Single cell enhancer activity distinguishes GABAergic and cholinergic lineages in embryonic mouse basal ganglia

Linda Su-Feher1,2,*, Anna N. Rubin1,2, Shanni N. Silberberg3, Rinaldo Catta-Preta3,2, Kenneth J. Lim4,2,5, A genes hera R. Ypsilanti1,*, Iva Zdilar1,2,*, Christopher S. McGinnis1,*, Gabriel L. McKinsey6, Thomas E. Rubino Jr.3,6, Michael J. Hawrylycz6,7,7, Zev J. Gartner4,6,7, Luis Puuelles8, Hongkui Zeng9,*, John L. R. Rubenstein1,2,3, and Alex S. Nord1,2,3

Enhancers integrate transcription factor signaling pathways that drive cell fate specification in the developing brain. We paired enhancer labeling and single-cell RNA-sequencing (scRNA-seq) to delineate and distinguish specification of neuronal lineages in mouse medial, lateral, and caudal ganglionic eminences (MGE, LGE, and CGE) at embryonic day (E)11.5. We show that scRNA-seq clustering using transcription factors improves resolution of regional and developmental populations, and that enhancer activities identify specific and overlapping GE-derived neuronal populations. First, we mapped the activities of seven evolutionarily conserved brain enhancers at single-cell resolution in vivo, finding that the selected enhancers had diverse activities in specific progenitor and neuronal populations across the GEs. We then applied enhancer-based labeling, scRNA-seq, and analysis of in situ hybridization data to distinguish transcriptionally distinct and spatially defined subtypes of MGE-derived GABAergic and cholinergic projection neurons and interneurons. Our results map developmental origins and specification paths underlying neurogenesis in the embryonic basal ganglia and showcase the power of scRNA-seq combined with enhancer-based labeling to resolve the complex relationships of neuronal specification underlying mouse brain development.

Significance

During brain development, neurons are generated by spatially and temporally distinct processes that remain to be fully characterized. The ganglionic eminences (GEs) in the embryonic subpallium give rise to GABAergic and cholinergic neuron lineages that form the basal ganglia or migrate to the cerebral cortex. Beyond a limited set of canonical RNA markers, the transcriptional states of GE progenitors and immature neurons cells remain poorly defined. We combine enhancer labeling, single-cell transcriptomics using transcription factor-anchored clustering, and integration with in situ hybridization data to distinguish emerging neuronal populations in embryonic mouse basal ganglia. Our results demonstrate the specificity of enhancer-based labeling at single-cell resolution and reveal developmental origins and specification processes of critical neuronal lineages.
BG-derived neuronal lineages at single-cell resolution. We profiled FACS enhancer-positive cells as well as unsorted or enhancer-negative cells to define enhancer activity in the context of embryonic BG populations. We used a transcription factor-anchored approach to derive cell clusters in our scRNA-seq data, which improved developmental relevance of identified clusters. Our study identified proliferative and postmitotic cells that are distinctly labeled by enhancers active in MGE, LGE, and CGE, and revealed specification pathways of enhancer-labeled and spatially defined populations of early BG-derived GABAergic and cholinergic lineages.

Results

Comparative Activity of Seven Enhancers in Embryonic Day 11.5 BG via scRNA-seq. To understand developmental origins of BG-arising neuronal populations, we profiled enhancer-labeled cells from embryonic day (E)11.5 MGE, LGE, or CGE for seven representative subpallial enhancer transgenic mouse lines (15) (Fig. 1A). The transgenic lines express GFP and CreERT2 with enhancer-driven divergent patterns in VZ, SVZ, and mantle (MZ) zones of the GEs, summarized in Fig. 1B. The objective of these experiments was threefold: first, to establish the utility and sensitivity to detect enhancer-driven reporter expression via scRNA-seq; second, to understand developmental origins of BG-arising neuronal populations in MGE (enhancers hs1538, hs1056, hs799, and hs192), LGE (hs841 and hs599), and CGE (hs841 and hs953) at E11.5; and third, to resolve developmental trajectories labeled by regionally distinct MGE progenitor (hs1538 and hs1056) and neuronal (hs799 and hs192) enhancers.

We performed BG microdissection and preparation of reporter-positive and ungated single cells from each enhancer line, as indicated in Fig. 1 A and B. Cell suspensions were either enriched for cells with transgene expression through FACS (SI Appendix, Fig. S1) or passed directly to the Fluidigm C1 system for capture and amplification of full-length mRNA from single cells. After sequencing and quality control (SI Appendix, Fig. S2 A–I), 1,027 cells were included for analysis, with ~594,000 reads and ~5,140 genes per cell on average.

We used a combination of FACS gating and transgene (CreERT2-ires-GFP or tdTomato) RNA expression to distinguish enhancer-labeled cells from cells with no enhancer activity: 315 cells were unsorted or FACS- and transgene negative; 712 cells were FACS+ or had detectable reporter transcripts. Enhancer-driven reporter transcripts were detectable at single-cell resolution across all enhancers, with variation in presence and transcript level captured via scRNA-seq (Fig. 1C). Five enhancers exhibited strong parallels between reporter protein GFP+ gating and transgene RNA expression to scRNA-seq (SI Appendix, Fig. S2A, Fig. S2M). The other two lines, dMGE-hs1538 and MGE-hs192, had weaker transgene RNA-seq sensitivity, with 30 to 40% of FACS reporter-positive cells having detectable transgene transcript. Nonetheless, even for enhancers with weaker transcriptional activity, reporter transgene was reliably detected via scRNA-seq in a substantial fraction of FACS-determined reporter-positive cells, demonstrating the overall utility of this approach and enabling comparison of enhancer-labeled cells. By analyzing both enhancer-positive and unlabeled cells, we tested how specific the enhancer activities were in cells within and across these BG lineages.

Transcription Factor Expression Organizes scRNA-seq Data by Proliferative State and BG Region. Next, we sought to define the cell types present in E11.5 BG via clustering using both enhancer-positive and unsorted cells. Using highly variable genes in scRNA-seq analysis is a common approach for feature selection (16); however, this method did not adequately separate regional and cell state identity in our data (SI Appendix, Fig. S3). As an alternative, we used a transcription factor (TF)-curated approach, with the rationale that TFs drive lineage

Fig. 1. Profiling enhancer-labeled single cells in E11.5 BG. (A) Schematic of dissection, with colors representing the activity of seven transgenic enhancer reporters characterized using C1 scRNA-seq. (B) Summary of the enhancers profiled, including dissection region, putative target gene, and representative GFP IHC imaging of enhancer transgenic reporters at E11.5, depicting activity within the ganglionic eminences. Red lines indicate microdissection boundaries. GFP IHC images reprinted with permission (15). (C) Violin plot of normalized transgene expression by enhancer. (D) Visualization of single cells by UMAP, colored by enhancer. Colors correspond to header colors in B. Enhancer-negative cells are depicted in gray.
specification and cell identity. We rooted this analysis using 689 TFs profiled for RNA ISH patterns at E11.5 and E13.5 in the Allen Developing Mouse Brain Atlas (ABA) (17) (SI Appendix, Figs. S4 A and B and Dataset S2). Of these TFs, 455 were expressed in our E11.5 scRNA-seq data, of which 292 (64.2%) had detectable ISH expression in the BG (SI Appendix, Fig. S4 C–E). We used these 455 TFs to define scRNA-seq cell identity and for visual representation via UMAP plots (Fig. 1 D–F). We compared TF-curated analysis to cluster identities using highly variable genes before and after performing regression analysis to reduce the influence of cell-cycle phase (SI Appendix, Figs. S3 A–E). Excluding non-TFs reduced the contribution of cell-cycle phase and confounding sources of variation (e.g., sequencing batch) while maintaining cell cycle (principal component [PC1]: 15.06% variance explained) and dissection (PC3: 4.48% variance explained) as the main sources of variance (SI Appendix, Fig. S3 F and G). Finally, we assessed clustering using only enhancer-positive cells, which largely reproduced results from the full set of enhancer-positive and unlabeled cells (SI Appendix, Fig. S3 H). In summary, we found that using TFs as features for clustering improved resolution of cell clusters by regional origin and proliferative state.

Having shown that scRNA-seq can reliably identify enhancer-positive cells and that our TF-curated approach enables separation across regional origin and proliferative states, we next modeled transcriptional differences across the full set of enhancer-positive and enhancer-negative cells. First, we modeled transitional cell states via diffusion mapping. Second, we examined differences across cluster-based transcriptional identities. We identified two major diffusion components (DC) corresponding to proliferative state (DC1) and MGE from LGE/CGE origin (DC2) (Fig. 2 A). DC1 captured the stem cell, progenitor, and neurogenic transition, with genes such as Hes1, Ccdn2a, Gadd45g, and Slc34a2 marking cells at various stages of this transition (Fig. 2 B). Lower values of DC2 were associated with MGE identity, marked by expression of MGE-specific genes, such as Nbc2-1, Lhx6, and Lhx8 (18) (Fig. 2 B). Higher values of DC2 were associated with LGE or CGE identity (Fig. 2 A and B). DC2 diversity was driven by expression of region-defining TFs, such as Nbc2-1 in the MGE and Pax6 in the CGE and LGE (Fig. 2 B). DC1 and DC2 distinguished cells labeled by different enhancers and indicate that these developmental enhancers are active across maturation states within the GEs.

We next performed clustering using TF-curated scRNA-seq expression, identifying 12 cell clusters that separated by proliferative state and regional or cell-type identity (Fig. 2 C and Dataset S3). We further used random forest classification to define informative transcripts that discriminate cells labeled by specific enhancers (Dataset S4). Cells labeled by enhancers dMGE-hs1538, MGE-hs1056, CGE-hs841, and LGE-hs841 primarily grouped into mitotic clusters (cl-1), cl-2, cl-3, cl-8, and cl-9, further separated by regional identity (MGE vs. non-MGE). Within these regional boundaries, VZ/SVZ-associated enhancers split across multiple clusters (Fig. 2 D), paralleling diffusion mapping results suggesting mitotic enhancers label multiple proliferative states. Compared to enhancers with progenitor activity, enhancers active in postmitotic cells (MGE-hs192, MGE-hs799, LGE-hs599, and CGE-hs953) were biased toward specific cell-type clusters within broader regional identities (Fig. 2 D). To characterize cell types that were differentially labeled by these enhancers, we performed differential gene-expression analysis using the full transcriptome of 17,015 expressed genes to identify differentially expressed (DE) genes for each TF-defined cluster (Fig. 2 E and Dataset S3).

Proliferative clusters cl-1, -3, -2, and -9 encompassed VZ and SVZ proliferating cells including those labeled by mitotic-associated enhancers MGE-hs1538, MGE-hs1056, and CGE- and LGE-hs841. Cells in cl-1 expressed neural stem cell markers including higher expression of Nefa, Nfib, and Olig2 (Fig. 2 E and F). cl-3 was associated with higher levels of intermediate progenitor markers such as Asb4 and Aset1. Proliferative LGE and CFE cells from h841 comprised the majority of cl-2 and cl-9 and expressed VZ markers including Lhx2 and Fzd8. Enhancer-negative LGE and LGE-hs841 mitotic cells additionally form clusters cl-8 and cl-11, which expressed genes such as Lrp4, Hap1, and Kitl. cl-10 was composed primarily of enhancer-negative cells from MGE dissections and expressed Lhx9 and Tbr1, indicative of cortical or diencephalic rather than BG identity. From random forest classification, markers that distinguished LGE, CGE, and MGE progenitor-associated enhancers recapitulated region-associated TFs from DC2 (Dataset S4). MGE rostroventral (hs1538) and caudoventral (hs1056) biased enhancers were distinguished by quantitative differences across TFs including Otx2 and Id4, identifying TF expression gradients that distinguished progenitor cells across MGE regional axes.

Compared to proliferative enhancers, postmitotic enhancers active in SVZ/MZ mapped to distinct transcriptional clusters corresponding to emerging neuron types. LGE-hs599 and CGE-hs953 are both represented in cl-4, which expressed higher levels of genes, including Sp8, Tiam2, and Tshz1 (Fig. 2 E and F), suggesting cl-4 is more immature than cl-5 and cl-7 and has not yet acquired strong LGE or CGE regional specificity. cl-5, composed predominantly of LGE-hs599 cells, expressed genes including Rhfox1, Tac1, and Zfp503. cl-7, composed predominantly of CGE-hs953 and enhancer-negative cells, expressed genes such as Six3, Tshz2, and Sema3a. Genes defining these clusters shared general markers of early GABAergic projection neurons and were consistent with fate mapping of hs599+ and hs953+ cells to projection neuron populations in the adult forebrain, including striatal medium spiny neurons (15, 19) and Sp8+ neurons in the amygdala (15). MGE MZ-associated enhancers h192 and h799 also exhibited markers of early neuronal fate commitment (Fig. 2 E and Datasets S3 and S4). cl-0, composed predominantly of h799+ cells, expressed early MGE-derived cortical IN lineage markers, including Maf, Mafb, Neph1, Calb1, and Snr. Conversely, cl-6, composed of primarily h192+ cells, expressed a wide range of TFs including Tle4 and Zic1, Zic3, and Zic4, suggestive of GABAergic and cholineric projection neuron commitment (20).

These experiments mapped the cell type-specific activity of seven evolutionarily conserved enhancers across E11.5 GEs. We found regionally separated but otherwise similar mitotic identities among enhancer-labeled progenitors across MGE, LGE, and CGE, and localized regional signatures within MGE via comparing dorsal hs1538+ cells with more ventral hs1056+ cells. Across GEs, enhancer-labeled neuronal cell types emerged from more general early postmitotic clusters. Postmitotic cells in MGE separated into signatures suggesting GABAergic projection, cholineric, and IN trajectories that were differentially labeled by h192 and h799. This initial survey mapped the activity of these enhancers to distinct cell populations and identified known and novel markers for maturation state and regional identity for progenitor and early born neuronal cell types.

**Dissection of Early Born Neuron Types and Positional Identity in MGE via Enhancer-Labeled 3’ scRNA-seq and ISH.** To evaluate the impact of deeper sampling and to comprehensively interrogate emerging neuronal types in E11.5 MGE labeled by
hs192 and hs799, we performed 3' scRNA-seq using the 10x Genomics Chromium system. FACS-purified MGE un gated and reporter-positive cells were dissected from embryos across three litters for each transgenic line and prepared for multi-plexed scRNA-seq using MULTISeq. (21) (Fig. 3). 4,001 single cells were used for downstream analysis. We used the same TF-curated approach to drive cell clustering. Of the 463 TFs detected in the 3' scRNA-seq MGE data, 421 (90.9%) were represented in both full-length and 3' scRNA-seq approaches (Fig. 3B and SI Appendix, Fig. S4 A and B). We identified 18 cell clusters in the 3' scRNA-seq data (Fig. 3C). Consistent with the full-length scRNA-seq data, using TFs as features reduced separation driven by cell-cycle phase (SI Appendix, Fig. S5 C and D). The primary component of transcriptional variation (PC1: 24.24% variance explained) in TF-curated data was associated with proliferative state (SI Appendix, Fig. S5D). Finally, we tested clustering using only enhancer-labeled cells (SI Appendix, Fig. S5E), finding that postmitotic cluster structure remained stable. Mitotic clusters were not captured by the enhancer-positive cells, as expected considering these two enhancers primarily label postmitotic cells. Overall, the MGE-derived 3' scRNA-seq and full-length scRNA-seq data had similar cluster topology and identified similar cell states, as determined by canonical correlation analysis (SI Appendix, Fig. S1).

Cells were defined as enhancer-positive for hs192 or hs799 if they had at least one UMI (unique molecular identifier) count for CreER12-IRES-GFP. Enhancer-positive cells from independent samples clustered together, demonstrating reproducibility across biological replicates and distinct distributions of enhancer-labeled hs192 vs. hs799 cells (Fig. 3B). As expected, most enhancer-labeled cells mapped to postmitotic clusters. hs799+ cells were substantially enriched (>50% of cluster composition) in cl-0, -10, and -13, and to a lesser extent in cl-2, -4, -8, and -11 (Fig. 3F). In contrast, hs192+ cells were broadly distributed, with the greatest representation in postmitotic clusters cl-7, -8, -14, -15, and -16, and decreased representation in cl-0, -2, -10, and -13 (Fig. 3F). Cells labeled by at least one of the two enhancers were present in all postmitotic clusters alongside un gated enhancer negative cells, indicating that these two enhancers exhibit differential activity across scRNA-seq-defined cell type trajectories, and that together they have activity across major emerging MGE-derived neuronal lineages at E11.5. The proportion of positive cells for each enhancer varied by cell cluster (Fig. 3F). These results are consistent with fate mapping data showing unequal numbers of CIN, GABAergic projection, and cholinergic neuronal descendants from cells having activity of hs192 or hs799 (15).

We next performed DE and gene ontology (GO) analysis by cluster, expanding to include all 18,088 expressed genes to identify TF and non-TF markers (Fig. 3 G and H and Datasets S3 and S5). Proliferative clusters (cl-1, -5, -6, and -12) corresponding to cells within the MGE VZ, were enriched in terms including “DNA replication” and “chromosome segregation” (Fig. 3H). In comparison, proliferative clusters cl-9 and cl-3, likely corresponding to cells within the SVZ, were enriched for terms such as “exit from mitosis” (Fig. 3H).
enhancer activity is first evident for both hs799 and hs192 among cells in late SVZ clusters based on transgene expression (Fig. 3D and E). Postmitotic clusters subdivided into MGE-derived maturing GABAergic and cholinergic trajectories (cl-0, -2, -4, -7, -8, -10, -11, -13, and -16) differentially labeled by hs192 and hs799, as described in detail below.

Some clusters were made up of postmitotic cells that appeared to originate outside the MGE (cl-14, -15, and -17), and were likely migrating through MGE at E11.5 or captured at dissection boundaries. Cl-14 cells were enriched for hs192+ cells and had properties of LGE-derived immature medium spiny neurons, which express Ebf1, Nrg1, and Zfp503 (Nolz1), but not Nkx2-1 (Fig. 3G). Cl-15 contained both hs192+ and hs799+ cells and may derive from preoptic area (POA), based on Hmx2 (Nkx2-2) and Hmx3 (Nkx5-1) expression. Cl-17 cells were mostly transgene-negative, and may originate from regions adjacent to subpallium (e.g., hypothalamus or prethalamic eminence). These non-MGE clusters are not further discussed.

Projecting scRNA-seq Cell Identities onto Developing Mouse MGE via ISH Data. Using transcriptional identities defined by scRNA-seq, we used ABA ISH data (17) from the 689 TFs to map the anatomical distribution of cell populations in the E11.5 telencephalon. We graded expression of these TFs in VZ, SVZ, and MZ of MGE and LGE by reviewing individual sections from E11.5 and E13.5 (Dataset S2): 332 genes had visually detectable mRNA ISH patterns in the BG, of which 283 were also detected in E11.5 scRNA-seq (SI Appendix, Fig. S4C). We examined E11.5, E13.5, and E15.5 ISH time points, as ISH expression patterns that were undetectable or ambiguous at E11.5 could resolve spatial or anatomical lineages at E13.5 and E15.5. From this, we identified sentinel genes that

Fig. 3. Regional and proliferative gradients within E11.5 MGE. (A) Schematic of dissections used for MULTI-seq. (B) Visualization of single cells from E11.5 MGE by UMAP (green: M, mitotic; orange: PM, postmitotic). (C) UMAP colored by enhancer (red: hs799; blue: hs192; gray: enhancer-negative). (D) Proportion representation of hs192, hs799, and enhancer-negative by cluster. (E) Representative DE genes by cluster. Color represents mean z-score of normalized expression across clusters. (H) Representative gene ontology biological process terms separating clusters. Color represents log(adjusted/expected).
could be used as spatiotemporal markers to presumptively assign cells or clusters to distinct neuroanatomical regions and specific neuronal trajectories on the basis of mRNA expression. Using these sentinel markers and our scRNA-seq data, we characterized emergent specification trajectories and regional distributions of proliferative (Fig. 4) and postmitotic (Fig. 5) populations in MGE.

MGE Progenitors Stratify by VZ to SVZ and Dorsoventral and Rostrocaudal Axes. Analysis of mitotic cells (Mki67+) in the developing MGE provided evidence for distinct progenitor stages and regional patterning (Fig. 4 A–C). We identified four mitotic clusters driven by DE genes associated with progenitor maturation steps, suggesting discrete histogenetic stages: VZ1 (neuropoietin), VZ2 (radial glia), SVZ1 (secondary progenitor 1), and SVZ2 (secondary progenitor 2).

The most immature VZ stage, the neuropoietin (VZ1), was represented by cells with highest expression of early mitotic markers such as Hes1 and Id4 (22) and correspond to cl-5 and cl-6 (Fig. 4D and SI Appendix, Fig. S7A). More mature VZ2 cells expressed Wnt7b, Id3, and Tryp1, and were mostly within cl-12 and cl-1 (Fig. 4E and SI Appendix, Fig. S7 F and G). VZ2 cells had high expression of Hes5 and Fabp7, which mark radial glia (22, 23) (SI Appendix, Fig. S7 E and H). Other genes marked both VZ zones, including Lhx2, Rest, and Rgce (SI Appendix, Fig. S7 B–D). Genes like Axl1, Gix2, and Dlx2 began expression in VZ, (Fig. 4I and SI Appendix, Fig. S7 J and K). SVZ organized into progressively more mature SVZ1 in cl-9 and SVZ2 in cl-3. SVZ1 identity was assigned to cl-9 based on overlapping Olig2 and Dlx2 expression (24) (Fig. 4 H and I). Gsx1 had perhaps the most specific SVZ1 periventricular expression by ISH and was largely confined to cl-9 (Fig. 4F). Cells in this cluster also expressed high levels of known SVZ genes (Axl1, Gix2, and Hes6) (25–27) and neurogenic transition markers (Btg2) (28) (SI Appendix, Fig. S7 J–L). Gadd45γ, a marker of intermediate progenitors in the cortex (29), was highest in cl-9. SVZ2 identity was linked to cl-3. This most mature progenitor stage was associated with loss of Olig2 and high Dlx1, Dlx2, and Dlx5 (24, 30) (Fig. 4 H–K). cl-3/5SVZ2 markers (e.g., Proxl, Sp9, and St18) were expressed in a distinct layer of cells superficial to SVZ1 markers (Fig. 4G and SI Appendix, Fig. S7 M and N). Insm1 andIsl1 appeared to be expressed in both SVZ1 and SVZ2 (SI Appendix, Fig. S7 O and P), cl-8 and cl-11 may represent the earliest stage of neuronal commitment as SVZ2 cells exit the cell cycle; these clusters are discussed further below.

Markers indicating dorsal MGE identity (e.g., Nkx6-2, Gli1, and Gli2) were expressed in a subset of cl-12 and cl-6 cells (Fig. 4L and SI Appendix, Fig. S8 A and B). Dach2, a novel marker in this category, was expressed in the dMGE by ISH (SI Appendix, Fig. S8C). Genes indicating ventral (v)MGE identity (Shh, Slit2, and Tal2) (31) were expressed in a subset of cl-5 and cl-1 cells (Fig. 4M and SI Appendix, Fig. S8 D and E), with highest expression in medial sagittal planes of ISH sections (SI Appendix, Fig. S8 D and E). Bcan, Dach1, and Slit1 may also mark vMGE progenitor identity (SI Appendix, Fig. S8 F–H). The rostral MGE had high Otx2 expression (31), whereas the caudal MGE was marked by high Nkx2l and Nkx2p expression (32) (Fig. 4 N and O and SI Appendix, Fig. S8M). The POA and preoptic hypothalamus (POH) are contiguous with the caudal mitotic zone in the MGE. cl-12 may represent a mixture of POA2 and POH progenitors, based on expression of Nkx6-2, Dlx1, and Pax6 (18) (SI Appendix, Fig. S8 B, O, and P). POA1 cells lack these markers and have higher expression of Etv1 (18); these cells may be intermixed with MGE cells within cl-1. The septum is contiguous with the rostral MGE, and the septal markers Fgf15 (33), Zic1, and Zic4 (34, 35) were also expressed by rostral MGE progenitors (SI Appendix, Fig. S8 J and P). Pou3f1 and Ctnmap2 were novel markers of rostral cells (SI Appendix, Fig. S8 K and L), whereas Ptx3 was a new caudal MGE marker (SI Appendix, Fig. S8M). We leveraged canonical correlation analysis of MGE cells across both C1 and 10x datasets and ISH data, showing general patterns were concordant across the datasets (SI Appendix, Fig. S6). Patterning markers, including Id4, Otx2, and Tgf712, distinguished hs1538 (rostrodorsal-biased) and hs1056 (caudodorsal-biased) cells (Dataset S4), validating the regional identity evident among enhancer-labeled MGE progenitors. For all markers represented in Fig. 4 and SI Appendix, Figs. S7 and S8, we show quantitative differences across clusters in a summary violin plot (SI Appendix, Fig. S7Q).

Emergence of MGE Neuronal Lineages Revealed by Differential Enhancer Labeling. Postmitotic clusters were identified as the precursors of distinct MGE-derived cell lineages: GABAergic INs destined for the cortex (CINs) or other structures, GABAergic projection neurons, and cholinergic neurons (Fig. 5A). Early-born neurons emerge in SVZ2 and make up the MZ, which can be further visualized via a topological projection map that preserves spatial relationships across the telencephalon (15, 36) (Fig. 5 B and C). The organization and cellular outputs of the BG primordia identified herein using this topological map of E11.5 telencephalon correlate with progenitor zone anatomical map defined in Flames et al. (18) (SI Appendix, Table S1). Importantly, some of the cell types identified by scRNA-seq appeared to emerge in different spatial subdomains of the MGE based on overlapping marker gene expression in these topological maps.

Previous scRNA-seq studies have captured some of these emergent neuronal identities but largely have not resolved the spatial organization underlying this process (11, 12). Nks2-1, Lhx6, and Lhx8 marked the MGE-derived postmitotic neurons (Fig. 3G). To analyze these relationships, we focused on major emergent IN, GABAergic projection neuron, and cholinergic lineages and the differential activity of hs192 and hs799 that distinguishes these specific emerging neuronal populations. cl-0 and a subset of cl-13 were composed primarily of hs799+ with only a few hs192+ cells present, and expressed immature IN markers. A subset of cl-8 included hs799+ and hs192+ that expressed MafB. cl-0 showed the most consistent expression of immature CIN markers, including Calb1, Calc2, Erbb4, Lmo3, Maf, Npy, Sall6, St1, and Zeb2 (Figs. 3G and 5 D–F and SI Appendix, Fig. S9 A–G). Earlier lineage markers, such as Lhx8, were reduced in these cells. Subsets of cl-13 and cl-8 cells were MafB+, but lacked Maf and other CIN markers. The small proportion of MafB+ hs192+ cells in cl-8 and cl-0 was consistent with our C1 data and in line with fate mapping of a subset of hs192-lineage cells to CINs (15). Sentinel genes for cells from cl-0 and cl-13 were expressed in a distinct area in the periventricular MZ of the caudal MGE based on ISH at E11.5. We found that Calb1 expression is largely located in the caudal Diagonal (Dg) region, as illustrated on the topological map of the telencephalon (Fig. 5 F). The caudal Dg also showed expression of Calc2, Erbb4, Npy, St1, Maf, and MafB; the latter two genes are perhaps the most specific markers of MGE-derived CINs (37, 38) (Fig. 5 G and H and SI Appendix, Fig. S9 A–D). Contiguous with this region, in more lateral ISH sections, were cells expressing CIN markers, presumably migrating to the
Fig. 4. Separation of mitotic progenitors in E11.5 MGE. (A) Visualization of single cells from E11.5 MGE by UMAP, colored by expression of mitotic marker Mki67. Dotted lines show proposed boundaries for cells from VZ1, VZ2, SVZ1, and SVZ2. (B) Region of UMAP covering the VZ cells, colored by TF-defined cell cluster to show proposed regional identities (also shown in Fig. 3E). (C, Upper) Coronal schematic showing positions (blue dotted lines) of parasagittal sections used to illustrate gene-expression patterns. (Lower) Reference images of parasagittal sections from the ABA with relevant subpallial anatomical domains labeled. (D–O) Gene-expression UMAPs and representative ISH on E11.5 medial parasagittal sections showing expression domains for mitotic cell identity markers. (D–G) Gene markers for VZ-SVZ developmental stages (Id4, Wnt7b, Gsx1, St18). Colored labels correspond to cluster identity. (H–K) UMAPs and ISH for Olig2 and Dlx2/1/5, genes that help define SVZ1 and SVZ2 identity. (L–O) Gene markers with known regional expression patterns (Gli1, Tal2, Otx2, Nr2f1) in MGE VZ correlate with spatially defined UMAP locations, corresponding to the schematic in C. Lateral and medial sections shown. (M) Tal2 expression is highest at the medial level; coronal inset shows this corresponds to ventral MGE (arrow). Additional genes marking these developmental stages and regional identities shown in SI Appendix, Figs. S7 and S8. (Scale bar, 500 μm.)
LGE and CGE on their way to the cortex (e.g., Sst and Calb1 in Fig. 5 D and F). Additionally, Adams5, Bend4, Dgap1, Kitl, and Rai2 may be novel markers of immature CINs (SI Appendix, Table S1) for abbreviations. Blue dotted lines (rostral and caudal) indicate positions of coronal fluorescent images shown in F, I, and L. (C) Schematic depicting broad anatomical domains at E15.5. Dienc: diencephalon. (D–L) Gene-expression UMAPs and representative ISH on E11.5 parasagittal sections from ABA showing markers for distinct postmitotic neuronal lineages. Arrows indicate regions of higher expression. Scale bars: low magnification, 500 μm; high magnification insets, 100 μm.

INs, including hst799+ cells that contribute to SST+CINs, as well as other populations.

The enhancer-labeled transcriptional signatures and corresponding ISH patterns suggested at least two distinct classes of GABAergic projection neurons differentially labeled by hst799 and hst192 and distinguished by Gbx1+ or Zic1+ expression, respectively, that originate in different parts of the MGE (Fig. 5 H and L). The first GABAergic projection class was preferentially labeled by hst799 and expressed Gbx1, with more...
restricted expression of Foxp2 and Gbx2, and corresponded to Shh

+ cl-15 cells and to subsets of cl-4 and cl-2 (Fig. 5 G–I and SI Appendix, Fig. S8D). Gbx1, Kitl, Lmo3, Sox6, Th, Tle4, Tbx22, Zeb2, and Zic1 (5, 37) are additional markers for cells from cl-2, cl-4, and cl-13 (Fig. 3G, 5L, and 6D and SI Appendix, Fig. S9 E–H, J, and K). These cells may map to less mature states in cl-10 and cl-1 (for cl-13 only) or cl-4 and cl-11 (for cl-2). On the topological map of the MGE, Gbx2 expression was in the rostral part of the pallidal subdivision (Fig. 5I). Anatomically, the E15.5 expression domains of Foxp2, Gbx2, Gbx1, and Tle1, determined by ISH, were nestled along the pallidum’s radial axis. Foxp2 expression was superficial, possibly in the ventral pallidum (39) (Fig. 5G). Gbx1 expression encompassed the ventral pallidum and the entire globus pallidus (5H). Gbx2 expression included the ventral pallidum and part of the globus pallidus (20) (Fig. 5I). Tle1 was expressed throughout the globus pallidus but in few cells in the ventral pallidum (5) (Fig. 5J). Foxp2+ cells become superficial pallidal projection neurons (i.e., ventral pallidum), whereas other cl-2 cells, and some cl-4 and cl-13 cells, may contribute to deeper pallidal structures (i.e., globus pallidus). Thus, we propose these populations that share increased hs799 activity branch from immature neuron types within cl-2, cl-4, and cl-13.

The second GABAergic lineage expressed Zic1, as well as more restricted expression of Zic3 and Zic4, and was enriched for clusters with high hs192 activity (Fig. 5L). ISH MGE expression of Zic1 TFs was restricted to the rostroventral MGE and overlapped little with Gbx2; Zic1 may not overlap with the more caudal CIN markers (Fig. 5 D–F and SI Appendix, Fig. S9 A–D). On the topological map, Zic1 expression was restricted to the rostral Dg region, and was continuous with Zic1 expression in the septum (Fig. 5L). As noted, Zic1 was also expressed in the VZ and SVZ of the rostral MGE and septum, suggesting that this progenitor zone may generate the Zic1+ postmitotic cells in cl-11 and cl-4. It is unclear from ISH how Zic1-associated GABAergic projection neurons are spatially organized within the GP, though the spatial organization appears more diffuse across the GP. These hs192-biased Zic1+ cells may constitute a distinct cell type within the GP.

Cells in cl-16 were exclusively enhancer hs192+ (i.e., no hs799+ cells were identified) and expressed definitive cholinergic marker genes. Among these, Nrk1l (40) was the most specific to this cluster (Fig. 5K). Although Nrk1l expression was weak in ISH at E11.5, scattered positive cells were visible in the striatum and GP by E15.5, consistent with the distribution of cholinergic INs (Fig. 5K). Cl-16 cells also expressed Gbx1, Gbx2, Isl1, and Zic4, all of which are cholinergic lineage markers (4, 20, 41–43), as well as Zic1, Zic2, and Zic5 (Fig. 5 H, I, and L and SI Appendix, Fig. S9 L and M). Fig15 appeared to be a novel marker of this population, but its expression was not maintained as cells mature (SI Appendix, Fig. S9 N). A subset of hs192+ cl-4 cells also expressed some of these markers (e.g., Zic2) and may represent a less mature state or another type of cholinergic neuron. Fig9, Fig17, and Nkx2-1/Zic4 fate mapping provides evidence that these cells arise from the junction of the rostromedialventral MGE with the septum (41, 44), consistent with the expression of Zic1.

The final postmitotic group, cl-7, included two subpopulations: the upper part was almost exclusively hs192+, while the lower part also contained hs799+ cells. The upper part of cl-7 was Nkx2-1+ and strongly expressed Sox3 and Sph8 (SI Appendix, Fig. S9 O and P), suggesting a CGE-derived CIN or LGE-derived olfactory bulb IN identity (45). The lower part of cl-7 was Nkx2-1+ and Lbx8+, and thus likely MGE-derived; these cells also expressed genes shared by cl-2 and cl-14, including Id4 and Tle4 (Figs. 4D and 6D). Earlier states of cl-7 cells may map back to cl-8 then SVZ2 cl-3 cells, with Nkx2-2 expression in these clusters indicating caudal MGE origin (SI Appendix, Fig. S8M).

Enhancer-labeled scRNA-seq and ABA ISH indicated specific spatial localizations for cell types with divergent transcriptomic identities. To verify this finding, we performed fluorescent IHC and ISH to map protein expression across coronal sections for four sentinel genes representative of these populations (Fig. 5 F, I, and L). We examined Calb1 for early CNs, with predicted expression in caudal Dg; Gbx1 and Gbx2 for GABAergic projection neurons in rostral Pallidium; and Zic1 for rostral Dg GABAergic projection and cholinergic populations. These experiments validated that early born neuronal populations indeed exhibit spatial segregation to specific rostrocaudal and dorsoventral zones of the MGE MZ. For all postmitotic markers represented in Fig. 5 and SI Appendix, Fig. S9, we show quantitative differences across clusters in a summary violin plot (SI Appendix, Fig. S9Q). Our experiments captured specification trajectories and associations between spatial and transcriptomic identity of E11.5 MGE-derived neuron populations labeled by hs192 and hs799.

Validation of MGE Enhancer-Labeled Populations. Lastly, we sought to independently validate spatial and transcriptomic differences between proliferative and postmitotic enhancer-labeled cell populations. We used ISH to map the highest expression of the hs1538-associated marker Id4 in the VZ of the dorsal and rostral MGE, where hs1538 is active (SI Appendix, Fig. S10). In postmitotic populations, we tested for overlap between hs192 and hs799 GFP+ enhancer-labeled cells and Tle4 and Zic1, two TFs that were enriched in hs192+ cells (Fig. 6). hs192 activity and Zic1 expression were highest in the rostroventral MGE and paraseptal region, whereas hs799 GFP expression was in the dorsal MGE (Fig. 6 A and B; see section series in SI Appendix, Fig. S11). For Tle4, scRNA-seq and ABA ISH show that it is restricted to specific cell clusters in the MGE. We performed immunofluorescent colabeling of TLE4 with hs192 GFP and hs799 GFP; quantification of the proportion of cells labeled by both TLE4 and GFP from each enhancer (Fig. 6D) showed significantly more hs192 GFP+ cells were TLE4+ compared to hs799 GFP+ cells (Fig. 6 E and F). Thus, ISH and colabeling experiments validate the results from scRNA-seq–based separation of enhancer-labeled MGE cells.

Discussion

Transcriptional profiling at single-cell resolution has transformed our understanding of the diversity of cell types in the brain. Here, we combined enhancer-based cell labeling, TF-anchored scRNA-seq clustering, and ISH-based spatial annotation to map the neurogenic landscape of embryonic mouse BG. Through this integrated approach, we quantified enhancer activity in specific cells in vivo and generated new insights regarding the specification paths for early GABAergic neurogenesis in the ganglionic eminences.

Enhancer activity is often tightly restricted to specific cell types and developmental stages (46–48), making enhancers potent tools for genetic labeling and manipulation (13–15). However, the specificity of enhancer action in vivo at the single-cell level has not been deeply explored. Our results demonstrate that scRNA-seq can capture reporter transcripts across histogenetic subtypes labeled by individual enhancers. By profiling both enhancer-positive cells and unlabeled cells, our scRNA-seq analysis revealed the onset and offset of enhancer
activity, as well as cell populations where each enhancer was active with fine-scale resolution. Our results highlight the specificity and complexity of developmental enhancer activity during regional and cell type specification, as epitomized here by the activity for hs192 and hs799 across emerging MGE neuroanatomical lineages. This study offers a new perspective of single cell-resolved in vivo activity for seven evolutionarily conserved neurodevelopmental enhancers.

Two recent studies (11, 12) used scRNA-seq to follow BG development and CIN lineage specification in the E12.5 to E14.5 embryonic mouse. Our results capture an earlier timepoint when alternative neuronal lineages originate, and we identified transcriptomic differences among progenitor and postmitotic cells across the GEs. We show that using a TF-anchored clustering approach improved resolution of developmentally relevant clusters in our E11.5 data. Furthermore, integrating enhancer labeling and ISH data, we assigned likely identities to four types of MGE progenitors; distinct MGE progenitor regions; and enhancer-labeled subtypes of maturing MGE-derived INs, GABAergic projection neurons, and cholinergic neurons that together mature to form pallidal structures or migrate to become INs in the cortex, amygdala, and striatum.

Progenitor cohorts with overlapping transcriptional states and region-specific signatures were labeled by spatially distinct enhancers with VZ activity (hs1538, hs1056, hs841) Figs. 2 and 3. Similar to studies at older ages (11, 12), we found at E11.5 that scRNA-seq resolves the maturation gradient of BG progenitor populations (Fig. 4A). Comparison of hs1538+ and hs1056+ enabled the discovery of rostrocaudal (Id4, Otx2, Tcf7l2, and Zic1) and dorsoventral (Nr2f1/2 and Tal2) axes for MGE progenitors. Interestingly, Id4 is both associated with neuroepithelial progenitors (49, 50) and biased toward rostro-dorsal MGE VZ, indicating potential differences in progenitor state composition across regional domains within E11.5 MGE VZ. Positional information in E11.5 MGE progenitors appears largely encoded by gradients of these and other TFs rather than the expression of specific TF domains. Combinations of TFs then activate region-specific enhancers such as hs1538, similar to the classic model of ectoderm patterning in Drosophila embryos (51) and as has been reported for the cortical VZ (14).

Early-born CGE, LGE, and MGE GABAergic projection, cholinergic, and inhibitory neuronal developmental trajectories are labeled by differential enhancer activities (hs953, hs599, hs192, and hs799, respectively) (Figs. 2, 3, and 5). Compared
to the distinct MGE signature, LGE and CGE cells had more similar transcriptomic identities, but diverged between h6599 and h6593 cells in later postmitotic neuron clusters. Comparison of MGE enhancers h6192 and h6799 in combination with ISH annotation defined three spatially distinct regions in the MGE MGZ that give rise to molecularly distinct cells. CNIs (expressing multiple markers) were detected in the caudal Dg region (Fig. 5F). Gbx1/2+ cells were mapped to the Pallidial region and Zic5+ cells to the rostral Dg regions (Fig. 5 J and I). We are intrigued by the possibility that Gbx+ and Zic+ cells contribute to distinct types of MGE GABAergic neurons, including within the GP, in addition to the cholinergic neurons already described (20, 41). Previous studies have not identified a specific MGE region giving rise to CNIs, nor spatially distinct zones for generation of pallidal neurons and CNIs. This contrast to our findings may be because our analysis focused on a younger age (E11.5) than most studies. Consistent with our results, Puelles et al. (36) previously showed that Sst expression begins in the Dg region. It is probable that as development proceeds, additional MGE regions also generate CNIs (11, 12, 15). Overall, the results from our experiments define emerging GABAergic neuron types and elucidate spatial and transcriptomic relationships in embryonic BG during mammalian brain development.

Beyond BG neurogenesis, our study has broader implications for the application of scRNA-seq to developing tissues. We demonstrate that anchoring scRNA-seq analysis on TFs reduces weight of cell cycle and technical sources of variation, improving the power of histogenetic cell-type classification. This is consistent with the vast body of literature defining TF gradients as master regulators of lineage specification. Our results highlight the value and need for curated approaches in scRNA-seq analysis (e.g., our focus on TF transcripts) and the utility of orthogonal data, in this case ISH and enhancer labeling, to understand complex developmental processes. In summary, this study represents a unique approach, pairing scRNA-seq with enhancer-driven reporter expression in vivo, and offers insight into GABAergic neurogenesis.

Methods

Resource Availability.

Lead contacts. Further information and requests for resources and reagents should be directed to and will be fulfilled by the corresponding authors, A.S.N. (asnord@ucdavis.edu) and J.L.R.R. (john.rukenstein@ucsf.edu).

Materials availability. The enhancer transgenic mouse lines used in this study have been previously published (15) and deposited to the Mutant Mouse Research and Resource Centers (MMRRC) repository.

Data and code availability. The datasets generated are available at the National Center for Biotechnology Information Gene Expression Omnibus (NCBI GEO). The analysis codes can be found on the Nord Lab Git Repository (https://github.com/NordNeurogenomicsLab/).

Experimental Model and Subject Details.

Mice. Embryos of either sex from the seven enhancer transgenic mouse lines were used, and all embryos of the correct genotype from a single litter were pooled as a single biological replicate for all sequencing experiments. All animal care, procedures, and experiments were conducted in accordance with the NIH guidelines and approved by the University of California, San Francisco Animal Care Committee’s regulations (Protocol AN180174-02).

Method Details.

scRNA-seq library generation and analysis. E11.5 transgene-positive embryos were identified by screening on a fluorescent microscope. MGs, LGs, or CGEs were dissected, pooled, and single-cell suspensions prepared. For gated samples, GFP+ and tdTomato+ cells were isolated by FACS. For ungated samples, DAPI-stained cells were excluded but GFP+ and GFP− cells were collected. For full-length transcriptome scRNA-seq, cells were loaded onto Fluidigm integrated fluidsics chips on the Fluidigm C1 system and cell lysis, reverse transcription and cDNA amplification were performed according to manufacturer protocol. One microliter of diluted cDNA per well was used for library preparation using the Nextera XT DNA Sample Prep Kit. For multiplexed 3′ scRNA-seq library generation (MULTI-seq), dissociated cells for each single-litter pooled MGE sample were labeled with barcoded lipid-modified oligonucleotides, as previously published (21), and processed on the Chromium 10x system. Single-cell cDNA libraries were generated using the Chromium Single Cell 3′ GEM, Library & Gel Bead Kit (v3) according to manufacturer’s instructions. After cDNA clean-up, supernatant containing the barcode library fraction was processed as described previously (21) and sequenced.

C1 libraries were sequenced on an Illumina HiSeq 4000 instrument using a single-end 50-bp protocol. Reads were uniquely aligned to a custom mouse genome including sequences for CreERT2, IRES, EGFP, and tdTomato, duplicates were removed, and gene counts were generated. 10x MULTI-seq libraries were sequenced on an Illumina NovaSeq SP instrument and processed using CellRanger (v3.0.2) (52) using the custom mouse genome. Samples were demultiplexed and negative or multiplet cells were removed using the MULTIx2.eq (21) pipeline. Raw and aligned read quality was assessed for 5′ or 3′ bias, exonic read distribution and GC content distribution. We used the R package Seurat (v3.2.2) (16, 53) for feature selection, clustering, and visualization. To define proliferative state and cell cycle phase, we used the Seurat function CellCycleScoring on a published list of cell cycle genes (54) to assign mitotic (M) or postmitotic (PM) state. We removed all cells from clusters with high miRNA expression and outlier cell clusters.

We extracted counts for 689 transcription factors scored for E11.5 and E13.5 BG and cortex (Dataset S2) using ISH data available in the ABA (17). Genes with expression in >10 cells were used. Principal components analysis (PCA) were used to define clusters. For full transcriptome analyses, the top 3,000 highly variable genes were used. We performed diffusion mapping via the R package destiny (55) on the TF-based C1 dataset. For DE analysis, we expanded to all expressed genes and compared clusters with Wilcoxon rank-sum test. Genes with adjusted P < 0.05 were considered statistically significant. We assessed DE along DCs using a generalized additive model via the gam function in R. We used canonical correlation analysis to combine C1 and 10x data using both TF-curated and full transcriptome datasets. The integrated dataset was used for PCA and the first 15 PCs were used to define 17 clusters. We used the R package topGO (v2.36.0) (56) to perform GO analysis by cluster and enhancer group on DE genes in the 10x dataset. We used random forest (57) to identify TFs important for classifying cells in pairwise comparisons of enhancers in the C1 data, using out-of-box error rates for all forests, mean decrease in accuracy, and mean decrease in node impurity scores for all genes.

ISH transcription factor scoring. Expression levels of TFs in the ABA (17) were estimated in subdomains of the subpallium for two ages: E11.5 and E13.5. The subpallium was divided into LGE, MGE, septum, and POA. The LGE and MGE were further subdivided into laminar subdomains: VZ/SVZ and MZ for E11.5; VZ, SV21, and SV22, and MZ for E13.5. Expression was annotated from sagittal sections in each subdomain.

Histology. E11.5 heads were drop-fixed in 4% paraformaldehyde, cryoprotected in 20% sucrose overnight, embedded in OCT (TissueTek), and frozen on dry ice. The 15-μm cryostat sections were cut onto SuperFrost slides and dried, rinsed in PBS, blocked for 1 h at room temperature in PBST, and incubated with primary antibody overnight. Sections were washed then incubated with Alexa Fluor-conjugated secondary antibodies. Finally, sections were washed and coverslipped with Fluorescence Mounting Medium (DAKO #503203). For FISH, antisense RNA probes have been described previously (26). ISH was performed on 15-μm cryostat sections as described previously (58) up until antibody blocking, with the addition of peroxidase quenching for 20 min after SSC washes. Low-magnification epifluorescent images were taken using a 4× or 10× objective, and confocal images with 20× air and 40× oil objectives. Raw images were preprocessed to adjust brightness/contrast and convert to 8-bit RGB. Confocal images were stitched to create composites. Cell counting was performed on single confocal image planes. Counts were summed from at least three rostrocaudal sections for each brain. Tle4+ cells were counted first with the GFP channel hidden, excluding cells in the VZ (designated by nuclear staining), then scored as positive or negative for overlap with GFP.

Data Availability. The data have been deposited in Github (https://github.com/NordNeurogenomicsLab/) and the NCBI GEO database, https://www.ncbi.nlm.nih.gov/geo (accession no. GSE199552). Some study data available (Requests for resources and reagents should be directed to and will be fulfilled by corresponding authors, A.S.N. and J.L.R.R.). Previously published data were used for this work. [The enhancer transgenic mouse lines used in this study have been previously published (1) and deposited to the MMRRC repository.]
ACKNOWLEDGMENTS. University of California, San Francisco (UCSF) and Davis (UCD) cores supported this work. Sequencing was carried out at the UCSF Center for Advanced Technology and HG Genomics Core, UCD DNA Core; confocal microscopy was carried out at the UCSF Nikon Imaging Center (15100017993); cell sorting was carried out at the UCSF Helen Diller Family Comprehensive Cancer Center Laboratory for Cell Analysis (P30CA82103) and UCSF Parnassus Flow Core (RRID:SCR_018206, NI3 P03 DK63720, 15100012822). L.S.F. was supported by the UCD Schwall Fellowship in Medical Research and Wener and Jacobsen Fellowship, and National Institute of General Medical Sciences (NIGMS) NIH T32 GM08799. R.C.P.- was supported by a Science Without Borders Fellowship (Conselho Nacional de Pesquisas, Brazil). This work was supported by the following research grants: National Institute of Mental Health (NIMH) R01 MH081880 and NIMH R37/R01 MH049428 (to J.J.R.R.); NIMH/NIGMS R35 GM119831 (to A.S.N.); Department of Defense W81XWH-10-1-0123 and W81XWH-13-1-0221 (to Z.J.G.); NIH IO1CA199315 and DP2 HD080351-01 (to Z.J.G.); NSC MCB-1330864 (to Z.J.G.); and the UCSF Center for Cellular Construction DB1548297 (to Z.J.G.). Z.J.G. is a Chan-Zuckerberg BioHub Investigator. L.P. is supported by Seneca Foundation (5672 Fundación Séneca, Autonomy Community of Murcia) Excellency Author affiliations: "Department of Psychiatry and Behavioral Sciences, University of California, Davis, CA 95616; 6Department of Neurobiology, Physiology, and Behavior, University of California, Davis, CA 95616; "Nina Ireland Laboratory of Developmental Neurobiology, Department of Psychiatry, University of California, San Francisco Medical School, San Francisco, CA 94143; 8Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143; 9Department of Pediatrics, University of California, San Francisco, CA 94143; "Allen Institute for Brain Science, Seattle, WA 98109; "Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA 94158; "Chan Zuckerberg Biohub, University of California, San Francisco, CA 94143; "Center for Cellular Construction, University of California, San Francisco, CA 94143; and 12Department of Human Anatomy and Psychobiology, IMIB-Arrixaca Institute, University of Murcia, Spain. M.M. McGregor et al.

Author contributions: L.S.F., A.N.R., S.N.S., M.J.H., C.T., Z.J.G., H.J.L., J.L.R.R., and A.S.N. designed research; L.S.F., A.N.R., S.N.S., I.Z., and G.L.M. performed research; M.J.H., C.T., and H.Z. contributed to the project’s hybridization data generation and access; L.S.-F., A.N.R., R.C.P., K.J.L., A.R.Y., C.S.M., T.E.R., L.P., J.L.R.R., and A.S.N. analyzed data; L.S.F., A.N.R., J.L.R.R., and A.S.N. wrote the paper.

L. Becari, R. Marco-Ferreres, P. Bovolenta, The logic of gene regulatory networks in early vertebrate forebrain patterning. Mech. Dev. 130, 99–111 (2013).

A. S. Nord, Learning about mammalian gene regulation from functional enhancer assays in the developing mammalian CNS. Neuron 84, 555–574 (2018).

T. Inoue, M. Ota, M. Ogawa, K. Mikoshiba, J. Aruga, Zic1 and Zic3 regulate medial forebrain cholinergic neurons. J. Neurosci. 30, 6063–6076 (2010).

K. Yun, A. Mantani, S. Garel, J. Rubenstein, M. A. Israel, Id4 regulates neural progenitor proliferation and differentiation in vivo. Development 134, 2837–2851 (2007).

A. Telezhnikova et al., Gfi1 promotes neurogenesis and opposes Fgf8 function during neocortical development. Neural Dev. 3, 17 (2008).

M. E. Hatten, N. Heintz, Brain lipid-binding protein (BLBP): A novel signaling system in the developing cerebral cortex. Neuron 29, 277–289 (2000).

A. Telezhnikova et al., Gfi1 promotes neurogenesis and opposes Fgf8 function during neocortical development. Neural Dev. 3, 17 (2008).

M. M. McGregor et al., 59 9 130

Acknowledgments. University of California, San Francisco (UCSF) and Davis (UCD) cores supported this work. Sequencing was carried out at the UCSF Center for Advanced Technology and HG Genomics Core, UCD DNA Core; confocal microscopy was carried out at the UCSF Nikon Imaging Center (15100017993); cell sorting was carried out at the UCSF Helen Diller Family Comprehensive Cancer Center Laboratory for Cell Analysis (P30CA82103) and UCSF Parnassus Flow Core (RRID:SCR_018206, NI3 P03 DK63720, 15100012822). L.S.F. was supported by the UCD Schwall Fellowship in Medical Research and Wener and Jacobsen Fellowship, and National Institute of General Medical Sciences (NIGMS) NIH T32 GM08799. R.C.P.- was supported by a Science Without Borders Fellowship (Conselho Nacional de Pesquisas, Brazil). This work was supported by the following research grants: National Institute of Mental Health (NIMH) R01 MH081880 and NIMH R37/R01 MH049428 (to J.J.R.R.); NIMH/NIGMS R35 GM119831 (to A.S.N.); Department of Defense W81XWH-10-1-0123 and W81XWH-13-1-0221 (to Z.J.G.); NIH IO1CA199315 and DP2 HD080351-01 (to Z.J.G.); NSC MCB-1330864 (to Z.J.G.); and the UCSF Center for Cellular Construction DB1548297 (to Z.J.G.). Z.J.G. is a Chan-Zuckerberg BioHub Investigator. L.P. is supported by Seneca Foundation (5672 Fundación Séneca, Autonomy Community of Murcia) Excellency.

Author affiliations: L.S.F., A.N.R., S.N.S., M.J.H., C.T., Z.J.G., H.J.L., J.L.R.R., and A.S.N. designed research; L.S.F., A.N.R., S.N.S., I.Z., and G.L.M. performed research; M.J.H., C.T., and H.Z. contributed to the project’s hybridization data generation and access; L.S.-F., A.N.R., R.C.P., K.J.L., A.R.Y., C.S.M., T.E.R., L.P., J.L.R.R., and A.S.N. analyzed data; L.S.F., A.N.R., J.L.R.R., and A.S.N. wrote the paper.