A single-cell atlas of spatial and temporal gene expression in the mouse cranial neural plate

Eric R. Brooks[1,2], Andrew R. Moorman[3], Bhaswati Bhattacharya[1], Ian Prudhomme[1], Max Land[3], Heather L. Alcorn[1], Roshan Sharma[3], Dana Pe'er[3], and Jennifer A. Zallen[1,*]

[1]HHMI and Developmental Biology Program, Sloan Kettering Institute
[2]North Carolina State University
[3]HHMI and Computational and Systems Biology Program, Sloan Kettering Institute
*For correspondence: zallenj@mskcc.org

Abstract

The formation of the mammalian brain requires regionalization and morphogenesis of the cranial neural plate, which transforms from an epithelial sheet into a closed tube that provides the structural foundation for neural patterning and circuit formation. Sonic hedgehog (SHH) signaling is important for cranial neural plate patterning and closure, but the transcriptional changes that give rise to the spatially regulated cell fates and behaviors that build the cranial neural tube have not been systematically analyzed. Here we used single-cell RNA sequencing to generate an atlas of gene expression at six consecutive stages of cranial neural tube closure in the mouse embryo. Ordering transcriptional profiles relative to the major axes of gene expression predicted spatially regulated expression of 870 genes along the anterior-posterior and mediolateral axes of the cranial neural plate and reproduced known expression patterns with over 85% accuracy. Single-cell RNA sequencing of embryos with activated SHH signaling revealed distinct SHH-regulated transcriptional programs in the developing forebrain, midbrain, and hindbrain, suggesting a complex interplay between anterior-posterior and mediolateral patterning systems. These results define a spatiotemporally resolved map of gene expression during cranial neural tube closure and provide a resource for investigating the transcriptional events that drive early mammalian brain development.

Introduction

Development of the mammalian brain requires coordination between the genetic programs that generate cell fate and the dynamic and spatially regulated cell behaviors that establish tissue structure. The cranial neural plate is an epithelial sheet that undergoes neuronal differentiation and structural remodeling to produce functionally distinct brain regions in the forebrain, midbrain, and hindbrain. The forebrain gives rise to the cerebral cortex, thalamus, and hypothalamus, the midbrain generates the visual and auditory systems, and the hindbrain forms the cerebellum and other structures that control cognitive, emotional, autonomic, and motor functions (Chen et al., 2017). The generation of these three domains is controlled by a cartesian landscape of transcriptional information that directs cell fate and behavior along the anterior-posterior and mediolateral axes of the cranial neural plate, guided by conserved signals in the WNT, BMP, SHH,

and FGF families and retinoic acid levels (Liu and Joyner, 2001a; Wurst and Bally-Cuif, 2001; Maden, 2007; Kicheva and Briscoe, 2023). However, the transcriptional programs that drive neural tube patterning and closure to provide a blueprint for the fine-scale organization of the mammalian brain are not well understood.

Single cell RNA sequencing (scRNA-seq) is a powerful tool for analyzing gene expression during development (Tam and Ho, 2020). This approach has been used to describe gene expression in several mouse organs (Cardoso-Moreira et al., 2019; Delile et al., 2019; de Soysa et al., 2019; Nowotschin et al., 2019; Soldatov et al., 2019; La Manno et al., 2021; Zalc et al., 2021) and entire mouse embryos at different stages (Argelaguet et al., 2019; Cao et al., 2019; Cheng et al., 2019; Pijuan-Sala et al., 2019; Mittnenzweig et al., 2021; Sampath Kumar et al., 2023). Region-specific transcriptional signatures in the mouse cranial neural plate are first detected between embryonic days 7.0 and 8.5 (Pijuan-Sala et al., 2019; La Manno et al., 2021; Lohoff et al., 2022; Sampath Kumar et al., 2023) and gene expression profiles in neuronal populations have been characterized in detail at later stages of brain development (Rosenberg et al., 2018; Li et al., 2019; Wizeman et al., 2019; Wang et al., 2020; Di Bella et al., 2021; La Manno et al., 2021; Ruan et al., 2021; Langlieb et al., 2023; Yao et al., 2023; Zhang et al., 2023). However, although the transcriptional changes that partition cranial neuroepithelial progenitors into distinct domains are beginning to be identified, significant questions remain. First, cranial neural plate differentiation and closure has generally been studied at single time points, and limited temporal information is available about transcriptional dynamics and cell-state transitions over time. Second, the spatially regulated transcriptional programs activated by morphogen patterning systems have not been systematically analyzed, as traditional methods for assaying spatial gene expression are relatively low-throughput. In particular, the transcriptional changes that pattern the cranial neuroepithelium and give rise to the forebrain, midbrain, and hindbrain have not been systematically analyzed. Third, the transcriptional basis of pathological signaling outcomes, such as the effects of increased Sonic hedgehog signaling on tissue patterning (Dessaud et al., 2008; Ribes and Briscoe, 2009; Goetz and Anderson, 2010) and cranial neural tube closure (Murdoch and Copp, 2010; Brooks et al., 2020), are not well understood.

To address these questions, we used scRNA-seq to analyze the spatial and temporal regulation of gene expression in the developing mouse cranial neural plate. We obtained gene expression profiles for 39,463 cells from the cranial region of mouse embryos at six stages from day 7.5 to day 9.0 of embryonic development, during which the cranial neuroepithelium undergoes dramatic changes in tissue patterning and organization, culminating in neural tube closure. Analysis of the 17,695 cranial neural plate cells in this dataset revealed distinct gene expression trajectories over time in the developing forebrain, midbrain, and hindbrain. In addition, we used this dataset to computationally reconstruct a high-resolution map of gene expression in the E8.5-9.0 cranial neural plate at late stages of neural tube closure. This map predicted spatially regulated expression for 870 genes along the anterior-posterior and mediolateral axes of the cranial neural plate, including 687 genes whose expression had not previously been characterized, and recapitulated the expression patterns of known genes with over 85% accuracy. Finally, we used scRNA-seq analysis to systematically investigate the consequences of activated SHH signaling, which disrupts neural tube patterning and closure. This analysis revealed region-specific

transcriptional responses to SHH signaling in the forebrain, midbrain, and hindbrain, suggesting complex interactions between anterior-posterior and mediolateral patterning systems. Together, these results define a spatially and temporally resolved map of gene expression during cranial neural tube closure and provide a resource for examining the early transcriptional events that drive mammalian brain development.

Results

Temporal evolution of region-specific transcriptional programs in the cranial neural plate

To characterize gene expression during cranial neural tube patterning and closure, we performed scRNA-seq analysis of manually dissected cranial regions from mouse embryos at six consecutive stages spanning embryonic days (E) 7.5 to 9.0 of development (Materials and methods). This dataset includes two stages before neural tube closure (0 somites and 1-2 somites, corresponding to E7.5-7.75), two stages during early closure as the neural plate bends and elevates (3 somites and 4-6 somites, E8.0-8.25), and two stages during late closure as the lateral borders meet to form a closed tube (7-9 somites and 10+ somites, E8.5-9.0) (Figures 1A and B). Using stringent criteria to filter out doublets and low-quality cells (Materials and methods), we obtained 39,463 cranial cells from 30 samples (3-5 embryos/replicate, 4-6 replicates/stage), with a median library depth of >42,000 transcripts (unique molecular identifiers or UMIs)/cell and >5,900 genes/cell (average 1.56 reads/UMI) (Table S1). Using the PhenoGraph clustering method (Levine et al., 2015) after correcting for cell-cycle stage (Materials and methods), we identified 29 transcriptionally distinct clusters in the cranial region that represent 7 cell types based on known markers: neural plate (17,695 cells), mesoderm (13,952 cells), non-neural ectoderm (2,252 cells), neural crest (2,993 cells), endoderm (2,062 cells), notochord (449 cells), and blood cells (60 cells) (Figure 1C, Figure 1—figure supplement 1, Tables S1 and S2). Thus, this dataset captures all major cranial populations in the mouse embryo.

To examine gene expression in the cranial neural plate in more detail, we reclustered the 17,695 cranial neural plate cells in our dataset and identified 15 clusters using PhenoGraph analysis (Figure 1—figure supplement 2A and B, Table S3). These clusters did not feature markers of neuronal or radial glial cells, consistent with the later onset of neurogenesis in the cranial neural plate (Götz and Huttner, 2005), although we observed modest expression of neuron-specific cytoskeletal regulators (Figure 1—figure supplement 2C). Clusters instead appeared to distinguish spatial and temporal properties of cells, as transcriptional profiles were strongly correlated with embryonic stage for all cranial populations (Figure 1D) and within the cranial neural plate (Figure 1E), and gene expression in the cranial neural plate was strongly correlated with anterior-posterior location (Figure 1F-J).

To disambiguate spatial and temporal features of transcriptional regulation, we used region-specific markers to computationally divide the cranial neural plate into four populations: the developing forebrain (*Otx2, Six3*) (6,060 cells), the developing midbrain and midbrain-hindbrain boundary (MHB, also known as rhombomere 1) (*En1, Fgf8, Pax2, Pax5, Pax8, Wnt1*) (5,306

cells), presumptive rhombomeres 2-5 (*Egr2*, *Gbx1, Gbx2, Hoxa2, Hoxb2, Hoxb3*), referred to as the hindbrain (5,568 cells), and the ventral midline (*Foxa2, Nkx6-1, Ptch1*, *Shh*), also known as the floor plate (761 cells). We then applied diffusion component analysis to identify gene expression trends in each population (Materials and methods). The top diffusion component (DC0) in each region correlated with embryo stage, suggesting that DC0 can be used to analyze gene expression over time (Figure 2A-C, Figure 2—figure supplement 1A-F). Consistent with the use of DC0 to represent developmental time, E-cadherin (*Cdh1*) transcript levels decreased and N-cadherin (*Cdh2*) transcript levels increased relative to DC0 in each region (Figure 2D-F), correlating with stage-specific changes in the corresponding proteins (Figure 2G). These results show that DC0 captures the known temporal shift in cadherin expression in the cranial neural plate (Kimura et al., 1995; Radice et al., 1997; Lee et al., 2007) and demonstrate that this transition is associated with a global change in transcriptional state.

As DC0 correlates with developmental stage, we investigated gene expression dynamics in the cranial neural plate by aligning the expression of each gene in the forebrain, midbrain, and hindbrain relative to DC0 for each region and clustering genes based on their expression along this axis (Figure 2—figure supplement 1G-I, Table S4) (Materials and methods). A number of genes were downregulated relative to DC0 throughout the cranial neural plate, including pluripotency factors (*Oct4/Pou5f1*), epithelial identity genes *(Cdh1, Epcam*), mitochondrial regulators (*Chchd10, Eci1*), epigenetic factors (*Dnmt3b*), and the Alzheimer's-associated gene *Apoe* (Figure 2H, Figure 2—figure supplement 2A). In addition, several genes were upregulated relative to DC0 throughout the cranial neural plate, including neurogenic factors (*Ndn, Pou3f2, Sox21*), cell division regulators (*Mycl, Mycn*), and genes involved in neural tube closure (*Cdh2*, *Palld*) (Figure 2H, Figure 2—figure supplement 2B). These results indicate that cells in the E7.5-E9.0 cranial neural plate transition from a pluripotent to a pro-neurogenic and morphogenetically active state.

In addition to changes in gene expression throughout the cranial neural plate, several genes were specifically upregulated in the forebrain, midbrain, hindbrain, or in multiple brain regions (Figure 2—figure supplement 2C, Figure 2—figure supplement 3, Table S4). These include genes encoding specific transcriptional regulators in the forebrain (*Arx, Barhl2, Emx2, Foxd1, Foxg1, Hopx, Lhx2, Six6*), midbrain (*En2*, *Pax8*), and hindbrain (*Egr2, Hoxa2, Msx3*), as well as transcriptional regulators expressed in multiple brain regions (*Irx2, Otx1, Pax6*) (Figure 2H, Figure 2—figure supplement 2C, Figure 2—figure supplement 3). In addition, several genes involved in cell-cell communication were upregulated in a region-specific fashion, including *Dkk1, Scube1, Scube2,* and *Wnt8b* in the forebrain, *Cldn10, Fgf17*, and *Spry1* in the midbrain, and *Fgf3*, *Plxna2*, *Robo2,* and *Sema4f* in the hindbrain (Figure 2—figure supplement 3). These results are consistent with the emergence of region-specific cell identities along the anterior-posterior axis of the E7.5-E9.0 cranial neural plate (Liu and Joyner, 2001a; Wurst and Bally-Cuif, 2001; Lohoff et al., 2022), and provide an opportunity to systematically define the transcriptional basis of these developmental patterning events.

Modular organization of gene expression along the anterior-posterior axis of the cranial neural plate

The emergence of distinct gene expression signatures in the presumptive forebrain, midbrain, and hindbrain suggests that this dataset can be used to identify factors that contribute to distinct cell fates and behaviors along the anterior-posterior axis. To characterize the spatial organization of transcription in the cranial neural plate, we pooled cranial neural plate cells from E8.5 (7-9 somites) and E8.75-9.0 (10+ somites) embryos, when strong regional patterns of gene expression are apparent, and we reapplied diffusion component analysis on the aggregated data (Materials and methods). The top diffusion component in the pooled E8.5-9.0 dataset (DC0) reproduced the anterior-posterior order of known markers of the forebrain (*Six3, Otx1, Otx2*), midbrain (*En1, Wnt1*), and hindbrain (*Gbx2*), indicating that this diffusion component correctly orders cells relative to the anterior-posterior axis (Figure 3A-C, Figure 3—figure supplement 1, Table S5). In addition, DC0 recapitulated the anterior-posterior order of markers of more refined transcriptional domains, such as the future telencephalon (*Foxg1, Lxh2*), diencephalon (*Barhl2*), midbrain-hindbrain boundary (*Fgf8*), and rhombomeres posterior to rhombomere 1 (*Hoxb1, Hoxb2, Egr2*) (Figure 3B and C). The second diffusion component in the pooled E8.5-9.0 dataset (DC1) also displayed a strong anterior-posterior signature, with DC1 separating genes expressed in the midbrain from genes expressed in the forebrain and hindbrain (Figure 3—figure supplement 1, Table S5). These results indicate that DC0 and DC1 correlate with anterior-posterior pattern, with DC0 ordering cells along the anterior-posterior axis.

To identify genes with spatially regulated transcriptional signatures along the anterior-posterior axis, we aligned all genes along DC0 and clustered them based on their dynamics along this axis (Materials and methods). Further, we used HotSpot (DeTomaso et al., 2019), a method to identify genes whose expression reflects statistically significant local similarity among cells (including spatial proximity), to distinguish genes with significant expression along DC0. This analysis identified 483 genes in 11 clusters that are predicted to be patterned relative to the anterior-posterior axis of the cranial neural plate (Figure 3D, Figure 3—figure supplement 2, Table S6). By comparing with markers for specific anterior-posterior domains (Figure 3C), we assigned these clusters to the future forebrain (clusters 0 and 7), midbrain (clusters 3 and 5), and hindbrain (clusters 4, 6, and 10), with some clusters spanning multiple domains (clusters 1, 2, 8, and 9) (Figure 3D). We note that although the midbrain emerged as a distinct region in our analysis, a subset of midbrain cells displayed gene expression signatures that overlap with the forebrain (cluster 8) and others displayed expression signatures that overlap with the hindbrain (clusters 1 and 9), likely corresponding to anterior and posterior domains within this structure. Based on the expression of *Fgf8* and *Gbx2*, which define the midbrain-hindbrain boundary (MHB), we refer to the midbrain population as midbrain/MHB. Together, these results provide a comprehensive molecular description of the modular organization of gene expression in the cranial neural plate and reveal that hundreds of genes are expressed in a small number of discrete domains along the anterior-posterior axis, corresponding to the metameric organization of the forebrain, midbrain, and hindbrain.

To assess the accuracy of the predicted expression patterns, we compared the predictions with published gene expression data in the Mouse Genome Informatics Gene Expression Database (http://informatics.jax.org) (Baldarelli et al., 2021). Of 483 genes predicted to be patterned along the anterior-posterior axis, 162 had published images of *in situ* hybridization experiments or protein immunostaining in the E8.5-9.0 cranial region (Theiler stages 12-14) (Theiler, 1989) that could be compared with the patterns predicted by our scRNA-seq analysis. We found that 141 of the 162 genes for which data were available (87%) were expressed in region-specific domains along the anterior-posterior axis that were consistent with the patterns predicted by our analysis (Table S7, Materials and methods). The remaining 21 genes were expressed in the predicted domain(s) in addition to one or more regions not predicted by our analysis, likely due to the increased sensitivity of *in situ* hybridization compared with scRNA-seq (Lohoff et al., 2022). Therefore, ordering gene expression relative to DC0 recapitulated the expression patterns of 141 known genes and predicted patterned expression along the anterior-posterior axis for an additional 321 genes whose expression in the developing cranial neural plate had not previously been characterized.

To identify region-specific transcriptional regulators that could contribute to these spatial patterns, we compared the list of genes with high information content relative to DC0 with transcriptional regulators predicted by the KEGG BRITE and Gene Ontology databases (Ashburner et al., 2000; Gene Ontology Consortium et al., 2023; Kanehisa et al., 2023) (Materials and methods). This approach identified 123 genes encoding transcription factors that are predicted to be spatially regulated along the anterior-posterior axis of the E8.5-9.0 cranial neural plate (54 shown in Figure 3E-G; see Table S8 for full list). These include *Otx2*, a homeodomain transcription factor that is required to form the forebrain and midbrain (Acampora et al., 1995; Matsuo et al., 1995; Ang et al., 1996), as well as genes required for specific aspects of development in the forebrain (*Foxg1, Hesx1, Rax, Six3)* (Xuan et al., 1995; Mathers et al., 1997; Dattani et al., 1998; Lagutin et al., 2003), midbrain (*En1*, *En2)* (Joyner et al., 1991; Wurst et al., 1994; Liu and Joyner, 2001b), and midbrain-hindbrain boundary (*Fgf8, Gbx2, Pax2, Pax5)* (Crossley et al., 1996; Favor et al., 1996; Torres et al., 1996; Wassarman et al., 1997; Millet et al., 1999; Schwarz et al., 1999; Li et al., 2002; Chi et al., 2003), Together, these results indicate that this dataset links known regulators of cell fate and behavior to defined anterior-posterior expression domains in the cranial neural plate and predicts spatially regulated expression for hundreds of previously uncharacterized genes.

Distinct patterns of gene expression along the mediolateral axis of the developing midbrain

In addition to patterning along the anterior-posterior axis, cell fate determination along the mediolateral axis, which becomes the dorsal-ventral axis, is essential for cranial neural tube closure, neuronal differentiation, and neural circuit formation (Shimamura et al., 1997; Agarwala et al., 2001; Hoch et al., 2009). However, although region-specific gene expression along the mediolateral axis has been systematically characterized for neuronal populations in the spinal cord (Delile et al., 2019; Sampath Kumar et al., 2023), the transcriptional cascades that govern

mediolateral patterning in the cranial neural plate are not well understood. To elucidate the molecular basis of this mediolateral patterning, we examined whether other diffusion components correlate with mediolaterally patterned gene expression in the E8.5-9.0 cranial neural plate. In contrast to DC0 and DC1, which correlate with markers of anterior-posterior pattern, DC2 was positively correlated with laterally expressed genes (*Pax3, Pax7, Tfap2a, Zic1*) and negatively correlated with markers of the ventral midline (floor plate) (*Foxa2, Nkx6-1, Ptch1, Shh*) (Figure 4A-C, Table S5). These results indicate that DC2 can be used to order genes with respect to the mediolateral axis.

To identify genes with mediolaterally patterned expression in the cranial neural plate, we aligned all genes expressed in the E8.5-9.0 cranial neural plate along DC2 and clustered them based on their expression along the axis (Materials and methods). This analysis identified 253 genes in 7 clusters that display high information content along DC2 and are predicted to be mediolaterally patterned (Figure 4D, Figure 4—figure supplement 1A and B, Table S6). These include four clusters (0, 4, 5, and 6) that contain genes predicted to have increased expression in the midline and two clusters (1 and 2) that contain genes predicted to be expressed in broader lateral domains. As genes in clusters 1 and 2 tended to be expressed in the midbrain/MHB region based on the UMAP representations, this suggested that DC2 primarily captures the mediolateral pattern in this region. We therefore excluded cluster 3, which contains genes expressed largely outside of this region, from further analysis and focused our analysis of DC2 on midbrain/MHB and midline cell populations.

To validate the predicted expression patterns, we compared gene expression along DC2 with expression data in the Mouse Genome Informatics Gene Expression Database. Of 222 genes predicted to be patterned along DC2 (excluding cluster 3), 65 genes had published expression data at these stages and 56 (86%) were consistent with the patterns predicted by our analysis (Table S7, Materials and methods). Thus, ordering gene expression relative to DC2 reproduced the expression of 56 known genes and predicted mediolaterally patterned expression for an additional 157 genes whose expression had not previously been characterized in the cranial neural plate.

Comparison of genes predicted to be patterned relative to DC2 with known transcriptional regulators, as we did for DC0, identified 78 transcriptional regulators that are predicted to be patterned along the mediolateral axis of the developing midbrain (38 shown in Figure 4E and F; see Table S8 for full list). These include genes that encode medially expressed effectors of SHH signaling (*Foxa2*, *Nkx6-1*, *Nkx2-2*, and *Nkx2-9*) (Sasaki et al., 1997; Briscoe et al., 1999; Sander et al., 2000; Dessaud et al., 2008; Ribes and Briscoe, 2009; Nishi et al., 2015). In addition, several genes in predicted laterally expressed clusters have been shown to promote dorsal cell fates (*Msx1*, *Msx*2) (Bach et al., 2003), suppress neuronal differentiation (*Hes1, Hes3*) (Hirata et al., 2001; Hatakeyama et al., 2004), and regulate cranial neural tube closure (*Gli3, Sp8, Zic2, Zic5)* (Nagai et al., 2000; Bell et al., 2003; Treichel et al., 2003; Inoue et al., 2004). A subset of genes were expressed in continuously increasing or decreasing patterns relative to DC2, reminiscent of gradients (Figure 4—figure supplement 1C-F). Together, these results indicate that this dataset

can be used to predict spatially regulated gene expression patterns along the mediolateral axis of the developing midbrain.

An integrated framework for analyzing cell identity in multiscale space

Patterning and morphogenesis of the cranial neural plate require the integration of positional information along the anterior-posterior and mediolateral axes. Nearly one-third of genes that were clustered relative to DC0, and nearly two-thirds of genes that were clustered relative to DC2, were also patterned along the orthogonal axis (Table S6), suggesting a high degree of overlap between anterior-posterior and mediolateral patterning systems. We therefore sought to develop a comprehensive two-dimensional framework to capture both anterior-posterior and mediolateral information. To analyze the full cartesian landscape of gene expression, we extended our approach to all high-eigenvalue diffusion components in the E8.5-9.0 cranial neural plate (Figure 3—figure supplement 1), after confirming that each correlated with subsets of known spatial markers in the neural plate (Materials and methods). This analysis identified 870 genes with high mutual information content with respect to anterior-posterior and mediolaterally correlated diffusion components (Figure 5A and B). Moreover, because we used the full diffusion space, the resulting clusters integrate more than one cartesian axis at once and capture more fine-grained spatial patterns. These include the genes identified in our analyses of DC0 (anterior-posterior pattern in Figure 3) and DC2 (mediolateral pattern in Figure 4), as well as additional genes that were not identified from analysis of a single diffusion component (Figure 5B, Table S6), suggesting these expression patterns reflect inputs from a combination of anterior-posterior and mediolateral spatial cues.

To capture both coarse- and fine-grained aspects of gene expression patterns in two dimensions, we clustered genes at different levels of stringency in their pairwise correlations (Materials and methods), referred to here as distance cutoffs (Figure 5C, Table S6). At the largest distance cutoff (lowest clustering stringency) of D=10, 870 genes were assigned to 15 clusters that are predicted to share broad features of anterior-posterior and mediolateral identity (Figure 5—figure supplement 1, Table S6). At an intermediate distance cutoff (D=6), individual clusters corresponded to more refined spatial domains, such as the early telencephalon (*Foxg1, Lhx2*), diencephalon (*Barhl2, Wnt8b*), and midbrain-hindbrain boundary (*Cldn10, Fgf8, Fgf17, Spry2*) (Figure 5D, Table S6). In addition, this distance cutoff separated midline transcripts enriched in the midbrain (*Foxa1, Foxa2*) from midline transcripts enriched in the forebrain (*Gsc, Nkx2-1, Nkx2-4*), and separated lateral transcripts enriched in the midbrain (*Pax5, Pax8*) from lateral transcripts that were more broadly expressed along the anterior-posterior axis (*Msx1, Msx2, Pax7, Zic1, Zic5*) (cluster 3) (Figure 5D, Table S6). By contrast, a smaller distance cutoff (higher clustering stringency) of (D=4) was needed to distinguish transcripts specific to rhombomeres 3 and 5 (*Egr2/Krox20*) or rhombomere 4 (*Hoxb1*) from general rhombomere markers (*Hoxa2, Hoxb2*) (Figure 5D, Table S6). Therefore, clustering genes at varying degrees of stringency provided an effective strategy to distinguish genes with different two-dimensional expression patterns.

Spatially regulated expression of genes involved in cell-cell communication in the cranial neural plate

Tissue morphogenesis and cranial neural tube closure require short- and long-range communication between cells mediated by secreted and transmembrane proteins (Harris and Juriloff, 2007, 2010). For example, cranial neural tube patterning and closure require short-range interactions mediated by Cdh2/N-cadherin (Radice et al., 1997) as well as long-range signaling by SHH, BMP, and WNT (Liu and Joyner, 2001a; Wurst and Bally-Cuif, 2001; Murdoch and Copp, 2010; Brooks et al., 2020; Kicheva and Briscoe, 2023). However, the full range of surface-associated and secreted signals that mediate communication between cells in the cranial neural plate has not been systematically characterized. To elucidate mechanisms of spatially regulated cell-cell communication in the cranial neural plate, we applied the mutual information method to identify cell-surface and secreted proteins that are predicted to be patterned relative to the anterior-posterior or mediolateral axes. We used the diffusion component framework to compare genes that display high information content along DC0 or DC2 with ligands and receptors in the Cellinker (Zhang et al., 2021) and CellTalkDB (Shao et al., 2020) databases and other secreted and transmembrane factors identified by sequence homology (Materials and methods). Using this approach, we identified 194 secreted or transmembrane proteins predicted to be patterned along the anterior-posterior or mediolateral axes or expressed in more complex patterns in the cranial neural plate (104 shown in Figure 6A-C, see Table S9 for full list).

This analysis revealed genes encoding spatially localized ligands and receptors known to have patterning or morphogenetic functions, such as *Shh* and its inhibitory receptor *Ptch1*, *Wnt* family ligands and their corresponding *Frizzled* receptors, *Fgf* ligands and receptors, and *Ephrin* ligands and their *Eph* receptors (Figure 6D-F, Figure 6—figure supplement 1A-E). This analysis also predicted patterned expression of two planar polarity genes required for cranial neural tube closure, *Celsr1* in the midbrain and hindbrain and *Vangl1* in the ventral midline (Figure 6A and B) (Curtin et al., 2003; Torban et al., 2008). This approach can also be used to predict the expression domains of other classes of genes, such as transcriptional regulators in the *Fox, Hes, Irx, Msx, Pax*, and *Zic* families (Figure 6—figure supplement 1F-K). Intersecting molecularly defined gene families with the spatial map predicted by scRNA sequencing may be generally useful for the targeted investigation of gene families that share sequence features but display a complex spatial relationship.

SHH signaling promotes distinct transcriptional programs in different cranial regions

Cell identity in the cranial neural plate is spatially and temporally regulated by conserved morphogens and growth factor signaling pathways, but how these signals influence the transcriptional programs that promote cranial neural tube patterning and morphogenesis has not been systematically examined. SHH is an essential regulator of mediolateral and dorsal-ventral patterning in the developing brain and spinal cord. In the spinal cord, loss of SHH signaling reduces or eliminates specific medial cell types, whereas excessive SHH signaling results in an

expansion of medial populations at the expense of lateral cell fates (Ingham and McMahon, 2001; McMahon et al., 2003; Dessaud et al., 2008; Ribes and Briscoe, 2009; Goetz and Anderson, 2010; Kicheva and Briscoe, 2023). In cranial tissues, excessive SHH signaling interferes with the patterned cell behaviors required for cranial neural tube closure (Murdoch and Copp, 2010; Brooks et al., 2020), and regulated SHH is important for determining spatially delimited neural fates (Fuccillo et al., 2004; Blaess et al., 2006; Balordi and Fishell, 2007; Xu et al., 2010; Tang et al., 2013). However, the transcriptional programs activated by ectopic SHH signaling in this region have not been systematically examined.

To characterize the gene expression changes that occur in response to excess SHH signaling in the cranial neural plate, we performed scRNA-seq analysis on embryos cultured with the small molecule Smoothened agonist SAG, which promotes ligand-independent activation of the SHH receptor Smoothened (Chen et al., 2002). Embryos cultured with SAG for 12 h between E8.0 and E8.5 displayed a lateral expansion of the expression of the known SHH target genes, *Foxa2* and *Nkx6-1* (Figure 7A), and disruption of neural fold elevation (Brooks et al., 2020), consistent with an increase in SHH pathway activity. Transcriptional profiles were obtained for 4,024 and 4,140 cells from dissected cranial regions of control and SAG-treated embryos, respectively, with a median library depth of >16,000 transcripts (UMIs)/cell and >4,300 genes/cell (average 1.57 reads/UMI) (Figure 7B, Table S1). Of these, we obtained 1,619 neural plate cells from control embryos and 1,401 neural plate cells from SAG-treated embryos. These cells were assigned to the forebrain, midbrain/MHB, or hindbrain using known markers and differentially expressed genes were identified in each region (Figure 7C-J, Table S10). Differentially expressed genes were identified by MAST analysis (Finak et al., 2015), which accounts for sample variation in scRNA-seq data when comparing log-transformed data (Materials and methods). Using a false discovery rate-adjusted p-value of <0.001, and a MAST hurdle transformation of the $\log_2$ fold-change value of >0.24 or <-0.24, we identified 166 genes that were significantly upregulated and 199 genes that were significantly downregulated in at least one region of the cranial neural plate in SAG-treated embryos (Table S11).

Several known SHH pathway genes were deregulated in SAG-treated embryos (highlighted in red in Figure 7D-J), consistent with an expansion of SHH signaling. The known SHH target gene *Ptch1* (Goodrich et al., 1996) was one of the top two upregulated genes in all regions of the cranial neural plate in SAG-treated embryos (Table S10, Table S11). Further, the known SHH targets *Gli1* (Lee et al., 1997) and *Foxa2* (Sasaki et al., 1997) were significantly upregulated but did not meet our MAST hurdle threshold, indicating the presence of additional SHH targets beyond our cutoffs. In addition to upregulated genes, several factors that are negatively regulated by SHH in other contexts were transcriptionally downregulated in the cranial neural plate of SAG-treated embryos, including the SHH co-receptors *Cdon* and *Gas1* and the transcriptional repressor *Gli3* (Marigo et al., 1996; Wang et al., 2000; Okada et al., 2006; Tenzen et al., 2006; Yao et al., 2006; Zhang et al., 2006; Allen et al., 2007; Goetz and Anderson, 2010). We also observed a previously unreported inhibition of *Scube2*, which encodes a cell surface protein that modulates SHH secretion (Woods and Talbot, 2005; Creanga et al., 2012). Therefore, our dataset captures known components of SHH signaling and identify other candidate transcriptional feedback mechanisms that may modulate SHH signaling in the cranial neural plate.

Nearly half of the transcripts that were differentially regulated in the cranial neural plate of SAG-treated embryos were identified as spatially regulated in our analysis of wild-type embryos, consistent with a strong effect of SHH on spatial patterning (Table S6, Table S10, Table S11). Notably, several transcripts were upregulated in a region-specific fashion in SAG-treated embryos, including the transcriptional regulators *Foxd1, Hesx1*, and *Maf* in the forebrain (Figure 7E), *Foxb1* in the forebrain and midbrain/MHB (Figure 7H), *Pou3f1* in the forebrain and hindbrain, and *Nkx6-1* and *Nkx6-2* in the midbrain/MHB and hindbrain (Figure 7J). In addition, SAG treatment led to the upregulation of several genes associated with medial clusters in wild type (highlighted in purple in Figure 7D-J) and the downregulation of several genes associated with lateral clusters (highlighted in blue). Downregulated genes included lateral markers expressed throughout the cranial neural plate (*Msx1, Pax3, Pax7, Zic2, and Zic5)* or in specific regions, such as *Pax8* in the midbrain/MHB, *Msx3* in the hindbrain, and *Zic1* in both domains. These results indicate that increased SHH signaling enhances medial cell fates and suppresses lateral cell fates in the cranial neural plate, similar to the effects of SHH signaling in the spinal cord (Dessaud et al., 2008; Ribes and Briscoe, 2009; Goetz and Anderson, 2010). Notably, a number of genes that were modulated by SAG treatment were associated with anterior-posterior or spatially patterned clusters in the cranial neural plate (highlighted in green and orange in Figure 7D-J). The effects of activated SHH signaling are consistent with the results of our spatial analysis showing that many genes are responsive to both anterior-posterior and mediolateral inputs (Figure 5, Table S6), and suggest a connection between anterior-posterior and mediolateral patterning systems.

The activation of SHH signaling by SAG also modified the expression of genes involved in other biological processes. For example, we observed changes in the expression of genes involved in retinoic acid signaling (*Crabp2* and *Cyp26b1*) and components of the WNT signaling pathway, which has been shown to function antagonistically to SHH in several contexts (Dessaud et al., 2008; Ribes and Briscoe, 2009; Ulloa and Martí, 2010; Borday et al., 2012). In particular, several genes encoding WNT ligands were downregulated in the forebrain (*Wnt7b* and *Wnt8b*) or midbrain/MHB (*Wnt1* and *Wnt5a*) of SAG-treated embryos (Figure 7E and F), and two components required for WNT signaling, *Axin2* and *Fzd10*, were downregulated in the midbrain/MHB (Figure 7F). In addition, the WNT inhibitor *Sfrp1* was upregulated in the hindbrain (Figure 7G). These results raise the possibility that SHH could antagonize WNT signaling both by inhibiting the expression of essential pathway components and enhancing the expression of a pathway repressor. Finally, SAG-treated embryos displayed a downregulation of several genes associated with neuronal function, including transcripts associated with neurodevelopmental disorders (*Irf2bpl)* (Marcogliese et al., 2018), epilepsy and autism spectrum disorders (*Cntnap2*) (Peñagarikano et al., 2011), epilepsy and ataxia (*Kcna1)* (Smart et al., 1998) and GABAergic neuron development and function (*Tal2*) (Bucher et al., 2000; Achim et al., 2013). These results provide clues to the molecular basis of the antagonism between SHH and WNT signaling, and suggest new potential connections between SHH signaling and human disease.

Discussion

In this study, we describe a spatially and temporally resolved dataset of single-cell gene expression profiles in the cranial neural plate of the mouse embryo during cranial neural tube patterning and closure. Analysis of gene expression in mouse cranial neuroepithelial cells revealed dynamic changes in gene expression over a 1.5-day window of development that captured the emergence of distinct transcriptional signatures in the future forebrain, midbrain, hindbrain, and midline, allowing the imputation of developmental trajectories in specific spatial domains. In addition, we used this single-cell dataset to reconstruct spatial patterns of gene expression during the final stages of neural tube closure, with computational ordering based on diffusion components providing a good approximation of the anterior-posterior and mediolateral axes. The spatial patterns predicted by our analysis displayed >85% agreement with previously reported gene expression studies, indicating that this approach can be used to predict the expression of previously unexplored genes. By extending this analysis to multiple dimensions, we identified refined transcriptional domains that integrate inputs from both anterior-posterior and mediolateral patterning systems. This dataset provides a comprehensive spatiotemporal atlas of transcriptional cell states during a critical window of development, revealing the complex spatial and temporal regulation of gene expression that accompanies the morphogenesis and patterning of the cranial neural plate during neural tube closure.

The present study expands on previous scRNA-seq datasets that capture broad spatial identities (Pijuan-Sala et al., 2019; La Manno et al., 2021; Zalc et al., 2021) or elements of anterior-posterior and mediolateral organization (Lohoff et al., 2022) in the mouse cranial neural plate. Because cell states are strongly dependent on the location of cells relative to signaling centers or other tissue features, the high density of spatial sampling provided by our dataset made it possible to define a diffusion component spanning the entire anterior-posterior axis of the cranial neural plate, as well as a diffusion component that recapitulates the mediolateral organization of midbrain cells. Other diffusion components identified in our analysis may be useful for exploring different aspects of spatial organization, such as mediolateral patterning in the forebrain (DC6, DC7), rhombomere-specific transcriptional programs (DC3, DC4, DC5), and gene expression signatures shared by spatially separated domains (DC8, DC9). Diffusion-based approaches have been used to define spatial and temporal cell orderings in other tissues, including the mouse embryonic endoderm (Nowotschin et al., 2019) and hematopoietic lineage (Setty et al., 2019). By combining diffusion component analysis with methods to interrogate spatial autocorrelation (DeTomaso et al., 2019; DeTomaso and Yosef, 2021), we were also able to identify genes expressed in more complex, two-dimensional patterns, an approach that may prove useful in other contexts. Although spatial transcriptomics is important for interrogating gene expression patterns in complex and heterogeneous samples such as tumors (Longo et al., 2021; Rao et al., 2021; Moffitt et al., 2022; Moses and Pachter, 2022), these observations suggest that for well-ordered tissues, increasing the sampling density can provide insight into the spatial and temporal dynamics of cell state and gene expression changes without the need for explicitly spatial approaches.

We took advantage of the ability to detect spatially restricted gene expression in the wild-type cranial neural plate to investigate how region-specific patterns change in the presence of activated

SHH signaling, which disrupts neural tube patterning and closure (Dessaud et al., 2008; Goetz and Anderson, 2010; Murdoch and Copp, 2010). Consistent with a critical role for SHH in mediolateral patterning, ectopic SHH signaling led to a significant increase in the expression of several components of the SHH signaling pathway, as well as an upregulation of medially expressed genes and a downregulation of genes encoding lateral cell fates in the cranial neural plate. This analysis also provides insight into the interaction between SHH and WNT signaling pathways, which often function antagonistically in neural patterning (Ulloa and Martí, 2010; Borday et al., 2012), and identify transcriptional changes that could underlie this antagonism. In particular, our data suggest that crosstalk between SHH and WNT pathways could involve transcriptional repression of WNT ligands and effectors and transcriptional activation of the pathway inhibitory factor *Sfrp1,* perhaps through a GLI-binding site in the *Sfrp1* promoter (Katoh and Katoh, 2006). Bulk RNA sequencing of whole E9.5 embryos reveals that embryos lacking the ciliary G protein-coupled receptor Gpr161, which display expanded SHH signaling (Mukhopadhyay et al., 2013), have reduced expression of WNT pathway genes, including several that overlap with our dataset (Kim et al., 2019). Together, these results raise the possibility that SHH activity could broadly antagonize WNT signaling at the transcriptional level.

Gene expression changed significantly along both mediolateral and anterior-posterior axes in SAG-treated embryos, raising the question of how SHH simultaneously influences both mediolateral and anterior-posterior patterning in the cranial neural plate. In one model, SHH activation could disrupt anterior-posterior patterning in part by inhibiting WNT and retinoic acid signaling, which are required to establish posterior neural fates (Yamaguchi, 2001; Maden, 2007). Alternatively, expanded SHH signaling may override anterior-posterior gene expression programs in favor of a mutually exclusive medial cell fate. Finally, SHH signaling could activate distinct gene expression programs in the forebrain, midbrain, and hindbrain, allowing a single morphogen to generate region-specific cell fates and tissue structures. Several mechanisms have been proposed to influence the targets of SHH signaling, including the presence of intermediary or accessory factors such as FOXC1 and FOXC2 in pharyngeal tissues (Yamagishi et al., 2003), SOX2 in the spinal cord (Peterson et al., 2012), and HAND2 in mandibular development (Elliott et al., 2020). In other cases, the SHH signaling response is modulated by the receptivity of cells to SHH, such as RFX4-dependent primary cilia formation in the central nervous system (Ashique et al., 2009) or the chromatin accessibility of SHH target sites in the limb bud (Lex et al., 2020). The activation of tissue-specific gene expression programs by SHH has also been observed in the chick coelomic epithelium (Arraf et al., 2020), raising the possibility that interaction with orthogonal patterning systems may be a general mechanism that shapes the outcome of SHH signaling.

The systematic profiling of gene expression in the cranial neural plate provides a resource for defining the transcriptional changes that underlie the acquisition of region-specific cell fates in the forebrain, midbrain, and hindbrain. These datasets can provide insight into cell differentiation mechanisms *in vivo* and inform efforts to induce the differentiation of specific cranial neuronal and neuron-adjacent cell populations from pluripotent stem cells *in vitro* (Tabar and Studer, 2014; Medina-Cano et al., 2022). Although the present analysis focused on the cranial neural plate, this dataset will be useful for analyzing dynamic changes in gene expression in several other cranial

cell populations that exhibit significant changes in transcriptional state during the developmental window analyzed. For example, this dataset includes nearly 14,000 cranial mesoderm cells that could help to shed light on the critical roles of the mesoderm tissue during cranial development (Chen and Behringer, 1995; Camus et al., 2000; Zohn et al., 2007; Zohn and Sarkar, 2012; Bildsoe et al., 2013). This dataset also adds earlier time points to previous studies of the cranial neural crest (Zalc et al., 2021; Williams et al., 2022), and can provide insight into development of the non-neural ectoderm, which contributes essential signaling molecules and forces required for cranial neural tube closure (Dickinson et al., 1995; Shimamura and Rubenstein, 1997; Molè et al., 2020; Maniou et al., 2021; Christodoulou and Skourides, 2022). Further analysis of gene expression during this critical stage of development will provide insight into the spatiotemporal delineation of diverse cell populations during neural tube closure and will provide a framework for systematically elucidating the transcriptional changes that contribute to the morphogenesis and patterning of the mammalian brain.

Materials and methods

Mouse handling and use

Mice used in this study were 8-12-week-old inbred FVB/NJ mice purchased from the Jackson Laboratory (strain #001800). Timed pregnant dams were euthanized by carbon dioxide inhalation followed by secondary cervical dislocation. All mice were bred and housed in accordance with PHS guidelines and the NIH Guide for the Care and Use of Laboratory Animals and an approved Institutional Animal Care and Use Committee protocol (15-08-13) of Memorial Sloan Kettering Cancer Center.

Embryo collection and cell dissociation

FVB embryos were dissected from the uterine horn on ice in pre-chilled DMEM/F-12 with Glutamax (ThermoFisher), hereafter referred to as DMEM, including resection of the yolk sac and amnion. Embryo stages were assigned as follows: 0 somites (E7.5), 1-2 somites (E7.75), 3 somites (E8.0), 4-6 somites (E8.25), 7-9 somites (E8.5), 10+ somites (E8.75-E9.0). Cranial tissues from E7.5-E8.0 embryos were collected by dissecting at the posterior limit of the cranial neural plate. Cranial tissues from E8.25-E9.0 embryos were collected by dissecting embryos just posterior to the otic sulcus. Due to differences in embryo size, a larger region of the hindbrain was included at earlier stages, which may account for a subset of transcripts that were present in early but not late hindbrain samples. Visible cardiac structures, including the early heart tube, were manually resected. Cranial tissues from 3-5 embryos were pooled per replicate; 4-6 replicates were analyzed per stage in wild-type embryos and 3 SAG-treated replicates and 2 control replicates were analyzed for the SAG treatment experiment (Table S1). Individual cells were dissociated by transferring pooled cranial tissues into 2 mL of pre-chilled phosphate buffered saline (PBS) containing 2.5% pancreatin (Sigma) and 0.5% trypsin (Sigma) and incubated 5 min on ice. Tissues were then washed 1 min in cold DMEM with 10% calf serum (Gibco), followed by a 1 min wash in cold DMEM alone. Pooled tissues were then transferred to a clean watch glass containing 200 µL of a 1:2 mixture of accutase (Sigma) and 0.25% trypsin in PBS and incubated 15 min at 37ºC, gently swirled, and incubated for an additional 15 min. After incubation, tissues

were returned to ice and 600 µL of DMEM with 20% calf serum in DMEM was added to each watch glass. Tissues were manually triturated on ice using tungsten needles for 5-10 min, then passed over a 40 µm FlowMi cell strainer (Sigma) and concentrated by 450 g centrifugation at 4ºC. After centrifugation, the supernatant was removed and the cells were resuspended in 50 µL PBS with 0.4% bovine serum albumin (BSA). 10 µL of cells were combined with 10 µL of Trypan blue and placed in a hemocytometer to analyze the proportion of single cells and dead cells. Pools with less than 85% singlets or 85% viable cells were discarded. Isolated cells were then encapsulated and barcoded using standard protocols on the Chromium V3 chemistry platform from 10X Genomics. 3' RNA-seq libraries were subsequently generated following standard 10X Genomics protocols.

SAG treatment

For treatment with Smoothened Agonist (SAG) (Sigma 912545-86-9), FVB embryos were dissected at E8.0 for culture following standard protocols in a 50:50 mixture of pre-warmed and gassed (37ºC, 5% CO_2) DMEM and whole embryo culture rat serum (Inotiv). After dissection, embryos were randomly assigned to treatment with 2 µM SAG in DMSO or an equivalent volume (0.02% v/v) of DMSO alone as a control and cultured 12 h in pre-warmed and gassed (37ºC, 5% CO_2) DMEM in 24-well Lumox plates (Sarstedt). After the culture period, rare embryos with signs of excessive cell death or developmental failure were discarded; the frequency of these embryos did not differ between control and SAG-treated conditions. After final resection of the yolk sac and amnion, embryo dissection, cell dissociation, and library preparation were performed as described.

Single-cell transcriptome sequencing

For replicates 3-23 and 40-43 (Table S1), cell suspensions were loaded on a Chromium instrument (10X Genomics) following the user guide manual for 3′ v3. In brief, cells were washed once with PBS containing 1% bovine serum albumin (BSA) and resuspended in PBS containing 1% BSA to a final concentration of 700–1,300 cells/µL. The viability of cells was above 80%, as confirmed with 0.2% (w/v) Trypan Blue staining (Countess II). Cells were captured in droplets. Following reverse transcription and cell barcoding in droplets, emulsions were broken and cDNA purified using Dynabeads MyOne SILANE followed by PCR amplification following the manual instructions.

For replicates 44-55 (Table S1), single cell suspensions were stained with Trypan blue and Countess II Automated Cell Counter (ThermoFisher) was used to assess both cell number and viability. Following quality control, the single cell suspension was loaded onto Chromium Next GEM Chip G (10X Genomics PN 1000120) and GEM generation, cDNA synthesis, cDNA amplification, and library preparation of up to 5,000 cells proceeded using the Chromium Next GEM Single Cell 3' Kit v3.1 (10X Genomics PN 1000268) according to the manufacturer's protocol. cDNA amplification included 11 cycles and 17-265 ng of the material was used to prepare sequencing libraries with 8-14 cycles of PCR.

After PicoGreen quantification and quality control by Agilent TapeStation, indexed libraries were pooled equimolar and sequenced on a NovaSeq 6000 in a PE28/91 or PE28/90 run, using the NovaSeq 6000 SP, S1, S2, or S4 Reagent Kit (100 or 200 cycles) (Illumina).

FASTQ alignment

FASTQ files were preprocessed using the using the Sequence Quality Control (SEQC) bioinformatics pipeline (Azizi et al., 2018) aligning reads to the mm38 mouse reference genome with default parameters for the 10x single-cell 3-prime library. The SEQC pipeline performs read alignment, multi-mapping read resolution, as well as cell barcode and UMI correction to generate a cells x genes count matrix. The pipeline was run without performing default cell filtering steps (our strategy for cell filtering is described below). Some gene names have been updated in subsequent genome revisions, but for data interoperability we have preserved the original gene symbols.

Ambient RNA correction and cell filtering

Aligned, unfiltered count matrices of each sample were processed using CellBender (Fleming et al., 2023) to remove ambient RNA contamination and identify empty droplets. The expected number of real cells input to CellBender's *remove-background* function was determined by SEQC (which uses the optima in the second derivative of library size to identify real cells from droplets with ambient RNA expression) and the total droplets used to estimate ambient background RNA was set to 30,000. Training was run for 100 epochs. All droplets that were estimated by CellBender with nonzero probability to be real cells were kept for downstream analysis. From the raw count matrix corrected for ambient RNA contamination, cells that did not pass the following filters were removed: (1) Library size > 1,000 UMIs. (2) Library complexity: we use SEQC's method to remove cells with low ratio of UMIs vs. genes expressed, with a 0.1 cutoff on the residual for the linear model fit. This step primarily removes high library-size cells with one predominant gene program, e.g. mitochondrial gene-enriched stressed cells. (3) Percent mitochondrial RNA < 20%.

After initial filtering, each sample showed a clear bimodality separating cells of high library size and number of genes expressed and cells of low to moderate library size. In addition, none of the cells from replicate 48 passed the filtering criteria and replicate 54 was removed from the dataset due to low library size (<100 UMIs per cell), so these samples are not reported.

We sought to assess whether the resulting bimodal distribution of library size was biologically driven. For this, we first fit a kernel density estimation model to the library size distributions to systematically assign cells to the higher quality (library size and number of genes expressed) or lower quality mode. We then proceeded to compute differentially expressed genes (DEGs) between the two classes. In order to control for the differences in library size between the two groups, we downsampled both groups to 1,000 UMIs (the lowest number of UMIs in the low-quality group) and then identified DEGs calculated between the high and low-quality cells using their raw expression matrices (Wilcoxon rank-sum, Benjamini-Hochberg correction). We found that most of the genes differentially upregulated in the low library size group were associated with ribosomal activity or cell stress and cell death (e.g. GM42418, which was the most significantly

upregulated). This evidence suggested that this low-quality cluster of cells likely comprised cells that were under stress and perhaps leaking RNA, resulting in a bias to the capture and amplification of highly expressed genes.

In order to systematically remove the above identified low quality cells with lower library size, we finely clustered the cells using PhenoGraph with a value of *k=8* (Levine et al., 2015). Cells of each cluster were reassigned as low- or high-quality based on whether the library size and no. genes distributions of their cluster were more likely to come from the high-quality group described above. Because the library sizes were approximately normal for each group and cluster, a Z-test was performed and p-value cutoff of $1e^{-10}$ was used to decide from which group each cluster was more likely to come. The motivation for filtering by cluster, rather than threshold values, was to avoid relying on a single hard cutoff; instead, we retained cells based on phenotypic similarity, which is captured by cell clustering.

In total, 39,463 cells in the wild-type dataset passed quality control filtering, with a median library depth of >42,000 transcripts (UMIs)/cell and >5,900 genes/cell (average 1.56 reads/UMI). In addition, 8,164 cells in the control and SAG treatment dataset passed quality control filtering, with a median library depth of >16,000 UMIs/cell and >4,300 genes/cell (average 1.57 reads/UMI).

Normalization and preprocessing

All cells within each dataset were first normalized to median library size and the natural log of normalized expression with pseudocount 1 was computed for each cell. Cell cycle influence was removed by regression with the Python package, fscLVM, using several cell-cycle associated GO gene sets provided in Table S12. For each dataset, feature selection was performed using the “highly_variable_genes” function in scanpy with "flavor = seurat_v3” and “n_top_genes = 3000”. All mouse mitochondrial genes, ribosomal genes, and genes expressed in fewer than 10 cells were excluded from feature selection. In addition, genes corresponding to each cell type and spatial axis delineation were compiled from literature and retained. Principal component analysis was performed on the log-transformed normalized expression matrix; the number of principal components was selected to explain 75% of the variance in the dataset. Cells were clustered with PhenoGraph with the number of nearest neighbors set to 30, after confirming the Rand Index is robust for nearby values (RI > 0.8 for an absolute difference of 5 and 10 in *k* values).

Cell typing

To identify major cell types in these scRNA-seq datasets, first, we manually annotated each PhenoGraph cluster based on the expression of literature derived gene signatures and identification of cell type-specific DEGs. More specifically, the clusters were first annotated through the following steps: (1) We identified cell type signatures enriched in each cluster by computing the average expression of each signature (see ‘Computing Gene Set Scores’ below) per cluster. The list of signatures compiled from the literature is provided in Table S13. The expression of cell type signatures in the wild-type dataset is shown in Figure 1—figure supplement 1A. (2) We examined differentially expressed genes (DEGs) per cluster using the R package MAST on log_2-transformed normalized counts. The DEGs of each cluster in the full cranial dataset are listed in Table S2. This provided a preliminary cell type assignment for each cluster.

Because these datasets capture developmental states at which cell types are not fully resolved, unique cell types were not consistently separated by cluster, especially in samples from early time points. For this reason, we refined the cluster-based annotations by implementing a semi-supervised classification using the *classify* method in the PhenoGraph package. Briefly, PhenoGraph classification began with a training set of labeled cells (e.g. whose cell type is known) and a test set of unlabeled cells (e.g. whose cell type is unknown). An absorbing Markov chain was computed from the dataset, and the absorption probabilities for each cell type label were calculated for all unlabeled cells; for each cell type and unlabeled cell, the absorption probability represents the likelihood that a random walk originating from any cell of that cell type will reach the unlabeled cell. All cells were then labeled according to their maximum absorption probability.

Beginning from cluster-based cell type annotations, a labeled training set of cells was created for each cell type using the following method: First, threshold values equal to the 20th percentile cell type signature scores were calculated for all cells previously given a cell type label. Cells whose score exceeded the threshold for their own cell type and were below the thresholds for all other cell types were used as training examples. All other cells were unlabeled and re-annotated using PhenoGraph's *classify* method, with a value of *k=30*. Cell type compositions of each sample are provided in Table S1. After classification in this manner, we discovered one PhenoGraph cluster of cells (cluster 28) which did not express any neural plate cell type markers listed above. However, we found this cluster differentially expressed multiple hemoglobin genes including *Hba-a1*, *Hba-x*, *Hbb-bh1*, *Hbb-y* (Table S2), so we re-annotated this cluster as "Blood Cells". Additionally, manual analysis of cluster 12 revealed relatively high levels of *Foxa2* expression and this cluster was retyped as endoderm.

Finally, to identify major spatial domains of the neural plate, each dataset was first subset to all neural plate cells and processed as described above. The same cell typing procedure was performed on the neural-plate-only datasets using cell type signatures compiled from the literature for each spatial domain (forebrain, midbrain, hindbrain, and midline) and the same input parameters. Cell types in the cranial neural plate were identified using the markers described in Table S13 and the DEGs associated with each cluster are listed in Table S3.

Computing per-cell geneset scores

For each cell type, we computed a signed signature score, as described in (DeTomaso et al., 2019), which uses sets of genes known to be expressed in that cell type (positive markers) as well as genes which are known not to be expressed in that cell type (negative markers). The score was calculated as the mean difference in expression of positive and negative gene sets, z-normalized using the expected mean and variance of a random signature with the same number of positive/negative genes. Where no negative genes are provided, the geneset score was simply the z-normalized mean expression of genes compared to a random signature with the same number of genes. This method was used to score spatially patterned gene clusters.

UMAP visualization

Each dataset was embedded in two dimensions with UMAP (McInnes et al., 2018), which was run on a k-nearest neighbor graph (k = 30) produced from its principal components using the *umap* function in the Scanpy package. The UMAP initializations were based on partition-based graph abstraction implemented in the Scanpy package using PhenoGraph clusters. Calculation of principal components and PhenoGraph clusters is as described above. Gene expression was visualized using imputed values unless otherwise specified. Gene expression was imputed with MAGIC, using parameters k=5 and t=3 (van Dijk et al., 2018).

Temporal analysis

To identify temporal gene expression trends in the neural plate, we first separated our complete neural plate dataset—consisting of all embryo stages–into three subsets, each containing only cells from the forebrain, midbrain, and hindbrain, respectively. This was done in order to minimize the influence of anterior-posterior patterned changes in gene expression during analysis. For each subset, a diffusion map was computed using the implementation in Palantir (Setty et al., 2019) on the log-normalized count matrix (natural log with pseudo-count 1). A *k*=30 value was used to determine the nearest-neighbor graphs. We verified that the first diffusion component (DC0) in each region was associated with changes in somite stage (Figure 2A-C). We further observed changes in E-cadherin (*Cdh1*) and N-cadherin (*Cdh2*) expression that were consistent with this interpretation (Figure 2D-F).

For each spatial subset of the data, expression trends of temporally regulated genes were computed for DC0 and clustered using Palantir. Briefly, for each gene, a gene trend was calculated by de-noising gene expression using MAGIC imputation and fitting a generalized additive model to its values of expression for cells ordered along a provided axis of change (this step is identical to Palantir except that DC0 was the analogue of the “pseudotime” axis used in Palantir). Gene trends were clustered with a value of *k=20* to produce 5, 5, and 6 gene trend clusters in the forebrain, midbrain/MHB, and hindbrain subsets, respectively.

Diffusion component analysis

We first sought to establish that cell-cell similarity in spatial position is well-described by the top diffusion components of the developed neural plate dataset. A diffusion map was computed using the implementation in Palantir (Setty et al., 2019) on the natural log-normalized count matrix. A *k*=30 value was used to determine the nearest-neighbor graph. The number of diffusion components was chosen based on the eigengap, selecting the components preceding the largest eigengap found within the first 40 eigenvalues. For the E8.5-9.0 cranial neural plate dataset, this resulted in the selection of 10 diffusion components (Figure 3—figure supplement 1).

Each diffusion component was correlated (Pearson correlation) with all genes and the top 100 most correlated genes (p-value < 0.01) were identified (Table S5). The diffusion components were found to strongly correlate with genes known to be spatially patterned with the neural plate. Using highly correlated genes known from the literature, each diffusion component was annotated with the specific spatial patterns in the neural plate it represents (see below). As diffusion components 0 and 2 were strongly correlated with markers of the anterior-posterior and mediolateral axes, respectively (Figures 3 and 4), we concluded that components 0 and 2 of the diffusion map,

selected by eigengap, provide information associated with spatial patterning of the neural plate; for downstream analysis, we used diffusion components 0 and 2 as proxies for the anterior-posterior and mediolateral axes, respectively.

In addition to the analyses of DC0, DC1, and DC2 discussed in the main text, analysis of the top 25 correlated genes revealed distinct spatial patterns for DC3-DC9. Genes expressed in rhombomeres 5 and 6 were positively correlated with DC3 (*Hoxb3*, *Mafb*, *Hoxd3*, *Hoxa3*, *Egr2*); genes expressed in rhombomeres 3 and 5 were positively correlated with DC4 (*Eg2*, *Cyp26b1*, *Hoxb3*, *Hoxa2*, *Mafb*, *Hoxa3*); genes expressed in the midbrain-hindbrain boundary and in rhombomeres 3 and 4 were positively correlated with DC5 (*Hoxb1*, *Foxd3*, *Epha2*, *Egr2*, *Wnt8a*, *Hoxb2*, *Fgf8*, *Fgf17*, *Otx1*, *Crabbp2*, *Mafb*); genes expressed in the forebrain and midbrain were positively correlated with DC6 (*Pax6*, *Emx2*, *Pax2*, *Pax8*, *Barhl2*, *Pax5*, *Fgf8*, *Nkx2-1*, *Six6*, *Hesx1*, *En2*, *En1*); genes expressed in the forebrain and midline were positively correlated with DC7 (*Nkx2-1, Gsc, Fgf8, Foxd1, Nkx2-9, Fexf2, Foxg1, Sp8, Ptch1*); genes expressed in the forebrain and rhombomeres 2-5 were positively correlated with DC8 (*Six6, Foxd1, Dkk1, Vax1, Meis2, Emx2, Lhx5, Epha7, Dmbx1, Foxg1, Barhl2*); and genes expressed in lateral domains of all anterior-posterior regions were positively correlated with DC9 (*Tfap2c, Pax3, Msx2, Msx1, Tfap2a, Zic5, Pax7, Foxb1, Msx3, Zic2*).

Selection of spatially informative genes using Hotspot

Having established that our diffusion map contains information associated with the spatial positioning of cells in the neural plate, we next sought performed feature selection to isolate genes that exhibit patterns of expression in physical space (treating the entire diffusion space in this analysis as an approximation of spatial position). Spatially informative genes were selected using the Hotspot procedure, as described in (DeTomaso and Yosef, 2021). Briefly, Hotspot identifies spatially informative genes in two steps: (1) A similarity graph is computed between cells as a k-nearest neighbor graph in a user-defined space. (2) Feature selection is performed to identify spatially autocorrelated genes, which are genes whose patterns of expression are well-represented by the similarity graph. To do so, we defined a “local autocorrelation” test statistic evaluated on each gene, defined as the sums of weighted pairwise products of nearby cells in the similarity graph. The resulting values were transformed into *Z* scores, given a null model in which the expression of a gene by each cell is independent of similar “local” cells, for significance calculation which we then used to filter genes.

To construct a similarity graph between cells, an embedding was calculated on diffusion components on three latent spaces: (1) A latent space comprising only the anterior-posterior axis-correlated diffusion component (DC0). (2) A latent space comprising only the mediolateral axis-correlated diffusion component (DC2). (3) A multiscale space embedding was calculated on the top 10 diffusion components of the developed neural plate dataset.

These latent spaces relate to the sections “Modular organization of gene expression along the anterior-posterior axis of the cranial neural plate”, “Mediolaterally patterned gene expression in the developing midbrain”, and “An integrated framework for analyzing cell identity in multiscale space”, respectively, within the main text. The three analyses relating to these sections are also

elaborated in the two sections which follow. In each case, the Hotspot open-source Python package was used to create a k-nearest neighbor graph on the latent space with a value of *k=30*, and to compute local autocorrelations of all 19,623 genes detected in this scRNA-seq dataset with respect to the kNN graph. Spatially informative genes were selected according to the following criteria: (1) To eliminate lowly expressed genes, all genes were removed whose natural log-transformed range of expression was less than or equal to 1. (2) All genes were kept whose False Discovery Rate (FDR) values, calculated by Hotspot from p-values associated with genes' spatial autocorrelations, were below 10^{-5}, and whose Z-scored spatial autocorrelation statistic values were greater than or equal to 10. The latter value was chosen by knee point, calculated on all genes.

Gene trend clustering

We first sought to classify genes within this set based on their primary spatial axis of patterning (anterior-posterior *or* mediolateral), and cluster genes with similar patterns of expression along each axis individually. To identify genes with anterior-posterior axis-dependent patterns of expression, the Hotspot package was run with all genes, a value of *k=30*, and a latent space comprising only the anterior-posterior axis correlated diffusion component (diffusion component 1); 483 genes, having a Z-scored spatial autocorrelation values calculated on DC0 greater than 10 and FDR values below 10^{-5}, were considered patterned along the anterior-posterior axis.

Expression trends of anterior-posterior axis-patterned genes were computed for DC0 and clustered using Palantir (Setty et al., 2019) Briefly, for each gene, a gene trend was calculated by de-noising gene expression using MAGIC imputation and fitting a generalized additive model to its values of expression for cells ordered along a provided axis of change (this step is identical to Palantir except that DC0 was the analog of the "pseudotime" axis used in Palantir). Because known anterior-posterior axis patterns of gene expression are not strongly recapitulated in the midline cells of the neural plate, only non-midline cells were used to compute anterior-posterior axis gene trends. Gene trends were clustered with a value of *k=20* to produce 11 gene trend clusters. The value of k was selected as being sufficiently low to distinguish known patterns of gene expression along the anterior-posterior axis.

Mediolateral axis-dependent patterns of gene expression were identified similarly, using the mediolateral axis-correlated diffusion component (DC2), resulting in 253 mediolateral axis-patterned genes with Z-scored autocorrelation values greater than 10 and FDR values below 10^{-5}. Mediolateral axis gene trends were calculated on midline and midbrain/MHB cells, which most strongly recapitulate mediolateral patterns of gene expression, and 7 gene trend clusters were identified for this axis using a value *k=20*. Visualization of the UMAP representations indicated that cluster 3 in this analysis contained genes that were highly correlated with non-midbrain/MHB anterior-posterior identities; we therefore excluded this cluster from further analyses.

Identification of gene clusters in multiscale space

Comparing the spatial autocorrelation values of genes in the anterior-posterior and mediolateral axes, a large fraction of genes that were patterned relative to either DC0 or DC2 were also patterned along the orthogonal axis (Table S6), suggesting that these groups contain genes that

respond to inputs from both anterior-posterior and mediolateral patterning systems. For this reason, we sought to integrate multiple spatially relevant DCs into one analysis. To do so, a multiscale space embedding was calculated on the top 10 diffusion components of the developed neural plate dataset. We chose to use multiscale space as it provides an appropriate Euclidean space to more accurately measure the phenotypic similarity/distance between cells. Using this latent space, the Hotspot package was run with all genes and a value of *k=30*. Of the 19,623 genes detected in the cranial neural plate scRNA-seq dataset, 870 genes were identified as spatially informative.

Spatially informative genes were grouped into modules by performing hierarchical clustering on the Pearson correlation coefficient matrix. Correlations were calculated from the log-normalized count matrix, and hierarchical clustering was performed using Euclidean distance and the Ward linkage criterion. The number of clusters was determined by a linkage distance threshold, above which clusters were not merged. This value was determined by finding the knee-point on the median inter-cluster pairwise correlation of genes, calculated for increasing numbers of clusters (Figure 5C), and corresponds to an average correlation of approximately 0.5 among genes belonging to the same cluster. We chose this method to compromise between the similarity of genes within a cluster and the total number of clusters considered. In total, 15 spatially patterned gene clusters were identified (Figure 5C).

Visualization of gene expression trends

To visualize gene trends along each DC axis, we use generalized additive models (GAMs) with cubic splines as smoothing functions as in Palantir (Setty et al., 2019). In our work, trends for a module score were fitted using a regression model on the DC values (x-axis) and module score values (y-axis). The resulting smoothed trend was derived by dividing the data into 500 equally sized bins along each DC and predicting the module score at each bin using the regression fit.

Validation of predicted gene expression patterns

Prediction of gene expression patterns were performed on cells from E8.5-9.0 embryos based on the relative expression of genes along DC0 (to predict the anterior-posterior expression domain), or DC2 (to predict the mediolateral domain). To compare predictions to known gene expression patterns, we mined the Mouse Genome Informatics Gene Expression Database (http://informatics.jax.org) for images of RNA expression by *in situ* hybridization or protein localization by immunostaining of wild type embryos Theiler Stages 12-14. Links to the corresponding images were computationally generated in Python. This code, which can be used to look up available gene expression images for any gene at any stage in the Mouse Genome Informatics Gene Expression Database, is available as an annotated iPython notebook containing the code and instructions on Github (https://github.com/ZallenLab). Links to the requested stages are provided even if no images are currently available. For each gene predicted by our analysis to be in an anterior-posterior or mediolateral cluster, we compared the gene expression pattern in the cranial region in the published images to the predicted gene expression and a Yes, No, or NA (not applicable) determination was made. Genes were assigned as Yes if the predicted gene expression domain(s) and no others in cranial tissues showed expression in MGI database. Genes were assigned as No if the gene was not expressed in the predicted region, or if significant

expression outside of the predicted domain(s) within the cranial region was present, even if the gene was also expressed in the predicted domain(s). If no suitable pictures were present, the gene was assigned as NA.

Examples of images that match the anterior-posterior patterns predicted by our analysis can be found in the indicated panels in the following references: de la Pompa et al., 1997 (Ascl1, Figure 5C); Zhao and Duester, 2009 (Axin2, Figure 3C); Tamplin et al., 2011 (Clu1, supplementary data); Vacalla and Theil, 2002 (Casz1, Figure 2B); Hadjantonakis et al., 1998 (Celsr1, Figure 2F); Lyn and Giguere, 1994 (Crabp1, Figure 2A); Tahayato et al., 2003 (Cyp26c1, Figure 2A); Lewis et al., 2007 (Dkk1, Figure 1J): Broccoli et al., 2002 (Dmbx1, Figure 2A; En1, Figure 2B); Dunty et al., 2008 (Dusp6, Figure 2J); Flenniken et al., 1996 (Efnb1, Figure 5A); Ishikawa et al., 2003 (Egr2, Figure 5I); Anselme et al., 2007 (Emx2, Figure 3O; Pax6, Figure 2H); Hirata et al., 2004 (Fezf2, Figure 1A); Wright and Mansour, 2003 (Fgf3, Figure 1A); Cajal et al., 2012 (Fgf8, Figure 9M and Hesx1, Figure 9D); Chung et al., 2006 (Foxg1, Figure 4A; Gbx2, Figure 2I); Waters et al., 2003 (Gbx1, Figure 2G); Lobe, 1997 (Hes3, Figure 1E); Bertrand et al., 2011 (HoxB1, Figure S1I); Wong et al., 1999 (Mab21l2, Figure 1B); Oulad-Abdelghani et al., 1997 (Meis2, Figure 4B); Liguori et al., 2003 (Nkx2-1, Figure 4B); Tian et al., 2002 (Otx1, Figure 5G); Furushima et al., 2007 (Shisa1, Figure 3B); Treichel et al., 2001 (Sp5, Figure 2F); Hallonet et al., 1998 (Vax1, Figure 3A); Mielcarek et al., 2009 (Vgll3, Figure 3A); Tamplin et al., 2008 (Wfdc2, supplement).

Examples of images that match the mediolateral patterns predicted by our analysis can be found in the indicated panels in the following references: Lin et al., 2000 (Arl4A, Figure 2A); Tamplin et al., 2011 (Calca, Spink1, Wif1, supplementary data); Durmus et al., 2006 (Cthrc1, Figure 1B); Bouchard et al., 2005 (En2, Figure 2C); Burke and Oliver, 2002 (Foxa2, Figure 2C); Lee and Fan, 2001 (Gas1, Figure 1C); Hui et al,. 1994 (Gli3, Figure 4B); Klymiuk et al., 2012 (Ifitm1, Figure 2A); Failli et al., 2002 (Lmx1a, Figure 1A); Foerst-Potts and Sadler, 1997 (Msx1, Figure 7A); Trainor et al., 2002 (Msx2, Figure 3D); Pabst et al., 1998 (Nkx2-2, Figure 4C); Danesh et al., 2009 (Nog, Figure 1D1); Shylo et al., 2020 (Shh, Figure 6D); Pryor et al., 2014 (Vangl1, Figure 2A).

Differential gene expression in SAG-treated embryos

For cells belonging to spatial domains of the neural plate (forebrain, midbrain/MHB, and hindbrain), DEGs were calculated between SAG-treated and control populations using the R package MAST (Finak et al., 2015) on $\log_2$-transformed normalized counts. Midline cells were not included in this analysis, due to the low cell number recovered from both control and SAG treated samples. The DEGs related to each spatial domain for SAG-treated and control embryos are listed in Tables S10 and S11.

Expression analysis of transcription factors and receptors

For the analysis of transcriptional regulators, spatially patterned genes in the cranial neural plate (Table S6) were compared with a list of *Mus musculus* genes associated with the transcription factor term in the KEGG BRITE database (Kanehisa et al., 2023) using Release 110.0 or the transcription regulator activity GO term (GO: 0140110) (Ashburner et al., 2000; Gene Ontology Consortium et al., 2023) using the May, 2022 release (Table S8). *Nrg1, Wnt3a*, and *Wnt4* were excluded based on sequence homology. For the analysis of ligands and receptors, spatially

patterned genes in the cranial neural plate were compared with a list of predicted mouse ligands and receptors from the Cellinker database (Zhang et al., 2021) and the CellTalk database (Shao et al., 2021) (Table S9). *Ada, Bex3, Lman1, Mip,* and *Mylk*) were excluded based on sequence homology. Additional predicted ligands (*Lgi1, Lgi2*) and receptors (*Cldn10, Emp1, Lrtm1, Pcdh8, Tnfrsf19,* and *Vangl1*) that did not appear in either database were added.

Immunofluorescence

Timed pregnant dams were euthanized and embryos collected in ice cold PBS and then fixed in 4% paraformaldehyde (PFA) (Fisher 50-980-494) for 1-2 h at room temperature or overnight at 4ºC. Embryos were then washed three times for 30 min each in PBS + 0.1% Triton X-100 (Fisher 327371000) (PBTriton) at room temperature followed by a 1 h incubation in blocking solution (PBTriton + 3% BSA) at room temperature. Embryos were then incubated in staining solution (PBTriton + 1.5% BSA) with primary antibodies overnight at 4º C, followed by three 30 min washes in PBTriton at room temperature. Embryos were then incubated in staining solution with Alexa Fluor-conjugated secondary antibodies (1:500) (ThermoFisher) for 1 h at room temperature, followed by three 30 min washes in PBTriton at room temperature. Embryos were stored in PBTriton at 4º C until imaging. Primary antibodies were rat anti-E-cadherin (1:300) (Sigma U3254), rabbit anti-FoxA2 (1:1,000) (abcam ab108422), rabbit anti N-cadherin (1:500) (Cell Signaling Technology D4R1H), and mouse anti-Nkx6-1 (1:50) (Developmental Studies Hybridoma Bank F55A10). Secondary Alexa fluor conjugated antibodies were added along with Alexa 546-conjugated phalloidin (Molecular Probes) at a 1:1,000 dilution. Embryos were washed three times for 30 min each in PBTriton at room temperature and imaged on a Zeiss LSM 900 confocal microscope.

Confocal imaging

Whole-mount imaging was performed by mounting embryos dorsal side down in PBTriton (PBS + 0.1% Triton X-100) in Attofluor cell chambers (ThermoFisher A7816), using a small fragment of broken coverglass (Corning 2850-18) with small dabs of vacuum grease (Dow Corning DC 976) to mount the embryo on a #1.5 coverglass (Fisher 12-546-2P). Embryos were imaged by confocal microscopy on inverted microscopes on a Zeiss LSM700 equipped with a Plan-NeoFluar 40x/1.3 objective, a Zeiss LSM900 with a Plan-NeoFluar 40x/1.3 objective, or a Leica SP8 with an HC PL Apo 40x/1.3 objective. Images were captured by tile-based acquisition of contiguous z-stacks of 50-150 µm depth with 0.6-1.2 µm optical slices and 0.3-0.5 µm z-steps. For tiled images, computational stitching was performed using 10% overlap per tile in Zen (Zeiss) or LAS-X (Leica) software.

Figure assembly

Heatmaps, UMAPs, and plots of normalized expression levels were generated from cleaned and normalized expression data using Scanpy software (Wolf et al., 2018). For the heatmaps, gene expression was normalized to the highest expression of that gene along the diffusion component or to the highest expression of that gene within the treatment group (control or SAG). Only forebrain, midbrain, and hindbrain cells were included in the E8.5-9.0 DC0 heatmaps (midline cells were excluded). Only midline and midbrain cells were included in the E8.5-9.0 DC2

heatmaps (forebrain and hindbrain cells were excluded). For the SAG analysis, heatmaps that show cells from specific regions or combinations of regions are indicated.

Maximum-intensity projections of image z-stacks were generated in Zen (Zeiss), LAS-X (Leica), or FIJI. Figures were assembled using Photoshop and Illustrator software (Adobe). Fluorescence images were contrast adjusted for display using the FIJI redistribution of ImageJ (Schindelin et al., 2012).

Acknowledgments
The authors thank Meg Distinti, Tarek Islam, and Matthew Schilling for technical assistance and Alex Joyner, Tom Schultheiss, and members of the Zallen lab for helpful feedback on the manuscript. ERB was supported by NIH/NINDS F32 fellowship NS098832 and startup funds from North Carolina State University. AM, ML, and RS were supported by the Alan and Sandra Gerry Metastasis and Tumor Ecosystems Center, Memorial Sloan Kettering Cancer Center. All sequencing was performed by the Sloan Kettering Integrated Genomics Operation Core, funded by NIH/NCI Cancer Center Support Grant P30 CA008748, Cycle for Survival, and the Marie-Josée and Henry R. Kravis Center for Molecular Oncology, Memorial Sloan Kettering Cancer Center. All Single Cell Genomics was performed by the Single Cell Analytics and Innovation Lab, Memorial Sloan Kettering Cancer Center. JAZ and DP are investigators of the Howard Hughes Medical Institute.

Competing interests
DP is on the scientific advisory board of Insitro. The other authors declare that no competing interests exist.

Ethics
All animal experiments were conducted in accordance with PHS guidelines and the NIH Guide for the Care and Use of Laboratory Animals and an approved Institutional Animal Care and Use Committee protocol (15-08-13) of Memorial Sloan Kettering Cancer Center.

Data availability
All single-cell datasets generated in this study are publicly available at the Gene Expression Omnibus, accession number GSE273804. Processed h5ad files to examine gene expression patterns in cellxgene (https://github.com/chanzuckerberg/cellxgene) or scanpy (https://github.com/scverse/scanpy) (Wolf et al., 2018) will be made available on request while under review and added to the Gene Expression Omnibus upon publication. Code to automatically generate customizable links to gene expression images in the Mouse Genome Informatics Gene Expression Database is available on Github (https://github.com/ZallenLab).

Figure legends

Figure 1. Construction of a single-cell RNA expression atlas during mouse cranial neural tube closure. (A) Images of mouse embryos spanning the stages of cranial neural tube closure.

Bars, 100 µm. (B) Schematics of cranial tissues collected for scRNA-sequencing (dashed boxes, left) and neural tube closure in the midbrain (right) at the indicated stages. The heart was removed at later stages. (C and D) UMAP projections of all cranial cell populations analyzed colored by cell type (C) or embryonic stage (D). (E) UMAP projections of cranial neural plate cells colored by embryonic stage. (F-I) UMAP projections of cranial neural plate cells colored by normalized expression of genes primarily associated with the forebrain (*Otx2*), midbrain and midbrain-hindbrain boundary (MHB) (*En1*), hindbrain (*Gbx2*), and ventral midline (*Shh*). (J) UMAP projection of cranial neural plate cells colored by neural plate region.

Figure 2. Temporal changes in gene expression reveal shared and region-specific transcriptional trajectories. (A-C) Cells from the forebrain (A), midbrain (B), and hindbrain (C) at progressive stages of neural plate development, plotted relative to the time-correlated diffusion component 0 (DC0) in each region. (D-F) Normalized expression of *E-cadherin* (*Cdh1*) and *N-cadherin* (*Cdh2*) plotted relative to the normalized time-correlated diffusion component (DC0) in each region. (G) E-cadherin protein (magenta) is lost from cell-cell junctions and N-cadherin protein (green) accumulates at cell-cell junctions between E7.75 and E8.5 in the mouse cranial neural plate. (H) UMAP projections of cranial neural plate cells colored by normalized gene expression, showing examples of genes that are downregulated or upregulated throughout the cranial neural plate (left) and genes that are specifically upregulated in the forebrain, midbrain/MHB, or hindbrain. Bars, 100 μm (left panels in G), 20 μm (middle and right panels in G).

Figure 3. Region-specific patterns of gene expression along the anterior-posterior axis. (A) UMAP projections of cranial neural plate cells from E8.5-9.0 embryos colored by their value along the anterior-posterior-correlated diffusion component (DC0) (left) or by the normalized expression of markers primarily associated with the forebrain (*Otx2*), midbrain (*En1*), and hindbrain (*Gbx2*). (B and C) Line plot (B) and heatmap (C) showing the normalized expression of known markers of anterior-posterior identity relative to DC0. DC0 correctly orders forebrain, midbrain, and hindbrain markers relative to their known positions along the anterior-posterior axis. (D) Heatmap showing the normalized expression of 483 genes with high information content relative to DC0. Gene expression profiles are grouped into 11 clusters based on similarities in expression along DC0. Examples of genes in each cluster are listed on the right. (E and F) Schematics showing the predicted expression of example transcriptional regulators along the anterior-posterior axis. Anterior (A), posterior (P). (G) Heatmaps showing the normalized expression of example transcriptional regulators relative to DC0. Heatmaps show one gene per row, one cell per column, with the cells in each row ordered by their value along DC0. Colored bars (right) show the cluster identity relative to DC0. Cells assigned to the forebrain, midbrain/MHB, or hindbrain are indicated at the top of each heatmap.

Figure 4. Patterned gene expression along the mediolateral axis of the developing midbrain. (A) UMAP projections of cranial neural plate cells from E8.5-9.0 embryos colored by their value along the mediolaterally correlated diffusion component (DC2) (left) or by the normalized expression of markers for the ventral midline (*Shh*), medial cells (*Ptch1*), or lateral cells (*Pax7*). (B and C) Line plot (B) and heatmap (C) showing the normalized expression of

known markers of mediolateral cell identity relative to DC2 in the midbrain/MHB. DC2 correctly orders midline, medial, and lateral markers relative to their known positions along the mediolateral axis. (D) Heatmap showing the normalized expression of 222 genes with high information content relative to DC2. Gene expression profiles are grouped into 7 clusters based on similarities in expression along DC2. Examples of genes in each cluster are listed on the right. Cluster 3 (not shown) displayed divergent UMAP patterns associated with anterior-posterior patterning and was excluded from further analysis. (E) Heatmaps showing the normalized expression of specific transcriptional regulators relative to DC2. (F) Schematics showing the predicted expression of example transcriptional regulators along the mediolateral axis of the midbrain/MHB. Heatmaps show one gene per row, one cell per column, with the cells in each row ordered by their value along DC0. Colored bars (right) show the cluster identity relative to DC2. Cells assigned to the midline or midbrain/MHB are indicated at the top of each heatmap.

Figure 5. Multiscale analysis of spatial gene expression in the mouse cranial neural plate. (A) Plot of genes by Z-Normalized Moran's I relative to DC0 and DC2 for all genes expressed in the E8.5-9.0 cranial neural plate. Genes meeting the cutoff for further analysis are colored according to cluster identity at D=10; other genes are shown in gray. (B) Spatially informative genes from (A) replotted by spatial autocorrelation (Moran's I) along DC0 and DC2, colored by cluster identity at D=10. (C) Dendrogram of spatially informative genes showing gene clusters at different distances corresponding to the average correlation among genes. Left, the number of clusters at each distance (black curve) and the average correlation between expression patterns within clusters (blue curve). Right, dendrogram showing the relationships between clusters, colored by cluster identity at D=10 as in A and B. (D) UMAP projections of E8.5-9.0 neural plate cells colored by normalized gene expression. Examples of genes that mark specific presumptive brain structures (the telencephalon, diencephalon, and midbrain-hindbrain boundary), midline or lateral domains, or subsets of rhombomeres are shown. Brackets and boxes indicate regions of increased gene expression. Spatial cluster identities at a stringency of D=6 (top and middle panels) or D=4 (bottom panels) are indicated.

Figure 6. Patterned expression of secreted and transmembrane proteins in the cranial neural plate. (A and B) Heatmaps showing the normalized expression of transmembrane and secreted proteins relative to DC0 (A) or DC2 (B). Heatmaps show one gene per row, one cell per column, with the cells in each row ordered by their value along DC0 or DC2. Colored bars (right) show the cluster identity relative to DC0 or DC2. Cells assigned to the forebrain (FB), midbrain/MHB (MB/MHB), hindbrain (HB), or midline are indicated at the top of each heatmap. (C) Schematics showing the predicted expression of example transmembrane and secreted proteins along the anterior-posterior axis. Anterior (A), posterior (P). (D-F) UMAP projections of cranial neural plate cells from E8.5-9.0 embryos colored by the normalized expression of a subset of *Wnt* ligands (D), *Frizzled* receptors (E), and *Ephrin* ligands and *Eph* receptors (F). Spatial cluster identities at D=10 are indicated.

Figure 7. Single-cell RNA sequencing reveals region-specific transcriptional programs activated by SHH signaling. (A) Maximum-intensity projections of the midbrain and anterior hindbrain regions of the cranial neural plate of 5-somite mouse embryos treated with vehicle

control (left) or with 2 μM Smoothened Agonist (SAG) stained for FOXA2 and NKX6-1. SAG-treated embryos display an expanded floor plate region. Bars, 100 μm. (B and C) UMAP projections of 1,619 and 1,409 cranial neural plate cells from control and SAG-treated embryos at E8.5, respectively, colored by treatment (B) or anterior-posterior domain (C). (D-J) Heatmaps showing the normalized expression of candidate SHH target genes that were upregulated or downregulated throughout the cranial neural plate (D) or in the indicated region(s) (E-J) in SAG-treated embryos. Heatmaps show one gene per row, one cell per column, with the cells in each row grouped by control (green) or SAG treatment (orange), indicated at the top of each heatmap. Genes involved in SHH signaling (red) and genes assigned to medial (purple), lateral (blue), anterior-posterior (green), or other spatial clusters (orange) are indicated. Genes that had a MAST hurdle value >0.24 or <-0.24 and a false discovery rate adjusted p-value of p<0.001 in at least one region were indicated as differentially expressed in all regions in which p<0.001 and the MAST hurdle value was >0.10 or <-0.10.

Supplementary figure legends

Figure 1—figure supplement 1. Assignment of cell identities in the mouse cranial region. (A) Dot plot showing the normalized expression of a subset of markers used for cell type determination. (B) UMAP projection of all cranial cells analyzed (39,463 cells), colored by PhenoGraph cluster. Assigned cell types are listed on the right. (C) Distribution of cell cycle stages in the dataset.

Figure 1—figure supplement 2. Assignment of cell identities in the mouse cranial neural plate. (A) Dot plot showing the normalized expression of a subset of markers used to assign cells to different neural plate regions. (B) UMAP projection of cranial neural plate cells (17,695 cells) reclustered in the absence of other cell types, colored by PhenoGraph cluster. Assigned neural plate regions are listed on the right. (C) UMAP projections of cranial neural plate cells colored by normalized gene expression. Markers for radial glial cells (*Gfap*) and immature neurons (*Neurod1, Dcx*) were not detected in the cranial neural plate at these stages, suggesting that cells are pre-neurogenic, although weak expression of neuronal cytoskeletal components (*Nefl, Tubb3/Tuj1, Mapt*) was observed.

Figure 2—figure supplement 1. Analysis of gene expression trends in the developing forebrain, midbrain, and hindbrain. (A-F) UMAP projections of neural plate cells separated by region based on known markers of anterior-posterior identity. Cells are colored by embryo stage (A-C) or by their value along the top diffusion component (DC0) for each region (D-F). Higher DC0 values correlate with later time points. (G-I) Gene expression profiles were clustered by differential expression along DC0 and the average normalized expression (solid blue line) +/- 1 standard deviation (dotted blue lines) and the expression of all individual genes in each cluster (gray lines) are shown.

Figure 2—figure supplement 2. Examples of genes that are temporally regulated throughout the cranial neural plate. (A-C) UMAP projections of cranial neural plate cells colored

by normalized gene expression. Genes that display globally decreasing expression (A), globally increasing expression (B), or increasing expression in a subset of domains (C) are shown.

Figure 2—figure supplement 3. Examples of genes that are upregulated in the forebrain, midbrain, or hindbrain. (A-C) UMAP projections of cranial neural plate cells colored by normalized gene expression. Genes that display increasing expression in the future forebrain (A), midbrain/MHB (B), or hindbrain (C) are shown.

Figure 3—figure supplement 1. Diffusion component analysis of E8.5-9.0 neural plate cells. UMAP projections of neural plate cells from E8.5-E 9.0 embryos colored by the value along the indicated diffusion component. Red indicates higher DC values and blue indicates lower DC values. The top 10 diffusion components calculated for the dataset are shown.

Figure 3—figure supplement 2. Clustering analysis reveals distinct patterns of gene expression along DC0 in the E8.5-9.0 cranial neural plate. (A and B) Gene expression profiles in the E8.5-9.0 neural plate were clustered by differential expression along DC0, which correlates with cell position along the anterior-posterior axis. (A) UMAP projections of E8.5-9.0 neural plate cells colored by the average gene expression for each cluster. Red indicates higher expression and blue indicates lower expression of the genes in each cluster. (B) The average normalized gene expression for each cluster (solid blue line) +/- 1 standard deviation (dotted blue lines) and all individual genes in each cluster (gray lines) are shown.

Figure 4—figure supplement 1. Clustering analysis reveals distinct patterns of gene expression along DC2 in the E8.5-9.0 cranial neural plate. (A and B) Gene expression profiles in the E8.5-9.0 neural plate were clustered by differential expression along DC2, which correlates with cell position along the mediolateral axis. (A) UMAP projections of E8.5-9.0 neural plate cells colored by the average gene expression for each cluster. Red indicates higher expression and blue indicates lower expression of genes in the cluster. (B) The average normalized gene expression for each cluster (solid blue line) +/- 1 standard deviation (dotted blue lines) and all individual genes in each cluster (gray lines) are shown. (C-F) Heatmaps (C and E) and line plots (D and F) showing the normalized expression of a subset of transcriptional regulators (C and D) or transmembrane and secreted proteins (E and F) that appear to show graded changes in expression relative to DC2. Heatmaps show one gene per row, one cell per column, with the cells in each row ordered by their value along DC2. Cells assigned to the midline or midbrain/MHB are indicated at the top of each heatmap.

Figure 5—figure supplement 1. Multiscale clustering analysis reveals spatial patterns of gene expression in the cranial neural plate. UMAP projections of E8.5-9.0 neural plate cells colored by the average expression of genes in each spatial cluster. Red indicates higher expression and blue indicates lower expression of the genes in each cluster.

Figure 6—figure supplement 1. Patterned expression of gene families. (A-K) UMAP projections of E8.5-9.0 neural plate cells colored according to relative expression of an individual member of the indicated gene family. (A) *Wnt* ligands, (B) *Frizzled* receptors, (C) *Eph/Ephrin*

signaling molecules, (D) *Fgf* ligands, (E) *Fgf* receptors, (F) *Fox* transcription factors, (G) *Hes* transcription factors, (H) *Irx* transcription factors, (I) *Msx* transcription factors, (J) *Pax* transcription factors, and (K) *Zic* transcription factors. Spatial cluster identities at a stringency of D=10 are shown.

Table legends

Table S1. Cell type and quality control metrics for wild-type and SAG treatment datasets. The total number of cells post-filtering, the number of cells assigned to each cell type, and quality metrics (median number of UMIs, median number of genes, median percent mitochondrial UMIs, median percent ribosomal UMIs, and mean reads/UMI) are shown for each replicate in the wild-type, SAG treatment, and SAG control datasets.

Table S2. PhenoGraph clustering of all cranial cells in E7.5-9.0 embryos. Gene expression for each PhenoGraph cluster in the full cranial cell dataset. The false-discovery rate (FDR) adjusted p-value, MAST hurdle value, and empirical $\log_2$ fold change are shown for each gene.

Table S3. PhenoGraph clustering of cranial neural plate cells in E7.5-9.0 embryos. Gene expression for each PhenoGraph cluster in the cranial neural plate are shown. The false-discovery rate (FDR) adjusted p-value, MAST hurdle value, and empirical $\log_2$ fold change are shown for each gene.

Table S4. Temporally regulated genes in the forebrain, midbrain, and hindbrain of E7.5-9.0 embryos. Gene correlation scores for the temporally correlated diffusion component (DC0) in each region are shown for the forebrain, midbrain/midbrain-hindbrain boundary, and hindbrain cranial neural plate populations. The cluster number, Pearson correlation coefficient, and normalized range of expression are shown for each gene.

Table S5. Diffusion component gene correlations in the E8.5-9.0 cranial neural plate. Gene correlations with each of the top 10 diffusion components in the E8.5-9.0 cranial neural plate are shown. The false discovery rate (FDR) adjusted p-value and Pearson correlation coefficient for each diffusion component are shown for each gene.

Table S6. Spatially patterned genes in the E8.5-9.0 cranial neural plate. All genes detected in cells from wild-type E8.5-9.0 embryos are shown. The results of HotSpot analysis are given, including the Geary's C spatial autocorrelation coefficient (C, MSD), the z-scored coefficient (Z, MSD), the p-value (Pval, MSD), the false discovery rate (FDR, MSD), and the normalized range of expression for each gene, in addition to the results of HotSpot analysis for the anterior-posterior (AP) axis-correlated diffusion component (DC0) and the mediolateral (ML) axis-correlated diffusion component (DC2). Cluster identities along the AP axis (DC0), ML axis (DC2), and at various 2D distance (D) values are shown.

Table S7. Comparison of predicted gene expression patterns to published data. For genes predicted to be spatially patterned in Table S6, the predicted expression of genes in anterior-

posterior (AP) or mediolateral (ML) clusters was compared with published images in the Mouse Genome Informatics Gene eXpression Database (GXD). In addition to the information in Table S6, the predicted and observed expression patterns and whether they agree is reported and a link to available GXD images of wild-type embryos from Theiler stages 12-14 is provided.

Table S8. Spatially patterned transcriptional regulators in the E8.5-9.0 cranial neural plate. Spatial (2D) cluster (distance D=10), AP cluster, and ML cluster of transcriptional regulators in the E8.5-9.0 cranial neural plate are shown.

Table S9. Spatially patterned secreted and transmembrane proteins in the E8.5-9.0 cranial neural plate. Spatial (2D) cluster (distance D=10), AP cluster, and ML cluster of secreted (ligands) and transmembrane (receptors) proteins in the dataset are shown.

Table S10. Differentially expressed genes in SAG-treated and control embryos in the E8.5 cranial neural plate. Results of MAST differential gene expression analysis in SAG-treated vs, control embryos separated by forebrain, midbrain/MHB, and hindbrain regions. The false discovery rate (FDR) adjusted p-value, MAST hurdle value, and empirical $\log_2$ fold change value are shown for each gene.

Table S11. Regional analysis of transcriptional changes upon SAG treatment. Genes with significant expression changes from Table S10 (FDR-adjusted p-value $p<0.001$) grouped by region, including all cranial neural plate regions (FB-MB-HB), two regions (FB-MB, FB-HB, MB-HB), or single regions (FB, MB, HB). Upregulated genes with a MAST hurdle value > 0.24 are highlighted in green. Downregulated genes with a MAST hurdle value < -0.24 are highlighted in red.

Table S12. Genes used for cell cycle normalization. Gene ontology (GO) terms associated with the cell cycle and the associated genes used in cell cycle normalization are listed.

Table S13. Genes used for cell typing. Genes used to assign cell/tissue type are listed, along with their positive or negative expression in that tissue and any expression overlap in other tissues.

References

Acampora D, Mazan S, Lallemand Y, Avantaggiato V, Maury M, Simeone A, Brûlet P. 1995. Forebrain and midbrain regions are deleted in Otx2-/- mutants due to a defective anterior neuroectoderm specification during gastrulation. *Development* **121**:3279–3290.

Achim K, Peltopuro P, Lahti L, Tsai H-H, Zachariah A, Astrand M, Salminen M, Rowitch D, Partanen J. 2013. The role of Tal2 and Tal1 in the differentiation of midbrain GABAergic neuron precursors. *Biol Open* **2**:990–997.

Agarwala S, Sanders TA, Ragsdale CW. 2001. Sonic hedgehog control of size and shape in midbrain pattern formation. *Science* **291**:2147–2150.

Allen BL, Tenzen T, McMahon AP. 2007. The Hedgehog-binding proteins Gas1 and Cdo cooperate to positively regulate Shh signaling during mouse development. *Genes Dev* **21**:1244–1257.

Ang SL, Jin O, Rhinn M, Daigle N, Stevenson L, Rossant J. 1996. A targeted mouse Otx2 mutation leads to severe defects in gastrulation and formation of axial mesoderm and to deletion of rostral brain. *Development* **122**:243–252.

Anselme I, Laclef C, Lanaud M, Rüther U, Schneider-Maunoury S. 2007. Defects in brain patterning and head morphogenesis in the mouse mutant Fused toes. *Dev Biol* **304**:208–220.

Argelaguet R, Clark SJ, Mohammed H, Stapel LC, Krueger C, Kapourani C-A, Imaz-Rosshandler I, Lohoff T, Xiang Y, Hanna CW, Smallwood S, Ibarra-Soria X, Buettner F, Sanguinetti G, Xie W, Krueger F, Göttgens B, Rugg-Gunn PJ, Kelsey G, Dean W, Nichols J, Stegle O, Marioni JC, Reik W. 2019. Multi-omics profiling of mouse gastrulation at single-cell resolution. *Nature* **576**:487–491.

Arraf AA, Yelin R, Reshef I, Jadon J, Abboud M, Zaher M, Schneider J, Vladimirov FK, Schultheiss TM. 2020. Hedgehog Signaling Regulates Epithelial Morphogenesis to Position the Ventral Embryonic Midline. *Dev Cell* **https://doi.org/10.1016/j.decel.2020.04.016**. doi:10.1016/j.devcel.2020.04.016

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**:25–29.

Ashique AM, Choe Y, Karlen M, May SR, Phamluong K, Solloway MJ, Ericson J, Peterson AS. 2009. The Rfx4 transcription factor modulates Shh signaling by regional control of ciliogenesis. *Sci Signal* **2**:ra70.

Azizi E, Carr AJ, Plitas G, Cornish AE, Konopacki C, Prabhakaran S, Nainys J, Wu K, Kiseliovas V, Setty M, Choi K, Fromme RM, Dao P, McKenney PT, Wasti RC, Kadaveru K, Mazutis L, Rudensky AY, Pe'er D. 2018. Single-Cell Map of Diverse Immune Phenotypes in the Breast Tumor Microenvironment. *Cell* **174**:1293-1308.e36.

Bach A, Lallemand Y, Nicola M-A, Ramos C, Mathis L, Maufras M, Robert B. 2003. Msx1 is required for dorsal diencephalon patterning. *Development* **130**:4025–4036.

Baldarelli RM, Smith CM, Finger JH, Hayamizu TF, McCright IJ, Xu J, Shaw DR, Beal JS, Blodgett O, Campbell J, Corbani LE, Frost PJ, Giannatto SC, Miers DB, Kadin JA, Richardson JE, Ringwald M. 2021. The mouse Gene Expression Database (GXD): 2021 update. *Nucleic Acids Res* **49**:D924–D931.

Balordi F, Fishell G. 2007. Hedgehog signaling in the subventricular zone is required for both the maintenance of stem cells and the migration of newborn neurons. *J Neurosci* **27**:5936–5947.

Bell SM, Schreiner CM, Waclaw RR, Campbell K, Potter SS, Scott WJ. 2003. Sp8 is crucial for limb outgrowth and neuropore closure. *Proc Natl Acad Sci U S A* **100**:12195–12200.

Bildsoe H, Loebel DAF, Jones VJ, Hor ACC, Braithwaite AW, Chen Y-T, Behringer RR, Tam PPL. 2013. The mesenchymal architecture of the cranial mesoderm of mouse embryos is disrupted by the loss of Twist1 function. *Dev Biol* **374**:295–307.

Blaess S, Corrales JD, Joyner AL. 2006. Sonic hedgehog regulates Gli activator and repressor functions with spatial and temporal precision in the mid/hindbrain region. *Development* **133**:1799–1809.

Borday C, Cabochette P, Parain K, Mazurier N, Janssens S, Tran HT, Sekkali B, Bronchain O, Vleminckx K, Locker M, Perron M. 2012. Antagonistic cross-regulation between Wnt and Hedgehog signalling pathways controls post-embryonic retinal proliferation. *Development* **139**:3499–3509.

Bouchard M, Grote D, Craven SE, Sun Q, Steinlein P, Busslinger M. 2005. Identification of Pax2-regulated genes by expression profiling of the mid-hindbrain organizer region. *Development* **132**:2633–2643.

Briscoe J, Sussel L, Serup P, Hartigan-O'Connor D, Jessell TM, Rubenstein JL, Ericson J. 1999. Homeobox gene *Nkx2.2* and specification of neuronal identity by graded Sonic hedgehog signalling. *Nature* **398**:622–627.

Broccoli V, Colombo E, Cossu G. 2002. Dmbx1 is a paired-box containing gene specifically expressed in the caudal most brain structures. *Mech Dev* **114**:219–223.

Brooks ER, Islam MT, Anderson KV, Zallen JA. 2020. Sonic hedgehog signaling directs patterned cell remodeling during cranial neural tube closure. *Elife* **9**. doi:10.7554/eLife.60234

Bucher K, Sofroniew MV, Pannell R, Impey H, Smith AJ, Torres EM, Dunnett SB, Jin Y, Baer R, Rabbitts TH. 2000. The T cell oncogene Tal2 is necessary for normal development of the mouse brain. *Dev Biol* **227**:533–544.

Burke Z, Oliver G. 2002. Prox1 is an early specific marker for the developing liver and pancreas in the mammalian foregut endoderm. *Mech Dev* **118**:147–155.

Cajal M, Lawson KA, Hill B, Moreau A, Rao J, Ross A, Collignon J, Camus A. 2012. Clonal and molecular analysis of the prospective anterior neural boundary in the mouse embryo. *Development* **139**:423–436.

Camus A, Davidson BP, Billiards S, Khoo P, Rivera-Pérez JA, Wakamiya M, Behringer RR, Tam PP. 2000. The morphogenetic role of midline mesendoderm and ectoderm in the development of the forebrain and the midbrain of the mouse embryo. *Development* **127**:1799–1813.

Cao J, Spielmann M, Qiu X, Huang X, Ibrahim DM, Hill AJ, Zhang F, Mundlos S, Christiansen L, Steemers FJ, Trapnell C, Shendure J. 2019. The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566**:496–502.

Cardoso-Moreira M, Halbert J, Valloton D, Velten B, Chen C, Shao Y, Liechti A, Ascenção K, Rummel C, Ovchinnikova S, Mazin PV, Xenarios I, Harshman K, Mort M, Cooper DN, Sandi C, Soares MJ, Ferreira PG, Afonso S, Carneiro M, Turner JMA, VandeBerg JL, Fallahshahroudi A, Jensen P, Behr R, Lisgo S, Lindsay S, Khaitovich P, Huber W, Baker J, Anders S, Zhang YE, Kaessmann H. 2019. Gene expression across mammalian organ development. *Nature* **571**:505–509.

Chen JK, Taipale J, Young KE, Maiti T, Beachy PA. 2002. Small molecule modulation of Smoothened activity. *Proc Natl Acad Sci U S A* **99**:14071–14076.

Chen VS, Morrison JP, Southwell MF, Foley JF, Bolon B, Elmore SA. 2017. Histology Atlas of the Developing Prenatal and Postnatal Mouse Central Nervous System, with Emphasis on Prenatal Days E7.5 to E18.5. *Toxicol Pathol* **45**:705–744.

Chen ZF, Behringer RR. 1995. twist is required in head mesenchyme for cranial neural tube morphogenesis. *Genes Dev* **9**:686–699.

Cheng S, Pei Y, He L, Peng G, Reinius B, Tam PPL, Jing N, Deng Q. 2019. Single-cell RNA-seq reveals cellular heterogeneity of pluripotency transition and X chromosome dynamics during early mouse development. *Cell Rep* **26**:2593-2607.e3.

Chi CL, Martinez S, Wurst W, Martin GR. 2003. The isthmic organizer signal FGF8 is required for cell survival in the prospective midbrain and cerebellum. *Development* **130**:2633–2644.

Christodoulou N, Skourides PA. 2022. Distinct spatiotemporal contribution of morphogenetic events and mechanical tissue coupling during Xenopus neural tube closure. *Development* **149**. doi:10.1242/dev.200358

Chung AC-K, Xu X, Niederreither KA, Cooney AJ. 2006. Loss of orphan nuclear receptor GCNF function disrupts forebrain development and the establishment of the isthmic organizer. *Dev Biol* **293**:13–24.

Creanga A, Glenn TD, Mann RK, Saunders AM, Talbot WS, Beachy PA. 2012. Scube/You activity mediates release of dually lipid-modified Hedgehog signal in soluble form. *Genes Dev* **26**:1312–1325.

Crossley PH, Martinez S, Martin GR. 1996. Midbrain development induced by FGF8 in the chick embryo. *Nature* **380**:66–68.

Curtin JA, Quint E, Tsipouri V, Arkell RM, Cattanach B, Copp AJ, Henderson DJ, Spurr N, Stanier P, Fisher EM, Nolan PM, Steel KP, Brown SDM, Gray IC, Murdoch JN. 2003. Mutation of Celsr1 disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. *Curr Biol* **13**:1129–1133.

Danesh SM, Villasenor A, Chong D, Soukup C, Cleaver O. 2009. BMP and BMP receptor expression during murine organogenesis. *Gene Expr Patterns* **9**:255–265.

Dattani MT, Martinez-Barbera JP, Thomas PQ, Brickman JM, Gupta R, Mårtensson IL, Toresson H, Fox M, Wales JK, Hindmarsh PC, Krauss S, Beddington RS, Robinson IC. 1998. Mutations in the homeobox gene HESX1/Hesx1 associated with septo-optic dysplasia in human and mouse. *Nat Genet* **19**:125–133.

de la Pompa JL, Wakeham A, Correia KM, Samper E, Brown S, Aguilera RJ, Nakano T, Honjo T, Mak TW, Rossant J, Conlon RA. 1997. Conservation of the Notch signalling pathway in mammalian neurogenesis. *Development* **124**:1139–1148.

de Soysa TY, Ranade SS, Okawa S, Ravichandran S, Huang Y, Salunga HT, Schricker A, Del Sol A, Gifford CA, Srivastava D. 2019. Single-cell analysis of cardiogenesis reveals basis for organ-level developmental defects. *Nature* **572**:120–124.

Delile J, Rayon T, Melchionda M, Edwards A, Briscoe J, Sagner A. 2019. Single cell transcriptomics reveals spatial and temporal dynamics of gene expression in the developing mouse spinal cord. *Development* **146**. doi:10.1242/dev.173807

Dessaud E, McMahon AP, Briscoe J. 2008. Pattern formation in the vertebrate neural tube: a sonic hedgehog morphogen-regulated transcriptional network. *Development* **135**:2489–2503.

DeTomaso D, Jones MG, Subramaniam M, Ashuach T, Ye CJ, Yosef N. 2019. Functional interpretation of single cell similarity maps. *Nat Commun* **10**:4376.

DeTomaso D, Yosef N. 2021. Hotspot identifies informative gene modules across modalities of single-cell genomics. *Cell Syst* **12**:446-456.e9.

Di Bella DJ, Habibi E, Stickels RR, Scalia G, Brown J, Yadollahpour P, Yang SM, Abbate C, Biancalani T, Macosko EZ, Chen F, Regev A, Arlotta P. 2021. Molecular logic of cellular diversification in the mouse cerebral cortex. *Nature* **595**:554–559.

Dickinson ME, Selleck MA, McMahon AP, Bronner-Fraser M. 1995. Dorsalization of the neural tube by the non-neural ectoderm. *Development* **121**:2099–2106.

Dunty WC Jr, Biris KK, Chalamalasetty RB, Taketo MM, Lewandoski M, Yamaguchi TP. 2008. Wnt3a/beta-catenin signaling controls posterior body development by coordinating mesoderm formation and segmentation. *Development* **135**:85–94.

Durmus T, LeClair RJ, Park K-S, Terzic A, Yoon JK, Lindner V. 2006. Expression analysis of the novel gene collagen triple helix repeat containing-1 (Cthrc1). *Gene Expr Patterns* **6**:935–940.

Elliott KH, Chen X, Salomone J, Chaturvedi P, Schultz PA, Balchand SK, Servetas JD, Zuniga A, Zeller R, Gebelein B, Weirauch MT, Peterson KA, Brugmann SA. 2020. Gli3 utilizes Hand2 to synergistically regulate tissue-specific transcriptional networks. *Elife* **9**. doi:10.7554/eLife.56450

Failli V, Bachy I, Rétaux S. 2002. Expression of the LIM-homeodomain gene Lmx1a (dreher) during development of the mouse nervous system. *Mech Dev* **118**:225–228.

Favor J, Sandulache R, Neuhäuser-Klaus A, Pretsch W, Chatterjee B, Senft E, Wurst W, Blanquet V, Grimes P, Spörle R, Schughart K. 1996. The mouse Pax2(1Neu) mutation is identical to a human PAX2 mutation in a family with renal-coloboma syndrome and results in

developmental defects of the brain, ear, eye, and kidney. *Proc Natl Acad Sci U S A* **93**:13870–13875.

Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M, Linsley PS, Gottardo R. 2015. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* **16**:278.

Fleming SJ, Chaffin MD, Arduini A, Akkad A-D, Banks E, Marioni JC, Philippakis AA, Ellinor PT, Babadi M. 2023. Unsupervised removal of systematic background noise from droplet-based single-cell experiments using CellBender. *Nat Methods* **20**:1323–1335.

Flenniken AM, Gale NW, Yancopoulos GD, Wilkinson DG. 1996. Distinct and overlapping expression patterns of ligands for Eph-related receptor tyrosine kinases during mouse embryogenesis. *Dev Biol* **179**:382–401.

Foerst-Potts L, Sadler TW. 1997. Disruption of Msx-1 and Msx-2 reveals roles for these genes in craniofacial, eye, and axial development. *Dev Dyn* **209**:70–84.

Fuccillo M, Rallu M, McMahon AP, Fishell G. 2004. Temporal requirement for hedgehog signaling in ventral telencephalic patterning. *Development* **131**:5031–5040.

Furushima K, Yamamoto A, Nagano T, Shibata M, Miyachi H, Abe T, Ohshima N, Kiyonari H, Aizawa S. 2007. Mouse homologues of Shisa antagonistic to Wnt and Fgf signalings. *Dev Biol* **306**:480–492.

Gene Ontology Consortium, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, Hill DP, Lee R, Mi H, Moxon S, Mungall CJ, Muruganugan A, Mushayahama T, Sternberg PW, Thomas PD, Van Auken K, Ramsey J, Siegele DA, Chisholm RL, Fey P, Aspromonte MC, Nugnes MV, Quaglia F, Tosatto S, Giglio M, Nadendla S, Antonazzo G, Attrill H, Dos Santos G, Marygold S, Strelets V, Tabone CJ, Thurmond J, Zhou P, Ahmed SH, Asanitthong P, Luna Buitrago D, Erdol MN, Gage MC, Ali Kadhum M, Li KYC, Long M, Michalak A, Pesala A, Pritazahra A, Saverimuttu SCC, Su R, Thurlow KE, Lovering RC, Logie C, Oliferenko S, Blake J, Christie K, Corbani L, Dolan ME, Drabkin HJ, Hill DP, Ni L, Sitnikov D, Smith C, Cuzick A, Seager J, Cooper L, Elser J, Jaiswal P, Gupta P, Jaiswal P, Naithani S, Lera-Ramirez M, Rutherford K, Wood V, De Pons JL, Dwinell MR, Hayman GT, Kaldunski ML, Kwitek AE, Laulederkind SJF, Tutaj MA, Vedi M, Wang S-J, D'Eustachio P, Aimo L, Axelsen K, Bridge A, Hyka-Nouspikel N, Morgat A, Aleksander SA, Cherry JM, Engel SR, Karra K, Miyasato SR, Nash RS, Skrzypek MS, Weng S, Wong ED, Bakker E, Berardini TZ, Reiser L, Auchincloss A, Axelsen K, Argoud-Puy G, Blatter M-C, Boutet E, Breuza L, Bridge A, Casals-Casas C, Coudert E, Estreicher A, Livia Famiglietti M, Feuermann M, Gos A, Gruaz-Gumowski N, Hulo C, Hyka-Nouspikel N, Jungo F, Le Mercier P, Lieberherr D, Masson P, Morgat A, Pedruzzi I, Pourcel L, Poux S, Rivoire C, Sundaram S, Bateman A, Bowler-Barnett E, Bye-A-Jee H, Denny P, Ignatchenko A, Ishtiaq R, Lock A, Lussi Y, Magrane M, Martin MJ, Orchard S, Raposo P, Speretta E, Tyagi N, Warner K, Zaru R, Diehl AD, Lee R, Chan J, Diamantakis S, Raciti D, Zarowiecki M, Fisher M, James-Zorn C, Ponferrada V, Zorn A, Ramachandran S, Ruzicka L, Westerfield M. 2023. The Gene Ontology knowledgebase in 2023. *Genetics* **224**. doi:10.1093/genetics/iyad031

Goetz SC, Anderson KV. 2010. The primary cilium: a signalling centre during vertebrate development. *Nat Rev Genet* **11**:331–344.

Goodrich LV, Johnson RL, Milenkovic L, McMahon JA, Scott MP. 1996. Conservation of the hedgehog/patched signaling pathway from flies to mice: induction of a mouse patched gene by Hedgehog. *Genes Dev* **10**:301–312.

Götz M, Huttner WB. 2005. The cell biology of neurogenesis. *Nat Rev Mol Cell Biol* **6**:777–788.

Hadjantonakis AK, Formstone CJ, Little PF. 1998. mCelsr1 is an evolutionarily conserved seven-pass transmembrane receptor and is expressed during mouse embryonic development. *Mech Dev* **78**:91–95.

Hallonet M, Hollemann T, Wehr R, Jenkins NA, Copeland NG, Pieler T, Gruss P. 1998. Vax1 is a novel homeobox-containing gene expressed in the developing anterior ventral forebrain. *Development* **125**:2599–2610.

Harris MJ, Juriloff DM. 2010. An update to the list of mouse mutants with neural tube closure defects and advances toward a complete genetic perspective of neural tube closure. *Birth Defects Res A Clin Mol Teratol* **88**:653–669.

Harris MJ, Juriloff DM. 2007. Mouse mutants with neural tube closure defects and their role in understanding human neural tube defects. *Birth Defects Res A Clin Mol Teratol* **79**:187–210.

Hatakeyama J, Bessho Y, Katoh K, Ookawara S, Fujioka M, Guillemot F, Kageyama R. 2004. Hes genes regulate size, shape and histogenesis of the nervous system by control of the timing of neural stem cell differentiation. *Development* **131**:5539–5550.

Hirata H, Tomita K, Bessho Y, Kageyama R. 2001. Hes1 and Hes3 regulate maintenance of the isthmic organizer and development of the mid/hindbrain. *EMBO J* **20**:4454–4466.

Hirata T, Suda Y, Nakao K, Narimatsu M, Hirano T, Hibi M. 2004. Zinc finger gene fez-like functions in the formation of subplate neurons and thalamocortical axons. *Dev Dyn* **230**:546–556.

Hoch RV, Rubenstein JLR, Pleasure S. 2009. Genes and signaling events that establish regional patterning of the mammalian forebrain. *Semin Cell Dev Biol* **20**:378–386.

Hui CC, Slusarski D, Platt KA, Holmgren R, Joyner AL. 1994. Expression of three mouse homologs of the Drosophila segment polarity gene cubitus interruptus, Gli, Gli-2, and Gli-3, in ectoderm- and mesoderm-derived tissues suggests multiple roles during postimplantation development. *Dev Biol* **162**:402–413.

Ingham PW, McMahon AP. 2001. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev* **15**:3059–3087.

Inoue T, Hatayama M, Tohmonda T, Itohara S, Aruga J, Mikoshiba K. 2004. Mouse Zic5 deficiency results in neural tube defects and hypoplasia of cephalic neural crest derivatives. *Dev Biol* **270**:146–162.

Ishikawa T-O, Tamai Y, Li Q, Oshima M, Taketo MM. 2003. Requirement for tumor suppressor Apc in the morphogenesis of anterior and ventral mouse embryo. *Dev Biol* **253**:230–246.

Joyner AL, Herrup K, Auerbach BA, Davis CA, Rossant J. 1991. Subtle cerebellar phenotype in mice homozygous for a targeted deletion of the En-2 homeobox. *Science* **251**:1239–1243.

Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* **51**:D587–D592.

Katoh Y, Katoh M. 2006. WNT antagonist, SFRP1, is Hedgehog signaling target. *Int J Mol Med* **17**:171–175.

Kicheva A, Briscoe J. 2023. Control of Tissue Development by Morphogens. *Annu Rev Cell Dev Biol* **39**:91–121.

Kim S-E, Lei Y, Hwang S-H, Wlodarczyk BJ, Mukhopadhyay S, Shaw GM, Ross ME, Finnell RH. 2019. Dominant negative GPR161 rare variants are risk factors of human spina bifida. *Hum Mol Genet* **28**:200–208.

Kimura Y, Matsunami H, Inoue T, Shimamura K, Uchida N, Ueno T, Miyazaki T, Takeichi M. 1995. Cadherin-11 expressed in association with mesenchymal morphogenesis in the head, somite, and limb bud of early mouse embryos. *Dev Biol* **169**:347–358.

Klymiuk I, Kenner L, Adler T, Busch DH, Boersma A, Irmler M, Fridrich B, Gailus-Durner V, Fuchs H, Leitner N, Müller M, Kühn R, Schlederer M, Treise I, de Angelis MH, Beckers J. 2012. In vivo functional requirement of the mouse Ifitm1 gene for germ cell development, interferon mediated immune response and somitogenesis. *PLoS One* **7**:e44609.

La Manno G, Siletti K, Furlan A, Gyllborg D, Vinsland E, Mossi Albiach A, Mattsson Langseth C, Khven I, Lederer AR, Dratva LM, Johnsson A, Nilsson M, Lönnerberg P, Linnarsson S. 2021. Molecular architecture of the developing mouse brain. *Nature* 1–5.

Lagutin OV, Zhu CC, Kobayashi D, Topczewski J, Shimamura K, Puelles L, Russell HRC, McKinnon PJ, Solnica-Krezel L, Oliver G. 2003. Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. *Genes Dev* **17**:368–379.

Langlieb J, Sachdev NS, Balderrama KS, Nadaf NM, Raj M, Murray E, Webber JT, Vanderburg C, Gazestani V, Tward D, Mezias C, Li X, Cable DM, Norton T, Mitra P, Chen F, Macosko EZ. 2023. The cell type composition of the adult mouse brain revealed by single cell and spatial genomics. *bioRxiv*. doi:10.1101/2023.03.06.531307

Lee CS, Fan CM. 2001. Embryonic expression patterns of the mouse and chick Gas1 genes. *Mech Dev* **101**:293–297.

Lee J, Platt KA, Censullo P, Ruiz i Altaba A. 1997. Gli1 is a target of Sonic hedgehog that induces ventral neural tube development. *Development* **124**:2537–2552.

Lee JD, Silva-Gagliardi NF, Tepass U, McGlade CJ, Anderson KV. 2007. The FERM protein Epb4.1l5 is required for organization of the neural plate and for the epithelial-mesenchymal transition at the primitive streak of the mouse embryo. *Development* **134**:2007–2016.

Levine JH, Simonds EF, Bendall SC, Davis KL, Amir E-AD, Tadmor MD, Litvin O, Fienberg HG, Jager A, Zunder ER, Finck R, Gedman AL, Radtke I, Downing JR, Pe'er D, Nolan GP. 2015. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. *Cell* **162**:184–197.

Lewis SL, Khoo P-L, Andrea De Young R, Bildsoe H, Wakamiya M, Behringer RR, Mukhopadhyay M, Westphal H, Tam PPL. 2007. Genetic interaction of Gsc and Dkk1 in head morphogenesis of the mouse. *Mech Dev* **124**:157–165.

Lex RK, Ji Z, Falkenstein KN, Zhou W, Henry JL, Ji H, Vokes SA. 2020. GLI transcriptional repression regulates tissue-specific enhancer activity in response to Hedgehog signaling. *Elife* **9**. doi:10.7554/eLife.50670

Li JYH, Lao Z, Joyner AL. 2002. Changing requirements for Gbx2 in development of the cerebellum and maintenance of the mid/hindbrain organizer. *Neuron* **36**:31–43.

Li Q, Cheng Z, Zhou L, Darmanis S, Neff NF, Okamoto J, Gulati G, Bennett ML, Sun LO, Clarke LE, Marschallinger J, Yu G, Quake SR, Wyss-Coray T, Barres BA. 2019. Developmental heterogeneity of microglia and brain myeloid cells revealed by deep single-cell RNA sequencing. *Neuron* **101**:207-223.e10.

Liguori GL, Echevarría D, Improta R, Signore M, Adamson E, Martínez S, Persico MG. 2003. Anterior neural plate regionalization in cripto null mutant mouse embryos in the absence of node and primitive streak. *Dev Biol* **264**:537–549.

Lin C-Y, Huang P-H, Liao W-L, Cheng H-J, Huang C-F, Kuo J-C, Patton WA, Massenburg D, Moss J, Lee F-JS. 2000. ARL4, an ARF-like Protein That Is Developmentally Regulated and Localized to Nuclei and Nucleoli*. *J Biol Chem* **275**:37815–37823.

Liu A, Joyner AL. 2001a. Early anterior/posterior patterning of the midbrain and cerebellum. *Annu Rev Neurosci* **24**:869–896.

Liu A, Joyner AL. 2001b. EN and GBX2 play essential roles downstream of FGF8 in patterning the mouse mid/hindbrain region. *Development* **128**:181–191.

Lobe CG. 1997. Expression of the helix-loop-helix factor, Hes3, during embryo development suggests a role in early midbrain-hindbrain patterning. *Mech Dev* **62**:227–237.

Lohoff T, Ghazanfar S, Missarova A, Koulena N, Pierson N, Griffiths JA, Bardot ES, Eng C-HL, Tyser RCV, Argelaguet R, Guibentif C, Srinivas S, Briscoe J, Simons BD, Hadjantonakis A-K, Göttgens B, Reik W, Nichols J, Cai L, Marioni JC. 2022. Integration of spatial and single-cell transcriptomic data elucidates mouse organogenesis. *Nat Biotechnol* **40**:74–85.

Longo SK, Guo MG, Ji AL, Khavari PA. 2021. Integrating single-cell and spatial transcriptomics to elucidate intercellular tissue dynamics. *Nat Rev Genet* **22**:627–644.

Lyn S, Giguère V. 1994. Localization of CRABP-I and CRABP-II mRNA in the early mouse embryo by whole-mount in situ hybridization: implications for teratogenesis and neural development. *Dev Dyn* **199**:280–291.

Maden M. 2007. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nat Rev Neurosci* **8**:755–765.

Maniou E, Staddon MF, Marshall AR, Greene NDE, Copp AJ, Banerjee S, Galea GL. 2021. Hindbrain neuropore tissue geometry determines asymmetric cell-mediated closure dynamics in mouse embryos. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/pnas.2023163118

Marcogliese PC, Shashi V, Spillmann RC, Stong N, Rosenfeld JA, Koenig MK, Martínez-Agosto JA, Herzog M, Chen AH, Dickson PI, Lin HJ, Vera MU, Salamon N, Graham JM Jr, Ortiz D, Infante E, Steyaert W, Dermaut B, Poppe B, Chung H-L, Zuo Z, Lee P-T, Kanca O, Xia F, Yang Y, Smith EC, Jasien J, Kansagra S, Spiridigliozzi G, El-Dairi M, Lark R, Riley K, Koeberl DD, Golden-Grant K, Program for Undiagnosed Diseases (UD-PrOZA), Undiagnosed Diseases Network, Yamamoto S, Wangler MF, Mirzaa G, Hemelsoet D, Lee B, Nelson SF, Goldstein DB, Bellen HJ, Pena LDM. 2018. IRF2BPL Is Associated with Neurological Phenotypes. *Am J Hum Genet* **103**:245–260.

Marigo V, Johnson RL, Vortkamp A, Tabin CJ. 1996. Sonic hedgehog differentially regulates expression of GLI and GLI3 during limb development. *Dev Biol* **180**:273–283.

Mathers PH, Grinberg A, Mahon KA, Jamrich M. 1997. The Rx homeobox gene is essential for vertebrate eye development. *Nature* **387**:603–607.

Matsuo I, Kuratani S, Kimura C, Takeda N, Aizawa S. 1995. Mouse Otx2 functions in the formation and patterning of rostral head. *Genes Dev* **9**:2646–2658.

McInnes L, Healy J, Melville J. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv [statML]*.

McMahon AP, Ingham PW, Tabin CJ. 2003. Developmental roles and clinical significance of hedgehog signaling. *Curr Top Dev Biol* **53**:1–114.

Medina-Cano D, Corrigan EK, Glenn RA, Islam MT, Lin Y, Kim J, Cho H, Vierbuchen T. 2022. Rapid and robust directed differentiation of mouse epiblast stem cells into definitive endoderm and forebrain organoids. *Development* **149**:dev200561.

Millet S, Campbell K, Epstein DJ, Losos K, Harris E, Joyner AL. 1999. A role for Gbx2 in repression of Otx2 and positioning the mid/hindbrain organizer. *Nature* **401**:161–164.

Mittnenzweig M, Mayshar Y, Cheng S, Ben-Yair R, Hadas R, Rais Y, Chomsky E, Reines N, Uzonyi A, Lumerman L, Lifshitz A, Mukamel Z, Orenbuch A-H, Tanay A, Stelzer Y. 2021. A single-embryo, single-cell time-resolved model for mouse gastrulation. *Cell* **184**:2825-2842.e22.

Moffitt JR, Lundberg E, Heyn H. 2022. The emerging landscape of spatial profiling technologies. *Nat Rev Genet* **23**:741–759.

Molè MA, Galea GL, Rolo A, Weberling A, Nychyk O, De Castro SC, Savery D, Fässler R, Ybot-González P, Greene NDE, Copp AJ. 2020. Integrin-Mediated Focal Anchorage Drives Epithelial Zippering during Mouse Neural Tube Closure. *Dev Cell* **52**:321-334.e6.

Moses L, Pachter L. 2022. Museum of spatial transcriptomics. *Nat Methods* **19**:534–546.

Mukhopadhyay S, Wen X, Ratti N, Loktev A, Rangell L, Scales SJ, Jackson PK. 2013. The ciliary G-protein-coupled receptor Gpr161 negatively regulates the Sonic hedgehog pathway via cAMP signaling. *Cell* **152**:210–223.

Murdoch JN, Copp AJ. 2010. The relationship between sonic Hedgehog signaling, cilia, and neural tube defects. *Birth Defects Res A Clin Mol Teratol* **88**:633–652.

Nagai T, Aruga J, Minowa O, Sugimoto T, Ohno Y, Noda T, Mikoshiba K. 2000. Zic2 regulates the kinetics of neurulation. *Proc Natl Acad Sci U S A* **97**:1618–1623.

Nishi Y, Zhang X, Jeong J, Peterson KA, Vedenko A, Bulyk ML, Hide WA, McMahon AP. 2015. A direct fate exclusion mechanism by Sonic hedgehog-regulated transcriptional repressors. *Development* **142**:3286–3293.

Nowotschin S, Setty M, Kuo Y-Y, Liu V, Garg V, Sharma R, Simon CS, Saiz N, Gardner R, Boutet SC, Church DM, Hoodless PA, Hadjantonakis A-K, Pe'er D. 2019. The emergent landscape of the mouse gut endoderm at single-cell resolution. *Nature* **569**:361–367.

Okada A, Charron F, Morin S, Shin DS, Wong K, Fabre PJ, Tessier-Lavigne M, McConnell SK. 2006. Boc is a receptor for sonic hedgehog in the guidance of commissural axons. *Nature* **444**:369–373.

Oulad-Abdelghani M, Chazaud C, Bouillet P, Sapin V, Chambon P, Dollé P. 1997. Meis2, a novel mouse Pbx-related homeobox gene induced by retinoic acid during differentiation of P19 embryonal carcinoma cells. *Dev Dyn* **210**:173–183.

Pabst O, Herbrand H, Arnold HH. 1998. Nkx2-9 is a novel homeobox transcription factor which demarcates ventral domains in the developing mouse CNS. *Mech Dev* **73**:85–93.

Peñagarikano O, Abrahams BS, Herman EI, Winden KD, Gdalyahu A, Dong H, Sonnenblick LI, Gruver R, Almajano J, Bragin A, Golshani P, Trachtenberg JT, Peles E, Geschwind DH. 2011. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **147**:235–246.

Peterson KA, Nishi Y, Ma W, Vedenko A, Shokri L, Zhang X, McFarlane M, Baizabal J-M, Junker JP, van Oudenaarden A, Mikkelsen T, Bernstein BE, Bailey TL, Bulyk ML, Wong WH, McMahon AP. 2012. Neural-specific Sox2 input and differential Gli-binding affinity provide context and positional information in Shh-directed neural patterning. *Genes Dev* **26**:2802–2816.

Pijuan-Sala B, Griffiths JA, Guibentif C, Hiscock TW, Jawaid W, Calero-Nieto FJ, Mulas C, Ibarra-Soria X, Tyser RCV, Ho DLL, Reik W, Srinivas S, Simons BD, Nichols J, Marioni JC, Göttgens B. 2019. A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature* **566**:490–495.

Pryor SE, Massa V, Savery D, Andre P, Yang Y, Greene NDE, Copp AJ. 2014. Vangl-dependent planar cell polarity signalling is not required for neural crest migration in mammals. *Development* **141**:3153–3158.

Radice GL, Rayburn H, Matsunami H, Knudsen KA, Takeichi M, Hynes RO. 1997. Developmental defects in mouse embryos lacking N-cadherin. *Dev Biol* **181**:64–78.

Rao A, Barkley D, França GS, Yanai I. 2021. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**:211–220.

Ribes V, Briscoe J. 2009. Establishing and interpreting graded Sonic Hedgehog signaling during vertebrate neural tube patterning: the role of negative feedback. *Cold Spring Harb Perspect Biol* **1**:a002014.

Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, Pun SH, Sellers DL, Tasic B, Seelig G. 2018. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* **360**:176–182.

Ruan X, Kang B, Qi C, Lin W, Wang J, Zhang X. 2021. Progenitor cell diversity in the developing mouse neocortex. *Proc Natl Acad Sci U S A* **118**. doi:10.1073/pnas.2018866118

Sampath Kumar A, Tian L, Bolondi A, Hernández AA, Stickels R, Kretzmer H, Murray E, Wittler L, Walther M, Barakat G, Haut L, Elkabetz Y, Macosko EZ, Guignard L, Chen F, Meissner A. 2023. Spatiotemporal transcriptomic maps of whole mouse embryos at the onset of organogenesis. *Nat Genet* **55**:1176–1185.

Sander M, Paydar S, Ericson J, Briscoe J, Berber E, German M, Jessell TM, Rubenstein JL. 2000. Ventral neural patterning by *Nkx* homeobox genes: *Nkx6.1* controls somatic motor neuron and ventral interneuron fates. *Genes Dev* **14**:2134–2139.

Sasaki H, Hui C, Nakafuku M, Kondoh H. 1997. A binding site for Gli proteins is essential for HNF-3beta floor plate enhancer activity in transgenics and can respond to Shh in vitro. *Development* **124**:1313–1322.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**:676–682.

Schwarz M, Alvarez-Bolado G, Dressler G, Urbánek P, Busslinger M, Gruss P. 1999. Pax2/5 and Pax6 subdivide the early neural tube into three domains. *Mech Dev* **82**:29–39.

Setty M, Kiseliovas V, Levine J, Gayoso A, Mazutis L, Pe'er D. 2019. Characterization of cell fate probabilities in single-cell data with Palantir. *Nat Biotechnol* **37**:451–460.

Shao X, Liao J, Li C, Lu X, Cheng J, Fan X. 2021. CellTalkDB: a manually curated database of ligand-receptor interactions in humans and mice. *Brief Bioinform* **22**. doi:10.1093/bib/bbaa269

Shao X, Lu X, Liao J, Chen H, Fan X. 2020. New avenues for systematically inferring cell-cell communication: through single-cell transcriptomics data. *Protein Cell* **11**:866–880.

Shimamura K, Martinez S, Puelles L, Rubenstein JL. 1997. Patterns of gene expression in the neural plate and neural tube subdivide the embryonic forebrain into transverse and longitudinal domains. *Dev Neurosci* **19**:88–96.

Shimamura K, Rubenstein JL. 1997. Inductive interactions direct early regionalization of the mouse forebrain. *Development* **124**:2709–2718.

Shylo NA, Emmanouil E, Ramrattan D, Weatherbee SD. 2020. Loss of ciliary transition zone protein TMEM107 leads to heterotaxy in mice. *Dev Biol* **460**:187–199.

Smart SL, Lopantsev V, Zhang CL, Robbins CA, Wang H, Chiu SY, Schwartzkroin PA, Messing A, Tempel BL. 1998. Deletion of the K(V)1.1 potassium channel causes epilepsy in mice. *Neuron* **20**:809–819.

Soldatov R, Kaucka M, Kastriti ME, Petersen J, Chontorotzea T, Englmaier L, Akkuratova N, Yang Y, Häring M, Dyachuk V, Bock C, Farlik M, Piacentino ML, Boismoreau F, Hilscher MM, Yokota C, Qian X, Nilsson M, Bronner ME, Croci L, Hsiao W-Y, Guertin DA, Brunet J-F, Consalez GG, Ernfors P, Fried K, Kharchenko PV, Adameyko I. 2019. Spatiotemporal structure of cell fate decisions in murine neural crest. *Science* **364**:eaas9536.

Tabar V, Studer L. 2014. Pluripotent stem cells in regenerative medicine: challenges and recent progress. *Nat Rev Genet* **15**:82–92.

Tahayato A, Dollé P, Petkovich M. 2003. Cyp26C1 encodes a novel retinoic acid-metabolizing enzyme expressed in the hindbrain, inner ear, first branchial arch and tooth buds during murine development. *Gene Expr Patterns* **3**:449–454.

Tam PPL, Ho JWK. 2020. Cellular diversity and lineage trajectory: insights from mouse single cell transcriptomes. *Development* **147**. doi:10.1242/dev.179788

Tamplin OJ, Cox BJ, Rossant J. 2011. Integrated microarray and ChIP analysis identifies multiple Foxa2 dependent target genes in the notochord. *Dev Biol* **360**:415–425.

Tamplin OJ, Kinzel D, Cox BJ, Bell CE, Rossant J, Lickert H. 2008. Microarray analysis of Foxa2 mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. *BMC Genomics* **9**:511.

Tang M, Luo SX, Tang V, Huang EJ. 2013. Temporal and spatial requirements of Smoothened in ventral midbrain neuronal development. *Neural Dev* **8**:8.

Tenzen T, Allen BL, Cole F, Kang J-S, Krauss RS, McMahon AP. 2006. The cell surface membrane proteins Cdo and Boc are components and targets of the Hedgehog signaling pathway and feedback network in mice. *Dev Cell* **10**:647–656.

Theiler K. 1989. The house mouse. Berlin, Heidelberg: Springer Berlin Heidelberg.

Tian E, Kimura C, Takeda N, Aizawa S, Matsuo I. 2002. Otx2 is required to respond to signals from anterior neural ridge for forebrain specification. *Dev Biol* **242**:204–223.

Torban E, Patenaude A-M, Leclerc S, Rakowiecki S, Gauthier S, Andelfinger G, Epstein DJ, Gros P. 2008. Genetic interaction between members of the Vangl family causes neural tube defects in mice. *Proc Natl Acad Sci U S A* **105**:3449–3454.

Torres M, Gómez-Pardo E, Gruss P. 1996. Pax2 contributes to inner ear patterning and optic nerve trajectory. *Development* **122**:3381–3391.

Trainor PA, Sobieszczuk D, Wilkinson D, Krumlauf R. 2002. Signalling between the hindbrain and paraxial tissues dictates neural crest migration pathways. *Development* **129**:433–442.

Treichel D, Becker MB, Gruss P. 2001. The novel transcription factor gene Sp5 exhibits a dynamic and highly restricted expression pattern during mouse embryogenesis. *Mech Dev* **101**:175–179.

Treichel D, Schöck F, Jäckle H, Gruss P, Mansouri A. 2003. mBtd is required to maintain signaling during murine limb development. *Genes Dev* **17**:2630–2635.

Ulloa F, Martí E. 2010. Wnt won the war: antagonistic role of Wnt over Shh controls dorso-ventral patterning of the vertebrate neural tube. *Dev Dyn* **239**:69–76.

Vacalla CMH, Theil T. 2002. Cst, a novel mouse gene related to Drosophila Castor, exhibits dynamic expression patterns during neurogenesis and heart development. *Mech Dev* **118**:265–268.

van Dijk D, Sharma R, Nainys J, Yim K, Kathail P, Carr AJ, Burdziak C, Moon KR, Chaffer CL, Pattabiraman D, Bierie B, Mazutis L, Wolf G, Krishnaswamy S, Pe'er D. 2018. Recovering Gene Interactions from Single-Cell Data Using Data Diffusion. *Cell* **174**:716-729.e27.

Wang B, Fallon JF, Beachy PA. 2000. Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb. *Cell* **100**:423–434.

Wang Q, Ding S-L, Li Y, Royall J, Feng D, Lesnar P, Graddis N, Naeemi M, Facer B, Ho A, Dolbeare T, Blanchard B, Dee N, Wakeman W, Hirokawa KE, Szafer A, Sunkin SM, Oh SW, Bernard A, Phillips JW, Hawrylycz M, Koch C, Zeng H, Harris JA, Ng L. 2020. The Allen Mouse Brain Common Coordinate Framework: A 3D reference atlas. *Cell* **181**:936-953.e20.

Wassarman KM, Lewandoski M, Campbell K, Joyner AL, Rubenstein JL, Martinez S, Martin GR. 1997. Specification of the anterior hindbrain and establishment of a normal mid/hindbrain organizer is dependent on Gbx2 gene function. *Development* **124**:2923–2934.

Waters ST, Wilson CP, Lewandoski M. 2003. Cloning and embryonic expression analysis of the mouse Gbx1 gene. *Gene Expr Patterns* **3**:313–317.

Williams RM, Lukoseviciute M, Sauka-Spengler T, Bronner ME. 2022. Single-cell atlas of early chick development reveals gradual segregation of neural crest lineage from the neural plate border during neurulation. *Elife* **11**. doi:10.7554/eLife.74464

Wizeman JW, Guo Q, Wilion EM, Li JY. 2019. Specification of diverse cell types during early neurogenesis of the mouse cerebellum. *Elife* **8**. doi:10.7554/eLife.42388

Wolf FA, Angerer P, Theis FJ. 2018. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**:15.

Wong RL, Chan KK, Chow KL. 1999. Developmental expression of Mab21l2 during mouse embryogenesis. *Mech Dev* **87**:185–188.

Woods IG, Talbot WS. 2005. The you gene encodes an EGF-CUB protein essential for Hedgehog signaling in zebrafish. *PLoS Biol* **3**:e66.

Wright TJ, Mansour SL. 2003. Fgf3 and Fgf10 are required for mouse otic placode induction. *Development* **130**:3379–3390.

Wurst W, Auerbach AB, Joyner AL. 1994. Multiple developmental defects in Engrailed-1 mutant mice: an early mid-hindbrain deletion and patterning defects in forelimbs and sternum. *Development* **120**:2065–2075.

Wurst W, Bally-Cuif L. 2001. Neural plate patterning: upstream and downstream of the isthmic organizer. *Nat Rev Neurosci* **2**:99–108.

Xu Q, Guo L, Moore H, Waclaw RR, Campbell K, Anderson SA. 2010. Sonic hedgehog signaling confers ventral telencephalic progenitors with distinct cortical interneuron fates. *Neuron* **65**:328–340.

Xuan S, Baptista CA, Balas G, Tao W, Soares VC, Lai E. 1995. Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. *Neuron* **14**:1141–1152.

Yamagishi H, Maeda J, Hu T, McAnally J, Conway SJ, Kume T, Meyers EN, Yamagishi C, Srivastava D. 2003. Tbx1 is regulated by tissue-specific forkhead proteins through a common Sonic hedgehog-responsive enhancer. *Genes Dev* **17**:269–281.

Yamaguchi TP. 2001. Heads or tails: Wnts and anterior-posterior patterning. *Curr Biol* **11**:R713-24.

Yao S, Lum L, Beachy P. 2006. The ihog cell-surface proteins bind Hedgehog and mediate pathway activation. *Cell* **125**:343–357.

Yao Z, van Velthoven CTJ, Kunst M, Zhang M, McMillen D, Lee C, Jung W, Goldy J, Abdelhak A, Baker P, Barkan E, Bertagnolli D, Campos J, Carey D, Casper T, Chakka AB, Chakrabarty R, Chavan S, Chen M, Clark M, Close J, Crichton K, Daniel S, Dolbeare T, Ellingwood L, Gee J, Glandon A, Gloe J, Gould J, Gray J, Guilford N, Guzman J, Hirschstein D, Ho W, Jin K, Kroll M, Lathia K, Leon A, Long B, Maltzer Z, Martin N, McCue R, Meyerdierks E, Nguyen TN, Pham T, Rimorin C, Ruiz A, Shapovalova N, Slaughterbeck C, Sulc J, Tieu M, Torkelson A, Tung H, Cuevas NV, Wadhwani K, Ward K, Levi B, Farrell C, Thompson CL, Mufti S, Pagan CM, Kruse L, Dee N, Sunkin SM, Esposito L, Hawrylycz MJ, Waters J, Ng L, Smith KA, Tasic B, Zhuang X, Zeng H. 2023. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *bioRxiv*. doi:10.1101/2023.03.06.531121

Zalc A, Sinha R, Gulati GS, Wesche DJ, Daszczuk P, Swigut T, Weissman IL, Wysocka J. 2021. Reactivation of the pluripotency program precedes formation of the cranial neural crest. *Science* **371**. doi:10.1126/science.abb4776

Zhang M, Pan X, Jung W, Halpern AR, Eichhorn SW, Lei Z, Cohen L, Smith KA, Tasic B, Yao Z, Zeng H, Zhuang X. 2023. Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624**:343–354.

Zhang W, Kang J-S, Cole F, Yi M-J, Krauss RS. 2006. Cdo functions at multiple points in the Sonic Hedgehog pathway, and Cdo-deficient mice accurately model human holoprosencephaly. *Dev Cell* **10**:657–665.

Zhang Y, Liu T, Wang J, Zou B, Li L, Yao L, Chen K, Ning L, Wu B, Zhao X, Wang D. 2021. Cellinker: a platform of ligand-receptor interactions for intercellular communication analysis. *Bioinformatics*. doi:10.1093/bioinformatics/btab036

Zhao X, Duester G. 2009. Effect of retinoic acid signaling on Wnt/β-catenin and FGF signaling during body axis extension. *Gene Expr Patterns* **9**:430–435.

Zohn IE, Anderson KV, Niswander L. 2007. The Hectd1 ubiquitin ligase is required for development of the head mesenchyme and neural tube closure. *Dev Biol* **306**:208–221.

Zohn IE, Sarkar AA. 2012. Does the cranial mesenchyme contribute to neural fold elevation during neurulation? *Birth Defects Res A Clin Mol Teratol* **94**:841–848.

Figure 1

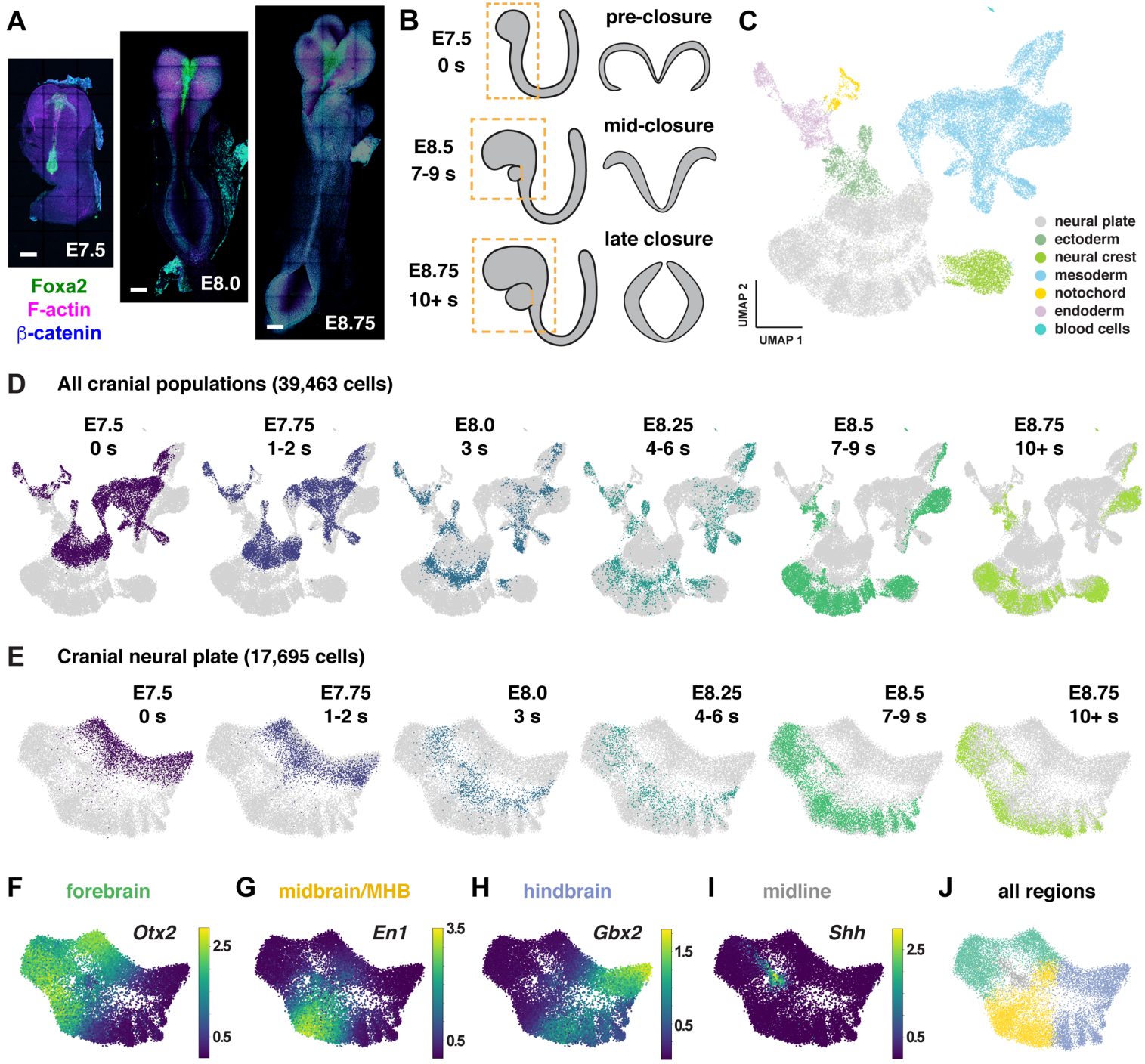
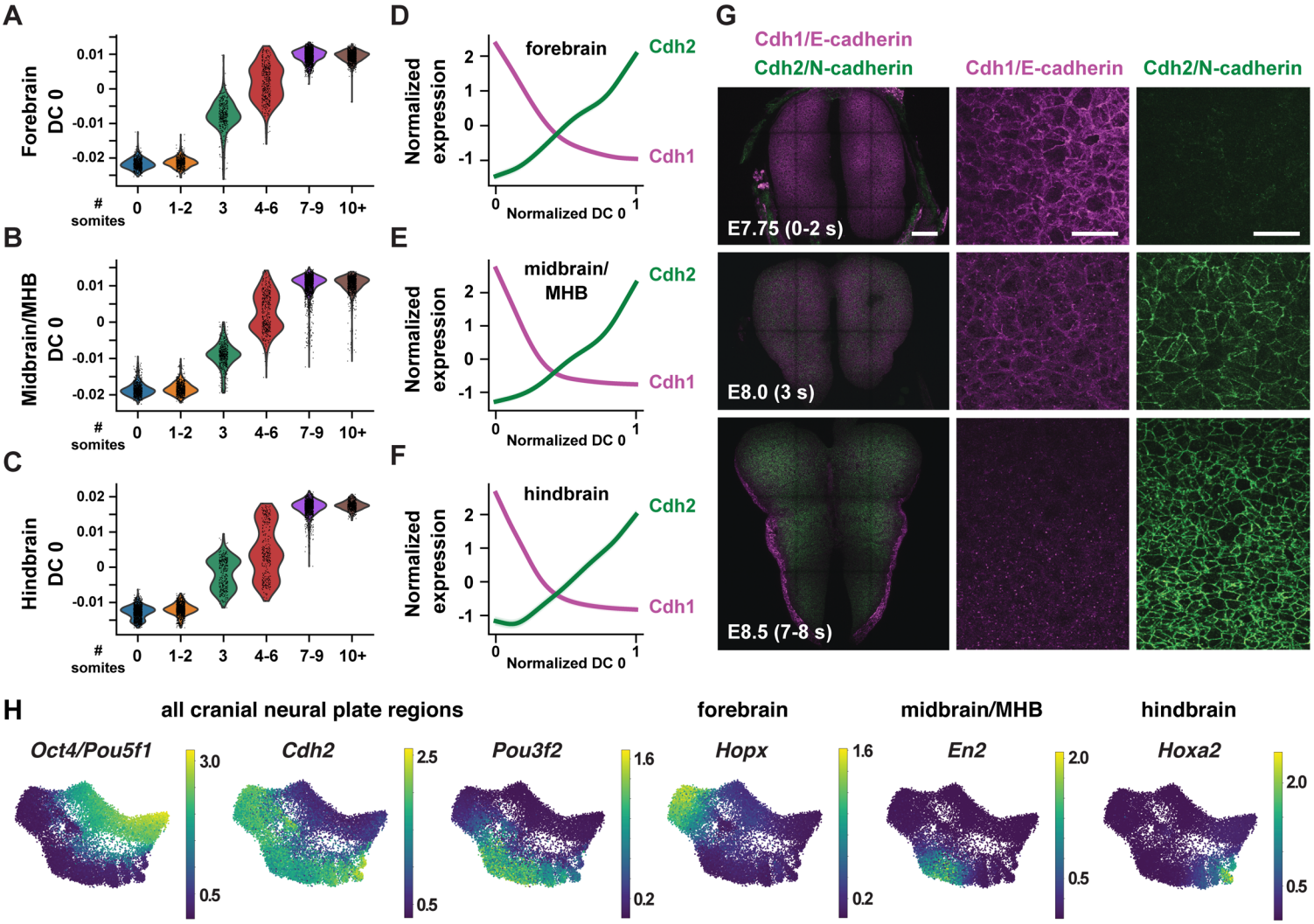

A
E7.5
E8.0
E8.75
Foxa2
F-actin
β-catenin
B
E7.5
0 s
pre-closure
E8.5
7-9 s
mid-closure
E8.75
10+ s
late closure
C
UMAP 2
UMAP 1
neural plate
ectoderm
neural crest
mesoderm
notochord
endoderm
blood cells
D
All cranial populations (39,463 cells)
E7.5
0 s
E7.75
1-2 s
E8.0
3 s
E8.25
4-6 s
E8.5
7-9 s
E8.75
10+ s
E
Cranial neural plate (17,695 cells)
E7.5
0 s
E7.75
1-2 s
E8.0
3 s
E8.25
4-6 s
E8.5
7-9 s
E8.75
10+ s
F
forebrain
Otx2
2.5
0.5
G
midbrain/MHB
En1
3.5
0.5
H
hindbrain
Gbx2
1.5
0.5
I
midline
Shh
2.5
0.5
J
all regions

Figure 2


A
Forebrain DC 0
0.01
0
-0.01
-0.02
somites
0
1-2
3
4-6
7-9
10+
B
Midbrain/MHB DC 0
0.01
0
-0.01
-0.02
somites
0
1-2
3
4-6
7-9
10+
C
Hindbrain DC 0
0.02
0.01
0
-0.01
somites
0
1-2
3
4-6
7-9
10+
D
forebrain
Cdh2
Cdh1
Normalized expression
2
1
0
-1
0
Normalized DC 0
1
E
midbrain/ MHB
Cdh2
Cdh1
Normalized expression
2
1
0
-1
0
Normalized DC 0
1
F
hindbrain
Cdh2
Cdh1
Normalized expression
2
1
0
-1
0
Normalized DC 0
1
G
Cdh1/E-cadherin
Cdh2/N-cadherin
Cdh1/E-cadherin
Cdh2/N-cadherin
E7.75 (0-2 s)
E8.0 (3 s)
E8.5 (7-8 s)

H
all cranial neural plate regions
forebrain
midbrain/MHB
hindbrain
Oct4/Pou5f1
3.0
0.5
Cdh2
2.5
0.5
Pou3f2
1.6
0.2
Hopx
1.6
0.2
En2
2.0
0.5
Hoxa2
2.0
0.5

Figure 3

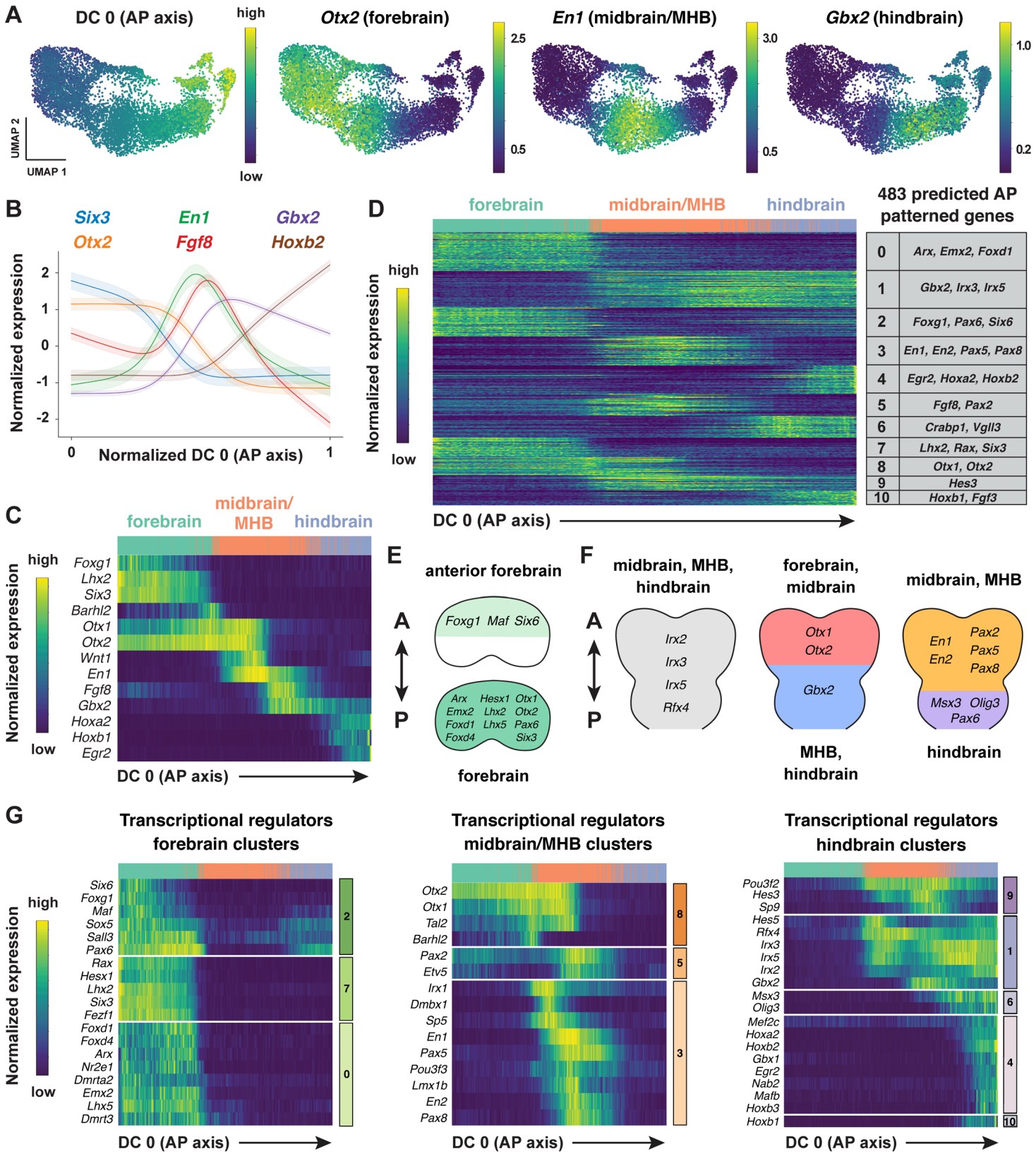

A
DC 0 (AP axis)
high
low
UMAP 2
UMAP 1
Otx2 (forebrain)
2.5
0.5
En1 (midbrain/MHB)
3.0
0.5
Gbx2 (hindbrain)
1.0
0.2
B
Six3
Otx2
En1
Fgf8
Gbx2
Hoxb2
Normalized expression
Normalized DC 0 (AP axis)
D
forebrain
midbrain/MHB
hindbrain
high
low
Normalized expression
DC 0 (AP axis)
483 predicted AP patterned genes
0 Arx, Emx2, Foxd1
1 Gbx2, Irx3, Irx5
2 Foxg1, Pax6, Six6
3 En1, En2, Pax5, Pax8
4 Egr2, Hoxa2, Hoxb2
5 Fgf8, Pax2
6 Crabp1, Vgll3
7 Lhx2, Rax, Six3
8 Otx1, Otx2
9 Hes3
10 Hoxb1, Fgf3
C
forebrain
midbrain/ MHB
hindbrain
high
low
Normalized expression
Foxg1
Lhx2
Six3
Barhl2
Otx1
Otx2
Wnt1
En1
Fgf8
Gbx2
Hoxa2
Hoxb1
Egr2
DC 0 (AP axis)
E
anterior forebrain
A
P
Foxg1 Maf Six6
Arx Hesx1 Otx1
Emx2 Lhx2 Otx2
Foxd1 Lhx5 Pax6
Foxd4 Six3
forebrain
F
midbrain, MHB, hindbrain
forebrain, midbrain
midbrain, MHB
A
P
Irx2
Irx3
Irx5
Rfx4
Otx1
Otx2
Gbx2
En1
En2
Pax2
Pax5
Pax8
Msx3 Olig3
Pax6
MHB, hindbrain
hindbrain
G
Transcriptional regulators forebrain clusters
Transcriptional regulators midbrain/MHB clusters
Transcriptional regulators hindbrain clusters
high
low
Normalized expression
Six6
Foxg1
Maf
Sox5
Sall3
Pax6
Rax
Hesx1
Lhx2
Six3
Fezf1
Foxd1
Foxd4
Arx
Nr2e1
Dmrta2
Emx2
Lhx5
Dmrt3
2
7
0
Otx2
Otx1
Tal2
Barhl2
Pax2
Etv5
Irx1
Dmbx1
Sp5
En1
Pax5
Pou3f3
Lmx1b
En2
Pax8
8
5
3
Pou3f2
Hes3
Sp9
Hes5
Rfx4
Irx3
Irx5
Irx2
Gbx2
Msx3
Olig3
Mef2c
Hoxa2
Hoxb2
Gbx1
Egr2
Nab2
Mafb
Hoxb3
Hoxb1
9
1
6
4
10
DC 0 (AP axis)

Figure 4

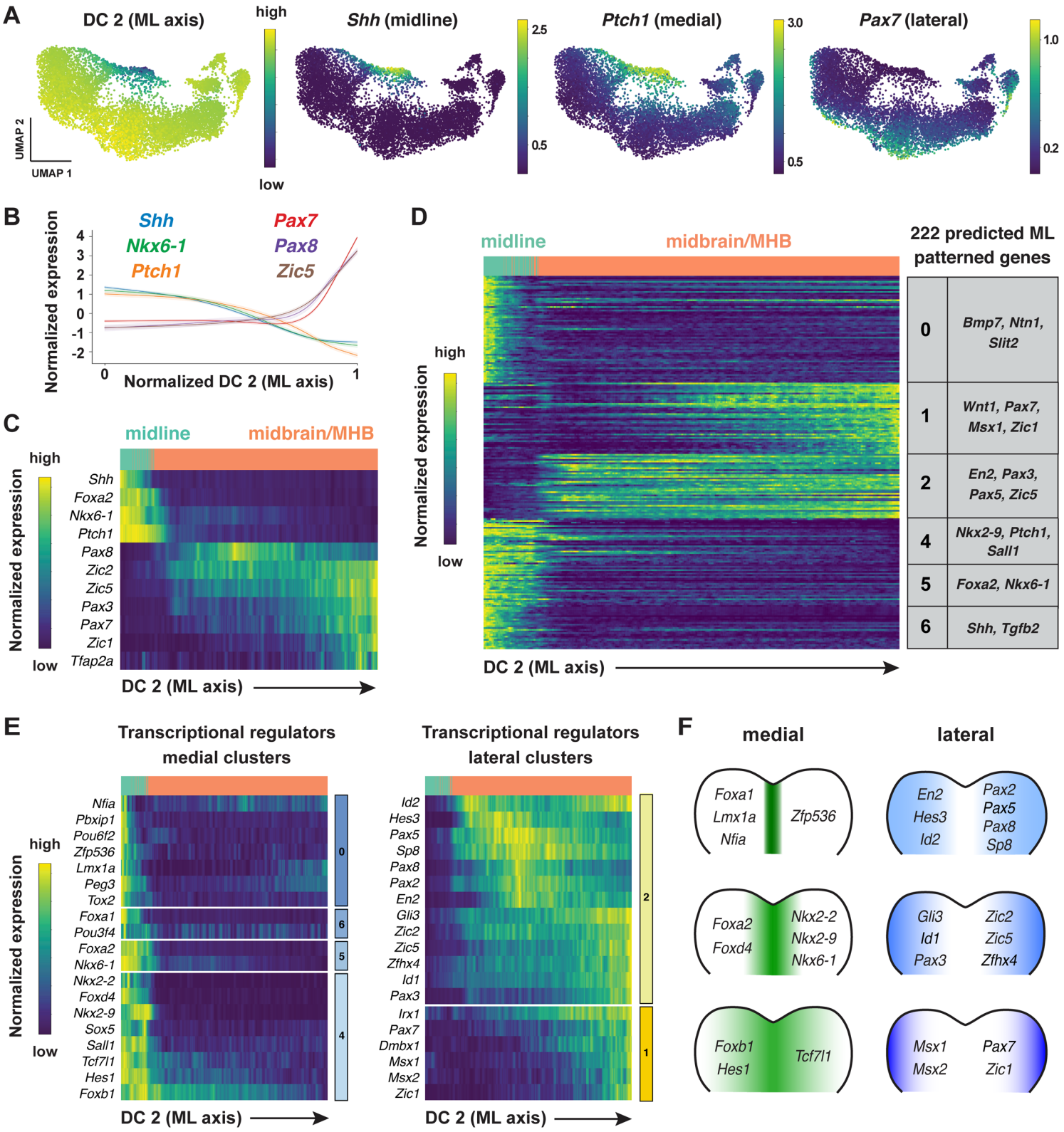

A
DC 2 (ML axis)
high
low
UMAP 2
UMAP 1
Shh (midline)
2.5
0.5
Ptch1 (medial)
3.0
0.5
Pax7 (lateral)
1.0
0.2
B
Shh
Nkx6-1
Ptch1
Pax7
Pax8
Zic5
Normalized expression
Normalized DC 2 (ML axis)
C
midline
midbrain/MHB
high
low
Normalized expression
Shh
Foxa2
Nkx6-1
Ptch1
Pax8
Zic2
Zic5
Pax3
Pax7
Zic1
Tfap2a
DC 2 (ML axis)
D
midline
midbrain/MHB
high
low
Normalized expression
DC 2 (ML axis)
222 predicted ML patterned genes
0 Bmp7, Ntn1, Slit2
1 Wnt1, Pax7, Msx1, Zic1
2 En2, Pax3, Pax5, Zic5
4 Nkx2-9, Ptch1, Sall1
5 Foxa2, Nkx6-1
6 Shh, Tgfb2
E
Transcriptional regulators medial clusters
high
low
Normalized expression
Nfia
Pbxip1
Pou6f2
Zfp536
Lmx1a
Peg3
Tox2
Foxa1
Pou3f4
Foxa2
Nkx6-1
Nkx2-2
Foxd4
Nkx2-9
Sox5
Sall1
Tcf7l1
Hes1
Foxb1
0
6
5
4
DC 2 (ML axis)
Transcriptional regulators lateral clusters
Id2
Hes3
Pax5
Sp8
Pax8
Pax2
En2
Gli3
Zic2
Zic5
Zfhx4
Id1
Pax3
Irx1
Pax7
Dmbx1
Msx1
Msx2
Zic1
2
1
DC 2 (ML axis)
F
medial
lateral
Foxa1
Lmx1a
Nfia
Zfp536
En2
Hes3
Id2
Pax2
Pax5
Pax8
Sp8
Foxa2
Foxd4
Nkx2-2
Nkx2-9
Nkx6-1
Gli3
Id1
Pax3
Zic2
Zic5
Zfhx4
Foxb1
Hes1
Tcf7l1
Msx1
Msx2
Pax7
Zic1

Figure 5

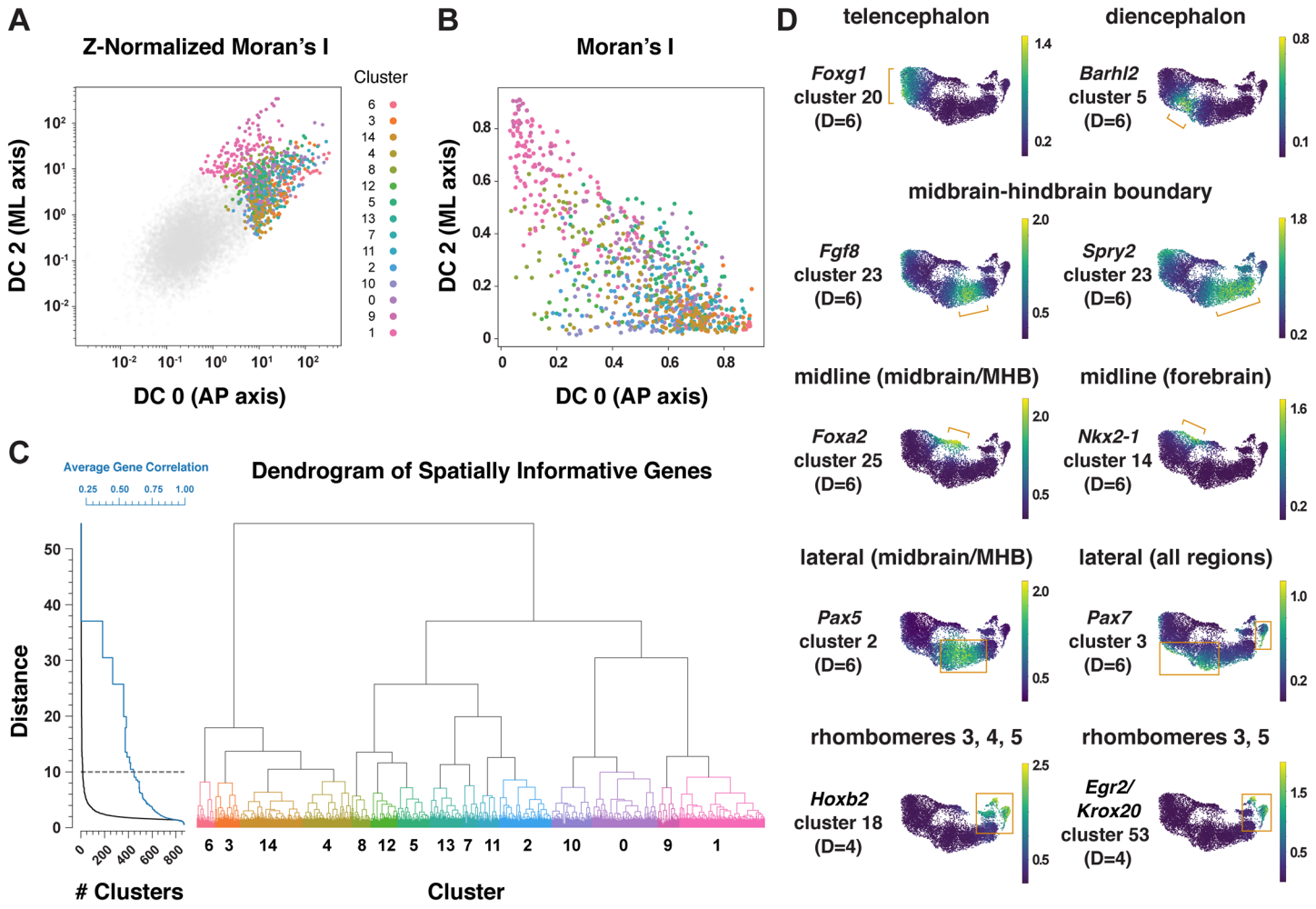

A
Z-Normalized Moran's I
Cluster
6
3
14
4
8
12
5
13
7
11
2
10
0
9
1
DC 2 (ML axis)
DC 0 (AP axis)
B
Moran's I
DC 2 (ML axis)
DC 0 (AP axis)
C
Average Gene Correlation
0.25 0.50 0.75 1.00
Dendrogram of Spatially Informative Genes
Distance
Clusters
Cluster
6 3 14 4 8 12 5 13 7 11 2 10 0 9 1
D
telencephalon
Foxg1
cluster 20
(D=6)
diencephalon
Barhl2
cluster 5
(D=6)
midbrain-hindbrain boundary
Fgf8
cluster 23
(D=6)
Spry2
cluster 23
(D=6)
midline (midbrain/MHB)
Foxa2
cluster 25
(D=6)
midline (forebrain)
Nkx2-1
cluster 14
(D=6)
lateral (midbrain/MHB)
Pax5
cluster 2
(D=6)
lateral (all regions)
Pax7
cluster 3
(D=6)
rhombomeres 3, 4, 5
Hoxb2
cluster 18
(D=4)
rhombomeres 3, 5
Egr2/
Krox20
cluster 53
(D=4)

Figure 6

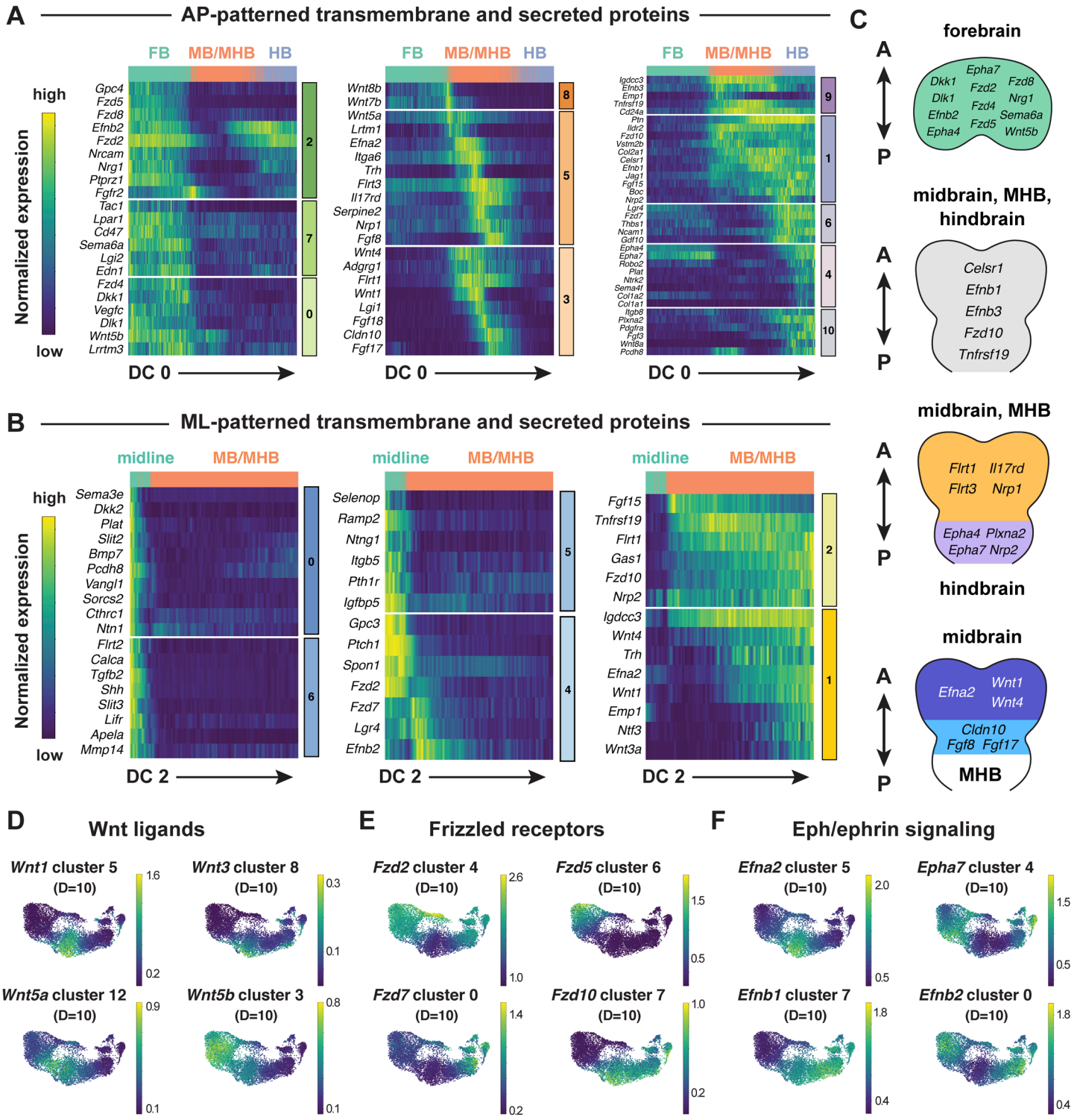
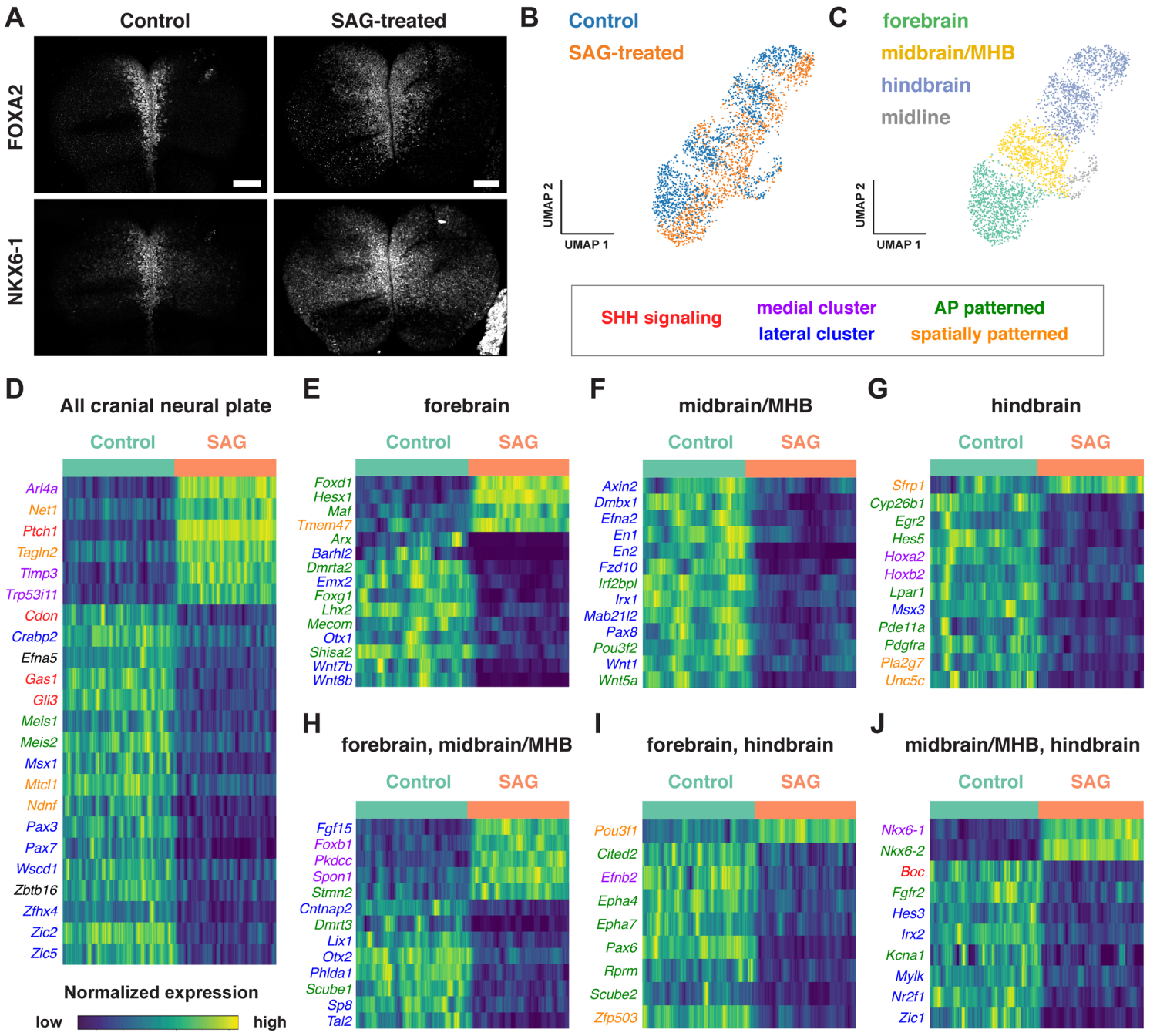

A
AP-patterned transmembrane and secreted proteins
FB
MB/MHB
HB
high
low
Normalized expression
Gpc4
Fzd5
Fzd8
Efnb2
Fzd2
Nrcam
Nrg1
Ptprz1
Fgfr2
Tac1
Lpar1
Cd47
Sema6a
Lgi2
Edn1
Fzd4
Dkk1
Vegfc
Dlk1
Wnt5b
Lrrtm3
Wnt8b
Wnt7b
Wnt5a
Lrtm1
Efna2
Itga6
Trh
Flrt3
Il17rd
Serpine2
Nrp1
Fgf8
Wnt4
Adgrg1
Flrt1
Wnt1
Lgi1
Fgf18
Cldn10
Fgf17
DC 0
B
ML-patterned transmembrane and secreted proteins
midline
MB/MHB
Sema3e
Dkk2
Plat
Slit2
Bmp7
Pcdh8
Vangl1
Sorcs2
Cthrc1
Ntn1
Flrt2
Calca
Tgfb2
Shh
Slit3
Lifr
Apela
Mmp14
Selenop
Ramp2
Ntng1
Itgb5
Pth1r
Igfbp5
Gpc3
Ptch1
Spon1
Fzd2
Fzd7
Lgr4
Efnb2
Fgf15
Tnfrsf19
Flrt1
Gas1
Fzd10
Nrp2
Igdcc3
Wnt4
Trh
Efna2
Wnt1
Emp1
Ntf3
Wnt3a
DC 2
C
forebrain
A
P
Epha7
Dkk1
Fzd2
Fzd8
Dlk1
Fzd4
Nrg1
Efnb2
Fzd5
Sema6a
Epha4
Wnt5b
midbrain, MHB, hindbrain
Celsr1
Efnb1
Efnb3
Fzd10
Tnfrsf19
midbrain, MHB
Flrt1
Il17rd
Flrt3
Nrp1
Epha4
Plxna2
Epha7
Nrp2
hindbrain
midbrain
Efna2
Wnt1
Wnt4
Cldn10
Fgf8
Fgf17
MHB
D
Wnt ligands
Wnt1 cluster 5 (D=10)
Wnt3 cluster 8 (D=10)
Wnt5a cluster 12 (D=10)
Wnt5b cluster 3 (D=10)
E
Frizzled receptors
Fzd2 cluster 4 (D=10)
Fzd5 cluster 6 (D=10)
Fzd7 cluster 0 (D=10)
Fzd10 cluster 7 (D=10)
F
Eph/ephrin signaling
Efna2 cluster 5 (D=10)
Epha7 cluster 4 (D=10)
Efnb1 cluster 7 (D=10)
Efnb2 cluster 0 (D=10)

Figure 7


A
Control
SAG-treated
FOXA2
NKX6-1
B
Control
SAG-treated
UMAP 2
UMAP 1
C
forebrain
midbrain/MHB
hindbrain
midline
UMAP 2
UMAP 1
SHH signaling
medial cluster
lateral cluster
AP patterned
spatially patterned
D
All cranial neural plate
Control
SAG
Arl4a
Net1
Ptch1
Tagln2
Timp3
Trp53i11
Cdon
Crabp2
Efna5
Gas1
Gli3
Meis1
Meis2
Msx1
Mtcl1
Ndnf
Pax3
Pax7
Wscd1
Zbtb16
Zfhx4
Zic2
Zic5
Normalized expression
low
high
E
forebrain
Control
SAG
Foxd1
Hesx1
Maf
Tmem47
Arx
Barhl2
Dmrta2
Emx2
Foxg1
Lhx2
Mecom
Otx1
Shisa2
Wnt7b
Wnt8b
F
midbrain/MHB
Control
SAG
Axin2
Dmbx1
Efna2
En1
En2
Fzd10
Irf2bpl
Irx1
Mab21l2
Pax8
Pou3f2
Wnt1
Wnt5a
G
hindbrain
Control
SAG
Sfrp1
Cyp26b1
Egr2
Hes5
Hoxa2
Hoxb2
Lpar1
Msx3
Pde11a
Pdgfra
Pla2g7
Unc5c
H
forebrain, midbrain/MHB
Control
SAG
Fgf15
Foxb1
Pkdcc
Spon1
Stmn2
Cntnap2
Dmrt3
Lix1
Otx2
Phlda1
Scube1
Sp8
Tal2
I
forebrain, hindbrain
Control
SAG
Pou3f1
Cited2
Efnb2
Epha4
Epha7
Pax6
Rprm
Scube2
Zfp503
J
midbrain/MHB, hindbrain
Control
SAG
Nkx6-1
Nkx6-2
Boc
Fgfr2
Hes3
Irx2
Kcna1
Mylk
Nr2f1
Zic1