Inhibitory neuron diversity originates from cardinal classes shared across germinal zones.

Christian Mayer[1,2,#], Christoph Hafemeister[2,#], Rachel C. Bandler[1,#], Robert Machold[1], Gord Fishell[1,*] and Rahul Satija[2,3,*]

[1]NYU Neuroscience Institute, Langone Medical Center, New York, NY, USA

[2]New York Genome Center, New York, NY, USA

[3]New York University, Center for Genomics and Systems Biology, New York NY, USA

[#] Equal contribution

* Corresponding authors

ABSTRACT

During embryogenesis, the ventral telencephalon gives rise to a remarkable variety of GABAergic cell types, from long-range projection neurons to locally-projecting interneurons. Our current understanding of how this diversity is generated is far from complete. Using Drop-seq to profile the single-cell transcriptomes of over 36,000 ventral telencephalic cells isolated from three distinct germinal zones (the ganglionic eminences), we found that regardless of their origin and ultimate fate, all lineages transit through one of three cardinal transcriptional programs that define the earliest branch points in cell fate determination. Thus, the ability of these germinal zones to generate distinct cells likely relies on the small number of genes that are differentially expressed within each of these regions.

INTRODUCTION

During cerebral development, a diverse ensemble of neurons originating from distinct germinal zones assembles to form local and long-range circuits. While excitatory (glutamatergic) pyramidal cells arise directly from the cortical primordium (pallium), inhibitory GABAergic neurons are derived from progenitor cells residing in the ventral telencephalon, primarily within the medial, caudal, and lateral ganglionic eminences (MGE, CGE, and LGE, respectively)[1-3]. Each ganglionic eminence (GE) is thought to give rise to non-overlapping functional groups of inhibitory neurons in the adult, including both projection neuron (PN) and interneuron (IN) cell types[4-7]. These cells are destined to populate a large number of divergent structures including

the cortex, striatum, globus pallidus, amygdala, and olfactory bulb. While a number of transcription factors have previously been shown to be critical for the development of inhibitory neurons [8-10], currently our understanding of the comprehensive transcriptional trajectories that direct cell fates within these populations is far from complete. For example, it is not known when fate restricted cardinal cell classes emerge during early cell differentiation, or by what genetic logic they are organized. Bulk RNA-sequencing does not allow for the resolution of cell type heterogeneity because it can only provide a population average[11,12]. The advent of single-cell RNA sequencing (scRNA-seq) and associated computational methods has been transformative for resolving developmental heterogeneity[13,14], particularly with new massively parallel technologies for transcriptomics[15,16]. Here we use Drop-seq to generate an unbiased view of the transcriptional landscape present in thousands of individual progenitor and precursor cells within the GEs and their subsequent developmental trajectories.

Transcriptional profiling of GE cells.

We manually dissected the MGE, CGE and LGE of E13.5 (MGE) and E14.5 (CGE, LGE) wild type mouse embryos, including the ventricular zone (VZ), subventricular zone (SVZ) and adjacent mantle zone (MZ) within each eminence (Fig. 1a). These embryonic stages correspond to periods of abundant neurogenesis within each of these structures[17-20]. After cell dissociation, we utilized Drop-seq[15] to sequence the transcriptomes of 10,123 single cells from the MGE, 11,956 from the CGE, and 14,658 from the LGE, using three independent biological replicates for each eminence. We first regressed out confounding sources of technical variation between single cells, including sequencing depth and library complexity (Supplementary Methods). We then used a curated gene set[22] to define a cell-cycle score for each cell in the dataset, allowing the assignment of a mitotic (M, S, G-phase) or postmitotic status (cell cycle score near zero) to every cell (Extended Data Fig. 1). This enabled us to perform latent variable regression to mitigate heterogeneity resulting from cell-cycle state, so that the downstream analysis would not be dominated by mitotic phase-specific gene expression[23].

To identify distinct groups of cells in our data, we applied a graph-based clustering algorithm after non-linear dimensional reduction with diffusion maps[24]. We identified small populations of excitatory neurons (*Neurod6* 2.6% of cells) and epithelial cells (*Igfbp7*; 0.7% of cells) (Fig. 1b), both of which were excluded from further analysis. The remaining 96.7% of cells were GE-

derived progenitors, many of which expressed GABAergic neural precursor marker genes (e.g., *Dlx1*; Fig. 1b). To our surprise, diffusion map analysis did not identify discrete clusters of progenitors corresponding to each eminence. Rather, the transcriptional heterogeneity of mitotic progenitors was distributed across a continuous developmental landscape (Fig. 1c). Supporting this observation, the expression of early, intermediate, and late marker genes was strongly associated with the top diffusion map coordinates (Fig. 1c). In addition, we obtained the same result using principal component analysis (Extended Data Fig. 2).

To establish a quantitative temporal account of differentiation programs within each eminence, we implemented an unsupervised procedure to 'order' single-cells based on their expression profiles. After calculating the diffusion map coordinates (DMC), we fit a principle curve through the resulting point cloud which summarizes the best path through this continuous landscape of differentiation[25]. We hypothesized that each cell's projection onto this curve therefore represents its progression along a continuous developmental trajectory (Fig. 1d). We validated the resulting maturation trajectory (MT) with multiple independent supervised analyses. First, we observed that cells early in the MT were mitotic, based on their previously calculated cell-cycle scores (Fig. 1e). Second, gene expression dynamics of canonical regulators along the MT strongly recapitulated known dynamics associated with neuronal maturation. More specifically, MT identified the temporal expression profile of early, intermediate, and late marker genes that roughly correlate with gene expression in the VZ, SVZ, and MZ, respectively (Fig. 1f-g). For instance, cells located early in MT showed high levels of the VZ marker genes *Nes* and *Fabp7*[26,27], cells located towards the middle of MT showed strong expression of the SVZ marker genes *Ccnd2*[28] and *Ascl1*[29], and cells located late in MT showed high levels of the MZ marker gene *Gad1*[30]. Therefore, our MT captures the spatiotemporal progression from highly proliferative radial glia progenitor cells in the VZ to neuronal precursor cells in the MZ.

A uniform sequence of gene expression in mitotic progenitors

To explore the dynamics of gene expression in the early stages of maturation, we computationally isolated mitotic progenitor cells using cell-cycle scores and the MT (Fig. 1e). We then identified the sequential patterns of gene expression characterizing the initial stages of cell differentiation in all three eminences (Extended Data Table 1). Using mutual information

to discover genes whose expression was dynamic across the MT, we found three robust and highly reproducible sequential waves of gene expression (Fig. 2a). These roughly correlated with the spatiotemporal progression from the VZ to the MZ (Extended Data Fig. 3) and validated the sensitivity of our approach. The transcriptional waves corresponded with known transitions between developmental stages, as cells transited from expressing self-renewing proliferative genes (e.g., *Fabp7*, *Nes*, *Hes1* and *Notch1*), to proneuronal genes (e.g., *Ascl1* and *Olig2*), to neurogenic genes (e.g., *Dlx1, Dlx2*, and *Dlx6,* and *Arx*; Fig. 2b). The unsupervised reconstruction of the MT confirms the ordering of numerous genetic pathways (Fig. 2b). For instance, we found that *Ascl1*, a transcription factor known to trigger cell-cycle exit and to promote differentiation[31], is one of the first genes to be upregulated in the second wave. Its upregulation is followed by some of its well-known targets, including the genes encoding the Notch ligand *Dll1*[32] and the homeobox gene *Dlx1*[33,34] (Fig. 2b). Thus, MT enables numerous genes to be placed with precision into a framework of a large number of developmentally regulated genes, many with yet unknown roles in neurogenesis. To our surprise, during development the vast majority of dynamically expressed genes followed a similar temporal sequence across all three eminences (Fig. 2a-b). This indicates that a common developmental program consisting of hundreds of genes uniformly directs the early mitotic stages of inhibitory neurogenesis. Notably, we obtained nearly identical results when we computed the trajectories for each GE independently (Extended Data Fig. 4).

The GEs are known to produce largely non-overlapping populations of inhibitory neurons, which can be clearly distinguished by subtype specific gene expression in adults[35,36]. Therefore, there must exist determinant factors that function within individual GEs alongside this common developmental program. To identify such determinant factors, we performed a differential expression analysis, focusing on cells at the earliest stage of MT (Supplementary Methods; Extended Data Table 2). Genes enriched in the MGE compared to the CGE/LGE included the transcription factors *Nkx2-1, Lhx6, Lhx8, Aff2, Zic1,* and *Lmo1* (Fig. 2c), many of which had been previously described as being upregulated in the MGE[10,37-40]. Transcription factor genes that were preferentially expressed in the CGE compared to the MGE included *Pax6, Nr2f1, Nr2f2, Nr2e1, Btg1, Lmcd1* and *Helt*, and those in the LGE compared to the MGE included *Pax6, Ebf1, Isl1, Foxp2, Lmcd1, Eno1* and *Hmga2.* Only a few transcription factors showed differential expression between the CGE and LGE, and these included *Nr2f2, Nr2f1,*

Ebf1, Dlx5, Zic1, and *Crebzf* (Fig. 2c-d). While these genes were differentially expressed even in the earliest progenitors, their temporal dynamics spanned the range of MT, with markers peaking both early (e.g. Nr2f1; CGE), and late (e.g. Lhx6; MGE). These genes and their maturation dynamics were largely consistent with ISH data (Extended Data Fig. 5). Taken together, our data supports a model where a small number of differentially expressed determinant factors controls eminence-specific subtype development and gradually directs divergent transcriptional programs.

Divergence of developmental programs

While we observed striking evidence for a universal MT traversed by mitotic cells, we next asked when cells diverge from this common developmental program. Using a bootstrapped minimum spanning tree (MST) approach (Fig. 3a) to detect potential fate bifurcations, we found that soon after cells became postmitotic the MT progressively divided into three branches (Fig. 3b; Supplementary Methods). We assigned cells to branches by traversing the final MST and annotating major splits (Fig. 3a, c). To identify ensembles of transcription factors that were expressed specifically within individual branches, we performed pairwise differential expression tests between branches (Extended Data Table 3). Genes marking branch 1 cells included well-known regulators of IN development and/or function, including *Maf, Npy, Arx, Dlx1, Tcf4, Sox6, Sox11* and *Dlx2*[41], reviewed in[42]. In contrast, genes enriched in branches 2 and 3 were largely devoid of IN marker genes, but were instead characterized by a large number of known PN marker genes, such as *Foxp1, Isl1, Ebf1, Meis2* and *Bcl11b* (Fig. 3d-e)[43]. These results provide evidence that after cell cycle exit, postmitotic cells from the GEs split into one IN and two PN branches. Notably, we saw no evidence for this bifurcation occurring in cycling progenitors, even when applying a supervised analysis with branch-dependent genes (Extended Data Fig. 6).

By E18.5 the majority of cells derived from the GEs have migrated either into the developing cortex or into distinct subcortical forebrain structures[44]. The vast majority of GE-derived neurons in the cortex ultimately become INs, whereas the majority of GE-derived neurons in subcortical brain structures become PNs[4,5,7,44]. Therefore, using a *Dlx6a:Cre* fate mapping strategy to collect GE-derived cells from the cortex and subcortex at E18.5, it is possible to enrich for interneurons and projection neurons, respectively. This allowed us to confirm that the

postmitotic branching we observed within the GEs reflects a split into IN and PN precursor cells. We trained a k-nearest-neighbor classifier to assign branch identity based on gene expression within the eminences, and applied this to our E18.5 cortex and subcortex data (Fig. 3f; Supplementary Methods). Consistent with our hypothesis, the majority of *Dlx6a:Cre;Ai9* positive cells from the cortex mapped to branch 1, while the majority of *Dlx6a:Cre;Ai9* positive cells from the subcortex mapped to branches 2 and 3 (Fig. 3g). Taken together, these results indicate that soon after GE cells become postmitotic, cardinal cell-classes emerge that likely represent IN and PN precursors.

We next sought to elucidate the underlying genetic logic of how each GE contributes to the generation of these cardinal cell-classes by determining their respective contributions to the three branches in our MST plot. Strikingly, each of the three resultant branches was composed of postmitotic cells from all three GEs (Fig. 3h-i). These results indicate that transcriptionally defined developmental branches demarcate the segregation of progenitors into IN and PN cardinal classes common to all GEs.

To search for further heterogeneity, we excluded the trunk and focused on the 3 branches, which represent the most differentiated GE cells at this time point. We then jointly clustered these cells using bootstrapped community detection (Supplementary Methods) and uncovered sources of heterogeneity that were highly reproducible across bootstraps, but below the cutoff for branch detection. This analysis reliably recapitulated the division of cells into three broad branches, but additionally, revealed that each of the three branches also contained multiple cell clusters (Fig. 4a-b). Strikingly, clusters within each branch were composed of cells from all three GEs (Fig. 4c), providing further evidence that the GEs produce common cardinal classes.

Previous studies provide evidence that after becoming postmitotic, GE progenitors begin to restrict their fate potentials[45]. To test if the cardinal classes present at E13.5/E14.5 prescribe the heterogeneity found later in development, we applied our clustering procedure to the E18.5 cortical and subcortical datasets (Extended Data Fig. 7). We then used a k-nearest-neighbor classifier to assign E18.5 neurons to each of the distinct cardinal cell-clusters identified at E13.5/E14.5 (Fig. 4d). The majority of neurons from each E18.5 cluster could be assigned to a single E13.5/E14.5 cluster. Conversely, each E13.5/E14.5 cluster gave rise to one or more

clusters within the E18.5 populations (Fig. 4e-f). These results demonstrate how fate restricted cardinal classes that exist within the GEs diversify with time to create further heterogeneity.

DISCUSSION

Our data reveals the transcriptional framework underlying the emergence of neuronal identity within the ventral forebrain. To our surprise, rather than finding evidence for distinct progenitor cell types[46,47], we found a high degree of transcriptional similarity among all early mitotic cells. As progenitors became postmitotic, our unsupervised approach revealed the emergence of a common set of cardinal classes within all three of the GEs marking the onset of cell type differentiation and specialization.

Instead of using candidate marker-based enrichment strategies to deconvolve heterogeneous populations, we took an unbiased strategy to comprehensively sequence GE cells and reconstruct their developmental trajectories. By performing low-coverage sequencing of tens of thousands of cells, our approach also prioritized breadth over a depth-based strategy, where small numbers of cells would be sequenced at very high coverage. Recent studies[21] have demonstrated that a breadth-based strategy as facilitated by Drop-seq is significantly more powerful for resolving cell types in complex tissue. Our precise and highly reproducible MT and branch detection demonstrate that droplet-based methods can reconstruct developmental processes as well. While we cannot preclude that deeper coverage, or more cells, would reveal additional sub-structures, it is clear that the prevalent sources of transcriptional heterogeneity within early development are common to all GEs.

While our unsupervised approach is very effective at detecting cardinal populations with highly correlated gene expression, it will not allow us to identify divergent populations that are driven by a single gene. Our findings support a model where a small number of differentially expressed genes work in concert with a common developmental program. These differentially expressed genes act as determinant factors to seed eminence-specific fates. This is consistent with the ability of single genes, such as *Nkx2-1* and *Zeb2*[48,49], to act as master regulators. Previous work studying the genesis of inhibitory cells in the ventral forebrain has indicated a role for birthdate, germinal zone of origin, mode of division, and lineage in shaping cell fate. The need to consider these factors in the context of our analysis is also evident. For instance,

while our analysis revealed three branches within the postmitotic cells derived from each GE, branch 2 from the MGE contained notably less cells. This likely reflects the fact that PNs from the MGE are largely born prior to E13.5[38]. Further studies will be needed to determine how single-cell gene expression maps to each of these parameters.

At what stage of differentiation do fate restricted cardinal classes emerge and by what genetic logic are they organized? One possibility could be that mitotic progenitors within each GE are already relegated to specific cardinal classes (Model 1; Extended Data Fig. 8). Alternatively, the generation of distinct cardinal classes might occur upon progenitors within each GE becoming postmitotic (Model 2; Extended Data Fig. 8). Our data revealed a third unexpected outcome, namely the existence of cardinal classes that are shared across all three GEs (Model 3; Extended Data Fig. 8). The high degree of similarity between the GEs early in development might be the result of evolutionary duplications of a single primordial eminence, which produced a common set of proto-neuronal cardinal classes. Further work will be needed to explore which neuronal subtypes are derived from common cardinal classes, what similarities exist between neurons derived from the same cardinal class but different GEs, and how this relates to their ultimate role within distinct neural circuits. Regardless, our results suggest that the common attributes of neural precursors that ultimately differentiate to participate in distinct circuits far outstrip their differences. As such, evolution has allowed the breadth of computational circuits within the brain to flourish exponentially.

REFERENCES

1. Anderson, S. A., Eisenstat, D. D., Shi, L. & Rubenstein, J. L. Interneuron migration from basal forebrain to neocortex: dependence on Dlx genes. *Science* **278,** 474–476 (1997).
2. Corbin, J. G., Nery, S. & Fishell, G. Telencephalic cells take a tangent: non-radial migration in the mammalian forebrain. *Nat. Neurosci.* **4,** 1177–1182 (2001).
3. Kriegstein, A. R. & Noctor, S. C. Patterns of neuronal migration in the embryonic cortex. *Trends Neurosci.* **27,** 392–399 (2004).
4. Nery, S., Fishell, G. & Corbin, J. G. The caudal ganglionic eminence is a source of distinct cortical and subcortical cell populations. *Nat. Neurosci.* **5,** 1279–1287 (2002).
5. Wichterle, H., Turnbull, D. H., Nery, S., Fishell, G. & Alvarez-Buylla, A. In utero fate mapping reveals distinct migratory pathways and fates of neurons born in the mammalian basal forebrain. *Development* **128,** 3759–3771 (2001).
6. Marin, O., Anderson, S. A. & Rubenstein, J. L. Origin and molecular specification of striatal interneurons. *J. Neurosci.* **20,** 6063–6076 (2000).
7. Fishell, G. & Rudy, B. Mechanisms of inhibition within the telencephalon: "where the wild things are". *Annu. Rev. Neurosci.* **34,** 535–567 (2011).

8. Sussel, L., Marin, O., Kimura, S. & Rubenstein, J. L. Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. *Development* **126,** 3359–3370 (1999).
9. Elias, L. A. B., Potter, G. B. & Kriegstein, A. R. A Time and a Place for Nkx2-1 in Interneuron Specification and Migration. *Neuron* **59,** 679–682 (2008).
10. Sandberg, M. *et al.* Transcriptional Networks Controlled by NKX2-1 in the Development of Forebrain GABAergic Neurons. *Neuron* **91,** 1260–1275 (2016).
11. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16,** 133–145 (2015).
12. Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C. & Teichmann, S. A. The technology and biology of single-cell RNA sequencing. *Mol. Cell* **58,** 610–620 (2015).
13. Trapnell, C. *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **32,** 381–386 (2014).
14. Trapnell, C. Defining cell types and states with single-cell genomics. *Genome Res.* **25,** 1491–1498 (2015).
15. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161,** 1202–1214 (2015).
16. Klein, A. M. *et al.* Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell* **161,** 1187–1201 (2015).
17. Inan, M., Welagen, J. & Anderson, S. A. Spatial and temporal bias in the mitotic origins of somatostatin- and parvalbumin-expressing interneuron subgroups and the chandelier subtype in the medial ganglionic eminence. *Cereb. Cortex* **22,** 820–827 (2012).
18. Miyoshi, G. *et al.* Genetic fate mapping reveals that the caudal ganglionic eminence produces a large and diverse population of superficial cortical interneurons. *J. Neurosci.* **30,** 1582–1594 (2010).
19. Miyoshi, G., Butt, S. J. B., Takebayashi, H. & Fishell, G. Physiologically distinct temporal cohorts of cortical interneurons arise from telencephalic Olig2-expressing precursors. *J. Neurosci.* **27,** 7786–7798 (2007).
20. van der Kooy, D. & Fishell, G. Neuronal birthdate underlies the development of striatal compartments. *Brain Res* **401,** 155–161 (1987).
21. Shekhar, K. *et al.* Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* **166,** 1308–1323.e30 (2016).
22. Tirosh, I. *et al.* Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* **352,** 189–196 (2016).
23. Buettner, F., Natarajan, K. N., Casale, F. P. & Proserpio, V. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nature* **33,** 155–160 (2015).
24. Haghverdi, L., Buettner, F. & Theis, F. J. Diffusion maps for high-dimensional single-cell analysis of differentiation data. *Bioinformatics* **31,** 2989–2998 (2015).
25. Petropoulos, S. *et al.* Single-Cell RNA-Seq Reveals Lineage and X Chromosome Dynamics in Human Preimplantation Embryos. *Cell* **165,** 1012–1026 (2016).
26. Lendahl, U., Zimmerman, L. B. & McKay, R. D. CNS stem cells express a new class of intermediate filament protein. *Cell* **60,** 585–595 (1990).
27. Anthony, T. E., Klein, C., Fishell, G. & Heintz, N. Radial glia serve as neuronal progenitors in all regions of the central nervous system. *Neuron* **41,** 881–890 (2004).
28. Petros, T. J., Bultje, R. S., Ross, M. E., Fishell, G. & Anderson, S. A. Apical versus Basal Neurogenesis Directs Cortical Interneuron Subclass Fate. *Cell Reports* **13,** 1090–1095

(2015).
29. Casarosa, S., Fode, C. & Guillemot, F. Mash1 regulates neurogenesis in the ventral telencephalon. *Development* **126,** 525–534 (1999).
30. Tamamaki, N. *et al.* Green fluorescent protein expression and colocalization with calretinin, parvalbumin, and somatostatin in the GAD67-GFP knock-in mouse. *J. Comp. Neurol.* **467,** 60–79 (2003).
31. Treutlein, B. *et al.* Dissecting direct reprogramming from fibroblast to neuron using single-cell RNA-seq. *Nature* **534,** 391–395 (2016).
32. Castro, D. S. *et al.* Proneural bHLH and Brn proteins coregulate a neurogenic program through cooperative binding to a conserved DNA motif. *Developmental Cell* **11,** 831–844 (2006).
33. Poitras, L., Ghanem, N., Hatch, G. & Ekker, M. The proneural determinant MASH1 regulates forebrain Dlx1/2 expression through the I12b intergenic enhancer. *Development* **134,** 1755–1765 (2007).
34. Schuurmans, C. *et al.* Sequential phases of cortical specification involve Neurogenin-dependent and -independent pathways. *EMBO J* **23,** 2892–2902 (2004).
35. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* (2016). doi:10.1038/nn.4216
36. Zeisel, A. *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347,** 1138–1142 (2015).
37. Batista-Brito, R. & Fishell, G. The developmental integration of cortical interneurons into a functional network. *Curr. Top. Dev. Biol.* **87,** 81–118 (2009).
38. Flandin, P. *et al.* Lhx6 and Lhx8 coordinately induce neuronal expression of Shh that controls the generation of interneuron progenitors. *Neuron* **70,** 939–950 (2011).
39. Neves, G. *et al.* The LIM homeodomain protein Lhx6 regulates maturation of interneurons and network excitability in the mammalian cortex. *Cereb. Cortex* **23,** 1811–1823 (2013).
40. Fogarty, M. *et al.* Spatial genetic patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult cortex. *J. Neurosci.* **27,** 10935–10946 (2007).
41. Batista-Brito, R., Machold, R., Klein, C. & Fishell, G. Gene Expression in Cortical Interneuron Precursors is Prescient of their Mature Function. *Cereb. Cortex* **18,** 2306–2317 (2008).
42. Butt, S. J. B. *et al.* Transcriptional regulation of cortical interneuron development. *J. Neurosci.* **27,** 11847–11850 (2007).
43. Stenman, J., Toresson, H. & Campbell, K. Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. *J. Neurosci.* **23,** 167–174 (2003).
44. Marin, O. & Rubenstein, J. L. R. Cell migration in the forebrain. *Annu. Rev. Neurosci.* **26,** 441–483 (2003).
45. McConnell, S. K. & Kaznowski, C. E. Cell cycle dependence of laminar determination in developing neocortex. *Science* **254,** 282–285 (1991).
46. Franco, S. J. *et al.* Fate-restricted neural progenitors in the mammalian cerebral cortex. *Science* **337,** 746–749 (2012).
47. La Manno, G. *et al.* Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells. *Cell* **167,** 566–580.e19 (2016).
48. McKinsey, G. L. *et al.* Dlx1&2-dependent expression of Zfhx1b (Sip1, Zeb2) regulates the fate switch between cortical and striatal interneurons. *Neuron* **77,** 83–98 (2013).

49. Butt, S. J. B. *et al.* The requirement of Nkx2-1 in the temporal specification of cortical interneuron subtypes. *Neuron* **59,** 722–732 (2008).

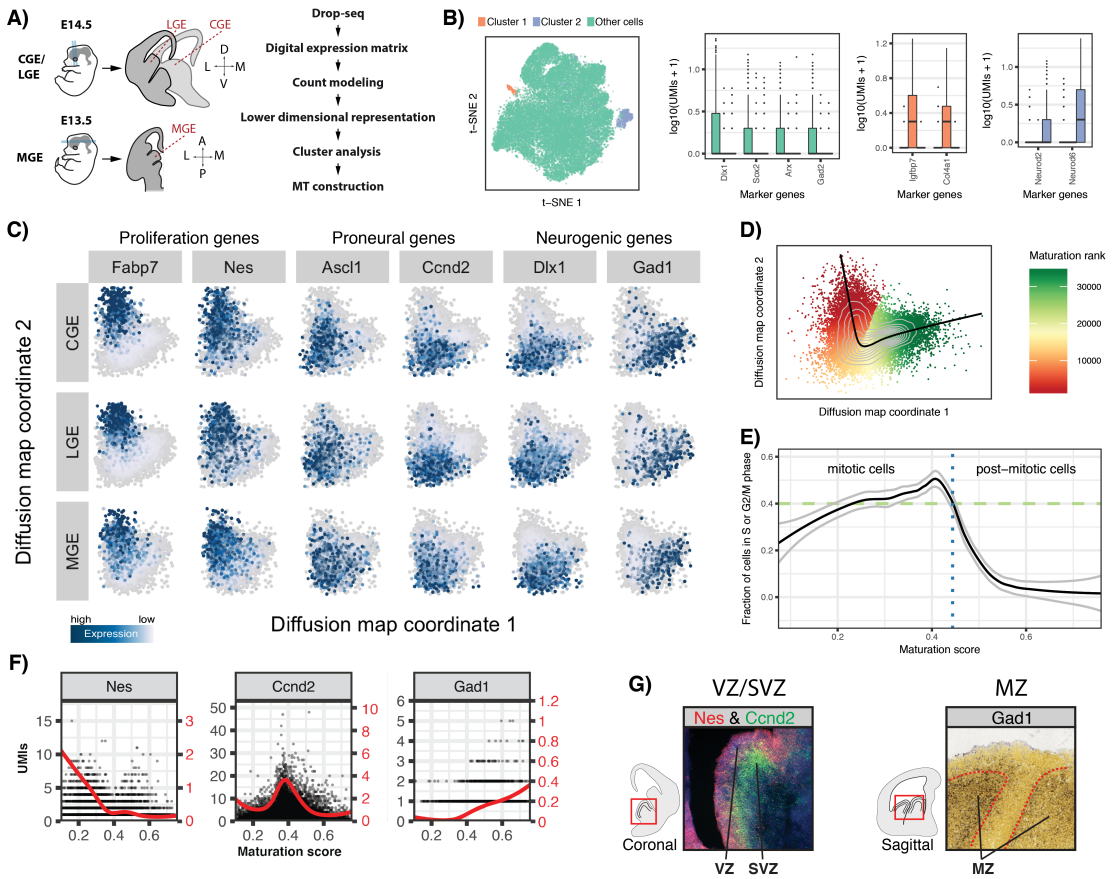

Figure 1. Transcriptional heterogeneity in the GEs

A) Schematic of experimental workflow. MGEs were dissected from horizontal brain sections, whereas the CGE/LGE were dissected from coronal brain sections, as indicated by the light blue lines. Axes: Dorsal (D), Ventral (V), Posterior (P), Anterior (A), Lateral (L), Medial (M). **B)** tSNE and box plots depicting the presence of distinct cell types in Drop-seq data. The vast majority of cells identified were GE precursors (green), while a small minority of cells represented excitatory neurons (blue) and vascular endothelial cells (orange) based on canonical marker expression. **C)** Diffusion map analysis of eminence datasets suggests a developmental continuum. Each eminence was analyzed independently, revealing nearly identical patterns. Cells are colored according to the expression of canonical regulators. **D)** A principal curve was fitted to the two dominant diffusion map coordinates to order cells along a maturation trajectory (MT). **E)** The percentage of actively cycling cells that exhibit a sharp transition (dotted-blue line), marking the shift during MT into a postmitotic phase. **F)** Black points: single-cell UMI count as a function of MT progression for a selection of genes. Red curve reflects the locally averaged trend line. **G)** In-situ hybridization (ISH) patterns of early, intermediate and late MT genes in the eminences that are highly expressed within anatomical boundaries of the Ventricular Zone (VZ), Subventricular Zone (SVZ), and Mantle Zone (MZ), respectively; The ISH image for Gad1 was taken from the Allen Brain Institute.

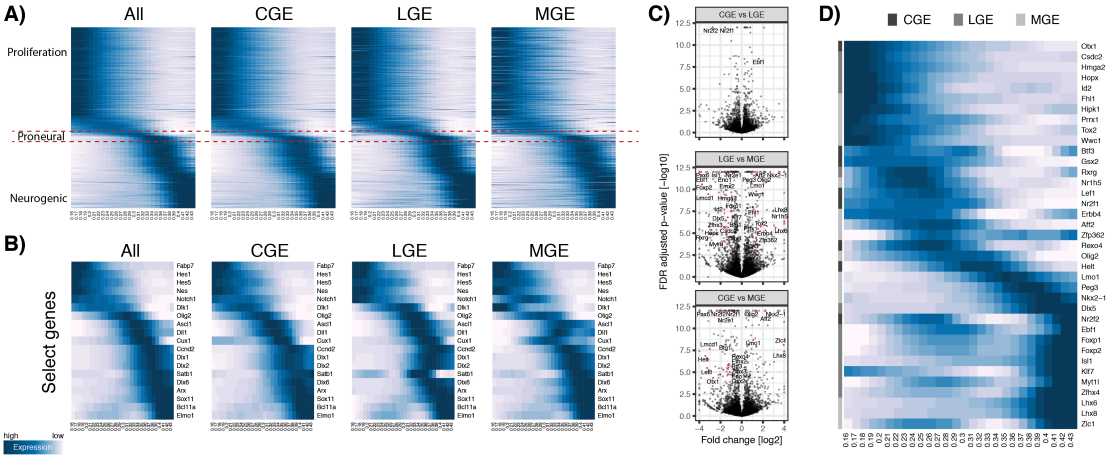

Figure 2. A common developmental program of gene expression functions in mitotic progenitors of all three GEs

Gene expression dynamics, based on local averaging of single cell data, plotted along MT for all **(A)** and select **(B)** developmentally regulated genes. These dynamics reveal sequential waves of gene expression that are consistent between eminences. **C)** Volcano plots depicting differential gene expression in early mitotic cells (MT < 0.3). Transcription factor gene names are annotated on the plot **D)** Heat-map of temporal dynamics for a representative selection of differentially expressed transcription factors in **(C)**. Trajectories for a gene are shown for the eminence where it is enriched, as indicated by the left-side label.

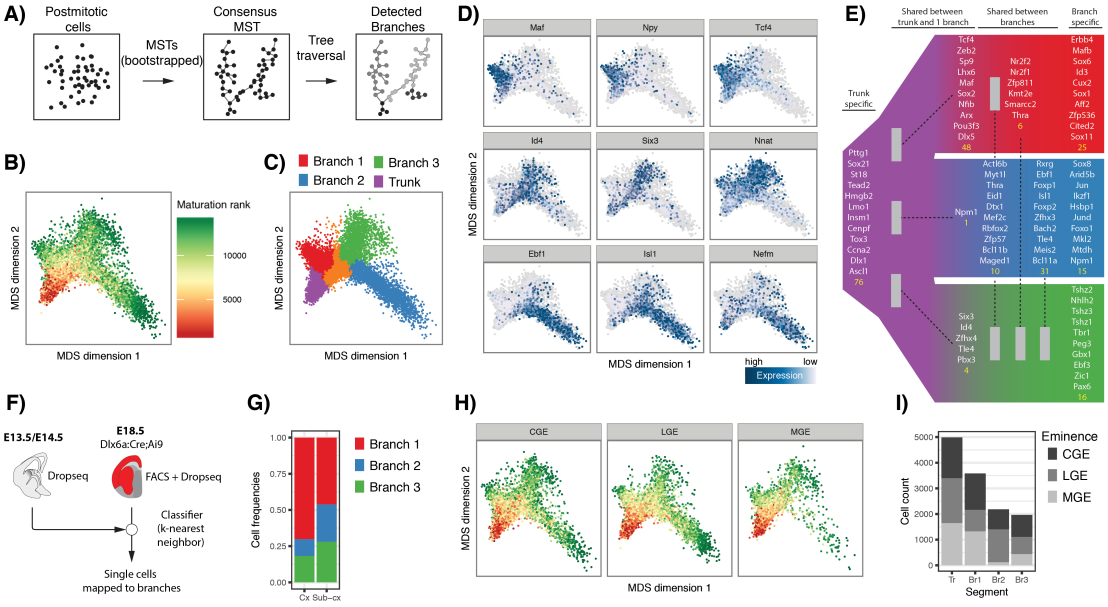

Figure 3. All eminences give rise to common cardinal cell-classes

A) Schematic depicting the Minimum Spanning Tree (MST) building and branch detection process. **B)** Multidimensional scaling (MDS) based on the consensus MST depicts a common trunk and three divergent developmental trajectories. **C)** By traversing the final MST and annotating major splits we assigned cells to the major branches; trunk, purple; branch 1, red; branch 2, blue; branch 3, green. **D)** Cells are colored according to their expression of branch-dependent genes. **E)** Diagram depicting differentially expressed genes between the trunk and branches of the MST. Yellow numbers indicate the total number of genes in each category. **F)** Schematic of experiment to confirm projection neuron (PN) and interneuron (IN) branches. **G)** Mapping and quantification of Dlx6a:Cre;Ai9 positive E18.5 subcortical and cortical cells to E13.5/E14.5 branches. **H)** Cells from each eminence can be found throughout the branching structure, with varying quantitative contributions **(I)** to each segment.

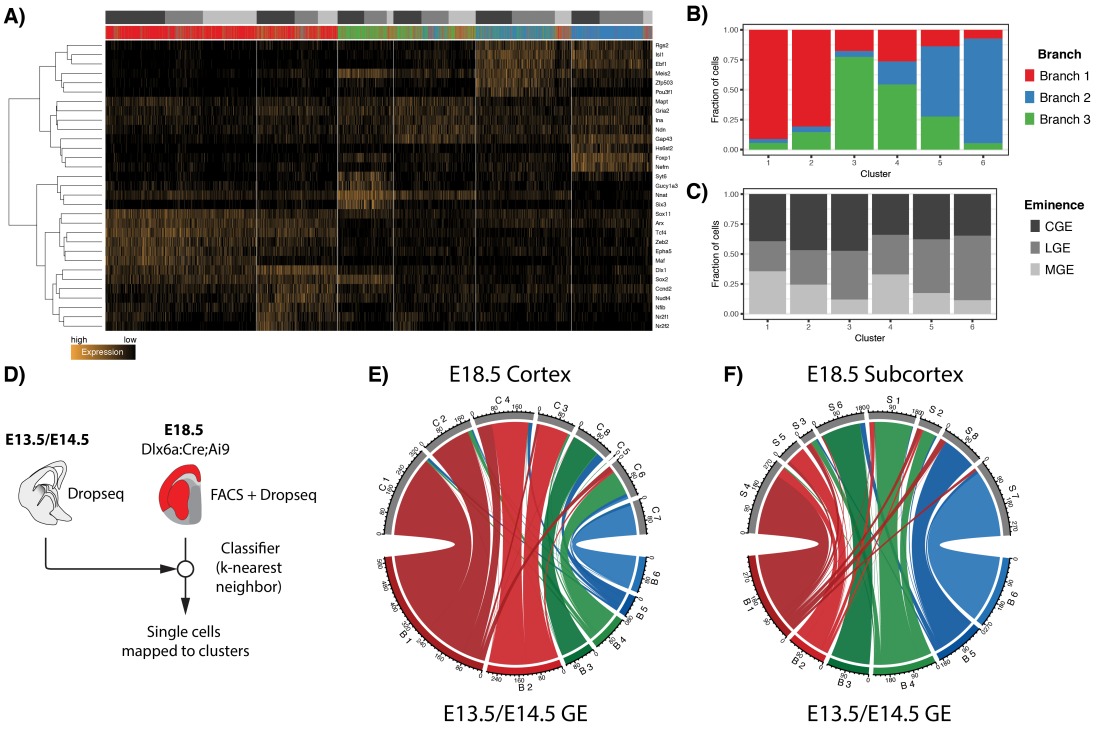

Figure 4. The cardinal classes of postmitotic GE progenitors are fate-restricted
A) Unsupervised clustering of cells in the three branches uncovers additional sources of heterogeneity ('cardinal clusters'). Heat map illustrates single-cell gene expression; top legend indicates branch ID and eminence of origin. **B & C)** Bar graphs depicting the contribution of each GE **(B)** and each branch **(C)** to the clusters shown in **(A)**. We identified that clusters segregate by branch but contain cells from all eminences. **D)** Schematic of experiment to test for potential fate restriction of cardinal classes. **E & F)** 'Circular' plots illustrating the mapping between E18.5 cortical **(E)** and subcortical **(F)** cells and the GE cardinal clusters. In the circular plots, each line connects an E18.5 cell with its designated E13.5/E14.5 cluster based on transcriptional similarity (KNN classifier).