DEVELOPMENTAL NEUROSCIENCE

Multiple parallel cell lineages in the developing mammalian cerebral cortex

Lucia Del-Valle-Anton[1], Salma Amin[1]†, Daniela Cimino[2], Florian Neuhaus[3], Elena Dvoretskova[3], Virginia Fernández[1], Yigit K. Babal[3], Cristina Garcia-Frigola[1], Anna Prieto-Colomina[1], Raquel Murcia-Ramón[1], Yuki Nomura[1], Adrián Cárdenas[1], Chao Feng[3], Juan Antonio Moreno-Bravo[1], Magdalena Götz[2], Christian Mayer[3], Víctor Borrell[1]*

Cortical neurogenesis follows a simple lineage: apical radial glia cells (RGCs) generate basal progenitors, and these produce neurons. How this occurs in species with expanded germinal zones and a folded cortex, such as human, remains unclear. We used single-cell RNA sequencing from individual cortical germinal zones in ferret and barcoded lineage tracking to determine the molecular diversity of progenitor cells and their lineages. We identified multiple RGC classes that initiate parallel lineages, converging onto a common class of newborn neuron. Parallel RGC classes and transcriptomic trajectories were repeated across germinal zones and conserved in ferret and human, but not in mouse. Neurons followed parallel differentiation trajectories in the gyrus and sulcus, with different expressions of human cortical malformation genes. Progenitor cell lineage multiplicity is conserved in the folded mammalian cerebral cortex.

INTRODUCTION

The cerebral cortex contains an extraordinary abundance and diversity of excitatory neuron (EN) types generated during embryonic development. Cortical neurogenesis is currently understood as a simple lineage process, as previously shown in mouse, where apical radial glia cells (aRGCs; also known as aRG or vRG) generate basal intermediate progenitor cells (IPCs) that produce neurons (*1*–*3*). Neuron abundance depends on the amplification of basal progenitors before neurogenesis, whereas neuron diversity is generated by variations of the aRGC population ground state over developmental time and across cortical space (*4*–*6*). However, unlike mouse, carnivores and primates have a substantial diversity of basal progenitor cell types with complex lineage relationships (*7*–*12*), which questions the simple lineage model. Here, we have addressed this fundamental question by analyzing the cortical germinal zones (GZs) at single-cell transcriptomic level with unprecedented layer resolution.

aRGCs are the primary type of cortical progenitor cell, constituting the ventricular zone (VZ), and basal progenitors constitute a secondary germinal layer, the subventricular zone (SVZ). In primates and carnivores, the SVZ is massively expanded and subdivided in inner and outer domains (ISVZ and OSVZ) (*10*, *13*). The OSVZ contains abundant IPCs and basal radial glia cells (bRGCs; also known as bRG or oRG), a second type of basal progenitor key for increased neuron production and cortex folding (Fig. 1A) (*10*, *12*, *14*). Single-cell RNA sequencing (scRNA-seq) studies have revealed the transcriptomic profile of these major progenitor cell classes in mouse and human cortex (*4*, *5*, *15*–*18*). However, molecularly characterizing the full diversity of cortical progenitors observed in primates and carnivores, as well as their lineage relationships, remains a major unresolved challenge.

Prior studies in ferret have reported bulk transcriptomic differences between cortical germinal layers (VZ, ISVZ, and OSVZ), between regions prospectively generating folds or fissures, and along developmental stages (*11*, *19*). These differences may reflect distinct compositions of progenitor cell classes and lineages (Fig. 1A). We therefore performed scRNA-seq analyses of individual microdissected cortical GZs in ferret to investigate their cellular diversity and infer the lineage relationships between identified cell classes. We find a diversity of RGCs in VZ and OSVZ greater than previously reported and their involvement in parallel neurogenic trajectories, both within and between these GZs. Using a progenitor cell barcoding approach, we confirm that these transcriptomic trajectories correspond to cell lineages in vivo. Parallel lineages converge onto a single major population of newborn neurons, in agreement with the notion that cortical neuron–type diversity is defined postmitotically (*4*, *17*, *20*). We observe similar traits in human and ferret cortex, but not in mouse, and propose that cortex development in species with the large and folded brain follows a multiple parallel lineage principle.

RESULTS

Single-cell resolution transcriptomic atlas of ferret cortical GZs

To elucidate the detailed complexity of cortical progenitor cell populations, we built a single-cell atlas of the ferret cortex GZs. Our sampling strategy was specifically designed to reveal features linked to cortex folding, analyzing separately two key GZs (VZ and OSVZ), two cortical regions [splenial gyrus (SG) and lateral sulcus (LS)], and two developmental stages that are critical for bRGC genesis and the initial formation of the OSVZ [embryonic day 34 (E34) and postnatal day 1 (P1); Fig. 1A] (*11*). Tissues were isolated individually from living brain slices by microdissection and processed for scRNA-seq (Fig. 1B). We analyzed each type of sample from three independent biological replicas (fig. S1), each replica profiled individually and containing cells from multiple embryos or one newborn ferret, thus producing an atlas that comprises the single-cell transcriptome of cortical GZs (Fig. 1C). Optimum quality controls for cell barcode filtering

[1]Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas & Universidad Miguel Hernández, Sant Joan d'Alacant 03550, Spain. [2]Institute of Stem Cell Research, Helmholtz Zentrum, Munich, Germany. [3]Max Planck Institute for Biological Intelligence, Munich, Germany.
*Corresponding author. Email: vborrell@umh.es
†Present address: Human Technopole, Milan, Italy.

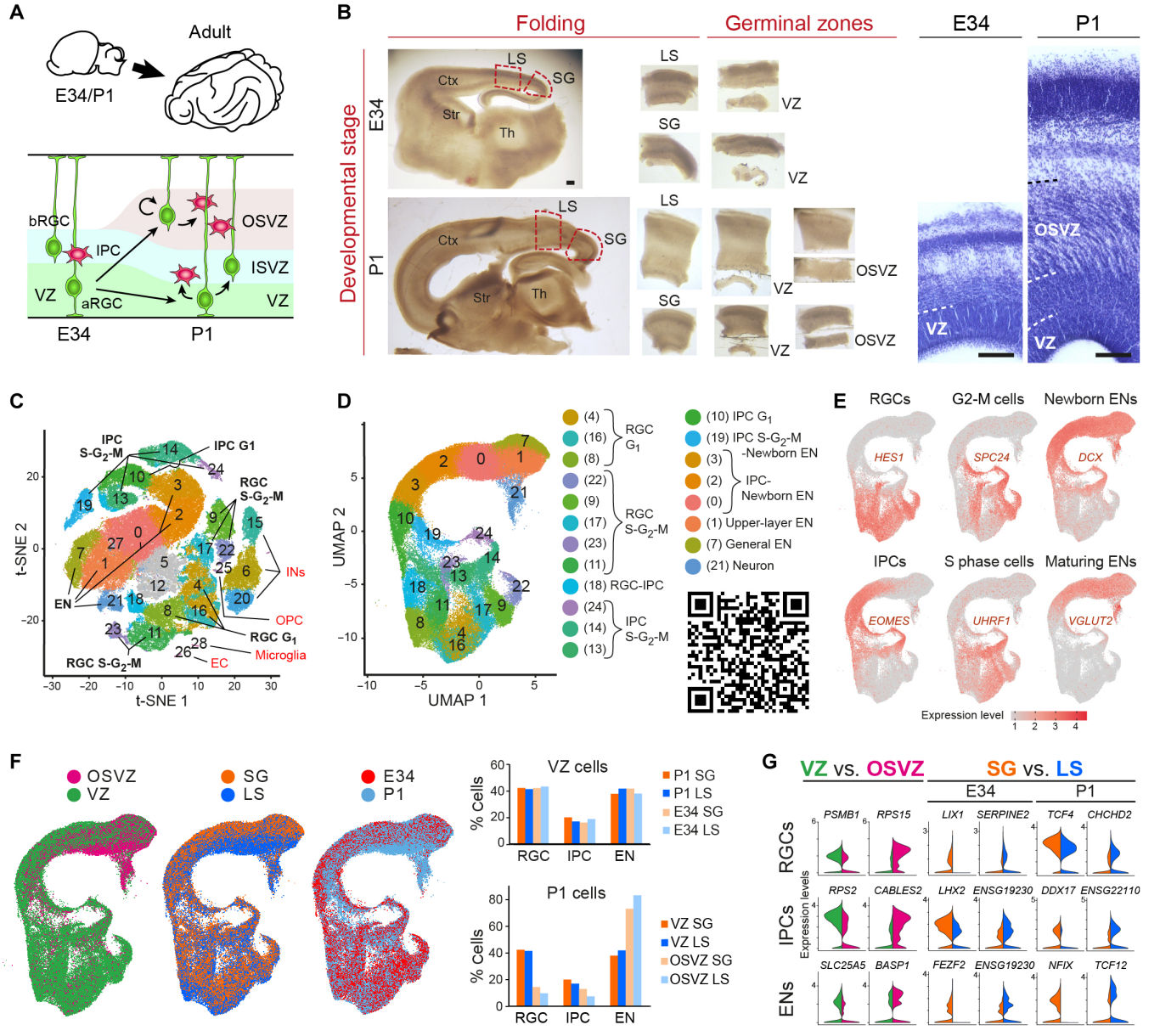

Fig. 1. Similar cellular composition of GZs across cortical areas and developmental stages. (**A**) Germinal layers and known major progenitor cell types and lineages in the developing ferret cortex. (**B**) Examples of live parasagittal tissue slice microdissections for scRNA-seq analyses and Nissl stains of the SG. Dashed lines indicate borders of GZs. Ctx, cortex; Th, thalamus; Str, striatum. Scale bars, 100 μm. (**C**) t-Distributed stochastic neighbor embedding (t-SNE) plot of the full dataset; clusters identified as in (D). Clusters in gray were only found in some litters and excluded. (**D** and **E**) Uniform Manifold Approximation and Projection (UMAP) and unbiased clustering of scRNA-seq data from cells in the EN lineage (D) and expression level of marker genes for the main cell classes (E). QR code, https://in.umh-csic.es/en/cortevo/, for access to expression of all transcripts. (**F**) UMAPs of cells from the indicated conditions and frequency of main cell types across conditions. (**G**) Violin plots for genes differentially expressed between conditions in the three main cell types. Adjusted *P* values range from 2.2×10^{-308} to 0.02 (see table S2).

(false discovery rate = 0.1%, iterations for Monte Carlo *P* value calculation = 100,000) and cell assessment (use of fixed and adaptive thresholds) were followed by analysis with Seurat-3 [with proven integration performance compared to other single-cell analytical methods (*21*, *22*)] and careful selection criteria for cluster resolution (evaluation of six potential resolutions from three different groups of progenitor cells; figs. S1 and S2) to lastly retain a total of 101,811 cells for further analyses (table S1).

A number of transcriptomically distinct cell clusters were identified corresponding to cell types involved in the lineage of ENs: RGCs, IPCs, and newborn ENs (Fig. 1C). Other, less abundant cell classes such as immature inhibitory neurons (INs), microglia, oligodendrocyte progenitor cells (OPCs), and endothelial cells (ECs) were discarded for further analyses (Fig. 1C and fig. S3). The transcriptome of progenitor cells is highly dynamic and oscillatory as they progress through the cell cycle (*23*, *24*). To minimize this effect

on cell clustering and identify bona fide distinct progenitor cell classes, we partially regressed cell cycle genes (fig. S2F). This distinguished cells in G_1 from those in S-G_2-M (Fig. 1C), which together with the expression of canonical marker genes let us identify multiple clusters of RGCs, IPCs, and ENs along a continuous transcriptional space (Fig. 1, D and E; see https://in.umh-csic.es/en/cortevo/ for the dataset of all transcripts). Despite the partial regression of cell cycle genes, their high expression still dominated the distinction of RGC and IPC clusters in S-G_2-M (fig. S3).

Cortical progenitor cells have different lineage behaviors between SG and LS (*10*) and between developmental stages (*11*) and have very different cell biology in VZ (aRGCs) and OSVZ (bRGCs; Fig. 1A) (*10*, *25*). However, all transcriptomic clusters of the main cortical cell types were found in both germinal layers, both cortical regions, and at both developmental stages (Fig. 1F and fig. S4, A and B). Even the relative abundance of the major cell classes (RGC, IPC, and EN) was similar between cortical regions and ages, with only RGCs exceeding in VZ compared to OSVZ and vice versa with ENs (Fig. 1F). We found only a reduced number of differentially expressed genes (DEGs) in RGCs, IPCs, and ENs between VZ and OSVZ and between SG and LS (Fig. 1G, figs. S5 and S6, and table S2).

Multiple RGC clusters are repeated in VZ and OSVZ

Despite the known differences in the biology of major cortical progenitor cell types between germinal layers, cortical regions, and developmental stages (*10*, *11*, *19*, *26*), these were not identifiable at the global transcriptomic level, so we next focused on individual cell clusters. The composition and cell abundance of IPC clusters were similar across all conditions, suggesting a remarkable similarity in IPC composition (fig. S4C and table S3). Even more notably, aRGCs and bRGCs were not identified as distinct cell clusters. Although we had microdissected VZ and OSVZ separately to allow us faithfully knowing which cells are presumptive aRGCs (VZ) and bRGCs (OSVZ), all RGC clusters were present in both layers (Fig. 2A). Similarly, all RGC clusters were present in SG and LS, as well as at E34 and P1. At the quantitative level, we found enriched abundance of cells belonging essentially only to clusters 8 and 16 (Fig. 2A and table S3). These corresponded to RGCs in G_1 and, thus, potentially were two different RGC kinds. In contrast, the third cluster of RGCs in G_1 (cluster 4) was similarly abundant across conditions (Fig. 2A), thus representing a third kind of RGCs that was common to all GZs, ages, and cortical regions. There was a twofold enrichment of cluster 8 cells in VZ and of cluster 16 cells in OSVZ (Fig. 2A and table S3), suggestive of these corresponding to aRGCs and bRGCs, respectively, although this was observed only in LS, not in SG. The distinction between clusters 16, 8, and 4 did not depend on the phase of cell cycle, as complete cell cycle regression did not alter their separation nor the clustering pattern of our entire dataset (fig. S7, A to C). This reinforced the notion that these were three distinct cell types, rather than distinct states of a single-cell type: RGα (cluster 16), RGβ (cluster 8), and RGγ (cluster 4).

To characterize in greater depth the heterogeneity of RGCs and distinguish aRGCs from bRGCs, we next subclustered RGCs from the primary clusters RGα (16) and RGβ (8) in our P1 samples (the only including both VZ and OSVZ). Cells from the primary clusters RGα (16) and RGβ (8) remained separate upon subclustering (Fig. 2B), confirming their distinct identity. Comparison between SG and LS showed that most subclusters were highly enriched in one or the other region, splitting in two parts each of the primary clusters RGα (16) and RGβ (8) (Fig. 2B and table S3). This revealed the existence of four major subclasses of RGCs: two enriched in LS (RGα_1 and RGβ_1) and two in SG (RGα_2 and RGβ_2; Fig. 2, B to D). These four transcriptomic subclasses still did not distinguish aRGCs from bRGCs, as most subclusters were similarly abundant in VZ and OSVZ, with only subcluster s6 being significantly enriched in this comparison (Fig. 2B). Subcluster s6 (part of RGβ_1) was highly enriched in VZ of LS [3.22 ± 0.05 $\log_2$ fold change (FC) compared to VZ of SG, $P < 0.0001$; 2.91 ± 0.20 $\log_2$FC compared to OSVZ of LS, $P = 0.011$] (Fig. 2B and table S3). This subcluster was a selective hot spot of *CRYAB* expression (Fig. 2, C, E, and F), a proposed marker of truncated RGCs (tRG) both in human and in ferret (*27*, *28*). This was in agreement with tRG being a VZ-specific cell type (Fig. 2, F and G, and fig. S6), thus representing a sixth RGC class (Fig. 2, C and D, and fig. S8A).

LS is located immediately rostral to SG, so the different abundance in tRG could simply reflect their slight difference in the anterior-posterior gradient of cortical maturation. A detailed analysis of tRG revealed that *CRYAB* mRNA levels in individual cells were much higher at P1 than E34, paralleled with a fourfold increase in the difference between LS and SG (Fig. 2K and table S3). The relative abundance of *CRYAB*$^+$ cells among RGCs in VZ was similar between LS and SG at E34, but threefold higher in LS at P1 (Fig. 2K and table S3). Anti–Alpha-crystalin B (CRYAB) immunostains confirmed the truncated morphology of tRGs, with a short basal process ending within the ISVZ, occasionally onto the blood vessel (Fig. 2, L and M). CRYAB-immunoreactive cells were virtually absent at E34 but abundant in newborn ferrets (P1 to P3) and highly enriched in LS compared to SG, in agreement with our transcriptomic results (Fig. 2, K to M). To further characterize tRGs in ferret, we sparsely labeled apical progenitors in vivo by green fluorescent protein (GFP) electroporation at P1 and examined those in both thick histological sections and clarified brains. Amid a majority of aRGCs spanning the full cortical thickness, tRGs displayed a short basal process ending in the ISVZ, frequently with a bend (Fig. 2, N and O, and fig. S8B).

To specifically identify bRGCs, we examined the expression of *HOPX*, commonly used as marker of this cell type (*5*, *12*, *15*, *27*, *29*, *30*). *HOPX* was expressed across all RGC subclusters derived from cluster 16, with a clear enrichment in VZ over OSVZ (Fig. 2, H to J). This expression pattern was similar to the human embryo cortex before gestational week 16 (GW16) (*15*), in line with prior studies questioning *HOPX* as a universal marker of bRGCs (*12*, *29*). Instead, *HOPX* expression labeled three RGC subclasses: RGα_1, RGα_2, and RGγ, present in both VZ and OSVZ (Fig. 2, C and J, and fig. S6). Together, we identified six major classes of RGCs (all in G_1): two enriched in LS (RGα_1 and RGβ_1), two enriched in SG (RGα_2 and RGβ_2), one class enriched in VZ of LS (tRG), and one class (RGγ) common to all conditions (Fig. 2D). These six cell populations were identified as transcriptomically different in both VZ and OSVZ (except for tRGs), while the two known different cell types aRGCs and bRGCs remained transcriptomically indistinguishable, supporting the identification of the former as different cell types.

To characterize the biological differences between SG- and LS-enriched RGC classes, we performed functional enrichment analysis (fig. S9). Compared to RGC classes abundant in LS (RGα_1 and RGβ_1), those in SG (RGα_2 and RGβ_2) were enriched in functions related to cell cycle progression, DNA replication, and cell division (fig. S9A), in agreement with the greater rates of cell proliferation previously observed in SG compared to LS (*10*, *31*). Within SG and LS, RGα was enriched in mechanisms promoting cell cycle progression and cell

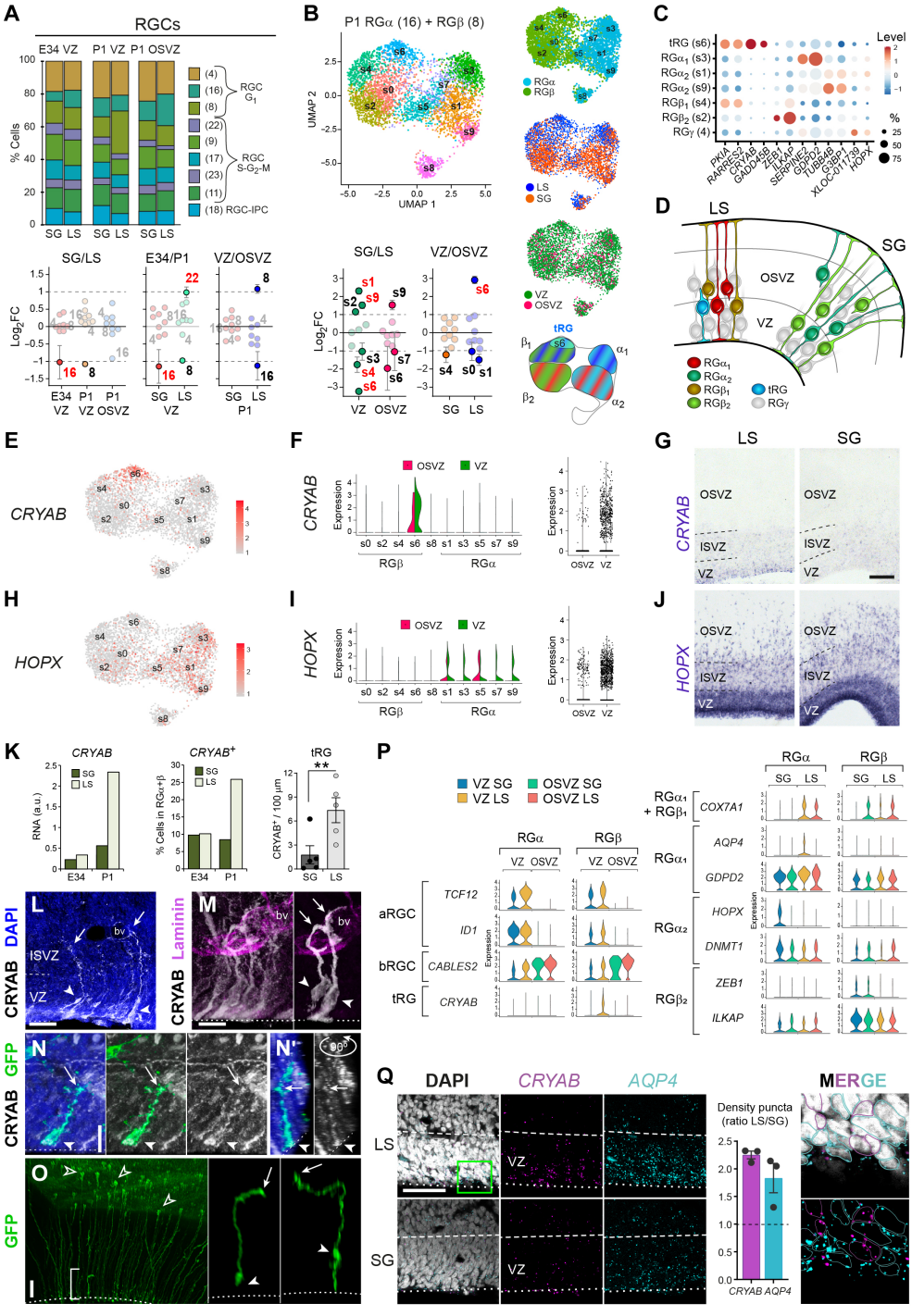

Fig. 2. Multiplicity of RGC types across GZs and cortical areas. (**A**) Frequency distribution of RGC clusters across conditions, and pair-wise cluster enrichment. Means ± SEM for clusters above ±1 $\log_2$FC is indicated (red indicates adjusted $P < 0.05$). (**B**) UMAP of P1 cells from clusters 16 (RGα) and 8 (RGβ) after unbiased subclustering, UMAPs of indicated features, pair-wise subcluster enrichment, and schema of RGC subclass distribution in UMAPs. (**C**) Gene expression across cell classes at P1 (subclusters indicated). (**D**) Subclasses of ferret RGCs in G_1 identified at P1, duplicated in VZ and OSVZ. (**E** to **J**) UMAPs, violin plots and ISH stains of *CRYAB* and *HOPX* expression in RGC subclusters and germinal layers at P1. (**K**) Mean *CRYAB* RNA levels, abundance (RNA), and density (protein; P3) of CRYAB$^+$ cells (means ± SEM; $^{**}P < 0.01$, *t* test). a.u., arbitrary units. (**L** to **N'**) tRGs in VZ of P1 (L and M) and P3 (N and N') ferrets identified with anti-CRYAB immunostains and morphology confirmed by GFP (N). bv, blood vessels. Arrows indicate the tip of the basal process, arrowheads indicate the soma, dotted lines indicate the apical VZ surface. (N') shows orthogonal view. (**O**) GFP-labeled RGCs in clarified P2 ferret brains, with tRGs shown in detail. Open arrowheads indicate basal endfoot arbors of aRGCs. (**P**) Violin plots for marker genes in RGC classes between conditions. *P* values range from 2.2×10^{-308} to 0.01 (see table S4). *ILKAP*, *GDPD2*, and *DNMT1* are top markers for RGβ and RGα. *HOPX* distribution defines RGα$_2$. (**Q**) RNAscope coexpression analysis at P1 for RGC marker genes between conditions. Green box in indicates area shown at high magnification in a single-plane confocal image. Cells expressing *CRYAB* or *AQP4* are indicated with colored circles. Scale bars, 100 µm (G and J), 60 µm (Q), 30 µm (L and O), 10 µm (M and N).

proliferation (amplificative RGCs), whereas RGβ was enriched in genes related to cell differentiation (neurogenic or differentiative RGCs; fig. S9, B and C).

Given that aRGCs and bRGCs were indistinguishable at global transcriptomic level, we lastly focused on the analysis of specific genes. Differential expression analysis between GZs revealed that genes preferentially expressed in aRGCs (VZ) were related to enhancing progenitor cell amplification, via regulation of transcription (*TCF12*, *ID1*, and *HMGN1*), mitochondrial function (*MICU2*), and cell proliferation (*YPEL3*). In contrast, genes preferentially expressed in bRGCs (OSVZ) were related to promotion of neurogenesis, via oxidative phosphorylation (*ATP5I*), negative regulation of cell cycle (*CABLES2*), and ribosomal activity (*RPS15*). Specific RGC classes were characterized by enriched expression of few genes, including *COX7A1*, *PKIA* (RGβ$_1$), *ZEB1*, *ILKAP* (RGβ$_2$), *AQP4, GDPD2* (RGα$_1$), *SNRPA1*, *DNMT1*, *HOPX* (RGβ$_2$), and *CRYAB* (tRG) (Fig. 2, C and P, and figs. S10 and S11). Notably, we also identified Growth Arrest And DNA Damage Inducible Beta (GADD45B) and Transmembrane Protein 196 (TMEM196) as previously unidentified tRG markers (fig. S11). The identity of cell classes as defined by these genes was confirmed by fluorescence in situ hybridization (FISH) analyses by means of RNAscope (Fig. 2Q and figs. S12 and S13).

Together, our results indicated the existence of six major classes of RGCs in the germinal layers of the developing ferret: three characteristic of LS (RGα$_1$, RGβ$_1$, and tRG), two of SG (RGα$_2$ and RGβ$_2$), and one ubiquitous or general (RGγ). RGC classes were distinct in their amplificative (RGα) versus differentiative (RGβ) profile and populated VZ and OSVZ similarly (except for tRGs), indicating that they existed in duplicate as aRGCs and bRGCs (Fig. 2D). *HOPX* marked three of the RGC classes common to VZ and OSVZ, so in ferret, it was not specific to bRGCs.

Three parallel RGC trajectories duplicated in VZ and OSVZ

Next, we used Monocle-3 to perform a transcriptomic trajectory analysis of our single-cell dataset (Fig. 3A). We identified three parallel trajectories from mitotic RGCs to ENs. The roots of these trajectories were two distinct classes of dividing RGCs, each including two transcriptomic cell clusters: mitotic RGCs class A (clusters 23 + 11)

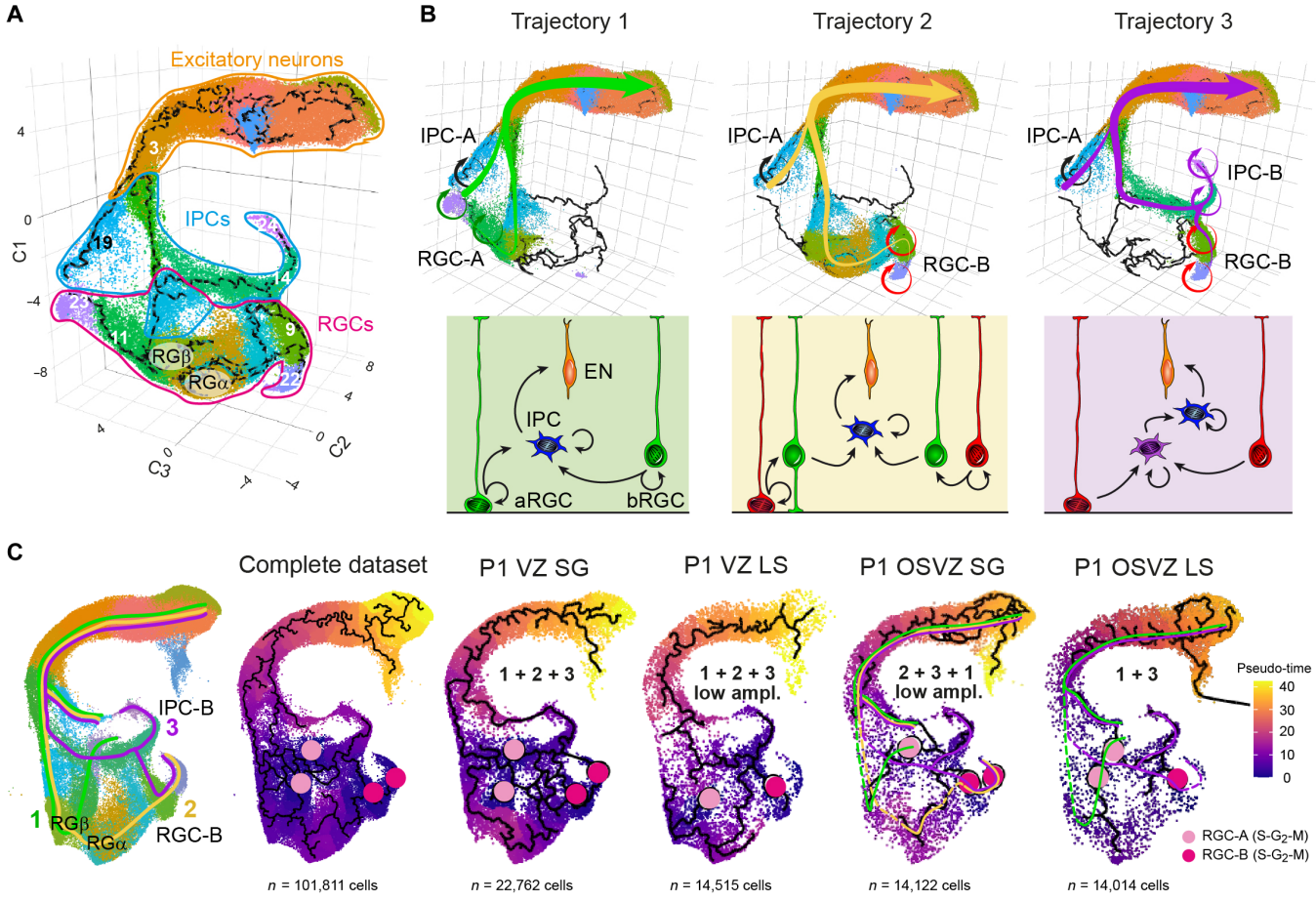

Fig. 3. Parallel cell trajectories converge into a single class of newborn neurons. (**A**) Three-dimensional (3D) UMAP and cell trajectories of the clustered scRNA-seq ferret dataset. Clustering is the same as in Fig. 1D and in (C). Numbers and greek letters identify key clusters. (**B**) 3D UMAPs showing the clusters involved in each of the three trajectories identified, each starting from a set of mitotic RGC clusters, and schematic drawings of summarized cell lineages. Circular arrows indicate mitotic cell clusters. (**C**) 2D UMAP of cell clusters indicating the three trajectories identified in the 3D UMAP [color-coded as in (B)], and pseudo-time and trajectory plots of the complete ferret dataset and individual conditions, indicating the RGCs of origin (pink and red circles) and trajectories present in each case (indicated with colored lines in OSVZ). The total number of cells in each condition is indicated.

and mitotic RGCs class B (clusters 22 + 9). Trajectory 1 stemmed from mitotic RGC-As, followed by RGβ (cluster 8), and then IPCs that transited through mitosis (IPC class A, cluster 19) to produce ENs (Fig. 3B). Trajectory 2 stemmed from mitotic RGC-Bs, leading to RGγ (cluster 4) and/or RGα (cluster 16), followed by RGβ, and then IPCs that lastly divided (IPC-A) producing ENs (Fig. 3B). Trajectory 3 stemmed again from RGC-Bs but now giving rise to mitotic IPC-Bs (clusters 24 + 14) and then the common IPCs that lastly divided (IPC-A) to generate ENs (Fig. 3B). The identification of these three trajectories did not depend on the cell cycle, as analysis of the same dataset but subject to complete cell cycle regression yielded the same result (fig. S7D). Each of the three trajectories had a distinctive pseudo-temporal transcriptomic signature at the level of progenitor cells. Trajectory 1 was characterized by expression of *PTTG1*, *TOX3*, and *LGALS1*; trajectory 2 was characterized by expression of *SLC1A3* (canonical RGC marker), *ZFP36L1*, and *SPARC*; and trajectory 3 was characterized by expression of *TOP2A*, *NSL1*, and *SPC25* (fig. S14 and table S5). All three trajectories converged onto a single cluster of ENs (cluster 3; Fig. 3A), suggesting that they are not specializations to produce distinct neuron types but parallel lineages producing a single type of newborn neuron.

The three trajectories occurred in VZ, both at E34 and P1, and in SG and LS (Fig. 3C and fig. S15). However, the trajectory in P1 VZ LS involved fewer mitotic RGC clusters, suggesting lower RGC amplification in agreement with previous histological studies (*10*). The same three trajectories were also traced in the OSVZ of SG, whereas LS displayed only trajectories 1 and 3 (Fig. 3C and fig. S15). This was again consistent with previous analyses showing that OSVZ progenitors are more proliferative in SG than LS (*10*). Together, our observations indicated that VZ and OSVZ of the developing ferret cortex contain three parallel transcriptional trajectories of apical and basal progenitor cells that converge onto a single transcriptomic class of newborn EN.

Parallel RGC lineages in the ferret cortex

To confirm the validity of the three observed transcriptomic trajectories as distinct cell lineages in ferret, we used the barcode lineage tracing technique TrackerSeq (*32*). This method involves integrating DNA barcodes into the genome of electroporated RGCs, allowing us to track the clonal relationships with their progeny over time (Fig. 4A). We used in utero electroporation of E34 ferret embryos to deliver the TrackerSeq barcode library to apical progenitors in the VZ of the cerebral cortex (putative RGCs). We allowed the transduced cells to develop a cell lineage for 3 days in vivo, then collected them by fluorescence-activated cell sorting (FACS) from a cell suspension and created single-cell sequencing libraries, for both transcriptomes and lineage barcodes (Fig. 4A and fig. S16). We recovered a total of 1038 multicell clones, composed of both mitotic and postmitotic cells (Fig. 4, B to F, and fig. S16), as expected from the timing of our experiment and consistent with previous work (*32*). In addition, in agreement with previous findings in mice (*32*), our results suggest that cortical RGCs in ferret only generate ENs, not inhibitory interneurons.

To examine these cell lineage results in the context of our previous extensive characterization of cell clusters, we annotated clusters in the TrackerSeq dataset based on the scRNA-seq data described above, using correlation-based label transfer (*33*). A majority of clones contained cells assigned to the upper-layer EN clusters (cluster 1: 472 clones, 636 cells), while much fewer contained RGα (cluster 16: 19 clones, 19 cells) or mitotic IPC-Bs (clusters 14 and 24: 38 clones, 39 cells) (Fig. 4E and fig. S16). We then evaluated the probability of identifying shared clone IDs among all pairs of cell clusters and quantified lineage coupling by calculating a *z* score for clone counts relative to a random distribution (see Materials and Methods) (*34*). These *z* scores revealed groups of cell clusters with similar lineage coupling signatures (Fig. 4G). By visualizing lineage coupling scores as a network in which nodes represent cell clusters, we were able to validate clonal lineages for our three inferred trajectories (Fig. 4, F and H, and fig. S16). Specifically, mitotic RGC-As (clusters 23 and 11) were coupled with RGβ (cluster 8) and IPC-A (cluster 19), which largely recovered trajectory 1. The second population of cell clusters that exhibited clonal coupling was composed of mitotic RGC-Bs (clusters 22 and 9) and cluster 17, then RGα (cluster 16), RGβ (cluster 8), and RGγ (cluster 4), all RGC in G_1, followed by RGC-IPC cluster 18 and IPC cluster 10, and lastly postmitotic IPCs-EN clusters 0, 1, and 3, which nicely recovered trajectory 2. We also observed clones containing RGC-Bs (clusters 22 and 9), coupled to mitotic IPC-Bs (clusters 24 and 14), then mitotic IPC-A (cluster 19), and ENs, recovering trajectory 3. Hence, the three transcriptomic trajectories were validated as cell lineages in vivo (Fig. 4, F and H). Notably, we further observed robust coupling scores directly between mitotic RGCs and mature ENs, but without IPCs, consistent with direct neurogenesis (Fig. 4H). This was observed in 4.3% of clones containing RGC-As (cluster 23) and 8.8% of clones containing RGC-Bs (clusters 9 and 22). Clones without IPCs were smaller (mean, 2.39; maximum, 4 cells) than clones containing IPCs (mean, 2.86 cells; maximum, 8 cells; Wilcoxon test, $P = 0.0452$). These maximum clone sizes were exactly the predicted output of direct and indirect neurogenesis, respectively, after three cell mitotic cell cycles. This supported the existence of direct neurogenesis in ferret, potentially representing yet additional parallel RGC lineages.

We then investigated whether the observed lineages starting from RGC-A and RGC-B give rise to transcriptomically distinct ENs. We grouped individual ENs based on their coupling to either RGC-A or RGC-B and performed a principal components (PC) analysis (PCA) and differential gene expression analysis to assess heterogeneity among these groups of ENs. We found no difference in PC space between ENs originating from distinct RGC populations (Fig. 4, I and J). Whereas we did identify some DEGs between these EN populations, their number was not significantly greater than expected by chance (Fig. 4K). Thus, our in vivo analyses of cell clones validated our above conclusions based on inferred cell trajectories: parallel RGC lineages coexist in the developing ferret cerebral cortex and converge onto a single transcriptomic class of early postmitotic neuron.

Multiple RGC classes are conserved in ferret and human but not in mouse

The unprecedented diversity of cortical RGC classes and lineages observed in ferret prompted us to investigate their conservation in species with a cortex large and folded (i.e., human) versus small and smooth (i.e., mouse). We integrated our ferret datasets with public scRNA-seq datasets from human and mouse cortex at equivalent stages of cortical development (Fig. 5A) (*16*, *18*). The integrated dataset cleanly clustered RGCs (*HES1*$^+$), IPCs (*EOMES*$^+$), ENs (*VGLUT2*$^+$), and inhibitory interneurons (*GAD65*$^+$) from the three species (Fig. 5, B and C, and fig. S17). Previous single-cell studies of the embryonic human cortex identify aRGCs and bRGCs based on presumed transcriptomic signatures, so we used cross-species comparison to identify these cell classes in ferret and mouse. Cells identified as vRGs (aRGCs)

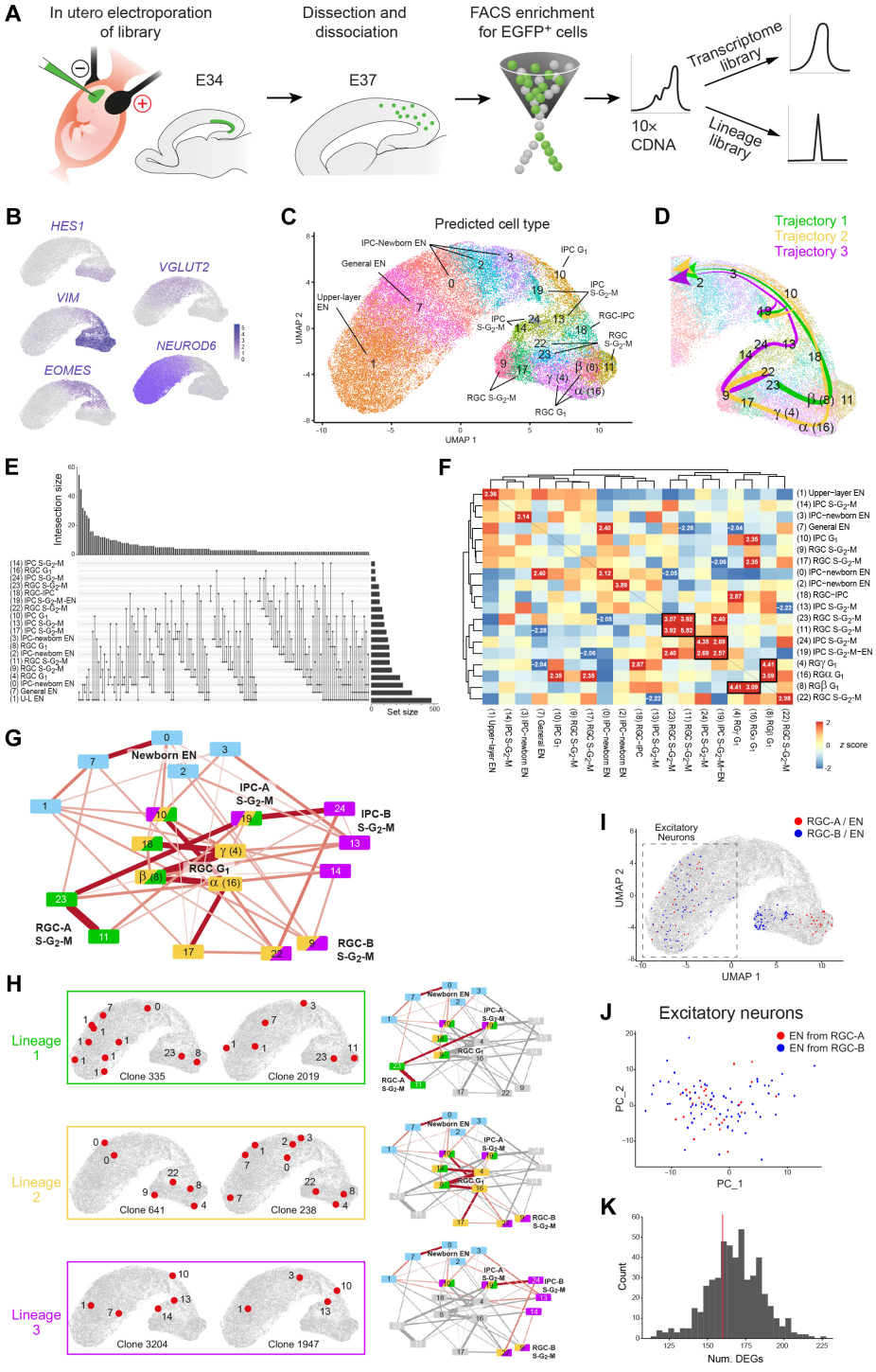

Fig. 4. Lineage relationships in the developing ferret cortex. (**A**) Lineage barcodes were delivered by electroporation at E34; by E37, EGFP-positive cells were extracted; and lineage barcode and transcriptome libraries were sequenced. (**B**) UMAP embedding of all sampled cells with expression of known marker genes (*HES1/VIM*: RGC; *EOMES*: IPC; *VGLUT2* and *NEUROD6*: EN). (**C**) Cell class annotation after transferring previous labels using correlation-based label transfer. (**D**) Arrangement of cell trajectories identified previously on the UMAP of the new library. (**E**) UpSet plot showing observed combinations of clones in different cell clusters, size of the interaction (top), and absolute number of clones per cell cluster (right). (**F**) Heatmap showing lineage coupling correlation scores between all cell clusters. *z* scores larger than ±2 are more than 2 SDs from random, hence high-confidence interactions. (**G**) Cell cluster network visualization, with edges representing lineage coupling correlation scores. Position of nodes was set manually. Interactions shown are those with *z* score > 0.5. (**H**) Position of cells in the UMAP embedding of example clones, as well as coupling subnetwork of clusters involved, for each of the three lineages. Interactions shown are those with a *z* score > 0.5. (**I**) UMAP highlighting cells from clones containing only ENs plus RGC-A (clusters 23 + 11) or ENs plus RGC-B (clusters 22 + 9). (**J**) PCA of ENs from (I) color-coded according to their clone type of origin. (**K**) Number of DEGs between EN types as in (J) (red line), and frequency distribution of DEGs following random distribution of clone cells (500 repetitions; gray bars).

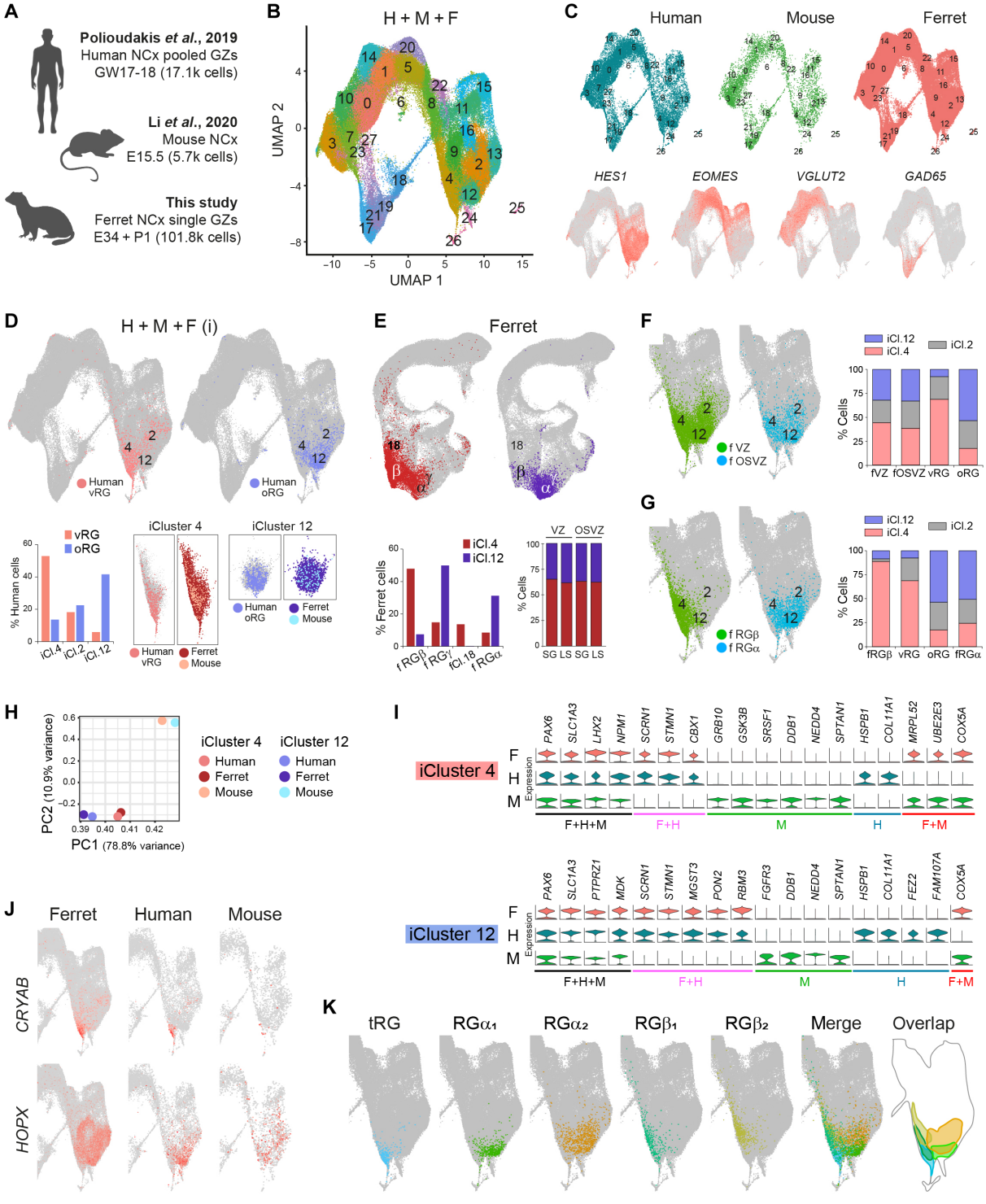

Fig. 5. Conserved RGC classes in ferret and human. (**A**) Sources of integrated scRNA-seq datasets (*16, 18*). NCx, neocortex. (**B**) UMAP of integrated human (H), mouse (M), and ferret (F) datasets and unbiased cell clustering. (**C**) Distribution of cells from each species within the integrated UMAP and expression of genes identifying major cell classes. (**D**) Integrated UMAP highlighting human cells identified as vRG and oRG in (*16*), frequency distribution in the integrated clusters, and subset of ferret and mouse cells belonging to the same integrated cluster. (**E**) UMAP of ferret dataset highlighting cells from iCl.4 (red) and iCl.12 (purple) and their frequency distributions among ferret clusters (f) and P1 sample types. (**F** and **G**) Details from the integrated UMAP highlighting P1 ferret G_1 RGCs collected from VZ and OSVZ or belonging to fRGα and fRGβ and their frequency distributions in the indicated integrated cluster compared to human vRG and oRG. (**H**) Reactome-based PCA of iCl.4 and iCl.12 in the three species. (**I**) Violin plots for genes expressed in the indicated integrated cluster in each species. Adjusted *P* values range from 2.2×10^{-308} to 0.002 (see table S6). (**J**) Integrated UMAPs (detail) of *CRYAB* and *HOPX* expression in the indicated species. (**K**) Individual and merged distribution of RGC classes in the integrated UMAPs (detail) and schematic of their overlap.

in (*16*) were mostly part of the integrated cluster 4 (iCl.4) and were largely segregated from those identified as oRGs (bRGCs), found mostly in iCl.12 (Fig. 5D and fig. S17B). Ferret cells from each of these integrated clusters were then identified in the Uniform Manifold Approximation and Projection (UMAP) of cells only from ferret, where they were found also segregated: iCl.4 cells mostly corresponded to the ferret fRGβ, and iCl.12 cells corresponded to fRGγ and fRGα (Fig. 5E). Although cells in iCl.4 and iCl.12 were presumed to correspond to aRGCs and bRGCs, respectively, this was not the case in ferret, as both clusters were present and similarly abundant in VZ and OSVZ (Fig. 5E).

To identify the cell clusters corresponding to aRGCs and bRGCs, we located RGCs in G_1 from ferret VZ and OSVZ in the species integrated UMAP. Ferret RGCs from VZ and OSVZ had a very similar contribution to iCl.12, iCl.2, and iCl.4 (Fig. 5F), demonstrating that these integrated RGC clusters do not separate aRGCs from bRGCs. This cell distribution across transcriptomic clusters was completely different from that of vRG and oRG cells as identified in the original human study (Fig. 5F), questioning the identity of the latter as aRGCs and bRGCs. Because ferret clusters fRGβ and fRGα distinguish two classes of RGCs, we next studied their distribution within the species integrated dataset. We found that the distribution of fRGβ and fRGα across clusters iCl.12, iCl.2, and iCl.4 was similar to that of human vRG and oRG cells, respectively (Fig. 5G). This led to the notion that human cells currently defined as vRGs and oRGs are in fact the two distinct RGα and RGβ found in ferret, present both in VZ and OSVZ. This notion was confirmed with the analysis of marker gene expression. In addition to *HOPX*, a number of genes have been used as canonical markers for human aRGC or bRGC (*5*, *15*, *16*). In our ferret samples, expression of these genes correlated mostly with RGC class (cluster) rather than with their layer of residence (fig. S18).

Reactome-based analyses indicated that cells within iCl.4 and iCl.12 were functionally much more similar between human and ferret than with mouse (Fig. 5H and fig. S17C). Interspecies differences involved signaling pathways regulating progenitor cell self-amplification and neurogenesis (i.e., *NOTCH*, *ROBO*, and *FGF*) and cell migration (*ERBB2* and *NETRIN*). In these integrated RGC clusters, canonical RGC genes were expressed similarly in the three species, but other genes were selectively expressed in human and ferret but not in mouse, or vice versa, in addition to some human- and ferret-specific genes (Fig. 5I, fig. S17D, and table S6). For example, high expression of *STMN1* in human and ferret iCl.4 and iCl.12 compared to that in mouse correlated with the greater self-renewal of the former (*35*). The patterns of expression of *CRYAB* and *HOPX* further highlighted differences in progenitor cell populations between human and ferret versus mouse cortex (Fig. 5J). Analysis of location of the multiple ferret RGC populations that we identified above, within the integrated multispecies dataset, showed very distinct distributions with minimal overlap between RGC classes (Fig. 5K). Together, our results strongly suggested that the embryonic human cerebral cortex contains a similar variety of RGC types as ferret (RGα, RGβ, and tRG) and that these may also be duplicated in VZ and OSVZ.

Parallel progenitor cell trajectories are conserved in ferret and human

To analyze the conservation of cell lineages across ferret, human, and mouse, we focused on cells from the glutamatergic lineage (Fig. 6A). All integrated RGC clusters were present in the three species, and some were highly enriched (two- to fourfold) in ferret and human compared to those in mouse, or vice versa (Fig. 6, B and C, and table S3). Within the species integrated dataset, we identified the ferret mitotic RGCs and IPC-A and IPC-B and reconstructed their three parallel trajectories as identified above (Fig. 6, D to F). Trajectory analysis of ferret cells alone, as well as human cells alone within the integrated UMAP, revealed strong conservation between human and ferret, containing the same mitotic RGC clusters of origin and largely the same three main trajectories (Fig. 6, G and H). In contrast, analysis of mouse cells alone within the integrated UMAP only identified the presence of one mitotic RGC-B cluster (corresponding to cluster 22) that initiated only trajectory 2. Monocle did not identify the other two trajectories nor the other two mitotic origins (cluster 17/RGC-A; cluster 23/RGC-B) that were at very low abundance in mouse (Fig. 6, C, G, and H). To test whether differences between ferret and human with mouse were of technical origin due to their very different sample sizes, we repeated these analyses after equalizing the three species for cell numbers. After downsampling the ferret and human datasets with two different methods, we still found a similar enrichment of specific RGC clusters in those versus mouse (fig. S19). However, the complexity of the ferret and human transcriptomic trajectories across the dataset was lost, while in mouse, it remained unchanged (fig. S19). Hence, downsampling to equalize cell numbers across species did not improve the characterization of cell populations nor their trajectories. Together, our analyses indicated a marked conservation of RGC classes between human and ferret, but less in mouse. Concomitantly, three main trajectories of progenitor cells were found in human and ferret, but our dataset only revealed one in mouse. Our analyses confirmed that the transcriptomic profile of aRGCs and bRGCs is remarkably similar, such that cell populations previously identified as aRGCs and bRGCs in human datasets may rather be distinct populations of RGCs that coexist in both VZ and OSVZ.

Distinct maturation of newborn neurons in gyrus and sulcus linked to human cortical disease

Our analyses of microdissected GZs in ferret identified newborn ENs in addition to progenitor cells, in agreement with histological evidence (*26*). The multiple parallel trajectories and lineages of progenitor cells identified above did not derive into early neuronal diversification, as might be expected, but converged onto a single class of IPC–newborn ENs. This was in agreement with the notion that cortical neuron–type diversity is defined postmitotically (*4*, *17*, *20*). Subclustering of ferret late IPCs and newborn ENs distinguished two main cluster groups, which we identified as "immature lower layer" and "mature upper layer" neurons, owing to the expression of marker genes (Fig. 7A and fig. S20). This was unexpected because at E34 and P1, only upper layer neurons are born (*36*). Pseudo-time analysis revealed that transcriptomic cell maturation progresses from "lower layer" to "upper layer" profiles, as in normal cortex development and consistent with the former being enriched in VZ and the later enriched in OSVZ (Fig. 7, B and C). This indicated that cortical neurons born at late stages and destined to upper cortical layers follow an early maturation process that recapitulates the transcriptomic progression of cortical neurogenesis. Many EN subclusters were differentially enriched in SG or LS, suggesting that distinct neuron subtypes are produced in these closely related cortical areas (Fig. 7, D and E, and table S2). Cell clusters enriched in the gyrus or sulcus segregated orthogonal to the pseudo-time trajectory, thus following parallel trajectories of maturation (Fig. 7C).

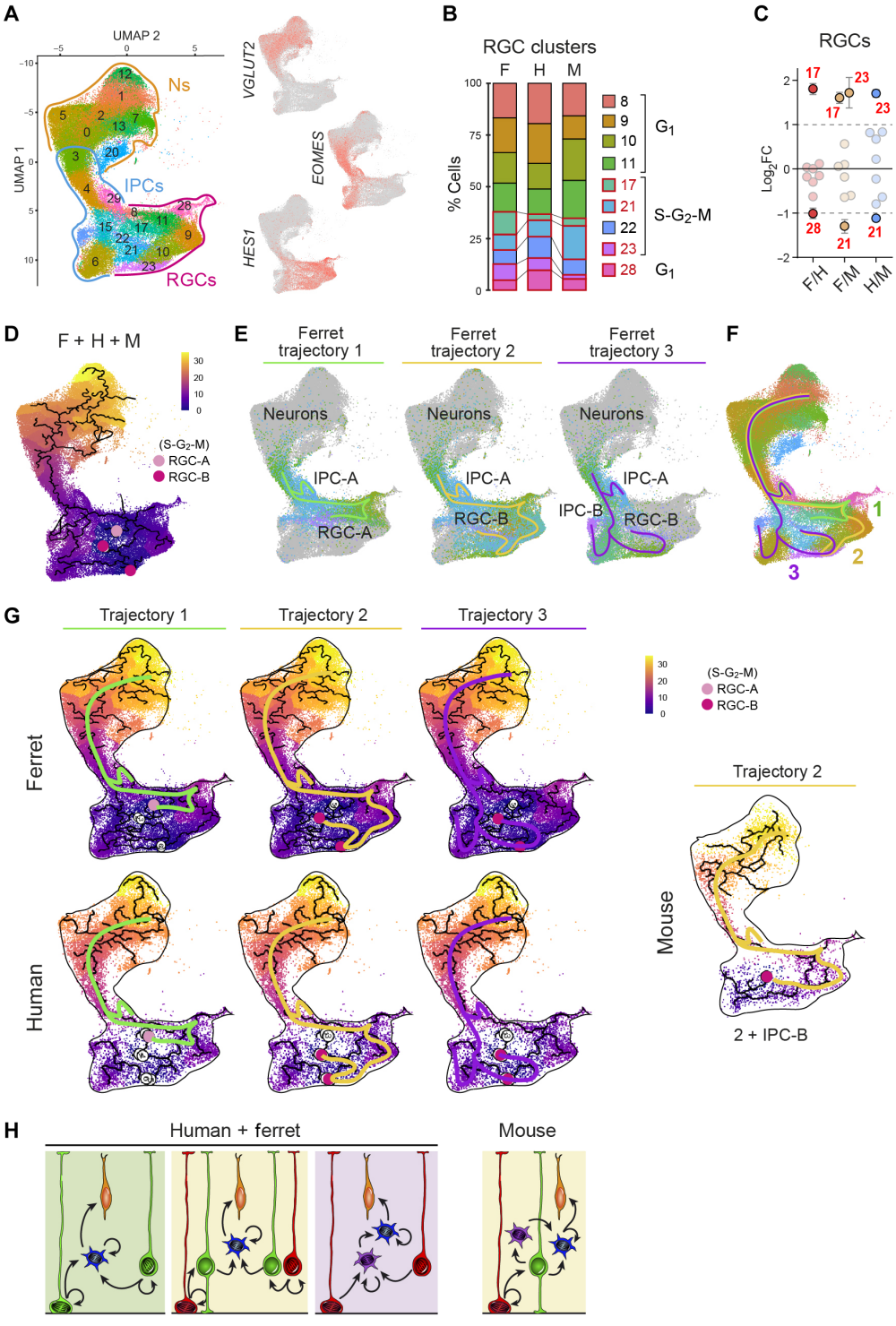

Fig. 6. Conserved and divergent RGC classes in ferret, human, and mouse. (**A**) UMAP of integrated human, mouse, and ferret (P1) datasets with unbiased cell clustering (showing only cells in the EN lineage) and expression level of marker genes for major cell types. (**B** and **C**) Frequency distribution of integrated RGC clusters within species (B) and pair-wise cluster enrichment between species (C). Clusters above ±2 FC are highlighted (means ± SEM). Red font indicates statistical significance (adjusted $P < 0.05$; see table S3). (**D**) Pseudo-time and trajectory plot for the multispecies integrated UMAP, indicating the mitotic RGCs of origin (pink and red circles). (**E** and **F**) UMAPs highlighting the progenitor cell clusters involved in each of the three trajectories (colored lines) identified in the integrated dataset (E) and summary of trajectories (F). (**G**) Pseudo-time and trajectory plots for ferret, human, and mouse cells within the integrated UMAP. All three trajectories existed in ferret and human, but only one in mouse. (**H**) Schematic drawings of summarized cell lineages found in human and ferret versus mouse.

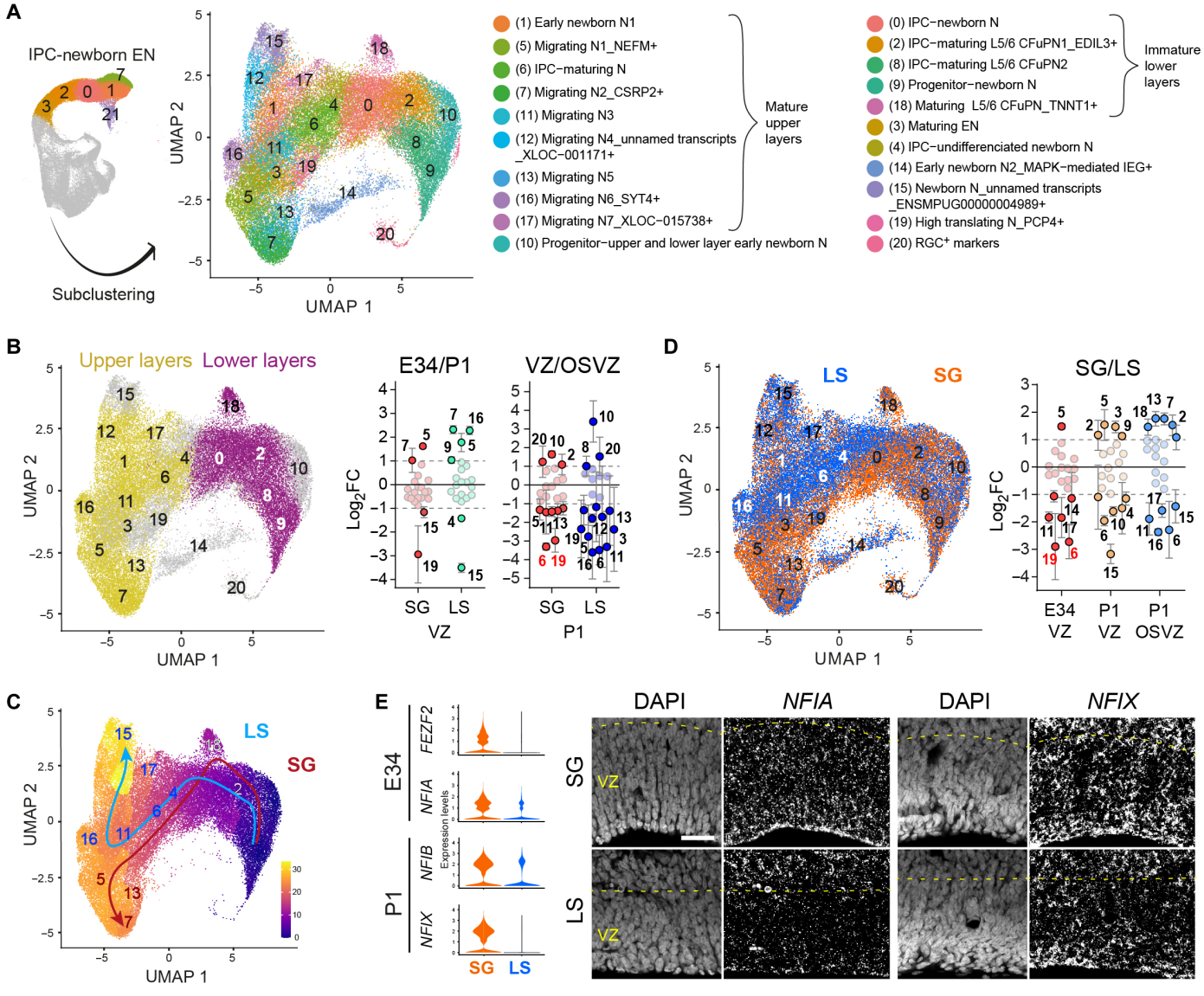

Fig. 7. Different newborn neuron classes in the gyrus and sulcus. (**A**) UMAP of subclustered IPC-newborn neurons from the ferret dataset. (**B** to **D**) UMAPs highlighting mature upper layers and immature lower layers cluster sets (B), pseudo-time (C), and cells from LS versus SG (D) and pair-wise cluster enrichment between conditions (B and D). Clusters above ±2 FC are indicated (means ± SEM); red font indicates statistical significance (adjusted $P < 0.05$; see table S3). In (C), cluster numbers indicated in blue are enriched in LS, and those in red are enriched in SG; arrows indicate estimated pseudo-time trajectory of cells from each region. (**E**) Violin plots for DEGs between immature neurons in SG and LS and RNAscope analysis at the corresponding ages. Adjusted *P* values range from 2×10^{-139} to 6.5×10^{-25} (see table S2). Scale bar, 30 μm.

Human malformations of cortical development (MCDs) are caused by mutations in developmentally relevant genes, but their expression and specific functions in normal development remain under debate (*37*). Analysis of our datasets revealed high levels of expression of MCD genes in ENs and mitotic RGCs (fig. S21, A and B). Expression of MCDs was higher in SG than LS in P1 VZ (fig. S21C), supporting their executive role in cortex folding. Genes linked to microcephaly were most highly expressed by RGCs and IPCs in mitosis (not in G_1), in agreement with their key role in mitotic spindle and neurogenesis. Genes linked to polymicrogyria and lissencephaly (LIS) were expressed across neuronal clusters, but at higher levels in SG compared to LS, in agreement with their relevance in cortex folding. This difference was especially prominent in lower-layer neurons regarding LIS genes (fig. S21, D and E).

DISCUSSION

The large size and intricate folding of the human cerebral cortex depend on the extraordinary proliferation and neuron production of its neural progenitor cells during development. To characterize the full diversity of cortical progenitor cell classes, their lineage relationships, and the phylogenetic conservation of these features, we have analyzed single-cell transcriptomes of the developing cerebral cortex in ferret, an established animal model of cortical expansion and folding. The unprecedented spatial resolution of our sampling, where individual

germinal layers from identified cortical areas were profiled separately, provided an unmatched power of cell class discrimination.

Expanded diversity of cortical progenitor cells

Previous work identified three main types of RGCs in the mammalian cerebral cortex: aRGCs and tRGs in VZ and bRGCs in OSVZ (*7*–*10*, *27*, *38*). In primates, bRGCs are further subdivided in several morphotypes, and evidence in ferret suggests the existence of a variety of aRGCs (*11*) and tRGs (*28*). Despite recent advances, this RGC diversity has remained elusive at the transcriptomic level. Our present analysis reveals the existence of six different transcriptomic classes of RGCs in VZ (RGα_1, RGα_2, RGβ_1, RGβ_2, RGγ, and tRG), five of which are repeated in OSVZ (Fig. 8A). aRGCs remain indistinguishable from bRGCs, which indicates that transcriptomic differences between the six classes of RGCs here identified are larger than between classical aRGCs and bRGCs, strongly supporting that the former correspond to bona fide different types of RGCs.

The integration of our ferret single-cell dataset with public datasets from mouse and human reveals a much greater transcriptional distance among progenitor cell types in ferret and human than those in mouse, such that the former display a diversity of progenitor cells that is missing in mouse. The original analysis of this mouse dataset alone, combined with in vivo cell lineage tracing, revealed a much greater diversity of progenitor subclasses than we identified here, which work in parallel to the production of cortical neuron types in mouse cortex. This suggests that in ferret and human, the diversity of progenitor cell classes, or states, and their lineages toward neuron production are likely even more complex than evidenced by our current, conservative analyses.

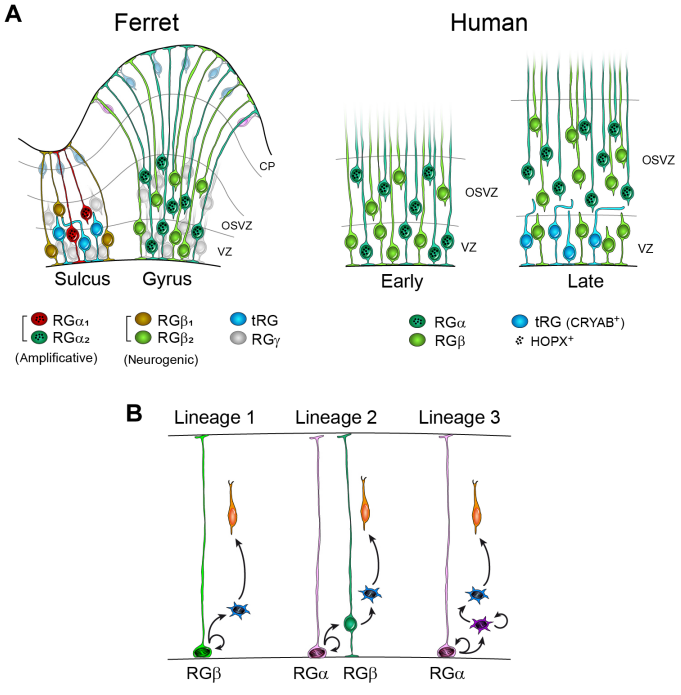

Fig. 8. RGC types and lineages in the cerebral cortex. Schematic summary of key findings and interpretations in this study. (**A**) Conserved diversity of RGC types in ferret and human cortex, including amplificative (RGα), neurogenic (RGβ), general (RGγ), and truncated (tRG). (**B**) Three distinct cell lineages operate in parallel during development.

Duplication of parallel progenitor cell lineages

Our analysis of single-cell transcriptomes revealed three trajectories of progenitor cells, which implicate distinct sets of RGC classes that operate in parallel and duplicated in VZ and OSVZ. The triplication of cell lineages provides a mechanism to amplify and diversify the cellular output, as well as to facilitate regulatory feedback, conferring robustness and evolvability to the system (*9*). We found that the three transcriptomic trajectories are shared in human and ferret, with a large and folded cortex, but our analyses did not reveal them in mouse, showing a single trajectory. The smaller number of cells analyzed in mouse may be a limiting factor, where larger cell numbers might reveal additional trajectories. Thus, it is possible that multiple trajectories are also present in mouse, albeit with a major predominance of a single trajectory, and that they were quantitatively enhanced during cortical evolution. Similarly, IPCs predominate in mouse over a very low abundance of bRGCs, whereas these are both very abundant in ferret and human. Whereas these lineages were observed both in gyrus and sulcus regions, the RGC classes involved were different. On the basis of the analysis of their transcriptomic profile, RGC classes prevalent in the gyrus are more amplificative than those in the sulcus, which are conversely more differentiative. Therefore, the final neuronal output from RGCs in the gyrus is much greater than those in the sulcus. By using TrackerSeq lineage tracing, we confirmed that the three transcriptomic trajectories correspond to bona fide progenitor cell lineages. This experiment also revealed that 4 to 8% of clones containing mitotic RGCs also included ENs, but not IPCs, suggestive of direct neurogenesis. The existence of direct neurogenesis was missed in our Monocle trajectory analysis, possibly due to the substantial transcriptional distance between the RGCs and ENs involved. The low frequency of this mode of neurogenesis as compared to the more complex parallel lineages, as well as its very reduced neuronal output, points to a limited contribution toward overall cortical neurogenesis in species with a large cortex.

The three parallel progenitor cell trajectories converge onto a single cluster of cells featured as late IPCs–newborn ENs. Similarly, our barcoded cell lineage analysis did not reveal substantial heterogeneity between groups of ENs coupled to either lineage. This suggests that different lineages with different types of progenitor cells may generate similar postmitotic neuronal precursor states, again conferring reliability to this process. Nevertheless, note that our clonally labeled cells were collected only 3 days after in utero electroporation and that lineage-specific transcriptomic heterogeneity may emerge later on during neuronal maturation. Accordingly, early newborn neurons may differ at epigenetic levels, or minor transcriptomic differences undetected in our analyses may have long-term effects in the neuron differentiation trajectory (*32*).

Specialized cell types illuminate cortical folding

Primary folding of the cerebral cortex is highly stereotyped and predefined by the pattern of progenitor cell proliferation (*10*, *14*, *31*) and a transcriptomic protomap (*19*, *39*, *40*) along the cortical GZs (*38*). Here, we identify the source of these two patterns: cellular heterogeneity between regions generating future gyrus and sulcus. The GZs of gyrus or sulcus are populated by distinct classes of RGCs, with different transcriptomic profiles and different amplificative versus neurogenic dynamics, underlying greater neuronal output in the gyrus. We

also find different classes of newborn neurons produced in adjacent gyrus versus sulcus, possibly owing to their different RGC classes of origin. These postmitotic neurons born across cortical space (SG or LS) follow parallel maturation trajectories, as observed for mouse RGCs across developmental time (*4*), but here simultaneously. Genes mutated in human malformations of cortex folding have enriched expression in immature neurons of SG, in agreement with the folding protomap (*19*). Future studies should address whether these differences are also relevant in the specialization of neural connectivity and function across cortical areas.

Newborn neurons use the scaffold of radial fibers from RGCs to migrate from VZ to the cortical plate (CP) (*41*). The greater abundance of bRGCs in prospective gyri compared to sulci expands that fiber scaffold into a fanned array, favoring the lateral dispersion of migrating neurons and cortical surface growth (*10*). Here, we observe tRGs, lacking such long radial fiber (*27*), enriched in the prospective LS sulcus as compared to the SG gyrus. This selective enrichment may magnify differences in the radial fiber scaffold between SG and LS, further promoting cortex folding (Fig. 8A). Enrichment of tRGs in LS is temporally coincident with the onset of formation of this particular sulcus; hence, this precise developmental timing may contribute to sulcus formation. Whether this mechanism is conserved in other folds of the ferret cortex will need confirmation in future work. In addition to their relevance in cortex folding, tRGs have been proposed to produce ependymal cells once neurogenesis is complete (*28*).

Specialization of cortical progenitors in evolution

Our analyses of individual GZs uncover an unprecedented diversity of cortical progenitor cells and lineages in ferret and human. While confirming the existence of tRG in ferret as in human (*27*, *28*), we find that aRGCs and bRGCs are transcriptomically indistinguishable in both species at this depth of analysis, including the expression of *HOPX* and other proposed marker genes. Instead, two major transcriptomic classes of RGCs populate both VZ and OSVZ: one expressing *HOPX* (RGα) and the other not (RGβ; Fig. 8A) (*12*, *29*). On the basis of histological evidence, this RGC composition found in ferret across development is likely conserved in human before GW16, but not after (*15*). After GW16, GZs of the human cortex seem to become further specialized, with the replacement in VZ of $HOPX^+$ RGα by tRGs, such that the RGC transcriptomic signature containing *HOPX* (RGα) remains only in the OSVZ (Fig. 8A).

The integration of our ferret dataset with those of human and mouse provides a framework for studying cellular mechanisms of cortical development and evolutionary complexification, where the simple lineages of mouse expanded into multiple parallel lineages (Fig. 8B). Integration with data from additional primate species will test this model and lay out the recent evolutionary trajectory of human cortex development.

MATERIALS AND METHODS

Animals

Pregnant pigmented ferrets (*Mustela putorius furo*) were obtained from Euroferrets (Denmark) and kept on a 12-hour:12-hour light:dark cycle at the Animal Research Facility (SEA) of the Universidad Miguel Hernández. All animals were treated according to Spanish and European regulations. Experimental protocols were approved by the Institutional Animal Care and Use Committee from the Spanish Council for Scientific Research.

Brain extraction and microdissection

For single-cell experiments, E34 ferret embryos were obtained by cesarean section of timed pregnant females under deep anesthesia with sodium pentobarbital (Dolethal and Vetoquinol). E34 ferret embryos and P1 kits were anesthetized and decapitated, the brains were extracted, telencephalic hemispheres were separated in cold Leibovitz's (1×) L15 medium with L-glutamine (Gibco) and embedded in 4% low melting agarose (SeaPlaque, Lonza) within square molds (Peel-A-Way, Polysciences). Agarose blocks containing each brain hemisphere were vibratome-cut (Leica VT1000S) in ice-cold L15 medium obtaining 300-μm-thick sagittal slices. Living cortical slices were further microdissected under a dissecting scope (Leica MS5) with microsurgical knives (MSP) in ice-cold L15 medium. The VZ of E34 embryos and P1 kits and the OSVZ of P1 kits were isolated from the prospective SG and LS. The prospective SG was recognized as the region where the caudal end of the cortex curves forming a 45° angle over the hippocampal region; LS was identified as the cortical region overlying the hippocampus immediately caudal to the dentate gyrus. The use of these landmarks as faithful indicators of LS and SG has been demonstrated previously by means of cell lineage tracing (*10*, *42*), dye tracing, and tracking of the radial fiber scaffold in the developing ferret cortex (*10*). We identified the VZ as the cell dense layer extending from the apical side of the cortex; the OSVZ was identified as the bottom half of a translucent zone between two cell dense layers (ISVZ and CP; Fig. 1B). The microdissected tissue pieces were transferred to 2-ml DNA LoBind tubes (Eppendorf) using ribonuclease (RNase)–free filter pipette tips (JETbiofil). Each E34 sample contained the microdissected regions of two to four embryos from a single litter ($n = 4$ litters). P1 samples were composed of one animal per litter ($n = 5$ litters).

Single-cell suspension, cell concentration, and sample viability

Tissue pieces were enzymatically dissociated for 20 to 30 min at 37°C using MACS Neural Tissue Dissociation Kit P (Miltenyi Biotec) according to the manufacturer's instructions. Cells were mechanically suspended using wide-tipped, fire-polished Pasteur pipettes (Thermo Fisher Scientific), followed by P1000 pipettes (Gilson). Cell suspensions were filtered to new 2-ml DNA LoBind tubes (Eppendorf) through 40-μm PluriStrainer Mini (PluriSelect) prewet with Hank's balanced salt solution (1×) without $CaCl_2/MgCl_2$/pH indicator (Gibco). Cells were centrifuged (Spectrafuge 24D, Labnet) twice for 5 min at 1000 rpm, and cell pellets were resuspended in 40 μl of phosphate-buffered saline (PBS) (pH 7.4) (1×) without $CaCl_2/MgCl_2$ (Gibco) containing 0.04% bovine serum albumin (Sigma-Aldrich). Cell concentration from homogeneous suspensions was measured, and cell death was estimated twice using trypan blue 0.4% stain (Thermo Fisher Scientific): first, on an automatic cell counter (Countess II FL, Thermo Fisher Scientific) and, second, using a hemocytometer (BLAUBRAND Neubauer, BRAND) on an inverted microscope (Leica DM IL). Debris-free suspensions with cell viability over 90% were used immediately for single-cell isolation.

Single-cell isolation and cDNA library preparation

Cell suspensions from 400 to 1200 cells/μl were loaded into the Chromium Single Cell Controller (10x Genomics) to target a recovery of 7000 to 10,000 cells per sample. Single-cell Gel bead-in-EMulsions (GEMs) were generated, cells were lysed, and the released RNAs were retrotranscribed (RT) and barcoded inside the GEMs according to

Single Cell 3′ Reagent Kits v2 and v3 (10x Genomics) protocols. Following GEM break, barcoded cDNAs were pooled and cleaned from leftover RT reagents using DynaBeads MyOne Silane Beads (Invitrogen) and amplified (8 to 14 cDNA amplification cycles) following the Single Cell 3′ Reagent Kits v2 and v3 (10x Genomics) protocols. Enzymatic fragmentation, size selection, and cleanup with SPRIselect Reagent kit (Beckman Coulter) were followed by library construction (total sample index cycles from 12 to 16 cycles), consistent with standard Illumina sequencing constructs and performed in line with Single Cell 3′ Reagent Kits v2 and v3 (10x Genomics) protocols. cDNA and library concentration and quality check were assessed using Bioanalyzer High Sensitivity DNA Kit (Agilent) and 2100 Expert Software (Agilent). Libraries for 18 samples (E34VZSG, $n = 3$; E34VZLS, $n = 3$; P1VZSG, $n = 3$; P1VZLS, $n = 3$; P1OSVZSG, $n = 3$; P1OSVZLS, $n = 3$) were generated.

Library sequencing

Libraries were sequenced at the Centre of Genomic Regulation (Barcelona, Spain) with an Illumina HiSeq2500 system: one sample per lane, paired-end [125–base pair (bp) R1 and R2 or 50-bp R1 and 100-bp R2] reads. Quality check on the raw sequence data was performed using FastQC-0.11.8 (*43*).

Genome alignment

Ferret gene annotation from (*44*), comprising Ensembl *M. putorius furo* annotation expanded using bulk RNA-seq data, was further enriched with National Center for Biotechnology Information (NCBI)'s RefSeq mitochondrial genes from *M. putorius furo* annotation release 101 (accession no. NC_020638.1) and transformed from a general feature format (GFF3) into gene transfer format (GTF) using Cufflinks-2.2.1 (*45*). Johnson's GenBank accessions were matched to RefSeq accessions through cthreepo-0.1.1 (*46*), and gene_name and transcript_name attibutes were added to GTF rows where they were missing by appending gene_id and transcript_id entries using Kotliar's code (*47*). Cell barcodes were corrected, scRNA-seq reads were mapped to NCBI ferret reference genome (MusPutFur1.0/GCF_000215625.1), and Unique Molecular Identifiers (UMIs) were collapsed and corrected by STAR-2.7.3a (*48*) adapting the parameters to v2 or v3 10x chemistry, depending on the sample.

Cell barcode filtering

Count matrices from each sample were loaded, and cell-containing droplets were identified and distinguished from empty droplets retaining ambient RNA by DropletsUtils-1.6.1 (*49*). To achieve this, barcode ranks from the total UMI count were calculated, and knee and inflexion points from the curve were established (fig. S1A). Increased niters argument to 100,000 and decreased false discovery rate to 0.1% were applied to all samples. A lower bound of 80 was implemented for replicates P1VZLS_1 and E34VZSG_1. Mitochondrial genes were removed before cell calling for replicates E34VZLS_2, P1VZSG_2, P1VZLS_3, P1OSVZSG_1, and P1OSVZLS_1 to promote potential unhealthy cell removal using A. Lun proposal implemented by Pickering's code (*50*). This reduced the differences between droplets containing broken cells and ambient RNA and sharpened the distinction from healthy cell–containing droplets. The automatic retention of droplets with large total UMI counts (determined by the knee point) was disabled for replicate P1VZLS_3. The distribution of *P* values for the categorized empty droplets was computed for each sample (fig. S1A).

Cell quality assessment

Seurat-3.1.5 software (*33*) was used to merge samples, add sample's metadata, and further select high-quality cells from the filtered cell barcodes. All cells included in this study had >500 UMIs (fig. S1C) and >250 genes (fig. S1D) per cell and <15% of reads mapping to mitochondrial genes (fig. S1G). Genes were included if they were expressed in a minimum of 10 cells. An adaptive threshold was applied for the maximum number of genes, which consisted on keeping cells with <3 SDs above the mean number of genes detected for each sample (fig. S1B). This cell selection resulted in a total of 16,227 genes from 111,291 high-quality cells (fig. S1, E and F) and retained for downstream scRNA-seq analyses.

Normalization, regression, and sample integration

Samples were preliminarily evaluated for the effects of biological (litter, age, dissecting area, and layer) and technical covariates (10x chemistry, sequencing batch, and sequencing depth). Cell cycle phase for differentiating cells was assessed by assigning S and G_2-M scores to normalized cell counts using an imported list of ferret cell cycle stage markers and Seurat's "CellCycleScoring" function. The merged object was split into a list divided by batch with each litter as an element, and normalization was performed along with regressing out cell cycle difference between S and G_2-M scores and mitochondrial mapping percentage variables by Seurat's "SCTransform" (*51*). A total of 5000 variable genes and scaled data output matrices containing variable and nonvariable genes (return.only.var.genes = FALSE) were selected for downstream analyses. Sample integration was performed using 5000 genes for anchoring (*33*).

Clustering and resolution

Seurat-3.1.5 software (*33*) was used to perform PCA of the integrated dataset. An elbow plot of the SD from the first 100 PCs was traced. The threshold identifying the majority of the sample's variation was established at 32 PCs ± 5 PCs. *K*-nearest neighbors graph construction for cell clustering was tested for 27, 32, and 37 PCs, together with different resolutions (0.5, 0.8, 1.1, 1.4, 1.7, and 2.1) using Seurat's "FindNeighbors" and "FindClusters" functions. Clustree-0.4.2 (*52*) enabled the visualization of the relationships between clusters by the increasing resolution. Resolutions were ruled out when cells comprising one cluster where coming from >2 clusters from the previous resolution or did not constitute an independent cluster in the following resolution. A total of 37 PCs and 1.1 resolution values were selected for cell clustering (fig. S2A). Results were visualized using t-distributed stochastic neighbor embedding (t-SNE) and UMAP dimensional reduction.

Cluster quality assessment

t-SNE plots were colored by metadata to check cell distribution using Seurat-3.1.5 software (*33*). Although the great majority of cells was shuffled and, therefore, free from confounding sources of unwanted variation, clusters 5, 12, 21, and 27 stood out for 10x chemistry (fig. S2B), sequencing batch (fig. S2C), and/or litter (fig. S2D) variables. Further examination of these clusters showed that clusters 5, 12, and 27 did not include cells from all biological replicates from the same condition but mostly cells from a single litter (fig. S2E). These three cell clusters were considered litter-specific and were not included in downstream analyses (Fig. 1C and fig. S3A). PCA confirmed the removal of S-G_2-M cell cycle phase differences while maintaining the distinction between cycling (S-G_2-M) and noncycling (G_1) cells (fig. S2F).

Cell type identification

Following quality controls, cluster composition was homogeneous across conditions, without clusters being specific to a certain age, layer, or region. Therefore, cluster markers were computed by "FindConservedMarkers" function in Seurat-3.1.5 (*33*). Briefly, this function firstly segregated cells by condition and then carried out differential gene expression analyses for each cluster against all others on the normalized data matrix from the object's RNA assay. Wilcoxon rank sum test was used together with the following parameters: Only positive genes were reported (only.pos = TRUE), genes detected in ≥10% of cells in either of the two compared cell groups were tested (min.pct = 0.1), and genes with a $\log_2$FC ≥ 0.25 of average expression between the two compared cell groups were checked (logfc.threshold = 0.25). Gene-level *P* values were calculated for each condition and a combined *P* value across them. Top 20 markers by average FC with the largest $P \leq 0.05$ across conditions were extracted for each cluster and evaluated to identify cell types. Canonical markers for known cell identities during cortical development were also explored (*4*, *5*, *15*, *16*, *20*, *44*, *53*–*64*). Identified cell types were assigned to the different clusters, and clusters 5, 12, and 27 were removed (as explained above: cluster quality assessment), remaining a total of 101,811 cells.

DEG analyses and cluster enrichment

DEGs between two cell groups were obtained by Seurat's function "FindMarkers" (*33*) on the normalized data matrix from the object's RNA assay. This function constructed a logistic regression model out of the expression from each gene to predict cell membership and used a likelihood ratio to compare the built model to a null one where the cell membership does not depend on gene expression. Cell phase S and G_2-M scores, mitochondrial mapping percentage, sequencing depth, 10x chemistry, sequencing batch, and litter were selected as latent variables. In addition, the following parameters were applied: Positive and negative genes were reported (only.pos = FALSE), genes detected in ≥10% of cells in either of the two compared cell groups were tested (min.pct = 0.1), and genes with a $\log_2$FC ≥ 0.25 of average expression between the two compared cell groups were checked (logfc.threshold = 0.25). The obtained DEG list was subsequently filtered by the following criteria: gene detected in ≥44.5% of cells from the first group (pct.1 ≥ 0.445), gene difference of ≥14.5% between the two groups (pct.1-pct.2 ≥ 0.145), $\log_2$FC ≥ 0.445 of the average expression between the two groups (avg_logFC ≥ 0.445), and adjusted $P \leq 0.05$ based on Bonferroni correction using the total number of genes in the dataset (p_val_adj ≤ 0.05). To identify cluster enrichment, the percentage of RGCs or IPCs per cluster for each replicate was normalized, and $\log_2$FC from the mean between replicates for each cluster was calculated.

Subset analyses

Distinct subsets of the dataset were selected from the RGC or IPC–newborn EN clusters. The RNA assay was kept by the "DietSeurat" function (*33*) to be reanalyzed as detailed above (see the "Normalization, regression, and sample integration," "Clustering and resolution," "Cluster quality assessment," "Cell type identification," and "DEG analyses and cluster enrichment" sections) with the following adjusted parameters: P1 RGC clusters 8 and 16: integration nfeatures = 4000, clustering dims = 1:26, resolution = 1.4, cluster marker identification by "FindAllMarkers," DEG-filtering criteria (a) for cluster 8 or 16 VZ-OSVZ analyses and (b) for cluster 8 or 16 SG-LS–enriched subcluster analyses (SG clusters 1, 2, and 9; LS clusters 3, 4, and 6): pct.1 ≥ 0.2, pct.1/pct.2 ≥ 1.45, p_val_adj ≤ 0.05, pct.2 ≤ 0.05 (for cluster 16 VZ-OSVZ comparison) and ≤ 0.1 (for cluster 8 LS-SG comparison), p_val ≤ 0.05 (for cluster 16 OSVZ-VZ comparison), and $\log_2$FC from the number of cells of each cluster per condition for each replicate. IPC-newborn neurons: normalization vars.to.regress = c("S.Score," "G_2M.Score," "percent.mt"), integration nfeatures = 4000, clustering dims = 1:33, resolution = 1.4, cluster marker identification by "FindAllMarkers," quality control (QC) cluster 21 removed, $\log_2$FC from the number of cells of each cluster per condition for each replicate, and canonical markers from (*17*, *60*, *65*–*72*). RGC with partial cell cycle regression: integration nfeatures = 4000. RGC with total cell cycle regression: normalization vars.to.regress = c("S.Score," "G_2M.Score," "percent.mt") and integration nfeatures = 4000.

Cell cycle regression analyses

Complete cell cycle regression was performed on the full dataset during the normalization step [vars.to.regress = c("S.Score," "G_2M.Score," "percent.mt")] to compare it with partial cell cycle regression output (see fig. S7B). Downstream analyses were carried out as detailed above (see the "Normalization, regression, and sample integration," "Clustering and resolution," and "Cluster quality assessment" sections) with the following adjusted parameters: clustering dims = 1:42; resolution = 1.4, and QC clusters 7, 17, and 30 removed. Likewise, complete cell cycle regression was performed on the RGC subset (for specific parameters, see the "Subset analyses" section).

Functional enrichment analyses

Pathway enrichment analyses were performed for the identified hub cell types using ReactomeGSA-1.8.0 (*73*), with Reactome database versions 79 and 80. For the subset of ferret P1 RGC clusters 8 and 16, ferret gene name nomenclature was standardized to capital letters. For the integrated multispecies dataset, concurrent high-confidence orthologs in the three species were selected. Mean gene expression of the RNA assay–normalized data matrix for the compared cell groups was calculated, and gene set variation analysis was performed. This analysis does not assign statistical significance for the enrichment of gene sets. Pathway-level expression values were sorted by the maximum difference between cell groups, and the first 25 pathways were selected for heatmap plotting. PCAs for the pathway-level expression values were run for the multispecies dataset.

Trajectory analyses

Cells belonging to the excitatory neuronal lineage (Fig. 1D) were selected. Clustering and partition metadata entries (monocle3_clusters, monocle3_partitions = 1) were assigned to the Seurat object for subsequent analyses with the Monocle package, and UMAP dimensional reduction for downstream three-dimensional (3D) rendering was carried out. SeuratWrappers-0.3.0 (*74*) was used to convert the RNA assay from the Seurat object into CellDataSet (cds) class. Trajectories were constructed by "learn_graph" function from Monocle3-0.2.2.0 (*75*–*77*). Cells of origin for each trajectory were identified as the center of RGC clusters with the highest G_2-M score using a graphical user interface. For pseudo-time analyses per condition, the Seurat object was split into conditions before cds construction. To further characterize the pseudo-time trajectories for each condition individually, the three replicates that constitute each condition were integrated as explained above (see the "Cell quality assessment"; "Normalization, regression, and sample integration": by replicate ("orig.ident");

"Clustering and resolution": E34VZSG dims = 1:25 resolution = 1.4, E34VZLS dims = 1:37 resolution = 1.4, P1VZSG dims = 1:32 resolution = 1.1, P1VZLS dims = 1:40 resolution = 1.1, P1OSVZSG dims = 1:24 resolution = 1.1, P1OSVZLS dims = 1:24 resolution = 2.1; "Cluster quality assessment" sections), and pseudo-time analyses were carried out. To study DEGs dynamics over pseudo-time trajectories, the common neuronal clusters that constitute the three trajectories identified in Fig. 3B were cut out from the analyses, and clusters specific to trajectory 1, 2, or 3 were selected. For this, DEGs were previously calculated in Seurat [filtering parameters: pct.1 ≥ 0.5, pct.2 ≤ 0.5, pct.1-pct.2 ≥ 0.345, avg_logFC ≥ 0.445, p_val_adj ≤ 0.05, pct.1-pct.2 ≥ 0.055 (for trajectory1-trajectory2 + 3 comparison), avg_logFC ≥ 0.345 (for trajectory 1–trajectory 2 + 3 comparison)], size factors were computed, and gene names were added to cds. Pseudo-time examination was also performed for the IPC–newborn EN subset, where clusters 14 and 20 were excluded as they were found in different partitions. Cells in the center of clusters 9 and 10 (cluster 3 in the full ferret dataset; see Fig. 7A) were designated as trajectory origin. To build pseudo-time trajectories for different species at comparable developmental stages, the integrated Seurat object with ferret P1, human GW17 to GW18, and mouse E15.5 cells (see the "Species dataset integration and DEG analyses" and "Cell downsampling" section) was analyzed as a whole or split into species metadata before cds construction.

Electroporation

E34 ferret embryos and P1 kits were electroporated following standard protocols as previously described (*11*, *42*). Briefly, for in utero electroporation, timed pregnant females were deeply anesthetized with isoflurane (ISOFLO, Ecuphar), the abdominal cavity was open, and the uterine horns were exposed. DNA solution (2.5 μl) was injected into the lateral ventricle using pulled borosilicate glass capillaries (1B120F-4, World Precision Instruments), and square electric pulses (75 V; 50-ms on, 950-ms off, and five pulses) were applied with an electric stimulator (Cuy21EDIT, Bex Co. Ltd.) using round electrodes (CUY650P7, Nepa Gene). Uterine horns were placed back into the abdominal cavity, suture was closed, and the pregnant female was allowed to fully recover on a heating pad before returning to the home cage. For postnatal electroporation, kits were deeply anesthetized with isoflurane, DNA solution (2.5 μl) was injected into the lateral telencephalic ventricle with a glass micropipette through the skull, and square electric pulses were passed through the head (75 V; 50-ms on, 950-ms off, and five pulses) using the same electrodes. For RGC labeling, plasmid concentrations were as follows: pCAG-Floxp-farnesylated–enhanced GFP (EGFP) (1 μg/μl) and pCAG-Cre (10 ng/μl). For barcode lineage tracing, plasmid concentrations were as follows: pEF1a-pBase (1 μg/μl; piggyBac transposase; a gift from R. Platt), TrackerSeq library (2 μg/μl) diluted in endo-free TE buffer [50 mM tris-HCl (pH 7.5) and 1 mM EDTA (pH 8.0)] and 0.002% Fast Green FCF (Sigma-Aldrich).

Analysis of barcoded cell lineages

TrackerSeq is a piggyBac transposon–based (*78*) library that was previously developed to be compatible with the 10x single-cell transcriptomic platform (*32*). It records the in vivo lineage history of single cells through the integration of multiple oligonucleotide sequences into the genome. Each of these individual lineage barcodes is a 37-bp-long synthetic nucleotide that consists of short random nucleotides bridged by fixed nucleotides. This design results in a library with a theoretical complexity of approximately 4.3 million lineage barcodes (16^8) with each barcode differing from another by at least 5 bp. The barcode library was prepared as described in detail in (*32*). We assessed the integrity of the TrackerSeq barcode library by sequencing to a depth of approximately 13 million reads to test whether any barcode was overrepresented. A total of 5×10^6 clusters of barcodes were identified, suggesting that the barcode library has a diversity that is at least in the 10^6 range.

E34 ferret embryos were electroporated with the piggyback-transposase and the TrackerSeq library. E37 embryos were obtained by cesarean section under deep anesthesia with sodium pentobarbital (Dolethal and Vetoquinol). Embryos were anesthetized and decapitated, the brains were extracted, telencephalic hemispheres were separated in ice-cold Leibovitz's (1×) L15 medium with L-glutamine (Gibco), and blocks of caudal cerebral cortex containing GFP^+ cells including the prospective SG and LS were obtained with microsurgical knives (MSP) in ice-cold L15 medium. Tissue pieces were enzymatically dissociated using papain (LK003150, Wortington) with deoxyribonuclease at 37°C for 20 min. Earle's balanced salt solution (from the Worthington kit) was added to the digested tissue, and the mixture was triturated with a 10-ml plastic pipette. Reconstituted Worthington inhibitor solution and deoxyribonuclease solution were added and gently mixed. Cells were pelleted by centrifugation at 300*g* for 5 min at room temperature, resuspended in ice-cold L15 medium containing 5% fetal bovine serum, and filtered to DNA LoBind tubes (Eppendorf) through 40-μm PluriStrainer Mini (PluriSelect). Positive cells were isolated by flow cytometry using a BD FACSAria II Cell Sorter (BD, FACSDiva software, version 6.1.3) with an 85-μm nozzle. EGFP-positive cells were collected into bovine serum albumin (Sigma-Aldrich)–precoated 2-ml DNA LoBind tube for downstream processing on the 10x Genomics Chromium platform. After sorting, 80,000 cells were divided into four samples and loaded onto 10x Genomics Chromium platform for GEMs and cDNA generation carrying cell- and transcript-specific barcode using the Chromium Next GEM Single Cell 3′ Reagent Kit v3.1 with Feature Barcoding technology (PN-1000121) following the manufacturer's protocol (document number CG000205, 10x Genomics). Uniquely barcoded RNA transcripts were reverse-transcribed. 3′ Gene Expression libraries were generated according to the manufacturer's user guide with use of Chromium Library v3.1 kit (PN-1000121) and Dual Index Kit TT Set A (PN-1000215) (10x Genomics). Libraries were quantified with Agilent Bioanalyzer. The TrackerSeq lineage libraries were retrieved from cDNA. Ten microliters of cDNA was amplified with Q5 polymerase (M094S, New England Biolabs) in a 50-μl reaction as described like in (*32*). Libraries were purified with a dual-sided selection using SPRIselect (B23318, Beckman Coulter) and quantified with Agilent Bioanalyzer. Transcriptome and TrakerSeq barcode libraries were sequenced on an Illumina NovaSeq at the Next-Generation Sequencing Facility of the Max Planck Institute of Biochemistry (Munich, Germany).

Fastq files for both transcriptome and barcoded sequences were obtained by bcl2fastq v2.20.0.422 (Illumina). For barcoded lineage preprocessing, a modified version of the TrackerSeq pipeline (*32*) was used (https://github.com/mayer-lab/TrackerSeq). Raw barcoded sequence reads were trimmed by BBMap v38.68 [BBMap (*79*); sourceforge.net/projects/bbmap/]. Whitelist of cell barcodes and barcode extraction from sequence reads were obtained by UMI-tools v1.1.2 (*80*). After that barcode extracted fastq files were processed to generate sparse matrix with cell barcodes in rows and

clone IDs in columns. For further analysis, only cell-UMI-BC triples with at least 10 reads and cell-BC pairs with at least nine UMI were selected. Instead of the previously described clustering approach using Jaccard distance (*32*, *34*), we used a network-based approach to assigning cells to their corresponding clones. Using the sparse matrix as input, an undirected network was built between cell barcodes (i.e., cells) as nodes, which are connected if at least one lineage barcode is present in both cells. After building the network, we calculated connected components. Each connected component represents one clone. This novel approach requires fewer computational resources compared to the orginal method, and, thus, it can be applied to large datasets.

For lineage transcriptome preprocessing, 10x compatible expression matrices were generated by STARsolo in STAR aligner (*48*) with specified parameters: --soloType CB_UMI_Simple, --soloCBstart 1, --soloCBlen 16, --soloUMIstart 17, --soloUMIlen 12, --soloCBmatchWLtype 1MM_multi_Nbase_pseudocounts, --soloUMIfiltering MultiGeneUMI_CR, --soloUMIdedup 1MM_CR, --clipAdapterType CellRanger4, --outFilterScoreMin 30, --outSAMtype BAM SortedByCoordinate, --outSAMattributes CR UR CY UY CB UB. These parameters are designed to ensure that STARsolo is as close as possible to the 10x compatible output of the Cell Ranger application. After obtaining the expression matrices, empty and pseudo-drop cells were filtered by DropletUtils using "emptyDrops" functions with 100,000 iterations and minimum UMI count of <500 (*49*). Last, clone IDs were matched with cell barcodes in transcriptome dataset and included into metadata of corresponding samples. Of 51,006 GFP-positive cells, we filtered 15,326 cells with lineage barcodes. We have excluded from our analysis a clone that consisted of approximately 4000 cells, as it was likely formed by low-quality cells present in our dataset. After these preprocessing and filtering steps, we obtained 11,933 single-cell clones and 1413 multicell clones.

Transcriptomic counts were preprocessed in a similar fashion as above. Briefly, cells were filtered for number of counts, number of features, and percentage of mitochondrial reads (nCount < 10,000, nFeature > 1000, percent.mt < 5), followed by log normalization. We used DoubletFinder (*81*) to exclude putative doublets from the dataset. We used the Seurat workflow to identify highly variable features, scale log-normalized counts, and regress out read depth and the difference between G_2M and S phase scores (*33*). We performed PCA and embedded cells using UMAP algorithm as implemented in Seurat (*33*, *82*–*84*). Expression of established marker genes was used to check for major cell types. Cells were annotated using correlation-based label transfer using the previous scRNA-seq experiment as reference (*33*). Before label transfer, we excluded all non-neuronal and inhibitory clusters from the reference set to enable a more precise label transfer. Lineage coupling *z* scores and correlations were calculated as implemented in (*32*). Lineage coupling *z* scores were visualized using Cytoscape (*85*).

To infer heterogeneity of ENs originating from distinct RGC populations (RGC-A or RGC-B), clones containing only cells in mature ENs (clusters 1 + 7) and cells in either RGC-A (clusters 23 + 11) or RGC-B (22 + 9) were identified. PCA was performed with scaled gene expression values from ENs and differentially expressed analysis was applied to ENs. Differentially expressed analysis yielded 160 DEGs. To check whether the number of DEGs is significantly more than you would expect by chance, we performed a permutation test by randomly sampling clone group labels.

Species dataset integration and DEG analyses

Raw counts and cell metadata from *Homo sapiens* (*16*) (GW17 to GW18, GZs microdisected away from the CP, and Drop-Seq technology) and *Mus musculus* (*18*) (E15.5, whole cortex, and ddSEQ technology) were downloaded. High-confidence orthologous genes between human-ferret and human-mouse were selected for downstream analyses: (a) Data frames containing gene names from ferret and downloaded human and mouse datasets were generated; (b) Ensembl datasets "hsapiens_gene_ensembl," "mpfuro_gene_ensembl," and "mmusculus_gene_ensembl" were obtained from BioMart database thru "useMart" function using biomaRt-2.42.1 (*86*, *87*); (c) homolog genes between human-ferret and human-mouse and their characteristics were extracted form Ensembl datasets in accordance with different ferret and mouse attributes by biomarRt "getLDS" and "getBM" functions (Ensembl 104: 11,615 ferret entries and 13,779 mouse entries); (d) high-confidence orthologs were excerpted from the homolog lists of genes in line with the Ensembl criteria: 1:1 orthologs; last common ancestor/minimum gene order conservation score/minimum whole-genome alignment coverage/minimum percentage of target gene identical to query human gene/minimum percentage of human query gene identical to target gene = *Bilateria*/_/_/25%/25%, *Chordata*/_/_/25%/25%, *Vertebrata*/_/_/25%/25%, *Gnathostomata*/_/_/25%/25%, *Euteleostomi*/50/50/25%/25%, *Sarcopterygii*/50/50/25%/25%, *Tetrapoda*/50/50/25%/25%, *Amniota*/50/50/25%/25%, *Mammalia*/75/75/50%/50%, *Theria*/75/75/50%/50%, *Eutheria*/75/75/50%/50%,*Boreoeutheria*/75/75/50%/50%,and*Euarchontoglires*/75/75/50%/50% (9671 ferret genes and 3366 mouse genes); (e) gene names from ferret and mouse datasets that were determined as high-confidence orthologs were replaced by human HUGO Gene Nomenclature Committee (HGNC) gene names or human Ensembl IDs (in the lack of gene names), and ferret and mouse genes without a high-confidence orthologue were added a f- or a m- prefix, respectively. Ferret Seurat object was cleaned by "DietSeurat" from the previously generated SCT, and integrated assays before genes from ferret and mouse Seurat objects were renamed using "RenameGenesSeurat" function from Seurat.Utils (*88*). The CP from the human Seurat object was discarded for downstream analyses, and the cell cycle scores were assigned to the mouse dataset as explained above (see the "Normalization, regression, and sample integration" section). Last, metadata from the different species was labeled to match between datasets before merging them. The merged object was split into a list divided by batch with each subject (i.e., human donor, mouse replicate, or ferret litter) as an element. Normalization, regression, and sample integration steps were performed as explained above. Clustering and resolution were established as 23 PCs and 1.1 following the process explained previously (see the "Clustering and resolution" section). The distribution of previous cluster labels given in (*16*) and (*18*) was explored on the integrated dataset as a control of integration (fig. S17B).

DEG analyses of the concurrent high-confidence orthologs in the three species for iCl.4 and iCl.12 were performed as explained above (see the "DEG analyses and cluster enrichment" section) with cell phase S and G_2-M scores, mitochondrial mapping percentage, and sequencing depth as latent variables. The following filtering criteria were applied: for up-regulated genes, pct.1 ≥ 0.445, pct.1-pct.2 ≥ 0.145, avg_logFC ≥ 0.445, and p_val_adj ≤ 0.05; for down-regulated genes, pct.2 ≥ 0.445, pct.2-pct.1 ≥ 0.145, avg_logFC ≤ −0.445, and p_val_adj ≤ 0.05. Common up-regulated genes between two of the species against the third one were obtained. In addition, DEG

analyses using "FindConservedMarkers" function and "species" as grouping variable as explained before (see the "Cell type identification" section) were carried out with Wilcoxon rank sum test and logistic regression model to acquire common up-regulated genes between the three species at the same time. An adjusted $P \leq 0.05$ (mouse_p_adj_val, human_p_adj_val, and ferret_p_adj_val ≤ 0.05) was used as subsequent filtering criteria. Stacked violin plots from the top 10 genes from each comparison were produced from the scaled data using Tang's code (*89*). Dataset integration from the different species including only P1 ferret cells was also performed as explained in this section with the following specific parameters: clustering dims = 1:24, resolution = 1.4, cluster marker identification by "FindAllMarkers," QC cluster 32 removed, and $\log_2$FC from the number of cells of each RG cluster per species for each replicate.

Pseudo-bulk expression analyses

Clusters from progenitors in G_1 cell phase and neurons were selected from the ferret dataset. A vector for canonical markers from human apical, basal, or pan-RGCs from scRNA-seq bibliography of cortical development was created (*5*, *15*, *16*). Average gene expression of the normalized and scaled data matrix from the RNA assay for the selected clusters coming from different isolated germinal layers was calculated. Heatmap with expression level values was plotted for each gene using pheatmap-1.0.12 (*90*).

Cell downsampling

Cells from our current P1 ferret (65,413 cells) and human available (17,136 cells) (*16*) datasets were downsampled to match mouse dataset (*18*) cell number (5777 cells) before being merged. To achieve this, two methods were carried out using "subset" function from base with different identity arguments and downsample cell numbers: (a) random downsampling—to randomly select cells from the set of all cells in the dataset so the total number of cells is 5777 cells. It produces a proportional decrease in all clusters. (b) Threshold downsampling—to randomly select cells from each identified cell cluster in the dataset so the maximum cells per cluster does not exceed 614 cells for human [that has 16 identified clusters in (*16*) dissected GZ] or 230 cells for ferret (that has 26 identified clusters in this study). It results in a reduction of size from the most populous clusters. Random downsampling resulted in a total of 17,331 cells (ferret 5777 cells + human 5777 cells + mouse 5777 cells). Threshold downsampling resulted in a total of 17,319 (ferret 5764 cells + human 5778 cells + mouse 5,777 cells). Random or threshold downsampled cells were subsequently merged and analyzed as detailed above (see the "Species dataset integration and DEG analyses" section) with the following adjusted parameters: random downsampling—clustering dims = 1:16, resolution = 1.4, cluster marker identification by "FindAllMarkers," and $\log_2$FC from the number of cells of each RG cluster per species for each replicate. Threshold downsampling—clustering dims = 1:19, resolution = 1.7, cluster marker identification by "FindAllMarkers," and $\log_2$FC from the number of cells of each RG cluster per species for each replicate.

Analysis of MCD genes

Databases from genes linked to MCDs were explored (*91*–*97*). Gene lists of markers for polymicrogyria ($n = 53$ genes), LIS ($n = 36$ genes), and/or microcephaly ($n = 10$ genes) (total genes without repetitions = 96) were supplied to "AddModuleScore" function from Seurat. This function calculated the enrichment of the provided gene list for each cell. Briefly, the function calculates the so-called gene score values, which are the average relative expression of the gene set minus the average relative expression of a randomly selected set of control genes. The control genes were selected by binning the expression level from all the genes in the dataset into 24 bins and then randomly choosing 100 genes for each gene in the provided gene list that belong to the same expression bin. Thus, the control group of genes has an equivalent distribution of the expression levels to the MCD gene and is 100 times larger. This minimizes the effect of the varying quality that each cell may have on the cell score. Feature maps highlighting the clusters that might be susceptible to MCD were plotted by "FeaturePlot_scCustom" function using scCustomize-0.7.0 (*98*).

Tissue processing for immunocytochemistry, conventional, and FISH

E34 ferret embryos were obtained by cesarean section of timed pregnant females upon deep anesthesia with sodium pentobarbital (Dolethal and Vetoquinol). E34 embryos and P1 to P6 kits were deeply anesthetized and perfused transcardially with ice-cold 4% paraformaldehyde (PFA; Sigma-Aldrich) in phosphate buffer (pH 7.4).

P1 and P3 brains for immunocytochemistry were extracted and postfixed in 4% PFA. Then, brains were cryoprotected in 30% sucrose (PanReac AppliChem) and sectioned. Brain sections were incubated with primary antibodies overnight (o.n.), followed by appropriate fluorescently conjugated secondary antibodies and counterstained with 4′,6-diamidino-2-phenylindole (DAPI; D9542, Sigma-Aldrich). Primary antibodies used were against GFP (1:1000; GFP-1020, Aves Lab), anti-alpha B crystallin (1:500; ab13496, Abcam), and anti-laminin (1:500; AB2034, Merck Millipore). Secondary antibodies were from Jackson ImmunoResearch [Alexa Fluor 488 anti-chicken immunoglobulin Y (IgY) (1:250; 703-545-155)] and from Invitrogen [Alexa Fluor 555 anti-mouse IgG (1:250; A31570) and Alexa Fluor 488 anti-rabbit IgG (1:250; A21206)].

For clarification experiments, electroporated brains were immunostained and cleared using the iDisco+ protocol as described in (*99*). Briefly, individual brain hemispheres were incubated at 37°C for 7 days (P2) or 10 days (P6) with anti-GFP antibodies (1:1000, GFP-1020, Aves Lab) and 4 days (P2) or 7 days (P6) with appropriate Alexa Fluor 647–conjugated secondary antibodies (1:500; 703-605-155, Jackson ImmunoResearch).

Brains for conventional ISH were extracted in ice-cold PBS (1×), split by the midline, and postfixed at 4°C for 72 hours in 4% PFA (E34) or 4% PFA and 30% sucrose (PanReac AppliChem) (P1). Brains for FISH by means of RNAscope were dissected and postfixed at 4°C for 6 days in 4% PFA (E34 and P1). Brains were then sagittal-sectioned and handled as follows: E34 ISH brains were embedded in 4% low melting agarose (SeaPlaque, Lonza) diluted in PBS (1×) within histology square molds (Peel-A-Way, Polysciences) and vibratome-cut (Leica VT1000S) in PBS (1×) at 50 μm in thickness; P1 ISH brains were frozen and cryotome-sectioned (Leica SM2010R) at 50 μm; E34 and P1 samples for FISH were frozen in embedding medium for frozen tissue (Neg-50, Epredia) by means of liquid nitrogen and isopentane (VWR), and cryostat-sectioned (Leica CM1860 UV) at 30 μm on Superfrost Plus slides (Epredia).

Constructs

For ISH, ferret RNA was extracted using RNeasy Mini Kit (QIAGEN) and RT by the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). Selected genes were amplified by polymerase chain reaction using GoTaq G_2 Flexi DNA Polymerase (Promega) with the

following primers: *CRYAB*, 5'-AGCTAGTGAAACAAGACCAGGA-3′ (forward) and 5′-ACCATGTTCATCCTGACGCT-3′ (reverse); *HOPX*, 5′-ACTCAGCTTCCAGATCCAGG-3′ (forward) and 5′-TTTTGACTCCCAGCCTGACT-3′ (reverse); *NDFIP1*, 5′-ATTTC-CAAAGCCGCCTTCCT-3′ (forward) and 5′-AAGACTGCCCCC TACTGACT-3′ (reverse); *PAX6*, 5′-AGCATGCAGAACAGTCA-CAGC-3′ (forward) and 5′-GGCAAACACATCTGGATAATGG-3′ (reverse); *PSMB1*, 5′-TCTCGCCCTACGCTTTCAAC-3′ (forward) and 5′-CACACTGCTTACTTCAGCCTTT-3′ (reverse); *SLC1A3*, 5′-CTTCCGCGCTACAAAACCAG-3′ (forward) and 5′-CAGAGGGG CGTACCACATTA-3′ (reverse); *EOMES*, 5′-ACTGGTTCCCACTG-GATGAGAC-3′ (forward) and 5′-CTTCGCTCTGTTGGGGT-GAAAGG-3′ (reverse). The resulting amplicons were purified with illustra GFX polymerase chain reaction DNA and Gel Band Purification Kit (GE Healthcare) and cloned into pGEM-T Easy Vector System (Promega).

In situ hybridization

Plasmids containing ferret clones were linearized and purified, and sense or antisense complementary RNA (cRNA) probes were transcribed with T7 or SP6 RNA polymerases (Roche) and labeled with digoxigenin (DIG) RNA labeling Mix (Roche) according to the manufacturer's instructions. Free-floating brain sections were mounted on Superfrost Plus slides (Epredia) in RNase-free (diethyl pyrocarbonate; Sigma-Aldrich) PBS and dried at 37°C for 4 hours. Sections were fixed with 4% PFA (Sigma-Aldrich) for 10 min and washed with RNase-free PBS before being treated with proteinase K (10 μg/ml) (Invitrogen) for 5 min. After a new step of fixation and permeabilization with RNase-free PBS-Tween, brain sections were prehybridized in a humidified chamber for 1 hour at 62°C in hybridization buffer [50% formamide (deionized) (Ambion); 10% dextran sulfate sodium salt from *Leuconostoc* spp. (Sigma-Aldrich); salt solution (1×) containing 200 mM NaCl, 10 mM tris-HCl, 1 mM tris base, 50 mM $NaH_2PO_4.2H_2O$, and 50 mM Na_2HPO_4 (pH 7.5); Denhardt's solution (1× from a 50× stock) (Sigma-Aldrich); and yeast tRNA (0.7 mg/ml; Invitrogen)]. Then, brain slices were hybridized in a humidified chamber with denatured DIG-labeled cRNA probes diluted in hybridization buffer o.n. at 62°C. Sections were then washed [50% formamide (Sigma-Aldrich), 0.5% Tween 20, and saline sodium citrate buffer (0.5× from a 20× stock) (Sigma-Aldrich)] and incubated with blocking solution [malelic acid buffer containing Tween 20 (1×), 10% sheep serum (Sigma-Aldrich), and 10% blocking reagent (Roche)] in a humidified chamber for 1 hour at room temperature. Next, sections were incubated with anti-DIG alkaline phosphatase–coupled Fab fragments in blocking solution (1:3500; Roche) o.n. at 4°C. Slides were thoroughly washed with malelic acid buffer containing Tween 20 (1×) and with prestaining NTMT buffer (1×) [100 mM NaCl, 100 mM tris-HCl (pH 9.5), 50 mM $MgCl_2$, and 0.5% Tween 20], before incubated in developing solution [nitrotetrazolium blue chloride (0.05 mg/ml; Sigma-Aldrich) and 5-bromo-4-chloro-3-indolyl phosphate *p*-toluidine salt (0.18 mg/ml; Roche) in NTMT buffer (1×)]. Once cRNAs were visible, brain sections were incubated in TE buffer, fixed for 5 min in 4% PFA, washed again with TE buffer, dehydrated, and coverslipped using Eukitt mounting medium (Fluka). Each probe was assayed on P1 ($n = 2$) and E34 ($n = 1$) animals.

Fluorescence in situ hybridization

Ferret probes were acquired from ACD biotechne catalog (https://acd-bio.com/species-common/domestic-ferret): Mp-NFIA-C2 (#586521), Mp-NFIX (#586541), Mp-PAX6-C3 (#513391), and 3-plex negative control probe (#320871); or designed upon request. The RNAscope Multiplex Fluorescent Detection Kit v2 (ACD), Pretreatment Reagents (ACD), RNAscope Multiplex TSA Buffer (ACD), and RNAscope Wash Buffer Reagents (ACD) were used according to the manufacturer's instructions. Briefly, slides from frozen tissue were washed in PBS (1×) for 5 min, dried at 60°C for 50 min, incubated in 4% PFA (Carl Roth) for 15 min at 4°C, and dehydrated [50, 70, and 2× 100% ethanol (Thermo Fisher Scientific and Merck Millipore)] at room temperature. After drying, hydrogen peroxide solution was added for 10 min and washed with RNase-free (diethyl pyrocarbonate, SERVA) water. For antigen accessibility, slides were incubated in 99°C antigen retrieval solution for 3 min, washed with RNase-free water, and dehydrated in 100% ethanol. Protease III was added for 13 min at 40°C on dried slides and washed with RNase-free water. For hybridization, C2 and C3 probes were diluted in C1 probes at a 1:50 ratio and incubated on the slides for 2 hours at 40°C. Slides were then washed in (1×) wash buffer (ACD) and stored in (5×) saline sodium citrate buffer o.n. at room temperature. After washing in (1×) wash buffer, target probe signals were amplified by sequential incubation with the following amplifiers with 2-min washes twice in (1×) wash buffer at room temperature in between: (a) Multiplex FL v2 AMP 1 for 30 min at 40°C, (b) Multiplex FL v2 AMP 2 for 30 min at 40°C, and (c) Multiplex FL v2 AMP 3 for 15 min at 40°C. Probes were then developed one by one in a three-step incubation with 2-min washes twice in (1×) wash buffer at room temperature in between: probe C1 [(a) Multiplex FL v2 HRP-C1 for 15 min at 40°C, (b) TSA Plus fluorophore for 30 min at 40°C, and (c) Multiplex FL v2 HRP blocker for 15 min at 40°C], probe C2 [(a) Multiplex FL v2 HRP-C2 for 15 min at 40°C, (b) TSA Plus fluorophore for 30 min at 40°C, (c) Multiplex FL v2 HRP blocker for 15 min at 40°C], and probe C3 [(a) Multiplex FL v2 HRP-C3 for 15 min at 40°C, (b) TSA Plus fluorophore for 30 min at 40°C, and (c) Multiplex FL v2 HRP blocker for 15 min at 40°C]. Before mounting with Aqua-Poly/Mount medium (Polysciences), sections were counterstained with DAPI (1:1000; Sigma-Aldrich). Probes were developed with the following detection kits diluted in Multiplex TSA Buffer (ACD): TSA Plus Cyanine 3 [(1:1500; Akoya Biosciences) for Mp-NFIX, Mp-AQP4, Mp-ZEB1, Mp-PAX6, and Mp-ID1 and (1:750; Akoya Biosciences) for Mp-GADD45B probes] and TSA Plus Cyanine 5 [(1:3000; Akoya Biosciences) for Mp-NFIA and (1:1500; Akoya Biosciences) for Mp-CRYAB, Mp–Homeodomain-Only Protein Homeobox (HOPX), Mp-ZEB1, and Mp-CABLES2 probes]. Each combination of probes was applied to the following number of animals: Mp-NFIX and Mp-NFIA to P1 ($n = 3$) and E34 ($n = 3$); Mp-ID1 and 2-plex assays to P1 ($n = 3$) and E34 ($n = 1$). Negative controls for each combination of probes were carried out on P1 ($n = 1$) and E34 ($n = 1$) animals.

Imaging

Images from immunolabeling were acquired with an Olympus FV10 confocal microscope using a 63× oil objective and FV10-ASW 4.2 software. Stacks of single-plane images were acquired with a 1-μm interval. To analyze clarified brains, 3D imaging acquisition was performed using ultramicroscope II using ImSpector Pro software (Miltenyi/LaVision BioTec), and images were generated using Imaris ×64 software (Oxford Instruments). Images for ISH were acquired using transmitted light (Zeiss Imager. Z2 microscope) with a 10× lens coupled to a digital camera (AxioCam MRm). For FISH, high-magnification images (40× objective) were acquired on a confocal laser scanning microscope (Zeiss

LSM710). Stacked images of five steps with a 2.5-μm interval were acquired.

FISH counting

Images for RGC probes were quantified through a generated macro using ImageJ-2.0.0-rc-65/1.51w (*100*), MorphoLibJ (*101*), and 3D ImageJ Suite (*102*) plugins. Briefly, layers were delineated and measured their surface. The last slice from the stack of images was discarded to avoid uneven effect of probe penetrance into the tissue. Then, images were grayscaled and split into channels to create individual channel binary masks using Gaussian Blur/Maximum filter difference and Default/Moments threshold for probes puncta. Particles of 1 pixel size were discarded. Puncta were identified in the 3D binary mask and counted using connected-component labeling and 3D Manager. For $PAX6^+$ cell counting, cell binary masks were created using Unsharp Mask plus Gaussian Blur filters on the nuclei, followed by Find Maxima for each slice. Then, cells outline and puncta binary masks were merged for each slice, and cells containing puncta from PAX6 probe alone or together with the other probe for each slice were counted.

HOPX sequence identification

The sequence under ENSMPUG00000030542 Ensembl ID in the ferret gene annotation from (*44*) was identified as HOPX. In this GTF, ENSMPUG00000030542 was simultaneously associated with the positive (three exons) and negative (two exons) DNA strands. Our database has >99% of alignment to the positive strand and <1% to the negative strand. It is the positive strand from the same location where NCBI's latest *M. putorius furo* annotation (ASM1176430v1.1/GCF_011764305.1) using Gnomon predicted the presence of HOPX (gene ID: 101691950) containing three exons with 100% identity with our transcript (*103*). ENSMPUG00000030542 gene position (GL897112: 1,462,992 to 1,492,582) (*104*) was searched in the UCSC Genome Browser (*105*) where HOPX RNAs for numerous organisms (e.g., human, mouse, and pig) were aligned to the positive strand with 82 to 86% identity (*106*–*108*). The protein sequence of human HOPX (GeneBank: AAH14225.1) and the ferret positive strand of ENSMPUG00000030542 are 91,78% homologous ($E = 6 \times 10^{-51}$) (*103*).

Statistics

Wilcoxon rank sum test and logistic regression model were used as described in previous sections (see the "Cell type identification," "DEG analyses and cluster enrichment," and "Species dataset integration and DEG analyses" sections). Prism software was used for the statistical analyses of cluster enrichment and gene score. For cluster enrichment and gene score analyses, analysis of variance (ANOVA) was followed by post hoc Bonferroni test, and significance was set at adjusted $P < 0.05$.

Single-cell data plotting

UMAPs and t-SNE plots from cells colored by different experimental features were randomly shuffled for plotting. UMAPs showing the expression level of marker genes and human malformation gene score values were generated from the normalized data slot of Seurat's RNA assay. Cells were plotted in order of expression with a minimum cutoff value of 10%.

Supplementary Materials

This PDF file includes:

Figs. S1 and S21

Legends for tables S1 to S6

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S6

REFERENCES AND NOTES

1. A. Llorca, O. Marín, Orchestrated freedom: New insights into cortical neurogenesis. *Curr. Opin. Neurobiol.* **66**, 48–56 (2021).
2. Y. Lin, J. Yang, Z. Shen, J. Ma, B. D. Simons, S. H. Shi, Behavior and lineage progression of neural progenitors in the mammalian cortex. *Curr. Opin. Neurobiol.* **66**, 144–157 (2021).
3. P. Gao, M. P. P. Postiglione, T. G. G. Krieger, L. Hernandez, C. Wang, Z. Han, C. Streicher, E. Papusheva, R. Insolera, K. Chugh, O. Kodish, K. Huang, B. D. D. Simons, L. Luo, S. Hippenmeyer, S.-H. H. Shi, Deterministic progenitor behavior and unitary production of neurons in the neocortex. *Cell* **159**, 775–788 (2014).
4. L. Telley, G. Agirman, J. Prados, N. Amberg, S. Fievre, P. Oberst, G. Bartolini, I. Vitali, C. Cadilhac, S. Hippenmeyer, L. Nguyen, A. Dayer, D. Jabaudon, S. Fièvre, P. Oberst, G. Bartolini, I. Vitali, C. Cadilhac, S. Hippenmeyer, L. Nguyen, A. Dayer, D. Jabaudon, Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364**, eaav2522 (2019).
5. T. J. Nowakowski, A. Bhaduri, A. A. Pollen, B. Alvarado, M. A. Mostajo-Radji, E. Di Lullo, M. Haeussler, C. Sandoval-Espinosa, S. J. Liu, D. Velmeshev, J. R. Ounadjela, J. Shuga, X. Wang, D. A. Lim, J. A. West, A. A. Leyrat, W. J. Kent, A. R. Kriegstein, Spatiotemporal gene expression trajectories reveal developmental hierarchies of the human cortex. *Science* **358**, 1318–1323 (2017).
6. V. Martinez-Cerdeno, S. C. Noctor, A. R. Kriegstein, The role of intermediate progenitor cells in the evolutionary expansion of the cerebral cortex. *Cereb. Cortex* **16**, i152–i161 (2006).
7. D. V. Hansen, J. H. Lui, P. R. L. L. Parker, A. R. Kriegstein, Neurogenic radial glia in the outer subventricular zone of human neocortex. *Nature* **464**, 554–561 (2010).
8. S. A. Fietz, I. Kelava, J. Vogt, M. Wilsch-Brauninger, D. Stenzel, J. L. Fish, D. Corbeil, A. Riehn, W. Distler, R. Nitsch, W. B. Huttner, M. Wilsch-Bräuninger, D. Stenzel, J. L. Fish, D. Corbeil, A. Riehn, W. Distler, R. Nitsch, W. B. Huttner, OSVZ progenitors of human and ferret neocortex are epithelial-like and expand by integrin signaling. *Nat. Neurosci.* **13**, 690–699 (2010).
9. M. Betizeau, V. Cortay, D. Patti, S. Pfister, E. Gautier, A. Bellemin-Ménard, M. Afanassieff, C. Huissoud, R. J. Douglas, H. Kennedy, C. Dehay, Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron* **80**, 442–457 (2013).
10. I. Reillo, C. De Juan Romero, M. Á. García-Cabezas, V. Borrell, A role for intermediate radial glia in the tangential expansion of the mammalian cerebral cortex. *Cereb. Cortex* **21**, 1674–1694 (2011).
11. M. Á. Martínez-Martínez, C. De Juan Romero, V. Fernández, A. Cárdenas, M. Götz, V. Borrell, A restricted period for formation of outer subventricular zone defined by Cdh1 and Trnp1 levels. *Nat. Commun.* **7**, 1–15 (2016).
12. N. Matsumoto, S. Tanaka, T. Horiike, Y. Shinmyo, H. Kawasaki, A discrete subtype of neural progenitor crucial for cortical folding in the gyrencephalic mammalian brain. *eLife* **9**, 1–26 (2020).
13. I. H. M. Smart, C. Dehay, P. Giroud, M. Berland, H. Kennedy, Unique morphological features of the proliferative zones and postmitotic compartments of the neural epithelium giving rise to striate and extrastriate cortex in the monkey. *Cereb. Cortex* **12**, 37–53 (2002).
14. M. Nonaka-Kinoshita, I. Reillo, B. Artegiani, M. Á. Martínez-Martínez, M. Nelson, V. V. Borrell, F. Calegari, Regulation of cerebral cortex size and folding by expansion of basal progenitors. *EMBO J.* **32**, 1817–1828 (2013).
15. A. A. Pollen, T. J. Nowakowski, J. Chen, H. Retallack, C. Sandoval-Espinosa, C. R. Nicholas, J. Shuga, S. J. J. Liu, M. C. Oldham, A. Diaz, D. A. Lim, A. A. Leyrat, J. A. West, A. R. Kriegstein, Molecular identity of human outer radial glia during cortical development. *Cell* **163**, 55–67 (2015).
16. D. Polioudakis, L. de la Torre-Ubieta, J. Langerman, A. G. Elkins, X. Shi, J. L. Stein, C. K. Vuong, S. Nichterwitz, M. Gevorgian, C. K. Opland, D. Lu, W. Connell, E. K. Ruzzo, J. K. Lowe, T. Hadzic, F. I. Hinz, S. Sabri, W. E. Lowry, M. B. Gerstein, K. Plath, D. H. Geschwind, L. de la Torre-Ubieta, J. Langerman, A. G. Elkins, X. Shi, J. L. Stein, C. K. Vuong, S. Nichterwitz, M. Gevorgian, C. K. Opland, D. Lu, W. Connell, E. K. Ruzzo, J. K. Lowe, T. Hadzic, F. I. Hinz, S. Sabri, W. E. Lowry, M. B. Gerstein, K. Plath, D. H. Geschwind, A single-cell transcriptomic atlas of human neocortical development during mid-gestation. *Neuron* **103**, 785–801.e8 (2019).
17. D. J. Di Bella, E. Habibi, R. R. Stickels, G. Scalia, J. Brown, P. Yadollahpour, S. M. Yang, C. Abbate, T. T. Biancalani, E. Z. Macosko, F. Chen, A. Regev, P. Arlotta, Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **595**, 554–559 (2021).

18. Z. Li, W. A. Tyler, E. Zeldich, G. Santpere Baró, M. Okamoto, T. Gao, M. Li, N. Sestan, T. F. Haydar, Transcriptional priming as a conserved mechanism of lineage diversification in the developing mouse and human neocortex. *Sci. Adv.* **6**, eabd2068 (2020).
19. C. De Juan Romero, C. Bruder, U. Tomasello, J. M. Sanz-Anquela, V. V. Borrell, J. M. Sanz-Anquela, V. V. Borrell, Discrete domains of gene expression in germinal layers distinguish the development of gyrencephaly. *EMBO J.* **34**, 1859–1874 (2015).
20. L. Telley, S. Govindan, J. Prados, I. Stevant, S. Nef, E. Dermitzakis, A. Dayer, D. Jabaudon, Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science* **351**, 1443–1446 (2016).
21. R. Chazarra-Gil, S. Van Dongen, V. Y. Kiselev, M. Hemberg, Flexible comparison of batch correction methods for single-cell RNA-seq using BatchBench. *Nucleic Acids Res.* **49**, e42 (2021).
22. H. T. N. Tran, K. S. Ang, M. Chevrier, X. Zhang, N. Y. S. Lee, M. Goh, J. Chen, A benchmark of batch-effect correction methods for single-cell RNA sequencing data. *Genome Biol.* **21**, 12 (2020).
23. H. Shimojo, T. Ohtsuka, R. Kageyama, Oscillations in notch signaling regulate maintenance of neural progenitors. *Neuron* **58**, 52–64 (2008).
24. S. Ohnuma, W. A. Harris, Neurogenesis and the cell cycle. *Neuron* **40**, 199–208 (2003).
25. E. Taverna, M. Götz, W. B. Huttner, The cell biology of neurogenesis: Toward an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell Dev. Biol.* **30**, 465–502 (2014).
26. I. Reillo, V. V. Borrell, Germinal zones in the developing cerebral cortex of ferret: Ontogeny, cell cycle kinetics, and diversity of progenitors. *Cereb. Cortex* **22**, 2039–2054 (2012).
27. T. J. Nowakowski, A. A. Pollen, C. Sandoval-Espinosa, A. R. Kriegstein, Transformation of the radial glia scaffold demarcates two stages of human cerebral cortex development. *Neuron* **91**, 1219–1227 (2016).
28. M. Bilgic, Q. Wu, T. Suetsugu, A. Shitamukai, Y. Tsunekawa, T. Shimogori, M. Kadota, O. Nishimura, S. Kuraku, H. Kiyonari, F. Matsuzaki, Truncated radial glia as a common precursor in the late corticogenesis of gyrencephalic mammals. *eLife* **12**, 1–35 (2023).
29. E. R. Thomsen, J. K. Mich, Z. Yao, R. D. Hodge, A. M. Doyle, S. Jang, S. I. Shehata, A. M. Nelson, N. V. Shapovalova, B. P. Levi, S. Ramanathan, Fixed single-cell transcriptomic characterization of human radial glial diversity. *Nat. Methods* **13**, 87–93 (2015).
30. S. Vaid, J. G. Camp, L. Hersemann, C. Eugster Oegema, A. K. Heninger, S. Winkler, H. Brandl, M. Sarov, B. Treutlein, W. B. Huttner, T. Namba, A novel population of Hopx-dependent basal radial glial cells in the developing mouse neocortex. *Development* **145**, dev169276 (2018).
31. A. Kriegstein, S. Noctor, V. Martinez-Cerdeno, Patterns of neural stem and progenitor cell division may underlie evolutionary cortical expansion. *Nat. Rev. Neurosci.* **7**, 883–890 (2006).
32. R. C. Bandler, I. Vitali, R. N. Delgado, M. C. Ho, E. Dvoretskova, J. S. Ibarra Molinas, P. W. Frazel, M. Mohammadkhani, R. Machold, S. Maedler, S. A. Liddelow, T. J. Nowakowski, G. Fishell, C. Mayer, Single-cell delineation of lineage and genetic identity in the mouse brain. *Nature* **601**, 404–409 (2022).
33. T. Stuart, A. Butler, P. Hoffman, C. Hafemeister, E. Papalexi, W. M. Mauck, Y. Hao, M. Stoeckius, P. Smibert, R. Satija, Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).
34. D. E. Wagner, C. Weinreb, Z. M. Collins, J. A. Briggs, S. G. Megason, A. M. Klein, Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science* **360**, 981–987 (2018).
35. K. Boekhoorn, V. van Dis, E. Goedknegt, A. Sobel, P. J. Lucassen, C. C. Hoogenraad, The microtubule destabilizing protein stathmin controls the transition from dividing neuronal precursors to postmitotic neurons during adult hippocampal neurogenesis. *Dev. Neurobiol.* **74**, 1226–1242 (2014).
36. C. A. Jackson, J. D. Peduzzi, T. L. Hickey, J. D. Peduzzi, T. L. Hickey, Visual cortex development in the ferret. I. Genesis and migration of visual cortical neurons. *J. Neurosci.* **9**, 1242–1253 (1989).
37. A. J. Barkovich, R. Guerrini, R. I. Kuzniecky, G. D. Jackson, W. B. Dobyns, A developmental and genetic classification for malformations of cortical development: Update 2012. *Brain* **135**, 1348–1369 (2012).
38. L. Del-Valle-Anton, V. Borrell, Folding brains: From development to disease modeling. *Physiol. Rev.* **102**, 511–550 (2022).
39. M. Bertacchi, A. L. Romano, A. Loubat, F. Tran Mau-Them, M. Willems, L. Faivre, P. Khau van Kien, L. Perrin, F. Devillard, A. Sorlin, P. Kuentz, C. Philippe, A. Garde, F. Neri, R. Di Giaimo, S. Oliviero, S. Cappello, L. D'Incerti, C. Frassoni, M. Studer, NR2F1 regulates regional progenitor dynamics in the mouse neocortex and cortical gyrification in BBSOAS patients. *EMBO J.* **39**, e104163 (2020).
40. S. Han, S. Okawa, G. A. Wilkinson, H. Ghazale, L. Adnani, R. Dixit, L. Tavares, I. Faisal, M. J. Brooks, V. Cortay, D. Zinyk, A. Sivitilli, S. Li, F. Malik, Y. Ilnytskyy, V. E. Angarica, J. Gao, V. Chinchalongporn, A. M. Oproescu, L. Vasan, Y. Touahri, L. A. David, E. Raharjo, J. W. Kim, W. Wu, W. Rahmani, J. A.-W. Chan, I. Kovalchuk, L. Attisano, D. Kurrasch, C. Dehay, A. Swaroop, D. S. Castro, J. Biernaskie, A. del Sol, C. Schuurmans, Proneural genes define ground-state rules to regulate neurogenic patterning and cortical folding. *Neuron* **109**, 2847–2863.e11 (2021).
41. R. L. Sidman, P. Rakic, Neuronal migration, with special reference to developing human brain: A review. *Brain Res.* **62**, 1–35 (1973).
42. V. Borrell, In vivo gene delivery to the postnatal ferret cerebral cortex by DNA electroporation. *J. Neurosci. Methods* **186**, 186–195 (2010).
43. S. Andrews, FastQC: A quality control tool for high throughput sequence data (2010); www.bioinformatics.babraham.ac.uk/projects/fastqc.
44. M. B. Johnson, X. Sun, A. Kodani, R. Borges-Monroy, K. M. Girskis, S. C. Ryu, P. P. Wang, K. Patel, D. M. Gonzalez, Y. M. Woo, Z. Yan, B. Liang, R. S. Smith, M. Chatterjee, D. Coman, X. Papademetris, L. H. Staib, F. Hyder, J. B. Mandeville, P. E. Grant, K. Im, H. Kwak, J. F. Engelhardt, C. A. Walsh, B.-I. I. Bae, Aspm knockout ferret reveals an evolutionary mechanism governing cerebral cortical size. *Nature* **556**, 370–375 (2018).
45. C. Trapnell, B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, L. Pachter, Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).
46. V. Kodali, cthreepo (version 0.1.1) [Source code] (2017); https://github.com/vkkodali/cthreepo.
47. D. Kotliar, [Source code] (2018); https://github.com/Hoohm/dropSeqPipe/issues/30#issuecomment-385957912.
48. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. A. T. L. Lun, S. Riesenfeld, T. Andrews, T. P. Dao, T. Gomes; participants in the 1st Human Cell Atlas Jamboree, J. C. Marioni, EmptyDrops: Distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. *Genome Biol.* **20**, 63 (2019).
50. A. Pickering, [Source code] (2020); https://github.com/MarioniLab/DropletUtils/issues/36#issue-564741989.
51. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 296 (2019).
52. L. Zappia, A. Oshlack, Clustering trees: A visualization for evaluating clusterings at multiple resolutions. *Gigascience.* **7**, giy083 (2018).
53. S. A. Yuzwa, M. J. Borrett, B. T. Innes, A. Voronova, T. Ketela, D. R. Kaplan, G. D. Bader, F. D. Miller, Developmental emergence of adult neural stem cells as revealed by single-cell transcriptional profiling. *Cell Rep.* **21**, 3970–3986 (2017).
54. M. Florio, M. Heide, A. Pinson, H. Brandl, M. Albert, S. Winkler, P. Wimberger, W. B. Huttner, M. Hiller, Evolution and cell-type specificity of human-specific genes preferentially expressed in progenitors of fetal neocortex. *eLife* **7**, e32332 (2018).
55. L. Lim, D. Mi, A. Llorca, O. Marin, O. Marín, Development and functional diversification of cortical interneurons. *Neuron* **100**, 294–313 (2018).
56. D. Mi, Z. Li, L. Lim, M. Li, M. Moissidis, Y. Yang, T. Gao, T. X. Hu, T. Pratt, D. J. Price, N. Sestan, O. Marín, Early emergence of cortical interneuron diversity in the mouse embryo. *Science* **360**, 81–85 (2018).
57. D. Mizrak, H. M. Levitin, A. C. Delgado, V. Crotet, J. Yuan, Z. Chaker, V. Silva-Vargas, P. A. Sims, F. Doetsch, single-cell analysis of regional differences in adult V-SVZ neural stem cell lineages. *Cell Rep.* **26**, 394–406.e5 (2019).
58. S. Vaid, W. B. Huttner, Transcriptional regulators and human-specific/primate-specific genes in neocortical neurogenesis. *Int. J. Mol. Sci.* **21**, E4614 (2020).
59. U. C. Eze, A. Bhaduri, M. Haeussler, T. J. Nowakowski, A. R. Kriegstein, Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat. Neurosci.* **24**, 584–594 (2021).
60. M. Safran, N. Rosen, M. Twik, R. BarShir, T. I. Stein, D. Dahary, S. Fishilevich, D. Lancet, The GeneCards suite, in *Practical Guide to Life Science Databases*, I. Abugessaisa, T. Kasukawa, Eds. (Springer Nature, 2022; pp. 27–56.
61. C. P. Wonders, S. A. Anderson, The origin and specification of cortical interneurons. *Nat. Rev. Neurosci.* **7**, 687–696 (2006).
62. V. Martínez-Cerdeño, C. L. Cunningham, J. Camacho, J. L. Antczak, A. N. Prakash, M. E. Cziep, A. I. Walker, S. C. Noctor, V. Martinez-Cerdeno, C. L. Cunningham, J. Camacho, J. L. Antczak, A. N. Prakash, M. E. Cziep, A. I. Walker, S. C. Noctor, Comparative analysis of the subventricular zone in rat, ferret and macaque: Evidence for an outer subventricular zone in rodents. *PLOS ONE* **7**, e30178 (2012).
63. B. Cauli, X. Zhou, L. Tricoire, X. Toussay, J. F. Staiger, Revisiting enigmatic cortical calretinin-expressing interneurons. *Front. Neuroanat.* **8**, 52 (2014).
64. M. B. Johnson, P. P. Wang, K. D. Atabay, E. A. Murphy, R. N. Doan, J. L. Hecht, C. A. Walsh, Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat. Neurosci.* **18**, 637–646 (2015).
65. M. Krishna-K, F. Nuernberger, C. Weth, Layer-specific expression of multiple cadherins in the developing visual cortex (V1) of the ferret. *Cereb. Cortex* **19**, 388–401 (2009).

66. B. J. Molyneaux, P. Arlotta, R. M. Fame, J. L. MacDonald, K. L. MacQuarrie, J. D. Macklis, Novel subtype-specific genes identify distinct subpopulations of callosal projection neurons. *J. Neurosci.* **29**, 12343–12354 (2009).
67. J. J. Rowell, A. K. Mallik, J. Dugas-Ford, C. W. Ragsdale, Molecular analysis of neocortical layer structure in the ferret. *J. Comp. Neurol.* **518**, 3272–3289 (2010).
68. H. Zeng, E. H. Shen, J. G. Hohmann, S. W. Oh, A. Bernard, J. J. Royall, K. J. Glattfelder, S. M. Sunkin, J. A. Morris, A. L. Guillozet-Bongaarts, K. A. Smith, A. J. Ebbert, B. Swanson, L. Kuan, D. T. Page, C. C. Overly, E. S. Lein, M. J. Hawrylycz, P. R. Hof, T. M. Hyde, J. E. Kleinman, A. R. Jones, Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell* **149**, 483–496 (2012).
69. K. Toma, C. Hanashima, Switching modes in corticogenesis: Mechanisms of neuronal subtype transitions and integration in the cerebral cortex. *Front. Neurosci.* **9**, 274 (2015).
70. B. B. Lake, R. Ai, G. E. Kaeser, N. S. Salathia, Y. C. Yung, R. Liu, A. Wildberg, D. Gao, H.-L. Fung, S. Chen, R. Vijayaraghavan, J. Wong, A. Chen, X. Sheng, F. Kaper, R. Shen, M. Ronaghi, J.-B. Fan, W. Wang, J. Chun, K. Zhang, Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* **352**, 1586–1590 (2016).
71. T. Nomura, W. Yamashita, H. Gotoh, K. Ono, Species-specific mechanisms of neuron subtype specification reveal evolutionary plasticity of amniote brain development. *Cell Rep.* **22**, 3142–3151 (2018).
72. S. Tutukova, V. Tarabykin, L. R. Hernandez-Miranda, The role of neurod genes in brain development, function, and disease. *Front. Mol. Neurosci.* **14**, 662774 (2021).
73. J. Griss, G. Viteri, K. Sidiropoulos, V. Nguyen, A. Fabregat, H. Hermjakob, ReactomeGSA - Efficient multi-omics comparative pathway analysis. *Mol. Cell. Proteomics* **19**, 2115–2125 (2020).
74. SatijaLab, seurat-wrappers (version 0.3.0) [Source code] (2007); https://github.com/satijalab/seurat-wrappers.
75. C. Trapnell, D. Cacchiarelli, J. Grimsby, P. Pokharel, S. Li, M. Morse, N. J. Lennon, K. J. Livak, T. S. Mikkelsen, J. L. Rinn, The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32**, 381–386 (2014).
76. X. Qiu, Q. Mao, Y. Tang, L. Wang, R. Chawla, H. A. Pliner, C. Trapnell, Reversed graph embedding resolves complex single-cell trajectories. *Nat. Methods* **14**, 979–982 (2017).
77. J. Cao, M. Spielmann, X. Qiu, X. Huang, D. M. Ibrahim, A. J. Hill, F. Zhang, S. Mundlos, L. Christiansen, F. J. Steemers, C. Trapnell, J. Shendure, The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**, 496–502 (2019).
78. S. Ding, X. Wu, G. Li, M. Han, Y. Zhuang, T. Xu, Efficient transposition of the piggyBac (PB) transposon in mammalian cells and mice. *Cell* **122**, 473–483 (2005).
79. B. Bushnell, J. Rood, E. Singer, BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE* **12**, e0185056 (2017).
80. T. Smith, A. Heger, I. Sudbery, UMI-tools: Modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
81. C. S. McGinnis, L. M. Murrow, Z. J. Gartner, DoubletFinder: Doublet detection in single-cell rna sequencing data using artificial nearest neighbors. *Cell Syst.* **8**, 329–337.e4 (2019).
82. Y. Hao, S. Hao, E. Andersen-Nissen, W. M. Mauck, S. Zheng, A. Butler, M. J. Lee, A. J. Wilk, C. Darby, M. Zager, P. Hoffman, M. Stoeckius, E. Papalexi, E. P. Mimitou, J. Jain, A. Srivastava, T. Stuart, L. M. Fleming, B. Yeung, A. J. Rogers, J. M. McElrath, C. A. Blish, R. Gottardo, P. Smibert, R. Satija, Integrated analysis of multimodal single-cell data. *Cell* **1–15**, (2021).
83. A. Butler, P. Hoffman, P. Smibert, E. Papalexi, R. Satija, Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).
84. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015).
85. P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
86. S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, W. Huber, BioMart and Bioconductor: A powerful link between biological databases and microarray data analysis. *Bioinformatics* **21**, 3439–3440 (2005).
87. S. Durinck, P. T. Spellman, E. Birney, W. Huber, Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
88. A. Vertesy, Seurat.Utils [Source code] (2007); https://github.com/vertesy/Seurat.utils.
89. M. Tang, [Source code] (2020); https://divingintogeneticsandgenomics.rbind.io/post/stacked-violin-plot-for-visualizing-single-cell-data-in-seurat/.
90. R. Kolde, (version 1.0.12) [Source code] (2018). https://github.com/raivokolde/pheatmap.
91. S. Bizzotto, F. Francis, Morphological and functional aspects of progenitors perturbed in cortical malformations. *Front. Cell. Neurosci.* **9**, 30 (2015).
92. M. Breuss, J. Morandell, S. Nimpf, T. Gstrein, M. Lauwers, T. Hochstoeger, A. Braun, K. Chan, E. R. Sánchez Guajardo, L. Zhang, M. Suplata, K. G. Heinze, K. Elsayad, D. A. Keays, The expression of tubb2b undergoes a developmental transition in murine cortical neurons. *J. Comp. Neurol.* **523**, 2161–2186 (2015).
93. M. W. Breuss, I. Leca, T. Gstrein, A. H. Hansen, D. A. Keays, Tubulins and brain development – The origins of functional specification. *Mol. Cell. Neurosci.* **84**, 58–67 (2017).
94. D. M. Romero, N. Bahi-Buisson, F. Francis, Genetics and mechanisms leading to human cortical malformations. *Semin. Cell Dev. Biol.* **76**, 33–75 (2018).
95. J. Ferent, D. Zaidi, F. Francis, Extracellular control of radial glia proliferation and scaffolding during cortical development and pathology. *Front. Cell Dev. Biol.* **8**, 578341 (2020).
96. R. Oegema, T. S. Barakat, M. Wilke, K. Stouffs, D. Amrom, E. Aronica, N. Bahi-Buisson, V. Conti, A. E. Fry, T. Geis, D. G. Andres, E. Parrini, I. Pogledic, E. Said, D. Soler, L. M. Valor, M. S. Zaki, G. Mirzaa, W. B. Dobyns, O. Reiner, R. Guerrini, D. T. Pilz, U. Hehr, R. J. Leventer, A. C. Jansen, G. M. S. Mancini, N. Di Donato, International consensus recommendations on the diagnostic work-up for malformations of cortical development. *Nat. Rev. Neurol.* **16**, 618–635 (2020).
97. Online Mendelian Inheritance in Man (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, 2020); https://omim.org/.
98. S. Marsh, scCustomize: An R package for custom visualization & analyses of single cell sequencing (2021); https://doi.org/10.5281/zenodo.5706430.
99. N. Renier, Z. Wu, D. J. Simon, J. Yang, P. Ariel, M. Tessier-Lavigne, iDISCO: A simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* **159**, 896–910 (2014).
100. J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: An open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
101. D. Legland, I. Arganda-Carreras, P. Andrey, MorphoLibJ: Integrated library and plugins for mathematical morphology with ImageJ. *Bioinformatics* **32**, 3532–3534 (2016).
102. J. Ollion, J. Cochennec, F. Loll, C. Escudé, T. Boudier, TANGO: A generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinformatics* **29**, 1840–1841 (2013).
103. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
104. X. Peng, J. Alföldi, K. Gori, A. J. Eisfeld, S. R. Tyler, J. Tisoncik-Go, D. Brawand, G. L. Law, N. Skunca, M. Hatta, D. J. Gasper, S. M. Kelly, J. Chang, M. J. Thomas, J. Johnson, A. M. Berlin, M. Lara, P. Russell, R. Swofford, J. Turner-Maier, S. Young, T. Hourlier, B. Aken, S. Searle, X. Sun, Y. Yi, M. Suresh, T. M. Tumpey, A. Siepel, S. M. Wisely, C. Dessimoz, Y. Kawaoka, B. W. Birren, K. Lindblad-Toh, F. Di Palma, J. F. Engelhardt, R. E. Palermo, M. G. Katze, The draft genome sequence of the ferret (*Mustela putorius furo*) facilitates study of human respiratory disease. *Nat. Biotechnol.* **32**, 1250–1255 (2014).
105. W. J. Kent, BLAT—The BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).
106. W. J. Kent, C. W. Sugnet, T. S. Furey, K. M. Roskin, T. H. Pringle, A. M. Zahler, D. Haussler, The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
107. K. D. Pruitt, T. Tatusova, D. R. Maglott, NCBI Reference Sequence (RefSeq): A curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **33**, D501–D504 (2005).
108. K. D. Pruitt, G. R. Brown, S. M. Hiatt, F. Thibaud-Nissen, A. Astashyn, O. Ermolaeva, C. M. Farrell, J. Hart, M. J. Landrum, K. M. McGarvey, M. R. Murphy, N. A. O'Leary, S. Pujar, B. Rajput, S. H. Rangwala, L. D. Riddick, A. Shkeda, H. Sun, P. Tamez, R. E. Tully, C. Wallin, D. Webb, J. Weber, W. Wu, M. DiCuccio, P. Kitts, D. R. Maglott, T. D. Murphy, J. M. Ostell, RefSeq: An update on mammalian reference sequences. *Nucleic Acids Res.* **42**, D756–D763 (2014).

Acknowledgments: Annotated data are available at https://in.umh-csic.es/en/cortevo/. We thank R. Borges-Monroy and C. A. Walsh for sharing ferret gene annotation file; S. Choudhary for help with gene annotation; F. Francis and S. Cappello for sharing MCD gene lists; and M. Florio for tips on single-cell suspension. We thank E. Sánchez-Pérez, C. Monzó-Gallego, and E. Llorens for ferret handling; A. J. Caler-Escribano, J.-P. López-Atalaya, A. Márquez-Galera, and R. Soler for technical support and advice; and A. Sempere for help with statistics. We also thank members of our laboratory for discussions and critical reading of the manuscript. **Funding:** This work was supported by ERC-AdG grant Neurocentro-885382 to M.G. and by Spanish State Research Agency grants SAF2015-69168-R, PGC2018-102172-B, and PDI2021-125618NB and "Severo Ochoa" Programme for Centers of Excellence in R&D (CEX-2021-00165-S) to V.B. Additional support was provided by Spanish State Research Agency FPI contract (BES-2016-077737) to L.D.-V.-A., "JdC incorporación" fellowship (IJC2020-044653-I) to V.F., and by Fundación Tatiana Pérez de Guzmán el Bueno fellowship to A.P.-C. **Authors contributions:** L.D.-V.-A. and V.B. conceived the project and designed experiments. L.D.-V.-A. and S.A. performed scRNA-seq experiments. C.F. prepared and quality-controlled the TrackerSeq library. E.D. performed the TrackerSeq experiments. F.N. and Y.K.B. analyzed the TrackerSeq data. D.C. performed the RNAscope experiments.

L.D.-V.-A. analyzed the scRNA-seq and RNAscope data. C.G.-F. and A.P.-C. cloned the ferret homologs. Y.N. and A.P.-C. performed the conventional ISH. V.F., A.C., and V.B. performed the electroporation experiments. R.M.-R. performed iDisco experiments. V.F. performed immunocytochemistry and analyzed iDisco experiments. J.A.M.-B., M.G., C.M., and V.B. provided reagents and resources. V.B. supervised experiments, provided funding, and wrote the manuscript with input from all other authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO SuperSeries accession number GSE234305. There are no restrictions on data availability.

Submitted 12 January 2024
Accepted 22 February 2024
Published 27 March 2024
10.1126/sciadv.adn9998