

A single-cell RNA-seq survey of the developmental landscape of the human prefrontal cortex

Suijuan Zhong^{1,2*}, Shu Zhang^{3*}, Xiaoying Fan^{3*}, Qian Wu^{1,2*}, Liying Yan^{3*}, Ji Dong³, Haofeng Zhang⁴, Long Li^{1,2}, Le Sun¹, Na Pan¹, Xiaohui Xu⁴, Fuchou Tang^{3,5,6}, Jun Zhang⁴, Jie Qiao^{3,5,6} & Xiaoqun Wang^{1,2,7}

The mammalian prefrontal cortex comprises a set of highly specialized brain areas containing billions of cells and serves as the centre of the highest-order cognitive functions, such as memory, cognitive ability, decision-making and social behaviour^{1,2}. Although neural circuits are formed in the late stages of human embryonic development and even after birth, diverse classes of functional cells are generated and migrate to the appropriate locations earlier in development. Dysfunction of the prefrontal cortex contributes to cognitive deficits and the majority of neurodevelopmental disorders; there is therefore a need for detailed knowledge of the development of the prefrontal cortex. However, it is still difficult to identify cell types in the developing human prefrontal cortex and to distinguish their developmental features. Here we analyse more than 2,300 single cells in the developing human prefrontal cortex from gestational weeks 8 to 26 using RNA sequencing. We identify 35 subtypes of cells in six main classes and trace the developmental trajectories of these cells. Detailed analysis of neural progenitor cells highlights new marker genes and unique developmental features of intermediate progenitor cells. We also map the timeline of neurogenesis of excitatory neurons in the prefrontal cortex and detect the presence of interneuron progenitors in early developing prefrontal cortex. Moreover, we reveal the intrinsic development-dependent signals that regulate neuron generation and circuit formation using single-cell transcriptomic data analysis. Our screening and characterization approach provides a blueprint for understanding the development of the human prefrontal cortex in the early and mid-gestational stages in order to systematically dissect the cellular basis and molecular regulation of prefrontal cortex function in humans.

The prefrontal cortex (PFC) covers the front of the frontal lobe of the brain in mammals and has important roles in memory, emotion, cognitive behaviour, decision-making and social behaviour^{1–4}. To analyse the molecular features of cells in the PFC during human brain development, we obtained 2,309 single cells from human embryonic PFCs at gestational weeks (GW) 8 to 26 (specifically, GW8, GW9, GW10, GW12, GW13, GW16, GW19, GW23 and GW26; three and two biological replicates at GW10 and GW23, respectively; Extended Data Fig. 1a, b). To classify the major cell types in the developing PFC, we performed *t*-distributed stochastic neighbour embedding (*t*-SNE) analysis using Seurat⁵ and identified six major clusters: neural progenitor cells (NPCs), excitatory neurons, interneurons, astrocytes, oligodendrocyte progenitor cells (OPCs) and microglia (Fig. 1a, Extended Data Figs 1c, d, 2a, b). Biological replicate samples were evenly distributed on the *t*-SNE plot (Extended Data Fig. 1c). To further analyse the subclusters within each cell type, we used random forest⁶ analysis to segregate cells into 35 distinct subtypes (Fig. 1b, Extended

Data Fig. 2c). By analysing differentially expressed genes among the clusters (Fig. 1b, Supplementary Table 1), we identified *SFRP1* and *RBFOX1* as markers of NPCs and excitatory neurons and verified this by immunofluorescence (Extended Data Fig. 2d). We found that NPCs, excitatory neurons and interneurons were sub-clustered in a development-dependent manner (Extended Data Fig. 3). We then reconstructed the developmental time course and lineage relationships using Monocle analysis^{7–9}. Microglia and interneurons were excluded because microglia are mesoderm-derived cells⁷ and interneurons are generated in the ganglionic eminence and migrate tangentially to the PFC^{10–12}. The remaining cells were distributed along pseudo-temporally ordered paths from NPCs to neurons (neuronal lineage) or to OPCs and astrocytes (glial lineage) (Fig. 1c, Extended Data Fig. 4a). Neurons developed from NPCs in early gestational weeks whereas OPCs and astrocytes differentiated from NPCs in later weeks (Fig. 1d, Extended Data Fig. 4b).

Using pseudo-time analysis, we found decreased SOX2 expression in the neuronal lineage but sustained high expression in the glial lineage (Fig. 1e, Extended Data Fig. 4). Next, we performed immunofluorescence staining of PAX6, SOX2 and NEUROD2 in embryos from GW8 to GW23. In the PFC from GW8 embryos, the majority of PAX6⁺SOX2⁺ cells were tightly localized in the ventricular zone. At GW12, the subventricular zone (SVZ) included a mixture of NPCs and a few neurons. As the PFC developed, the ventricular zone narrowed, whereas the NPC-enriched outer SVZ (oSVZ) expanded at GW16. At GW23, many of the SOX2⁺ cells in the oSVZ did not express PAX6 (Extended Data Fig. 5a, b), which suggests that the progenitor cells had developed into other lineages, such as oligodendrocytes and astrocytes, in the PFC, consistent with single-cell analysis showing that the majority of OPCs and astrocytes appeared in GW23 and GW26.

Progenitors of microglia arise from peripheral mesodermal tissue, but not from the neuroectoderm¹³. We observed that microglia appeared early in the PFC and were present throughout PFC development (Fig. 1a). Microglia were present outside the ventricular zone at GW8, when they had amoeboid or star-like morphology (Fig. 1f, Extended Data Fig. 5c). Microglia migrated to the SVZ and intermediate zone at GW12 and GW16, and then penetrated to the ventricular zone and cortical plate at GW19 and GW23 (Fig. 1f, g, Extended Data Fig. 5d). Microglia in the intermediate zone exhibited multi-directionally orientated processes, and those in the ventricular zone and SVZ had a simpler morphology at GW19 (Fig. 1f). The early appearance and stable population of microglia in the developing PFC, especially in the intermediate zone, support a model in which microglia are involved in regulating neuronal apoptosis, neurogenesis and synaptic pruning in development^{14–16}.

To reveal the diversity of NPCs in human PFC, we identified nine subclusters by gene ontology analysis of the differentially expressed

¹State Key Laboratory of Brain and Cognitive Science, CAS Center for Excellence in Brain Science and Intelligence Technology (Shanghai), Institute of Biophysics, Chinese Academy of Sciences, Beijing, 100101, China. ²University of Chinese Academy of Sciences, Beijing, 100049, China. ³Beijing Advanced Innovation Center for Genomics, College of Life Sciences, Department of Obstetrics and Gynecology, Third Hospital, Peking University, Beijing, 100871, China. ⁴Obstetrics and Gynecology, Medical Center of Severe Cardiovascular of Beijing Anzhen Hospital, Capital Medical University, Beijing, 100029, China. ⁵Biomedical Institute for Pioneering Investigation via Convergence and Center for Reproductive Medicine, Ministry of Education Key Laboratory of Cell Proliferation and Differentiation, Beijing, 100871, China. ⁶Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, 100871, China. ⁷Beijing Institute for Brain Disorders, Beijing, 100069, China.
*These authors contributed equally to this work.

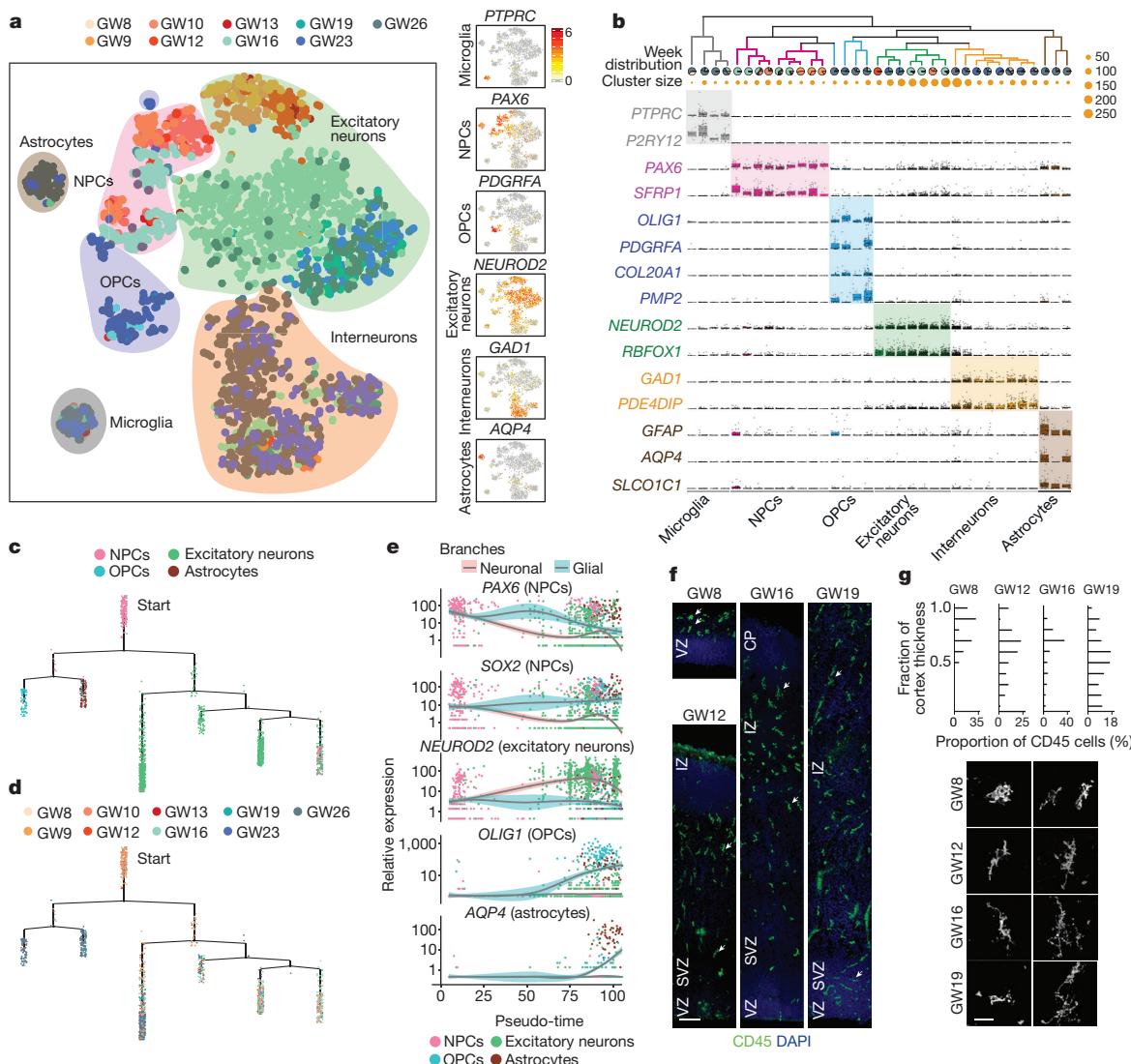


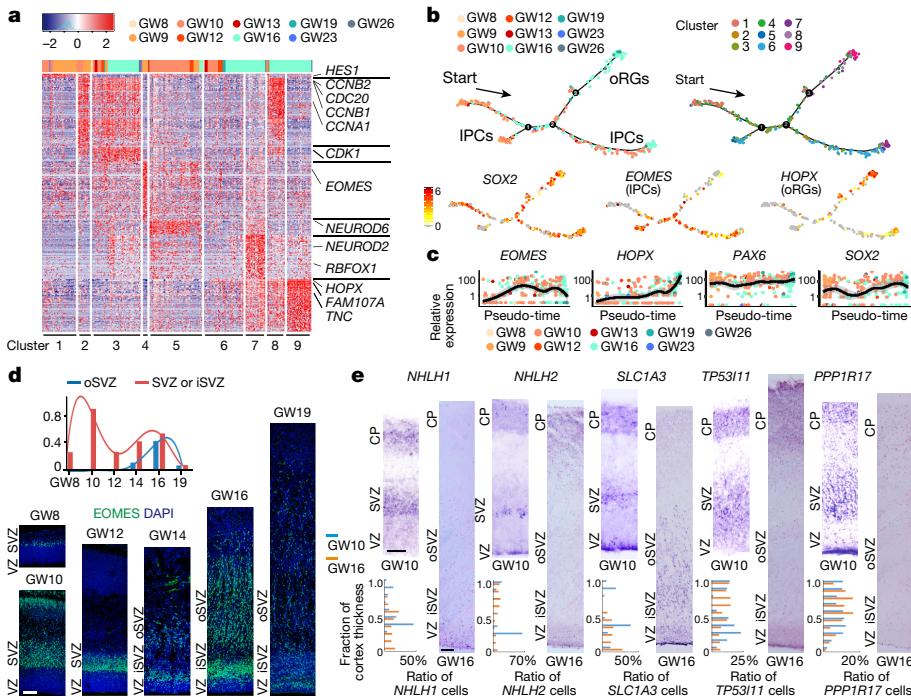
Figure 1 | Molecular diversity of cells from the developing human PFC. **a**, Visualization of major classes of cells using t-SNE. Dots, individual cells; colour, gestational weeks (GW); colour contours, cell types. Right, expression of known markers; grey, no expression; yellow-red, relative expression. **b**, Hierarchical clustering analysis of 35 subclasses. Microglia: $n = 5, 38, 10$ and 15 cells, left to right; NPCs: $n = 28, 23, 44, 60, 54, 21, 15, 38$ and 7 cells, left to right; OPCs: $n = 22, 28, 12$ and 55 cells, left to right; excitatory neurons: $n = 53, 103, 149, 163, 189, 142$ and 258 cells, left to right; interneurons: $n = 261, 120, 37, 64, 24, 63, 80$ and 52 cells, left to right; astrocytes: $n = 47, 12$ and 17 cells, left to right. Pie charts, distribution over gestational weeks; orange dots, subclass size. **c, d**, Single-cell trajectories by Monocle analysis showing development of the PFC.

genes (Fig. 2a, Extended Data Fig. 6a). *HOPX*, *FAM107A* and *TNC*, which have been identified as markers of outer radial glia (oRG)^{17,18}, were expressed in clusters 7, 8 and 9, while *EOMES* was expressed in clusters 3, 4 and 5, which suggests that the NPCs in these clusters were intermediate progenitor cells (IPCs) (Extended Data Fig. 6b). Using Monocle analysis, we found that the developmental trajectory of NPCs followed three major paths: early and late paths from ventricular radial glia (vRG) to IPC and a late path to oRG cells (Fig. 2b). To identify genes regulating radial glial cells in symmetric, neurogenic or gliogenic divisions¹⁹, we analysed the differentially expressed genes in *EOMES*⁺ NPCs from GW8 to GW10 and from GW16 to GW19 (Extended Data Fig. 6c). We found that *HMGAA2* was expressed in SOX2⁺ and Ki-67⁺ cells in the ventricular zone and SVZ of GW10 PFC but not in NPCs from GW16 PFC (Extended Data Fig. 6d). In addition, several differentially expressed genes in *EOMES*⁺ NPCs

The distance from a cell to the root corresponds to pseudo-time. Branched trajectories are plotted as a two-dimensional tree layout. **e**, Expression of known markers with pseudo-time. The lines with pink and blue shadows represent the neuronal and glial branches, respectively. The shadow represents the confidence interval around the fitted curve. $n = 1,540$ cells. **f**, The distribution and morphology of microglia in the PFC. Scale bars, 100 μm (left) and 20 μm (right). **g**, Quantification of the distribution of microglia in **f**. CD45⁺ cells were distributed across ten evenly spaced bins from the ventricular zone to the cortical plate. The experiment was repeated three times independently with similar results. See Extended Data Figs 1–5.

from GW16 to GW19 were oRG markers, and some were known neuronal markers (Extended Data Fig. 6c). Therefore, we co-stained GW16 PFC for *HOPX*, *SOX2* and *NEUROD1*, and observed that some *HOPX*⁺*SOX2*⁺ cells also expressed *NEUROD1* in oSVZ (Extended Data Fig. 6d), indicating that these progenitors might undergo neurogenic cell division and the resulting neurons may retain markers from the progenitor cells. Consistent with this hypothesis, the gene ontology results indicated that the oRG cells in cluster 8 (GW9 to GW16) were actively proliferating, whereas the oRG cells in cluster 7 (GW16 to GW26) were undergoing neuronal differentiation (Extended Data Fig. 6a).

IPCs act as transient amplifying progenitor cells, enabling the rapid generation of neurons during cortical development^{20–23}. To illustrate IPC development, we assessed the relative expression of *EOMES* in pseudo-time (Fig. 2b, c) and the proportion of *EOMES*⁺ cells at different gestational times, and found that they peaked during GW10



and GW16 (Extended Data Fig. 6e). To validate these results, we immunostained cortices of the PFC from GW8 to GW19 (Fig. 2d). We observed early increases in the number of IPCs at GW8, just above the thick ventricular zone. At GW10, the number of IPCs increased rapidly in the expanded SVZ. Thereafter, IPCs were not generated actively until GW16, when they were abundant in the iSVZ and oSVZ. Quantification corroborated the two peaks of IPC proliferation at GW10 and GW16, primarily located in the SVZ and oSVZ, respectively (Fig. 2d). These results suggest that accumulated IPCs at GW10 might originate from the vRG cells, whereas the majority of IPCs, especially those located in the oSVZ, might originate from the oRG cells that are present at GW16. To further characterize the differences between IPCs generated at early and late development stages, we compared differentially expressed genes in *EOMES*⁺ NPCs from GW10 and GW16. IPCs from GW10 and GW16 were enriched for genes involved in stem cell proliferation and neuronal differentiation, respectively (Extended Data Fig. 7a). In addition, we observed some *EOMES*⁺ cells that expressed early neuronal markers in an intermediate state between cell cycle-active NPCs and cell cycle-inactive *EOMES*⁻ neurons (Extended Data Fig. 7b). To identify markers for IPCs, we selected genes that were highly expressed in clusters 3, 4 and 5, including *NHLH1*, *NHLH2*, *TP53I11*, *SLC1A3* and *PPP1R17* (Extended Data Fig. 7c, d). We used RNA *in situ* hybridization (ISH) to validate these candidates. *NHLH1*- and *SLC1A3*-expressing cells were primarily localized in the SVZ at GW10, whereas they expanded from the ventricular zone into the oSVZ in GW16 (Fig. 2e). *NHLH2*, *TP53I11* and *PPP1R17*, in addition to expression in the SVZ at GW10 and in the oSVZ at GW16, also exhibited some cortical plate localization (Fig. 2e). These results demonstrate the heterogeneity of IPCs and identify new markers for IPCs that are expressed at different developmental time points.

The two main types of neurons in the brain are excitatory neurons and interneurons^{3,24,25}. Excitatory neurons could be further divided into seven clusters by random forest analysis (Extended Data Fig. 8a), and these were categorized into three groups according to their developmental stage (Fig. 3a–c, Extended Data Fig. 8b). Gene ontology terms enriched in each of the adjacent two periods showed that gene sets related to regulation of cell migration were more enriched at GW13 than at GW12, suggesting that GW13 is a critical time for migration of newly formed neurons. Neurons from GW16 PFC showed increased expression of genes for cell fate commitment; whereas neurons from

Figure 2 | Molecular signature of neural progenitor cells of the developing human PFC. **a**, Heat map showing gene expression in the NPC subclasses segregated using random forest classification. Top, the distribution of each subclass in gestational weeks. **b**, Cell lineage relationship of NPCs in the PFC. Monocle recovered a branched single-cell trajectory. **c**, Expression of NPC markers, ordered by Monocle analysis in pseudo-time. The shadow represents the confidence interval around the fitted curve. **d**, **e**, *In situ* hybridization of IPC-specific genes in the PFC at GW10 and GW16. The cells expressing IPC markers were distributed across ten evenly spaced bins going from the ventricular zone to the cortical plate (CP). The experiments in **c–e** were repeated three times independently with similar results. Scale bars, 100 μm (**d, e**). See Extended Data Figs. 6, 7.

GW19 PFC exhibited increased expression of functional genes, such as those involved in calcium import. Genes related to axonogenesis were expressed in late stage PFC (GW19 to GW26); this was followed by expression of genes for synapse formation from GW23 to GW26 (Fig. 3c, Extended Data Fig. 8c), suggesting that the initial formation of neural connections takes place between GW19 and GW26.

To confirm this hypothesis, we examined the electrophysiological properties of neurons in different layers of the PFC at GW23 and GW26. We analysed 53 neurons from three GW23 PFC samples. No action potentials were detected in these neurons (Extended Data Fig. 9a–c). Next, we examined 58 neurons from the cortical plate and subplate of three GW26 PFC biological samples. In the upper layer neurons (SATB2⁺FOXP2⁻), action potentials were only barely initiated, with a low level of outward potassium and inward sodium channel currents (Fig. 3d–f, Extended Data Fig. 9d). Deeper layer neurons (SATB2⁺FOXP2⁺) exhibited stronger action potentials and patterned firing (Fig. 3d–f, Extended Data Fig. 9d). There were no differences between whole-cell membrane Na⁺-K⁺ currents in upper- or deep-layer neurons, but deep-layer neurons exhibited more complex 3D morphology (Fig. 3d, g–i, Extended Data Fig. 9d–f). In addition, both spontaneous EPSCs (excitatory postsynaptic currents) and IPSCs (inhibitory postsynaptic currents) were detected in several deep-layer neurons (Fig. 3j), suggesting that these cells were beginning to receive excitatory and inhibitory presynaptic inputs. Subplate neurons also showed more mature electrophysiology than the cortical plate neurons (Fig. 3d–h, Extended Data Fig. 9d). Consistent with their mature electrophysiological properties, the subplate neurons exhibited more complex 3D dendrite morphology than neurons located in the upper or deep layers (Fig. 3d, i, Extended Data Fig. 9f).

Cortical interneurons originate from the ganglionic eminence and migrate tangentially to the neocortex, where excitatory neurons are generated^{10–12}. However, whether a subgroup of interneurons is generated locally in the PFC remains unclear. Random forest analysis clusters PFC interneurons into eight subgroups (Extended Data Fig. 10a). Groups of cells expressing interneuron progenitor markers, such as *TTF1*, *LHX6* and *DLX1*, were detected at an early timepoint in a pseudo-time alignment by Monocle analysis, and persisted throughout development (Fig. 4a–c). *CALB2*⁺ and *SST*⁺ interneurons appeared early in pseudo-time, and were followed by *CALB1*⁺, *CCK*⁺ and *VIP*⁺ interneurons (Fig. 4a–c, Extended Data Fig. 10b).

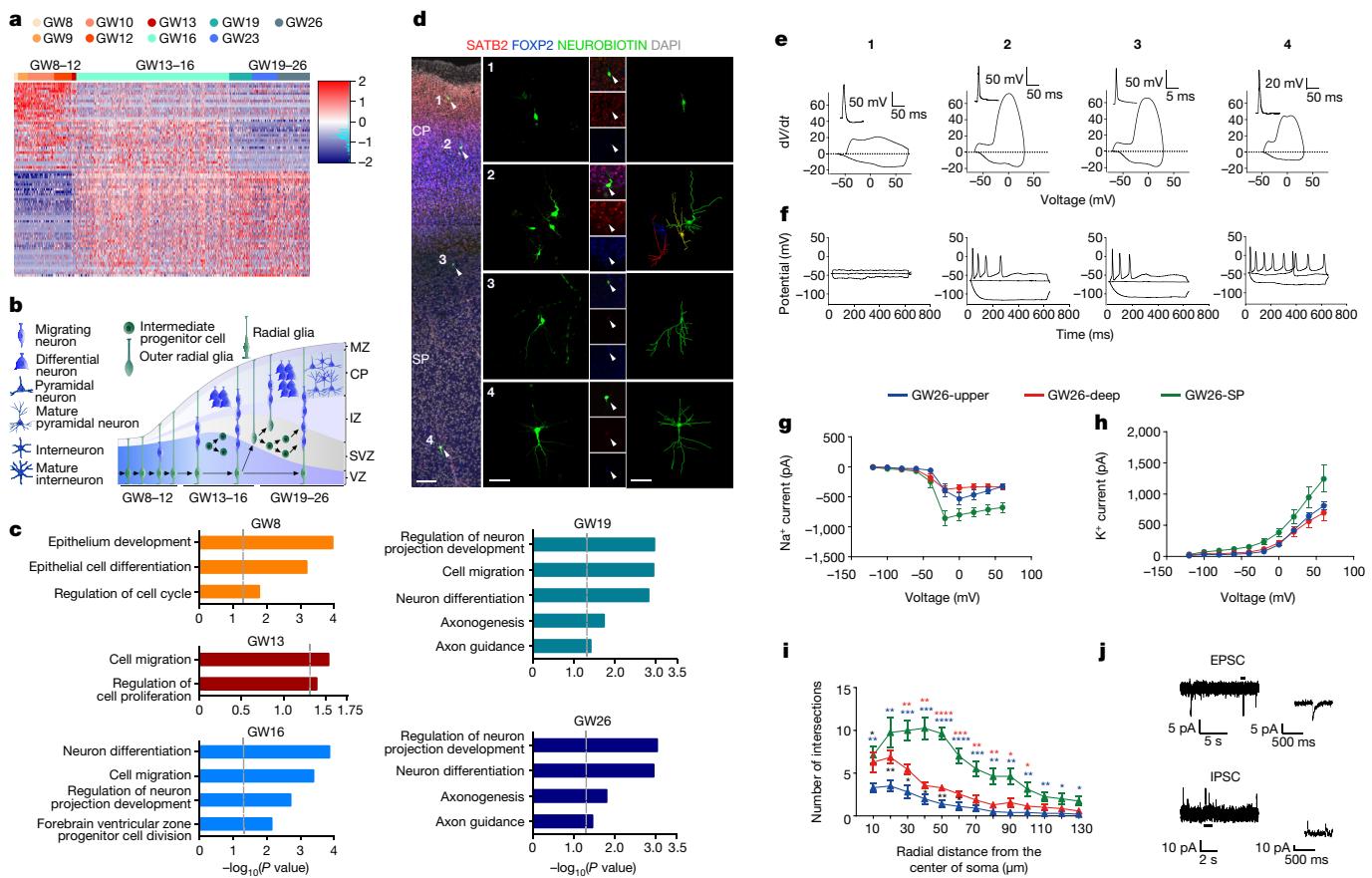


Figure 3 | Dynamics of neurogenesis in the developing human PFC.

a, Heat map shows differentially expressed genes in excitatory neurons from the three developmental stages. **b**, Scheme showing the timing of neurogenesis in the developing PFC. MZ, marginal zone; IZ, intermediate zone. **c**, The enriched gene ontology terms show the cell properties of the PFC at different weeks. $n = 1,057$ neurons.

d, Immunostaining after whole-cell patch-clamp recordings of PFC neurons at GW26. SP, subplate. Right, 3D reconstructions of traced cells. Scale bar, 50 μm . Representative images are shown from $n = 26$ neurons of three independent replicates. **e**, **f**, Whole-cell patch-clamp recordings including action potential (insets in **e**), phase plot (**e**) and evoked action potential (**f**) of the PFC neurons in **d** at GW26. **g**, **h**, Quantification of

Na^+ current (**g**) and K^+ current (**h**) in neurons from different locations in GW26 PFC. $n = 42$ neurons from three independent replicates. Upper versus deep, $P > 0.05$; upper versus subplate, $P < 0.0001$; deep versus subplate, $P < 0.0001$; two-way analysis of variance (ANOVA).

i, Quantification of neuron dendrite morphology by Sholl analysis. $n = 25$ from three independent replicates. Asterisks indicate P values for comparisons between neurons from upper and deep layers (black), deep layer and subplate (red), upper layer and subplate (blue): * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$; two-tailed t -test.

j, Representative example of whole-cell patch-clamp recordings of spontaneous EPSC and IPSC of deeper layer neurons. Data in **g-i** are mean \pm s.e.m. See Extended Data Figs 8, 9.

No parvalbumin (PV)⁺ interneurons were detected, suggesting that this neuron subtype might develop after GW26 (Fig. 4c). To assess whether interneurons are generated locally in the developing human PFC, we examined the presence and localization of interneuron progenitors by ISH and immunofluorescence. We observed a few TTF1⁺ cells in the anterior dorsal PFC above the classical neural progenitor zone at GW10 (Fig. 4d, Extended Data Fig. 10c). By contrast, TTF1⁺ cells were enriched in the medial ganglionic eminence (MGE), but not in the cortex close to the MGE (Fig. 4d). There were more TTF1⁺ cells in the dorsal PFC at GW17, which may originate from the ganglionic eminence (Fig. 4d). Furthermore, cell cycle analysis revealed that very few *TTF1*, *LHX6*, *DLX1* or *DLX2*-expressing cells in the developing PFC were actively progressing through the cell cycle (Extended Data Fig. 10d). Notably, *SST* mRNA could be detected in the PFC as early as GW7 by PCR with reverse transcription and in GW8 by RNA sequencing (RNA-seq) analysis (Fig. 4c, Extended Data Fig. 10e). This early appearance of *SST*-expressing cells suggests that they are unlikely to have migrated from the ganglionic eminence. However, using ISH, we also detected *SST* mRNA in the ventricular zone and subventricular zone at GW10, although *SST* protein was undetectable at this stage (Extended Data Fig. 10f, g). Together, our data indicate that although cells expressing interneuron progenitor markers were

detected in the dorsal PFC at an early stage, they may not have entered the cell cycle.

Since excitatory neurons and interneurons build up circuits cooperatively, we next investigated whether the development of these two types of neurons was synchronized. Excitatory neuron development peaked at GW16, whereas interneuron development peaked at GW26. We therefore selected neurons at GW16 and GW26 for pathway analysis by gene-set enrichment analysis (GSEA). Axon guidance signals, including axon attraction, repulsion and outgrowth, were more active in excitatory neurons than in interneurons at GW16; this was reversed at GW26 (Fig. 4e, f). The neurotrophin signalling pathway, which has roles in neuronal differentiation and prevention of cell death, exhibited a similar pattern (Extended Data Fig. 10h). Additionally, gene enrichment analysis showed that Notch signals were more involved in regulating biological activities of NPCs than those of neurons (Extended Data Fig. 10i). Together, these results suggest that excitatory neurons mature earlier than interneurons, and that this maturation process is regulated by extrinsic and intrinsic signals.

Our single-cell-resolution data illustrate the complex diversity of cell types in the developing human PFC that underlies the sophisticated cognitive function of humans. Notably, we found new IPC markers that enabled us to identify diverse neuron subtypes that are

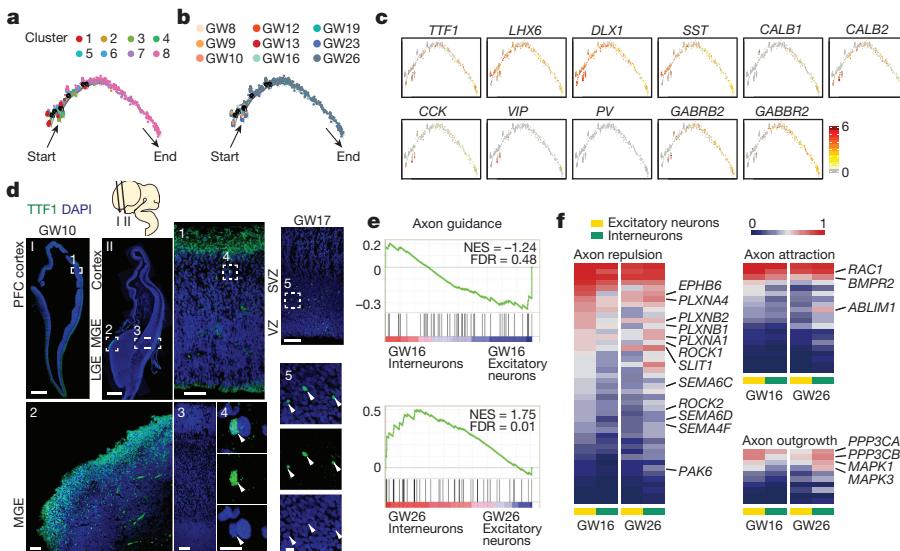


Figure 4 | Interneuron development and signalling pathways that regulate neuronal maturation. **a, b**, Single-cell trajectories of interneurons in the PFC by Monocle analysis. **c**, Expression of GABA receptors and genes known to be expressed in interneuron progenitors and interneuron subtypes are mapped to the single-cell trajectory plot. **d**, Immunofluorescence staining of TTF1 in the PFC (section I) and MGE-LGE (section II) at GW10 and GW17. LGE, Lateral ganglionic eminence. Scale bars, 500 µm (I, II), 100 µm (top right (GW17), 3), 50 µm (1, 2) and 10 µm (4, 5). $n = 4$ independent replicates per gestational week. **e**, GSEA of axon guidance signalling pathway in excitatory neurons and interneurons at GW16 and GW26. $n = 663$ excitatory neurons; $n = 485$ interneurons. NES, normalized enrichment score; FDR, false discovery rate. **f**, Mean expression of genes involved in the axon guidance signalling pathway in excitatory neurons and interneurons at GW16 and GW26. See Extended Data Fig. 10.

likely to contribute to the cellular basis of elaborate circuit formation in humans. Some neurological disorders and social cognition deficits, such as autism spectrum disorder and schizophrenia, have been linked to an imbalance of excitatory and inhibitory neurons (E/I ratio) in the PFC^{26–28}. Our data indicate that interneurons appear and mature later than excitatory neurons, and that this pattern is intrinsically regulated. Thus, transcriptome profiling of thousands of single cells in the developing PFC is a powerful tool for investigating the mechanisms behind neurological diseases and exploring potential therapies.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 15 September 2017; accepted 7 February 2018.

Published online 14 March 2018.

- Roth, G. & Dicke, U. Evolution of the brain and intelligence in primates. *Prog. Brain Res.* **195**, 413–430 (2012).
- O’Rahilly, R. & Müller, F. Significant features in the early prenatal development of the human brain. *Ann. Anat.* **190**, 105–118 (2008).
- Rakic, P. Evolution of the neocortex: a perspective from developmental biology. *Nat. Rev. Neurosci.* **10**, 724–735 (2009).
- Masserotti, G., Gascón, S. & Götz, M. Direct neuronal reprogramming: learning from and for development. *Development* **143**, 2494–2510 (2016).
- Satija, R., Farrell, J. A., Gennert, D., Schier, A. F. & Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
- Lake, B. B. et al. Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
- Trapnell, C. et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
- Qiu, X. et al. Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
- Qiu, X. et al. Single-cell mRNA quantification and differential analysis with Census. *Nat. Methods* **14**, 309–315 (2017).
- Ma, T. et al. Subcortical origins of human and monkey neocortical interneurons. *Nat. Neurosci.* **16**, 1588–1597 (2013).
- Sandberg, M. et al. Transcriptional networks controlled by NFKX2-1 in the development of forebrain GABAergic neurons. *Neuron* **91**, 1260–1275 (2016).
- Radonjić, N. V. et al. Diversity of cortical interneurons in primates: the role of the dorsal proliferative niche. *Cell Rep.* **9**, 2139–2151 (2014).
- Kierdorf, K. et al. Microglia emerge from erythromyeloid precursors via Pu-1 and Irf8-dependent pathways. *Nat. Neurosci.* **16**, 273–280 (2013).
- Prinz, M. & Priller, J. Microglia and brain macrophages in the molecular age: from origin to neuropsychiatric disease. *Nat. Rev. Neurosci.* **15**, 300–312 (2014).
- Hong, S. et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. *Science* **352**, 712–716 (2016).
- Vasek, M. J. et al. A complement-microglial axis drives synapse loss during virus-induced memory impairment. *Nature* **534**, 538–543 (2016).
- Pollen, A. A. et al. Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
- Lewitus, E., Kelava, I. & Huttner, W. B. Conical expansion of the outer subventricular zone and the role of neocortical folding in evolution and development. *Front. Hum. Neurosci.* **7**, 424 (2013).

- Noctor, S. C., Martínez-Cerdeño, V., Ivic, L. & Kriegstein, A. R. Cortical neurons arise in symmetric and asymmetric division zones and migrate through specific phases. *Nat. Neurosci.* **7**, 136–144 (2004).
- Homem, C. C., Repic, M. & Knoblich, J. A. Proliferation control in neural stem and progenitor cells. *Nat. Rev. Neurosci.* **16**, 647–659 (2015).
- Kowalczyk, T. et al. Intermediate neuronal progenitors (basal progenitors) produce pyramidal-projection neurons for all layers of cerebral cortex. *Cereb. Cortex* **19**, 2439–2450 (2009).
- Noctor, S. C., Martínez-Cerdeño, V. & Kriegstein, A. R. Contribution of intermediate progenitor cells to cortical histogenesis. *Arch. Neurol.* **64**, 639–642 (2007).
- Vasistha, N. A. et al. Cortical and clonal contribution of Tbr2 expressing progenitors in the developing mouse brain. *Cereb. Cortex* **25**, 3290–3302 (2015).
- Lui, J. H., Hansen, D. V. & Kriegstein, A. R. Development and evolution of the human neocortex. *Cell* **146**, 18–36 (2011).
- Bandler, R. C., Mayer, C. & Fishell, G. Cortical interneuron specification: the juncture of genes, time and geometry. *Curr. Opin. Neurobiol.* **27**, 17–24 (2017).
- Shi, G. X., Han, J. & Andres, D. A. RIN GTPase couples nerve growth factor signaling to p38 and b-Raf/ERK pathways to promote neuronal differentiation. *J. Biol. Chem.* **280**, 37599–37609 (2005).
- Canitano, R. & Pallagrosi, M. Autism spectrum disorders and schizophrenia spectrum disorders: excitation/inhibition imbalance and developmental trajectories. *Front. Psychiatry* **8**, 69 (2017).
- Bicks, L. K., Koike, H., Akbarian, S. & Morishita, H. Prefrontal cortex and social cognition in mouse and man. *Front. Psychol.* **6**, 1805 (2015).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank members of the Wang and Tang laboratories for discussions. This work was supported by National Basic Research Program of China (2014CB964600), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDA16020601), the National Natural Science Foundation of China (NSFC) (91732301, 31371100, 31771140), National Key Research and Development Program of China (2017YFA0103303, 2017YFA0102601), Shanghai Brain-Intelligence Project from STCSM (16JC1420500), Newton Advanced Fellowship (NA140246) to X.W. and Youth Innovation Promotion Association CAS to Q.W.

Author Contributions Q.W., X.F., J.Q., F.T. and X.W. conceived the project, designed the experiments and wrote the manuscript. S.Zho. and X.F. performed RNA-seq. S.Zha. and J.D. analysed the data. J.Z., L.L., L.S., H.Z., L.Y. and X.X. prepared the samples. S.Zho. and Q.W. performed immunofluorescence, *in situ* hybridization and imaging. L.S. and N.P. performed the electrophysiology experiments. All authors edited and proofread the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing interests. Readers are welcome to comment on the online version of the paper. Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to X.W. (xiaozunwang@ibp.ac.cn), J.Q. (jie.qiao@263.net), J.Z. (drzhangj@outlook.com) or F.T. (tangfuchou@pku.edu.cn).

Reviewer Information *Nature* thanks H. Song and the other anonymous reviewer(s) for their contribution to the peer review of this work.

METHODS

Ethics statement. The human embryo collection and research analysis was approved by the Reproductive Study Ethics Committee of Peking University Third Hospital (2012SZ-013 and 2017SZ-043) and Beijing Anzhen Hospital (2014012x). The informed consent was designed as recommended by the ISSCR guidelines for fetal tissue donation. Informed consent for fetal tissue procurement and research was obtained from the patient after her decision to legally terminate her pregnancy but before the abortive procedure. Fetal cortical tissue samples were collected after the donor patients signed an informed consent document that was in strict observance of the legal and institutional ethical regulations for elective pregnancy termination specimens at Peking University Third Hospital and Beijing Anzhen Hospital, Capital Medical University. All samples used in these studies had not been involved in any other procedures. All the protocols were in compliance with the 'Interim Measures for the Administration of Human Genetic Resources' administered by the Chinese Ministry of Health.

Tissue sample collection and dissection. Fetal brains were collected in ice-cold artificial cerebrospinal fluid containing 125.0 mM NaCl, 26.0 mM NaHCO₃, 2.5 mM KCl, 2.0 mM CaCl₂, 1.0 mM MgCl₂, 1.25 mM NaH₂PO₄ at a pH of 7.4 when oxygenated (95% O₂ and 5% CO₂). The prefrontal cortex was dissected and placed in hibernate E medium (Invitrogen, A1247601). The PFC tissue was first digested in brain tissue digestion medium 1 (BTDM1, 2 mg/ml collagenase IV (Gibco, 17104-019) and 10 U/μl DNase I (NEB, M0303L) in hibernate E medium). After brief pipetting to roughly separate the tissue into small pieces, brain tissue was then digested in digestion medium 2 (BTDM2, 1 mg/ml papain (Sigma, P4762) and 10 U/μl DNase I in hibernate E medium). The tissue was vortexed and kept at 37 °C for 5 min. The cell suspension was pipetted further to disperse it into single cells and centrifuged at 300g for 5 min to obtain a cell pellet. The digestion medium was carefully removed and the cell pellet was resuspended in 500 μl hibernate E medium and kept on ice.

Single cells were picked into cell lysis buffer by mouth pipette and reverse transcription was performed as described for Smartseq2^{29,30}, except that we used barcoded reverse transcription primers instead of the oligdT30VN primer. Amplification was performed using the Smartseq2 protocol with minor modifications. The PCR products with different barcodes were pooled together for purification and library construction³¹. The libraries were processed on the Illumina platform for sequencing of 150 bp pair-end reads.

Processing of raw single-cell RNA-seq data. The read2 data of pooled cells were split into single-cell data using the barcode sequences contained in the first 8 bp. The next 8 bp were recorded as unique molecular identifiers (UMIs). The read1 data were then sorted by the header of the corresponding read2 data and added to the UMI information. The poly(A) tail sequence and template switch oligo (TSO) sequences were then trimmed from the read1 sequences. Quality control was performed by removing reads with adapter contaminants (length <37 bp) and low-quality bases ($N > 10\%$). Subsequently, TopHat³² (version 2.0.12) was used to align clean reads to the hg19 human transcriptome downloaded from UCSC³³. Uniquely aligned reads were counted using the 'htseq-count' tool of HTSeq³⁴. Transcripts for the same genes that shared the same UMI sequence were merged, thus the number of different UMIs for each gene was considered as the transcript count of that gene without PCR bias for each sequenced cell.

The transcript counts of each cell were normalized to transcript per million (TPM), where TPM is the transcript count of each gene divided by the sum of transcript counts of that cell, multiplied by one million. TPM values were then normalized by $\log((\text{TPM}/10) + 1)$ for subsequent analysis. TPM was divided by 10 because the sequencing depth of 64% of single-cell samples contained less than 1,000,000 reads. Overall 2,394 individual cells were collected for single-cell cDNA amplification and 2,309 cells passed the quality control criteria. On average, there were 1.1 million mapped reads and 2,654 detected genes for each cell.

Identification of cell types and subtypes by nonlinear dimensional reduction and random forests. The Seurat⁵ package (v.1.2.1) implemented in R was applied to identify major cell types among 2,394 single cells from the PFC. Only cells that expressed more than 1,000 genes were considered, and only genes with normalized expression level greater than 1 that were expressed in at least three single cells were included, leaving 17,854 genes across 2,333 samples for clustering analysis. After initial clustering, two clusters enriched in haemoglobin genes and microglia-specific genes (PTPRC, CSF1R, AIF1) were removed before the second clustering analysis. The haemoglobin genes HBA1, HBA2, HBB, HBD, HBE1, HBG1, HBG2, HBM, HBQ1 and HBZ, and immune cells found in the second clustering analysis were also excluded from subsequent analyses. Following the second clustering analysis, 2,309 cells excluding immune cells and cells enriched in haemoglobin genes were mapped by t-SNE (Fig. 1a). PAX6, NEUROD2, GAD1, PDGFRA, AQP4 and PTPRC were used as markers to identify the major cell types in the brain: neural progenitor cells (NPCs), excitatory neurons, interneurons, oligodendrocyte progenitor cells (OPCs), astrocytes and microglia, respectively.

Random forests⁶ was used to identify subtypes of the major cell types in human embryonic PFC. We used about 1,000 genes, identified as over-dispersed genes by 'mean.var.plot' of Seurat, as feature genes for selection by random forests. The number of folds in the cross-validation of random forest analysis was varied between 5 and 10.

Identification of differentially expressed genes among clusters. Genes that were differentially expressed in each cluster were identified using the Seurat function 'find_all_markers' and tested by 'roc' and DESeq2³⁵. The ROC test returns the 'classification power' for any individual gene (ranging from 0 (random) to 1 (perfect)); positive 'avg_diff' means upregulation in the corresponding clusters, whereas negative 'avg_diff' means downregulation. DESeq2 returns 'log2Fold-Change' (higher absolute value means higher fold change of corresponding clusters) and *P* value of each tested gene. We used DAVID^{36,37} (<https://david.ncifcrf.gov/home.jsp>) and Metascape³⁸ (<http://metascape.org>) to perform biological process enrichment analysis with the differentially expressed genes of each cluster.

Constructing single cell trajectories in the PFC. The Monocle^{7–9} package was used to analyse single cell trajectories in order to discover developmental transitions. We used differentially expressed genes identified by Seurat to sort cells in pseudo-time order. The actual gestational time of each cell informed us of the start point of the pseudo-time in the first round of 'orderCells'. We then set this state as the root_state argument and called 'orderCells' again. 'DDRTree' was applied to reduce dimensions and the visualization functions 'plot_cell_trajectory' or 'plot_complex_cell_trajectory' were used to plot the minimum spanning tree on cells.

Pathways analysis. GSEA³⁹ was applied to identify a priori-defined gene sets that show statistically significant differences between two given clusters (<http://www.broadinstitute.org/gsea/index.jsp>). We used the expression file containing about 5,000 variable genes as input, and implied gene sets of KEGG⁴⁰ pathways and Gene Ontology⁴¹ (<http://www.geneontology.org/>), which were collected in Molecular Signatures Database (MSigDB)^{39,42}.

We calculated the mean expression of 'axon guidance pathway' genes in excitatory and inhibitory neurons respectively from GW16 and GW26 PFC. The mean expression values of both types of neurons at the same gestational time were compared and tested by *t*-test for significance. Significantly differentially expressed genes were plotted as heat maps.

The average gene expression level for each stage (early, middle and late stages) in the Notch pathway was plotted in the heat map and the ratio of expressed genes to all the genes in Notch pathway was plotted as a histogram at the top of the heat map. **Cell-cycle analysis.** We applied a cell-cycle-related gene set with 46 genes for G1/S phase and 54 genes for G2/M phase of the cell cycle. We added three genes, CDK4, CDK2 and CDK6, based on the core set of cell-cycle-related genes from Tirosh *et al.*^{43,44}. We defined G1/S and G2/M states of each cell by comparing the average expression of the two gene sets.

Immunofluorescent staining. Tissue samples were fixed overnight in 4% paraformaldehyde, cryoprotected in 30% sucrose and embedded in optimal cutting temperature medium (Thermo Scientific). Thin 40-μm cryosections were collected on superfrost slides (VWR) using a Leica CM3050S cryostat. For immunohistochemistry, heat-induced antigen retrieval was performed in 10 mM sodium citrate buffer, pH 6. Primary antibodies: rabbit anti-FOXP2 (1:500, Abcam ab16046), mouse anti-SATB2 (1:250, Abcam ab51502), mouse anti-CD45 (1:100, Abcam ab8216), goat anti-AIF1 (1:200, Abcam ab5076), goat anti-SOX2 (1:250, Santa Cruz sc-17320), rabbit anti-PAX6 (1:500, BioLegend 901301), rabbit anti-SFRP1 (1:500, Abcam ab126613), chicken anti-EOMES (1:500, Millipore AB15894), mouse anti-RBFOX1 (1:500, Abcam ab183348), rabbit anti-TTF1 (1:300, Abcam ab86023), Neurobiotin tracer (1:1,000, Vectorlabs SP-1120-50), rabbit anti-NEUROD2 (1:500, Abcam ab104430), mouse anti-NEUROD1 (1:100, Abcam ab60704), rabbit anti-HMGA2 (1:100, Abcam ab207301), rabbit anti-HOPX (1:1,000, Santa cruz sc-30216), mouse anti-Ki-67 (1:100, BD 550609). Primary antibodies were diluted in blocking buffer containing 10% donkey serum, 0.5% Triton X-100 and 0.2% gelatin in PBS at pH 7.4. Binding was visualized using an appropriate Alexa Fluor 488, Alexa Fluor 594 or Alexa Fluor 647 fluorophore-conjugated secondary antibody (Life Technologies). Cell nuclei were stained using DAPI (Life Technologies). Images were collected using an Olympus FV1000 confocal microscope.

In situ hybridization. The *in situ* hybridization protocol has been described previously⁴⁵. In brief, probes complementary to target human mRNA used for RNA *in situ* hybridization were cloned from primary human fetal cortical cDNA samples, reverse-transcribed using PrimeScript II 1st Strand cDNA Synthesis Kit (Takara) with oligo dT primers; or from RNA samples isolated from GW7 or GW17 human cortex using SV Total RNA Isolation System (Promega). Specific genes were amplified using the following primers: *NHL1* forward, CTTCGCCAGCTGCGCAAG; *NHL1* reverse, ATTCCCAGCTACTGAGATTG; *NHL2* forward, GAGCTCCGCAAATTGCTGC; *NHL2* reverse, GCTCAGATCCTCCCCACAGT;

TP53I11 forward, CGTGTGCTGTGATCTCCTGG; *TP53I11* reverse, GAACTC AGGCTGGTTGCAG; *SLC1A3* forward, CTAGGCCTCAGTGTCTCATCT; *SLC1A3* reverse, GCGTCTTGACTGGATATTCC; *PPP1R17* forward, TGTGACATTGCTCAGGGACG; *PPP1R17* reverse, CTCACTAGCTAACCCACACTC; *SST* forward, CCTAGAGTTGACCAGCCACTC; *SST* reverse, AGTTCTAATGCAAGGGTCTCGP; *TFI* forward, CGACTTTCTTAACAACCTGGC; *TFI* reverse, AAAAGACTGACGCCGCAAATAC. Primers specific for target genes of interest were designed using Primer3 and amplified by PCR using Q5 High-Fidelity DNA Polymerase (NEB). PCR products of the predicted band sizes were extracted from gels and ligated into the Hieff Clone Plus One Step Cloning Kit (Yeason). Ligation products were transfected into Trans5 α Chemically Competent *Escherichia coli* (Transgene). Cloned sequences were confirmed by sequencing. Digoxigenin-labelled RNA probes for *in situ* hybridization were generated by linearizing the pSPT18 Vector and transcribing the probe *in vitro* using T7 or SP6 RNA Polymerase (Roche) in the presence of DIG-RNA Labelling Mix (Roche). Fetal brain sections (30 μ m thick) were hybridized with RNA probes, at a final concentration of 500 ng/ml, overnight at 64.5 °C in hybridization solution (50% formamide, 10% dextran sulphate, 0.2% tRNA (Invitrogen), 1 \times Denhardt's solution (Sigma), 1 \times salt solution (containing 0.2 M NaCl, 0.01 M Tris, 5 mM NaH₂PO₄, 5 mM Na₂HPO₄, 5 mM EDTA pH 7.5)). After the sections were washed, alkaline phosphatase-coupled anti-digoxigenin Fab fragments (Roche) were applied. For visualization of the labelled cRNAs, the sections were incubated in the dark in NBT–BCIP solution (Roche). Images were taken with a Leica SCN400 (Leica Microsystems).

Electrophysiological recording. Coronal slices containing PFC (500 μ m) from GW23 or GW26 embryos were prepared using a vibratome (VT1200S, Leica) in oxygenated (95% O₂ and 5% CO₂) ice-cold sucrose-based artificial cerebrospinal fluid (s-ACSF, 234 mM sucrose, 2.5 mM KCl, 26 mM NaHCO₃, 1.25 mM NaH₂PO₄, 11 mM D-glucose, 0.5 mM CaCl₂ and 10 mM MgSO₄). The slices were kept in an incubating chamber filled with oxygenated ACSF (126 mM NaCl, 3 mM KCl, 1.2 mM NaH₂PO₄, 2.4 mM CaCl₂, 1.3 mM MgSO₄, 26 mM NaHCO₃, 10 mM D-glucose) at 34 °C for 30 min. After a recovery period of at least 60 min at room temperature, an individual slice was transferred to a recording chamber and was continuously superfused with oxygenated ACSF (4 ml/min) at room temperature.

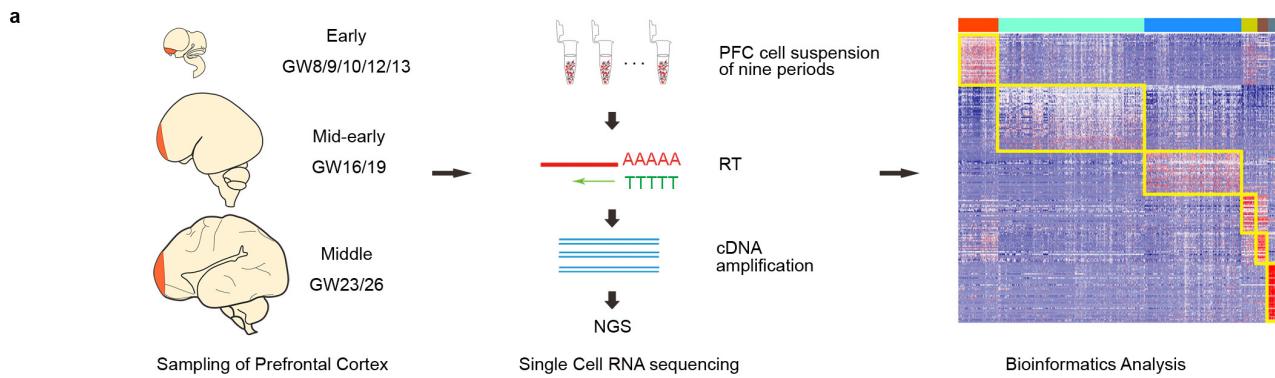
Whole-cell patch-clamp recordings were performed with a patch pipette solution containing (in mM) 140 K-gluconate, 5 KCl, 0.2 EGTA, 2 MgCl₂ and 10 HEPES. The electrode was filled with ~0.5% of neurobiotin and 0.1% Lucifer yellow, which was diffused into the recording neuron under the whole-cell recording for further confirmation staining. Electrode resistance ranged from 10 to 13 M Ω . Tight seals (>1 G Ω) were obtained on cell bodies before rupturing the membrane with negative pressure.

The whole-cell recording was conducted with an MultiClamp 700B amplifier (Axon Instruments). Currents were typically digitized at 100 kHz, and macroscopic records were filtered at 2 kHz. Electrophysiological data were analysed with ClampFit software (version 10.0; Axon Instruments) and GraphPad Prism (v.6.0; GraphPad Software). A single action potential was evoked by injecting short

(2–4 ms) depolarizing current pulses (in the current clamp). Data were accepted for analysis only in cases in which the recorded neurons exhibited a resting potential <−10 mV throughout the experiment. Spiking properties were calculated based on the response to a 500-ms current pulse at −60 pA, 0 pA and 60 pA, respectively. To detect spontaneous glutamate receptor-mediated excitatory postsynaptic currents (sEPSCs) or inhibitory postsynaptic currents (sIPSCs), the cells were monitored at a holding potential of −70 mV or 0 mV, respectively.

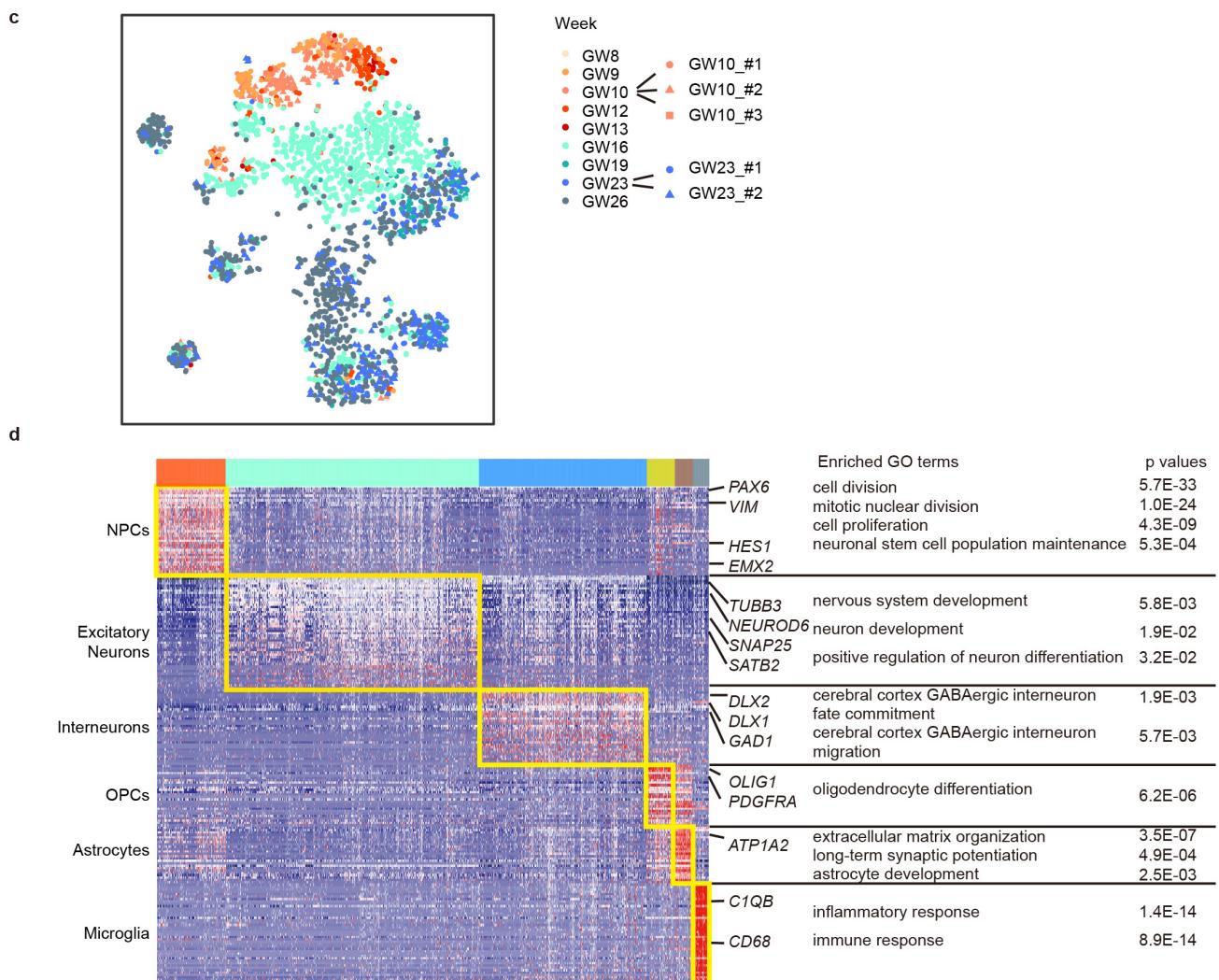
Data availability. The single-cell RNA-seq data used in this study have been deposited in the Gene Expression Omnibus (GEO) under the accession number GSE104276. Raw image files used in the figures that support the findings of this study are available from the corresponding authors upon reasonable request.

29. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc.* **9**, 171–181 (2014).
30. Guo, F. *et al.* The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell* **161**, 1437–1452 (2015).
31. Li, L. *et al.* Single-cell RNA-seq analysis maps development of human germline cells and gonadal niche interactions. *Cell Stem Cell* **20**, 891–892 (2017).
32. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
33. Hinrichs, A. S. *et al.* The UCSC Genome Browser Database: update 2006. *Nucleic Acids Res.* **37**, D590–D598 (2006).
34. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
35. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
36. Huang, W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
37. Huang, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
38. Tripathi, S. *et al.* Meta- and orthogonal integration of influenza ‘omics’ data defines a role for UBR4 in virus budding. *Cell Host Microbe* **18**, 723–735 (2015).
39. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
40. Ogata, H. *et al.* KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **27**, 29–34 (1999).
41. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).
42. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
43. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352**, 189–196 (2016).
44. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
45. Palop, J. J., Roberson, E. D. & Cobos, I. Step-by-step *in situ* hybridization method for localizing gene expression changes in the brain. *Methods Mol. Biol.* **670**, 207–230 (2011).



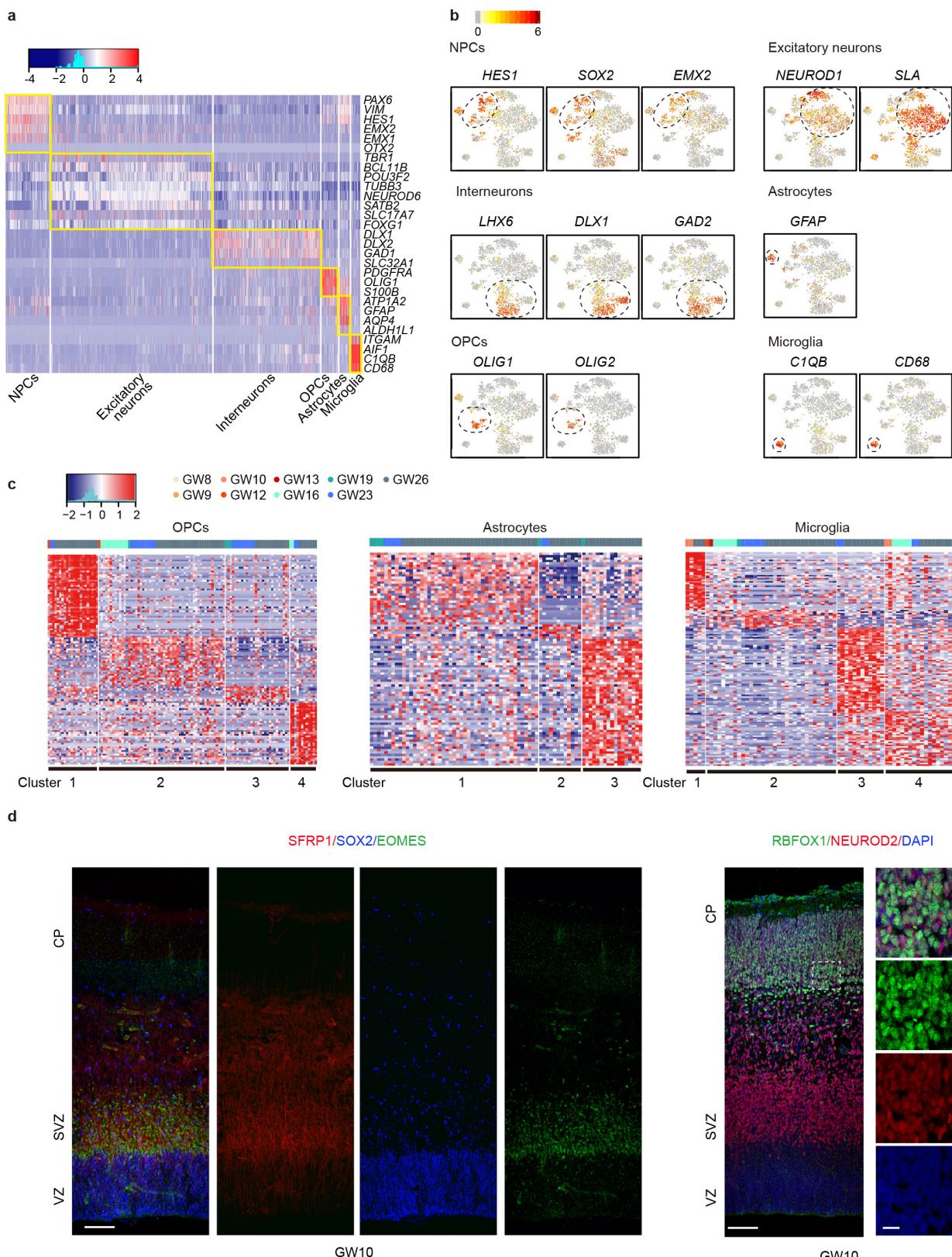
b

	GW8	GW9	GW10_01	GW10_02	GW10_03	GW12	GW13	GW16	GW19	GW23_01	GW23_02	GW26	Sum
Gender	Female	Female	Male	Female	Female	Male	Female	Female	Female	Male	Female	Female	2,394
Sequenced cells	23	88	48	95	48	88	24	789	120	143	181	747	2,309
Filtered cells	23	88	47	92	47	85	24	776	120	132	176	699	2,309



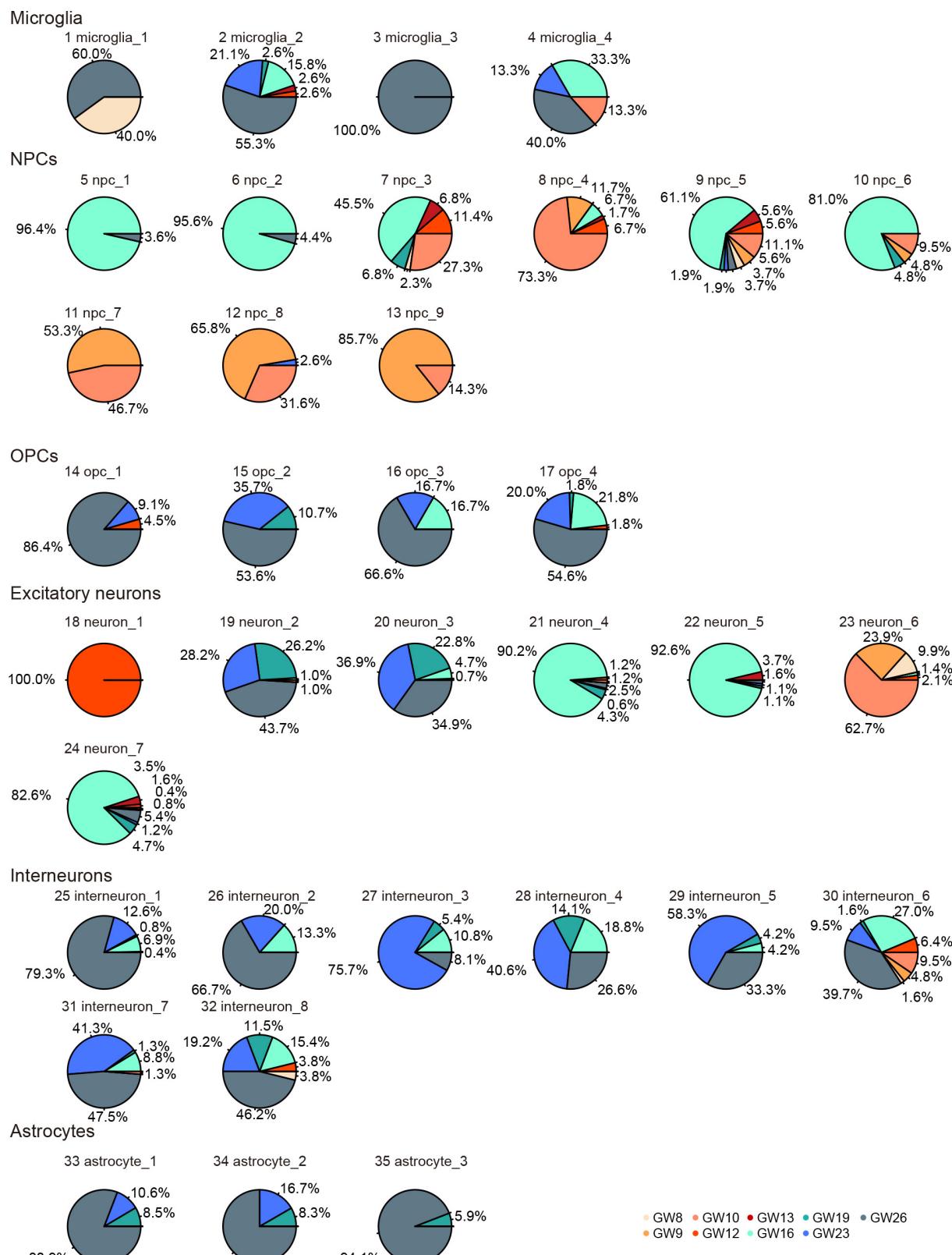
Extended Data Figure 1 | Single-cell RNA-seq information and molecular diversity of single cells. **a**, Experimental workflow for single-cell RNA-seq of human developing PFC. **b**, Table summarizing PFC sampling. **c**, t-SNE plots of cells in the PFC. $n = 3$ (GW10) and $n = 2$ (GW23) independent biological samples. No obvious differences in distribution were observed among the different batches at the same

development stages. Each colour represents the gestational week, and the colour contours correspond to the cell types. Expression of known markers is shown using the same layout (grey, no expression; yellow-red, relative expression). **d**, Heat map shows blocks of genes enriched in each cell type. Right, specific genes related to each type are highlighted with enriched gene ontology terms. $n = 2,309$ cells.

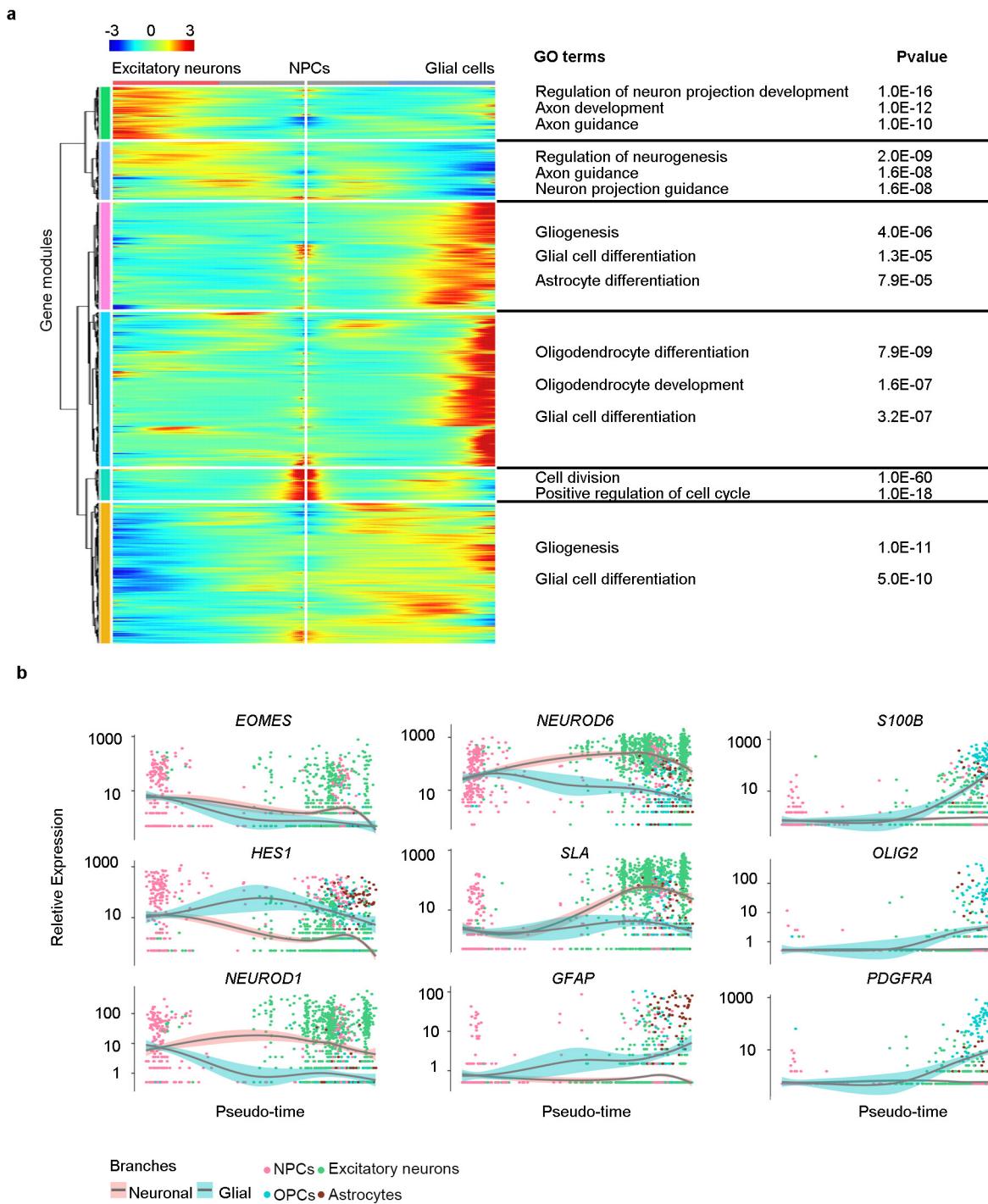

Extended Data Figure 2 | Molecular diversity of subgroups of cells.

a, Heat map showing the expression level and identity of genes in six major cell types in the PFC. Expression of genes known to be expressed in each cell type is shown to the right of each heat map panel. **b**, Expression of known markers is shown using the same layout as in Fig. 1a (grey, no expression; yellow–red, relative expression). **c**, Heat maps show the

subclasses of OPCs, astrocytes and microglia. The genes were organized into clusters. Top chart, gestational weeks. **d**, Immunostaining for new markers of NPCs (SFRP1) and excitatory neurons (RBFOX1) at GW10 in the PFC. Scale bars, 100 µm (left), 100 µm (middle) and 10 µm (right). $n = 3$ independent replicates per gestational week.

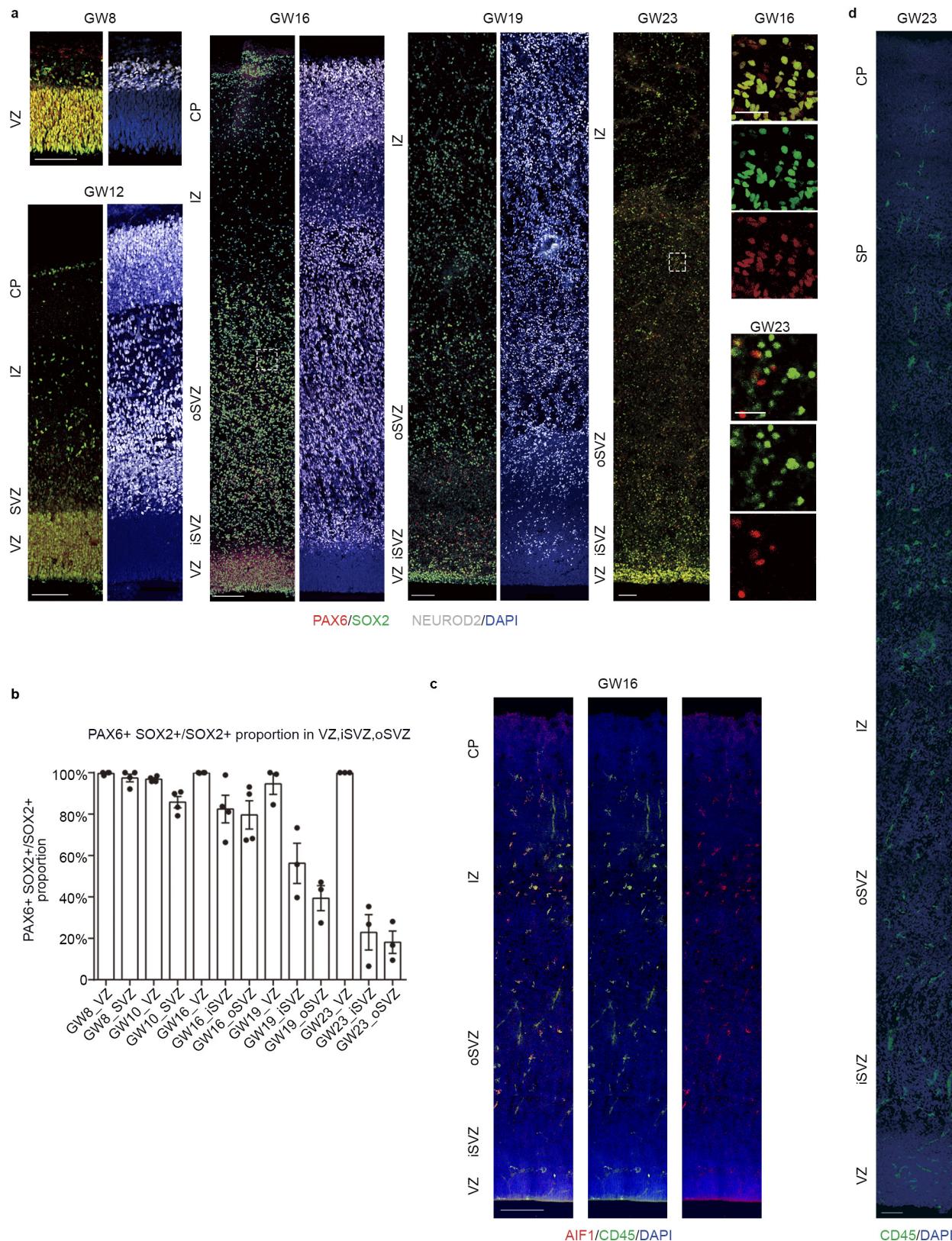


Extended Data Figure 3 | Pie chart of the distribution of the 35 subclasses of the six cell types across gestational time.



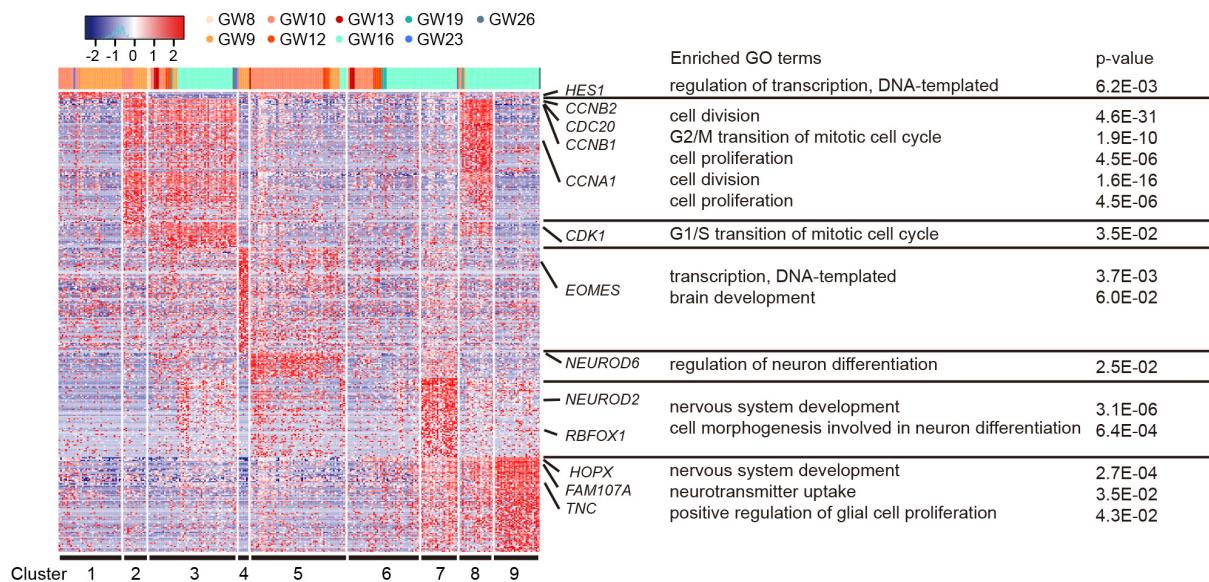
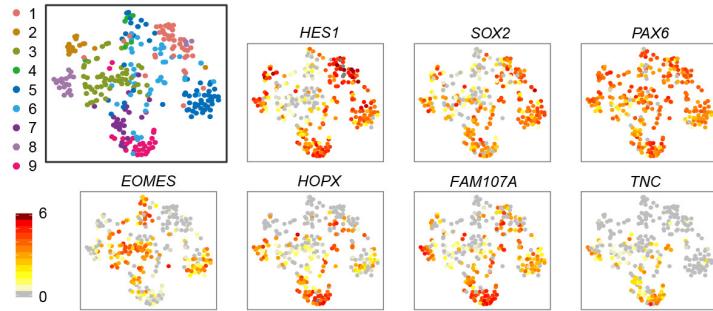
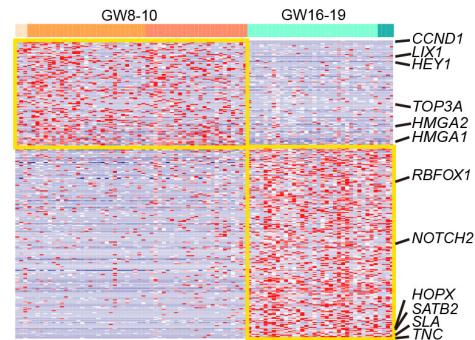
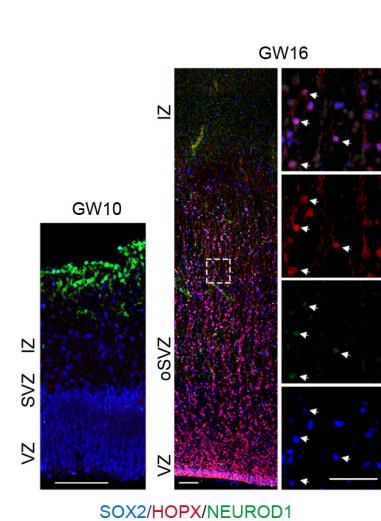
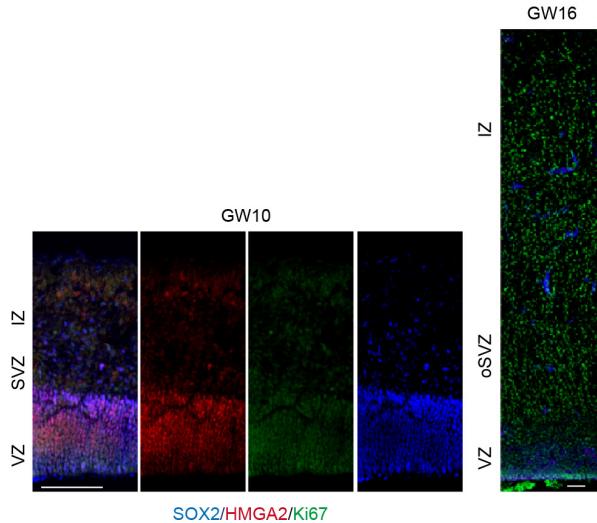
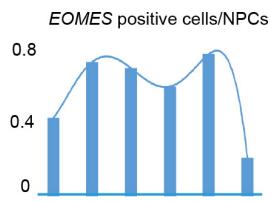
Extended Data Figure 4 | Relative expression of known markers of each cell type. **a**, The bifurcation of gene expression along two branches is clustered hierarchically into six modules. Gene ontology analysis of each module reflected the processes controlling neuronal and glial fate commitment. In this heat map, columns are points in pseudo-time, rows are genes and the middle (NPCs) is the beginning of pseudo-time. One

lineage goes from the middle of the heat map to the right (glial cells) while the other lineage goes to the left (excitatory neurons). $n = 1,540$ cells. **b**, The markers (*EOMES*, *HES1* for NPCs; *NEUROD1*, *NEUROD6*, *SLA* for excitatory neurons; *GFAP*, *S100B* for astrocytes; *OLIG2*, *PDGFRA* for OPCs) were ordered by Monocle analysis in pseudo-time as in Fig. 1c; the shadow indicates the confidence interval around the fitted curve.



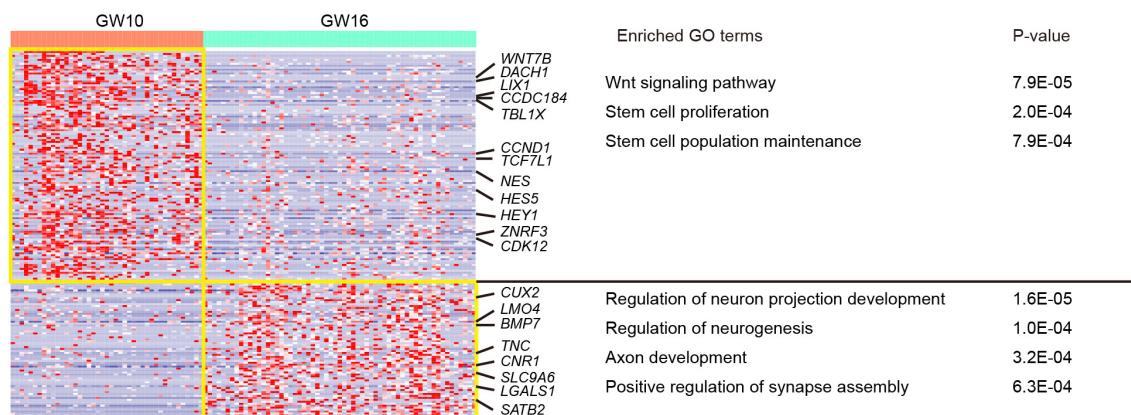
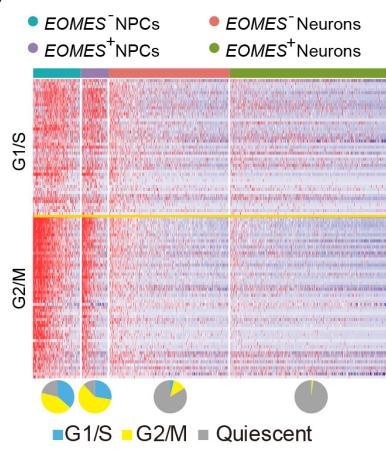
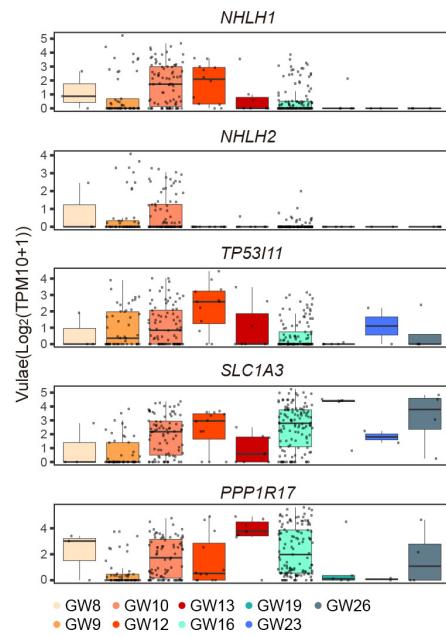
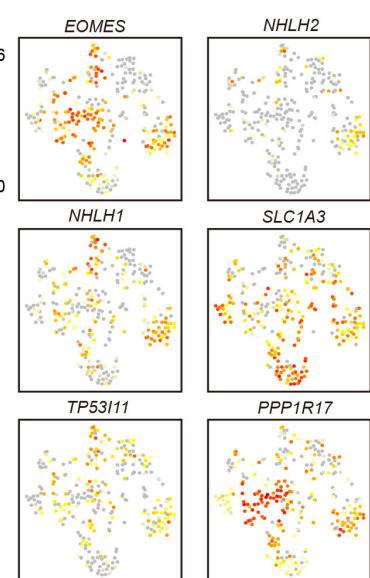
Extended Data Figure 5 | Immunostaining of neural progenitor cells and microglia in the developing PFC. **a**, Immunostaining for known markers (PAX6 and SOX2 for NPCs, NEUROD2 for excitatory neurons) at GW8, GW12, GW16, GW19 and GW23, showing the position of NPCs and excitatory neurons. Scale bars, 100 μ m (left) and 25 μ m (right). $n = 3$ for GW 19 and GW23, $n = 4$ for GW8, GW12 and GW16. **b**, Bar

charts showing the proportion of PAX6⁺SOX2⁺ cells in SOX2⁺ cells in the ventricular zone, iSVZ and oSVZ in the PFC at GW8, GW12, GW16, GW19 and GW23 relative to the images in **a**. Data are mean \pm s.e.m. **c**, Co-staining for AIF1 and CD45 to label microglia in the PFC at GW16. **d**, Immunostaining for the microglia marker CD45 in the PFC at GW23. Scale bar, 100 μ m. $n = 3$ independent replicates per gestational week.

a**b****c****d****e**

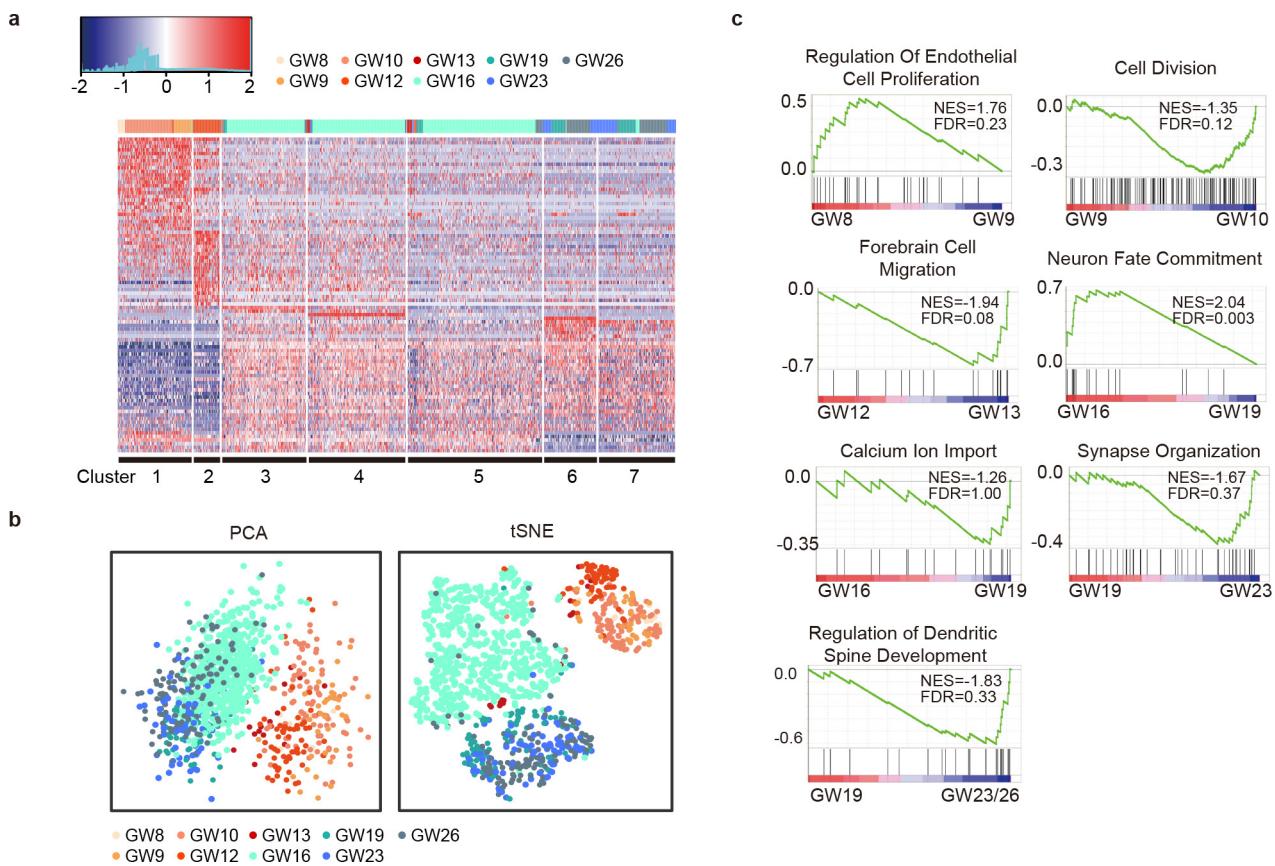
Extended Data Figure 6 | Subclasses and Monocle analysis of neural progenitor cells in the developing PFC. **a**, Heat map of differentially expressed genes of subclasses in NPCs. Expression of known genes in each type is shown on the right of each heat map panel with enriched gene ontology terms. The graph on the top shows the distribution of each subclass across gestational week. $n = 288$ cells. **b**, Visualization of nine major classes of NPCs using t-SNE (colour on the left, subtype of NPCs) with known marker expression (right and bottom: grey, no expression;

yellow-red, relative expression). **c**, Heat map of differentially expressed genes between GW8 to GW10 and GW16 to GW19 in *EOMES*⁻ NPCs. The graph at the top shows the distribution of each subclass across gestational weeks. **d**, Immunostaining for different stage markers of GW10 (HMGA2) and GW16 (HOPX) in the PFC. Scale bars, 100 μ m. $n = 3$ independent replicates per gestational week. **e**, Histogram showing the ratio of *EOMES*⁺ cells to all NPCs across gestational weeks quantified by RNA-seq data.

a**b****c****d**

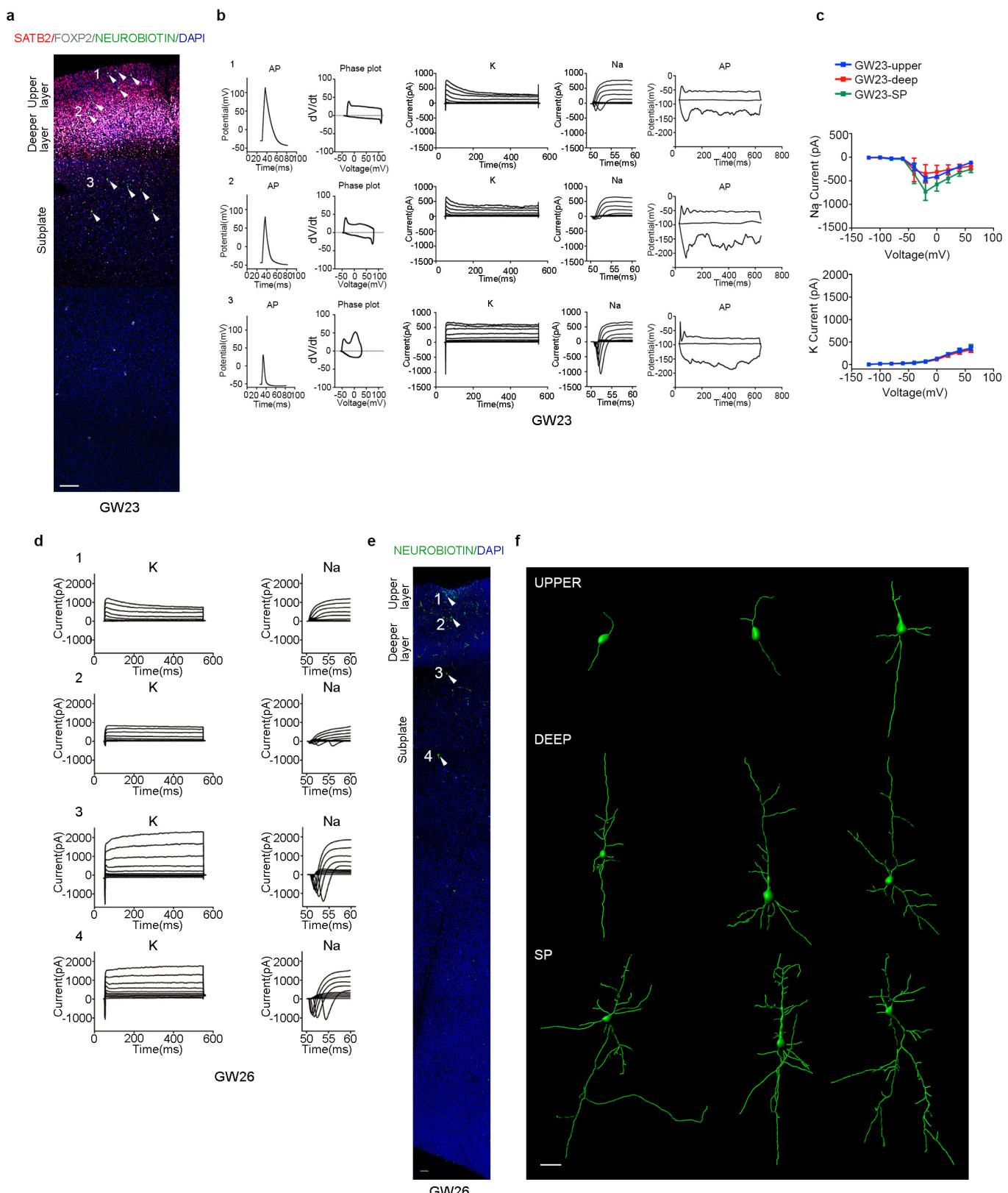
Extended Data Figure 7 | Expression of markers of IPCs in the developing PFC. **a**, Heat map of differentially expressed genes in *EOMES*⁺ NPCs at GW10 and GW16. Expression of known genes at each stage is shown to the right of each heat map panel with enriched gene ontology terms. $n = 288$ cells. **b**, Heat map of expression of cell-cycle genes in all NPCs and neurons. The pie chart at the bottom indicates the ratios of cell types at each cell-cycle stage. **c**, Gene expression of novel markers of IPCs

across developmental time corresponding to human cortical neurogenesis. NPCs: GW8, 3 cells; GW9, 50 cells; GW10, 84 cells; GW12, 12 cells; GW13, 7 cells; GW16, 123 cells; GW19, 5 cells; GW23, 2 cells; GW26, 4 cells. **d**, The expression of novel markers of IPCs is shown using the same layout as Extended Data Fig. 2b (grey, no expression; yellow-red, relative expression). $n = 288$ NPCs.



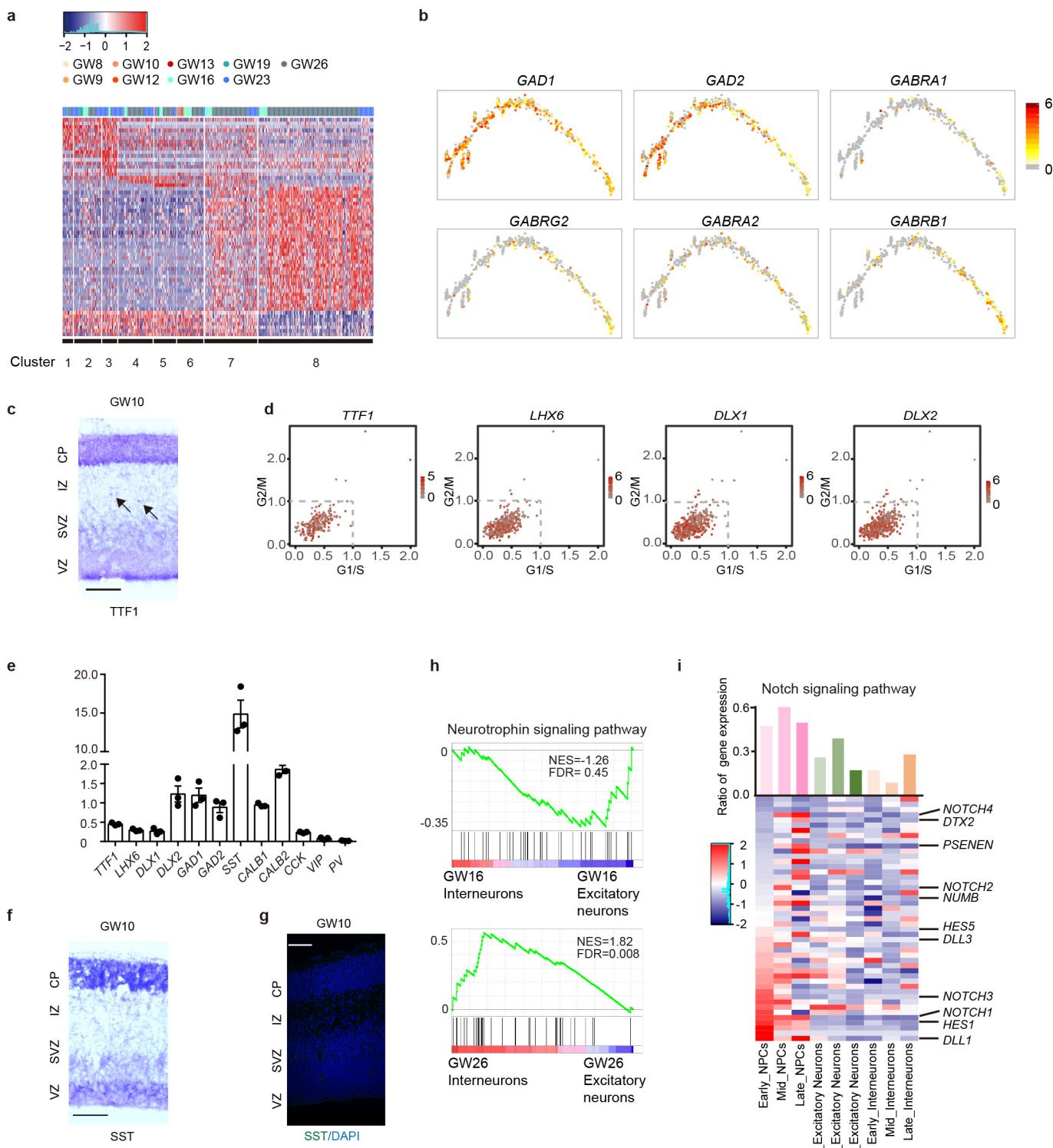
Extended Data Figure 8 | Subclasses of excitatory neurons in the developing PFC. **a**, Heat map of differentially expressed genes in different subtypes of excitatory neurons. The bar chart at the top shows the gestational week. **b**, Principal component analysis and t-SNE were used to sort excitatory neurons into subgroups. Each dot represents a single cell.

The colour shows the gestational week. **c**, GSEA shows genes related to key functions of excitatory neurons in the PFC across multiple developmental time points corresponding to human cortical neurogenesis. $n = 1,057$ excitatory neurons.



Extended Data Figure 9 | Whole-cell patch-clamp recordings in the developing PFC. **a**, Immunostaining after whole-cell patch-clamp recordings of the PFC at GW23. Scale bar, 100 μ m. $n = 3$ independent replicates. **b**, Whole-cell patch-clamp recordings of the PFC at GW23. **c**, Quantification of Na^+ current (top) and K^+ current (bottom) of neurons located in upper layer, deep layer or subplate of GW23 PFC. $n = 31$ neurons from three independent replicates. Data are mean \pm s.e.m.

All $P > 0.05$, two-way ANOVA. **d**, Whole-cell patch-clamp recordings of the PFC neurons at GW26. **e**, Immunostaining after whole-cell patch-clamp recordings of the PFC neurons at GW26. Scale bar, 100 μ m. **f**, Three-dimensional reconstructions of neurons are shown. Scale bar, 30 μ m. Nine representative images are shown from $n = 26$ neurons from three independent replicates.



Extended Data Figure 10 | Subclasses of interneurons and signal pathways regulating neurogenesis. **a**, Heat map of illustrated subclasses of interneurons. The bar chart on the top shows the gestational week. **b**, The expression of known markers of interneurons mapped to the Monocle analysis (Fig. 4a) shows the pseudo-time course of interneurons during development (grey, no expression; yellow-red, relative expression). **c**, *In situ* hybridization of *TTF1* shows the position of interneuron progenitor cells in PFC at GW10. Scale bar, 100 μm. *n* = 3 independent replicates. **d**, Pattern of marker gene expression in interneuron progenitor cells mapped onto the cell cycle plot for all interneurons. Interneurons with high expression of *TTF1*, *LHX6*, *DLX1* and *DLX2* exhibit low

expression of cell cycle genes. **e**, Quantification of specific markers for interneuron progenitor cells and interneurons of PFC at GW7 by reverse transcription-PCR analysis. *n* = 3 independent replicates. **f**, *In situ* hybridization of *SST* in PFC at GW10. Scale bar, 100 μm. **g**, Immunostaining of *SST* in PFC at GW10. Scale bar, 100 μm. *n* = 3 independent replicates. **h**, GSEA enrichment plot of the KEGG neurotrophin signalling pathway. *n* = 663 excitatory neurons; *n* = 485 interneurons. **i**, Mean expression of Notch signalling pathway genes in NPCs, excitatory neurons and interneurons at different stages of development. The bar chart at the top represents the ratio of expressed genes to all genes in the Notch pathway.

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

► Experimental design

1. Sample size

Describe how sample size was determined.

We collected prefrontal cortices from embryonic stages available under the regulations by Reproductive Study Ethics Committee. To exclude batch effect for analysis, we collected two to three replicates for three stages of samples. Final dataset scale was determined according to the quality control criteria as described in the methods.

2. Data exclusions

Describe any data exclusions.

Cells detected with less than 1,000 genes and genes with normalized expression level less than 1 or expressed less than 3 single cells were removed out before initial clustering analysis.

3. Replication

Describe whether the experimental findings were reliably reproduced.

All replications were consistent for data results.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

The samples were allocated into each experimental groups based on the gestational stage. See methods 'Tissue sample collection and dissection'.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

The investigators were blinded to group allocation during data collection and analysis.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

TopHat(version 2.0.12) was applied to align clean reads to the hg19 human transcriptome downloaded from UCSC. The Seurat (version 1.2.1) package implemented in R was applied to identify major cell types. randomForest (version4.6.12) was applied to identify subtypes. The Monocle(version2.6.1) package was applied to analyze cell lineage developmental relationships. GSEA was applied to identify priori defined gene sets that show statistically significant differences between two given clusters.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No restrictions.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Description,supplier,catalog number,lot number,dilutions,species reactivity,application; Rabbit polyclonal to FOXP2,Abcam,ab16046,GR191926-1,1:500,"Mouse, Rat, Human", "WB, IHC-Fr,IHC-Fr, ICC/IF, IHC-P";
Mouse monoclonal [SATBA4B10] to SATB2,Abcam,ab51502,GR70015-8,1:250, "Mouse, Rat, Human, Zebrafish","IP, ICC/IF, IHC-Fr,IHC-Fr, IHC-P";
Mouse monoclonal [MEM-28] to CD45,Abcam,ab8216,GR302332-1,1:100,"Human","Flow Cyt, IP, IHC-P, WB, ICC/IF";
Goat polyclonal to AIF1,Abcam,ab5076, 1428158,1:200,"Rabbit, Guinea pig, Cow, Dog, Human, Pig, Common marmoset","Electron Microscopy, IHC-Fr, IHC-P, WB, IHC-FrFl, ICC, ICC/IF";
Rabbit monoclonal [EPR7003] to SFRP1,Abcam,ab126613,GR82085-19,1:500, "Human", "WB, IHC-P, ICC/IF";
Mouse monoclonal [1G10] to A2BP1 / RBFOX1,Abcam,ab183348,GR298959-2,1:500, "Mouse, Rat, Cow, Human","ICC/IF, WB";
Rabbit polyclonal to TTF1,Abcam,ab86023,785740,1:300,"Mouse, Human";"ICC/IF, WB";
Rabbit polyclonal to NEUROD2,Abcam,ab104430,GR94291-4,1:500,"Mouse, Human", "IHC-Fr, IP, IHC-P, WB";
Rabbit Polyclonal to PAX6,BioLegend,901301,B201255,1:500,"Human, Mouse, Mammalian","WB, IHC, IF";
Chicken polyclonal to EOMES, Millipore AB15894, 2697506, 1:500,"Mouse, Rat, Human, Avian"," WB, IHC";
Goat polyclonal to SOX2,Santa Cruz,sc-17320,H1406,1:250,"Mouse, Rat, Human", "WB, IP, IF, IHC, IHP, ELISA";
Mouse monoclonal to NEUROD1, Abcam,ab60704,GR3183945-2,1:100,"Mouse, Human, Apterodon leptorhynchus", "IHC-Fr, IHC-P, WB, ELISA, ICC/IF, Flow Cyt" ;
Rabbit monoclonal [EPR18114] to HMGA2, Abcam, ab207301, 1:100,"Human","IHC-P, ICC/IF" ;
Rabbit polyclonal to HOPX, Santa Cruz,sc-30216, D1615,1:1000,"Mouse, Rat, Human", "WB, IP, IF, IHC, IHP, ELISA";
Mouse monoclonal [B56] to Ki67, BD Biosciences, 550609, 19679, 1:100,"Human, Mouse, Rat, Rhesus", "Flow Cyt, IHC-Fr".

10. Eukaryotic cell lines

- a. State the source of each eukaryotic cell line used.
- b. Describe the method of cell line authentication used.
- c. Report whether the cell lines were tested for mycoplasma contamination.
- d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No eukaryotic cell lines were used.

No eukaryotic cell lines were used.

No eukaryotic cell lines were used.

No commonly misidentified cell lines were used.

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

- Provide details on animals and/or animal-derived materials used in the study.

No animals were used.

Policy information about [studies involving human research participants](#)

12. Description of human research participants

- Describe the covariate-relevant population characteristics of the human research participants.

Human pre-frontal cortex of gestational weeks 8-26 (GW 8, 9, 10, 12, 13, 16, 19, 23 and 26 weeks) were obtained from the aborted embryos under the agreement of puerperae.