RESOURCE/METHODOLOGY

Neural stem cell epigenomes and fate bias are temporally coordinated during mouse cortical development

Yonatan Shapira,[1,2,4] Florian Noack,[3,4] Silvia Vangelisti,[3] Faye Chong,[3] Aviezer Lifshitz,[1,2] Amos Tanay,[1,2] and Boyan Bonev[3]

[1]Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot 7610001, Israel; [2]Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 7610001, Israel; [3]Research Unit Brain Epigenomics, Helmholtz Center Munich, Munich 81377, Germany

During cortical development, neural stem cells (NSCs) combine self-renewal with the sequential production of different subtypes of projection neurons as well as glia cells. How the NSC epigenome accommodates this over time remains unresolved. Here, we address this gap by multimodal epigenomic profiling of mouse cortical development across six time points and five embryonic days. Single-cell gene expression and temporal modeling reveal that NSC self-renewal is not homeostatic, showing progressively stronger astrocytic preference over time. Chromosome accessibility, DNA methylation, and Hi-C show that this process involves major reorganization of the NSC epigenome. A model combining transcription factor motif affinities with epigenetic features, as well as integration of the results with a reporter assay in vivo, show that activation of the NSC neuronal fate regulatory program may be affected by a changing epigenome. Collectively, our findings uncover temporal epigenomic reprogramming that underlies the evolving differentiation potential of NSCs, providing insights into the intrinsic and extrinsic mechanisms that pattern cortical lineages.

[*Keywords*: brain development; computational modeling; epigenetics; single-cell omics; in vivo MPRA]

Supplemental material is available for this article.

Received June 11, 2025; revised version accepted December 22, 2025.

The cerebral cortex is the region of the brain responsible for cognitive function, sensory perception, and consciousness. It contains an unparalleled variety of neuron subtypes with unique molecular, morphological, and connectivity features, which are generated in a precise temporal sequence from neural progenitors (Molyneaux et al. 2007; Greig et al. 2013; Lodato and Arlotta 2015; Govindan and Jabaudon 2017; Koo et al. 2023). Fate-mapping studies have revealed that although the majority of neural stem cells (NSCs) are multipotent, a small proportion (10%–20%) may be precommitted to a certain fate early on in development (Gao et al. 2014; Llorca et al. 2019). Furthermore, refinement of subtype identity in postmitotic neurons has also been observed (Govindan and Jabaudon 2017), which can enhance or counteract lineage initiation in progenitor cells. Single-cell RNA-seq in mouse (Telley et al. 2016; Di Bella et al. 2021; La Manno et al. 2021) and human (Pollen et al. 2015; Li et al. 2020; Eze et al. 2021; Trevino et al. 2021; Ziffra et al. 2021; Braun et al. 2023) fetal brains have extensively profiled the remarkable diversity of cell types established during development, but the precise molecular mechanisms involved in lineage specification are still not fully resolved.

There is increasing evidence that chromatin remodeling and epigenetic regulation are essential for determining fate choices in the cortex (Pereira et al. 2010; Morimoto-Suzki et al. 2014; Telley et al. 2019) . Transplantation studies have revealed that the lineage potential of cortical progenitors to generate different types of neurons is primarily cell-autonomous (Shen et al. 2006; Oberst et al. 2019), but external stimuli such as bioelectric membrane

[4]**These authors contributed equally to this work.**
Corresponding authors: boyan.bonev@helmholtz-munich.de, amos.tanay@weizmann.ac.il
Article published online ahead of print. Article and publication date are online at http://www.genesdev.org/cgi/doi/10.1101/gad.353090.125. Freely available online through the *Genes & Development* Open Access option.

properties or Wnt signaling can overwrite the inherent developmental programs (Vitali et al. 2018; Oberst et al. 2019). Additional evidence was provided by the phenotypic defects in the timing and rate of neuronal differentiation in Polycomb mutants (Pereira et al. 2010; Morimoto-Suzki et al. 2014), disruption of the barrel cortex upon neuron-specific conditional *Ctcf* deletion (Hirayama et al. 2012), and impaired neuronal migration following mutations in the cohesin-loading factor Nipbl (associated with Cornelia De Lange syndrome) (van den Berg et al. 2017). Furthermore, the transcription factor (TF) Pax6 has been shown to interact directly with chromatin remodelers (Ninkovic et al. 2013), Neurog2 and Ascl1 have been proposed to act as pioneering TFs (Aydin et al. 2019), and Lhx2 and Ldb1 are essential for enhancer–promoter contacts in olfactory sensory neurons (Monahan et al. 2019).

Epigenetic regulation has also been implicated in the transition from a neurogenic to a gliogenic program in NSCs (Hirabayashi and Gotoh 2010; Amberg et al. 2019). For example, gliogenic promoters such as GFAP and S100β are selectively demethylated during this process (Takizawa et al. 2001; Fan et al. 2005), while proneurogenic genes become repressed via Polycomb (Hirabayashi et al. 2009; Morimoto-Suzki et al. 2014). Furthermore, global chromatin compaction, possibly mediated by changes in the expression of high-mobility group proteins, has been linked to the transition from neurogenic to gliogenic divisions during cortical development (Kishi et al. 2012; Kuwayama et al. 2023). However, whether such a switch represents a gradual transition or an abrupt change in cell identity remains unclear.

We and others have shown previously that cellular identity in neuronal differentiation is established by the complex interplay between transcriptional regulators, *cis*-regulatory elements, and the chromatin landscape, all within the physical constraints imposed by 3D nuclear architecture (Bonev 2017; Baizabal et al. 2018; Song et al. 2020; Noack et al. 2022, 2023; Zenk et al. 2024; Mannens et al. 2025). This coordinated remodeling is not only limited to development but also important for neuronal reprogramming (Pereira et al. 2024) and disease (Dileep et al. 2023; Xiong et al. 2023). However, it remains unclear whether and how epigenetic mechanisms contribute to the progressive restriction of lineage potential in neural stem cells, to their ability to generate different type of projection neurons, and to the switch to gliogenesis.

Here we characterized, in parallel, changes in the transcriptional and epigenetic states of neural stem cells (NSCs) and their differentiation fate biases across developmental time. We comprehensively profiled the transcriptional and epigenetic landscapes of mouse cortical development at single-cell resolution over a critical window in cortical development. Using differentiation flow models, we inferred NSC differentiation rates and specification preferences over time. We uncovered coordinated genome-wide interplay between chromosome accessibility, DNA methylation, 3D genome organization, and enhancer activity in purified populations of NSCs across the same developmental window. Collectively, our data show a temporally dynamic NSC transcriptome and epigenome and suggest potential mechanisms that may promote the plasticity and flexibility of NSCs. Inferred gene expression across metacell states can be viewed interactively at https://apps.tanaylab.com/MCV/mmcortex.

Results

A metacell flow model identifies lineage trajectories during mouse corticogenesis

To quantitatively and comprehensively understand self-renewal and diversification of NSCs during mouse cortical development, we profiled the single-cell transcriptome and the chromatin accessibility landscape in parallel, sampling two biological replicates daily across a critical window of mouse cortical development (embryonic days 13–18 [E13–E18]) (Fig. 1A). Focusing first on the scRNA-seq data, which comprised 42,172 cells across all time points, we inferred metacells and a manifold structure over them (Fig. 1B; Supplemental Fig. S1A,B; Baran et al. 2019; Ben-Kiki et al. 2022). We corrected ambient noise using MCnoise (Materials and Methods) and excluded cell types of noncortical origin, such as interneurons and microglia (Supplemental Fig. S1C; Materials and Methods).

Next, we inferred differentiation flows using the metacell flow algorithm (Fig. 1C; Materials and Methods; Mittnenzweig et al. 2021; Mayshar et al. 2023). Flows link the observed metacell distributions over time (Fig. 1D) such that metacells in time point t feed metacells with maximal transcriptional similarity in time point $t+1$, with flows constrained by estimated proliferation rates and observed distributions (Fig. 1E; Supplemental Fig. S1D,E). NSCs were observed to give rise to both NSCs and IPCs at time points E13–E15 and at later time points also fed astrocytes and OPCs. IPCs flowed at all time points into IPCs and the immature neuron states iCPNs/iCfuPNs and iCPN_early/late. In our data, proliferation rates were estimated to vary between 0 for postmitotic neurons and two doublings per 24 h for NSCs (Materials and Methods; Sahara et al. 2020).

We initially annotated metacells using known marker genes for NSCs, intermediate progenitor cells (IPCs), corticofugal projection neuron (CfuPN) and callosal projection neuron (CPN) subtypes, astrocytes, and oligodendrocyte precursors (OPCs) (Fig. 1F). The inference of differentiation flows is agnostic to cell type annotation, allowing us to use flows to define refined subsets of cycling IPCs, immature neurons (iCPNs/CfuPNs), corticofugal-biased immature neurons (iCfuPNs), and callosal-biased early (iCPN_early) and late (iCPN_late) immature neurons (Materials and Methods). Overall, the output of the manifold and flow model is 1093 annotated metacells, and inferred probability mass flows between them across the six time points.

When examining cell type fractions over time, we noticed that NSCs, IPC/IPC_cycs, and CfuPNs (containing subcerebral projection neurons [SCPNs] and corticothalamic projection neurons [CthPNs]) cell type fractions

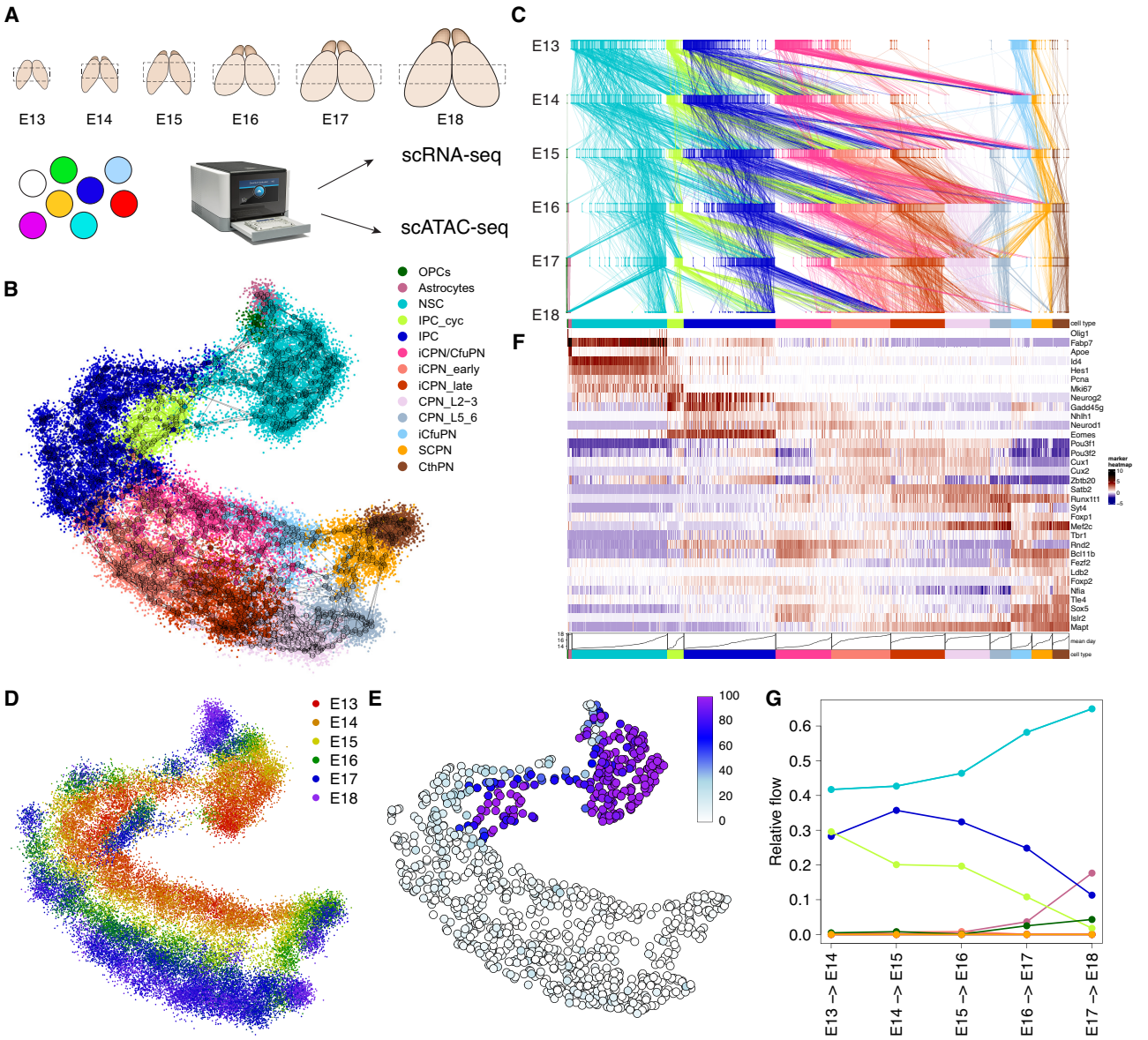

Figure 1. Comprehensive identification of lineage trajectories in mouse cortical development. (*A*) Schematic representation of the single-cell experimental approach. The mouse somatosensory cortex was sampled at 24 h intervals between E13 and E18 and used for scRNA-seq and scATAC-seq assays. (*B*) Two-dimensional UMAP projection of metacells (circles are colored by cell type), edges of metacell similarity graph, and single cells (dots are colored by cell type). (*C*) Visualization of metacell temporal flow model; vertical bars are metacell fractions in time points, and diagonal lines represent flows between metacells. (*D*) Single cells positioned by 2D projection of metacells and colored by time point. (*E*) Metacells are colored by cell cycle score, representing the percentage of cells classified as proliferating in each metacell. (*F*) Heat map of relative expression of marker genes across the metacell manifold. (*G*) Relative temporal outflow from NSCs per time point (e.g., E13 to E14, E14 to E15, etc.), aggregated and colored by cell type.

decreased (Supplemental Fig. S1E) while CPN proportion increased, consistent with the transition from mostly deep layer to exclusively upper layer neurogenesis. Additionally, astrocytes and OPCs emerged in significant numbers at E17, indicative of the NSC fate shift from producing neurons to astrogliogenesis. To further capture the trends in proliferation and differentiation, we extracted the total probability mass flows aggregated over cell types at each time point (Supplemental Fig. S1F). NSC maintenance rates—defined as the flow fraction going into NSC metacells in time $t+1$ of the total flow outgoing from NSC metacells at time t—varied from 42% (E13 → E14) to 62% (E17 → E18) (Fig. 1G). This is consistent with asymmetric neurogenic proliferation of NSCs, whereby in each division, one daughter cell remains an NSC, and another differentiates into an IPC. NSCs switched over time from a complete dominance of neurogenic IPC/IPC_cyc differentiation to later onset of astrocyte and oligodendrocyte differentiation at E17 (Fig. 1G; Greig et al. 2013).

NSC fate bias changes with time and is synchronized with the cell cycle

To uncover the temporal dynamics of fate bias in stem cells, we first analyzed groups of genes covarying between NSC metacells (Supplemental Fig. S2A,B). This unsupervised analysis identified gene clusters associated with the cell cycle as well as genes that exhibit increased or decreased expression over time (Supplemental Table S1) but did not uncover any other programs. To understand NSC fate predispositions and stemness signature, we also identified a neurogenesis/IPC-biased gene module and a gliogenesis/astrocyte-biased gene module (Fig. 2A, top/middle; Materials and Methods) to which we added genes enriched in NSCs relative to both astrocytes and IPCs but not correlated with cell cycle phase, defining a core "stemness" gene module (Fig. 2A, bottom; Materials and Methods).

Next, we profiled the activity of IPC, astrocyte, and stemness gene programs across cell types and over time (Fig. 2B). We observed that the expression of the IPC program remained consistently low in NSCs across time, while the expression of astrocyte-linked genes increased by a factor of ~8 between E13 and E18, reaching levels similar to those in astrocytes. In contrast, the stemness signature of NSCs decreased almost twofold over time but remained at higher levels than in other cell types.

To investigate the link between differentiation fate and the cell cycle, we phased single-cell profiles into G0, G1, S, G2, and M fractions based on gene expression (Fig. 2C; Supplemental Fig. S2C–E; Materials and Methods). For IPCs and neuronal types, >96% of the cells were assigned to the G0 (60.7%) or G1 (35.6%) phases (Supplemental Fig. S2C). For NSCs, we observed a slowing down of proliferation with time, with an increasing fraction of NSCs linked with the G0 stage (increasing from 9% to 32%) (Fig. 2D). These results are consistent with the lengthening of the NSC cell cycle observed during cortical development (Calegari et al. 2005; Telley et al. 2019).

To further dissect the relationship between cell cycle phase and cell fate, we analyzed cell cycle properties of single cells across the trajectory of NSC-to-IPC differentiation (Fig. 2E,F,H), defined by NSC- and IPC-specific signatures over 19 bins (Fig. 2G). We observed that when the population shifts from predominantly NSCs to mostly IPCs/IPC_cycs (bins 8–12) (Fig. 2G), the fraction of cells found in M phase also switches from lower to higher relative to S (from bin 13 to 17) (Fig. 2E,F), together with an overall convergence to a nonproliferating state (Fig. 2H). Such enrichment of mitotic states, starting from specific stages of IPC program activation, suggests coordination or synchronization between differentiation and the cell cycle. In contrast, grouping cells along the NSC-to-astrocyte trajectory (Fig. 2I–L) did not reveal a similar trend; there is no enrichment of M-phase over S-phase cells at any bin (Fig. 2I), and acquisition of astrocyte transcriptional signatures (Fig. 2K) co-occurs with reduction in the cycling fraction (Fig. 2L), consistent with a gradual convergence of the NSC program toward the astrocyte program (Supplemental Fig. S2F–I).

In summary, our model suggests that NSCs progressively lose their proliferative capacity while gradually activating an astrocyte program, where there is a possible synchronization of fate selection and the mitotic cycle.

Distinct temporal trends of neural chromatin remodeling from a slowly shifting NSC state

Changes in NSC differentiation fates over time may be driven by changing external signals, shifting epigenome landscape, or a combination of the two. To gain insight into the epigenomic signatures underlying these dynamics, we generated scATAC-seq libraries from 28,229 cells isolated from the same samples used for scRNA-seq (Fig. 1A; Supplemental Fig. S3A–C; Materials and Methods). We developed an algorithm to map scATAC profiles over the scRNA-seq metacells using matching gene variation (Supplemental Fig. S3D,E; Materials and Methods). This resulted in an inferred accessibility score over metacells for ~120,000 genomic hotspots (peaks). Among these, ~19,000 peaks were proximal (<1 kb) to an annotated transcription start site (TSS), while the remaining ~102,000 peaks were defined as TSS-distal, referred to here as CREs.

Next, we grouped TSSs and CREs into 10 and 60 clusters, respectively, based on accessibility patterns across metacells (Supplemental Table S2; Materials and Methods). As expected, TSS clusters showed little variability across metacells (Fig. 3A), consistent with previous results (Noack et al. 2022). In contrast, CRE clusters were broadly separated into a constitutively accessible group (~34,000 loci; class I) (Fig. 3B) and clusters with variable accessibility (classes II–VI). Within the variable groups, we identified two key groups associated with stem and progenitor cell activation. Specifically, class II CREs were accessible in NSCs but lost activity immediately, or with some delay, upon transition to IPCs, while class III CREs transiently gained accessibility in IPCs, followed by either rapid or slow decline in iCPNs. Additional CRE clusters corresponded to regulatory elements active in more differentiated neuronal states: Class IV was biased toward the corticofugal neural states (iCfuPNs/SCPNs/CthPNs), class V was biased toward CPN states, and class VI displayed a late pan-neuronal activation pattern. Interestingly, analysis of motif enrichment in these CRE clusters uncovered a rich combinatorial structure (Fig. 3B): CRE classes III–V that were enriched in intermediate progenitors (IPCs/IPC_cycs) and immature neurons (iCPNs/CfuPNs and iCPN_early/late) were characterized by a unique combination of motifs enriched in NSCs (NFIB and EMX1) and motifs specific to IPCs and neurons (NEUROD/G and MEF2C, respectively).

To dissect the accessibility landscape underlying the major cortical fate decision differentiating CPNs from CfuPN lineages, we decoupled subtype specification from the overall pan-neuronal maturation process. We identified 911 CREs with CfuPN-specific late (i.e., low in IPCs) activation and 515 CREs with specific activation in CPNs (Fig. 3C, top). We then traced the accessibility of these loci in their respective differentiation trajectories

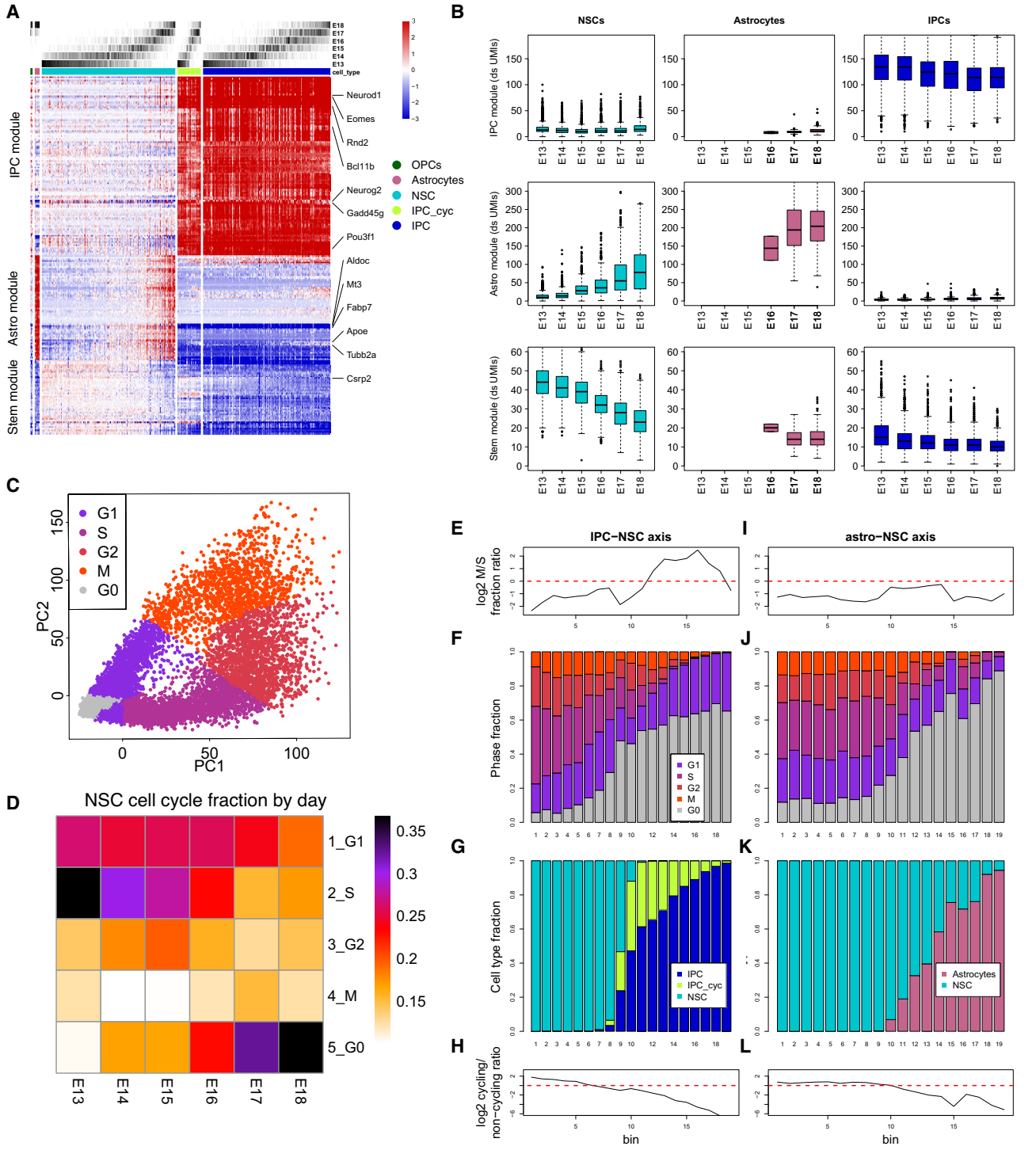

Figure 2. Temporal dynamics of NSC fate bias in the developing cortex. (*A*) Gene expression across metacells, relative to mean expression in NSCs and of IPCs, astrocytes, and stem gene modules. The *top* annotation, *above* "cell_type," shows the fraction of cells from each time point in each metacell. Key marker genes are also labeled. (*B*) Gene expression (total UMIs) of gene modules per cell type in downsampled single cells (3000 UMIs/cell) per time point. (*C*) The first two principal components of the cell cycle in all cells in the manifold, colored by inferred cell cycle stage. (*D*) Fractions of NSCs in different cell cycle stages per time point. (*E–L*) Ratio and scaled histogram of cell cycle phase and cell type fraction across the NSC transition into IPCs (*E–H*) or astrocytes (*I–L*). Single cells were divided into 19 bins in NSC → IPC and NSC → astrocyte trajectories. (*E,I*) Log_2 ratio of fraction of cells in M over S phase per bin. (*F,J*) Fractions in cell cycle stages per bin. (*G,K*) Cell types per bin. (*H,L*) Log_2 ratio of cycling versus noncycling cells per bin.

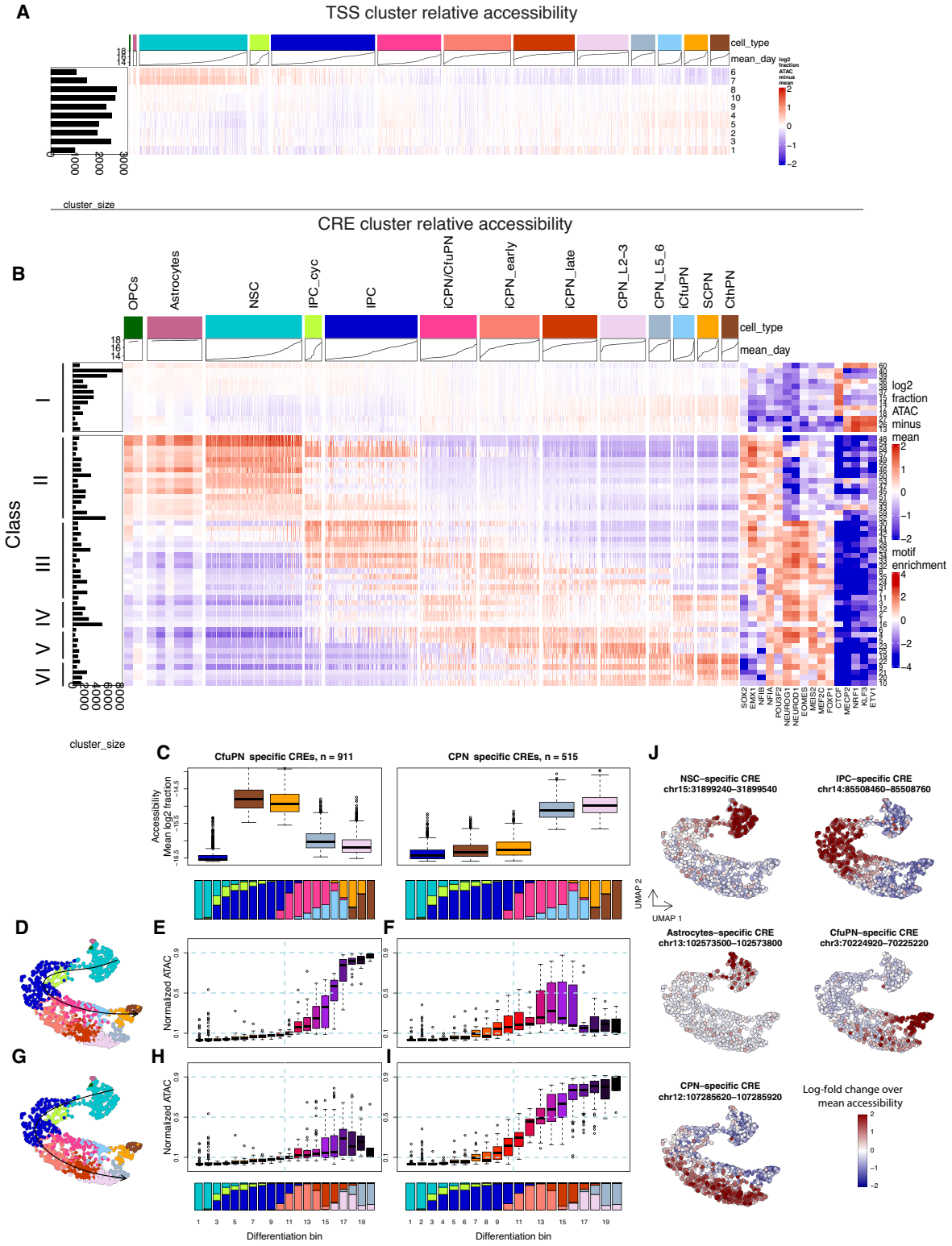

Figure 3. Widespread remodeling of the epigenetic landscape in projection neuron subtype trajectories. (*A*) Heat map of deviation from mean ($\log_2$ fraction) accessibility across metacells in TSS-proximal CRE clusters. Metacells are ordered by cell type (colors; *top*) and mean metacell day (line plots; *top*). Numbers at the *left* represent CRE cluster size. (*B*) Same as *A* but for TSS-distal CRE clusters. (*Left*) CRE cluster class. (*Right*) Heat map of motif enrichment ($\log_2$ of observed over expected high-affinity sequences) for selected TF PWMs. (*C*) Mean accessibility of CfuPN-specific (*left*) and CPN-specific (*right*) CREs across select cell types. (*D*) Approximate CfuPN trajectory across the manifold. (*E*) Mean normalized accessibility of CfuPN-specific CREs across differentiation bins in the CfuPN trajectory. (*F*) Mean normalized accessibility of CPN-specific CREs across differentiation bins in the CfuPN trajectory. (*G*) Approximate CPN trajectory across the manifold. (*H*) Mean normalized accessibility of CfuPN-specific CREs across differentiation bins in the CPN trajectory. (*I*) Mean normalized accessibility of CPN-specific CREs across differentiation bins in the CPN trajectory. (*J*) Examples of cell type-specific CREs visualized as fold change over their respective mean accessibility.

over the transcriptional and ATAC manifold (Fig. 3D–F [CfuPN trajectory], G–I [CPN trajectory]). We found that CfuPN-specific CREs were activated late in the CfuPN trajectory (Fig. 3E), with minimal representation in the immature states. In contrast, CPN-specific CREs were activated earlier in both trajectories (Fig. 3F,I) and subsequently suppressed in the later stages of the CfuPN trajectory (Fig. 3F). For illustration purposes, we provide examples of individual cell type-specific CREs and their accessibility patterns across metacells (Fig. 3J).

Overall, for at least several days after the final cell division, we observed progressive refinement of the two lineages associated with epigenome rewiring predominantly in distal regulatory elements. Furthermore, our data suggest that differentiation of deep-layer neurons is correlated with a repression of a CPN-associated epigenetic landscape in immature neurons, potentially mediated by the repressive TFs of the CfuPN lineage such as Fezf2 (Lodato et al. 2014; Tsyporin et al. 2021) and Sox5 (Kwan et al. 2008).

Accessibility gain in NSCs is accompanied by progressive loss of CpG methylation

We next extended our epigenomic profiling of NSC fate bias and performed methyl-Hi-C on sorted Pax6$^+$/Tbr2$^-$ NSCs isolated from the mouse embryonic somatosensory cortex every day between E13 and E17 as described previously (Fig. 4A; Materials and Methods; Noack et al. 2022). This comprehensive profiling allowed us to quantify differential methylation at 99,078 well-covered TSSs and CREs. TSS-proximal CREs were almost completely unmethylated (86% with $m < 0.05$) (Supplemental Fig. S4A), consistent with previous reports (Noack et al. 2022). In contrast, constitutively accessible distal CREs (e.g., class I) (Fig. 3A) exhibited slightly higher, but still low, methylation levels (46% with $m < 0.05$ and 81% at $m < 0.3$) (Supplemental Fig. S4A). Distal CREs with lineage-specific accessibility exhibited more heterogeneous methylation patterns (Fig. 4B), reflecting the dynamic regulation of these elements during differentiation.

NSC-specific CREs were predominantly unmethylated (82% with $m < 0.05$) or lowly methylated (99.5% with $m < 0.3$ across time points). Conversely, IPC CRE clusters included more sites with high methylation ($m < 0.3$, only 66%), and most neuron-specific CREs were highly methylated (17% with $m < 0.3$). In general, CREs were biased toward methylation loss between E13 and E17 in NSCs (Supplemental Fig. S4B), while methylation in TSSs remained largely unchanged. Interestingly, astrocyte-specific CREs (defined based on gaining accessibility exclusively upon astrocyte differentiation) showed remarkable loss of methylation in NSCs over time (from 43% with $m < 0.3$ at E13 to 95% at E17) (Fig. 4B). These results suggest that loss of DNA methylation in NSCs may be associated with priming, and initial high levels of DNA methylation at these CREs might prevent premature activation of the astrocytic program.

Accessibility and methylation are traditionally considered anticorrelated, a trend that we also observed in our data (Supplemental Fig. S4C). Consistent with this general pattern, CREs gaining accessibility over time in NSCs frequently became demethylated, indicating concomitant regulation (Fig. 4C,D). In contrast, loci that lost accessibility in NSCs over time remained lowly methylated even at E17, suggesting that accessibility and methylation can be temporally decoupled within the same lineage (Fig. 4E,F). Moreover, further analysis of cell type-averaged accessibility of temporally regulated CREs in NSCs revealed that regions that gained accessibility ("opening") retained high accessibility in astrocytes and OPCs, while CREs that lost accessibility ("closing") were NSC-specific (Fig. 4G). Consistent with these findings, the genome-wide chromatin accessibility landscape of late NSCs was more similar to astrocytes than to early NSCs (Supplemental Fig. S4E).

Interestingly, expression of genes encoding DNA methylation and demethylation enzymes was partially correlated during the developmental time window that we examined (Fig. 4H). Most notably, the methylation writer Dnmt3a was highly expressed in NSCs and progressively downregulated upon differentiation, while the demethylation factor Tet3 displayed the opposite trend. These complementary patterns suggest that shifts in the balance of methylation and demethylation activity may contribute to the dynamic remodeling of chromatin accessibility in NSCs. However, this mechanism alone cannot fully account for the observed changes, as CREs losing accessibility did not acquire de novo methylation (Fig. 4F) despite continued Dnmt3a expression, implying that other repressive processes may enforce loss of chromatin accessibility independent of DNA methylation.

Collectively, these data suggest that NSC fate bias is correlated with demethylation at astrocytic regulatory elements, which are progressively remodeled toward activation in late NSCs.

Hi-C uncovers 3D chromosomal remodeling in NSCs over time

To compare the epigenome dynamics at the linear genome level with changes in higher-order chromatin structure, we next examined 3D genome organization. We focused on chromatin insulation (defined as the proportion of contacts crossing a particular genomic region compared with adjacent loci), as we and others have previously linked changes in chromatin insulation to changes in gene expression in mouse neural differentiation (Beagan et al. 2017; Bonev 2017; Noack et al. 2022). Overall, insulation scores were highly conserved across developmental stages (Supplemental Fig. S5A). Nevertheless, we identified 335 genomic regions with insulation changes between E13 and E17 (Fig. 5A; Materials and Methods), including 100 regions that gained insulation ("insulating") and 235 regions that lost insulation over time ("deinsulating") (Fig. 5A). *Hapln1* is an example of a deinsulating locus positioned at the border of two TADs (Fig. 5B,C) that undergoes gradual compaction (Fig. 5C) and becomes expressed only in astrocytes (Fig. 5D). In contrast, the *Tnc* locus is an example of an "insulating" locus

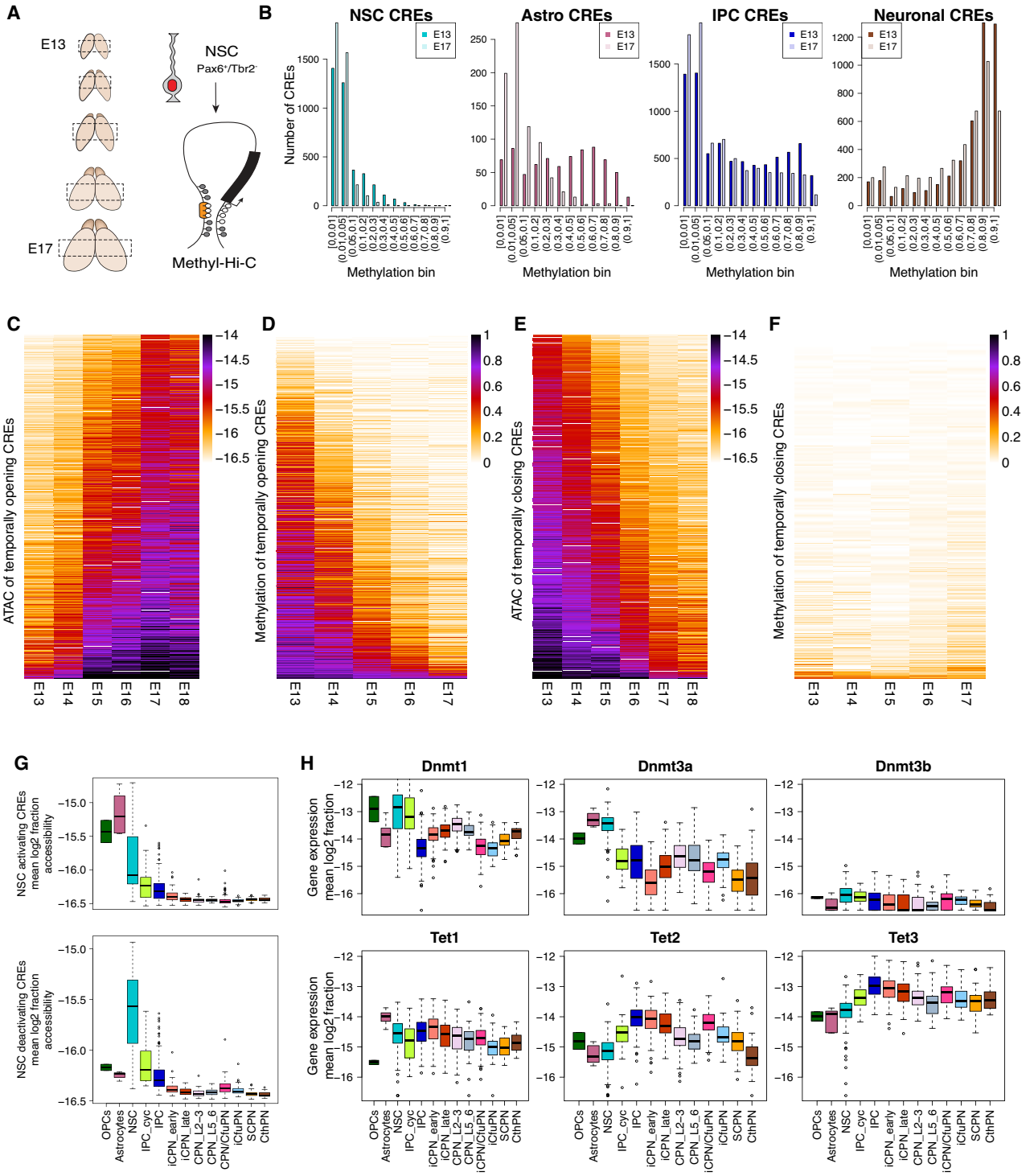

Figure 4. Distinct DNA methylation dynamics are associated with the changes in NSC temporal fate bias at CREs. (*A*) Schematic representation of the experimental approach to profile 3D genome organization and DNA methylation in NSCs. The somatosensory cortex was extracted from E13–E17 embryos. Dissociated NSCs were isolated by FACS and subjected to the methyl-Hi-C assay. (*B*) Histogram showing the number of cell type-specific CREs (as determined by the ATAC manifold) in methylation bins in the first (E13) and last (E17) time points assayed. (*C*) Heat map of chromatin accessibility of temporally activating CREs in NSCs. (*D*) Heat map of DNA methylation of temporally activating CREs in NSCs. (*E*) Heat map of chromatin accessibility of temporally deactivating CREs in NSCs. (*F*) Heat map of DNA methylation of temporally deactivating CREs in NSCs. (*G*) Box plots showing cell type-averaged accessibility of temporally activating (*top*) or deactivating (*bottom*) CREs in NSCs. Units are mean $\log_2$ fraction, averaged across metacells (and CREs) within cell types. For example, if Tet1's expression in astrocytes is −14, then the average fraction of UMIs in astrocyte metacells is one out of 2^{14}. (*H*) Box plots showing mean expression of DNA de/methylation enzyme genes in metacells by cell type.

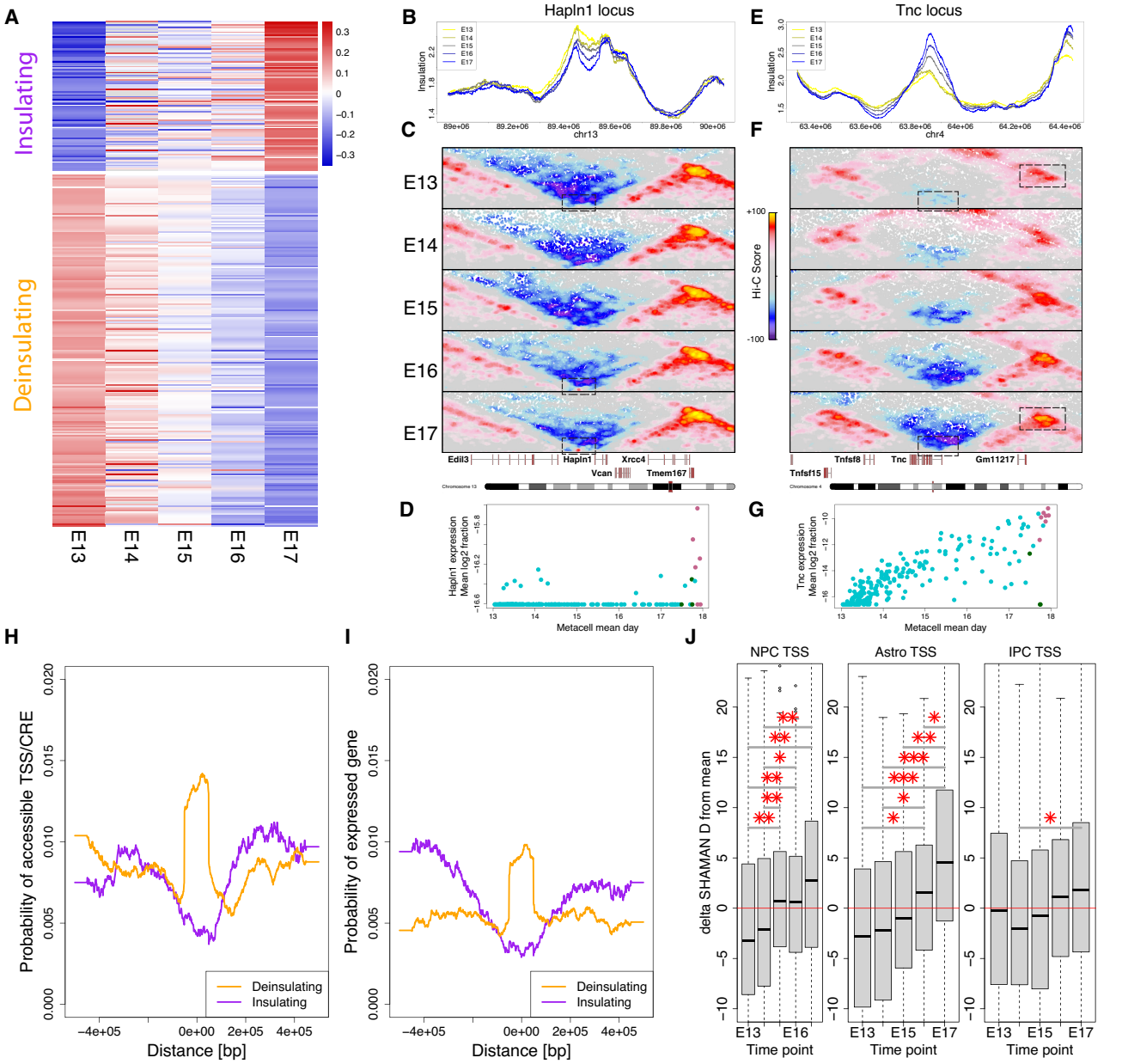

Figure 5. Three-dimensional epigenome rewiring suggests potential CRE cooperativity in NSC epigenome remodeling. (*A*) Heat map showing deviation from mean insulation score in regions with increasing (insulating) or decreasing (deinsulating) insulation scores in NSCs across time. (*B*) Insulation scores across time in an ~1 Mb window around the *Hapln1* gene. (*C*) Normalized interaction frequencies (SHAMAN scores) in the same window as *B* across time. (*D*) Hapln1 expression in NSC, astrocyte, and OPC metacells across time. (*E–G*) Same as *B–D* but for the *Tnc* gene. (*H,I*) Rolling average (window size = 60 kb) probability of neighboring accessible CREs (*H*) or expressed genes (*I*) in NSCs around insulating (purple) and deinsulating (orange) regions. (*J*) Deviation from mean SHAMAN scores (D) between CREs that are temporally activating in NSCs and TSSs of cell type-specific genes. Two-dimensional intervals for SHAMAN D score calculation were generated between the CREs and 10 kb diameter windows around the TSSs. In each 2D interval, SHAMAN D scores were calculated, and the mean was subtracted for each TSS group separately.

positioned at a subTAD boundary (Fig. 5E,F) where chromatin insulation gradually increases along with *Tnc* expression, also reaching peak levels in astrocytes (Fig. 5G).

To assess the relationship between insulation and chromatin accessibility, we quantified the distribution of CREs and genes around insulating and deinsulating sites. We found that both accessible CREs and TSSs of genes expressed in NSCs were significantly enriched around deinsulating compared with insulating regions ($P < 5 \times 10^{-4}$) (Fig. 5H,I; Supplemental Fig. S5B). However, despite these local enrichments, changes in insulation did not show a consistent genome-wide correlation with changes in accessibility or transcription at nearby CREs or genes (data not shown).

We next focused on the chromosomal contacts of CREs that gain accessibility over time in NSCs. We observed that these CREs engaged in significantly stronger interactions with astrocyte TSSs (as well as NSC TSSs) across multiple time points (Fig. 5J) while exhibiting weak or nonsignificant contacts with IPC TSSs.

Collectively, our data on both insulation dynamics and CRE contacts suggest that, in addition to gradual changes in DNA methylation, NSCs' chromosomal architecture evolves over time to form a potentially astrocytic-biased architecture. This is accompanied by decreased potential for generating IPCs and an increased probability for conversion toward astrocytes, potentially contributing to the fate switch of NSCs at the end of neurogenesis.

Modeling IPC chromatin accessibility using TF affinities and epigenomics

To dissect how *cis*-regulatory sequences and epigenetic context shape NSC differentiation, we developed a machine learning model to predict the relative CRE accessibility in IPCs compared with NSCs. First, we used ICEQREAM (Bercovich et al. 2025), a sequence-based quantitative modeling approach (Materials and Methods) to generate sequence models that best explain IPC versus NSC CRE specificity (Supplemental Fig. S6A). Model performance plateaued at ∼16 candidate motifs (Supplemental Fig. S6B), defining a compact yet informative set of sequence features that captured the transcriptional logic underlying lineage-specific chromatin accessibility. These motif-specific features were then integrated with further parameters defining the epigenetic baseline state, including mean NSC accessibility and methylation level at each CRE. In addition, we incorporated regional activity metrics by summing up accessibility (within a 50 kb window) and transcription (within a 0.5 Mb window) around each CRE.

Based on these features, our model achieved surprising accuracy, with an R^2 of 0.64 on held-out data (Fig. 6A), and an R^2 of 0.41 when applied to a generic, non-cortex-related set of CREs (Supplemental Fig. S6C; Materials and Methods). We used SHAP analysis to assign a contribution score for each feature in each model and quantified its importance for predictive accuracy. The most important model features (Fig. 6B) were the baseline NSC ATAC level, an E-box motif (potentially linked to bHLH TFs such as NEUROG/NEUROD), a T-box motif (potentially associated with EOMES), and mean NSC methylation level. These top features, however, provided only around half of the total SHAP values, with additional features such as Fos::Jun and SOX motifs (Noack et al. 2022) contributing to the prediction of the IPC–NSC accessibility difference.

Understanding the mechanism underlying a prediction necessitates careful stratification of the model's features, which in many cases are highly codependent. This is specifically important when the model considers a set of CREs that are preselected to be active in either NSCs or their progenies. For example, the E-box motif was highly correlated with NSC methylation levels in this CRE set (Supplemental Fig. S6E), but this effect was due to the enrichment of E-box elements in IPC-specific CREs that were methylated in NSCs prior to their activation. Only when correcting for E-box affinity by stratification was a negative correlation between DNA methylation and IPC accessibility (Fig. 6C) observed, consistent with the model suggested earlier (Fig. 4).

Finally, we asked whether the local epigenetic context affects the activity of lineage-specific CREs. Analysis of local CRE E-box sequence affinity conditioned on the presence of proximal CREs with high E-box and/or T-box sequence affinities within a 20 kb radius (Fig. 6D) corroborated our previous findings (Fig. 5). Specifically, CREs with medium–high or high E-box affinity were further activated in the vicinity of additional high-affinity CREs. This effect was found to be most pronounced when considering neighboring elements within 20 kb (Supplemental Fig. S6F). Moreover, the number of active neighboring elements—without filtering for their sequence content—was also predictive of higher accessibility in the IPC state (Fig. 6E).

Overall, these results point to a complex and multivariate regulation of lineage transitions in the cortex. While the initial epigenetic state and specific transcription factor motifs play predominant roles, our results suggest that the local epigenetic context and synergistic interactions among multiple CREs may further contribute to ensuring robust cell fate transitions.

In vivo reporter assays validate sequence-intrinsic models of CRE accessibility

To validate and further dissect autonomous versus context-dependent CRE regulation, we selected 11,905 sequences representing elements with NSC and/or IPC accessibility (Materials and Methods) referred to here as eCREs (episomal *cis*-regulatory elements) to distinguish them from CREs in endogenous context. We subjected them to an in vivo massively parallel reporter assay (MPRA) as described previously (Materials and Methods; Noack et al. 2022). Episomal libraries were electroporated into embryonic cortices between E12 and E16, and NSC and IPC cell populations were isolated via FACS 24 h after each electroporation (Fig. 7A; Materials and Methods). The resulting sequencing libraries were processed using MPRAflow (Gordon et al. 2020) and scored using MPRAnalyze (Ashuach et al. 2019), revealing that 9101 eCREs (76%) were found to contribute significantly to reporter expression in at least one time point.

After normalization and exclusion of two libraries (Materials and Methods), we observed distinct activity patterns in NSCs and IPCs across different sampling days (Fig. 7B). Cross-correlation of chromatin accessibility with reporter activity showed that sequences associated with differential accessibility in NSCs or IPCs within their native chromosomal context retained their cell type-specific activator autonomously and out of such context (Fig. 7C).

Next, we adapted our IPC–NSC accessibility prediction model to rely on sequence alone by fixing all epigenetic

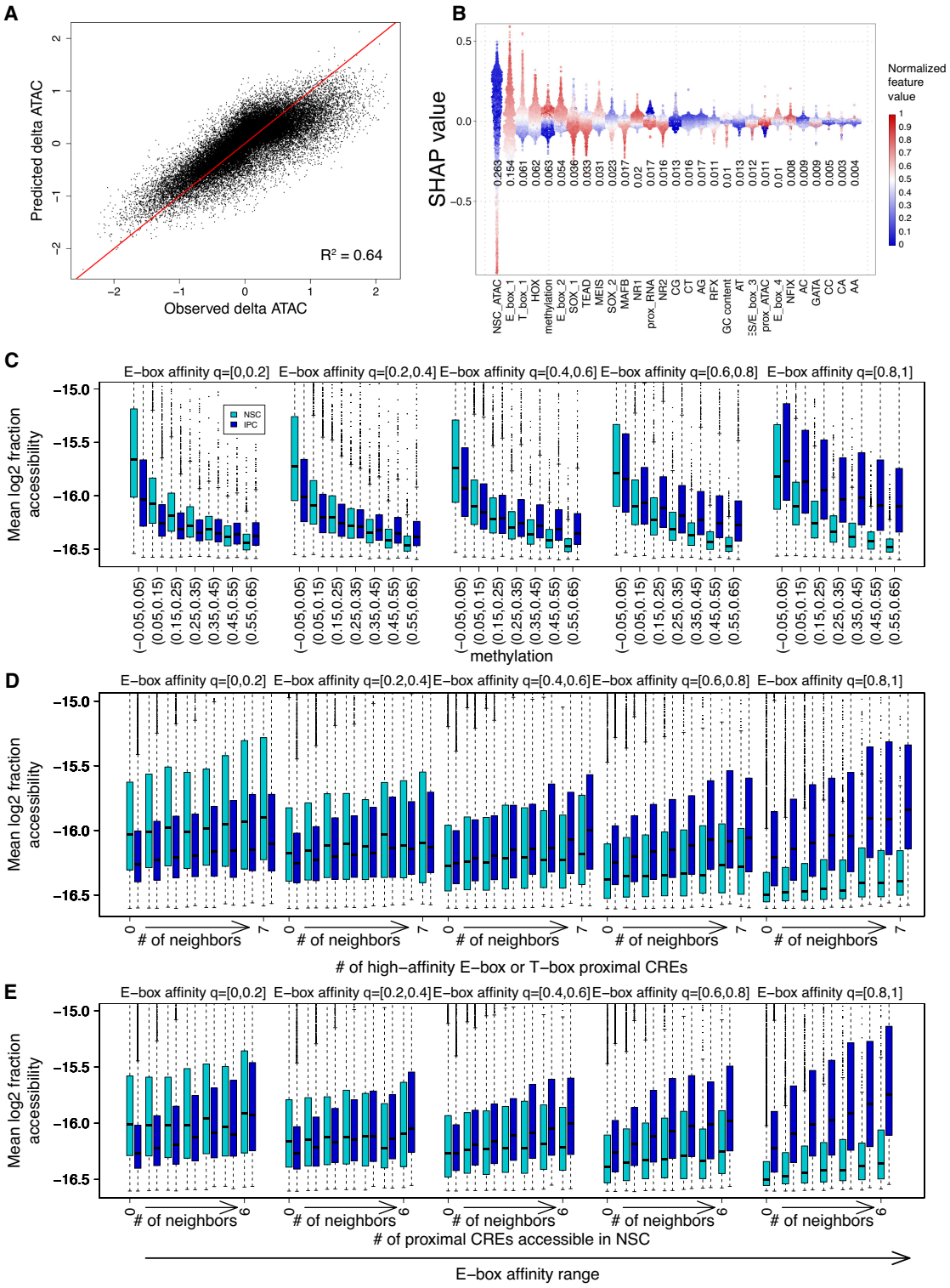

Figure 6. Machine learning model quantifying the contribution of TF affinities and epigenomes in NSC lineage transitions. (*A*) Scatter plot of mean IPC versus mean NSC chromatin accessibility as predicted by the XGBoost sequence and epigenomic model versus observed difference. (*B*) Distribution of SHAP values for main features of the XGBoost model. Text labels are the fraction of the sum of each feature's absolute SHAP value from the total absolute SHAP value across all features. Points are colored by the linearized feature value. (*C*) Mean NSC and IPC accessibility for all TSS-distal CREs, stratified by E-box energy and methylation. Each group of bars represents a different range of affinities for E-box (e.g., Neurog2) binding, with low affinities at the *left* and high affinities at the *right*. The energies within each group are stratified by ranges of methylation in NSCs, as specified in the axis labels. (*D*) Same as *C* but stratified by the number of neighboring CREs with high E-box/T-box affinity instead of by methylation. The energies within each group are stratified by number of neighbors with high E-box (e.g., NEUROG2) or T-box (e.g., EOMES) affinity. (*E*) Same as *C* but stratified by the number of neighboring CREs that are accessible in NSCs and/or IPCs instead of by methylation. The energies within each group are stratified by number of neighbors accessible in IPCs.

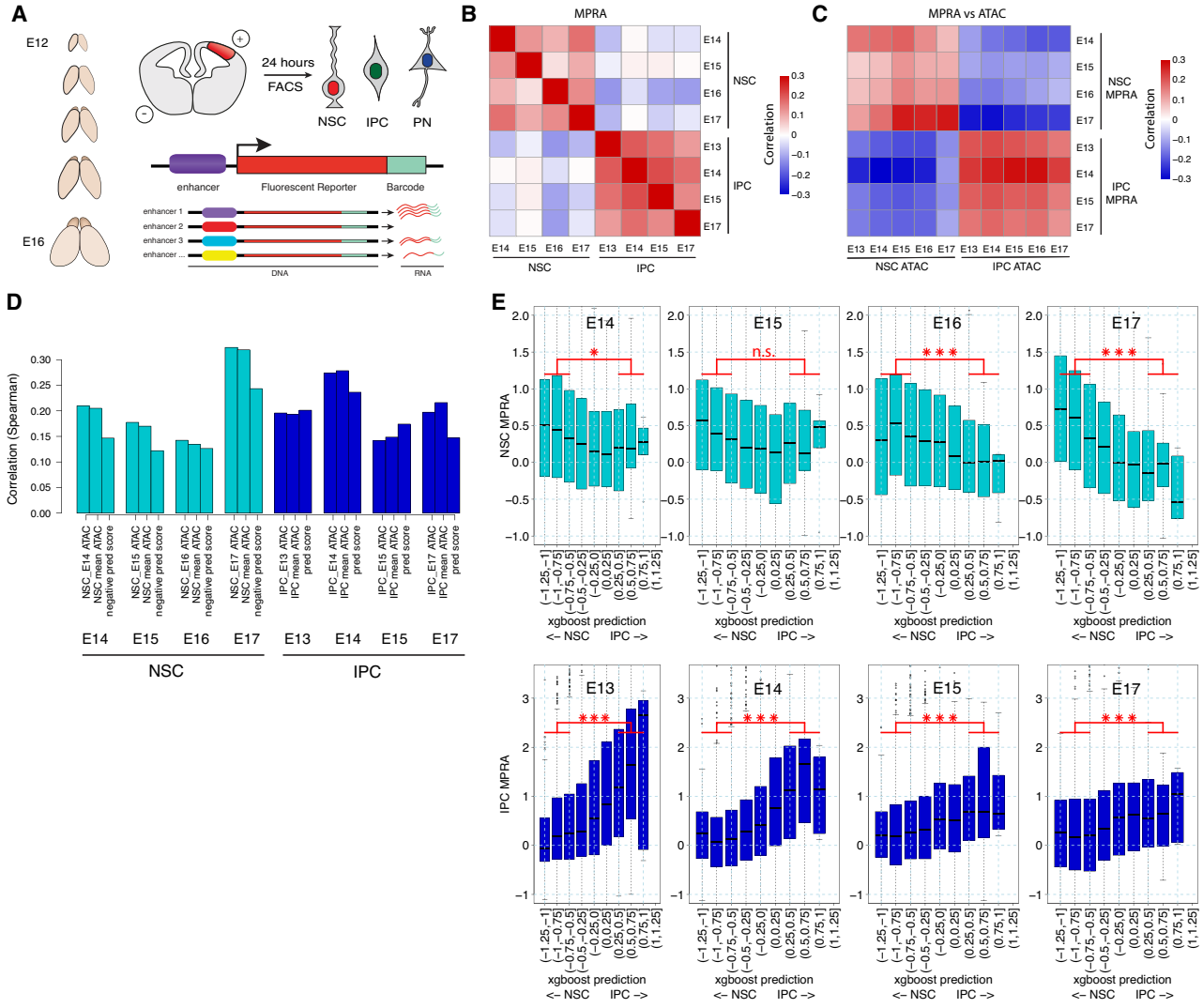

Figure 7. Massively parallel reporter assays in vivo decouple context-dependent CRE activation from autonomous CRE activation. (*A*) Schematic representation of the in vivo MPRA experiment. (*B*) Heat map of correlations between normalized MPRA scores across time points and cell types. (*C*) Heat map of correlations between normalized MPRA scores (rows) and chromatin accessibility (ATAC) of cognate CREs in the respective cell types (columns). (*D*) Bar plots depicting the correlation between MPRA scores and chromatin accessibility from same time point and cell type, mean ATAC of the respective cell type across all time points, and predicted score from the XGBoost sequence model. (*E*) MPRA scores versus binned predicted sequence score by cell type and time point. Bars and asterisks denote pooled bins that were compared and the significance of two-sided KS. (*) $P < 5 \times 10^{-2}$, (***) $P < 5 \times 10^{-3}$.

parameters to match an unmethylated and open context as provided by the reporter system (Materials and Methods). We then predicted an IPC–NSC activity score for each experimentally active sequence. This predicted activity score correlated with the MPRA signal at levels comparable with those using time point-specific and cell type-averaged accessibility (Fig. 7D). Moreover, stratification of profiled elements by their predicted NSC versus IPC sequence potential (as in Fig. 6) validated the model by showing that scores are positively correlated with IPC activity and negatively correlated with NSC activity at all time points (Fig. 7E).

Taken together, the in vivo reporter assays validate our CRE sequence model and further highlight the autonomous sequence potential of CREs. Notably, the temporal shift in NSC CRE regulation indicates that the fate bias change toward astrocytic identity is driven by changes in *trans*-factors' activity, occurring in parallel with gradual NSC epigenetic programming (i.e., DNA methylation and chromosomal conformation).

Discussion

In this study, we dissected the process of neuronal stem cell differentiation and fate determination from E13 to E18, an essential window of cortical development in mice. Our analysis combined single-cell gene expression

and chromatin accessibility data together with cell type-specific, genome-wide profiling of 3D genome organization, DNA methylation, and in vivo activity of 11,905 CREs, representing one of the most comprehensive data sets characterizing mammalian cortical development. Standardized sampling over time enabled us to infer the temporal differentiation dynamics from NSCs toward more differentiated progenitor states, distinct projection neuron subtypes, and glia. Single-cell analysis revealed that NSCs gradually lose their transcriptomic stem cell signature while converging toward an astrocytic program by E17. In addition, we demonstrated that the NSC mitotic cycle is synchronized specifically with IPC regulatory activation, an effect not observed during the NSC-to-astrocyte transition.

Across multiple scales and modalities, we demonstrated that corticogenesis involves extensive cell type-specific remodeling of CREs in concert with the increased activity of specific TFs that regulate stem cell differentiation. Beyond the canonical differentiation trajectories between cell types, our temporally resolved analysis revealed gradual epigenetic remodeling over absolute time within cell types. In NSCs, progressive opening of CREs was largely inversely correlated with DNA methylation levels. However, loss of accessibility in early NSC CREs was not accompanied by a corresponding increase in DNA methylation, though we cannot exclude a related change in DNA hydroxymethylation (5hmC). While 5hmC can represent a substantial fraction of cytosine modifications in mature neuronal subtypes and is strongly associated with active regulatory elements and gene expression (Mellén et al. 2017; Stoyanova et al. 2021), its levels are generally lower in proliferating progenitors and early postmitotic cortical neurons, which constitute the primary focus of our analyses. Thus, although cell type-specific and stage-specific accumulation of 5hmC at regulatory elements could contribute to an apparent decoupling between bulk DNA methylation and chromatin accessibility, the major effects observed here likely reflect alternative repressive mechanisms in dividing progenitors and early-born neurons (such as Polycomb) (Morimoto-Suzki et al. 2014; Telley et al. 2019). Future studies that directly resolve 5mC and 5hmC will be important to further dissect their respective contributions during later stages of neuronal maturation.

In addition to changes in the linear epigenetic landscape, we also observed significant dynamics in 3D genome organization. Chromosomal contact analysis revealed that NSCs gradually remodel specific TAD insulation anchors, a process that correlates with local CRE activity and transcriptional output. Furthermore, interactions between NSC CREs and astrocyte-specific TSSs become stronger prior to the switch to astrogliogenesis, suggesting that epigenetic priming within the regulatory landscape may facilitate this fate transition. Collectively, as we have shown across multiple modes of genomic regulation, the transcriptomic and epigenomic NSC state gradually drifts toward the astrocyte state over time.

Whether the epigenetic remodeling that we observed represents a defining characteristic of a restricted NSC differentiation potential remains an open question. Consistent with this notion, both Polycomb group and high-mobility group proteins have been shown to play an instructive role in NSC lineage progression and the switch to astrogliogenesis (Hirabayashi et al. 2009; Nagao et al. 2014). Our mechanistic modeling of CRE activity, enabled by the comprehensive and multimodal nature of our data and a novel machine learning approach, provides critical insights into this process. We identified key TF motifs as strong predictors of chromatin accessibility dynamics during the NSC-to-IPC transition, yet other epigenetic features such as DNA methylation and chromosomal context (total activation potential of the regions surrounding the CRE) contribute further to this process. Notably, these features align with the epigenetic dynamics that we directly profiled using methyl-Hi-C, highlighting previously unrecognized correlations in the temporal regulation of cortical development.

Finally, we further validated and assessed the autonomous capacity of CRE sequences to respond to TFs by performing cell type-specific MPRAs on >11,000 sequences across five developmental days in vivo. This data set shows that the *trans*-factor environment of NSCs and IPCs, possibly modulated by external signaling cues, is sufficient to drive appropriate episomal activity of cell type-specific CREs. Moreover, it highlights the distinction between CREs with autonomous regulatory potential and those that require synergistic activation within a genomic context, providing a new perspective on the mechanism of CRE activation in the context of NSC lineage specification.

Transcriptional atlases provide a crucial blueprint for understanding tissue developmental dynamics. Our data show that inference of such dynamics can rely on careful single-cell temporal sampling, metacell-based differentiation flow modeling, and in-depth single-cell and population-averaged epigenetic profiling. In the future, modeling of tissue development would be further enhanced by spatiotemporal profiling, which can augment the observed temporal behaviors with an essential context describing intercellular signals and interactions. Going beyond single-cell trajectories (e.g., pseudotime over transcriptional manifolds) and into models that describe cell dynamics and are aware of epigenetic history and tissue context will be essential for understanding the transition between stem cell multipotency and self-renewal and the formation of organized, balanced, and functional brain (or any other) tissue.

Materials and methods

Experimental model

Time-mated pregnant C57BL/6JRj mice were purchased from Janvier Laboratories and housed under standard conditions in compliance with local regulations set by the Regierung Oberbayern, Germany. Mouse embryos were used irrespective of sex. All experiments were conducted in accordance with national guidelines and approved by

local authorities (Regierung Oberbayern, Germany; ROB-55.2-2532.Vet_02-19-175).

scRNA-seq and scATAC-seq

Somatosensory cortices were isolated from E13, E14, E15, E16, E17, and E18 embryos after removal of the meninges and dissociated using a papain-based neural dissociation kit (Miltenyi Biotec 130-092-628) following the manufacturer's protocol. Library preparation for scRNA-seq (v3, 10x Genomics) and scATAC-seq (v1.1, 10x Genomics) was performed according to the manufacturer's instructions, targeting a recovery of 6000 nuclei per sample.

Fluorescence-activated cell sorting (FACS)

FACS was performed as described previously (Noack et al. 2022), following intracellular staining. A detailed protocol for the immunoFACS is available at https://www.protocols.io/view/immunofacs-b2a2qage. In brief, dissociated cells were fixed for 10 min at room temperature using 1% freshly prepared formaldehyde in PBS (Thermo Fisher 28908) with slow rotation, and the reaction was quenched by adding glycine (Invitrogen 15527013) to reach a final concentration of 0.2 M. Following fixation, the cells were centrifuged at 500*g* for 5 min at 4°C, washed once with a buffer composed of 1% BSA and 0.1% RNasin Plus RNase inhibitor (Promega N261A) in PBS, and then incubated for 10 min at 4°C in a permeabilization solution containing 0.1% freshly prepared saponin (Sigma-Aldrich SAE0073), 0.2% BSA (Thermo Fisher 15260-037), and 0.1% RNasin Plus RNase inhibitor in PBS. After permeabilization, the buffer was removed by centrifuging the cells at 2500*g* for 5 min at 4°C, followed by staining with antibodies against Pax6 (1:40; BD Bioscience 561664), Eomes (1:33; BD Bioscience 566749), and Tubb3 (1:14; BD Bioscience 560394) in a staining buffer (0.1% saponin, 1% BSA, 0.1% RNasin Plus RNase inhibitor in PBS) for 1 h at 4°C with slow rotation. The cells were subsequently washed twice with the permeabilization buffer, once with a wash buffer containing DAPI (1:1000; Thermo Fisher 62248), and finally with wash buffer without DAPI. Each washing step involved a 5 min incubation at 4°C with slow rotation, and cells were centrifuged at 2500*g* for 5 min at 4°C between washes. After the final wash, cells were resuspended in PBS with 1% BSA and 1% RNasin Plus RNase inhibitor, filtered through a 40 µm cell strainer (Thermo Fisher 15342931), and promptly subjected to FACS. Sorted cells were either directly used for nucleotide isolation (MPRA) or flash-frozen and stored at –80°C (methyl-Hi-C and in situ Hi-C).

Methyl-Hi-C

Somatosensory cortices were isolated from E13, E14, E15, E16, and E17 embryos after removal of the meninges and dissociated using a papain-based neural dissociation kit (Miltenyi Biotec 130-092-628) following the manufacturer's protocol. Each biological replicate represented a pool of four to six littermates from separate mothers. Methyl-Hi-C and low-input in situ Hi-C on the sorted NSCs were conducted as reported previously (Noack et al. 2022), and the detailed experimental procedures are available at https://www.protocols.io/view/methylhic-bif2kbqe and https://www.protocols.io/view/in-situ-hi-c-brd4m28w, respectively. Briefly, frozen pellets of fixed cells were thawed on ice and then lysed using 0.2% Igepal-CA630 (Sigma-Aldrich I3021). The cells were subsequently permeabilized with 0.5% SDS (Invitrogen AM9823) and digested overnight at 37°C using 200 U of DpnII (New England Biolabs R0543). Next, the sticky ends were filled in by incubating the nuclei for 4 h at room temperature with DNA polymerase I (New England Biolabs M0210) in the presence of a nucleotide mix containing biotin-14-dATP (Life Technologies 195245016) in DpnII buffer, followed by a proximity ligation step of at least 6 h at 16°C. Afterward, the nuclei were lysed, and the proximity-ligated DNA underwent reverse cross-linking overnight at 68°C, was purified by ethanol precipitation, and then sheared to ~550 bp fragments using a Covaris S220 sonicator. An end repair was performed after sonication by incubating the samples with T4 DNA polymerase (New England Biolabs M0203) for 4 h at 20°C. Before bisulfite conversion, ~0.01% of sheared and biotinylated fully methylated pUC19 (Zymo Research D5017) and unmethylated λ DNA (Promega D1521) was added to the samples. Bisulfite conversion was then carried out using the EZ DNA Methylation Gold kit (Zymo Research D5005) followed by library construction with the Accel-NGS Methyl-Seq DNA library kit (Swift Bioscience 30024) according to the manufacturer's instructions up to the adapter ligation step. Afterward, biotin pull-down was performed using MyOne streptavidin T1 beads (Thermo Fisher 65602) with five washes using a buffer containing 0.05% Tween-20 (Sigma-Aldrich P9416) and two additional washes with low-TE water. Finally, libraries were amplified on the streptavidin beads using EpiMark Hot Start Taq (New England Biolabs M0490) along with Methyl-Seq indexing primers (Swift Bioscience 36024) using the following program: 30 sec at 95°C; 14 cycles of 15 sec at 95°C, 30 sec at 61°C, and 60 sec at 68°C; followed by 5 min at 68°C; and a final hold at 10°C. The streptavidin T1 beads were then magnetically pelleted, and the libraries present in the supernatant were purified with 0.6× AMPure XP beads (Agencourt A63881) to achieve an average fragment size of ~500 bp.

Massive parallel reporter assay (MPRA)

The MPRA design followed the approach previously described by Noack et al. (2022), with the following specifications. The MPRA plasmid pool included 997 scrambled control sequences, which had matched GC content and were prescreened to minimize the presence of expressed TF motifs, and 11,905 putative episomal *cis*-regulatory elements (eCREs) derived from CREs with temporally increasing, decreasing, or stable accessibility patterns in NSCs, IPCs, or neurons (~1000 per cell type and pattern

combination). All sequences were centered on the accessibility peak and resized to 266 bp.

A detailed protocol for MPRA plasmid pool generation is available at https://www.protocols.io/view/mpra-plasmid-pool-preparation-bxchpit6.

We electroporated the MPRA library into the cortices of mouse embryos at E12, E13, E14, E15, and E16 and collected the embryos 24 h after. We then removed the meninges, dissected the somatosensory cortex, and dissociated the cells using a papain-based neural dissociation kit (Miltenyi Biotec 130-092-628) according to the manufacturer's protocol. Dissociated cells were stained for Pax6, Tbr2, and Tubb3, and specific populations were sorted using FACS as described above.

In utero electroporation

In utero electroporation was carried out according to the previously reported protocol (Noack et al. 2022). In summary, pregnant mice with time-mated plugs were anesthetized with isoflurane, the uterus was exposed, and 1–3 µL of plasmid DNA in PBS was injected into the telencephalic lumen. This injection was immediately followed by five pulses at 35 V, each lasting 100 msec with 1 sec intervals, delivered via platinum electrodes (NepaGene CUY650P1). At either 24 or 48 h after electroporation, the pregnant females were sacrificed, the embryos' brains were dissected, and the electroporation efficiency was evaluated using a fluorescent stereomicroscope.

scRNA-seq mapping and filtering

scRNA-seq data were processed using a standard CellRanger pipeline (v3.1.0) to generate count matrices per library. Raw count matrices were then filtered by removing cells if they had >15% of UMIs mapping to the mitochondrial genome, 0.5% of reads mapping to hemoglobin genes, <1500 UMIs, or >30,000 UMIs.

Doublets were removed using DoubletFinder (McGinnis et al. 2019) with 30 principal components, log normalization of gene expression (scale 10,000), 2000 genes, and clustering resolution of 0.6. A linear model was fit to 10x Genomics data on expected doublet fractions in proportion to cells loaded in a chip, and 0.75 times the expected fraction per sample was used as a parameter for the calculation of the Poisson exponent in the DoubletFinder pipeline. The pK/BCMVN parameter was taken as the maximum in the values searched per sample. pN = 0.25 for all samples. Out of 60,217 cells, 2046 (3.4%) were classified as doublets.

Metacell analysis and filtering

Metacell (Baran et al. 2019; Ben-Kiki et al. 2022) analysis was performed as described with the following parameter tuning: Feature genes were selected using the functions mcell_gset_filter_varmean and mcell_gset_filter_cov with default parameters. The feature gene set was split into 96 clusters, and clusters were removed if they contained the "undesired variance" genes ("Isg15," "Wars," "Ifit1," "Mki67," "Pcna," "Hist1h," "Smc4," "Mcm3," "Top2a," "Fos," "Hsp90ab1," "Hspa1a," "Hif1a," "Xist," and "Tsix") or had >10 genes in the top 100 feature genes correlated with at least two of the UV genes. UV genes and all ribosomal genes were removed from the feature gene set.

Metacells were then constructed using default parameters. Metacells were selected for removal based on a gene footprint score >1.5 for any of the nonneural/noncortical markers ("Reln," "Lhx5," "Gad2," "Sst," "Lhx6," "Nrxn3," "Gsx2," "Dlx2," "Aif1," "C1qb," "Hexb," and "Igfbp7"). All single cells belonging to such metacells were removed from subsequent analyses. Following filtering of noncortical/nonneural cells, metacells were again constructed as above.

Metacell UMAP and noise cleanup

We removed cell cycle-correlated genes (see also "NSC Gene Module Analysis" below; Supplemental Fig. S2) and constructed a metacell–metacell similarity graph with the mcell_mgraph_logistic function with mcell_mgraph_max_confu_deg = 7 and otherwise default parameters. UMAP was constructed using 10 neighbors, spread = 1, and min_dist = 0.5. We used MCNoise (https://www.github.com/tanaylab/mcnoise) to remove background gene expression on the RNA metacell manifold. The parameters used were num_of_genes_cluster = 13, number_of_mc_clusters = 8, genes_min_diff = 4, thr_max_value = –2, min_number_of_gmctypes_per_batch = 10, and min_number_of_batches_per_gmctypes = 5; three gene clusters (2, 4, and 13) were excluded from the analysis. The ambient noise levels of all the batches were stable for threshold = –3, and those estimations were used for downstream analysis.

Inference of metacell temporal flows

Flows were generated using the metacell.flow package (Mittnenzweig et al. 2021; Mayshar et al. 2023). This involved inference of optimal flow values given the estimated densities of metacells per time point, estimated proliferation rate per metacell, and evaluation of differentiation costs (or distances) over the manifold graph.

1. Metacell frequencies per day were estimated directly from the data by tabulating the time label of all cells and summing across metacells each day.
2. Cell cycle scores (fraction of nonproliferating cells) for metacells were calculated using thresholds of 0.0025 M-gene fraction and 0.001 S-gene fraction (Mittnenzweig et al. 2021). An estimate for the proliferation rate (or the fold change in number of cells per metacell per time step) was calculated while taking into account two factors: the cell cycle score (or at what fraction of the maximal rate each metacell is cycling) and an "IPC score" proportional to the concentration of the *Eomes* mRNA, because it is known that IPCs divide at a slower rate relative to NSCs (Calegari et al. 2005;

Sahara et al. 2020). The formula for the proliferation rate was

$$PR_{mc_i} = 2^{\frac{\text{max.div}}{\text{day}} \times CC.\text{score}_i \times C_{\text{IPC}i}}$$

where the per-metacell cell cycle score $CC.\text{score}_i = \frac{|\{m_i > m_0 \cup s_i > s_0\}|}{n_i}$ was the proportion of cells in each metacell surpassing the thresholds for M- and S-phase UMI fractions, the IPC score was $C_{\text{IPC}i} = 1 - \frac{1}{2}\text{plogis}(v_{\text{Eomes}i}|\mu = 0.1, \sigma = 0.01)$, where $v_{\text{Eomes}i}$ is the *Eomes* expression (*metacell* e_gc vector) rescaled between 0 and 1, and $\frac{\text{max.div}}{\text{day}} = 2$, corresponding to a 12 h cell cycle duration for NSCs.

3. The construction of metacell differentiation costs was done using the metacell graph described above (used also for UMAP construction).

Given the estimation of 1–3, we built a network model using the function mcell_mctnet_from_mgraph with parameters off_capacity_cost2 = 10^4, max_flow_tolerance = 0.04, flow_tolerance = 0.04, and capacity_var_factor = 0.2 (and otherwise default). We then solved the flow problems using the function mcell_new_mctnetflow.

Metacell cell type annotation

Metacells were annotated for cell types based on markers known from the literature. Annotation for intermediate states (namely, various classes of immature neurons) utilized marker genes and observed flows from the temporal flow model. We note that metacell construction and temporal flow model inference are agnostic to annotation.

NSC gene module analysis

We selected all genes that have maximum expression ($\log_2$ of RNA frequency) of at least –14, and a $\log_2$ fold change of at least 2 between the fifth and 95th quantiles in NSC metacells. The Spearman correlation matrix of these genes' expression across NSCs was hierarchically clustered, and the dendrogram was cut with $k = 8$.

The IPC module was composed of genes with a fold change of at least 2 between mean expression in IPCs and mean expression in NSCs, astrocytes, and oligodendrocytes.

The astrocyte gene module was composed of genes with a fold change of at least 2 between mean expression in astrocytes and mean expression in NSCs, IPCs, and oligodendrocytes.

The stem gene module was composed of genes with a fold change of at least 1 between mean expression in NSCs and mean expression in astrocytes and IPCs and had a correlation <0.5 between their expression across NSC and IPC_cyc metacells and the expression of the cell cycle genes Top2a, Mki67, Mcm4, and Pcna.

The NSC gene module was composed of genes with a fold change of at least 1 between mean expression in NSCs and mean expression in IPC_cycs and astrocytes.

Single-cell cell cycle analysis

The scRNA matrix was downsampled to 3,000 UMIs/cell. Single-cell module scores for NSC gene modules 1, 3, 6, and 7 and IPC, astrocyte, and stem gene modules were the sum of UMIs from each module. We centered the matrix of single-cell cell cycle gene modules and calculated principal components for it. The first two principal components were retained as the cell cycle coordinates, and we rotated them by 140° for convenience. For downstream analyses, we assigned as "cycling" those single cells that had total UMIs in cell cycle modules exceeding the 80th percentile. The radial cell cycle phase was calculated for cycling cells as $\phi = -\arctan\frac{PC2 - \langle PC2\rangle}{PC1 - \langle PC1\rangle}$, where angled brackets denote the mean. We used this radial phase to calculate the smoothed phase using the principal_curve function in R with periodic_lowess smoothing on the rotated PC1/2, with cells ordered by the radial phase. Principal curve coordinates were smoothed in a moving average window of 50 points. The phase assigned to each cell was the nearest principal curve point (Euclidean distance). The vector of cells' phases was binned into 12 equally spaced bins. Bins 1, 4, and 12 corresponded to G1 stage, bins 5–7 corresponded to S stage, bins 8 and 9 corresponded to G2 stage, bins 10 and 11 corresponded to M stage, and bins 2 and 3 corresponded to G0 stage. While the G0 label was assigned due to a lack of cell cycle signature, we do not claim that every cell with this label was indefinitely postmitotic, particularly those in progenitor cell types.

Single-cell ordering across differentiation axes

To compute the IPC differentiation axis, a principal curve was calculated for the NSC and IPC gene module scores for all single NSCs, IPCs, and IPC_cycs. Cells with >30 UMIs in the astrocyte module were removed from downstream analysis. The developmental axis coordinate for each cell was the principal curve point that was closest to it (Euclidean distance). Cells were then assigned to 19 equally spaced bins across the axis.

To compute the astrocyte differentiation axis, a principal curve was calculated for the NSC and astrocyte gene module scores for all single NSCs and astrocytes. Cells with >30 UMIs in the IPC module were removed from downstream analysis. The developmental axis coordinate for each cell was the principal curve point that was closest to it (Euclidean distance). Cells were then assigned to 19 equally spaced bins across the axis.

Low-level filtering of scATAC profiles

Matrices summing ATAC UMIs in single cells in promoter regions ([–2000,200] bp from the TSS) and gene bodies (–2 kb from the TSS and along the gene body) and 500 bp

around all peaks were constructed. Only scATAC profiles with total UMI counts above the 10th and under the 98th percentiles in all matrices were retained. This removed 14,389 out of 42,141 (34%) of scATAC profiles. In addition, we filtered doublets using AMULET (Thibodeau et al. 2021) on each batch separately with default parameters.

Filtering of noncortical scATAC profiles

We downsampled the single-cell gene body accessibility matrix B_{gm} to the minimal library size (7603 UMIs). We then selected a multimodal feature gene set that was the union of RNA feature genes and genes variable in their gene body accessibility. We calculated the correlation between the deviation from the mean $(B_{gi} - \frac{1}{N}\sum_{i=1}^{N} B_{gi})$ and the normalized expression of the multimodal feature gene set across RNA metacells (containing noncortical states; i.e., as above in "Metacell Analysis and Filtering"). The gene body matrix was clustered based on this correlation, using k-means with $k = 185$ to generate ATAC microclusters. We calculated the correlation between mean microcluster accessibility and RNA metacell expression, assigned as noncortical ATAC microclusters those whose top 5 correlated RNA metacells were designated as noncortical in "Metacell Analysis and Filtering" above, and removed all 1563 (5.6% of total) scATAC profiles belonging to them.

TSS enrichment score

Only TSSs of genes expressed in the RNA metacell manifold were considered. The "flank" coverage was calculated as the mean marginal UMI count between 900 and 1000 bp (in each direction) from the TSSs. The TSS enrichment score was the fold change of mean coverage in 1 bp resolution in 900 bp around the TSS in each direction, over the flank coverage.

Peak calling

UMIs were pooled from all cells and averaged in 20 bp bins. Contiguous genomic intervals with coverage exceeding 195 UMIs, corresponding to ~99th percentile coverage, were identified. The intervals were centered on the position with maximal coverage, and uniform-sized peaks were generated by extending from the maximal point and 140 bp on each side. Overlapping peaks were unified and recentered until convergence. This yielded a peak set containing 122,400 peaks that were used for all subsequent analyses. We note that peak sets with higher sensitivity were considered but not used, with the goal of a conservative high-specificity set of epigenetic hotspots for downstream analysis. We classified peaks as TSSs if their distance to an annotated transcription start site was <1 kb. All other peaks were denoted as CREs.

Generation of ATAC microclusters (ATAC-MCs)

After filtering low-quality and noncortical cells, we downsampled the matrix of promoter accessibility to the minimum library size of 2023 UMIs. For each time point, we clustered the cells into 30 clusters using k-means based on the correlations between promoter accessibility and fold change of gene expression from the median across metacells on 1475 filtered feature genes (essentially identical to those used to generate RNA metacells).

Matching scATAC profiles to RNA metacells

We defined a linear problem that encodes an optimal assignment of scATAC profiles to RNA metacells. scATAC profiles from time point t can only be assigned to RNA metacells containing at least one cell from t, since cells for the scRNA-seq and scATAC-seq were sampled isogenically. We calculated the correlation between ATAC-MCs and the fold-change of gene expression from the median across RNA metacells for feature genes (as defined above).

We then constructed a bipartite graph connecting each ATAC-MC with the top 10 correlated RNA metacells, and similarly each RNA metacell with the top 10 correlated ATAC-MCs. This resulted in an unbalanced graph (with some ATAC-MCs and RNA MCs becoming high-degree hubs). We performed sampling of 10 edges for each MC and each ATAC-MC and combined them to improve balancing. This defined a set of edges $(i \text{ and } j) \in E$, where i is an ATAC-MC and j is an RNA MC.

We can now formulate the ATAC-MC-to-MC assignment problem linearly, searching for assignment values for each edge x_{ij} and potential additional overflow/underflow x_j^{s1}, x_j^{s2}, x_j^{s3}, and x_j^{s4}. We solved

$$\text{argmin}_x \sum_{(i,j)\in E} -\rho_{ij} x_{ij} + \sum_{o\in 1..4} w_o x_j^{so},$$

where ρ_{ij} is the Spearman correlation linking the accessibility and RNA profiles of the respective ATAC-MCs and RNA MCs, and the overflow/underflow weights w_0 are parameters (here $w_1 = -10$, $w_2 = -1$, $w_3 = 1$, and $w_4 = 10$) approximating a convex cost for imperfect assignment as described below.

This optimization is subject to the following constraints:

1. $\forall ij$, $x_{ij} \geq 0$ (assigned frequencies are nonnegative).
2. $\sum_{i,j} x_{ij} = 1$ (the total of assigned frequencies is forced to be 1).
3. $\forall i$, $\sum_j x_{ij} = p_i^{\text{ATAC}}$, where p_i^{ATAC} is the observed ATAC-MC frequency
4. $\forall j$, $\sum_i x_{ij} = \sum_o x_j^{so}$ (the total assigned frequency to each MC equals the total "frequency capacity" of each metacell).
5. $\forall j$, $x_j^{s1} \leq 0.8 p_j^{\text{RNA}}$, $x_j^{s2} \leq 0.2 p_j^{\text{RNA}}$, $x_j^{s3} \leq 0.2 p_j^{\text{RNA}}$, and $x_j^{s4} \leq 2 p_j^{\text{RNA}}$, where p_j^{RNA} is the metacell frequency for

the respective time points. These constraints define ranges of assigned frequency for each metacell and are matched with overflow/underflow weights (w_0, see above) such that when $x_j^{s1} + x_j^{s2}$ are maximized (negative), cost is minimized. Together with constraint 4, this encourages the total assignment to an RNA metacell to equal p_j^{RNA}.

We solved the linear problem with the "lpsymphony" R package for each embryonic day separately, with the solution giving assigned frequencies from each ATAC-MC to each RNA MC.

Resampling ATAC metacells given RNA assignments

Given x_{ij}, all cells belonging to ATAC-MC^i were randomly assigned to metacells according to the proportions $x_{ij}/\Sigma_j x_{ij}$. This gave an assignment of each valid scATAC profile in the data to an RNA MC. An scATAC matrix for all cells was constructed from the fragments.tsv output of the CellRanger ARC program using the mcATAC package (https://www.github.com/tanaylab/mcatac) functions write_scc_reads_from_fragments_file (to create an ScCounts object) and scc_extract (to convert ScCounts → scATAC matrix). Cells were removed if they had <6000 or >60,000 UMIs. We then aggregated scATAC profiles according to the RNA MC assignments using the mcATAC functions scc_project_on_mc and mcc_to_mcatac.

CRE group annotation

We separated the peak set into TSSs (<1 kb from the TSS of a gene expressed in the RNA manifold) and CREs (all other peaks). Both classes were clustered using *K*-means (TSSs: $k = 10$; CREs: $k = 60$). A subset of CRE clusters was identified as having substantially lower variability across the ATAC manifold along with qualitatively different sequence content relative to CRE clusters outside of the subset. CREs belonging to clusters in this subset were denoted as "constitutive CREs."

Calculation of putative transcription factor binding motif affinity

Sequence affinities were calculated for PWMs of interest as described previously (Schuettengruber et al. 2014).

Selection of neuron branch-specific CREs

For CfuPN- and IPC-specific CREs the difference between mean CfuPN (CthPN and SCPN) accessibility and CPN (CPN_L5-6 and CPN_L2-3) accessibility was >1, and the absolute difference between mean CfuPN accessibility and IPC accessibility was <1.

For CPN- and IPC-specific CRE, the difference between mean CfuPN accessibility and CPN accessibility was <–1, and the absolute difference between mean CPN accessibility and IPC accessibility was <1.

For pan-neuronal CREs, the absolute difference between mean CfuPN accessibility and CPN accessibility was <1.

For CfuPN-specific CREs, the difference between mean CfuPN accessibility and CPN accessibility was >1, and the difference between mean CfuPN accessibility and IPC accessibility was >1 and not in the "pan-neuronal CRE" group.

For CPN-specific CREs, the difference between mean CfuPN accessibility and CPN accessibility was <–1, and the difference between mean CPN accessibility and IPC accessibility was >1 and not in the "pan-neuronal CRE" group.

Normalization of ATAC signal

The following normalization was used for comparing trends in Figure 3, E–I: For each CRE, the (punctured) marginal coverage was calculated in windows of 10 kb radius around the CRE, and accessibility (UMIs/metacell) was normalized by dividing it by the punctured marginal coverage while setting the minimum to quantile 0.1 of the punctured marginal coverage and then linearly transforming the data to the range [0,1].

Kinetics of branch-specific CREs across neuronal trajectories

Trajectories were defined as all metacells belonging to defined cell types. For CfuPNs, this included NSCs, IPCs, IPC_cycs, iCPNs/CfuPNs, iCfuPNs, SCPNs, and CthPNs. For CPNs, this included NSCs, IPCs, IPC_cycs, iCPN_early, iCPN_late, CPN_L5-6, and CPN_L2-3. We defined an NSC-specific gene module as all genes with normalized expression twofold higher in NSCs relative to IPCs and glial cell types and defined a neuron-specific gene module as all genes with mean expression across all mature neuron cell types twofold higher than mean expression in IPCs. The NSC score was the summed expression of NSC genes; likewise for the neuron score. Metacells were ordered and binned ($n = 20$) by the difference between the neuron and NSC scores. Theorder and binning were used to derive kinetics of CfuPN- and CPN-specific CREs and cell type fractions that are plotted in Figure 3, D–I.

Identification of temporally increasing/decreasing CREs in NSCs

We derived the per-day accessibility profile of NSCs by multiplying the raw metacell accessibility matrix of NSCs (CREs × metacells) by the matrix of their per-day composition (metacells × days). We transformed this matrix to log fractions by normalizing by total UMIs per day and taking the $\log_2$ (plus pseudocount of 1×10^{-5}) of the normalized matrix. Temporally increasing CREs were those with a E17–E13 ΔATAC ≥ 1, and temporally decreasing CREs had E17–E13 ΔATAC ≤ -1 intersected with the list of CREs with robust methylation signal (as defined below).

Methylation coverage filtering

We counted 5mC reads in all CREs of the ATAC manifold and discarded those with <20 reads/CRE in all time points, retaining ~99,000 out of 122,000 CREs.

Identification of cell type-specific CREs for methylation analysis

NSCs were identified by mean accessibility (log fraction) in NSCs ≥0.75 of the mean in IPCs and ≥0.75 of the mean in astrocytes.

IPCs were identified by mean accessibility (log fraction) in IPCs ≥0.75 of the mean in NSCs and ≥0.75 of the mean in mature neural cell types (CthPN, SCPN, CPN_L2-3, and CPN_L5-6).

Neurons were identified by mean accessibility (log fraction) in mature neuronal cell types ≥0.75 of the mean in IPCs.

Astrocytes were identified by mean accessibility (log fraction) in astrocytes ≥0.75 of the mean in NSCs and ≥0.75 of the mean in oligodendrocytes.

Hi-C—insulation score

Scores in each time point (E13–E17) were calculated in 250 kb windows, with 1 kb step size, as described previously (Bonev 2017). High-variance region seeds were 1 kb intervals in which the difference between E17 and E13 insulation was >0.25 or smaller than –0.35. Neighboring seeds were merged, and the resulting regions <5 kb were discarded. Mean insulation was calculated for each region in each time point, and the negative of the insulation score was used for visualization (such that higher score = higher insulation).

Neighboring CRE enrichment in differential insulation groups

CREs with mean accessibility ≥–15.5 in NSCs were selected, their distances from high-variance insulation regions were calculated, and the distribution of neighbors in 1 kb distance bins from each type of high-variance insulation region (increasing/decreasing across time points) was calculated. We smoothed the distribution in 100 kb windows and divided it by the number of regions in each group to get the distributions plotted in Figure 5H. To check for significant differences in proximal CRE abundance, all neighboring CREs' distances were binned into 50 kb-sized disjoint bins, and the distribution of the normalized number of neighbors in each bin was compared between insulating and deinsulating regions.

Normalized SHAMAN D scores between temporally activating CREs in NSCs and cell type-specific TSSs

Cell type-specific TSSs were selected similarly to the procedure used to select gene modules, where NSC-specific genes were selected based on their differential expression relative to astrocytes and IPCs, IPC-specific genes were differentially expressed relative to NSCs and mature neurons, and astrocyte-specific genes were differentially expressed relative to NSCs and OPCs (all fold changes of 1). Temporally activating CREs in NSCs were chosen as those whose accessibility per time point had a correlation of at least 0.5 with the time point index and with a fold change of at least 0.75 between minimum and maximum accessibility per time point. Distances were calculated between temporally activating CREs and cell type-specific TSSs and cut at 1 Mb radius. Two-dimensional intervals for SHAMAN *D* score calculation were generated between the CREs and 10 kb diameter windows around the TSSs. In each 2D interval, SHAMAN *D* scores were calculated, and the mean was subtracted for each TSS group separately.

Preparation of training data for accessibility prediction model

We used ICEQREAM (Bercovich et al. 2025) to perform sequence regression on the IPC–NSC ATAC difference, with parameters max_motif_num = 16, n_prego_motifs = 4, and spat_bin_size = 10. For each CRE, ATAC UMIs from all CREs within the distance (1 and 50 kb) of that CRE from NSC scATAC profiles were summed. For each CRE, RNA UMIs from all genes within the distance (1 and 500 kb) of that CRE from NSC scRNA profiles were summed. All nonconstitutive CREs with sufficient methylation coverage were split into 11 equally sized groups, and XGBoost (R package) was run with cross-validation on 10 of these groups with the parameters "md" = 3, "eta" = 0.3, "nr" = 250, obj = "reg:squarederror", and em = "rmse." SHAP values were extracted using the option predcontrib = TRUE. The R^2 shown here is the one obtained by averaging over the R^2 in each of the held-out groups.

Testing ATAC predictive model of general enhancers

All proximal and distal enhancers based on ENCODE cCREs (Moore et al. 2020) were collected and unified into a fixed length and no overlap by union and centering. The new interval set was merged with the cortex interval set, and total ATAC UMIs in NSCs and IPCs were summed for the combined interval set and normalized. All other sequence and epigenomic features (nucleotide and dinucleotide content, methylation, and proximal ATAC and RNA UMIs) were calculated exactly the same as the cortex interval set, and the XGBoost regression model was fit with the same hyperparameters and 10-fold cross-validation.

Stratification of CREs by E-box affinity, number of proximal elements with high E-box/T-box, and methylation

Distances between all CREs were calculated. E-box and T-box affinity (prego-inferred "E_box_1" and "T_box_1" PWMs shown in Supplemental Fig. S6A) was considered "high" when above the 0.8 quantile. The number of neighboring high-affinity E-box/T-box elements per CRE was

capped at five each and seven total. Active CREs were considered those with mean NSC or IPC ATAC > −16, and the number of neighboring active CREs per CRE was capped at 6.

MPRA signal cleanup and normalization

Only eCREs that were assigned a *P*-value of <0.1 by the MPRAnalyze pipeline in at least one time point were used. Out of these, we selected eCREs whose original genomic intervals overlapped with the CRE set of our ATAC manifold and were above quantile 0.75 in terms of mean MPRA activity in NSCs or IPCs. MPRAnalyze MAD scores for all time points and cell types were centered (mean = 0, SD = 1). Two libraries with low QC (correspondence of MPRA and ATAC signal) were removed.

Generating XGBoost model predictions for MPRA library sequences

Sequence affinities for PWMs obtained from the ICE-QREAM model were calculated for all filtered eCREs. We assumed a value of 0 methylation and intermediate-level ATAC (mean value of NSC ATAC) for all CREs. Dinucleotide content was recalculated, and the rest of the features were imputed as "GC content" = 0.05, "prox_ATAC" = 10, and "prox_RNA" = 10 for all CREs. This aimed to emulate the activated state of the CRE sequence.

Data availability

The sequencing data generated in this study can be accessed at the GEO database with the accession numbers GSE292318, GSE292319, GSE292320, and GSE292321. The E14 data (scRNA, scATAC, and methyl-Hi-C) are available from GEO with accession number GSE155677.

Code availability

The code used for generating the analysis and all of the figures is available at https://www.github.com/tanaylab/mmcortex. Methyl Hi-C data were processed using the JuiceMe pipeline, available at https://www.github.com/aidenlab/JuiceMe. The R package to compute the expected tracks and the Hi-C scores is available at https://www.github.com/tanaylab/shaman.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

Sequencing was performed at the Helmholtz Munich by the Next-Generation Sequencing Core Facility and cell sorting was done at the Flow Cytometry Core Facility at the Biomedical Center, Ludwig-Maximilians-Universität. Work in the group of B.B. was supported by the Network of European Funding for Neuroscience Research (ERA-NET NEURON) grant "Molecular and circuits bases of epileptogenic mosaic cortical malformations" (MOSAIC) and European Research Council Consolidator Grant (EpiCortex, 101044469). Work in the group of A.T. was supported by a European Research Council Advanced Grant (Cells2Tissues) and the Israeli Science Foundation. Y.S. thanks all Tanay laboratory members for technical help and useful discussions.

Author contributions: F.N. and S.V. performed the experiments. Y.S., F.C., A.T., and B.B. analyzed the data. A.L. established the MCView interface. B.B. and A.T. supervised the project. Y.S., A.T., and B.B. wrote the manuscript with input from all authors.

References

Amberg N, Laukoter S, Hippenmeyer S. 2019. Epigenetic cues modulating the generation of cell-type diversity in the cerebral cortex. *J Neurochem* **149:** 12–26. doi:10.1111/jnc.14601

Ashuach T, Fischer DS, Kreimer A, Ahituv N, Theis FJ, Yosef N. 2019. MPRAnalyze: statistical framework for massively parallel reporter assays. *Genome Biol* **20:** 183. doi:10.1186/s13059-019-1787-z

Aydin B, Kakumanu A, Rossillo M, Moreno-Estellés M, Garipler G, Ringstad N, Flames N, Mahony S, Mazzoni EO. 2019. Proneural factors Ascl1 and Neurog2 contribute to neuronal subtype identities by establishing distinct chromatin landscapes. *Nat Neurosci* **22:** 897–908. doi:10.1038/s41593-019-0399-y

Baizabal J-M, Mistry M, García MT, Gómez N, Olukoya O, Tran D, Johnson MB, Walsh CA, Harwell CC. 2018. The epigenetic state of PRDM16-regulated enhancers in radial glia controls cortical neuron position. *Neuron* **98:** 945–962.e8. doi:10.1016/j.neuron.2018.04.033

Baran Y, Bercovich A, Sebe-Pedros A, Lubling Y, Giladi A, Chomsky E, Meir Z, Hoichman M, Lifshitz A, Tanay A. 2019. Metacell: analysis of single-cell RNA-seq data using K-nn graph partitions. *Genome Biol* **20:** 206. doi:10.1186/s13059-019-1812-2

Beagan JA, Duong MT, Titus KR, Zhou L, Cao Z, Ma J, Lachanski CV, Gillis DR, Phillips-Cremins JE. 2017. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res* **27:** 1139–1152. doi:10.1101/gr.215160.116

Ben-Kiki O, Bercovich A, Lifshitz A, Tanay A. 2022. Metacell-2: a divide-and-conquer metacell algorithm for scalable scRNA-seq analysis. *Genome Biol* **23:** 100. doi:10.1186/s13059-022-02667-1

Bercovich A, Lifshitz A, Eldar M, Cheng S, Stok Ranen R, Stelzer Y, Tanay A. 2025. IceQream: quantitative chromosome accessibility analysis using physical TF models. *Nat Commun* **16:** 8984. doi:10.1038/s41467-025-63925-x

Bonev B, Mendelson Cohen N, Szabo Q, Fritsch L, Papadopoulos GL, Lubling Y, Xu X, Lv X, Hugnot J-P, Tanay A, et al. 2017. Multiscale 3D genome rewiring during mouse neural development. *Cell* **171:** 557–572.e24. doi:10.1016/j.cell.2017.09.043

Braun E, Danan-Gotthold M, Borm LE, Lee KW, Vinsland E, Lönnerberg P, Hu L, Li X, He X, Andrusivová Ž, et al. 2023. Comprehensive cell atlas of the first-trimester developing human brain. *Science* **382:** eadf1226. doi:10.1126/science.adf1226

Calegari F, Haubensak W, Haffner C, Huttner WB. 2005. Selective lengthening of the cell cycle in the neurogenic subpopulation of neural progenitor cells during mouse brain development. *J*

Neurosci **25:** 6533–6538. doi:10.1523/JNEUROSCI.0778-05.2005

Di Bella DJ, Habibi E, Stickels RR, Scalia G, Brown J, Yadollahpour P, Yang SM, Abbate C, Biancalani T, Macosko EZ, et al. 2021. Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **595:** 554–559. doi:10.1038/s41586-021-03670-5

Dileep V, Boix CA, Mathys H, Marco A, Welch GM, Meharena HS, Loon A, Jeloka R, Peng Z, Bennett DA, et al. 2023. Neuronal DNA double-strand breaks lead to genome structural variations and 3D genome disruption in neurodegeneration. *Cell* **186:** 4404–4421.e20. doi:10.1016/j.cell.2023.08.038

Eze UC, Bhaduri A, Haeussler M, Nowakowski TJ, Kriegstein AR. 2021. Single-cell atlas of early human brain development highlights heterogeneity of human neuroepithelial cells and early radial glia. *Nat Neurosci* **24:** 584–594. doi:10.1038/s41593-020-00794-1

Fan G, Martinowich K, Chin MH, He F, Fouse SD, Hutnick L, Hattori D, Ge W, Shen Y, Wu H, et al. 2005. DNA methylation controls the timing of astrogliogenesis through regulation of JAK-STAT signaling. *Development* **132:** 3345–3356. doi:10.1242/dev.01912

Gao P, Postiglione MP, Krieger TG, Hernandez L, Wang C, Han Z, Streicher C, Papusheva E, Insolera R, Chugh K, et al. 2014. Deterministic progenitor behavior and unitary production of neurons in the neocortex. *Cell* **159:** 775–788. doi:10.1016/j.cell.2014.10.027

Gordon MG, Inoue F, Martin B, Schubach M, Agarwal V, Whalen S, Feng S, Zhao J, Ashuach T, Ziffra R, et al. 2020. lentiMPRA and MPRAflow for high-throughput functional characterization of gene regulatory elements. *Nat Protoc* **15:** 2387–2412. doi:10.1038/s41596-020-0333-5

Govindan S, Jabaudon D. 2017. Coupling progenitor and neuronal diversity in the developing neocortex. *FEBS Lett* **591:** 3960–3977. doi:10.1002/1873-3468.12846

Greig LC, Woodworth MB, Galazo MJ, Padmanabhan H, Macklis JD. 2013. Molecular logic of neocortical projection neuron specification, development and diversity. *Nat Rev Neurosci* **14:** 755–769. doi:10.1038/nrn3586

Hirabayashi Y, Gotoh Y. 2010. Epigenetic control of neural precursor cell fate during development. *Nat Rev Neurosci* **11:** 377–388. doi:10.1038/nrn2810

Hirabayashi Y, Suzki N, Tsuboi M, Endo TA, Toyoda T, Shinga J, Koseki H, Vidal M, Gotoh Y. 2009. Polycomb limits the neurogenic competence of neural precursor cells to promote astrogenic fate transition. *Neuron* **63:** 600–613. doi:10.1016/j.neuron.2009.08.021

Hirayama T, Tarusawa E, Yoshimura Y, Galjart N, Yagi T. 2012. CTCF is required for neural development and stochastic expression of clustered Pcdh genes in neurons. *Cell Rep* **2:** 345–357. doi:10.1016/j.celrep.2012.06.014

Kishi Y, Fujii Y, Hirabayashi Y, Gotoh Y. 2012. HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. *Nat Neurosci* **15:** 1127–1133. doi:10.1038/nn.3165

Koo B, Lee K-H, Ming G, Yoon K-J, Song H. 2023. Setting the clock of neural progenitor cells during mammalian corticogenesis. *Semin Cell Dev Biol* **142:** 43–53. doi:10.1016/j.semcdb.2022.05.013

Kuwayama N, Kujirai T, Kishi Y, Hirano R, Echigoya K, Fang L, Watanabe S, Nakao M, Suzuki Y, Ishiguro K, et al. 2023. HMGA2 directly mediates chromatin condensation in association with neuronal fate regulation. *Nat Commun* **14:** 6420. doi:10.1038/s41467-023-42094-9

Kwan KY, Lam MMS, Krsnik Z, Kawasawa YI, Lefebvre V, Šestan N. 2008. SOX5 postmitotically regulates migration, postmigratory differentiation, and projections of subplate and deep-layer neocortical neurons. *Proc Natl Acad Sci* **105:** 16021–16026. doi:10.1073/pnas.0806791105

La Manno G, Siletti K, Furlan A, Gyllborg D, Vinsland E, Mossi Albiach A, Mattsson Langseth C, Khven I, Lederer AR, Dratva LM, et al. 2021. Molecular architecture of the developing mouse brain. *Nature* **596:** 92–96. doi:10.1038/s41586-021-03775-x

Li Z, Tyler WA, Zeldich E, Santpere Baró G, Okamoto M, Gao T, Li M, Sestan N, Haydar TF. 2020. Transcriptional priming as a conserved mechanism of lineage diversification in the developing mouse and human neocortex. *Sci Adv* **6:** eabd2068. doi:10.1126/sciadv.abd2068

Llorca A, Ciceri G, Beattie R, Wong FK, Diana G, Serafeimidou-Pouliou E, Fernández-Otero M, Streicher C, Arnold SJ, Meyer M, et al. 2019. A stochastic framework of neurogenesis underlies the assembly of neocortical cytoarchitecture. *eLife* **8:** e51381. doi:10.7554/eLife.51381

Lodato S, Arlotta P. 2015. Generating neuronal diversity in the mammalian cerebral cortex. *Annu Rev Cell Dev Biol* **31:** 699–720. doi:10.1146/annurev-cellbio-100814-125353

Lodato S, Molyneaux BJ, Zuccaro E, Goff LA, Chen H-H, Yuan W, Meleski A, Takahashi E, Mahony S, Rinn JL, et al. 2014. Gene co-regulation by Fezf2 selects neurotransmitter identity and connectivity of corticospinal neurons. *Nat Neurosci* **17:** 1046–1054. doi:10.1038/nn.3757

Mannens CCA, Hu L, Lönnerberg P, Schipper M, Reagor CC, Li X, He X, Barker RA, Sundström E, Posthuma D, et al. 2025. Chromatin accessibility during human first-trimester neurodevelopment. *Nature* **647:** 179–186. doi:10.1038/s41586-024-07234-1

Mayshar Y, Raz O, Cheng S, Ben-Yair R, Hadas R, Reines N, Mittnenzweig M, Ben-Kiki O, Lifshitz A, Tanay A, et al. 2023. Time-aligned hourglass gastrulation models in rabbit and mouse. *Cell* **186:** 2610–2627.e18. doi:10.1016/j.cell.2023.04.037

McGinnis CS, Murrow LM, Gartner ZJ. 2019. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst* **8:** 329–337.e4. doi:10.1016/j.cels.2019.03.003

Mellén M, Ayata P, Heintz N. 2017. 5-Hydroxymethylcytosine accumulation in postmitotic neurons results in functional demethylation of expressed genes. *Proc Natl Acad Sci* **114:** E7812–E7821. doi:10.1073/pnas.1708044114

Mittnenzweig M, Mayshar Y, Cheng S, Ben-Yair R, Hadas R, Rais Y, Chomsky E, Reines N, Uzonyi A, Lumerman L, et al. 2021. A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell* **184:** 2825–2842.e22. doi:10.1016/j.cell.2021.04.004

Molyneaux BJ, Arlotta P, Menezes JRL, Macklis JD. 2007. Neuronal subtype specification in the cerebral cortex. *Nat Rev Neurosci* **8:** 427–437. doi:10.1038/nrn2151

Monahan K, Horta A, Lomvardas S. 2019. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **565:** 448–453. doi:10.1038/s41586-018-0845-0

Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shoresh N, Adrian J, Kawli T, Davis CA, Dobin A, Kaul R, et al. 2020. Expanded encyclopaedias of DNA elements in the human and mouse genomes. *Nature* **583:** 699–710. doi:10.1038/s41586-020-2493-4

Morimoto-Suzki N, Hirabayashi Y, Tyssowski K, Shinga J, Vidal M, Koseki H, Gotoh Y. 2014. The polycomb component Ring1B regulates the timed termination of subcerebral

projection neuron production during mouse neocortical development. *Development* **141:** 4343–4353. doi:10.1242/dev.112276

Nagao M, Lanjakornsiripan D, Itoh Y, Kishi Y, Ogata T, Gotoh Y. 2014. High mobility group nucleosome-binding family proteins promote astrocyte differentiation of neural precursor cells. *Stem Cells* **32:** 2983–2997. doi:10.1002/stem.1787

Ninkovic J, Steiner-Mezzadri A, Jawerka M, Akinci U, Masserdotti G, Petricca S, Fischer J, von Holst A, Beckers J, Lie CD, et al. 2013. The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. *Cell Stem Cell* **13:** 403–418. doi:10.1016/j.stem.2013.07.002

Noack F, Vangelisti S, Raffl G, Carido M, Diwakar J, Chong F, Bonev B. 2022. Multimodal profiling of the transcriptional regulatory landscape of the developing mouse cortex identifies Neurog2 as a key epigenome remodeler. *Nat Neurosci* **25:** 154–167. doi:10.1038/s41593-021-01002-4

Noack F, Vangelisti S, Ditzer N, Chong F, Albert M, Bonev B. 2023. Joint epigenome profiling reveals cell-type-specific gene regulatory programmes in human cortical organoids. *Nat Cell Biol* **25:** 1873–1883. doi:10.1038/s41556-023-01296-5

Oberst P, Fièvre S, Baumann N, Concetti C, Bartolini G, Jabaudon D. 2019. Temporal plasticity of apical progenitors in the developing mouse neocortex. *Nature* **573:** 370–374. doi:10.1038/s41586-019-1515-6

Pereira JD, Sansom SN, Smith J, Dobenecker M-W, Tarakhovsky A, Livesey FJ. 2010. Ezh2, the histone methyltransferase of PRC2, regulates the balance between self-renewal and differentiation in the cerebral cortex. *Proc Natl Acad Sci* **107:** 15957–15962. doi:10.1073/pnas.1002530107

Pereira A, Diwakar J, Masserdotti G, Beşkardeş S, Simon T, So Y, Martín-Loarte L, Bergemann F, Vasan L, Schauer T, et al. 2024. Direct neuronal reprogramming of mouse astrocytes is associated with multiscale epigenome remodeling and requires Yy1. *Nat Neurosci* **27:** 1260–1273. doi:10.1038/s41593-024-01677-5

Pollen AA, Nowakowski TJ, Chen J, Retallack H, Sandoval-Espinosa C, Nicholas CR, Shuga J, Liu SJ, Oldham MC, Diaz A, et al. 2015. Molecular identity of human outer radial glia during cortical development. *Cell* **163:** 55–67. doi:10.1016/j.cell.2015.09.004

Sahara S, Kodama T, Stevens CF. 2020. A common rule governing differentiation kinetics of mouse cortical progenitors. *Proc Natl Acad Sci* **117:** 15221–15229. doi:10.1073/pnas.1916665117

Schuettengruber B, Oded Elkayam N, Sexton T, Entrevan M, Stern S, Thomas A, Yaffe E, Parrinello H, Tanay A, Cavalli G. 2014. Cooperativity, specificity, and evolutionary stability of polycomb targeting in *Drosophila*. *Cell Rep* **9:** 219–233. doi:10.1016/j.celrep.2014.08.072

Shen Q, Wang Y, Dimos JT, Fasano CA, Phoenix TN, Lemischka IR, Ivanova NB, Stifani S, Morrisey EE, Temple S. 2006. The timing of cortical neurogenesis is encoded within lineages of individual progenitor cells. *Nat Neurosci* **9:** 743–751. doi:10.1038/nn1694

Song M, Pebworth M-P, Yang X, Abnousi A, Fan C, Wen J, Rosen JD, Choudhary MNK, Cui X, Jones IR, et al. 2020. Cell-type-specific 3D epigenomes in the developing human cortex. *Nature* **587:** 644–649. doi:10.1038/s41586-020-2825-4

Stoyanova E, Riad M, Rao A, Heintz N. 2021. 5-Hydroxymethylcytosine-mediated active demethylation is required for mammalian neuronal differentiation and function. *eLife* **10:** e66973. doi:10.7554/eLife.66973

Takizawa T, Nakashima K, Namihira M, Ochiai W, Uemura A, Yanagisawa M, Fujita N, Nakao M, Taga T. 2001. DNA methylation is a critical cell-intrinsic determinant of astrocyte differentiation in the fetal brain. *Dev Cell* **1:** 749–758. doi:10.1016/S1534-5807(01)00101-0

Telley L, Govindan S, Prados J, Stevant I, Nef S, Dermitzakis E, Dayer A, Jabaudon D. 2016. Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. *Science* **351:** 1443–1446. doi:10.1126/science.aad8361

Telley L, Agirman G, Prados J, Amberg N, Fièvre S, Oberst P, Bartolini G, Vitali I, Cadilhac C, Hippenmeyer S, et al. 2019. Temporal patterning of apical progenitors and their daughter neurons in the developing neocortex. *Science* **364:** eaav2522. doi:10.1126/science.aav2522

Thibodeau A, Eroglu A, McGinnis CS, Lawlor N, Nehar-Belaid D, Kursawe R, Marches R, Conrad DN, Kuchel GA, Gartner ZJ, et al. 2021. AMULET: a novel read count-based method for effective multiplet detection from single nucleus ATAC-seq data. *Genome Biol* **22:** 252. doi:10.1186/s13059-021-02469-x

Trevino AE, Müller F, Andersen J, Sundaram L, Kathiria A, Shcherbina A, Farh K, Chang HY, Pașca AM, Kundaje A, et al. 2021. Chromatin and gene-regulatory dynamics of the developing human cerebral cortex at single-cell resolution. *Cell* **84:** 5053–5069.e23. doi:10.1016/j.cell.2021.07.039

Tsyporin J, Tastad D, Ma X, Nehme A, Finn T, Huebner L, Liu G, Gallardo D, Makhamreh A, Roberts JM, et al. 2021. Transcriptional repression by FEZF2 restricts alternative identities of cortical projection neurons. *Cell Rep* **35:** 109269. doi:10.1016/j.celrep.2021.109269

van den Berg DLC, Azzarelli R, Oishi K, Martynoga B, Urbán N, Dekkers DHW, Demmers JA, Guillemot F. 2017. Nipbl interacts with Zfp609 and the integrator complex to regulate cortical neuron migration. *Neuron* **93:** 348–361. doi:10.1016/j.neuron.2016.11.047

Vitali I, Fièvre S, Telley L, Oberst P, Bariselli S, Frangeul L, Baumann N, McMahon JJ, Klingler E, Bocchi R, et al. 2018. Progenitor hyperpolarization regulates the sequential generation of neuronal subtypes in the developing neocortex. *Cell* **174:** 1264–1276.e15. doi:10.1016/j.cell.2018.06.036

Xiong X, James BT, Boix CA, Park YP, Galani K, Victor MB, Sun N, Hou L, Ho L-L, Mantero J, et al. 2023. Epigenomic dissection of Alzheimer's disease pinpoints causal variants and reveals epigenome erosion. *Cell* **186:** 4422–4437.e21. doi:10.1016/j.cell.2023.08.040

Zenk F, Fleck JS, Jansen SMJ, Kashanian B, Eisinger B, Santel M, Dupré J-S, Camp JG, Treutlein B. 2024. Single-cell epigenomic reconstruction of developmental trajectories from pluripotency in human neural organoid systems. *Nat Neurosci* **27:** 1376–1386. doi:10.1038/s41593-024-01652-0

Ziffra RS, Kim CN, Ross JM, Wilfert A, Turner TN, Haeussler M, Casella AM, Przytycki PF, Keough KC, Shin D, et al. 2021. Single-cell epigenomics reveals mechanisms of human cortical development. *Nature* **598:** 205–213. doi:10.1038/s41586-021-03209-8

Neural stem cell epigenomes and fate bias are temporally coordinated during mouse cortical development

Yonatan Shapira, Florian Noack, Silvia Vangelisti, et al.

Genes Dev. published online April 22, 2026
Access the most recent version at doi:10.1101/gad.353090.125