lined up (Fig 5D and E), supporting a more mature state of ExN5. GO analysis of the pseudogenes revealed that ExN1/2-enriched genes were largely clustered in functions of mRNA splicing, while ExN4/5-enriched genes were more related to synaptic transmission, supporting gradual neuronal maturation and primary synaptic construction during the processing from ExN1 to ExN5 (Fig 5E). Meanwhile, ExN5 showed more prominent expression of cortical-layer hallmark genes across layers 5–6 (Fig EV4B), coinciding with the inside-out sequential birth of cortical neurons. *GRIN2B* is a NMDA-type glutamate receptor subunit and was one of the pseudogenes enriched in ExN5 (Fig 5F). RNAscope fluorescent *in situ* hybridization confirmed high expression of *GRIN2B* in deep layers of cortical plate (CP) in CTX (Fig 5G). Pseudotime analysis also revealed that excitatory neurons from PFC and OL origins followed identical developmental trends (Fig 5H). Together, these findings suggest that cortical Glu excitatory neurons of different lobes gradually adopt their layer identities along with neuronal maturation and neural network integration (Fig 5I).

Interneurons in PFC and OL at GW12 and GW14 were categorized into two major clusters, which could be largely classified into SST and PV INs of a MGE origin, and VIP, CCK, and *RELN* INs of a CGE origin (Lake et al, 2018) (Fig 5A and B, EV4A and C–E). The percentage of SST INs within the entire IN population gradually decreased from GW12 to adult, while the percentage of PV INs plateaued at GW14 (Fig EV4F). In addition, the percentage of CGE-originated INs increased from GW12 to GW14 and plateaued in adult (Fig EV4F). These results suggest that the migration and maturation of SST INs precede those of PV INs, and MGE INs precede CGE INs to populate the CTX (Fig EV4E and F).

Generally, INs in CTX are born in the ventral telencephalon, after which they tangentially migrate and populate the CTX. At GW12 and GW14, majority of *GAD1*- and *SST*-expressing CTX INs were clustered with INs in MGE (Fig EV4G). Pseudotime analysis aligned INs in order from MGE origins to those located in CTX (Fig EV4H). Heatmap and GO analyses showed that INs remained in MGE highly expressed genes related to neuron fate specification, while those migrated to CTX expressed genes related to neuronal synaptic plasticity (Fig EV4H), indicating gradual neuronal maturation of INs along with long-distance migration from MGE to CTX.

Regional patterning and neurogenesis are two branching points for generating distinct neuronal subtypes of MGE

Uniform Manifold Approximation and Projection plot showed that 22,164 neurons extracted from MGE at GW12 and GW14 were clustered into immature neurons and neurons with clear subtype identities, that is, SST, PV INs and ChAT neurons (Fig 6A), and featured genes in these neuronal subtypes were largely consistent with a previous report (Lake et al, 2018) (Fig EV5A). SST and PV INs harbored differential expression of potassium channels and glutamate ionotropic receptors, with *GRIK1* and *GRIK2* highly expressed in SST INs and *KCNC2* and *GRIA4* preferentially expressed in PV INs (Fig 6B), highlighting distinct electrophysiological and synaptic properties in these neurons (Kudo et al, 2011; Frankel et al, 2014; Mennesson et al, 2019). We also found that ChAT neurons at the same developmental stage expressed higher amounts of progenitor genes, such as *LHX8*, *SHH*, and *NKX2.1* (Fig 6B), indicating a more immature state and a likely dependency on prolonged SHH signaling for their progenitor maintenance and subsequent neuronal differentiation.

To map the origins of each neuronal subtype within MGE, we performed unbiased PCA to probe the divergences among these closely related subpopulations. The SST and PV subpopulations could be clearly separated at the eNB stage, whereas the ChAT and GABA progenitors could be distinguished much earlier at the RG stage (Figs 6C and EV5B and C). Sequential projections of RGs/IPC_{div} to eNBs, immature neurons, and then PV, SST, and ChAT neuronal subtypes in UMAP plots with canonical correlation analysis (CCA) showed that RGs/IPC_{div} were already segregated into two major clusters, which mapped to PV/SST INs and ChAT neuronal fates, respectively, while PV and SST neuronal fates could only be separated in eNBs and onward (Figs 6D and EV5D–F).

These results suggest that ChAT and GABA lineage fates follow a “linear model” and are predetermined in RGs when MGE regional patterning is accomplished. In contrast, SST and PV lineages adopt

a “branched model” for identity determination, and their subtype fates are determined much later when neurogenesis is initiated at the eNB stage (Fig 6D). Heatmap analysis also identified DEGs

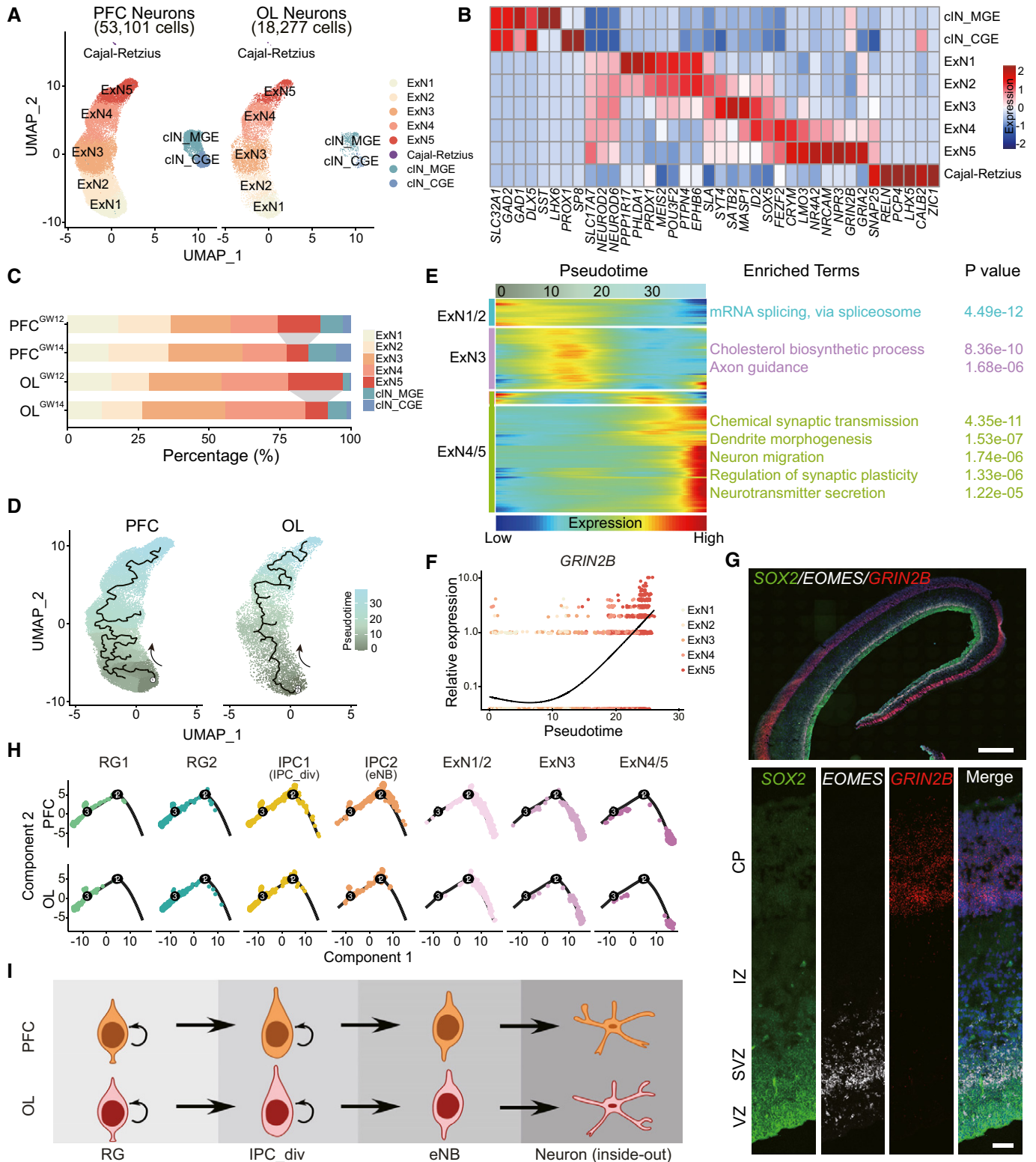


Figure 5.

Figure 5. Mapping developmental trajectories of diversified cortical neurons in the developing CTX.

- A Visualization of neurons in PFC and occipital lobe (OL), respectively. Cortical Glu neurons are classified into five subclusters, marked by ExN1–5. Cortical interneurons contain two subclusters derived from MGE or CGE. Cajal–Retzius cells are separated from other neuronal subtypes in both PFC and OL. ExN, excitatory neuron; cINs, cortical interneurons.
- B Heatmap showing featured DEGs in each subcluster related to panel A.
- C Cell percentage of each subcluster in PFC and OL at GW12 and GW14.
- D Single-cell trajectories of Glu neurons in PFC and OL by Monocle analysis.
- E Gene expression patterns of Glu neurons according to the pseudotime alignment. Enriched functional terms and *P* values for each gene group are listed on the right.
- F Expression of *GRIN2B*, a typical marker of ExN5, along with the pseudotime.
- G RNAscope fluorescent *in situ* hybridization showing expression of *SOX2*, *EOMES*, and *GRIN2B* mRNAs in coronal sections of the developing human brain at GW13. Scale bars, 1 mm for the upper panel and 50 μ m for the lower panel.
- H, I Pseudotime analysis by Monocle 2 (Panel H) and schematic diagram (Panel I) revealing Glu neurons of PFC and OL followed a common developmental route.

between PV and SST, as well as GABA and ChAT lineages at different developmental stages after their corresponding branching points (Figs 6E and EV5G–I).

Mapping of branch points of sister lineages is essential as these cues are useful for generation and enrichment of progenitors fated to become functional lineages, which is a major prerequisite for manufacturing precise cell sources for cell therapy (Ma *et al*, 2012; Liu *et al*, 2013; Yue *et al*, 2015). We searched surface hallmark genes distinguished in GABA and ChAT progenitors in RG and found that *ERBB4* was expressed in GABA but not in ChAT progenitors (Fig 6E). To validate, we dissected the progenitor cells lining the lateral ventricle at the MGE region at GW12–GW14 and applied fluorescence-activated cell sorting (FACS) to enrich *ERBB4*⁺ vs. *ERBB4*[−] progenitors (Fig 6F). Quantitative polymerase chain reaction (qPCR) analysis revealed that both *ERBB4*⁺ and *ERBB4*[−] progenitors expressed *NKX2.1*, reinforcing their MGE identity (Fig 6G). FACS-enriched *ERBB4*⁺ and *ERBB4*[−] progenitors were then plated to generate neuronal cultures. Immunocytochemistry experiments showed that *ERBB4*⁺ progenitors predominantly generated GABA INs, whereas *ERBB4*[−] ones were prone to produce ChAT neurons (Fig 6H and I). These results were in line with fate-mapping studies and suggest that *ERBB4*⁺/*NKX2.1*⁺ progenitors and *ERBB4*[−]/*NKX2.1*⁺ progenitors of MGE belong to two distinct subpopulations, and they are fated to become GABA INs and ChAT neurons at later developmental stages, respectively.

Discussion

Here, we present a conserved map of developmental trajectory of both dorsal and ventral human telencephalons for generating distinct neuronal subtypes (Fig 7). Generating variable human telencephalic neurons from corresponding regional progenitors follows temporal cell fate transition from NE to RGs to IPCs_{div}, and then to eNBs. *PAX6*-positive NE cells respond to patterning morphogens and regionalize into CTX or MGE progenitors, which are linearly destined to become Glu neurons or GABA/ChAT neurons. Among MGE progenitors, *ERBB4*⁺ and *ERBB4*[−] ones are largely predetermined to differentiate into GABA INs and ChAT neurons, respectively. RGs harbor featured genes responsible for cell adhesion, cell proliferation, and regulation of Notch signaling pathway, such as *VIM*, *HES1*, and *SFRP1*. *INSM1* and *ELAVL4* are expressed in both IPCs_{div} and eNBs, and they are essential for the delamination of

neural progenitor cells from the ventricular surface or the anatomical placement of neurons. *ASCL1* is highly enriched in IPCs_{div}, while *DCX* and *SOX11* mark eNBs effectively. RGs and IPCs_{div} remain in cell cycle, and their potent expansion might be responsible for the pattern of gyrification in humans. *DCX*-expressing eNBs are largely neuroblasts and are in the process of cell cycle exit and neurogenesis.

Our proposed human telencephalon developmental roadmap also reveals that after regional patterning of NE cells, both dorsal and ventral telencephalic RGs could also be classified into multiple subclusters. These subclusters are characterized with either pseudotime stages or lineage identities with variable potencies for generating different neuronal subtypes. We therefore propose that regional patterning is the primary and major branching point for generating distinct RGs, which are fated to major cortical and MGE neuronal identities. The secondary branching point occurs during the transition from cycling IPCs_{div} to postmitotic eNBs, and neuronal populations, such as MGE-originated PV and SST INs, are further diversified at this key neurogenesis stage. Meanwhile, during neuronal maturation, migration, and neural circuit integration, functional diversities of each specific neuronal subtypes, such as layered Glu neurons, will arise to form a delicate and balanced functional network.

ASCL1, a proneural gene, is uniformly expressed in IPCs_{div} in both CTX and MGE in humans. The involvement of *ASCL1* in generation of GABA INs from ventral telencephalic progenitors has been extensively studied in rodent (Casarosa *et al*, 1999; Peltopuro *et al*, 2010). Recent studies showed that *ASCL1* was expressed in VZ and SVZ in fetal cortex of monkey and human being, which co-expressed *PAX6* and *EOMES* (Ma *et al*, 2013; Alzu'bi & Clowry, 2019), consistent with our current study. In the current study, we demonstrate the uniform expression of *ASCL1* in the IPCs_{div} at SVZ region of both dorsal and ventral telencephalons at the single-cell level. More importantly, KO and rescue experiments show that *ASCL1* is essential for the transition from RG to IPCs in both dorsal and ventral telencephalic progenitors in humans. We have demonstrated that both the dorsal and ventral telencephalons are specified from common *PAX6*-expressing NE in humans (Zhang *et al*, 2010). It is therefore unsurprising that both dorsal and ventral telencephalons adopt similar developmental programs, including *ASCL1*-triggered cell fate transition from RGs to IPCs. It has been reported that mutations in the *CPAP* gene, which acts downstream of *ASCL1*, are responsible for primary autosomal microcephalies, characterized by severely reduced brain sizes (Garcez *et al*, 2015). Whether

Figure 6. Regional patterning and neurogenesis are two branching points for generating distinct neuronal subtypes in MGE.

- A UMAP plot showing 22,164 neurons extracted from MGE at GW12 and GW14 clustered into immature neurons and neurons with clear subtype identities, that is, SST INs, PV INs and ChAT neurons. PV, parvalbumin neurons; SST, somatostatin neurons; ChAT, cholinergic neurons.
- B Violin plots showing expression of specific marker genes in 1,939 PV INs, 2,224 SST INs, 713 ChAT neurons, and 17,288 immature neurons from 3 biological replicates of MGE at GW12/14.
- C Principal component analysis (PCA) showing the divergences of progenitors fated for SST, PV, and ChAT neurons in eNBs (upper panel) and RGs (lower panel).
- D Fate mapping of ChAT, SST, and PV lineages in developmental stages of neuron, eNB, IPC_div, and RG through canonical correlation analysis (CCA; left panel). Schematic diagram on the right panel illustrating two cell lineage branching models. In the “branched” model, common progenitors generate multiple lineages in eNBs during neurogenesis, while in the “linear” model, lineage fates have already been fixed in RGs once they are regionalized from NE.
- E Heatmap showing DEGs specific for GABA- or ChAT-biased progenitors in MGE RGs.
- F Schematic diagram showing the experimental procedures of fluorescence-activated cell sorting (FACS) of subpopulations of MGE progenitors for differentiation potency analyses.
- G Gene expressions of *ERBB4* and *NKX2.1* in *ERBB4*⁺ and *ERBB4*⁻ progenitor cells enriched with FACS. 3 isolated MGE samples at GW12-GW14. Data are presented as mean ± SD. Unpaired *t*-test, **P* < 0.05; NS, nonsignificant.
- H Immunocytochemistry experiments showing that *ERBB4*⁺ MGE progenitors mostly produce GABA neurons, while *ERBB4*⁻ MGE progenitors are prone to differentiate into ChAT neurons. Scale bars, 50 μm for the upper panel and 70 μm for the lower panel.
- I Quantification analyses of percentages of neuronal subtypes generated from *ERBB4*⁺ and *ERBB4*⁻ MGE progenitors. 18–72 cells in at least three independent experiments. Data are presented as mean ± SD. Unpaired *t*-test, ****P* < 0.001.

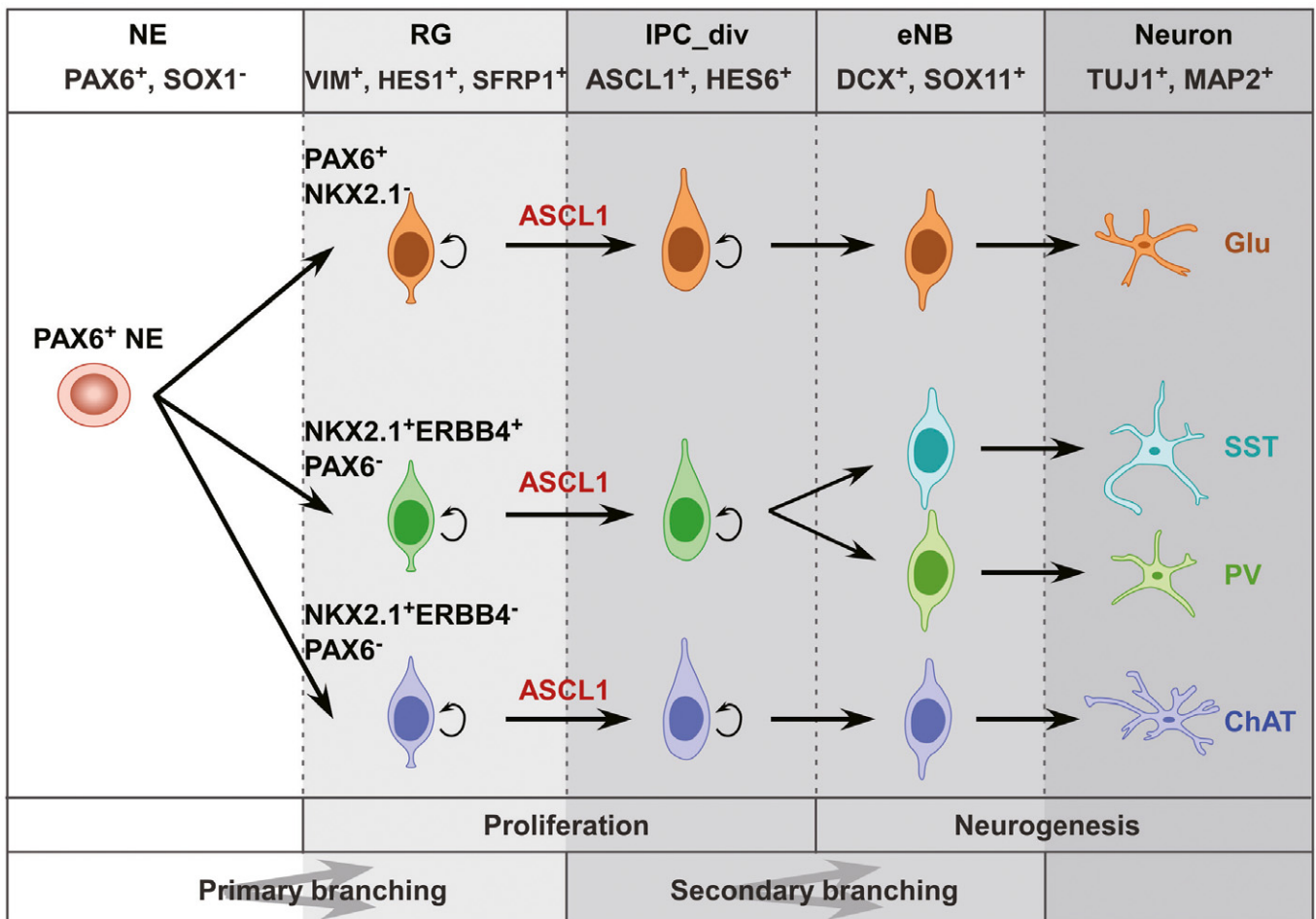


Figure 7. Developmental scheme for generating distinct neuronal subtypes in early human telencephalon.

In both dorsal and ventral telencephalons, the developmental trajectory of regional progenitors follows temporal cell fate transition from RGs to IPCs_div, then to eNBs and ultimately proceed to neurons with characteristic genes expressed at each stage. *ASCL1* is a major player in driving RGs to IPCs_div in both MGE and CTX. Regional patterning is the primary and major branching point for generating distinct RGs, which are fated to major cortical and MGE neuronal identities. The secondary branching point occurs during the transition from cycling IPCs_div to postmitotic eNBs, and neuronal populations, such as MGE-originated PV and SST INs, are further diversified at this key neurogenesis stage. Functional diversities of each specific neuronal subtypes, such as layered Glu neurons, are obtained during subsequent neuronal maturation, migration, and neural circuit integration.

ASCL1-governed transition of RGs to IPCs in human telencephalon development accounts for the unique features of the human brain either structurally or functionally, such as the formation of expanding oSVZ and the human pattern of gyrification, requires extensive studies.

As invaluable cell sources, hESCs and human-induced pluripotent stem cells (hiPSCs) can be efficiently differentiated into dorsal and ventral telencephalic progenitors and then corresponding neurons. In addition, these *in vitro*-specified regional progenitors hold great promise in replenishing neuronal loss and curing neurological disorders, including stroke, Alzheimer's disease, Huntington's disease, and epilepsy (Andres *et al*, 2011; Ma *et al*, 2012; Liu *et al*, 2013; Cunningham *et al*, 2014; Yue *et al*, 2015; Upadhyaya *et al*, 2019). Here, we provide a thorough developmental roadmap for early human telencephalic development and describe the molecular controls for progenitor proliferation and differentiation, as well as the branching points for generating diversified neuronal subtypes. These molecular switches guiding temporal progressing and lineage branching could guide the construction of sophisticated technologies for generating, expanding, and enrichment of specific telencephalic progenitors and therefore distinct neuronal subtypes for either disease studies or cell replacement therapy.

Materials and Methods

Ethics statement

Human embryonic brains from normal aborted fetuses were obtained from Shanghai First Maternity and Infant Hospital and Shanghai Baoshan Luodian Hospital with agreement of the donors and approval of related ethical review and informed consent documents. All the procedures were approved by the Ethics Committee of School of Medicine, Tongji University, and complied with the fundamental guidelines for the proper conduct of Interim Measures for the Administration of Human Genetic Resources and related activities in academic research institutions under the jurisdiction of the Chinese Ministry of Health. Detailed information of samples is listed in Appendix Table S1.

Tissue dissection and single-cell RNA-seq library preparation

The PFC, OL, and MGE from GW12 and GW14 embryonic brains were dissected according to the defined anatomical locations shown in Fig EV1A and B. All dissection procedures were carried out on ice. For making single-cell suspensions, 10× Genomics® Sample Preparation Demonstrated Protocol was followed with slight modifications. Briefly, dissected tissues were washed in chilled Hibernate E®/B27®/GlutaMAX™ (HEB) medium (BrainBits®) and then cut into small clumps using ophthalmic scissors. After removing HEB medium, 1 ml prewarmed papain solution (2 mg/ml; BrainBits®) was added into the tissue samples and incubated at 37°C for 15 min with repeated gentle swirls. After adequate washing with HEB medium, tissues were gently pipetted up and down for 15 times. The supernatant containing dispersed cells was then collected and centrifuged at 400 rcf for 3 min. Cell pellet was then resuspended in 1 ml cold 1× PBS without calcium nor magnesium and filtered through a 30-µm cell strainer (Miltenyi, 130-098-458). The cell

concentration and viability were then analyzed on Countess® II Automated Cell Counter, and all samples subjected to library construction and scRNA-seq showed cell viability larger than 85%.

Each sample was adjusted to 1,000 cells/µl, and library construction was run by using the Chromium Single Cell 3' Library and Gel Bead Kit v2 (120237, 10× Genomics). Briefly, single cells were isolated in emulsions using a microfluidic platform, and each single-cell emulsion was barcoded with a unique set of oligonucleotides. The GemCode Platform was used to carry out reverse transcription within each single-cell emulsion, which was amplified to construct a library. The manufacturer's protocol was used with a target capture of 10,000 cells for the 3' gene expression samples. Each sample was processed on an independent Chromium Single Cell A Chip (10× Genomics) and subsequently run on a thermocycler (Eppendorf). The libraries were then sequenced with 150-bp paired-end reads by using the Illumina HiSeq X Ten or NovaSeq 6000.

Processing of raw single-cell RNA-seq data

Raw sequencing reads were examined by quality metrics and analyzed using the Cell Ranger pipeline (10× Genomics). Briefly, the "mkfastq" command was used to generate FASTQ files and the "count" command was used to generate raw feature-barcode matrices aligned to human reference genome GRCh38. To ensure PCR-amplified transcripts were counted only once, only single UMIs were counted for gene expression level. To minimize batch effects, we used cellranger aggr pipeline to equalize the read depth between samples and recompute the matrices. In this way, the combined cell × UMI count matrixes of GW12 and GW14 of PFC, OL, and MGE were generated for downstream analyses.

Quality control

To ensure our analyses was based on high-quality cells, we implemented a series of quality control measurements. The cells of low quality that had UMIs < 1,000, unique detected genes < 500, and mitochondrial content > 5% were first filtered out. For the remaining cells, we filtered additional outliers. Considering the possibilities of duplicates, any cells that had three standard deviations above the mean for the total number of UMIs were removed, and the cells showed an unusually high or low number of detected genes given their number of UMIs by fitting a loess curve were also removed. We also filtered gene features, and any gene expressed by < 5 cells was removed. This resulted in a total of 150,350 cells, containing 27,833 cells from the PFC^{GW12}, 25,276 cells from the MGE^{GW12}, 11,833 cells from the OL^{GW12}, 46,671 cells from the PFC^{GW14}, 24,561 cells from the MGE^{GW14}, and 14,176 cells from the OL^{GW14}.

Single-cell RNA-seq data clustering and analysis

Clustering and analysis were first performed separately for PFC^{GW12}, OL^{GW12}, MGE^{GW12}, PFC^{GW14}, OL^{GW14}, and MGE^{GW14} datasets. We used R packages Seurat v3.1.2 (Butler *et al*, 2018; Stuart *et al*, 2019) to manage our data and created Seurat object. To reduce cell-to-cell variation due to technical factors including the number of molecules detected in each cell, we used scransform (SCT) algorithms (<https://github.com/ChristophH/scransform>) (Hafemeister & Satija, 2019) to perform the normalization and variance stabilization.

Briefly, the Pearson residuals from “regularized negative binomial regression” were proposed where cellular sequencing depth was utilized as a covariate in a generalized linear model, which successfully removed the influence of technical characteristics from downstream analyses while preserving biological heterogeneity. For detecting the highly variable genes (HVGs), other sources of variability such as percent of mitochondrial genes and cell cycle (CC) were unnecessary and regressed out. Here, CC scoring for each cell was made by CellCycleScoring function in Seurat, using a published list of CC genes (Tirosh *et al*, 2016). Then, the expression level of 3,000 HVGs in the cells was scaled and centered along each gene and was conducted to PCA. We then assessed the number of principal components (PCs) to be included in downstream analysis by (i) plotting the cumulative standard deviations accounted for each PC using the function ElbowPlot in Seurat to identify the “knee” point at a PC number after which successive PCs explain diminishing degrees of variance, and (ii) exploring primary sources of heterogeneity in the datasets using the PCHotmap function in Seurat. Resulting cells were reclustered and visualized in two-dimensional coordinates by UMAP. Cell clusters were annotated manually on the basis of known markers. For combined visualization, the datasets of PFC^{GW12}, OL^{GW12}, MGE^{GW12}, PFC^{GW14}, OL^{GW14} and MGE^{GW14} were integrated to a single object using FindIntegrationAnchors and IntegrateData function, and standard Seurat integrated analysis was performed. We used the Seurat subset function on the deep analysis of progenitor or neuron subsets. The subset command was used with the option “do.clean” set to “TRUE”. A new analysis was performed on the subsets using the SCTransform, RunPCA, and RunUMAP functions. New UMAP plots were generated for the subpopulations of progenitors or neurons.

Marker gene identification and cell-type annotation

Cell-type-specific gene expression signatures were identified by using Seurat function FindAllMarkers. In brief, we took one cluster of cells and compared with the rest clusters of cells, using a binomial model. For any given comparison, we only considered genes that met three filtering thresholds: a log₂ average expression difference > 0.25, at least 25% of cells that express the gene in either population, and at least 25 cells per group compared. Cell clusters were annotated using canonical markers of known cell types.

Identification of differentially expressed genes among cell clusters

We used Seurat’s FindMarkers function (Wilcoxon rank sum test) to detect differentially expressed genes between two groups of cells. The average log₂ fold change (LFC) of gene abundances was calculated. *P* values were adjusted based on Bonferroni correction. We selected adjust *P* value < 1E-5 and |LFC| > 0.25 as the threshold to judge the significance of differentially expressed genes.

Constructing single-cell trajectories

The Monocle (version 2.12.0) (Qiu *et al*, 2017) package was used to analyze single-cell trajectories in order to discover developmental transitions. We estimated size factors and dispersion using the default functions. We detected the differentially expressed genes by

performing differentialGeneTest function that uses a likelihood-ratio test to compare a vector generalized additive model (VGAM) using a negative binomial family function to a reduced model in which one parameter of interest has been removed. In practice, the model of “Seurat_clusters” defined previously by Seurat was fit. Genes were termed as significantly differentially expressed with a *q* value (Benjamini–Hochberg-corrected *P* value) < 0.01 and then were subjected to sort cells in pseudotime order. The actual developmental stage of each cell informed us of the start point of the pseudotime in the first round of “orderCells”. We then set this state as the root_-state argument and called “orderCells” again. DDRTree function was applied to reduce dimensions resulting in the minimum spanning tree on cells.

Classification of embryonic neurons

To classify embryonic neurons, we utilized a publicly available dataset of adult cortical INs (Lake *et al*, 2018). We first found all the shared HVGs in both embryonic and adult datasets. Then, we performed RF feature selection within the shared HVGs that best represented each cell type for all IN cell types defined by Zhang and colleagues (Lake *et al*, 2018), referred hereon as the adult cell-type features. For defining embryonic neurons by utilizing the adult cell-type features, we first conducted canonical correlation analysis (CCA) on embryonic and adult single-cell datasets using the adult cell-type features. Then, we performed UMAP plots to reduce the dimension of embryonic and adult cell data into the same two-dimensional space. Subsequently, we used the two UMAP coordinates for adult cells to conduct *k*-nearest neighbors analysis (knn) of the cell types and reassign cell identities for adult cells (*k* = 30) using the knn.cv function from R package FNN. Briefly, we calculated the average distance (*u*) and standard deviation (σ) of each cell with each of its 30 neighboring cells and removed any neighbor that was more than $u + \sigma$ away. Among the remaining neighbors, we counted the identities represented by the neighbors. A cell was assigned the identity represented by the majority, and at least 10, of its neighbors; in case of ties, the cell remains unassigned. Through this process, we were able to confidently reestablish the cells to 5 cell types for the adult dataset. The cells unable to assign were removed from the downstream analysis. Subsequently, we used the same knn approach on embryonic single cells, using adult cells as neighbors (*k* = 5) and assigned adult identities to embryonic cells. We also used this approach to assign the embryonic GABA INs into subtypes according to the adult dataset.

Mapping neuronal lineage branching points in embryonic MGE

The evolutionary relationship of PV, SST, and ChAT neurons of MGE origin in each developmental state of mitotic RG (IPC_div) and postmitotic eNB (neuron) was further analyzed in order to match their branching points. In brief, the Pearson correlations were calculated using R’s prcomp among each neuronal type and each cell of one state, and then, the values were subjected to the PCA with the first two dimensions displayed using R’s biplot. We found that PV and SST neurons were separated from eNBs, while GABA and ChAT neurons differed from RG cells. Next, we used genetic fate-mapping strategies in combination with 10× Genomics Chromium System for scRNA-seq to trace MGE neuronal fates at states of neuron, eNB, IPC_div, and RG one by one. In brief, we began with mapping PV,

SST, and ChAT to immature neurons. The mapping strategy was performed using the integration tool (Stuart *et al*, 2019) built in Seurat. The principle of integration was based the canonical correlation analysis, followed by L2 normalization of the canonical correlation vectors, to project the datasets into a subspace defined by shared correlation structure across datasets. In the shared space, we identified pairs of mutual nearest neighbors across PV, SST, ChAT, and immature neuron datasets. Cells that shared biological state across datasets were retained and served as “anchors” to guide dataset integration. Subsequently, we continued to track the fate of cells to eNB, IPC_div, and RG sequentially by using the same integration approach.

Functional annotation of genes

Metascape (<http://metascape.org>) integrates several authoritative data resources, such as Gene Ontology (GO), the Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, and Canonical Pathways, so it could execute pathway enrichment and biological process annotation to provide comprehensive and detailed information for defined clusters of genes (Zhou *et al*, 2019). We considered pathways and processes as statistically significant with $P < 0.01$, a minimum count of 3, and an enrichment factor > 1.5 .

Protein–protein interaction network construction

The online STRING database for retrieving interactive genes (Szklarczyk *et al*, 2019) provides protein–protein interaction information of genes. Only protein associations with combined confidence score above 0.5 were used to assign weights to each network link. The line shape indicated the predicted mode of action between proteins referring to <https://string-db.org/> (version 11.0).

Data visualization

Data visualization was mainly carried out through various functions in R package Seurat. All UMAP plots were generated by DimPlot function. FeaturePlot function was used to display a two-dimensional expression map for marker genes. DoHeatMap function helped generate heatmaps showing the cluster-specific genes in each cell with the color encoded the Z -scored expression level. Violin plots were performed through VlnPlot function. Dot plots of Fig EV4B were made by DotPlot function, in which the color intensity represented the average Z -scored expression levels and the dots size represented the fraction of expressed cells. For mapping single-cell trajectory, we used plot_cell_trajectory function in R package Monocle to generate a two-dimensional location of each cell. We also used function of plot_pseudotime_heatmap in Monocle to perform heatmaps showing the gene expression in each cell ordered in pseudotime. The dendrogram of Fig EV2E was generated using plot.as.phylo function in R package hclust. The scatter plot of differentially expressed genes and the gene function maps were generated using the R package ggplot2.

RNAscope fluorescent *in situ* hybridization

RNAscope fluorescent *in situ* hybridization was done using RNAscope Fluorescent Multiplex Kit V2 (323100, Advanced Cell

Diagnostics, Inc.) according to the manufacturer’s instructions. Probes were designed by and purchased from Advanced Cell Diagnostics (Newark, CA, USA) to label human brain tissues for mRNAs encoding SOX2 (400871-C3), EOMES (429691-C2), ASCL1 (459721), and GRIN2B (485851).

Embryonic brain samples were fixed with 4% PFA overnight and then sequentially cryoprotected with 20% and 30% sucrose for 24 h each. Coronal sections (10- μ m-thick) were cut on a cryostat and mounted onto glass slides. After baking at 60°C for 1 h followed by brief washing with PBS, brain sections were treated with hydrogen peroxide for 10 min at room temperature, and washed twice with double-distilled water. Then, the sections were treated by target repair reagent for 5 min at boiling water bath, washed once with double-distilled water and incubated for 3 min in 100% ethanol, and air-dried at room temperature. A hydrophobic barrier was created around the section using an ImmEdge™ hydrophobic barrier pen. After the barrier had completely dried, the sections were treated with protease III for 7.5 min at 40°C using an oven designed with HybEZ Hybridization System (Advanced Cell Diagnostics), washed twice with double-distilled water, and incubated with probe solutions for 2 h at 40°C. The sections were washed twice with wash buffer, followed by the incubation of fluorescent multiplex detection reagents to amplify the hybridization signals for probes on target RNAs. Three amplification reagents were used, AMP 1, AMP 2, and AMP 3, each at 40°C for 30, 30, and 15 min, respectively, with two washes of 2 min with wash buffer after each AMP treatment.

The signal was developed by treating the sections in sequence with the HRP reagent corresponding to each channel at 40°C for 15 min, followed by the TSA Plus fluorophore assigned to the probe channel at 40°C for 30 min, and HRP blocker at 40°C for 15 min, again with two washes after each of the incubation steps. In those steps, SOX2-C3 signal was incubated with TSA Plus FITC and detected with 488-nm laser, ASCL1 and GRIN2B were incubated with TSA Plus Cy3 and detected with 520-nm laser, and EOMES-C2 was incubated with TSA Plus Cy5 and detected with 590-nm laser. Following nucleus counterstaining with DAPI for 1 min at room temperature, sections were mounted with Fluoromount-G (Southern Biotech) and stored at 4°C until imaged. Stained sections were imaged by microscopy (SLIDEVIEW VS200, OLYMPUS) with a 20 \times objective and then laser confocal microscope (TCS SP8, Leica) with a 63 \times objective.

Human embryonic stem cell culture and neural differentiation

Both hESC culture and neural differentiation were performed as described previously (Zhang *et al*, 2010; Chi *et al*, 2016a, 2016b; Fang *et al*, 2017; Chen *et al*, 2018b; Ma *et al*, 2019). The hESCs (WA09, WiCell) were expanded every 5 days on a feeder layer. Neural differentiation was performed through embryoid body (EB) formation. Briefly, hESCs were broken into small clumps and suspended in hESC culture medium (hESCM, 392.5 ml of DMEM/F12, 100 ml of knockout serum replacer, 5 ml of MEM non-essential amino acid solution, 2.5 ml of 200 mM L-glutamine solution, and 3.5 μ l of 14.2 M β -mercaptoethanol) for 4 days to form EBs. The EBs were then switched to neural induction medium (NIM, 489 ml of DMEM/F12, 5 ml of N2 supplement, 5 ml of MEM non-essential amino acid solution, and 1 ml of 1 mg/ml heparin) to direct cells

toward a NE fate for 2 days. At day 6, cell aggregates were plated on laminin-coated culture surface and neural tube-like rosettes could be seen at days 14–17. For ventralization, a combination of SHH (250 ng/ml, R&D Systems) and the smoothed activator purmorphamine (0.3 μ M, Stemgent) was added to the NE cells at days 10–17 (Li *et al*, 2009; Zhang *et al*, 2010; Chi *et al*, 2016b). The neural progenitors were maintained in suspension culture as neurospheres in NIM at day 18 through day 25. At day 25, neurospheres were broken into single cells or small clusters and then plated on precoated coverslips or plates supplied with brain-derived neurotrophic factor (BDNF), glial cell line-derived neurotrophic factor (GDNF), and insulin-like growth factor (IGF) in neural differentiation medium (NDM, 485 ml of Neurobasal, 5 ml of N2 supplement, 10 ml of B27 supplement, 1 μ g/ml laminin, 0.1 μ M cAMP, and 200 μ g/l ascorbic acid).

Generation of transgenic cell lines

ASCL1-KO hESC line was constructed through CRISPR/Cas9-mediated homologous recombination (HR). The targeting donor vector comprised a left homology arm, followed by *Ef1 α* promoter-driven BSD resistance gene, and a right homology arm. The length of both arms was about 800 bp. The gRNA targeting sequence was GCGCCGGGCTAAAACAAAC. hESCs were pretreated with 1 mM Y27632 for at least 3 h, digested into single cells with trypsin, and electroporated with 5 μ g Cas9, 5 μ g gRNA targeting plasmid, and 40 μ g targeting donor plasmid. After electroporation, cells were plated on MEF feeders and cultured in hESCM supplied with Y27632. Medium was changed on the next day, and blasticidin (1 mg/ml) was added to remove the non-transduced cells. After 2 weeks, existing colonies with single-cell origin were subjected to genomic DNA PCR validation. For amplifying the 5' arm and 3' arm of *ASCL1*, the following primers were used: *ASCL1*-left-F: TGCGTCCGTCTCAGATCTCTGGGAGTATCCAGAGCACTC; *ASCL1*-left-R: GCTGACCGTCTCCGTTCGCAAACCTCCATTACAGTGGG; *ASCL1*-right-F: TGCGTCCGTCTCGGACAGCTGACTCTAAGCTTTCTTTGAC; *ASCL1*-right-R: GCTGACCGTCTCAAATTGAGAGGCCGAGTAGGATTCTAATC.

ASCL1-Res hESCs were established on the background of *ASCL1*-KO through TALEN-mediated HR in the *AAVS1* locus. The rescue (Res) donor plasmid was constructed by inserting *ASCL1* open reading frame downstream of the TET promoter followed by CAG-rtTA expression cassette in the backbone plasmid containing *AAVS1* HR arms and puromycin selection element (modified from Addgene #52344). *ASCL1*-KO hESCs were digested into single cells after Y27632 pretreatment for 3 h. Cells were then electroporated with plasmids of *AAVS1*-TALENs (5 μ g for both left and right) and *ASCL1*-Res donor plasmids (40 μ g). Puromycin (0.5 μ g/ml) was supplied for 3 days for positive selection. For amplifying the cDNAs of *ASCL1*, the following primers were used: *ASCL1*-Sall1-F: ATCGGTCGACGCCGTGCGCATGGAAAGCTCTGC; *ASCL1*-Mlu1-R: ATCGACGCGTTCAGAACCAGTTGGTGAAGTCG.

RNA-seq and data processing

RNA-seq and data processing were conducted according to our previous studies (Chi *et al*, 2016a; Liu *et al*, 2016, 2019; Fang *et al*, 2017; Ma *et al*, 2019). Briefly, total RNA was isolated and used to

prepare a library, which was then sequenced on an Illumina HiSeq 2500 platform.

Sequencing reads from each sample were mapped to the human reference genome (hg19 version) using HISAT2.1.0. The mapped reads were further analyzed using Cuffdiff v1.3.0, and the expression levels for each transcript were quantified as reads per kilobase of transcript per million mapped reads (RPKM).

For differential expression analysis, sequencing counts at the gene level were obtained using HTSeq v0.9.1. R package DESeq2 was then used to identify DEGs between different conditions. To assess the significance of differential gene expression, the *P* value threshold was set at 0.05 and the fold change was set at 2.

FACS sorting of ERBB4⁺ and ERBB4⁻ MGE progenitors for neuronal culture

Medial ganglionic eminence tissues were dissected from normally aborted human fetuses at GW12-GW14 and cut into small clumps, which were then digested with accutase into single cells and resuspended with blocking buffer (PBS containing 3% BSA) for 30 min at room temperature. Prepared single cells were further incubated with APC-conjugated ERBB4 antibody for 1 h. FACS analysis was performed with BD FACSAria II, and fluorescence was excited with a 640-nm laser. Both ERBB4⁺ and ERBB4⁻ MGE progenitors were plated onto the Matrigel-precoated coverslips to generate neurons in NDM with GDNF, IGF, and BDNF before harvesting for immunostaining.

Immunohistochemistry

Cells were fixed in 4% paraformaldehyde for 10 min at room temperature. Immunofluorescence labeling was performed by blocking the samples with 10% donkey serum in PBST (0.2% Triton X-100 in phosphate-buffered saline) for 1.5 h at room temperature, followed by incubation with primary antibodies overnight at 4°C. After washing, samples were incubated with fluorescently conjugated secondary antibodies (1:1,000; Jackson, West Grove, PA, USA) overnight at 4°C. Nuclei were counterstained with Hoechst 33258 for 10 min at room temperature. Stained samples were washed and cover-slipped with Fluoromount-G[®] (SouthernBiotech, USA). Images were collected using a Leica TCS SP8 confocal microscope. Detailed information of antibodies is listed in Appendix Table S2.

mRNA extraction and qPCR

Total RNA was isolated using a TRIzol kit (Invitrogen), and RNA concentration was determined by NanoDrop 2000c (Thermo Scientific). Total RNA (1 μ g) was reverse-transcribed into cDNA using SuperScript III (Invitrogen) cDNA Synthesis Kit with random primers. cDNA was then used as template for the qPCR (Bio-Rad, CFX Connect Real-Time System). qPCRs were performed in a mixture containing cDNA, primers, and 1 \times SYBR GREEN PCR Master mix. Expression levels of the mRNA were calculated through the comparative Ct method. The housekeeping gene glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) was amplified as an internal control for gene expression analysis. The primer oligonucleotides used for qPCR are listed in Appendix Table S3.

Western blotting

Cell pellets were collected and lysed in lysis buffer with protease inhibitor cocktail (Sigma). The particulate fraction was removed by centrifugation. Proteins (20 µg) were separated on 10% SDS-PAGE and subjected to immunoblotting analysis. Primary antibodies used in this study are listed in Appendix Table S2.

Statistical analyses

Data are presented as means ± SEM. Unpaired two-tailed Student's *t*-test was used. Statistical significance was considered at *P* value below 0.05: **P* < 0.05, ***P* < 0.01, and ****P* < 0.001.

Data availability

The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive (Genomics, Proteomics & Bioinformatics 2021) in National Genomics Data Center (Nucleic Acids Res 2021), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences, under accession number HRA000858 that are publicly accessible at <https://ngdc.cnbc.ac.cn/gsa-human>.

Expanded View for this article is available online.

Acknowledgements

This work was supported by grants from the National Key Research and Development Program of China (Grant no. 2018YFA0108000 and 2019YFA0110300), the National Natural Science Foundation of China (Grant No. 8205020, 32000689, 31400934, 31771132, 31872760, 31801204, and 31800858), the Science and Technology Commission of Shanghai Municipality (19JC1415100 and 21140902300), the Shanghai Municipal Education Commission (C120114), China Postdoctoral Science Foundation (Grant No. 2017M621526), the Fundamental Research Funds for the Central Universities, the Major Program of Development Fund for Shanghai Zhangjiang National Innovation Demonstration Zone (Stem Cell Strategic Biobank and Clinical Translation Platform of Stem Cell Technology, ZJ2018-ZD-004), and the Research Funds for Shanghai Baoshan Luodian Hospital (19-A-9).

Author contributions

XZ conceived and designed the project. LM, YH, and NL performed most of the experiments. YD performed the bioinformatics analyses. BF, ShuZ, ShaZ, XX, ZZ, and LL helped to set up the genetic engineering system and performed Western blot, immunostaining, ISH, and genomic DNA PCR experiments. XJZ, XL, WH, and ZW helped to obtain clinical samples. CJ helped with bioinformatics analyses. LM, YD, YH, NL, and XZ wrote the manuscript. All authors discussed the results and commented on the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Alzu'bi A, Clowry GJ (2019) Expression of ventral telencephalon transcription factors ASCL1 and DLX2 in the early fetal human cerebral cortex. *J Anat* 235: 555–568

- Andres RH, Horie N, Slikker W, Keren-Gill H, Zhan KE, Sun G, Manley NC, Pereira MP, Sheikh LA, McMillan EL *et al* (2011) Human neural stem cells enhance structural plasticity and axonal transport in the ischaemic brain. *Brain* 134: 1777–1789
- Bormuth I, Yan K, Yonemasu T, Gummert M, Zhang M, Wichert S, Grishina O, Pieper A, Zhang W, Goebbels S *et al* (2013) Neuronal basic helix–loop–helix proteins neurod2/6 regulate cortical commissure formation before midline interactions. *J Neurosci* 33: 641–651
- Butler A, Hoffman P, Smibert P, Papalexi E, Satija R (2018) Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 36: 411–420
- Casarosa S, Fode C, Guillemot F (1999) Mash1 regulates neurogenesis in the ventral telencephalon. *Development* 126: 525–534
- Chen M, Puschmann TB, Marasek P, Inagaki M, Pekna M, Wilhelmsson U, Pekny M (2018a) Increased neuronal differentiation of neural progenitor cells derived from phosphovimentin-deficient mice. *Mol Neurobiol* 55: 5478–5489
- Chen Z, Ren X, Xu X, Zhang X, Hui Yi, Liu Z, Shi L, Fang Y, Ma L, Liu Y *et al* (2018b) Genetic engineering of human embryonic stem cells for precise cell fate tracing during human lineage development. *Stem Cell Rep* 11: 1257–1271
- Chenn A, Walsh CA (2002) Regulation of cerebral cortical size by control of cell cycle exit in neural precursors. *Science* 297: 365–369
- Chi L, Fan B, Zhang K, Du Y, Liu Z, Fang Y, Chen Z, Ren X, Xu X, Jiang C *et al* (2016a) Targeted differentiation of regional ventral neuroprogenitors and related neuronal subtypes from human pluripotent stem cells. *Stem Cell Rep* 7: 941–954
- Chi L, Fan B, Feng D, Chen Z, Liu Z, Hui Yi, Xu X, Ma L, Fang Y, Zhang Q *et al* (2016b) The dorsoventral patterning of human forebrain follows an activation/transformation model. *Cereb Cortex* 27: 2941–2954
- Cunningham M, Cho J-H, Leung A, Savvidis G, Ahn S, Moon M, Lee P, Han J, Azimi N, Kim K-S *et al* (2014) hPSC-derived maturing GABAergic interneurons ameliorate seizures and abnormal behavior in epileptic mice. *Cell Stem Cell* 15: 559–573
- DeBoer EM, Azevedo R, Vega TA, Brodtkin J, Akamatsu W, Okano H, Wagner GC, Rasin MR (2014) Prenatal deletion of the RNA-binding protein HuD disrupts postnatal cortical circuit maturation and behavior. *J Neurosci* 34: 3674–3686
- Fan X, Dong JI, Zhong S, Wei Y, Wu Q, Yan L, Yong J, Sun LE, Wang X, Zhao Y *et al* (2018) Spatial transcriptomic survey of human embryonic cerebral cortex by single-cell RNA-seq analysis. *Cell Res* 28: 730–745
- Fan X, Fu Y, Zhou X, Sun LE, Yang M, Wang M, Chen R, Wu Q, Yong J, Dong JI *et al* (2020) Single-cell transcriptome analysis reveals cell lineage specification in temporal-spatial patterns in human cortical development. *Science Advances* 6: eaaz2978
- Fang Y, Liu Z, Chen Z, Xu X, Xiao M, Yu Y, Zhang Y, Zhang X, Du Y, Jiang C *et al* (2017) Smad5 acts as an intracellular pH messenger and maintains bioenergetic homeostasis. *Cell Res* 27: 1083–1099
- Farkas LM, Haffner C, Giger T, Khaitovich P, Nowick K, Birchmeier C, Pääbo S, Huttner WB (2008) Insulinoma-associated 1 has a panneurogenic role and promotes the generation and expansion of basal progenitors in the developing mouse neocortex. *Neuron* 60: 40–55
- Fiddes IT, Lodewijk GA, Mooring M, Bosworth CM, Ewing AD, Mantalas GL, Novak AM, van den Bout A, Bishara A, Rosenkrantz JL *et al* (2018) Human-specific NOTCH2NL genes affect notch signaling and cortical neurogenesis. *Cell* 173: 1356–1369.e1322
- Fietz SA, Kelava I, Vogt J, Wilsch-Bräuninger M, Stenzel D, Fish JL, Corbeil D, Riehn A, Distler W, Nitsch R *et al* (2010) OSVZ progenitors of human and

- ferret neocortex are epithelial-like and expand by integrin signaling. *Nat Neurosci* 13: 690–699
- Fishell G, Kepecs A (2020) Interneuron types as attractors and controllers. *Annu Rev Neurosci* 43: 1–30
- Frankel WN, Mahaffey CL, McGarr TC, Beyer BJ, Letts VA (2014) Unraveling genetic modifiers in the *gria4* mouse model of absence epilepsy. *PLoS Genet* 10: e1004454
- Gao XL, Tian WJ, Liu B, Wu J, Xie W, Shen Q (2020) High-mobility group nucleosomal binding domain 2 protects against microcephaly by maintaining global chromatin accessibility during corticogenesis. *J Biol Chem* 295: 468–480
- Garcez PP, Diaz-Alonso J, Crespo-Enriquez I, Castro D, Bell D, Guillemot F (2015) *Cenpj/CPAP* regulates progenitor divisions and neuronal migration in the cerebral cortex downstream of *Ascl1*. *Nat Commun* 6: 6474
- Gopinathan G, Foyle D, Luan X, Diekwisch TGH (2019) The Wnt antagonist *SFRP1*: a key regulator of periodontal mineral homeostasis. *Stem Cells Dev* 28: 1004–1014
- Gulacsi A, Anderson SA (2006) *Shh* Maintains *Nkx2.1* in the MGE by a *Gli3*-independent mechanism. *Cereb Cortex* 16: i89–i95
- Gunhaga L, Marklund M, Sjödal M, Hsieh J-C, Jessell TM, Edlund T (2003) Specification of dorsal telencephalic character by sequential Wnt and FGF signaling. *Nat Neurosci* 6: 701
- Hafemeister C, Satija R (2019) Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 20: 296
- Hallonet M, Hollemann T, Wehr R, Jenkins NA, Copeland NG, Pieler T, Gruss P (1998) *Vax1* is a novel homeobox-containing gene expressed in the developing anterior ventral forebrain. *Development* 125: 2599
- Hansen DV, Lui JH, Parker PRL, Kriegstein AR (2010) Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* 464: 554–561
- Hodge RD, Bakken TE, Miller JA, Smith KA, Barkan ER, Graybuck LT, Close JL, Long B, Johansen N, Penn O et al (2019) Conserved cell types with divergent features in human versus mouse cortex. *Nature* 573: 61–68
- Johnson MB, Sun X, Kodani A, Borges-Monroy R, Girsakis KM, Ryu SC, Wang PP, Patel K, Gonzalez DM, Woo YM et al (2018) *Aspm* knockout ferret reveals an evolutionary mechanism governing cerebral cortical size. *Nature* 556: 370–375
- Kepecs A, Fishell G (2014) Interneuron cell types are fit to function. *Nature* 505: 318–326
- Kudo T, Loh DH, Kuljis D, Constance C, Colwell CS (2011) Fast delayed rectifier potassium current: critical for input and output of the circadian system. *J Neurosci* 31: 2746–2755
- Kuwajima T, Soares CA, Sitko AA, Lefebvre V, Mason C (2017) *Sox*C transcription factors promote contralateral retinal ganglion cell differentiation and axon guidance in the mouse visual system. *Neuron* 93: 1110–1125.e1115
- Laclef C, Métin C (2018) Conserved rules in embryonic development of cortical interneurons. *Semin Cell Dev Biol* 76: 86–100
- Lake BB, Ai R, Kaeser GE, Salathia NS, Yung YC, Liu R, Wildberg A, Gao D, Fung H-L, Chen S et al (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352: 1586–1590
- Lake BB, Chen S, Sos BC, Fan J, Kaeser GE, Yung YC, Duong TE, Gao D, Chun J, Kharchenko PV et al (2018) Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat Biotechnol* 36: 70–80
- Li X-J, Zhang X, Johnson MA, Wang Z-B, Lavaute T, Zhang S-C (2009) Coordination of sonic hedgehog and Wnt signaling determines ventral and dorsal telencephalic neuron types from human embryonic stem cells. *Development (Cambridge, England)* 136: 4055–4063
- Liu Y, Weick JP, Liu H, Krencik R, Zhang X, Ma L, Zhou G-M, Ayala M, Zhang S-C (2013) Medial ganglionic eminence-like cells derived from human embryonic stem cells correct learning and memory deficits. *Nat Biotechnol* 31: 440–447
- Liu L, Liu XU, Ren X, Tian Y, Chen Z, Xu X, Du Y, Jiang C, Fang Y, Liu Z et al (2016) *Smad2* and *Smad3* have differential sensitivity in relaying TGF β signaling and inversely regulate early lineage specification. *Sci Rep* 6: 21602
- Liu L, Chen Z, Zhang X, Li S, Hui Y, Feng H, Du Y, Jin G, Zhou X, Zhang X (2019) Protection of ZIKV infection-induced neuropathy by abrogation of acute antiviral response in human neural progenitors. *Cell Death Differ* 26: 2607–2621
- Lodato S, Arlotta P (2015) Generating neuronal diversity in the mammalian cerebral cortex. *Annu Rev Cell Dev Biol* 31: 699–720
- Lui JH, Hansen DV, Kriegstein AR (2011) Development and evolution of the human neocortex. *Cell* 146: 18–36
- Ma L, Hu B, Liu Y, Vermilyea Scott C, Liu H, Gao L, Sun Y, Zhang X, Zhang S-C (2012) Human embryonic stem cell-derived GABA neurons correct locomotion deficits in quinolinic acid-lesioned mice. *Cell Stem Cell* 10: 455–464
- Ma T, Wang C, Wang L, Zhou X, Tian M, Zhang Q, Zhang Y, Li J, Liu Z, Cai Y et al (2013) Subcortical origins of human and monkey neocortical interneurons. *Nat Neurosci* 16: 1588–1597
- Ma L, Wang Y, Hui YI, Du Y, Chen Z, Feng H, Zhang S, Li N, Song J, Fang Y et al (2019) WNT/NOTCH pathway is essential for the maintenance and expansion of human MGE progenitors. *Stem Cell Rep* 12: 934–949
- Mayer C, Hafemeister C, Bandler RC, Machold R, Batista Brito R, Jaglin X, Allaway K, Butler A, Fishell G, Satija R (2018) Developmental diversification of cortical inhibitory interneurons. *Nature* 555: 457–462
- Mennesson M, Rydgren E, Lipina T, Sokolowska E, Kuleskaya N, Morello F, Ivakine E, Voikar V, Risbrough V, Partanen J et al (2019) Kainate receptor auxiliary subunit NETO2 is required for normal fear expression and extinction. *Neuropsychopharmacology* 44: 1855–1866
- Methot L, Hermann R, Tang Y, Lo R, Al-Jehani H, Jhas S, Svoboda D, Slack RS, Barker PA, Stifani S (2013) Interaction and antagonistic roles of NF- κ B and *Hes6* in the regulation of cortical neurogenesis. *Mol Cell Biol* 33: 2797–2808
- Metzakopian E, Lin W, Salmon-Divon M, Dvinge H, Andersson E, Ericson J, Perlmann T, Whitsett JA, Bertone P, Ang SL (2012) Genome-wide characterization of *Foxa2* targets reveals upregulation of floor plate genes and repression of ventrolateral genes in midbrain dopaminergic progenitors. *Development* 139: 2625–2634
- Mi DA, Li Z, Lim L, Li M, Moissidis M, Yang Y, Gao T, Hu TX, Pratt T, Price DJ et al (2018) Early emergence of cortical interneuron diversity in the mouse embryo. *Science (New York, NY)* 360: 81–85
- Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K et al (2014) Transcriptional landscape of the prenatal human brain. *Nature* 508: 199–206
- Nowakowski TJ, Bhaduri A, Pollen AA, Alvarado B, Mostajo-Radji MA, Di Lullo E, Haeussler M, Sandoval-Espinosa C, Liu SJ, Velmeshev D et al (2017) Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* 358: 1318–1323
- Peltopuro P, Kala K, Partanen J (2010) Distinct requirements for *Ascl1* in subpopulations of midbrain GABAergic neurons. *Dev Biol* 343: 63–70
- Pollen A, Nowakowski T, Chen J, Retallack H, Sandoval-Espinosa C, Nicholas C, Shuga J, Liu S, Oldham M, Diaz A et al (2015) Molecular identity of human outer radial glia during cortical development. *Cell* 163: 55–67

- Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C (2017) Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods* 14: 979–982
- Rakic P (2009) Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* 10: 724–735
- Silbereis C, Pochareddy S, Zhu Y, Li M, Sestan N (2016) The cellular and molecular landscapes of the developing human central nervous system. *Neuron* 89: 248–268
- Spruck CH, de Miguel MP, Smith APL, Ryan A, Stein P, Schultz RM, Lincoln AJ, Donovan PJ, Reed SI (2003) Requirement of Cks2 for the first metaphase/anaphase transition of mammalian meiosis. *Science* 300: 647–650
- Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM, Hao Y, Stoekius M, Smibert P, Satija R (2019) Comprehensive Integration of single-cell data. *Cell* 177: 1888–1902.e1821
- Sussel L, Marin O, Kimura S, Rubenstein JL (1999) Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. *Development* 126: 3359–3370
- Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P et al (2019) STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47: D607–D613
- Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G et al (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352: 189–196
- Upadhyay D, Hattiangady B, Castro OW, Shuai B, Kodali M, Attaluri S, Bates A, Dong YI, Zhang S-C, Prockop DJ et al (2019) Human induced pluripotent stem cell-derived MGE cell grafting after status epilepticus attenuates chronic epilepsy and comorbidities via synaptic integration. *Proc Natl Acad Sci USA* 116: 287–296
- Wamsley B, Fishell G (2017) Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nat Rev Neurosci* 18: 299–309
- Wang C, Liu F, Liu Y-Y, Zhao C-H, You Y, Wang L, Zhang J, Wei B, Ma T, Zhang Q et al (2011) Identification and characterization of neuroblasts in the subventricular zone and rostral migratory stream of the adult human brain. *Cell Res* 21: 1534–1550
- Wang F, Flanagan J, Su N, Wang LC, Bui S, Nielson A, Wu X, Vo HT, Ma XJ, Luo Y (2012) RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J Mol Diagn* 14: 22–29
- Wang L, Meece K, Williams DJ, Lo KA, Zimmer M, Heinrich G, Martin Carli J, Leduc CA, Sun L, Zeltser LM et al (2015) Differentiation of hypothalamic-like neurons from human pluripotent stem cells. *J Clin Invest* 125: 796–808
- Wilson SW, Rubenstein JLR (2000) Induction and dorsoventral patterning of the telencephalon. *Neuron* 28: 641–651
- Yue W, Li Y, Zhang T, Jiang M, Qian Y, Zhang M, Sheng N, Feng SU, Tang KE, Yu X et al (2015) ESC-derived basal forebrain cholinergic neurons ameliorate the cognitive symptoms associated with Alzheimer's disease in mouse models. *Stem Cell Rep* 5: 776–790
- Zaoui K, Boudhria Z, Khalifé P, Carmona E, Provencher D, Mes-Masson A-M (2019) Ran promotes membrane targeting and stabilization of RhoA to orchestrate ovarian cancer cell invasion. *Nat Commun* 10: 2666
- Zeng H, Shen E, Hohmann J, Oh S, Bernard A, Royall J, Glattfelder K, Sunkin S, Morris J, Guillozet-Bongaarts A et al (2012) Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* 149: 483–496
- Zhang X, Huang CT, Chen J, Pankratz MT, Xi J, Li J, Yang Y, LaVaute TM, Li X-J, Ayala M et al (2010) Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* 7: 90–100
- Zhao X, Rouhiainen A, Li Z, Guo S, Rauvala H (2020) Regulation of neurogenesis in mouse brain by HMGB1. *Cells* 9: 1714
- Zhong S, Zhang S, Fan X, Wu Q, Yan L, Dong JI, Zhang H, Li L, Sun LE, Pan NA et al (2018) A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex. *Nature* 555: 524–528
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK (2019) Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10: 1523